MODELING AND ANALYSIS OF SUBSTRATE COUPLING IN INTEGRATED CIRCUITS

by

Ranjit Gharpurey

Memorandum No. UCB/ERL M95/47

28 June 1995

MODELING AND ANALYSIS OF SUBSTRATE COUPLING IN INTEGRATED CIRCUITS

by

Ranjit Gharpurey

Memorandum No. UCB/ERL M95/47

28 June 1995

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

Abstract

Modeling and Analysis of Substrate Coupling in Integrated Circuits

by

Ranjit Gharpurey

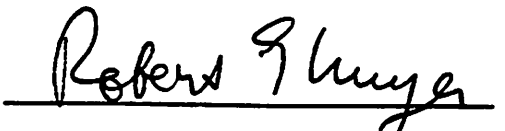
Doctor of Philosophy in Engineering- Electrical Engineering and Computer Sciences

University of California at Berkeley

Professor Robert G. Meyer, Chair

Substrate coupling in integrated circuits is the process whereby parasitic current flow in the substrate electrically couples devices in different parts of the circuit. Higher levels of integration and higher frequencies of operation make the coupling more pronounced in modern circuit realizations. High levels of integration are desirable in several applications for reducing the overall power dissipation, reducing the number of components and lowering costs. Portable radio-receivers are an example of such an application. Electrical coupling in the substrate leads to undesirable interaction between devices which can degrade circuit performance. The degradation can manifest itself in several ways. In mixed analog-digital circuits, for example, the switching-noise generated by digital circuits can be coupled to the sensitive analog circuits through the substrate.

Performance degradation due to substrate coupling can be addressed at the circuit-design stage by including accurate substrate models in circuit simulations. An efficient and elegant technique to model substrate coupling is presented in this dissertation. The technique uses a combination of the classical Green function approach and the Fast Fourier Transform. The speed of this technique makes it suitable for optimization of circuit layout for minimization of substrate coupling related effects. The nature of substrate coupling in different types of substrates has been analyzed. The effectiveness of isolation schemes, such as guard rings, in different types of substrates has also been presented. Several effects of substrate coupling on circuit performance have been identified and remedies have been suggested. Experimental verification of the substrate models has been performed.



Prof. Robert G. Meyer
Thesis Committee Chairman

Table of Contents

Table of Contents iii

Acknowledgments vi

Chapter 1: Introduction 1

Chapter 2: Injection, Reception and Transmission of Substrate Noise 2

2.0: Injection of Substrate Noise 4

2.0.1: Injection Mechanisms in Bipolar Junction Transistors and Diodes 4

2.0.2: Current Injection due to MOS Devices 5

2.0.3: Current Injection Mechanisms in Passive On-Chip Components 7

2.1: Reception of Substrate Noise 8

2.2: Transmission of Substrate Noise 9

2.2.0: Simple Substrate Models 10

2.2.1: High-Frequency Characterization of Substrates 11

Chapter 3: Modeling of Substrate Coupling in Integrated Circuits 13

3.0: Approximations made in the Analysis 13

3.0.1: The Electrostatic Assumption 14

3.0.2: The Assumption of Linearity 15

3.0.3: The Assumption of Equipotentials 15

3.1: A Numerical Technique for Extraction of Substrate Models 16

3.1.1: The Laplace Equation and Kirchoff's Laws 16

3.1.2: Direct Techniques for Solution of the Linear System 20

3.1.3: Simple Iterative Techniques 21

3.1.4: Modeling the Frequency-Dependence of Substrates 24

3.2: An Integral-Equation-based approach to Substrate Modeling 26

3.2.1: Techniques for Solution of Electrostatic Problems 26

3.2.2: Use of the Green Function for Parasitic Extraction 33

3.2.3: Extraction of Substrate Parasitics 36

3.2.4: Matrix Inversion 42

3.2.5: Improvements to the Model 45

3.3: Comparison of Techniques 51

Chapter 4: Substrate Coupling in Integrated Circuits 54

4.0: Substrate Coupling in Different Substrate Types 54

4.0.1: Variation of Model Elements as Functions of Position and Contact Area 56

4.0.2: Isolation between Single-Ended Contacts with Capacitive Injection and Reception Mechanisms 60

4.0.3: Improvement in Isolation in Differential Circuits with Capacitive Injection and Reception Mechanisms 69

4.0.4: Dependence of Substrate Coupling on Load Impedance 75

4.0.5: Behavior of Substrate Coupling with Non-Capacitive Reception and Injection Mechanisms 76

4.1: Guard Rings and Substrate Taps 78

4.1.1: Guard Rings in Different Substrate Types 78

4.1.2: Optimal Sizing and Placement of Guard Rings in Single-Ended Circuits 82

4.1.3: Guard Rings in Differential Circuits 95

4.1.4: The Use of Dual Guard Rings 97

4.2: Effects of Substrate Coupling on Circuit Performance 98

4.2.1: Power Loss in the Substrate 99

4.2.2: Degradation of Circuit Performance due to Substrate-Coupled Noise 102

4.2.3: Change in Circuit Bandwidth 108

4.2.4: Change in Circuit Gain 110

4.2.5: Catastrophic Consequences of Substrate Coupling 111

4.3: Conclusion 112

Chapter 5: Experimental Verification 113

5.1: Test-Chip Description 113

5.2: Measurement Procedure 113

5.3: Simulation Procedure 114

5.4: Discussion of Results 114

Chapter 6: Conclusions and Future Directions 118

References 119

Appendix A3.1: The Electrostatic Green Function in the Substrate Medium 122

Appendix A3.2: Derivation of the Coefficients of Potential 127

Appendix A3.3: The Discrete Cosine Transform 132

References 133

Acknowledgments

I wish to express my gratitude to my advisor, Prof. Robert G. Meyer, for his guidance and encouragement during the course of my doctoral studies. I am also grateful to Prof. Paul R. Gray for his encouragement and suggestions. I would like to thank the members of my qualifying committee, Prof. Paul R. Gray, Prof. Bernhard E. Boser and Prof. Stanley Mandelstam for their time.

My studies at Berkeley have been made greatly stimulating and enjoyable by the faculty members I could interact with and learn from. I am indebted to them. I am grateful to Prof. Chenming Hu, for his exceptional support and kindness during my first two semesters at Berkeley.

I am thankful to William Mack of Philips Semiconductors for his generous help with the experimental part of this work and for sharing his intuition and knowledge with me. Thanks are also due to Sheng Lee of Philips Semiconductors for the experimental test-chips he provided, and to Kevin McAdams, also of Philips Semiconductors, for his help with the test-board fabrication. I benefited greatly from my interaction with Brian Brandt, John Fattaruso, Venu Gopinathan, Sami Kiriaki and Pradeep Shah of the analog group at Texas Instruments, Dallas, during the summer of 1993, and I thank them for their time, interest and encouragement.

It was a great pleasure to interact with the members of my research group Pramote Piriyapoksombut, Keng Fong, Cynthia Keys, Robert Chu, Chris Hull, Joo Leong Tham and Nhat Nguyen. Special thanks are due to Seema Butala and Sovarong Leang for having been great friends and project partners and to fellow graduate students Thomas Cho, Cormac Conroy, Edward Liu, Sekhar Narayanaswami, Keith Onodera, Jeff Ou, Todd Weigandt, Naiyavudhi Wongkomet and Darrin Young. I am especially grateful to Edoardo Charbon, for his encouragement and interest in my work. Thanks are due to Heather Brown, Dianna Bolt and Tom Boot for helping me with administrative matters.

I am thankful to my undergraduate faculty at the Indian Institute of Technology, Kharagpur, in particular Prof. B. N. Chatterjee, Prof. S. K. Lahiri, Prof. J. C. Biswas and Prof. N. B. Chakraborty for their help and guidance during my undergraduate years.

I am grateful to my wife, Anupa, for her support and advice during my Ph.D. and to my father and my

sister for their encouragement. I am indebted to my mother, Usha Gharpurey, for her love and support and for being a source of strength and inspiration.

This work was supported by the Advanced Research Projects Agency under Grant J-FBI-92-150.

Chapter 1: Introduction

The growth of the personal communications market in recent years has led to a great demand for low-power radio-frequency circuits with high levels of integration for portability and compact size. Certain types of circuits have traditionally been built on separate substrates in order to minimize parasitic interaction between them, for example low-noise amplifiers and switching circuits such as dividers are traditionally built on separate substrates. Integrating noise-generating circuits and low-noise circuits on the same substrate is being currently viewed as a major challenge by circuit designers.

Higher levels of integration have several associated advantages and disadvantages. An obvious advantage is the reduced package count. This leads to lowered costs and reduced sizes. The power dissipation can also be reduced as fewer pads and interconnect lines need to be driven, thereby avoiding the associated capacitance. It may also be possible to improve the high-frequency response of circuits or even extend the frequency range of circuit performance, as the package interconnect parasitics often degrade the frequency response at the high-frequency end of the application.

A major disadvantage of integration is the increased interaction between circuits. This interaction can appear in two major ways. It can occur due to the significant mutual inductance and capacitance which exist between any two bond-wires and pins in a package. The second method for interaction is through the common substrate shared by the circuits. The isolation provided by the substrates is non-ideal. Currents can flow through the substrate due to the nonzero dielectric constants and conductivities of the substrate materials and couple circuits located in different parts of the substrate.

The problem of substrate coupling is addressed in this research. The sources of coupling are identified first, and the problem of efficient modeling of substrate coupling is discussed next. As will be discussed later, substrate coupling is a layout related effect, and the layout needs to be optimized in order to minimize substrate related noise. Thus any simulation strategy for substrate noise estimation must be suitable for optimization tasks. Such techniques are presented in Chapter 3. Substrate coupling in different substrate types and circuit effects of substrate coupling are discussed in Chapter 4. Experimental results are shown subsequently.

Chapter 2: Injection, Reception and Transmission of Substrate Noise

The cross-sections of two typical substrates currently being used in integrated-circuit manufacturing are shown in Fig. 1. The substrates shown here are the low impedance and high impedance bulk substrates. Fig. 1a depicts a high-impedance substrate. These substrates are composed of a lightly-doped bulk region which is about 200-400 μm thick and a thin epi-layer which has a lower resistivity. The low-impedance substrate consists of a lightly-doped p or n-type epi-layer grown on a heavily-doped p or n-type bulk. The bulk is typically 100-400 μm thick, and the epi-layer thickness varies from 5 to 15 μm. The epi-layer thickness is limited by the upward diffusion of dopants during any high-temperature processing steps. Typical bulk and epi-layer resistivities are shown in Fig. 1b.

The mechanisms for substrate noise injection and reception in different devices and the transmission mechanisms in substrates are discussed in this chapter.

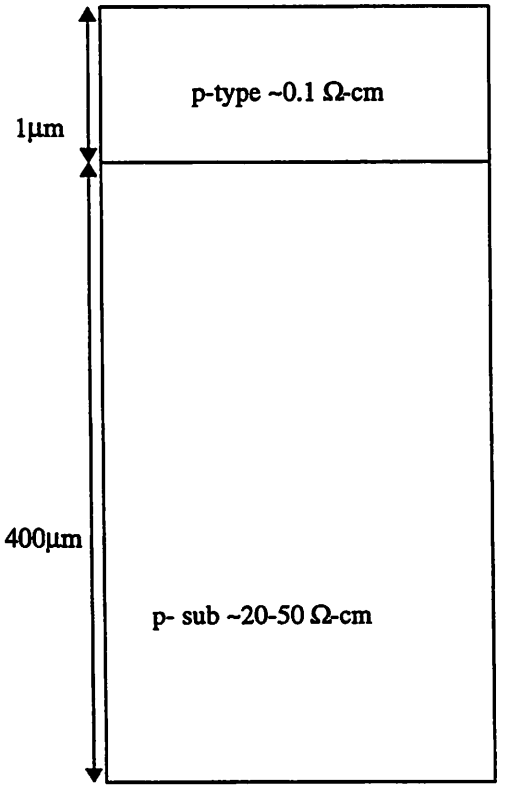


Figure 1a: High-Resistivity Substrate

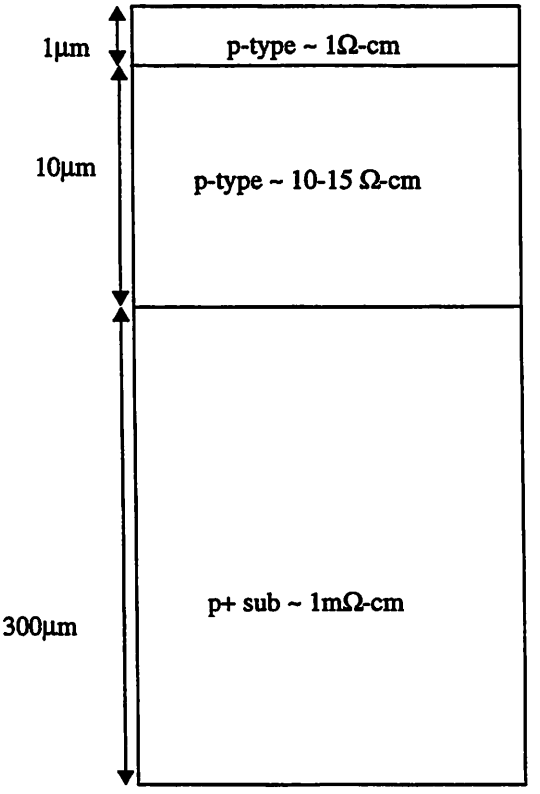


Figure 1b: Low-Resistivity Substrate

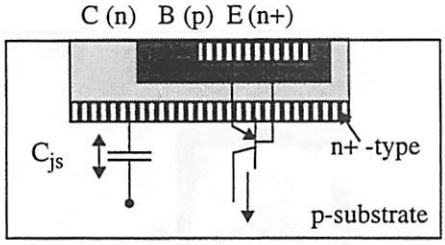


Figure 2a: NPN Transistor

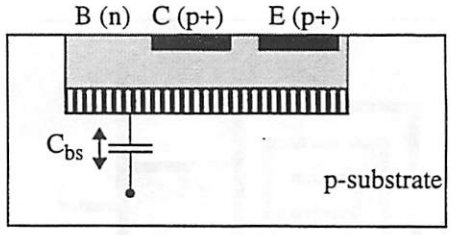


Figure 2b: Lateral PNP Transistor

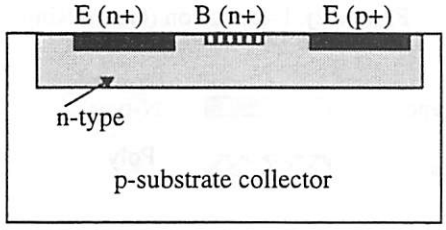


Figure 2c: Vertical PNP Transistor

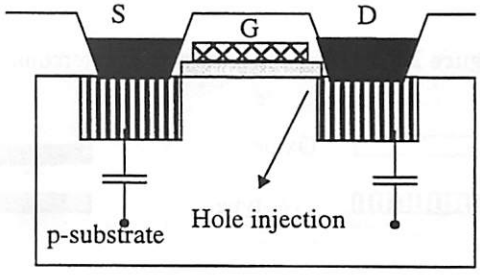


Figure 2d: NMOS Transistor

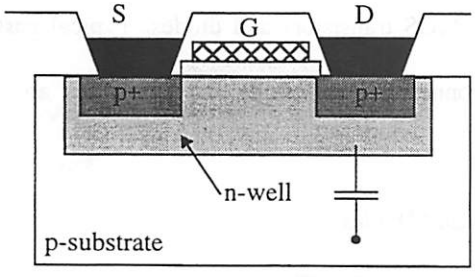


Figure 2e: PMOS Transistor

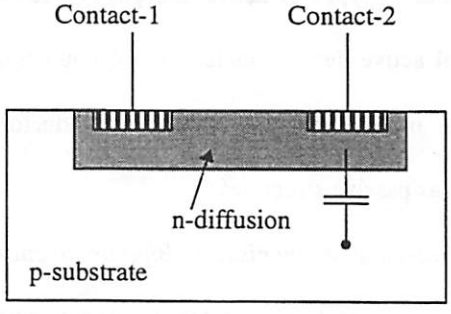


Figure 2f: Diffused Resistor

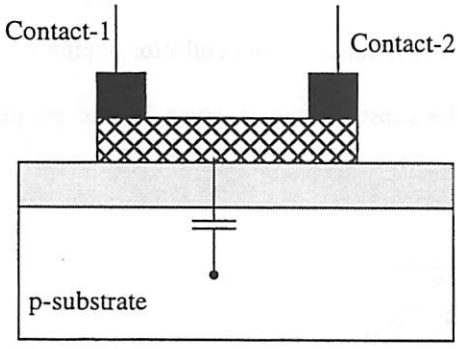


Figure 2g: Poly Resistor

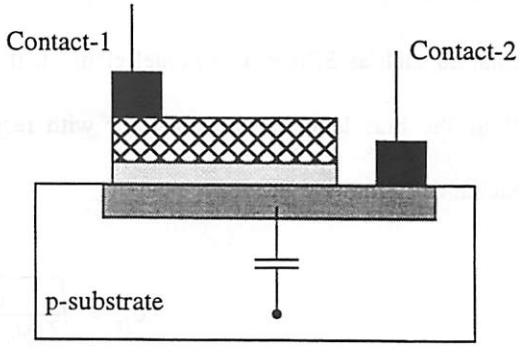


Figure 2h: Gate Oxide Capacitor

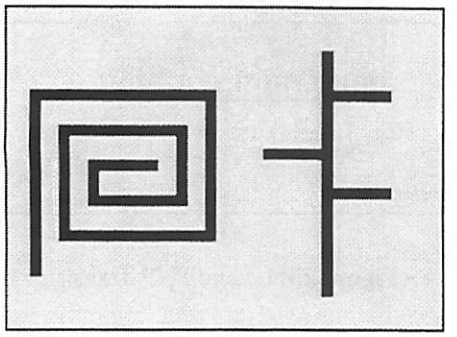


Figure 2i: A Spiral Inductor and an Interconnect

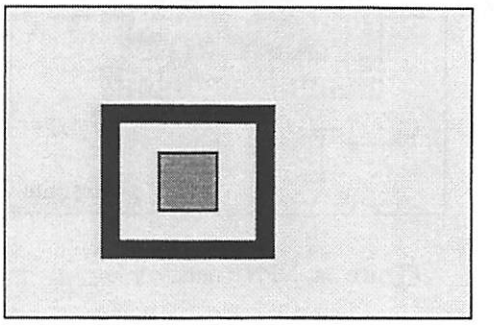


Figure 2j: P-Diffusion (Guard Ring)

2.0: Injection of Substrate Noise

Different types of active and passive devices used in silicon integrated circuits are shown in Fig. 2. Typical active devices include bipolar junction transistors, MOS transistors and diodes. Typical passive devices include resistors, capacitors, inductors and interconnects. Local wells and diffusions are also shown as passive structures.

2.0.1: Injection Mechanisms in Bipolar Junction Transistors and Diodes

The cross-section of a bipolar npn transistor is shown in Fig. 2a. These devices interact with the substrate through the collector-to-bulk pn junction capacitance (C_{js}). This capacitance is included in circuit simulators such as SPICE as a model element. It depends on the substrate and collector doping levels, as well as the bias level of the collector with respect to the substrate and is given by the pn junction capacitance formula

$$C_{js} = \sqrt{\frac{q\varepsilon}{2(\psi_{bi} + V_{cs})}\left(\frac{N_C N_S}{N_C + N_S}\right)} \qquad (1)$$

where N_C and N_S are the collector and the substrate doping levels respectively (cm^{-3}), ψ_{bi} is the built-in

potential of the junction and V_{cs} is the collector-to-substrate bias voltage. This formula is valid for abrupt pn junctions. Capacitance expressions for more complex doping profiles are discussed in [1].

Another mechanism for noise injection into the substrate from a bipolar transistor is observed when the device approaches the saturation region of operation. When the base of the transistor begins to be forward biased with respect to the collector, the parasitic pnp transistor shown in Fig. 2a begins to enter its forward active region of operation. The base of the npn device acts as the emitter of the pnp, the collector as its base and the p-substrate as its collector. The parasitic pnp device is in the cutoff region when the npn transistor is in the forward-active mode. It can be expected that the forward gain of this device is small, since the parasitic base (the npn collector) thickness is quite large.

Lateral pnp transistors inject noise into the substrate through the base-to-substrate capacitance (Fig. 2b). In vertical pnp devices the substrate is the collector node (Fig. 2c). Hence these devices can act as significant substrate noise injectors, and a sufficiently low impedance path must be provided near the device in order to collect the current [2].

Diodes are typically fabricated by tying together the base and the collector of bipolar npn transistors. Substrate current injection in these devices takes place through the collector-to-substrate junction capacitance.

2.0.2: Current Injection due to MOS Devices

MOS devices are shown in Fig. 2d and 2e. The figures depict an n-well process. In such a process, NMOS devices interact directly with the substrate through the source/drain-to-substrate capacitance, modeled as CJ0 and CJSW in SPICE. Additionally, hot-electron effects also cause injection of majority-carriers into the substrate. Hot-electron effects are observed when the field in the depleted drain-end of the transistor becomes large enough to cause impact ionization and generate electron-hole pairs. The dependence of the hot-electron induced substrate current I_{sub} on the device operating current is given by the following semi-analytical expression.

$$I_{sub} = K_1 (V_{ds} - V_{dsat}) I_d \exp\left(-\frac{K_2}{(V_{ds} - V_{dsat})} \right) \tag{2}$$

where I_d is the drain current, V_{ds} is the drain-to-source voltage and V_{dsat} is the drain-to-source voltage at saturation. K_1 and K_2 are semi-empirical constants. K_2 depends on t_{ox} (the oxide thickness) and x_j (the drain junction depth) as

$$K_2 \propto t_{ox}^{1/3} x_j^{1/2} \tag{3}$$

This phenomenon is discussed extensively in [3] and the references contained there. Eq. (3) is derived by considering the exponential dependence of the carrier ionization-coefficients on the inverse of the electrical field in the channel. Integrating the substrate current generated per unit length, over the length of the channel, results in (3). Recent experimental evidence suggests that hot-electron induced substrate currents are the dominant cause of substrate noise in NMOSFETs up to at least one hundred megahertz [4]. Shorter device channel lengths in future technologies are likely to worsen this problem due to increased channel fields and smaller t_{ox} and x_j.

The nature of current injection due to capacitive coupling and avalanche induced currents is different because hot-electron induced currents are always injected into the substrate. In a switching CMOS inverter, hot-electron induced current will be injected into the substrate during both the 0-1 and 1-0 transitions, while the capacitive component of the current will reverse direction during the two edges. As a consequence, hot-electron induced currents will possess large even-harmonics of the fundamental switching frequency and a DC component, while the capacitive currents will possess large odd-harmonics, and no DC component. The presence of a DC component in any substrate current can be potentially very harmful to circuit operation. In addition to causing a drift in threshold voltages, it can also lead to an increase in minority-carrier injection into the substrate due to partial forward-biasing of device-to-substrate junctions.

For small-signal analysis, the effect of the hot-electron induced current can be modeled as a drain-to-body transconductance g_{db} [2] given by

$$g_{db} = \frac{\partial I_{dub}}{\partial V_D} = \frac{K_2 I_{sub}}{(V_{ds} - V_{dsat})^2} \tag{4}$$

The major effect of this parameter on small-signal circuit analysis is that this term appears in parallel with

the r_o of the device and tends to lower the output impedance of the transistor.

Hot-electron induced substrate currents in PMOS devices are considerably smaller than in comparably sized NMOS devices due to a lower hole ionization-coefficient. Further, PMOS devices in the process shown here are built in a locally grounded well. Thus it may be expected that PMOS devices cause lower substrate bounce than comparably sized NMOS devices. This is indeed the case as long as the n-well has a very good ground contact. If the well potential is allowed to vary with respect to the substrate potential, the well acts as a large injector, with a large reverse-biased well-to-substrate capacitance and can cause significant substrate noise injection.

In addition to the specific forms of substrate injection present in the active devices discussed above and in Section 2.0.1, the reverse-biased pn junctions formed by these devices with the substrates also exhibit a steady DC leakage current. This current consists of carriers which are swept across the depletion barrier in the direction of the electric field. Electrons are injected into the n-region and holes into the p-region under the action of the field. Hence the substrate current induced by this mechanism is a majority-carrier drift current.

2.0.3: Current Injection Mechanisms in Passive On-Chip Components

The passive components in typical processes are shown in Fig. 2f-k. These include resistors, capacitors, inductors and local diffusions.

Resistors in modern processes are either poly-type or diffused. Poly resistors have a comparatively smaller parasitic capacitance to the substrate. Thus diffusion resistors inject more noise into the substrate than poly resistors for the same dimensions. If one end of the resistance is connected to an AC ground then the current injected into the substrate at low-frequencies, due to a voltage V_{in} applied at the other end is given by

$$I = \sqrt{\frac{j\omega C}{R}}\tanh\left(\frac{\sqrt{j\omega RCL}}{2}\right)V_{in} \qquad (5)$$

where C is the per unit length capacitance of the resistor, R is the per unit length resistance and L is the length of the resistance. This equation predicts a $f^{0.5}$ dependence of the injected substrate current at high

frequencies. This formula is obtained by modeling the resistor as a dissipative transmission line.

Capacitors can be either poly-to-poly, metal-to-poly or poly-to-substrate types. Metal-to-metal capacitors have the largest ratio of the parasitic capacitance to the substrate for a given capacitance. Hence if these devices are used for implementing large on-chip capacitors, they can act as significant substrate noise-injectors.

On-chip inductors and interconnects inject noise into the substrate through the parasitic oxide capacitance with the substrate. The substrate parasitic can lead to lowering of the inductor quality factor. Thus the substrate loss must be modeled to obtain an accurate prediction of inductor performance.

Local diffusions in the substrate can be p or n-type. N-type diffusions inject noise through a reverse bias capacitance. P-type diffusions are often used as substrate taps or guard rings. They serve to tie down the substrate to a desired potential. If designed improperly these diffusions can inject very high levels of noise into the substrate, as they act as wide ground-planes on the substrate and any voltage bounce on these diffusions is conveyed throughout their extent on the chip through a very low impedance path. Guard rings will be discussed in detail in Chapter 4.

2.1: Reception of Substrate Noise

The reception of noise by most devices on the surface takes place through capacitive sensing. This is true of bipolar transistors, capacitors, resistors and interconnect lines.

The junction with the substrate in lateral pnp devices is formed by the n-type base region. If the pnp device is used in a gain stage, then the base of the device must be carefully shielded, or connected to a low impedance node. Otherwise the substrate noise will be amplified by the gain of the circuit.

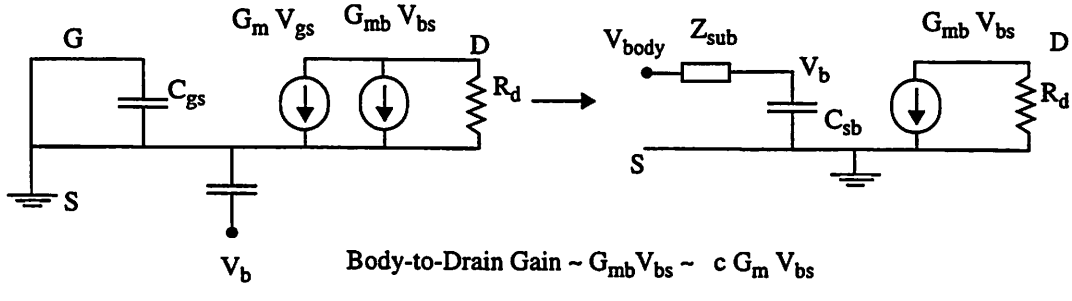


Figure 3: Body Effect in MOSFETs

In addition to capacitive pickup through the source and drain depletion junctions, MOS devices also exhibit a more severe form of substrate interaction due to the body effect. The threshold voltage of an MOS transistor is a strong function of the substrate potential. For a uniform surface impurity concentration N_A, the dependence of the threshold voltage is given by [2]

$$V_t = V_{t0} + \frac{\sqrt{2q\varepsilon N_A}}{C_{ox}}\left(\sqrt{2\phi_f + V_{SB}} - \sqrt{2\phi_f}\right) \tag{6}$$

where V_t is the threshold voltage, ε is the silicon dielectric permittivity, C_{ox} is the per unit area oxide capacitance, $2\phi_f$ is the surface inversion potential and V_{SB} is the source-to-body potential. The body effect makes the drain current dependent on the substrate potential. This effect can be represented by a linearized model parameter g_{mb} in the small signal device model. The drain current depends on the body-to-source potential as $g_{mb}v_{bs}$. As can be seen from Fig. 3, with the gate and the source shorted, a gain stage exists between the substrate to the drain-node of the MOS transistor. By making suitable approximations it can be shown that [2]

$$\frac{g_{mb}}{g_m} = \frac{\sqrt{2q\varepsilon N_A}}{2C_{ox}\sqrt{2\phi_f + V_{SB}}} \tag{7}$$

where g_m is the small-signal transconductance of the device. The parameter g_m relates the drain current to the gate-to-source voltage. In typical processes the above ratio varies from 0.1 to 0.3. The parasitic body-to-drain gain is thus lower than the gate-to-drain gain by a factor of 14-20 dB only.

The body effect in MOSFETs makes these devices especially vulnerable to substrate noise reception. While the capacitive pickup, exhibited by most other devices, becomes significant only at relatively high frequencies (above a MHz), the body effect can be an issue at low frequencies.

2.2: Transmission of Substrate Noise

Substrates act as the media for coupling of noise from one device to another. Thus in order to understand the phenomenon of substrate coupling, it is essential to determine the dominant mechanisms for current flow in the substrates and their dependence on the various substrate material parameters. Efficient models for the substrate, which use this information, are developed in the next chapter.

2.2.0: Simple Substrate Models

If the substrate is modeled as a lossy dielectric, then the distributed form of Ohm's law shown below is applicable to the substrate.

$$J = (\sigma + j\omega\varepsilon)E \tag{8}$$

where J is the current density in the substrate (A/cm^2), E is the electric field (V/cm), σ is the conductivity and ε is the dielectric permittivity of silicon. For a first order calculation of the substrate impedance, using typical material resistivities (1mΩ-cm to 20Ω-cm), it may be adequate to consider the substrate to be simply a distributed resistance. This is certainly true for frequencies up to 2GHz, since the susceptance term '$\omega\varepsilon$' is approximately 75Ω-cm at this frequency. The conductivity σ depends on the carrier concentration 'p' and the mobility 'μ_p' as $\sigma=pq\mu_p$ and hence is a function of the doping profiles in the substrate.

The above model considers current flow only due to drift (field induced) currents. This model would be sufficient for low-level majority-carrier conduction.

The flow of minority-carriers is considerably more complex than the simple majority-carrier case considered above. The motion of the excess carriers is dictated by the following transport equations.

$$\frac{\partial n_1}{\partial t} = D_n \nabla^2 n_1 + \nabla \bullet J_n - \frac{n_1}{\tau} \tag{9a}$$

$$\frac{\partial p_1}{\partial t} = D_p \nabla^2 p_1 - \nabla \bullet J_p - \frac{p_1}{\tau} \tag{9b}$$

D_p and D_n are the diffusion constants, J_n and J_p are the drift current terms ($=\sigma_{n,p}E$), τ is the carrier recombination time and p_1 and n_1 are the excess carrier concentrations over the thermal equilibrium values.

Minority-carriers, once injected into the substrate, can exist for long periods of time (carrier lifetime) and cause significant local variations in conductivity. However a large injection of minority-carriers into the substrate usually indicates a fault condition, as this occurs when a device-to-substrate junction is turned on. Hence to model substrate cross-talk we only consider the drift-induced substrate currents given by (8).

Computation of minority-carrier flow in the substrate requires the solution of the time-dependent-

diffusion-equations (9a and 9b). This can be computationally expensive even for a single device. The solution of (9a) and (9b) in large-scale substrate simulations is not currently feasible due to computational constraints.

2.2.1: High-Frequency Characterization of Substrates

At frequencies above 4-5 GHz, the error in the electrostatic assumption made above may not be tolerable. At these frequencies a correct model of the substrate can be obtained only by solving Maxwell's equations in the substrates. The problem is made more complex by the fact that coupling through the substrate may no longer be the dominant parasitic. Considerable energy flow takes place through the oxide layers and even through the package. Thus at these frequencies it is insufficient to consider the substrate as a distinct medium for energy flow. However, it can be expected that the results of the simple first-order calculations will provide good indicators for design even at frequencies where some of the assumptions of the model begin to fail. For the purpose of fast parasitic extraction, it often becomes necessary to sacrifice

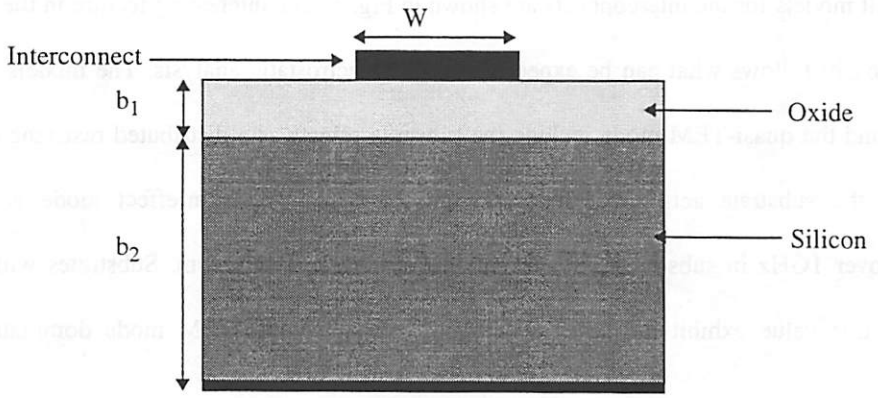
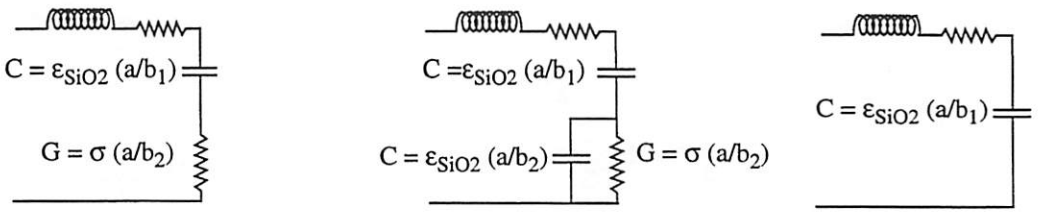


Figure 4: Interconnect in an IC



a. Slow-wave mode b. Quasi-dielectric mode c. Skin-effect mode

Figure 5: Transmission Modes in an Interconnect Line

accuracy by ignoring some of the more detailed features of the governing equations.

In a related problem, the behavior of an interconnect over the SiO_2-Si system (Fig. 4) is considered. This problem is not the same as the substrate problem considered here but sheds light on the high-frequency behavior of the substrate.

Depending on the frequency of the signal, three different modes can be observed in the substrate [5]. These modes are the quasi-TEM modes, the slow-wave modes and the skin-effect modes. The quasi-TEM modes exist when the susceptibility of the silicon layer is much higher than its conductivity. The SiO_2-Si multilayered media behave like a multilayered dielectric. The skin-effect modes also appear at high frequencies, when the vertical dimensions of the Si layers become comparable to the skin depth. At frequencies where these modes are not dominant, the substrate can be modeled as a distributed resistance. In [5] the authors consider a wide interconnect line to simplify the analysis, and use the transverse resonance technique to analyze the fundamental TM mode.

The circuit models for the interconnects are shown in Fig. 5. The interesting feature in the model is that the model closely follows what can be expected from an electrostatic analysis. The models for the slow-wave mode and the quasi-TEM mode include the substrate effects as a distributed resistance. In the skin-effect mode the substrate acts as a thin lossy ground-plane. The skin-effect mode is dominant at frequencies over 1GHz in substrates with resistivities less than 0.01 Ω–cm. Substrates with resistivities greater than this value exhibit the other two modes, with the quasi-TEM mode dominating at higher frequencies.

Chapter 3: Modeling of Substrate Coupling in Integrated Circuits

Efficient techniques are presented in this chapter for modeling the effects of the finite and distributed substrate impedance in circuit simulators. It is desirable to predict the degree of coupling in any circuit at the design stage. Further, as will be discussed later, substrate coupling is closely related to circuit layout. In many design tasks layout optimization may be required to minimize the substrate parasitic. Thus a substrate modeling tool should ideally be efficient enough to be used in optimization tasks.

Certain approximations are necessary in most analyses in order to obtain reasonable computation time or to reduce memory requirements on the computer. Trade-offs between accuracy and simulation time are inevitable, as low computation times are usually achieved by ignoring some of the second-order effects in the simulation. Some of the approximations made in the substrate simulator are discussed in the first section of the chapter.

Two simulation techniques are presented in the subsequent sections. The first is a numerical technique and the second uses an integral-equation formulation for extracting the substrate model. In both the techniques a lumped equivalent macromodel representing the substrate is extracted by solving the differential equations representing the medium. The lumped macromodel relates the voltage and the current vectors at the substrate contacts.

A comparison of the techniques in terms of memory, computation time and accuracy is made at the end of this chapter. It is shown in this section that the Integral-Equation technique is preferable since the substrate simulator should be suitable to be used in layout optimization. In fact this technique can be significantly faster than the purely numerical (finite-difference) based methods and is chosen as the basis for the simulator applications shown in the next chapter.

3.0: Approximations made in the Analysis

Several approximations must be made in order to extract the macromodel in a reasonable amount of time. The approximations involved, their validity and the point at which the approximations fail are discussed below.

3.0.1: The Electrostatic Assumption

The first approximation involved is that of considering only the electrostatic limit of the Maxwell's equations. This approximation is accurate at low frequencies but it fails at frequencies above 4-5GHz in typical silicon substrates as discussed in the previous chapter. This approximation is valid in integrated circuit substrates because the dimensions of the substrates are typically much smaller than the smallest electrical wavelengths. Thus distributed effects are not observed in the substrates.

Typically the first deviations from the electrostatic model will occur when the vertical dimensions of the substrates become comparable to the skin depth in the medium. If the substrate conductivity is much larger than the susceptibility, then the electrostatic model of the substrate simply involves resistors. The skin-effect makes the resistance between two contacts on the surface frequency-dependent and also leads to the appearance of a significant inductive reactance term in the contact-to-contact impedance. An example of this behavior in circular metal wires is shown in [6]. At very low frequencies the current is almost uniformly distributed over the cross-section of the wire and the resistive and the reactive parts of the per unit length impedance are given by

$$R_{lf} = \frac{\rho}{\pi r^2} \quad X_{lf} = \frac{\omega\mu}{8\pi} \tag{10}$$

and at very high frequencies the impedance terms are given by

$$R_{hf} = X_{hf} = \left(\sqrt{\frac{\omega\mu\rho}{2}}\right) / (2\pi r) \tag{11}$$

where ω is the frequency in rad-s^{-1}, ρ is the resistivity, μ is the magnetic permeability and r is the radius of the wire.

The low-frequency values accurately represent the impedance until the frequency at which the radius of the wire equals the skin depth. Beyond this frequency, the resistive term tends to increase. The self-inductive term increases as $f^{0.5}$, which implies that the inductance of the wire decreases. This decrease is due to constriction of the current towards the circumference of the wire. The key inference is that the DC formulas suffice for first-order impedance estimates at most frequencies of interest.

The computational simplification achieved from the electrostatic assumption is enormous. Solution of

the Maxwell equations in the substrate involves the solution of two inhomogeneous wave equations [7]. In the electrostatic approximation the scalar potential in the substrate satisfies the Laplace equation, and very efficient numerical techniques for parasitic extraction can be developed.

The electrostatic assumption also simplifies the boundary conditions to be used in the problem. The simplest boundary conditions in the substrate shown in Fig. 1 would be Neumann (zero normal E-field) on the top and the sides of the substrate, and Dirichlet (zero potential) on the backplane to model the header metal present in most IC packages. These boundary conditions tend to fail at very high frequencies when the package and the oxide isolation layers become significant paths for energy flow. At even higher frequencies the free-space boundary conditions may be required, as there could be radiation from the substrate. As may be expected, the solution of the Maxwell's equations for a complex structure as the IC substrate is far from trivial and may even be impossible on present day workstations due to computation-time and memory constraints.

3.0.2: The Assumption of Linearity

The conductivity of the silicon substrate is dependent on the electric field in the substrate. Thus the current-field relationship is nonlinear in silicon. This effect becomes significant at high-fields or at high current densities. The injection of minority-carriers can also make the conductivity time, location and field dependent. As discussed in the previous chapter minority-carrier leakage is avoided by reverse-biasing the device-substrate junctions.

The conductivity of the silicon layers is assumed to be a constant, independent of the field. It is also assumed to be isotropic. Several nonlinear effects are considered in device simulators. These tools, however, are completely unsuitable for parasitic extraction due to the large computation times involved in even one or two device problems.

3.0.3: The Assumption of Equipotentials

The device-to-substrate junctions are treated as equipotential contacts with the surface. A more accurate model of the junction would consider the device-to-substrate junction as a depleted semiconductor region (a dielectric). This model is difficult to implement in a fast substrate noise simulator,

since the extent of the depletion region depends on the voltage across the junction.

The magnitude of the error from this approximation is reduced considerably due to the small dimensions of the devices compared to the substrate. Thus, while the instantaneous variations of the voltage across the reverse-biased junction may change the value of the depletion capacitance considerably, the change in the value of the substrate model impedance values is expected to be small (Fig. 6). The nonlinear behavior of the junction capacitors is included in circuit simulators such as SPICE.

