Switching Noise and Timing Characteristics in Nanoscale Integrated Circuits

by

Emre Salman

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
School of Engineering and Applied Sciences

University of Rochester
Rochester, New York

2009

Dedication

This work is dedicated to my parents, Nevin Salman and Ayhan Teoman Salman.

Curriculum Vitae

Emre Salman was born in Ankara, Turkey in 1981. He received the B.S. degree in microelectronics engineering from Sabancı University, Istanbul, Turkey in 2004, and the M.S. degree in electrical and computer engineering from the University of Rochester, Rochester, NY in 2006, where he is completing the Ph.D. degree in electrical engineering.

He was a co-op student at ST Microelectronics, Istanbul, Turkey between October 2003 and May 2004, where he worked on the design and verification of a clock and data recovery circuit for a multichannel fiber-optic transceiver. During summer 2005, he was a research and development intern at Synopsys, Mountain View, CA, where he worked on pessimism reduction in static timing analysis (STA) and cell library characterization. During summers 2006 and 2007, he was with Freescale Semiconductor, Tempe, AZ, where he worked on noise reduction techniques at the circuit and physical levels in large scale mixed-signal circuits, computationally efficient substrate noise analysis, and signal isolation methodologies with application to

monolithic transceivers in CMOS and BiCMOS technologies.

His primary research expertise is in the field of noise and timing characteristics in mixed-signal and high performance synchronous digital circuits. Specifically, he is interested in the physical codesign of clock and power distribution networks, the interdependence between noise and timing, noise estimation in large scale circuits, and signal isolation methodologies to enhance signal and power integrity.

Publications

Journal Papers

1. E. Salman, E. G. Friedman, R. M. Secareanu, and O. L. Hartin, "Worst Case Power/Ground Noise Estimation Using an Equivalent Transition Time for Resonance," *IEEE Transactions on Circuits and Systems - I: Regular Papers* (in press).

2. 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* (in press).

3. E. Salman, E. G. Friedman, R. M. Secareanu, and O. L. Hartin, "Identification of Dominant Noise Source and Parameter Sensitivity for Substrate Coupling," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems* (in press).

4. E. Salman, A. Dasdan, F. Taraporevala, K. Küçükçakar, and E. G. Friedman, "Exploiting Setup-Hold Time Interdependence in Static Timing Analysis," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 26, No. 6, pp. 1114–1125, June 2007.

Conference Papers

5. 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*, May 2009 (in press).

6. S. Köse, E. Salman, and E. G. Friedman, "Shielding Methodologies in the Presence of Power/Ground Noise," *Proceedings of the IEEE International Symposium on Circuits and Systems*, May 2009 (in press).

7. S. Köse, E. Salman, Z. Ignjatovic, and E. G. Friedman, "Pseudo-Random Clocking to Enhance Signal Integrity," *Proceedings of the IEEE International SOC Conference*, pp. 47–50, September 2008.

8. 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.

9. E. Salman, R. Jakushokas, E. G. Friedman, R. M. Secareanu, and O. L. Hartin, "Input Port Reduction for Efficient Substrate Extraction in Large Scale ICs," *Proceedings of the IEEE International Symposium on Circuits and Systems*, pp. 376–379, May 2008.

10. E. Salman, E. G. Friedman, R. M. Secareanu, and O. L. Hartin, "Equivalent Rise Time for Resonance in Power/Ground Noise Estimation," *Proceedings of the IEEE International Symposium on Circuits and Systems*, pp. 2422–2425, May 2008.

11. E. Salman, E. G. Friedman, R. M. Secareanu, and O. L. Hartin, "Dominant Substrate Noise Coupling Mechanism for Multiple Switching Gates," *Proceedings of the IEEE International Symposium on Quality Electronic Design*, pp. 261–266, March 2008.

12. E. Salman, E. G. Friedman, R. M. Secareanu, and O. L. Hartin, "Substrate Noise Reduction Based on Noise Aware Cell Design," *Proceedings of the IEEE International Symposium on Circuits and Systems*, pp. 3227–3230, May 2007.

13. E. Salman, E. G. Friedman, and R. M. Secareanu "Substrate and Ground Noise Interactions in Mixed-Signal Circuits," *Proceedings of the IEEE International SOC Conference*, pp. 293–296, September 2006.

14. E. Salman, A. Dasdan, F. Taraporevala, K. Küçükçakar, and E. G. Friedman, "Pessimism Reduction in Static Timing Analysis Using Interdependent Setup and Hold Times," *Proceedings of the IEEE International Symposium on Quality Electronic Design*, pp. 159–164, March 2006.

15. E. Salman, H. Akın, Ö. Gürsoy, A. Ergintav, I. Tekin, A. Bozkurt, and Y. Gürbüz "Design of a 3.2 mW PLL Based Clock and Data Recovery Circuit in 90 nm CMOS Technology," *Proceedings of the Mediterranean Microwave Symposium*, September 2005.

Patent Disclosures

16. R. M. Secareanu, O. L. Hartin, and E. Salman, "Apparatus and Method for Reducing Noise in Mixed-Signal Circuits and Digital Circuits," United States patent pending.

17. A. Dasdan, E. Salman, F. Taraporevala, and K. Küçükçakar, "Characterizing Sequential Cells Using Interdependent Setup and Hold Times, and Utilizing the Sequential Cell Characterization in Static Timing Analysis" United States Patent, No. 7,506,293.

Acknowledgments

The years I spent at the University of Rochester have been full of unforgettable memories. Thanks to many people who deserve my highest gratitude.

First of all, I am deeply grateful to my academic advisor, Professor Eby G. Friedman. I wish I could find a word to better describe his role in this process since he has been much more than just an academic advisor. His mentorship in my research, motivation, dedication, consideration, and generosity were the primary reasons that made me greatly enjoy my last years as a student. He not only gave me the opportunity to complete my doctoral degree, but also provided valuable insight and vision for my professional career. I always feel how lucky I am in having Professor Eby G. Friedman as my advisor. Thank you Professor.

I would like to thank Professors Michael Huang, Paul Ampadu, and Jim Zavislan for serving on my proposal and defense committees. The valuable comments they provided have greatly improved my dissertation. I would also like to express my appreciation to the University of Rochester, and specifically, the Department of

Electrical and Computer Engineering for providing such a friendly and encouraging environment with all the required tools to support high quality academic research.

I am grateful to Dr. Radu M. Secareanu of Freescale Semiconductor for the collaboration, friendship, and technical discussions we enjoyed during my summer internships and over the phone during my academic years. He contributed significantly to my academic research by providing valuable feedback. I would also like to thank Dr. Olin L. Hartin for providing me with the opportunity of two summer internships at Freescale Semiconductor. I would also like to express my gratitude to Marie Burnham and Dr. Len Borucki for their advice on my studies, listening to my presentations, and very sincere friendship during my two warm summers in Arizona.

I thank Dr. Ali Dasdan for his mentorship, collaboration, and suggestions during my internship at Synopsys. I am also grateful to Feroze Taraporevala and Dr. Kayhan Küçükçakar for giving me the opportunity to work at Synopsys, where I had a very fruitful start to my doctoral studies.

I also acknowledge the previous and current members of the High Performance VLSI/IC Design and Analysis Laboratory: Dr. Guoqing Chen, Dr. Mikhail Popovich, Dr. Vasilis Pavlidis, Jonathan Rosenfeld, Renatas Jakushokas, Ioannis Savidis, and Selçuk Köse for their companionship and support. I also thank RuthAnn Williams for her friendship, motivating discussions, and constant help on the administrative side.

I am grateful to all of my friends in Turkey and those in the United States for making it easier to be so far away from home. I thank my relatives for their understanding and support. Last, but not least, I express my deepest appreciation to my parents Nevin Salman and Ayhan Teoman Salman to whom I owe not only my success but every step of my life.

This research was supported in part by the Semiconductor Research Corporation under Contract No. 2004-TJ-1207, the National Science Foundation under Contract No. CCF-0541206, grants from the New York State Office of Science, Technology & Academic Research to the Center for Advanced Technology in Electronic Imaging Systems, and by grants from Intel Corporation, Eastman Kodak Corporation, and Freescale Semiconductor Corporation.

Abstract

Continuous progress in the design and manufacturing of integrated circuits (ICs) has enabled the integration of more than two billion transistors on the same die with clock frequencies well above several gigahertz. These improvements have triggered the era of system-on-chip (SoC) and system-in-package (SiP), drastically changing the classical understanding of noise in complex ICs. Traditionally, device noise has been the primary concern for analog ICs while digital ICs have typically been considered to be relatively immune to noise. This situation has changed significantly due to denser integration and faster signal transition times. Specifically, *switching noise* has become a primary design criterion for both mixed-signal and high performance synchronous digital ICs.

Voltage fluctuations on the power/ground nodes of a circuit, *i.e.*, power/ground noise, is a type of switching noise affecting both mixed-signal and digital ICs. A methodology is proposed to accurately estimate the worst case power/ground noise in an inductive power/ground distribution network with a decoupling capacitor. In

mixed-signal ICs, power/ground noise affects the highly sensitive analog/RF blocks through the monolithic substrate, degrading critical performance parameters such as gain, bandwidth, dynamic range, total harmonic distortion, and phase noise. Several approaches are presented to efficiently model and alleviate substrate noise coupling in mixed-signal ICs. The proposed analysis process determines the noise characteristics of a circuit, thereby identifying appropriate noise isolation techniques. A methodology is also proposed to reduce noise by incorporating noise-aware standard cells. The proposed methodologies and algorithms are validated with industrial circuits, exhibiting significant improvement in computational efficiency while maintaining sufficient accuracy in the noise voltage. A significant reduction in substrate coupling noise is also demonstrated.

In synchronous digital ICs, switching noise affects the timing characteristics of a circuit by generating additional delay uncertainty, possibly degrading system performance or causing a circuit to fail. Interdependent setup and hold times are characterized and exploited to compensate for delay uncertainty, producing a more robust circuit tolerant to switching noise. The proposed algorithms are demonstrated on industrial circuits, verifying the efficiency of exploiting interdependence in reducing delay uncertainty. The research presented in this dissertation provides methodologies and algorithms for designing both mixed-signal and synchronous digital ICs with superior noise performance and enhanced signal integrity.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

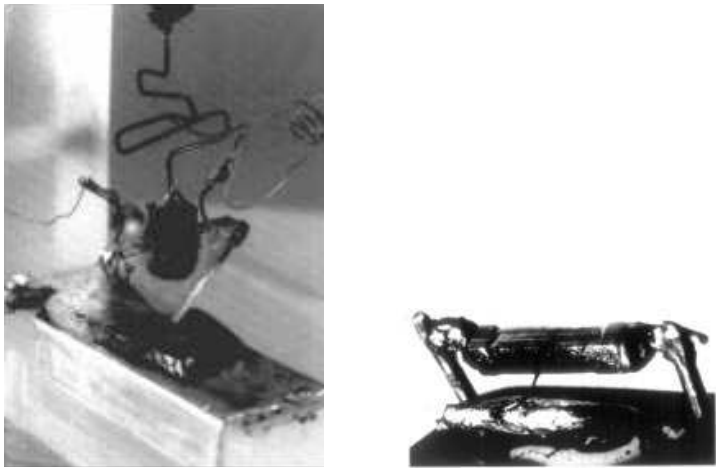
> "*I would like to describe a field, in which little has been done, but in which an enormous amount can be done in principle. [...] Furthermore, a point that is most important is that it would have an enormous number of technical applications. What I want to talk about is the problem of manipulating and controlling things on a small scale. [...] For instance, wires should be 10 or 100 atoms in diameter, and the circuits should be a few thousand angstroms across. [...] There is nothing that I can see in the physical laws that says that the computer elements cannot be made enormously smaller than they are now. In fact, there may be certain advantages...*"
>
> From Richard P. Feynman's talk, "*There is Plenty of Room at the Bottom,*" at California Institute of Technology, December 1959.

The ever continuing journey started with the discovery of a device called a *transresistor*, *i.e.*, transistor, at Bell Laboratories by John Bardeen, Walter Brattain, and William Shockley about a decade before Feynman gave this speech. The discovery of the transistor, similar to other discoveries, was the result of continual efforts throughout the history of engineering and science. The theoretical progress achieved in quantum mechanics during the 1920's played an important role in understanding solid-state electronics [1]. In 1926, Julius E. Lilienfeld conceptually invented a

device whose conductivity would change by applying a voltage, thereby achieving amplification [2], [3], [4]. In 1935, Oskar Heil described the principle of a field effect transistor [5]. Whether Lilienfeld and Heil had ever built such a device, however, is unclear [6].

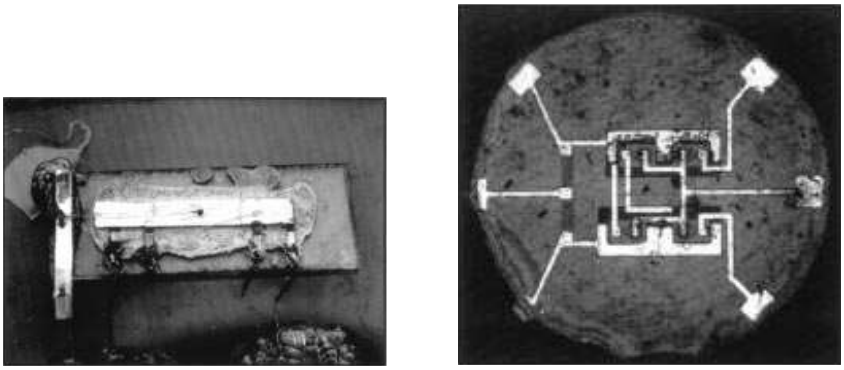
The first realization of the transistor was driven by the need to replace conventional vacuum tubes with a device that would be smaller in size, consume less power, and behave more reliably. In early 1946, Mervin Kelly of Bell Labs created a group including Bardeen, Brattain, and Shockley to work on solid-state devices [1]. By 1947, Bardeen and Brattain created the first operational point-contact transistor, as shown in Fig. 1.1(a) [1], which was announced by Bell Labs in 1948 [7]. The same year, Shockley developed the theory of a bipolar junction transistor [8] which was successfully built in 1950, as shown in Fig. 1.1(b) [1]. Bardeen, Brattain, and Shockley were later awarded the Nobel prize in physics in 1956 for this discovery [1].

In the early 1950s, Bell Labs sold the license of the transistor for $25,000 [1] which brought various companies into play. Jack Kilby of Texas Instruments developed the first integrated circuit (IC) in 1958 using discrete wires as interconnections. The circuit consisted of a transistor, a capacitor, and a resistor, as illustrated in Fig. 1.2(a) [1]. In the mean time, a significant improvement took place at Bell Labs. John Atalla discovered silicon-dioxide as the natural insulator of silicon, allowing a better interface between silicon and silicon-dioxide [9]. Finally, in 1959, Jean Hoerni



(a) (b)

Figure 1.1: Photographs of the first transistor: (a) point-contact transistor by Bardeen and Brattain, (b) bipolar junction transistor by Shockley.



(a) (b)

Figure 1.2: Photographs of the first integrated circuits: (a) built by Texas Instruments in 1958 using discrete wires as interconnections, (b) built by Fairchild Semiconductor using a planar process technology.

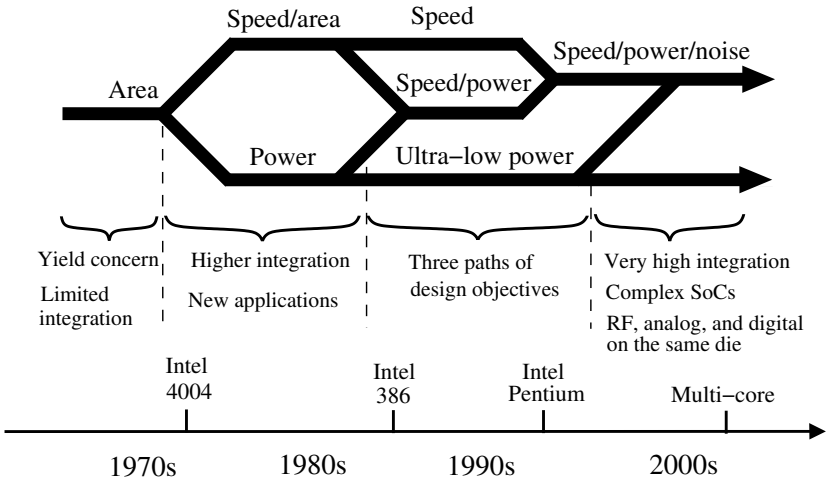


Figure 1.3: Evolution of design objectives for integrated circuits.

of Fairchild Semiconductor introduced the planar process for manufacturing transistors on a substrate [10]. Bob Noyce of Fairchild Semiconductor used this process to build the first monolithic IC in 1959. An early example of an IC built by Fairchild Semiconductor using a planar process is shown in Fig. 1.2(b) [1].

From 1960 to now, microelectronic IC technology has progressed enormously with simultaneous advances in fabrication technology, devices, and design methodologies. Several billions of devices can now be integrated onto the same die, achieving 100,000 times more performance as compared to 1960 [11]. The primary *design objectives* driving this advancement have also evolved during this time frame. This evolution of the design objectives is depicted in Fig. 1.3, as described in [12]. In the 1960's and 1970's, yield was the primary concern due to limited integration density, and consequently, area was the primary design objective. A system contained a large number of interconnected ICs on a printed circuit board (PCB) where the system

speed was primarily determined by the inter-chip communication [12].

In the 1980's, the integration density significantly increased, and the speed bottleneck shifted from the inter-chip communication to the intra-chip communication. Circuit speed became an important design objective. In the meantime, a new class of handheld applications such as calculators and wrist watches emerged, making power consumption another primary parallel design objective [12].

During the 1990's, the design objectives could be categorized under three paths. Speed was the primary objective of the first path where additional power consumption could be tolerated for the sake of higher performance. The second path focused on consumer electronics where high speed was not required, and ultra-low power was the primary design objective. Finally, there were those applications where both speed and power had to be simultaneously considered. Speed and power became the primary design focus of this last design choice [12].

During the 2000's, the integration density has further increased, allowing complex system-on-chips (SoC) where analog, RF, and digital circuits are built on the same die. This highly dense integration of various functionalities, higher clock frequencies (faster signal transition times), and reduced power supply voltages caused another primary design objective, noise, to emerge. Simultaneous optimization of speed, power, and noise has therefore become the primary focus in the IC design process. This co-optimization has enormously complicated the design process due to tradeoffs

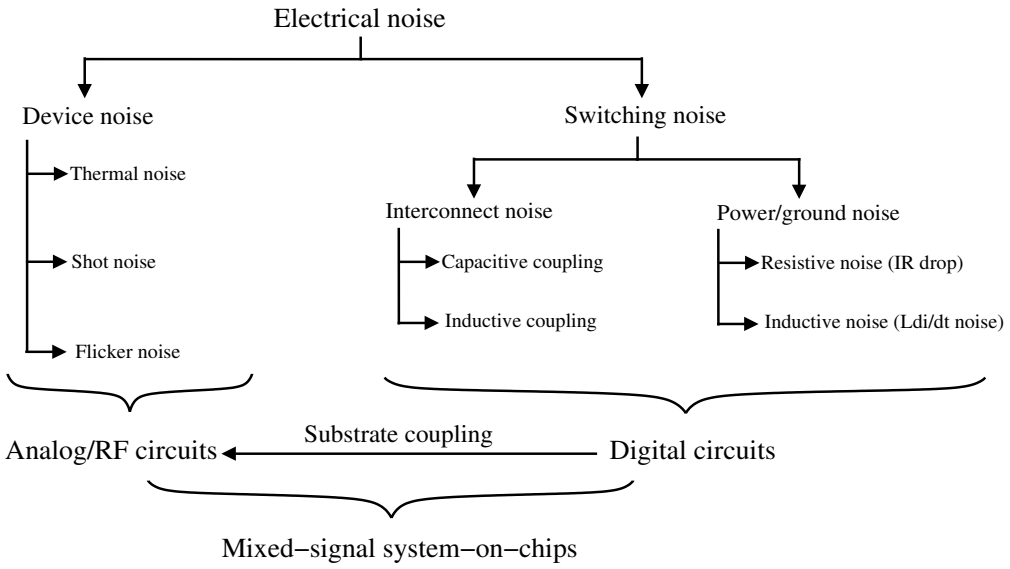


Figure 1.4: Classification of noise in electrical circuits. Device noise is a primary concern for analog/RF circuits. Switching noise originates in digital circuits and affects the analog/RF circuits through substrate coupling.

among these design objectives, requiring new design methodologies.

An overview of noise in electrical circuits is provided in Section 1.1. An outline of the dissertation is presented in Section 1.2.

1.1 Noise in Electrical Circuits

In electrical circuits, noise refers to undesired fluctuations in current or voltage that can directly interfere with the functional signal or indirectly degrade system performance. Noise in electrical circuits can broadly be classified under two primary categories, as shown in Fig. 1.4. The first category, device noise, is the intrinsic noise of the devices while the second category, switching noise, is the induced noise due

to the switching activity of a digital signal. These two types of noise are described, respectively, in Sections 1.1.1 and 1.1.2.

1.1.1 Active and Passive Device Noise

The intrinsic noise of a device can be represented as a random signal, *i.e.*, future values of the signal cannot be predicted from past values [13]. Consequently, device noise is typically analyzed through statistical characteristics such as the average noise power over a long time, power spectrum density, and probability density function [13]. Three primary types of device noise exist: (1) thermal noise, (2) shot noise, and (3) flicker noise. These noise types are briefly described in the following subsections.

Thermal Noise

Thermal noise, also known as Johnson or Nyquist noise, occurs in a conductor due to the random motion of the charge carriers. This random motion is generated by the thermal agitation of the electrons that are in "thermal equilibrium with the molecules" [14]. The power spectral density $S_v(f)$ (V^2/Hz) of thermal noise is therefore proportional to the temperature and independent of frequency, *i.e.*, the mean noise power per unit bandwidth is constant at all frequencies,

$$S_v(f) = 4k_bTR, \tag{1.1}$$

where k_b is the Boltzmann's constant, T is the absolute temperature, and R is the resistor of the conductor. This type of noise source is called *white* noise due to the quality that the noise power is constant with frequency [14]. Also note that thermal noise is independent of any applied voltage.

An MOS transistor also generates thermal noise due to the channel resistance. This noise is typically modeled as a current source I_n between the source and drain nodes with a power spectral density,

$$S_i(f) = 4k_bT\gamma g_m, \tag{1.2}$$

where g_m is the transconductance of the transistor and the coefficient γ is approximately 2/3 for long channel devices and increases as the channel length decreases [13]. The resistive gate terminal of an MOS transistor also produces thermal noise. Additional contacts and gate folding are two layout techniques to reduce thermal noise [13].

Shot Noise

In electronic devices, shot noise refers to fluctuations in the current due to the discrete nature of charges. As opposed to thermal noise, shot noise is dependent upon the mean current flow or bias in a semiconductor [15]. The power spectral density $S_i(f)$ (A^2/Hz) of shot noise is

$$S_i(f) = 2qI, \tag{1.3}$$

where q is the magnitude of an electronic charge and I is the DC bias current. Shot noise is typically a dominant noise source in diodes and modeled with a parallel current source [16].

Flicker Noise

Flicker noise, also known as $1/f$ noise, occurs in semiconductors due to various reasons such as energy traps resulting in generation and recombination of charge carriers [15]. The power spectral density $S_i(\omega)$ of flicker noise has an inverse frequency dependence [15],

$$S_i(\omega) = K\frac{I^\gamma}{\omega^\alpha}, \tag{1.4}$$

where K is a constant depending upon the device characteristics, I is the DC current, and ω is the angular frequency. Note that γ and α are exponents with typical values of, respectively, between 1 and 2, and between 0.8 and 1.2 [15].

The flicker noise of an MOS transistor is modeled as a voltage source at the gate terminal with a power spectral density $S_v(f)$ [16],

$$S_v(f) = \frac{K}{WLC_{ox}f}, \tag{1.5}$$

where W, L, and C_{ox} represent, respectively, the width of the transistor, channel length, and gate capacitance per unit area.

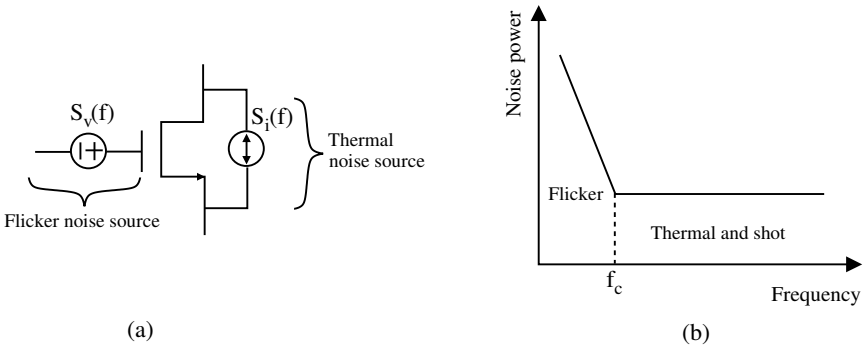


Figure 1.5: Thermal and flicker noise in an NMOS transistor: (a) Thermal noise is modeled as a current source between the drain and source, and flicker noise is modeled as a voltage source at the gate of the transistor, (b) variation of noise power as a function of frequency.

Note that flicker noise and thermal noise are the two dominant noise sources in MOS transistors. An NMOS transistor with these noise sources is illustrated in Fig. 1.5(a). The overall frequency dependence of the noise power is depicted in Fig. 1.5(b). Flicker noise dominates at lower frequencies, and thermal noise (and partly shot noise) determine the noise floor at higher frequencies. The crossover frequency f_c is called the corner frequency, describing the frequency band of flicker noise [13].

1.1.2 Switching Noise

Switching noise is caused by a logic transition, *i.e.*, high-to-low or low-to-high, of a digital signal. The switching noise is typically two to three orders greater than the device noise [17]. In mixed-signal circuits, therefore, switching noise coupling to the analog/RF circuits is critical.

As opposed to device noise, switching noise is not strictly a random process since this noise is dependent upon the switching activity of the circuit. Practically, however, switching noise can be treated as a random phenomenon due to several reasons [17]. For example, a large number of nodes switches at different locations across the die where the cumulative noise behaves as random. Furthermore, the time constant of each node (the time required for the noise to dissipate through the ground or power network) is different [17], contributing to the randomness of the noise. Note however that if multiple dominant noise sources are synchronized, the overall noise may not be random [17].

Two types of switching noise exist: (1) interconnect noise or crosstalk and (2) power/ground noise, as shown in Fig. 1.4. Both crosstalk and power/ground noise have deleterious effects on the operation and performance of a circuit, as described in Chapters 2 and 3.

1.2 Dissertation Outline

Switching noise in synchronous digital integrated circuits is the primary focus of Chapter 2. Two types of switching noise, *i.e.*, power/ground noise and interconnect noise, are introduced. The impedance characteristics of a power distribution network are analyzed. Existing approaches to model interconnect noise are summarized. The

effects of switching noise on digital circuits, *i.e.*, functional failure, glitch power consumption, and delay uncertainty, are discussed. Finally, existing design techniques to reduce power/ground noise and interconnect noise are described.

Switching noise in mixed-signal/analog integrated circuits is described in Chapter 3. The switching noise coupling mechanisms within analog/RF circuits are explained with a primary focus on substrate coupling. The noise injection mechanisms into the substrate are described. Substrate modeling and extraction techniques, and circuit level substrate noise analysis methodologies are summarized. The effects of substrate noise on analog/RF circuits are discussed with specific examples of several commonly used circuit blocks. Existing noise isolation methodologies to reduce substrate noise coupling are also reviewed.

A time domain analysis of power/ground noise is described in Chapter 4. Specifically, the non-monotonic dependence of switching noise on rise time is presented. An approximate solution for the worst case rise time producing the maximum power and ground noise is developed. The effect of the parasitic resistance of the power/ground network on the worst case rise time and noise is also investigated. The efficiency of several noise reduction techniques are compared through a sensitivity analysis.

The effect of the substrate network on ground noise is the focus of Chapter 5. Specifically, current propagation along the ground network and through the substrate is affected by the substrate contacts. The interaction of these two networks is

investigated. The contribution of the substrate to the total ground noise is shown to be significant.

The power/ground noise model developed in Chapter 4 is extended in Chapter 6 to generate a macromodel for substrate coupling. The macromodel is used to identify the dominant substrate noise coupling mechanism for multiple switching gates at early stages of the design process. The regions illustrating the dominant noise coupling mechanisms are generated as a function of multiple parameters. Several noise reduction techniques are also proposed depending upon the particular noise mechanism that is dominant.

In Chapter 7, a high-level substrate noise analysis methodology is proposed for large scale mixed-signal circuits. Rather than homogeneous extraction of the entire substrate, a fine extraction is applied to those regions where the dominant current flow occurs, while a coarse extraction is achieved for the remaining regions. This heterogeneous extraction of the substrate significantly improves the computational complexity of the extraction process. An algorithm with linear time complexity is proposed to identify these regions within the substrate. The proposed methodology achieves a significant reduction in the number of extracted substrate resistors while maintaining sufficient accuracy.

Vertical current propagation within a substrate is typically not desired since the vertical current paths reduce the efficiency of the substrate contacts and guard rings.

A methodology is described in Chapter 8 to reduce the noise injected into the substrate by minimizing vertical current propagation within the substrate. This reduction is achieved by designing noise-aware standard cells. Limitations of the existing substrate biasing schemes are removed while achieving more than a 60% reduction in the peak-to-peak substrate noise.

Switching noise affects the timing characteristics of synchronous digital integrated circuits by introducing additional delay uncertainty. The interdependence of timing constraints, *i.e.*, setup and hold times, is described in Chapter 9. An efficient algorithm is proposed to reduce significant pessimism in static timing analysis by exploiting these interdependent setup and hold times. Industrial circuits are used to validate the proposed methodology.

Interdependent setup and hold times of a register are exploited in Chapter 10 to reduce delay uncertainty. Specifically, the increase or decrease in the data path delay due to power supply noise is compensated by exploiting the interdependence relationship, thereby producing a more robust circuit. An algorithm is proposed to determine the appropriate (setup, hold) pair to improve the tolerance of a circuit to power supply noise while increasing the maximum operating frequency.

The research presented in this dissertation is concluded in Chapter 11, summarizing the significance of the results. Directions for future study are proposed in Chapter 12. Specifically, a unified noise analysis methodology is described to investigate

the tradeoffs among the integrated circuit, package, and board level characteristics. The codesign of the power grid, clock network, and critical data paths is also proposed the exploit the interactions among these networks. Finally, the effect of substrate isolation techniques on power integrity is presented as a future research topic to improve the signal integrity and timing characteristics of a circuit.

Chapter 2

Switching Noise in Synchronous Digital Integrated Circuits

The continuous progress in complementary metal-oxide-semiconductor (CMOS) technology has enabled the integration of more than two billion transistors on a single integrated circuit. The density of the devices within a circuit has dramatically increased, making the related issues of *noise* and *noise coupling* among various circuit elements a primary concern for nanoscale integrated circuits.

Traditionally, digital circuits consisting of static CMOS gates have been treated as inherently immune to noise due to relatively high noise margins [18], [19]. Consequently, the concept of *noise* in an integrated circuit has long been associated with analog, RF, and dynamic CMOS circuits due to the higher sensitivity of these types of circuits [20]. This situation, however, has dramatically changed in the last decade. Three fundamental reasons have stimulated this change:

- Noise margins have been reduced significantly due to scaling of the power supply

and threshold voltages, making digital logic circuits more sensitive to noise.

- The operating frequencies have substantially increased, placing more stringent constraints on the timing requirements of the critical paths. The sensitivity of the circuit delay to noise has therefore become more significant. Furthermore, higher clock rates have enabled faster on-chip signal transitions, exacerbating the magnitude of the noise coupling.

- The delay of the interconnects has become comparable or greater than the delay of the logic gates in deep submicrometer technologies. As such, the impact of interconnect noise on the signal characteristics, *i.e.*, *signal integrity*, and on system performance has become significant.

As a result of these three reasons, noise in digital integrated circuits has started to receive considerable attention.

The simultaneous operation of an enormous number of devices at an aggressive clock frequency with reduced physical distances among the interconnects cause two primary types of *switching noise* in a synchronous digital integrated circuit: power/ground noise and interconnect noise. Power and ground noise refers to voltage fluctuations, respectively, on the power and ground distribution networks. The simultaneous switching of a large number of logic gates requires a significant amount of current drawn from the power supply. This current flows through the parasitic

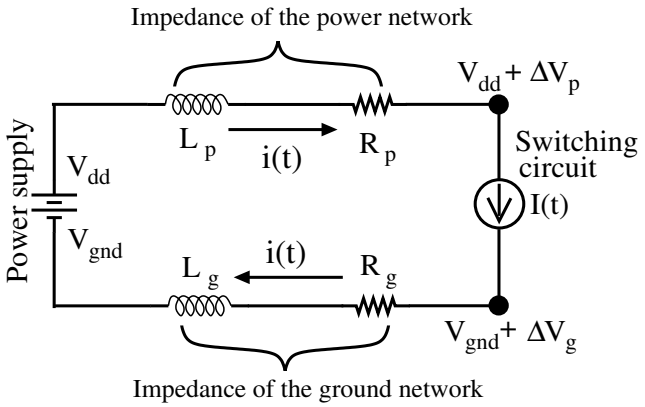


Figure 2.1: Power and ground noise due to the switching current flowing through the power and ground network impedances.

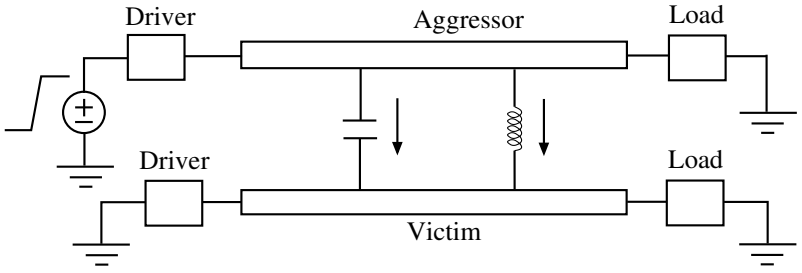


Figure 2.2: Interconnect noise due to capacitive and inductive coupling from a switching aggressor node to a victim node.

impedance of the power distribution network, causing both static and dynamic voltage fluctuations [21], [22]. Similarly, the current flowing from the switching logic gates to the reference ground of the power supply causes voltage fluctuations on the ground distribution network, referred to as ground noise or ground bounce [23]. Noise generation on the power and ground nodes is illustrated in Fig. 2.1.

The interconnect noise or *crosstalk* refers to a voltage induced on a *victim* node due to capacitive and/or inductive coupling from a switching *aggressor* node, as depicted in Fig. 2.2. Electromagnetic coupling has long been a primary concern for

radio frequency (RF) circuits, microwave circuits, and high speed printed circuit (PC) boards [24]. The effects of on-chip crosstalk in high speed digital integrated circuits have started to become significant in the last decade due to the asymmetric scaling between the vertical and lateral dimensions, and faster signal transitions in deep submicrometer technologies. Specifically, the lateral dimensions have been scaled to enhance performance and achieve higher density, while the vertical dimensions have not changed considerably [25]. Coupling among interconnects in a digital integrated circuit has therefore increased significantly. Similarly, the speed of on-chip signal transitions and length of interconnects have increased, exacerbating on-chip inductive effects [26].

A brief overview of switching noise in synchronous digital integrated circuits is provided in this chapter. In Section 2.1, the problem of power distribution and power/ground noise is described. The use of decoupling capacitances to lower the power noise of a power distribution network is also discussed. Capacitive and inductive crosstalk among on-chip interconnects is reviewed in Section 2.2. Existing models to efficiently analyze coupling noise are summarized. The deleterious effects of switching noise on digital integrated circuits, such as logical failure, glitch power consumption, and greater delay uncertainty, are presented in Section 2.3. Existing design techniques to reduce switching noise are discussed in Section 2.4. Finally, the chapter is summarized in Section 2.5.

2.1 Power/Ground Noise

Reliable distribution of power and ground voltages in a high performance integrated circuit is a challenging task due to the high current demands and reduced operating voltages [12]. The issue of power delivery is described in Section 2.1.1. The impedance characteristics of a power distribution system is reviewed in Section 2.1.2. The effects of a decoupling capacitance on the noise and impedance characteristics are investigated in Section 2.1.3.

2.1.1 Power and Ground Distribution Networks

Power and ground voltages should be reliably distributed to satisfy functionality and performance while considering the overall resources devoted to these distribution networks. A power distribution system consists of the power supply; interconnect networks at multiple levels beginning with the voltage regulator, including the printed circuit board and package, and the integrated circuit; decoupling capacitances at these multiple levels; vias connecting different layers of conductors; and the power and ground pads [12]. A simplified circuit representation of this system is illustrated in Fig. 2.3. Each level of a conductor is modeled as a series RL circuit where R is the parasitic resistance and L is the parasitic inductance of the interconnect. The decoupling capacitances are represented as C. The superscripts p and g refer to the power and ground conductors, respectively. The superscript C refers to the parasitic

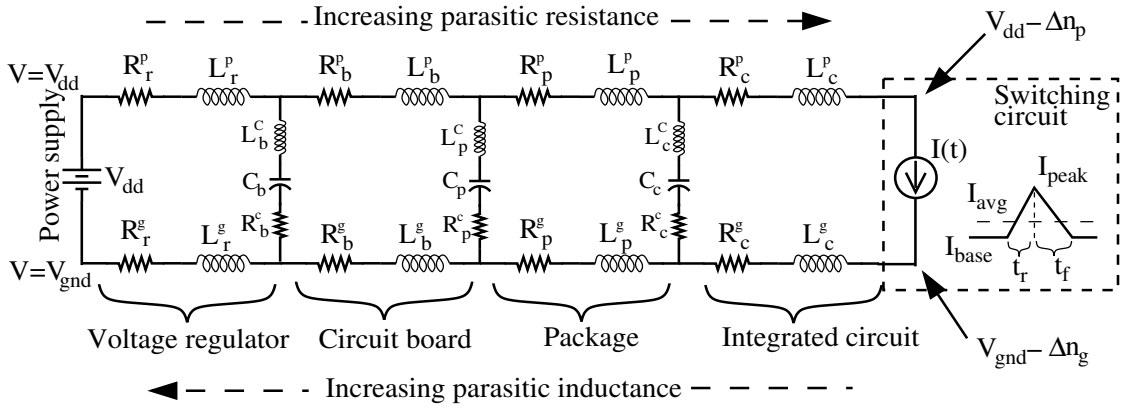


Figure 2.3: Simplified circuit level representation of a power distribution system.

impedance characteristics of the decoupling capacitance. The conductors at various levels are represented by the subscripts where r, b, p, and c refer, respectively, to the voltage regulator, printed circuit board, package, and integrated circuit. The switching circuit is represented by a linear current source that mimics the current characteristics at the load.

The parasitic resistance typically increases from the power supply to the integrated circuit due to the increasing aspect ratio of the conductors. Alternatively, the parasitic inductance typically decreases in the same direction due to the increasing interconnect density, as illustrated in Fig 2.4. Note however that this behavior has changed in recent years due to advances in packaging technology, resulting in lower parasitic inductances at the package level (such as a flip-chip technology), and faster on-chip signal transitions, all causing the on-chip wires to exhibit significant inductive behavior [27]. That is, the on-chip parasitic inductance can be comparable to

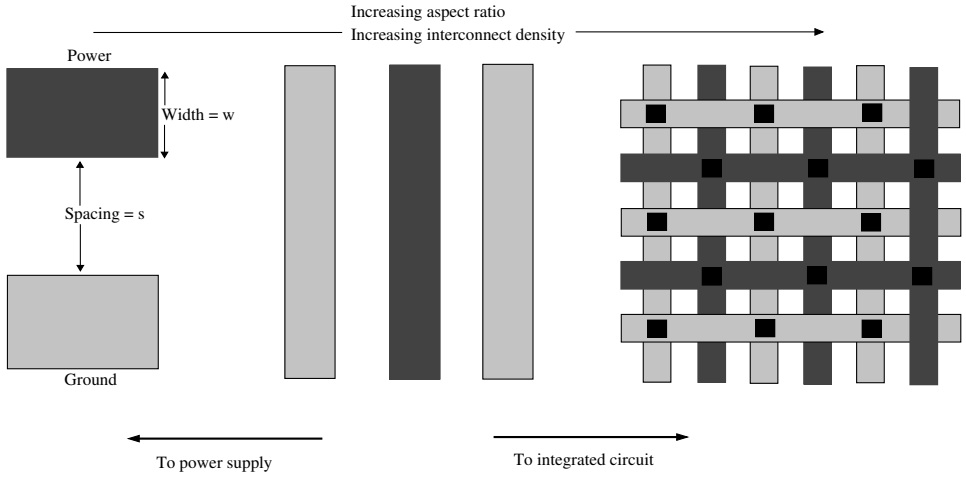


Figure 2.4: Change in the physical characteristics of the power/ground interconnects from the power supply to the integrated circuit.

or greater than the package inductance depending upon the type of package, on-chip signal characteristics, and on-chip power distribution network.

The available range of voltage, referred to here as the *voltage headroom*, of the switching circuits is reduced due to the parasitic impedance of the power and ground distribution networks. Specifically, the voltage headroom of the power supply $V_{dd} - V_{gnd}$ is reduced to $V_{dd} - V_{gnd} - (\Delta p + \Delta g)$ where Δp and Δg represent, respectively, the power and ground noise, as illustrated in Fig. 2.3. Assuming the decoupling capacitances are negligible, the power noise is

$$\Delta p = I(t)R^p_{tot} + L^p_{tot}\frac{\partial i}{\partial t}. \tag{2.1}$$

Note that the ground noise is similarly determined by replacing R^p_{tot} and L^p_{tot}, respectively, by R^g_{tot} and L^g_{tot}. The first term in (2.1), commonly referred to as the IR drop, is proportional to the magnitude of the current. The second term, referred to as $L\ \partial i/\partial t$ noise, is proportional to the *rate of change* in the current. Consequently, the total peak noise does not necessarily occur when the IR drop or $L\ \partial i/\partial t$ noise is individually the greatest. These two types of noise should therefore be analyzed simultaneously to avoid pessimism when estimating the peak noise. Typically, two design metrics should be satisfied when characterizing the power noise in the time domain [28].

- The maximum average noise within a clock cycle. This metric determines the amount of variation in the delay of a critical path. That is, the maximum operating frequency of the circuit under power supply variations is dependent upon this parameter.

- The maximum peak noise that can be tolerated to ensure accurate functionality of the circuit, as determined by the noise margins. A power noise greater than the maximum noise can change the logic state of a node, causing a circuit to fail.

The time domain representation of the peak and average power noise, and a related clock waveform are illustrated in Fig. 2.5 for a synchronous digital circuit. A significant number of registers switches simultaneously with the rising edge of the clock,

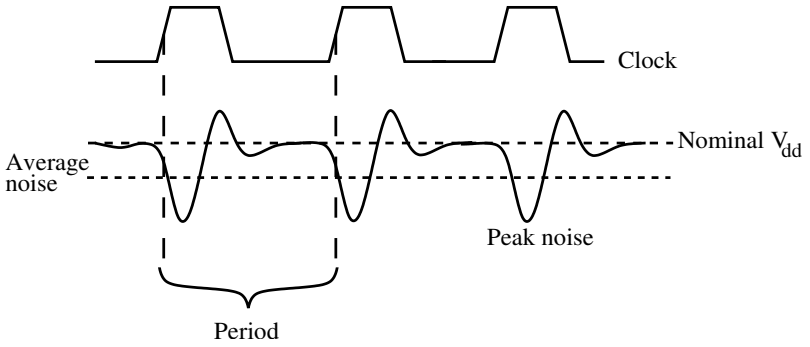


Figure 2.5: Clock waveform and the corresponding peak and average power noise in the time domain.

causing a considerable drop in the power supply voltage during this transition time.

A power distribution system can also be analyzed in the frequency domain with a *target impedance* over a specific range of operating frequency [29]. The impedance characteristics of a power distribution system is described in the following section.

2.1.2 Impedance Characteristics of a Power Distribution Network

The target impedance of a power distribution network is determined from the nominal power supply voltage, maximum tolerable noise on the load circuit as a per cent of the nominal power supply voltage, and the current sourced by the load circuit from the power distribution network [30],

$$Z_{target} = \frac{Nominal\ voltage * Tolerable\ noise\ (\%)}{Current}. \tag{2.2}$$

The target impedance should be satisfied over a wide range of frequency where the circuit operates. Note that this frequency range is determined by the power spectral density of the current transients rather than the clock frequency, where significant high frequency components exist due to the fast switching events. The average current is estimated from the power consumption of the circuit at a specific clock frequency,

$$I = \frac{Power\ consumption}{Nominal\ voltage}. \tag{2.3}$$

Replacing (2.3) in (2.2) permits the target impedance to be described as a function of the power and nominal voltage,

$$Z_{target} = \frac{Nominal\ voltage^2 * Tolerable\ noise\ (\%)}{Power\ consumption}. \tag{2.4}$$

As determined by (2.4), the target impedance is reduced quadratically as the nominal voltage is scaled. Furthermore, the power consumption increases due to the higher integration densities, requiring further reductions in the target impedance. A recent 16 core microprocessor, implemented in 65 nm CMOS technology, consumes 250 W at a supply voltage of 1.2 V [31]. According to (2.4), the impedance of the power distribution network for this microprocessor should be less than 0.3 mΩ over a wide frequency range. The target impedance should therefore be aggressively decreased to satisfy power/ground noise constraints in deep submicrometer technologies.

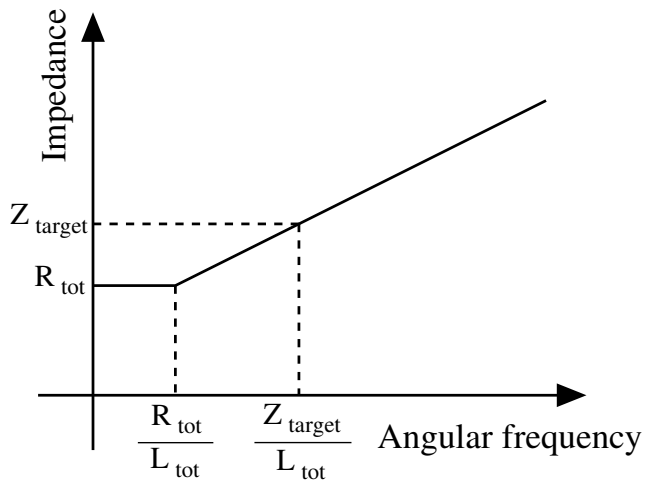


Figure 2.6: Impedance characteristics of a power distribution network with no decoupling capacitance.

The magnitude of the output impedance of a power distribution system with no decoupling capacitances is $|R_{tot} + j\omega L_{tot}|$. The dependence of this impedance on frequency is illustrated in Fig. 2.6. The impedance is dominated by the total resistance R_{tot} for frequencies $\omega < R_{tot}/L_{tot}$. R_{tot} should therefore be smaller than the target impedance to satisfy specific noise constraints at low frequencies. Alternatively, the total inductance L_{tot} starts to dominate the impedance for frequencies $\omega > R_{tot}/L_{tot}$. At $\omega = Z_{target}/L_{tot}$, the impedance exceeds the target impedance, as illustrated in Fig. 2.6. The maximum frequency of a power distribution network with no decoupling capacitances is therefore determined by $\omega = Z_{target}/L_{tot}$.

Decoupling capacitances have traditionally been used at multiple levels of the power distribution network to increase the frequency at which the output impedance exceeds the target impedance [32], [33]. The effect of the decoupling capacitance on

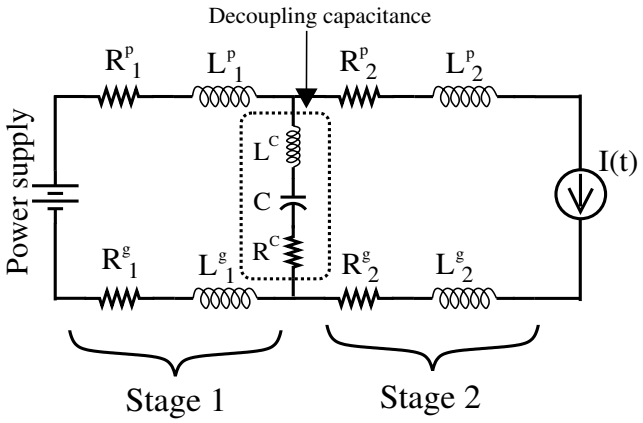


Figure 2.7: A circuit model of a power distribution system with a decoupling capacitance. Stage 1 represents the impedance between the decoupling capacitance and the power supply. Stage 2 represents the impedance between the decoupling capacitance and the switching circuit.

the power noise and impedance characteristics is described in the following section.

2.1.3 Decoupling Capacitance

A decoupling capacitance within a power distribution network refers to an intentional or intrinsic capacitance placed across power and ground conductors to reduce power supply noise by temporarily providing charge to load circuits during switching events. A simplified circuit representation of a power distribution system with a decoupling capacitance is illustrated in Fig. 2.7. The decoupling capacitance is represented by a capacitance C in series with an effective series resistance (ESR) R^C and effective series inductance (ESL) L^C. Note that these parasitic elements of a decoupling capacitance play an important role in determining the efficiency of the capacitance [34].

A decoupling capacitance divides the power distribution system into two stages, as illustrated in Fig. 2.7. Stage 1 represents the impedance between the decoupling capacitance and the power supply. Stage 2 represents the impedance between the decoupling capacitance and the switching circuit.

The amount of switching current provided by the decoupling capacitance and the power supply is dependent upon the frequency (or more precisely, the rise time) of the current transients. This current distribution as a function of frequency is illustrated in Fig. 2.8. At frequencies sufficiently lower than the resonant frequency $\omega_{res} = 1/(LC)$, the power supply provides the overall switching current due to the high impedance of the decoupling capacitance. Alternatively, at frequencies higher than the resonant frequency, the switching current is provided by the decoupling capacitance, bypassing the parasitic inductances L_1^p and L_1^g and resistances R_1^p and R_2^g. The power noise decreases at higher frequencies due to the reduced overall parasitic impedance. At the resonant frequency $\omega \approx 1/[C(L_1^p + L_1^g)]$, both the power supply and decoupling capacitance provide significant current, giving rise to a resonant behavior [12]. The impedance of the power distribution network is greatest at the resonant frequency, causing the power noise to increase at that frequency.

The effect of a decoupling capacitance on the impedance characteristics of a power distribution network is illustrated in Fig. 2.9. The decoupling capacitance increases the maximum frequency from Z_{target}/L_{tot} to $Z_{target}/(L_2^p + L_2^g + L^C)$ where $L_2^p + L_2^g + L^C$

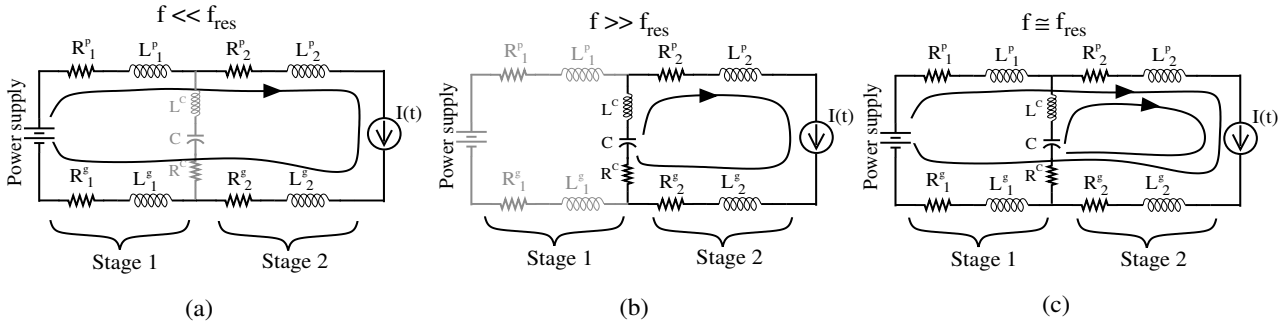


Figure 2.8: The switching current distribution as a function of frequency in a power distribution network with a decoupling capacitance: (a) At frequencies $f << f_{res}$, most of the current is provided by the power supply. The decoupling capacitance does not affect the power noise. (b) At frequencies $f >> f_{res}$, most of the current is provided by the decoupling capacitance, bypassing the parasitic impedance of stage 1. The decoupling capacitance reduces the power noise. (c) At frequencies close to the resonant frequency, both the power supply and decoupling capacitance provide current, giving rise to resonant behavior. The power noise increases due to the high impedance of the power distribution network.

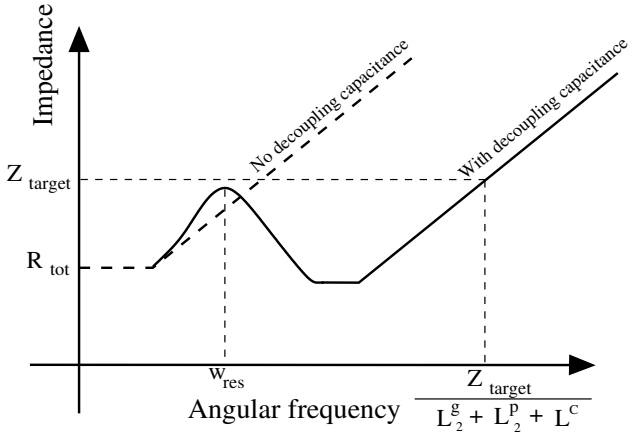


Figure 2.9: Impedance characteristics of a power distribution network with a decoupling capacitance.

is the parasitic inductance of stage 2. Note, however, that at the resonant frequency $\omega \approx 1/[C(L_1^p + L_1^g)]$, the total impedance with a decoupling capacitance is larger than the impedance without a decoupling capacitance due to the parallel combination of the capacitance and inductance of stage 1, forming an LC tank circuit. Specifically, the impedance is increased by the quality factor Q (or inverse damping factor) of this tank circuit,

$$Q = 1/\zeta = [\frac{1}{(R_g^1 + R_p^1 + R^C)}\sqrt{\frac{(L_g^1 + L_p^1)}{C}}]. \tag{2.5}$$

Special consideration should therefore be given to satisfying the target impedance of a power distribution network with a decoupling capacitance at the resonant frequency. A time domain analysis of resonant behavior is provided in Chapter 4 where an equivalent rise time that corresponds to resonant frequency is presented.

As illustrated in Fig. 2.9, the maximum frequency over which the target impedance is not exceeded has been increased with the use of a decoupling capacitance. A single decoupling capacitance however is not sufficient to satisfy the target impedance over a wide range of frequency due to the following reasons.

- The decoupling capacitance has to be sufficiently large to store sufficient charge for the current transients and to achieve a sufficiently high damping factor to ensure that the impedance at the resonant frequency is smaller than the target impedance.

- The decoupling capacitance has to be sufficiently close to the load circuit to minimize the impedance of stage 2 since the maximum frequency where the decoupling capacitance is effective is determined by this impedance.

- The first two conditions cannot be simultaneously achieved due to physical area constraints within the integrated circuit and the distributed nature of a large decoupling capacitance [35].

Consequently, each level of the power distribution network is hierarchically composed of distributed decoupling capacitances where the board decoupling capacitance is the largest and the on-chip decoupling capacitance is the smallest. The board level decoupling capacitance provides charge when the response time of the power supply is insufficient for fast current transients. Similarly, the on-chip decoupling capacitance provides charge when the current transients are faster than the response time of the package level decoupling capacitance. This *hierarchical* placement of decoupling capacitors can be designed to satisfy the target impedance over a broader range of frequencies.

2.2 Interconnect Coupling Noise: Crosstalk

Crosstalk (or coupling) among various conductors is a major concern in high performance digital integrated circuits due to higher integration densities, faster on-chip signal transitions, and the increasing significance of interconnects on system performance and reliability [36], [37]. An overview of interconnect coupling noise is provided in this section. The scaling characteristics of on-chip interconnects are reviewed in Section 2.2.1. Capacitive and inductive coupling are discussed as two separate coupling mechanisms, respectively, in Sections 2.2.2 and 2.2.3.

2.2.1 Ideal vs. Realistic Interconnect Scaling

Improvements in CMOS process technology are based on not only scaling the devices, but also scaling the interconnect. The dimensions of two parallel conductors are illustrated in Fig. 2.10. L_{int}, W_{int}, and T_{int} refer, respectively, to the length, width, and thickness of the conductor. W_{sp} refers to the spacing between two parallel conductors located on the same metalization layer and H is the distance between the conductor and the substrate or the thickness of the dielectric between two adjacent metal layers. The aspect ratio of an interconnect is $AR = T_{int}/W_{int}$.

Based on ideal scaling, each dimension of an interconnect scales by the same scaling factor S ($S > 1$) except the length of the global interconnects [38], [39]. Scaling the length of the line is dependent upon the type of interconnect, *i.e.*, local vs.

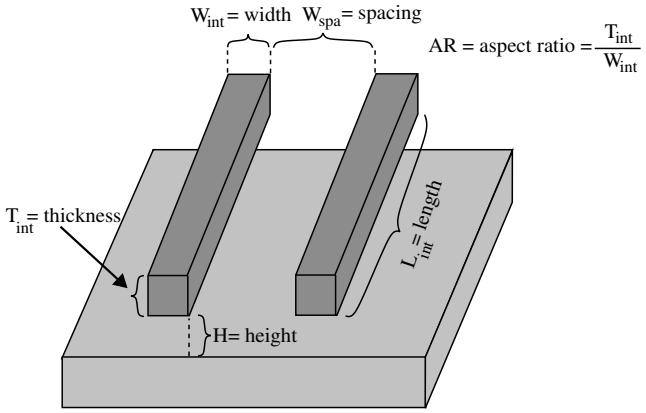


Figure 2.10: Physical dimensions of two parallel conductors located on the same metalization layer.

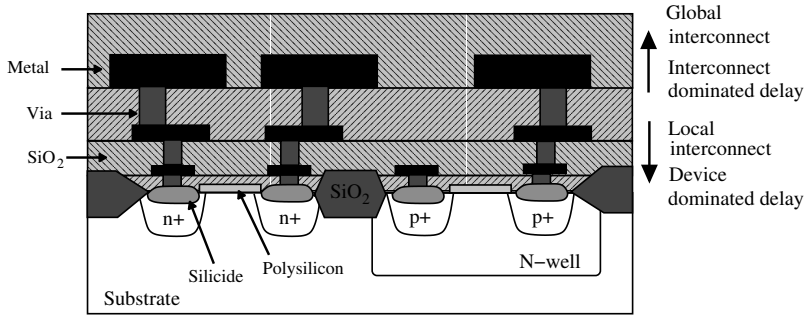


Figure 2.11: Local versus global interconnects. Global interconnects are typically located on the upper metal layers while the local interconnects are closer to the devices.

global, as illustrated in Fig. 2.11. The local interconnects typically transmit signals at the device level and within a gate or block in a circuit. The gate delays typically dominate at this level. Alternatively, the global interconnects typically transmit signals across the integrated circuit. Examples of global interconnect include clock and power/ground distribution networks and global signal lines. The interconnect delay dominates the gate delay in these global interconnects. Due to an increase in the overall chip area, the length of the global interconnects typically grows by a factor

Table 2.1: Scaling characteristics of local and global interconnects for four different scaling scenarios: ideal, quasi-ideal, constant resistance, and constant thickness. S is the scaling factor for the local interconnects and S_c is the scaling factor for the die size.

Interconnect parameters	Ideal scaling		Quasi-ideal scaling		Constant resistance		Constant thickness	
	Local	Global	Local	Global	Local	Global	Local	Global
Length (L_{int})	$1/S$	S_c	$1/S$	S_c	$1/S$	S_c	$1/S$	S_c
Width (W_{int})	$1/S$	$1/S$	$1/S$	$1/S$	$1/\sqrt{S}$	$1/\sqrt{S}$	$1/S$	$1/S$
Thickness (T_{int})	$1/S$	$1/S$	$1/\sqrt{S}$	$1/\sqrt{S}$	$1/\sqrt{S}$	$S_c\sqrt{S}$	1	1
Height (H)	$1/S$	$1/S$	$1/\sqrt{S}$	$1/\sqrt{S}$	$1/\sqrt{S}$	$1/\sqrt{S}$	$1/S$	$1/S$
Spacing (W_{spa})	$1/S$	$1/S$	$1/S$	$1/S$	$1/\sqrt{S}$	$1/\sqrt{S}$	$1/S$	$1/S$
Aspect ratio ($AR = T_{int}/W_{int}$)	1	1	$S/\sqrt{S}$	$S/\sqrt{S}$	1	S_cS	S	S
Resistance ($R_{int} = (\rho L_{int})/(W_{int}T_{int})$)	S	S_cS^2	$\sqrt{S}$	$S_cS\sqrt{S}$	1	1	1	SS_c
Coupling capacitance ($C_c = (\epsilon L_{int}T_{int})/W_{spa}$)	$1/S$	S_c	$1/\sqrt{S}$	$S_c\sqrt{S}$	$1/S$	$S^2S_c^2$	1	S_cS
Ground capacitance ($C_g = (\epsilon L_{int}W_{int})/H$)	$1/S$	S_c	$1/(S\sqrt{S})$	$S_c/\sqrt{S}$	$1/S$	S_c	$1/S$	S_c
C_c/C_g	1	1	S	S	1	S_cS^2	S	S

S_c, referred to as the chip scale factor, as described in [38] and [39].

The scaling characteristics of the local and global interconnect dimensions are listed in Table 2.1 for four different cases: ideal, quasi-ideal, constant resistance, and constant thickness. In ideal scaling, each dimension of the interconnect is scaled by the identical scaling factor S. Quasi-ideal scaling is developed to represent a more realistic scenario [39]. According to constant resistance scaling, local and global interconnect resistances remain the same. Finally, the constant thickness scaling assumes that the thickness of the interconnects remains the same with future technologies.

The total resistance of a line is determined by $R = (\rho L_{int})/A$ where ρ is the resistivity of the material and $A = W_{int}T_{int}$ is the cross sectional area of the interconnect. According to ideal scaling, the line resistance increases by S for local interconnects

and S^2S_c for global interconnects. This dramatic increase in the resistance, particularly the global interconnects, causes non-ideal scaling to differ from ideal scaling. Based on quasi-ideal scaling, vertical dimensions such as H and T_{int} are scaled by a factor $\sqrt{S}$ rather than S. Consequently, the resistance of the local and global interconnects scales, respectively, by $\sqrt{S}$ and $S_cS\sqrt{S}$. For the case of constant resistance, the width also scales by $\sqrt{S}$. Note that for global interconnects, the thickness should be *increased* by a factor of $S_c\sqrt{S}$ to maintain a constant resistance. Finally, for the case of constant thickness, the resistance of the local and global interconnects scales, respectively, by one (1) and S_cS.

Note that the aspect ratio of the global interconnects increases in all cases other than ideal scaling, as listed in Table 2.1. The implications of this increase of the aspect ratio on the interconnect capacitance and capacitive coupling are discussed in the following section.

2.2.2 Capacitive Coupling

The coupling capacitance among the on-chip interconnects and the dependence of this capacitance on the switching characteristics of the input signals is described in this section. The noise due to the coupling capacitance and some of the primary issues in modeling this noise are also discussed.

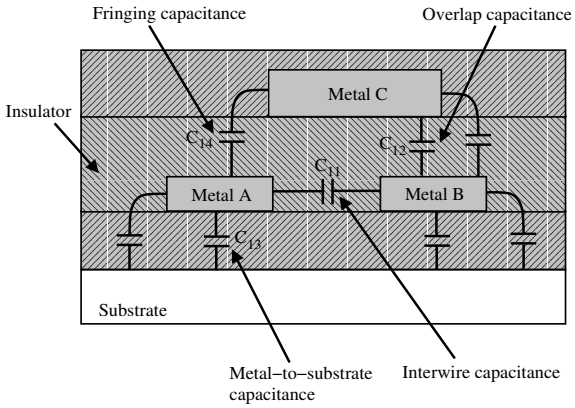


Figure 2.12: Capacitive components among three metal wires.

Scaling Characteristics of Coupling Capacitance

The different capacitive components among the three metal wires are illustrated in Fig. 2.12. Metal lines A and B are located on the same layer. Metal line C is located on an upper layer adjacent to metal lines A and B. The sidewall capacitance C_{11} between metal lines A and B, referred to as interwire or coupling capacitance, is,

$$C_{11} = \frac{\epsilon L_{int} T_{int}}{W_{spa}}, \tag{2.6}$$

where ϵ is the dielectric constant of the insulator, and L_{int}, T_{int}, and W_{spa} are defined in Fig. 2.10. The parallel plate capacitance C_{13} between a metal line and the substrate, referred to as metal-to-substrate or ground capacitance, is

$$C_{13} = \frac{\epsilon L_{int} W_{int}}{H}, \tag{2.7}$$

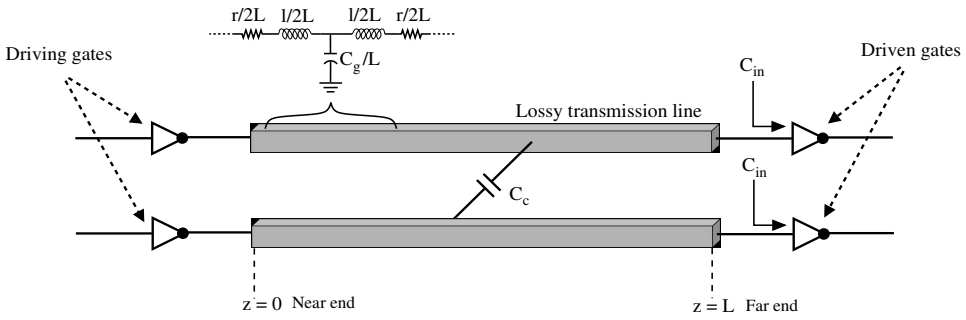


Figure 2.13: Capacitively coupled interconnects driven by CMOS inverters. The interconnects are represented as lossy transmission lines.

where H is defined in Fig. 2.10. The ratio of coupling capacitance C_{11} to the ground capacitance C_{13} is

$$\frac{C_{11}}{C_{13}} = \frac{T_{int}H}{W_{int}W_{spa}} = AR\frac{H}{W_{spa}}. \tag{2.8}$$

As listed in Table 2.1, for global interconnects, this ratio increases by a factor of S for quasi-ideal scaling and by a factor of S_cS^2 for constant resistance, making capacitive coupling a significant signal integrity issue in nanoscale integrated circuits [40].

Dependence of Coupling Capacitance on Switching Characteristics

Two capacitively coupled global interconnects, driven by CMOS inverters, are illustrated in Fig. 2.13. The global interconnects, represented as lossy transmission lines, are modeled as a distributed RLC impedance. Note that the interconnect delay is dominant in these long interconnects where the interconnect ground capacitance C_g is much greater than the input capacitance C_{in} of the driven gate. The total interconnect capacitance is therefore crucial in satisfying timing requirements.

Table 2.2: The dependence of coupling capacitance on the signal switching characteristics.

	Signal transition	Effective coupling capacitance	Total capacitance of each interconnect
Interconnect 1	Switching	$\approx C_c$	$\approx C_g + C_c$
Interconnect 2	Nonswitching		
Interconnect 1	In phase	≈ 0	$\approx C_g$
Interconnect 2			
Interconnect 1	Out of phase	$\approx 2 \times C_c$	$\approx C_g + 2 \times C_c$
Interconnect 2			

The amount of interconnect coupling capacitance is dependent upon the nature of the signal transitions [41], [42], [43], [44]. That is, if only one of the interconnects is switching, the total capacitance of this interconnect is determined by $C_g + C_c$. If both interconnects are driven by *in-phase* signals (switching in the same direction), the effective coupling capacitance is approximately zero and the total capacitance of each interconnect is approximately equal to the ground capacitance C_g. Alternatively, if the signals on each interconnect are *out-of-phase* (switching in the opposite direction), the effective coupling capacitance approximately doubles to $2 \times C_c$. The total capacitance of the lines can therefore be approximated as $C_g + 2 \times C_c$, as listed in Table 2.2. Consequently, if the signals on both interconnects are in-phase, the total capacitance is lower, reducing the delay of the interconnect. Alternatively, out-of-phase signals increase the total capacitance, degrading the delay of the interconnect. This dependence of the total capacitance and delay on the switching characteristics also causes the overall *delay uncertainty* to increase, making the circuit timing analysis process quite challenging.