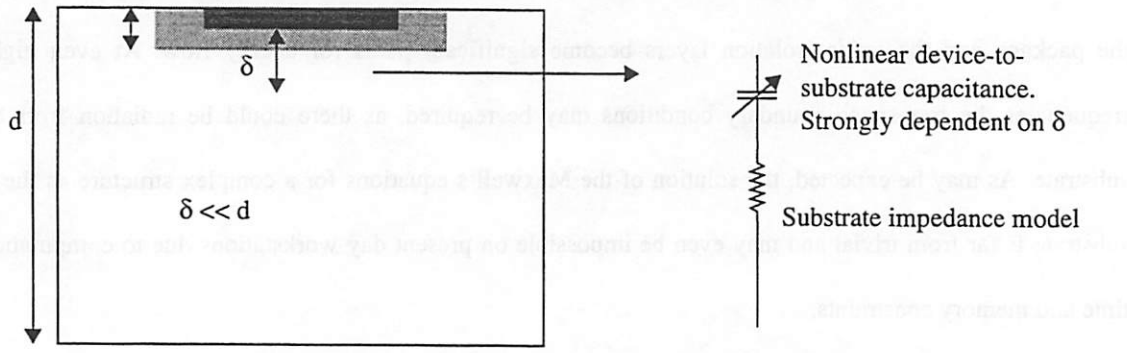


Figure 6: Bias Dependence of the Substrate Impedance

3.1: A Numerical Technique for Extraction of Substrate Models

In the electrostatic case, extraction of the substrate macromodel requires the solution of the Laplace equation in the substrate. Solution of partial differential equations by the use of numerical techniques has been studied extensively and has been presented by several authors ([8], [9]). These techniques usually involve approximating the differential equations of the system by difference equations. The resulting matrix is then solved using a method which is appropriate for the matrix size involved. Different iteration techniques, suitable for solution of the matrices in this case will be discussed in this section.

3.1.1: The Laplace Equation and Kirchoff's Laws

In the electrostatic case, the potential ϕ in the medium can be obtained by the solution of the Laplace equation subject to the given boundary conditions.

$$\nabla^2\varphi = 0 \qquad (12)$$

The boundary conditions may be of the Dirichlet (potential specified), Neumann (field specified) or the mixed types, where the potential is specified over part of the boundary and the field over the rest.

In the rectangular coordinate system, the Laplace equation is given by

$$\frac{\partial^2 \varphi}{\partial x^2} + \frac{\partial^2 \varphi}{\partial y^2} + \frac{\partial^2 \varphi}{\partial z^2} = 0 \tag{13}$$

Since substrate geometries conform to this coordinate system, the Laplace equation will be considered in this form. Rewriting the second order derivative in x in the difference form one obtains

$$\frac{\partial^2 \varphi}{\partial x^2} = \frac{\varphi_{i+1,j,k} + \varphi_{i-1,j,k} - 2\varphi_{i,j,k}}{\Delta_i^2} \tag{14}$$

and similarly for the other partial derivative terms. Δ_i is the step size in the X-direction. $\phi_{i,j,k}$ is the potential at the (i, j, k) point on the grid.

As can be expected, Kirchoff's laws for linear circuit analysis provide the above relation too, if the substrate is idealized as a linearized resistive mesh as shown in Fig. 7 below.

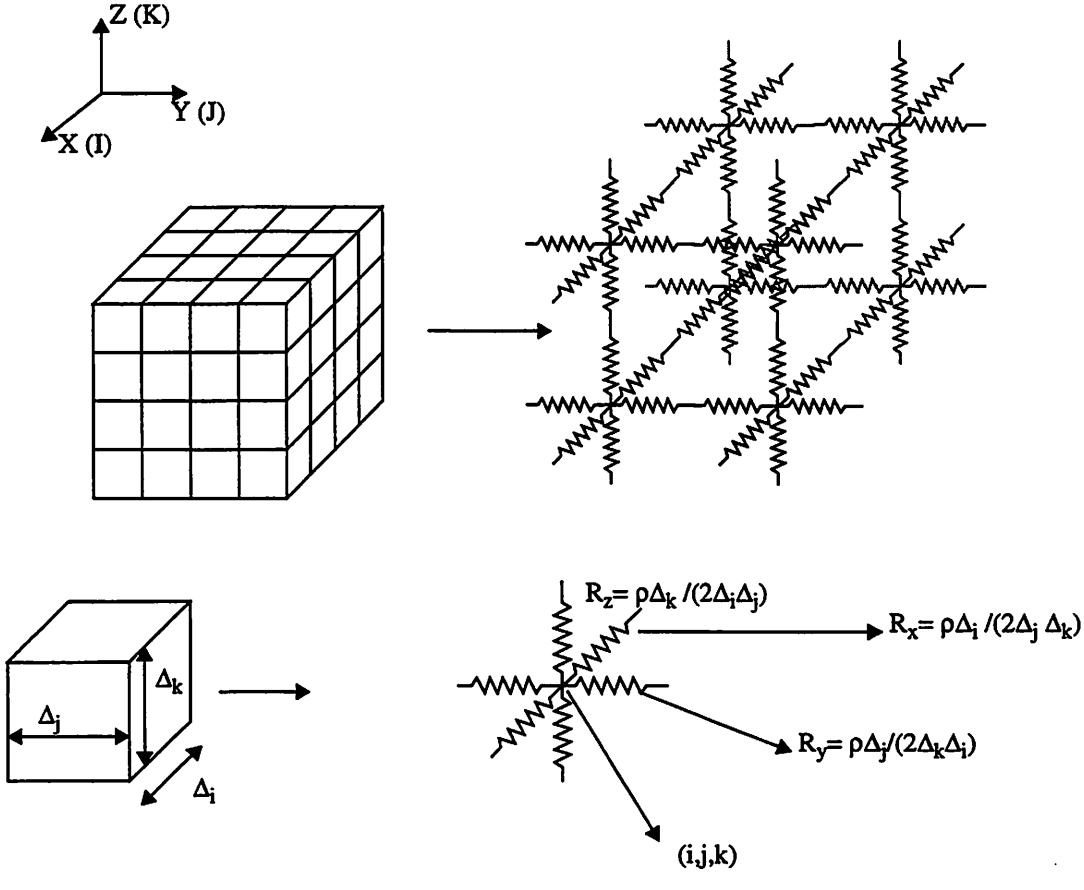


Figure 7: Representation of the Substrate by a Mesh

Each of the faces of the subdivisions of the substrate is considered to be equipotential. If the currents at node (i, j, k) are summed in the grid shown in Fig. 7, then the equation for the grid voltages reads

$$2\Delta_j\Delta_k\frac{\varphi_{i,j,k}-\varphi_{i+1,j,k}}{\rho\Delta_i}+2\Delta_j\Delta_k\frac{\varphi_{i,j,k}-\varphi_{i-1,j,k}}{\rho\Delta_i}+2\Delta_k\Delta_i\frac{\varphi_{i,j,k}-\varphi_{i,j+1,k}}{\rho\Delta_j}+2\Delta_k\Delta_i\frac{\varphi_{i,j,k}-\varphi_{i,j-1,k}}{\rho\Delta_j}+$$
$$2\Delta_i\Delta_j\frac{\varphi_{i,j,k}-\varphi_{i,j,k+1}}{\rho\Delta_k}+2\Delta_i\Delta_j\frac{\varphi_{i,j,k}-\varphi_{i,j,k-1}}{\rho\Delta_k}=0 \qquad (15)$$

Dividing throughout by $2\Delta_i\Delta_j\Delta_k/\rho$, we get the Laplace equation shown earlier. Rewriting the above equation in terms of the potential ϕ_{ijk} we get

$$-\frac{\varphi_{i-1,j,k}}{R_x}-\frac{\varphi_{i,j-1,k}}{R_y}-\frac{\varphi_{i,j,k-1}}{R_z}+\left(\frac{2}{R_x}+\frac{2}{R_y}+\frac{2}{R_z}\right)\varphi_{i,j,k}-\frac{\varphi_{i+1,j,k}}{R_x}-\frac{\varphi_{i,j+1,k}}{R_y}-\frac{\varphi_{i,j,k+1}}{R_z}=0 \qquad (16)$$

The Kirchoff's law is solved in the substrate to reduce the distributed resistance of the substrate to lumped equivalents. The surface contacts are considered to be equipotentials as discussed in Section 3.0.3. Thus the model for the three-contact problem shown in Fig. 8a, in the presence of a backplane, consists of six resistors as shown in Fig. 8b. In order to extract the individual resistors shown in Fig. 8b, one contact is set at a unit potential, while the others are at ground. The inverse of the current exiting at any of the grounded contacts yields the resistance from the contact at unit potential to that contact.

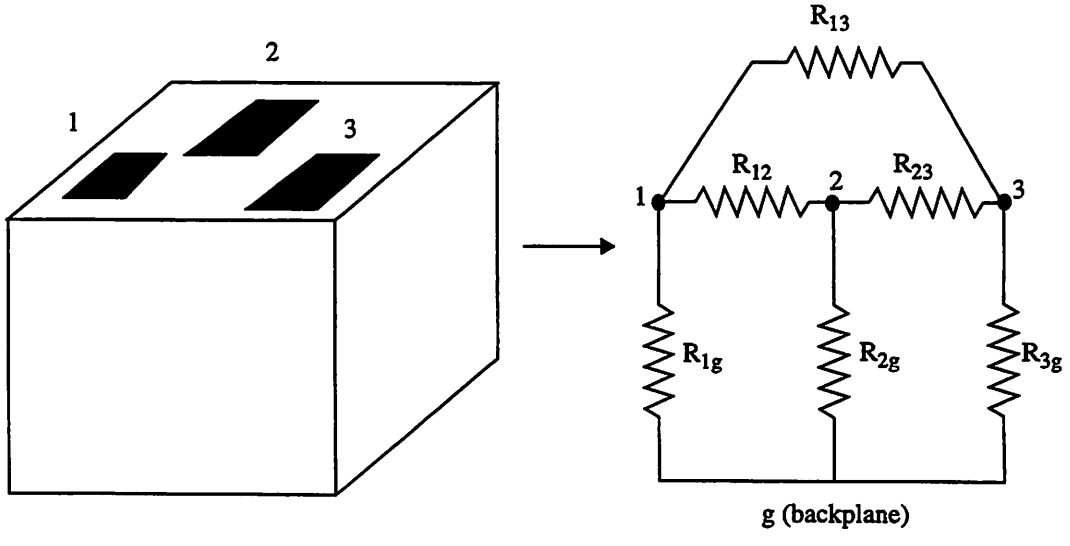


Figure 8a: Three-Contact Substrate Problem

Figure 8b: Model of the Substrate Problem

The steps involved in the extraction of the substrate model can be summarized as follows:

1] Determine a satisfactory grid representing the substrate as in Fig. 7.

2] Set up the node voltage equations as in (16).

3] Set one contact at 1V and all others at zero.

4] Solve the system of equations for the node voltages.

5] Determine the currents exiting each contact at ground potential and calculate the resistance from contact -to-contact.

The node voltage equations can be written in a matrix form

$$[A][\varphi] = [b] \tag{17}$$

where the elements of [A] are the conductances associated with the nodes and [b] is the source voltage matrix. The nodes of the grid which lie in the region of the surface contact, form the source matrix [b]. [ϕ] is the matrix of the node voltages.The matrix [A] is sparse. In fact, no more than seven elements in any row can be nonzero. It is also diagonally-dominant, symmetric and positive definite.

Determining a suitable grid for the substrate is a difficult problem. If the number of the surface contacts is large or the dimensions of the contacts are small, then using a uniform grid may not be practical. For example, if the minimum contact size is 5μm on a 1000*1000*500μm^3 substrate, the number of grid points, assuming a uniform grid, will be at least 200*200*100 that is four million. This can lead to prohibitive matrix sizes, depending on the technique used for the matrix solution.

A common strategy to deal with this problem is to make the grid dense in the areas where the variation in the potential is expected to be large and make the grid coarse where the potential variation is small. For example, the grid may be made dense around sharp corners, or in the vicinity of contacts, and coarse in regions away from the contacts.

This scheme helps reduce the size of the matrices involved. The problem with the technique is that it is difficult to determine beforehand the regions where the field is high. Some heuristic guidelines based on distance to the contacts etc. may be developed and implemented in the simulation tool. Another approach is to use progressively denser grids until the error is tolerable. This may not always be a good solution

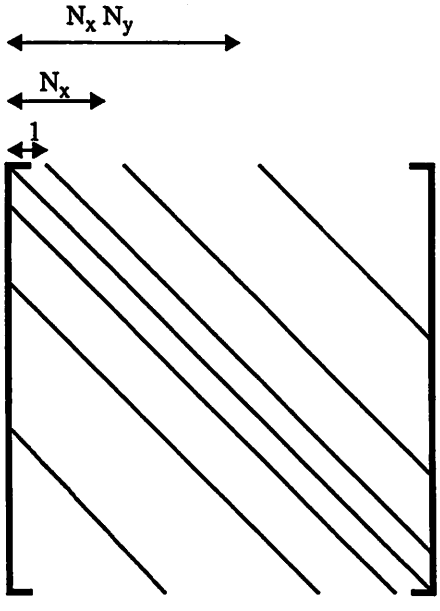
since it may require considerable computation too.

The next step in the process is the solution of (17). This is the computationally intensive step in the problem. Several algorithms exist for the solution of sparse linear systems of equations. Some of these techniques will be discussed next.

3.1.2: Direct Techniques for Solution of the Linear System

Direct techniques for solution of a linear system of equations [A][X]=[b] include Gaussian elimination and LU factorization. These techniques are unsuitable for solving the substrate problem. Although substrate conductance matrices are sparse as stated earlier, the matrices can be of a very large order and the solution techniques used must be able to exploit sparsity in order to avoid excessive demands on the computer storage.

If the grid used to represent the substrate has N_x, N_y and N_z points in the X, Y, Z directions respectively, and the grid is numbered in a (X, Y, Z) sequence, then the matrix A has the structure shown below.



The bandwidth of the matrix is $2N_xN_y$. If LU factorization is used on this matrix, [L] and [U] will both have a bandwidth of N_xN_y with all the upper entries of L and all the lower entries of U equal to zero. However LU factorization will not preserve sparsity within the band. The total storage requirement will grow from approximately $7N_xN_yN_z$ to $(N_xN_y)N_xN_yN_z$, which may be prohibitive. In such cases, the

iterative techniques presented below are found to be more suitable.

3.1.3: Simple Iterative Techniques

Iterative techniques used in the solution of sparse matrices have been discussed extensively in the past [9]. Iterative techniques have the advantage that they work with the system matrix, without modifying the matrix significantly. If the matrix is sparse to begin with, the sparsity is preserved throughout the solution procedure. Iterative techniques generally start with an initial approximation to the solution vector. The technique converges towards the solution vector, through a sequence of iterative steps. The convergence is asymptotic and an infinite number of iterations are required to converge to the solution vector. In practice, however, a fairly good approximation to the correct solution can be achieved by truncating the iteration once a pre-decided error threshold is reached.

In this section, four iterative techniques are presented. These techniques typify the solution strategies employed in a wide cross-section of iterative schemes. The convergence analysis for these and other schemes have been presented in [8], [9] and [10].

Jacobi's Method and The Gauss-Siedel Relaxation Scheme

The simplest iterative technique for solution of [A][X]=[b] is the Jacobi's method. In this method a sequence of vectors approximating the solution vector are obtained as follows. The i-th unknown x_i can be rewritten in terms of a_{ij} and b_i as

$$x_i = \frac{-\sum_{j=1, j\neq i}^{N} a_{ij}x_j + b_i}{a_{ii}}; (a_{ii} \neq 0) \tag{18}$$

If $[X]^0$ is a first approximation to the solution vector [X], then the elements of the (k+1)-th approximation vector $[X]^{k+1}$ can be written in terms of the previous vectors as

$$x_i^{k+1} = \frac{-\sum_{j=1, j\neq i}^{N} a_{ij}x_j^k + b_i}{a_{ii}} \tag{19}$$

If $[X]^k$ -> [X] as $k \rightarrow \infty$, then [X] is the solution vector of the above system of equations. This technique is slow to converge and other techniques are usually preferred for faster convergence.

If the matrix A is split as A= D(L+I+U) where [D] is a diagonal matrix with $[D]_{ii} = a_{ii}$ then the above equation can be rewritten as

$$[X]^{k+1} = -(L+U)[X]^{k} + D^{-1}[b] \tag{20}$$

The matrix (L+U) is referred to as B and is known as the iteration matrix of the technique. Each iterative method has a unique iteration matrix. If λ_m is the eigenvalue of B with the largest modulus (the spectral-radius), then it can be shown that after k-iterations, where k is large,

$$[X]^{k} - [X] \propto \lambda_m{}^{k} \tag{21}$$

If k_1 more iterations are required to reduce the amplitude of the dominant eigenvector by a factor 10^{-n}, then we must have

$$|\lambda_m|^{k_1} \leq 10^{-n} \tag{22}$$

or

$$k_1 \geq -\frac{n}{\log(|\lambda_m|)} \tag{23}$$

The term in the denominator is referred to as the asymptotic rate-of-convergence of the technique. The asymptotic rate-of-convergence can be determined for several iterative schemes. The closer the spectral-radius is to the unit-circle in the complex plane, the slower is the rate of convergence. While the asymptotic rate-of-convergence is a good figure of merit for comparing different techniques, it should be used with caution [9]. Other convergence estimators are discussed in [8].

The spectral-radius is difficult to estimate in most problems. An estimate for the spectral-radius for the solution of the Laplace equation in two-dimensions with Dirichlet boundary conditions is presented in [8] for Jacobi's method. The grid considered there is uniform in the X and the Y directions with I and J divisions in the two dimensions respectively. For this problem, for large I and J the spectral-radius varies as

$$\lambda_m = 1 - \frac{1}{4}\left(\frac{1}{I^2} + \frac{1}{J^2}\right)\pi^2 \tag{24}$$

The asymptotic convergence rate is given by

$$R = \frac{1}{4}\left(\frac{1}{I^2} + \frac{1}{J^2}\right)\pi^2 \tag{25}$$

The spectral-radius rapidly approaches unity and the convergence rate tends to zero, as the minimum grid spacing becomes smaller.

The Gauss-Siedel method is obtained from the Jacobi's method by using the most updated vector values that is

$$x_i^{k+1} = \frac{-\sum_{j=1}^{j<i} a_{ij}x_j^{k+1} - \sum_{j=i+1}^{j\le N} a_{ij}x_j^{k+1} + b_i}{a_{ii}} . \tag{26}$$

The (k+1)-th update can be expressed in terms of the D, L and U matrices as

$$[X]^{k+1} = -L[X]^{k+1} - U[X]^k + D^{-1}[b] \tag{27}$$

which implies

$$[X]^{k+1} = -(I+L)^{-1}U[X]^k + (I+L)^{-1}\left(D^{-1}[b]\right) \tag{28}$$

Thus the iteration matrix for this method is $(I+L)^{-1}U$. The Gauss-Siedel method has twice the rate of convergence of the Jacobi method for the two-dimensional Laplace equation considered previously.

These methods are of academic interest as they are slow to converge. However these techniques are important as more efficient techniques can be derived from these by simple modifications. One such important class of techniques is presented next.

The Simultaneous-Over-Relaxation Method

This method is considerably faster than the previous two methods and is a variant the Gauss-Siedel scheme. If $[X']^{k+1}$ is the (k+1)-th vector estimate computed using the Gauss-Siedel scheme, then the (k+1)-th term for this method ($[X]^{k+1}$) is computed by the following formula

$$[X]^{k+1} = [X]^k + \beta\left([X']^{k+1} - [X]^k\right) \tag{29}$$

The method converges for β between 0 and 2. Values between 1 and 2 lead to an acceleration of the

technique. The convergence is the fastest at an optimum value of β, but it is very difficult to determine the precise value beforehand.

This technique was implemented for solving the substrate mesh equations. It was found that for typical problems with approximately one hundred thousand node points, the optimum value of the relaxation parameter was between 1.9 and 1.95.

The Conjugate Gradient Algorithm

The advantage of iterative algorithms over direct techniques such as Gaussian elimination lies in the fact that these techniques exploit sparsity well. A direct method with a similar property is the Conjugate Gradient algorithm [10]. This technique tries to find the solution of [A][x]=[b] by minimizing the function $(1/2)[x^T][A][x]-[b][x]$. In this technique, as in the iterative techniques discussed above, an approximate vector $[x^k]$ is iterated until it converges to [x]. The advantage of the technique is that it is bound to converge in at most N iterations where N is the size of matrix A.

Several additional techniques exist for efficient solution of sparse matrices. Many of these rely on determining a suitable iteration matrix for the problem. One such technique uses a combination of an iterative method and the Conjugate Gradient Algorithm discussed above [11]. The method is faster than SOR but requires more storage. This technique was used for solving the substrate problem recently [12].

The iterative schemes discussed above are the only practical techniques for solving the large matrices which result from the finite differences technique. However, despite all of the simplifying assumptions made earlier in Section 3.0, the solution of substrate problems with more than ten-to-twenty contacts becomes impractical as discussed later. Due to this reason the finite-differences technique was not used for substrate parasitic extraction.

3.1.4: Modeling the Frequency-Dependence of Substrates

The problem of the determining the frequency response of the substrate has been analyzed in [12]. Before discussing the technique presented there, the frequency response of two different types of substrates is analyzed. These are the uniform (single-resistivity), and the low-impedance substrates shown in Fig. 1b. For a uniform substrate, the frequency dependent model for the substrate can be determined

trivially, since the substrate has a single RC time constant. Therefore in these substrates, a model such as in Fig. 8b should be modified to include a parallel capacitance per resistance. The value of this capacitance is such that the RC product equals the dielectric relaxation time-constant of the substrate ($\rho\varepsilon$). The case of the low-impedance substrate can also be analyzed similarly, if the low-resistivity bulk is treated as a single node. In that case, the resistors in the model would again be replaced by an RC parallel combination such that the RC product equals the relaxation time constant of the high-resistivity epi-layer.

The substrate model behavior with respect to frequency is quite complex for more complicated substrate cross-sections, as the different layers of silicon have different time constants. A technique known as Asymptotic Waveform Evaluation (AWE) has recently been applied to the problem in [12]. AWE constructs an approximate representation of the frequency response of a circuit by selecting an appropriate low-order transfer function to model the circuit.

AWE determines the frequency response of a circuit by a two-step process. The first involves extracting the moments of the circuit and the second involves matching these moments to a low-order transfer function. The second process is done by use of the Pade approximation. A discussion of this technique is presented in [13].

Once a low order transfer function has been found for the circuit, the time response with an arbitrary time input such as a step or an exponential, may be found by multiplying the s-domain representation of the input with the transfer function and performing an inverse Laplace transform.

While this technique is elegant in its approach, the utility of its use in substrate modeling problems, as discussed in [12], is questionable. The principal reason is that with typical substrate doping levels, the substrate impedance in silicon is dominated by the conductance and not the susceptance up to 4-5GHz. For substrates with higher doping levels (or higher conductivity), this frequency is higher. However the appearance of the skin-effect modes in high-conductivity substrates at these frequencies complicates the modeling problem. This issue is not addressed in [12]. Thus for most substrate modeling tasks, the additional computational complexity of AWE over a simple electrostatic model may not be justifiable. The authors reach a similar conclusion in [14].

3.2: An Integral-Equation-based approach to Substrate Modeling

In this section a fast and efficient technique for substrate modeling problems is presented. This technique relies on the electrostatic Green function in the substrate medium and the Fast Fourier Transform algorithm. Green functions have been used extensively for capacitance extraction in dielectric media in the past. This section begins with a discussion of classical techniques for determining the potential and the fields in electrostatic problems. The method of conformal mapping is discussed at first. The Green function technique is discussed subsequently. Discussion of this technique is divided into two subsections. The first presents the concept of Green function as applied to electrostatic problems and demonstrates two schemes for determining the Green functions in the medium. These are the method of images and the separation-of-variables (SOV) techniques. The utility of the Green function for capacitance extraction is discussed next. The Green function for the electrostatic substrate model is presented next. A substrate parasitic extractor using this form of the Green function and the Fast Fourier Transform (FFT) algorithm is discussed in the subsequent section.

The following section discusses the matrix operations performed in the simulator. The Sherman-Morrison formula for matrix inversion can be applied to the substrate problem to handle optimization related tasks efficiently. Improvements to the technique are discussed at the end of this section.

3.2.1: Techniques for Solution of Electrostatic Problems

Electrostatic problems appear in many practical modeling tasks. Examples include capacitor estimation on PC boards, or derivation of the current-voltage relations in semiconductor devices. These problems involve the solution of the Laplace equation or the Poisson equation in the medium, subject to the given boundary conditions.

The differential form of the Gauss's law relates the potential function ϕ to the free charge density ρ through the following partial differential equation

$$\nabla \bullet \varepsilon (\nabla \phi) = -\rho \qquad (30)$$

The dielectric constant ε is a measure of the polarizability of the medium [7] and can be considered to be an isotropic quantity in most modeling problems. If, in addition, the material is homogeneous as well, the

above equation can be rewritten to obtain the Poisson equation shown on the next page.

$$\nabla^2\Phi = -\frac{\rho}{\varepsilon} \tag{31}$$

In addition to a knowledge of the dielectric constant of the medium and the free-charge density, the boundary conditions in the medium must be specified for a solution of the problem. These can be of three types. In the first type the potential is specified on the boundary (Dirichlet conditions), in the second case the field (gradient of the potential) is specified on the boundary (Neumann conditions) and in the third case potential is specified over part of the boundary and the field over the remaining portion (the mixed boundary conditions). The uniqueness theorem of electrostatics [7] assures the existence of a unique solution to the Poisson equation in the presence of these three types of boundary conditions. The fourth possibility, in which both the potential and its gradient is specified over the entire boundary, does not possess a unique solution in the case of closed boundaries.

In the absence of free charges the Poisson equation is transformed into the Laplace equation with the right hand side of (31) being set identically to zero. The problem central to electrostatics is to determine the solution of (31) given the charge density ρ.

3.2.1.1: Conformal Mapping

The conformal mapping technique is useful in two-dimensional electrostatic problems. Due to this limitation it has not been a popular modeling procedure, except in certain cases where the behavior of the potential function may be essentially considered two-dimensional even though the problem is three-dimensional, for example for calculating fringing fields in a parallel plate capacitor [6].

This technique involves determining an analytic function F of the complex variable z, which maps the complex Z plane (or the XY plane) onto another plane W such that W=F(z). Since F is analytic, the real and imaginary parts of W, say u and v respectively, must satisfy the Cauchy-Riemann conditions

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} \quad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x} \tag{32}$$

Consequently, u and v must satisfy the Laplace equation and can be used to represent the potential and field functions in the Z plane.

The analyticity of F also assures that the angle between any two intersecting curves in the XY plane will be preserved in the u-v plane as well. If F is chosen appropriately, it is possible to transform a problem involving uniform fields in the u-v plane, to the XY plane, such that the boundary conditions are satisfied in the XY plane. The angle preserving property of the transformation guarantees the orthogonality of the potential and field curves (i.e. between the constant u and constant v lines) in the XY plane as well.

The main problem in this technique is determining the function F with the necessary properties. Experience and intuition are often the guiding factors in this process. One class of transformations which has been found very useful in some microstrip related problems is the Schwarz transformations for polygons.

3.2.1.2: Integral Solutions for Potential

The image-based techniques and the SOV technique can both be related to the potential of a point charge in the medium. The potential of a point charge in a medium is known as the Green function of the medium. In this section, the concept of the Green functions as applicable to electrostatic problems is discussed. Green functions are useful as solution techniques in many problems such as full-electromagnetic calculations [15] or diffusion problems.

Consider two arbitrary scalar fields ϕ and ψ defined in a volume V bounded by a surface S. The divergence of the vector fields $\phi\nabla\psi$ and $\psi\nabla\phi$ is given by

$$\nabla\bullet(\phi\nabla\psi) = \phi\nabla^2\psi + \nabla\psi\bullet\nabla\phi \tag{33}$$

and

$$\nabla\bullet(\psi\nabla\phi) = \psi\nabla^2\phi + \nabla\psi\bullet\nabla\phi \tag{34}$$

respectively.

Subtracting (34) from (33) one has

$$\nabla\bullet(\phi\nabla\psi) - \nabla\bullet(\psi\nabla\phi) = \phi\nabla^2\psi - \psi\nabla^2\phi \tag{35}$$

Integrating both sides of (35) over volume V and applying the divergence theorem one has

$$\int_V (\phi\nabla^2\psi - \psi\nabla^2\phi)\, dv = \oint_S (\phi\nabla\psi - \psi\nabla\phi) \bullet \hat{n} ds = \oint_S \left(\phi\frac{\partial\psi}{\partial n} - \psi\frac{\partial\phi}{\partial n}\right) ds \tag{36}$$

where $\hat{n}$ is the unit outward normal vector to the surface S enclosing the volume V.

Let ϕ be the potential resulting from a localized charge density $\rho(r')$, and ψ be the potential due to a point charge placed at a point r'. The potential functions are related to the charge densities by the Poisson equation. Thus

$$\nabla^2\phi = -\frac{\rho}{\varepsilon} \tag{37}$$

and

$$\nabla^2\psi(r, r') = -\frac{\delta(r - r')}{\varepsilon} \tag{38}$$

Resubstituting the Poisson equation into (36) and using r' as the integration variable, one has

$$\phi(r) = \int_V \rho(r')\,\psi(r, r')\, d^3r' + \varepsilon\oint_S \left(\psi\frac{\partial\phi}{\partial n} - \phi\frac{\partial\psi}{\partial n}\right) ds' \tag{39}$$

$\phi(r)$ is the potential at point r due to the charge $\rho(r')$. $\psi(r, r')$ is the potential in the medium due to a point charge placed at r' and is known as the Green function. It depends on the coordinates of the observation point r and also the source point r'. The above is a convolution relationship, which relates the known density $\rho(r')$ to the potential function $\phi(r)$ by convolution with the Green function.

If the Green function is known, (39) provides a technique for determining the potential at any point in the volume V due to a known arbitrarily distributed charge density. The method of images and the SOV technique are two different approaches to evaluating the Green function.

In the absence of any boundaries, that is in the free-space case, the function $\psi(r, r')$ reduces to $1/(4\pi\varepsilon|\vec{r} - \vec{r}'|)$, where r is the observation point and r' is the source point. In this case, if the field term in the surface integral falls faster than r^{-1} at large distances from the source, then as the surface S is taken to infinity, the surface integral term vanishes and the potential function $\phi(r)$ is related to the charge density through the well known relation

$$\phi(r) = \int_V \frac{\rho(r')}{4\pi\varepsilon|\vec{r}-\vec{r}'|} d^3r' \qquad (40)$$

The function $\psi(r, r')$, in the presence of finite boundaries can be thought off as consisting of two parts, the first is the free-space Green function and the second is the potential due to charges outside the volume V, distributed so that the required boundary conditions are satisfied on S [7].

The Method of Images

The method of images is a simple technique to determine the Green function. In this method, point charges are distributed in space in order to achieve the required boundary conditions.

To illustrate this technique we consider a simple example where the Green function in the volume enclosed by two ground-planes inclined at an angle θ, is to be found. (Fig. 9)

If $2\pi/\theta$ is an integral value (e.g. six in Fig. 9), then the image charges for Dirichlet boundary conditions can be found by dividing the entire plane into $2\pi/\theta$ divisions and placing image charges as shown in Fig. 9. The potential at any point 'r' in the enclosed region can be found by summing the potential contributions due to each point charge. The interpretation of the Green function given in the last paragraph of the previous subsection can be clearly understood in this example. The charges outside of the volume enclosed by the planes A and B are placed so as to obtain the necessary boundary conditions on A and B.

The method of images can be extended to problems with dielectrics as well. Shown in Fig. 10 are the image charges required to represent the potential due to a charge Q inside the medium of dielectric

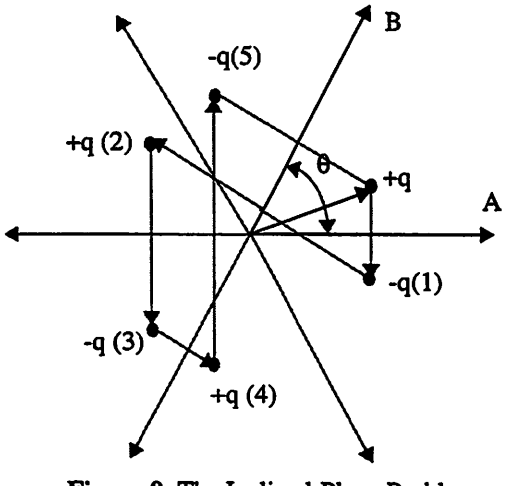


Figure 9: The Inclined-Plane Problem

constant ε_2 in the presence of a ground-plane. The image charges shown give the potential at any point inside the medium 2 by the formula

$$\Phi(r) = \frac{1}{4\pi\varepsilon_2} \sum_{n=1}^{\infty} \frac{q_n}{|r - r_n'|} \tag{41}$$

where r'_n is the location of the n-th image charge. The images are determined by first finding the image at the dielectric interface, next reflecting the images in the ground-plane and so forth.

In the presence of multiple dielectrics, the method of images leads to several nested infinite-series. In fact the method becomes unusable beyond three or four dielectrics, as the image charges in an N-layer medium consist of N-1 nested summations, with each summation index going to infinity. This problem is

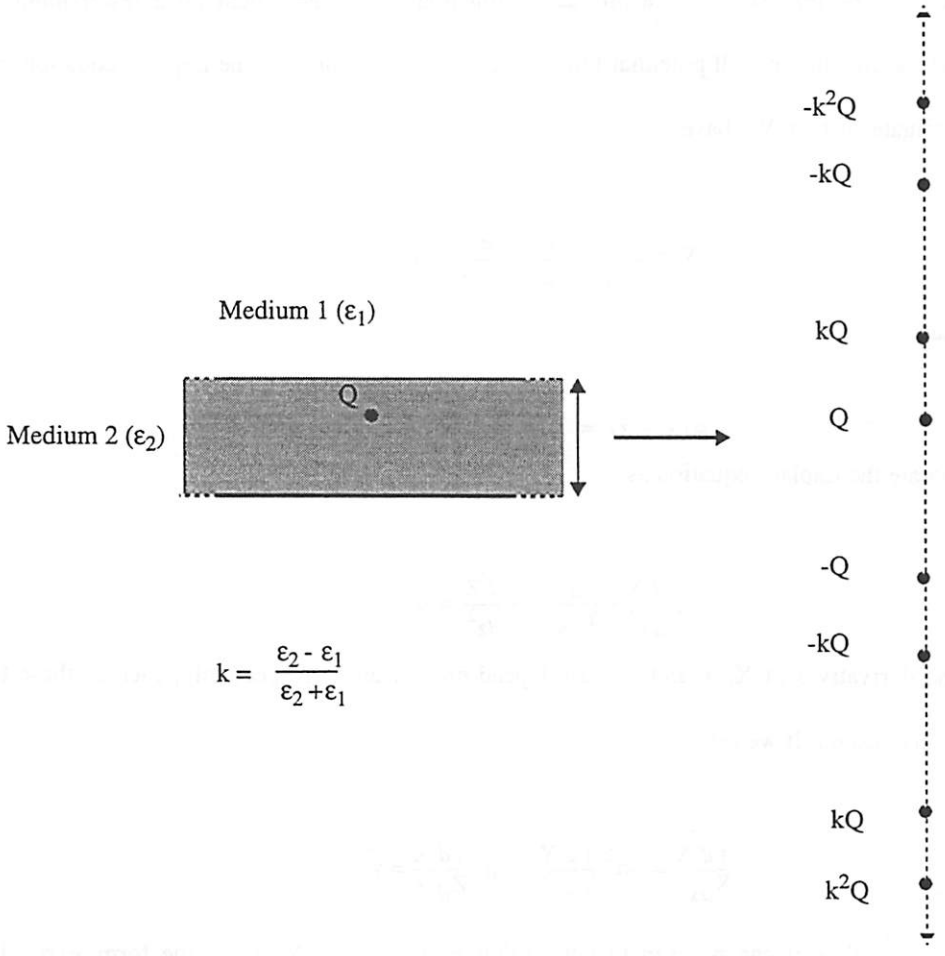


Figure 10: Two-Layer Dielectric Problem

worse in finite-boundary problems, as the images are required to match the boundary conditions on the side faces of the medium which adds more levels of nested summations.

Ref. [16] demonstrates a scheme to avoid the multiple levels of nesting. In this reference, the authors replace the surface of each dielectric layer with a charge sheet to model its effect. This technique however increases the size of the matrix relating the potential and the charge and requires higher inversion time.

The Separation-of-Variables Technique

The separation-of-variables technique is a classical tool of electromagnetic theory which has been utilized for analyzing problems long before the advent of modern digital computers. In this method, the partial differential equation representing the action of the Laplacian operator on the potential is reduced to ordinary differential equations in the coordinate system of choice. The technique relies on the assumption that the potential can be represented as a product of functions, each dependent on a fewer number of coordinate variables than the overall potential function. For example, consider the Laplace equation in the rectangular coordinate system. We have

$$\nabla^2\phi = \frac{\partial^2\phi}{\partial x^2}+\frac{\partial^2\phi}{\partial y^2}+\frac{\partial^2\phi}{\partial z^2} = 0 \tag{42}$$

If we assume that

$$\phi(x, y, z) = X(x)\,Y(y)\,Z(z) \tag{43}$$

then we can separate the Laplace equation as

$$\frac{1}{X}\frac{d^2X}{dx^2}+\frac{1}{Y}\frac{d^2Y}{dy^2}+\frac{1}{Z}\frac{d^2Z}{dz^2} = 0 \tag{44}$$

Since the second derivatives of X, Y and Z must depend on x, y and z respectively, each of these terms must be equal to a constant. If we set

$$\frac{1}{X}\frac{d^2X}{dx^2} = -\alpha^2;\frac{1}{Y}\frac{d^2Y}{dy^2} = -\beta^2;\frac{1}{Z}\frac{d^2Z}{dz^2} = \gamma^2 \tag{45}$$

where $\alpha^2+\beta^2 = \gamma^2$, then it can be seen that the solutions for X and Y are of the form $\exp(\pm i\alpha x)$ $\exp(\pm i\beta y)$ and that for Z is of the form $\exp(\pm\gamma z)$. Using this information and the knowledge of the

boundary conditions, it is possible to find a solution for ϕ from X, Y and Z.

The SOV technique in different coordinate systems is discussed in [7]. The potential in the cylindrical coordinate system can be expanded in terms of Bessel functions. In spherical coordinates, the potential can be expressed in terms of Legendre polynomials. These solutions are not discussed here. Substrate geometries in semiconductors conform to the rectangular system. Further, expansions based on Bessel functions or the Legendre polynomials are computationally time consuming, and these functions are not an attractive choice as basis functions.

In order to compute the potential due to a point charge in the rectangular coordinate system, one can resort to two techniques. The first technique involves integrating over the point source discontinuity, and matching potential and field boundary conditions in the plane of the point charge. The second technique involves expressing the potential as an eigenfunction summation [7]. The second method can be considerably slower to converge than the former and is not discussed. A discussion of the first technique is reserved until Section 3.2.3, in which the Green function is derived for the substrate problem using this technique.

3.2.2: Use of the Green Function for Parasitic Extraction

Once the Green function of the medium has been determined by using either of the techniques discussed above, it is possible to compute the potential resulting from any known charge distribution in the substrate. The use of the Green function for parasitic extraction is demonstrated by using the example in Fig. 11. Fig. 11 considers a single-layer dielectric of infinite lateral extent and a finite thickness 'd'. Two contacts are defined on the surface, and the equivalent capacitances between the two surface contacts are to be determined. It is assumed that the Green function G(r, r') is already known.

If the contacts are treated as equipotential boundaries on part of the surface, the problem would be a mixed boundary value problem, with the normal field specified on part of the surface and the potential on the rest. This class of problems possesses a unique solution to the Laplace equation. However, very few mixed-boundary value problems are amenable to direct solution techniques, and a purely theoretical solution often requires making approximations of treating certain quantities as perturbations etc. In many

cases, a rigorous solution may not even exist. Several mixed-boundary value problems are discussed in [17].

The Green function technique solves this problem by finding charge distributions on the contacts which provide an approximately constant potential on the surface of the contacts. A unit charge is distributed on a contact and the potential resulting from this charge distribution on the same contact and all other contacts is determined by means of the integral relation

$$\phi(r) = \int_V \rho(r')\, G(r, r')\, d^3r' + \varepsilon \oint_S \left(G \frac{\partial \phi}{\partial n} - \phi \frac{\partial G}{\partial n} \right) ds' \tag{46}$$

In the case of the mixed-boundary problem treated here or for Dirichlet boundaries the second term can be set to zero and the integral equation reduced to

$$\phi(r) = \int_V \rho(r')\, G(r, r')\, d^3r' \tag{47}$$

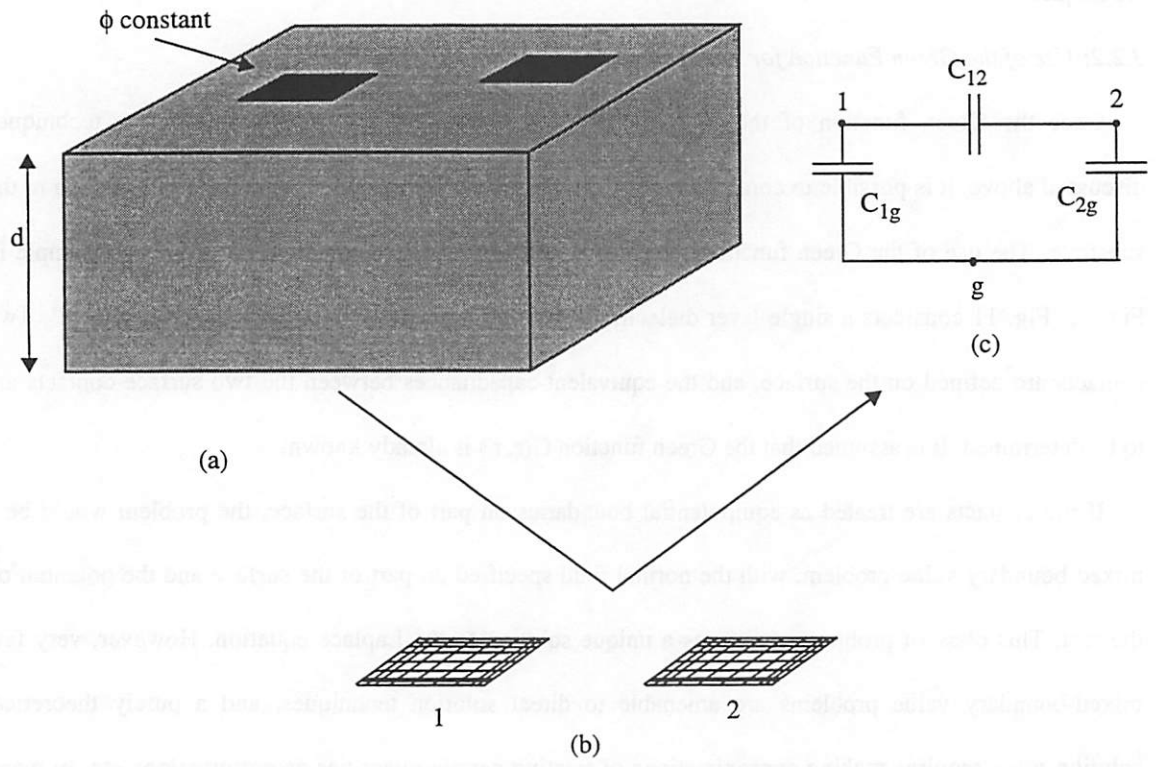


Figure 11: Green Function based Extraction of the Substrate Macromodel

This simplification is not possible in Neumann boundary problems [7].

The potential at any point of the contact can be taken as representative of the contact potential, or the average potential over the volume of the contact can be treated as the contact potential. If the second definition is chosen then the integral equation in (47) is replaced by

$$\bar{\phi}_i = \frac{1}{V_i} \int_{V_i} \int_{V_j} \rho G dv_j dv_i \tag{48}$$

where V_i and V_j are the volumes of contact-i and contact-j respectively and ρ is the charge distribution on contact-j. $\bar{\phi}_i$ is the average potential on contact-i induced by the charge distribution ρ on contact-j. If a uniform charge distribution is chosen over the volume of contact-j such that $\rho = Q_j / V_j$, then (48) becomes

$$\bar{\phi}_i = \frac{Q_j}{V_j V_i} \int_{V_i} \int_{V_j} G dv_j dv_i \tag{49}$$

By considering all combinations of contacts, in an N-contact problem, a matrix of coefficients relating the potentials on various contacts to the charges Q is generated such that

$$[\bar{\Phi}] = [P]\,[Q] \tag{50}$$

The matrix [P] is known as the 'coefficient-of-potential' matrix. The inverse of this matrix relates the charge vector to the potential vector and is the matrix of capacitive coefficients. It is also known as the 'coefficient-of-induction' matrix. If we represent this matrix by [c], we have

$$[Q] = [c]\,[\Phi] \tag{51}$$

The charge on any contact-i can be expressed in terms of contact potentials and the contact-to-contact capacitance as

$$Q_i = C_{i0}(\Phi_i) + C_{i1}(\Phi_i - \Phi_1) + \ldots + C_{iN}(\Phi_i - \Phi_N) \tag{52}$$

C_{i0} is the capacitance between contact-i and the ground point. Expanding the i-th row of (51) and comparing with (52) we have

$$C_{im} = -c_{im}; C_{i0} = c_{ii} + \left(\sum_{m=1}^{N} c_{im} \right)_{m \neq i} \tag{53}$$

A single contact can be divided into sub-contacts for improved accuracy. The charge on the contact is then found by setting all the sub-contacts to unit potential. A typical charge distribution on contacts in the two-contact problem of Fig. 11a is shown in Fig. 11b, where the charge density is depicted by the density of the hatched lines. Usually it is necessary to subdivide the contacts to obtain accurate results. Different authors use different subdivision schemes. The simplest scheme is to use uniform divisions. However, it is usually a non-optimal solution and a more efficient solution is to use finer grids near the edges of the contacts and coarse grids in the interior of the contact [18].

It must be clarified that while in the numerical scheme discussed earlier the entire substrate is meshed, this technique requires only the contacts to be meshed. This results in significantly smaller matrices. These matrices are however dense, compared to the numerical scheme which results in sparse matrices.

A uniform charge density has been used in this section. This is not necessary and spline approximations [19] can also be used. However using splines makes the evaluation of (48) more complicated.

3.2.3: Extraction of Substrate Parasitics

The integrated-circuit substrate can be treated as a distributed resistance at low frequencies. Under this approximation the substrate can be treated as the specific case of a generalized multilayered dielectric problem. In this case, we regard the substrate as an equivalent dielectric composed of several layers, each of which has a dielectric constant proportional to the resistivity of the corresponding resistive substrate layer. The resistive problem and the dielectric problem are equivalent in the electrostatic limit. In the resistive case the parasitics consist of resistors and relate the contact voltages and the contact currents. In the dielectric case the parasitics are capacitors and relate contact potentials and contact charges. The boundary conditions between layers in the resistive case require that the normal component of the current density at the interface of two neighboring layers, say 'i' and 'i+1', be equal, that is

$$J_n = \sigma_i E_i = \sigma_{i+1} E_{i+1} \tag{54}$$

The boundary conditions at the dielectric interface between two dielectric layers require that the normal component of the electrical displacement be equal, that is

$$D_n = \varepsilon_i E_i = \varepsilon_{i+1} E_{i+1} \tag{55}$$

In both the resistive and the capacitive cases, the potential satisfies the Laplace equation. The parasitic capacitance and the resistance extracted in the two cases are also in the same proportion. The capacitance between any two contacts 'i' and 'j' is defined as the ratio of the charge on contact-j to the potential of contact-i, which can be set to unity. By use of Stoke's theorem that can be shown to be

$$C = -\frac{1}{\varepsilon}\oint_S E \bullet ds \tag{56}$$

where S is the surface area of the contact and E is the electric field in the medium.

In a similar manner the resistance between the contacts can be defined to be

$$\frac{1}{R} = -\sigma\oint_S E \bullet ds \tag{57}$$

where σ is the medium conductivity.

By treating a generalized dielectric problem, it is possible to include the susceptance of the substrate by considering complex dielectric constants and determine the contact-to-contact impedance in the frequency domain.

Fig. 12a depicts the cross-section of the equivalent dielectric substrate. The substrate consists of N layers of dielectric constants ε_k (k= -N to 0) of thicknesses d_k respectively. The bottom plane is considered to be an ideal ground contact at potential zero, and the normal electric field is considered zero on the side faces and on the surface.

The derivation of the Green function is shown in the Appendix A3.1. Here we consider only the case when both the point charge and the point of observation are in the same dielectric layer ε_N. To a first order the contacts are assumed to be planar and on the surface with z=z'=0.

The Green function is then given by

$$G(x, y;x', y')_{z = z' = 0} = (G_0)_{z = z' = 0} + \left(\sum_{m = 0}^{\infty} \sum_{n = 0}^{\infty} f_{mn} \cos\left(\frac{m\pi x}{a}\right) \cos\left(\frac{m\pi x'}{a}\right) \cos\left(\frac{n\pi y}{b}\right) \cos\left(\frac{n\pi y'}{b}\right) \right) \quad (58)$$

where $C_{mn}=0$ for (m=n=0), $C_{mn}=2$ for m=0 or n=0 but $m \neq n$ and $C_{mn} = 4$ for all (m>0) and (n>0). G_0 is the term representing the m=n=0 case and is shown in (62).

$$f_{mn} = \frac{1}{ab\gamma_{mn}\varepsilon_N} \times \frac{\beta_N \tanh(\gamma_{mn}d) + \Gamma_N}{\beta_N + \Gamma_N \tanh(\gamma_{mn}d)} \quad (59)$$

where

$$\gamma_{mn} = \sqrt{\left(\frac{m\pi}{a}\right)^2 + \left(\frac{n\pi}{b}\right)^2} \quad (60)$$

β_N and Γ_N are computed recursively from equations

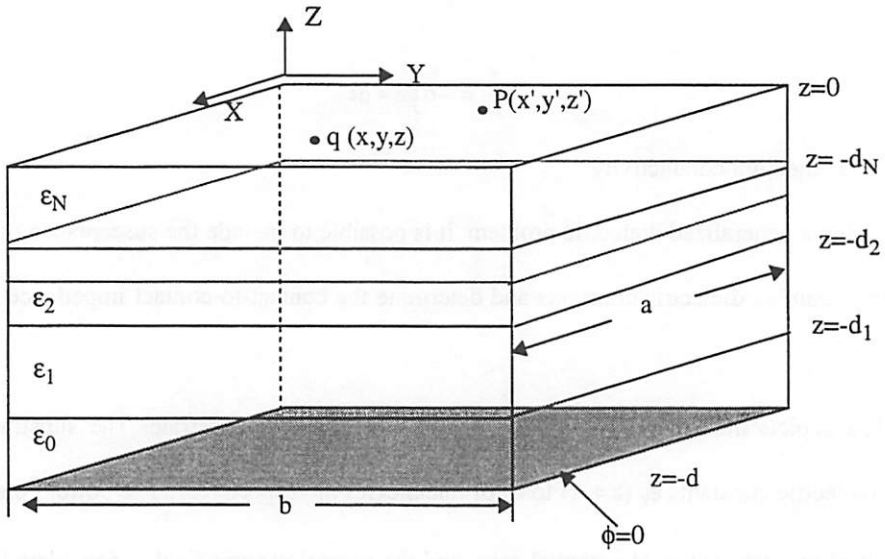


Figure 12a: Substrate Cross-Section

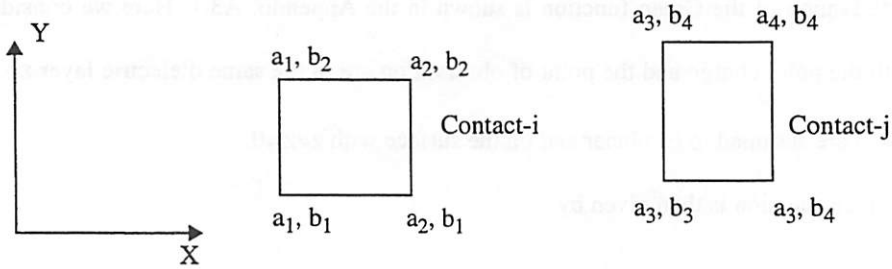


Figure 12b: Contact Coordinates

$$\begin{bmatrix}\beta_k\\ \Gamma_k\end{bmatrix} = \begin{bmatrix}\frac{\varepsilon_{k-1}}{\varepsilon_k}\cosh^2(\theta_k) - \sinh^2(\theta_k) & , \left(\frac{\varepsilon_{k-1}}{\varepsilon_k} - 1\right)\cosh(\theta_k)\sinh(\theta_k)\\ \left(1 - \frac{\varepsilon_{k-1}}{\varepsilon_k}\right)\cosh(\theta_k)\sinh(\theta_k) & , \cosh^2(\theta_k) - \frac{\varepsilon_{k-1}}{\varepsilon_k}\sinh^2(\theta_k)\end{bmatrix}\begin{bmatrix}\beta_{k-1}\\ \Gamma_{k-1}\end{bmatrix} \tag{61}$$

where $1 \le k \le N$ and $\theta_k = \gamma_{mn}(d-d_k)$, $\beta_0 = 1$ and $\Gamma_0 = 0$.

For m=n=0

$$G_0 = \frac{1}{ab\varepsilon_N} \times \frac{\Gamma_N}{\beta_N} \tag{62}$$

where

$$\begin{bmatrix}\beta_k\\ \Gamma_k\end{bmatrix} = \begin{bmatrix}\frac{\varepsilon_{k-1}}{\varepsilon_k} & 0\\ \left(\frac{\varepsilon_{k-1}}{\varepsilon_k} - 1\right)d_k & 1\end{bmatrix}\begin{bmatrix}\beta_{k-1}\\ \Gamma_{k-1}\end{bmatrix} \tag{63}$$

where $\beta_0 = 1$ and $\Gamma_0 = d$.

If surface contacts are considered, then the volume integrals in (49) are replaced by equivalent surface integrals, that is

$$\bar{\phi}_i = \frac{Q_j}{S_j S_i} \int_{S_i}\int_{S_j} G\,ds_j\,ds_i \tag{64}$$

The above definition is used to find $p_{ij} = \bar{\phi}_i/Q_j$ for any two contacts i and j defined on the surface by X and Y coordinates (a_1, a_2), (b_1, b_2) and (a_3, a_4) and (b_3, b_4) respectively, as shown in Fig. 12b. 'p_{ij}' is then given by

$$p_{ij} = \frac{\Gamma_N}{ab\varepsilon_N\beta_N} + \left\{\sum_{m=0}^{\infty}\sum_{n=0}^{\infty} f_{mn}C_{mn}\frac{a^2b^2}{m^2n^2\pi^4} \times \frac{\left(\sin\left(m\pi\frac{a_2}{a}\right) - \sin\left(m\pi\frac{a_1}{a}\right)\right)\left(\sin\left(m\pi\frac{a_4}{a}\right) - \sin\left(m\pi\frac{a_3}{a}\right)\right)}{(a_2 - a_1)(a_4 - a_3)}\right.$$

$$\left.\frac{\left(\sin\left(n\pi\frac{b_2}{b}\right) - \sin\left(n\pi\frac{b_1}{b}\right)\right)\left(\sin\left(n\pi\frac{b_4}{b}\right) - \sin\left(n\pi\frac{b_3}{b}\right)\right)}{(b_2 - b_1)(b_4 - b_3)}\right\} \tag{65}$$

Three-dimensional contacts, where the Z-dimension must be included are considered in Appendix

A3.2. The three-dimensional nature of the contacts will be of importance only when the side-wall areas of the contacts are comparable to the bottom-wall areas. In most applications, this is not the case and a simple two-dimensional calculation should suffice. However, the three-dimensional structure can be included with a small computational penalty by using the method discussed in A3.2.

3.2.3.1: Fast Computation of the Coefficient-of-Potential Matrix

The doubly infinite series in (65) is slow to converge, requiring large upper limits M and N for the indices m and n to achieve an acceptable error. For a single layer, we have from (59)

$$f_{mn} = \frac{\tanh(\gamma_{mn} d)}{ab\gamma_{mn}\varepsilon_0} \tag{66}$$

To consider the worst possible case for convergence, consider the case in (65) where $m\pi(((a_{4(2)}-a_{3(1)})/a) \ll 1$, $n\pi(((b_{4(2)}-b_{3(1)})/b) \ll 1$ and $\gamma_{mn}d > 1$. This will be the case for small contact areas and large substrate dimensions (a, b) and (d > a, b). In this case the series for p_{ij} can be approximated by

$$p_{ij} = \sum_{m=0}^{M}\sum_{n=0}^{N} C_{mn}\frac{1}{ab\gamma_{mn}\varepsilon_0}.$$

This series begins to converge once the conditions $m\pi(((a_{4(2)}-a_{3(1)})/a) \ll 1$ and $n\pi(((b_{4(2)}-b_{3(1)})/b) \ll 1$ are no longer satisfied at sufficiently large values of m and n. Before that however, the series above behaves as a divergent series. Hence in order to get accurate results, the upper limits M and N must be sufficiently large.

For J contacts on the surface, the size of the [P] matrix is J^2. If each element of the series in (65) is computed J^2 times using a direct computation approach, then this method will not possess much advantage over the numerical scheme. In fact this problem has traditionally led to a rejection of the series form of the Green function for parasitic extraction. However, this problem can be eliminated by use of the Discrete Cosine Transform as described below.

By use of the identity

$$\sin\left(m\pi\frac{a_i}{a}\right)\sin\left(m\pi\frac{a_j}{a}\right) = 0.5\left(\cos\left(m\pi\frac{(a_i - a_j)}{a}\right) + \cos\left(m\pi\frac{(a_i + a_j)}{a}\right)\right) \tag{67}$$

(65) can be rewritten as a sum of 64 terms of the form

$$\sum_{m=0}^{\infty}\sum_{n=0}^{\infty} k_{mn}\cos\left(m\pi\left(\frac{a_{1,2}\pm a_{3,4}}{a}\right)\right)\cos\left(n\pi\left(\frac{b_{1,2}\pm b_{3,4}}{b}\right)\right) \tag{68}$$

where $k_{mn} = \frac{a^2b^2}{m^2n^2\pi^4}f_{mn}C_{mn}$.

Let the ratios of the contact coordinates and the substrate dimensions be expressed as ratios of integers such that

$$\frac{a_k}{a} = \frac{p_k}{P};\frac{b_k}{b} = \frac{q_k}{Q} \tag{69}$$

Then (68), summed to upper limits P-1 and Q-1, can be expressed as

$$\sum_{m=0}^{P-1}\sum_{n=0}^{Q-1} k_{mn}\cos\left(m\pi\left(\frac{p_{1,2}\pm p_{3,4}}{P}\right)\right)\cos\left(n\pi\left(\frac{q_{1,2}\pm q_{3,4}}{Q}\right)\right) \tag{70}$$

The two-dimensional Discrete Cosine Transform (DCT) of any series k_{mn} is defined as

$$K_{pq} = \sum_{m=0}^{P-1}\sum_{n=0}^{Q-1} k_{mn}\cos\left(m\pi\frac{p}{P}\right)\cos\left(n\pi\frac{q}{Q}\right) \tag{71}$$

The DCT can be computed in a very efficient manner by use of the Fast Fourier Transform algorithm as shown in Appendix A3.3.

Eq. (70) cannot directly be considered to be the DCT of k_{mn} because of the presence of the terms $(p_{1,2}+p_{3,4})$ and $(q_{1,2}+q_{3,4})$. These indices have an upper limit going up to 2P and 2Q respectively unlike the indices in (71) which are restricted to P-1 and Q-1. However, the properties of the DCT shown in (72a-c) allow the computation of (70) from the DCT of k_{mn} as expressed in (71).

$$K_{(2P-p,\,q)} = K_{p,\,q} \tag{72a}$$

$$K_{(p,\,2Q-q)} = K_{p,\,q} \tag{72b}$$

$$K_{(2P-p,\,q)} = K_{p,\,q} \tag{72c}$$

By use of the DCT, computation of the [P] matrix proceeds in the sequence of steps shown below.

1] Input substrate data. This includes the thickness and dielectric constants of the various dielectric

layers composing the substrate.

2] Compute f_{mn} as shown in (59)-(63) and k_{mn}.

3] Compute the DCT of k_{mn}.

4] Read in contact X and Y coordinates and compute the equivalent integer representation of the coordinates as in (69).

5] Compute (65) using (68).

6] Generate matrix [P] by computing all possible combinations of contacts.

7] Invert [P] to generate [c] as in (51) and extract contact-to-contact parasitic as in (53).

8] If recomputation needs to be done with different contact coordinates go to step 4 above.

Once the DCT of k_{mn} has been computed, it can be stored as a matrix and need not be recomputed for different surface contact configurations. This feature makes this method particularly suitable for optimization related tasks, where the designer wishes to experiment with the layout of the contacts. An updated [P] matrix can be generated by simply accessing the appropriate elements of the DCT array.

The new [P] matrix, resulting from varying the contact coordinates, differs from the old one only in those elements which correspond to the displaced contacts. If only a few contacts are displaced, then the inverse of the new [P] matrix can be computed from that of the old one in a very efficient manner by use of the Sherman-Morrison formula as discussed in 3.2.4.2. The DCT computation itself is very fast, as it relies on the FFT algorithm.

As stated in the paragraph preceding (48), it is possible to regard the potential at any point on the contact as the contact potential, or one may average the potential over the volume of the contact. Using the average potential as is done in (65) results in a symmetric [P] matrix. Symmetric matrices have several advantages over asymmetric matrices with regards to computational efficiency. One advantage is that only one half of a symmetric matrix needs to be computed. It is also easier to invert a symmetric matrix as discussed below.

3.2.4: Matrix Inversion

Once the [p] matrix has been computed from the 2D FFT array, the matrix must be inverted to obtain

the [c] matrix. The matrix is first decomposed using the symmetric Choleski LU-decomposition technique [19], followed by an inversion of the L matrix. The capacitors are then extracted from the inverse of the [L] matrix.[1]

3.2.4.1: The Symmetric Choleski Decomposition

This technique for LU decomposition is convenient for symmetric matrices. The U matrix is a transpose of the L matrix. The L matrix is determined by solving the following equations for the elements of the matrix.

$$\begin{bmatrix} l_{11} & 0 & \cdots & 0 \\ l_{21} & l_{22} & \cdots & 0 \\ | & | & \cdots & | \\ l_{n1} & l_{n2} & \cdots & l_{nn} \end{bmatrix} \begin{bmatrix} l_{11} & l_{21} & \cdots & l_{n1} \\ 0 & l_{22} & \cdots & l_{n2} \\ | & | & \cdots & | \\ 0 & 0 & \cdots & l_{nn} \end{bmatrix} = \begin{bmatrix} p_{11} & p_{21} & \cdots & p_{n1} \\ p_{21} & p_{22} & \cdots & p_{n2} \\ | & | & \cdots & | \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{bmatrix} \tag{73}$$

The solution to the above set of equations can be found by the following algorithm.

$$l_{kj} = \frac{p_{kj} - \sum_{n=1}^{j-1} l_{kn} l_{jn}}{l_{jj}} ; j = 1, \ldots, k-1 \tag{74}$$

$$l_{kk} = \left(p_{kk} - \sum_{n=1}^{k-1} l_{kn}^2 \right)^{\frac{1}{2}} ; k = 1, \ldots, N \tag{75}$$

where N is the size of matrix A. The above decomposition requires approximately $N^3/6$ multiplications.

Since [L] is a lower triangular matrix, so is $[L]^{-1}$. The inverse of [L] can be computed in a straightforward manner [19] and requires approximately $N^3/6$ multiplications. The matrix of capacitive coefficients [c] is the product of $[L]^{-1}$ and $\{[L]^T\}^{-1}$. For a single calculation of the capacitances, it is unnecessary to compute [c], as shown below. Avoiding computation of [c] reduces the total multiplication count in the algorithm.

In order to extract the contact-to-contact capacitance from any one contact (assume 'i') to all other contacts, all the subdivisions of contact-i are set to unit potential, while all other contact subdivisions are

1. If [P] is not symmetric then regular LU decomposition must be used, and both [L] and [U] need to be inverted. Thus symmetric [P] matrices are advantageous.

set to zero. The capacitance connected between contact-i and contact-j is then given by the ratio of the total charge on contact-j to the potential on contact-i which is set to unity.

Let contact-i have N_1 subdivisions, and contact-j have M_1 subdivisions. Further, let the charge on the n-th subdivision of 'j' be given by Q_n^j, while the element of [c] representing interaction between the m-th division of 'i' and the n-th division of 'j' be given by c_{mn}^{ij}. Then the capacitance between contact-i and contact-j is given by

$$C_{ij} = \sum_{n=1}^{N_1} Q_n^j = \sum_{n=1}^{N_1} \left(\sum_{m=1}^{M_1} c_{mn}^{ij} \right) = \sum_{i,j} [c]\,[\varphi] \Big|_{\varphi^{j \neq i} = 0,\ \varphi_m^i = 1} \tag{76}$$

where ϕ_m^i is the potential on the m-th subdivision of contact-i and $[c]=([L]^T)^{-1}[L]^{-1}$.

If the expression for [c] is used in the calculation, it can be seen that it is unnecessary to compute [c] explicitly. The matrix $[L]^{-1}$ is first multiplied by the vector $[\phi]$, which extracts an N by N_1 sub-matrix from $[L]^{-1}$. This sub-matrix is then multiplied with $([L]^T)^{-1}$ to obtain C_{ij}.

3.2.4.2: The Sherman-Morrison Formula

In certain substrate parasitic extraction tasks, the circuit designer may wish to experiment with the layout of a few devices. If only a few devices are displaced, then only those elements of the [p] matrix, which relate the displaced contacts with the other contacts, are changed. In order to recompute the substrate macromodel in such a case, it is unnecessary to perform the LU decomposition and the inversion steps by the use of the Sherman-Morrison formula.

Consider the layout in Fig. 13. If contact-i shown in the figure is moved to a new location, then only those rows and columns of [p] which involve the subdivisions of contact-i, will be affected. We shall assume that the contact-i is transformed to contact-i* in a sequence of steps (1 to 4 in Fig. 13) in which each of the subdivisions is moved to the new coordinates.

Consider the displacement of i_1 to i_1*. This displacement causes one row and one column of the [p] matrix to change. Let the changes to the row be represented by a vector δp_r and the changes to the column by δp_c. δp_c is the transpose of δp_r. Let [p] be changed to $[p]+\delta p_r$, represented by [p'], where the elements of

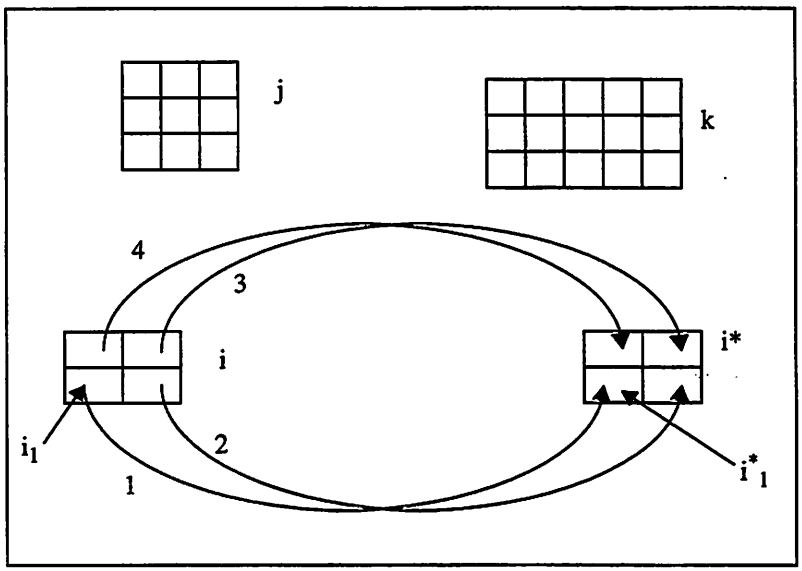


Figure 13: Layout-Optimization Problem

δp_r are added to the appropriate row of [p] which we refer to as row-k. The inverse of [p'] is related to the $[p]^{-1}$ or [c] by the following relation.

$$[p']^{-1} = [c] - \frac{[c]_k([c] \bullet \delta p_r)}{1 + \delta p_r \bullet [c]_k} \tag{77}$$

where $[c]_k$ is the k-th column of matrix [c]. The above is the Sherman-Morrison formula. A similar approach is followed to compute the effect of δp_c on the inverse of the matrix.

The procedure is repeated for displacements 2-4 as well. Each step of the Sherman-Morrison formula requires $2N^2$ operations. This is a significant improvement over the N^3 order of the matrix inversion process.

3.2.5: Improvements to the Model

3.2.5.1: Lateral Variations in the Dielectric Constant

The Green function technique is limited to planar structures and it is usually difficult to develop analytical models for handling lateral variations in the dielectric constant. If these variations are localized and small, then their boundaries can be treated as equipotentials. The more general case of large regions of lateral variation in the dielectric constant requires a combination of analytical and numerical techniques.

An example of a large lateral variation is a deep and wide oxide trench on the surface of the substrate.

The algorithm for tackling problems with lateral resistivity variations is explained by considering a specific example shown in Fig. 14a. The macromodel for two contacts shown above is to be extracted, given a large area of dielectric constant ε^* between the two contacts. As shown in Fig. 14b, where only the top layer of the substrate has been shown, the problem can be considered to be a superposition of two

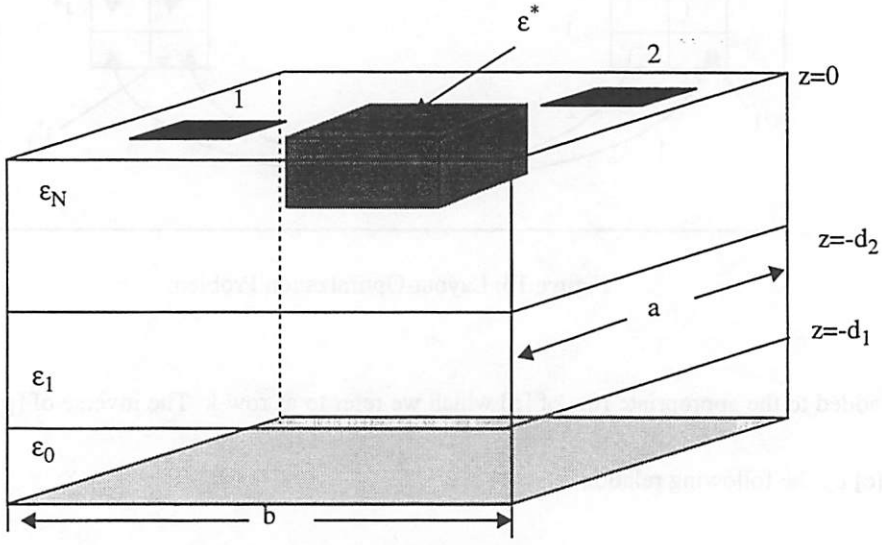


Figure 14a: Lateral Variation in the Dielectric Constant

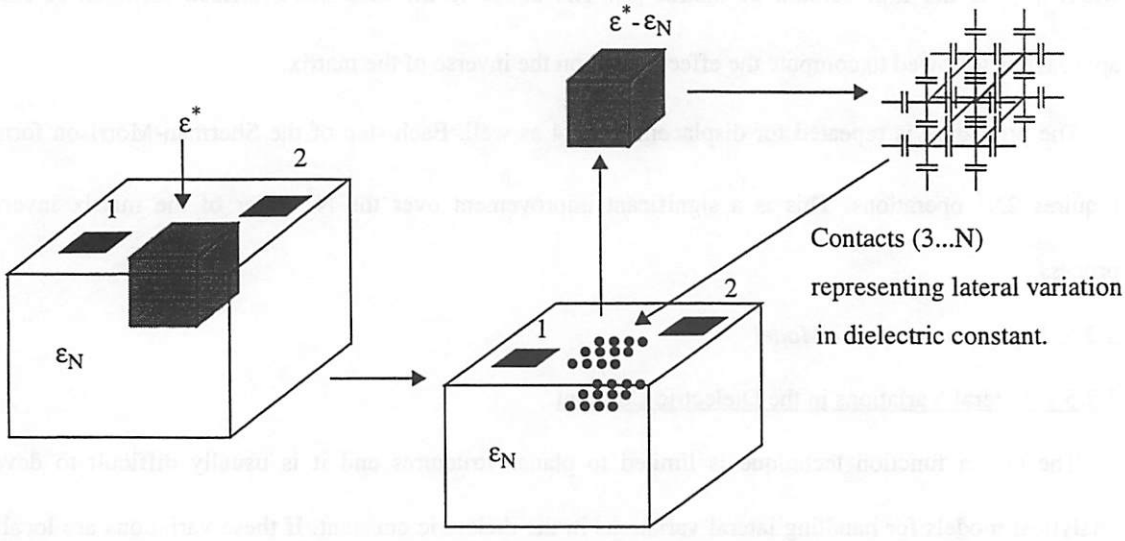


Figure 14b: A Technique to Model Lateral Dielectric Constant Variations

problems.

The first problem consists of the original contacts and several equivalent contacts (3-N) in the region where the lateral variation in the dielectric constant was located. The dielectric constant of the entire top-layer is assumed to be ε_N and the capacitance mesh for contacts 1-N are computed using the Green function technique. In the second problem the equivalent capacitance matrix for a material of dielectric constant ε^*-ε_N is calculated. This is a simple grid representation of this material. This grid is then placed in parallel with contacts 3-N in the circuit simulator in order to compute the effect of the lateral variation in the dielectric constant.

This method should be used only in places where the accuracy may be seriously compromised if the lateral variation is not considered. This is so because the additional contacts used to represent the lateral variations increase the size of the [P] matrix and also increase the matrix size in the circuit simulator.

3.2.5.2 Reducing the Complexity of the Substrate Problem

The number of model elements generated by the substrate simulator grows approximately as N(N+1)/2, for N contacts on the surface. The substrate model is included in a circuit simulator. With the $O(N^2)$ growth of the substrate-model impedance matrix size, it is easy to see that the substrate sub-circuit can slow down the circuit simulator operation for large values of N. For example, the substrate simulator will generate approximately 500,000 resistors for a 1000 contact problem. The substrate simulator will itself be slow because the inversion of the dense [P] matrix is an $O(N^3)$ operation.[2] It is thus necessary to explore algorithms which can reduce the size of the impedance matrices extracted by the substrate simulator.

One such algorithm, which is used to make the admittance matrix [y] sparse, is presented in this section. Consider the hypothetical problem shown in Fig. 15a below. The substrate shown in the figure has N contacts on the surface. Each contact is connected to ground by an arbitrary impedance. This impedance can include the reactance of the substrate capacitance and any external impedance connected to the contact on the surface of the substrate. When all the impedance values are zero, that is, all the contacts are at perfect grounds, the only current-flow path that can exist between any pair of contacts i and j is modeled

2. After the first inversion is performed, optimization will not be slow, if the Sherman-Morrison formula is employed, as discussed in Section 3.2.4.2.

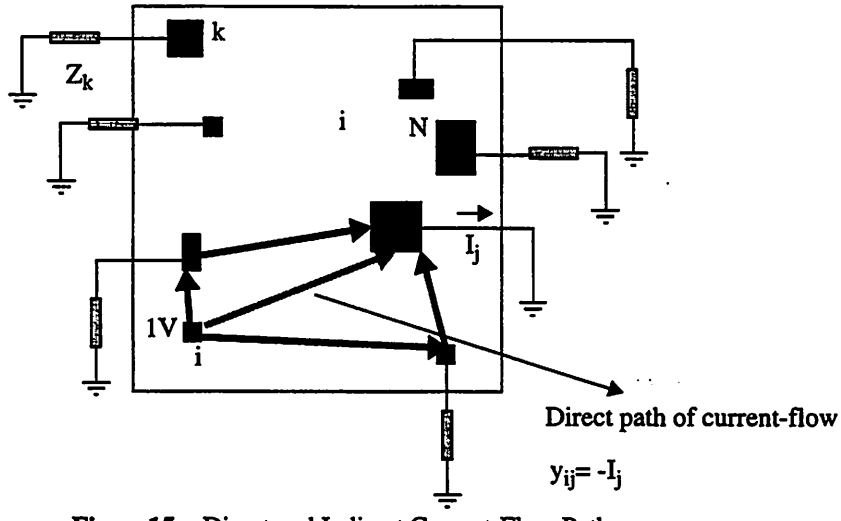


Figure15a: Direct and Indirect Current-Flow Paths

by the coefficient y_{ij} of the admittance matrix, depicted by the shaded line in the figure. If, however, the impedances Z_k shown in Fig. 15a are nonzero, the dominant current path will not be the direct-path y_{ij}, but will consist of the intermediate paths between the contacts as shown in the figure in solid lines. It is possible that the current flowing through the intermediate paths exceeds the direct-path current. In such a case, the coefficient y_{ij} can be set to zero in the matrix [y], without significantly compromising accuracy. Hence if all impedance values Z_k are known, then for any row 'i' of the matrix [y], all the elements y_{ij}, for which the indirect current paths dominate, can be set to zero.

Quantitatively, when contact-j is at ground and contact-i is at unit potential, the voltages at all other contacts are given by the following equation, where for simplicity, the contact-impedances have been considered to be resistive (R_k).

$$\begin{bmatrix} \left(y_{11}+\frac{1}{R_1}\right) & y_{12} & \cdots & y_{1k} & \cdots & y_{1N} \\ y_{21} & \left(y_{22}+\frac{1}{R_2}\right) & \cdots & y_{2k} & \cdots & y_{2N} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ y_{k1} & y_{k2} & \cdots & \left(y_{kk}+\frac{1}{R_k}\right) & \cdots & y_{kN} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ y_{N1} & y_{N2} & \cdots & y_{Nk} & \cdots & \left(y_{NN}+\frac{1}{R_N}\right) \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \\ \cdots \\ V_k \\ \cdots \\ V_N \end{bmatrix} = \begin{bmatrix} -y_{1i} \\ -y_{2i} \\ \cdots \\ -y_{ki} \\ \cdots \\ -y_{Ni} \end{bmatrix} \tag{78}$$

Once the voltages V_k are computed from the above equation, the current I_j can be computed from the following equation

$$I_j = y_{ji} + \sum_{k=1}^{N} y_{jk} V_k \Big|_{k \neq i,\, k \neq j} \tag{79}$$

'y_{ji}' can be ignored or set to zero whenever the following condition is satisfied

$$\left| \sum_{k=1}^{N} y_{jk} V_k \right|_{k \neq i,\, k \neq j} > |y_{ji}| \tag{80}$$

The following strategy has been employed to implement the above scheme (Fig. 15b). To find the interactions between any contact-I and other contacts, the surface of the substrate is partitioned with contact-I at the center. All interactions in the partition which includes contact-I, called the central partition,

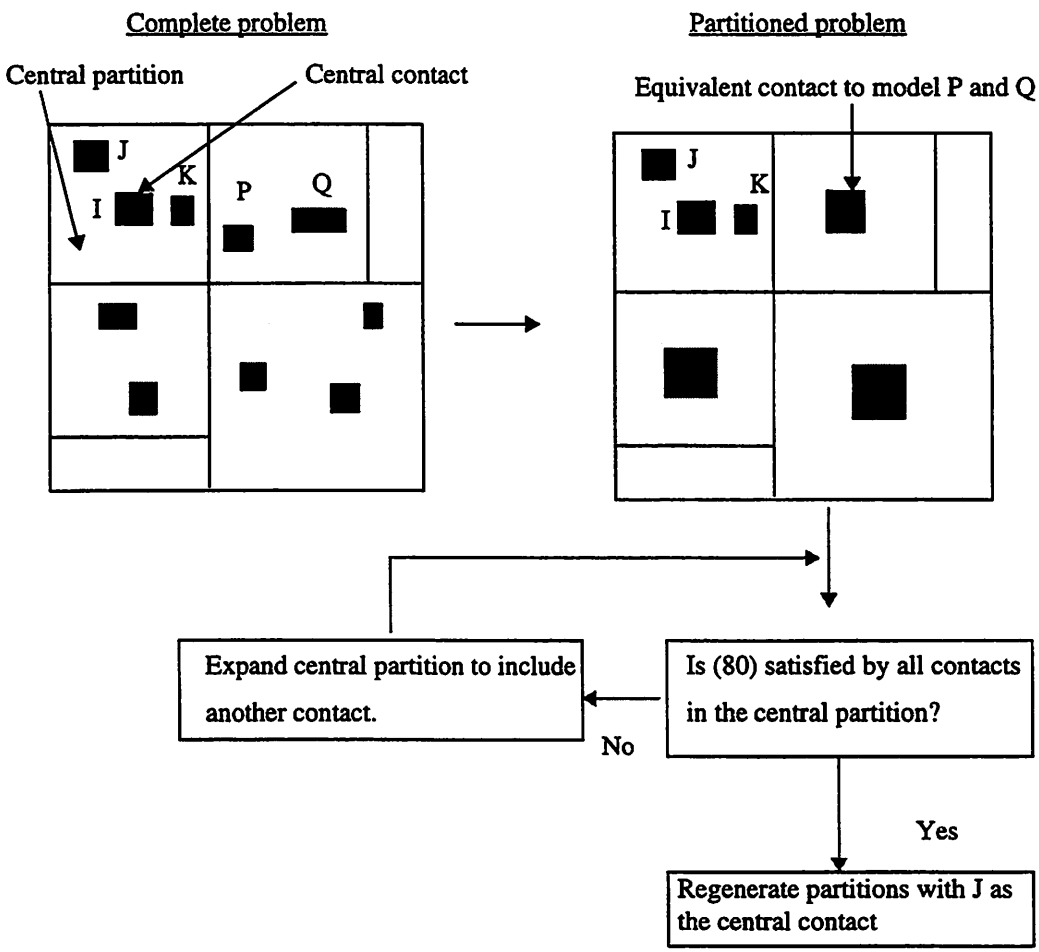


Figure 15b: Partitioning of the Substrate Problem

are computed explicitly. The interaction between I (called the central contact) and the contacts lying outside the central partition are not computed explicitly. Instead equivalent contacts, placed at the center of the other partitions, are used to model the influence of the contacts belonging to those partitions on the contacts in the central partition. The area of an equivalent contact in a partition is set equal to the sum of the areas of all contacts lying in that partition. The central partition is expanded to include contacts until the condition expressed in (80) is satisfied[3]. This step is iterated for all contacts, by treating each contact as a central contact and by generating new partitions to compute direct interactions for each contact. These steps are shown in Fig. 15b.

An exponential-square partitioning scheme is shown in Fig. 15b. The central partition and the eight partitions in its immediate neighborhood are of minimum area. The subsequent partitions grow by powers of three for distant contacts (Fig. 15c). Similar partitioning schemes have been used in the past in algorithms which employ the multipole approach [20].

The size of the [P] matrix to be inverted is reduced considerably by the use of the equivalent contacts to model distant contacts. A reduction in the size of [P] lowers the substrate simulator inversion time

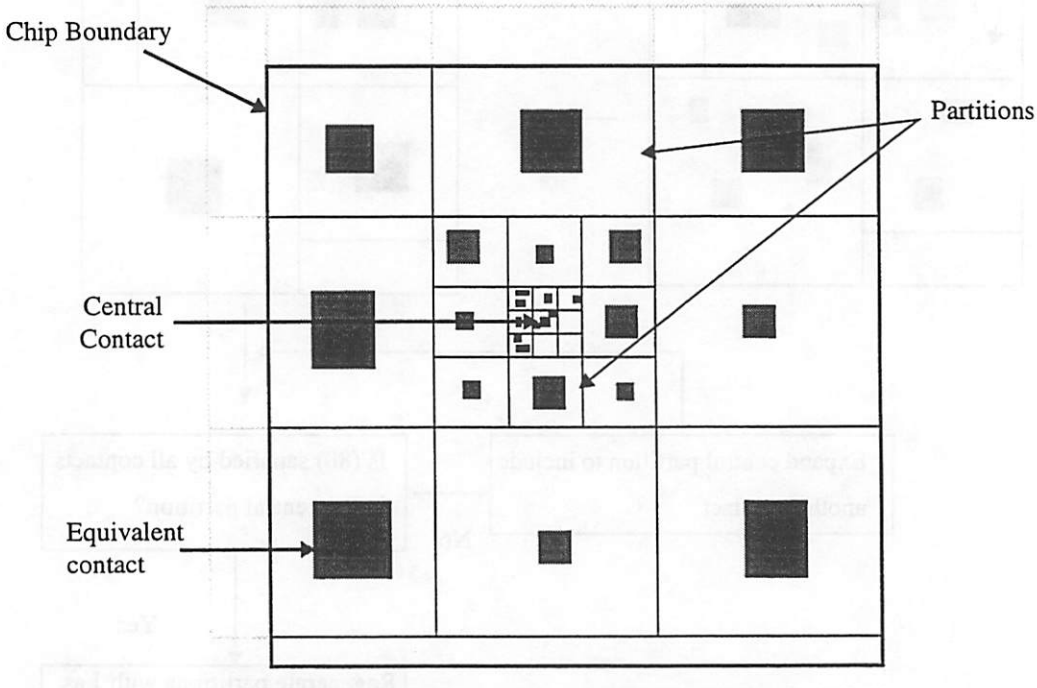


Figure 15c: An Exponential Partitioning Scheme

3. A fixed ratio of the LHS and the RHS of (80) can also be used as the threshold.

significantly. This technique has been applied to large problems, for example, the partial admittance matrix of a 2500 contact problem was extracted in 67 minutes on a DEC Alpha workstation.