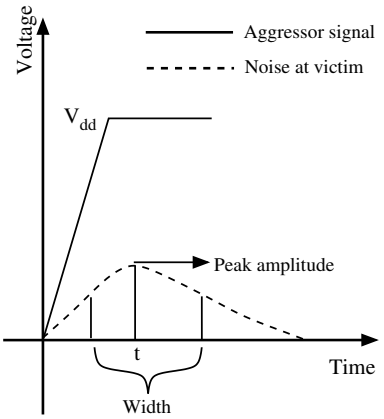


Figure 2.14: Coupling noise induced on the victim line due to signal switching on the aggressor line.

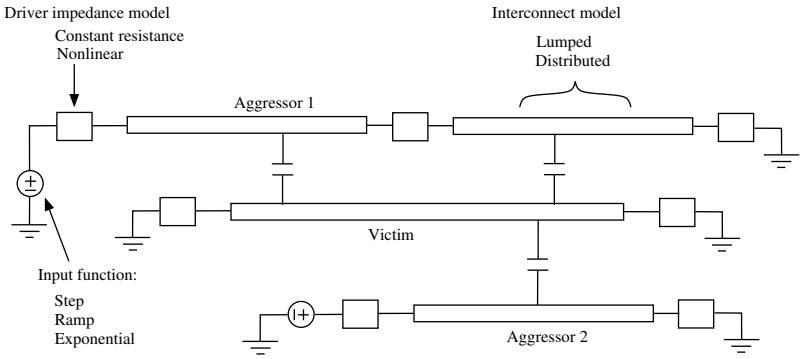


Figure 2.15: Primary modeling issues for coupling noise: Type of input excitation, impedance model of the driver, interconnect model, and location and number of aggressors affecting the victim line.

Modeling Capacitively Coupled Noise

A signal switching on one of the interconnects, referred to as an aggressor, induces coupling noise on the neighboring nonswitching interconnect, referred to as a victim, as illustrated in Fig. 2.2. Estimating the peak, width, and timing of this coupling noise is important to guarantee accurate logical operation (see Fig. 2.14). Four primary issues exist in modeling the coupling noise, as illustrated in Fig. 2.15.

- The impedance of the gate driving the interconnect changes with respect to the input voltage and operating regime of the transistor. This resistance is often approximated by a constant linear resistance to reduce the complexity of the transistor models. A transistor operating in the saturation region is commonly used since the I-V curve is close to constant in this region and the resistance is highest, providing a worst case value of the transistor impedance.

- The noise on the victim line depends strongly upon the type of input excitation. While an exponential function results in higher accuracy, the complexity of the analysis is much greater. A step input excitation provides efficient analysis, but exhibits a greater error for signals with slow transition times. A saturated ramp function is the most commonly used input excitation to model coupling noise since this function offers a reasonable tradeoff between accuracy and complexity.

- The interconnects can either be modeled as distributed RLC transmission lines or lumped elements such as π or L networks. Although a lumped model is simpler, the distributed nature of an interconnect becomes more pronounced in long global lines [27].

- The location and relative position of the aggressor and victim interconnects is more significant in complex integrated circuits where multiple aggressors can affect the coupling noise on a victim line, as illustrated in Fig. 2.15.

Various closed-form expressions and modeling techniques have been proposed to analyze the characteristics of coupling noise. A distributed RC model has been developed by Sakurai in [41] by approximating the driver gates with a linear resistor and assuming a step input excitation. A lumped RC model has been proposed in [43] where inputs can have arbitrary transition times and directions. A lumped π model with a ramp input excitation is used in [44]. The dependence of the delay uncertainty on the input slew has also been investigated. An upper bound on the coupling noise is proposed in [45], known as the Devgan metric. An enhancement to the Devgan metric has been proposed in [46] to improve the accuracy for fast input signals. Expressions have also been developed [47] for capacitively coupled distributed RC networks with an arbitrary number of aggressors.

These research results do not consider the inductive behavior of on-chip interconnects. As demonstrated in [48], on-chip inductance can significantly affect the signal behavior under certain conditions. Interconnect coupling noise in the presence of mutual inductance is discussed in the following section.

2.2.3 Inductive Coupling

According to the definition of magnetic flux, loop inductance is a constant relating the magnetic flux induced in a loop to a current in another loop [12], [49], as illustrated in Fig. 2.16. The mutual inductance L_{ij} is

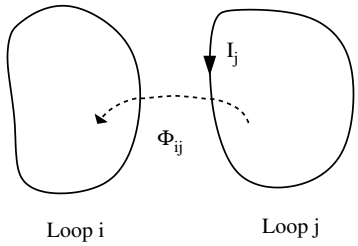


Figure 2.16: The loop inductance relates the magnetic flux in loop i induced by the current in loop j.

$$L_{ij} = \frac{\phi_{ij}}{I_j}, \tag{2.9}$$

where ϕ_{ij} is the magnetic flux in loop i induced by the current I_j flowing through loop j. Note that if the loops are the same ($i = j$), L_{ii} is defined as the self-inductance of the loop.

The effect of inductive coupling on the overall crosstalk noise has been investigated, demonstrating the need to consider mutual inductance among the interconnects to accurately estimate coupling noise and delay. Modeling the mutual inductance of on-chip interconnects, however, is a challenging process due to the following reasons:

- Inductive coupling originating due to magnetic fields is a long range phenomenon as opposed to capacitive coupling [49] which is caused by local electric fields, as illustrated in Fig. 2.17. The mutual inductance of nonadjacent interconnects is therefore non-negligible.

- As illustrated in Fig. 2.16, the loop inductance is defined for a closed loop. The extraction of self-and mutual inductances for complex on-chip interconnects is

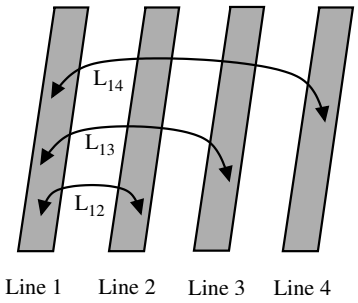


Figure 2.17: The long range effect of inductive coupling where the mutual inductance among nonadjacent interconnects is not negligible.

challenging since identifying the current return paths is not simple. The concept of a partial inductance has been introduced to overcome this difficulty [50], [51], where the loop is divided into multiple segments and the current return loop for each segment is assumed to be at infinity.

- A system with inductance can exhibit an underdamped response as opposed to capacitively dominant systems that are overdamped. The underdamped response can produce oscillations, overshoots, and undershoots. An illustrative example of coupling noise in the presence of mutual inductance is depicted in Fig. 2.18, as shown in [52]. These high frequency effects caused by an underdamped response make coupling noise analysis significantly more difficult.

Closed-form expressions for the peak inductive coupling noise is provided in [53] assuming that the coupling capacitance and mutual inductance are less than 30% of the self-capacitance and self-inductance, respectively. A traveling wave based waveform

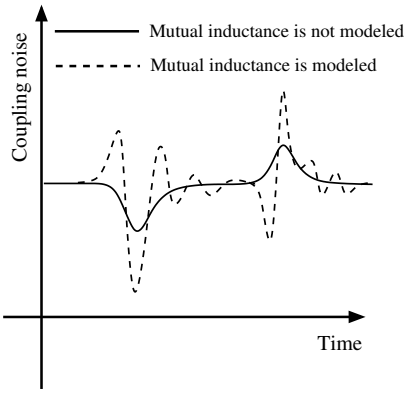


Figure 2.18: The effect of mutual inductance on coupling noise characteristics.

approximation is proposed in [54] for inductance prominent multicoupled interconnects, where the low frequency characteristics of the transient signal is approximated in the frequency domain and a time domain approximation is used to determine the high frequency characteristics. A transmission line based approach for inductive coupling is described in [55] to estimate the coupling noise waveform, peak noise, and dependence of the noise on relevant physical design parameters such as the line width and spacing.

2.3 Effects of Switching Noise in Digital ICs

The noise generated by a switching circuit affects the operation of the circuit in various ways. The possibility of a noise induced logical failure is described in Section 2.3.1. The effect of switching noise on the power consumption of a circuit is investigated in Section 2.3.2. Finally, the timing variation and delay uncertainty

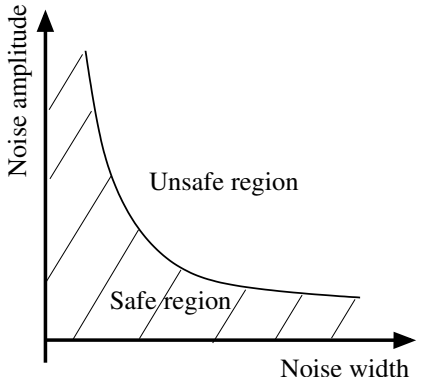


Figure 2.19: Typical noise tolerance curve of a gate based on the noise amplitude and width.

caused by switching noise is discussed in Section 2.3.3.

2.3.1 Functional Failure

Failure criteria of a circuit due to switching noise are dependent upon both the noise and the circuit characteristics. Specifically, the *noise tolerance* of a gate has to be sufficiently high to ensure that the noise at any node of the gate does not change the logical state of the gate and the following gates. Katopis defines the noise tolerance as "the maximum noise pulse amplitude at a given noise pulsewidth that an infinite chain of gates can withstand without noise amplification occurring" [56]. A typical noise tolerance curve based on this definition is depicted in Fig. 2.19. *Noise immunity* is defined as a special case of noise tolerance where the noise source is only applied at the input of the gate. Alternatively, noise tolerance is called *noise margin* if the noise source is applied both at the input and the power supply of the gate [56].

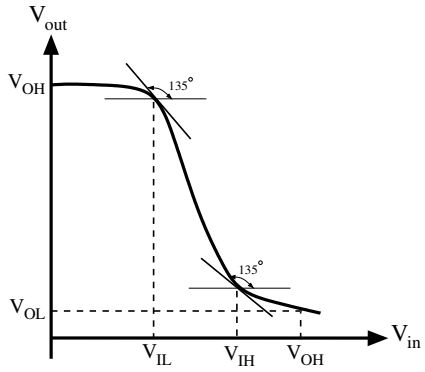


Figure 2.20: Typical DC voltage transfer characteristics of an inverter.

Two metrics exist to determine whether a specific noise waveform causes functional failure in a circuit: (1) Static noise margins based on a DC voltage transfer curve. (2) Dynamic noise margins based on the time domain characteristics of the noise.

Static Noise Margin

Traditionally, logical failure analysis caused by noise is based on *static* noise margins, obtained through a DC voltage transfer characteristics (VTC) of the gate, as shown in Fig 2.20 for an inverter [57]. Two critical points are identified on the VTC curve where the gain of the inverter is equal to unity, *i.e.*, $\partial V_{out}/\partial V_{in} = -1$. The input voltages satisfying this condition are called V_{IL} and V_{IH}. V_{IL} is defined as the maximum input voltage that can produce a logic high at the output. Similarly, V_{IH} is defined as the minimum input voltage that can produce a logic low at the output. V_{OH} and V_{OL} are defined, respectively, as the maximum output voltage when the output level is logic high and the minimum output voltage when the output level is

logic low. These two voltages determine the voltage swing of the gate. Noise margins based on static voltage transfer characteristics, also referred to as unity gain based noise margins, are determined by

$$NM_L = V_{IL} - V_{OL}, \tag{2.10}$$

$$NM_H = V_{OH} - V_{IH}, \tag{2.11}$$

where NM_L and NM_H are the noise margins when the input voltage is, respectively, at level low and at level high. These noise margins determine the amount of perturbation in the signal levels that can be tolerated by a digital logic gate.

Note that the gain of an inverter is higher than one (1) for an input voltage between V_{IL} and V_{IH}. Any perturbation of the input signal in this region is amplified at the output. A steeper voltage transfer characteristic reduces this region, improving the static noise margins of a gate.

Dynamic Noise Margin

The time domain characteristics of the noise such as the rise time and noise width is not considered in static noise margins. Digital gates are usually less sensitive to noise with high frequency components as compared to DC noise since the digital gates behave as low pass filters. The response of the gate therefore depends upon

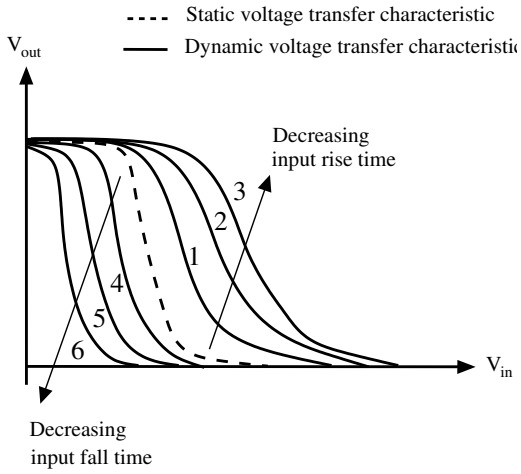


Figure 2.21: Dynamic voltage transfer characteristics of a CMOS inverter for various input transitions.

the frequency content of the noise. The concept of dynamic noise margins has been developed to consider these frequency dependencies, thereby reducing the inherent pessimism of static noise margins [58].

The dependence of the noise margins on the input rise time and output load capacitance has been demonstrated in [58]. An example of dynamic VTC is depicted in Fig. 2.21 for a CMOS inverter. The dashed curve represents the static VTC of the inverter. The curves numbered as 1, 2, and 3 represent the dynamic VTC for a rising input where 1 is the slowest transition and 3 is the fastest transition. Similarly, curves 4, 5, and 6 represent the dynamic VTC for a falling input where 4 is the slowest transition and 6 is the fastest transition.

A static VTC approximates the dynamic behavior only if the input transitions are slow, as illustrated in Fig 2.21. Note that NM_L of the dynamic VTC is higher than

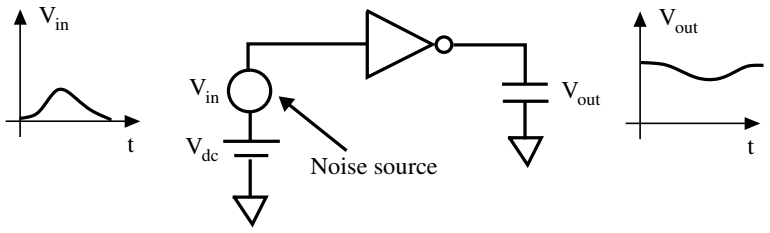


Figure 2.22: Time domain DC noise sensitivity of a CMOS inverter.

the static VTC for rising input transitions due to a greater V_{IL}. Similarly, NM_H of the dynamic VTC is higher than the static VTC for falling input transitions due to a smaller V_{IH}. A static VTC is therefore a conservative approach in estimating noise margins in logic gates.

A time domain DC noise sensitivity has been proposed in [59] as a failure criterion, where the transient characteristics of the noise have been considered. For a CMOS inverter, as shown in Fig. 2.22, the sensitivity of the output voltage $V_{out}(t)$ to DC variations at the input is,

$$S(t) = \frac{\partial V_{out}(t)}{\partial V_{dc}}. \tag{2.12}$$

If the magnitude of $S(t)$ is greater than one, the gate is considered to be *noise unstable* and the corresponding noise voltage $V_{in}(t)$ at the input violates the dynamic noise margins of the gate. The pessimism associated with static noise margins is reduced with dynamic noise margins at the expense of increased computational complexity due to the requirement to calculate the time domain sensitivities.

A *latch transition* based criterion is presented in [60] to further reduce the pessimism of the *unity gain* based criterion introduced in [59]. The proposed latch transition criterion is based on the observation that the noise can cause failure in the circuit only if this noise propagates to the input of a memory element and is able to change the state of this memory element. Rather than enforcing noise margins at each stage, only the memory elements are evaluated to determine if the state of the memory element has changed due to the propagated noise at the input. Two primary difficulties exist for the application of this criterion. The first difficulty is to accurately model the noise propagation at the input of a memory element. A model of this propagation is difficult due to the nonlinear behavior of the gates along the path. The second difficulty is to determine whether the noise at the input of a memory element can change the state of the memory element. The result depends not only on the noise characteristics, but also the relative alignment of the noise pulse with the active edge of the clock signal [61], as illustrated in Fig. 2.23.

2.3.2 Glitch Power Consumption

A glitch refers to a spurious transition at the input or output of a logic gate due to two reasons: (1) The switching times of the input signals produce a temporary voltage spike at the output before the signal reaches steady state, as illustrated in Fig. 2.24(a) [62], [63], or (2) capacitive or inductive coupling induces interconnect

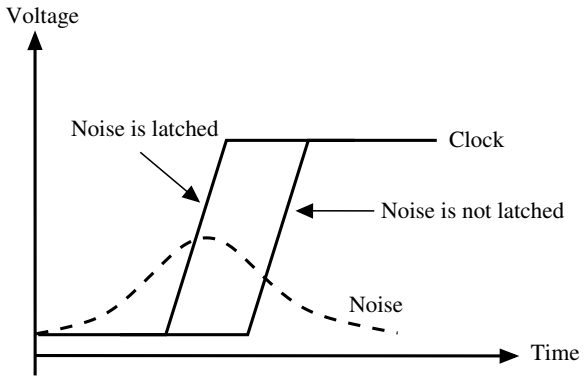


Figure 2.23: Alignment of the noise pulse with the clock signal to determine whether the noise is latched into the memory element.

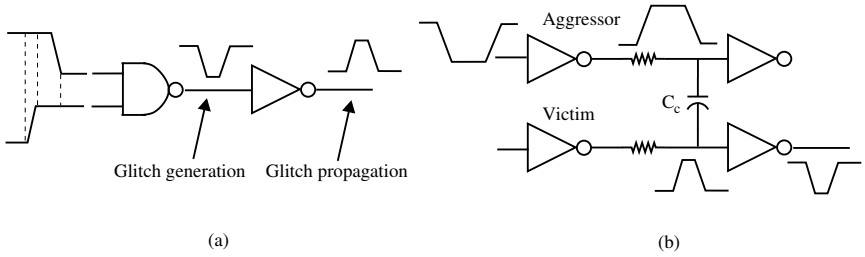


Figure 2.24: Generation and propagation of a glitch due to two different reasons: (a) Misalignment of the switching times of the input signals. (b) Capacitive coupling among interconnects.

noise, causing a voltage spike [64], as illustrated in Fig. 2.24(b).

These spurious voltage spikes or glitches dissipate unnecessary power. If the peak magnitude of the glitch is greater than the threshold voltage, the transistor turns on momentarily, dissipating extra dynamic power. Alternatively, if the peak magnitude of the glitch is smaller than the threshold voltage, the glitch contributes to the static power consumption due to increased leakage current.

2.3.3 Increased Delay Uncertainty

Delay uncertainty refers to a variation in the delay of a gate or an interconnect due to various reasons such as process variations, fluctuations in the power supply voltage, and operating temperature. Delay uncertainty (DU) can be quantified as the difference between the minimum and maximum delays along a path [44],

$$DU = tD_{max} - tD_{min}. \tag{2.13}$$

An alternative definition of delay uncertainty is the ratio of the maximum delay to the minimum delay [65],

$$DU = \frac{tD_{max}}{tD_{min}}. \tag{2.14}$$

Within a timing path, the arrival times of a data or clock signal at the input of a memory element such as a register vary due to this delay uncertainty, as illustrated in Fig. 2.25. The uncertainty in the data path delay tD, clock launch path delay tL, and clock capture path delay tC is due to variations both in the gate and interconnect delays. Note that for clock paths, the delay tends to be interconnect dominated whereas the data paths are typically device dominated. The delay uncertainty in interconnects and gates are described in the following subsections.

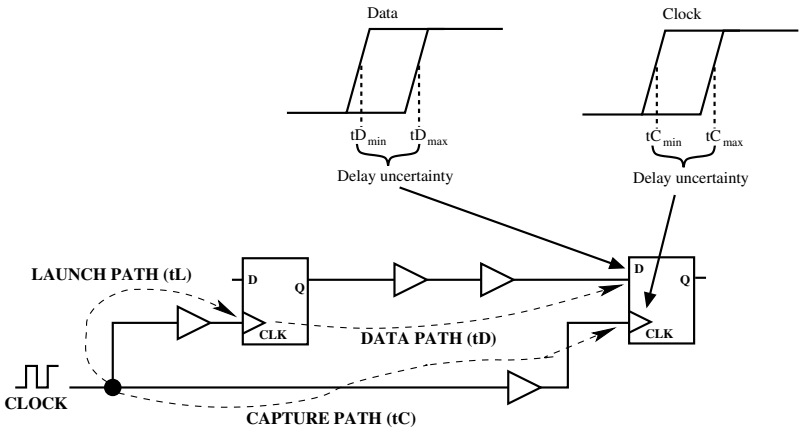


Figure 2.25: Variation in the arrival time of the data and clock signals, described as delay uncertainty.

Delay Uncertainty in Coupled Interconnects

The effective coupling capacitance between two interconnects is dependent upon the switching characteristics of the signals due to the Miller effect [66], as described in Section 2.2.2. Two capacitively coupled interconnects are shown in Fig. 2.26(a). Based on Miller effect, this coupling circuit can be approximated with a decoupling circuit using the Miller factor α [66], as illustrated in Fig. 2.26(b). The total effective capacitance of the victim interconnect is $C_g + \alpha C_c$. The value of the Miller factor α depends upon the switching directions, rise/fall times, and switching times of the aggressor and victim signals. Traditionally, α is assumed to be two if the aggressor and victim switch in opposite directions, *i.e.*, out-of-phase. Alternatively, if the aggressor and victim are in-phase, α is assumed to be zero, as listed in Table 2.2. However, values of α higher than two and smaller then zero have been reported [67], [68]. In [66],

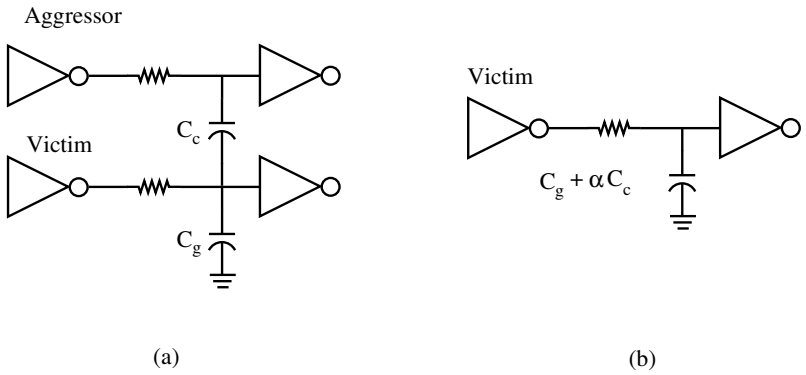


Figure 2.26: Transformation of a capacitively coupled interconnect into a decoupled interconnect using the Miller factor: (a) Two capacitively coupled interconnects, (b) Decoupled interconnect using the Miller factor.

an upper bound of three and a lower bound of minus one have been demonstrated if the switching times of the aggressor and victim are different.

This dependence of the total interconnect capacitance on the switching characteristics of the signals significantly changes the delay of the interconnect dominated timing paths. For a global interconnect where the interconnect capacitance C_{int} is much greater than the load capacitance C_L, the delay t_{int} is given by [69]

$$t_{int} = C_{int} \times (0.7 \times R_{tr} + 0.4 \times R_{int}), \tag{2.15}$$

where R_{tr} and R_{int} are, respectively, the resistance of the driver transistor and interconnect resistance. A worst case analysis considers the highest value of α to determine the total interconnect capacitance when evaluating the delay of the critical paths. Alternatively, a pessimistic analysis of the fast paths requires the lowest value of α. The

increasing difference between these two extrema has forced the timing analysis process to efficiently consider the effects of coupling capacitances on the signal arrival times [70], [71].

Delay Uncertainty in Logic Gates

The delay of a logic gate is dependent upon the drain current flowing through the gate to charge (during a low-to-high transition at the output) or discharge (during a high-to-low transition at the output) the load of a gate. Based on Sakurai's α-power law model [72], the drain current I_D is a function of the power supply voltage V_{DD},

$$I_D = \begin{cases} 0 & \text{if } V_{GS} \leq V_{TH} \text{ (cutoff region)} \\ (I'_{D0}/V'_{D0})V_{DS} & \text{if } V_{DS} < V'_{D0} \text{ (triode region)} \\ I'_{D0} & \text{if } V_{DS} \geq V'_{D0} \text{ (saturation region)} \end{cases}$$

where

$$I'_{D0} = I_{D0}(\frac{V_{GS} - V_{TH}}{V_{DD} - V_{TH}})^{\alpha}, \tag{2.16}$$

$$V'_{D0} = V_{D0}(\frac{V_{GS} - V_{TH}}{V_{DD} - V_{TH}})^{\alpha/2}. \tag{2.17}$$

I_{D0} refers to the drain current at $V_{GS} = V_{DS} = V_{DD}$. V_{D0} is the drain saturation voltage at $V_{GS} = V_{DD}$. Note that depending upon the number of power and ground pins and the type of package, the power/ground noise can appear during the input

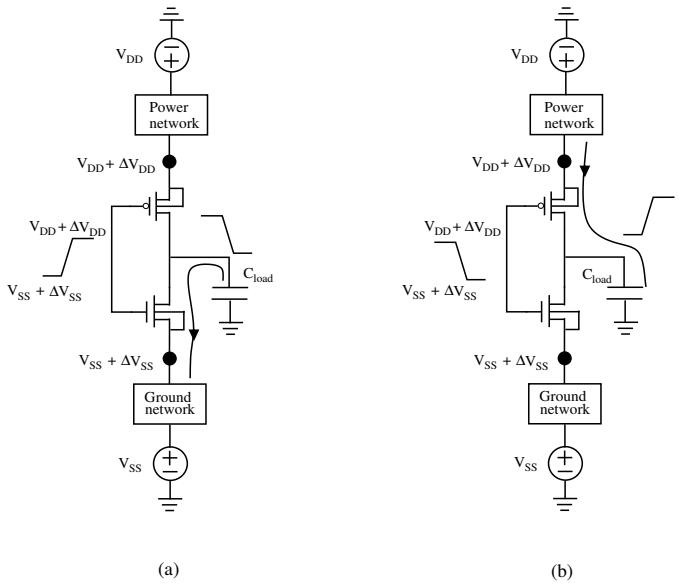


Figure 2.27: Power noise ΔV_{DD} and ground noise ΔV_{SS} appear at the input of the gate assuming a single power and ground pin: (a) High-to-low transition of the output. (b) Low-to-high transition of the output.

transition of the gates, as depicted in Fig. 2.27. That is, the high and low levels of the input transitions are, respectively, $V_{DD} + \Delta V_{DD}$ and $V_{SS} + \Delta V_{SS}$ where ΔV_{DD} and ΔV_{SS} refer, respectively, to the power and ground noise. The V_{GS} term in (2.16) and (2.17) is therefore also dependent on the power/ground noise.

An analysis of the variation in delay in a buffer circuit due to DC power/ground noise has been described in [73]. The effects of *differential mode noise* $\Delta V_{DD} - \Delta V_{SS}$ and *common mode noise* $\Delta V_{DD} + \Delta V_{SS}$ have been investigated. Expressions have been developed to quantify the change in delay as a function of multiple parameters such as the rise time, load capacitance, alpha power law parameter [72], and threshold voltage. Furthermore, the effect of different load models such as a lumped RC load,

π model load, distributed RC load, and an effective capacitance load on the delay variation characteristics has been discussed [73].

The effect of power noise on the timing characteristics of device dominated and interconnect dominated paths has been investigated in [28]. The *average power noise* rather than the peak noise is shown to be the fundamental metric to determine the timing variations within a data path. That is, "from a timing standpoint, a short pulse with a large amplitude can have the same effect as a wider pulse with a smaller amplitude" [28].

Integrating the effects of power/ground noise in static timing analysis methodologies has also received attention [74], [75]. Traditionally, a worst case voltage drop is considered for each power node along the path, resulting in a highly pessimistic analysis. A voltage-aware static timing analysis methodology has been proposed in [74]. The proposed technique considers the mismatches among the power supplies of successive gates on a data path to accurately estimate the worst case delay. A vectorless analysis based on spatial and temporal superposition of the voltage drops has been presented in [75]. Correlations among the switching currents of various circuit blocks are incorporated into the analysis to consider spatial power noise variations. An average improvement of 17% is demonstrated over a traditional pessimistic analysis.

2.4 Techniques to Reduce Switching Noise in Digital ICs

Existing methodologies to alleviate the effects of switching noise in digital integrated circuits are described in this section. Techniques to reduce power/ground noise and interconnect coupling noise are summarized, respectively, in Sections 2.4.1 and 2.4.2.

2.4.1 Power/Ground Noise Reduction

Several techniques to reduce power/ground noise are summarized in this section. These techniques include the use of decoupling capacitors, skew and slew rate control, spread spectrum clock generation, and low impedance packaging techniques.

Decoupling capacitors

Decoupling capacitors have been introduced in Section 2.1.3 to satisfy target impedance requirements while reducing peak power/ground noise. A decoupling capacitor stores charge, providing part of the switching current while bypassing the impedance on the power supply portion of the power distribution network. The efficiency of on-chip decoupling capacitors in reducing the noise is strongly dependent upon the placement and size of these capacitors [32], [76], [77], [78], [79], [80].

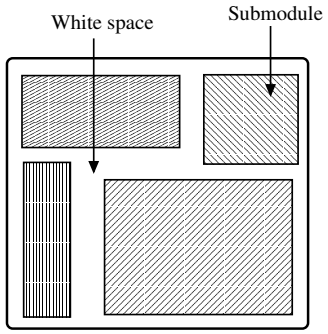


Figure 2.28: Illustration of white space utilized for decoupling capacitor placement.

A simple charge based allocation strategy has been proposed in [32]. The size of the decoupling capacitor is determined based on the current demand I and fraction n of the ripple voltage allowed,

$$C_{decap} = \frac{\partial Q}{\partial V} = \frac{I}{f \times V_{DD} \times n}. \tag{2.18}$$

The application of this strategy, however, is infeasible due to the increasing complexity of power distribution networks and the requirement for a system of hierarchical decoupling capacitors [12], [35]. An excessive noise based allocation has been proposed in [76]. The peak noise is analyzed assuming no decoupling capacitors. The amount of decoupling capacitance is determined at the next step based on *excessive noise* to prevent unnecessarily large decoupling capacitors. The *white space* among the submodules is utilized to place the decoupling capacitors, as illustrated in Fig. 2.28. The final location of the decoupling capacitors is therefore limited by the circuit floorplan. A placement methodology has been proposed in [77], [78] based on the effective radii

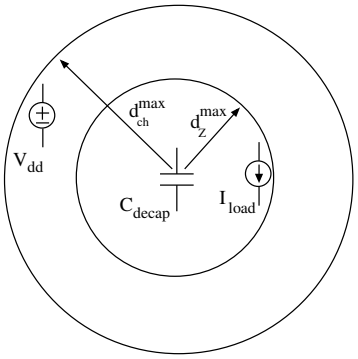


Figure 2.29: Effective radii of the decoupling capacitor as determined by the target impedance and charge time.

of the decoupling capacitor, as illustrated in Fig 2.29. A design space is characterized to satisfy the target impedance, as determined by d_Z^{max} in Fig. 2.29, while restoring the charge on the decoupling capacitor within a target charge time, as determined by d_{ch}^{max} in Fig. 2.29.

Another class of high level techniques to determine decoupling capacitor sizing and placement is based on applying a sensitivity evaluation [79], [80]. The sensitivity of a capacitor to the power noise is determined for various capacitor sizes. In [79], the time integral of the noise at each node is used as a figure of merit to determine the *adjoint sensitivities*. Note that an adjoint sensitivity analysis enables the sensitivity of a performance parameter to be determined for various circuit parameters [81]. The proposed technique requires knowledge of the voltage waveforms, increasing the computational complexity. A partitioning based approach has been proposed in [80] to improve the complexity where the circuit is divided into smaller subcircuits and each subcircuit is separately optimized. While achieving a system level optimization,

these high level techniques usually fail to consider the physical characteristics and constraints within the circuit such as the effective series resistance and effective series inductance of the decoupling capacitor.

Skew and Slew Rate Control

Skew control refers to intentional skews introduced to the signal or clock paths to reduce the peak currents in order to decrease the overall simultaneous switching activity. This technique has been proposed in [23] for output drivers by inserting a chain of buffers along the signal path. More than a 60% reduction in the ground bounce is demonstrated using this skew control methodology. Clock skew optimization methodologies have also been proposed to shape the supply current by dividing the synchronous clock into multiple subclocks with relative skew [82], [83], [84]. Peak currents have been reduced by a factor of two. An example of supply current modification in both the time and frequency domains is illustrated in Fig. 2.30, as depicted in [84]. Note that these intentional skews should not violate the timing requirements of the circuit. These approaches are therefore limited by the available slack in the timing paths, making this methodology strictly dependent upon the local timing constraints.

Slew rate control has also been proposed to reduce $L\ \partial i/\partial t$ drops by increasing the rise and fall times of the current transients. Those parts of the circuit operating at a

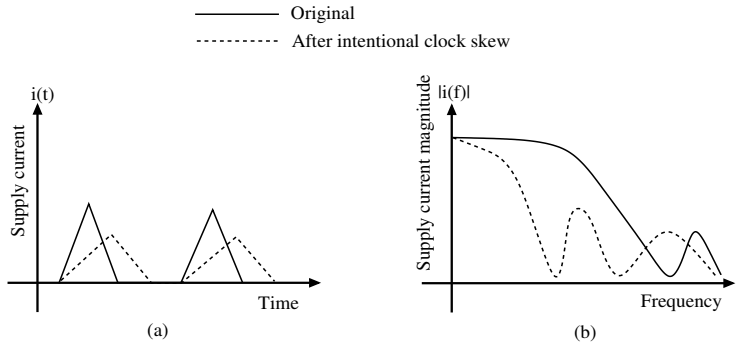


Figure 2.30: Supply current before and after intentional clock skew to reduce power noise: (a) In the time domain. (b) In the frequency domain.

faster speed than necessary are slowed down, increasing the rise and fall times [85], [86]. This technique however is not feasible in deep submicrometer technologies where longer rise and fall times are prohibitive in high performance circuits.

Spread Spectrum Clock Generation

Electromagnetic interference (EMI) is a typical concern for high speed synchronous digital systems where the amount of radiated emissions should be less than a specific value. The energy of periodic signals with fast transition times is concentrated around the harmonic frequencies within a narrow band [87].

Spread spectrum clocking techniques have been introduced to alleviate electromagnetic interference by frequency modulating the system clock signal [88], [89]. The peak amplitude at the harmonic frequencies is reduced by spreading the energy over a wider bandwidth, as depicted in Fig. 2.31. In the time domain, this frequency modulation results in clock waveforms with slightly different periods in each cycle.

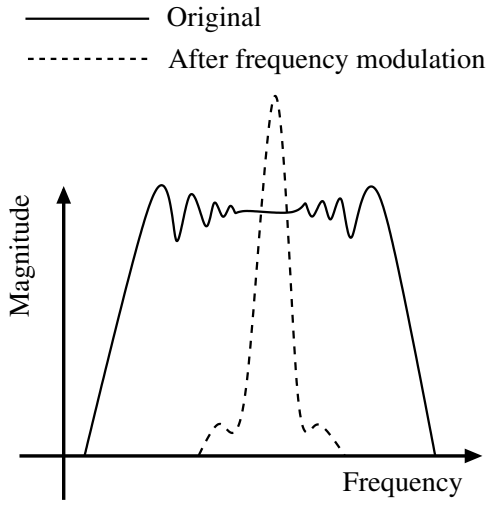


Figure 2.31: Reduction in the peak amplitude of a harmonic due to frequency modulation of the trapezoidal clock signal.

This approach has been adopted to reduce power/ground noise, as described in [84]. The spreading is maximized in the vicinity of the circuit resonance. A 14.5 dB reduction in the peak amplitude of the resonant frequency has been demonstrated [84].

Note that spread spectrum clocking to reduce power/ground noise exhibits a strong tradeoff between the amount of attenuation and the overall system speed. That is, as the modulation index increases, the reduction in noise also increases. The modulation index, however, cannot be chosen arbitrarily large due to the timing constraints of the circuit.

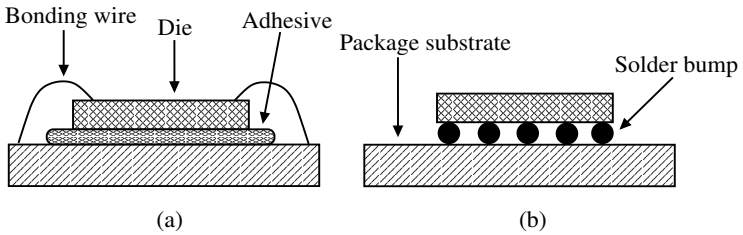


Figure 2.32: Two different bonding techniques to connect the die pads and package: (a) Wire bonding. (b) Solder bump bonding using flip-chip technique.

Package Characteristics

The characteristics of the package and board technology play a significant role in the overall noise performance of the integrated circuit. The length, width, thickness, and spacing of the interconnects, the number of power/ground planes, the number of pads, and type of materials used within the package determine the electrical characteristics of the circuit by affecting the parasitic impedances of the package, specifically, the inductance, resistance, and capacitance.

The *package bandwidth* is an important parameter characterizing a packaging technology, as determined by these parasitic impedances. Several other significant parameters of a package are the rate of heat removal to satisfy thermal constraints, stress limits on the mechanical connections, and the number of I/O pins [15].

Two particular die attachment techniques, wire bonding and flip-chip packages with solder bump bonding, can have a significant effect on the noise performance of an integrated circuit. These two types of bonding mechanisms are illustrated in Fig. 2.32. Wire bonding connects the die pad with the package substrate by

a relatively long wire, as illustrated in Fig. 2.32(a). This structure exhibits poor electrical performance, primarily due to the large mutual and self-inductance of the bonding wire. Furthermore, the number of pins are limited. The primary advantage of wire bonding is low cost. Alternatively, solder bump bonding using a flip-chip technique achieves enhanced electrical performance. The die is flipped and aligned to the pads with the solder bumps, as illustrated in Fig. 2.32(b). The height of the bump is considerably smaller than the length of the bond wires, significantly reducing the parasitic inductance. Furthermore, the entire die surface is utilized for the pad, providing a large number of available pins. The primary disadvantages of the flip-chip technique are higher cost, poor thermal conduction ability, and difficult to inspect and test.

2.4.2 Reduction of Interconnect Coupling Noise

Several techniques to reduce interconnect coupling noise are reviewed in this section. These techniques include shield insertion, repeater insertion, signal rerouting and wire reordering, wire sizing and spacing, and gate sizing.

Shield Insertion

Shield insertion refers to a technique where a power or ground line is placed between the aggressor and victim signal lines to reduce the capacitive and inductive

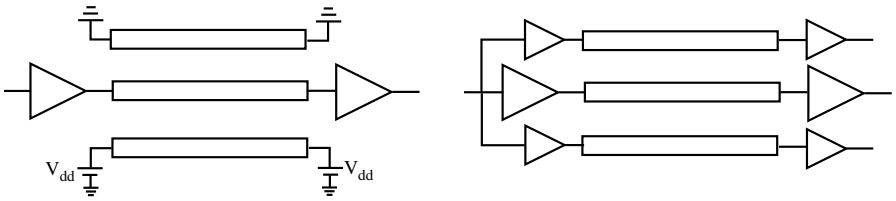


Figure 2.33: Shield insertion to reduce coupling among interconnects: (a) Passive shielding where the shield lines are connected to power or ground. (b) Active shielding where the shield lines are driven with the same input signal to reduce the effective coupling capacitance.

coupling, as depicted in Fig 2.33(a). If a power or ground line is in the vicinity of the signal line, the signal line can be placed close to this existing power or ground line. Alternatively, a power or ground line is explicitly placed near the signal line for the purpose of shielding. This type of shield insertion is called passive shielding [90]. The signal line is isolated from the switching neighbor lines, reducing any capacitive coupling. The inductive coupling is also decreased since the shield line provides a closer current return path, reducing the mutual inductance [91].

An active shield insertion methodology has also been proposed to exploit the Miller effect [92], [93]. That is, the shield line switches in the same direction as the signal line, reducing the effective coupling capacitance, as illustrated in Fig. 2.33(b). A 16% improvement in performance is achieved as compared to passive shielding.

Both active and passive shielding consumes additional area. A shield placement methodology is proposed in [90] to minimize the area occupied by the shield lines while satisfying noise coupling constraints. Active shielding also consumes more power due

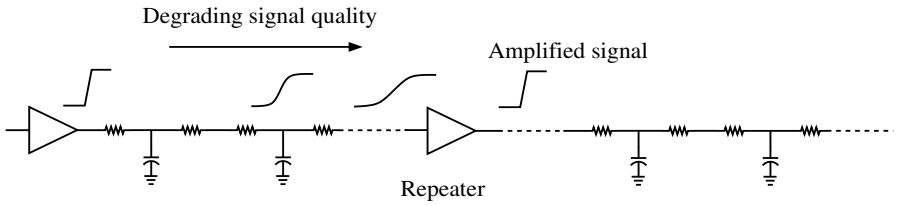


Figure 2.34: Repeater insertion to reduce the interconnect delay by restoring the signal along a line.

to the additional switching activity of the shield lines. The power consumption due to the coupling capacitance is reduced however. An active shielding scheme is proposed in [94] to lower the power consumption.

Repeater Insertion

The delay of a global interconnect grows quadratically with the length of the interconnect due to a linear increase in both the parasitic resistance and capacitance of the line. Repeater insertion has been commonly used to overcome this quadratic dependency, reducing the delay by dividing the interconnect into smaller subsections. A repeater is typically an inverting or noninverting buffer placed at specific locations along the interconnect, as illustrated in Fig. 2.34. The amplifying nature of inverters is exploited to restore the signal while reducing the overall delay of the line. Traditionally, delay and power have been the primary design metrics to optimize the size, number, and location of the repeaters. For an interconnect with constant wire width, Bakoglu characterizes the optimal number of repeaters k and uniform size h for each

repeater to obtain the minimum delay in an interconnect [39],

$$k = \sqrt{\frac{R_{int}C_{int}}{2.3R_0C_0}}, \tag{2.19}$$

$$h = \sqrt{\frac{R_0C_{int}}{R_{int}C_0}}, \tag{2.20}$$

where R_{int} and C_{int} are, respectively, the total interconnect resistance and capacitance. R_0 and C_0 are, respectively, the output resistance and input capacitance of the minimum size inverter.

An algorithm is presented in [95] to minimize the Elmore delay [96], [97] of an interconnect represented as an RC tree network. The proposed algorithm determines the optimal location and size of the repeaters with polynomial time complexity.

A repeater design methodology is proposed in [98] where the delay of an RC interconnect driven by an inverter is determined based on the α-power law model [72]. The proposed technique determines the optimum number of uniformly sized repeaters where the error of the model is less than 16% as compared to SPICE.

The effect of inductance on repeater insertion methodologies has been investigated in [99]. The quadratic dependence of delay on line length for an RC line approaches a linear dependence for an RLC line. The optimum number of repeaters is therefore fewer in the presence of inductance.

Inserting repeaters not only reduces the delay of the global interconnects, but also

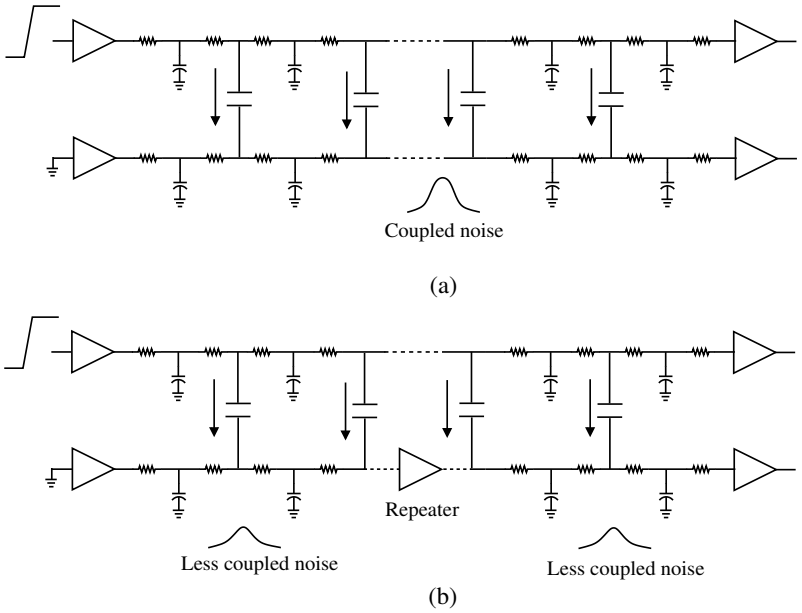


Figure 2.35: Repeater insertion to reduce capacitive coupling noise: (a) No repeaters exist. The total coupling capacitance is high. (b) The repeater reduces the coupling capacitance by dividing the interconnect into multiple (two in this example) sections.

lessens the capacitive coupling noise among interconnects, as depicted in Fig. 2.35. The coupling noise is proportional to the length of the two parallel interconnects. This parallel portion is decreased by inserting repeaters, thereby reducing the coupling noise.

The optimal delay algorithm proposed in [95] for repeater insertion has been extended to consider coupling noise among interconnects [100]. A repeater insertion algorithm has been proposed for a tree structured RC interconnect to optimize delay while satisfying coupling noise constraints [100]. The Devgan noise metric [45] is used to evaluate whether the noise conditions are satisfied.

A methodology for simultaneous shield and repeater insertion is proposed in [101]

to reduce the coupling noise, also based on the Devgan noise metric. Global routing of the power/ground and signal wires is simultaneously considered, utilizing the power/ground wires as shield lines. Improvements of up to 53% and 28% are achieved as compared to, respectively, only repeater insertion and only shield insertion.

Signal Rerouting and Wire Ordering

Interconnect routing and topology optimization have long been a focus of interest in the integrated circuit design process [102]. The area, delay, and power of these interconnects are the primary design objectives that are traditionally considered during the routing process. Given the connectivity information, the wire length is minimized to achieve the highest performance, while reducing the area and switching power. Minimum spanning trees and Steiner minimal trees are frequently applied to this interconnect routing optimization process [103].

Due to the increasing importance of interconnect coupling noise, signal routing and wire ordering methodologies have been reconsidered by including crosstalk as a design constraint in addition to the traditional constraints of area, delay, and power. Specific focus has been placed on the relative position of the wires, *e.g.*, the spacing and length of the overlap between the victim and aggressor lines [104], [105], [106], [107].

Traditionally, the interconnects are routed in an integrated circuit through a two

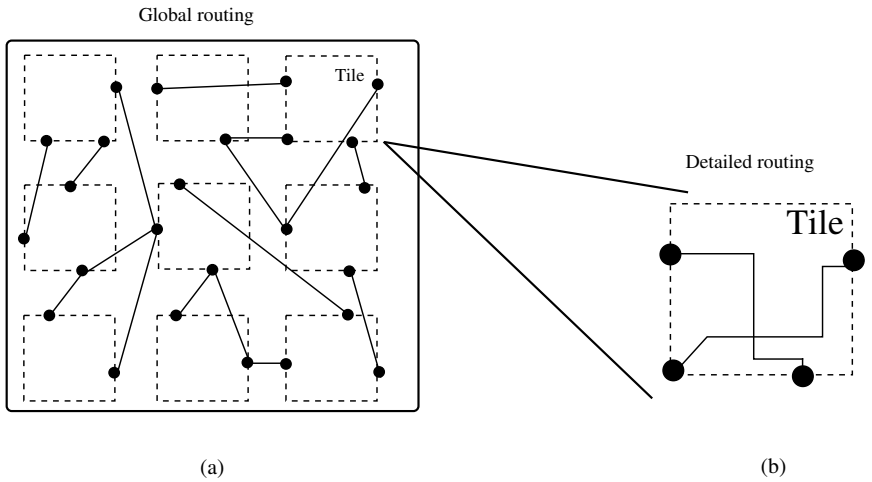


Figure 2.36: Routing the interconnects in an integrated circuit: (a) Global routing. (b) Detailed routing.

step process: *global routing* followed by *detailed routing*, as depicted in Fig. 2.36. In the global routing phase, the overall area is divided into tiles. For each net, the paths through the tiles are determined. The detailed routing of the nets within the tiles is achieved in the second phase [108].

Algorithms and methodologies have been proposed to reduce coupling noise at the detailed routing phase [104], [105]. A modified greedy channel routing algorithm has been described in [106] to consider capacitive coupling noise where the parasitic resistance of the wires is neglected. Overlap among the wires and the drive strength along the adjacent wires are exploited. On average, a 22% reduction in coupled noise is achieved as compared to a conventional greedy router. Typically, a more accurate estimation of the crosstalk noise is obtained at the detailed routing phase, but the flexibility in allocating resources is usually limited by the global routing process [107].

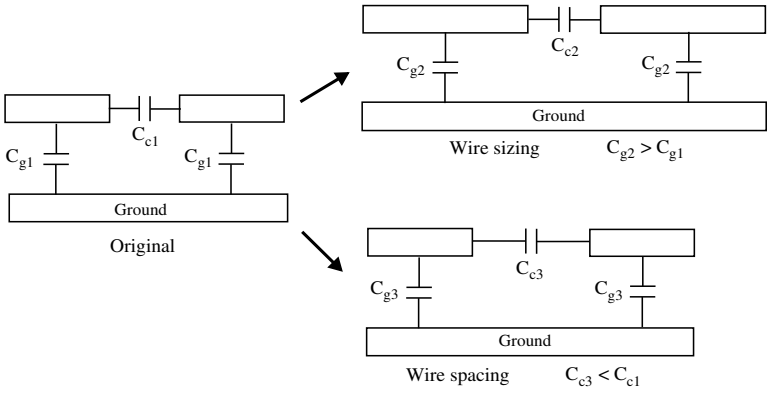


Figure 2.37: Wire sizing and spacing to reduce coupling noise, respectively, by increasing the ground capacitance and decreasing the coupling capacitance.

A global routing methodology has been introduced in [107] to satisfy crosstalk constraints. The proposed technique exploits the flexibility in allocating resources during the global routing process. An extended global routing methodology is proposed in [109] for *RLC* lines where the mutual inductive coupling among the interconnects is included. The primary limitation in considering the crosstalk at the global routing phase is insufficient circuit level information to accurately estimate the coupling noise.

Wire Sizing and Spacing

Increasing the width of the wire and spacing between two wires are commonly used design techniques to reduce coupling noise, as illustrated in Fig. 2.37. Increasing the width of the lines increases the ground capacitance and reduces the parasitic resistance. An increase in the ground capacitance of the aggressor line decreases

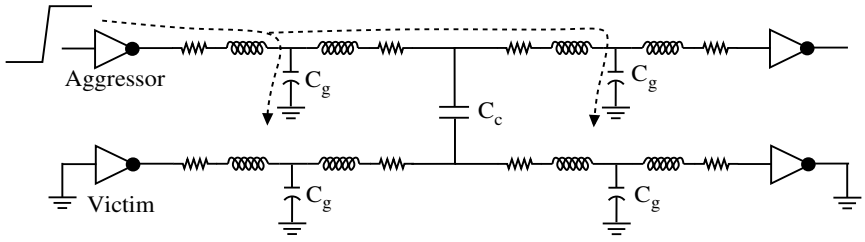


Figure 2.38: An increase in the width of the aggressor line increases the ground capacitance. A higher ground capacitance behaves as a filter, reducing the coupling noise.

the coupling noise due to lower impedance to the ground. The increased ground capacitance behaves as a filter, reducing the coupling noise, as illustrated in Fig. 2.38. The disadvantages of widening the line are additional area and possibly an increase in delay due to the additional interconnect capacitance.

The effect of increasing the line width has been investigated in [55] for both capacitive and inductive coupling. If only capacitive coupling exists, the coupling noise is reduced due to the higher ground capacitance. Alternatively, if only inductive coupling exists, the coupling noise increases with wider lines. The total coupling noise is therefore shown to be relatively insensitive to the line width in the presence of both capacitive and inductive coupling, as depicted in Fig. 2.39.

Increasing the spacing between the aggressor and victim lines reduces the coupling noise by decreasing the coupling capacitance, as illustrated in Fig 2.37. The reduction in inductive coupling however is not as significant due to the fact that the mutual inductance decreases logarithmically with distance as compared to the linear reduction of the coupling capacitance [52], [55]. For sufficiently wide spacing, the coupling

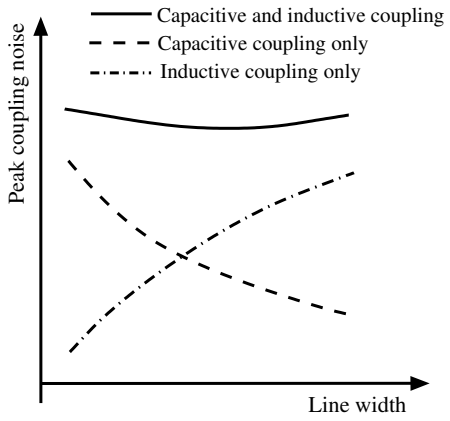


Figure 2.39: The effect of line width on coupling noise in the presence of capacitive and inductive coupling.

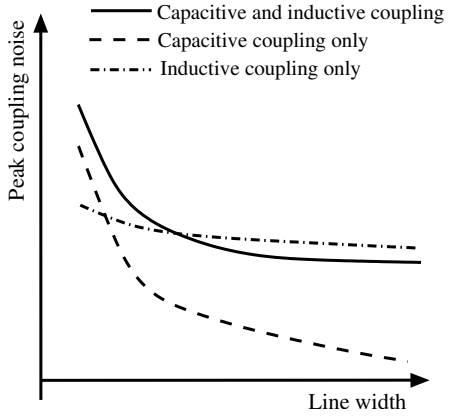


Figure 2.40: The effect of spacing on coupling noise in the presence of capacitive and inductive coupling.

noise is dominated by inductive coupling, as depicted in Fig. 2.40 [55]. Increasing the spacing is therefore an efficient technique only below a certain threshold where the mutual inductance starts to dominate the overall noise coupling.

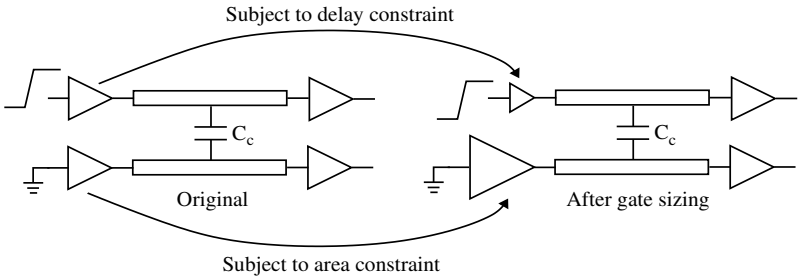


Figure 2.41: Gate sizing to reduce coupling noise. The driver of the aggressor and victim are, respectively, downsized and upsized to reduce coupling noise.

Gate Sizing

Gate sizing is a commonly used technique to exploit the tradeoff between speed and power in integrated circuits. Those gates located along the critical paths are sized larger to satisfy critical delay constraints. Alternatively, the remaining gates are sized smaller to reduce power consumption.

The relative size of the aggressor and victim drivers also affects the coupling noise and coupling noise induced delay variation [110], [111]. Gate sizing to reduce coupling noise is depicted in Fig. 2.41. A decrease in the size of the driver in the aggressor reduces the coupling noise since the ability of the aggressor to induce noise on the victim line is degraded. Downsizing the driver on the aggressor, however, slows down the signal path. Adjusting the size of the aggressor driver to reduce crosstalk is therefore subject to delay constraints. Alternatively, increasing the size of the victim driver reduces the coupling noise since the victim is more effectively connected to a stable voltage such as power or ground. Increasing the size of the driver on the

victim, however, increases the overall area, making this technique subject to area constraints [110], [111].

A gate sizing algorithm has been proposed in [112]. The algorithm is applied once the routing phase has been completed. Each line is considered as both an aggressor and a victim since increasing the size of a victim driver can produce a negative effect on a different line for which the victim behaves as an aggressor. The gates should therefore be sized to not introduce additional coupling noise on one line while reducing the noise on another line.

2.5 Chapter Summary

An overview of switching noise in synchronous digital integrated circuits has been presented in this chapter. The characteristics of switching noise can be summarized as follows:

- Noise in synchronous digital integrated circuits has become a primary design concern due to decreasing noise margins, increasing clock frequencies, faster on-chip signal transitions, and the significant role of the interconnects on the overall system performance and signal integrity

- Two primary types of switching noise exist in a synchronous digital integrated circuit: power/ground noise and interconnect coupling noise

- Power/ground noise refers to both static and dynamic voltage fluctuations on the power and ground nodes due to large amounts of switching current flowing through the RLC parasitic impedances of the power and ground distribution networks

- Interconnect coupling noise refers to a voltage induced on a victim line due to capacitive and inductive coupling from a switching aggressor line

- A power distribution network consists of interconnect networks at multiple physical levels such as the voltage regulator, printed circuit board, package, and integrated circuit

- A power distribution network can be analyzed in the time and frequency domains to determine whether the noise constraints can be satisfied

- In the time domain, the peak noise and average noise within a clock period are the two primary metrics. The peak noise should be less than the noise margin to guarantee accurate functionality. The average noise should be sufficiently low to not introduce any timing violations

- In the frequency domain, the impedance of the power distribution network should be less than the target impedance over a wide frequency range as determined by the transition time of the signals

- Hierarchical decoupling capacitors across multiple physical levels are commonly used to increase the frequency range of a power distribution network and reduce the power/ground noise

- The length of the global interconnects grows by a factor of S_c (chip scale factor) due to the trend of increasing die size

- Practical scaling differs from ideal scaling to alleviate the dramatic increase in interconnect resistance. The thickness and height of the interconnect do not scale at the same ratio as the length and width of the interconnect

- The ratio of the coupling capacitance to the ground capacitance has been increasing, making capacitive coupling a significant signal integrity issue

- The effective capacitance of coupled interconnects is dependent upon the type of signal transitions where in-phase switching reduces the effective coupling capacitance and out-of-phase switching increases the effective coupling capacitance, as determined by the Miller effect

- Inductive coupling is a long range phenomenon making the mutual inductance of nonadjacent interconnects not negligible

- Efficient and sufficiently accurate modeling of the coupling noise is a significant issue. The resistance of the driver transistor, type of input excitation,

the parasitic impedances of the interconnect, and the mutual capacitances and inductances are some of the primary considerations when modeling the interconnect noise

- Switching noise can cause a circuit failure, increase the glitch power consumption, and introduces delay uncertainty within the circuit
- The failure criterion is traditionally based on static noise margins as obtained through DC voltage transfer characteristics. Static noise margins do not consider the time domain characteristics of the signal
- Dynamic noise margins consider the effects of rise time and pulse width on the failure criterion, thereby reducing the pessimism associated with static noise margins
- The variation of the effective coupling capacitance in terms of the transition type causes delay uncertainty in coupled interconnects
- The power/ground noise causes delay uncertainty in logic gates by changing the switching current
- The placement and size of the decoupling capacitors play an important role in determining the efficiency of these capacitors
- Skew control has been proposed to reduce power/ground noise by decreasing

the simultaneous switching activity and therefore the noise

- Slew rate control has been proposed to reduce $L\ \partial i/\partial t$ noise by increasing the transition time of the current transients. This technique however is not feasible for advanced technologies due to stringent performance requirements

- Spread spectrum clocking techniques, initially proposed to reduce electromagnetic interference, have been exploited to lower power/ground noise by frequency modulating the clock signal

- Advanced packaging technologies such as flip chip bonding can be exploited to reduce the overall self and mutual inductance, resulting in lower power/ground noise

- Active and passive shield insertion is commonly used to reduce capacitive and inductive coupling by inserting additional isolation between the aggressor and victim lines while providing a closer current return path

- Inserting repeaters along an interconnect reduces coupling noise by decreasing the length of the adjacent interconnect sections

- The quadratic dependence of delay on line length approaches a linear dependence in the presence of inductance, reducing the optimum number of repeaters for an RLC line

- Signal routing and wire ordering algorithms can be exploited to reduce coupling noise while considering delay, area, and power constraints
- Increasing the width of a wire reduces the capacitive coupling by increasing the ground capacitance. The inductive coupling however increases with wire width, making the total coupling noise relatively insensitive to wire width
- Increasing the spacing between the aggressor and victim lines reduces the coupling noise due to lower coupling capacitance. The reduction in inductive coupling however is not as significant due to the long range effect of the inductance
- Decreasing the size of the driver on the aggressor reduces the coupling noise due to a lower drive strength of the aggressor at the expense of increased delay
- Increasing the size of the driver on the victim reduces the coupling noise due to a more effective connection to ground at the expense of increased area

Chapter 3

Switching Noise in Mixed-Signal/Analog Integrated Circuits

The types and effects of switching noise in digital circuits have been reviewed in Chapter 2. In mixed-signal circuits, the switching noise generated by the digital circuit can propagate through various media and affect the analog/RF blocks within the same monolithic substrate, as illustrated in Fig. 3.1. Several functions are integrated

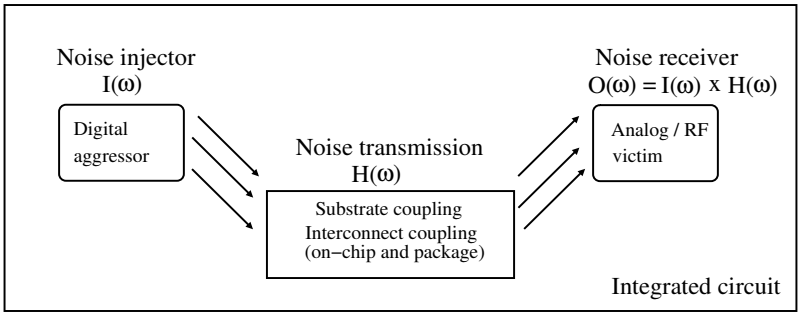


Figure 3.1: Noise is generated by the aggressor digital circuit, transmitted through the substrate and interconnect at the circuit and package levels, and received by the analog/RF victim circuit.

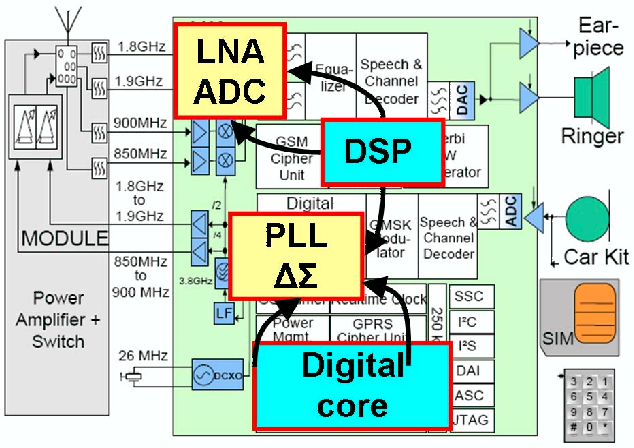


Figure 3.2: Mixed-signal circuit where the noisy digital circuits coexist with the sensitive analog/RF circuits.

on these system-on-chips (SoCs) [113], [114], [115]. Highly sensitive analog/RF circuits such as low noise amplifiers (LNA), voltage controlled oscillators (VCO), analog filters, and high precision analog-to-digital (ADC) and digital-to-analog converters (DAC) coexist with noisy digital circuits within the same integrated circuit, as depicted in Fig. 3.2 [116]. This highly dense integration places stringent signal integrity and isolation constraints on the sensitive analog/RF circuits [117]. The closer physical distance between the digital and analog/RF circuits exacerbates the noise coupling between these circuit blocks.

A common technique to reduce noise coupling is to allocate separate power and ground networks/pads for the digital and analog/RF circuits [118]. Although significant isolation is achieved with dedicated power/ground networks, capacitive and inductive coupling among the interconnects at the circuit and package level remains.

Furthermore, the common substrate behaves as a conductive RC path between the noisy digital circuits and the sensitive analog/RF circuits.

The magnitude of the switching noise is typically much greater than the inherent noise of the transistors caused by shot, thermal, and flicker noise [17]. The efficient analysis of the switching noise coupling and corresponding noise reduction techniques are therefore a primary concern for large scale mixed-signal integrated circuits, such as integrated transceivers [115].

An overview of switching noise in mixed-signal integrated circuits is provided in this chapter. In Section 3.1, the noise coupling mechanisms are reviewed. An overview of substrate coupling is provided in Section 3.2. The effects of substrate coupling noise in mixed-signal circuits are described in Section 3.3. Techniques to alleviate substrate coupling are reviewed in Section 3.4. Finally, the chapter is summarized in Section 3.5.

3.1 Switching Noise Coupling Mechanisms in Mixed-Signal/Analog ICs

In a mixed-signal environment, the switching noise couples to a victim circuit through two primary mechanisms. The first mechanism is a specific case of interconnect coupling, as described in Chapter 2, where the victim is an analog/RF circuit.

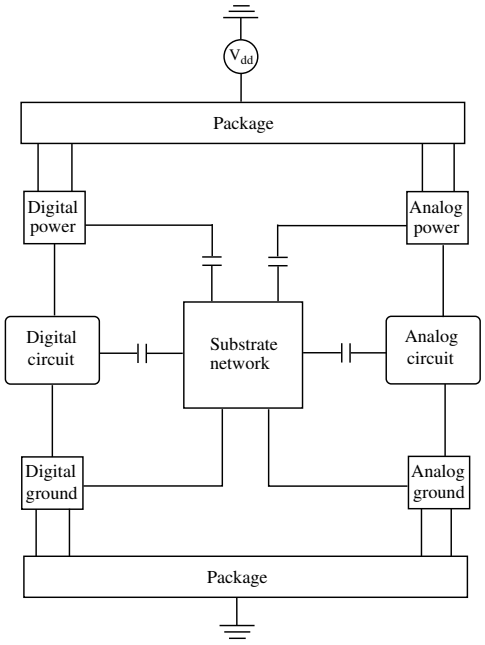


Figure 3.3: Interactions among the digital, analog/RF circuits, and substrate in a mixed-signal environment with dedicated power/ground networks.

The second mechanism is coupling through the monolithic substrate. Interactions among the digital circuit, analog/RF circuit, and the substrate are depicted in Fig. 3.3 for a mixed-signal environment where digital and analog/RF circuits have dedicated power/ground networks.

Interconnect coupling exists both at the integrated circuit and package levels due to coupling capacitance and mutual inductance among the on-chip interconnects and power/ground pads. The conductive substrate is shared by both digital and analog/RF circuits, forming a medium for noise propagation. These two types of mechanisms are described, respectively, in Sections 3.1.1 and 3.1.2.

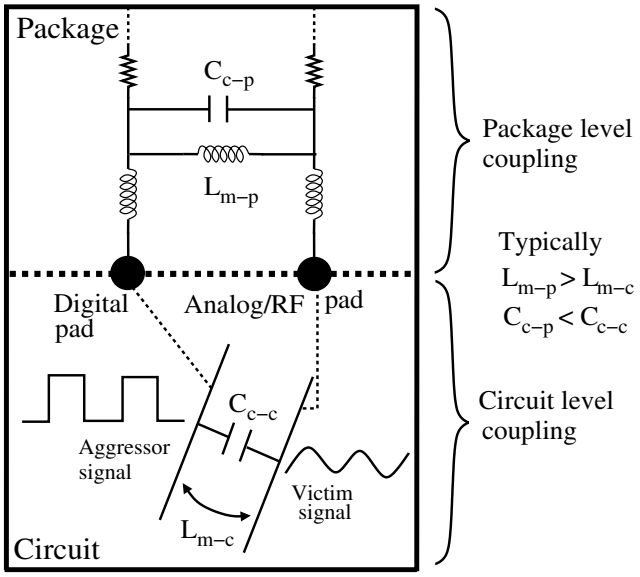


Figure 3.4: Capacitive and inductive coupling at the circuit and package levels. The mutual inductance typically dominates at the package level while the coupling capacitance typically dominates at the circuit level.

3.1.1 Interconnect Coupling

The switching noise generated by an aggressor digital circuit couples into the sensitive analog/RF circuits through capacitive and inductive coupling. At the circuit level, the aggressor is a noisy power/ground line within the digital circuit or a digital signal with a high switching activity factor. At the package level, noise coupling occurs among the digital and analog/RF pads, as illustrated in Fig. 3.4.