The partition technique shown above has not been optimized. More robust and accurate schemes need to be researched. There are some problems with this form of partitioning. One problem is that since the partitions are changed for each contact, the size of the equivalent contacts in non-central partitions are also different. Thus the boundary conditions can be different for each contact. If contact-i is a central contact, with contact-j inside the central partition, then the y_{ij} obtained from this partition can, in general, be different from the value y_{ji} obtained when 'j' is the central contact and 'i' is within the central partition.

The capability to generate equivalent contacts to simplify the inversion is made possible by the use of the DCT. This feature is not possible if the Green function is computed explicitly.

3.3: Comparison of Techniques

Two techniques have been presented in this chapter for computing the substrate macromodels. The first technique is purely numerical while the second utilizes a combination of numerical and analytical methods. In this section, a general comparison of the techniques will be carried out. The advantages and the disadvantages of each technique will be listed and discussed.

The primary advantage of the integral-equation technique is the speed of computation, which was found to be significantly superior to the numerical technique. Simultaneous-Over-Relaxation was employed in the purely numerical technique. A comparison of the computation time required in two problems is shown in the tables below. The first problem is that of a single square contact on a uniform substrate. The computation time for the numerical and the DCT based techniques are compared. The computation time in the numerical scheme is small for a small number of grid points, but becomes large as the number of points is increased. It can be seen from Table 1 that the accuracy of the result in the numerical technique is poor for a small number of grid points. Table 1 also lists the computation time if the [P] matrix is computed explicitly without using the DCT. As can be seen the computation time is large. For J contacts the time required in the direct computation will increase as J^2. Thus the Green function must be computed by using the DCT.

The second problem considered is that of four contacts laid out in a row. The time required to compute the ten element macromodel is shown in Table 2. The computation time of the DCT based technique is again seen to be superior.

The time required to compute the DCT was seventy five seconds for a 512*512 point DCT. This is a setup time and is not a part of the actual computation time. All of the above simulations were performed on an IBM RS 6000 workstation.

The memory requirement is defined as the number of double precision words required in each technique. For achieving good accuracy, the memory requirement of the numerical technique is seen to be very large. The major memory utilization in the DCT based technique is for the storage of the DCT elements. Thus the memory requirement for the DCT technique is nearly the same for both the problems simulated. Another advantage of the DCT technique can be realized by generating the [P] matrix directly from the DCT array during computation. By adopting this procedure the [P] matrix does not have to be stored explicitly. The computational penalty is small, since the calculation of the p-coefficients from the DCT array simply requires sixty four additions, as discussed in Section 3.2.3.1.

Technique	Memory	Computation time (s)	Resistance (Ω)
FD (3920 pts.)	19k	15	233
FD (6480 pts.)	32k	35	273
FD (13520 pts.)	67k	75	298
FD (35280 pts.)	176k	180	311
FD (109520 pts.)	547k	674	318
Green function	10k	190	345
DCT	263k	1	340

Table 1: Resistance of a single contact to the backplane using the three techniques

Technique	Computation time (s)	Memory
DCT	2	263k
FD (50000 pts.)	1174	250k

Table 2: Four-contact problem using the DCT and the FD Techniques

A problem with the numerical technique is that the number of mesh points rises rapidly with the number of contacts, since the X and Y extent of each contact must be meshed fine enough to provide reasonable accuracy. The mesh extends on the areas of the chip not occupied by the contacts as well. In small problems with few contacts, it may be possible to make the mesh coarse in regions far away from the contacts. This will most probably not be possible in a large problem with several hundred contacts. To illustrate the problem we consider a hypothetical case consisting of hundred surface contacts, distributed on the surface such that the X and Y coordinates of no two contacts overlap. Let the X and Y dimensions of all contacts be meshed into five points along the two axes. Let the region between any two contacts be also similarly meshed into five grid lines in the X and Y directions. The number of X and Y grid lines will thus each be in the order of a thousand. The number of points on the surface will be in the order of a million. Further, to simulate the vertical extent of substrate, if we conservatively use twenty planes, then the total number of grid points will be in the order of several million. If the resistivity varies significantly in the vertical direction, then the number of Z-planes required will increase. The computation time for problems of this magnitude will always be large, regardless of the technique used to solve the system of equations. The power of the DCT based technique lies in the fact that meshing needs to be done only in the region of the contacts, not in the bulk.

Another disadvantage of the numerical technique is that for optimization, if a single contact is varied, then the entire problem has to be recomputed. This is not a problem with the DCT technique since the Sherman-Morrison formula can be utilized very efficiently for this process.

An advantage of the purely numerical technique is its versatility. The technique can be used to model lateral variations in resistivity without any overheads, unlike the DCT based method.

Chapter 4: Substrate Coupling in Integrated Circuits

In this chapter, the effects of substrate coupling in integrated circuits are considered. The chapter begins with a discussion of substrate coupling in different substrate types. Two types of substrates are analyzed (Fig. 1). These substrates have been chosen as they are widely used in industry. Other substrate cross-sections can be simulated by using the technique as well. The conclusions regarding the nature of coupling are behavioral and the exact magnitude of coupling must be simulated for individual cases.

Substrate taps and guard rings are commonly used in mixed analog-digital designs. Typical improvements in isolation, which may be expected from the use of these grounding schemes in different substrate types are presented in the following section. An attempt is made to provide simple guidelines for effective guard ring layout. The cases where guard rings may not be useful are also discussed. Since guard rings can take up a significant amount of chip area, this knowledge is useful. It is also possible that the use of guard rings can actually worsen isolation. This issue is also addressed in this chapter.

Substrate coupling can have different types of effects in different circuit applications. In general these effects may be divided into two types. The first class of effects involve those in which one circuit influences the performance of another on the same substrate. The second class of effects are those in which the finite substrate impedance effects the performance of a single circuit block, without interaction from other circuits. An example of the first class of interaction is the noise injection into a low-noise amplifier (LNA) in the presence of a switching ring oscillator on the same substrate. The lowering of the noise figure and the bandwidth of an LNA due to the substrate parasitic belong to the second class of effects. Several such circuit-level effects of substrate coupling are discussed by using the simulator described in the previous chapter and the circuit simulator SPICE.

4.0: Substrate Coupling in Different Substrate Types

Shown in Fig. 1 are two substrates which are studied and compared in this section. The figure is repeated here for convenience (Fig. 16). The first substrate consists of a thin surface buried-p region of 1-2μm thickness and a highly resistive bulk region of resistivity in the range of 10-30Ω-cm. This substrate is used in certain BiCMOS processes. The low-resistivity surface-layer serves as a channel-stop layer for

MOSFETS, reduces CMOS latch-up and raises the surface inversion potential under the field oxide. The second substrate is used in CMOS processes. The epi-layer thickness is approximately 5-15μm with a resistivity of about 5-20Ω-cm. Some processes have a thin channel-stop implant at the surface, approximately 0.5-1μm thick, with a lower resistivity in the range of 0.5-1Ω-cm. The bulk region is heavily doped, with a resistivity of about 10-100mΩ-cm, and a thickness in the range of 100-400μm. This substrate is called the low-resistivity substrate.

The simulation technique discussed in Chapter 3 is applied to study the substrate coupling in these substrates. The behavior of the substrate model elements in the low- and high-resistivity substrates is presented in Section 4.0.1. The isolation provided by the two substrates between single-ended contacts is discussed in the next section. Differential circuits are known to provide better isolation than single-ended circuits. Typical improvements in isolation, achievable on the two types of substrates, are discussed in Section 4.0.3. Substrate coupling is strongly dependent on the load-impedances used on-chip. An analysis of this dependence is presented in Section 4.0.4. As discussed in Chapter 2, some devices inject and receive substrate noise by non-capacitive mechanisms. Substrate coupling caused by these effects is discussed in Section 4.0.5.

In the simulations discussed below, different values have been used for layer resistivities as shown in Fig. 16. These values are indicated in each of the simulations, and referred to as $\rho h_{(1\text{-}3)}$ and $\rho l_{(1\text{-}4)}$ in the

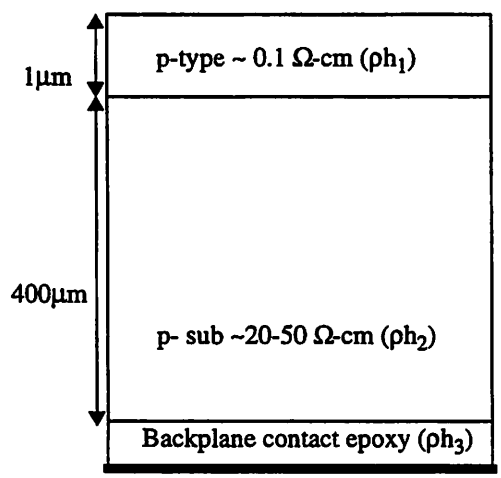


Figure 16a: High-Resistivity Substrate

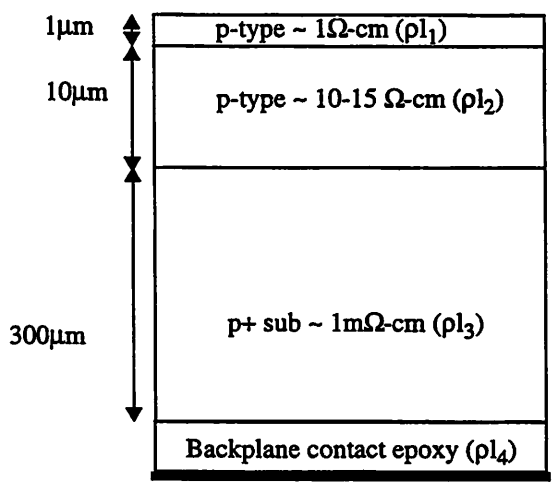


Figure 16b: Low-Resistivity Substrate

case of the high- and low-resistivity substrates respectively. The bottom layer in both substrates is the epoxy material used to attach the die to the package. A high-resistivity epoxy is used in some packages, while a conductive material is used in others. The value of the resistivity used in the particular experiment is indicated. The thick line at the bottom of each figure illustrates the die header metal. Other relevant variables, such as contact-to-contact distances and junction capacitance values, are also mentioned in each section.

4.0.1: Variation of Model Elements as Functions of Position and Contact Area

This experiment demonstrates the behavior of the substrate model elements as functions of area and the distance between them. This section merely serves to illustrate the behavior of model elements. Circuit related issues are discussed in later sections.

The resistance of a square surface-contact to the backplane is shown in Fig. 17a and Fig. 17b, for a high and a low-resistivity substrate respectively. The contact is placed at the center of the substrate. The

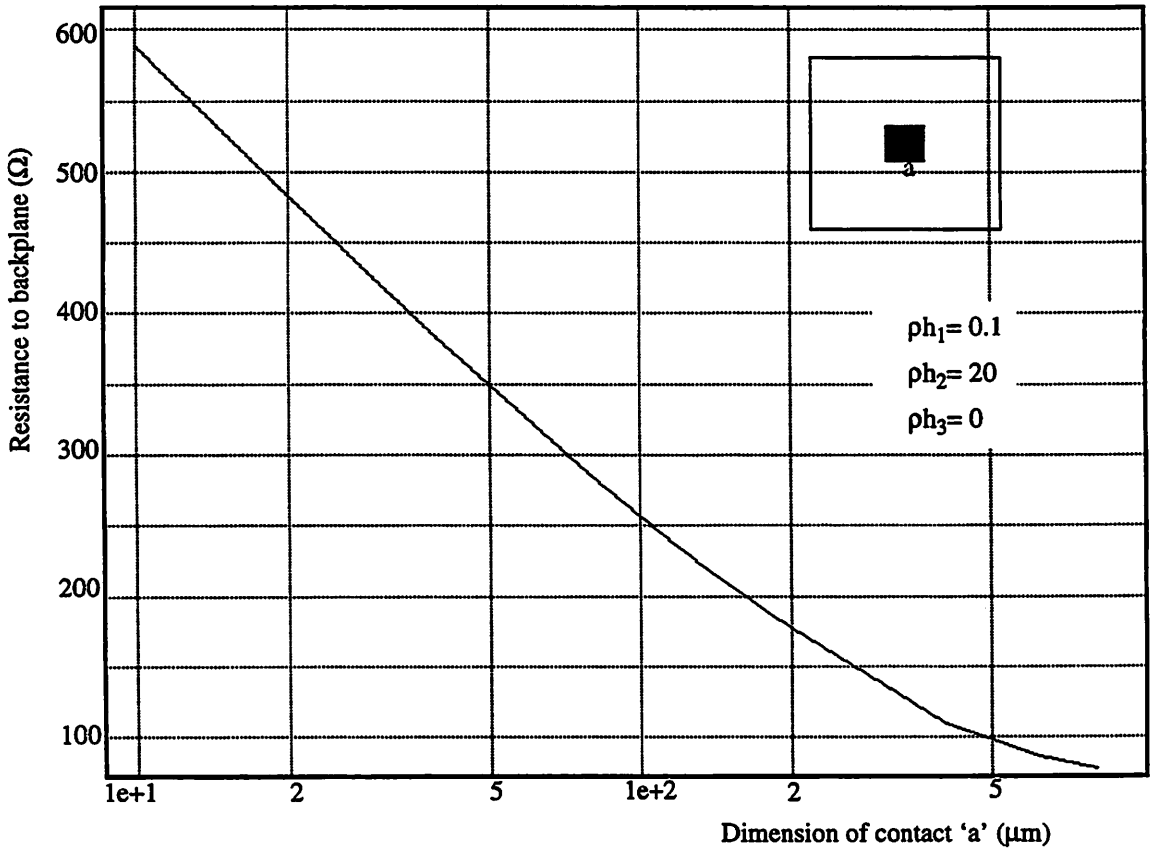


Figure 17a: Resistance to the Backplane in a High-Resistivity Substrate

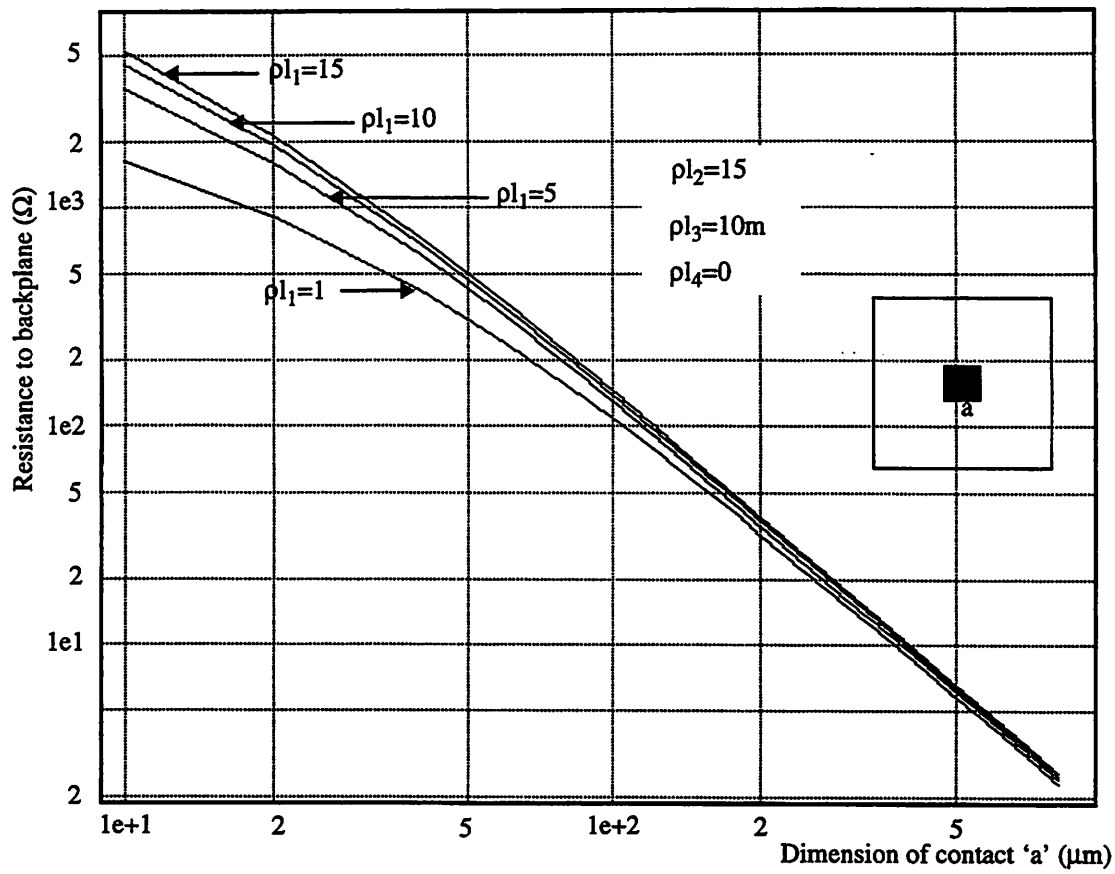


Figure 17b: Resistance to the Backplane in a Low-Resistivity Substrate

variation of the backplane resistance as a function of the contact dimension is shown in each figure.

The resistance in the high-resistivity substrate shows a weak logarithmic dependence on the dimension of the contact, while the resistance in the low-resistivity substrate is a much stronger function of the contact dimension ($\ln(R) \propto -\ln(a)$). In the case of the high-resistivity substrates, the presence of the low-resistivity epi-layer on the surface of the high-resistivity bulk region tends to increase the fringing fields. Thus the effective contact area appears to be much greater than the physical contact area. This explains the weak dependence of the resistance on the contact area. In the low-resistivity substrates, the presence of the thick low-resistivity bulk region very close to the surface reduces fringing fields. Thus a nearly inverse dependence of the contact resistance on the contact area is observed. The $\rho l_1=1$ case does

not satisfy the inverse dependence very well for small contact areas, since the presence of the low-resistivity epi-layer increases the fringing effect.

The dependence of three model-elements, in a two-contact problem, as a function of the distance between the contacts is shown in Fig. 18c-e. The layout of the problem is shown in Fig. 18a. Fig. 18b

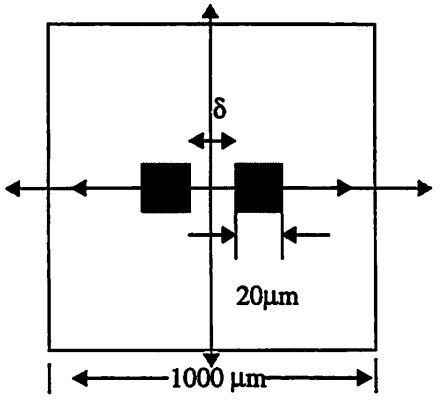


Figure 18a: Contact Layout

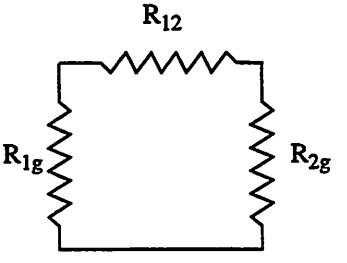


Figure 18b: Substrate Model

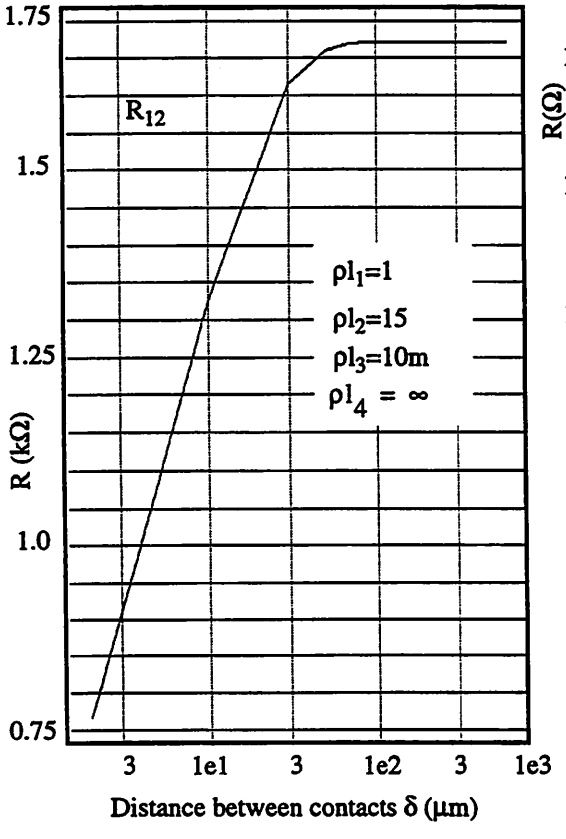


Figure 18c: Substrate without Backplane

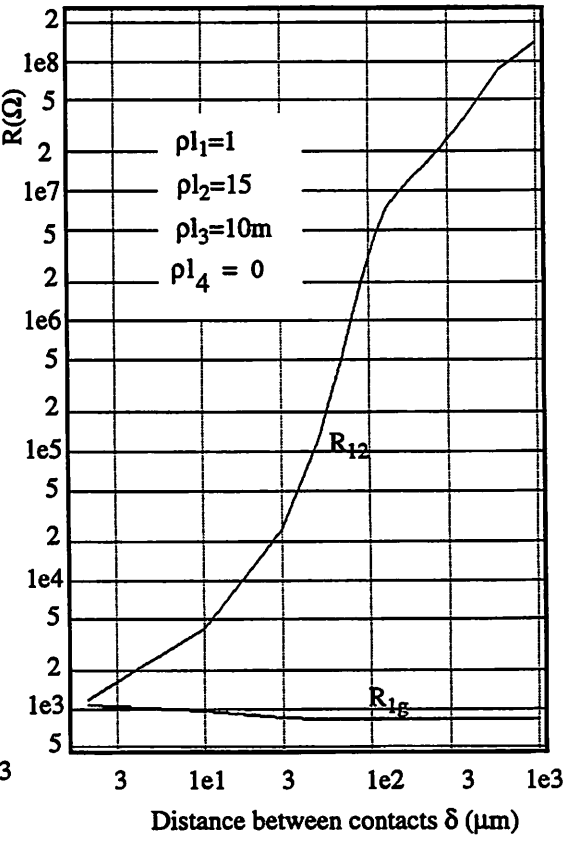


Figure 18d: Substrate with Backplane

depicts the model for the two-contact problem. The three resistance elements as a function of the distance between the contacts for the two substrate types are shown in Fig. 18c-e. Contacts 1 and 2, shown in Fig. 18a, are moved apart along a central axis as indicated in the figure.

The contacts are assumed to be equidistant from the center for all values of δ. Hence the resistors R_{1g} and R_{2g} in Fig. 18b are equal, and only one of them is shown in the figures. The element R_{12} in a low-resistivity substrate, with an infinitely large value of the epoxy resistivity, is shown in Fig. 18c [4]. R_{1g} and R_{2g} are omitted from this graph since they are infinitely large. R_{12} increases until the distance between the two contacts is about 40μm, and then becomes independent of distance. The current flow in this case is mostly through the low-resistivity bulk and not through the surface as the surface resistivity is high. Thus the behavior shown in Fig. 18c can be expected. The implications of this behavior for isolation are discussed later.

The model elements for a low-resistivity substrate with epoxy resistivity of zero are shown in Fig.

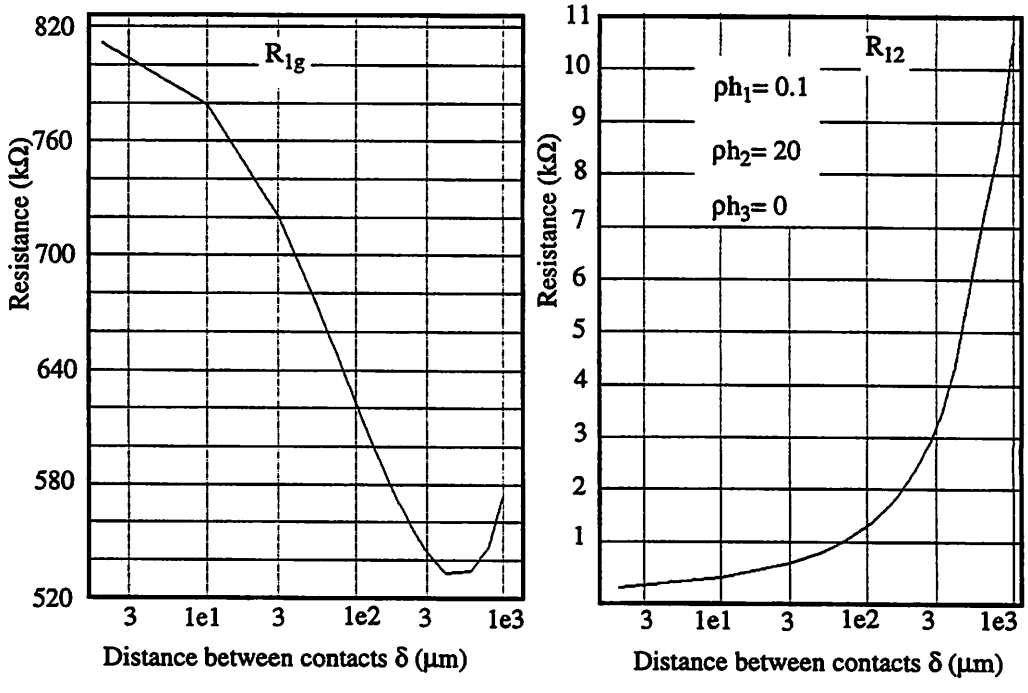


Figure 18e: Model-Element Variation in a High-Resistivity Substrate

4. An infinitely large epoxy resistivity is approximated in the tool by using a very large value of ρl_4 compared to $\rho l_{(1\text{-}3)}$.

18d. The value of R_{12} is seen to rise rapidly to a very large value, as a function of distance. The presence of the low-resistivity region very close to the surface leads to this dependence. The bulk region acts as a very good ground-plane and prevents lateral flow of substrate current. R_{1g} is small and asymptotes to a constant value after 50μm separation.

The behavior of the model elements for a high-resistivity substrate is shown in Fig. 18e. The resistance between the two contacts increases monotonically as can be expected. The resistance to the backplane is not a monotonic function of separation. It is large when the two contacts are close, which implies that a significant portion of the current flow is in the surface region. As the distance between the contacts increases, most of the current flow is to the backplane, which causes a decrease in the resistance to the backplane. When the contacts are brought near the edge of the substrate, the effective resistivity of the substrate increases, as the resistivity beyond the edge is infinite. This leads to the increase in the backplane resistance for a large δ in Fig. 18e.

The examples considered in this section are simple. These examples merely serve to illustrate that the behavior of the model elements is physically correct and can be intuitively explained.

4.0.2: Isolation between Single-Ended Contacts with Capacitive Injection and Reception Mechanisms

In this section, the isolation between single-ended contacts on the same substrate will be examined. Isolation between contacts is defined as the ratio of the voltage swing on the receiver contact to the voltage swing on the injector contact. The isolation between two contacts depends on several factors besides the substrate model elements. These include the contact-to-substrate capacitance, the backplane contact impedance, the frequency of operation and the load impedance connected to the receiver contact. The receiver contact load may be, for example, the load resistance connected to the collector of a BJT. The collector of the device acts as a receiver due to the collector-to-substrate capacitance. The load has a significant influence on the magnitude of the isolation since the voltage on the receiver node depends on the value of the load. The variation of substrate coupling as a function of the load impedance will be examined in Section 4.0.4. For determining all the other dependencies of substrate noise, the receiver impedance will be considered to be a 50Ω resistance. The definition of isolation is illustrated in Fig. 19a,

with all the above mentioned factors included. The backplane impedance has been assumed to be inductive. This represents the inductance of a bond-wire connected to the backplane to achieve a ground contact. In order to simulate substrates without a ground contact, ρh_3 and ρl_4 can be made very large compared to $\rho h_{(1\text{-}2)}$ and $\rho l_{(1\text{-}3)}$ respectively (Fig. 16). The surface-layout of the problem simulated is shown in Fig. 19b.

In order to be accurate, the bond-wire inductance between the node shown as contact-2 and the output at V_{rcv} must be included in the model. However, this study is simply behavioral. A complete model must be used for a specific simulation. By default, the two contacts are assumed to be squares of 50μm sides.

Low-Resistivity Substrates

The isolation between two surface contacts for a low-resistivity substrate (for $\rho l_4 = \infty$ and $\rho l_4 = 0$)[5] is shown in Fig. 19c and Fig. 19d for two values of the contact-to-substrate capacitance 0.3pF and 0.8pF respectively. If these junctions were formed by NMOS devices with a combined device drain and source area of 2500 μm^2, then the per unit area depletion capacitance values for these processes would be 1.2×10^{-4} and 3.2×10^{-4} F/m^2 respectively, which are fairly typical numbers. The parameters varied in

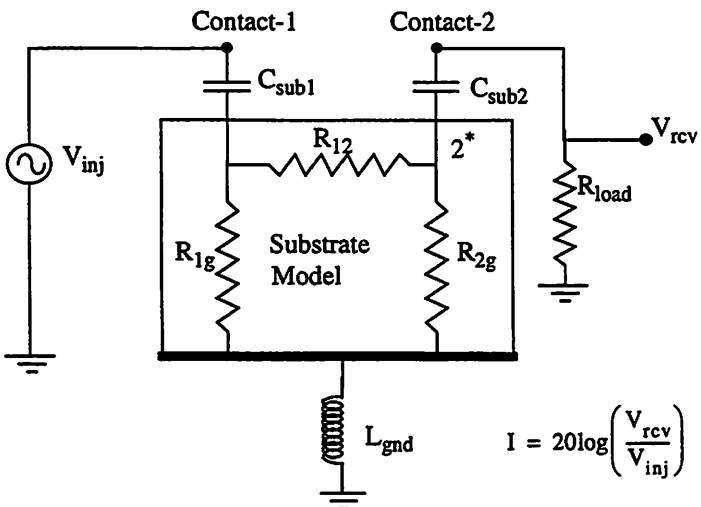


Figure 19a: Substrate Model

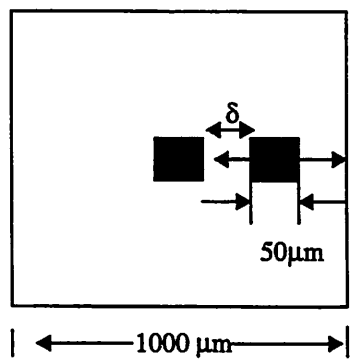


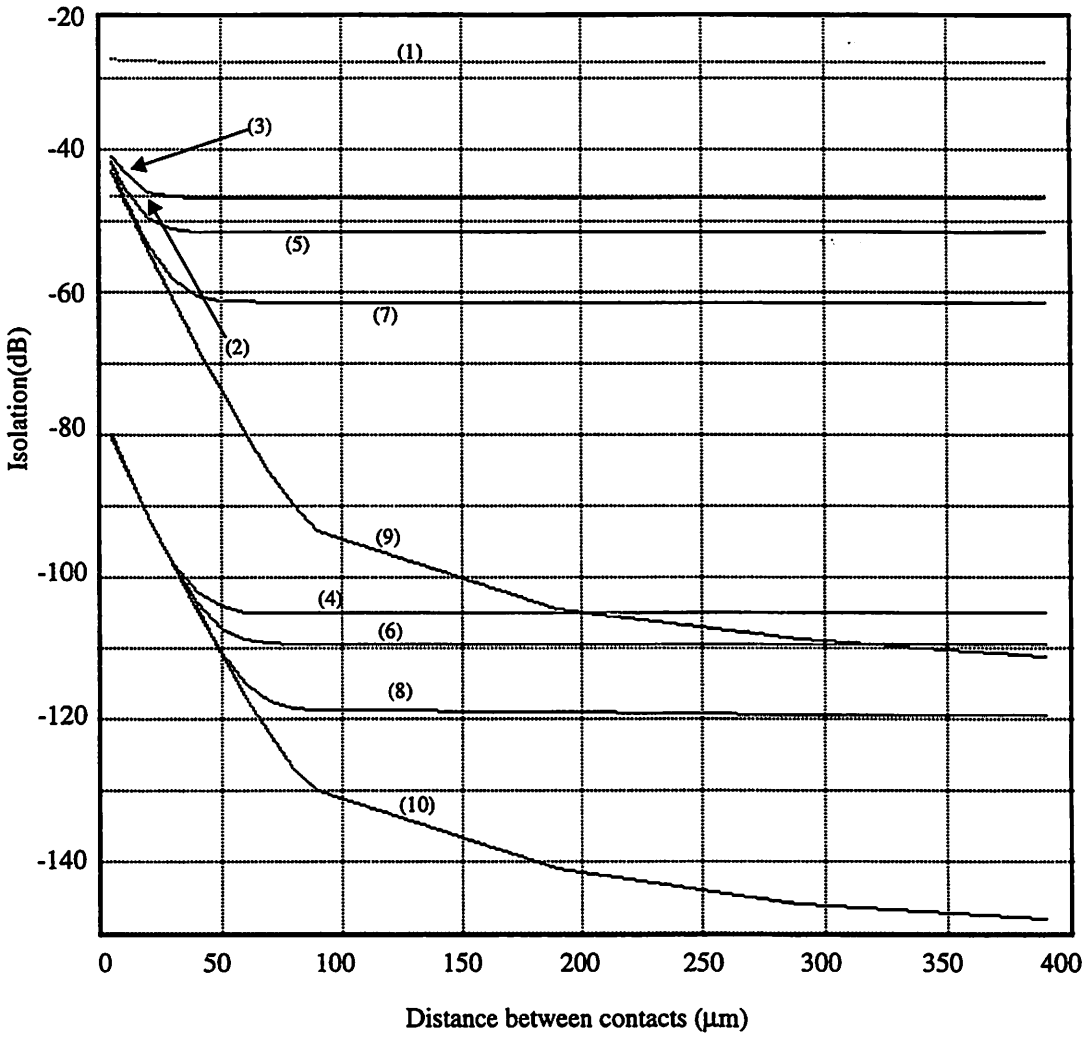
Figure 19b: Contact Layout

5. The case $\rho l_4 = \infty$ cannot be simulated by using a very large backplane impedance with the model values for the $\rho l_4 = 0$ case. The two problems have different boundary conditions. Setting $\rho l_4 = \infty$ simulates the absence of a backplane (Neumann boundary condition). In the latter case, the backplane is always an equipotential face, regardless of the value of L_{gnd}.

this study are L_{gnd}, the distance between the two nodes and the frequency (0.1GHz and 1GHz).

Curves (1) and (2), in Fig. 19c and Fig. 19d, represent the isolation for the case $\rho l_4 = \infty$ at 100MHz and 1GHz respectively. The isolation is seen to be almost independent of the distance between the two contacts. This is due to the fact that the resistance R_{12} is much smaller than the reactance of C_{sub1} and C_{sub2} at the frequencies shown here. Thus V_{rcv} is related to V_{inj} by the ratio of the resistor R_{12} and the reactance of the series combination of C_{sub1} and C_{sub2}. This is not a favorable environment for applications requiring high isolation since increasing the separation between the injector and receiver contacts is not helpful.

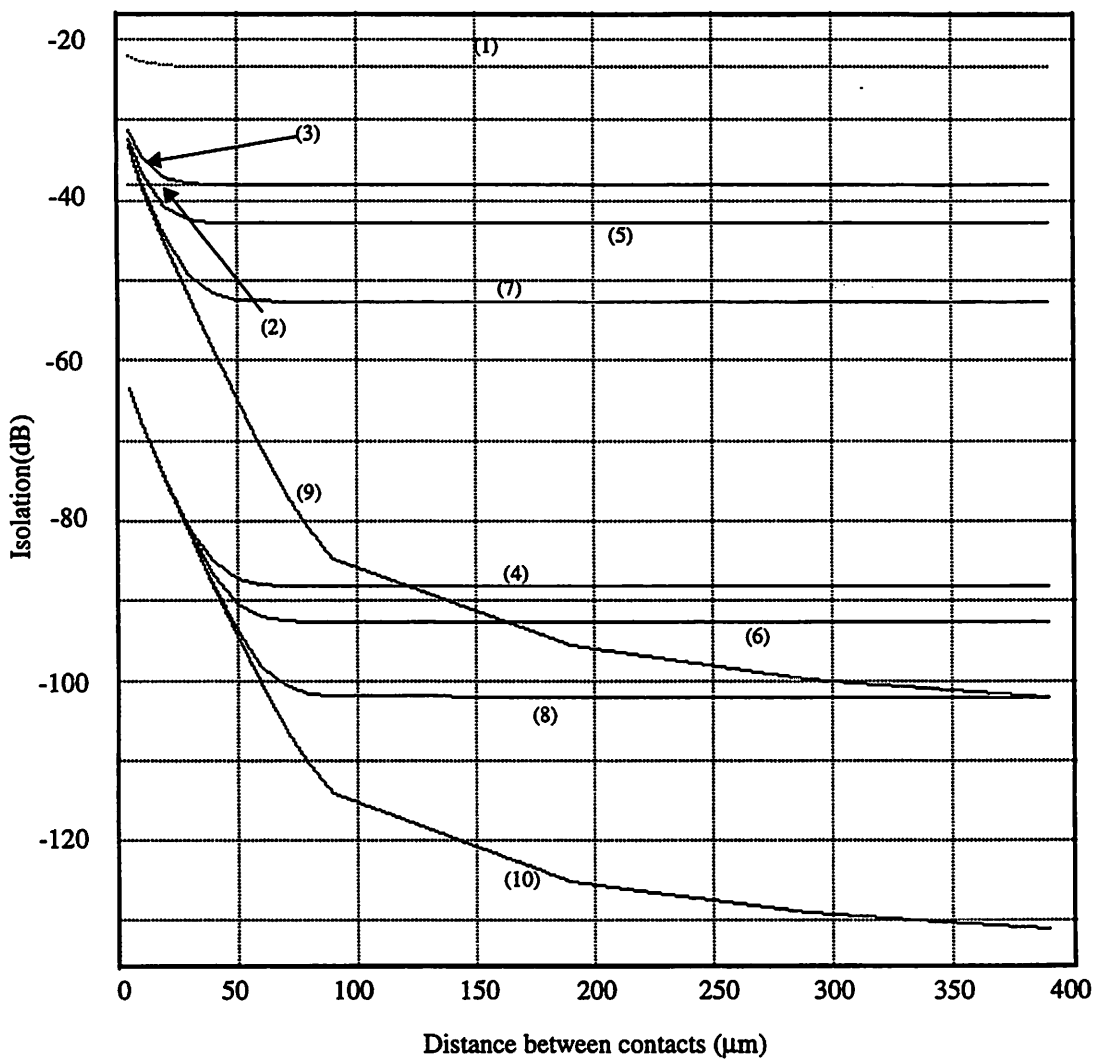
For the case $\rho l_4 = 0$ and $L_{gnd} \neq 0$, (curves 3-8), the isolation becomes independent of distance after a certain critical distance. Su et al. [21] found this critical distance to be four times the epitaxial thickness in their experiments. In fact, this distance is a function of the backplane impedance and varies from 2.5-5 times the epitaxial layer thickness (10μm) for L_{gnd} in the range of 1-5nH at 1GHz. For $L_{gnd} = 0$, the isolation continues to increase with increasing distance.