Sensitive analog/RF signals are usually placed sufficiently far from the noisy digital circuits such as the output drivers to reduce on-chip noise coupling [119]. Area and floorplan constraints, however, may require the analog/RF circuits be placed physically close to the noisy circuits. Shielding can be used if the additional ground

capacitance can be tolerated by the analog/RF signal. The number and distribution of the power/ground pads also play an important role in reducing the coupling noise. A greater number of power/ground pads and a larger separation between the power/ground pads of the digital and analog/RF circuits reduces the parasitic and mutual inductance. The second and more dominant switching noise coupling mechanism is the substrate coupling, as discussed in the following subsection.

3.1.2 Substrate Coupling

The common substrate forms a medium for noise transmission between the digital and analog/RF circuits in system-on-chips, as shown in Fig. 3.5 for two different types of substrates: (1) a high resistivity, *lightly doped substrate*, also referred to as a *bulk type substrate*, and (2) an *epi type substrate* where a high resistivity layer, referred to as an epi layer, is deposited over a low resistivity substrate [120]. While reducing the risk of latch-up, epi type substrates are typically less noise tolerant as compared to bulk type substrates due to the low resistive bulk [121], as further discussed in the following sections.

The current propagation path within a lightly doped bulk type substrate is illustrated in Fig. 3.5(a). A significant amount of the substrate current flows near the surface due to the lower substrate impedance between the aggressor and victim circuits. Alternatively, for epi type substrates, the substrate current flow is dominant

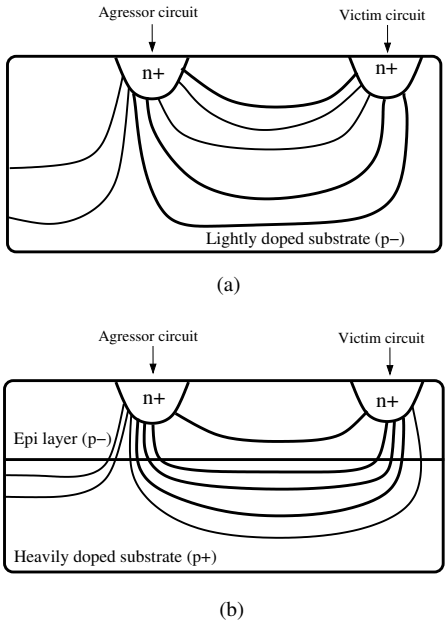


Figure 3.5: Current flow within the substrate: (a) high resistivity bulk type substrate, (b) epi type substrate where a high resistivity epi layer is deposited over a low resistivity substrate.

within the heavily doped bulk since the resistivity of the epi layer is significantly greater than the resistivity of the heavily doped bulk. Note that the type of substrate significantly affects the appropriate noise reduction technique and the model of the current propagation path to properly analyze the substrate noise at the victim circuit, as discussed in the following section.

3.2 Substrate Noise Coupling in Mixed-Signal ICs

An overview of substrate noise coupling in mixed-signal circuits is provided in this section. The primary noise injection mechanisms are described in Section 3.2.1.

Substrate extraction techniques are summarized in Section 3.2.2. Compact models to characterize the substrate impedance are discussed in Section 3.2.3. Finally, high level substrate noise analysis methodologies for large scale circuits are reviewed in Section 3.2.4.

3.2.1 Substrate Noise Injection Mechanisms

Noise couples into the substrate through three primary mechanisms: (1) impact ionization, (2) coupling from the reverse biased source/drain junction capacitances of the transistors during switching, and (3) coupling from the power/ground networks of the aggressor digital circuit [122], [123]. These noise injection mechanisms are described in the following subsections.

Impact Ionization

Impact ionization occurs due to the high electric field within the depletion region of the MOS transistor. Those electrons with sufficient energy create additional electron-hole pairs by a process known as impact ionization [124], [125]. For an NMOS transistor where the gate voltage is positive, the holes created due to impact ionization are collected by the substrate, resulting in current injected into the substrate. The total current I_{impact} produced by the impact ionization process is approximated

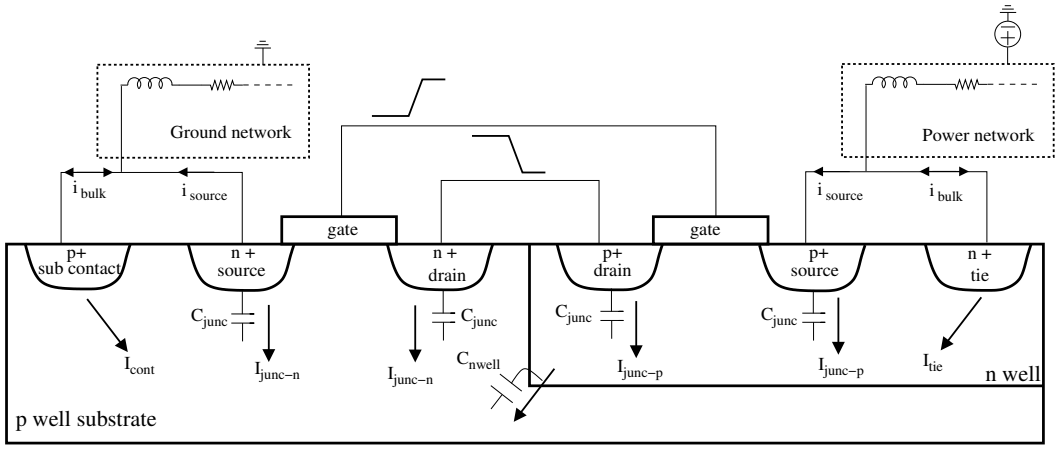


Figure 3.6: Noise injection mechanisms into the substrate for a CMOS inverter. I_{junc-n} and I_{junc-p} represent the source/drain junction coupling for, respectively, the NMOS and PMOS transistors. I_{cont} and I_{tie} represent, respectively, the current injected into the substrate from the substrate contact on the ground network and N-well tie on the power network.

as [123], [124]

$$I_{impact} \approx K_1(V_{DS} - V_{D0})I_{DS}e^{-(K_2/V_{DS}-V_{D0})}, \tag{3.1}$$

where K_1 and K_2 are technology dependent parameters, V_{DS} is the voltage difference between the drain and source, and V_{D0} is the drain saturation voltage. The substrate current generated by impact ionization is negligible as compared to source/drain coupling and power/ground coupling [122], [126].

Capacitive Coupling Through Source/Drain Junctions

Noise coupling from the source/drain junction capacitances and the power/ground networks for a CMOS inverter is depicted in Fig. 3.6. Current is injected into the substrate through the junction capacitances formed by the reverse biased p-n structures,

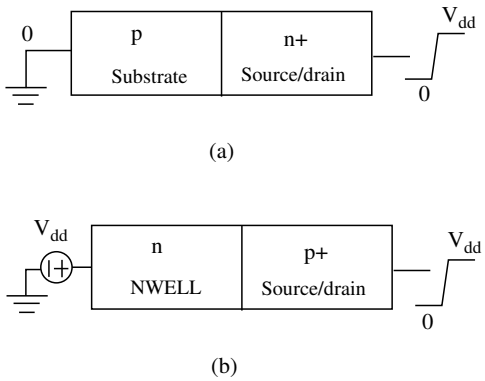


Figure 3.7: Reverse biased pn junctions form a depletion capacitance, thereby injecting noise into the substrate and N-well: (a) NMOS transistor, (b) PMOS transistor.

referred to as source/drain coupling. For an NMOS transistor, the p type substrate (connected to the ground network) and an n type source/drain forms a reverse biased p-n junction, as shown in Fig. 3.7(a). Similarly, for a PMOS transistor, the n type well (N-well) (connected to the power network) and a p type source/drain forms a reverse biased p-n junction, as illustrated in Fig. 3.7(b). The junction (or depletion) capacitance per unit area C_{junc} due to these reverse biased p-n junctions is determined from the parallel plate capacitance expression,

$$C_{junc} = \frac{\epsilon_s}{W}, \tag{3.2}$$

where ϵ_s is the permittivity of silicon and W is the width of the depletion layer. Assuming the full depletion approximation [127], the junction capacitance is

$$C_{junc} = \sqrt{\frac{q\epsilon_s N_A N_D}{2(N_A + N_D)(V_{bi} - V_j)}}, \tag{3.3}$$

where q is the unit electronic charge, N_A and N_D are, respectively, the doping concentrations of the substrate and source/drain regions (for an NMOS transistor), and V_{bi} and V_j are, respectively, the built-in voltage of the depletion region [127] and reverse biased junction voltage.

Coupling from Noisy Power/Ground Networks

The third noise injection mechanism is coupling from the power/ground networks, as illustrated in Fig. 3.6. Typically, the substrate is biased by placing substrate contacts connected to a nonideal ground network. The ground noise couples into the substrate through these substrate contacts. Similarly, the N-well is biased by placing ties (N-well taps) connected to the nonideal power network. The power noise first couples into the N-well through these ties, and finally into the substrate through the N-well capacitance, as illustrated in Fig. 3.6. The N-well behaves as capacitive isolation, reducing coupling of the power noise at low frequencies [114]. Power/ground coupling is typically assumed to be the dominant noise source for the substrate [122], [126] since the switching current is an order of magnitude greater than the capacitive coupling through the source/drain junctions. The dominant noise source however is dependent upon multiple parameters such as the circuit size, parasitic impedances of the power/ground network, decoupling capacitance, and rise time. A detailed discussion of the dominant noise source for substrate coupling is provided in Chapter 5.

3.2.2 Substrate Extraction Techniques

Substrate extraction refers to a process where a distributed RC equivalent circuit is obtained from the substrate volume based on solving electromagnetic differential equations within the substrate. Ampere's law [49], [128] relates the magnetic field around a closed loop to the electric current passing through the loop,

$$\Delta \times H = J + \frac{\partial D}{\partial t}, \tag{3.4}$$

where H, J, and D represent, respectively, the magnetic field intensity, current density, and electric flux density (also referred to as the electric displacement vector). Applying the divergence operation to both sides of (3.4),

$$\Delta \bullet \Delta \times H = \Delta \bullet J + \frac{\partial \Delta \bullet D}{\partial t}. \tag{3.5}$$

Since the divergence of a curl operation is zero,

$$\Delta \bullet J + \Delta \bullet \frac{\partial D}{\partial t} = 0. \tag{3.6}$$

Using the relations $J = (1/\rho)E$ and $D = \epsilon E$, (3.6) is rewritten as

$$\frac{1}{\rho}\Delta \bullet E + \epsilon \frac{\partial}{\partial t}(\Delta \bullet E) = 0, \tag{3.7}$$

where ρ and ϵ represent, respectively, the sheet resistivity and permittivity of the semiconductor, and E is the electric field.

Excluding the diffusion/active areas, the substrate is typically considered as "layers of uniformly doped semiconductor material of varying doping densities" [129], [130]. Neglecting the magnetic fields, the substrate can be represented from (3.7) [129]. Two primary techniques exist to determine the substrate impedances from (3.7): (1) the finite difference method and (2) the boundary element method. These two methods are described in the following subsections.

Finite Difference Method (FDM)

Equation (3.7) can be spatially discretized using the *finite difference method* (FDM) [129], [131] where the substrate is represented as a collection of rectangular prisms (or cuboids), as depicted in Fig. 3.8. A three-dimensional (3-D) mesh consisting of lumped RC impedances is generated to model the substrate. The value of these RC impedances within the rectangular prisms is obtained from Gauss' law [128] and the divergence theorem [132]. According to these theorems, (3.7) reduces to

$$\sum_j [\frac{V_i - V_j}{R_{ij}} + C_{ij}(\frac{\partial V_i}{\partial t} - \frac{\partial V_j}{\partial t})] = 0, \tag{3.8}$$

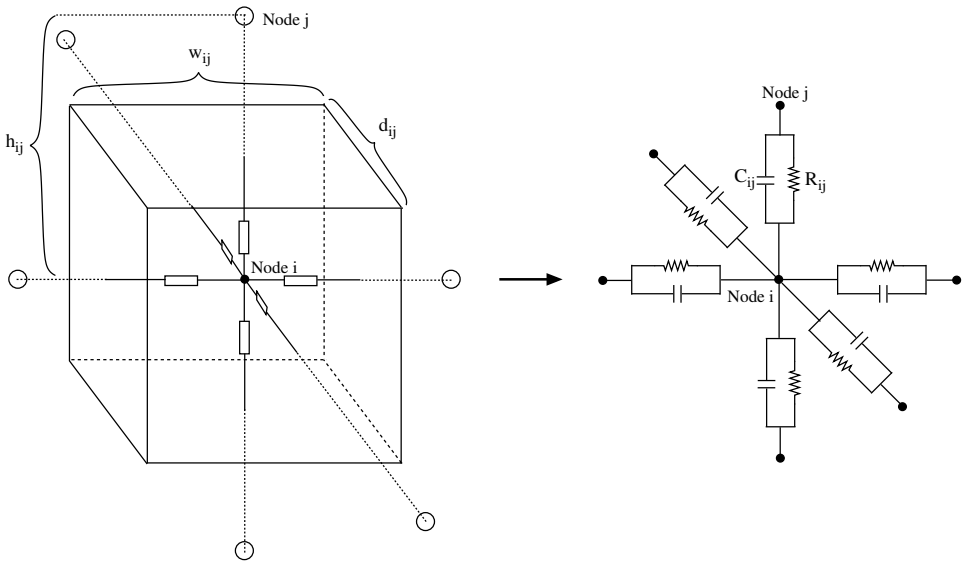


Figure 3.8: Representation of the substrate as a collection of rectangular prisms to discretize (3.7).

where the resistance R_{ij} between node i and j is

$$R_{ij} = \rho \frac{h_{ij}}{w_{ij} d_{ij}}, \tag{3.9}$$

and the capacitance between node i and j is

$$C_{ij} = \epsilon \frac{w_{ij} d_{ij}}{h_{ij}}. \tag{3.10}$$

Although the non-uniformities distributed throughout the substrate can be included using FDM [133], the overall accuracy is highly dependent upon the number of cuboids within the substrate, *i.e.*, the density of grids or the resolution of the discretization process. Typically, fine grids are used for those regions where the gradient of the

electric field is high to achieve sufficient accuracy [129]. Alternatively, coarser grids are used in the remaining regions to reduce the overall computational complexity [129]. The gradient of the electric field is higher in the vertical direction due to the varying doping densities in this direction. A common practice is therefore to use a finer grid in the vertical direction and a coarser grid in the lateral direction [129].

The primary limitation of FDM to discretize the substrate is the increasing number of nodes with larger circuits, making the analysis computationally inefficient. Traditional matrix solutions such as LU factorization [134] is prohibitive for large 3-D mesh networks.

Boundary Element Method (BEM)

An alternative technique to discretize (3.7) in integral form is the *boundary element method* (BEM) [135], [136]. Rather than discretizing the overall substrate, BEM discretizes only specific ports such as the substrate contacts, N-well ties, and the active/diffusion regions, which are essentially two-dimensional (2-D) structures. BEM therefore reduces a 3-D extraction process into a 2-D extraction process [130], significantly improving the computational complexity. The drawback of discretizing only the ports is the inability to consider the non-uniform structures within the substrate, such as channel stop implants.

In the boundary element method, the substrate impedance between two ports is

determined based on *Green's function* [137], which implicitly considers the substrate boundary conditions [130]. Assuming a quasi-static approximation, *i.e.*, the dielectric characteristics of the substrate are negligible over the frequency range of interest [138], [139], (3.7) reduces to the Laplace equation,

$$\Delta^2 \phi = 0, \tag{3.11}$$

where ϕ is the electrostatic potential. Applying Green's Function $G(r, r')$ to (3.11) determines the potential ϕ at a point r in the substrate medium due to a current $J(r')$ injected into another point r', also in the substrate [117],

$$\phi(r) = \int_V J(r')G(r, r')\partial V. \tag{3.12}$$

Once an appropriate Green's function is obtained using mathematical techniques [140], [135], [136], (3.12) is solved by discretizing each port into a set of panels, as depicted in Fig. 3.9. Three ports are shown within the substrate in Fig. 3.9(a). Note that these ports can be source/drain junctions of the transistors, substrate contacts, or N-well ties. Each port is discretized into multiple panels (four in this example), as illustrated in Fig. 3.9(b). The voltage-current relationship among each panel is represented in a matrix form,

$$\phi = Zi, \tag{3.13}$$

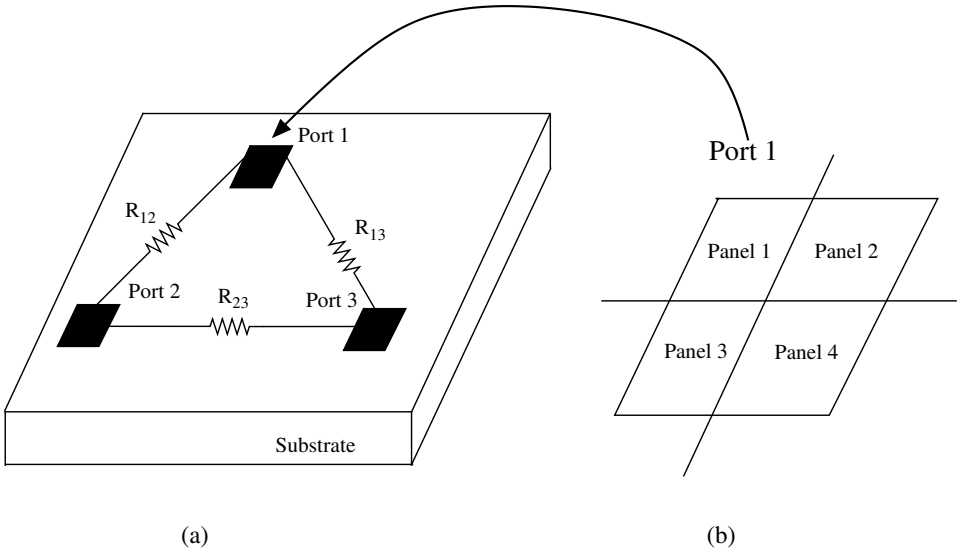


Figure 3.9: Discretization of each port on the substrate into panels in the boundary element method: (a) ports on the substrate and the resistances among the ports, (b) discretization of port 1 into four panels.

where Z is the impedance matrix and each element of the matrix z_{ij} is

$$z_{ij} = \frac{1}{S_i S_j} \int_{S_i} \int_{S_j} G(r, r') \partial a' \partial a, \tag{3.14}$$

where S_i and S_j are the area of, respectively, panels i and j. The impedance matrix Z is inverted to obtain the admittance matrix Y. Finally, the impedance between two ports on the substrate (such as R_{12} in Fig. 3.9(a)) is determined by summing the corresponding elements in the admittance matrix and taking the reciprocal of the result.

Note that for BEM, the size of the resulting matrix characterizing the substrate is significantly smaller, yet highly dense, as compared to FDM where the size of the

matrix is huge, but sparse. Several different techniques have been proposed to obtain a more efficient solution of the algebraic equations produced by FDM or BEM to reduce an RC network. These techniques include moment matching [132], [141], a fast Fourier transform algorithm [142], a fast eigendecomposition technique [133], a numerically stable Green Function [143], and a combination of BEM and FDM techniques [144]. The primary limitation of both of these approaches (FDM and BEM), however, is the increase in computational complexity with the size of the circuit, prohibiting the efficient analysis of large scale mixed-signal circuits. The second class of substrate modeling methods is the use of macromodels, as discussed in the following subsection.

3.2.3 Compact Substrate Models

Compact substrate models refer to analytic expressions characterizing the substrate impedance between two ports using two types of parameters [145]: (1) technology related parameters such as substrate doping profiles and (2) physical geometry related parameters such as size, area, and distance between the ports based on the circuit layout, as depicted in Fig. 3.10. For an epi type substrate where the heavily doped bulk is connected to a grounded backside contact, as shown in Fig. 3.11, the low resistivity bulk is represented as a single equipotential node [120]. This characteristic is exploited to represent the substrate as an impedance matrix including both the self and mutual impedances [145]. The self impedance between the port and back

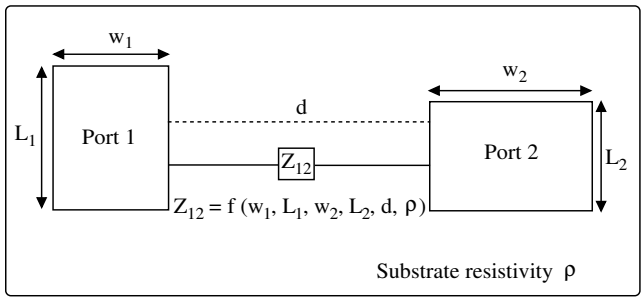


Figure 3.10: Substrate impedance between two ports represented in a compact form as a function of technology and geometry related parameters.

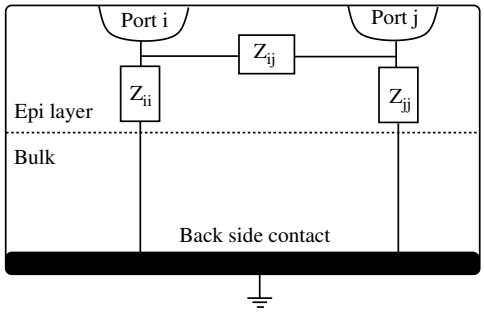


Figure 3.11: Epi type substrate with a grounded backside contact. Z_{ii} and Z_{jj} are the impedances between, respectively, port i and the backside contact, and port j and the back side contact. Z_{ij} is the impedance between the ports.

side contact Z_{ii} is [146]

$$Z_{ii} = \frac{1}{K_1 Area + K_2 Perimeter + K_3}, \tag{3.15}$$

where K_1, K_2, and K_3 are technology dependent fitting parameters. The mutual impedance between two ports is provided for two different regions [145]: (1) *far field region* and (2) *near field region* [145]. The mutual impedance $(Z_{ij})_{ff}$ for the far field region is

$$(Z_{ij})_{ff} = 1/(Z_0 e^{-\beta x}). \tag{3.16}$$

The mutual impedance $(Z_{ij})_{nf}$ for the near field region is

$$(Z_{ij})_{nf} = 1/(Z_{01} e^{-\gamma_1 x_1} + Z_{02} e^{-\gamma_2 x_2}). \tag{3.17}$$

The variation of the coupling strength or the substrate impedance between two ports is dependent upon the relative size of the ports and the separation between the two ports. That is, if the separation is sufficiently large as compared to the port size, the ports are treated as lumped elements, making the spreading effects negligible. In this far field region, the electric field varies in proportion to $1/d$ where d is the distance between the two ports [49]. Alternatively, if the two ports are close, the spreading effects of the current are more significant where the shape, size, and alignment of the ports play an important role in determining the impedance. In this near field

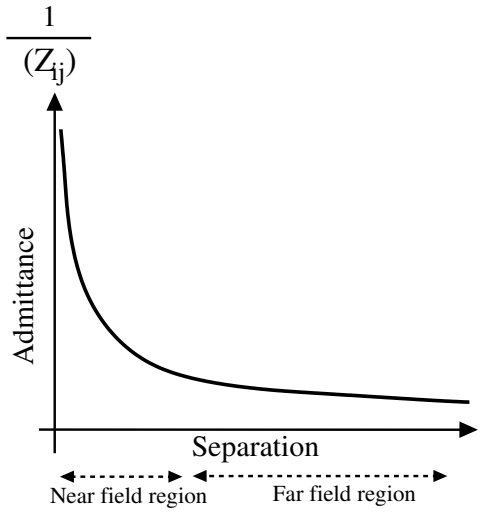


Figure 3.12: Nonlinear variation of mutual substrate admittance $1/(Z_{ij})$ with respect to the separation between the ports.

region, the electric field varies in proportion to $1/d^3$ [49]. This nonlinear behavior of the mutual impedance Z_{ij} with respect to the separation is depicted in Fig. 3.12, as shown in [145].

For a lightly doped substrate where the bulk cannot be represented as a single equipotential node, a resistance based formulation is developed in [145]. If the two ports are sufficiently close, *i.e.*, near field configuration, the substrate current tends to flow close to the surface along the path between the ports. Alternatively, for a far field configuration, the current tends to flow deeper in the substrate. Hence, if the separation increases, the coupling strength initially decreases in an exponential fashion and finally saturates as the separation further increases rather than completely diminishes. Based on this behavior, in lightly doped substrates, the substrate

resistance is expressed as [145]

$$R_{ij} = \beta[ln(d_{ij} + 1]^{\alpha_1}(s_i + s_j)^{\alpha_2}(p_i + p_j)^{\alpha_3}, \tag{3.18}$$

where d_{ij} is the separation between the two ports, s_i and s_j are the area of, respectively, ports i and j, and p_i and p_j are the perimeter of, respectively, ports i and j. The terms β, α_1, α_2, and α_3 are process dependent fitting parameters.

These macromodels characterizing the substrate impedance are computationally more efficient as compared to FDM and BEM which discretize the substrate based on Maxwell's equations. The disadvantages of these macromodels are the requirement to use process-dependent fitting parameters obtained through empirical data and scaling these models for different geometries.

3.2.4 High Level Substrate Noise Analysis

The accurate and efficient estimation of the substrate coupling noise, and functional verification of the circuit in the presence of this noise have become important design issues. Estimating substrate coupling noise in a large scale circuit such as a transceiver is, however, a challenging task since the circuit activity, power/ground network, and substrate network should be simultaneously considered, significantly increasing the computational complexity of the estimation process, as illustrated in Fig. 3.13. An analysis of substrate noise therefore requires not only a model of the

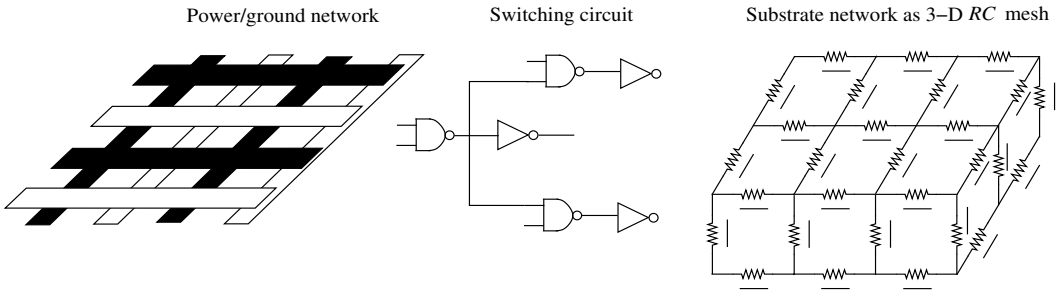


Figure 3.13: Complexity requirements of substrate noise analysis in a large scale circuit. The power/ground networks, switching circuit, and substrate network have to be simultaneously considered.

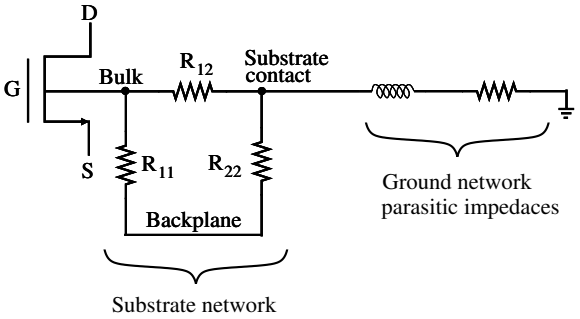


Figure 3.14: Schematic level representation of the substrate impedances to analyze substrate noise.

substrate network, but also the circuit switching activity and the parasitic impedance of the power/ground networks. That is, a high level substrate noise analysis methodology to reduce the computational complexity while achieving reasonable accuracy is a primary concern [147], [148], [149], [150].

A schematic based analysis methodology has been proposed in [149] to reduce the number of elements obtained from the post-layout extraction process. The substrate impedances obtained through compact models using (3.15), (3.16), and (3.17) are back annotated to the schematic, as depicted in Fig. 3.14. The transistor level simulation

of a large scale circuit including the back annotated substrate resistance of every port, however, is not feasible for large scale circuits due to the nonlinear nature of the device models. A common alternative approach is to precharacterize each standard cell in a library with the switching and bulk current profile of each cell [123], [148]. These current profiles are represented as piecewise linear current sources during the substrate noise analysis process to model the aggressor circuit. The nonlinear device models are therefore not considered, significantly decreasing the analysis time. Note that these approaches based on library characterization are only applicable to standard cell based aggressor circuits.

A methodology is proposed in [148] for accurately estimating the switching current drawn by a digital block as a function of the input slew and output load capacitance. Two different techniques are introduced: an input pattern dependent scheme for high accuracy and a pattern independent scheme for high computational efficiency. These current profiles are used within a reduced model of a digital block to analyze the substrate noise, as illustrated in Fig. 3.15. R_m is the average channel resistance of the switching transistors. C_m is the overall gate and interconnect capacitance within the block. C_c and R_c model the decoupling capacitance between the power and ground nodes. C_p is the parasitic capacitance of the power/ground networks. C_{nw} is the N-well capacitance and R_{sub} is the substrate resistance. The proposed methodology in [148] focuses on accurate generation of the current waveforms. Efficiently modeling a

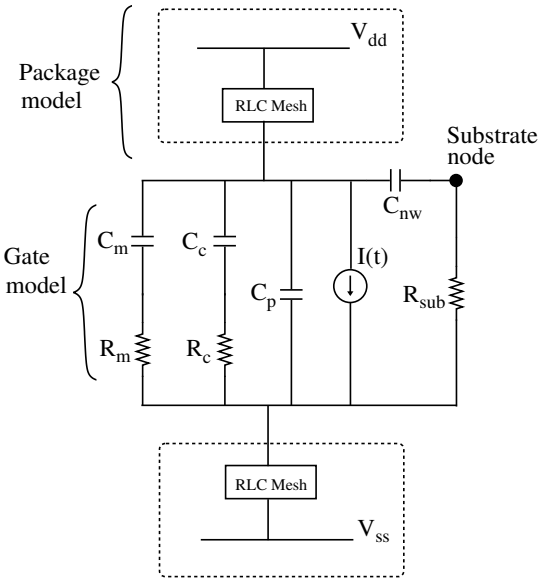


Figure 3.15: Reduced model of a digital block to analyze substrate noise including the parasitic impedance of the package.

substrate network in a large scale circuit, however, remains a primary research issue.

A high-level simulation methodology is described in [147] by generating a macromodel for each standard cell within a circuit. Each macromodel contains two piecewise linear current sources to represent the switching current and bulk current of the gate. These macromodels are connected to the parasitic impedance of the package and substrate network to analyze the substrate noise, as depicted in Fig. 3.16. The approach described in [147], however, is challenging, particularly for bulk type substrates where the substrate cannot be represented by a single equipotential node. Shorting all of the substrate contacts to a single node, as suggested in [147], is not an effective approach for those packaging techniques where the pad inductance is relatively low (such as a flip-chip package), and the on-chip inductance is significant. Furthermore, various

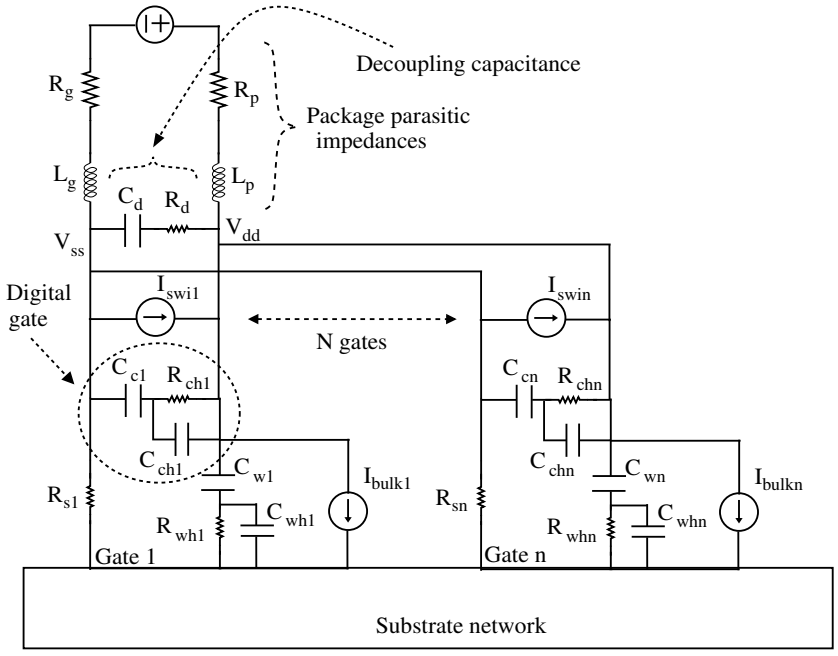


Figure 3.16: High level model to analyze substrate noise where a macromodel is generated for each standard cell.

blocks in a complex SoC exhibit different switching activities, causing on-chip ground bounce.

The accurate and efficient analysis of substrate noise in a large scale mixed-signal circuit is a challenging task, specifically for high resistivity bulk type substrates. The accuracy is highly dependent upon the model of the noise injection mechanism by the aggressor blocks and the model of the noise propagation process within the substrate. Sufficiently accurate models typically exhibit higher computational complexity, making the analysis process prohibitively complex for large scale circuits. A methodology is described in Chapter 6 to efficiently analyze substrate noise generated by a large aggressor block by reducing the number of input ports before the substrate extraction

process is initiated.

3.3 Effects of Substrate Noise in Mixed-Signal ICs

In a mixed-signal environment, substrate noise can affect a victim circuit through two primary mechanisms: (1) variation of the threshold voltage through the body effect and (2) capacitive coupling to the signal and power/ground lines and source/drain nodes of the transistors. The voltage fluctuations in the bulk of the transistors due to substrate noise modify the threshold voltage of these transistors through the *body effect* [127]. Assuming a uniform surface doping concentration, the dependence of the threshold voltage V_t on the bulk potential is [127]

$$V_t = V_{t0} + \frac{\sqrt{2qN_A\epsilon_{si}}}{C_{ox}}(\sqrt{|-2\phi_f + V_{sb}|} - \sqrt{|2\phi_f|}), \tag{3.19}$$

where V_{t0} is the threshold voltage at zero substrate bias, N_A is the substrate doping density, ϵ_{si} is the dielectric permittivity of the substrate, C_{ox} is the gate oxide capacitance per unit area, $2\phi_f$ is the surface inversion potential, and V_{sb} is the voltage difference between the source and bulk.

As described by (3.19), an increase in the body voltage will reduce the threshold voltage of an NMOS transistor. Alternatively, a decrease in the body voltage increases the threshold voltage. A change in the threshold voltage changes the drain current

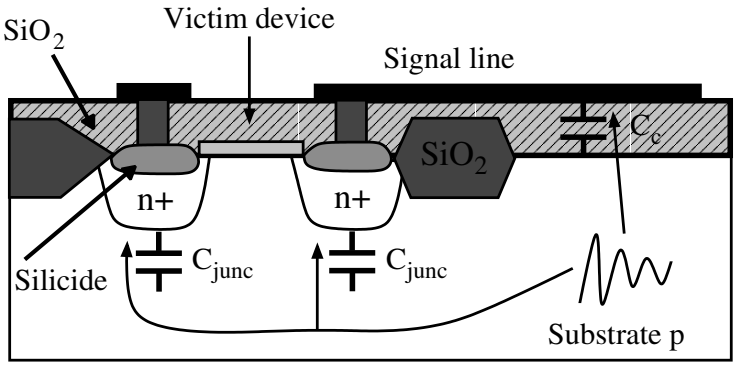


Figure 3.17: Substrate noise capacitively coupling to the signal and power/ground lines and source/drain nodes of the victim transistor.

of the transistor. This change in the drain current results in a mismatch among the transistors, causing significant performance degradation in an analog/RF circuit.

The second mechanism for the substrate to affect the victim circuit is through capacitive coupling, as depicted in Fig. 3.17. The substrate noise couples to the first metal layer through the oxide capacitance, degrading the analog/RF signal and disturbing the analog power/ground voltage. Similarly, the substrate noise couples to the source/drain node through the reverse biased junction capacitance. Note that the capacitive coupling to the victim circuit is relatively dominant at higher frequencies whereas the body effect is more dominant at low to medium frequencies [151].

The effect of the substrate noise on specific performance parameters for various types of analog/RF circuits is described in the following sections. Specifically, the effect of substrate noise on a low noise amplifier is discussed in Section 3.3.1. Performance degradation in a phase locked loop due to substrate noise is described in

Section 3.3.2. Finally, the effect of substrate noise on a sigma-delta modulator is explained in Section 3.3.3.

3.3.1 Effect of Substrate Noise on a Low Noise Amplifier

A low noise amplifier (LNA) is one of the most sensitive building blocks within the receiver circuitry in wireless communication systems. If the substrate noise is sufficiently high, the bias current of the amplifier is modified, degrading the gain of an LNA [152]. A more common situation is to observe the high frequency spectral components of the switching noise at the input and output of the LNA [152], [153].

If the switching events of the digital circuit are periodic, the spectrum of the substrate noise consists of discrete frequency components at integer multiples of the clock frequency [153], referred to as harmonic tones. Furthermore, substrate noise can mix with an RF input signal, producing intermodulation terms with frequency components between the integer multiples of the clock frequency [154], referred to as intermodulation (IM) tones. If the harmonic or intermodulation tones are located within the frequency band of the RF signal, the performance of the LNA is degraded. If these tones are located outside the RF signal band, an appropriate filter can be used to remove the noise from the system, as illustrated in Fig. 3.18 [153].

A typical effect of the in-band noise tones at the output of the LNA is an increase in the bit error rate due to distortion. Furthermore, the mixer receives distorted

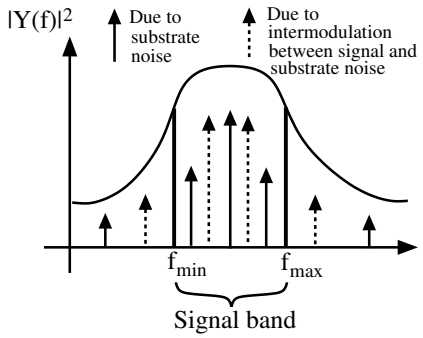


Figure 3.18: Noise tones located in the frequency band of the signal degrade performance. Those noise tones located outside the band can be filtered using an appropriate filter.

information from the LNA, transferring the noise throughout the system [152].

3.3.2 Effect of Substrate Noise on a Phase Locked Loop

A phase locked loop (PLL) is a common circuit used in integrated circuits for various functions. In high speed digital circuits, such as microprocessors and wireless communications, a PLL is used as a frequency synthesizer to generate an on-chip clock signal and to synchronize this clock with the system clock. In data communications, such as serial links, a PLL is utilized to recover the clock signal from the incoming distorted data signal [155]. In optical communications, a PLL is used within a clock and data recovery circuit to regenerate the clock signal and resample the data signal by using the recovered clock.

In these applications, clock jitter (or phase noise in the frequency domain), defined as "the random temporal variation of the phase" [156] is one of the most significant

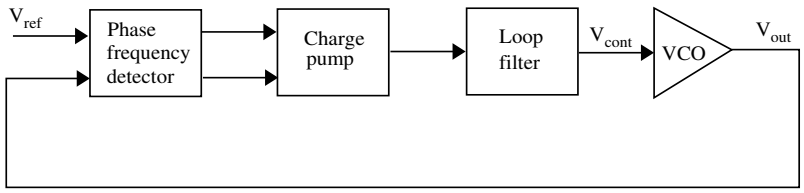


Figure 3.19: Architecture of a charge pump PLL consisting of a phase frequency detector, charge pump, loop filter, and a voltage controlled oscillator (VCO).

performance parameters. Various mechanisms contribute to the phase noise and jitter of a PLL such as device noise (thermal, flicker, and shot noise), device mismatch, and switching noise from digital circuits on the same die. The amount of jitter due to device noise and mismatch is typically negligible as compared to the jitter due to the switching noise, *i.e.*, power/ground and substrate noise [157].

A widely used type of PLL architecture is a charge pump PLL due to the greater acquisition range of this circuit [13]. The fundamental building blocks comprising a charge pump PLL are the phase and frequency detector (PFD), charge pump, loop filter, and a voltage controlled oscillator (VCO), as illustrated in Fig. 3.19. A PLL is a closed loop architecture, exhibiting a low pass filter characteristic from input V_{ref} to the output V_{out}, as illustrated in Fig. 3.20(a). Note that the input noise transfer function is the same as the noise transfer function. The high frequency input noise components are therefore suppressed due to the low pass characteristic of the PLL. Alternatively, a high pass filter characteristic is observed from the input of the voltage controlled oscillator (VCO) to the output of the PLL, as illustrated in Fig. 3.20(b) [13]. Fast jitter components generated within the VCO therefore propagate to the output,

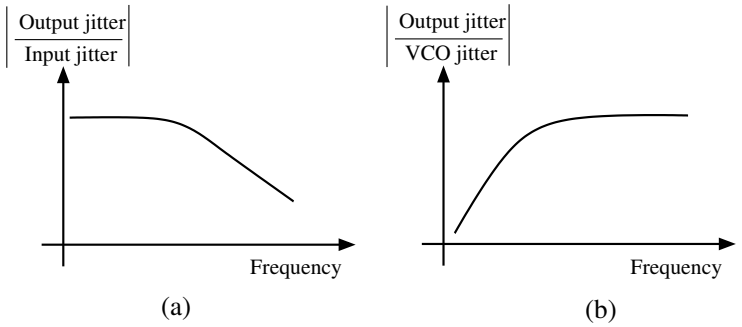


Figure 3.20: Jitter transfer function within a PLL: (a) PLL output to PLL input, (b) PLL output to VCO input.

making the VCO a critical block in terms of jitter generation and propagation.

In a ring VCO, substrate noise introduces jitter by modifying the voltage dependent drain-to-bulk junction capacitances of the transistors in each delay stage [157]. Another mechanism is the direct coupling to the control voltage of the VCO, changing the oscillation frequency [156]. For an LC-VCO, substrate noise may capacitively couple to the inductor, generating spurious frequency components at the output of the VCO [158].

A 40 dB increase has been observed in the phase noise of a ring VCO manufactured in an epi type substrate due to the substrate noise [156]. An increase in jitter from 100 ps to 500 ps has also been reported when the digital circuit surrounding the PLL is activated [159].

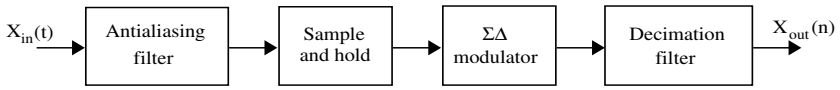


Figure 3.21: Analog to digital converter with a ΣΔ modulator to enhance SNR

3.3.3 Effect of Substrate Noise on a Sigma-Delta Data Converter

Analog-to-digital and digital-to-analog data converters are extensively utilized in mixed-signal integrated circuits. Two primary types of data converters are Nyquist-rate and oversampling converters. Oversampling converters achieve higher signal-to-noise ratio (SNR) and exhibit several other advantages over Nyquist-rate converters such as lower in-band noise characteristics [16]. Oversampling converters operate at a much greater speed than a Nyquist-rate converter and typically use noise shaping by utilizing sigma-delta (ΣΔ) modulators to further enhance SNR. An analog to digital converter with a ΣΔ modulator is illustrated in Fig. 3.21.

Enhanced noise characteristics achieved by noise shaping significantly relax the design specifications of the analog blocks [16].

ΣΔ modulators are typically implemented with switched capacitor circuitry [16] which requires a clock signal to perform the sample-and-hold operation. This clock signal suffers from jitter due to substrate noise, producing "sampling uncertainties" [160]. Up to a 40 dB decrease in the SNR of the ΣΔ modulator due to substrate noise has been demonstrated [160].

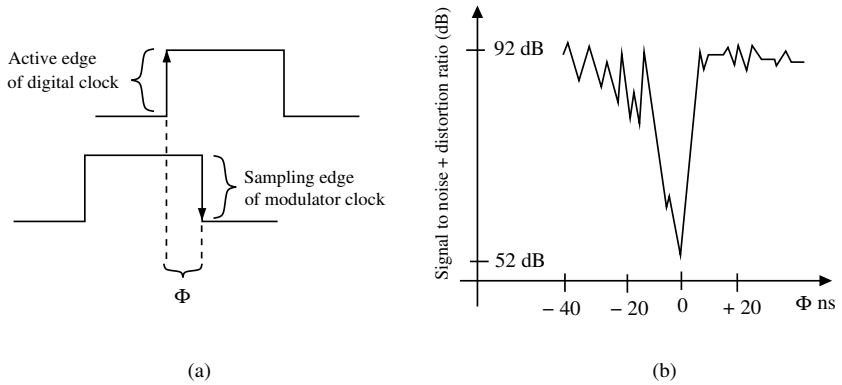


Figure 3.22: Sensitivity of the SNDR to the time difference Φ between the active edge of the digital clock and the sampling edge of the modulator clock: (a) digital clock and sampling clock of the modulator, (b) variation of SNDR with respect to Φ.

The sensitivity of the signal-to-noise and distortion ratio (SNDR) to the time difference between the clock of the aggressor digital circuit and the sampling clock of the $\Sigma\Delta$ modulator has also been investigated [161]. Significant degradation in the SNDR has been observed if the active edge of the digital clock coincides with the sampling edge of the modulator clock, as depicted in Fig. 3.22 [161]. As illustrated in Fig. 3.22(b), if $\Phi = 0$, the data is sampled by the modulator at the same time as the rising edge of the digital clock signal. SNDR decreases from approximately 92 dB to 52 dB, corresponding to a significant performance degradation. The timing of the sampling process with respect to the digital clock has a critical effect on the overall performance of the data converter.

3.4 Techniques to Reduce Substrate Noise in Mixed-Signal/Analog ICs

A typical design practice in mixed-signal circuits is to provide separate power and ground pads, and distribution networks for the digital and sensitive analog/RF circuits to prevent simultaneous switching noise from directly coupling into the power and ground nodes of the sensitive circuits. The common substrate medium, however, forms a conductive path between the aggressor and sensitive circuits, allowing noise transmission.

Various methodologies exist to suppress substrate noise coupling in mixed-signal integrated circuits. These techniques can be broadly classified under three primary categories [117].

- The first category focuses on reducing the magnitude of the switching noise. These techniques include the use of decoupling capacitors between the power and ground networks of the aggressor circuit, skew and slew rate control, spread spectrum clock generation, shielding to reduce mutual inductance, packaging techniques with low impedance characteristics, and asynchronous circuit design. Note that these techniques have already been described in Chapter 2.

- The second category reduces the substrate noise by modifying the noise transfer medium between the aggressor and victim circuits with the use of guard

rings, deep N-well (or triple well) isolation, higher resistivity substrate, greater distance between the aggressor and victim circuits, and/or silicon-on-insulator (SOI) technology.

- Finally, the third category reduces the effect of the substrate noise by minimizing the sensitivity of the victim circuits to noise with techniques such as differential signaling.

Physical separation of the aggressor and victim circuits is described in Section 3.4.1. The use of guard rings and deep N-well isolation is explained, respectively, in Sections 3.4.2 and 3.4.3. The advantages of silicon-on-insulator technology in reducing substrate noise are discussed in Section 3.4.4. Finally, the use of differential signaling is described in Section 3.4.5.

3.4.1 Physical Separation

The amount of noise current flowing from the aggressor towards the victim is a function of the substrate impedance between these two points. The variation of this impedance as a function of the physical separation is strongly dependent upon the type of substrate.

As depicted in Fig. 3.5(b), for epi type substrates, most of the substrate current flows through the low resistance bulk. Alternatively, lightly doped (high resistivity) substrates have a more uniform current flow, as illustrated in Fig. 3.5(a). Hence,

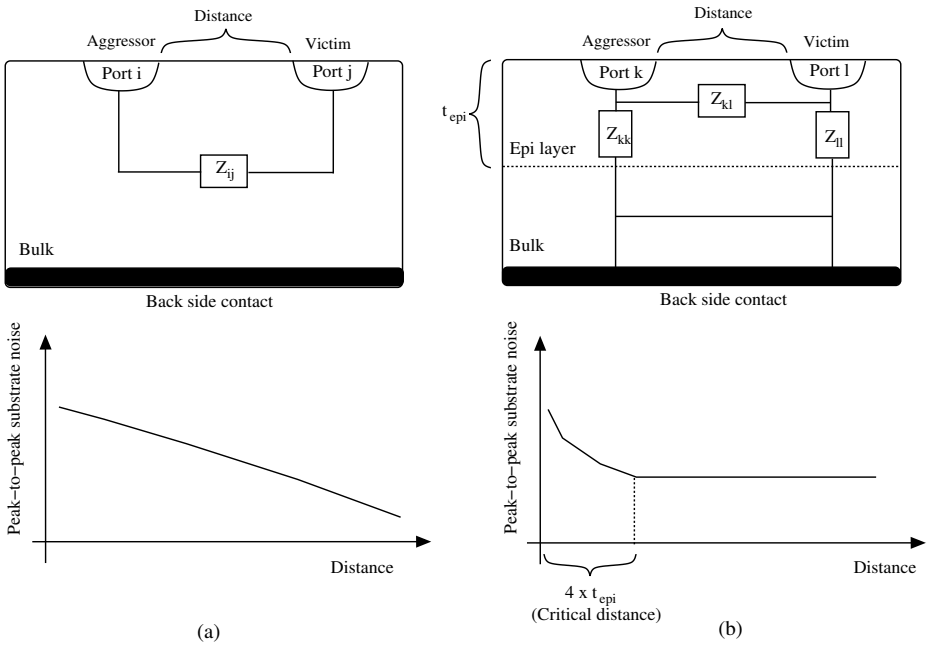


Figure 3.23: The effect of physical separation on the substrate noise characteristics: (a) for bulk type high resistivity substrates, (b) for epi type low resistivity substrates.

the efficiency of increasing the physical distance between the aggressor and victim is dependent upon the type of substrate. The variation of the peak-to-peak substrate noise as a function of distance is shown in Fig. 3.23 for both epi and bulk type substrates [120]. In bulk type substrates, the noise is reduced approximately linearly as the distance between the aggressor and victim is increased due to the reduction of the substrate impedance Z_{ij}, as shown in Fig. 3.23(a). Alternatively, in an epi type substrate, the noise initially decreases until a critical distance is reached. Any further increase in the distance does not affect the noise, as depicted in Fig. 3.23(b), due to the vertical current propagation through impedances Z_{kk} and Z_{ll} rather than the lateral impedance Z_{kl}. The critical distance beyond which the noise remains constant

is shown to be approximately 4 x t_{epi} where t_{epi} is the thickness of the epitaxial layer [120]. Note that this critical distance can vary between 2.5 x t_{epi} and 5 x t_{epi}, depending upon the impedance of the backside contact [136]. For epi type substrates, therefore, increasing the physical distance between the aggressor and victim is not an efficient isolation strategy.

3.4.2 Guard Rings

A guard ring refers to the p+ substrate contacts (or n+ taps for N-well) placed around the aggressor or victim and connected to a ground network (or power network for the N-well), as illustrated in Fig. 3.24. The purpose of a guard ring is to provide 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 an aggressor, victim, or both the aggressor and victim.

A guard ring on the p- substrate should be biased with a dedicated ground network and ground pad rather than the circuit ground to prevent additional noise injection due to ground bounce [120], [159], [162]. Assuming both analog and digital grounds exist within a circuit, biasing the ring with the digital ground causes additional noise injection whereas biasing the ring with the analog ground directly couples the substrate noise into the sensitive analog ground. A dedicated ground network should therefore be used to bias the rings. Increased noise has been observed if the guard

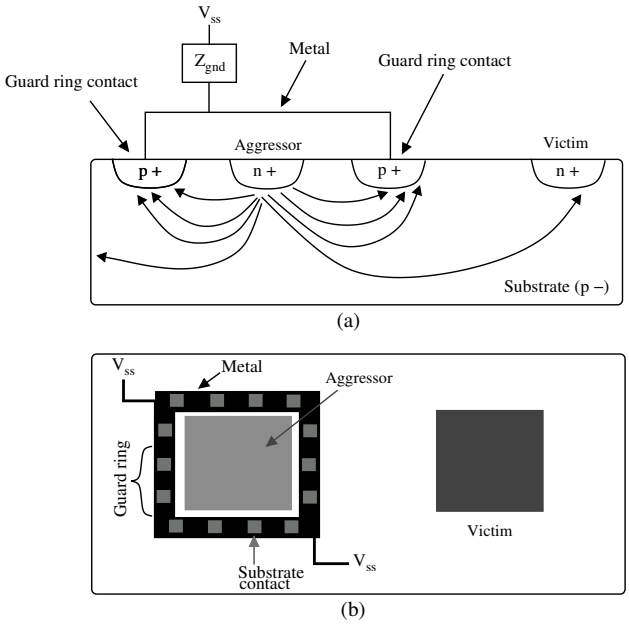


Figure 3.24: The use of guard rings around an aggressor circuit to provide a low impedance path for the injected noise current within the substrate: (a) Cross-sectional view, (b) Top view.

ring is biased with a noisy ground network [120]. Note that a detailed discussion on existing substrate biasing techniques and a biasing methodology to reduce substrate noise is provided in Chapter 7.

The efficiency of a guard ring is dependent upon several parameters. The proximity of the ring to the aggressor or victim plays an important role in both epi and bulk type substrates by determining the amount of noise current absorbed by the ring. Placing the ring as close as possible to the aggressor minimizes the vertical current paths that bypass the ring, thereby lowering the noise. Similarly, as the distance between the ring and victim decreases, the bulk of the victim device is better isolated from the noisy substrate. The variation of the substrate noise as a function

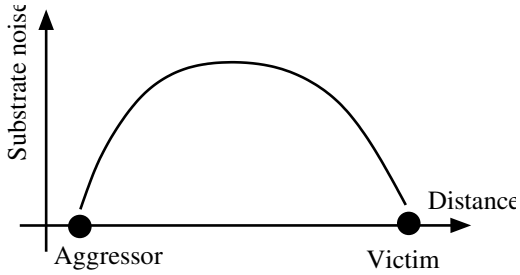


Figure 3.25: Variation of the substrate noise as a function of the location of the guard ring.

of the location of the guard ring is illustrated in Fig. 3.25 [159]. Note that if a ring is placed close to a sensitive victim, the parasitic impedances of the ring, *i.e.*, the resistance and inductance of the on-chip ground network used to bias the ring, and the inductance and capacitance (both self and mutual) of the bonding pad of this ground network, should be sufficiently small for the ring to be effective and to not inject additional noise.

Similar to the physical separation, the efficiency of the guard rings is also dependent upon the type of substrate. For epi type substrates, a modest reduction in noise is achieved due to the existence of a low impedance bulk. Only a 30% reduction has been demonstrated using guard rings on an epi type substrate [120] since the dominant current propagation path occurs within the low resistance bulk, bypassing the ring. Alternatively, for bulk type substrates, the p+ guard rings are more effective due to the more uniform current flow. A reduction of the substrate noise by an order of magnitude has been demonstrated in bulk type substrates [120].

The width of the guard ring is also an important parameter for enhancing the

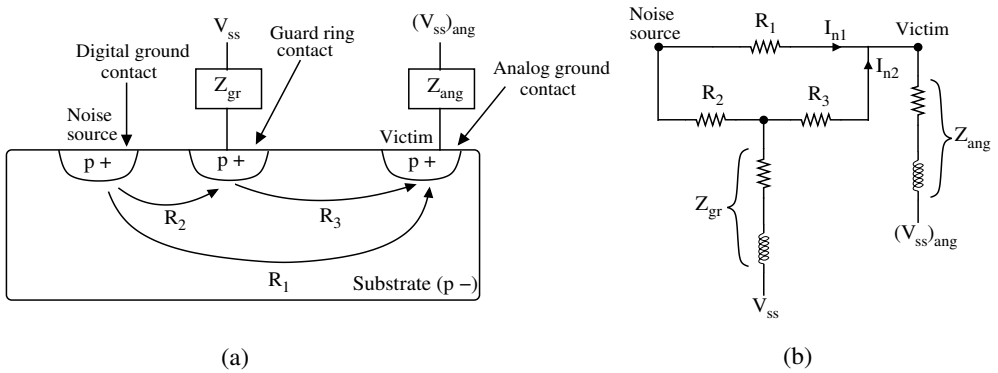


Figure 3.26: Noise source, guard ring, and victim: (a) physical representation, (b) circuit representation.

overall efficiency of the ring. Typically, a greater width, *i.e.*, larger substrate contacts and a wider metal line, reduces the substrate impedance between the ring and noise sources on the substrate, thereby enhancing the noise efficiency of the ring. The isolation achieved by the ring, however, remains constant after a certain width beyond which the parasitic impedance of the guard ring starts to dominate the overall substrate noise [158]. Referring to Fig. 3.26, two paths exist for the noise current to reach the victim. The first path I_{n1} is through R_1 and the second path I_{n2} is through R_2 and R_3. An increase in the width of the ring causes a greater current to flow through R_2. At a certain width, as determined by the operating frequency, I_{n2} becomes equal to I_{n1} due to the impedance of the guard ring Z_{gr}. A further increase in width does not lower the overall noise, as depicted in Fig. 3.27 [158]. Unnecessarily wide guard rings, therefore, consume additional area without improving the substrate noise characteristics.

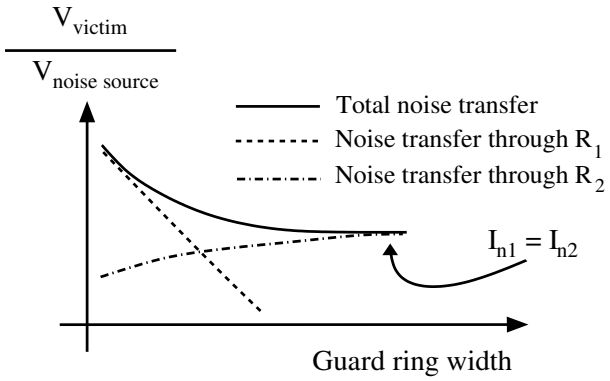


Figure 3.27: Variation of noise transfer from aggressor to victim as a function of the guard ring width. R_1, R_2, I_{n1}, and I_{n2} are shown in Fig. 3.26.

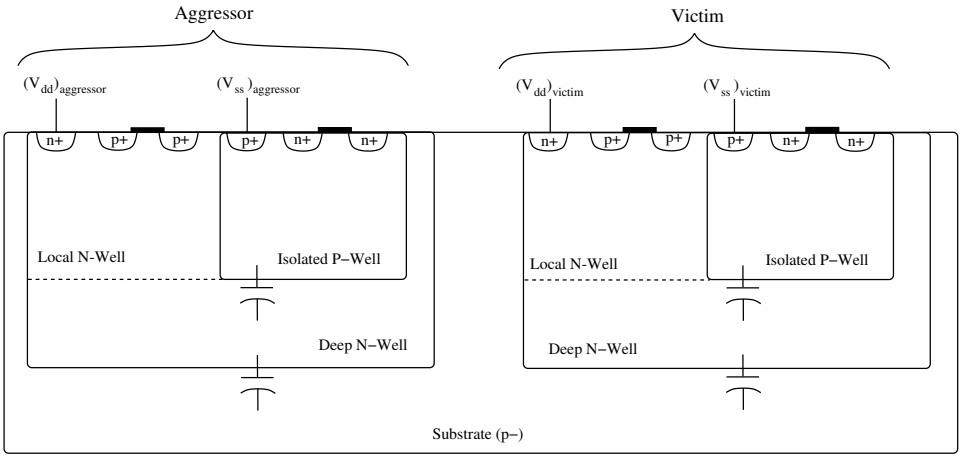


Figure 3.28: Use of deep N-well to capacitively isolate NMOS transistors from the noisy substrate.

3.4.3 Deep N-Well Isolation

A deep N-well or triple well isolation is a process dependent technique to reduce substrate noise coupling. The NMOS transistors are capacitively isolated from the noisy substrate by implanting a deep N-well between the p- substrate and local (or isolated) P-well, as depicted in Fig 3.28. Note that both the aggressor and victim are

located within different N-well regions which are separated by the p- substrate [117]. Alternatively, a deep N-well can be placed only under the aggressor or only under the victim. Isolation is achieved by the reverse biased junction capacitances formed between the isolated P-well and N-well, and between the N-well and p- substrate, as illustrated in Fig. 3.28. Deep N-well isolation is therefore more effective at low and medium frequencies where the impedance of the junction capacitances is relatively high.

A recommended technique to increase the efficiency of a deep N-well is to divide a large deep N-well into smaller sections, thereby reducing the junction capacitances [163]. For example, rather than placing the deep N-well below several victim blocks, a greater reduction is achieved if a separate deep N-well is placed below each victim block.

Typically, a deep N-well placed below the aggressor or victim is connected to the local N-well, where the PMOS transistors are formed, as depicted in Fig 3.28. Due to this connection, the power noise coupled into the local N-well couples into the p-substrate through a larger junction capacitance, degrading the ability to isolate the PMOS transistors. Similarly, the noise present in the p- substrate couples into the local N-well through a larger junction capacitance as compared to the case where no deep N-well is present. Dividing a deep N-well into smaller sections is therefore crucial in reducing the junction capacitances, thereby preventing the deleterious effect of a

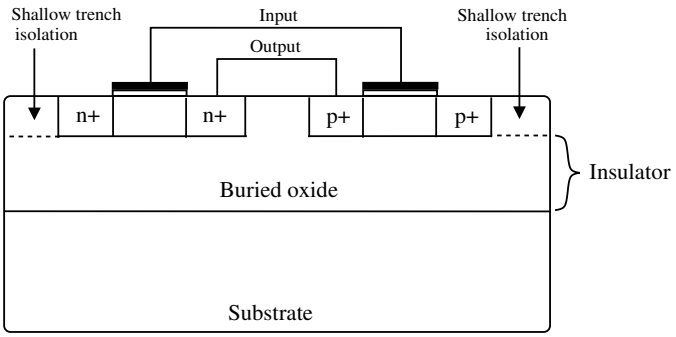


Figure 3.29: Cross section of a CMOS inverter built on SOI technology.

deep N-well on PMOS transistors.

3.4.4 Silicon-on-Insulator (SOI) Technology

Silicon-on-insulator (SOI) [164] refers to a technology where an insulating substrate is used for the transistors rather than the silicon. The devices are fabricated within a silicon layer formed over an insulating film, as depicted in Fig. 3.29 for a CMOS inverter where the insulator is silicon dioxide. Historically, SOI technology has been developed due to the requirement for radiation tolerant circuitry [165]. Specifically, memory circuits built in SOI are more resistant to cosmic rays and radioactive material [165]. Later, SOI technology has been extended to general purpose CMOS circuits due to speed and power advantages [164], [166]. Elimination of the junction capacitances of the devices and the absence of the body effect in stacked transistors have enabled SOI technology to achieve higher performance under the same power constraints [167].

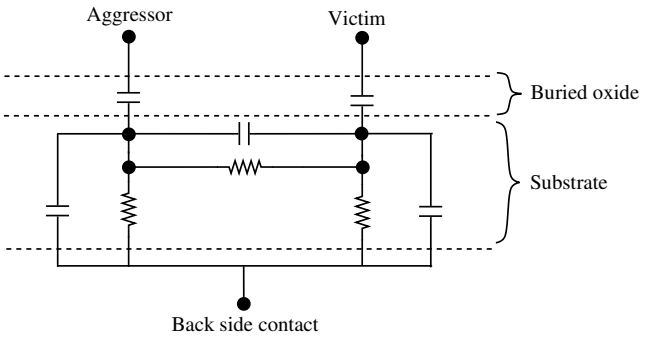


Figure 3.30: Equivalent circuit illustrating substrate noise coupling in SOI technology.

In SOI technology, coupling from the silicon layer to the substrate is significantly reduced due to the insulating oxide layer located between the silicon layer and the substrate [168], [169]. An equivalent circuit illustrating substrate noise coupling in SOI technology is shown in Fig. 3.30 [170]. The benefits of SOI technology in reducing the substrate noise have been investigated and compared with conventional CMOS [170]. At low frequencies, SOI technology with a low resistivity substrate and a floating back side contact achieves 40 dB and 25 dB reduction as compared to, respectively, an epi type silicon substrate and a high resistivity silicon substrate. If the back side of the SOI wafer is grounded, the reduction improves to, respectively, 55 dB and 40 dB. Above 300 MHz, however, the noise behavior of SOI technology and conventional CMOS becomes similar due to the lower impedance of the buried oxide. Alternatively, for a high resistivity SOI technology, the noise advantage is maintained up to 10 GHz, making high resistivity SOI a lower noise technology [170]. Note that the frequency where the crossover occurs is highly dependent upon the thickness of

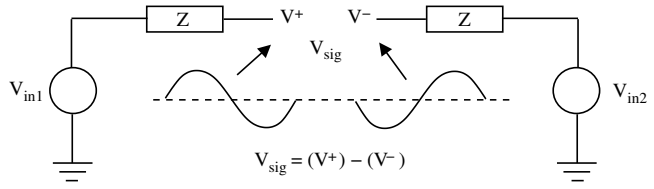


Figure 3.31: Differential signaling where the voltage refers to the difference between two out-of-phase signals rather than with reference to ground.

the buried oxide layer.

3.4.5 Differential Signaling

Differential signaling is a technique to reduce the sensitivity of analog/RF circuits to switching noise by increasing the common mode rejection ratio (CMRR) and power supply rejection ratio (PSRR).

A voltage essentially refers to a *difference* of the two values rather than an absolute value. In single ended circuits, the voltage of an input signal is defined with respect to the reference ground which can either be on-chip or off-chip [162]. The noise coupled to the input signal and the noise variations on the ground node of the transmitting or receiving circuit is not rejected. Alternatively, in differential circuits, the voltage of an input signal refers to the difference between two complementary signals that are out-of-phase, as illustrated in Fig. 3.31 [13]. Differential circuits exhibit higher immunity to switching noise such as from the power/ground and substrate due to greater CMRR. In differential circuits, CMRR is highly dependent upon the symmetry of the layout of the analog circuits [115]. The noise behaves as common mode only if

both pairs within the differential circuit are fully symmetric.

The efficiency of differential signaling in suppressing the substrate noise is also dependent upon the type of substrate. Low resistivity, epi type substrates provide increased differential mode noise isolation as compared to high resistivity bulk type substrates [136]. This behavior is due to the vertical current propagation within epi type substrates, reducing the effect of the specific location of the contacts, and thereby achieving enhanced symmetry. Alternatively, for bulk type substrates, the specific location of the contacts is important due to the lateral current flow, and thereby achieving symmetry is more difficult. Also note that the common mode substrate noise coupled to a device can mix with the input signal, resulting in differential mode intermodulation noise which cannot be rejected by differential signaling [154].

3.5 Chapter Summary

An overview of switching noise in mixed-signal integrated circuits has been presented in this chapter with a primary focus on substrate coupling. The characteristics of substrate coupling in mixed-signal integrated circuits can be summarized as follows:

- Superior performance and lower cost are the two primary driving factors for highly integrated system-on-chips where sensitive analog/RF blocks coexist on the same die with aggressor digital blocks

- Switching noise couples to sensitive blocks through two primary mechanisms: (1) capacitive and inductive coupling among the interconnects at the circuit and package level, and (2) coupling through the monolithic substrate
- The magnitude of the switching noise is typically much greater than the inherent noise of the transistors such as shot, thermal, and flicker noise
- The type of substrate significantly affects the model of the current propagation path and appropriate noise reduction techniques
- In a lightly doped (high resistivity) substrate, a significant amount of substrate current flows in the lateral direction near the surface due to the lower substrate impedance at the surface as compared to the deeper regions within the substrate
- In a heavily doped (low resistivity) substrate with an epi layer above, a significant amount of the substrate current flows vertically to the lower impedance bulk
- Noise is injected into the substrate through three primary mechanisms: (1) impact ionization, (2) coupling from the source/drain junction capacitances, and (3) coupling from the power/ground networks of the aggressor digital circuit
- A distributed equivalent RC circuit model is obtained from extracting the substrate by solving electromagnetic differential equations

- The finite difference method (FDM) and boundary element method (BEM) are two primary numerical techniques to extract the substrate volume
- The primary limitation of both FDM and BEM is the increase in computational complexity with the size of the circuit, prohibiting the efficient analysis of large scale circuits
- Compact substrate models are also used to characterize the substrate impedance between two ports as a function of technology and geometric related parameters
- These macromodels are computationally more efficient as compared to FDM and BEM, but require process-dependent fitting parameters. Furthermore, some accuracy is lost while scaling these models for different geometries
- Accurate and efficient estimation of the substrate coupling noise, and functional verification of the circuit in the presence of this noise are important design issues
- Estimating substrate coupling noise in a large scale circuit is a challenging task since the circuit activity, power/ground network, and substrate network should be simultaneously considered
- In an epi type substrate, the low resistive bulk is represented as a single node, making a model of the substrate network significantly less complicated
- In high resistivity substrates, the bulk cannot be represented as a single node,

complicating the model of the substrate network

- Substrate noise affects the victim through two primary mechanisms: (1) variation of the threshold voltage through the body effect and (2) direct capacitive coupling to the signal and power/ground lines, and the source/drain nodes of the transistors

- The bit error rate of a low noise amplifier (LNA) increases due to substrate noise if the harmonic or intermodulation tones of the noise are located within the RF signal band

- In a phase locked loop (PLL) circuit, the jitter caused by the power/ground and substrate noise dominates the jitter caused by device noise and mismatch

- In a PLL, the substrate noise introduces jitter by changing the voltage dependent drain capacitances of the transistors within the voltage controlled oscillator (VCO) and by capacitively coupling to the control voltage of the VCO

- The signal-to-noise ratio (SNR) of a sigma-delta data converter is significantly degraded due to substrate noise if the active edge of the digital clock coincides with the sampling clock of the sigma-delta converter

- Techniques that reduce switching noise, *i.e.*, power/ground noise and interconnect noise, as presented in Chapter 2, also reduce substrate noise

- Increasing the physical distance between the aggressor and victim is an efficient technique to reduce substrate noise for high resistivity substrates since the current flows more uniformly in high resistivity substrates
- Increasing the physical distance between the aggressor and victim is inefficient to reduce substrate noise for epi type substrates due to the vertical current flow
- Guard rings should be biased with a dedicated ground network and ground pad rather than the circuit ground to be effective in reducing substrate noise
- Guard rings achieve significant isolation in high resistivity substrates due to the lateral current flow. Alternatively, guard rings achieve only a modest noise reduction in epi type substrates due to the vertical current propagation path, bypassing the ring
- Increasing the width of the ring enhances isolation until a specific width beyond which any further increase in width consumes additional area without lowering the substrate noise
- A deep N-well or triple well achieves noise isolation by placing the NMOS devices within a local P-well surrounded by a deep N-well
- The efficiency of a deep N-well decreases with higher frequency due to the lower impedance of the N-well capacitances

- Dividing a large N-well into smaller sections improves the efficiency of the noise isolation due to lower N-well capacitances
- Noise isolation of silicon-on-insulator technology is superior as compared to conventional bulk type substrates due to the buried oxide layer between the silicon layer and the substrate
- Differential analog circuits achieve higher common mode and power supply noise rejection, reducing the sensitivity of the circuit to common mode noise
- Differential circuits on an epi type substrate achieve enhanced common mode noise rejection as compared to a high resistivity substrate due to the current propagation paths, resulting in enhanced symmetry of the contact locations within the epi type substrate
- Common mode substrate noise can mix with the input signal, generating differential mode intermodulation noise which cannot be rejected by differential signaling

Chapter 4

Equivalent Transition Time for Resonant Frequency

In a power distribution network, the decoupling capacitors are often used to reduce power/ground noise by temporarily providing charge to the load circuits during switching events, as described in Chapter 2. Several factors such as the placement, size, and recharge time of the decoupling capacitors should be considered for efficiency [171], [172], [173]. Another important consideration for decoupling capacitors is the LC resonance. A decoupling capacitor C_d forms an LC tank circuit with the parasitic inductance L_g of the interconnect. The impedance of this LC circuit is maximum at the resonant frequency $f_{res} = 1/(2\pi\sqrt{L_g C_d})$.

The impedance characteristics of a power distribution system have been investigated with particular focus on the resonant behavior [174], [175]. The corresponding transition time that produces the maximum noise in the time domain, however, has not received much attention. Faster signal transitions (smaller transition times) are

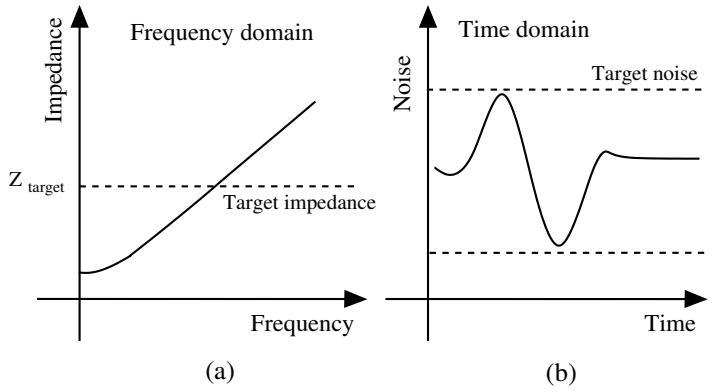


Figure 4.1: Target impedance and noise: (a) target impedance is not satisfied in the frequency domain, (b) although the target noise is satisfied in the time domain.

typically assumed to produce the worst case noise. This assumption, however, is not valid for an inductive power distribution network with a decoupling capacitor.

Although computationally more efficient, a frequency domain analysis is typically pessimistic as compared to a time domain analysis [176] since the frequency range over which the target impedance should be satisfied is not well defined. As illustrated in Fig. 4.1 [176], the peak-to-peak noise can be within the tolerable range although the target impedance is not satisfied. A frequency domain analysis therefore may cause overdesign of the power/ground network [176]. Alternatively, a time domain analysis produces an estimate of the power/ground noise that is less pessimistic.

A primary issue in time domain analysis is the difficulty in considering resonant behavior. The fastest transition, *i.e.*, smallest transition time of the current transients, is typically assumed to produce the greatest noise. This assumption is not valid, as described in this chapter, due to the non-monotonic behavior of the power/ground

noise. In [126], two extreme cases are considered in the analysis of ground noise: when the switching time is much smaller or much greater than the inverse resonant frequency. The maximum noise, however, occurs at an intermediate transition time rather than at one of the two extreme cases, as shown in this chapter.

The non-monotonic behavior of power/ground noise with respect to the transition time t_r is investigated for an inductive power distribution network with a decoupling capacitor. A time domain solution is provided for the transition time that produces resonant behavior, thereby maximizing the power/ground noise.

The sensitivity of the ground noise to the decoupling capacitance C_d and parasitic inductance L_g is also evaluated as a function of the transition time t_r. Increasing the decoupling capacitance is shown to efficiently reduce the noise for $t_r \leq 2\sqrt{L_g C_d}$. Alternatively, reducing the parasitic inductance L_g is shown to be effective for $t_r \geq 2\sqrt{L_g C_d}$.

The rest of the chapter is organized as follows. The equivalent circuit model to estimate the peak-to-peak power/ground noise is described in Section 4.1. The non-monotonic noise behavior with respect to the transition time is investigated in Section 4.2. Finally, the chapter is summarized in Section 4.3.

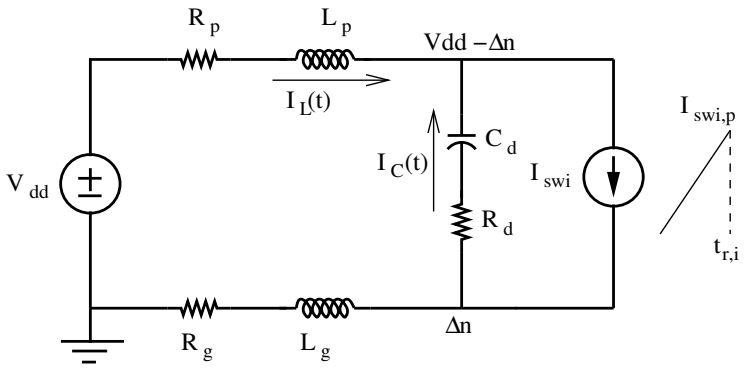


Figure 4.2: Equivalent circuit model to estimate power supply noise and ground bounce. R_p, L_p, and R_g, L_g represent the power and ground impedances, respectively. C_d is the decoupling capacitor and R_d is the effective series resistance of the capacitor. The load circuit is represented by a current source with a transition time $t_{r,i}$ and peak current $I_{swi,p}$.

4.1 Power/Ground Noise Model

An equivalent circuit model to investigate the noise behavior in terms of the transition time is shown in Fig. 4.2, where R_p, L_p, and R_g, L_g represent the power and ground impedances, respectively. C_d is the decoupling capacitor and R_d is the effective series resistance of the capacitor. Note that C_d is the summation of the intentional decoupling capacitance and the intrinsic capacitance of the non-switching gates within a circuit [177]. Techniques to estimate this capacitance are described in [178]. The load circuit is represented by a current source with a transition time $t_{r,i}$ and peak current $I_{swi,p}$. This triangular approximation of the current waveform is reasonable since a large number of registers switches simultaneously with the clock signal in synchronous digital circuits [179], [180], [181]. Note that this model does not consider the feedback effect of the power noise since in this model the current is

independent of this noise. Practically, however, the transistor current is affected by the power noise since this current is dependent upon the power supply [182]. Also note that the impedance between the decoupling capacitor and the current load is negligible, assuming that the decoupling capacitor is placed sufficiently close to the switching circuit [172].

The current provided by the decoupling capacitance $I_C(t)$ and the current flowing through the parasitic inductance $I_L(t)$ from the power supply are, respectively,

$$I_C(t) = -C_d \frac{\partial V_C}{\partial t}, \tag{4.1}$$

$$I_L(t) = \frac{1}{L_p} \int_0^t V_L(t) \partial t, \tag{4.2}$$

where $V_C(t)$ and $V_L(t)$ are, respectively,

$$V_C(t) = V_{dd} - 2\Delta n(t) + I_C(t) R_d, \tag{4.3}$$

$$V_L(t) = \Delta n(t) - I_L(t) R_p. \tag{4.4}$$

Note that the parasitic resistance and inductance of the power and ground networks are assumed to be equal due to the symmetry of these two networks [183], *i.e.*, $R_p = R_g$ and $L_p = L_g$.

A ramp function is assumed for the noise $\Delta n(t)$,

$$\Delta n(t) = \frac{V_{gnd,p}}{t_{r,v}} t, \tag{4.5}$$

where $V_{gnd,p}$ is the peak noise voltage and $t_{r,v}$ is the transition time of the noise spike [179], [184]. Practically, the noise can be better approximated with an exponential function (due to the discharge from the decoupling capacitor) at the expense of more complicated analytic solutions. The error introduced by approximating the noise as a ramp function is described in Section 4.2.

Replacing (4.3) in (4.1) and (4.4) in (4.2), and taking the derivative with respect to time results in the following differential equations,

$$I_C(t) = \frac{2C_d V_{gnd,p}}{t_{r,v}} - R_d C_d \frac{\partial I_C(t)}{\partial t}, \tag{4.6}$$

$$\frac{\partial I_L(t)}{\partial t} = \frac{V_{gnd,p} t}{L_g t_{r,v}} - \frac{R_g}{L_g} I_L(t). \tag{4.7}$$

Solving these differential equations with the initial conditions $I_C(0) = 0$ and $I_L(0) = 0$ produces the inductive and capacitive current, respectively,

$$I_C(t) = V_{gnd,p} \frac{2C_d}{t_{r,v}} (1 - e^{-t/(R_d C_d)}), \tag{4.8}$$

$$I_L(t) = V_{gnd,p}[\frac{t}{t_{r,v}R_g} - \frac{L}{t_{r,v}R_g^2}(1 - e^{-t/\frac{L_g}{R_g}})]. \tag{4.9}$$

These currents can be rewritten as

$$I_C(t) = G_C(t)V_{gnd,p}, \tag{4.10}$$

$$I_L(t) = G_L(t)V_{gnd,p}, \tag{4.11}$$

where $G_C(t)$, the conductance of the capacitance path, and $G_L(t)$, the conductance of the inductance path, are given, respectively, by

$$G_C(t) = \frac{2C_d}{t_{r,v}}(1 - e^{-t/(R_dC_d)}), \tag{4.12}$$

$$G_L(t) = \frac{t}{t_{r,v}R_g} - \frac{L_g}{t_{r,v}R_g^2}(1 - e^{-t/\frac{L_g}{R_g}}). \tag{4.13}$$

Note that these conductances are both a function of the transition time $t_{r,v}$. Hence, the amount of current provided by the decoupling capacitor as determined by (4.10), and the amount of current provided by the power supply as determined by (4.11) are both dependent upon the transition time. Specifically, as the transition time becomes smaller, $G_C(t)$ increases and $G_L(t)$ decreases. This situation is illustrated in Fig. 4.3 where (4.12) and (4.13) are plotted as a function of transition time. The

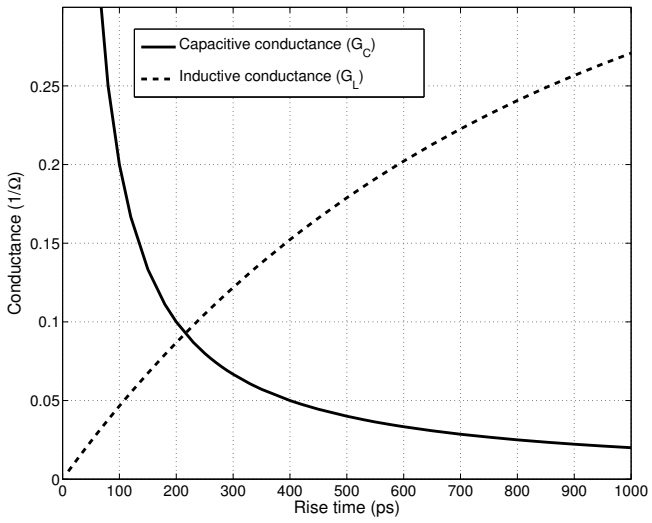


Figure 4.3: Capacitive and inductive conductance as a function of rise time where $I_{swi,p} = 11.5$ mA, $L_g = 1$ nH, $C_d = 10$ pF, $R_g = 2.2\ \Omega$, $R_d = 0.1\ \Omega$, and $t = t_{r,v}$.

operating parameters are $I_{swi,p} = 11.5$ mA, $L_g = 1$ nH, $C_d = 10$ pF, $R_g = 2.2\ \Omega$, $R_d = 0.1\ \Omega$, and $t = t_{r,v}$. Note that in the frequency domain, the admittance of an inductance $1/j\omega L$ decreases with frequency. A conductance refers to the time domain correspondence of the admittance. Since a decrease in transition time corresponds to an increase in frequency, the inductive conductance decreases. Similarly, the admittance of a capacitance $j\omega C$ increases with frequency. In the time domain, therefore, the capacitive conductance increases with decreasing transition time, as illustrated in Fig. 4.3.

The capacitive current $I_C(t)$, therefore, increases with decreasing transition time. Alternatively, the inductive current $I_L(t)$ increases with longer transition times. Intuitively, a smaller transition time corresponds to a higher frequency, where the

impedance of the capacitance is smaller and the inductance is higher. The capacitance is, therefore, more effective at smaller transition times and becomes less effective as the transition time increases since a majority of the switching current is provided by the power supply at higher transition times. The implications of this dependence are discussed in Section 4.2.

Assuming the peak ground noise $V_{gnd,p}$ occurs when the switching current reaches the maximum current $I_{swi,p}$ [179], *e.g.*, $t_{r,v} = t_{r,i} = t_r$, the summation of the capacitive and inductive currents at $t = t_r$ is equal to the peak switching current $I_{swi,p}$ of the load circuit,

$$I_{swi,p} = I_C(t_r) + I_L(t_r). \tag{4.14}$$

From (4.10), (4.11), and (4.14), the peak ground noise $V_{gnd,p}$ at $t = t_r$ can be expressed as

$$\frac{1}{V_{gnd,p}} = \frac{G_C(t_r)}{I_{swi,p}} + \frac{G_L(t_r)}{I_{swi,p}}. \tag{4.15}$$

Replacing (4.12) and (4.13) in (4.15) produces the peak noise voltage,

$$V_{gnd,p} = \frac{I_{swi,p} R_g^2 t_r}{2C_d R_g^2 (1 - e^{-t_r/(R_d C_d)}) - L_g (1 - e^{-t_r/\frac{L_g}{R_g}}) + R_g t_r}. \tag{4.16}$$

Note that if the capacitive current is much greater than the inductive current, *e.g.*,

(4.10) $\gg$ (4.11) or (4.12) $\gg$ (4.13), the second term in (4.15) can be neglected without a significant loss in accuracy, guaranteeing the pessimism of the expression. In this case, the peak ground noise is approximated by

$$V_{gnd,p} \approx I_{swi,p}/G_C(t_r). \tag{4.17}$$

Alternatively, if the inductive current is much greater, the first term in (4.15) can be neglected and the peak noise is estimated as

$$V_{gnd,p} \approx I_{swi,p}/G_L(t_r). \tag{4.18}$$

It is, however, important to note that the maximum peak noise occurs when the inductive and capacitive currents are approximately equal, as described in the following section.

If the circuit is underdamped, *i.e.*, the damping factor is smaller than one, oscillations occur due to a parallel combination of the parasitic inductance and decoupling capacitor. In this case, the peak-to-peak ground noise voltage $V_{gnd,pp}$ is

$$V_{gnd,pp} = V_{gnd,p}[1 + e^{-\pi\zeta/\sqrt{1-\zeta^2}}], \tag{4.19}$$

where ζ is the damping factor,

$$\zeta = [(2R_g + R_d)/2]\sqrt{C_d/2L_g}. \tag{4.20}$$

4.2 Non-monotonic Noise Behavior

The impedance of a parallel LC circuit is maximum at the resonant frequency, $\omega = 1/\sqrt{LC}$. At this frequency, both the capacitive and inductive paths carry a significant amount of current, giving rise to resonant behavior. Similarly, in the time domain, there exists a transition time at which the capacitive and inductive currents are close and the peak-to-peak noise is maximum. The *worst case transition time* producing the maximum noise is described in Section 4.2.1. The sensitivity of the noise to the decoupling capacitance and parasitic inductance as a function of transition time is explained in Section 4.2.2.

4.2.1 Worst Case Transition Time

The capacitive and inductive currents and the corresponding ground noise are plotted as a function of transition time in Fig. 4.4 using (4.8) and (4.9) for the currents, and (4.19) for the noise voltage where $I_{swi,p} = 11.5$ mA, $L_g = 1$ nH, $C_d = 10$ pF, $R_g = 2.2\ \Omega$, and $R_d = 0.1\ \Omega$. Equation (4.19) is also compared with SPICE in

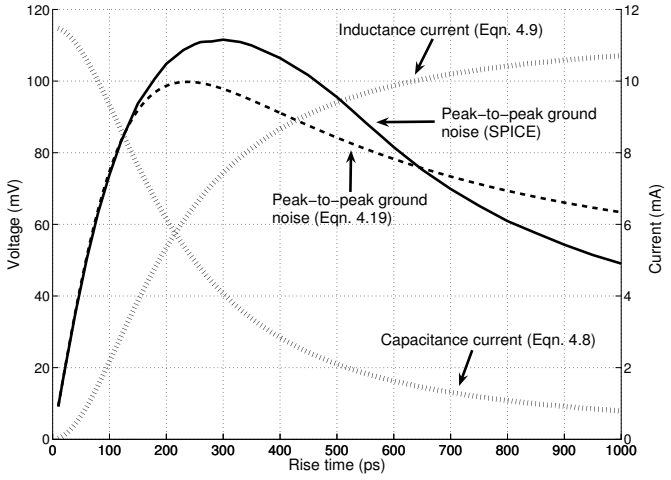


Figure 4.4: Comparison of peak-to-peak ground noise as a function of the transition time obtained from SPICE simulations and equation (4.19) for $I_{swi,p} = 11.5$ mA. The ground network impedances are $L_g = 1$ nH, $C_d = 10$ pF, $R_g = 2.2\ \Omega$, and $R_d = 0.1\ \Omega$. The dotted lines depict the estimated capacitive and inductive currents as a function of transition time.

Fig. 4.4.

The model accurately captures the non-monotonic dependence of noise on transition time, exhibiting a maximum error of 12.5%. Note that this error is due to approximating the noise as a ramp function which is a better assumption for smaller transition times, producing a smaller error.

As shown in Fig. 4.4, for sufficiently small transition times, the capacitive current dominates, and the inductance does not affect the ground noise. As the transition time increases, the capacitive current decreases and the inductive current increases. The peak noise occurs at a transition time where these currents are approximately equal. This specific transition time $t_{r,wc}$ is referred to as the *worst case transition time* or

the *equivalent transition time for resonance* (identical to the resonant behavior in the frequency domain). If the transition time further increases, the noise decreases due to lower $L\ \partial i/\partial t$ noise, making the capacitance less significant. The assumption of fast transients as the worst case scenario for noise can be overly optimistic in a circuit with sufficient decoupling capacitance. This conclusion is similar to reducing the resonant frequency with a larger capacitance. Increasing the decoupling capacitance, therefore, has the drawback of reducing the resonant frequency, or similarly, increasing the worst case transition time. Note that the intrinsic capacitance between the power and ground networks due to the non-switching gates contributes to the overall decoupling capacitance. For sufficiently large circuits, this capacitance can be significant [177], [178]. That is, the assumption of fast current transients as the worst case scenario is overly optimistic for large scale circuits.

According to Fig. 4.4, an expression for the worst case transition time $t_{r,wc}$ can be developed by equating (4.12) with (4.13) at $t = t_r$ and solving for t_r,

$$G_C(t_{r,wc}) - G_L(t_{r,wc}) = 0. \tag{4.21}$$

A closed form solution, however, does not exist due to the exponential terms in $G_C(t_{r,wc})$ and $G_L(t_{r,wc})$. Assuming $R_d \to 0$,

$$e^{-t_{r,wc}/(R_d C_d)} \rightarrow 0. \tag{4.22}$$

Replacing (4.22) in (4.12) gives

$$G_C(t_{r,wc}) \approx \frac{2C_d}{t_{r,wc}}. \tag{4.23}$$

Similarly, $G_L(t_{r,wc})$ can be expanded using a Taylor series expansion to

$$G_L(t_{r,wc}) = \frac{1}{R_g} - \frac{L_g}{t_{r,wc}{R_g}^2}[1 - (1 - \frac{t_{r,wc}R_g}{L_g} + \frac{1}{2!}(\frac{t_{r,wc}R_g}{L_g})^2 + ...)], \tag{4.24}$$

$$G_L(t_{r,wc}) = \frac{1}{2!}\frac{t_{r,wc}}{L_g} - \frac{1}{3!}\frac{(t_{r,wc})^2{R_g}^1}{{L_g}^2} + \frac{1}{4!}\frac{(t_{r,wc})^3{R_g}^2}{{L_g}^3}... \tag{4.25}$$

Assuming $R_g \rightarrow 0$ in (4.25) results in

$$G_L(t_{r,wc}) \approx \frac{t_{r,wc}}{2L_g}. \tag{4.26}$$

The worst case transition time $t_{r,wc}$ at which these conductances are approximately equal is determined from (4.23) and (4.26) as

$$\frac{2C_d}{t_{r,wc}} = \frac{t_{r,wc}}{2L_g} \Longrightarrow t_{r,wc} = 2\sqrt{L_g C_d}. \tag{4.27}$$

Table 4.1: Comparison of the peak-to-peak ground noise analytically obtained at $t_{r,wc} = 2\sqrt{L_g C_d}$ with the peak-to-peak noise obtained by SPICE for different parasitic ground network impedances.

$I_{swi,p}$ mA	L_g (nH)	C_d (pF)	R_g (ohm)	R_d (ohm)	Ground noise at $t_r = 2\sqrt{L_g C_d}$ (mV)	Maximum ground noise (SPICE) (mV)	Error (%)
11.5	0.25	10	2.2	0.1	44	46.2	4.8
11.5	0.5	10	2.2	0.1	66.3	67.8	2.2
11.5	1	10	2.2	0.1	98.5	99.8	1.3
11.5	1	15	2.2	0.1	78.3	79.6	1.6
11.5	1	20	2.2	0.1	66.3	67.8	2.2
11.5	1	20	4	0.1	59.4	64.8	8.3
11.5	1	20	6	0.1	55.1	70.2	21.5
11.5	1	20	6	1	54	68.6	21.3
11.5	1	20	6	2	52.8	67.2	21.4

The ground noise is obtained at $t_{r,wc} = 2\sqrt{L_g C_d}$ for different parasitic impedances of the ground network and compared with the maximum noise obtained at the same impedance. These results are listed in Table 4.1. The error of $t_{r,wc} = 2\sqrt{L_g C_d}$ in estimating the maximum noise is greater with increasing R_g and R_d, but is sufficiently small within the practical values of these resistances, as listed in Table 4.1.

The effect of the parasitic resistance R_g and the effective series resistance of the decoupling capacitance R_d on the worst case transition time is further illustrated, respectively, in Figs. 4.5 and 4.6. Increasing R_g reduces the noise until a specific transition time is reached due to additional damping. Beyond this transition time, however, the noise increases due to a greater IR drop along the ground network, making the decoupling capacitance ineffective. Alternatively, an increase in R_d results in decreased noise at higher transition times due to the increased damping and higher noise at smaller transition times where the decoupling capacitance is effective.

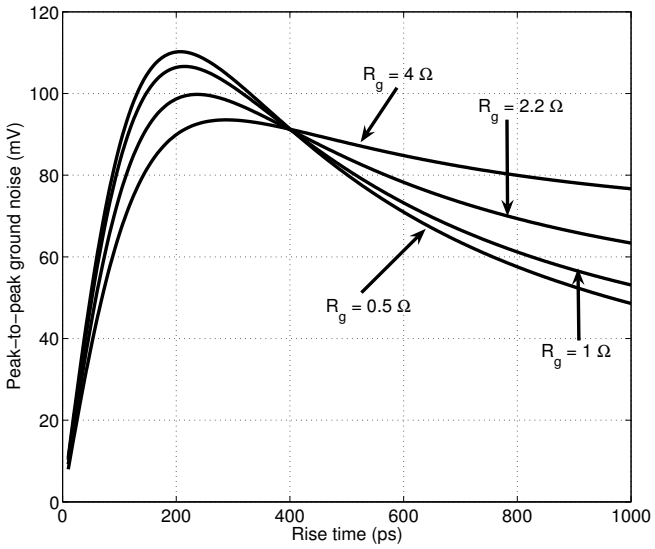


Figure 4.5: Peak-to-peak ground noise for different values of R_g when $I_{swi,p} = 11.5$ mA, $L_g = 1$ nH, $C_d = 10$ pF, and $R_d = 0.1\ \Omega$

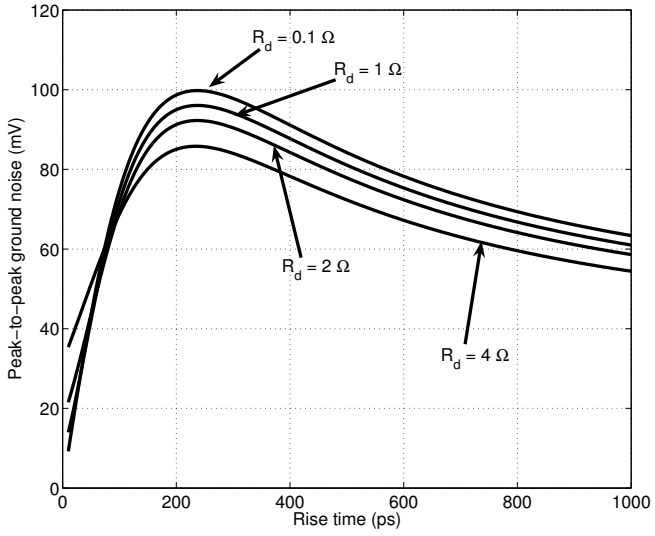


Figure 4.6: Peak-to-peak ground noise for different values of R_d when $I_{swi,p} = 11.5$ mA, $L_g = 1$ nH, $C_d = 10$ pF, and $R_g = 2.2\ \Omega$

4.2.2 Noise Sensitivity as a Function of Transition Time

The normalized sensitivity of the ground noise as a function of transition time is determined in this section to evaluate the efficacy of reducing the parasitic inductance and increasing the decoupling capacitance on reducing the ground noise. The normalized sensitivity of the ground noise to a parameter p_i is determined by

$$S_{p_i}^{V_{gnd,pp}} = \lim_{\Delta p_i \to 0} \frac{\frac{\Delta V_{gnd,pp}}{V_{gnd,pp}}}{\frac{\Delta p_i}{p_i}} = \frac{p_i}{V_{gnd,pp}} \frac{\partial V_{gnd,pp}}{\partial p_i}. \tag{4.28}$$

The normalized sensitivity of the ground noise as a function of transition time, as determined by (4.28), is shown in Fig. 4.7. The sensitivity of the noise to the decoupling capacitance is high at small transition times and decreases with increasing transition time. Alternatively, the sensitivity of the noise to the parasitic inductance is low at small transition times and increases with longer transition times. Increasing the decoupling capacitance is therefore effective in reducing the noise for $t_r \leq 2\sqrt{L_g C_d}$. Alternatively, reducing the parasitic inductance is effective for $t_r \geq 2\sqrt{L_g C_d}$. This behavior is due to the changing ratio of the capacitive and inductive currents with respect to the transition time, as shown in Fig. 4.4. The effect of the decoupling capacitance and parasitic inductance on the ground noise is listed, respectively, in Tables 4.2 and 4.3 for different transition times. At $t_r = 70$ ps, doubling the decoupling capacitance reduces the noise by 51.2%, and only 25.7% when $t_r = 400$ ps.

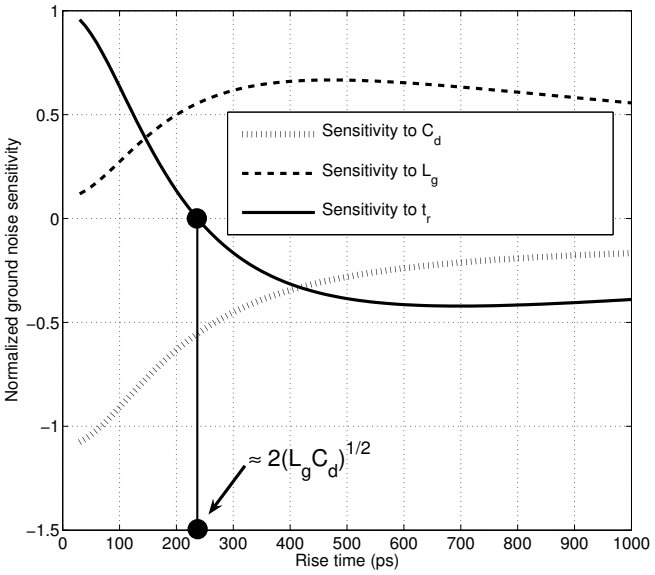


Figure 4.7: Normalized sensitivity of the ground noise as a function of transition time when $I_{swi,p} = 11.5$ mA, $L_g = 1$ nH, $C_d = 10$ pF, $R_g = 2.2\ \Omega$, and $R_d = 0.1\ \Omega$

Table 4.2: Effect of the decoupling capacitance on reducing the peak-to-peak ground noise. $L_g = 1\ nH$.

Transition time (ps)	$C_d = 10$ pF	$C_d = 20$ pF	Reduction
70	57.8	28.2	51.2%
$200 = 2\sqrt{L_g C_d}$	98.7	59.5	39.7%
400	91.1	67.7	25.7%
800	69.4	58.7	15.4%

Table 4.3: Effect of the parasitic inductance on reducing the peak-to-peak ground noise. $C_d = 10\ pF$.

Transition time (ps)	$L_g = 1$ nH	$L_g = 0.5$ nH	Reduction
70	57.8	48.9	15.4%
$200 = 2\sqrt{L_g C_d}$	98.7	67.7	31.4%
400	91.1	58.7	35.6%
800	69.4	48.2	30.5%

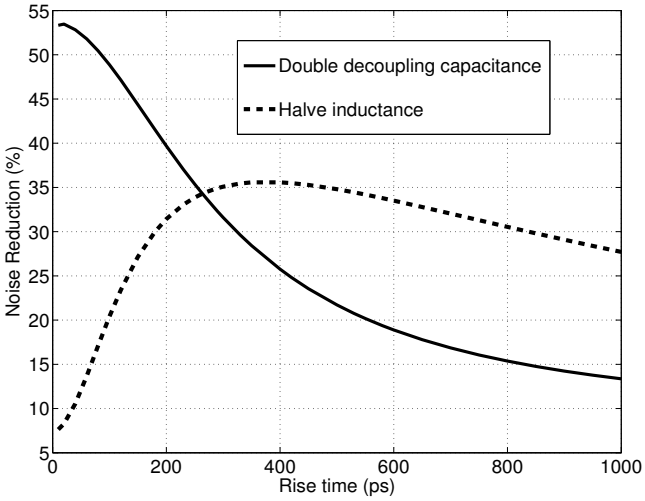


Figure 4.8: Comparison of the noise reduction obtained by doubling the decoupling capacitance and halving the inductance.

Halving the parasitic inductance, however, reduces the noise by only 15.4% when $t_r = 70$ ps, and 35.6% when $t_r = 400$ ps. Note that the sensitivity to the transition time crosses over at zero when $t_r \approx 2\sqrt{L_g C_d}$, demonstrating the non-monotonic dependence, as described in Section 4.2.1. Placing additional decoupling capacitance is therefore more effective in reducing the noise at smaller rise times. Alternatively, reducing the parasitic inductance is more effective at higher transition times. This behavior is illustrated in Fig. 4.8.

4.3 Conclusions

The non-monotonic dependence of the power/ground noise on the transition time is shown for an inductive power distribution network with a decoupling capacitance.

The power/ground interconnect is modeled as a series RL impedance. The decoupling capacitance is modeled as a capacitance in series with a resistance. The model captures the non-monotonic dependence of noise on the transition time with sufficient accuracy as compared to SPICE. The *worst case transition time* or the *equivalent transition time for resonance* producing the maximum peak-to-peak noise is also presented based on this model. The worst case power/ground noise is shown to be significantly inaccurate if determined assuming fast switching characteristics. The effect of the parasitic line resistance R_g and effective series resistance R_d of the decoupling capacitor on the noise is also investigated. Increasing R_g reduces the noise at smaller transition times due to additional damping since the decoupling capacitance is effective. At higher transition times, however, the noise increases due to a larger IR drop along the inductive power/ground path. Alternatively, an increase in R_d results in lower noise at higher transition times by providing additional damping and higher noise at lower transition times due to a larger IR drop along the capacitive path. The sensitivity of the noise on the decoupling capacitance and parasitic inductance is also investigated. The decoupling capacitance is shown to efficiently reduce the noise for transition times smaller than twice the LC time constant, $t_r \leq 2\sqrt{L_g C_d}$. Alternatively, reducing the parasitic inductance is effective for transition times greater than twice the LC time constant, $t_r \geq 2\sqrt{L_g C_d}$.

Chapter 5

Substrate and Ground Noise Interactions in Mixed-Signal Circuits

As the speed and density of a circuit increase, the average current required to charge the total capacitive load also increases. Due to the parasitic resistance and inductance of the on-chip power distribution network, these fast changes in current produce significant voltage fluctuations on the power nodes, affecting overall signal integrity, as described in Chapter 2.

Techniques to reduce power noise such as the use of on-chip decoupling capacitors have been developed [12], [30], [185]. Models of a power distribution system and the dependence of power noise on different circuit parameters have also been investigated [12], [29], [186], [175]. These papers, however, do not consider the interaction of the power distribution network with the substrate. If this interaction is not considered, the power noise analysis process can be significantly inaccurate.

Power supply coupling noise is an important source of substrate noise [120], [187]. This source of noise has been analyzed treating the substrate noise as the primary concern [178], [188]. Coupling of the substrate noise into the power supply and ground rails, however, has not received much attention.

In [189], the effect of the substrate on the power noise is analyzed at the IC level. The system is modeled as a linear network, neglecting junction coupling and impact ionization mechanisms. Furthermore, a resistive model is used to characterize the on-chip ground distribution network, neglecting the on-chip inductance. As shown in [190], however, on-chip inductive noise will become more significant with technology scaling.

The interaction of the substrate with an inductive on-chip ground distribution network is the focus of this chapter. The contribution of the substrate to the total ground noise is shown to be significant. For a CMOS inverter, the substrate can reduce negative peak ground noise by 49% during the high-to-low output transition. The substrate, however, increases the positive peak ground noise by 72% during the low-to-high output transition. The effect of the substrate should therefore not be neglected if the inductance of the on-chip ground distribution network is non-negligible. Furthermore, conventional triangle or trapezoid type current demand estimations of the nonlinear circuits are shown to be significantly inaccurate if the ground lines exhibit inductive behavior.

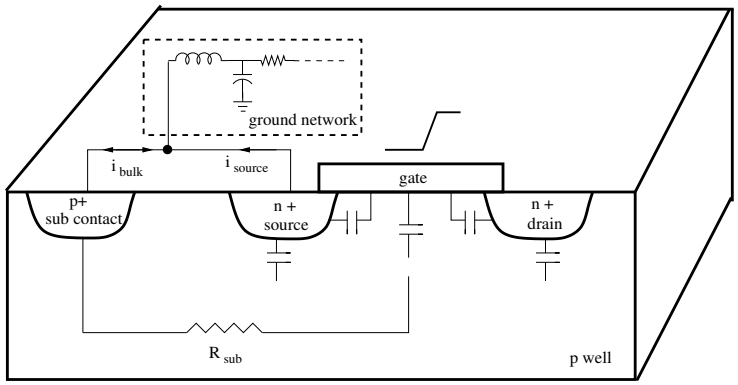


Figure 5.1: Interaction of the substrate and on-chip ground distribution network within an NMOS transistor. The current flowing through the ground network is the summation of the source and bulk currents.

The rest of the chapter is organized as follows. The interaction of the substrate with the ground network is described in Section 5.1. The substrate model, on-chip ground distribution network model, and the superposition methodology are described in Section 5.2. Simulation results illustrating the contribution of the substrate to the total ground noise are presented in Section 5.3. The chapter is summarized in Section 5.4.

5.1 Background

During the switching activity of the transistor, noise couples into the substrate through the junction capacitances of the transistor. The coupled noise propagates through the substrate, interacting with the ground distribution network through the substrate contact. This interaction is illustrated in Fig. 5.1. Note that the substrate

contact is connected to the ground network. The source and bulk currents therefore flow to the ground network.

Although the bulk current of the substrate is small as compared to the switching current of the transistor, the di/dt levels of the bulk current are sufficient to affect the overall ground noise. Therefore, if the parasitic inductance of the ground network is considered, the effect of the substrate on the ground noise cannot be neglected. Furthermore, the junction capacitances of the transistors, the substrate resistance, and the parasitic impedance of the power supply interconnect form an RLC circuit that contributes to the resonance of the power nodes. The effect of the substrate should therefore be considered when analyzing the on-chip power supply noise.

The purpose of this chapter is to quantify the effect of the substrate on the ground noise if the ground lines exhibit inductive behavior. A transistor level approach is presented to consider each of the substrate coupling mechanisms. The total on-chip ground noise is modeled as the superposition of two different noise voltages. The first noise voltage is caused by the source current of the transistor, and the second noise voltage is caused by the bulk current of the substrate. The individual contributions of the source and bulk currents to the overall ground noise are summed based on this superposition principle.

As shown in Figs. 5.1 and 5.4(a), the total current that flows through the ground

network is

$$i(t) = i_{source}(t) + i_{bulk}(t). \tag{5.1}$$

The bulk current in (5.1) is usually neglected due to the relatively small magnitude as compared to the source current. This assumption is acceptable if the ground network is modeled as a resistive-only network. If the on-chip inductance is considered, however, the total ground noise is

$$\Delta V = i_{source}(t)R + \frac{di_{source}(t)}{dt}L + i_{bulk}(t)R + \frac{di_{bulk}(t)}{dt}L, \tag{5.2}$$

where R and L are the parasitic resistance and inductance of the on-chip ground lines, respectively.

Although the absolute magnitude of the bulk current is smaller than the source current, the bulk current significantly affects the total ground noise due to the last term in (5.2), which represents the rate of change of the bulk current.

The effect of the substrate is quantified in this chapter using superposition of the source and bulk related noise voltages. Note that in (5.2), the first two terms and the last two terms represent, respectively, the source related ground noise and the bulk related ground noise. Transistor level simulation results illustrate that the substrate contributes, on average, 56% of the peak ground noise. The effect of the substrate should therefore be considered for accurate analysis of the ground noise.

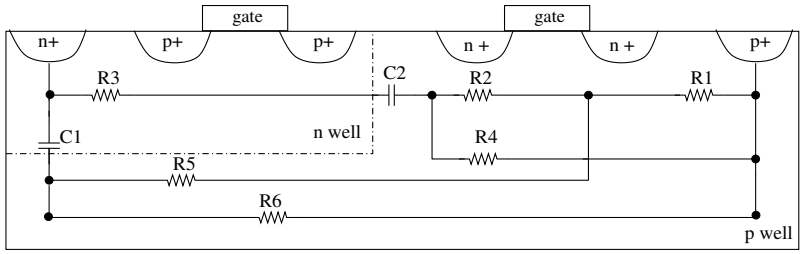


Figure 5.2: A bulk-type substrate model of a CMOS structure.

5.2 Circuit Models and Superposition Technique

A bulk-type (high-ohmic) substrate is assumed due to isolation advantages and applicability to mixed-signal circuits. The substrate model used in the simulations is shown in Fig. 5.2 for a CMOS inverter. R_1 and R_3 represent the resistance between the bulk nodes and the substrate contacts, and R_2 and R_5 represent the resistance between the N-Well and the bulk node of the NMOS transistor. R_4 and R_6 represent the resistance between the N-Well and the substrate contact for the ground connection. C_1 and C_2 represent the N-Well capacitance.

An impedance model composed of ten distributed RLC (T cell) [191] cells is used to model the on-chip ground distribution network, as shown in Fig. 5.3. The total resistance, inductance, and capacitance of the interconnect are assumed in this analysis to be 1 Ω, 200 pH, and 200 fF, respectively.

Simulations are based on a 0.18 μm CMOS technology using SPICE. A BSIM3v3 model is used to characterize the transistors since this model includes the effects of capacitive coupling and impact ionization as sources of substrate noise [192]. The

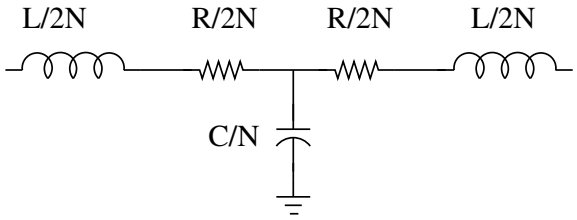


Figure 5.3: One stage of a distributed RLC interconnect model for an on-chip ground distribution network.

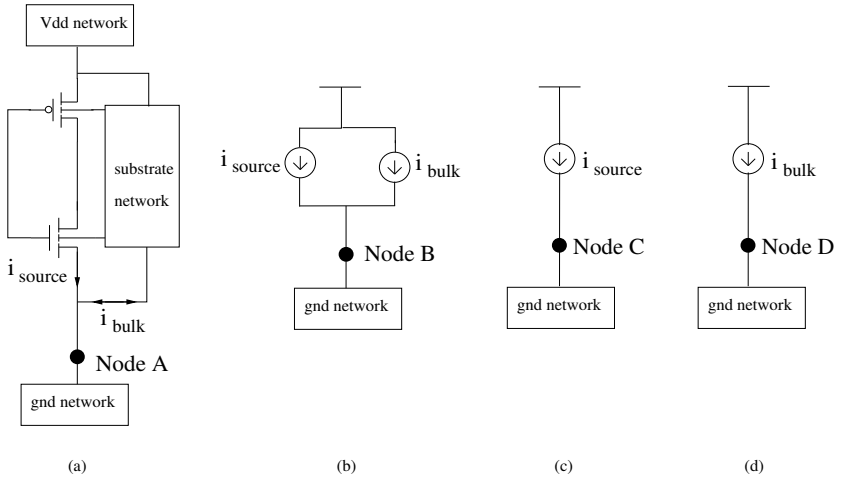


Figure 5.4: Illustration of the superposition technique to analyze the effect of the substrate on the on-chip ground noise. The overall ground noise is modeled as a superposition of two noise voltages: source related ground noise and bulk related ground noise: (a) The CMOS inverter including the substrate network and the ground distribution network used to analyze the total ground noise, (b) summation of the PWL approximation of the source current and bulk current which flows through the same ground network, (c) PWL approximation of the source current to estimate the source related ground noise, and (d) PWL approximation of the bulk current to estimate the bulk related ground noise.

simulation model illustrating the superposition methodology is shown in Fig. 5.4. A CMOS inverter is analyzed in Fig. 5.4(a) with the substrate and ground distribution network models. The voltage at Node A represents the on-chip ground noise. The source and bulk currents are individually approximated using piecewise linear (PWL)

current sources, as shown in Fig. 5.4(b). The voltage at Nodes A and B is compared in order to evaluate the accuracy of these estimates of the current demand. If the voltage at Nodes A and B is sufficiently similar, the voltage at Nodes C and D is analyzed. Note that the voltage at Nodes C and D represents, respectively, the source related ground noise and bulk related ground noise. The superposition of the source and bulk currents therefore supports the analysis of the individual contributions of the source and substrate to the total ground noise.

5.3 Simulation Results

PWL current estimation results are provided in Section 5.3.1. The results illustrating the effect of the substrate on the ground noise are presented in Section 5.3.2.

5.3.1 PWL Current Estimation

Estimating the current demand of the nonlinear circuits using linear current sources is a common approach to analyze power supply noise. Nonlinear circuits are replaced by these linear current sources in order to reduce the simulation time of the system [12]. If the inductance of the power and ground lines is considered, however, estimating the current demand becomes a difficult process. Conventional triangle or trapezoidal type current demand approximations [22] do not provide the necessary accuracy due to the inductance.

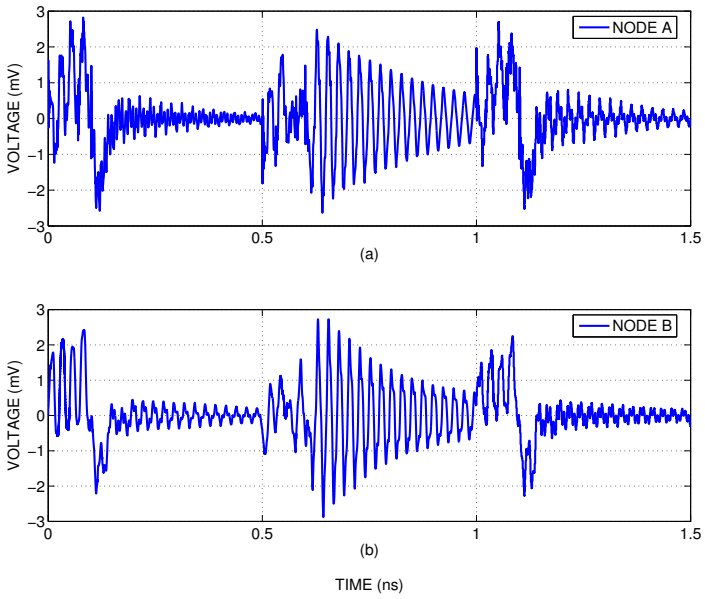


Figure 5.5: Comparison of the voltage at Nodes A and B (see Fig. 5.4) illustrating the accuracy of the current estimations. The maximum error for the peak voltage is less than 10% within one period.

In this analysis, estimates of the source and bulk currents are obtained using an iterative process. The number of time instances included in the PWL approximations is increased in each iteration until sufficient accuracy is achieved. The different rates of change in the current waveform should therefore be considered if the ground network exhibits inductive behavior.

The voltage at Nodes A and B as shown in Figs. 5.4(a) and 5.4(d) is compared in Fig. 5.5 to evaluate the accuracy of the approximations. The amount of error for the peak voltage is less than 10% at any time within one period. This model therefore provides sufficient accuracy to individually analyze the source and bulk

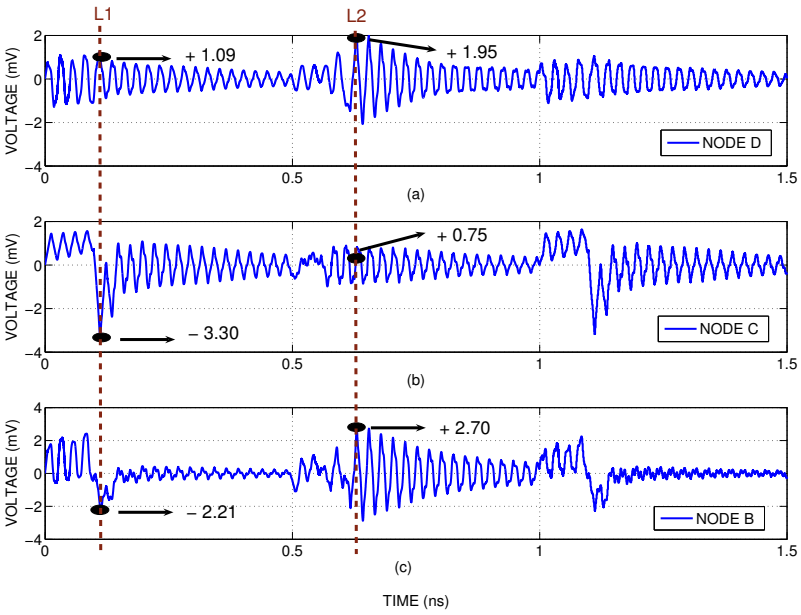


Figure 5.6: Comparison of the voltage at Nodes B, C and D (see Fig. 5.4). Nodes B, C, and D illustrate, respectively, the total ground noise, source related ground noise, and bulk related ground noise.

current contributions. Note that more than 50% error is obtained if a conventional triangle type approximation is used.

5.3.2 Effect of the Substrate on Ground Noise

In order to quantify the effect of the substrate on ground noise, the voltage at Nodes B, C, and D is compared in the time domain in Fig. 5.6.

The voltage at Nodes C and D represents, respectively, the ground noise due to the source and bulk currents. The voltage at Node B represents the total ground noise. The summation of the voltage waveforms shown in Figs. 5.6(a) and 5.6(b) results in the waveform shown in Fig. 5.6(c). Two different time instances are marked by

the dashed vertical lines, L1 and L2, to illustrate the effect of the substrate on the ground noise. At L1, the total ground noise is -2.21 mV. At this time, the noise voltage induced by the source current is -3.30 mV and the noise voltage induced by the bulk current is +1.09 mV. The substrate can therefore be used to lower the total ground noise since the source and bulk currents are temporally out-of-phase. At L2, the total ground noise is +2.70 mV, the bulk related ground noise is +1.95 mV, and the source related ground noise is +0.75 mV. At this time, therefore, the substrate degrades the total ground noise by 72% since the bulk current is in-phase with the source current. Furthermore, the resonance at Node B is primarily caused by the bulk current rather than the source current.

The proposed methodology can also be applied to a NAND and NOR gate in order to generalize the results, as listed in Table 5.1. The positive and negative peak noise is listed for the high-to-low and low-to-high output transitions. The effect of the substrate and source are individually listed for each gate. The contribution of the substrate varies between 22% and 95%, exhibiting an average contribution to the peak ground noise voltage of 56%.

5.4 Summary

The effect of the substrate on the ground noise is analyzed for an inverter, NAND, and NOR gate. The substrate is modeled as a resistive network and the on-chip

Table 5.1: Absolute (abs) and relative (rel) contribution of the source and bulk currents to the ground noise voltage for three gates: inverter (INV), NAND, and NOR. The negative and positive peak noise is listed for both the high-to-low and low-to-high output transitions.

GND Noise (mV)	High-to-low (output)		
	INV	NAND	NOR
Negative peak noise	-2.21	-4.44	-2.16
Bulk contribution: abs (rel)	+1.09 (-49.3%)	+0.98 (-22%)	+0.69 (-31.9%)
Source contribution: abs (rel)	-3.30 (149.3%)	-5.42 (122%)	-2.85 (131.9%)
Positive peak noise	+2.40	+3.73	+2.97
Bulk contribution: abs (rel)	+1.0 (41.6%)	+1.95 (52.3%)	+1.61 (54.2%)
Source contribution: abs (rel)	+1.40 (58.4%)	+1.78 (47.7%)	+1.36 (45.8%)
	Low-to-high (output)		
	INV	NAND	NOR
Negative peak noise	-2.86	-3.48	-2.77
Bulk contribution: abs (rel)	-2.07 (72.4%)	-1.39 (39.9%)	-2.18 (78.7%)
Source contribution: abs (rel)	-0.79 (27.6%)	-2.09 (60.1%)	-0.59 (21.3%)
Positive peak noise	+2.70	+2.57	+2.94
Bulk contribution: abs (rel)	+1.95 (72.2%)	+2.46 (95.7%)	+1.8 (61.2%)
Source contribution: abs (rel)	+0.75 (27.8%)	+0.11 (4.3%)	+1.14 (38.8%)

ground network is modeled as a distributed RLC impedance. The total ground noise is treated as the superposition of two noise voltages: the source related ground noise and bulk related ground noise. It is shown that the bulk current significantly affects the ground noise due to the inductance of the ground network since the rate of change of the bulk current is not negligible. Specifically, if the source and bulk currents are in-phase, the substrate degrades the ground noise. If the source and bulk currents are out-of-phase, the substrate improves the ground noise. Based on this analysis, the substrate contributes, on average, 56% of the peak ground noise. Two primary conclusions can therefore be drawn if the ground lines exhibit inductive behavior: (1) the substrate should not be neglected in an analysis and model of the ground noise,

and (2) triangle or trapezoid type current demand estimation of the nonlinear circuits is not sufficiently accurate since these models do not capture different rates of change of the current waveforms.

Chapter 6

Dominant Substrate Noise Coupling with Multiple Switching Gates

Substrate coupling continues to be a primary concern for mixed-signal systems-on-chips (SoCs) where sensitive analog/RF circuits often coexist with aggressor digital circuits on the same substrate, as described in Chapter 3. The substrate noise can affect a sensitive circuit by modifying the threshold voltage of the transistors, or by coupling into the signal and power/ground rails, degrading the precision required by sensitive circuitry [120], [193], or causing a circuit to fail [130].

The baseband digital circuit injects noise into the substrate through three primary mechanisms [122]: (1) coupling from the source/drain junction capacitances of the transistors during switching, (2) coupling from the power and ground networks of the digital circuit, and (3) impact ionization, which is negligible as compared to the first two mechanisms, as shown in [122]. The relative contributions of the first

two mechanisms, however, have not been quantified in a sufficiently accurate manner. For large scale circuits, it is usually assumed that power/ground coupling dominates source/drain coupling [126], [194]. The validity of this assumption, however, depends upon several parameters in the circuit such as the number of simultaneously switching gates, rise time, decoupling capacitance, and parasitic inductance of the power/ground rails. A simple, yet physically intuitive macrolevel model is presented in this chapter to identify the dominant substrate coupling mechanism at the early stages of the design process, while simultaneously considering multiple parameters such as the number of switching gates, rise time of the current load, package and on-chip parasitic impedance of the power/ground networks, and on-chip decoupling capacitance.

The rest of the chapter is organized as follows. The limitations of the previous works are summarized in Section 6.1. Models to estimate the peak-to-peak substrate noise are presented in Section 6.2. These expressions are used in Section 6.3 to identify the dominant noise generation mechanism. In Section 6.4, a sensitivity analysis validating the effects of these parameters on the substrate noise is presented. The design implications of these results are discussed in Section 6.5, and the chapter is summarized in Section 6.6.

6.1 Previous Work

The existing work on dominant noise analysis fails to collectively consider the various physical parameters in a sufficiently accurate model. Different noise injection mechanisms are investigated and quantified in [122], but the primary emphasis is on the generation rather than the propagation of the noise throughout the substrate. In [195], the value of inductance at which the power/ground coupling dominates is provided for a scaled buffer circuit. The effects of the other parameters, however, have not been investigated. In [126], the impact of technology scaling on different noise generation mechanisms is analytically examined. To determine the dominant noise mechanism, the substrate resistance between the source and the victim is linearly scaled with the number of switching gates to analyze the ground coupling. This procedure, however, produces highly pessimistic results for ground coupling since an additional contact in the substrate does not linearly decrease the substrate resistance due to the mesh structure of the substrate. Furthermore, in [126], two extreme cases are considered in analyzing ground noise: when the rise time is much smaller or much greater than the inverse of the resonant frequency. The maximum noise, however, typically occurs at an intermediate rise time rather than at one of the two extreme cases, as shown in this chapter.

A macrolevel model is presented to evaluate the dominant substrate coupling mechanism in the early stages of the design process, while considering multiple circuit

parameters such as the number of simultaneously switching gates, rise time, on-chip decoupling capacitance, package and on-chip parasitic inductance and resistance, substrate resistance, substrate contact density, and the physical distance between the aggressor and victim blocks. Identification of the dominant noise coupling mechanism helps in comparing various substrate noise reduction techniques to determine the preferable technique. Furthermore, the sensitivity of the substrate noise to various parameters is evaluated as a function of rise time and number of switching gates. Design implications of the dominant noise source and sensitivity analysis are also discussed.

6.2 Substrate Coupling Model to Estimate Noise

Coupling from the noisy ground network and source/drain junction coupling are considered to be the two primary noise generation mechanisms since coupling from the power network is isolated due to the n-well capacitance. Specifically, ground coupling dominates the power coupling until a sufficiently high frequency is reached, beyond which both mechanisms affect the noise similarly, as described in [114].

A high resistivity non-epi substrate is assumed to provide enhanced isolation, making the model applicable to mixed-signal circuits. Note that the model of the substrate is resistive since the dielectric characteristics are negligible for frequencies below about 10 GHz for a high resistivity substrate, as described in [139]. Models

for ground bounce coupling and source/drain junction coupling for a single switching gate, and for multiple gates are described, respectively, in Sections 6.2.1 and 6.2.2. Validation of the model is described in Section 6.2.3.

6.2.1 Substrate Coupling for a Single Switching Gate

Noise on the ground network resistively couples into the substrate through the substrate contacts. The ground noise is quantified, assuming that the substrate network does not affect the ground noise due to the high impedance of the substrate as compared to the ground network. An expression for the peak-to-peak ground noise $(V_{gnd})_{pp}$ has been developed in Chapter 4 as a function of package and on-chip parasitic impedances L_g and R_g, on-chip decoupling capacitor C_d, effective series resistance of the capacitor R_d, current source rise time $(t_r)_i$, and peak current $(I_{swi})_p$. This expression is used in this section to estimate substrate noise at the victim node, as illustrated in Fig 6.1. In Fig. 6.1, the substrate resistance between the contact and bulk of the device is represented by R_{cb}. R_{dist} represents the equivalent substrate resistance between the bulk and the victim node of the sensitive analog circuit. R_{vc} is the equivalent substrate resistance between the victim node and the analog contact. Note that the victim node refers to the bulk node within the victim device. R_{ang} and L_{ang} represent the parasitic impedance of the analog ground network.

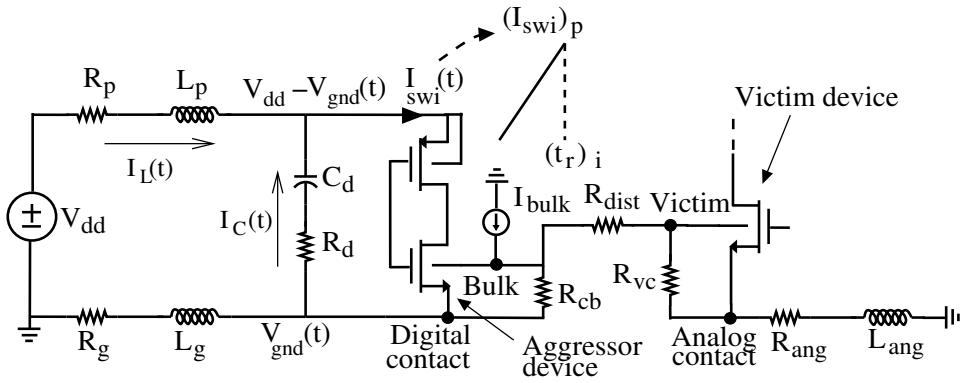


Figure 6.1: Equivalent model to estimate two types of substrate coupling mechanisms for a single switching gate: (1) ground coupling and (2) source and drain junction coupling.

The substrate noise at the victim node due to ground coupling can be approximated as

$$(V_{s-gnd})_{pp} \approx \frac{(V_{gnd})_{pp}}{R_{cb} + R_{dist} + R_{vc}}(R_{ang} + R_{vc} + \frac{L_{ang}}{t_r}), \tag{6.1}$$

where $(V_{gnd})_{pp}$ is determined by (4.19).

Noise couples into the substrate through the source/drain junction capacitance of the devices during switching. This noise source is modeled as a current source from within the bulk of a device with a peak current of $(I_{bulk})_p$ and a rise time of t_r (which is assumed to be equal to the rise time of the switching current). The substrate noise at the victim node due to source/drain junction coupling can be approximated as

$$(V_{s-bulk})_p \approx (I_{bulk})_p \frac{R_{cb}}{R_{cb} + R_{dist} + R_{vc}}.(R_{ang} + R_{vc} + \frac{L_{ang}}{t_r}). \tag{6.2}$$

The total noise at the victim node is the summation of (6.1) and (6.2),

$$(V_{s-total})_{pp} \approx (V_{s-gnd})_{pp} + (V_{s-bulk})_{p}. \tag{6.3}$$

6.2.2 Substrate Coupling Model for Multiple Switching Gates

The model introduced in the previous section for a single gate is extended to analyze the effect of simultaneously switching gates on the substrate noise characteristics. Each macromodel for a switching gate consists of two current sources, I_{swi-g} and I_{bulk-g}, to represent the switching and bulk currents, respectively, and a substrate resistance R_{cb} between the contact and bulk, assuming the gate has a substrate contact.

These gates are connected as shown in Fig. 6.2 to obtain a model of substrate coupling for multiple gates, assuming the aggressor consists of standard cells. For a given number of switching gates n, L and M gates are placed in the horizontal and vertical directions, respectively, such that $L\ x\ M = n$ and the resulting rectangle is as close as possible to a square in terms of the physical layout of the aggressor circuit. The bulk node of each gate located along the horizontal direction is connected through a substrate resistance R_{bb}. The bulk regions of the gates located along the vertical direction which share the same local ground line are vertically connected through the resistance 2 x R_{cb}.

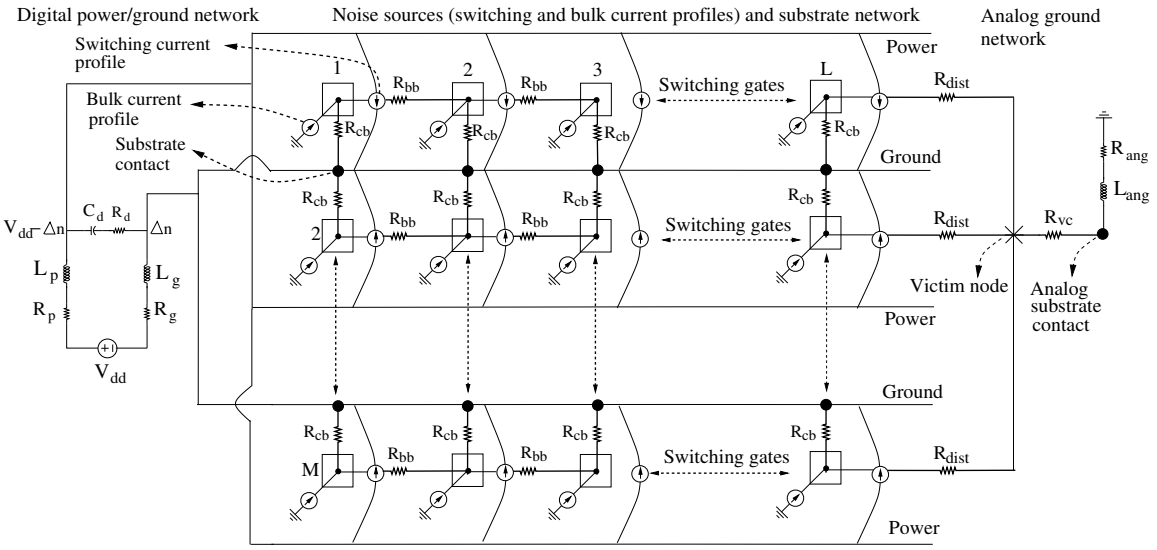


Figure 6.2: Equivalent circuit model to estimate substrate noise for multiple switching gates.

The ground noise $(V_{gnd})_{pp}$ at each substrate contact location is determined from (4.19) where the total peak current scales to $n(I_{swi-g})_p$. Note that the switching gates are assumed in this analysis to be identical. The peak-to-peak substrate noise at the victim node $(V_{victim})_{pp}$ is the summation of the noise due to each contact and bulk current source,

$$(V_{victim})_{pp} = [(V_{gnd})_{pp}TF_{c1} + ... + (V_{gnd})_{pp}TF_{cn}]$$
$$+[(I_{bulk1})_{pp}TF_{ib1} + ... + (I_{bulkn})_{pp}TF_{ibn}], \quad (6.4)$$

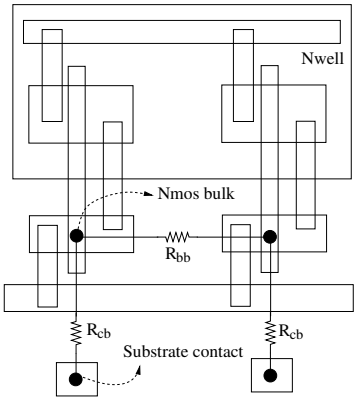


Figure 6.3: Layout of two inverters to extract R_{bb} and R_{cb}.

where $TF_{c1}, ..., TF_{cn}$ represent the voltage noise transfer function from the corresponding contact location to the victim node, and $TF_{ib1}, ..., TF_{ibn}$ represent the current noise transfer function from the corresponding bulk current source to the victim node. These transfer functions are determined from the resistive substrate network, as illustrated in Fig. 6.2. This model is used to quantify various noise sources and evaluate the dominant coupling mechanism as a function of multiple parameters.

6.2.3 Extraction of Parameters and Model Validation

An industrial 90 nm CMOS technology with a lightly doped (non-epi type) substrate is used to extract the parameters applied in this model. An inverter with NMOS size, $W/L = 0.31$ μm / 0.1 μm, and PMOS size, $W/L = 0.44$ μm / 0.1 μm, is used in all of the analyses presented in this paper. The layout of the two cells, as shown in Fig. 6.3, is extracted using Assura and SubstrateStorm [196]. Related parameters are listed in Table 6.1. The peak switching and bulk currents are obtained when the cell

Table 6.1: Extracted parameters characterizing an inverter.

Parameter	Value
$(W/L)_{nmos}$	0.31 μm / 0.1 μm
$(W/L)_{pmos}$	0.44 μm / 0.1 μm
$(I_{swi-g})_p$	57.5 μA
$(I_{bulk-g})_p$	6.7 μA
R_{bb}	16.8 kΩ
R_{cb}	10.7 kΩ
R_{dist}	40 kΩ
R_{vc}	660 Ω

is driven by a ramp input with a 100 ps rise and fall time which drives an identical gate. The substrate resistances R_{dist} and R_{vc} are similarly extracted assuming the victim node is located 100 μm from the aggressor circuit, and placed within a $p+$ guard ring with 15 analog substrate contacts.

At a certain number of switching gates, the estimated peak-to-peak substrate noise is characterized by (6.4). This expression is compared with SPICE in Fig. 6.4, where $n = 200$, $L_g = L_{ang} = 1$ nH, $C_d = 10$ pF, $R_g = R_{ang} = 2.2$ Ω, and $R_d = 0.1$ Ω. The model accurately captures the nonmonotonic dependence of the substrate noise on rise time, exhibiting a maximum error of 18.4%. Note that this error is due to approximating the noise as a ramp function (which is a better assumption for smaller rise times) and the effect of feedback on the nonlinear devices, which is not captured in the model.