Curve (1):	ρl_1=1	ρl_2=15	ρl_3=10m	$\rho l_4 = \infty$	f=1GHz
Curve (2):	ρl_1=1	ρl_2=15	ρl_3=10m	$\rho l_4 = \infty$	f=0.1GHz
Curve (3):	ρl_1=1	ρl_2=15	ρl_3=10m	$\rho l_4 = 0$	f=1GHz, L_{gnd}=5nH
Curve (4):	ρl_1=1	ρl_2=15	ρl_3=10m	$\rho l_4 = 0$	f=0.1GHz, L_{gnd}=5nH
Curve (5):	ρl_1=1	ρl_2=15	ρl_3=10m	$\rho l_4 = 0$	f=1GHz, L_{gnd}=3nH
Curve (6):	ρl_1=1	ρl_2=15	ρl_3=10m	$\rho l_4 = 0$	f=0.1GHz, L_{gnd}=3nH
Curve (7):	ρl_1=1	ρl_2=15	ρl_3=10m	$\rho l_4 = 0$	f=1GHz, L_{gnd}=1nH
Curve (8):	ρl_1=1	ρl_2=15	ρl_3=10m	$\rho l_4 = 0$	f=0.1GHz, L_{gnd}=1nH
Curve (9):	ρl_1=1	ρl_2=15	ρl_3=10m	$\rho l_4 = 0$	f=1GHz, L_{gnd}=0
Curve (10):	ρl_1=1	ρl_2=15	ρl_3=10m	$\rho l_4 = 0$	f=0.1GHz, L_{gnd}=0

C_{sub}=0.3pF

Figure 19c: Variation of Isolation in Low-Resistivity Substrates



Curve (1):	ρl_1=1	ρl_2=15	ρl_3=10m	$\rho l_4 = \infty$	f=1GHz
Curve (2):	ρl_1=1	ρl_2=15	ρl_3=10m	$\rho l_4 = \infty$	f=0.1GHz
Curve (3):	ρl_1=1	ρl_2=15	ρl_3=10m	$\rho l_4 = 0$	f=1GHz, L_{gnd}=5nH
Curve (4):	ρl_1=1	ρl_2=15	ρl_3=10m	$\rho l_4 = 0$	f=0.1GHz, L_{gnd}=5nH
Curve (5):	ρl_1=1	ρl_2=15	ρl_3=10m	$\rho l_4 = 0$	f=1GHz, L_{gnd}=3nH
Curve (6):	ρl_1=1	ρl_2=15	ρl_3=10m	$\rho l_4 = 0$	f=0.1GHz, L_{gnd}=3nH
Curve (7):	ρl_1=1	ρl_2=15	ρl_3=10m	$\rho l_4 = 0$	f=1GHz, L_{gnd}=1nH
Curve (8):	ρl_1=1	ρl_2=15	ρl_3=10m	$\rho l_4 = 0$	f=0.1GHz, L_{gnd}=1nH
Curve (9):	ρl_1=1	ρl_2=15	ρl_3=10m	$\rho l_4 = 0$	f=1GHz, L_{gnd}=0
Curve (10):	ρl_1=1	ρl_2=15	ρl_3=10m	$\rho l_4 = 0$	f=0.1GHz, L_{gnd}=0

C_{sub}=0.8pF

Figure 19d: Variation of Isolation in Low-Resistivity Substrates

The isolation is very sensitive to the backplane impedance in these substrates. For a constant contact-to-contact distance, the variation in the substrate isolation is very large over the range of backplane inductance considered. Providing a small backplane impedance is critical in these substrates. If we assume that the reactance of the C_{sub1} and C_{sub2} is small, the load resistance is small and R_{12} can be ignored (Fig. 18d), then the isolation between the contacts is governed by the following impedance

$$Z_{12}'(j\omega) = R_{1g} + R_{2g} - j\frac{R_{1g} \times R_{2g}}{\omega L_{gnd}} \qquad (81)$$

This is the impedance looking into the substrate between contacts 1 and 2 (Fig. 19e). The assumptions made here are not restrictive and are applicable in many cases where substrate coupling is significant. The sensitivity $S_{L_{gnd}}^{|Z_{12}|}$ is given by

$$S^{|Z_{12}|}_{L_{gnd}} = \frac{((\Delta|Z_{12}|)/|Z_{12}|)}{((\Delta L_{gnd})/L_{gnd})} = -\frac{(R_{1g} \| R_{2g})^2}{(R_{1g} \| R_{2g})^2 + (\omega L_{gnd})^2} \qquad (82)$$

where $(R_{1g} \| R_{2g}) = (R_{1g} \times R_{2g})/(R_{1g} + R_{2g})$.

The sensitivity is large for $R_{1g} \| R_{2g} \gg (\omega L_{gnd})$ which implies that for small values of L_{gnd}, a small change in L_{gnd} will cause a large fractional change in the isolation. This behavior is evident in Fig. 19c and

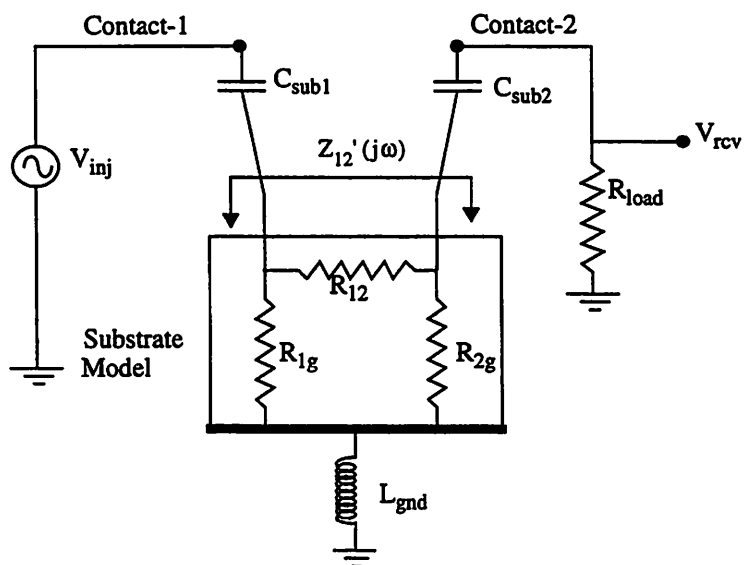


Figure 19e: Equivalent Substrate-Impedance Model

Fig. 19d.

The backplane inductance becomes a more severe problem for isolation between large contacts. This is so because the resistance to the backplane falls almost inversely with contact area. Thus $Z'_{12}(j\omega)$ at high frequencies falls inversely with area too which implies that the high-frequency isolation worsens almost linearly with area. The reactance of C_{sub} also falls as the inverse of contact area. Thus the isolation will degrade considerably with area in these substrates.

The isolation values are seen to be very good for small ground-plane inductance values (curves (7)-(10)), especially at 100MHz. This is expected since as discussed in the previous section the low-resistivity bulk restricts lateral current flow.

Comparison of Fig. 19c and Fig. 19d shows that the isolation degrades as the capacitance C_{sub} increases.

High-Resistivity Substrates

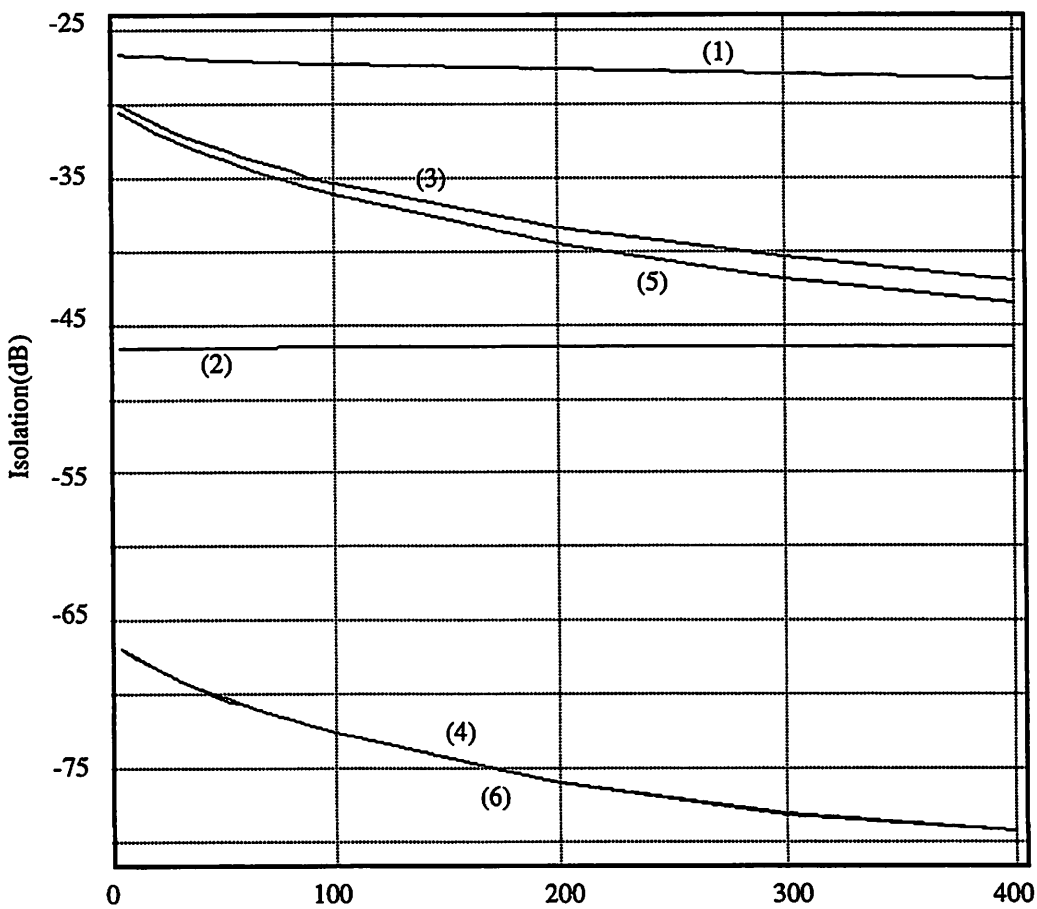
The isolation between two surface contacts for a high-resistivity substrate (for $\rho h_3 = \infty$ and $\rho h_3 = 0$) is shown in Fig. 19f and Fig. 19g. C_{sub} has been set to 0.3pF in Fig. 19f and to 0.8pF in Fig. 19g.

Curves (1) and (2) depict the isolation for $\rho h_3 = \infty$. In this case, the isolation is seen to be weakly dependent on the distance between the contacts at 1GHz. This is so because the value of R_{12} is comparable to the reactance of C_{sub} at this frequency. At 100MHz, the capacitive reactance is much larger than R_{12}. Thus the isolation is independent of distance as in the previous section.

For $\rho h_3 = 0$ the isolation is seen to be weakly dependent on the value of the backplane inductance, unlike that in the low-resistivity substrates. This is due to the significant surface conduction in these substrates. Curves (4) and (6) almost coincide in both Fig. 19f and Fig. 19g, which implies that almost all of the coupling takes place through the surface-layers. Since a significant part of the current flows at the surface, substrate taps can be expected to be more efficient in these substrates compared to the low-resistivity substrates. This is indeed the case as will be discussed later.

The effective substrate impedance $Z'_{12}(j\omega)$ in these substrates is a much weaker function of contact area. Thus the isolation does not degrade as rapidly as in the previous case with area. This is especially true for larger areas, when the capacitive reactance of C_{sub} is small, and the isolation between the two contacts is dominated by $Z'_{12}(j\omega)$.

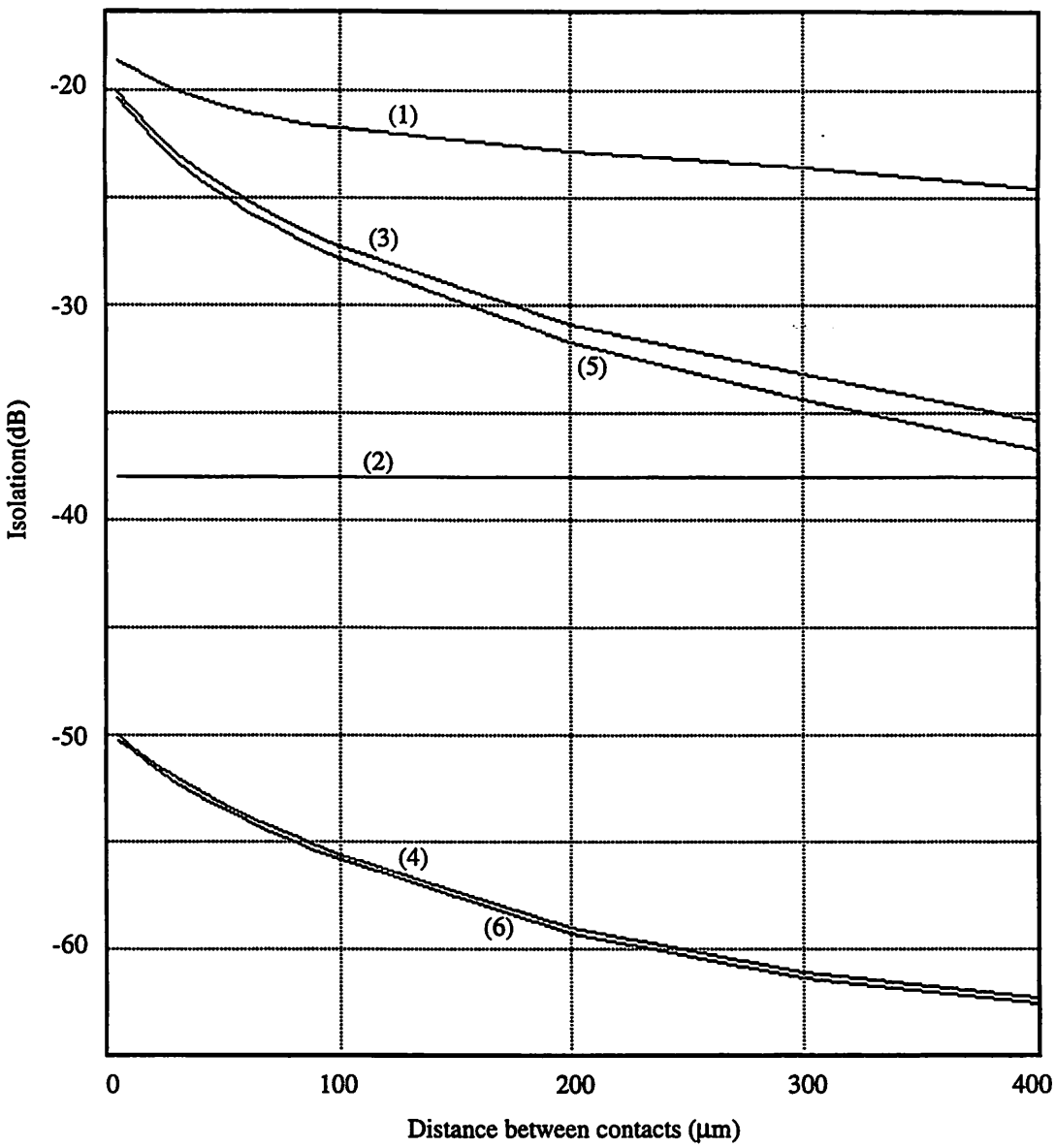
Another consequence of the surface conduction is that the isolation continues to improve as the distance between contacts increases, even for finite values of the backplane impedance.



Curve (1):	ρh_1=0.1	ρh_2=20	$\rho h_3 = \infty$	f=1GHz
Curve (2):	ρh_1=0.1	ρh_2=20	$\rho h_3 = \infty$	f=0.1GHz
Curve (3):	ρh_1=0.1	ρh_2=20	$\rho h_3 = 0$	f=1GHz, L_{gnd}=5nH
Curve (4):	ρh_1=0.1	ρh_2=20	$\rho h_3 = 0$	f=0.1GHz, L_{gnd}=5nH
Curve (5):	ρh_1=0.1	ρh_2=20	$\rho h_3 = 0$	f=1GHz, L_{gnd}=3nH
Curve (6):	ρh_1=0.1	ρh_2=20	$\rho h_3 = 0$	f=0.1GHz, L_{gnd}=3nH

C_{sub}=0.3pF

Figure 19f: Isolation in High-Resistivity Substrates



Curve (1):	ph_1=0.1	ph_2=20	$ph_3 = \infty$	f=1GHz
Curve (2):	ph_1=0.1	ph_2=20	$ph_3 = \infty$	f=0.1GHz
Curve (3):	ph_1=0.1	ph_2=20	$ph_3 = 0$	f=1GHz, L_{gnd}=5nH
Curve (4):	ph_1=0.1	ph_2=20	$ph_3 = 0$	f=0.1GHz, L_{gnd}=5nH
Curve (5):	ph_1=0.1	ph_2=20	$ph_3 = 0$	f=1GHz, L_{gnd}=3nH
Curve (6):	ph_1=0.1	ph_2=20	$ph_3 = 0$	f=0.1GHz, L_{gnd}=3nH

C_{sub}=0.8pF

Figure 19g: Isolation in High-Resistivity Substrates

4.0.3: Improvement in Isolation in Differential Circuits with Capacitive Injection and Reception Mechanisms

Differential circuits are often preferred over single-ended circuits in noisy environments. This is so because the noise appears as a common-mode signal on the differential outputs. The differential noise signal is typically several orders of magnitude smaller than what would be observed in a single-ended implementation of the circuit. The use of differential circuits in applications requiring a high power-supply-rejection-ratio (PSRR) is a typical example.

In this section, the use of differential circuits to reduce substrate noise is discussed. As in the previous section, the two types of substrates shown in Fig. 16 are used to study the improvement in substrate noise isolation by the use of differential circuits.

The layout of the structure considered is shown in Fig. 20a below. The differential circuit consists of two $200 \times 10\,\mu m^2$ contacts placed at the center of the substrate-surface. The surface injector is a single $50 \times 50\,\mu m^2$ contact. The injector contact is varied along an axis as shown in the figure. The equivalent circuit model is shown in Fig. 20b. If the injector contact is moved along the central Y-axis instead of the central X-axis, the circuit will be perfectly balanced with respect to the substrate, and the differential

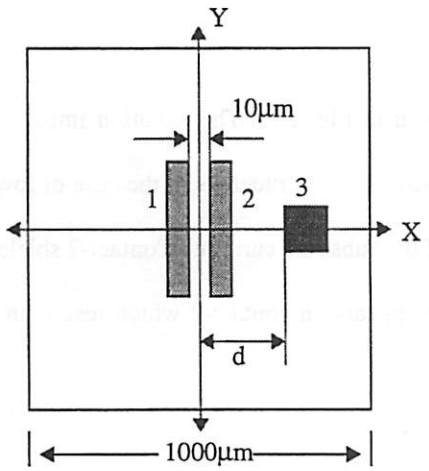
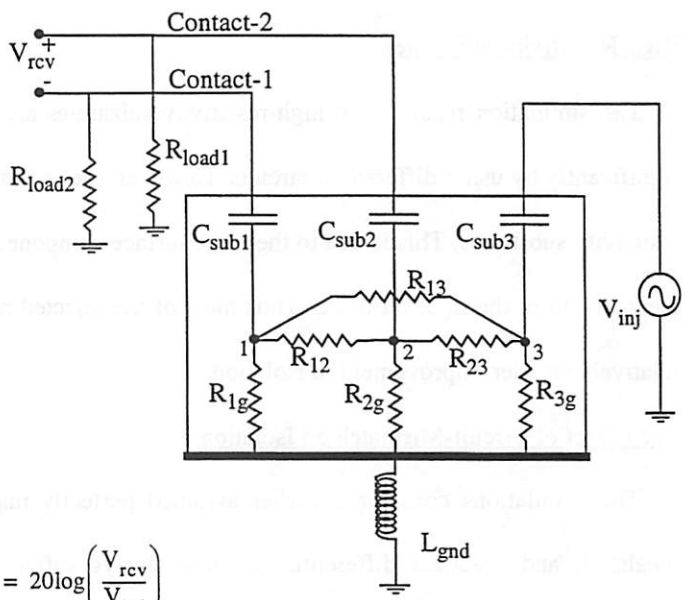


Figure 20a: Differential Receiver Layout



$$I = 20\log\left(\frac{V_{rcv}}{V_{inj}}\right)$$

Figure 20b: Substrate Model

substrate-noise-isolation will be infinite. The asymmetric axis-orientation used in this simulation results in the worst-case differential substrate-noise-isolation. The differential receiver contacts in Fig. 20b are assumed to be identical.

Low-Resistivity Substrates

The results of the simulations, for a low-resistivity substrate, are shown in Fig. 20c. Two substrates with $\rho l_4 = \infty$ and $\rho l_4 = 0$ were used in this simulation. The isolation in Fig. 20c falls rapidly with distance between the contacts and is insensitive to the backplane impedance. As the distance between the contacts increases, the direct coupling terms R_{23} and R_{13} increase rapidly in curves (3-10). Thus the coupling takes place almost entirely through the low-resistivity bulk region. As discussed earlier, the presence of the low-resistivity bulk region close to the surface reduces the fringing fields. Another consequence of the reduction in the fringing effect is that the resistances to the backplane are almost independent of the location of the contact on the surface of the substrate. Two contacts with the same area will have nearly the same resistance to the backplane, regardless of their contact coordinates. Consequently, voltage excursions of the bulk region or the backplane are conveyed to the two contacts of the differential circuit with an equal amplitude. This leads to the excellent differential isolation seen in Fig. 20c.

High-Resistivity Substrates

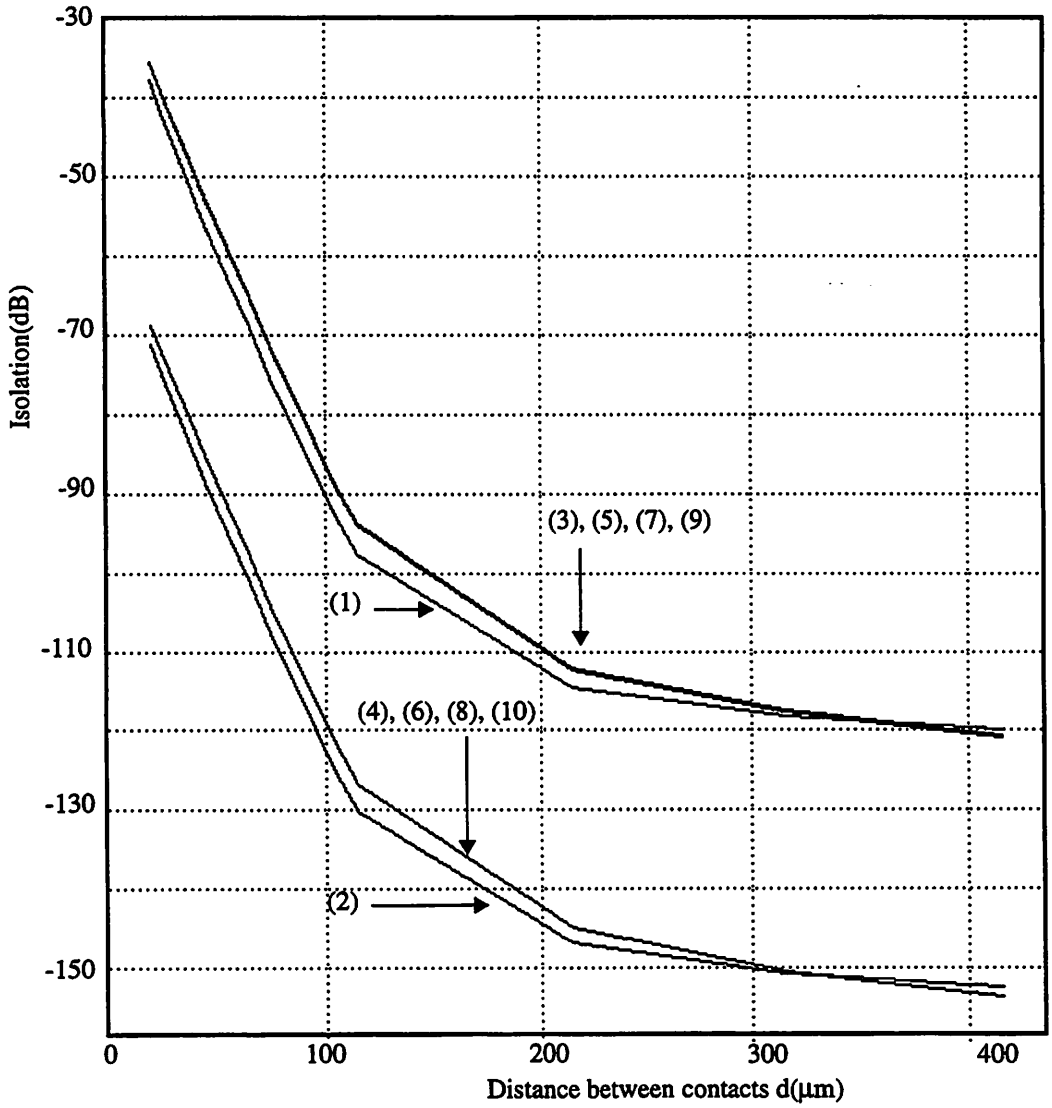
The simulation results with high-resistivity substrates are shown in Fig. 20d. The isolation improves significantly by using differential circuits. However, the improvement is not as much as in the case of low-resistivity substrates. This is due to the large surface component of the substrate currents. Contact-2 shields contact-1 from the injected noise. Thus most of the injected noise appears on contact-2 which results in a relatively smaller improvement in isolation.

The Effect of Circuit-Mismatch on Isolation

The simulations considered earlier assumed perfectly matched differential circuits. This situation is idealized, and practical differential circuits always suffer from component mismatch. The effect of component mismatch on differential-circuit substrate-isolation is presented in Fig. 20e in low-resistivity

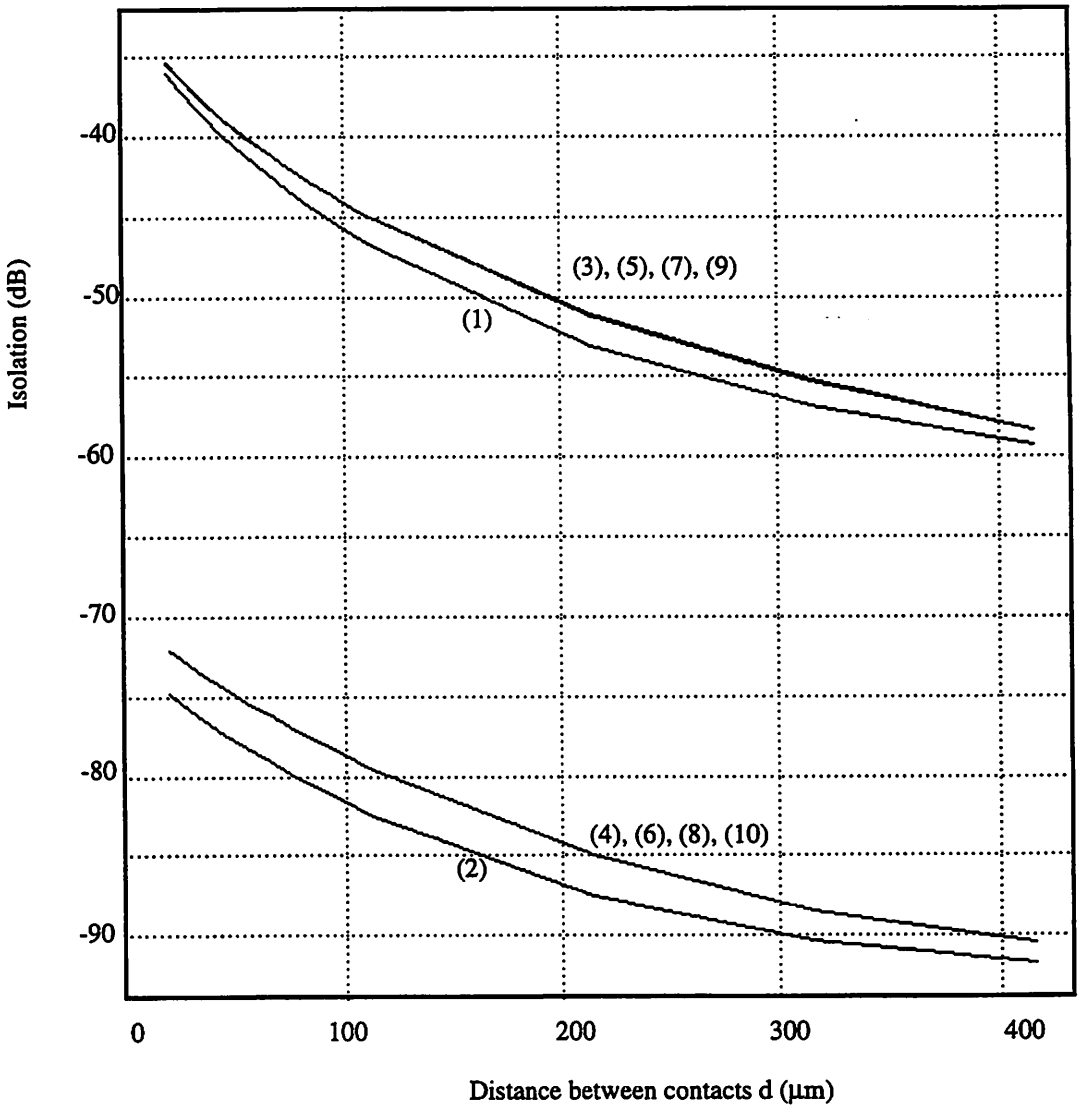
substrates. The layout of Fig. 20a is considered but the substrate-capacitors in the receiver-pair are assumed to be mismatched by five percent.

It can be observed from Fig. 20e that a slight mismatch in the differential pair can significantly reduce the isolation. The degradation is especially severe for distant contacts. The isolation in Fig. 20c is very large due to the idealized situation assumed there. Mismatches between circuit elements lead to common-mode substrate injection and reception, which make differential circuits exhibit single-ended behavior.



Curve (1):	ρl_1=1	ρl_2=15	ρl_3=10m	$\rho l_4 = \infty$	f=1GHz
Curve (2):	ρl_1=1	ρl_2=15	ρl_3=10m	$\rho l_4 = \infty$	f=0.1GHz
Curve (3):	ρl_1=1	ρl_2=15	ρl_3=10m	$\rho l_4 = 0$	f=1GHz, L_{gnd}=5nH
Curve (4):	ρl_1=1	ρl_2=15	ρl_3=10m	$\rho l_4 = 0$	f=0.1GHz, L_{gnd}=5nH
Curve (5):	ρl_1=1	ρl_2=15	ρl_3=10m	$\rho l_4 = 0$	f=1GHz, L_{gnd}=3nH
Curve (6):	ρl_1=1	ρl_2=15	ρl_3=10m	$\rho l_4 = 0$	f=0.1GHz, L_{gnd}=3nH
Curve (7):	ρl_1=1	ρl_2=15	ρl_3=10m	$\rho l_4 = 0$	f=1GHz, L_{gnd}=1nH
Curve (8):	ρl_1=1	ρl_2=15	ρl_3=10m	$\rho l_4 = 0$	f=0.1GHz, L_{gnd}=1nH
Curve (9):	ρl_1=1	ρl_2=15	ρl_3=10m	$\rho l_4 = 0$	f=1GHz, L_{gnd}=0
Curve (10):	ρl_1=1	ρl_2=15	ρl_3=10m	$\rho l_4 = 0$	f=0.1GHz, L_{gnd}=0

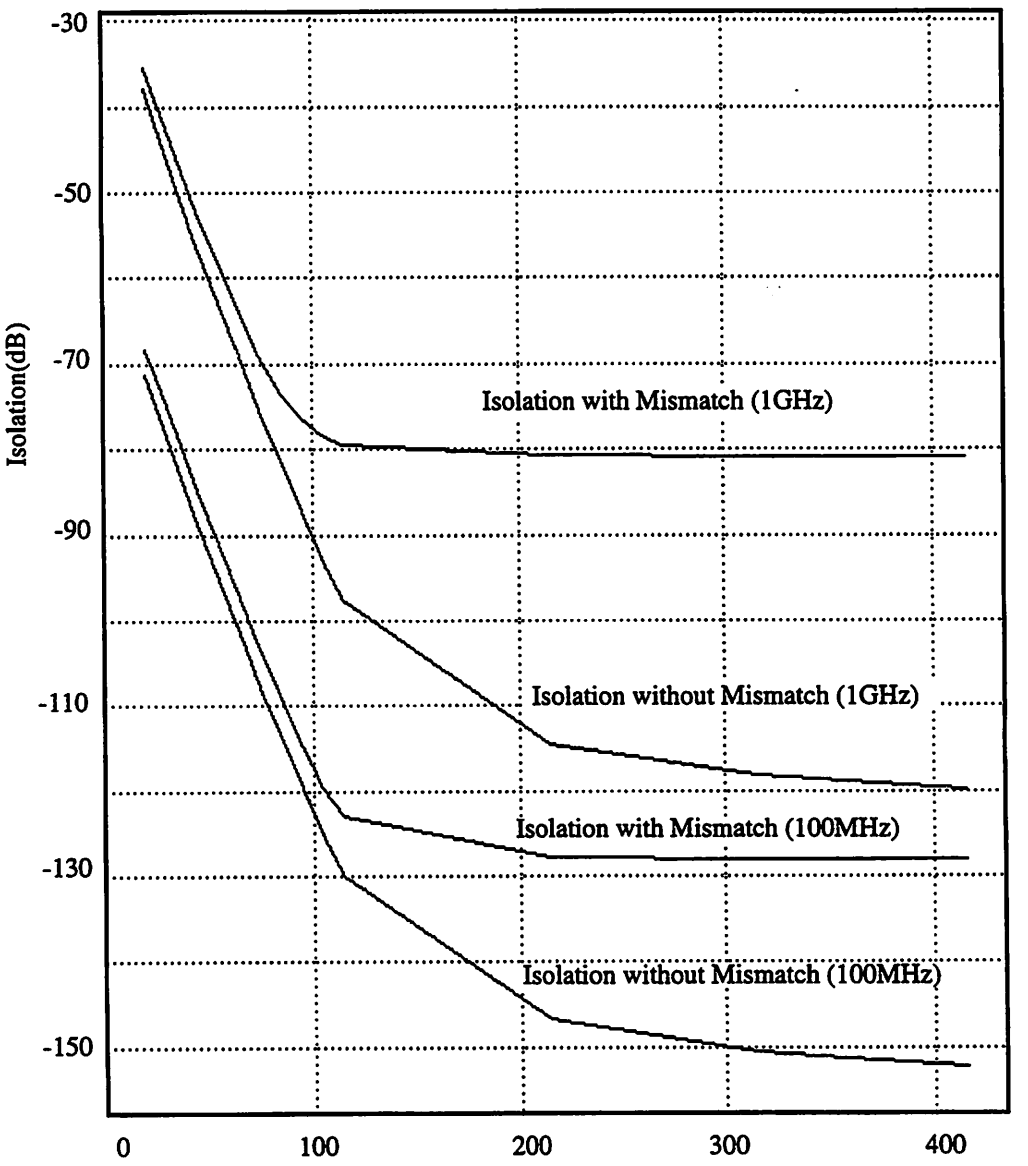
Figure 20c: Differential Isolation in Low-Resistivity Substrates



Curve (1):	ρh_1=0.1	ρh_2=20	$\rho h_3 = \infty$	f=1GHz
Curve (2):	ρh_1=0.1	ρh_2=20	$\rho h_3 = \infty$	f=0.1GHz
Curve (3):	ρh_1=0.1	ρh_2=20	$\rho h_3 = 0$	f=1GHz, L_{gnd}=5nH
Curve (4):	ρh_1=0.1	ρh_2=20	$\rho h_3 = 0$	f=0.1GHz, L_{gnd}=5nH
Curve (5):	ρh_1=0.1	ρh_2=20	$\rho h_3 = 0$	f=1GHz, L_{gnd}=3nH
Curve (6):	ρh_1=0.1	ρh_2=20	$\rho h_3 = 0$	f=0.1GHz, L_{gnd}=3nH
Curve (7):	ρh_1=0.1	ρh_2=20	$\rho h_3 = 0$	f=1GHz, L_{gnd}=1nH
Curve (8):	ρh_1=0.1	ρh_2=20	$\rho h_3 = 0$	f=0.1GHz, L_{gnd}=1nH
Curve (9):	ρh_1=0.1	ρh_2=20	$\rho h_3 = 0$	f=1GHz, L_{gnd}=0nH
Curve (10):	ρh_1=0.1	ρh_2=20	$\rho h_3 = 0$	f=0.1GHz, L_{gnd}=0nH

C_{sub}=0.8pF

Figure 20d: Differential Isolation in High-Resistivity Substrates



$\rho l_1=1 \quad \rho l_2=15 \quad \rho l_3=10m \quad \rho l_4=0$

$L_{gnd}=1nH$

Figure 20e: Effect of Mismatch on Differential Isolation

4.0.4: Dependence of Substrate Coupling on Load Impedance

In all of the earlier examples the load impedance at the receiver has been assumed to be 50Ω. The load impedance on-chip may be lower or higher. In this section, an earlier simulation is repeated with different values of the load resistance. The example shown here treats only the low-resistivity substrates, as the observations made here are valid for both types of substrates. Fig. 21 shows the isolation between contacts 1 and 2 in Fig. 19a for a fixed distance of 90μm between the contacts. The L_{gnd} is fixed at 3nH and C_{sub} at 0.8pF.

Usually the load resistance value is chosen such that the capacitive reactance of the device-to-substrate capacitance is larger than the load in the frequency band of interest[6]. In other words, the pole frequency at the load due to the substrate capacitance is much higher than the highest frequency of interest. In such a case, it is reasonable to assume that the load impedance does not effect the voltage division ratio in the

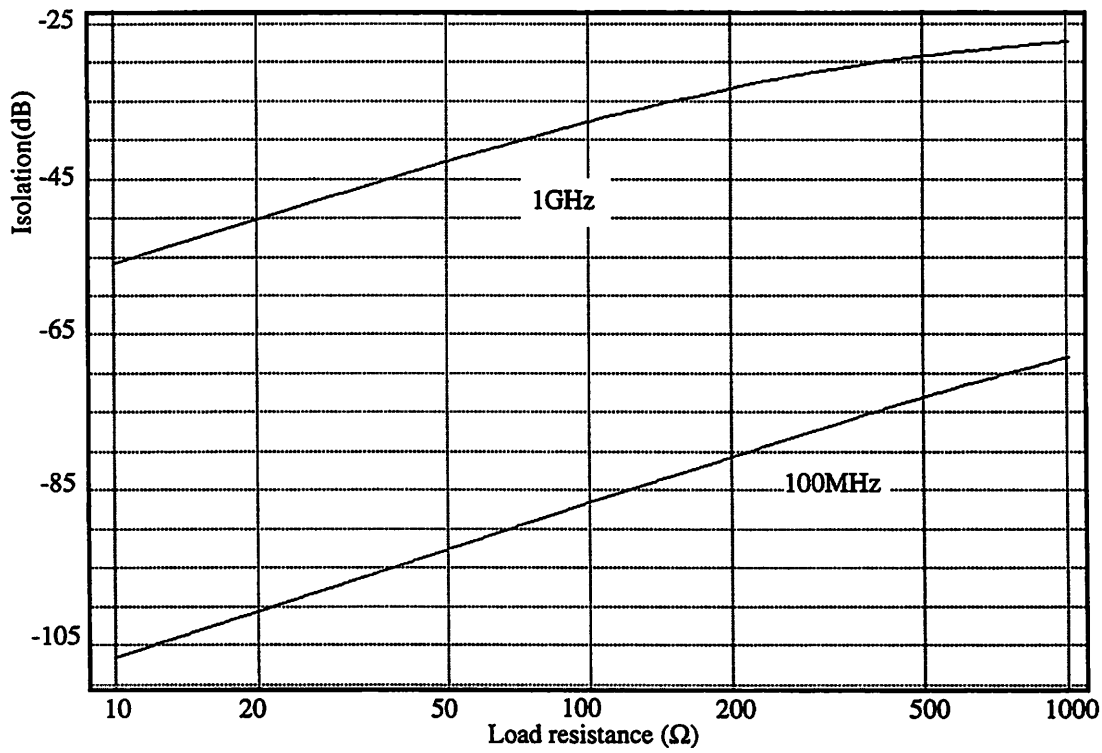


Figure 21: Variation of Isolation with Load-Resistance

6. The impedance of the load is the effective value of impedance seen looking into the device from the substrate. This definition of load includes the effective value of the load, including the effect of any feedback which may be present around the circuit.

substrate macromodel. Referring to Fig. 19a, we observe that the voltage at the internal substrate node (2*) is independent of the load impedance to the first order. The voltage at the load (V_{rcv}) can be calculated using the voltage division ratio of the load resistance and the substrate capacitance. If the above assumption regarding the load pole frequency is used, then V_{rcv} for different load values can be calculated by simply changing the voltage division ratios between nodes 2^* and contact-2 in Fig. 19a. This implies that for most frequencies of interest the behavior of substrate noise is the same as in Fig. 19c-g, scaled up or down by the load value.

The above approximation is valid if the received substrate noise is in the same frequency band as the circuit itself. If the frequency of the substrate noise signal is higher than the receiver circuit's load-C_{sub} pole, then the value of the load changes the degree of coupling. This behavior is seen in Fig. 21. The pole frequency with the highest load value of 1kΩ is approximately 200MHz. Thus at 100MHz the isolation scales linearly with the load resistance. This is not the case at 1GHz. The pole frequency equals 1GHz for a load resistance of 200Ω. Deviations from a linear increase at 1GHz are seen in Fig. 21 for load resistances greater than 200Ω.

4.0.5: Behavior of Substrate Coupling with Non-Capacitive Reception and Injection Mechanisms

The behavior of substrate coupling as a function of distance, L_{gnd}, and frequency which has been studied in the previous sections assumes capacitive injection and reception mechanisms. In this section, the behavior of substrate coupling with non-capacitive means of reception and injection will be considered. Such mechanisms include the body effect and hot-carrier injection due to avalanching in MOSFETs.

A linearized model of an NMOS device including the body effect is shown in Fig. 22a. The substrate model remains the same regardless of the mechanism of injection, since the carrier transport is by drift, and the medium can be considered to be a multilayered impedance.

The signal path from the body to the drain is traced in Fig. 22b, with the gate connected to ground. The gain from V_b to V_d is given by the following equation

$$\frac{V_d}{V_b} = \frac{(C_{db}s - g_{mb})R_d}{1 + R_d C_{db} s} \tag{83}$$

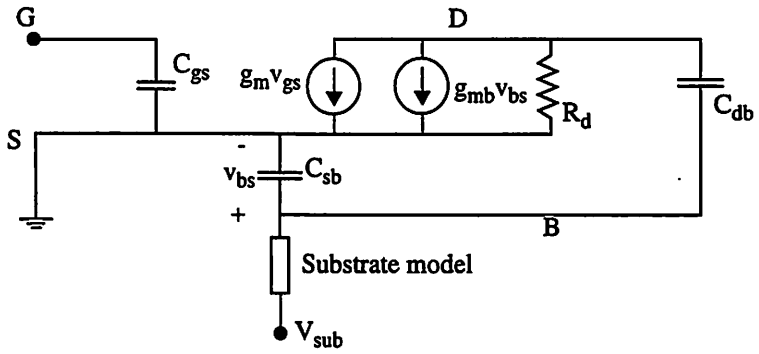


Figure 22a: Small-Signal Model of a MOSFET

The input impedance looking into node-B is given by

$$Z_{in} = \frac{1}{C_{db}s\left(\frac{1 + g_{mb}R_d}{1 + R_d C_{db}s}\right) + C_{sb}s} \tag{84}$$

In the absence of the body effect, the noise reception mechanism would be identical to that discussed in earlier sections. With a nonzero body effect parameter g_{mb}, the only difference that will be observed in the reception of substrate noise is that there will be noise-reception at very low frequencies as well. The behavior of substrate noise with respect to the other parameters (distance, area and backplane impedance) will be the same as the capacitive case.

As above, the substrate model remains unchanged regardless of the capacitive or non-capacitive nature of the injection mechanism. The only difference between these mechanisms is that the non-capacitive injection mechanisms are effective at DC and low-frequencies as well.

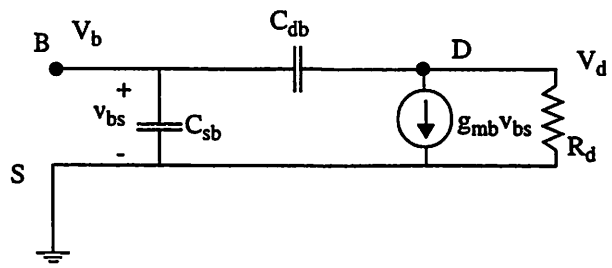


Figure 22b: Body-to-Drain Gain in a MOSFET

MOS devices possess non-capacitive injection and reception mechanisms. Therefore substrate-coupled noise can be important at low-frequencies too. In fact it is possible for the MOS device to be more sensitive to substrate noise at low-frequencies than at high frequencies. The input impedance looking into node-B in Fig. 22b is very large at low frequencies. Thus the voltage division ratios inside the substrate are not influenced by the device at low-frequencies. The factor g_{mb} can also be large enough that the substrate-to-drain gain is significant ($g_{mb}R_d$). As in the earlier cases, differential circuit topologies and good backplane contacts are effective techniques for noise reduction.

Non-capacitive injection and reception mechanisms are not a major problem in bipolar technology since the only significant injection and reception mechanisms are capacitive. The parasitic pnp transistor in Fig. 2a could become forward-active and inject noise into the substrate (Chap. 2), but that can be avoided by proper biasing.

4.1: Guard Rings and Substrate Taps

4.1.1: Guard Rings in Different Substrate Types

Guard rings and substrate taps are often used to reduce substrate coupling. In this section, the effectiveness of guard rings in different substrate types will be discussed. General guidelines for effective guard ring layout will be proposed.

A guard ring is shown in Fig. 23a. The ring is a surface-region heavily doped with the majority-carrier dopant and is intended to form a Faraday shield around any sensitive devices which need to be protected from substrate noise. The ring could also be placed around the noise injectors. The operation of the guard ring is shown in Fig. 23b. The model for the substrate with and without the guard ring is shown. The guard ring is effectively a current sink. By placing a current sink between the two contacts, the direct coupling model element R_{12} between contacts 1 and 2 is increased. Thus with the ring connected to ground, the isolation between the two contacts increases. L_{gr} represents the inductance of the bond-wire and the pin used to ground the ring. L_{gnd} is the backplane impedance connected to the substrate.

Contact-2 is effectively shielded from contact-1 by the ring, if the presence of the ring makes R_{12} large and the reactances of L_{gnd} and L_{gr} (X_{gnd} and X_{gr} respectively) are small at the frequency of interest.

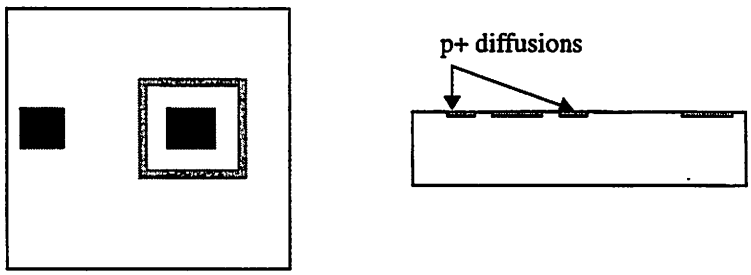


Figure 23a: Guard Ring Layout and Cross-Section

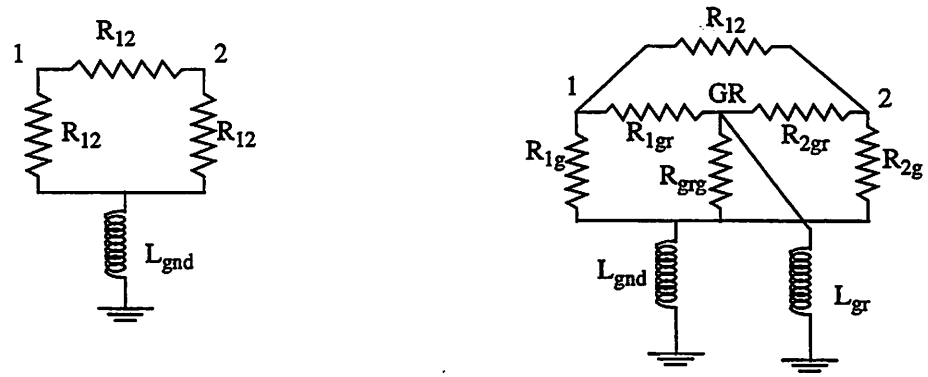


Figure 23b: Equivalent Substrate Model without and with a Guard Ring

If X_{gnd} is large, then the current flow through the backplane can dominate and if X_{gr} is large, then the dominant current flow will be through the resistors R_{1gr} and R_{2gr}. Thus care must be taken to keep these inductances small for the guard ring to be useful.

Guard Rings in Low-Resistivity Substrates

The effect of placing a 10μm wide guard ring between two square contacts of area 2500μm^2 is examined here. The layout is shown in Fig. 23c. In Fig. 23d, the isolation between the two contacts as a function of distance is examined at two frequencies 100MHz and 1GHz. This simulation is similar to the simulations presented in Section 4.0.2. C_{sub} has been chosen to be 0.8pF. L_{gr} is stepped between 0 and 3nH, while L_{gnd} is varied from 0 to 5nH.

Fig. 23d and Fig. 19d show the isolation for the two contacts, with the same set of parameters, with the guard ring included in Fig. 23d. On comparing the isolation in these two figures it can be seen that the improvement in isolation for the same separation of the contacts and the same value of L_{gnd} is in the range

of 7-10dB. The isolation is a weak function of the guard ring inductance L_{gr}.

Large gains in substrate isolation are achieved only by lowering the backplane inductance. This behavior can be expected in low-resistivity substrates because current flow in these substrates is mostly through the bulk, while the guard ring is an effective current sink for only the surface component of the current.

For an ideally grounded guard ring at 1GHz, the isolation seems to be improving when the contacts come closer together for an L_{gnd} of 3nH and 5nH. This behavior is observed because the guard ring acts as a good ground for contacts close to it, especially since at high frequencies the backplane impedance is large and surface currents tend to increase slightly.

In low-resistivity substrates, the most effective way of improving isolation is to provide a very good ground contact to the backplane. Surface isolation structures similar to guard rings are not very effective in these substrates.

Guard Rings in High-Resistivity Substrates

A similar simulation was performed for the case of high-resistivity substrates as was done in the above

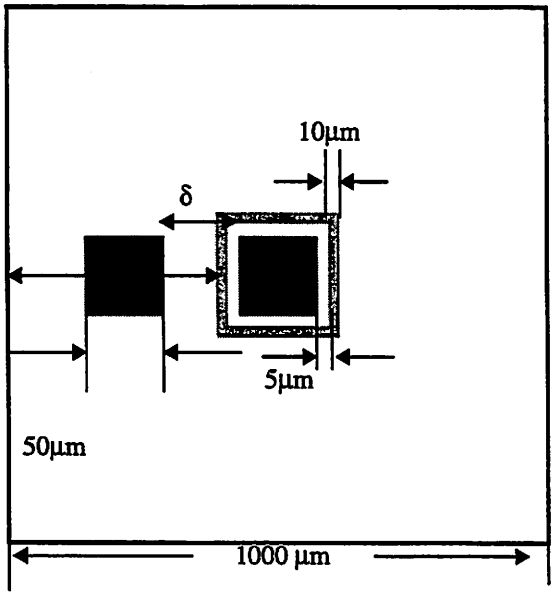
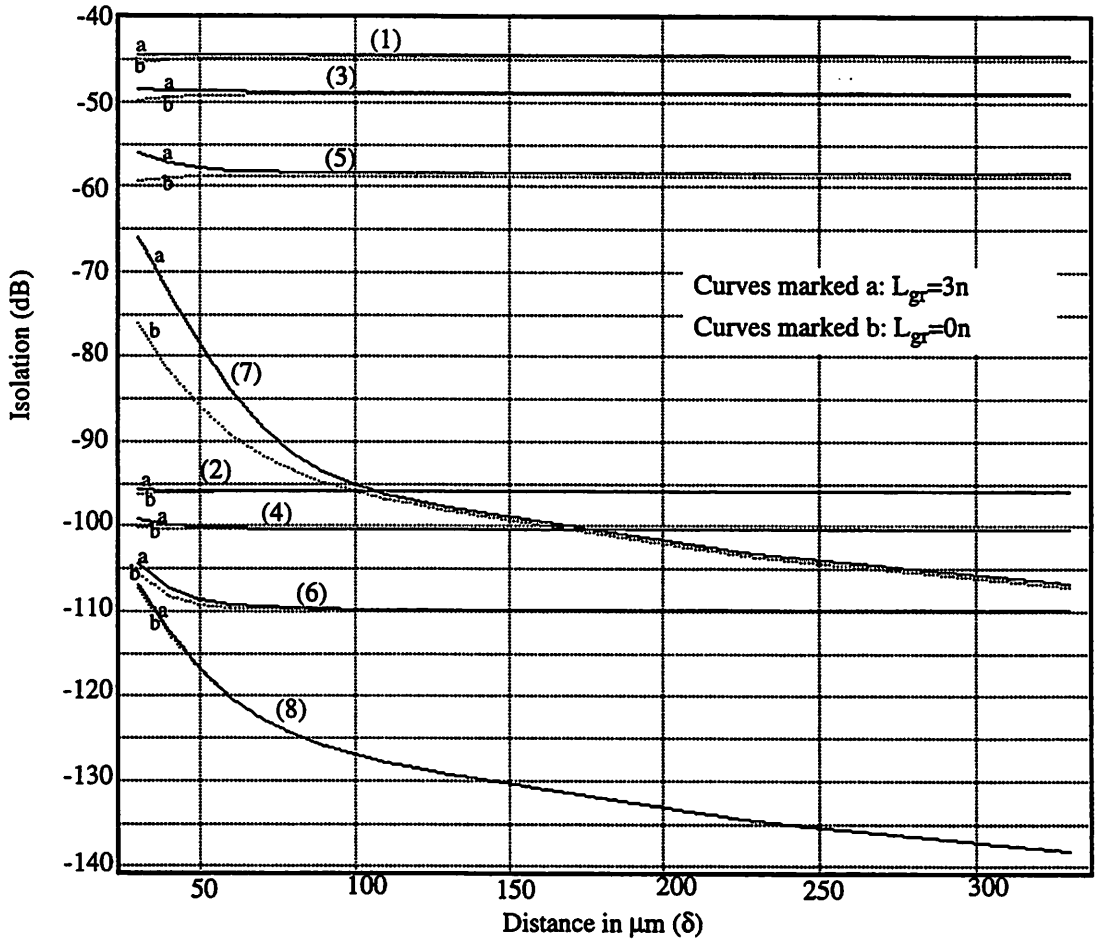


Figure 23c: Guard Ring Layout

subsection. The layout is shown in Fig. 23c. The results of the simulation are shown in Fig. 23e. Compared to the previous case, guard rings in high-resistivity substrates are very effective. Comparing Fig. 23e with Fig. 19g, we observe that large improvement in isolation is obtained with the ring. It can also be seen that lowering the value of L_{gr} improves the isolation by large numbers. As was discussed earlier, a large



Curve (1):	pl_1=1	pl_2=15	pl_3=10m	pl_4 = 0	f=1GHz, L_{gnd}=5nH
Curve (2):	pl_1=1	pl_2=15	pl_3=10m	pl_4 = 0	f=0.1GHz, L_{gnd}=5nH
Curve (3):	pl_1=1	pl_2=15	pl_3=10m	pl_4 = 0	f=1GHz, L_{gnd}=3nH
Curve (4):	pl_1=1	pl_2=15	pl_3=10m	pl_4 = 0	f=0.1GHz, L_{gnd}=3nH
Curve (5):	pl_1=1	pl_2=15	pl_3=10m	pl_4 = 0	f=1GHz, L_{gnd}=1nH
Curve (6):	pl_1=1	pl_2=15	pl_3=10m	pl_4 = 0	f=0.1GHz, L_{gnd}=1nH
Curve (7):	pl_1=1	pl_2=15	pl_3=10m	pl_4 = 0	f=1GHz, L_{gnd}=0
Curve (8):	pl_1=1	pl_2=15	pl_3=10m	pl_4 = 0	f=0.1GHz, L_{gnd}=0

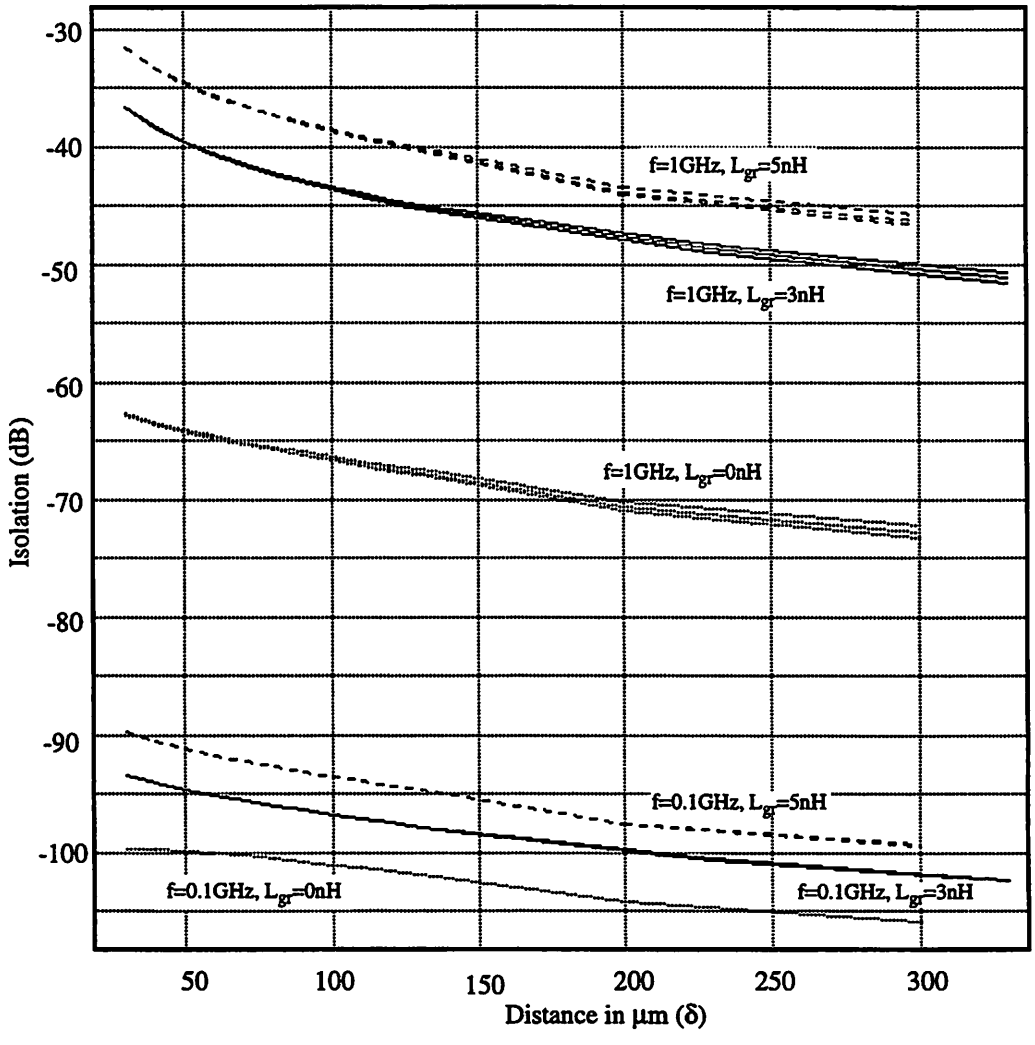
C_{sub}=0.8pF

Figure 23d: Isolation with a Guard Ring on a Low-Resistivity Substrate

fraction of the substrate current flows at the surface in these substrates. Hence the guard rings act as very good current sinks. The following study of guard rings will be restricted to this type of substrates only.

4.1.2: Optimal Sizing and Placement of Guard Rings in Single-Ended Circuits

Guard rings were shown to be effective means to reduce substrate noise in high-resistivity substrates in the previous section. In this section the optimal sizing of guard rings in these substrates, in order to



ρh_1=0.1 ρh_2=20 ρh_3 = 0

C_{sub}=0.8pF, L_{gnd} = 0, 1n, 3n, 5n

Figure 23e: Isolation with a Guard Ring on a High-Resistivity Substrate

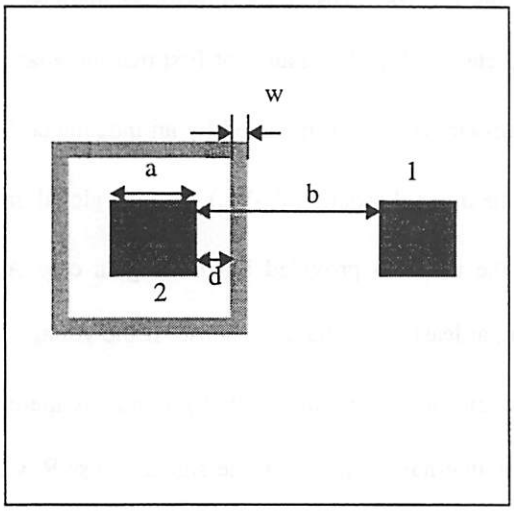


Figure 24a: Guard Ring Layout

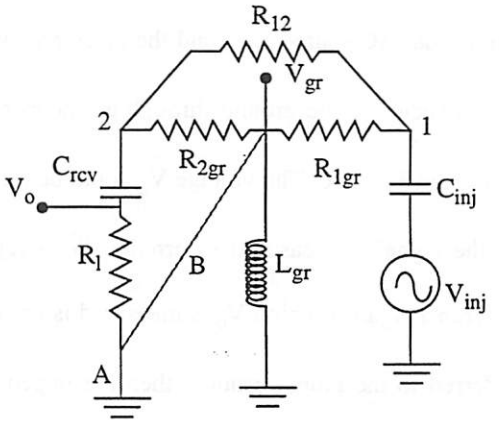


Figure 24b: Guard Ring Grounding Schemes

minimize the total noise appearing at the receiver node, is discussed. This section deals with single-ended circuits. Differential circuits are considered in the next section.

4.1.2.1: Guard Rings with Different Grounding Schemes

A single guard ring is assumed to be laid-out around the receiver contact as shown in Fig. 24a below. The injector and the receiver are assumed to be capacitively coupled to the substrate. The substrate

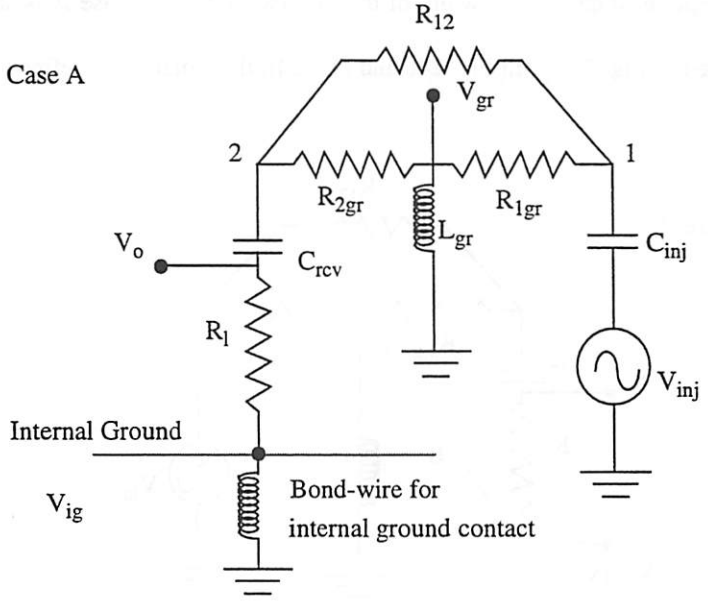


Figure 24c: Guard Ring with an External Ground Connection

contacts are shown by 1 and 2, for the injector and the receiver. The noise-injector is modeled by a sinusoidal AC source V_{inj}, and the noise received is depicted by V_o. We assume at first that the guard ring is connected to the ground through an independent bond-wire (Case A), modeled by an inductance L_{gr} as shown in Fig. 24c. The voltage V_o could be referred to the internal ground (V_o-V_{ig}) or to the global ground if the signal is measured externally. With regards to the isolation provided by the ring in case A, the reference against which V_o is measured is not important, at least in a behavioral sense. If the voltage V_o is referred to the global ground, then the impedance between the internal and global grounds, is merely an addition to R_l. On the other hand, if V_o is referred to the internal ground, then the signal across R_l will be directly proportional to the value of V_o referred to the global ground. Only the former case, when the signal is measured with reference to the global ground, is considered here. A detailed analysis of case A is reserved for the next section, where isolation with noise injection through the guard ring is considered.

When the ring is connected to the internal ground (Case B), it becomes necessary to identify the reference potential as in Fig. 24d below. These two cases are considered explicitly in the simulations.

The variation of isolation as a function of guard ring width and the bond-wire ground inductance L_{gr} is considered in Fig. 24e-g. The substrate parameters are mentioned in the figures. The dimensions a, b and d, shown in Fig. 24a are kept constant and the width of the ring (w) is varied. Case A is shown in Fig. 24e, while case B is considered in Fig. 24f-g. In Fig. 24e and Fig. 24f, the isolation is defined as the ratio of V_o

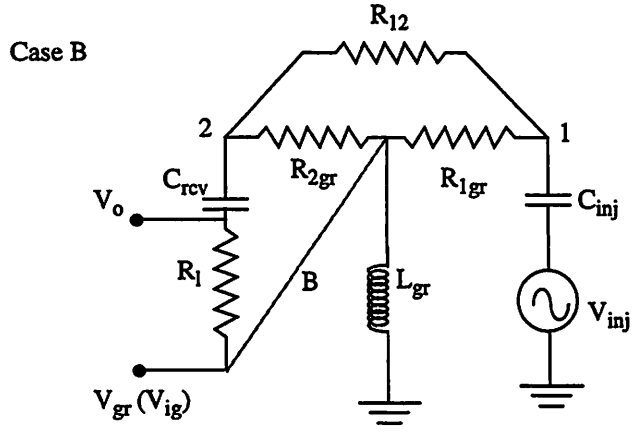


Figure 24d: Guard Ring with an Internal Ground Connection

and V_{inj} in dB. The isolation in Fig. 24g is defined to be $(V_o-V_{gr})/V_{inj}$ in dB.

We observe from Fig. 24e-f, that increasing the thickness of the guard ring from 2 to 34μm, provides a small improvement in the isolation but only for the case L_{gr}=0. For other values of L_{gr}, the isolation actually worsens. Thus in both these cases, it is advisable to use thin guard rings.

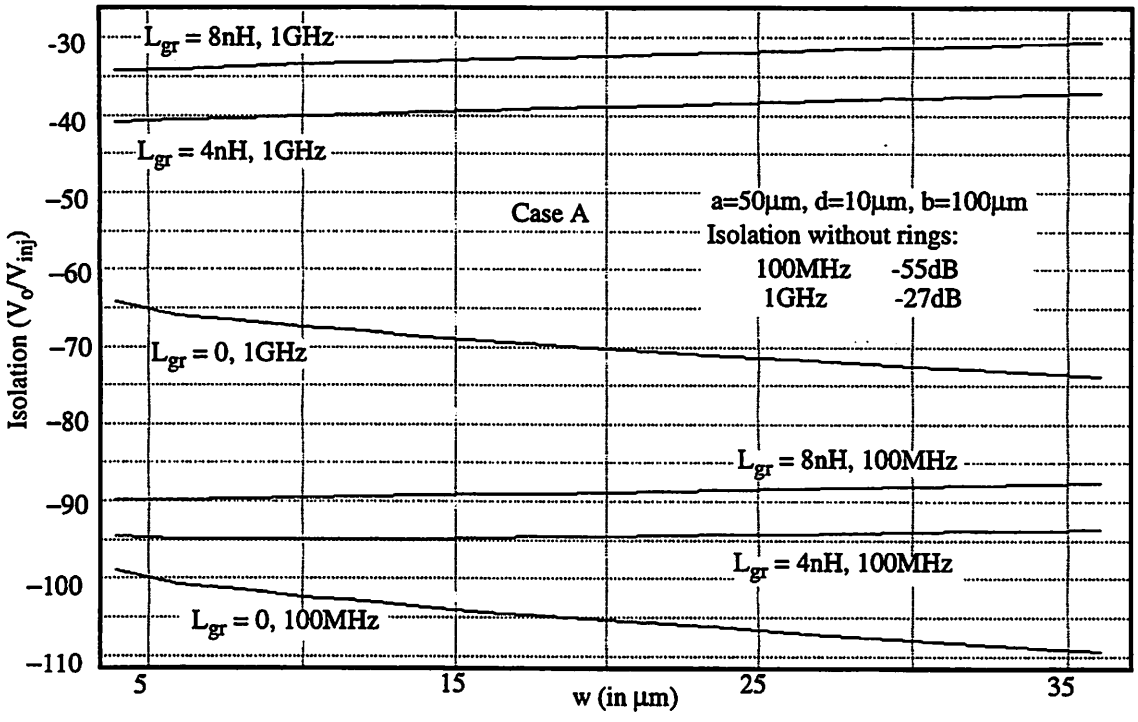


Figure 24e: Isolation with an Externally Grounded Guard Ring

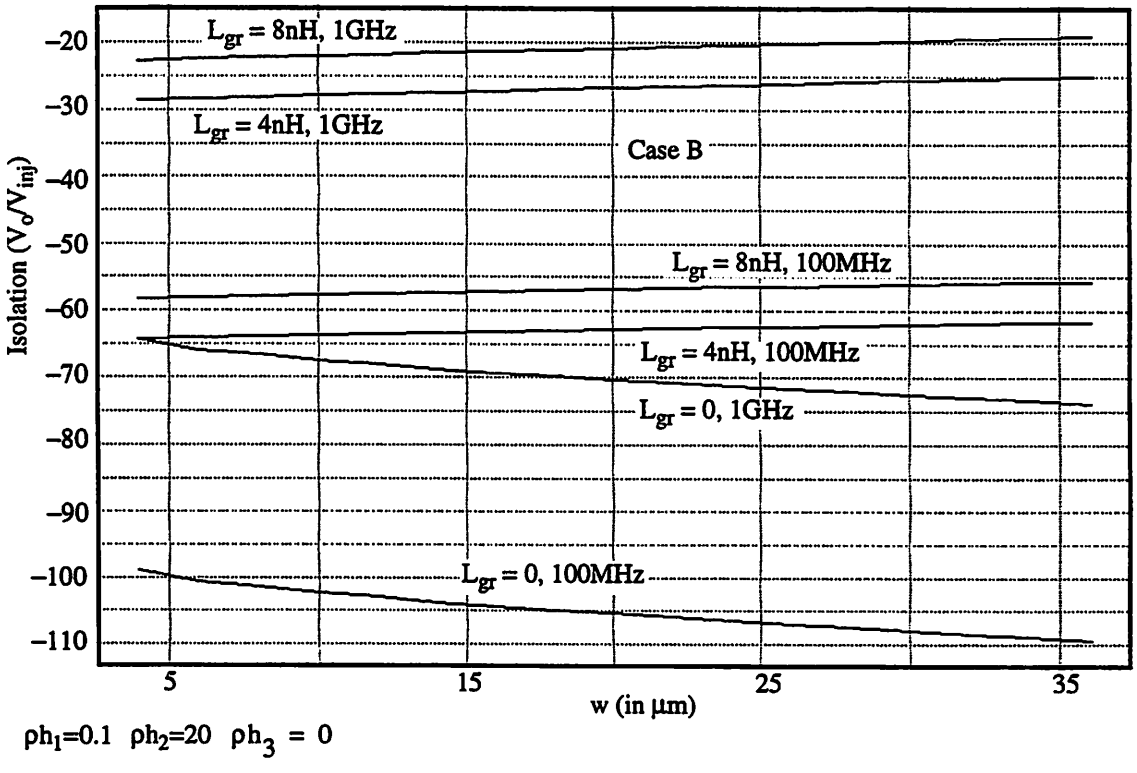


Figure 24f: Isolation with an Internally Grounded Guard Ring

Further examination of Fig. 24e and Fig. 24f reveals that the isolation in Fig. 24e is significantly better, which leads us to the conclusion that if the signal V_o is to be referenced to an external ground, the guard ring must be connected to ground through an independent bond-wire. It can further be observed that the isolation in Fig. 24f, for large values of the bond-wire inductance, is worse than that without the guard ring at all. Thus the guard ring must be connected to an internal ground, only if the signal is measured with respect to the internal ground. It can be seen from Fig. 24g that the isolation provided by an internally grounded guard ring is excellent, if the signal is measured with reference to the internal ground. Further, the isolation is practically independent of the value of the bond-wire inductance.

C_{rcv} and C_{inj} have been assumed to be 0.8pF for the simulations.

4.1.2.2: Optimization of Guard Ring Widths

Guard rings are often connected to ground by the use of a separate ground-pin (case A in the previous

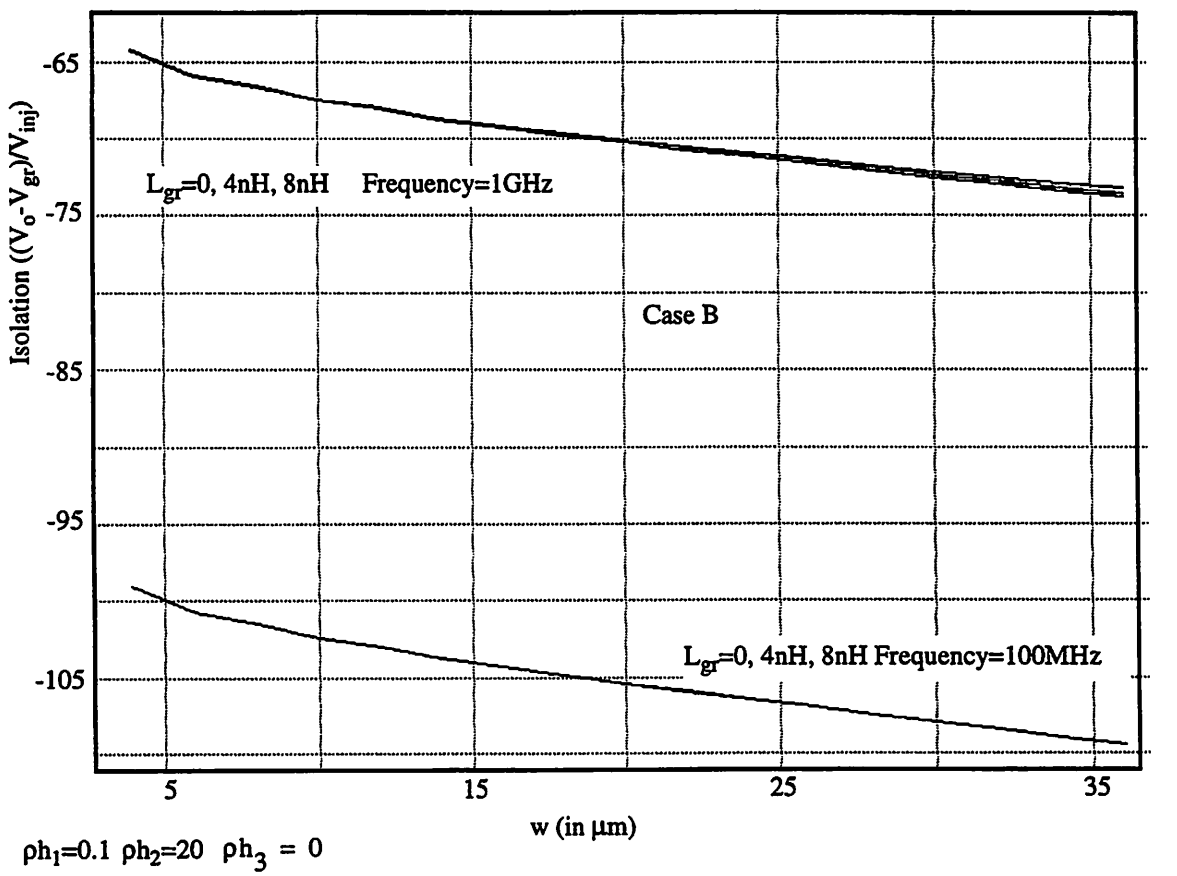


Figure 24g: Isolation with an Internally Grounded Guard Ring

section). The associated bond-wire can pickup noise from adjoining bond-wires, and inject noise into the circuit. The optimization of the guard ring size with a noisy ground-connection is treated in this section. Only case A is considered here. In case B, the influence of noise on the guard ring bond-wire can be minimized by referring the receiver voltage to the internal ground.

To obtain an understanding of optimization problem, the substrate model with $\rho h_3 = \infty$ will be considered. This is done because the substrate model is simpler for this case. In Fig. 25a below, the lengths a and d are kept constant. The width of the ring is varied and the optimum size of the ring is determined. Contact-1 is the injector and contact-2 is the receiver. The injector voltage is shown as V_{inj}. The bond-wire used to ground the guard ring can also pickup noise from adjacent pins due to bond-wire mutual inductances. This noise is modeled by the generator V_{inj1}. The problem at hand is to reduce the influence of V_{inj} on V_2, while simultaneously keeping the effect of V_{inj1} low[7].

We assume, for ease of analysis, that the impedances Z_{inj} and Z_{rcv} are small. These impedances model the reactances of the device-to-substrate capacitors. Hence node-1 in Fig. 25b is voltage driven and node-2 sees the resistance R_l. This approximation will be valid at very high frequencies.The isolation between nodes-1 and 2 is then determined by the reduced model shown in Fig. 25c.

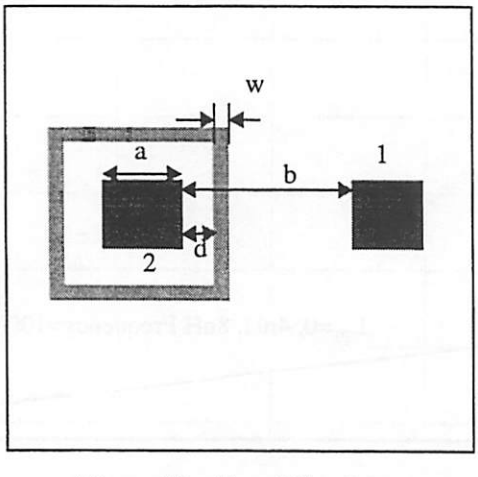


Figure 25a: Guard Ring Layout

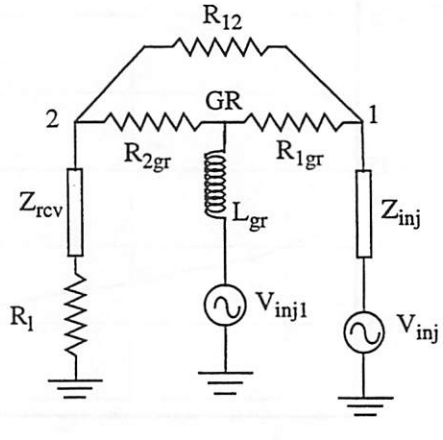


Figure 25b: Substrate Model

7. Not placing the ring avoids the injection of V_{inj1} altogether. In some applications, where V_{inj1} is large, it is possible that adding the ring actually increases the noise.

Under the approximations stated earlier and provided that R_l is small, the total output voltage at 2 can be shown to be

$$V_2 \approx R_l \times \left\{ \left(\frac{1}{R_{12}} + \frac{j\omega L}{R_{1gr}R_{2gr} + j\omega L(R_{1gr} + R_{2gr})} \right) V_{inj} + \left(\frac{R_{1gr}}{R_{1gr}R_{2gr} + j\omega L(R_{1gr} + R_{2gr})} \right) V_{inj1} \right\} \qquad (85)$$

Let us first consider the case when V_{inj1} is zero. The effect of increasing width (w) is that R_{12} increases and R_{1gr} and R_{2gr} decrease. At low frequencies the isolation between the contacts is determined by R_{12} and therefore for low-frequency operation a wide ring is beneficial. At high frequencies, however, the second term in the bracket starts becoming significant. In fact at very high frequencies, the isolation is governed by the sum (R_{1gr}+R_{2gr}). If this sum is small at these frequencies, the isolation can actually be worse than that achievable without the ring. Further, as R_{1gr} and R_{2gr} are reduced, the frequency at which the ring begins to lose its advantages is also lowered ($\omega \sim (R_{1gr}R_{2gr})/L(R_{1gr}+R_{2gr})$). It must be mentioned however, that this pole frequency is usually very large for typical values of L_{gr}. If V_{inj1} is nonzero, we observe a potentially more serious problem with a very wide guard ring. At low frequencies the noise term V_{inj1} appears at the output scaled by the ratio R_l/R_{2gr}. This ratio can be significant if R_{2gr} is too small[8].

An area-efficient way to increase R_{12} while keeping 'w' small, is to make 'd' small in Fig. 25a. However making 'd' small also decreases R_{2gr}, which leads to a higher dependence of V_2 on V_{inj1}. If V_{inj1} is significant, then 'w' and 'd' must both be made large. Another option is to place the ring around the

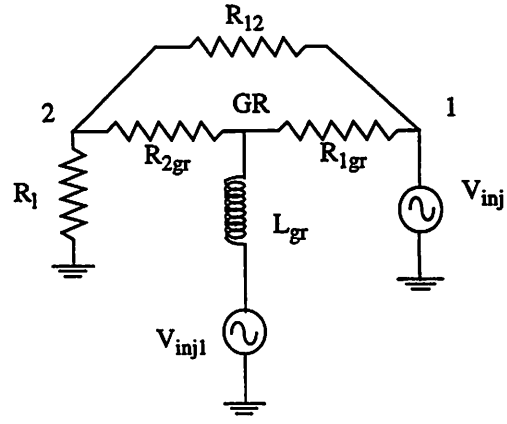


Figure 25c: Modeling Bond-Wire Noise in a Guard Ring

8. It is assumed that R_l is always smaller than all the other resistors.

injector contact. It can be seen from (85) that the dependence of V_2 on V_{inj} is symmetric in R_{1gr} and R_{2gr}. Thus increasing R_{2gr} and reducing R_{1gr} simultaneously has no effect on V_2 if V_{inj} is the only noise source. However the effect of V_{inj1} on V_2 is significantly reduced by using this arrangement.

The alternate scheme of placing the ring around the noise-injectors may not always be preferable to the scheme shown in Fig. 25a. The principal cause of the appearance of the noise V_{inj1} is the mutual inductance which is present between the pins and bond-wires in a package. A usual design practice in mixed analog-digital circuits is to place the analog and digital portions of the circuit in different parts of the substrate. The pads are also laid out such that the analog and digital bond-wires are placed in different sections of the package. If layout constraints force the bond-wire used to ground the ring to be located in proximity to the digital bond-wires, then it is possible that the magnitude of V_{inj1} may increase significantly compared to the other case. The ideal scheme for isolating the sensitive circuits from the noisy circuits is to place the ring around the noisy circuits and lay out the bond-pads such that the guard ring bond-wire is placed along with the analog-section bond-wires.

The above discussion is not definitive but is meant to shed light on some of the trade-offs involved in the design of optimal guard rings. In the rest of the section results from simulations are presented.

The noise appearing at contact-2 of Fig. 25b for a one-Volt swing on contact-1 is shown in Fig. 25d-g. The dimensions of the contacts and the ring and the substrate data are presented in the figures. The width of the ring is swept from 4μm to 32μm in increments of 2μm and the output signal is measured. C_{sub} is assumed to be 0.8pF in all the simulations.

Fig. 25d and Fig. 25e depict the substrate noise coupling in substrates without backplanes. Fig. 25f and Fig. 25g depict the case with ideally grounded backplanes. In each figure the isolation at two frequencies, 100MHz and 1GHz is shown. The guard ring bond-wire inductance is stepped from 0 to 8nH in steps of 2nH. The isolation without the guard rings at the two frequencies is mentioned in each graph.

A common conclusion which can be made from each of the graphs is that at high frequencies (~1GHz), thin guard rings provide better isolation than wide rings. The evidence of the existence of an optimum guard ring width is found in the isolation curves at 100MHz in all the curves for different values of the

inductance to ground. In Fig. 25d, for example, the optimal guard ring width for a 2nH ground path inductance is 15μm. The minimum is shallow. Thus the best rule for sizing guard rings is to use minimum width rings and ensure a very good ground connection to the ring. The presence of the ideally grounded backplane improves the isolation by 5-10dB compared to the case without the backplane. This suggests that the primary path to ground is provided by the surface ground rings and not by the backplanes in high-resistivity substrates.

It was mentioned earlier in this section that it is possible to place a guard ring around either the receiver or the injector contact. The dependence of the received noise on the noise terms V_{inj} and V_{inj1} is shown in Fig. 25h and Fig. 25i respectively, with the ring placed around the injector and the receiver in each figure. The layout parameters and the substrate parameters used are shown in the figures. The bond-wire inductance is assumed to be 4nH. The upper curve in Fig. 25h is the isolation from V_{inj} without the guard ring. It can be seen that the noise isolation provided by the ring, from the substrate injector is nearly the same, regardless of whether the ring is placed around the injector or the receiver. The isolation from V_{inj1} (Fig. 25i) is better by approximately 20dB in the case when the ring is placed around the injector rather than around the receiver.

As a specific example, consider the case when the bond-wire adjacent to the guard ring bond-wire carries a current of 1mA at 1GHz and has a mutual inductance of 1nH with the guard ring bond-wire. V_{inj1} will thus be 6.3mV. If the ring is laid-out around the receiver, then the voltage appearing on the receiver due to V_{inj1} will be 1.41mV. If the ring is laid around the injector, the noise at the receiver due to V_{inj1} will be 0.14mV. These voltages will scale linearly with the magnitude of the current in the adjacent bond-wire. For a 1V swing at 1GHz at the injector, the noise appearing at the receiver is 5.623mV. If the guard ring is placed around the injector, then the noise induced by V_{inj1} will exceed that caused by V_{inj} for current values exceeding 35mA. If the ring is placed around the receiver, the noise caused by V_{inj1} will dominate for current amplitudes greater than 3.5mA. [9]

9. The noise without the guard ring can be shown to be 22mV. Therefore it is advantageous to use a guard ring in this example.