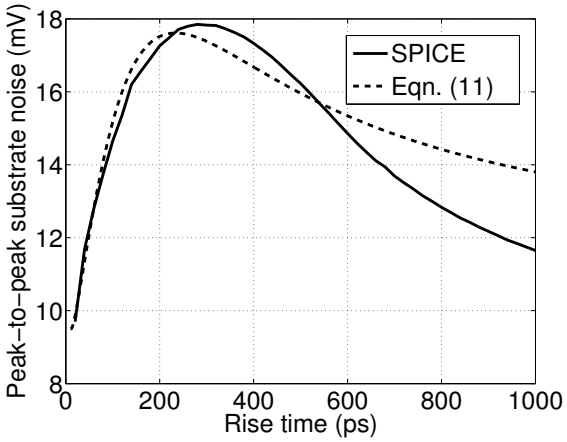


Figure 6.4: Comparison of peak-to-peak substrate noise as a function of the rise time obtained from SPICE simulations and (6.4).

6.3 Dominant Substrate Noise Coupling Mechanism

The models and expressions for ground and source/drain coupling are used in this section to evaluate the dominant substrate noise generation mechanism. The boundary conditions are determined where the ground coupling exceeds the source/drain coupling. These conditions are presented as a function of the parasitic inductance L_g, decoupling capacitor C_d, rise time t_r, and number of switching gates n.

The extraction and simulation of large scale circuits to determine the dominant noise generation mechanism is not feasible due to the high computational requirements since the logic gates, power/ground network, and substrate must all be considered together [129], [133], [136]. A sufficiently accurate model which includes the

effects of multiple parameters on the dominant noise is therefore required. The models and expressions presented in this paper are used to compare source/drain coupling with ground coupling, thereby providing improved understanding of the behavior of these two noise generation mechanisms as a function of multiple parameters.

Based on the model shown in Fig. 6.2, a specific number of switching gates exists beyond which ground coupling exceeds source/drain coupling. This number depends primarily upon the rise time, parasitic inductance, and decoupling capacitance. The effect of the number of switching gates, decoupling capacitance, and parasitic inductance on the dominant noise generation mechanism is explained, respectively, in Sections 6.3.1 and 6.3.2.

6.3.1 Effect of Circuit Size on the Dominant Noise Coupling Mechanism

As a greater number of gates simultaneously switch, the ground noise on each substrate contact increases due to the additional supply current. The ground coupling component of the substrate noise therefore increases with a larger number of switching gates. Furthermore, each switching gate injects noise from the junction capacitances, increasing the source/drain junction coupling mechanism. Alternatively, a contact filters the noise injected by the source/drain junction coupling and ground coupling from the other contacts, reducing the overall substrate noise.

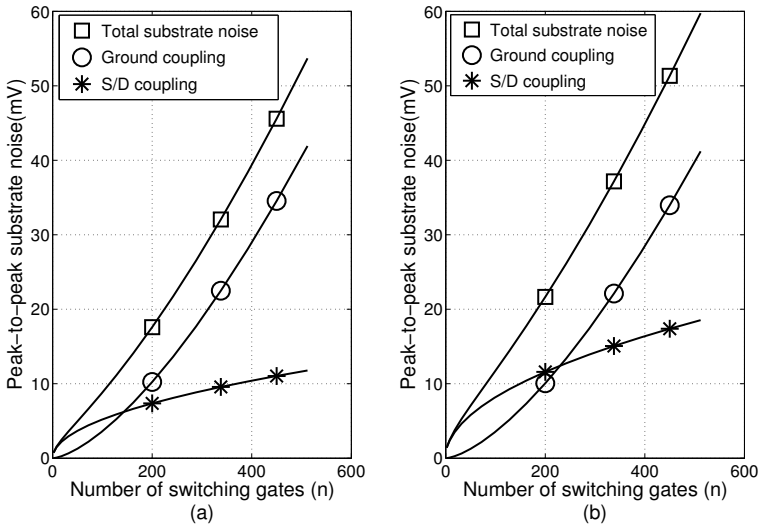


Figure 6.5: Number of simultaneously switching gates vs. substrate noise as predicted by (6.4) when $(t_r)_i = 250$ ps, $L_g = 1$ nH, $C_d = 10$ pF, $R_g = 2.2\ \Omega$, $R_d = 0.1\ \Omega$, $R_{bb} = 16.8$ kΩ, $R_{cb} = 10.7$ kΩ, $R_{dist} = 40$ kΩ, $R_{sc} = 660\ \Omega$, $R_{ang} = 2.2\ \Omega$, and $L_{ang} = 1$ nH. (a) Each gate has a substrate contact, (b) two gates share one substrate contact.

An example of source/drain coupling, ground coupling, and the total noise vs. number of switching gates is shown in Fig. 6.5, as predicted based on the model illustrated in Fig. 6.2. For a small number of switching gates, source/drain coupling dominates the ground coupling. For a larger number of switching gates, the ground coupling increases at a faster rate as compared to the source/drain coupling due to an increase in the overall supply current and number of contacts. The noise injected from the source/drain coupling is primarily filtered by these contacts rather than propagated towards the sense node. The source/drain coupling component of the substrate noise is therefore primarily produced by those gates closest to the sense node. At a certain number of switching gates, the ground coupling becomes larger than the

source/drain coupling. Note that this crossover number is higher in Fig. 6.5(b) where the two gates share one contact as opposed to Fig. 6.5(a) where a contact exists for each gate.

Ground coupling starts to dominate source/drain coupling beyond this crossover point. For large scale circuits with a significant number of switching gates, ground coupling is expected to be the dominant substrate noise generation mechanism. The source/drain coupling is effective only for those small number of gates which are sufficiently close to the sense node. For localized noise analysis, however, the effect of source/drain coupling cannot be neglected. Note that the specific number of switching gates where the crossover occurs is highly dependent on the rise time, parasitic inductance, and decoupling capacitance, as explained in the following section.

6.3.2 Effect of Rise Time, Inductance, and Capacitance on Dominant Noise Coupling Mechanism

The peak-to-peak ground noise is a function of the rise time of the current load, parasitic inductance of the ground network, and the decoupling capacitance in the circuit, as specified by (4.19). Correspondingly, these parameters determine the dominant substrate noise generation mechanism by affecting the number of switching gates at which ground coupling surpasses source/drain coupling. These crossover points are numerically determined at each rise time using (6.4) to quantify and compare those

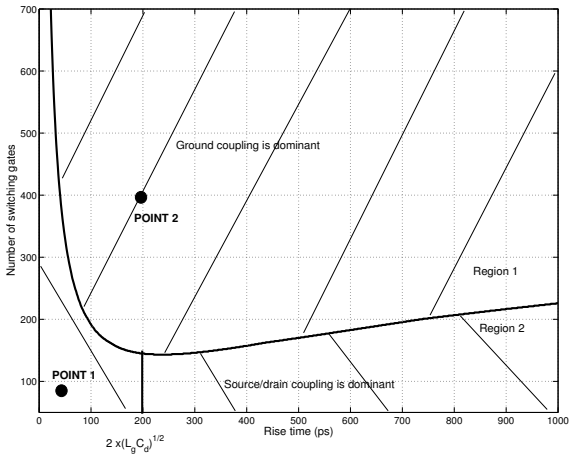


Figure 6.6: For each rise time, the number of switching gates for which source/drain coupling is equal to ground coupling is determined from (6.4). Regions 1 and 2 represent, respectively, the area where ground and source/drain coupling is dominant. The operating parameters are $L_g = 1$ nH, $C_d = 10$ pF, $R_g = 2.2\ \Omega$, $R_d = 0.1\ \Omega$, $R_{bb} = 16.8$ kΩ, $R_{cb} = 10.7$ kΩ, $R_{dist} = 40$ kΩ, $R_{sc} = 660\ \Omega$, $R_{ang} = 2.2\ \Omega$, and $L_{ang} = 1$ nH.

regions where ground and source/drain coupling are dominant. The results are illustrated in Fig. 6.6. At each rise time, the number of switching gates at which ground coupling is equal to source/drain coupling is illustrated. Hence, the area above the curve represents the region where ground coupling is dominant (region 1) and, correspondingly, source/drain coupling is dominant under the curve (region 2). Note that this graph is obtained for a specific value of decoupling capacitance and parasitic inductance.

For sufficiently small rise times, the ground noise is relatively low since the decoupling capacitance is effective. The number of switching gates where the crossover occurs is therefore the greatest for small rise times. This crossover point decreases

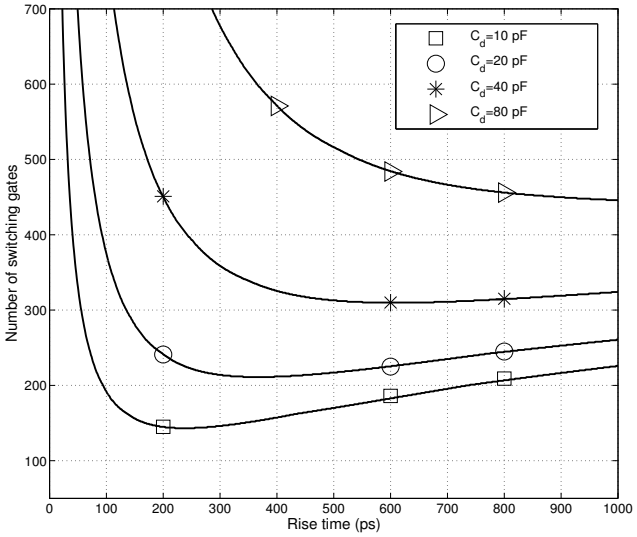


Figure 6.7: Effect of decoupling capacitance on the dominant noise coupling mechanism. In this figure, the data shown in Fig. 6.6 are obtained for different decoupling capacitances while the other circuit parameters are maintained the same.

as the rise time increases and is smallest at $t_r \approx 2\sqrt{(L_g C_d)}$ where the ground noise is greatest, maximizing the area of region 1. As the rise time further increases, the ground noise decreases due to lower $L\ \partial i/\partial t$ noise, increasing the area of region 2. Note that for small rise times or, equivalently, at higher operating frequencies, source/drain coupling becomes the significant noise injection mechanism.

The same graph is obtained for different decoupling capacitances and parasitic inductances to evaluate the effect of these parameters on the dominant noise generation mechanism, as illustrated in Figs. 6.7 and 6.8, respectively. As the parasitic inductance decreases or the decoupling capacitor increases, the area of region 1 decreases while the area of region 2 increases. For example, at $t_r = 300$ ps, the number of

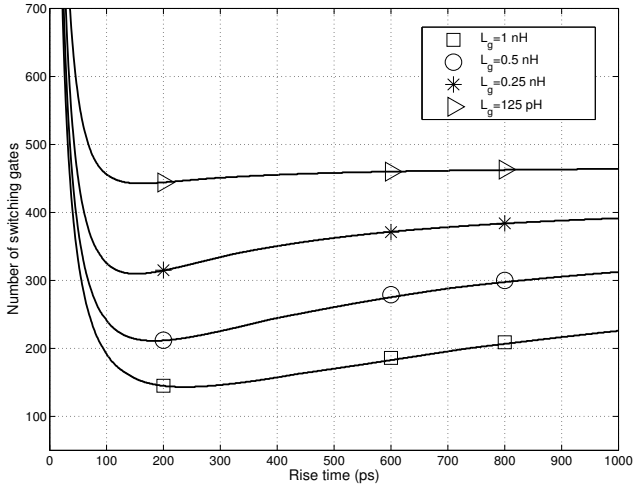


Figure 6.8: Effect of parasitic inductance on the dominant noise coupling mechanism. In this figure, the data shown in Fig. 6.6 are obtained for different parasitic inductances while the other circuit parameters are maintained the same.

switching gates n where the ground coupling is equal to the source/drain coupling is 146 for $C_d = 10$ pF. Alternatively, if $C_d = 40$ pF and 80 pF, n increases, respectively, to 359 and 677, as shown in Fig 6.7. Similarly, at the same rise time, ground coupling and source drain coupling are equal at $n = 146$ for $L_g = 1$ nH and at $n = 334$ and 451 for $L_g = 0.25$ nH and 125 pH, respectively, as shown in Fig 6.8. Thus, for circuits with flip-chip packages and sufficiently high decoupling capacitance, source/drain coupling cannot be neglected and may be the dominant substrate noise generation mechanism.

6.4 Parameter Sensitivity

As described in the previous section, the dominant noise injection mechanism is determined by multiple circuit parameters. Correspondingly, the sensitivity of noise to these parameters varies with respect to the operating point and the dominant noise source. As such, a particular circuit level noise reduction technique may be more efficient as compared to other techniques for a certain set of operating points. The normalized noise sensitivity as a function of rise time and number of switching gates is evaluated in this section based on the model illustrated in Fig. 6.1. The normalized sensitivity S of the substrate noise $(V_{s-total})_{pp}$ to a parameter p_i is

$$S_{p_i}^{(V_{s-total})_{pp}} = \lim_{\Delta p_i \to 0} \frac{\frac{\Delta (V_{s-total})_{pp}}{(V_{s-total})_{pp}}}{\frac{\Delta p_i}{p_i}} = \frac{p_i}{(V_{s-total})_{pp}} \frac{\partial (V_{s-total})_{pp}}{\partial p_i}, \tag{6.5}$$

where $V_{s-total}$ (the total substrate noise at the victim node) is given by (6.3). For multiple switching gates, the resistance R_{cb} is scaled by n where n is the number of switching gates tied to a substrate contact. Alternatively, R_{dist} remains the same, assuming that the analog circuit is sufficiently far from all of the switching gates.

The normalized sensitivity of the substrate noise as a function of rise time and number of switching gates, as determined by (6.5), is shown, respectively, in Figs. 6.9(a) and 6.9(b). The sensitivity of the noise to the decoupling capacitance is high at small rise times and decreases with increasing rise time. Alternatively, the sensitivity to the

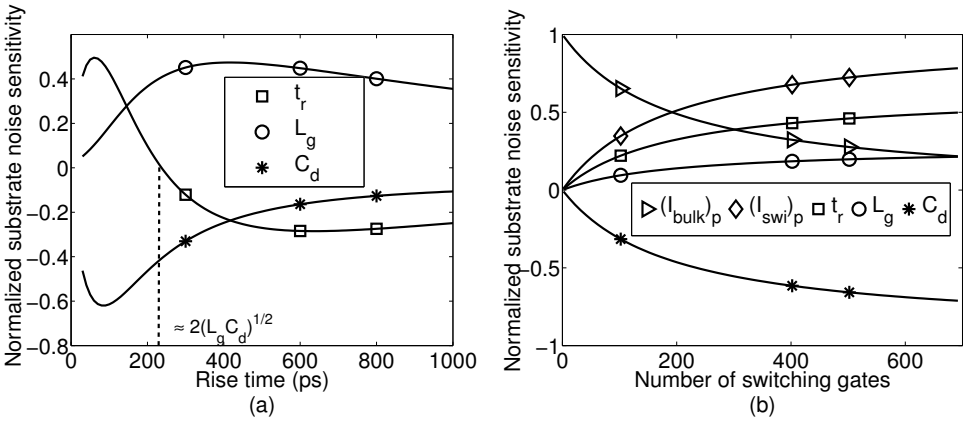


Figure 6.9: Substrate noise sensitivity when $L_g = 1$ nH, $C_d = 10$ pF, $R_g = 2.2\ \Omega$, $R_d = 0.1\ \Omega$, $R_{cb} = 10.7\ k\Omega$, $R_{dist} = 40\ k\Omega$, $R_{vc} = 660\ \Omega$, $R_{ang} = 2.2\ \Omega$, and $L_{ang} = 1$ nH, (a) As a function of rise time when $n = 400$, (b) As a function of the number of switching gates when $t_r = 100$ ps.

parasitic inductance is low at small rise times and increases with longer rise times. This behavior is due to the rise time dependent ratio of the switching current sourced by the decoupling capacitance and the power supply through the parasitic inductance. Note that the sensitivity to the rise time crosses over at zero when $t_r \approx 2\sqrt{(L_g C_d)}$, demonstrating the non-monotonic dependence of noise on the rise time, as shown in Fig 6.4.

The sensitivity to the switching current, parasitic inductance, decoupling capacitance, and rise time increases with a larger number of switching gates, as shown in Fig. 6.9(b), since the ground coupling starts to dominate for large scale circuits. For a small number of switching gates, the sensitivity to the total bulk current is sufficiently high, increasing the ability of the substrate contacts to reduce noise in small scale circuits, as described in the following section.

6.5 Design Implications

The design implications of the proposed macrolevel model to identify the dominant noise source and parameter sensitivities are discussed in this section. Specifically, the efficiency of increasing the substrate contact density, reducing the package and on-chip parasitic inductance, and placing additional on-chip decoupling capacitance is compared as a function of the rise time and number of switching gates. The noise reduction achieved by these techniques is listed in Table 6.2. This comparison can be used to determine the preferable noise reduction technique at early stages of the design process, as further described in the following subsections.

6.5.1 Increasing Substrate Contact Density

For those cases where source/drain coupling dominates, increasing the number of substrate contacts or placing a p+ guard ring around the aggressor circuit achieves enhanced noise reduction as compared to reducing the parasitic inductance or increasing the decoupling capacitance. Alternatively, if ground coupling is the dominant coupling mechanism, placing additional decoupling capacitance and reducing the parasitic inductance are more efficient techniques. This comparison is illustrated by points 1 and 2 in Fig 6.6, which represent, respectively, the dominance of ground coupling and source/drain coupling. For point 2, the peak-to-peak substrate noise is reduced by 31% by doubling the number of substrate contacts. Lowering the parasitic

Table 6.2: Effect of the decoupling capacitance, parasitic inductance, and substrate contact density on reducing the peak-to-peak substrate noise at various operating points.

Rise time (ps)	Number of switching gates 50		
	$C_d = 10$ pF	$C_d = 20$ pF	Reduction
70	4.5 mV	4.1 mV	8.9%
200	4.9 mV	4.5 mV	8.2%
400	4.8 mV	4.5 mV	6.3%
800	4.6 mV	4.4 mV	4.3%
	$L_g = 1$ nH	$L_g = 0.5$ nH	Reduction
70	4.5 mV	4.4 mV	2.2%
200	4.9 mV	4.6 mV	6.1%
400	4.8 mV	4.4 mV	8.3%
800	4.6 mV	4.3 mV	6.5%
	$d_{sc} = \frac{0.5}{gate}$	$d_{sc} = \frac{1}{gate}$	Reduction
70	6.6 mV	4.5 mV	31.8%
200	7.1 mV	4.9 mV	31%
400	6.9 mV	4.8 mV	30.4%
800	6.7 mV	4.6 mV	31.3%
Rise time (ps)	Number of switching gates 200		
	$C_d = 10$ pF	$C_d = 20$ pF	Reduction
70	13.5 mV	10.4 mV	22.9%
200	17.5 mV	13.5 mV	22.8%
400	16.7 mV	14.3 mV	14.4%
800	14.4 mV	13.3 mV	7.6%
	$L_g = 1$ nH	$L_g = 0.5$ nH	Reduction
70	13.5 mV	12.6 mV	6.7%
200	17.5 mV	14.3 mV	18.3%
400	16.7 mV	13.4 mV	19.8%
800	14.4 mV	12.3 mV	14.6%
	$d_{sc} = \frac{0.5}{gate}$	$d_{sc} = \frac{1}{gate}$	Reduction
70	17.7 mV	13.5 mV	23.7%
200	21.6 mV	17.5 mV	19%
400	20.7 mV	16.7 mV	19.3%
800	18.5 mV	14.4 mV	22.2%
Rise time (ps)	Number of switching gates 700		
	$C_d = 10$ pF	$C_d = 20$ pF	Reduction
70	55.5 mV	34.4 mV	38%
200	83.6 mV	56 mV	33%
400	78 mV	61.5 mV	21.2%
800	62.6 mV	55.1 mV	12%
	$L_g = 1$ nH	$L_g = 0.5$ nH	Reduction
70	55.5 mV	49.2 mV	11.4%
200	83.6 mV	61.7 mV	26.2%
400	78 mV	55.2 mV	29.2%
800	62.6 mV	47.7 mV	23.8%
	$d_{sc} = \frac{0.5}{gate}$	$d_{sc} = \frac{1}{gate}$	Reduction
70	63 mV	55.5 mV	11.9%
200	90.5 mV	83.6 mV	7.6%
400	85 mV	78 mV	8.2%
800	70 mV	62.6 mV	10.6%

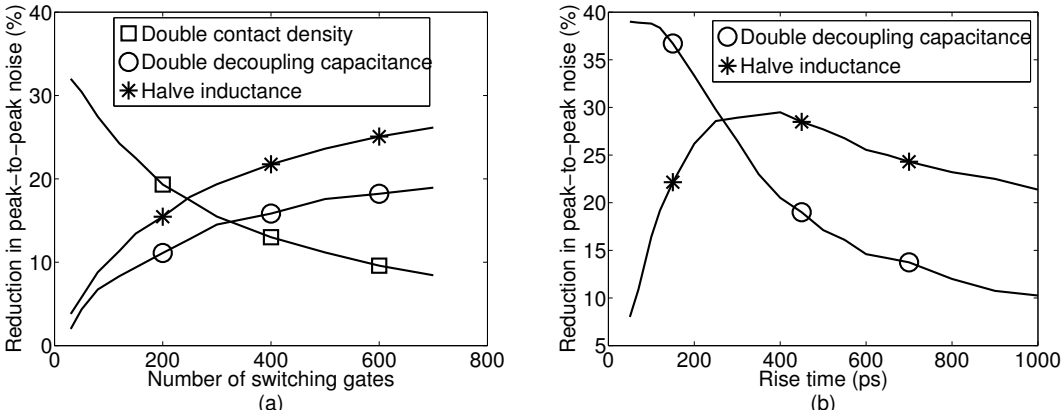


Figure 6.10: Comparison of noise reduction techniques when $L_g = 1$ nH, $C_d = 10$ pF, $R_g = 2.2\ \Omega$, $R_d = 0.1\ \Omega$, $R_{ang} = 2.2\ \Omega$, and $L_{ang} = 1$ nH: (a) As a function of the number of switching gates at $t_r = 400$ ps. (b) As a function of the rise time when $n = 700$.

inductance by a factor of four reduces the noise by only 3.5%. Similarly, increasing the decoupling capacitance by a factor of four reduces the noise by 10.5%. Alternatively, for point 1, where ground coupling is dominant, doubling the number of substrate contacts achieves a 12.1% reduction in noise while reducing the parasitic inductance and increasing the decoupling capacitance, each by a factor of four, reduces the noise by, respectively, 34.1% and 42.8%. The efficiency of increasing the substrate contact density is compared with reducing the parasitic inductance and increasing the decoupling capacitance in Fig. 6.10(a), demonstrating the significance of the number of contacts on small scale circuits where source/drain coupling is dominant.

6.5.2 Increasing Decoupling Capacitance vs. Reducing Parasitic Inductance

The efficiency of placing additional decoupling capacitance and reducing the parasitic inductance is a strong function of rise time, as illustrated by the sensitivities shown in Fig. 4.7(a). The efficiency of these two techniques is compared in Fig 6.10(b). At $t_r = 70$ ps, doubling the decoupling capacitance achieves a 39% reduction in the peak-to-peak substrate noise where $2\sqrt{L_g C_d} = 200$ ps. Halving the parasitic inductance, however, achieves a reduction of only 11%. Alternatively, at $t_r = 800$ ps, halving the parasitic inductance achieves enhanced noise reduction of 23%, while doubling the decoupling capacitance reduces the noise by 12%. Specifically, increasing the decoupling capacitance is effective for $t_r \ll 2\sqrt{L_g C_d}$, while reducing the parasitic inductance is effective for $t_r \gg 2\sqrt{L_g C_d}$.

6.6 Summary

A substrate coupling model for multiple switching gates is presented for macrolevel analysis of the various substrate noise coupling mechanisms. The model accurately captures the effects of multiple parameters, as validated by SPICE. The proposed model identifies the dominant noise source at the early stages of the design process as a function of multiple parameters. Specifically, the regions where ground coupling

and source/drain coupling dominate can be determined based on this model. Ground coupling tends to dominate if a larger number of gates are switching. For a sufficiently high decoupling capacitance and low parasitic inductance, such as a flip-chip package, source/drain coupling is shown to be the dominant noise coupling mechanism. Identification of the dominant noise source and parameter sensitivity analysis are used to determine the most efficient noise reduction technique.

Chapter 7

Efficient Substrate Noise Analysis in Large Scale Mixed-Signal Circuits

Reasonably accurate and computationally efficient estimation of the substrate noise is an important design issue. Overestimating the substrate noise can result in the use of overly conservative noise isolation techniques, consuming excessive resources. Alternatively, underestimating the noise can cause a circuit to fail. Efficient and reasonably accurate analysis of the substrate coupling noise at the boundary of a sensitive block is, however, a challenging task due to the high computational complexity of the substrate extraction process, as described in Chapter 3. A high level analysis methodology that exhibits low computational complexity while achieving reasonable accuracy is therefore necessary.

A methodology is introduced in this chapter to efficiently analyze the substrate

noise generated by an aggressor block. The methodology is based on identifying voltage domains within the substrate of an aggressor circuit. A voltage domain represents a region within the substrate that is biased with approximately the same voltage by substrate contacts, and is therefore effectively shorted. These voltage domains are determined from the difference in the transient voltage among the substrate contacts. Each domain is represented by a single equivalent contact, thereby reducing the overall number of input ports that need to be extracted. For n number of input ports, a minimum of $(n^2 - n)/2$ number of substrate resistors are required to model the substrate, assuming a quasi-static approximation. A reduction in the number of input ports, therefore, quadratically reduces the number of extracted substrate resistors.

A linear time algorithm is proposed to determine the voltage domains within a substrate by analyzing the transient voltage differences among the substrate contacts. Those contacts exhibiting a voltage difference smaller than a specified value are merged, and an equivalent contact is placed at the geometric mean of the merged contacts. The impedance of the ground network is updated to maintain accuracy. This methodology achieves more than four orders of magnitude reduction in the number of substrate resistors as compared to a detailed extraction of the circuit at the expense of 24% error in the peak-to-peak substrate noise voltage.

The rest of the chapter is organized as follows. The concept of voltage domains within the substrate to reduce computational complexity is described in Section 7.1.

The proposed methodology is introduced in Section 7.2. Simulation results of two different circuits are described in Section 7.3. Limitations of the methodology are discussed in Section 7.4. Finally, the chapter is summarized in Section 7.5.

7.1 Voltage Domains Within the Substrate

In a mixed-signal circuit, a common approach to bias the substrate of a digital block is to connect the substrate to the digital ground network with substrate contacts. Due to the parasitic impedance of the ground network, each substrate contact has an $IR + L\ \partial i/\partial t$ voltage bounce which resistively couples into the substrate. As such, if the voltage variation between a set of substrate contacts is sufficiently small, the corresponding area of the substrate is effectively short-circuited by these contacts, as illustrated in Fig. 7.1.

The transient voltage difference between two contacts C_1 and C_2 is determined by

$$V_{C1} - V_{C2} = V_{12} = i(t)_{12}R_{12} + L_{12}\frac{\partial i(t)_{12}}{\partial t}, \tag{7.1}$$

where $i(t)_{12}$ is the transient current of the ground network flowing from C_1 to C_2 injected by the switching gates. R_{12} and L_{12} are, respectively, the parasitic resistance and inductance of the ground network between C_1 and C_2. Referring to Fig. 7.1, the transient voltage difference among the contacts C_1, C_2, and C_3 and among C_4, C_5,

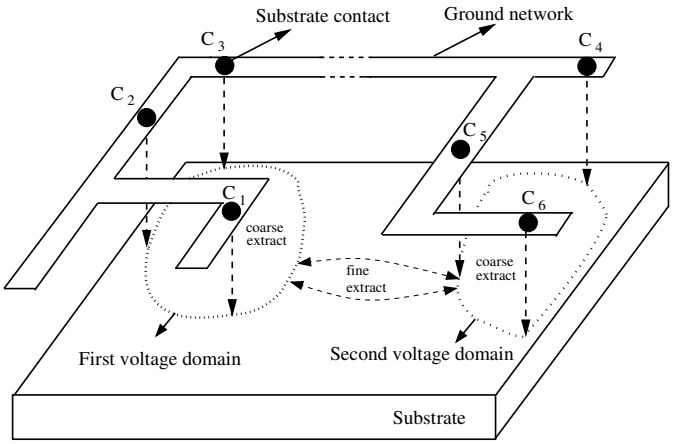


Figure 7.1: Identifying the voltage domains within the substrate. Assuming $V_{C1} \approx V_{C2} \approx V_{C3}$ and $V_{C4} \approx V_{C5} \approx V_{C6}$, two voltage domains are created by the first and last three contacts. A coarse extraction is performed within each domain to reduce the computational complexity, followed by a fine extraction of those domains where the dominant current flow occurs.

and C_6 is assumed to be sufficiently small such that, respectively, $V_{C1} \approx V_{C2} \approx V_{C3}$ and $V_{C4} \approx V_{C5} \approx V_{C6}$. As a result, the corresponding area biased by the first three contacts determines the first voltage domain on the substrate and, similarly, the last three contacts determine the second voltage domain. Since the voltage variations within a domain are sufficiently small, the dominant current flow occurs *among* these voltage domains. The small spatial voltage differences within the ground network can therefore be exploited to reduce the overall number of input ports for extraction.

An algorithm is described in this chapter to identify these voltage domains on the substrate. An equivalent contact is created for each domain while neglecting the other ports within that domain. The number of input ports is reduced, significantly improving the computational complexity of the extraction process.

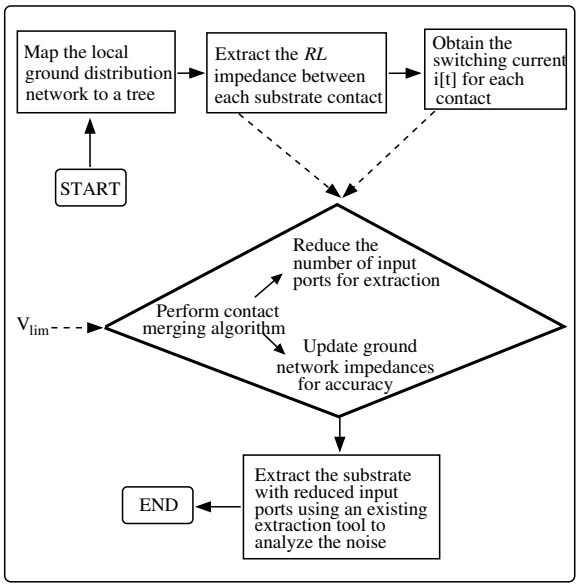


Figure 7.2: Flowchart summarizing the proposed methodology to analyze substrate noise in a large scale mixed-signal circuit.

7.2 Substrate Noise Analysis Methodology

The proposed methodology for efficiently estimating the substrate noise generated by an aggressive digital circuit consists of five steps, as described below and also illustrated in the flowchart shown in Fig. 7.2.

- **Step 1.** The local ground distribution network is extracted to obtain the parasitic resistance and inductance between each substrate contact. In a digital integrated circuit, the power/ground network is designed in a hierarchical approach, as shown in Fig. 7.3 [12], [197]. The upper metal layers span the entire die, forming a regular global mesh. The purpose of these upper metal layers is to distribute power/ground while minimizing the ohmic and inductive losses

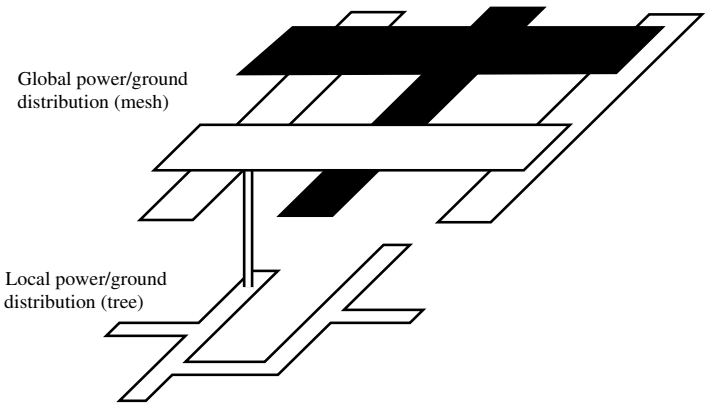


Figure 7.3: In a typical power/ground distribution network, upper metal layers span the entire die, forming a global mesh. The power/ground network is distributed to the devices in each macroblock through the local distribution network, represented as a tree.

throughout the die. Power is supplied to the devices in each macroblock by means of the local power distribution network which can be generally represented as a tree. Note that the substrate contacts are located on the first metal layer, and therefore, are part of the local ground network. The proposed approach assumes an ideal ground for those points where the local tree is connected to the global mesh. Consequently, the local ground network can be mapped to a tree data structure where each substrate contact is a node and the ideal ground is the root of the tree.

- **Step 2.** The current injected into each substrate contact by the switching circuit is characterized over a specific time window for a specific set of input vectors generating sufficient switching activity. For a large digital block, these current

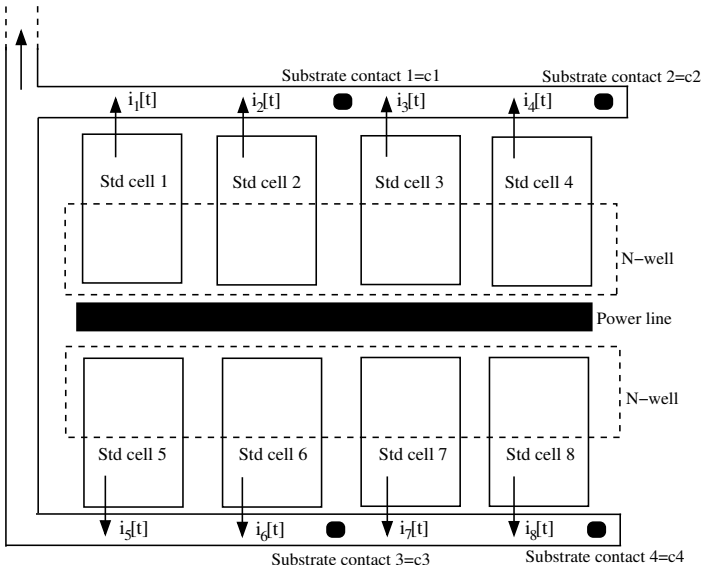


Figure 7.4: The current injected into each substrate contact by the switching circuit is characterized over a specific time window. The current injected by those cells located between two contacts is shifted to the previous contact such that $i_{c1}[t] = i_1[t] + i_2[t]$, $i_{c2}[t] = i_3[t] + i_4[t]$, $i_{c3}[t] = i_5[t] + i_6[t]$, and $i_{c4}[t] = i_7[t] + i_8[t]$.

profiles can be obtained by pre-characterizing each standard cell within each library followed by a behavioral simulation of the circuit to extract the switching time of each cell [123], [148]. The current injected by those cells located between two contacts is shifted to the previous contact to prevent overly optimistic results, as illustrated in Fig. 7.4. The total switching current injected into the first substrate contact (c_1) is equal to $i_1[t] + i_2[t]$. Similarly, for the second substrate contact (c_2), the injected current is equal to $i_3[t] + i_4[t]$. The transient voltage difference between c_1 and c_2 is therefore approximated as $(i_3[t] + i_4[t])Z(12)$ where $Z(12)$ is the RL impedance between the contacts, as determined by Step 1. Note that this approximation is pessimistic since the average current flowing

through this impedance is, in reality, less than $i_3[t] + i_4[t]$.

- **Step 3.** The proposed algorithm is performed to determine the voltage domains based on the data obtained from the first two steps and an additional parameter V_{lim} that defines the condition when to merge a set of contacts. The ground network impedances are updated based on the algorithm to maintain sufficient accuracy. Note that V_{lim} provides flexibility to exploit accuracy versus complexity tradeoffs, as described in Section 7.3.

- **Step 4.** For each voltage domain determined in Step 3, an equivalent contact with the same physical size is placed at the geometric mean of the merged contacts. All of the remaining ports into the substrate within that voltage domain, such as the source/drain regions of the devices, are neglected to reduce the computational complexity.

- **Step 5.** The substrate is extracted with these equivalent contacts which are also connected to the updated ground network. The resulting netlist is analyzed to determine the substrate noise at the sense node located around the sensitive block.

A theoretical analysis of the methodology is provided in Section 7.2.1. The contact merging algorithm identifying the voltage domains and creating an equivalent contact for each domain is described in Section 7.2.2. Bounds on the error in estimating the

substrate noise are discussed in Section 7.2.3. The time complexity of the contact merging algorithm is analyzed in Section 7.2.4.

7.2.1 Theoretical Analysis

Each substrate contact behaves as a noise source, injecting ground noise into the substrate. These contacts are considered as input ports for the extraction of the substrate. In the best case, $(n^2 - n)/2$ number of substrate resistances are required to model the substrate with n number of input ports. If, however, the noise on a set of input ports is sufficiently close, these ports identify a voltage domain that can be represented by an equivalent port, reducing the overall number of input ports that need to be extracted.

This reduction in the number of input ports is demonstrated by a simple example, as illustrated in Fig. 7.5, where the substrate is modeled as a two-dimensional resistive mesh. Points A and B represent two substrate contacts with noise voltages, respectively, V and $V - \epsilon$. Note that these noise voltages are due to the switching current flowing through the parasitic impedance of the ground network. R_c represents the substrate contact resistance. K is a location on the substrate where the distance among the points A, B, and K is identical. The sense node is connected to the analog ground with a resistance R_a. This circuit is modeled with equivalent resistances, as shown in Fig. 7.6(a), where R_s is the equivalent substrate resistance among points

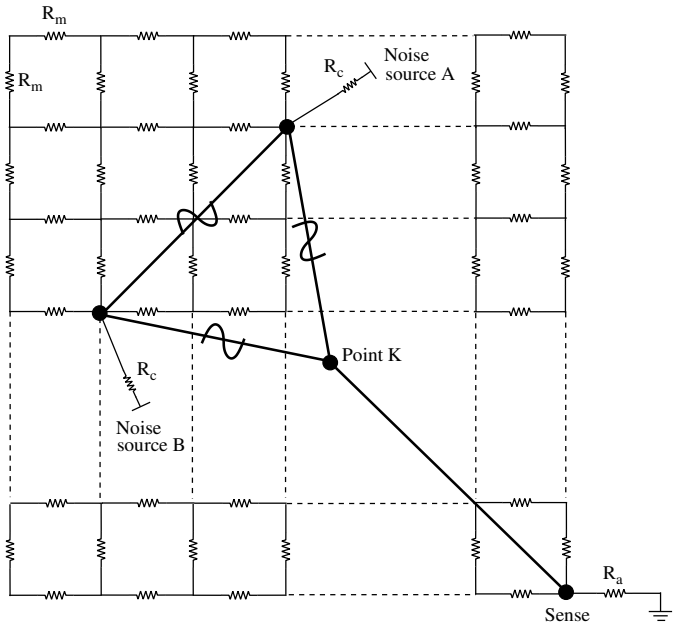


Figure 7.5: Two substrate contacts A and B acting as noise sources, where the substrate is represented as a 2-D resistive mesh.

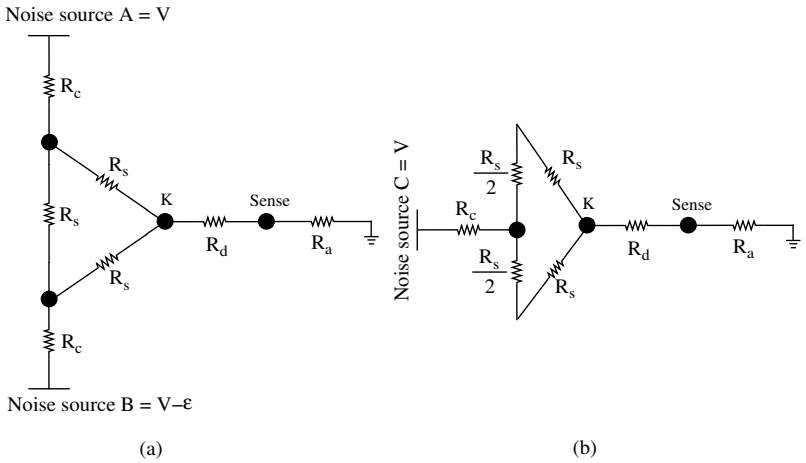


Figure 7.6: Equivalent circuit to analyze the impact of two noise sources with similar noise voltages on the substrate: (a) Before merging. (b) After merging.

A, B, and K, and R_d is the equivalent substrate resistance between point K and the sense node. Based on Fig. 7.6(a), the noise V_{ab}^{s} at the sense node due to two noise

sources A and B is

$$V_{ab}^{s} = \frac{R_a(2V - \epsilon)}{2R_a + 2R_d + R_s + R_c}. \tag{7.2}$$

Assuming the two noise sources A and B are merged into a single noise source C, as shown in Fig. 7.6(b), the noise V_{eqn}^{s} at the sense node due to this equivalent noise source is

$$V_{eqn}^{s} = \frac{4R_a V}{4R_a + 4R_d + 3R_s + 4R_c}. \tag{7.3}$$

The error in merging the two noise sources into an equivalent noise source is

$$E = |V_{ab}^{s} - V_{eqn}^{s}|. \tag{7.4}$$

The contact resistance R_c is typically in the range of ohms and therefore much smaller than the substrate resistance R_s which is in the range of kilo-ohms for a bulk type substrate [198]. Similarly, R_s is much smaller than R_d assuming that the sense node is far from the noise sources. The error can therefore be approximated as

$$E \approx \frac{R_a \epsilon}{2R_a + 2R_d}. \tag{7.5}$$

Since R_a is also much smaller than R_d, the error can be further simplified to

$$E \approx \frac{R_a \epsilon}{2R_d}. \tag{7.6}$$

Note that ϵ is the voltage difference between the substrate contacts A and B that are merged into a single noise source C. This voltage difference is scaled by $R_a/2R_d$ where $R_d \gg R_a$ to determine the error at the sense node on the substrate. The error therefore grows with increasing ϵ (or V_{lim} in the proposed algorithm) and decreases with increasing physical distance between the noise sources and the sense node.

As an example, the substrate is modeled as a 25 x 25 mesh where the the unit resistance R_m and the substrate contact resistance R_c are determined, respectively, as 3 KΩ and 18 Ω for a 90 nm CMOS technology with $V_{dd} = 1.2$ volts with a bulk type substrate. Assuming the ground bounce is within 15% of V_{dd}, two substrate contacts with noise voltages, respectively, 80 mV and 100 mV, are placed on the substrate as illustrated in Fig. 7.5. The resistance R_a is extracted as 1.4 KΩ using SubstrateStorm assuming the sense node is placed within a p+ guard ring with ten analog substrate contacts. The circuit is analyzed using SPICE and the noise at the sense node is determined as 12.8 mV. If these two noise sources are merged, the noise at the sense node is 13.2 mV with a single equivalent noise source. Noise sources with different noise voltages and the error at the sense node obtained from SPICE simulations and (7.4) are listed in Table 7.1. Note that the equivalent resistances R_s and R_d are determined from the mesh, respectively, as 3.4 KΩ and 8.54 KΩ. The SPICE results obtained from simulating the mesh are compared with (7.4). Note that the error increases with increasing ϵ.

Table 7.1: Error comparison due to merging contacts obtained by SPICE and (7.4) for different noise voltages.

	V_a (mV)	V_b (mV)	Noise at sense (mV)	Error (mV) SPICE	Error (mV) Eqn. (7.4)
Before merge	80	100	12.8	0.4	0.37
After merge	100		13.2		
Before merge	80	120	14.3	1.5	1.41
After merge	120		15.8		
Before merge	80	160	17.1	3.9	3.5
After merge	160		21		

This example demonstrates that the noise sources with approximately equal voltages can be represented with a single noise source with sufficiently small error due to the mesh structure of the substrate. The proposed algorithm to reduce the overall number of input ports *before* extracting the substrate by exploiting this characteristic is described in the following section.

7.2.2 Contact Merging Algorithm

The extracted local ground network (including the substrate contacts) of the aggressor circuit is mapped to a tree data structure where each substrate contact and intersection are represented as nodes. The root of the tree is represented by the ideal ground where the local tree is connected to the global mesh. Each node in the tree is characterized by seven elements:

- $i_{[t]}(node)$ represents the switching current profile *injected* into the node by the switching gates. Note that each current profile includes the switching time

information which is obtained at a specific time window and stored in an array at discrete time points. This timing information is obtained through a gate level behavioral simulation of the digital circuit.

- $i_{out[t]}(node)$ represents the total switching current profile flowing from the node towards the parent of the node.
- $R(node)$ represents the parasitic resistance of the ground interconnect between the node and the parent of the node.
- $L(node)$ represents the parasitic inductance of the ground interconnect between the node and the parent of the node.
- $C_x(node)$ represents the x_{th} child of the node.
- $nc(node)$ represents the number of children of the node.
- $V_{diff}(node)$ represents the peak value of the transient voltage difference between the node and the parent of the node.

An example structure is shown in Fig. 7.7 to illustrate these elements, and the inputs and outputs of the algorithm.

The algorithm traverses the entire tree starting from the leaf nodes to evaluate the voltage difference $V_{diff}(node)$ between each node and parent. If the peak value of this transient voltage difference is smaller than the limit voltage V_{lim}, which is an

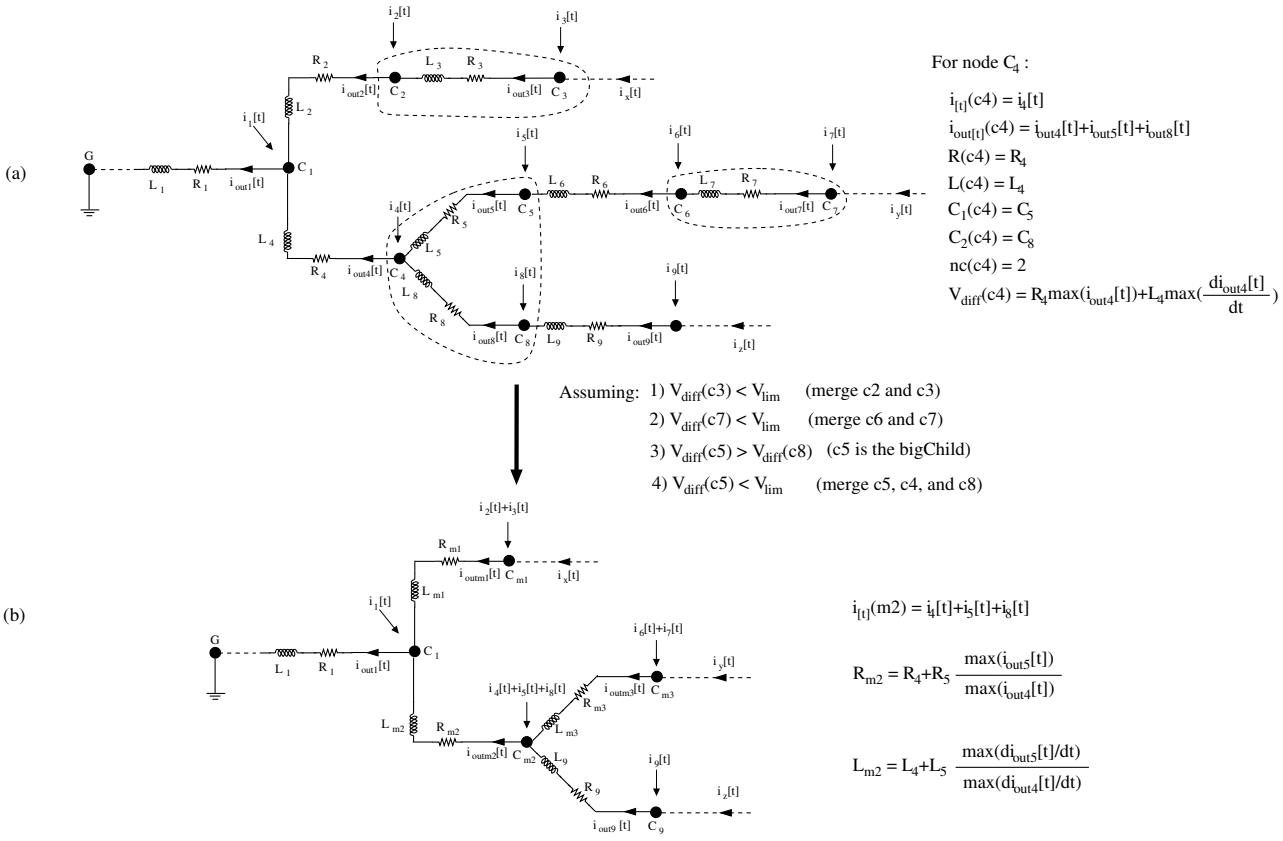


Figure 7.7: First metal layer of the digital ground network mapped to a tree where each node represents a substrate contact and the root is assumed to be the ideal ground where the local tree is connected to the global mesh: (a) Before merging. All seven elements are illustrated for node C_4, (b) After merging. Note that $i_{[t]}(node)$, $R(node)$, and $L(node)$ are updated as shown for node C_{m2} to maintain the voltage on the node with the least error.

input parameter, those nodes are merged into a single node and the node voltage is updated by modifying the resistance and inductance to maintain the absolute voltage with the least error. Note that rather than calculating the absolute voltages to determine whether the nodes should be merged, the voltage *difference* between the nodes, based on the current profile and the parasitic impedances, is sufficient. The analysis of the voltage difference rather than the absolute voltages significantly reduces the complexity and memory requirements of the algorithm.

Pseudo-code of the proposed recursive algorithm MERGE-CONTACTS is provided in Fig. 7.8. The algorithm starts with the *root* and V_{lim}, as specified by the user, for the first and second arguments, respectively. In lines 1-4, MERGE-CONTACTS is recalled for each node in the tree until the leaves are reached. The peak voltage difference $V_{diff}(node)$ between each leaf and the parent is calculated in line 21. In lines 6-13, the child with the greatest voltage difference between the parent is identified and called *bigChild*. If this voltage difference is smaller than V_{lim}, all of the children and the parent are merged into one node and the resistance, inductance, and the switching current profile of the merged node are updated in lines 15-17 to maintain the original transient voltage with the least error. The procedure for updating the impedances is illustrated in Fig. 7.9 for a simpler case. Assuming the peak voltage difference between C_1 and C_2 is less than V_{lim}, these two nodes are merged into C_m. The resistance and inductance of C_m are incremented, respectively, by ΔR and ΔL

MERGE-CONTACTS(node, V_{lim})

1. **if** node =! leaf
2. **for** x = 1:1:nc(node)
3. MERGE-CONTACTS($C_x(node)$, V_{lim})
4. **end**
5. $i_{out[t]}(node) = i_{[t]}(node) + i_{out[t]}(C_1(node)) + ... + i_{out[t]}(C_k(node))$
6. $V_{(diff-max)} = V_{diff}(C_1(node))$
7. bigChild = $C_1(node)$
8. **for** x=2:1:nc(node)
9. **if** $V_{diff}(C_x(node)) > V_{(diff-max)}$
10. $V_{(diff-max)} = V_{diff}(C_x(node))$
11. bigChild = $C_x(node)$
12. **end**
13. **end**
14. **if** $V_{(diff-max)} < V_{lim}$
15. $R(node) = R(node) + R(bigChild)\frac{max(|i_{out[t]}(bigChild)|)}{max(|i_{out[t]}(node)|)}$
16. $L(node) = L(node) + L(bigChild)\frac{max(|\partial(i_{out[t]}(bigChild))/\partial t|)}{max(|\partial(i_{out[t]}(node))/\partial t|)}$
17. $i_{[t]}(node) = i_{[t]}(node) + i_{[t]}(C_1(node)) + ... + i_{[t]}(C_k(node))$
18. Correct(node, bigChild)
19. **end**
20. **end**
21. $V_{diff}(node) = R(node)max(i_{out[t]}(node)) + L(node)max(\frac{\partial(i_{out[t]}(node))}{\partial t})$

Figure 7.8: Pseudo-code to merge the substrate contacts on the ground network based on spatial transient voltage differences.

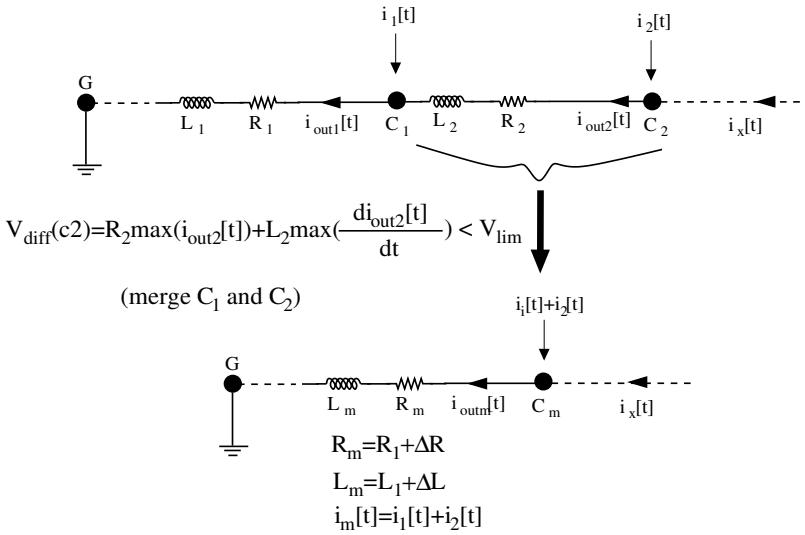


Figure 7.9: Illustration of impedance updated after merging two nodes. The resistance and inductance are incremented, respectively, by ΔR and ΔL to maintain the original voltage with the least error.

to compensate for the voltage loss caused by merging C_1 and C_2 such that

$$i_{out2}[t]R_2 + L_2\frac{\partial i_{out2}[t]}{\partial t} = \Delta R i_{outm}[t] + \Delta L\frac{\partial i_{outm}[t]}{\partial t}. \tag{7.7}$$

Since i_{outm} is equal to i_{out1}, ΔR and ΔL are given by, respectively,

$$\Delta R = R_2\frac{max(|i_{out2}[t]|)}{max(|i_{out1}[t]|)}, \tag{7.8}$$

$$\Delta L = L_2\frac{max(|\partial i_{out2}[t]/\partial t|)}{max(|\partial i_{out1}[t]/\partial t|)}. \tag{7.9}$$

Note that the algorithm maintains the peak value of the absolute voltage after merging. As such, the maximum value of the currents are considered when updating the

impedance. Another option is to consider the rms value rather than the maximum value. The rms value, however, produces a larger error in the substrate noise. Assuming $i_x[t] = 0$ in Fig. 7.9, (7.8) and (7.9) can be rewritten as

$$\Delta R = R_2 \frac{max(|i_2[t]|)}{max(|i_1[t] + i_2[t]|)}, \tag{7.10}$$

$$\Delta L = L_2 \frac{max(|\partial i_2[t]/\partial t|)}{max(|\partial (i_1[t] + i_2[t])/\partial t|)}. \tag{7.11}$$

The updated resistance and inductance of the merged node C_m are therefore, respectively,

$$R_m = R_1 + \Delta R, \tag{7.12}$$

$$L_m = L_1 + \Delta L. \tag{7.13}$$

If the node has more than one child, merging and updating the impedance of the merged node is achieved based on *bigChild*. For example, in Fig. 7.7(a), C_4 has two children, C_5 and C_8. Assuming $V_{diff}(c5)$ is greater than $V_{diff}(c8)$ and less than V_{lim}, C_5 is identified as the *bigChild*, and C_4, C_5, and C_8 are merged into one node C_{m2}. The resistance and inductance of the merged node are updated based on C_5, as shown in Fig. 7.7(b).

Since the nodes are merged based on the voltage difference of *bigChild*, a correction is required to maintain the original transient voltage for other children to prevent

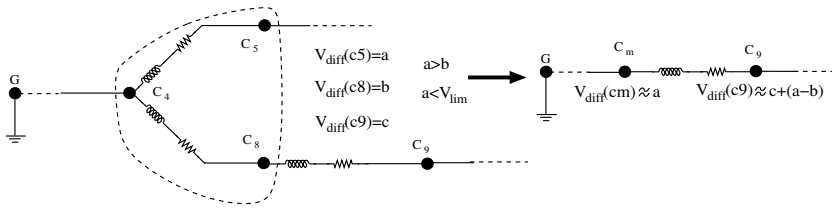


Figure 7.10: Illustration of the *Correct* function. After the nodes C_4, C_5, and C_8 are merged based on $V_{diff}(c5)$, $V_{diff}(c9)$ shifts by $(a - b)$ which is compensated by the *Correct* function.

error accumulation, which is achieved by the *Correct* function in line 18. An example of this process is shown in Fig. 7.10, illustrating the requirement for this function.

Assuming that $V_{diff}(c5) = a$ is greater than $V_{diff}(c8) = b$, and less than V_{lim}; nodes C_4, C_5, and C_8 are merged where *bigChild* is C_5. After merging, the impedance of the merged node C_m is adjusted to make $V_{diff}(cm)$ approximately equal to a. After merging, the new parent of C_9 is C_m, and $V_{diff}(c9)$ shifts by $a - b$. Note that the error for C_9 accumulates with additional merging. In order to prevent this error accumulation, the impedance of C_9 is updated by the *Correct* function to compensate for the error $a - b$. Pseudo-code for the *Correct* function is shown in Fig. 7.11. The voltage required to compensate this error is calculated in lines 2 and 3. The resistance and inductance are correspondingly updated in lines 5 and 6. The upper bounds on the error due to contact merging are discussed in the following section.

Correct(node, bigChild)
1. **for** k = 1:1:nc(node)
2. $V_{corr}(R) = \max[R(bigChild)\ i_{out[t]}(bigChild)$
   - $R(C_k(node))\ i_{out[t]}(C_k(node))]$
3. $V_{corr}(L) = \max[L(bigChild)\frac{\partial i_{out[t]}(bigChild)}{\partial t}$
   - $L(C_k(node))\frac{\partial i_{out[t]}(C_k(node))}{\partial t}]$
4. **for** p = 1:1:$nc(C_k(node))$
5. $R(C_p(C_k(node))) = R(C_p(C_k(node)))$ - $\frac{V_{corr}(R)}{max[|i_{out[t]}(C_p(C_k(node)))|]}$
6. $L(C_p(C_k(node))) = L(C_p(C_k(node)))$ - $\frac{V_{corr}(L)}{max[|\frac{\partial i_{out[t]}(C_p(C_k(node)))}{\partial t}|]}$
7. **end**
8. **end**

Figure 7.11: Pseudo-code of the *Correct* function to prevent error accumulation after merging a set of contacts.

7.2.3 Bounds on Error

As described in the previous section, the decision as to which contacts to merge is achieved based on the transient voltage difference among the contacts over a specific time window. The maximum voltage difference over this time window is compared with V_{lim} to determine whether the contacts can be merged. The resistance and inductance of a merged contact are updated, respectively, in (7.12) and (7.13) to maintain the peak value of the absolute voltage on the ground network. Referring to Fig. 7.9, the error $E[t]$ on the ground network due to merging contacts C_1 and C_2 into an equivalent contact C_m is determined by

$$E[t] = |V_{c2}[t] - V_{cm}[t]|, \tag{7.14}$$

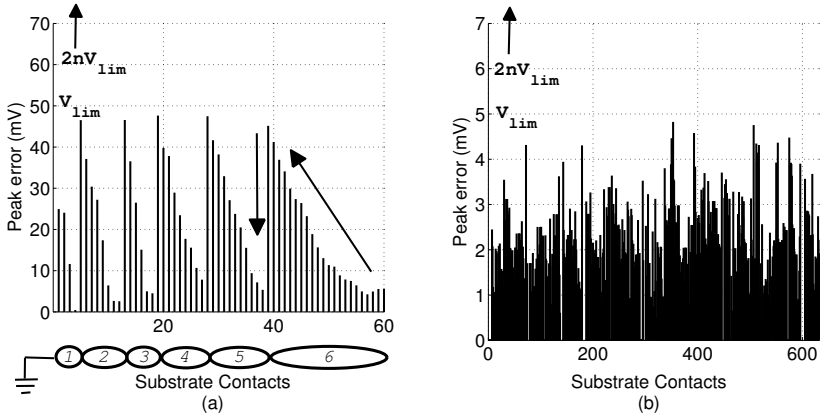


Figure 7.12: Maximum error in the voltage of each contact due to merging: (a) Ground network is composed of a single line with 60 substrate contacts, $V_{lim} = 50$ mV. Six voltage domains are identified. (b) Ground network is composed of a tree with 632 substrate contacts, $V_{lim} = 5$ mV. 105 voltage domains are identified.

where V_{c2} and V_{cm} are the voltages across, respectively, C_2 and C_m. For a single merge operation, the upper bound for this error is $2 \times V_{lim}$ (see Appendix A). For n number of merges resulting in n number of voltage domains, the error is bounded by $2n \times V_{lim}$, assuming the error introduced by each merge accumulates. In practice, however, the error does not reach this value since the error values accumulate only if the maximum error of each merge occurs at the same time. Furthermore, the maximum error does not occur at the peak voltage since the algorithm maintains the peak value of the voltage after merging. An example of the maximum error for each contact is illustrated in Fig. 7.12 for two arbitrary ground networks. In Fig. 7.12(a), the ground network is composed of a single line with 60 substrate contacts. V_{lim} is equal to 50 mV, generating six voltage domains, as illustrated in Fig. 7.12(a). The peak error for each contact is below the upper bound. Note that the error is

sufficiently small at the beginning of a voltage domain and gradually increases to V_{lim} as additional contacts are merged, as illustrated by the arrow in Fig. 7.12(a). When the maximum number of contacts to be merged is reached, the error exhibits a sudden decrease and starts to rise again for the following voltage domain. In Fig. 7.12(b), the error of each contact is depicted for a tree structured ground network consisting of 632 substrate contacts where $V_{lim} = 5$ mV. Note that the voltage error at the substrate contact determined from (7.14) and illustrated in Fig. 7.12 is scaled by the substrate resistance between the contact and the sense node to determine the error in the substrate noise at the sense node, as described in Section 7.2.1.

7.2.4 Complexity Analysis

In order to evaluate the time complexity of MERGE-CONTACTS, the total number of substrate contacts is N, where each contact is represented as a node in the tree. In the worst case, for N number of nodes in the tree, $N - 1$ merges can be achieved. For each merge, lines 5-17 shown in Fig. 7.8 require a time proportional to the number of discrete time points in the current waveforms t. Line 18 requires time proportional to $k \times l$, where k is the number of children of the node and l is the number of grandchildren. The time complexity of MERGE-CONTACTS is therefore $O(N \times (kl + t))$, which reduces to linear time complexity $O(N)$ since k, l, and t are constants. The algorithm, as implemented in Matlab, has been performed for various

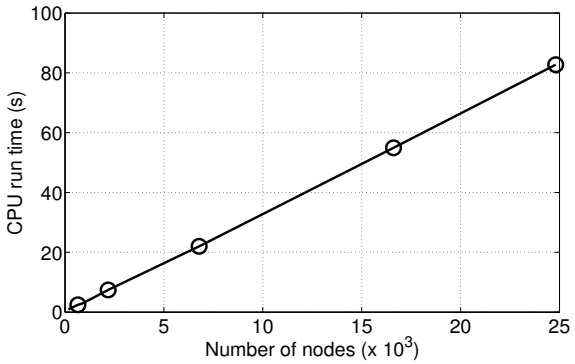


Figure 7.13: Dependency of CPU run time on the number of nodes, demonstrating the linear complexity of the algorithm.

number of nodes. The dependence of the CPU run time on the number of nodes is illustrated in Fig. 7.13, demonstrating the linear complexity of the algorithm.

7.3 Simulation Results

The proposed CONTACT-MERGE algorithm has been implemented in Matlab to evaluate the proposed methodology. Two different aggressor circuits have been analyzed to quantify the computational complexity and accuracy of the methodology. The results obtained from the first and second circuits are described, respectively, in Sections 7.3.1 and 7.3.2. The proper selection of V_{lim} for the algorithm is discussed in Section 7.3.3.

7.3.1 Circuit 1

The first circuit consists of a 4-bit carry select adder, a control unit, and scaled buffers at the output, designed in a 0.18 μm CMOS technology on a bulk type substrate. The parasitic resistance of the ground network is determined from the sheet and via resistances. The sheet resistances are 95 mΩ and 80 mΩ for the first and second metal layer, respectively, and the via resistance is 2 Ω. Note that the parasitic inductance of the ground network is neglected for this first circuit. The switching current waveform ($i_{[t]}(node)$) for each contact is obtained through a transistor level simulation for a specific time window since this circuit has been designed in a full custom design methodology.

The CONTACT-MERGE algorithm is performed to identify the voltage domains on the substrate. Four different voltages (0.05 volts, 0.1 volts, 0.25 volts, and 0.4 volts) are used for V_{lim} to investigate the complexity versus accuracy tradeoff. For $V_{lim} = 0.1$ volts, nine voltage domains are identified. If V_{lim} is increased to 0.25 volts, three voltage domains are determined. Each of these domains is represented by an equivalent substrate contact placed at the geometric mean of the merged contacts. These domains are illustrated in Fig. 7.14. The dashed lines represent the first metal layer and the solid lines represent the second metal layer of the ground network. The substrate contacts are represented by circles. The diamonds represent the intersection of two metal lines in the ground network. Note that the switching current for these

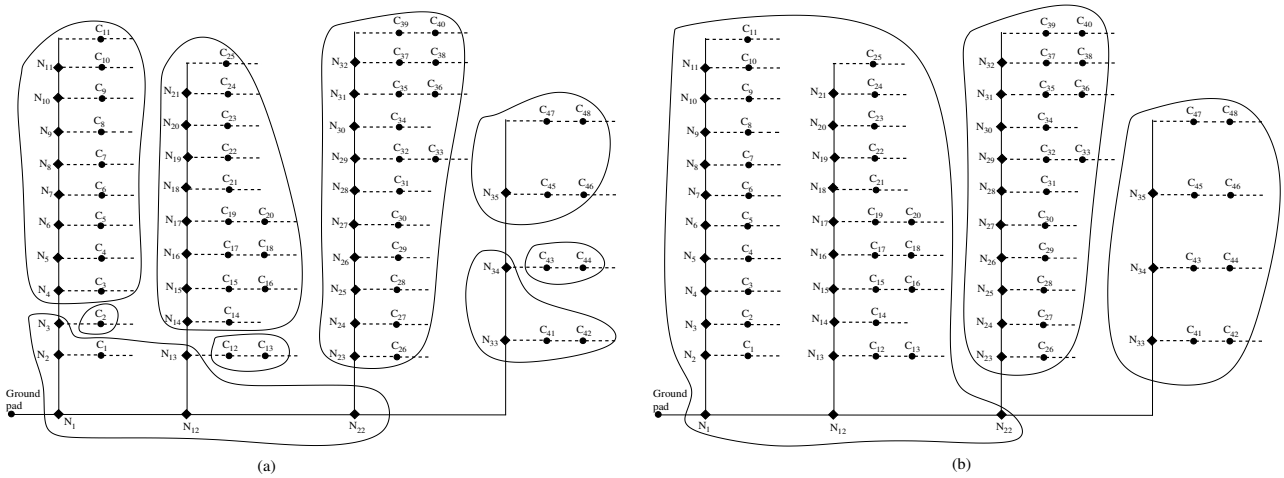


Figure 7.14: Voltage domains on the substrate as determined by the CONTACT-MERGE algorithm: (a) $V_{lim} = 0.1$ volts. Nine voltage domains are identified. (b) $V_{lim} = 0.25$ volts. Three voltage domains are identified.

nodes represented by a diamond is zero since these nodes do not represent substrate contacts.

As shown in Fig. 7.14(a), a fewer number of nodes are merged for those contacts closer to the ground pad. This behavior is due to the large voltage difference among these contacts since the overall switching current flowing among these contacts is relatively high. Similarly, depending upon the switching activity, a fewer number of substrate contacts are merged for those blocks that inject higher current. Contacts C_{41} to C_{48}, shown in Fig. 7.14(a), belong to those output buffers sinking higher current as compared to the other blocks in the circuit. Three different voltage domains are therefore determined for these substrate contacts.

The substrate is extracted for the pre- and post-merging cases using SubstrateStorm [196]. The noise is observed using Spectre at the sense node located 60 μm from the

Table 7.2: Original and updated $R(node)$ values of C_{41} to C_{48} for different values of V_{lim}.

<table>
<tr><th></th><th colspan="4">$R(node)$ (Ω)</th></tr>
<tr><th></th><th>Original</th><th>$V_{lim} = 0.05$ volts</th><th>$V_{lim} = 0.1$ volts</th><th>$V_{lim} = 0.25$ volts</th></tr>
<tr><td>C_{41}</td><td>4.3</td><td rowspan="2">8.2</td><td rowspan="4">16.7</td><td rowspan="11">19</td></tr>
<tr><td>C_{42}</td><td>4</td></tr>
<tr><td>N_{33}</td><td>10.9</td><td>10.9</td></tr>
<tr><td>N_{34}</td><td>7</td><td rowspan="5">14</td></tr>
<tr><td>C_{43}</td><td>4.3</td><td rowspan="2">3.1</td></tr>
<tr><td>C_{44}</td><td>4</td></tr>
<tr><td>N_{35}</td><td>7</td><td rowspan="5">15.8</td></tr>
<tr><td>C_{45}</td><td>4.3</td><td rowspan="2">8.2</td></tr>
<tr><td>C_{46}</td><td>4</td></tr>
<tr><td>C_{47}</td><td>11.3</td><td rowspan="2">15</td></tr>
<tr><td>C_{48}</td><td>4</td></tr>
</table>

nearest substrate contact. Note that the parasitic resistance among the substrate contacts on the ground network ($R(node)$) and the current profile of each contact ($i_{[t]}(node)$) are updated after merging based on the CONTACT-MERGE algorithm, as described in Section 7.2.2. As an example, the value of the original and updated $R(node)$ after merging are listed in Table 7.2 for C_{41} to C_{48}. For example, for $V_{lim} = 0.1$ volts, C_{45} to C_{48} are merged into one node, and the updated resistance of this node is determined as 15.8 Ω. Similarly, for $V_{lim} = 0.05$ volts, C_{41} and C_{42} are merged into one node with an updated resistance of 8.2 Ω.

The time domain noise waveforms observed at the sense node before and after merging are compared in Fig. 7.15 at two different time intervals. The waveform shape and peak magnitude of the substrate noise at the sense node after merging into nine contacts match the original noise voltage with a peak-to-peak error of 11% in the noise voltage. Note that the error increases to 70% if V_{lim} is increased to 0.4

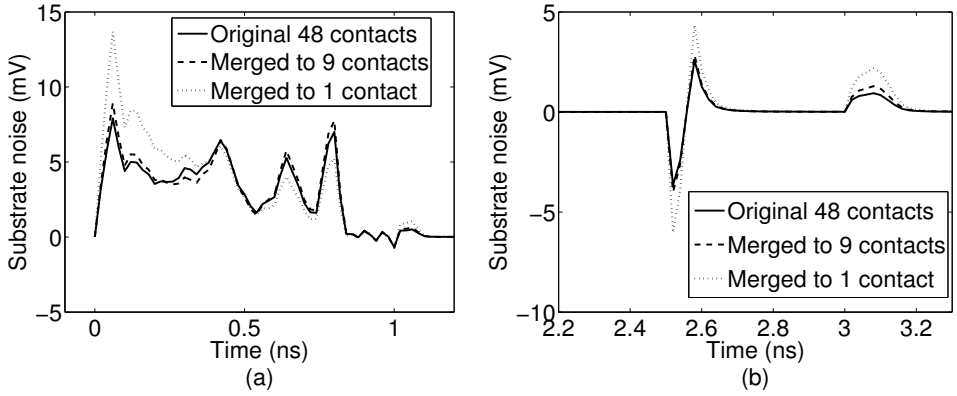


Figure 7.15: Comparison of the substrate noise at the sense node before and after merging for two different time intervals: (a) From 0 to 1.2 ns. (b) From 2.2 ns to 3.2 ns. The solid line represents the original circuit with 48 substrate contacts. The dashed and dotted lines represent, respectively, the reduced network with nine substrate contacts ($V_{lim} = 0.1$ volts) and a single substrate contact ($V_{lim} = 0.4$ volts). The peak-to-peak error in the noise voltage is 11% for nine substrate contacts and increases to 70% for a single substrate contact. The error in the rms noise voltage over one period is 6% for nine contacts and increases to 28% for a single contact.

volts, merging all of the contacts into a single contact. The error in the rms noise over one period is 6% for nine contacts and increases to 28% for a single contact.

The frequency domain characteristics are illustrated in Fig. 7.16. The ratio of the estimated power of the noise to the original power is less than 1 dB at the fundamental frequency (200 MHz), and four higher harmonics when $V_{lim} = 0.1$ volts (the number of contacts is reduced to nine). For $V_{lim} = 0.4$ volts (where the number of contacts is reduced to one), this ratio increases to 1.6 dB at the fundamental frequency. Note that in this case the ratio further increases at the higher harmonics, for example, 6 dB at 1 GHz.

Considering the number of substrate resistances, SubstrateStorm extracts 1128

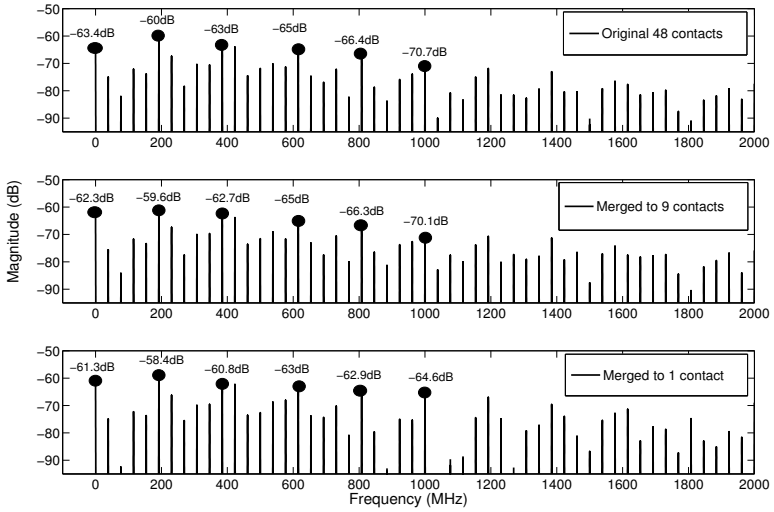


Figure 7.16: Comparison of the spectrum of the substrate noise at the sense node before and after merging into nine contacts ($V_{lim} = 0.1$ volts) and a single contact ($V_{lim} = 0.4$ volts).

resistors in the original system with 48 substrate contacts. Alternatively, when the number of contacts is reduced to nine, the number of extracted substrate resistances is 36, corresponding to a 31X reduction in extracted resistors. The dependence of the error and number of extracted substrate resistors on V_{lim} is shown in Fig. 7.17 to illustrate the complexity versus accuracy tradeoff.

Note that a rapid reduction in the number of resistors is achieved with a relatively small V_{lim}. Increasing V_{lim} above 0.1 volts marginally improves the complexity while introducing additional error at the sense node. These results are listed in Table 7.3 for four different values of V_{lim}.

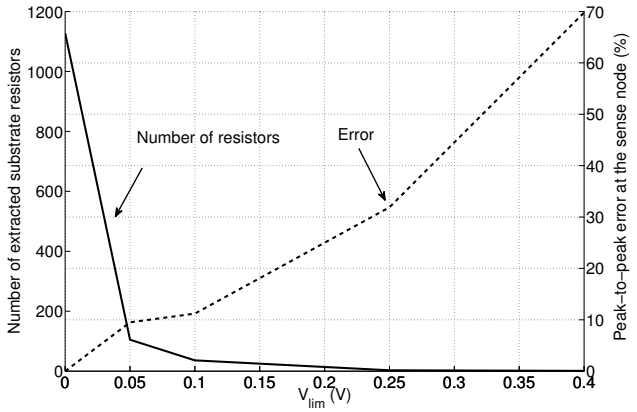


Figure 7.17: Number of extracted substrate resistors and the error in the peak-to-peak noise voltage at the sense node as a function of V_{lim} for circuit 1. The solid and dashed curves represent, respectively, the number of resistors and the error.

Table 7.3: Reduction in the number of extracted substrate resistors, Peak-to-peak (P-P) and RMS substrate noise at the sense node, and the corresponding error in the substrate noise for different values of V_{lim} in Circuit 1.

V_{lim} (V)	Number of substrate contacts	Number of substrate resistors	Reduction	Noise			Error		Estimated / Original
				P-P (mV)	RMS (mV)	At 200 MHz (dB)	P-P (%)	RMS (%)	At 200 MHz (dB)
0	48	1128	–	11.6	0.68	-60	–	–	–
0.05	15	105	11x	10.5	0.61	-60.9	9.5	10.3	-0.9
0.1	9	36	31x	12.9	0.72	-59.6	11.2	5.9	0.4
0.25	3	3	376x	15.3	0.78	-58.9	31.9	14.7	1.1
0.4	1	1	1128x	19.7	0.87	-58.4	69.8	27.9	1.6

7.3.2 Circuit 2

The second circuit is an aggressor digital core located close to a sensitive block in an industrial transceiver circuit, designed in a 90 nm CMOS technology on a bulk type substrate. The circuit contains approximately 200 standard cells. This second circuit is used for two purposes: to examine the behavior of the algorithm for a standard cell based circuit and to evaluate the effect of the parasitic inductance of the ground network on the accuracy of the methodology.

The parasitic resistance between the nodes on the ground network is determined from the sheet resistance (72 mΩ) of the metal lines. The parasitic inductance is extracted using Q3D Extractor [199]. For the vertical ground line with a 1 μm width, the parasitic inductance is 0.79 pH/μm. For the horizontal lines with a width of 0.14 μm, the inductance is extracted as 1.14 pH/μm. The switching current waveform for each contact ($i_{[t]}(node)$) is determined by characterizing each individual gate in the library for various combinations of input patterns. The timing information and input switching pattern of each gate are extracted from a behavioral simulation of the circuit.

The CONTACT-MERGE algorithm is performed with five different values of V_{lim}: 3, 5, 10, 15, and 20 millivolts. Eleven and six voltage domains are determined, respectively, for $V_{lim} = 5$ millivolts and $V_{lim} = 10$ millivolts. These voltage domains are illustrated in Fig. 7.18, where the simplified ground network is obtained from the

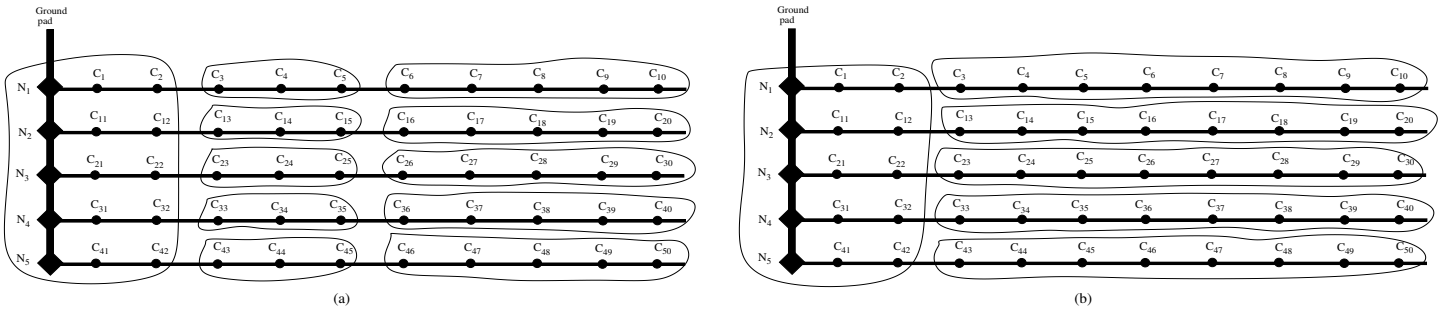


Figure 7.18: Voltage domains on the substrate as determined by the CONTACT-MERGE algorithm: (a) V_{lim} = 5 mV. Eleven voltage domains are identified. (b) V_{lim} = 10 mV. Six voltage domains are identified.

Table 7.4: Original and updated $R(node)$ and $L(node)$ values of C_{43} to C_{49} for various values of V_{lim}.

	$R(node)$ (Ω)				$L(node)$ (pH)			
	Original	$V_{lim}=3$ (mV)	$V_{lim}=5$ (mV)	$V_{lim}=10$ (mV)	Original	$V_{lim}=3$ (mV)	$V_{lim}=5$ (mV)	$V_{lim}=10$ (mV)
C_{43}	5.4	5.7	10.3	19.2	11.9	12.6	22.9	45
C_{44}	0.3				0.7			
C_{45}	4.6	6			10.3	13.4		
C_{46}	1.4				3.1			
C_{47}	3.7	7.5	8.9		8.2	19.2	22.2	
C_{48}	1.6				3.5			
C_{49}	3.8				8.3			

physical layout of the circuit. Note that since this block is standard cell based, the contacts are placed at regular locations to align the contacts in each row. Also note that the density of the substrate contacts is relatively high as compared to regular digital blocks due to the presence of a near-by sensitive circuit.

The substrate is extracted for the pre- and post-merging cases with SubstrateStorm. The parasitic resistance $R(node)$ and inductance $L(node)$ of each merged contact on the ground network are updated. These updated values are listed in Table 7.4 for contacts C_{43} to C_{49}.

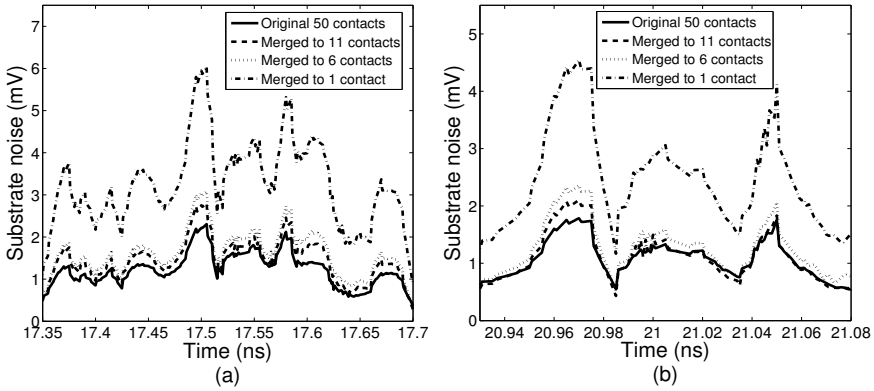


Figure 7.19: Comparison of the substrate noise at the sense node before and after merging for two different time intervals: (a) From 17.35 ns to 17.70 ns. (b) From 20.93 ns to 21.08 ns. The solid line represents the original circuit with 50 substrate contacts. The dashed, dotted, and dash-dotted lines represent, respectively, the reduced network with eleven substrate contacts (V_{lim} = 5 millivolts), six substrate contacts (V_{lim} = 10 millivolts), and a single substrate contact (V_{lim} = 20 millivolts). The peak-to-peak error in the noise voltage is 22% for eleven substrate contacts, 34% for six contacts, and increases abruptly to 160% for a single substrate contact.

The substrate noise is observed at a sense node located 20 μm from the digital circuit on the side of the sensitive analog block. The time domain noise waveforms observed at the sense node before and after merging are compared in Fig. 7.19 for two different time intervals.

The peak-to-peak error in the substrate noise voltage is 22% for eleven substrate contacts, 34% for six contacts, and increases abruptly to 160% for a single substrate contact. The error in the rms noise over a 10 ns timing window is 7% for eleven contacts, 21% for six contacts, and increases to 134% for a single contact.

The frequency domain characteristics are illustrated in Fig. 7.20. At the fundamental frequency (1 GHz), the ratio of the estimated power of the noise to the

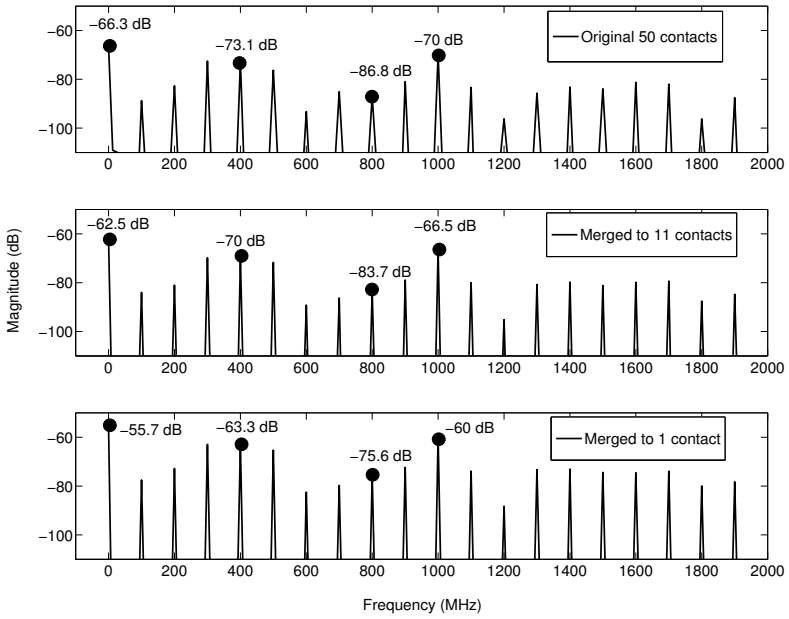


Figure 7.20: Comparison of the spectrum of the substrate noise at the sense node before and after merging into eleven contacts ($V_{lim} = 5$ millivolts) and a single contact ($V_{lim} = 20$ millivolts).

Table 7.5: Reduction in the number of extracted substrate resistors, substrate noise at the sense node, and the corresponding error in the substrate noise for different values of V_{lim} in Circuit 2.

V_{lim} (mV)	Number of substrate contacts	Number of substrate resistors	Reduction	Noise			Error		Estimated / Original
				P-P (mV)	RMS (mV)	At 1 GHz (dB)	P-P (%)	RMS (%)	At 1 GHz (dB)
0	50	1225	–	2.31	0.7	-70	–	–	–
3	17	136	9x	2.81	0.63	-67.5	21.6	10	2.5
5	11	55	22x	2.82	0.75	-66.5	22.1	7.1	3.5
10	6	15	82x	3.1	0.85	-65.5	34.2	21.4	4.5
15	3	3	408x	5.61	1.78	-59	142.8	154.3	11
20	1	1	1225x	6	1.64	-60	159.7	134.3	10

original power is 3.5 dB when $V_{lim} = 5$ millivolts (the number of contacts is reduced to eleven). For $V_{lim} = 20$ millivolts (where the number of contacts is reduced to one), this ratio increases to 10 dB. Note that the ratio remains approximately the same at the harmonics of the fundamental frequency.

Considering the number of substrate resistances, SubstrateStorm extracts 1225 resistors in the original system with 50 substrate contacts. The number of extracted substrate resistors reduces to 15 when $V_{lim} = 15$ millivolts, corresponding to a 82X reduction. These results are listed in Table 7.5 where the error and the reduction in the number of extracted substrate resistors are summarized for various values of V_{lim}. The tradeoff between the reduction in the number of extracted resistors and the accuracy of the substrate noise voltage is further illustrated in Fig. 7.21.

The complete layout of the circuit including all of the devices is extracted using Assura and SubstrateStorm to compare the results obtained by the proposed methodology with a fully extracted set of impedances. The noise waveforms at the sense node

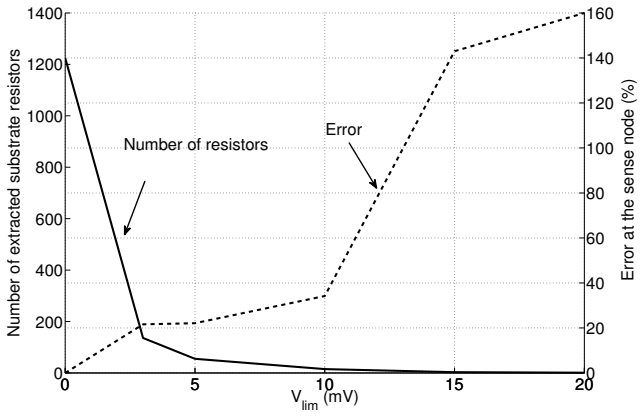


Figure 7.21: Number of extracted substrate resistors and the error in the peak-to-peak noise voltage at the sense node as a function of V_{lim} for circuit 2. The solid and dashed curves represent, respectively, the number of resistors and the peak-to-peak error at the sense node.

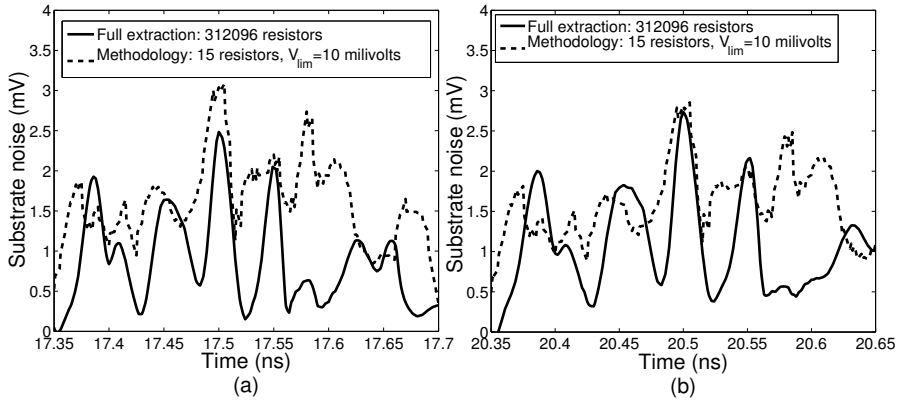


Figure 7.22: Comparison of the substrate noise at the sense node obtained by simulating the fully extracted circuit and application of the methodology when $V_{lim} = 10$ millivolts: (a) From 17.35 ns to 17.7 ns. (b) From 20.35 ns to 20.65 ns. The solid and dashed lines represent, respectively, the full extraction and methodology determined noise.

obtained by simulating the fully extracted circuit and applying the methodology when $V_{lim} = 10$ millivolts are compared in Fig. 7.22 for two different time intervals.

The fully extracted circuit with 200 gates consists of 312,096 resistors for the

substrate and 55,856 junction and well capacitances. The full extraction and simulation of the transient noise for this circuit on a dual core 64 bit Sun machine with Linux operating system requires approximately six hours. Alternatively, the proposed methodology reduces the number of extracted substrate resistors to 15 (for $V_{lim} = 10$ millivolts), achieving more than four orders of magnitude reduction, and requires negligible time. The peak-to-peak error of the methodology in estimating the substrate noise voltage is 24.1%, as illustrated in Fig. 7.22. The limitation of the methodology in terms of run time is the requirement to pre-characterize each cell in the library for various input switching patterns and to perform a gate level simulation of the circuit to extract the required timing information.

Note that the reduction achieved by the methodology is expected to increase for larger scale circuits due to the increasing number of substrate contacts. Substrate contacts are usually placed based on latch-up constraints to achieve a specific contact density [200]. A common practice is to increase the density of the contacts near those aggressor blocks that can potentially affect the sensitive circuit. An aggressor digital block may therefore require a significant number of substrate contacts, where a reduction in the number of contacts quadratically reduces the number of extracted substrate resistances. Note that this methodology achieves a reduction of more than four orders of magnitude in extracted substrate resistances as compared to a fully extracted circuit. This greater reduction is due to the increasing number of input

ports in the full extraction due to the source/drain junctions of the devices which are neglected in this methodology.

7.3.3 Proper Selection of V_{lim}

The proposed methodology requires a reasonable selection of V_{lim} to obtain sufficiently accurate results while reducing the overall number of extracted substrate resistors, as discussed in the previous section. Choosing a V_{lim} greater than the proper value significantly increases the peak-to-peak and rms error. Alternatively, an unnecessarily small V_{lim} limits the reduction in the extracted substrate resistors, and therefore the computational efficiency.

A reasonable value of V_{lim} depends strongly on the absolute value of the peak ground bounce on the local ground distribution network of the aggressor circuit. Choosing V_{lim} as 20% to 30% of the peak ground bounce generally produces reasonable results. A technique is proposed in this section to properly decide the value of V_{lim}.

As listed in Tables 7.3 and 7.5, the peak-to-peak and rms values of the estimated noise voltage is significantly greater for the extreme case when all of the contacts are merged into one contact. As V_{lim} is reduced or the number of contacts is increased, the estimated noise voltages decrease, and ultimately saturate. The value where this saturation starts to occur is a good choice for V_{lim}. The corresponding peak-to-peak noise obtained at this V_{lim} is reasonably accurate with respect to the fully extracted

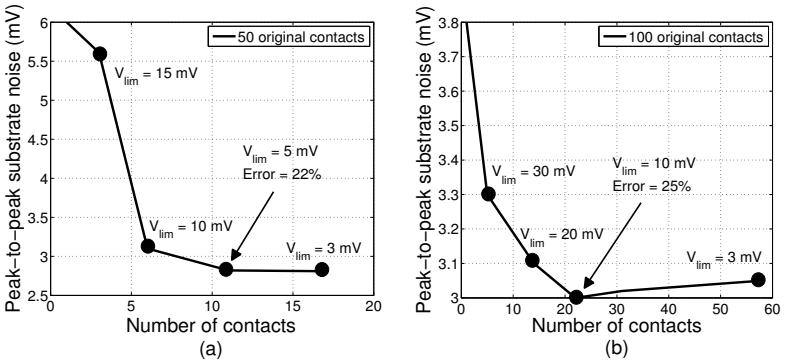


Figure 7.23: Variation of the peak-to-peak noise with respect to the number of contacts after merging: (a) For the second circuit with 50 original contacts. (b) An extension of the second circuit with 100 original contacts.

noise voltage. The variation of the peak-to-peak noise voltage with respect to the number of contacts is shown in Fig. 7.23 for the second circuit and an extension of the second circuit with 100 contacts.

As illustrated in Fig. 7.23, the peak-to-peak noise initially exhibits a rapid decrease, ultimately saturating as V_{lim} is decreased or the number of contacts is increased. For Fig. 7.23(a), at eleven contacts, the peak-to-peak error in the noise voltage is 22%. Similarly, for Fig. 7.23(b), the error is 25% at 22 contacts. The value of V_{lim} where the estimated peak-to-peak noise voltage saturates therefore produces sufficiently accurate results for this methodology. Note that the time complexity of the proposed algorithm is linear, allowing these iterations to be performed in a reasonable amount of time. V_{lim} can therefore be properly selected using this iterative methodology.

7.4 Limitations

As described in Section 7.2, the proposed algorithm requires current profiles for each substrate contact obtained during a specific time window. For a different window or with a different set of input vectors, these current profiles may change, affecting the outcome of the merging process. This dependence of the methodology on the timing window and input vector can be decreased by applying a statistical approach. Alternatively, the algorithm can be performed multiple times over different timing windows, resulting in slightly different substrate networks. This solution is computationally possible since the algorithm operates in linear time.

The proposed algorithm is independent of the location of the sensitive node where the noise is observed, assuming that the sense node is sufficiently far from the aggressor block. The dependence of the error on the specific location of the sense node is also an issue requiring further study. A larger reduction in the error and the number of extracted resistors are possible if the location of the sense node is considered when locating the equivalent contact after merging. Also note that if the sense node is sufficiently close to the aggressor block, the physical distance among the contacts is comparable to the distance between a contact and the sense node. In this case, the location of the equivalent contact plays a significant role in the accuracy of the algorithm. Determining the specific location of the equivalent contact relative to the location of the sense node can therefore also be a focus of future study.

The current version of the algorithm assumes that the local ground distribution network of the aggressor block can be represented as a tree. This assumption may not always be valid if the aggressor block has multiple connections to the global ground network, creating a mesh. One practical approach to handle a mesh structure is to assume that the connections to the upper metal layers can be treated as a quiet ground pad, and to partition the gates where each pad sinks the current of a subclass of the gates, as proposed in [76] for placing decoupling capacitors.

7.5 Summary

A methodology is proposed for the efficient analysis of substrate noise coupled to a sensitive block in large scale mixed-signal circuits. The substrate of the aggressor circuit is partitioned into voltage domains where each domain represents a region within the substrate biased with approximately the same voltage by the substrate contacts. Each of these voltage domains is therefore effectively shorted. A single equivalent input port to the substrate is generated for each domain, neglecting all of the remaining ports. The reduction in the number of input ports significantly reduces the number of extracted substrate resistors. An algorithm is described to determine these voltage domains by merging those contacts exhibiting a voltage difference smaller than a specified value, and generate an equivalent contact which is placed at the geometric mean of the merged contacts. The ground network impedance is updated to maintain

the accuracy of the noise voltage at the sense node. Simulation results demonstrate a reduction of more than four orders of magnitude in the number of extracted substrate resistors as compared to a fully extracted layout while introducing 24% error in the peak-to-peak value of the substrate noise voltage at the sense node.

Chapter 8

Substrate Noise Reduction Based on Noise Aware Cell Design

Simultaneous switching noise caused by the parasitic inductance (di/dt noise) and transient IR drops caused by the resistance of the ground network directly affect the substrate through the substrate contacts, as described in Chapter 3. Employing a dedicated substrate bias by separating the ground network of the contacts from the ground network of the digital circuit reduces substrate noise coupling [126], [201], [202]. This technique, known as Kelvin biasing, reduces substrate noise at the cost of lower device reliability, increased power/ground noise, and additional metal resources. These limitations make the use of Kelvin biasing impractical.

An alternative substrate biasing methodology is introduced in this chapter based on modifying the design of the standard cell library. Standard cells with dedicated substrate contacts are proposed for those cells that generate significant noise. These dedicated contacts are connected to a separate ground to provide isolation from the

noisy ground network. The proposed methodology achieves more than a 60% reduction in substrate noise while removing the limitations of the Kelvin biasing scheme.

The primary drawback of the proposed technique is the increased area due to the additional contacts and the separate ground network. This increased area, however, is minimized because these modified cells are only used in the primary noise generating blocks within a circuit. The additional ground network and contacts are therefore only required in these blocks.

The rest of the chapter is organized as follows. The substrate biasing schemes are reviewed in Section 8.1. The proposed methodology is described in Section 8.2. Simulation results are presented in Section 8.3. These results are discussed and compared with Kelvin biasing in Section 8.4. Finally, the chapter is summarized in Section 8.5.

8.1 Substrate Biasing Schemes

In a mixed-signal circuit, several different techniques exist to bias the substrate. The conventional approach is to connect the substrate contacts to the ground network of the digital circuitry. This technique, however, injects significant noise into the substrate since the digital ground suffers from simultaneous switching noise.

Kelvin biasing, as illustrated in Fig. 8.1, has been proposed to reduce noise injection by biasing the substrate with a dedicated ground network. This technique removes the resistive connection of the substrate with the noisy digital ground, thereby

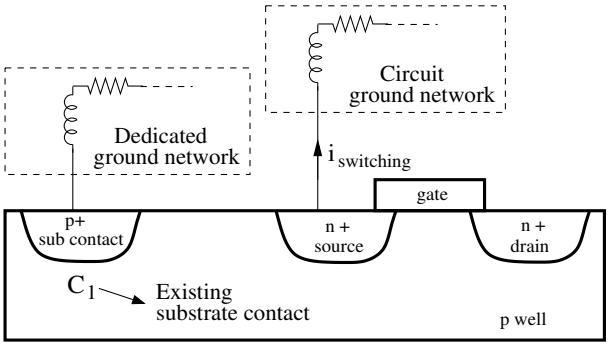


Figure 8.1: Kelvin biasing scheme to reduce substrate noise. The substrate contacts are connected to a dedicated ground network rather than the ground network of the digital circuit to prevent ground noise coupling into the substrate.

reducing substrate noise. Kelvin biasing, however, has significant limitations. Since the substrate is biased separately, the devices suffer from the body effect which can have a significant effect on high performance circuits. Another drawback is the increased power/ground noise since the n-well capacitance cannot be exploited as a decoupling capacitance. A reduction by a factor of three in the total decoupling capacitance is reported in [201] if Kelvin biasing is employed.

While reducing ground coupling into the substrate, Kelvin biasing is ineffective in reducing source/drain junction coupling. The overall reduction in substrate noise is, therefore, limited. Furthermore, Kelvin biasing requires an additional ground network, making the technique impractical in terms of the required metal resources.

Another technique to bias the substrate is a backside contact, which increases the cost and requires specialized packaging. The use of an analog ground rather than a digital ground has also been suggested to bias the substrate [201]. If an analog

ground is used, however, substrate noise can couple into the analog ground network, degrading performance.

8.2 Proposed Methodology

The design of a standard cell with a dedicated substrate contact is explained in Section 8.2.1. An analysis of the substrate noise reduction mechanism using these standard cells is described in Section 8.2.2.

8.2.1 Standard Cell with a Dedicated Substrate Contact

In the design of a digital integrated circuit, the placement of the substrate contacts is usually achieved after the place-and-route phase of the design flow is completed. The latch-up design rules determine the minimum distance among the contacts.

A standard cell design approach is proposed in this paper where each cell in the library has a dedicated substrate contact. This dedicated substrate contact is placed in close proximity to the cell, as determined by technology based design rules. Conventional and *aggressor* standard cells are illustrated in Fig. 8.2. Note that these aggressor cells are in addition to existing conventional cells in the library, where the choice between a conventional and aggressor cell is made depending upon several factors such as the switching activity of the digital block and the physical distance between the digital and sensitive analog blocks within a circuit.

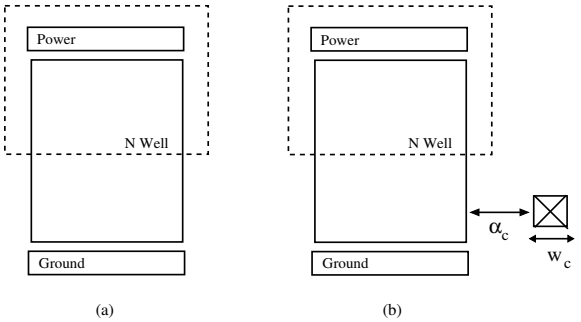


Figure 8.2: Standard cell: (a) conventional cell, (b) aggressor cell with a dedicated substrate contact. α_c is the minimum distance between the contact and the diffusion, and w_c is the width of the contact.

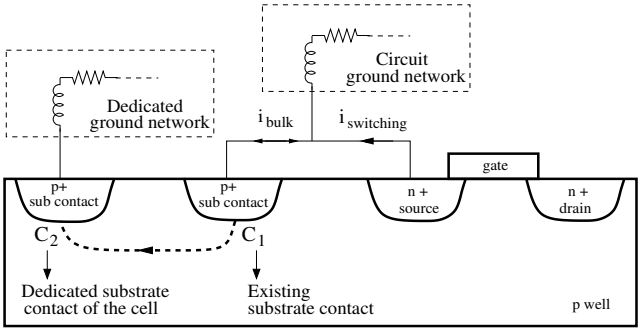


Figure 8.3: The effect of the dedicated contact on reducing substrate noise. The injected noise from the noisy contact C_1 is filtered through the dedicated contact C_2 rather than propagated into the substrate.

The physical representation of a cell with a dedicated substrate contact is shown in Fig. 8.3. C_1 represents an existing substrate contact placed according to latch-up design rules and C_2 represents the dedicated substrate contact of the cell. Note that C_1 is connected to the ground network of the digital circuit; a separate ground network, however, is required for the dedicated contacts in order to isolate these contacts from the noisy ground network. Noise reduction is achieved through the low impedance path between C_1 and C_2. The injected noise from the noisy contact C_1 is

filtered through C_2 rather than propagated into the substrate.

The drawback of providing a dedicated substrate contact is an increase in circuit area. This increase in area can be expressed as

$$\Delta A = L(\alpha_c + w_c)n, \tag{8.1}$$

where L is the length of the standard cell, α_c is the minimum distance between the n-type and p-type diffusion regions, w_c is the width of the contact, and n is the number of aggressor cells with a dedicated substrate contact. Note that these types of cells are only used in aggressor digital blocks that are identified as major noise sources. The number n, therefore, is usually a small fraction of the total number of cells, lessening the increase in area.

8.2.2 Analysis of Noise Reduction Mechanism

The dedicated contacts connected to a separate ground network significantly reduce the noise current propagating through the substrate. This reduction is due to the change in the impedance seen by the switching current.

Equivalent circuit models to analyze substrate noise through ground coupling are shown in Fig. 8.4. A circuit model for conventional substrate biasing is shown in Fig. 8.4(a). The switching current is divided based on the impedance of the ground

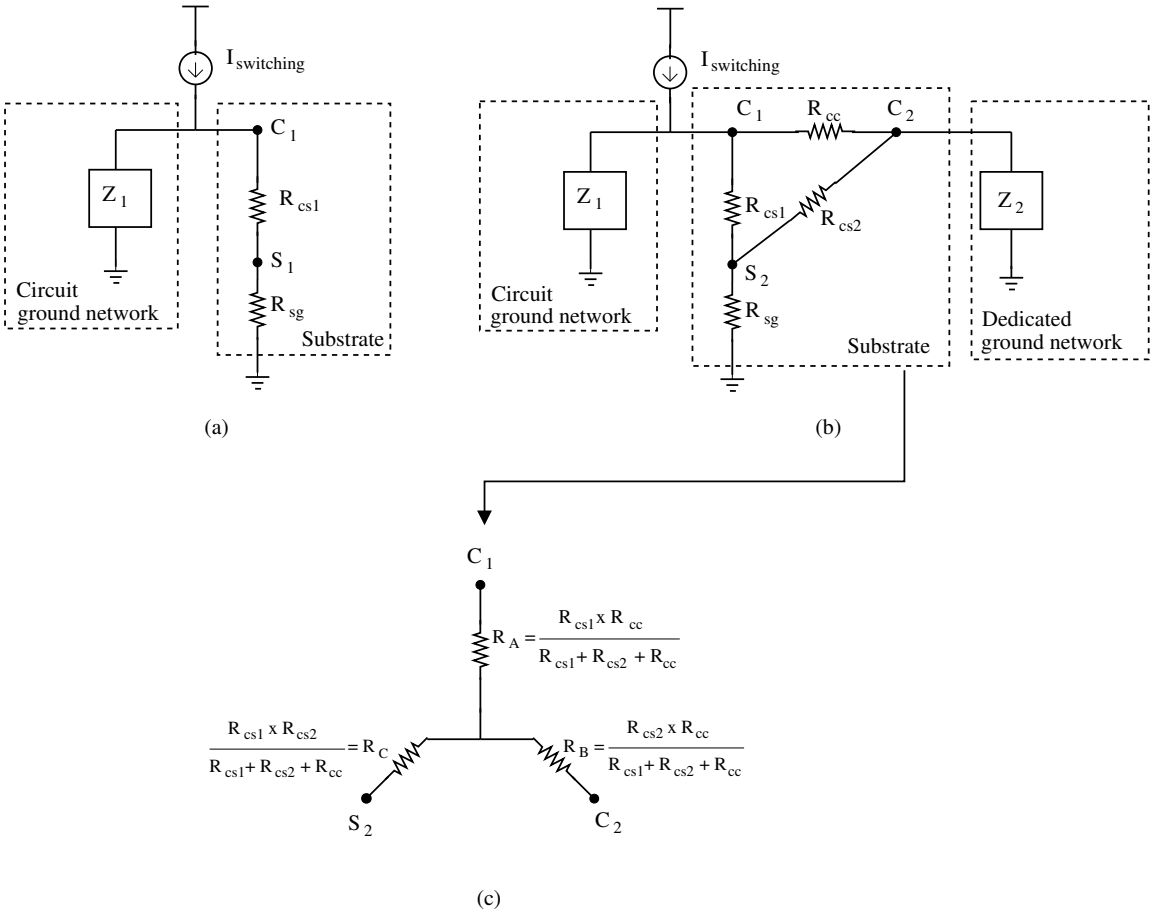


Figure 8.4: Equivalent circuit models to analyze substrate noise through ground coupling: (a) Circuit model for the conventional scheme. I_s is the switching current of the circuit. C_1 is the conventional substrate contact, Z_1 is the ground network impedance, and S_1 is the substrate location where the noise is analyzed, (b) Circuit model for the proposed scheme. C_2 is the dedicated substrate contact of the standard cell and Z_2 is the impedance of the additional ground network to which the dedicated contacts are connected, (c) Transformation of the mesh formed by the contacts C_1 and C_2, and node S_2.

and substrate networks. For Fig. 8.4(a), the noise voltage at sense node S_1 is

$$v_{s1}(\omega) = \frac{I_s|Z_1(\omega)|}{R_{cs1} + R_{sg} + |Z_1(\omega)|} R_{sg}, \tag{8.2}$$

where I_s is the switching current, $Z_1(\omega)$ is the ground network impedance, R_{cs1} is the substrate resistance between contact C_1 and the sense node S_1, and R_{sg} is the substrate resistance between the sense node and ground, which is usually the analog ground of the circuit.

A circuit model for the proposed substrate biasing scheme is shown in Fig. 8.4(b). The dedicated substrate contact C_2 is connected to a separate ground network with impedance Z_2. Note that the mesh formed by the contacts C_1 and C_2, and node S_2 can be transformed using resistances R_A, R_B, and R_C, as illustrated in Fig. 8.4(c). After this transformation, the noise voltage at the sense node S_2 can be expressed as

$$v_{s2}(\omega) = \left[\frac{I_s|Z_1(\omega)|}{[(R_C + R_{sg}) \backslash\backslash (R_B + |Z_2(\omega)|)] + R_A + |Z_1(\omega)|} \right]$$

$$x \left[\frac{R_B + |Z_1(\omega)|}{R_{sg} + R_C + R_B + |Z_2(\omega)|} \right] R_{sg}, \tag{8.3}$$

where I_s is the switching current, $Z_1(\omega)$ is the impedance of the circuit ground network, $Z_2(\omega)$ is the impedance of the dedicated ground network, and R_A, R_B, and R_C are the substrate resistances after transformation of the mesh.

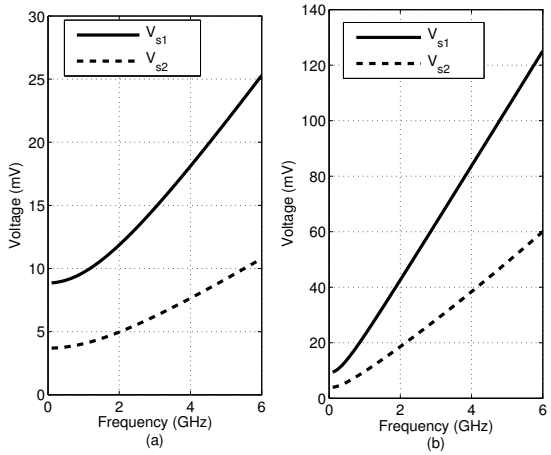


Figure 8.5: Analytic prediction of the noise voltages V_{s1} and V_{s2} with (a) a flip-chip package, (b) a bond-wire package.

For practical values of substrate resistances and assuming $Z_1 = Z_2$, the noise voltages predicted by (8.2) and (8.3) are illustrated as a function of frequency in Fig. 8.5. The solid line represents the noise voltage for a conventional scheme and the dashed line represents the noise voltage for the proposed scheme. More than a 50% reduction in noise voltage is predicted from an analytic analysis based on equivalent circuit models.

Note that the proposed scheme represents a *localized* guard ring placement methodology, as illustrated in Fig. 8.6. In conventional guard rings, the contacts are placed around the aggressor block, as illustrated in Fig. 8.6(a). The efficacy of conventional guard rings, however, is limited due to the vertical propagation paths of the current throughout the substrate. A portion of the noise current can flow deeper into the substrate, thereby bypassing the guard ring, making the isolation less effective. This

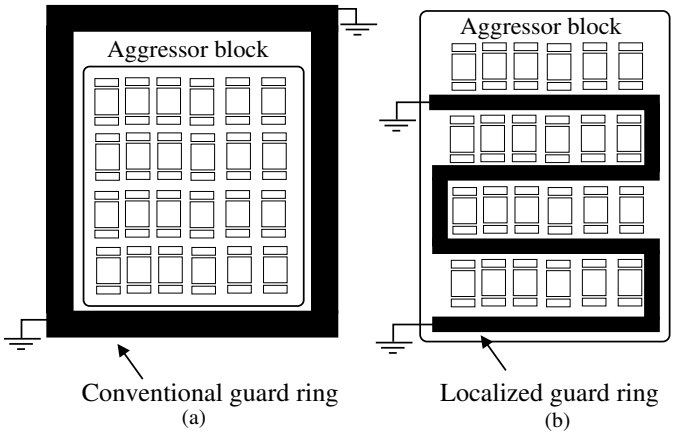


Figure 8.6: Simplified representation of a (a) conventional guard ring around an aggressor block, (b) proposed localized guard ring within an aggressor block achieved with noise aware standard cells.

inefficiency is significant, particularly in large aggressor blocks, since the noise current is more likely to spread throughout the substrate until the current reaches the guard ring surrounding the block. In the proposed scheme, as illustrated in Fig. 8.6(b), the guard ring is localized within the aggressor block due to noise aware standard cells rather than placing the ring around the block. The localized guard ring minimizes the vertical current propagation withing the substrate, thereby improving the efficiency of conventional guard rings.

8.3 Simulation Results

A noise generator circuit, as shown in Fig. 8.7, has been designed in a 90 nm double-well CMOS technology with a bulk type (non-epi) substrate. The circuit consists of four chains of scaled buffers driven by input signals with 70 ps rise and

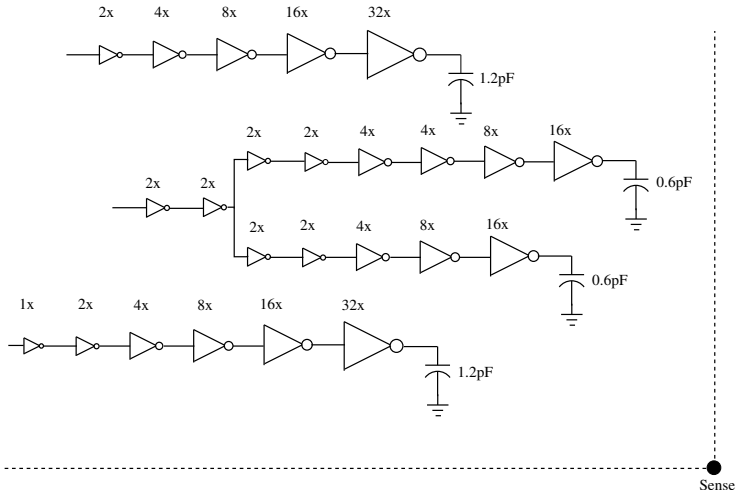


Figure 8.7: Noise generator circuit consisting of four chains of scaled buffers to evaluate the proposed substrate biasing methodology.

fall times.

Three different versions of the circuit have been realized. The first circuit utilizes the conventional technique with regular standard cells and the substrate contacts are placed according to technology-dependent latch-up constraints. The second circuit utilizes a Kelvin biasing technique where the contacts of the first circuit are connected to a separate ground network, as shown in Fig. 8.1. The third circuit utilizes the proposed methodology where each cell has a dedicated substrate contact, assuming all cells are aggressors. These dedicated contacts are connected to a separate ground network as shown in Fig. 8.3. Note that the contacts of the first circuit remain the same for the third circuit.

The layout and substrate impedances of the three circuits are extracted, respectively, using Assura RCX [196] and SubstrateStorm [196], while the overall netlist

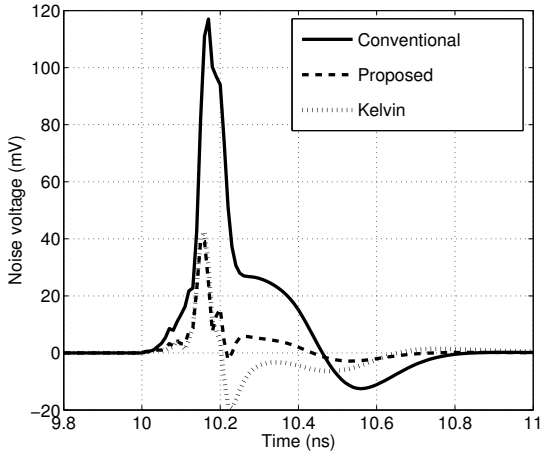


Figure 8.8: Substrate noise voltage at the sense node for the conventional, Kelvin, and proposed schemes.

is simulated using Spectre [196]. For each circuit, the substrate noise voltage is examined at the sense node which is located 25 μm away from the closest contact, as shown in Fig. 8.7. Note that the parasitic impedances of the ground network of the digital circuit and the dedicated ground network are assumed to be the same.

The substrate noise voltage waveforms for the conventional, Kelvin, and proposed techniques are shown in Fig. 8.8. The ground network has a distributed parasitic impedance of $R = 3\ \Omega$ and $L = 160$ pH with a bond-wire package exhibiting a lumped parasitic impedance of $R = 0.2\ \Omega$ and $L = 1$ nH.

The peak-to-peak substrate noise voltage for the three techniques are listed in Table 8.1 for different ground network impedances and package types. Significant noise reduction over conventional and Kelvin biasing schemes is achieved using the proposed methodology. The drawback is the increase in area due to the additional

Table 8.1: Comparison of peak-to-peak substrate noise voltage for the conventional, Kelvin, and proposed techniques. On-chip interconnect parasitic impedances: $R = 3$ Ω and $L = 160$ pH. Flip-chip parasitic impedances: $R = 0.1$ Ω and $L = 60$ pH. Bond-wire parasitic impedances: $R = 0.2$ Ω and $L = 1$ nH.

On-chip interconnect / package type	Peak-to-peak substrate noise (mV)			Noise reduction over conventional	Noise reduction over Kelvin
	Conventional	Kelvin	Proposed		
R only / flip-chip	32	20	9	72%	55%
RL / flip-chip	47	29	16	66%	45%
R only / bond-wire	116	58	39	66%	33%
RL / bond-wire	129	61	42	67%	31%

contacts. For this example, the additional required area is 82 μm^2, corresponding to a 12% increase.

8.4 Discussion

The proposed methodology achieves a greater reduction in noise as compared to Kelvin biasing. This result occurs because Kelvin biasing only reduces the ground coupling noise generation mechanism. The proposed technique, however, reduces both ground and the junction coupling mechanisms. The greater reduction in noise as compared to the Kelvin technique with a smaller ground parasitic impedance, as listed in the last column in Table 7.3, supports this conclusion.

In addition to a greater reduction in noise, the proposed methodology removes the limitations of Kelvin biasing. The voltage difference between the source and body of the transistors for Kelvin biasing is as large as 160 mV. For the proposed technique, this voltage difference is several millivolts. The body effect is, therefore, smaller and

the likelihood of latch-up is reduced.

The proposed scheme exploits the n-well capacitance as a decoupling capacitance to reduce power/ground noise. Kelvin biasing, however, cannot exploit the n-well capacitance because the circuit and the substrate have separate bias networks.

Another major limitation of Kelvin biasing is the requirement for an additional ground network, making this approach impractical considering limited metal resources. Alternatively, the proposed methodology only requires an additional ground network for the aggressor cells with a dedicated contact.

8.5 Summary

A substrate biasing scheme is presented in this chapter based on modifying the design of a standard cell library. Cells with dedicated substrate contacts are proposed for the noise generating digital blocks. The substrate is biased with the ground network of the digital circuit through the existing substrate contacts. The dedicated substrate contacts within the aggressor cells, however, are connected to a separate ground network. A low impedance path in the substrate is thereby created between the noisy contacts and the dedicated contacts. Simulation results show, on average, a reduction in substrate noise of 68% (9.87 dB) over the conventional technique and 41% (4.71 dB) over the Kelvin biasing scheme. Furthermore, the limitations of Kelvin biasing are removed using the proposed substrate biasing methodology.

Chapter 9

Exploiting Setup-Hold Time Interdependence in Timing Analysis

The effects of switching noise in mixed signal circuits have been studied in Chapters 4, 5, 6, 7, and 8 with a primary focus on worst-case power/ground noise estimation and substrate coupling. Switching noise is a significant concern not only for mixed-signal circuits, but also for synchronous digital ICs, as described in Chapter 2. Specifically, power/ground noise affects the timing characteristics of a circuit, producing delay uncertainty. The relationship between switching noise and the timing behavior of a circuit is therefore important to guarantee proper functionality at the target clock frequency. An interdependent characterization methodology for timing constraints, *i.e.*, setup and hold times, is described in this chapter. These interdependent timing constraints are exploited to reduce delay uncertainty, as described in Chapter 10.

The timing characteristics of a synchronous circuit is typically verified by means of static timing analysis (STA) tools. The STA tools rely on data described in the cell libraries to analyze the circuit. The characterization of the individual cells in cell libraries is, therefore, highly critical in terms of the accuracy of the STA results [203], [204], [205], [206]. Specifically, the setup time and hold time constraints of the sequential cells are used to verify the timing of a synchronous circuit. Inaccurate characterization of timing constraints causes the STA results to be either highly optimistic or pessimistic. Both cases should be avoided as the optimistic case can cause a fabricated circuit to fail whereas the pessimistic case unnecessarily degrades circuit speed.

The over-optimism or pessimism in STA is primarily due to *independent* characterization of the timing constraints although these constraints (including CLK-to-Q delay) are *interdependent*. The constraints should, therefore, be characterized interdependently to remove optimism or pessimism in STA. In [207], a timing constraint characterization is proposed that minimizes the sum of the CLK-to-Q delay and the setup time. In [208], the CLK-to-Q delay of a sequential cell is modeled considering the dependence between the CLK-to-Q delay and setup time. A 50 to 60 ps decrease on the clock period is shown if this dependence is considered during STA. These approaches, however, do not consider the interdependence between the setup time and hold time. An approach for interdependent characterization is proposed in [209].

A solution is offered in [210] considering the dependence between the setup time, hold time, and CLK-to-Q delay to determine the maximum operating frequency of a sequential cell. These approaches, however, do not exploit the interdependence in STA.

When the interdependence is considered during constraint characterization of a sequential cell, multiple valid constraint pairs, which are interchangeable, are obtained. These multiple pairs can be utilized in STA to significantly reduce timing violations and improve negative slack. Multiple constraint pairs, however, are currently not exploited.

A comprehensive methodology is proposed in this chapter to overcome the over optimism and pessimism of the current approaches. The methodology consists of two phases. In the first phase, an interdependent characterization of setup and hold times is described, resulting in multiple constraint pairs. In the second phase, an efficient algorithm with linear time complexity is presented to integrate the interdependence into an STA tool. The algorithm exploits multiple constraint pairs by dynamically switching between pairs in order to remove violations. The methodology is validated on high performance industrial circuits and an industrial STA tool. In particular, STA results demonstrate up to a 53% reduction in the number of constraint violations as well as up to a 48% reduction in the worst negative slack.

The rest of the chapter is organized as follows. Related background is provided in

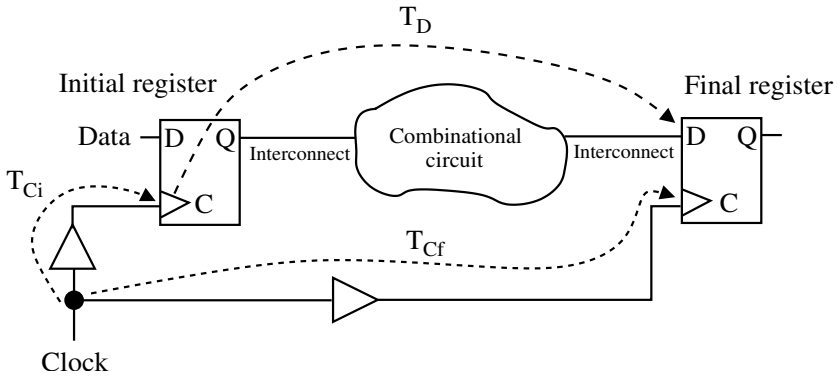


Figure 9.1: Simple synchronous circuit consisting of a combinational logic and two registers.

Section 9.1. The problem formulation, the concept of interdependence, and current approaches are discussed in Section 9.2. The proposed characterization methodology and algorithm are described in Section 9.3. STA results are presented in Section 9.4. Finally, the chapter is summarized in Section 4.3.

9.1 Background

The timing characteristics of synchronous circuits are reviewed in Section 9.1.1. Setup-hold times and skews are described in Section 9.1.2.

9.1.1 Timing Characteristics of Synchronous Circuits

A simple synchronous digital circuit consisting of two sequentially-adjacent registers and a combinational circuit between these registers is shown in Fig. 9.1. Two inequalities should be satisfied for this circuit to function properly. Referring to

Fig. 9.1, the first inequality is

$$T_{Cf} + T_{CP} \geq T_{Ci} + T_D + T_S, \tag{9.1}$$

where T_{Ci} and T_{Cf} are the delay for the clock signal to arrive, respectively, at the *initial* and *final* registers. Note that T_{Ci} and T_{Cf} are also referred to, respectively, as the delay of the clock launch path and clock capture path. T_{CP} is the clock period, T_D is the data path delay consisting of the clock-to-Q delay of the initial register, the logic delay of the combinational circuit, and the interconnect delay. T_S is the setup time of the final register. Note that (9.1) determines the maximum speed of the circuit, making this inequality important for the critical paths within a circuit.

The second inequality that needs to be satisfied is

$$T_{Ci} + T_D \geq T_{Cf} + T_H, \tag{9.2}$$

where T_H is the hold time of the final register. This inequality guarantees that no race condition exists, *i.e.*, the data is not latched within the final register at the same clock edge. Note that (9.2) is relatively more important for those timing paths where the data path delay is sufficiently small, such as a shift register or counter.

9.1.2 Setup / Hold Times and Skews

Inequalities (9.1) and (9.2) require a difference called a *skew* to be larger than or equal to a number called a *timing constraint*. These inequalities, therefore, can be rewritten as

$$\text{Setup skew} \geq T_S, \tag{9.3}$$

$$\text{Hold skew} \geq T_H, \tag{9.4}$$

where setup skew and hold skew are, respectively,

$$\text{Setup skew} = T_{Cf} + T_{CP} - (T_{Ci} + T_D), \tag{9.5}$$

$$\text{Hold Skew} = T_{Ci} + T_D - T_{Cf}. \tag{9.6}$$

These skews and constraints are defined more formally below.

Definition 1 *Setup skew* refers to the amount of time a change in the data signal arrives at the D input of a sequential cell before the arrival of the latching edge of the clock signal.

Definition 2 *Hold skew* refers to the amount of time the data signal is stable at the D input of a sequential cell after the arrival of the latching edge of the clock signal.

Definition 3 *Setup time* refers to the setup skew necessary for the clock to reliably capture the data.

Definition 4 *Hold time* refers to the hold skew necessary for the clock to reliably store the data.

Definition 5 A *setup time violation* refers to the situation where (9.1) or (9.3) is violated.

Definition 6 A *hold time violation* refers to the situation where (9.2) or (9.4) is violated.

Note the important difference between setup-hold skews and setup-hold times: Setup and hold skews refer to *any* time difference between the data and clock signals whereas the setup and hold times refer to the minimum required time difference to reliably capture and store the data. It is therefore important to decrease the setup and hold times while reliably capturing and storing the data.

9.2 Problem Formulation

Existing characterization approaches and the corresponding limitations are summarized in Section 9.2.1. Interdependence between setup and hold times is explained in Section 9.2.2. Note that the numerical results of the rest of the chapter are obtained through HSPICE simulations with BSIM4/BSIM3 models under typical corner conditions. A 90 nm industrial cell library is used for the simulations..

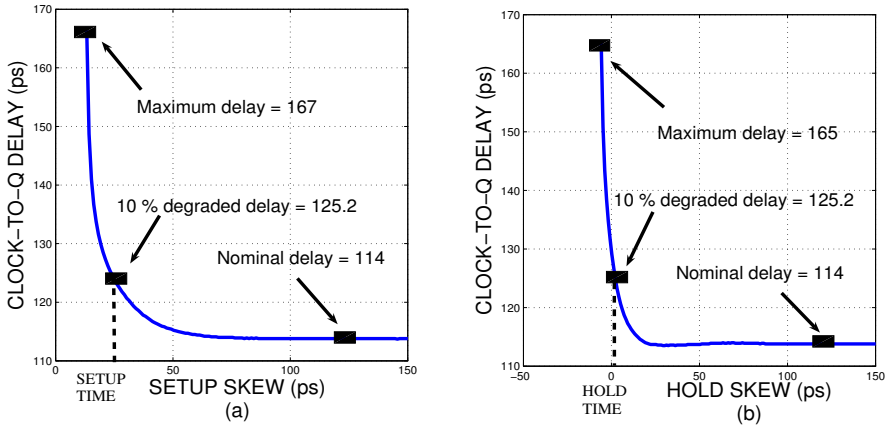


Figure 9.2: Independent constraint characterization for sequential cells: (a) setup skew vs. CLK-to-Q delay for setup time characterization, (b) hold skew vs. CLK-to-Q delay for hold time characterization.

9.2.1 Existing Characterization Approaches

A common approach to characterize setup time is to examine the setup skew versus CLK-to-Q delay relationship [207], [211] at a fixed hold skew, which is called here the *counterpart skew*. The process is similar for the hold time. These approaches are shown in Fig. 9.2.

According to [207], three regions can be determined for both of these plots: *stable*, *metastable*, and *failure* regions. The stable region is defined as the region in which the CLK-to-Q delay is independent of the setup or hold skew. As the skew decreases, the CLK-to-Q delay starts to rise in an exponential fashion [17]. If the skew is excessively small, the sequential cell fails to latch the data. This region is called the failure region. The region between the stable and failure regions is referred to as the metastable region.

The setup and hold times cannot fall in the failure region since the sequential cell is unable to latch the data in that region. The setup (hold) time is usually set to the setup (hold) skew where the stable region crosses over into the metastable region. There are different approaches to identify this *crossover point*, as listed in [207]. In some approaches, the crossover point is the time where a certain amount of degradation in the CLK-to-Q delay occurs. For example, a 10% degradation is assumed in this study. In some other approaches, the crossover point is the time where the sum of the setup skew and CLK-to-Q delay is minimized.

9.2.2 Interdependence Between Setup and Hold Times

The setup and hold times are not independent [209]; rather, these constraints are a function of the counterpart skews (hold skew for setup time and setup skew for hold time). These dependences are shown in Fig. 9.3. Note that the setup time decreases as the hold skew increases and the hold time decreases as the setup skew increases. Thus, the smallest setup and hold times occur when the counterpart skews are the largest.

Existing characterization approaches typically ignore the interdependence of the setup and hold times. This strategy leads to two main issues:

Issue 1. Ignoring the interdependence of the setup and hold times results in either over-optimism or pessimism depending upon the assumption on the counterpart skew.

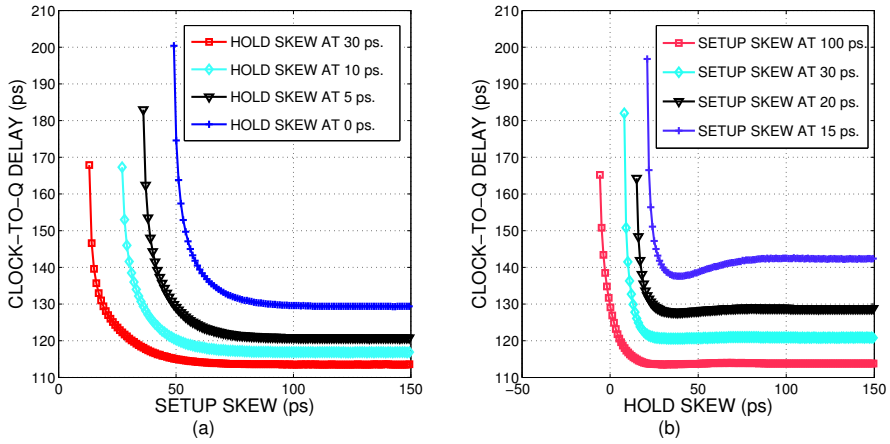


Figure 9.3: Constraint characterization for sequential cells at different counterpart skews: (a) setup skew vs. CLK-to-Q delay at different hold skews, (b) hold skew vs. CLK-to-Q delay at different setup skews.

If the counterpart skews are assumed to be unnecessarily large, the resulting setup and hold times are optimistic. If, however, the data waveform does not satisfy large counterpart skews, optimistic setup and hold times cause the circuit to fail despite not violating any of the individual constraints. This situation is illustrated in Fig. 9.4. In step A, the setup time is characterized as S at a sufficiently large hold skew. In step B, the hold time is characterized as H at a sufficiently large setup skew. In step C, the data waveform satisfies both of these constraints. The sequential cell, however, violates the minimum pulse width constraint due to over-optimism introduced during the characterization steps of A and B.

Alternatively, if the counterpart skews are assumed to be unnecessarily small, the resulting setup and hold times are pessimistic. An example to illustrate pessimism is shown in Fig. 9.5. This example is proposed in [209] to overcome the over-optimism.

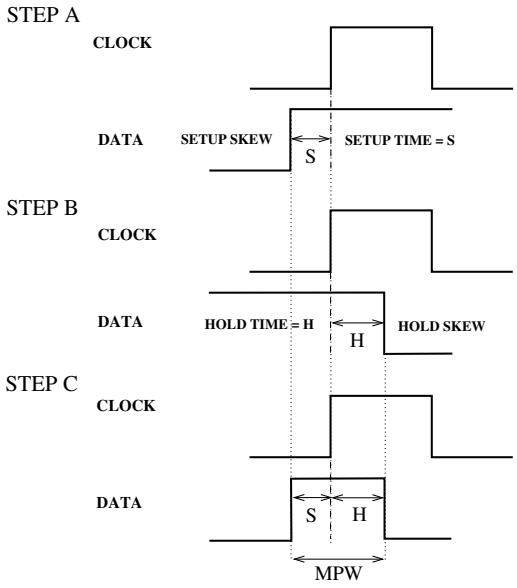


Figure 9.4: Illustration of over-optimistic constraint characterization. Step A: Setup time characterization at sufficiently large hold skew. Step B: Hold time characterization at sufficiently large setup skew. Step C: Minimum data pulse width (MPW) for which the sequential cell fails.

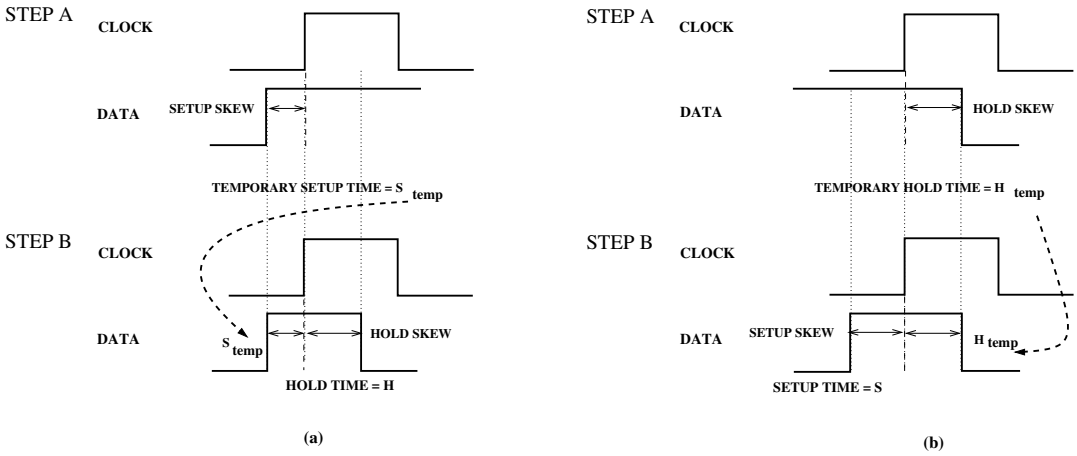


Figure 9.5: Illustration of pessimistic constraint characterization: (a) Hold time characterization. Step A: A temporary setup time characterization at sufficiently large hold skew. Step B: Hold time characterization when the setup skew is equal to the temporary setup time found in Step A. (b) Setup time characterization. Step A: A temporary hold time characterization at sufficiently large setup skew. Step B: Setup time characterization when the hold skew is equal to the temporary hold time found in Step A.

The hold time characterization is shown in Fig. 9.5(a). In step A, a temporary setup time, which is used in step B, is characterized at a sufficiently large hold skew. In step B, this temporary setup time is used as the counterpart skew to characterize the hold time. The hold time is therefore highly pessimistic. The process is similar for setup time characterization shown in Fig. 9.5(b) which produces a highly pessimistic result. Note that the temporary setup and hold times are used only to overcome over-optimism of the characterization process and do not reflect the final setup and hold times. Both of the over-optimistic and pessimistic cases should be avoided as the optimistic case can cause circuit failures after fabrication whereas the pessimistic case can cause false violations during STA.

Issue 2. The interdependence between the setup time and hold time can be used to improve the timing analysis results of a circuit. Therefore, if this dependence is considered but not exploited in STA, an opportunity to reduce the number of timing violations and improve the slack is lost.