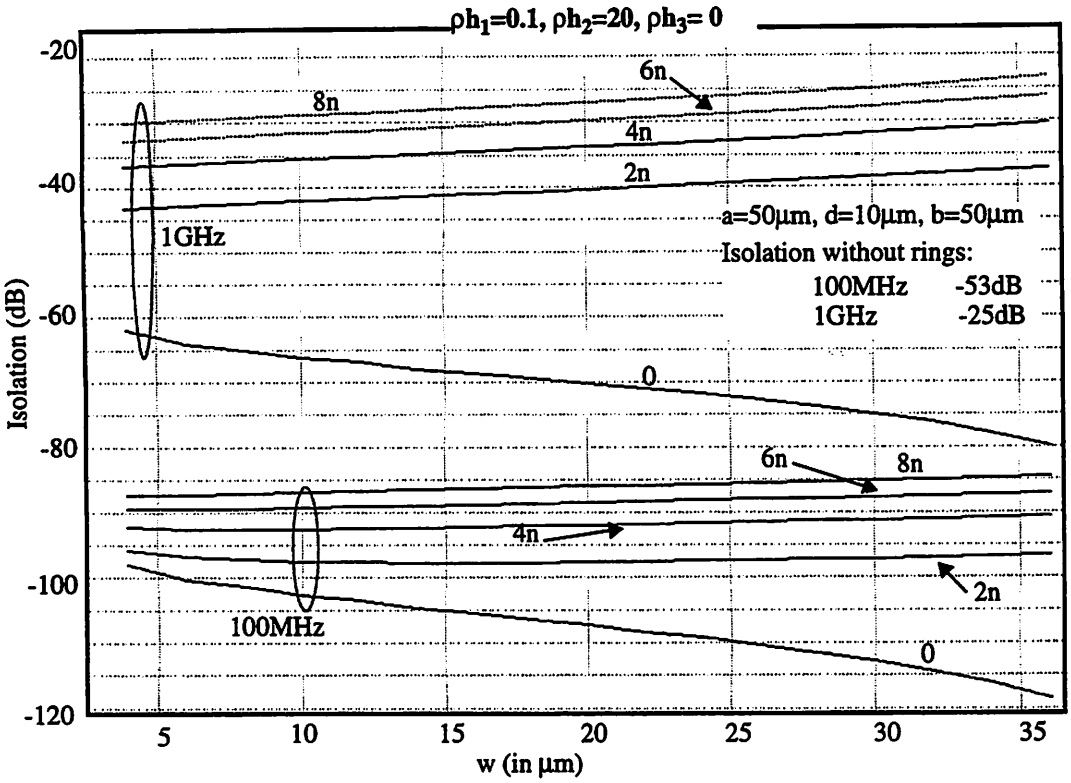


Figure 25d: Guard Ring Isolation as a Function of Width

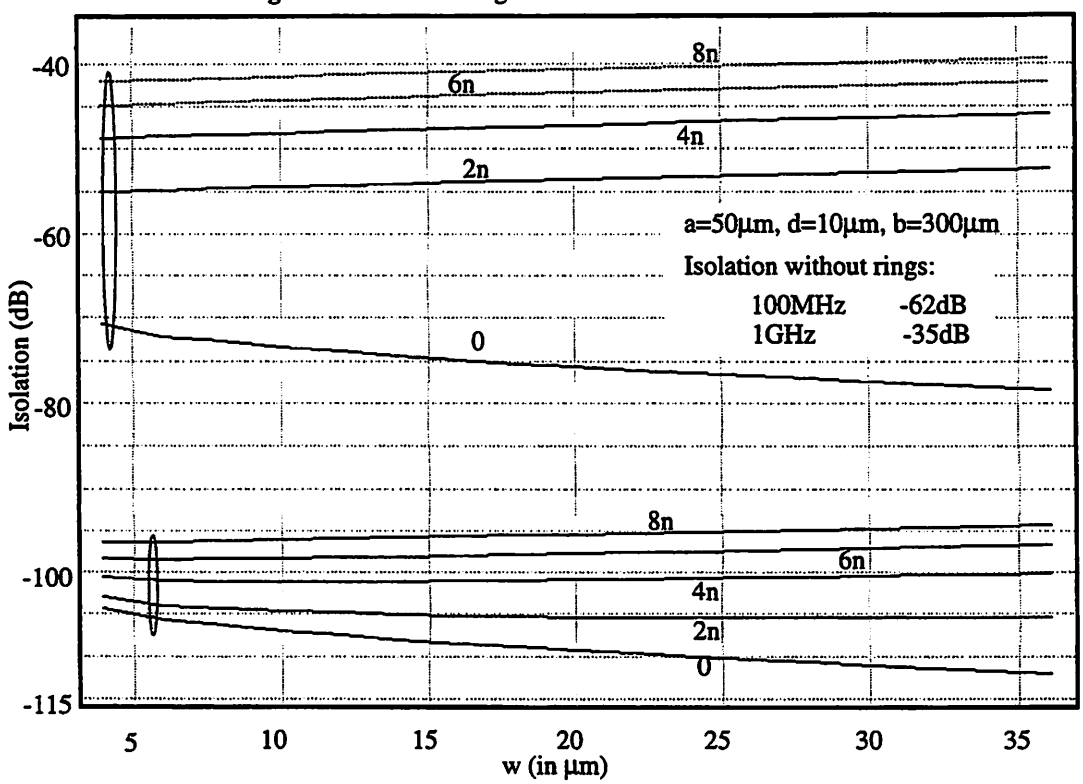


Figure 25e: Guard Ring Isolation as a Function of Width

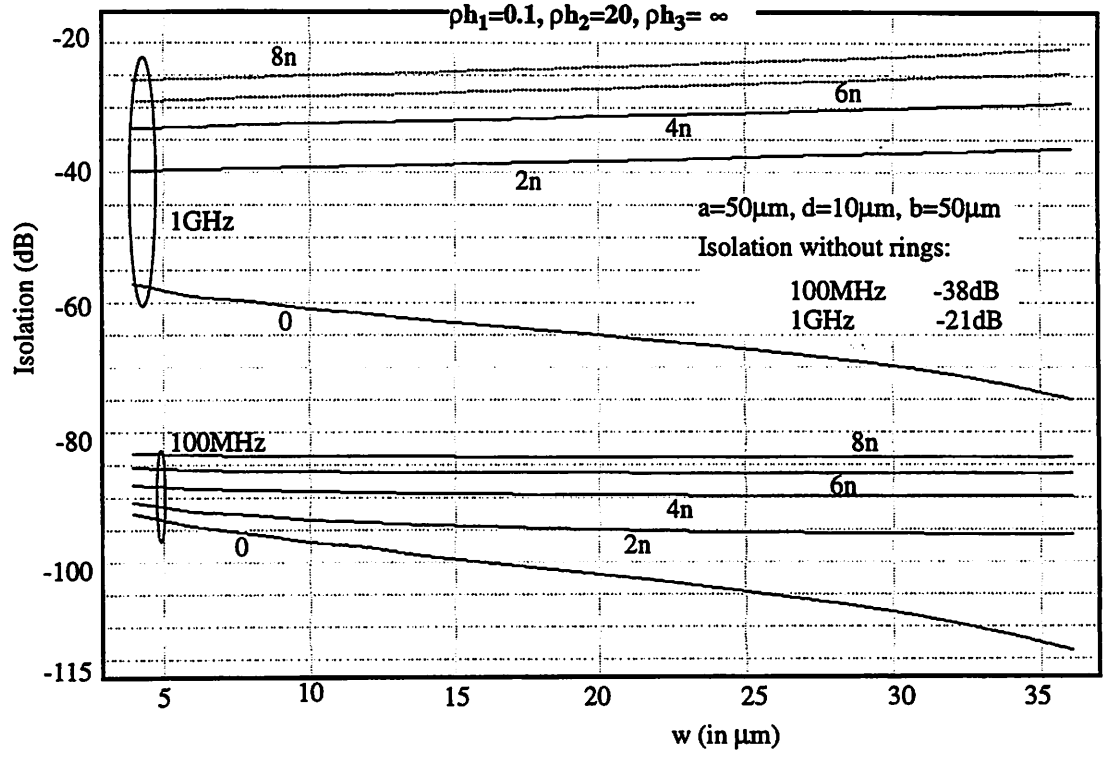


Figure 25f: Guard Ring Isolation as a Function of Width

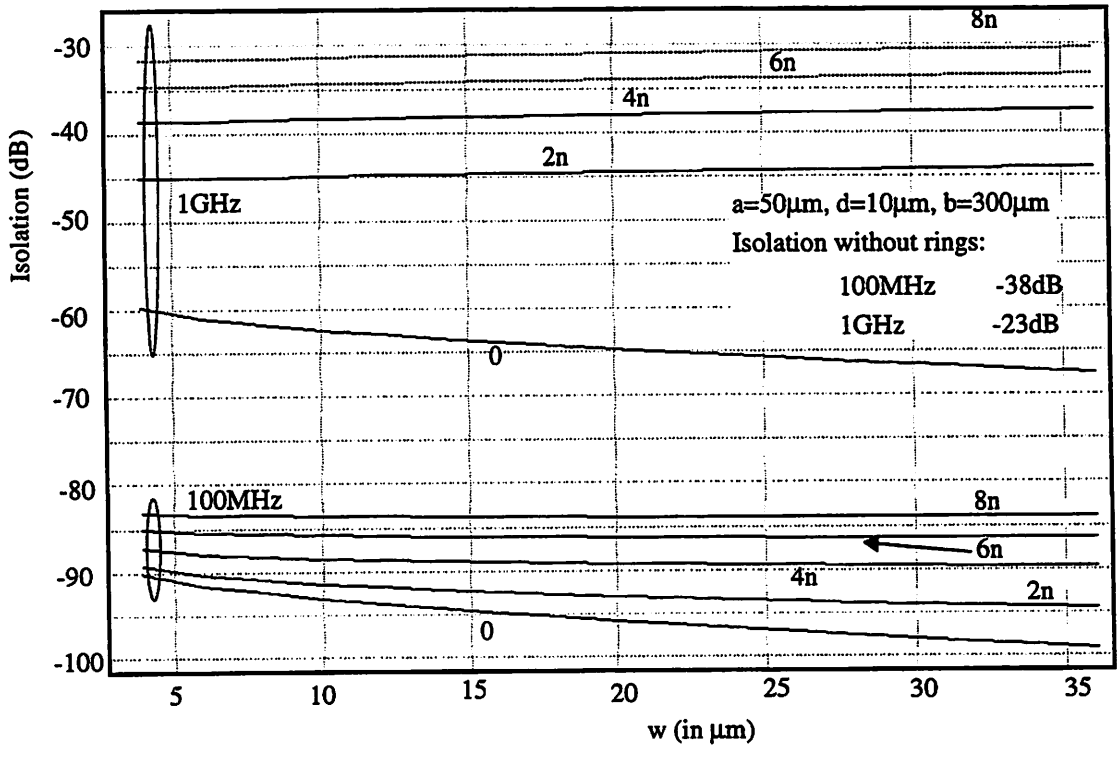


Figure 25g: Guard Ring Isolation as a Function of Width

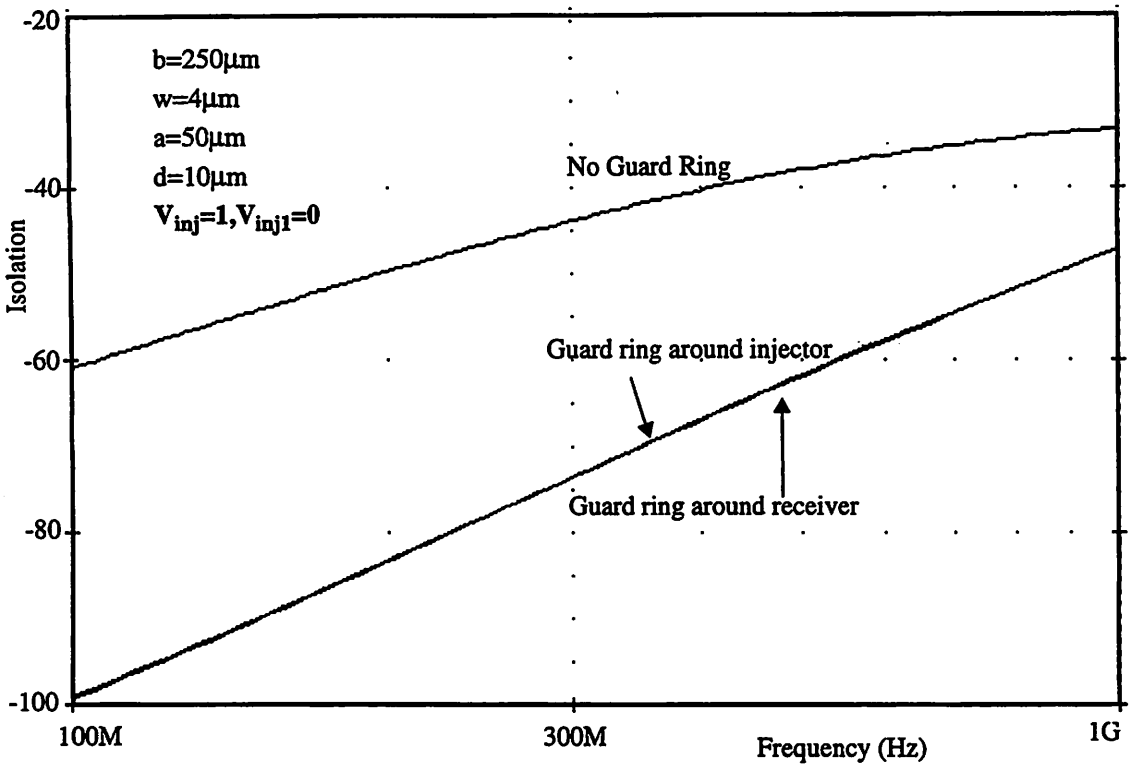


Figure 25h: Isolation with Substrate Injection Only

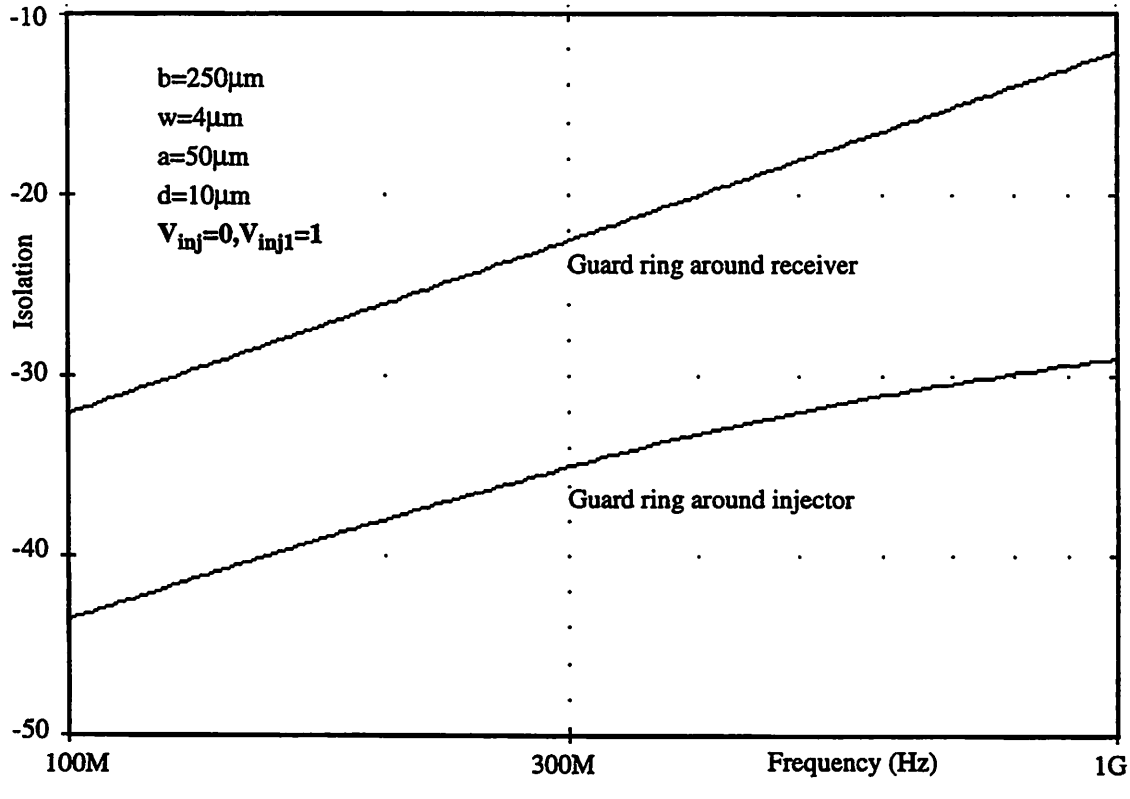


Figure 25i: Isolation with Guard Ring Bond-Wire Noise only

4.1.3: Guard Rings in Differential Circuits

It was mentioned earlier that the improvement in isolation in differential circuits in high-resistivity substrates was not as significant as in low-resistivity substrates. The reason for the excellent differential isolation in low-resistivity substrates was found to be that the heavily-doped bulk, which lies in close proximity to the surface, acts as a good ground-plane. As a result, noise signals coupled to the substrate appear as common mode signals on the differential outputs. Guard rings in high-resistivity substrates also act as ground-planes in close proximity to the devices. Hence it can be expected that differential isolation in high-resistivity substrates will be improved in a similar fashion by the use of guard rings. The effect of guard rings in differential circuits is studied in this section. Differential receivers with single-ended injectors are considered here.

We consider the differential substrate-contacts, without and with a surrounding guard ring shown below in Fig. 26a and Fig. 26b respectively. For the purpose of simulation, the guard ring shown in Fig. 26b is connected to ground through the bond-wire inductance L_{gr}, which is stepped over three values, 0, 4nH and 8nH. The width of the ring is stepped from 10μm to 48μm in steps of 2μm. The results of the simulation, that is, the differential and common-mode isolation values are shown in Fig. 26c.

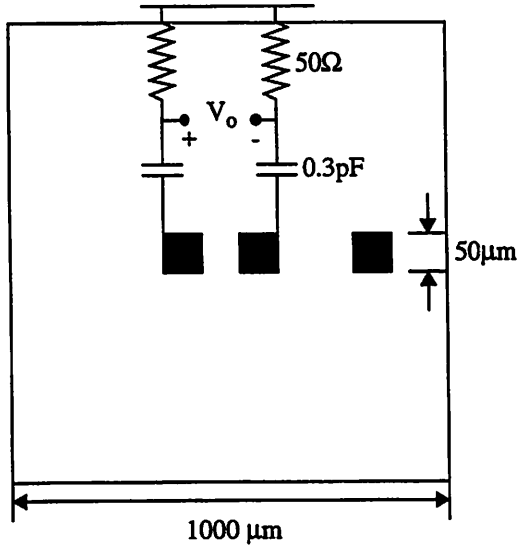


Figure 26a: Layout without Guard Ring

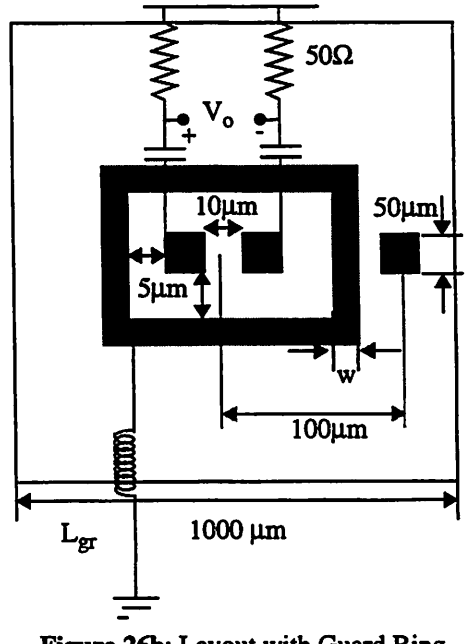


Figure 26b: Layout with Guard Ring

It can be observed from the figure that the differential mode-isolation improves slightly by the use of thick guard rings. Unlike the isolation in the single-ended case in the previous section, the isolation value is independent of the value of the bond-wire inductance. Thus the isolation improves by a significant amount, and is independent of the impedance to ground. Common-mode isolation, however, does worsen with increasing ring thickness, for nonzero values of the bond-wire inductance. Thus in differential circuits, as in the single-ended case, it is advisable to use thin guard rings.

The relative independence of the differential-mode isolation on the value of L_{gr} is similar to that seen earlier in low-resistivity substrates in Section 4.0.3. In both cases, the presence of an equipotential region

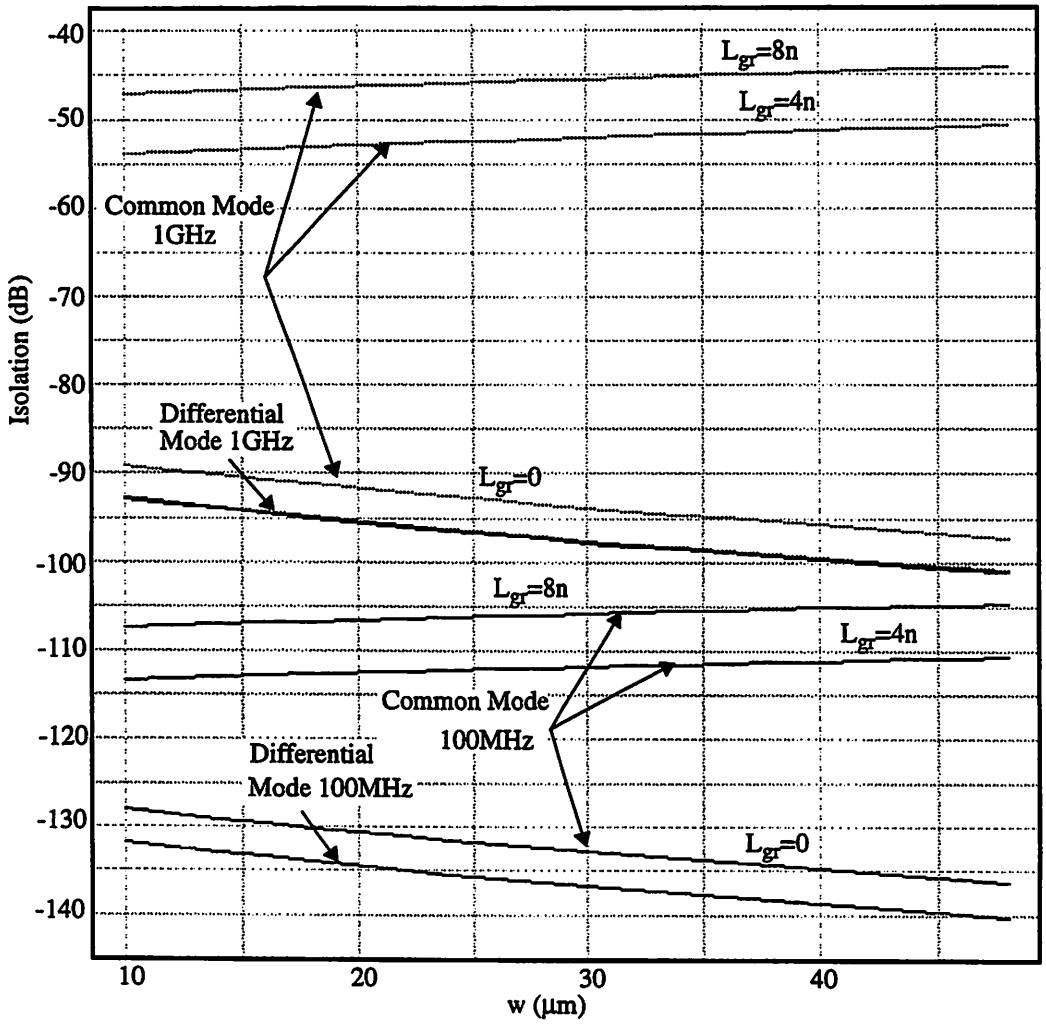


Figure 26c: Isolation in a Differential Circuit with a Guard Ring

in proximity to the circuits makes the noise a common-mode signal and is rejected. The presence of circuit mismatches will degrade isolation, as discussed in Section 4.0.3.

Circuits with differential injectors and receivers will show similar improvements in isolation in the presence of a guard ring around the receiver or the injector. These circuits have not been discussed separately in this section since the mechanism which leads to the improvement in isolation is the same as that discussed above.

4.1.4: The Use of Dual Guard Rings

The effect of one guard ring placed around the injector or the receiver has been discussed in the previous sections. It is also possible to place guard rings around both the injector and receiver contacts. The isolation between two single-ended contacts, with a guard ring placed around each contact is examined in this section. The layout is shown in Fig. 27. The results of the simulation are shown in Table 3. The isolation at 100MHz to 1GHz, without guard rings, with a guard ring around the receiver contact only and with two guard rings, is shown in the table.

The guard rings shown in Fig. 27 can be connected to ground separately or together. In the latter case,

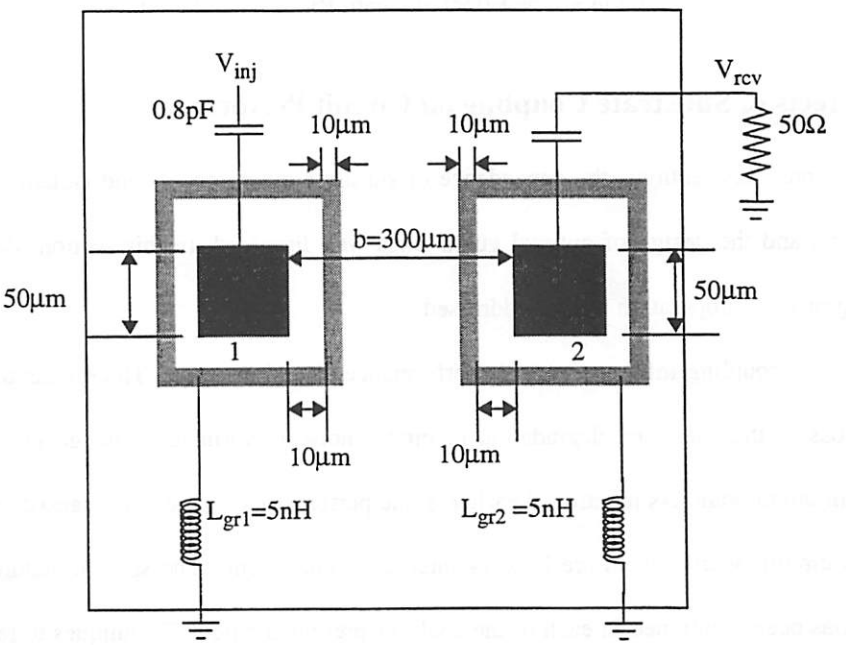


Figure 27 : Contact and Guard Ring layout

however, the two guard rings appear to act as one large ring, and all the conclusions of Section 4.1.2 are applicable. Hence the guard rings are assumed to be grounded separately. The relevant layout-data is shown in Fig. 27. The guard rings are assumed to be connected to ground through independent bond-wires. The bond-wires are modeled by 5nH inductors. The injector and receiver contacts are squares of 50μm sides. The substrate capacitance is assumed to be 0.8pF.

The improvement in isolation by placing two rings is seen to be significant, especially at low frequencies. Using two guard rings requires two package pins which may not always be possible. If, however, two pins are available, then using two guard rings is advisable.

Case	Frequency	Isolation(dB)
No Guard Rings	100MHz	-62
	1GHz	-34
Guard Ring around receiver only	100MHz	-99
	1GHz	-45
Two Guard Rings	100MHz	-130
	1GHz	-57

Table 3: The Effect of Guard Rings on Isolation

4.2: Effects of Substrate Coupling on Circuit Performance

In the previous sections, the dependence of substrate macromodels and isolation on various substrate parameters and the design of optimal guard rings was discussed. In this section, the effects of substrate coupling on circuit operation will be addressed.

Substrate coupling influences circuit performance in various ways. The effects discussed here include power loss in the substrate, degradation of circuit noise performance, change in circuit bandwidth and change in circuit gain. As mentioned earlier, some performance changes are caused externally by coupling between circuits while others are local or internal to one circuit. The specific nature of the performance change has been mentioned in each of the analyses presented below. Techniques to reduce the degradation caused by substrate coupling are also mentioned for each effect considered. Finally some catastrophic

effects of substrate coupling which can lead to a complete failure of the circuit are also mentioned.

4.2.1: Power Loss in the Substrate

Silicon substrates are lossy in nature and hence are modeled as distributed resistors as discussed in Chapter 3. The loss in the substrate lowers the efficiency of circuit components and must be minimized. Substrate power-loss can be caused by current-injection into the substrate by active and passive devices and is an effect internal to the device. The current-injection mechanisms of various devices were discussed in Chapter 2. Capacitive current-injection takes place in passive devices such as resistors, capacitors, inductors and interconnects and also in active devices through junction capacitors formed by the substrate.

The performance of inductors is significantly influenced by the presence of the substrate-parasitics. This is due to a large oxide parasitic-capacitance to the substrate, which strongly couples the inductor to the substrate. The functional dependence of substrate power-loss on parameters such as the parasitic-capacitance and the substrate parameters is similar in all devices coupled capacitively to the substrate. Hence only inductors are discussed in this section.

Inductors are implemented on-chip in planar spiral forms (Fig. 28a) and are increasingly being used in radio-frequency circuits for personal-communication applications. A first-order model of the inductor is shown in Fig. 28b where the inductance is depicted by L. The series loss in the inductor is modeled by R_{sr}, C_{s1} and C_{s2} represent the effective oxide-capacitance to the substrate, while R_{s1} and R_{s2} model the

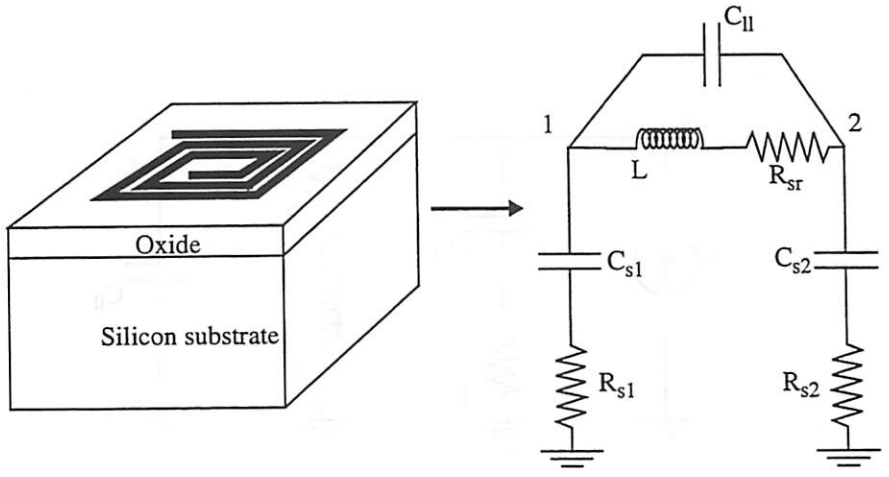


Figure 28a: Spiral Inductor

Figure 28b: Inductor Model

substrate parasitic. The inter-turn capacitance of the inductor is modeled by C_{ll}. This model of the inductor is valid if the inductor is used below its self-resonance frequency. More detailed inductor models are presented in [22].

The three major parameters of interest in the inductor are the value of the inductance L, the self-resonance frequency of the inductor and the quality factor or 'Q' of the inductor. The quality factor is defined as the per-unit-cycle energy stored to the power loss in the inductor. Typical Q factors of inductors fabricated in silicon processes are in the range of 3 to 10 [22]. The power-loss in the inductor is determined by the series and substrate resistors shown in Fig. 28b. The series loss is determined by the resistivity of the material used to fabricate the inductor and the length of the inductor winding. The substrate-loss term depends on the resistivity of the substrate. The substrate model-elements shown in Fig. 28b can be computed using the simulator described in Chapter 3. The dependence of substrate loss on substrate resistance is considered here.

For ease of analysis, we consider the case when one end of the inductor is connected to an AC ground (Fig. 28c). The inductor is assumed to be driven by a current source i_{in}. Simple circuit analysis shows that the loss in the substrate depends on R_{s1} in Fig. 28c by the following behavioral equation

$$\text{Loss} = \frac{f_1(\omega) R_{s1}}{f_2(\omega) R_{s1}^2 + f_3(\omega) R_{s1} + f_4(\omega)} \tag{86}$$

where f_1, f_2, f_3 and f_4 are functions of frequency 'ω' and depend on L, C_{ll}, C_{s1} and R_{sr}.

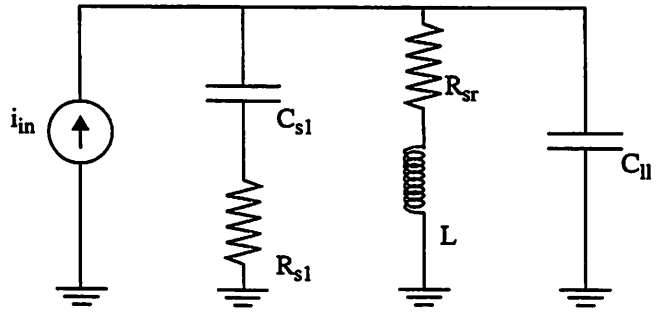


Figure 28c: Current-Driven Spiral Inductor

The above dependence suggests that the loss in the substrate is minimized either by a very large R_{s1} or by a small R_{s1}[10]. In fact it can be shown that the substrate loss is maximized for the following value of R_{s1} at any frequency ω.

$$R_{s1max} = \frac{1}{\omega C_{s1}} \times \sqrt{\frac{\left(1-\omega^2 L(C_{s1}+C_{ll})\right)^2 + \omega^2 R^2_{sr}(C_{s1}+C_{ll})^2}{\left(1-\omega^2 LC_{ll}\right)^2 + \omega^2 R^2_{sr} C^2_{ll}}} \quad (87)$$

When the inductor is used at a frequency much smaller that the self-resonance frequency of the structure, R_{s1max} is approximately equal to $1/\omega C_{s1}$. C_{s1} can be estimated from the size of the inductor and the thickness of the oxide. To minimize the substrate loss in the inductor, the substrate underneath the inductor can be doped to make R_{s1} either very large or very small compared to R_{s1max}.

The expression for the Q of the inductor is not derived here. The design-rule discussed here will suffice at frequencies below the self-resonance of the inductor. Experimental work is required to verify the above result.

Similar analyses can be performed for other devices too. In all cases, the substrate loss is maximized when the substrate resistance is nearly equal to the magnitude of the reactance of the device-to-substrate capacitance.

The substrate injection in some devices can contain a DC component, for example the avalanche induced substrate injection in NMOSFETs, the leakage current in diodes formed with the substrate or through the parasitic pnp transistor in npn BJTs (Chapter 2). These injection mechanisms lead to a DC power loss in the substrate. The leakage current in substrate junctions can be minimized by suitably adjusting the process to reduce the maximum fields in the depletion regions. Avalanche injection in NMOSFETs and the current levels in the parasitic pnp devices in npn BJTs can be reduced by biasing the devices such that the substrate injection is minimized [3]. This may not always be possible since the device bias-levels are decided by more important circuit issues. Therefore the most practical way to reduce the DC I^2R loss is to provide a low-resistance path to ground. DC current injection into the substrate can lead

10. It is assumed that R_{s1} is large enough so that the inductance L is not significantly reduced by an induced image in the substrate.

to more harmful effects and these are discussed later.

4.2.2: Degradation of Circuit Performance due to Substrate-Coupled Noise

In this section, the changes in the performance of circuits due to substrate noise are considered. Substrate noise injection may be internal to a circuit or a device or may be external, coupled from one circuit to another. Low noise-power-levels are important in several circuits such as low-noise-amplifiers (LNAs), oscillators, mixers and high-resolution AD converters. Noise injection is important in passive components which have a large substrate parasitic capacitance such as inductors or long interconnects, as well. An amplifier consisting of a single bipolar transistor is considered below to illustrate the internal and external mechanisms for noise injection.

Noise-Figure Degradation in a Low-Noise Amplifier

The noise mechanisms in the substrate which can lead to a possible degradation of the noise figure of the amplifier are shown in Fig. 29a. The substrate is lossy and is modeled by resistors R_{sbp} and R_{s1}, which represent the bond-pad to substrate-ground and the collector-to-substrate-ground model-elements respectively. The thermal noise in a resistor R can be modeled as an Additive White Gaussian noise (AWGN) of power-spectral-density 4kTR [2]. The noise sources modeling the thermal noise of R_{sbp} and R_{s1} are shown in Fig. 29a as $\overline{V}^2_{n2}$ and $\overline{V}^2_{n1}$ respectively. C_{sbp} is the bond-pad capacitance and C_{s1} is the collector-to-substrate capacitance of the BJT. $\overline{V}^2_{nc}$ models the noise due to other switching devices which may be present on the same substrate and is considered later. Experimental work studying the effect of the substrate loss on the noise-figure of bipolar transistor amplifiers was recently presented in [23].

The effects defined as being internal to the circuit are considered first, that is $\overline{V}^2_{nc}$ is assumed to be zero. The noise source $\overline{V}^2_{n2}$ is amplified by the transistor and adds to the noise power-spectral-density (PSD) at the output and will therefore be more harmful to circuit performance than $\overline{V}^2_{n1}$. The expression for the output PSD without the substrate noise sources is derived in [2]. The power-spectral density of the noise contribution at the output due to $\overline{V}^2_{n2}$ is given by the following expression

$$\frac{d\overline{v}^2_o}{df} = \frac{4kTg_m^2R_{c1}^2R_s^2r_\pi^2\omega^2C_{sbp}^2R_{sbp}}{\left(\left(a-\omega^2(b+cR_{sbp})\right)^2+\omega^2(d+eR_{sbp})^2\right)} \tag{88}$$

where $a = R_s + r_\pi + r_b$, $b = R_s r_\pi r_b C_\pi C_{sbp}$, $c = (R_s + r_b) r_\pi (C_\pi C_{sbp})$, $d = (r_\pi + r_b) R_s C_{sbp} + (R_s + r_b) r_\pi C_\pi$ and $e = (R_s + r_\pi + r_b) C_{sbp}$. In the above equations, r_b is the base resistance, R_s is the source resistance, r_π is the small-signal base input resistance and C_π is the small-signal base-to-emitter capacitance. The total impedance of C_{sbp} and R_{sbp} is assumed to be large, so that the noise contribution of the other sources is not changed by the loading of C_{sbp} and R_{sbp}. It can be shown that the contribution of R_{sbp} to the output noise is maximized when

$$R_{sbp} = \frac{1}{\omega}\sqrt{\frac{a^2+\omega^4b^2+\omega^2d^2-\omega^2 2ab}{\omega^2c^2+e^2}} \tag{89}$$

If r_b is small compared to R_s and r_π, and C_{sbp} is small compared to C_π, then R_{sbp} in (89) is given by the simple expression below.

$$R_{sbp} = \frac{1}{\omega C_{sbp}} \tag{90}$$

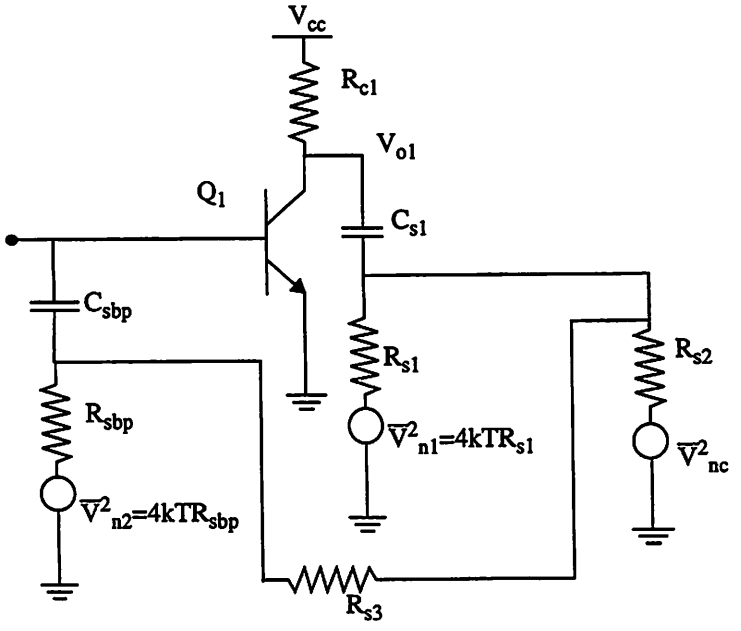


Figure 29a: Noise Sources in a Single-Transistor

Thus, as in the previous section, if both the frequency of operation and the bond-pad capacitance are known, then the substrate resistance must be either much greater than or much smaller than the reactance of the capacitance. The latter case can be achieved by the use of a bond-pad shown in Fig. 29b. In this figure a bond-pad with a grounded low-impedance shield underneath is shown ([23], [24]). If the bond-pads are made from the second or third-layer metallization then the shield can be formed by a lower metal layer or by a n+-shield underneath the bond-pad. Shields implemented in the n+-layer are preferable to p+ -shields because the depletion capacitance between the n+-shield and the p-type substrate isolates the shield from the substrate.