In [209], the first issue is resolved by considering this interdependence. However, *only one* interdependent pair of setup and hold times is considered. The interdependence therefore is not exploited to improve the slacks in STA. A comprehensive methodology is proposed in this chapter that solves both issues and offers reliable integration into an industrial STA tool, demonstrating up to a 53% reduction in the number of constraint violations as well as up to a 48% reduction in the worst negative

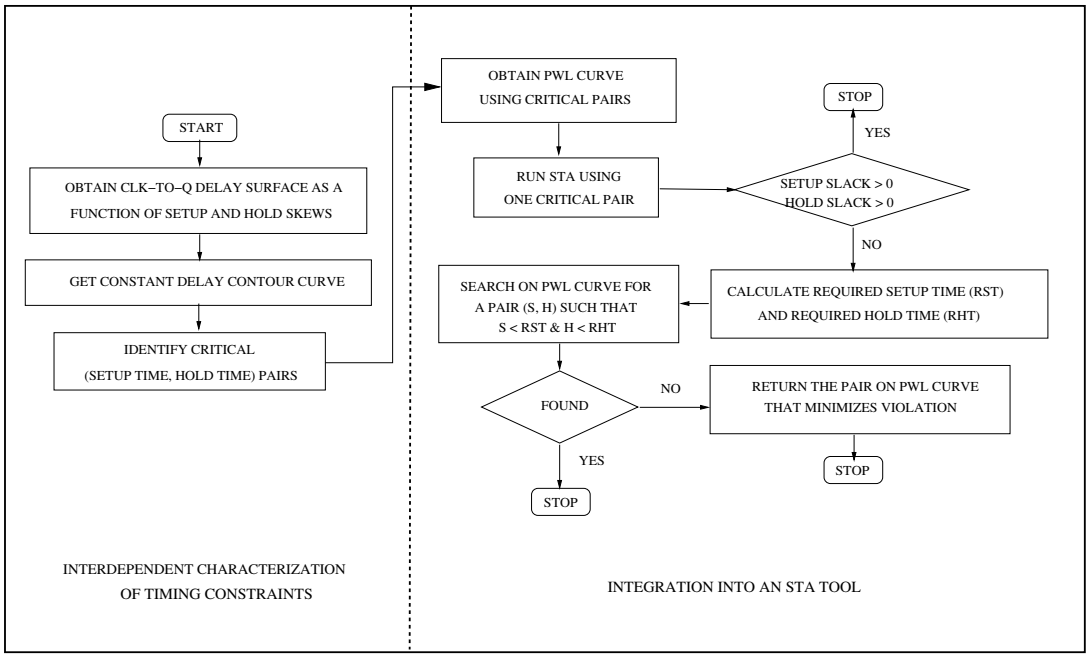


Figure 9.6: Flowchart summarizing the proposed methodology including interdependent characterization of timing constraints and integration into an STA tool.

slack.

9.3 Proposed Methodology

The proposed methodology is described in five steps below and also illustrated in the flowchart shown in Fig. 9.6. Each step is explained in a separate subsection. Steps 1 to 4 describe the first phase of the methodology related to the interdependent characterization of the timing constraints. Step 5 describes the second phase related to the efficient integration into an STA tool where the interdependence is exploited to remove timing violations.

- **Step 1.** For each data and clock slew combination, the circuit is simulated to obtain CLK-to-Q delay surfaces, each a function of *independently varying* setup and hold skews.

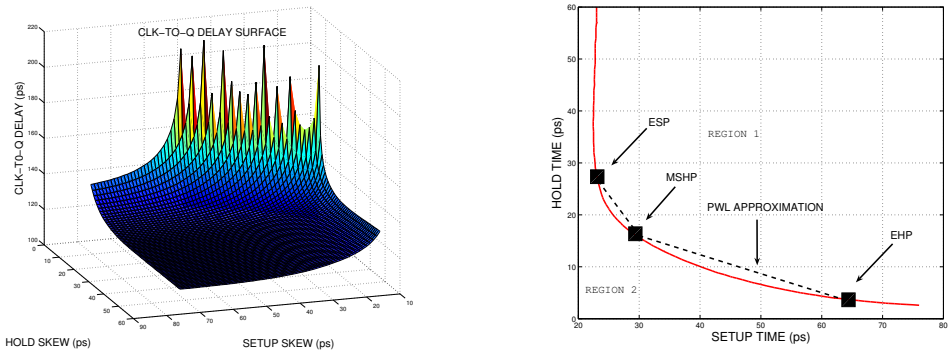
- **Step 2.** A contour, set to a constant CLK-to-Q delay, is obtained for each surface. Each point on the contour represents an *interdependent* pair of setup and hold times.

- **Step 3.** Critical pairs are identified on each contour. A cell library is created for each of these pairs.

- **Step 4.** At least two of the critical pairs are used to obtain a piecewise linear (PWL) curve. This curve approximates the contour with slightly pessimistic setup and hold times allowing a potentially infinite number of pairs.

- **Step 5.** A cell library is used for STA. If a violation occurs, the other libraries and the PWL curve are utilized to resolve the violation using the proposed algorithm. The process of violation resolution is therefore automated within STA.

Note that in Step 3, the creation of a cell library per critical pair is proposed to only simplify the proposed technique. In practice, a cell library per critical pair is not required, and the details are provided in Section 9.3.4.



(a) (b)

Figure 9.7: Interdependent constraint characterization of a sequential cell: (a) CLK-to-Q delay surface as a function of independently varying setup skew and hold skew, (b) The contour at a 10%-degraded CLK-to-Q delay. The contour includes the critical pairs as well as a piecewise linear (PWL) approximation. Regions 1 and 2 are the pessimistic and optimistic regions, respectively.

The limitations of this methodology, *e.g.*, increase in library characterization time and STA runtime, are discussed in Section 9.3.6.

9.3.1 CLK-to-Q Delay Surface

For a specific data and clock slew combination, the CLK-to-Q delay is a function of both the setup skew and hold skew. A typical delay surface is shown in Fig. 9.7(a). Note that in this figure, the CLK-to-Q delay increases when the skews either independently or simultaneously decrease. The multiple peaks on the surface mark the boundary beyond which the skews are excessively small and the sequential cell can no longer latch the data.

A delay surface is generated by independently varying the setup and hold skews. Independent variations allow the generation of the *actual* delay surface and simplify the library characterization process at the expense of additional circuit simulations.

Every point on the CLK-to-Q delay surface corresponds to a skew pair, denoted as (setup skew, hold skew). If a particular pair on this surface is identified as the final (setup time, hold time) pair, Issue 1 in Section 9.2.2 is resolved because the setup and hold times at this point are now interdependent.

Different approaches exist to select a final pair on the surface depending upon the definition of the crossover point [209]. Irrespective of the approach used, it is highly likely that there will be multiple final pairs satisfying this definition.

9.3.2 Constant Delay Contour

A definition of a common crossover point is a specific per cent degradation in the CLK-to-Q delay. Once the CLK-to-Q delay surface in three dimensions is obtained, all of the final pairs can be extracted from the constant delay contour as a per cent of the crossover point.

The contour obtained at a 10%-degraded CLK-to-Q delay is depicted in Fig. 9.7(b). Each (setup time, hold time) pair on this contour is interdependent and valid. Furthermore, any pair in region 1 is also valid with additional pessimism whereas any pair in region 2 is invalid, as the pairs in this region are optimistic.

Two important conclusions can be drawn from this contour:

1) Rather than individual and independent setup and hold times, there are multiple and interdependent (setup time, hold time) pairs. Any pair can be chosen depending upon the potential to remove timing violations.

2) As indicated in Fig. 9.7(b), the setup and hold times are inversely proportional, which can also be verified by least-square regression analysis. Hence, a small setup time can be obtained at the expense of a large hold time (or vice versa).

9.3.3 Critical Pairs on Contour

The following pairs on the contour are defined as critical pairs because these pairs are appropriate candidates to include in a cell library. For each pair X, the notation $X = (s, h)$ where $s[X] = s$ and $h[X] = h$ is used.

Definition 7 P is defined as the set of all (s, h) pairs on the contour where s is the setup time and h is the hold time.

Definition 8 S and H are defined as the set of all setup times s and hold times h on the contour, respectively.

Definition 9 The *minimum setup pair MSP* is defined as the pair (s, h) in P such that s is minimum in S. More formally,

$$MSP = (s, h) \in P \; such \; that \; s = \min_{\forall s \in S}(s). \tag{9.7}$$

Definition 10 The *minimum hold pair MHP* is defined as the pair (s, h) in P such that h is minimum in H. More formally,

$$MHP = (s, h) \in P \text{ such that } h = \min_{\forall h \in H}(h). \tag{9.8}$$

The setup (hold) time of MHP (MSP) can be impractically large to minimize the corresponding hold (setup) time. If a slightly larger, but bounded increase (controlled by a parameter ϵ) in hold (setup) time is allowed, the setup (hold) time of MHP (MSP) can be reduced to an acceptable level. This reduction is achieved by the effective hold (setup) pair.

Definition 11 The *effective setup pair ESP* is formally defined as

$$ESP = (s, h) \in P \text{ such that } s = s[MSP] + \epsilon_s * |s[MSP]|, \tag{9.9}$$

where ϵ_s is a user-controlled, nonnegative parameter.

Definition 12 The *effective hold pair EHP* is formally defined as

$$EHP = (s, h) \in P \text{ such that } h = h[MHP] + \epsilon_h * |h[MHP]|, \tag{9.10}$$

where ϵ_h is a user-controlled, nonnegative parameter.

Definition 13 The *minimum setup-hold pair MSHP* is defined as the pair (s, h) in P such that the summation of s and h is the minimum. More formally,

$$MSHP = (s, h) \in P \text{ such that } s + h = \min\{\sum_{\forall (s,h) \in P} (s + h)\}. \tag{9.11}$$

Note that MSHP corresponds to the minimum data pulse width that can be captured and stored by the sequential cell.

The distinction between the minimum and effective pairs can be illustrated by evaluating the minimum pulse width of the data signal. The minimum pulse width of the data signal is determined by summing the setup and hold times.

The variation of the minimum pulse width with respect to the setup and hold times is shown in Fig. 9.8. If the minimum constraints are used rather than effective constraints, the minimum pulse width increases significantly. Note that, at zero ϵ_s and ϵ_h, the effective constraints are equal to the minimum constraints.

9.3.4 Piecewise Linear (PWL) Approximation of Contour

In order to fully exploit this interdependence, the STA tool should use at least two (setup time, hold time) pairs on the contour. Since library characterization is expensive in time and memory, it may be impractical to generate more than two or three pairs. These pairs may, however, be insufficient to remove all violations.

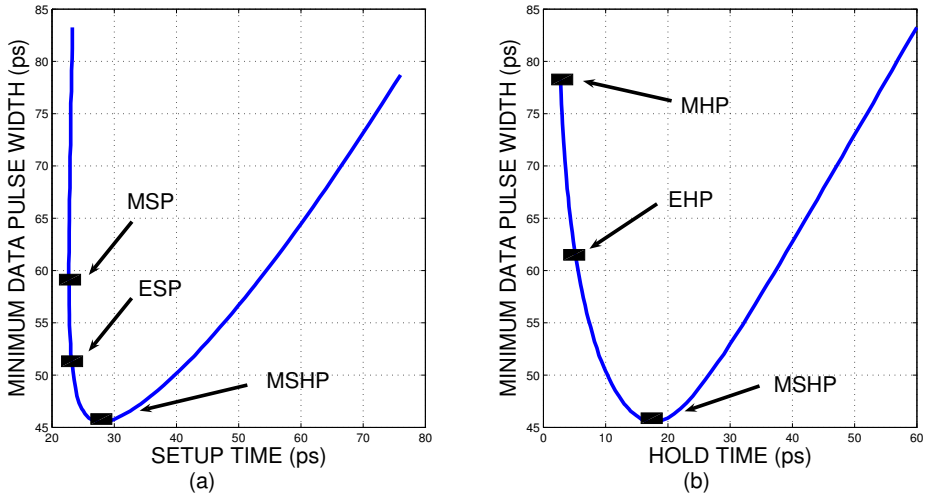


Figure 9.8: The relationship between the timing constraints and the minimum pulse width: (a) setup time vs. minimum pulse width, (b) hold time vs. minimum pulse width.

An improvement is to generate critical pairs and connect these critical pairs using a PWL curve to approximate the contour. For example, an approximation with two line segments, *i.e.*, linear, can be obtained by connecting ESP and EHP, and an approximation with three line segments can be obtained by connecting ESP, MSHP, and EHP.

To avoid optimism in PWL approximation, the contour should be convex with respect to PWL approximation. That is, as illustrated in Fig. 9.7(b), the line segments of the PWL approximation of the contour should remain in region 1.

The linear representation of the contour at three different data and clock slew pairs is shown in Fig. 9.9. The slews are determined with respect to 10% and 90% thresholds of the signal voltages. Each linear curve is obtained using ESP and EHP

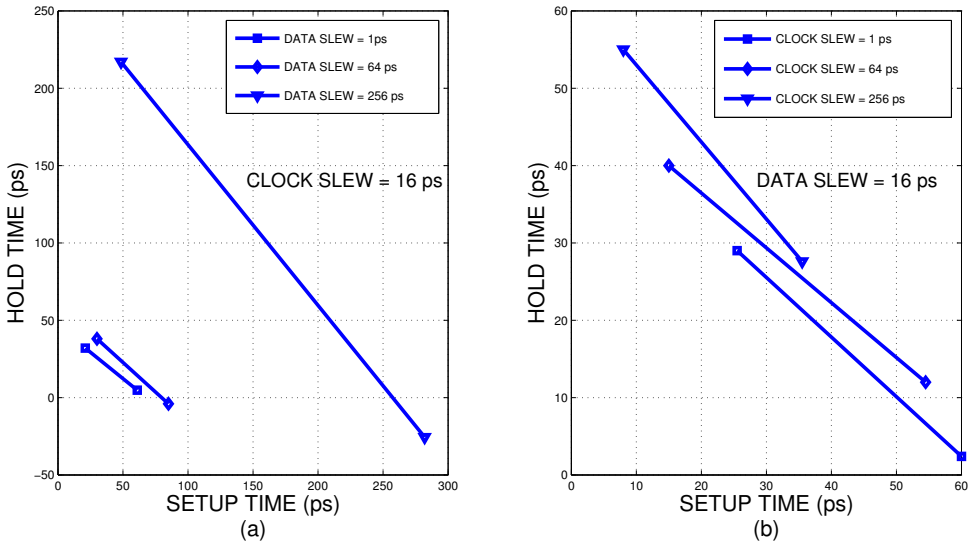


Figure 9.9: Linear representations using two pairs, ESP and EHP, where $\epsilon_s = 0$ and $\epsilon_h = 0.2$: (a) at different data slews, (b) at different clock slews.

on the contour when $\epsilon_s = 0$ and $\epsilon_h = 0.2$. Note that the number of critical pairs used in the PWL curve represents a tradeoff between accuracy and complexity.

To represent critical pairs in cell libraries for sequential cells, the following approach is proposed. Current cell libraries generally contain two lookup tables for setup time, and two tables for hold time (one table is defined for rising edge data signal and the other table for falling edge data signal). Therefore, if two critical pairs are used, there are two possible options depending upon the flexibility of the library format. If a slight modification is allowed on the library format, the proposed methodology does not require more tables: the existing tables for setup and hold times should be modified to contain interdependent (setup time, hold time) pairs instead of independent setup time and hold times. The library, therefore, still contains

```
FIND-BEST-PAIR(Ordered pairs P)
1.  Select a valid pair (s_0, h_0) from P
2.  Calculate hold slack = HS (as in Section 9.1.1)
3.  Calculate setup slack = SS (as in Section 9.1.1)
4.  if (HS ≥ 0 and SS ≥ 0) then
5.    return < (s_0, h_0), found >
6.  Calculate maximum required setup time = RST = s_0 + SS
7.  Calculate maximum required hold time = RHT = h_0 + HS
8.  for each (s_i, h_i) in P where i = 1, ··· , (|P| − 1) do
9.    if (s_i ≤ RST) then
10.     h = ((h_i − h_{i−1}) / (s_i − s_{i−1})) * (RST − s_{i−1}) + h_{i−1}
11.     if (h ≤ RHT) then
12.       return < (RST, h), found >
13. return < (RST, RHT), not found >
```

Figure 9.10: The algorithm FIND-BEST-PAIR to determine the (setup time,hold time) pair that removes the violation using the pairs P on the PWL approximation of the contour curve. This algorithm is performed only for sequential cells with violations.

four tables all of which consist of (setup time, hold time) pairs. If the current library format cannot be modified, the proposed methodology requires four additional tables: two tables for setup time, and two tables for hold time to sufficiently represent (setup time, hold time) pairs.

9.3.5 Integration of Interdependent Characterization into STA

The algorithm FIND-BEST-PAIR shown in Fig. 9.10 is proposed to exploit the interdependence of setup and hold times in STA.

The Algorithm

FIND-BEST-PAIR reads in the PWL representation P of the contour as an input. This representation is obtained using critical (setup time, hold time) pairs as described in Section 9.3.4. The critical pairs are sorted in descending order of setup times where two successive pairs imply a line segment. The PWL representation P therefore contains a set of connected line segments.

At line 1 of FIND-BEST-PAIR, the (setup time, hold time) pair with the largest setup time is selected from the input. Note that this pair can be any of the critical pairs on P, but a practical approach is to use *EHP*, as hold times are typically more critical. The setup and hold slacks are determined as described in Section 9.1.1. Both slacks are checked for violations. If both slacks are nonnegative, the algorithm terminates, returning the pair as the "best" pair. If one or both of the slacks are negative, these slacks are used to compute the *required setup time* (RST) and *required hold time* (RHT) to remove the violations. The loop at line 8 determines if such a pair exists in P. This line enables the tool to dynamically switch between interdependent (setup time, hold time) pairs to determine the "best" pair. If a pair is found that can remove the violations, that pair is returned as the "best" pair at line 12. If no such pair exists, the required setup and hold times are returned at line 13 with a warning that no solution is possible. Note that the required setup and hold times can be used to search for a pair that minimizes the violations.

The loop at line 8 iterates over each pair (s_i, h_i) in P (except the first pair since the first pair has already been checked before the loop) in order to determine if the line segment from (s_i, h_i) to (s_{i-1}, h_{i-1}) contains any pairs that can resolve both setup and hold time violations. The condition at line 9 determines if s_i is smaller than or equal to the RST. If this condition is satisfied, line 10 computes the hold time h such that the pair (RST, h) is on the line segment from (s_{i-1}, h_{i-1}) to (s_i, h_i). The computation of the hold time h is achieved using the equation of the line segment. If the condition at line 11 is also satisfied, the pair (RST, h) can resolve both violations. The algorithm terminates, returning this pair as the "best" pair at line 12. If line 12 is not reached during the iterations of the loop at line 8, the algorithm terminates with no solution, indicated by the returned value at line 13.

Examples

The behavior of the algorithm is illustrated for two different cases in Fig. 9.11. The situation where FIND-BEST-PAIR determines a solution and removes the violation is illustrated in Fig. 9.11(a). The situation where FIND-BEST-PAIR fails to determine a solution is illustrated in Fig. 9.11(b). For both figures, the original contour curve is represented by the function $h(s)$, the hold time as a function of the setup time. Note that, for simplicity, only two pairs are used in both cases for the PWL approximation. Therefore, only one line segment, defined by the pairs (s_0, h_0) and (s_1, h_1) exists as

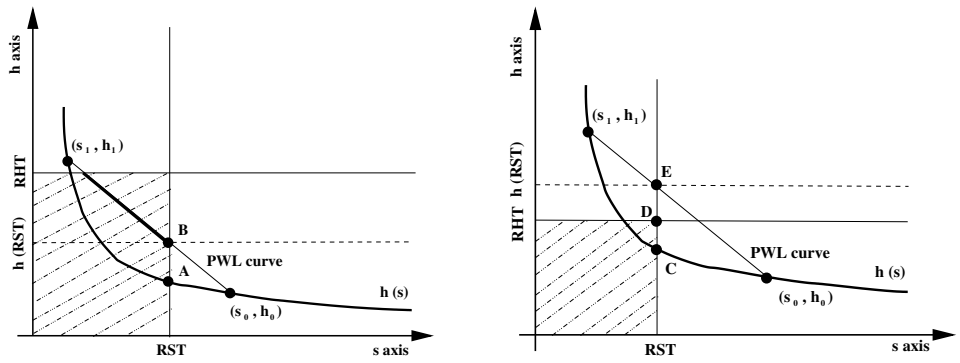


Figure 9.11: Illustration of the cases for which FIND-BEST-PAIR (a) determines a solution and (b) cannot determine a solution. The actual contour curve is represented as $h(s)$; the hold time as a function of the setup time. RST is the required setup time, and RHT is the required hold time. For (a), the algorithm returns the point B as the "best pair," the bold region of the PWL curve represents all of the pairs that can be returned. For (b), the algorithm returns the point D, because the PWL curve does not intersect with the shaded region, indicating that a solution is not possible.

shown in Fig. 9.11.

The shaded regions in Fig. 9.11 represent all of the pairs whose setup time is smaller than or equal to the *required setup time* (RST) and whose hold time is smaller than or equal to the *required hold time* (RHT). The pairs that are at the intersection of the shaded region and the PWL approximation can therefore resolve both violations. For Fig. 9.11(a), FIND-BEST-PAIR exits at line 12, returning the pair B since pair B is at the intersection. For Fig. 9.11(b), FIND-BEST-PAIR exits at line 13, returning the pair D because the PWL approximation does not intersect with the shaded region. In this case, therefore, no pair exists that can remove the setup and hold time violations.

Pairs A and C shown, respectively, in Figs. 9.11(a) and 9.11(b) illustrate the tradeoff between accuracy and complexity in terms of the number of pairs included in the PWL approximation. Pair A is on the actual contour in Fig. 9.11(a) and produces a larger hold slack value. Pair A, however, cannot be returned because it is not included in the PWL approximation. Similarly, for Fig. 9.11(b), pair C is also on the actual contour and can remove both violations since it is inside the shaded region. Pair C, however, cannot be returned because it is not included in the PWL approximation. A tradeoff therefore exists between accuracy and complexity in terms of the number of pairs included in the PWL approximation.

As shown in Fig. 9.11(a), FIND-BEST-PAIR returns the pair that minimizes the hold time. However, any pair that is at the intersection of the PWL curve and the shaded region (shown as the darker portion in Fig. 9.11(a)) can be returned. Note that adapting the algorithm to return any other valid pair is possible.

Complexity Analysis

In order to evaluate the time complexity of FIND-BEST-PAIR, the total number of the sequential cells in the circuit is assumed to be N, which is a small fraction of the total number of cells in the circuit. FIND-BEST-PAIR executes for each sequential cell that violates a setup or hold time constraint. In the worst case, all of the N sequential cells have a timing violation. FIND-BEST-PAIR, therefore, executes N

times in the worst case. For each execution, the algorithm requires time proportional to the number of pairs in the input P because the loop at line 8 iterates P times and each iteration takes constant time. The time complexity of FIND-BEST-PAIR is therefore $O(|P|)$, and the complexity of resolving N violations using this algorithm is $O(N|P|)$. In practice, P is expected to be two or three pairs. FIND-BEST-PAIR therefore executes in constant time and the N iterations of FIND-BEST-PAIR execute in $O(N)$ time. This time complexity is optimal since an optimal algorithm for violation resolution should execute in constant-time per violation.

9.3.6 Limitations

The primary limitations of the proposed methodology are twofold: 1) the constraint characterization time increases for sequential cells, and 2) the STA runtime increases if there is a timing violation. The first limitation is due to the generation of the delay surfaces, and the second limitation is due to the use of multiple constraints during STA.

The second limitation is not significant as the STA runtime increases only if there is a timing violation after analyzing the first (setup time, hold time) pair. Furthermore, if there is a violation, any other resolution method will also increase the overall time. Since the current resolution methods are not automated as proposed in this study, the time for violation resolution may actually decrease with the proposed methodology.

Note that the second limitation can be mitigated by using linear approximation.

In order to quantify the first limitation, assume N_s and N_h denote the number of setup skews and hold skews to sweep over for characterization. For the over-optimistic characterization approach, illustrated in Fig. 9.4, the total number of simulations to characterize both setup and hold times is equal to $N_s + N_h$. For the pessimistic characterization approach, illustrated in Fig. 9.5, the total number of simulations doubles to $2(N_s + N_h)$ due to additional steps. For the proposed characterization approach, the total number of simulations increases to $N_s N_h$ in the worst case to generate the entire CLK-to-Q delay surface, illustrated in Fig. 9.7(a).

To reduce the number of simulations for the proposed approach, a simple heuristic is to reduce the number of skews to be swept by eliminating those skew pairs that do not change the delay surface. For example, the locations of critical pairs on the contour in Fig. 9.7(b) indicate that it is not necessary to sweep the skew pairs within region 1.

9.4 STA Results

A 90 nm library is used as a template to generate three new cell libraries: library 1, library 2, and library 3. As discussed in Section 9.3 and Section 9.3.4, separate libraries are generated to simplify the evaluation of the methodology with an industrial STA environment. The sequential cells of each library are characterized using

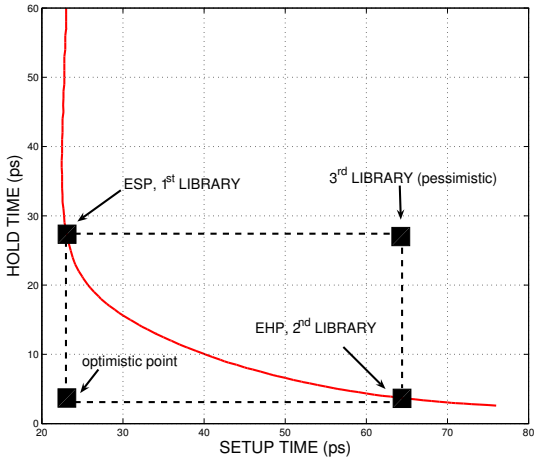


Figure 9.12: Constant delay contour curve illustrating the characterization points of the three prototype libraries.

HSPICE with BSIM4/BSIM3 models under typical corner conditions.

The library characterization points for these three libraries are illustrated on a contour at 10%-degraded CLK-to-Q delay in Fig. 9.12. Both library 1 and 2 are on the contour: library 1 is at ESP and library 2 is at EHP. Library 3 is not on the contour; this last library uses setup times from EHP and hold times from ESP, and therefore, is an example of independent and pessimistic characterization. The optimistic point that results from using relatively large counterpart skews is also shown in the figure. Note that the contour in Fig. 9.12 represents a single data and clock slew combination of 25 ps each. For STA results, the same contour is obtained for each data and clock slew combination that exists in the cell library. Different gains are obtained by using library 1 or library 2 over library 3 depending upon these slew combinations.

Table 9.1: Absolute (abs) and Relative (rel) improvements of two circuits with respect to library 3. WNS is the worst negative slack. ΔWNS is the increase in WNS, and ΔN is the decrease in the number of violations.

Circuit	Period (ps)	Library	WNS (ps) setup	WNS (ps) hold	Number of violations setup	Number of violations hold
Circuit A	1500	3	-1003	-488	361	1868
		1	-790	-488	285	1868
		2	-1003	-307	361	1684
Circuit B	2500	3	-766	-26	1977	1
		1	-397	-26	924	1
		2	-766	0	1977	0

Circuit	Period (ps)	Library	ΔWNS (ps): abs (rel) setup	ΔWNS (ps): abs (rel) hold	ΔN: abs (rel) setup	ΔN: abs (rel) hold
Circuit A	1500	1	213 (21.2%)	-	76 (21.0%)	-
		2	-	181 (37.1%)	-	184 (9.9%)
Circuit B	2500	1	369 (48.2%)	-	1053 (53.3%)	-
		2	-	26 (100.0%)	-	1 (100.0%)

An industrial sign-off STA tool, PrimeTime, is used to evaluate each prototype library on two industrial circuits: circuit A and circuit B. Both circuits are networking cores with nearly 20K cells. The clock frequencies of circuit A and circuit B are set to 666 MHz and 400 MHz, respectively.

From STA, the smallest negative slack value, referred to as the worst negative slack (WNS), and the number of violations are obtained for each of the endpoints. The STA results are listed in Table 9.1. Each row corresponds to one simulation with one circuit and one library.

WNS and the number of violations from library 3 are treated as a baseline, and the absolute and relative improvements are determined in WNS and number of violations with respect to library 3. Improvements in WNS and number of violations correspond, respectively, to an increase in WNS and a decrease in the number of violations. Note

that library 1 illustrates improvements in the setup time without affecting the hold time and library 2 illustrates improvements in the hold time without affecting the setup time. This result is because library 1 is characterized at ESP and library 2 is characterized at EHP, as shown in Fig. 9.12.

As listed in Table 9.1, the improvement in setup WNS is 369 ps (or 48.2%). This improvement corresponds to nearly 15% of the clock period. The improvement in hold WNS is 181 ps (or 37.1%). In terms of the number of violations, the improvement in the setup case is 53.3% and in the hold case is 9.9%. Note that for hold time improvements, the case represented by the last row, where the only hold time violation is removed, is ignored. The improvement in setup WNS provides a 14% and 15% increase in performance for circuits A and B, respectively. Furthermore, the improvement in the hold WNS reduces the required circuit modifications to remove the hold time violations.

These improvements can also be illustrated by means of slack histograms over all the endpoints rather than a single number like WNS. Histograms for the two circuits are shown in Figs. 9.13 and 9.14. For both histograms, there is a shift towards the positive side, indicating improvements in most of the slack values. The baseline is the slacks from library 3.

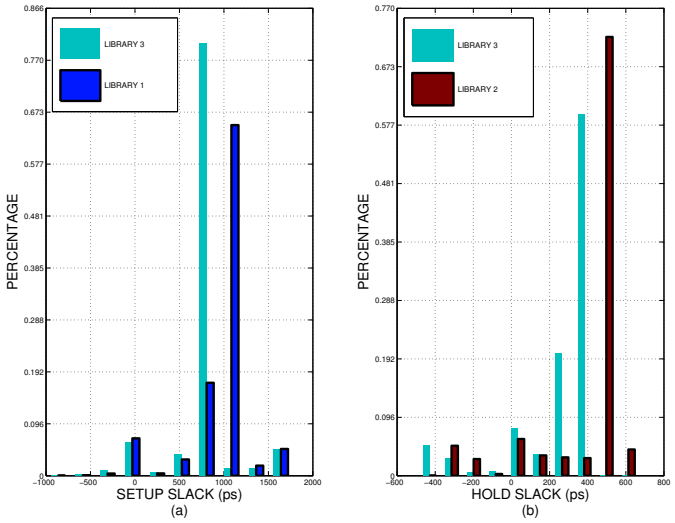


Figure 9.13: Slack histograms for circuit A: (a) setup slack histograms of library 3 and library 1, (b) hold slack histograms of library 3 and library 2.

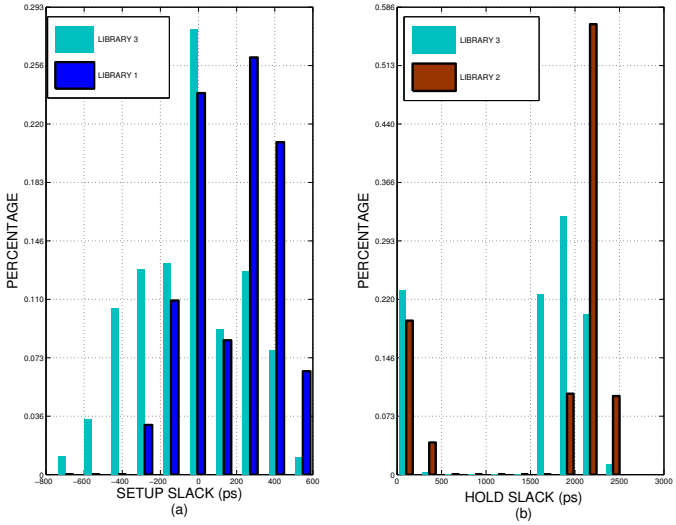


Figure 9.14: Slack histograms for circuit B: (a) setup slack histograms of library 3 and library 1, (b) hold slack histograms of library 3 and library 2.

9.5 Summary

A two-phase methodology is presented to exploit setup and hold time interdependence in static timing analysis. The issues related with independent characterization, *i.e.*, the over-optimism and pessimism, are discussed and illustrated. Interdependent constraint characterization to remove these problems is proposed in the first phase of the methodology. In the second phase, an efficient algorithm is presented to integrate this interdependence into an STA tool, exploiting multiple constraint pairs to reduce timing violations. The proposed algorithm automates the violation resolution process in a linear time complexity. The methodology is validated using industrial circuits and an industrial STA tool. The results show up to a 53% reduction in the number of constraint violations as well as up to a 48% reduction in the worst negative slack, corresponding to a 15% decrease in the clock period. These interdependent setup and hold time constraints are exploited to reduce delay uncertainty in the following chapter.

Chapter 10

Reducing Delay Uncertainty Using Interdependent Timing Constraints

The interdependence of timing constraints, *i.e.*, setup and hold times, has been described in Chapter 9. Significant pessimism has been reduced by exploiting these interdependent timing constraints in static timing analysis. The interdependence of the setup and hold times is important not only during the analysis (or verification) phase, but also during the design phase of a circuit.

Exploiting interdependence to reduce *delay uncertainty* in high performance synchronous digital ICs is explored in this chapter. As described in Chapter 2, delay uncertainty can be caused by switching noise in these circuits. Specifically, a variation in the power supply voltage will change the delay of a logic gate, producing delay uncertainty [28], [73]. A circuit should therefore have sufficient tolerance against these variations to function properly while satisfying a target clock frequency [212], [213], [214].

A methodology is proposed in this chapter for designing a robust circuit exhibiting higher tolerance to process and environmental variations. This higher tolerance is achieved by exploiting the interdependence between the setup and hold times, significantly reducing the delay uncertainty caused by circuit variations. An algorithm is proposed to determine the interdependent setup-hold pair of a register during the design process. A data path designed with the proposed setup-hold pair improves the overall tolerance to variations, producing a more robust circuit.

A methodology is evaluated for several technologies to determine the overall reduction in delay uncertainty. Power supply noise is treated as the source of the variation. The results of the case study demonstrate the significance of setup-hold interdependence in reducing delay uncertainty, thereby compensating for process and environmental variations.

The rest of the chapter is organized as follows. The importance of interdependence on the design process is explained in Section 10.1. A methodology to reduce delay uncertainty and compensate for variations is described in Section 10.2. A case study is presented in Section 10.3 to evaluate the efficacy of exploiting this interdependence in reducing delay uncertainty due to power noise. Finally, the chapter is summarized in Section 10.4.

10.1 Problem Formulation

A methodology to interdependently characterize timing constraints has been described in Chapter 9. The advantages of using interdependent setup-hold pairs in static timing analysis have also been reviewed. The computational efficiency of interdependent setup-hold time characterization can be improved through state transition equations [215]. Interdependent setup-hold times have also been exploited in the statistical static timing analysis process [216].

Interdependence of the timing constraints not only reduces pessimism in timing analysis, but also provides an opportunity to improve the tolerance of a circuit to process and environmental variations. A methodology is proposed in this chapter to reduce the delay uncertainty caused by process and environmental variations. The interdependence between the setup and hold times is exploited to design a more robust circuit which exhibits additional tolerance to variations.

The contour curve illustrated in Fig. 9.7(b) in Chapter 9 can be approximated as a linear line using two critical pairs: *minimum setup pair* (MSP) and *minimum hold pair* (MHP). MSP and MHP refer, respectively, to a pair on the contour with the minimum setup time and minimum hold time. Note that this approximation is accurate since any point above the curve, *i.e.*, in region 1, is a valid pair. An approximation of the contour using two critical pairs is illustrated in Fig. 10.1 where the pairs MSP and MHP are represented, respectively, as ($T_{S,min}$, $T_{H,max}$) and

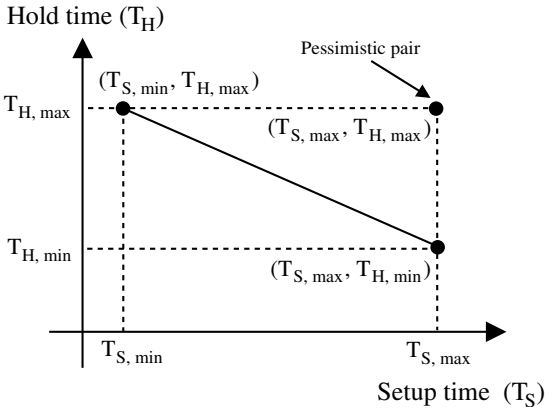


Figure 10.1: Linear approximation of the contour curve using two pairs: ($T_{S,min}$, $T_{H,max}$) and ($T_{S,max}$, $T_{H,min}$).

($T_{S,max}$, $T_{H,min}$).

According to (9.1) and (9.2), at a specific clock period, the delay of the data path T_D is dependent upon both the setup and hold times. Referring to Fig. 9.1, the delay of the data path should satisfy

$$T_{Cf} + T_H - T_{Ci} \leq T_D, \tag{10.1}$$

$$T_D \leq T_{Cf} + T_{CP} - (T_{Ci} + T_S), \tag{10.2}$$

where (10.1) and (10.2) determine, respectively, the lower and upper bounds of the data path delay.

The allowable range of T_D is minimized if the pessimistic pair ($T_{S,max}$, $T_{H,max}$) is used, causing the circuit to be overdesigned. Which specific (setup, hold) pair should be chosen to design the circuit is unclear even if the interdependence is known since

multiple valid pairs are available. For example, if the pair ($T_{S,min}$, $T_{H,max}$) is used, the lower bound constraint of the data path delay is difficult to satisfy since the hold time is large. Hence, the data path delay should be increased by inserting additional stages, dissipating unnecessary power. Alternatively, if the pair ($T_{S,max}$, $T_{H,min}$) is used, the upper bound constraint on the data path delay is difficult to satisfy since the setup time is large. Consequently, the data path delay should be lowered by inserting an additional register to satisfy the target frequency, also causing unnecessary power consumption. Furthermore, both pairs ($T_{S,min}$, $T_{H,max}$) and ($T_{S,max}$, $T_{H,min}$) exhibit low tolerance to process and environmental variations since the range of valid setup times $T_{S,max} - T_{S,min}$ and hold times $T_{H,max} - T_{H,min}$ is not exploited.

It is therefore important to determine the appropriate (setup time, hold time) pair during the design process that lowers power consumption, satisfies the required delay, and increases the robustness of the circuit to achieve higher tolerance to process and environmental variations. A methodology is described in this chapter to determine the appropriate (setup, hold) pair during the design process to improve the tolerance of a circuit to variations while increasing the maximum operating frequency.

10.2 Methodology to Reduce Delay Uncertainty and Compensate for Variations

An efficient technique to obtain the interdependent setup-hold time characteristics is explained in Section 10.2.1. A procedure to reduce delay uncertainty and compensate for variations is described in Section 10.2.2. The effect of variations on the interdependent characterization process is discussed in Section 10.2.3. The amount of compensation achieved by the proposed methodology is determined in Section 10.2.4. Finally, the relationship between setup-hold time interdependence and the permissible range is reviewed in Section 10.2.5.

10.2.1 Obtaining Linear T_S vs. T_H Relationship

The first step to reduce delay uncertainty is to obtain the linear T_S vs. T_H relationship for a register. Two critical pairs, MSP and MHP, should be identified to define this line, as illustrated in Fig. 10.1. The generation of the CLK-to-Q delay surface to obtain these critical pairs is inefficient since a large number of simulations is required. An alternative procedure is proposed to obtain MSP and MHP without generating the CLK-to-Q delay surface. Referring to Fig. 10.1, the four points defining MSP and MHP, *i.e.*, $T_{S,min}$, $T_{H,max}$, $T_{H,min}$, and $T_{S,max}$, are obtained as follows where the corresponding waveforms of the clock and data signals are illustrated in

Fig. 10.2.

- $T_{S,min}$: The register is simulated to obtain the CLK-to-Q delay as a function of setup skew at a sufficiently high hold skew. The setup skew corresponding to a 10% degradation in delay is chosen as $T_{S,min}$.

- $T_{H,max}$: The register is simulated to obtain the CLK-to-Q delay as a function of hold skew at *setup skew* $= T_{S,min}$. The hold skew corresponding to a 10% degradation in delay is chosen as $T_{H,max}$.

- $T_{H,min}$: The register is simulated to obtain the CLK-to-Q delay as a function of hold skew at a sufficiently high setup skew. The hold skew corresponding to a 10% degradation in delay is chosen as $T_{H,min}$.

- $T_{S,max}$: The register is simulated to obtain the CLK-to-Q delay as a function of setup skew at *hold skew* $= T_{H,min}$. The setup skew corresponding to a 10% degradation in delay is chosen as $T_{S,max}$.

Note that the critical pairs are obtained with significantly fewer simulations using this technique as compared to generating the complete CLK-to-Q delay surface. The linear relationship can be represented by the critical pairs as $T_H = f(T_S)$ or, equivalently,

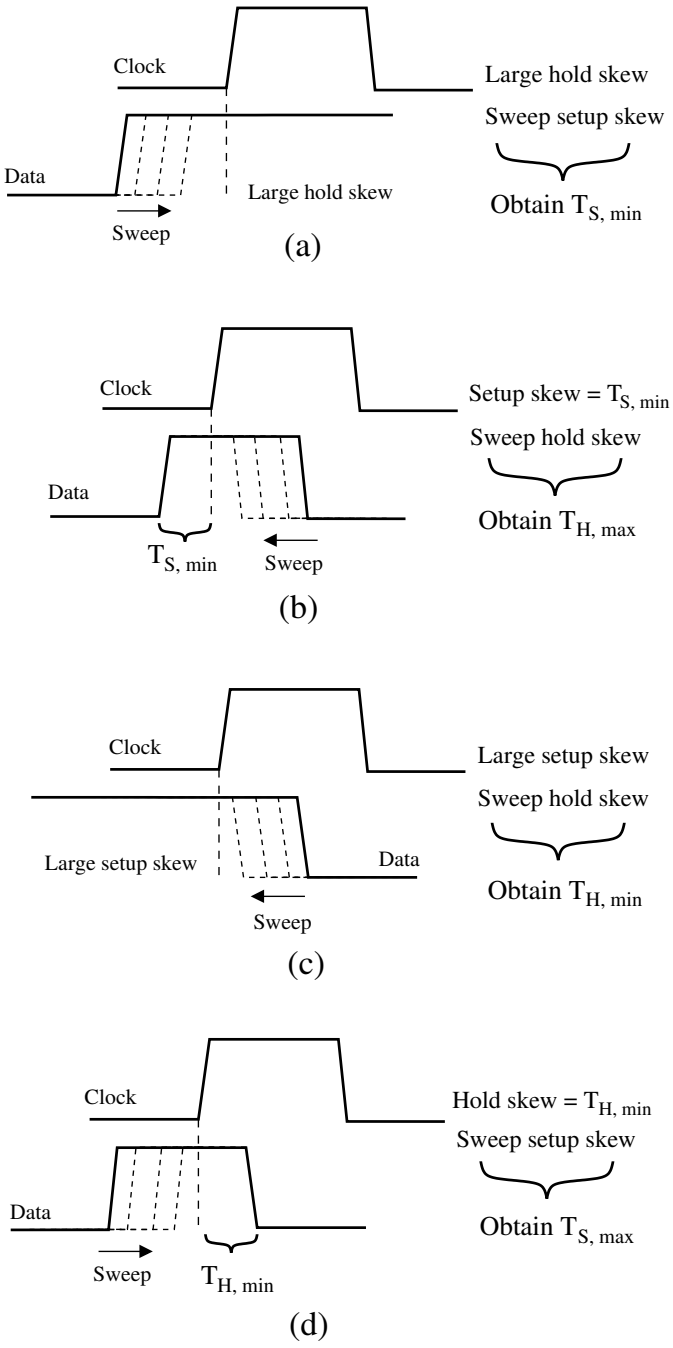


Figure 10.2: Efficiently obtaining the critical setup-hold pairs of a register without generating a three-dimensional CLK-to-Q delay surface: (a) $T_{S,min}$, (b) $T_{H,max}$, (c) $T_{H,min}$, (d) $T_{S,max}$.

$T_S = f^{-1}(T_H)$,

$$T_H = f(T_S) = \frac{T_S T_{H,r} - T_{H,max} T_{S,max} + T_{H,min} T_{S,min}}{-T_{S,r}}$$
$$\text{for} \quad T_{S,min} < T_S < T_{S,max}, \tag{10.3}$$

$$T_S = f^{-1}(T_H) = \frac{-m T_H T_{S,r} - T_{H,min} T_{S,min} + T_{H,max} T_{S,max}}{T_{H,r}}$$
$$\text{for} \quad T_{H,min} < T_H < T_{H,max}, \tag{10.4}$$

where the range of valid setup times $T_{S,r}$ and hold times $T_{H,r}$ are, respectively,

$$T_{S,r} = T_{S,max} - T_{S,min}, \tag{10.5}$$

$$T_{H,r} = T_{H,max} - T_{H,min}. \tag{10.6}$$

The inverse proportionality between the setup and hold times can be exploited to reduce delay uncertainty, as described in the following subsection.

10.2.2 Procedure to Reduce Delay Uncertainty

The available range of setup and hold times $T_{s,max} - T_{S,min}$ and $T_{H,max} - T_{H,min}$, respectively, can be exploited during the design process to reduce delay uncertainty,

improving the tolerance of a circuit to variations.

For a critical path, an increase in the delay of a data path ΔT_D due to a variation causes the frequency to be decreased to satisfy (10.2). This increase in the data path delay produces additional slack in (10.1), *i.e.*, hold skew. This additional slack in the hold skew can be exploited to increase the hold time in (10.1) by ΔT_{Hold} where $\Delta T_{Hold} = \Delta T_D$. An increase in the hold time enables a decrease in the setup time by $\Delta T_S = f^{-1}(\Delta T_H)$ due to the interdependence, as illustrated in Fig 10.1. The effect of the variation, *i.e.*, the decrease in frequency, can therefore be compensated by exploiting a lower setup time.

Similarly, for a timing path sensitive to a race condition, referred to as a *short path*, a decrease in the delay of the data path by ΔT_D can cause a hold time violation. Since the delay of the data path is reduced, the additional slack in (10.2), *i.e.*, setup skew, can be exploited by increasing the setup time by ΔT_S where $\Delta T_S = \Delta T_D$. An increase in the setup time supports a decrease in the hold time by $\Delta T_H = f(\Delta T_S)$, potentially resolving the violation. The delay uncertainty due to a variation is therefore reduced by exploiting interdependent setup-hold times. This procedure is summarized in the flowchart shown in Fig. 10.3.

Note that the variation in the delay of the clock launch path ΔT_{Ci} and clock capture path ΔT_{Cf} are assumed to be equal in this approach. Two sequentially-adjacent registers of a critical path are often placed physically close to each other, where

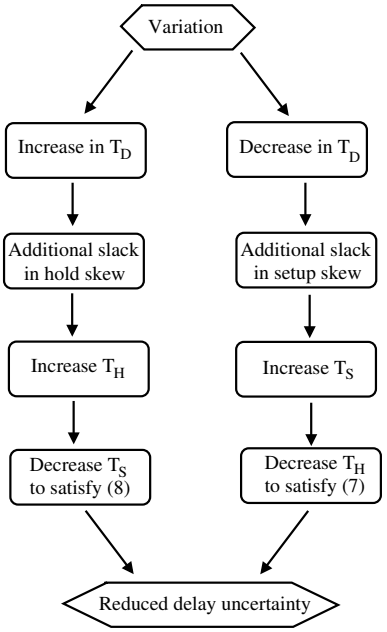


Figure 10.3: Flow diagram to reduce delay uncertainty by exploiting interdependent setup-hold times.

the difference between the clock launch and capture paths is typically minor, *i.e.*, a balanced clock tree, thereby validating this assumption. For those cases where this assumption is not accurate, *i.e.*, $\Delta T_{Ci} \neq \Delta T_{Cf}$ (non-zero clock skew), the variation in the delay of the clock path may either enhance or degrade the delay uncertainty depending upon the sign of $\Delta T_{Cf} - \Delta T_{Ci}$, as described in Section 10.2.4.

Another important consideration for this procedure is the effect of variations on the interdependent setup-hold characteristics. This variation in the critical setup-hold pairs is sufficiently low as compared to the valid range of setup times and hold times, enabling the proposed procedure. This behavior is further described in the following subsection.

10.2.3 Setup-Hold Time Characterization Under Variations

Process and environmental variations can also affect the critical pairs of a register, *i.e.*, $T_{S,min}$, $T_{H,max}$, $T_{S,max}$, and $T_{H,min}$. The effect of a variation on these critical pairs, however, is sufficiently small as compared to the valid range of setup times $T_{S,r}$ and hold times $T_{H,r}$. This behavior is primarily due to a stronger dependence of the CLK-to-Q delay on the setup skew and hold skew when these skews are reduced, as described in Chapter 9. For example, at a critical pair ($T_{S,min}$, $T_{H,max}$), the CLK-to-Q delay is primarily determined by $T_{S,min}$, lowering the effect of variations on this pair. Similarly, at a critical pair ($T_{S,max}$, $T_{H,min}$), $T_{H,min}$ has a relatively stronger effect on the CLK-to-Q delay.

To further illustrate this behavior, critical pairs are obtained for different power supply voltages in a 90 nm CMOS technology where the power noise is the source of the variation, as depicted in Fig. 10.4. The maximum variation in the power supply is assumed to be 10% of the nominal voltage. The variation in the critical pairs is compared with the valid range of setup and hold times. Specifically, the variation in the critical pairs caused by a 10% increase or decrease in the power supply voltage is listed, respectively, in Tables 10.1 and 10.2. As listed in these tables, $T_{S,r}$ and $T_{H,r}$ are sufficiently higher than the variation of the critical pairs, making interdependence an effective mechanism to reduce delay uncertainty.

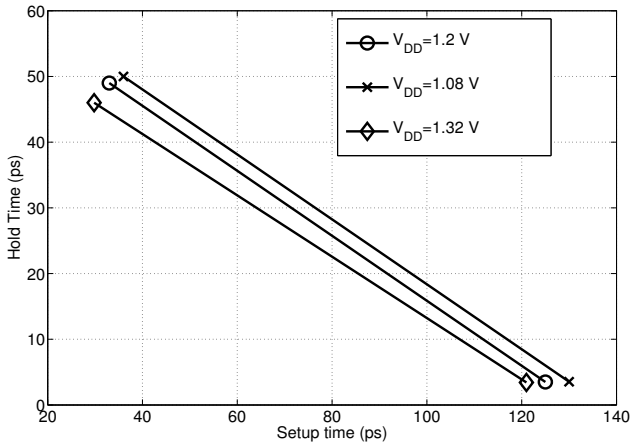


Figure 10.4: Variation of critical pairs as a function of the power supply voltage for a 90 nm CMOS technology.

Table 10.1: Variation of the critical points due to a 10% decrease in the power supply voltage.

Critical points	V_{DD}=1.2 V	V_{DD}=1.08 V	Variation (ps)	Variation (%)
$T_{S,min}$ (ps)	33	36	3	9.1
$T_{H,max}$ (ps)	49	50	1	2
$T_{S,max}$ (ps)	125	130	5	4
$T_{H,min}$ (ps)	3.5	3.54	0.04	1.1
$T_{S,r}$ (ps)	92	94	2	2.2
$T_{H,r}$ (ps)	45.5	46.5	1	2.2

Table 10.2: Variation of critical points due to a 10% increase in the power supply voltage.

Critical points	V_{DD}=1.2 V	V_{DD}=1.32 V	Variation (ps)	Variation (%)
$T_{S,min}$ (ps)	33	29.8	3.2	9.7
$T_{H,max}$ (ps)	49	46	3	6.1
$T_{S,max}$ (ps)	125	121	4	3.2
$T_{H,min}$ (ps)	3.5	3.43	0.07	2
$T_{S,r}$ (ps)	92	91.2	0.8	0.9
$T_{H,r}$ (ps)	45.5	42.57	2.93	6.4

10.2.4 Amount of Compensation

The compensation in delay variation (or the reduction in delay uncertainty) is dependent upon three primary factors: (1) the range of the valid setup times $T_{S,r}$ and hold times $T_{H,r}$, (2) the (setup, hold) pair used to determine the data path delay, and (3) the effect of the variations on the clock launch and capture paths.

If a register has a greater range of valid setup times and hold times, this register is more effective in reducing delay uncertainty. Note however that this type of register may exhibit other tradeoffs such as higher power consumption and CLK-to-Q delay.

As described in Section 10.1, the specific (setup, hold) pair used to determine the data path delay can lower the overall power consumption while satisfying the target frequency and achieving a higher tolerance to variations. The middle point of the setup-hold line ($T_{S,mid}$, $T_{H,mid}$),

$$T_{S,mid} = \frac{T_{S,min} + T_{S,max}}{2}, \tag{10.7}$$

$$T_{H,mid} = \frac{T_{H,max} + T_{H,min}}{2}, \tag{10.8}$$

results in a highest tolerance since the setup and hold times exhibit the maximum flexibility to variations, as illustrated in Fig. 10.5.

Choosing all of the data paths to satisfy the pair ($T_{S,mid}$, $T_{H,mid}$) may cause, however, overdesign of some specific paths. For example, for a critical path, $T_{S,mid} -$

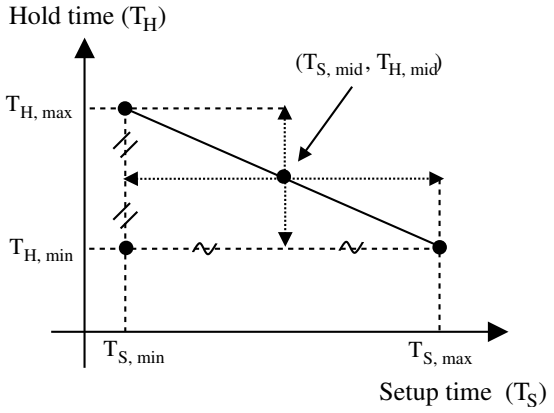


Figure 10.5: A data path designed at the pair $(T_{S,mid}, T_{H,mid})$ achieves the highest tolerance to variations.

$T_{S,min}$ may be greater than the worst case variation in the data path delay. In this case, a smaller setup time than $T_{S,mid}$ can be chosen for the data path. Similarly, for a short path, $T_{H,mid} - T_{H,min}$ may be greater than the worst case variation, suggesting a hold time smaller than $T_{H,mid}$ is possible.

An algorithm is proposed to prevent this overdesign where the selection of the (setup, hold) pair depends upon the type of data path. This technique prevents an overly conservative design while achieving a higher tolerance to variations. Pseudo-code of the proposed algorithm CHOOSE-PAIR is shown in Fig. 10.6.

The critical points $T_{S,min}$, $T_{S,max}$, $T_{H,min}$, and $T_{H,max}$ are obtained in the first line, as described in Section 10.2.1. The worst case variation in the delay of the data path is determined in the second line. In lines 3-10, a (setup, hold) pair is determined for a worst case path where the upper bound of the data path delay, *i.e.*, (10.2), is critical. The minimum valid setup time $T_{S,min}$ is increased by the worst case variation

CHOOSE-PAIR
1. Obtain $T_{S,min}$, $T_{S,max}$, $T_{H,min}$, and $T_{H,max}$ as in Section 10.2.1
2. Obtain worst case variations in data path delay $T_{D,max}$ and $T_{D,min}$
3. **if** data path is a critical path **then**
4. $T_S = T_{S,min} + (T_{D,max} - T_{D,nom})$
5. **if** $T_S \leq T_{S,max}$ **then**
6. find $T_H = f(T_S)$ using (10.3)
7. **else**
8. $T_S = T_{S,max}$
9. $T_H = f(T_{S,max}) = T_{H,min}$
10. **else if** data path is sensitive to a race condition (short path) **then**
11. $T_H = T_{H,min} + (T_{D,nom} - T_{D,min})$
12. **if** $T_H \leq T_{H,max}$ **then**
13. find $T_S = f^{-1}(T_H)$ using (10.4)
14. **else**
15. $T_H = T_{H,max}$
16. $T_S = f^{-1}(T_{H,max}) = T_{S,min}$
17. **else**
18. $T_S = (T_{S,min} + T_{S,max})/2$
19. $T_H = f(T_S) = (T_{H,min} + T_{H,max})/2$
20. **end**

Figure 10.6: Pseudo-code to determine the appropriate (setup, hold) pair that prevents the overdesign of a circuit while achieving a higher tolerance to variations.

$T_{D,max} - T_{D,nom}$ in the delay of the data path. If the resulting setup time T_S is valid, *i.e.*, $T_S \leq T_{S,max}$, T_S is chosen as the setup time. The corresponding hold time is determined by (10.3). If however T_S is outside the valid range, $T_{S,r}$ is insufficient to tolerate the variation. In this case, $T_{S,max}$ is chosen as the setup time, providing the maximum tolerance for the critical path.

Alternatively, if the data path is sensitive to a race condition (short path) where the lower bound of the data path delay, *i.e.*, (10.2), is critical, T_H is initially determined in line 11 as the summation of the minimum hold time $T_{H,min}$ and worst case variation $T_{D,nom} - T_{D,min}$. If this hold time is valid, *i.e.*, $T_H \leq T_{H,max}$, T_H is chosen as the hold time. The corresponding setup time is determined by (10.4). If the hold time is not within this valid range, $T_{H,max}$ is chosen as the hold time since $T_{H,max}$ provides the maximum tolerance for a short path.

For all of the remaining data paths where either (10.1) or (10.2) is not critical, $T_{S,mid}$ and $T_{H,mid}$ are chosen, respectively, as the setup and hold times. At this point, the setup and hold times exhibit the greatest tolerance to delay variations.

The amount of variation tolerated by this methodology is also dependent upon the variation in the delay of the clock launch path ΔT_{Ci} and clock capture path ΔT_{Cf}. Specifically, if $\Delta T_{Ci} = \Delta T_{Cf}$, these variations compensate each other, maintaining the validity of the proposed algorithm. In this case, the variation in the delay of the clock paths does not affect the amount of tolerance. If however $\Delta T_{Ci} > \Delta T_{Cf}$, less

variation can be tolerated by a critical path since the delay of the clock launch path is increased, delaying the data signal from leaving the register. For a short path, however, additional variation can be tolerated.

Alternatively, if $\Delta T_{Cf} > \Delta T_{Ci}$, additional variation can be tolerated for a critical path since the clock launch path is relatively faster than the clock capture path. For a short path, however, less variation can be tolerated. The effect of the variation on the clock path delays can be considered in the proposed algorithm by replacing line 4 with $T_S = T_{S,min} + (T_{D,max} - T_{D,nom}) + (T_{Ci} - T_{Cf})$ and line 11 with $T_H = T_{H,min} + (T_{D,nom} - T_{D,min}) + (T_{Cf} - T_{Ci})$.

10.2.5 Increase in Permissible Range

A *permissible range* has been defined determining the valid range of clock skew $T_{Ci} - T_{Cf}$ [217]. Rearranging (10.1) and (10.2), the clock skew should satisfy

$$T_H - T_D \leq T_{Ci} - T_{Cf}, \tag{10.9}$$

$$T_{Ci} - T_{Cf} \leq T_{CP} - T_D - T_S, \tag{10.10}$$

where the lower and upper bounds of the clock skew are determined, respectively, by (10.9) and (10.10). A methodology to tolerate skew variations by exploiting this permissible range has also been described [212].

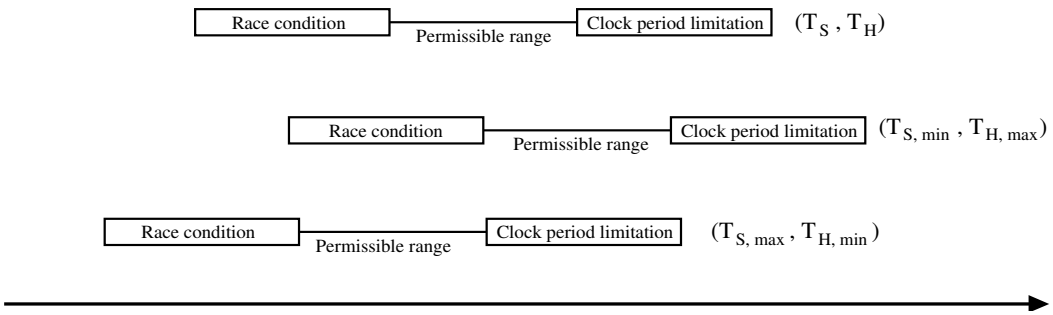


Figure 10.7: Exploiting interdependent setup and hold times within a permissible range of the clock skew. A shift in the permissible range is achieved. This shift is dependent upon the specific (setup, hold) pair, providing additional flexibility to tolerate clock skew.

The methodology proposed in this chapter further increases the effective permissible range by exploiting the interdependent setup and hold times. Specifically, the lower and upper bounds of the clock skew can change depending upon the specific (setup, hold) pair, as illustrated in Fig. 10.7. For example, if the pair $(T_{S,min}, T_{H,max})$ is used, the permissible range shifts to the right since both bounds increase. Alternatively, if the pair $(T_{S,max}, T_{H,min})$ is used, the permissible range shifts to the left since both bounds decrease, as shown in Fig. 10.7. The interdependence of the setup and hold times therefore provides additional flexibility in exploiting the permissible range. Note that the variation in the data path delay rather than the clock skew is the primary focus of this chapter.

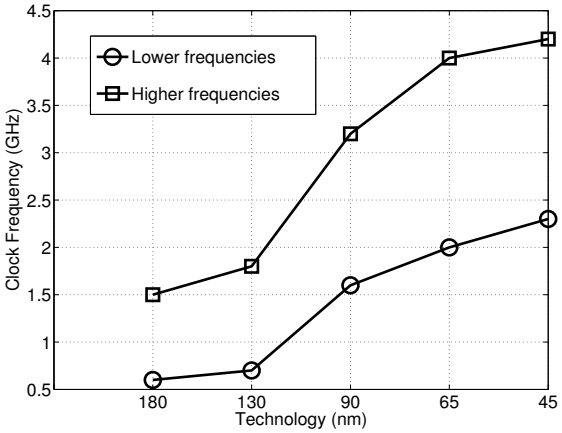


Figure 10.8: Target clock frequency for each technology node. Higher frequencies represent the upper bound of the frequency while the lower frequencies represent the lower bound of the frequency.

10.3 Case Study

The efficacy of setup-hold time interdependence to compensate power supply variations is evaluated in this section. Four CMOS technology generations are considered: 180 nm, 90 nm, 65 nm, and 45 nm. An industrial model is used for the 180 nm, 90 nm, and 65 nm CMOS technologies. For the 45 nm CMOS technology, a predictive model is used [218], [219]. Two clock frequencies are considered for each technology based on the data published in [220], as illustrated in Fig. 10.8. Higher frequencies represent the upper bound on the frequency while the lower frequencies represent the lower bound on the frequency.

The interdependent setup-hold time characteristics for each technology is described in Section 10.3.1. The dependence of these characteristics on technology

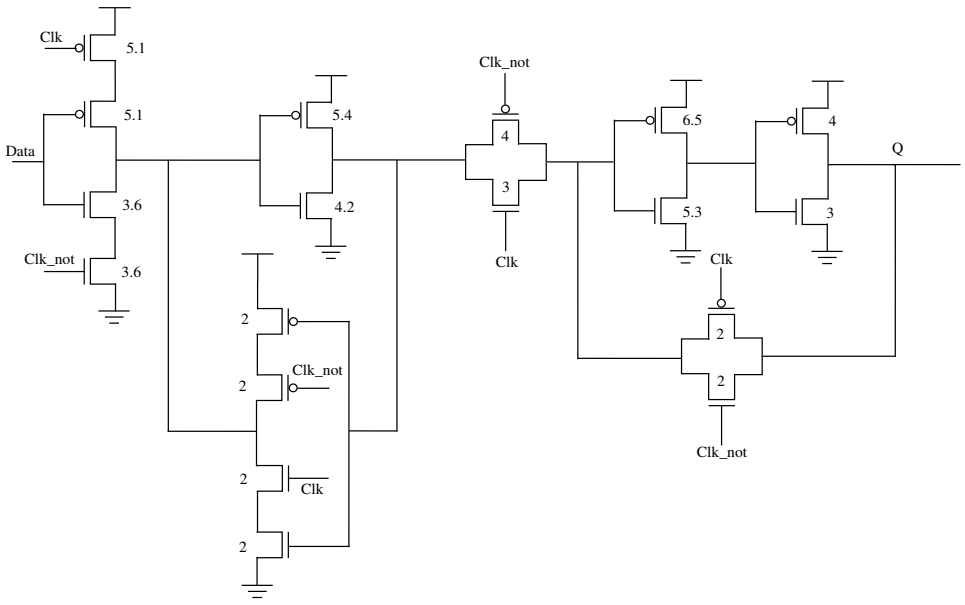


Figure 10.9: Master-slave rising edge triggered register to generate T_S vs. T_H relationship for each technology. The numbers represent the W/L ratio for each transistor.

is also discussed. The variation in the delay caused by the power noise is quantified as a function of technology in Section 10.3.2. Finally, the efficacy of setup-hold time interdependence in tolerating this delay variation is evaluated in Section 10.3.3.

10.3.1 Interdependent T_S vs T_H Relationship

A master-slave type register is used to illustrate the T_S vs. T_H relationship for each technology node. A schematic of the register is illustrated in Fig. 10.9, where the width W of the transistors is scaled to maintain a constant aspect ratio W/L for each technology. The register has been simulated to obtain the critical setup-hold pairs, as described in Section 10.2.1, where the signal transition times are assumed

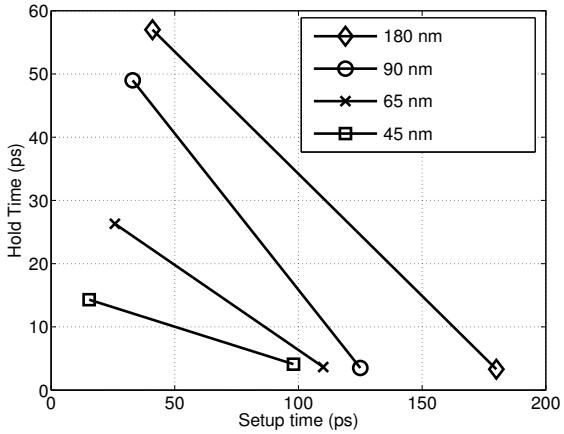


Figure 10.10: Interdependent setup-hold time characteristics for four technologies.

to be 10% of the clock period. The T_S vs. T_H relationship for each technology is illustrated in Fig. 10.10, where each line is

$$\text{180 nm: } T_H = -0.386T_S + 72.839 \text{ for 41 ps} \leq T_S \leq 180 \text{ ps}, \tag{10.11}$$

$$\text{90 nm: } T_H = -0.494T_S + 65.32 \text{ for 33 ps} \leq T_S \leq 125 \text{ ps}, \tag{10.12}$$

$$\text{65 nm: } T_H = -0.269T_S + 33.255 \text{ for 25.8 ps} \leq T_S \leq 110 \text{ ps}, \tag{10.13}$$

$$\text{45 nm: } T_H = -0.123T_S + 16.202 \text{ for 15.4 ps} \leq T_S \leq 98 \text{ ps}. \tag{10.14}$$

The range of valid setup times $T_{S,r} = T_{S,max} - T_{S,min}$ and range of valid hold times $T_{H,r} = T_{H,max} - T_{H,min}$ scale with technology, as shown in Fig. 10.10. These critical points, CLK-to-Q delay of the register, and power supply voltage are listed in Table 10.3 for each CMOS technology.

Note the behavior of $T_{S,r} = T_{S,max} - T_{S,min}$ (range of valid setup times) as a

Table 10.3: Power supply voltage, clock-to-Q delay, and critical points $T_{S,min}$, $T_{H,max}$, $T_{S,max}$, $T_{H,min}$, $T_{S,r}$, and $T_{H,r}$ for each technology.

CMOS Technology	V_{DD} (V)	Clock-to-Q delay (ps)	$T_{S,min}$ (ps)	$T_{H,max}$ (ps)	$T_{S,max}$ (ps)	$T_{H,min}$ (ps)	$T_{S,r}$ (ps)	$T_{H,r}$ (ps)
180 nm	1.8	172	41	57	180	3.3	139	53.7
90 nm	1.2	86.4	33	49	125	3.5	92	45.5
65 nm	1.1	57	25.8	26.3	110	3.6	84.2	22.7
45 nm	1.0	29.8	15.4	14.3	98	4.1	82.6	10.2

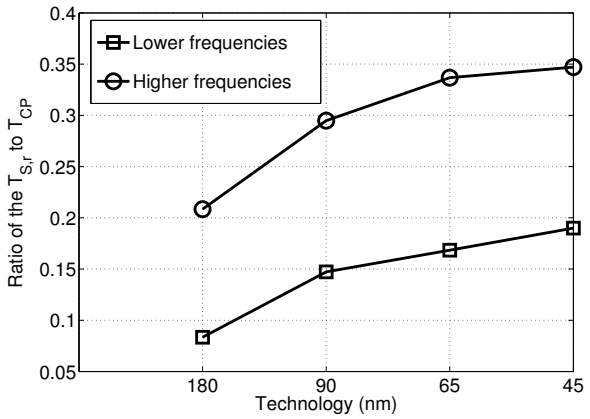


Figure 10.11: Ratio of the range of valid setup times $T_{S,r}$ to clock period T_{CP}.

function of technology. The ratio of the range of valid setup times to the clock period ($T_{S,r}/T_{CP}$) increases as the technology advances, as illustrated in Fig. 10.11. Specifically, for the 45 nm CMOS technology, the range of valid setup times is approximately 20% of the clock period at the lower frequencies. For the higher frequencies, this ratio increases to 35%. The interdependence of the setup-hold times is therefore more able to tolerate variations in deep submicrometer technologies, where the difference between the maximum and minimum setup time is a significant fraction of the clock period.

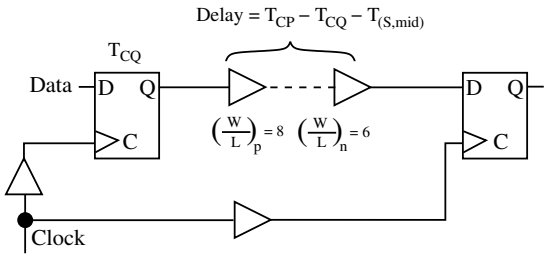


Figure 10.12: Critical path to determine the increase in the delay of the data path due to a decrease in V_{DD}.

10.3.2 Delay Variation due to Power Noise

The effect of power supply variations, *i.e.*, power noise, on delay is evaluated in this section. These variations are compared with the valid range of setup times $T_{S,r}$ and hold times $T_{H,r}$, thereby determining the ability of exploiting this interdependence to reduce delay uncertainty. The clock period corresponding to each technology is determined from Fig. 10.8. A critical path is designed for each technology where the register shown in Fig. 10.9 is synchronized. This critical path is designed to evaluate the efficacy of exploiting the interdependence relationship in compensating a drop in the power supply voltage since a decrease in V_{DD} increases the delay of the data path. Identical inverters are used in the combinational circuit. A specific number of inverters is inserted between the initial and final register until the delay of the data path satisfies (10.2), as illustrated in Fig. 10.12. The aspect ratio W/L of the PMOS and NMOS transistors are, respectively, eight and six.

A timing path sensitive to a race condition (a short path) can also be generated by abutting the registers, as illustrated in Fig. 10.13. This timing path is designed to

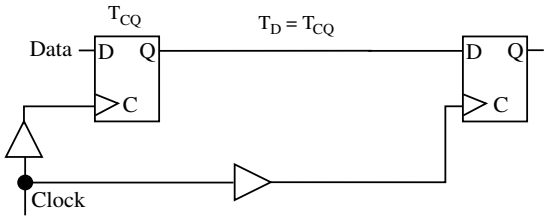


Figure 10.13: A short path to determine the decrease in the delay of the data path due to an increase in V_{DD}.

Table 10.4: Increase in the delay of the critical data path (shown in Fig. 10.12) caused by a 10% drop in the power supply voltage.

Technology (nm)	Frequency (GHz)	V_{DD1} (V)	T_{D1} (ps)	V_{DD2} (V)	T_{D2} (ps)	ΔT_D (ps)
180	1.5	1.8	540	1.62	597.3	57.3
	0.6	1.8	1555.4	1.62	1707.9	152.5
90	3.2	1.2	231.3	1.08	255	23.7
	1.6	1.2	536.2	1.08	586.9	50.7
65	4	1.1	180.5	0.99	201.7	21.2
	2	1.1	428.4	0.99	475.4	47
45	4.2	1	177.3	0.9	201.6	24.3
	2.3	1	373.8	0.9	425.4	51.6

evaluate the efficacy of exploiting the interdependence relationship in compensating an increase in the power supply voltage since an increase in V_{DD} reduces the delay of the data path.

The synchronous circuits shown in Figs. 10.12 and 10.13 are simulated with SPICE, where the power supply voltage is varied by 10%. The increase in the delay of the critical data path caused by this voltage drop is listed in Table 10.4 for both lower and higher frequencies. Similarly, the decrease in the delay of the short path is listed in Table 10.5, where the power supply voltage is increased by 10%. These variations are compared with the valid range of setup times $T_{S,r}$ and hold times $T_{H,r}$,

Table 10.5: Decrease in the delay of the data path shown in Fig. 10.13 due to a 10% increase in the power supply voltage.

Technology (nm)	V_{DD1} (V)	T_{D1} (ps)	V_{DD2} (V)	T_{D2} (ps)	ΔT_D (ps)
180	1.8	168.8	1.98	151.8	17
90	1.2	84.3	1.32	76.2	8.1
65	1.1	55.7	1.21	49.8	5.9
45	1	29.2	1.1	26.4	2.8

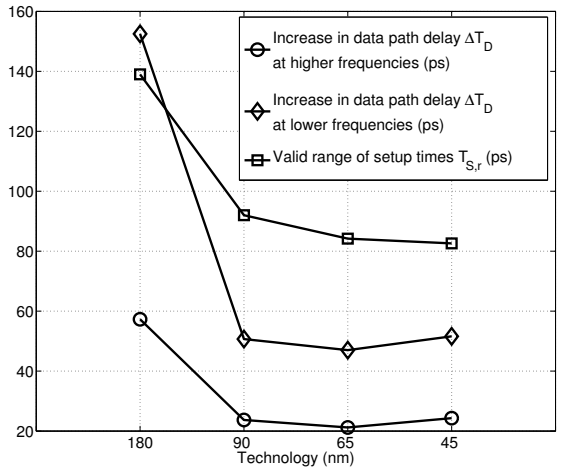


Figure 10.14: Comparison of the increase in the delay of a critical data path with $T_{S,r}$ to evaluate the efficacy of exploiting the interdependence relationship.

respectively, in Figs. 10.14 and 10.15 to evaluate the efficacy of exploiting the interdependence relationship, as described in the following section.

10.3.3 Compensation of Delay Variation

As illustrated in Fig. 10.14, the interdependence relationship can be used to compensate for delay variations since the range of valid setup times is higher than the increase in the data path delay except for the 180 nm CMOS technology operating

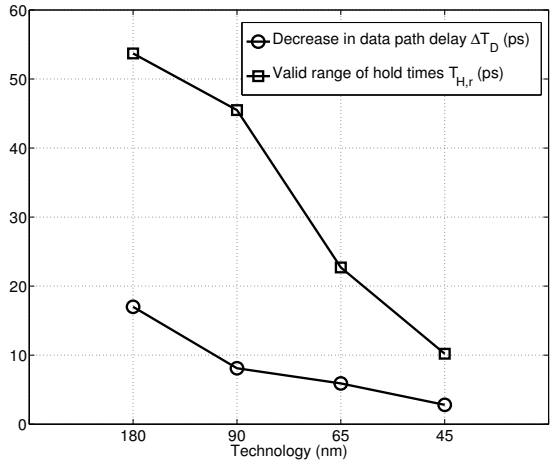


Figure 10.15: Comparison of the decrease in the delay of a short path with $T_{H,r}$ to evaluate the efficacy of exploiting the interdependence relationship.

at 600 MHz. At this frequency, the delay of the data path is relatively large, causing a higher absolute variation in this delay.

Note that the difference between the valid range of setup times and variation in delay is larger for the higher frequencies since the delay of the data path is lower for these frequencies. Exploiting the interdependence relationship is therefore more effective in reducing the delay uncertainty of a critical path operating at higher frequencies. Also note that the absolute variation in delay due to power supply noise somewhat saturates beyond the 130 nm technology node. This behavior is primarily due to the use of multicore processors where the increase in clock frequency is relatively low, as illustrated in Fig. 10.8.

For a short path, the range of valid hold times is larger than the decrease in data path delay for each technology, as illustrated in Fig. 10.15. The difference between

these two values, however, decreases for more deeply scaled technologies. For a short path, therefore, the interdependence relationship is more effective in reducing delay uncertainty in older technologies. This behavior is due to the significant decrease in the valid range of hold times with scaled technologies.

The procedure described in Section 10.2.2 is performed assuming the data paths are designed at the middle point ($T_{S,mid},\ T_{H,mid}$) of the interdependent setup-hold line. For example, for the 90 nm CMOS technology operating at 3.2 GHz, the delay of the worst case data path increases by 23.7 ps due to a drop in power supply voltage, as listed in Table 10.4. The hold time of 26.3 ps is increased by 23.7 ps to 50 ps. Since 50 ps is larger than the maximum hold time at this technology, the hold time is increased to 49 ps. An increase in the hold time enables a decrease in the setup time from 79 ps to 33 ps, tolerating 46 ps of delay uncertainty. Note that this delay uncertainty is larger than the initial variation of 23.7 ps, achieving about 100% compensation in the critical path.

Similarly, for a short path, the decrease in the delay of the data path is 8.1 ps, as listed in Table 10.5. The setup time can therefore be increased from 79 ps to 87.1 ps. The corresponding hold time is therefore reduced from 26.3 ps to 22.3 ps, as determined by (10.12), tolerating 4 ps of delay uncertainty. Since the variation in the delay of the data path is 8.1 ps, interdependence can compensate approximately 50% of the delay uncertainty of a short path. The results of this procedure for other

Table 10.6: Compensation of delay uncertainty caused by power noise for a critical data path.

Technology (nm)	Critical data path				
	(T_{S1}, T_{H1}) (ps)	Frequency (GHz)	ΔT_D (ps)	(T_{S2}, T_{H2}) (ps)	Compensation (%)
180 nm	(110.5, 30.2)	1.5	57.3	(41,57)	100
		0.6	152.5	(41, 57)	45.6
90 nm	(79, 26.3)	3.2	23.7	(33,49)	100
		1.6	50.7	(33, 49)	90.7
65 nm	(67.9, 14.9)	4	21.2	(25.8, 26.3)	100
		2	47	(25.8, 26.3)	89.5
45 nm	(56.7, 9.2)	4.2	24.3	(15.4, 14.3)	100
		2.3	51.6	(15.4, 14.3)	80.1

Table 10.7: Compensation of delay uncertainty caused by power noise for a short path.

Technology (nm)	Short path			
	(T_{S1}, T_{H1}) (ps)	ΔT_D (ps)	(T_{S2}, T_{H2}) (ps)	Compensation (%)
180 nm	(110.5, 30.2)	17	(127.5, 7.1)	41.8
90 nm	(79, 26.3)	8.1	(87.1, 22.3)	50
65 nm	(67.9, 14.9)	5.9	(73.8, 13.4)	26.2
45 nm	(56.7, 9.2)	2.8	(59.5, 8.9)	10.7

technology nodes are listed in Tables 10.6 and 10.7 for, respectively, a worst case data path and a short path.

As listed in Table 10.6, delay uncertainty caused by power supply noise in a critical data path can be compensated by up to 100% for the higher frequencies. For the lower frequencies, more than 80% compensation is achieved in the more deeply scaled technologies. Alternatively, as listed in Table 10.7, for a short path, the compensation is lower due to the relatively smaller slope of the function $T_H = f(T_S)$ as compared to $T_S = f^{-1}(T_H)$, as illustrated in Fig. 10.10. This reduction in delay uncertainty is achieved by designing the data path at the middle setup-hold point ($T_{S,mid}$, $T_{H,mid}$),

exploiting the interdependence between the setup and hold times.

10.4 Summary

A methodology is proposed to improve the tolerance of a circuit to process and environmental variations. The interdependence between the timing constraints is exploited to achieve a more robust circuit, significantly reducing the delay uncertainty caused by variations. An algorithm is proposed to appropriately determine the (setup, hold) pair during the design process to improve the tolerance of a circuit to variations while achieving a target frequency. A technique is also proposed to efficiently obtain the critical (setup, hold) pairs of a register.

A case study is described to evaluate the efficacy of the proposed methodology in reducing the delay uncertainty due to power supply noise. Four CMOS technologies (180 nm, 90 nm, 65 nm, and 45 nm) are evaluated to compare the compensation achieved by the proposed technique. The significance of the setup-hold interdependence relationship in reducing delay uncertainty, thereby improving the tolerance of a circuit to delay variations, is validated by these results.

Chapter 11

Conclusions

Noise has long been associated with classical analog/RF circuits and, to a lesser degree, certain types of digital circuits such as dynamic CMOS logic. Analog/RF circuits typically consist of highly sensitive blocks with critical performance characteristics such as signal-to-noise ratio, bandwidth, gain, gain-bandwidth product, dynamic range, total harmonic distortion, and jitter. Device noise such as thermal, shot, and flicker noise has traditionally been the primary concern for analog/RF circuits. Dynamic CMOS has also received considerable attention due to reduced noise margins, increasing the probability of a logical failure. Other CMOS logic families have typically been assumed relatively immune to noise due to higher noise margins.

The classical understanding of noise in integrated circuits has significantly changed over the last decade for several reasons. Improvements in manufacturing technologies have enabled the integration of more than two billion transistors on the same monolithic die, while improving overall performance and reducing cost.

This dense integration has started the era of system-on-chip (SoC) and system-in-package (SiP) where circuits with different functionalities and even technologies are integrated within the same die or package. The significant advantages demonstrated by these architectures are possible at the expense of greater noise coupling among various blocks where noise analysis and isolation techniques can be highly expensive. Specifically, the noise caused by the switching of a digital signal, *i.e.*, switching noise, has emerged as a primary design objective not only in mixed-signal ICs, but also high performance synchronous digital ICs.

Switching noise both in mixed-signal and synchronous digital circuits is the primary subject of this dissertation. Voltage fluctuations on the power/ground nodes of a circuit, *i.e.*, power/ground noise, is a type of switching noise affecting both mixed-signal and synchronous digital ICs. A methodology has been proposed to accurately estimate the worst case power/ground noise in an inductive power/ground distribution network with a decoupling capacitor. The non-monotonic behavior of noise as a function of transition time has been elucidated. The efficacy of several noise reduction techniques such as placing a decoupling capacitance and reducing parasitic inductance has also been evaluated.

In a mixed-signal circuit, the monolithic substrate forms a conducting medium between the aggressive digital and sensitive analog blocks, degrading signal integrity. Two primary concerns related to substrate coupling noise exist. The first issue is the

computationally efficient analysis of the substrate noise for a large scale circuit such as a transceiver. The analysis process determines the noise characteristics of a circuit, thereby identifying appropriate noise isolation techniques. An accurate estimate of the required level of isolation prevents overdesign of a circuit, significantly improving certain design criteria such as area and power dissipation. The efficacy of these isolation strategies can also be evaluated through efficient analysis techniques. The second issue is the development and application of effective noise isolation strategies to alleviate substrate coupling while considering other design parameters such as area, power consumption, and speed.

Several approaches have been presented to efficiently model and alleviate substrate noise coupling in mixed-signal circuits. The interaction between the ground and substrate networks has been investigated. The substrate is shown to significantly affect the ground noise by changing the current propagation path. An intuitive model has been proposed for investigating several substrate noise coupling mechanisms. The proposed model can be used at early stages of the design process to identify the dominant noise source.

Accurate and efficient estimation of substrate noise in a large scale circuit is a difficult process since the power/ground networks, substrate network, and the switching circuit should be simultaneously considered. A methodology has been proposed for efficiently estimating the noise while maintaining sufficient accuracy. The similarity

of the ground voltages was exploited to reduce the computational complexity of the substrate extraction process. The methodology has been verified with an industrial circuit, exhibiting significant reduction in analysis time while maintaining reasonable accuracy in the noise voltage. A methodology has also been proposed to reduce substrate noise in mixed-signal circuits. The noise was reduced by incorporating noise-aware standard cells where a localized guard ring around an aggressor circuit is utilized.

For synchronous digital circuits, the effect of switching noise on the timing characteristics is a primary concern. Specifically, additional delay uncertainty produced by the switching noise can cause a timing violation, degrading system performance or causing a functional failure. A methodology has been proposed to compensate for delay uncertainty, producing a robust circuit tolerant to power supply variations. Interdependent setup-hold times were exploited to improve the tolerance of a circuit to power/ground noise. The interdependence of the timing constraints has been demonstrated, significantly reducing pessimism in the static timing analysis process. Several algorithms have also been proposed to exploit the interdependence during both the design and analysis phases. The methodology was demonstrated on industrial circuits. The efficacy of the interdependence in reducing delay uncertainty was also evaluated for several technology nodes, verifying the proposed algorithms.

To conclude, several methodologies have been proposed in this dissertation to

analyze and reduce the effects of switching noise in mixed-signal and synchronous digital ICs. The substrate is a significant medium for noise transmission in mixed-signal ICs. For synchronous digital ICs, the effect of switching noise on the timing characteristics has been investigated. Industrial circuits were used to demonstrate the proposed methodologies and algorithms. These methodologies provide intuition and guidelines for designing complex integrated circuits with superior noise performance and enhanced signal integrity.

Chapter 12

Future Work

As described in previous chapters, noise management in complex systems-on-chip and systems-in-package is a challenging process. Sensitive analog/RF blocks suffer from the switching noise generated by the digital blocks and propagated across the common monolithic substrate. Furthermore, the critical paths of the high speed digital blocks suffer from delay uncertainty caused by power/ground noise.

These problems are exacerbated by the ever increasing density of on-chip devices and interconnects, higher clock frequencies, and faster signal transition times. Efficient methodologies are therefore required to enhance the signal integrity of a circuit, satisfying the switching noise constraints of each block. Computationally efficient methodologies are also needed to determine whether these noise reduction and isolation strategies are sufficient. Several research topics are proposed in this chapter to further investigate switching noise and develop an efficient approach to manage noise in complex ICs.

Switching noise such as power/ground noise depends upon not only the on-chip parasitic impedances, but also the package and board impedance characteristics. Typically, noise is independently analyzed at each level without considering the interactions among these levels, resulting in low accuracy in the estimated noise. A unified noise analysis methodology is proposed in Section 12.1 where the circuit, package, and board are simultaneously integrated to estimate power/ground noise.

The power grid is one of the most critical issues to enhance overall signal and power integrity. Traditionally, a target impedance is treated as the primary criterion when designing the power grid. More specific design criteria are proposed in Section 12.2 to develop a timing-aware power grid where the interactions among the power/ground network, clock network, and critical data signals are exploited.

Several noise isolation strategies exist to alleviate substrate coupling such as guard rings, triple well isolation, and use of a P-well block. These mechanisms can reduce substrate coupling at the expense of greater power/ground noise due to longer inductive current return paths. The effect of these mechanisms on power integrity has not received much attention. Research in this area is proposed in Section 12.3 to investigate the effects of substrate noise isolation methodologies on power integrity. Finally, the chapter is summarized in Section 12.4.

12.1 Unified Noise Analysis Methodology: Chip-Package-Board

Traditionally, the on-chip parasitic resistance is assumed to be higher than the package and board resistance since the thickness and width of the metal are smaller. Similarly, the parasitic inductance of the package and board is assumed to be higher than the on-chip inductance due to the geometry of the interconnect and return paths. The validity of these assumptions, however, should be reconsidered due to several reasons.

The length of the on-chip global interconnects has been increasing, transition times are significantly faster, and improvements in packaging technology have resulted in advanced packages with lower parasitic inductance and multiple layers. As a result of these developments, simple assumptions about the relative magnitude of the parasitic impedances within the integrated circuit, package, and board are no longer accurate. A unified noise analysis methodology is therefore required. A methodology that considers multiple levels within a system can improve the overall accuracy of the analysis process. Furthermore, the opportunities offered by advanced packaging structures can be better utilized and the corresponding tradeoffs more easily comprehended.

A primary objective of this analysis methodology is to efficiently integrate this process into an overall design flow, as illustrated in Fig. 12.1. Achieving this analysis

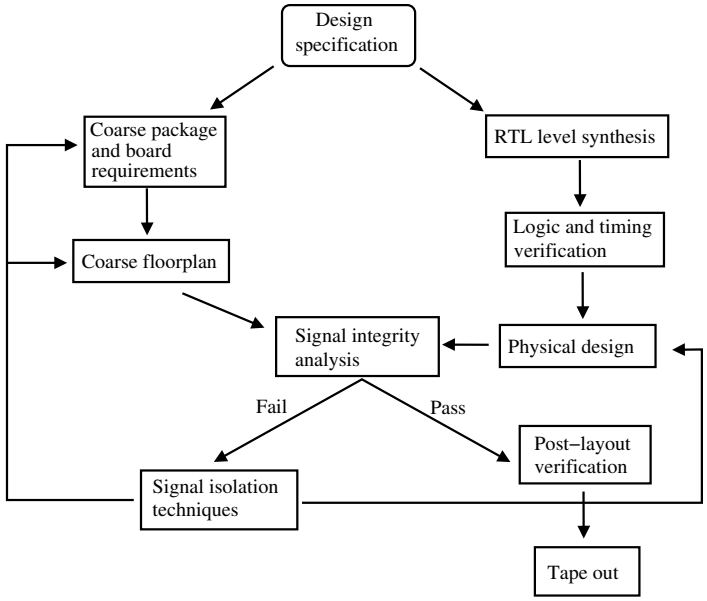


Figure 12.1: Integration of the signal integrity analysis process into a traditional IC design flow.

process in a computationally efficient manner is the primary challenge. Iterative simulation methodologies can be developed to reduce the computational complexity. For example, the integrated circuit and package can be initially analyzed individually. The power/ground noise obtained from this analysis can be used to further analyze the integrated circuit and board together, reducing the overall computational complexity. A unified noise analysis methodology would be useful in determining the optimum design space to achieve enhanced signal and power integrity while minimizing the resources devoted to these techniques.

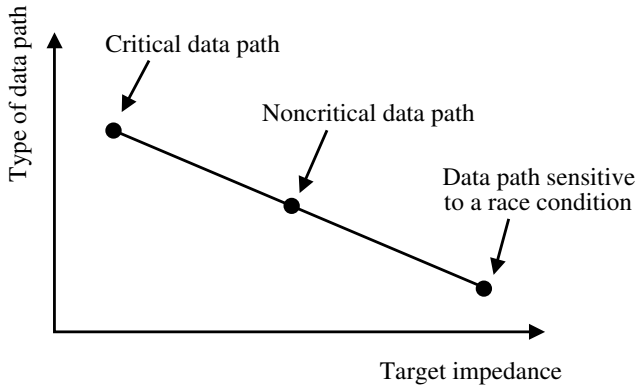


Figure 12.2: Multiple target impedances are defined dependent upon the type of data path to prevent an overly conservative power grid.

12.2 Timing Aware Power Grid Design

Existing methodologies to design a power grid do not consider the characteristics of the clock distribution network and critical data paths. The power/ground network within a high performance IC is typically composed of a homogeneous grid of several layers to satisfy a *single* target impedance while also satisfying specific electromigration constraints.

This homogeneous power grid can cause an overly conservative design which utilizes significant metal resources. Furthermore, power noise can significantly affect the delay of the critical paths. A heterogeneous power grid is proposed where multiple target impedances are defined rather than a single impedance, as illustrated in Fig. 12.2. Specifically, for those portions of the grid that supply power to the critical data paths, a smaller target impedance is specified, requiring greater metal resources. Alternatively, if the data path is not critical, additional noise can be tolerated which

utilizes less metal resources, thereby resulting in a higher target impedance. Also, if the data path is sensitive to a race condition, higher noise can be tolerated since a decrease in the power supply voltage increases the delay of the data path. For those cases where higher noise can be tolerated, the metal used by the power grid is lower, saving significant resources.

The interaction of the power grid with the clock network should also be considered. Specifically, power noise affects the clock delays, and therefore, the local clock skew. For those portions of the power grid that supply voltage to the clock path synchronizing the initial register of a critical data path, the target impedance should be low. Alternatively, the target impedance of the clock path synchronizing the final register of a critical path can be higher to exploit negative skew in the critical path. For a data path sensitive to a race condition (a short path), the target impedance should be high for the clock path driving the initial register and low for the clock path driving the final register to satisfy the hold time constraint. The simultaneous design of the power grid, clock network, and data signals should therefore be considered to improve the overall timing and signal integrity characteristics of a circuit.

12.3 Effect of Substrate Isolation on Power Integrity

Isolation of the noisy substrate from the sensitive analog/RF circuits is a primary design concern for mixed-signal circuits. Several classical techniques to achieve this isolation are guard rings around the aggressor and sensitive blocks, the use of a P-well layer to increase the substrate resistivity, and a triple well structure to achieve capacitive isolation.

These isolation strategies reduce substrate noise coupling at the expense of additional area, processing complexity, and I/O pads to bias these structures. Another important issue with these techniques is the effect of isolation on the inductive current return paths. This effect is particularly important for the high speed digital portions of a circuit where the parasitic loop inductance of the power grid can be substantial. For example, a large number of guard rings reduces the impedance of the substrate. A reduction in the impedance permits the substrate to behave as a return path for the inductive current. If the inductive current returns through the substrate, the current loop is significantly wider, increasing the loop inductance. Alternatively, a P-well layer increases the resistivity of the substrate, reducing the number of return current paths through the substrate. The loop inductance is therefore lower. An analysis of these effects can be used to evaluate substrate noise isolation techniques

in mixed-signal circuits where the inductive voltage drop within the digital blocks is of primary concern.

12.4 Summary

Several research topics have been described in this chapter that enhance the signal and power integrity of complex integrated circuits. The importance of a unified noise analysis methodology simultaneously considering the integrated circuit, package, and board characteristics has been discussed. Tradeoffs among these levels can be identified by means of a unified noise analysis process, supporting enhanced noise characterization. The simultaneous design of the clock, power, and data networks has also been discussed. Interactions among the power grid, clock distribution network, and the critical data paths can be effectively exploited to simultaneously satisfy timing, noise, and variability constraints.

Noise management and reduction will remain as a primary challenge for complex systems-on-chip and systems-in-package. Significant research is therefore required to enhance the physical codesign of power, clock, and data networks due to stringent timing and signal integrity constraints.

Bibliography

[1] W. F. Brinkman, D. E. Haggan, and W. W. Troutman, "A History of the Invention of the Transistor and Where It Will Lead Us," *IEEE Journal of Solid-State Circuits*, Vol. 32, No. 12, pp. 1858-1865, December 1997.

[2] J. E. Lilienfield, "Method and Apparatus for Controlling Electric Currents," U.S. Patent No. 1,745,175, October 1926.

[3] J. E. Lilienfield, "Amplifier for Electric Currents," U.S. Patent No. 1,877,140, December 1928.

[4] J. E. Lilienfield, "Device for Electric Current," U.S. Patent No. 1,900,018, March 1928.

[5] O. Heil, "Improvements in or Relating to Electrical Amplifiers and Other Control Arrangements and Devices," British Patent No. 439457, March 1935.

[6] R. G. Arns, "The Other Transistor: Early History of the Metal-Oxide-Semiconductor Field-Effect Transistor," *Engineering Science and Education Journal*, Vol. 7, No. 5, pp. 233-240, October 1998.

[7] J. Bardeen and W. H. Brattain, "The Transistor, a Semiconductor Triode," *Physical Review*, Vol. 74, No. 2, pp. 230–231, July 1948.

[8] W. Shockley, "The Theory of pn Junctions in Semiconductors and pn Junction Transistors," *Bell System Technical Journal*, Vol. 28, No. 3, pp. 435–489, July 1949.

[9] M. Atalla, E. Tannenbaum, and E. J. Scheinber, "Stabilization of Silicon Surfaces by Thermally Grown Oxides," *Bell System Technical Journal*, Vol. 38, No. 3, pp. 749–783, May 1959.

[10] J. A. Hoerni, "Planar Silicon Transistors and Diodes," *IEEE Transactions on Electron Devices*, Vol. 8, No. 2, pp. 655–657, March 1961.

[11] J. D. Meindl, "Retrospect and Prospect," *Proceedings of the IEEE International Conference on Very Large Scale Integration* (keynote speech), October 2007.

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

[13] B. Razavi, *Design of Analog CMOS Integrated Circuits*. McGraw-Hill, 2001.

[14] W. R. Bennett, *Electrical Noise*. McGraw-Hill, 1960.

[15] W. K. Chen, *The VLSI Handbook*. IEEE Press, 2000.

[16] D. Johns and K. Martin, *Analog Integrated Circuit Design*. Wiley, 1997.

[17] M. Shoji, *Theory of CMOS Digital Circuits and Circuit Failures*. Princeton University Press, 1992.

[18] J. T. Wallmark, "Noise Spikes in Digital VLSI Circuits," *IEEE Transactions on Electron Devices*, Vol. ED-29, No. 3, pp. 451–458, March 1982.

[19] K. L. Shepard and V. Narayanan, "Noise in Deep Submicron Digital Design," *Proceedings of the IEEE/ACM International Conference on Computer-Aided Design*, pp. 524–531, November 1996.

[20] P. Larsson and C. Svensson, "Noise in Digital Dynamic CMOS Circuits," *IEEE Journal of Solid-State Circuits*, Vol. 29, No. 6, pp. 655–662, June 1994.

[21] 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.

[22] P. Larsson, "Power Supply Noise in Future IC's: A Crystal Ball Reading," *Proceedings of the IEEE Custom Integrated Circuits Conference*, pp. 467–474, May 1999.

[23] P. Heydari and M. Pedram, "Ground Bounce in Digital VLSI Circuits," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, Vol. 11, No. 2, pp. 180–193, April 2003.

[24] I. Catt, "Crosstalk (Noise) in Digital Systems," *IEEE Transactions on Electronic Computers*, Vol. EC-16, No. 6, pp. 743–763, December 1967.

[25] L. Gal, "On-Chip Crosstalk - the New Signal Integrity Challenge," *Proceedings of the IEEE Custom Integrated Circuits Conference*, pp. 251–254, May 1995.

[26] A. Deutsch *et al.*, "On-Chip Wiring Design Challenges for Gigahertz Operation," *Proceedings of the IEEE*, Vol. 89, No. 4, pp. 529–555, April 2001.

[27] A. Deutsch *et al.*, "When are Transmission Line Effects Important for On-Chip Interconnections," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 45, No. 10, pp. 1836–1846, October 1997.

[28] 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.

[29] A. V. Mezhiba and E. G. Friedman, "Impedance Characteristics of Power Distribution Grids in Nanoscale Integrated Circuits," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, Vol. 12, No. 11, pp. 1148–1155, November 2004.

[30] 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.

[31] G. Konstadinidis *et al.*, "Implementation of a Third-Generation 16-Core 32-Thread Chip-Multithreading SPARC Processor," *Proceedings of the IEEE International Solid-State Circuits Conference*, pp. 84–85, February 2008.

[32] L. D. Smith, "Decoupling Capacitor Calculations for CMOS Circuits," *Proceedings of the IEEE Conference on Electrical Performance of Electronic Packaging*, pp. 101–105, November 1994.

[33] C. R. Paul, "Effectiveness of Multiple Decoupling Capacitors," *IEEE Transactions on Electromagnetic Compatibility*, Vol. 34, No. 2, pp. 130–133, May 1992.

[34] T. Roy, L. Smith, and J. Prymak, "ESR and ESL of Ceramic Capacitor Applied to Decoupling Applications," *Proceedings of the IEEE Conference on Electrical Performance of Electronic Packaging*, pp. 213–216, October 1998.

[35] M. Popovich, E. G. Friedman, R. Secareanu, and O. L. Hartin, "Efficient Placement of Distributed On-Chip Decoupling Capacitors in Nanoscale ICs," *Proceedings of the IEEE/ACM International Conference on Computer-Aided Design*, pp. 811–816, November 2007.

[36] S. Bothra, B. Rogers, M. Kellam, and C. M. Osburn, "Analysis of the Effects of Scaling on Interconnect Delay in ULSI Circuits," *IEEE Transactions on Electron Devices*, Vol. 40, No. 3, pp. 591–597, March 1993.

[37] D. Sylvester, C. Hu, O. S. Nakagawa, and S. Oh, "Interconnect Scaling: Signal Integrity and Performance in Future High-Speed CMOS Designs," *Proceedings of the IEEE Symposium on VLSI Technology*, pp. 42–43, June 1998.

[38] K. C. Saraswat and F. Mohammadi, "Effect of Scaling of Interconnections on the Time Delay of VLSI Circuits," *IEEE Transactions on Electron Devices*, Vol. ED-29, No. 4, pp. 645–650, April 1982.

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

[40] R. M. Dang and N. Shigyo, "Coupling Capacitances for Two-Dimensional Wires," *IEEE Electron Device Letters*, Vol. EDL-2, No. 8, pp. 196–197, August 1981.

[41] T. Sakurai, "Closed-Form Expressions for Interconnection Delay, Coupling, and Crosstalk in VLSI's," *IEEE Transactions on Electron Devices*, Vol. 40, No. 1, pp. 118–124, January 1993.

[42] K. T. Tang and E. G. Friedman, "Delay and Noise Estimation of CMOS Logic Gates Driving Coupled Resistive-Capacitive Interconnections," *Integration, the VLSI Journal*, Vol. 29, No. 2, pp. 131–165, September 2000.

[43] W. Chen, S. K. Gupta, and M. A. Breuer, "Analytic Models for Crosstalk Delay and Pulse Analysis Under Non-Ideal Inputs," *Proceedings of the IEEE International Test Conference*, pp. 809–818, November 1997.

[44] A. B. Kahng, S. Muddu, and D. Vidhani, "Noise and Delay Uncertainty Studies for Coupled *RC* Interconnects," *Proceedings of the IEEE International ASIC/SOC Conference*, pp. 3–8, September 1999.

[45] A. Devgan, "Efficient Coupled Noise Estimation for On-Chip Interconnects," *Proceedings of the IEEE/ACM International Conference on Computer-Aided Design*, pp. 147–153, November 1997.

[46] P. Heydari and M. Pedram, "Capacitive Coupling Noise in High-Speed VLSI Circuits," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 24, No. 3, pp. 478–488, March 2005.

[47] A. Vittal, L. H. Chen, M. Marek-Sadowska, K. Wang, and S. Yang, "Crosstalk in VLSI Interconnections," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 18, No. 12, pp. 1817–1824, December 1999.

[48] 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.

[49] S. Ramo, J. Whinnery, and T. Duzer, *Fields and Waves in Communication Electronics*. John Wiley and Sons Inc., 1984.

[50] E. Rosa, "The Self and Mutual Inductance of Linear Conductors," *Bulletin of the National Bureau of Standards*, Vol. 4, pp. 301–344, 1908.

[51] A. E. Ruehli, "Inductance Calculations in a Complex Integrated Circuit Environment," *IBM Journal of Research and Development*, Vol. 16, No. 5, pp. 470–481, September 1972.

[52] Y. Massoud *et al.*, "Managing On-Chip Inductive Effects," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, Vol. 10, No. 6, pp. 789–798, December 2002.

[53] K. T. Tang and E. G. Friedman, "Peak Noise Prediction in Loosely Coupled Interconnect," *Proceedings of the IEEE International Symposium on Circuits and Systems*, pp. 541–544, June 1999.

[54] S. Shin, Y. Eo, W. R. Eisenstadt, and J. Shim, "Analytical Models and Algorithms for the Efficient Signal Integrity Verification of Inductance-Effect-Prominent Multicoupled VLSI Circuit Interconnects," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, Vol. 12, No. 4, pp. 395–407, April 2004.

[55] K. Agarwal, D. Sylvester, and D. Blaauw, "Modeling and Analysis of Crosstalk Noise in Coupled *RLC* Interconnects," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 25, No. 5, pp. 892–901, May 2006.

[56] G. A. Katopis, "Delta-I Noise Specification for a High-Performance Computing Machine," *Proceedings of the IEEE*, Vol. 73, No. 9, pp. 1405–1415, September 1985.

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

[58] J. M. Zurada, Y. S. Joo, and S. V. Bell, "Dynamic Noise Margins of MOS Logic Gates," *Proceedings of the IEEE International Symposium on Circuits and Systems*, pp. 1153–1156, May 1989.

[59] K. L. Shepard and K. Chou, "Cell Characterization for Noise Stability," *Proceedings of the IEEE Custom Integrated Circuits Conference*, pp. 91–94, May 2000.

[60] V. Zolotov *et al.*, "Noise Propagation and Failure Criteria for VLSI Designs," *Proceedings of the IEEE/ACM International Conference on Computer-Aided Design*, pp. 587–594, November 2002.

[61] E. G. Friedman, "Latching Characteristics of a CMOS Bistable Register," *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, Vol. CAS-40, No. 12, pp. 902–908, December 1993.

[62] M. Favalli and L. Benini, "Analysis of Glitch Power Dissipation in CMOS ICs," *Proceedings of the IEEE International Symposium on Low Power Design*, pp. 123–128, April 1995.

[63] L. Benini *et al.*, "Glitch Power Minimization by Selective Gate Freezing," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, Vol. 8, No. 3, pp. 287–298, June 2000.

[64] K. Aingaran *et al.*, "Coupling Noise Analysis for VLSI and ULSI Circuits," *Proceedings of the IEEE International Symposium on Quality Electronic Design*, pp. 485–489, March 2000.

[65] M. H. Chowdhury *et al.*, "Performance Analysis of Deep Submicron VLSI Circuits in the Presence of Self and Mutual Inductance," *Proceedings of the IEEE International Symposium on Circuits and Systems*, pp. 197–200, May 2002.

[66] P. Chen, D. A. Kirkpatrick, and K. Keutzer, "Miller Factor for Gate-Level Coupling Delay Calculation," *Proceedings of the IEEE/ACM International Conference on Computer-Aided Design*, pp. 68–74, November 2000.

[67] G. Yee, R. Chandra, V. Ganesan, and C. Sechen, "Wire Delay in the Presence of Crosstalk," *Proceedings of the IEEE/ACM International Workshop on Timing Issues in the Specification and Synthesis of Digital Systems*, pp. 79–84, June 2000.

[68] A. B. Kahng, S. Muddu, and E. Sarto, "On Switch Factor Based Analysis of Coupled *RC* Interconnects," *Proceedings of the IEEE/ACM Design Automation Conference*, pp. 79–84, June 2000.

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

[70] T. Chen and A. Hajjar, "Statistical Timing Analysis of Coupled Interconnects Using Quadratic Delay-Change Characteristics," *IEEE Transactions on*

Computer-Aided Design of Integrated Circuits and Systems, Vol. 23, No. 12, pp. 1677–1683, December 2004.

[71] K. Agarwal, T. Sato, Y. Cao, D. Sylvester, and C. Hu, "Efficient Generation of Delay Change Curves for Noise-Aware Static Timing Analysis," *Proceedings of the IEEE International Symposium on Quality Electronic Design*, pp. 499–503, March 2003.

[72] T. Sakurai and A. R. Newton, "Alpha-Power Law MOSFET Model and its Applications to CMOS Inverter Delay and Other Formulas," *IEEE Journal of Solid-State Circuits*, Vol. 25, No. 2, pp. 584–594, April 1990.

[73] 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.

[74] D. Kouroussis, R. Ahmadi, and F. N. Najm, "Voltage-Aware Static Timing Analysis," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 25, No. 10, pp. 2156–2169, October 2006.

[75] S. Pant, D. Blaauw, V. Zolotov, S. Sundareswaran, and R. Panda, "Vectorless Analysis of Supply Noise Induced Delay Variation," *Proceedings of the IEEE/ACM International Conference on Computer-Aided Design*, pp. 184–191, November 2003.

[76] S. Zhao, K. Roy, and C. 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.

[77] H. Chen, J. Fang, and W. Shi, "Effective Decoupling Radius of Decoupling Capacitor," *Proceedings of the IEEE Symposium on Electrical Performance of Electronic Packaging*, pp. 277–280, October 2001.

[78] M. Popovich, E. G. Friedman, M. Sotman, A. Kolodny, and R. M. Secareanu, "Maximum Effective Distance of On-Chip Decoupling Capacitors in Power Distribution Grids," *Proceedings of the IEEE/ACM Great Lakes Symposium on VLSI*, pp. 173–179, April/May 2006.

[79] 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.

[80] H. Li *et al.*, "Partitioning-Based Approach to Fast On-Chip Decoupling Capacitor Budgeting and Minimization," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 25, No. 11, pp. 2402–2412, November 2006.

[81] S. W. Director and R. A. Rohrer, "The Generalized Adjoint Network and Network Sensitivities," *IEEE Transactions on Circuit Theory*, Vol. 16, No. 3, pp. 318–323, August 1969.

[82] P. Vuillod, L. Benini, A. Bogliolo, and G. De Micheli, "Clock-Skew Optimization for Peak Current Reduction," *Proceedings of the IEEE International Symposium on Low Power Electronic Design*, pp. 265–270, August 1996.

[83] A. Vittal, H. Fa, F. Brewer, and M. Marek-Sadowska, "Clock Skew Optimization for Ground Bounce Control," *Proceedings of the IEEE/ACM International Conference on Computer-Aided Design*, pp. 395–399, November 1996.

[84] M. Badaroglu *et al.*, "Digital Ground Bounce Reduction by Supply Current Shaping and Clock Frequency Modulation," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 24, No. 1, pp. 65–76, January 2005.

[85] K. Leung, "Controlled Slew Rate Output Buffer," *Proceedings of the IEEE Custom Integrated Circuits Conference*, pp. 5.5.1–5.5.4, May 1988.

[86] D. T. Cox *et al.*, "VLSI Performance Compensation for Offchip Drivers and Clock Generation," *Proceedings of the IEEE Custom Integrated Circuits Conference*, pp. 14.3.1–14.3.4, May 1989.

[87] H. Ott, *Noise Reduction Techniques in Electronic Systems.* New York: Wiley, 1988.

[88] F. Lin and D. Y. Chen, "Reduction of Power Supply EMI Emissions by Switching Frequency Modulation," *IEEE Transactions on Power Electronics*, Vol. 9, No. 1, pp. 132–137, January 1994.

[89] K. B. Hardin, J. T. Fessler, and D. R. Bush, "Spread Spectrum Clock Generation for the Reduction of Radiated Emissions," *Proceedings of the IEEE International Symposium on Electromagnetic Compatibility*, pp. 227–231, August 1994.

[90] P. Saxena and S. Gupta, "On Integrating Power and Signal Routing for Shield Count Minimization in Congested Regions," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 22, No. 4, pp. 437–445, April 2003.

[91] 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.

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

[93] S. V. Dijk and D. Hely, "Reduction of Interconnect Delay by Exploiting Crosstalk," *Proceedings of the IEEE European Solid-State Circuits Conference*, pp. 301–304, September 2001.

[94] 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.

[95] L. P. van Ginneken, "Buffer Placement in Distributed *RC* Tree Networks for Minimal Elmore Delay," *Proceedings of the IEEE International Symposium on Circuits and Systems*, pp. 865–868, May 1990.

[96] W. C. Elmore, "The Transient Response of Damped Linear Networks," *Journal of Applied Physics*, Vol. 19, pp. 55–63, January 1948.

[97] J. Rubinstein, P. Penfield, and M. A. Horowitz, "Signal Delay in *RC* Tree Networks," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 2, No. 3, pp. 202–211, July 1983.

[98] 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 Signal Processing*, Vol. 45, No. 5, pp. 607–616, May 1998.

[99] Y. I. Ismail and E. G. Friedman, "Effects of Inductance on the Propagation Delay and Repeater Insertion in VLSI Circuits," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, Vol. 8, No. 2, pp. 195–206, April 2000.

[100] C. J. Alpert, A. Devgan, and S. T. Quay, "Buffer Insertion for Noise and Delay Optimization," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 18, No. 11, pp. 1633–1645, November 1999.

[101] T. Zhang and 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.

[102] J. Cong, L. He, C. Koh, and P. H. Madden, "Performance Optimization of VLSI Interconnect Layout," *Integration, the VLSI Journal*, Vol. 21, No. 1-2, pp. 1–94, November 1996.

[103] A. B. Kahng and G. Robbins, *On Optimal Interconnections for VLSI*. Kluwer Academic Publishers, 1994.

[104] T. Gao and C. L. Liu, "Minimum Crosstalk Channel Routing," *Proceedings of the IEEE/ACM International Conference on Computer-Aided Design*, pp. 692–696, November 1993.

[105] D. A. Kirkpatrick and A. L. Sangiovanni-Vincentelli, "Techniques for Crosstalk Avoidance in the Physical Design of High-Performance Digital Systems," *Proceedings of the IEEE/ACM International Conference on Computer-Aided Design*, pp. 616–619, November 1994.

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

[107] H. Zhou and D. F. Wong, "Global Routing with Crosstalk Constraints," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 18, No. 11, pp. 1683–1688, November 1999.

[108] T. Y. Ho, Y. W. Chang, S. J. Chen, and D. T. Lee, "Crosstalk and Performance Driven Multilevel Full-Chip Routing," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 24, No. 6, pp. 869–878, June 2005.

[109] J. Xiong and L. He, "Extended Global Routing With *RLC* Crosstalk Constraints," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, Vol. 13, No. 3, pp. 319–329, March 2005.

[110] T. Xiao and M. Marek-Sadowska, "Gate Sizing to Eliminate Crosstalk Induced Timing Violation," *Proceedings of the IEEE International Conference on Computer Design*, pp. 186–191, September 2001.

[111] I. H. Jiang, Y. W. Chang, and J. Y. Jou, "Crosstalk-Driven Interconnect Optimization by Simultaneous Gate and Wire Sizing," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 19, No. 9, pp. 999–1010, September 2000.

[112] M. R. Becer *et al.*, "Postroute Gate Sizing for Crosstalk Noise Reduction," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 23, No. 12, pp. 1670–1677, December 2004.

[113] T. Kadoyama *et al.*, "A Complete Single-Chip GPS Receiver with 1.6-V 24-mW Radio in 0.18-μm CMOS," *IEEE Journal of Solid-State Circuits*, Vol. 39, No. 4, pp. 562–568, April 2004.

[114] A. Koukab, K. Banerjee, and M. Declercq, "Analysis and Optimization of Substrate Noise Coupling in Single-Chip RF Transceiver Design," *Proceedings of the IEEE/ACM International Conference on Computer-Aided Design*, pp. 309–316, November 2002.

[115] S. Kiaei, D. Allstot, K. Hansen, and N. K. Verghese, "Noise Considerations for Mixed-Signal RF IC Transceivers," *ACM Journal on Wireless Networks*, Vol. 4, pp. 41–53, January 1998.

[116] H. Eul, "ICs for Mobile Multimedia Communications," *Proceedings of the IEEE International Solid-State Circuits Conference*, pp. 21–39, February 2006.

[117] 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.

[118] G. H. Warren and C. Jungo, "Noise, Crosstalk, and Distortion in Mixed Analog/Digital Integrated Circuits," *Proceedings of the IEEE Custom Integrated Circuits Conference*, pp. 12.1.1–12.1.4, May 1988.

[119] G. A. S. Machado (Editor), *Low Power HF Microelectronics*. Institution of Electrical Engineers, 1996.

[120] 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.

[121] 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.

[122] J. Briaire and K. S. Krisch, "Principles of Substrate Crosstalk Generation in CMOS Circuits," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 19, No. 6, pp. 645–653, June 2000.

[123] E. Charbon, P. Miliozzi, L. P. Carloni, A. Ferrari, and A. Sangiovanni-Vincentelli, "Modeling Digital Substrate Noise Injection in Mixed-Signal IC's," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 18, No. 3, pp. 301–310, March 1999.

[124] T. Y. Chan, P. K. Ko, and C. Hu, "A Simple Method to Characterize Substrate Current in MOSFET's," *IEEE Electron Device Letters*, Vol. EDL-5, No. 12, pp. 505–507, December 1984.

[125] A. G. Chynoweth, "Ionization Rates for Electrons and Holes in Silicon," *Physical Review Letters*, Vol. 109, No. 5, pp. 1537–1540, March 1958.

[126] M. Badaroglu *et al.*, "Evolution of Substrate Noise Generation Mechanisms With CMOS Technology Scaling," *IEEE Transactions on Circuits and Systems I: Regular Papers*, Vol. 53, No. 2, pp. 296–305, February 2006.

[127] S. M. Sze, *Physics of Semiconductor Devices*. John Wiley and Sons Inc., 1981.

[128] D. K. Cheng, *Field and Wave Electromagnetics*. Addison-Wesley, 1989.

[129] 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.

[130] N. K. Verghese, D. J. Allstot, and M. A. Wolfe, "Verification Techniques for Substrate Coupling and their Application to Mixed-Signal IC Design," *IEEE Journal of Solid-State Circuits*, Vol. 31, No. 3, pp. 354-365, March 1996.

[131] 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.

[132] N. K. Verghese, D. J. Allstot, and S. Masui, "Rapid Simulation of Substrate Coupling Effects in Mixed-Mode ICs," *Proceedings of the IEEE Custom Integrated Circuits Conference*, pp. 18.3.1-18.3.4, May 1993.

[133] J. P. Costa, M. Chou, and L. M. Silveria, "Efficient Techniques for Accurate Modeling and Simulation of Substrate Coupling in Mixed-Signal IC's," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 18, No. 5, pp. 597-607, May 1999.

[134] T. L. Pillage, R. A. Rohrer, and C. Visweswariah, *Electronic Circuit and System Simulation Methods*. McGraw-Hill, 1994.

[135] 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, March 1994.

[136] 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.

[137] G. F. Roach, *Green's Functions*. Cambridge University Press, 1970.

[138] H. Li *et al.*, "Comprehensive Frequency Dependent Substrate Noise Analysis Using Boundary Element Methods," *Proceedings of the IEEE/ACM International Conference on Computer-Aided Design*, pp. 2–9, November 2002.

[139] H. Lan, Z. Yu, and R. W. Dutton, "A CAD-oriented Modeling Approach of Frequency-Dependent Behavior of Substrate Noise Coupling for Mixed-Signal IC Design," *Proceedings of the IEEE International Symposium on Quality Electronic Design*, pp. 195-200, March 2003.

[140] J. D. Jackson, *Classical Electrodynamics*. Wiley, 1962.

[141] L. T. Pillage and R. A. Rohrer, "Asymptotic Waveform Evaluation for Timing Analysis," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 9, No. 4, pp. 352-366, April 1990.

[142] E. Charbon, R. Gharpurey, R. G. Meyer, and A. Sangiovanni-Vincentelli, "Substrate Optimization Based on Semi-Analytical Techniques," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 18, No. 2, pp. 172-190, February 1999.

[143] A. M. Niknejad, R. Gharpurey, and R. G. Meyer, "Numerically Stable Green Function for Modeling and Analysis of Substrate Coupling in Integrated Circuits," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 17, No. 4, pp. 305-315, April 1998.

[144] E. Schrik and N. P. van der Meijs, "Combined BEM/FEM Substrate Resistance Modeling," *Proceedings of the IEEE/ACM Design Automation Conference*, pp. 771-776, June 2002.

[145] 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.

[146] 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, September 2002.

[147] M. Badaroglu *et al.*, "SWAN: High-Level Simulation Methodology for Digital Substrate Noise Generation," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, Vol. 14, No. 1, pp. 23-33, January 2006.

[148] S. M. Reddy and R. Murgai, "Accurate Substrate Noise Analysis Based on Library Module Characterization," *Proceedings of the IEEE International Conference on VLSI Design*, January 2006.

[149] P. Birrer, S. K. Arunachalam, M. Held, K. Mayaram, and T. S. Fiez, "Schematic-Driven Substrate Noise Coupling Analysis in Mixed-Signal IC Designs," *IEEE Transactions on Circuits and Systems I: Regular Papers*, Vol. 53, No. 12, pp. 2578–2587, December 2006.

[150] O. Valorge *et al.*, "A Simple Way for Substrate Noise Modeling in Mixed-Signal ICs," *IEEE Transactions on Circuits and Systems I: Regular Papers*, Vol. 53, No. 10, pp. 2167-2177, October 2006.

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

[152] D. Leenaerts and P. Vreede, "Influence of Substrate Noise on RF Performance," *Proceedings of the IEEE European Solid-State Circuits Conference*, pp. 328–331, September 2000.

[153] M. Xu, D. K. Su, D. K. Shaeffer, T. H. Lee, and B. A. Wooley, "Measuring and Modeling the Effects of Substrate Noise on the LNA for a CMOS GPS Receiver," *IEEE Journal of Solid-State Circuits*, Vol. 36, No. 3, pp. 473–485, March 2001.

[154] S. Hazenboom, T. S. Fiez, and K. Mayaram, "A Comparison of Substrate Noise Coupling in Lightly and Heavily Doped CMOS Processes for 2.4 GHz LNAs," *IEEE Journal of Solid-State Circuits*, Vol. 41, No. 3, pp. 574–587, March 2006.

[155] B. Razavi, *Phase-Locking in High-Performance Systems.* Wiley-Interscience, 2003.

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

[157] F. Herzel and B. Razavi, "A Study of Oscillator Jitter Due to Supply and Substrate Noise," *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, Vol. 46, No. 1, pp. 56–62, January 1999.

[158] 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.

[159] R. J. Welch and A. T. Yang, "Substrate Coupling Analysis and Simulation for an Industrial Phase Locked Loop," *Proceedings of the IEEE International Symposium on Circuits and Systems*, pp. 94–97, May 1998.

[160] P. Heydari, "Characterizing the Effects of the PLL Jitter Due to Substrate Noise in Discrete Time Delta-Sigma Modulators," *IEEE Transactions on Circuits and Systems I: Regular Papers*, Vol. 52, No. 6, pp. 1073–1085, June 2005.

[161] T. Blalack and B. A. Wooley, "The Effects of Switching Noise on an Oversampling A/D Converter," *Proceedings of the IEEE International Solid-State Circuits Conference*, pp. 200–201, February 1995.

[162] 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.

[163] 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.

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

[165] M. Alles and S. Wilson, "Thin Film Silicon-On-Insulator: An Enabling Technology," *Semiconductor International*, Vol. 20, No. 4, pp. 67–74, April 1997.

[166] A. G. Aipperspach, D. H. Allen, D. T. Cox, N. V. Phan, and S. N. Storino, "A 0.2-μm, 1.8-V, SOI, 550-MHz, 64-b PowerPC Microprocessor with Copper Interconnects," *IEEE Journal of Solid-State Circuits*, Vol. 34, No. 11, pp. 1430-1435, November 1999.

[167] P. T. Tran, "Overview of Fully Depleted Silicon-On-Insulator (SOI) Technology," *Proceedings of the 15th Biennial University/Government/Industry Microelectronics Symposium*, pp. 370–371, June 2003.

[168] J. P. Colinge, J. P. Eggermont, D. Flandre, P. Francis, and P. G. A. Jespers, "Potential of SOI for Analog and Mixed Analog-Digital Low-Power Applications," *Proceedings of the IEEE International Solid-State Circuits Conference*, pp. 194–195, February 1995.

[169] K. Joardar, "A Simple Approach to Modeling Cross-Talk in Integrated Circuits," *IEEE Journal of Solid-State Circuits*, Vol. 29, No. 10, pp. 1212-1219, October 1994.

[170] J. P. Raskin, A. Viviani, D. Flandre, and J. P. Colinge, "Substrate Crosstalk Reduction Using SOI Technology," *IEEE Transactions on Electron Devices*, Vol. 44, No. 12, pp. 2252–2261, December 1997.

[171] C. Y. Yeh and M. M. Sadowska, "Timing-Aware Power-Noise Reduction in Placement," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 26, No. 3, pp. 526-541, March 2007.

[172] 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.

[173] H. Yamamoto and J. A. Davis, "Decreased Effectiveness of On-Chip Decoupling Capacitance in High-Frequency Operation," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, Vol. 15, No. 6, pp. 649–659, June 2007.

[174] P. Larsson, "Resonance and Damping in CMOS Circuits with On-Chip Decoupling Capacitance," *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, Vol. 45, No. 8, pp. 849-858, August 1998.

[175] H. H. Chen and J. S. Neely, "Interconnect and Circuit Modeling Techniques for Full-Chip Power Supply Noise Analysis," *IEEE Transactions on Components, Packaging, and Manufacturing Technology*, Vol. 21, No. 3, pp. 209-215, August 1998.

[176] J. Chen and L. He, "Efficient In-Package Decoupling Capacitor Optimization for I/O Power Integrity," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 26, No. 4, pp. 734–738, April 2007.

[177] M. Sotman, M. Popovich, A. Kolodny, and E. G. Friedman, "Leveraging Symbiotic On-Die Decoupling Capacitance," *Proceedings of the IEEE Topical Meeting on Electrical Performance of Electronic Packaging*, pp. 111-114, October 2005.

[178] M. Badaroglu *et al.*, "Digital Circuit Capacitance and Switching Analysis for Ground Bounce in ICs With a High-Ohmic Substrate," *IEEE Journal of Solid-State Circuits*, Vol. 39, No. 7, pp. 1119-1130, July 2004.

[179] P. Larsson, "di/dt Noise in CMOS Integrated Circuits," *Analog Integrated Circuits and Signal Processing*, Vol. 14, No. 2, pp. 113-129, September 1997.

[180] M. Badaroglu *et al*, "Clock Skew Optimization Methodology for Substrate Noise Reduction With Supply Current Folding," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 25, No. 6, pp. 1146-1154, June 2006.

[181] 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.

[182] R. Senthinathan, G. Tubbs, and M. Schuelein, "Negative Feedback Influence in Simultaneously Switching CMOS Outputs," *Proceedings of the IEEE Custom Integrated Circuits Conference*, pp. 5.4.1-5.4.5, May 1988.

[183] R. Panda *et al.*, "Model and Analysis for Combined Package and On-Chip Power Grid Simulation," *Proceedings of the IEEE International Symposium on Low Power Electronics and Design*, pp. 179–184, July 2000.

[184] Y. Yang and J. R. Brews, "Design Trade-Offs for the Last Stage of an Unregulated Long-Channel CMOS Off-Chip Driver with Simultaneous Switching Noise and Switching Time Considerations," *IEEE Transactions on Components, Packaging, and Manufacturing Technology*, Vol. 19, No. 3, pp. 481-486, August 1997.

[185] S. H. Hashemi, P. A. Sandborn, D. Disko, and R. Evans, "The Close Attached Capacitor: A Solution to Switching Noise Problems," *IEEE Transactions on Components, Hybrids, and Manufacturing Technology*, Vol. 15, No. 6, pp. 1056-1063, December 1992.

[186] 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.

[187] M. Heijningen, "Analysis and Experimental Verification of Digital Substrate Noise Generation for Epi-Type Substrates," *IEEE Journal of Solid-State Circuits*, Vol. 35, No. 7, pp. 1002-1008, July 2000.

[188] R. M. Secareanu *et al.*, "Substrate Coupling in Digital Circuits in Mixed-Signal Smart-Power Systems," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, Vol. 12, No. 1, pp. 67-78, January 2004.

[189] R. Panda, S. Sundareswaran, and D. Blaauw, "Impact of Low-Impedance Substrate on Power Supply Integrity," *IEEE Design and Test of Computers*, Vol. 2, No. 3, pp. 16-22, May 2003.

[190] 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.

[191] T. Sakurai, "Approximation of Wiring Delay in MOSFET LSI," *IEEE Journal of Solid-State Circuits*, Vol. 18, No. 4, pp. 418-426, August 1983.

[192] *Berkeley Short-Channel IGFET Model 3, Version 3.0*, University of California, Berkeley, 1995.

[193] C. Soens *et al.*, "Performance Degradation of LC-Tank VCOs by Impact of Digital Switching Noise in Lightly Doped Substrates," *IEEE Journal of Solid-State Circuits*, Vol. 40, No. 7, pp. 1472-1481, July 2005.

[194] R. Murgai *et al.*, "Sensitivity-Based Modeling and Methodology for Full-Chip Substrate Noise Analysis," *Proceedings of the IEEE Design, Automation and Test Conference*, pp. 610-615, February 2004.

[195] B. E. Owens *et al.*, "Simulation and Measurement of Supply and Substrate Noise in Mixed-Signal ICs," *IEEE Journal of Solid-State Circuits*, Vol. 40, No. 2, pp. 382-391, February 2005.

[196] "Assura RCX^{TM}, SubstrateStorm^{TM}, Spectre^{TM} tools." [Online]. Available: http://www.cadence.com.

[197] M. Zhao, R. V. Panda, S. S. Sapatnekar, and D. Blaauw, "Hierarchical Analysis of Power Distribution Networks," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 21, No. 2, pp. 159-168, February 2002.

[198] S. Donnay and G. Gielen, *Substrate Noise Coupling in Mixed-Signal ASICs*. Springer, 2003.

[199] "Q3D Extractor^{TM} tool," [Online]. Available: http://www.ansoft.com.

[200] R. R. Troutman, *Latch-up in CMOS Technology: The Problem and Its Cure*. Springer, 1986.

[201] M. Felder and J. Ganger, "Analysis of Ground-Bounce Induced Substrate Noise Coupling in a Low Resistive Bulk Epitaxial Process: Design Strategies to Minimize Noise Effects on a Mixed-Signal Chip," *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, Vol. 46, No. 11, pp. 1427–1436, November 1999.

[202] M. Badaroglu *et al.*, "Methodology and Experimental Verification for Substrate Noise Reduction in CMOS Mixed-Signal ICs with Synchronous Digital Circuits," *IEEE Journal of Solid-State Circuits*, Vol. 37, No. 11, pp. 1383-1395, November 2002.

[203] W. Roethig, "Library Characterization and Modeling for 130 nm and 90 nm SOC Design," *Proceedings of the IEEE International SOC Conference*, pp. 383-386, September 2003.

[204] D. Patel, "CHARMS: Characterization and Modeling System for Accurate Delay Prediction of ASIC Designs," *Proceedings of the IEEE Custom Integrated Circuits Conference*, pp. 9.5.1 - 9.5.6, May 1990.

[205] R. W. Phelps, "Advanced Library Characterization for High-Performance ASIC," *Proceedings of the IEEE International ASIC Conference*, pp. 15-3.1 - 15-3.4, September 1991.

[206] M. A. Cirit, "Characterizing a VLSI Standard Cell Library," *Proceedings of the IEEE Custom Integrated Circuits Conference*, pp. 25-7.1 - 25-7.4, May 1991.

[207] V. Stojanovic and V.G. Oklobdzija, "Comparative Analysis of Master-Slave Latches and Flip-Flops for High-Performance and Low-Power Systems," *IEEE Journal of Solid-State Circuits*, Vol. 34, No. 4, pp. 536-548, April 1999.

[208] A. M. Jain and D. Blaauw, "Modeling Flip Flop Delay Dependencies in Timing Analysis," *Proceedings of the ACM/IEEE TAU Workshop*, February 2004.

[209] G. Rao and E. K. Howick, "Apparatus for Optimized Constraint Characterization with Degraded Options and Associated Methods." US Patent No. 6,584,598 B2, June 24 2003.

[210] E. K. Howick, "Conquering the High-Frequency Domain with Predictable Sequential Models," *Proceedings of the Electrical and Physical Design Conference*, January 2002.

[211] N. Weste and D. Harris, *CMOS VLSI Design.* Addison Wesley, 2004.

[212] J. L. Neves and E. G. Friedman, "Optimal Clock Skew Scheduling Tolerant to Process Variations," *Proceedings of the IEEE/ACM Design Automation Conference*, pp. 623–628, June 1996.

[213] J. Semiao *et al.*, "Improving Tolerance to Power Supply and Temperature Variations in Synchronous Circuits," *Proceedings of the IEEE Conference on Design and Diagnostics of Electronic Circuits and Systems*, pp. 1–6, April 2007.

[214] S. C. Chang, C. T. Hsieh, and K. C. Wu, "Resynthesis for Delay Variation Tolerance," *Proceedings of the IEEE/ACM Design Automation Conference*, pp. 814–819, July 2004.

[215] S. Srivastava and J. Roychowdhury, "Independent and Interdependent Latch Setup/Hold Time Characterization via Newton-Raphson Solution and Euler Curve Tracking of State-Transition Equations," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 27, No. 5, pp. 817–830, May 2008.

[216] S. Hatami, H. Abrishami, and M. Pedram, "Statistical Timing Analysis of Flip-Flops Considering Codependent Setup and Hold Times," *Proceedings of the IEEE/ACM Great Lakes Symposium on VLSI*, pp. 101–106, May 2008.

[217] E. G. Friedman, "Clock Distribution Networks in Synchronous Digital Integrated Circuits," *Proceedings of the IEEE*, Vol. 89, No. 5, pp. 665–692, May 2001.

[218] Y. Cao, T. Sato, D. Sylvester, M. Orshansky, and C. Hu, "New Paradigm of Predictive MOSFET and Interconnect Modeling for Early Circuit Design," *Proceedings of the IEEE Custom Integrated Circuits Conference*, pp. 201-204, May 2000.

[219] "Predictive Technology Model (PTM)." [Online]. Available: http://www.eas.asu.edu/ ptm.

[220] IEEE Circuits and Systems Society, *Proceedings of the IEEE International Solid-State Circuits Conference*.

Appendix A

An Upper Bound on the Error

Referring to Fig. 7.9, C_1 and C_2 are merged into C_m if

$$V_{diff}(c2) = R_2 max(i_2[t]) + L_2 max(\frac{\partial i_2[t]}{\partial t}) < V_{lim}. \tag{A.1}$$

Note that $i_x[t]$ is assumed to be zero to simplify the analysis. The error $E[t]$ due to this merging is determined from (7.14). Using (7.12) and (7.13), this error can be rewritten as

$$\begin{aligned} E[t] &= |i_2[t]R_2 + L_2\frac{\partial i_2[t]}{\partial t} - (i_1[t] + i_2[t])R_2\frac{max(|i_2[t]|)}{max(|i_1[t] + i_2[t]|)} \\ &\quad -L_2\frac{max(|\partial i_2[t]/\partial t|)}{max(|\partial(i_1[t] + i_2[t])/\partial t|)}\frac{\partial(i_1[t] + i_2[t])}{\partial t}|. \end{aligned} \tag{A.2}$$

Two cases need to be investigated to determine the upper bound for this expression: (1) The first two terms are greater than zero and the last two terms are smaller than

zero, and (2) the first two terms are smaller than zero and the last two terms are greater than zero. Assuming the first case holds, the error is

$$E[t] = i_2[t]R_2 + L_2\frac{\partial i_2[t]}{\partial t} + (i_1[t] + i_2[t])R_2\frac{max(|i_2[t]|)}{max(|i_1[t] + i_2[t]|)}$$

$$+L_2\frac{max(|\partial i_2[t]/\partial t|)}{max(|\partial(i_1[t] + i_2[t])/\partial t|)}\frac{\partial(i_1[t] + i_2[t])}{\partial t}, \tag{A.3}$$

where each term is greater than zero. Since

$$\frac{(i_1[t] + i_2[t])}{max(|i_1[t] + i_2[t]|)} \leq 1, \tag{A.4}$$

and

$$\frac{\partial(i_1[t] + i_2[t])/\partial t}{max(|\partial(i_1[t] + i_2[t])/\partial t)} \leq 1, \tag{A.5}$$

the error expression is

$$E[t] \leq i_2[t]R_2 + L_2\frac{\partial i_2[t]}{\partial t} + R_2 max(|i_2[t]|) + L_2 max(|\partial(i_2[t])/\partial t|). \tag{A.6}$$

The summation of the first two terms and the last two terms on the right side of the inequality is smaller than V_{lim}, as determined by (A.1). The upper bound of the error

due to merging C_1 and C_2 into a single contact is therefore $2 \times V_{lim}$,

$$E[t] \leq 2V_{lim}. \tag{A.7}$$

Note that the same bound holds for the second case where the first two terms in (A.2) are smaller than zero and the last two terms are greater than zero.