The contribution of R_{s1} to the output noise power-spectral density is given by the following equation

$$\frac{d\bar{v}^2}{df}_o = \frac{4kTR_{c1}^2\omega^2C_{s1}^2R_{s1}}{1+(R_{s1}+R_{c1})^2\omega^2C_{s1}^2} \tag{91}$$

This contribution is maximized for the following value of R_{s1}

$$R_{s1} = \frac{\sqrt{\left(1+\omega^2C_{s1}^2R_{c1}^2\right)}}{\omega C_{s1}} \tag{92}$$

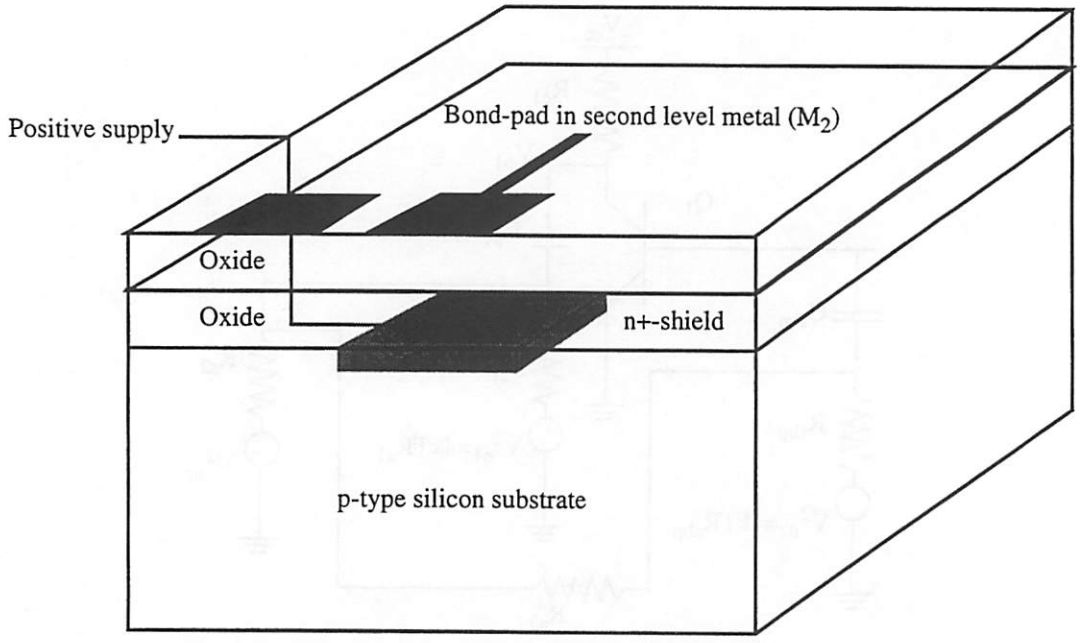


Figure 29b: Bond-Pad with Low-Impedance Shield

Once again the noise is maximized when the substrate resistance is approximately equal to the reactance of the collector-to-substrate capacitance.

The influence of the switching noise on the output noise in the circuit is more difficult to model as it is difficult to make any generalizations regarding the nature of this signal. The switching 'noise' may in fact be deterministic and not random at all. Even if the switching activity is random, it is difficult to determine a priori the distribution or the frequency dependence of the noise. In specific cases, circuit simulations, which incorporate knowledge of the switching activity, can be carried out.

In order to illustrate some of the trade-offs involved in LNA designs, in the presence of switching noise, we consider the example shown in Fig. 29c. A tuned LNA is shown in the figure. The load impedance is tuned to achieve a center-frequency of 900MHz. It is assumed that the inductor and the bond-pad used in the LNA are shielded. Thus the only noise sources contributing to the output noise are the device noise sources and the switching noise coupled to the output. The relevant device parameters and the inductor Q are shown in the figure. The input voltage source is assumed to be ideal.

All the switching noise-sources are combined together as a single large noise-source of area $500*500\mu m^2$, coupled to the substrate through a capacitance of 40pF[11]. The switching noise source is

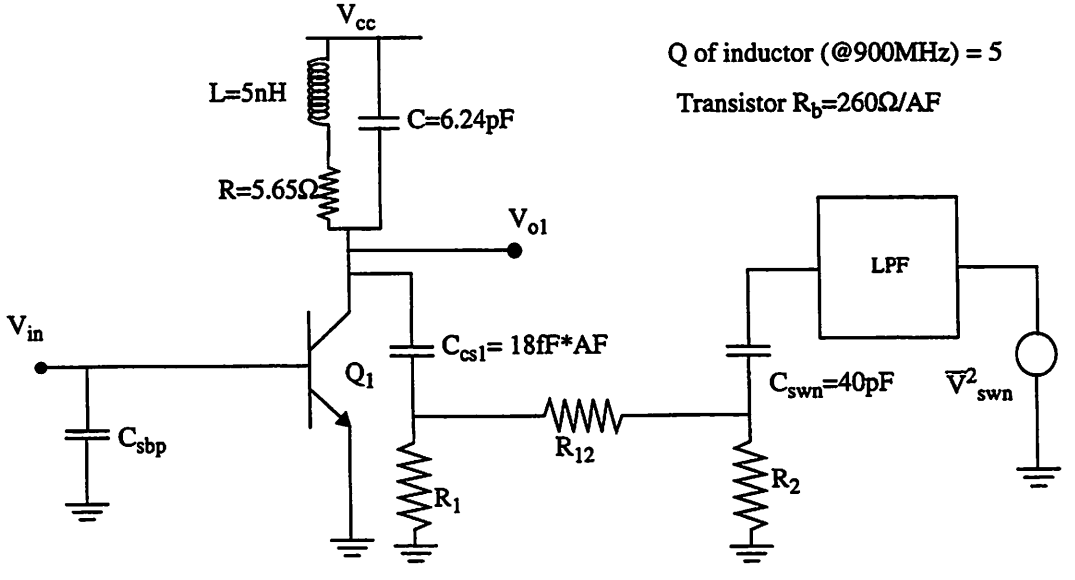


Figure 29c: Modeling of Switching Noise Sources

11. Typical device-parameter values have been assumed. These number may vary, depending on the process.

assumed to be AWGN, and is filtered by a low-pass filter in order to restrict the bandwidth of the noise. The layout of the LNA and the switching noise-source is shown in Fig. 29d.

The area of transistor Q_1 in Fig. 29c is stepped from twice to twelve-times the minimum device size. This ratio is called AF in Fig. 29c and Fig. 29e. The device bias is adjusted in order to achieve a gain of 5 (14dB) at the center frequency. The total noise at the output of the amplifier, with an ideal voltage source driving the input, in shown in Fig. 29e. Curve 'a' in Fig. 29e shows the output noise voltage considering only the device noise sources. Curves 'b', 'c' and 'd' show the total output noise when the total switching noise-voltages in a 1GHz bandwidth are 10mV, 100mV and 1V respectively. The total output noise in curve 'a' decreases as the device area is increased. This is due to the decrease in the value of the base resistance r_b. Curve 'b' shows the existence of an optimum device area. The output noise decreases at first as the area is increased and then increases for device areas greater than five times the minimum device size. The initial decrease is due to the decrease in the base resistance, while the increase is due to the increase in the collector-to-substrate capacitance of the device. The output-noise is dominated by the switching noise in curves 'c' and 'd' and the output-noise is a monotonically increasing function of the device area.

In order to decrease the influence of the switching noise on the output noise of the amplifier, guard rings and backplane contacts can be employed. The optimum device area for minimizing the output noise

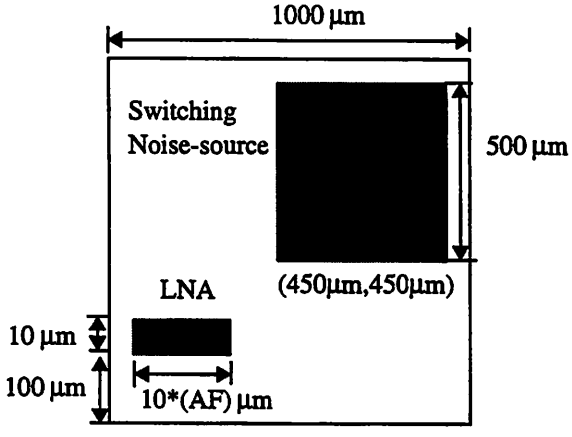


Figure 29d: Layout of LNA and Switching-Noise Generator

can then be recomputed.

Oscillators are susceptible to substrate-noise in a similar manner. A very important specification for an oscillator is the phase-noise performance [25]. The phase-noise of an oscillator can be degraded by the thermal noise of the substrate resistor or by the external switching noise coupled into the circuit. As in the LNA, the thermal noise of the substrate can be reduced by doping the substrate suitably, and external noise can be reduced by the use of guard rings in high-resistivity substrates or good backplane contacts in low-resistivity substrates.

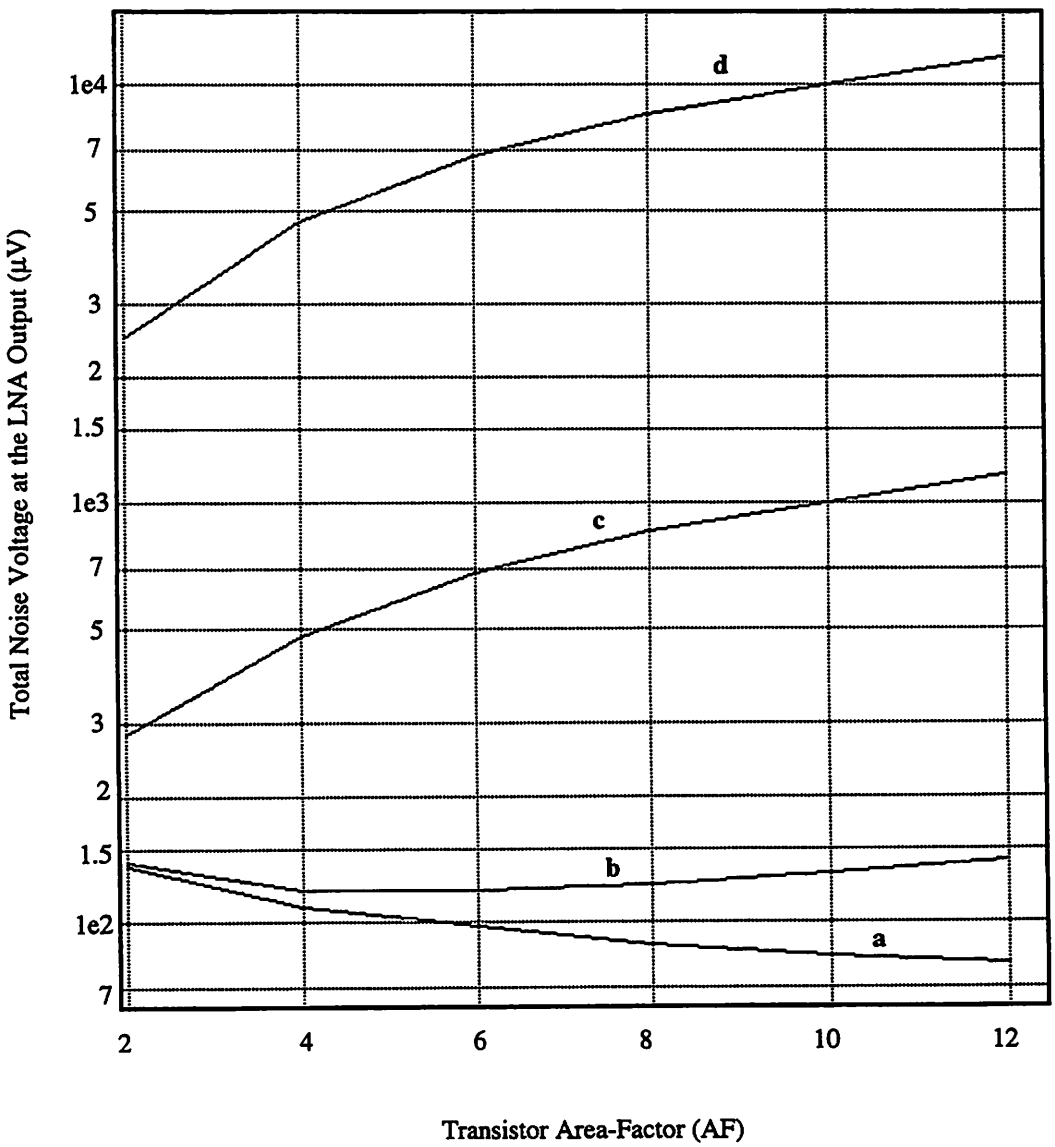


Figure 29e: Total Output Noise vs. Device Area

4.2.3: Change in Circuit Bandwidth

Circuit bandwidth is a critical parameter in amplifier design. Circuit bandwidth can be changed in at least two ways due to the substrate-impedance. The substrate acts as a feedback path, which can lead to significant changes in circuit-bandwidth. Yet another mechanism is the change in the device-to-substrate depletion capacitance of devices caused by a change in the substrate bias resulting from the flow of DC currents in the substrate. As discussed in Section 2.0.1, DC currents are injected into the substrate by the avalanche-induced substrate currents in NMOSFETs, the leakage currents in device-to-substrate pn junctions and the parasitic pnp transistor in npn BJTs. Experimental evidence of the change in depletion capacitance in NMOSFETs was shown in [26]. These two mechanisms are discussed below.

4.2.3.1: The Substrate as a Feedback Path

In order to examine the feedback path which exists through the substrate, we consider the BJT amplifier shown in Fig. 30. The partial circuit diagram of a cascade of common-emitter amplifiers is shown in the figure. C_{sbp} is the capacitance between the bond-pad and the substrate and R_{sbp} is the substrate resistance from the bond-pad to the ground. C_{sbp} is typically in the range of 0.5pF to 1pF while R_{sbp} can vary from a few ohms to a few kΩ, depending on the substrate type used. $C_{s1(2)}$ and $R_{s1(2)}$ are the

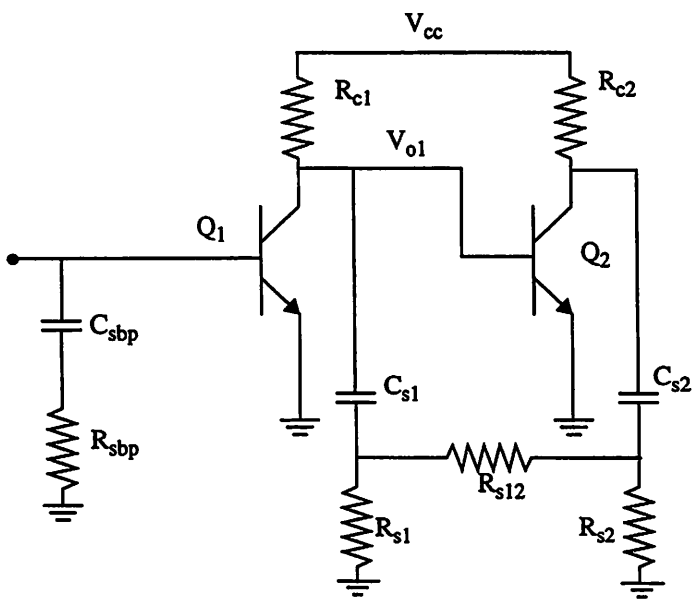


Figure 30: Feedback through the Substrate

collector-to-substrate capacitances and the substrate resistances respectively for $Q_{1(2)}$. R_{s12} models the interaction of $Q_1$and Q_2 through the substrate.

If the substrate model is not included in the circuit simulator, then the collector-to-substrate capacitors in Fig. 30 will be connected to the ground. With the substrate model included, a feedback path between the collector and the base of transistor Q_2 can be seen from the figure. If R_{s1} and R_{s2} are large compared to the reactance of C_{s1} and C_{s2} at the frequencies of interest, and if R_{s12} is small, then the series combination of C_{s1} and C_{s2} appear as a Miller-multiplied capacitance at the collector of Q_1[12]. This effect can be significant if Q_2 provides a large gain. For example, if C_{s1} and C_{s2} are 0.2pF each and Q_2 has a gain of 10 then the Miller-multiplied term at the collector of Q_1 will be 1pF. This capacitance will add directly to the C_π of Q_2 and will lead to a large fractional change in the bandwidth of the circuit.

Possible solutions to this problem include the use of a guard ring around the collector of Q_1, the use of a good backplane contact to reduce R_{s1} and R_{s2}, the use of differential circuits and the use of feedback around the amplifier. The use of feedback may not always be sufficient, depending on the application. The reduction in the bandwidth of a series-series feedback amplifier is discussed in [27], where the inclusion of the substrate macromodel in the circuit simulator is shown to cause a reduction in the circuit bandwidth from 1.9GHz to 1.8GHz, which represents a five percent change.

The substrate acts as a feedback path in MOS amplifiers as well and the lossy Miller effect discussed above can be seen there as well. A DC feedback path can also exist in MOS amplifiers due to the body effect and the drain-to-body transconductance (g_{db}) caused by hot-electron induced holes (Chapter 2). These effects can be studied in specific cases by using the simulator discussed earlier.

4.2.3.2: Change in Depletion Capacitance

If a DC current is injected into the substrate by any mechanism, the substrate potential will change due to the voltage drop in the substrate. The change in the substrate potential will vary across the substrate, which can be predicted by considering the complete substrate-macromodel in the circuit simulator. The

12. This can be the case, for example, in a high-resistivity substrate if the transistors are in close proximity, have large substrate capacitance, and the substrate is connected to ground by the use of small surface substrate contacts.

depletion capacitance of a reverse-biased pn junction depends on the DC bias value by the following relationship

$$C_j = \frac{C_{j0}}{\left(1 + \frac{V_{bias}}{\psi_{bi}}\right)^{\frac{1}{n}}} \tag{93}$$

where V_{bias} is the DC bias in the substrate, ψ_{bi} is the built-in potential of the junction and 'n' is an empirical exponent. 'n' depends on the doping profile. Using the above equation, it can be shown that a small change ΔV_{bias} in the substrate potential will cause the following change in the junction capacitance

$$\Delta C_j = -\frac{C_j}{n}\frac{((\Delta V_{bias})/\psi_{bi})}{\left(1 + \frac{V_{bias}}{\psi_{bi}}\right)} \tag{94}$$

Thus if V_{bias} is initially zero, n is assumed to be[13] 2 and ψ_{bi} is assumed to be 0.7V, then a 0.1V increase in V_{bias} changes C_j by approximately seven percent. If the bandwidth is dependent on the junction capacitance with the substrate, then this change would represent a seven percent change in the bandwidth of the circuit.

A circuit technique to reduce this problem is to make the bandwidth dependent on a passive capacitor built on the chip. A processing technique to reduce the change in the junction capacitance values is to connect the substrate region in the vicinity of the device to ground, which could be done by using a guard ring.

4.2.4: Change in Circuit Gain

The change in the substrate bias can cause a change in the depletion capacitance which exists between the device and the substrate. In MOS transistors a DC change in the threshold voltage V_t can also result. A change in the threshold voltage will lead to a change in the gain of the device and also the bias quantities in the circuit. As discussed in [2], a change of ΔV_{bias} in the substrate potential changes the drain-current I_d of a MOSFET by the following relation

13. An abrupt junction is assumed.

$$\Delta I_d = g_{mb}\Delta V_{bias} \tag{95}$$

The change in the bias level for each circuit or device can be obtained from the circuit simulator.

4.2.5: Catastrophic Consequences of Substrate Coupling

It was shown in Section 4.2.2 that when the substrate acts as a negative feedback path, the bandwidth of the circuit is changed due to the lossy-Miller effect. A much more serious consequence is observed when the substrate appears to be a positive feedback path.

As an example, the cascade of common-emitter stages considered in Fig. 30 is considered once again, but with one more transistor added to the chain, as shown in Fig. 31 below. As can be seen from this circuit-diagram, a positive feedback path now exists through the collector of Q_3 and the base of Q_2. The isolation in high-resistivity substrates is in the range of -45dB to -35dB (Fig. 19g). Thus if Q_2 and Q_3 have a combined gain greater than 40dB, it is possible to observe a loop-gain greater than unity through the substrate. This could lead to oscillations of the circuit and will lead to circuit-failure.

Similar positive feedback paths through the substrate can also exist in MOS amplifiers even at low-frequencies as discussed earlier. It is thus very important to ensure that sufficient isolation is provided

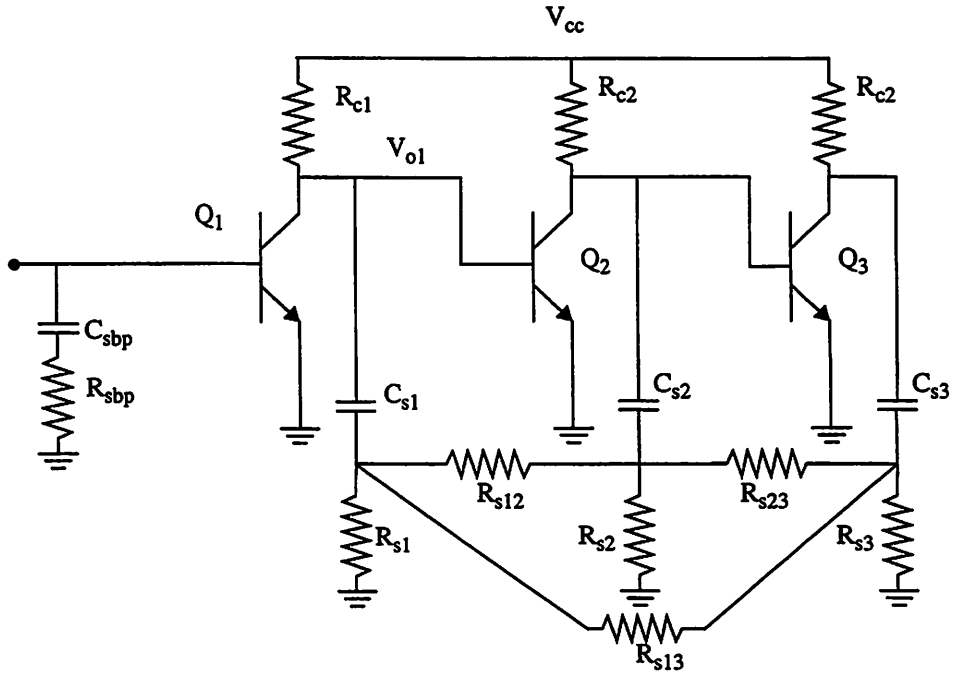


Figure 31: Positive Feedback through the Substrate

through the substrate, when high-gain amplifiers are implemented, or when high-gain paths exist in a circuit. Latch-up in CMOS circuits is a classic case of positive feedback through the substrate and has been extensively studied in the past [1]. Guard rings and backplane contacts can be used to reduce the loop-gain through the substrate. Low-resistivity substrates are in fact popular choices for CMOS fabrication because of their latch-up suppressing properties. In these substrates the contact-to-contact coupling term is very large which reduces the loop-gain of the positive feedback path causing latch-up (Fig. 18d). The utility of the substrate simulator is apparent in this situation, for predicting, identifying and solving the problem.

Yet another catastrophic failure can be caused when a device-to-substrate gets forward-biased by a voltage spike, or by large DC currents injected into the substrate by any of the three mechanisms identified earlier. Turning on of a substrate diode can cause a large current flow into the substrate and can lead to permanent destruction of the IC.

4.3: Conclusion

This chapter has been devoted to the circuit effects of substrate coupling and techniques to suppress them. Substrate isolation in different substrate types was discussed extensively. It was shown that low-resistivity substrates can provide excellent differential mode isolation and can provide good single-ended isolation in the presence of a good ground-plane contact. The 'distance-independent-isolation' property of these substrates was discussed in detail. Isolation in high-resistivity substrates was shown to be relatively independent of the backplane impedance. Guard rings were found to be effective in these substrates, and the optimal sizing of guard rings in these substrates was discussed. It was demonstrated that the isolation provided by narrow guard rings is as good and in some cases better than that provided by wide guard rings. It was shown that guard rings provide very good isolation in differential circuits on high-resistivity substrates. Several circuit effects of substrate coupling such as bandwidth reduction, noise-figure degradation, substrate power-loss and circuit-failure mechanisms were discussed in the last part of the chapter. It was shown that in many circuit applications, it is important to know a priori the effects of substrate coupling and to take the necessary design steps to avoid these effects.

Chapter 5: Experimental Verification

Experimental verification of the modeling techniques discussed in Chapter 3 is presented in this chapter. DC measurements were performed on a chip consisting of several p+-to-p contacts. The experimental results are presented and correlated against the predictions of the simulator. A good match is observed between the simulated and experimental results. This experiment is used to verify the predictions of the simulator rather than explore the circuit effects of substrate coupling. An analysis of the sources-of-error is also performed in this chapter.

5.1: Test-Chip Description

A die-photo of the test-chip is shown in Fig. 32a. The cross-section of the test-chip substrate is shown in Fig. 32b. The substrate is of the high-resistivity type discussed earlier in Chapter 4. It consists of two layers, a low-doping high-resistivity bulk of resistivity 20Ω-cm and a low-resistivity surface buried-p region of resistivity 0.1Ω-cm. The test-chip consists of twenty ohmic substrate-contacts (p+-to-p). The minimum contact size is 2μm by 2μm, while the largest contact size is that of the center square contact, which is 125μm to a side. There are four rings surrounding the center contact. The innermost ring is a large p+-contact while the outer rings consist of several individual p+ substrate-taps connected in parallel on the surface by metal interconnects. The contact-to-contact distances vary from 2μm to 100μm. The die measured 950μm by 875μm. A gold ohmic-contact was made to the back surface of the die and was connected to ground during the testing.

5.2: Measurement Procedure

Two experiments were performed on the test-chip. In the first experiment, a DC bias of 0.1V was applied at one contact, and the voltage appearing at another contact, with all other contacts at ground, was measured. Two values of DC bias were chosen for this purpose (0.1V and 1V) in order to ensure that the junction nonlinearity or high-field effects in the bulk do not interfere with the measurement. It was found that some of the contacts exhibited weak nonlinearity at 1V and therefore data was extracted at 0.1V bias only. The linearity around zero-bias was verified by using an IV curve-tracer. For any pair of contacts I and J, the voltage measured at contact-J (V_J) with 0.1V applied at contact-I is given by the following

equation

$$V_J = \frac{0.1R_{J0}}{R_{IJ}+R_{J0}} \tag{96}$$

where R_{IJ} is the resistance between contacts I and J and R_{J0} is the resistance between contact-J and ground. The ratio of V_J and V_I is shown in Fig. 32c.

The admittance matrix was experimentally determined, in the second experiment, and compared against simulated data. This extraction was performed on two test-chips. One contact was kept at a fixed DC bias of 0.1V as above and the current exiting the other contacts was measured.

5.3: Simulation Procedure

The two experiments discussed above were repeated on the simulator as well. The DCT technique discussed in the previous sections was used for the simulation. A minimum feature size of approximately 1μm and junction-depths of 0.5 μm were used in the simulation. The substrate model is shown in Fig. 32b.

5.4: Discussion of Results

Experimental and simulated results are compared in the plots shown in Fig. 32c and Fig. 32d. The data from the voltage-ratio experiment is shown in Fig. 32c. Voltage-ratios greater than 0.5 imply a strong coupling between two contacts. It can be seen from the figure that the match between the measured and experimental data is to within fifteen percent for voltage-ratios greater than 0.5. The match is good for voltage-ratios larger than 0.1. Below a ratio of 0.1 however, the matching between the experimental and the simulated values degrades considerably.

Fig. 32d shows the current exiting the contacts at ground with the center square contact at 0.1V. The current varies from 2mA to 40mA which implies that the smallest contact-to-contact resistance is 25Ω. The match between experimental and simulated data is seen to be very good for this case.

Sources-of-Error

One of the sources-of-error in these experiments is the approximate substrate-profile used in the simulations. The exact values of the bulk and the epitaxial layer resistivities can vary from one test-chip to another. Further, the step-profile assumed in Fig. 32b is an approximation to the physical case of a

gradually varying profile.

The approximate nature of the resistivity profile can be identified as a possible source-of-error from the data points for voltage-ratios less than 0.1 in Fig. 32c. The simulated values are observed to be consistently lower than the experimentally measured data, which implies that the potential falls-off faster in the simulation than in the physical substrate. The potential roll-off is directly dependent on the substrate profile. Further, while the exact match between the experimental and the simulated data degrades at low

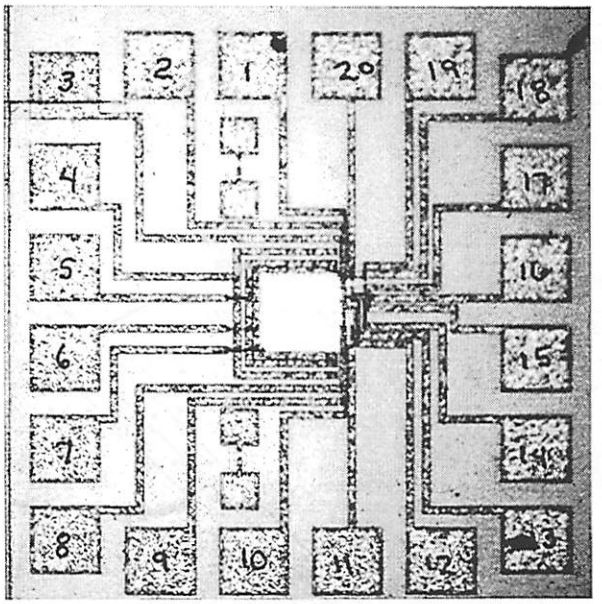


Figure 32a: Die-Photograph

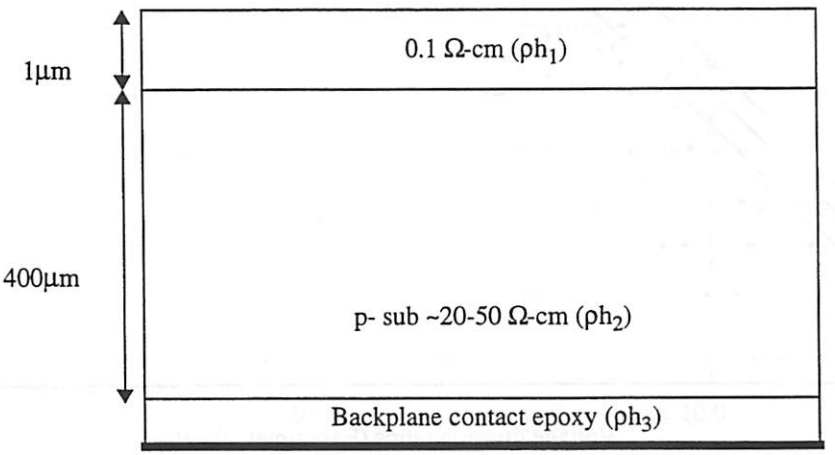


Figure 32b: Substrate Cross-Section

voltage-ratios, the two sets of data show a good correlation, which indicates the presence of a systematic error. An approximate substrate-profile will cause a systematic mismatch between the data.

The experimental data is noisy at small voltage-ratios. This possibly could be measurement-related noise. This issue needs to be investigated further. Parasitics not included in the simulator, for example, the contact resistances, are yet another source-of-error. These parasitics especially impact measurements between contacts with small contact-to-contact resistances.

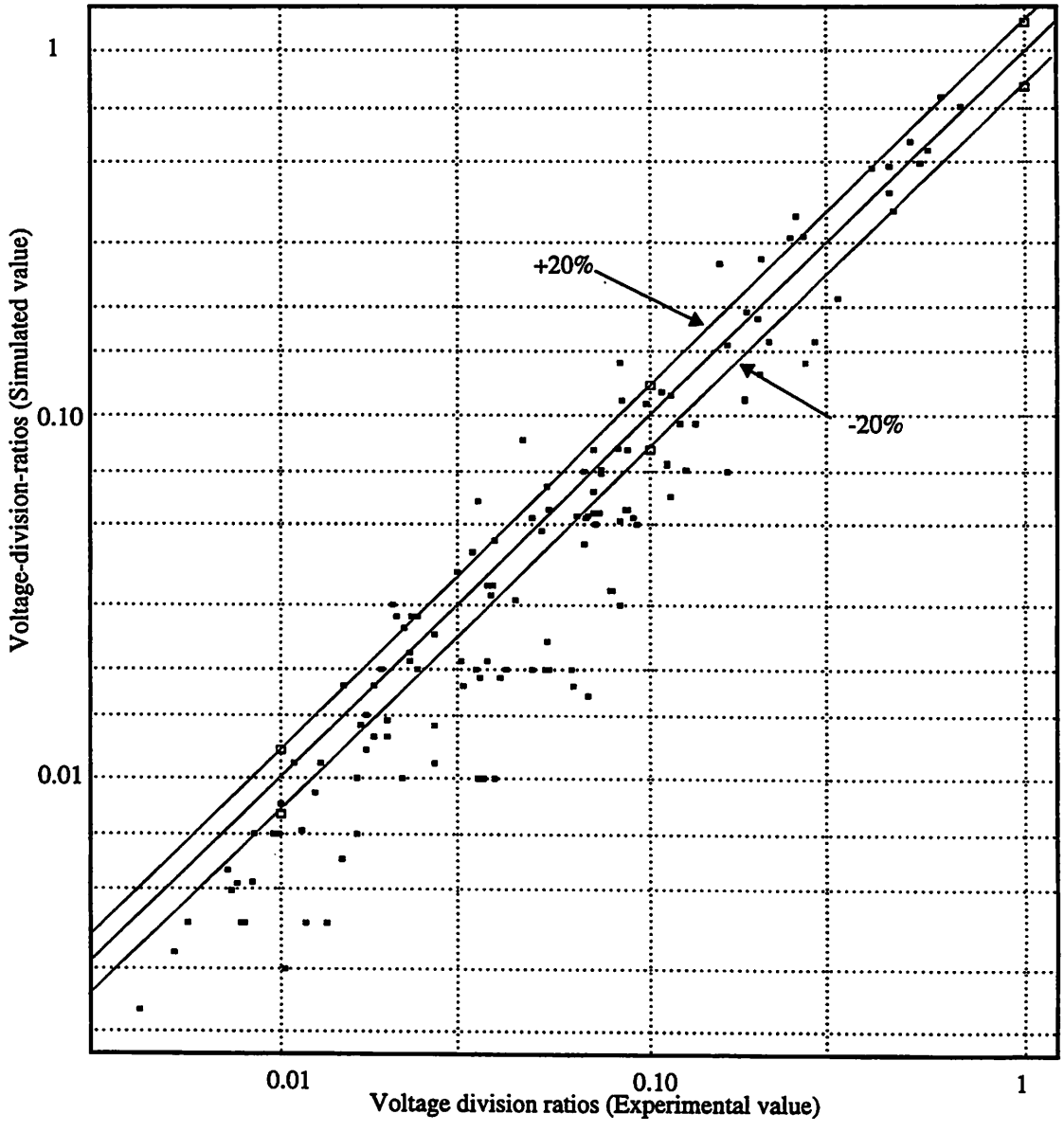


Figure 32c: Simulated vs. Experimental Results

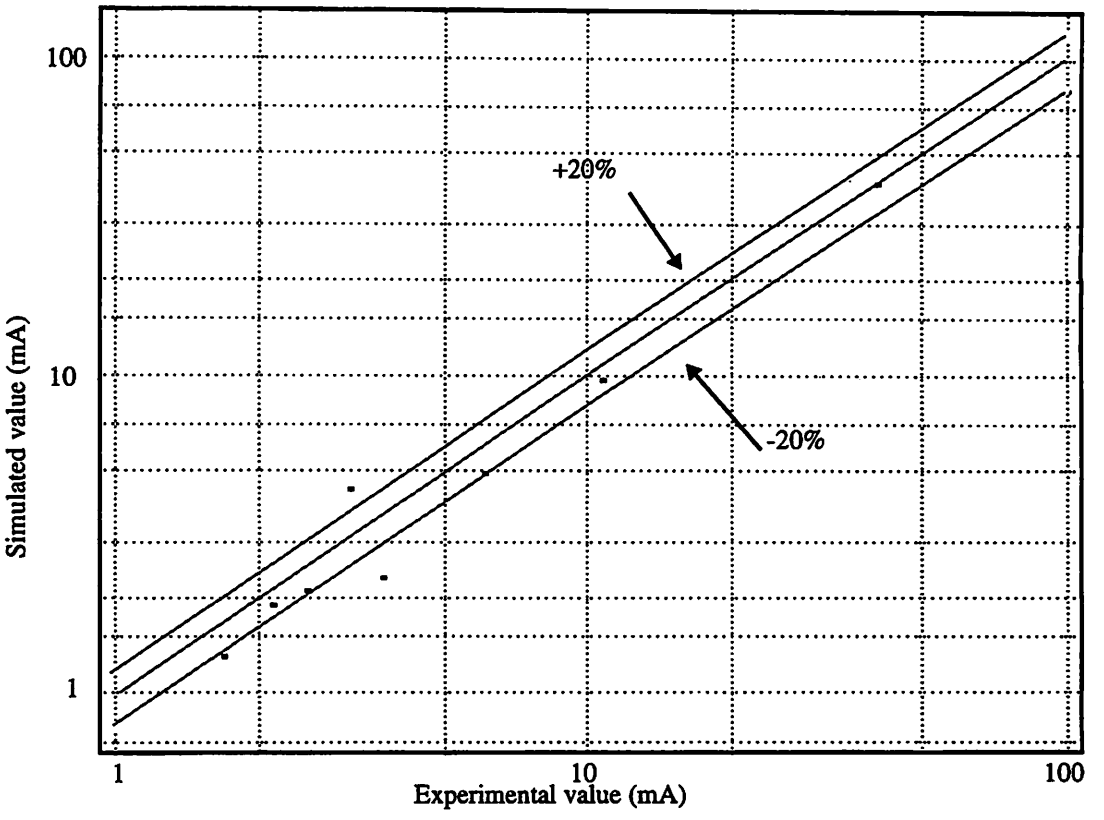


Figure 32d: Simulated vs. Experimental Results

The good match obtained between the measured and simulated values implies that the simulator is accurate. Greater accuracy can be achieved by the use of more accurate substrate-profiles. The simulation technique can be extended to several vertical layers with ease, without major computational overhead, which is one of the attractive features of the technique. The test-chip used for experimental verification provided a severe test for the capabilities of the simulator. This was due to the wide variation in the contact-size and the contact-to-contact distance in addition to the small dimensions of the minimum-size contacts and the small distance between the closest contacts.

Chapter 6: Conclusions and Future Directions

Substrate coupling in integrated circuits is a challenging design problem. This phenomenon is one of the major hurdles in the design of highly integrated mixed analog-digital circuits. The problem will become a significant issue in future generations of high-speed digital circuits as well. This thesis addresses efficient modeling of substrate coupling and examines the related isolation and circuit issues. A fast, memory-efficient and accurate modeling scheme has been presented. The use of this scheme makes it possible to optimize layout in order to minimize substrate coupling at the design stage. Experimental verification of the models has been obtained. Optimal guard ring sizing in the presence of external noise sources has been addressed. Several circuit-level effects of substrate coupling and their solutions have also been discussed.

There are several problems which deserve further attention. The capabilities of the simulator can be enhanced in several ways. Lateral resistivity-variations were addressed briefly in Chapter 3. The ability to incorporate these variations in a memory-efficient manner is an interesting problem. More efficient partitioning techniques can be explored (Section 3.2.5.2), to reduce the size of the substrate matrices.

DC experimental verification was presented in this thesis. Experiments involving simple structures have been presented in [21] and [28]. More work needs to be done on the circuit-level aspects of substrate coupling at high-frequencies. Several theoretical predictions on guard rings were made in Chapter 4 which need to be experimentally verified at high-frequencies.

The magnitude of substrate coupling depends strongly on the type of the package used in the IC. This dependence can be caused by several factors. For example, the value of the bond-wire inductances or the availability of good backplane contacts in a package will influence the degree of coupling. A comparison of substrate coupling in new packaging technologies such as Multichip-Modules also needs to be performed.

References

[1] S. M. Sze, *Physics of Semiconductor Devices*, New York: Wiley, c1981.

[2] P. R. Gray and R. G. Meyer, *Analysis and Design of Analog Integrated Circuits*, New York: Wiley, c1984.

[3] C. Hu, *VLSI Electronics: Microstructure Science*, Chapter 3, Vol. 18, New York: Academic Press, 1981.

[4] R. B. Merrill, W. M.Young and K. Brehmer, *Effect of Substrate Material on Crosstalk in Mixed Analog/Digital Integrated Circuits*, IEDM technical digest: International Electron Devices Meeting, San Francisco, CA, pp. 433-436, December 11-14, 1994.

[5] M. Hasegawa, N. Furukawa and M. Yanay, *Properties of Microstrip Lines on Si-SiO_2 System*, IEEE Transactions on Microwave Theory and Techniques, vol. MTT-19, pp. 869-881, 1971.

[6] S. Ramo, J. Whinnery and T. Van Duzer, *Fields and Waves in Communication Electronics*, New York: Wiley, c1994.

[7] J. D. Jackson, *Classical Electrodynamics*, 2d ed. New York: Wiley, c1975.

[8] D. M. Young, *Iterative Solution of Large Linear Systems*, New York, Academic Press, 1971.

[9] R. S. Varga, *Matrix Iterative Analysis*, Englewood Cliffs, N.J., Prentice-Hall, 1962.

[10] L. A. Hageman and D. M. Young, *Applied Iterative Methods*, New York: Academic Press, 1981.

[11] J. A. Meijerink and H. A. van der Vorst, *An Iterative Solution Method for Linear Systems of which the Coefficient Matrix is a Symmetric M-Matrix*, Mathematics of Computation, Vol. 31, pp. 148-162, January 1977

[12] N. K. Verghese, D. J. Allstot and S. Masui, *Rapid Simulation of Substrate Coupling in Mixed-Mode ICs*, Proc. IEEE Custom Integrated Circuits Conference, pp. 18.3.1-18.3.4, May 1993.

[13] X. Huang, V. R. Raghavan and R. A. Rohrer, *AWEsim: A Program for the Efficient Analysis of Linearized Circuits*, IEEE International Conference on Computer Aided Design, 11-15 November, 1990, Washington, DC, USA, Digest of Technical Papers, pp. 534-537.

[14] B. R. Stanisic, N. K. Verghese, L. R. Carley and D. J. Allstot, *Addressing Substrate Coupling in*

Mixed-Mode ICs: Simulation and Power Distribution Synthesis, IEEE Journal of Solid State Circuits, Vol. 29, No. 3, pp. 226-238, March 1994.

[15] Chen-to Tai, *Dyadic Green Functions in Electromagnetic Theory*, Scranton, Intext Educational Publishers, 1971.

[16] C. Wei, R. F. Harrington, J. R. Mautz and T. K. Sarkar, *Multiconductor Transmission Lines in Multilayered Dielectric Media*, IEEE Transactions on Microwave Theory and Techniques, Vol.32, No.4, pp. 439-449, April 1984.

[17] I. N. Sneddon, *Mixed Boundary Value Problems in Potential Theory*, Amsterdam, North-Holland Pub. Co.; New York, Wiley, 1966.

[18] W. T. Weeks, *Calculation of Coefficients of Capacitance of Multiconductor Transmission Lines in the presence of a Dielectric Interface*, IEEE Transactions on Microwave Theory and Techniques, Vol.18, No.1, pp. 35-43, January 1970.

[19] G. Dahlquist and A. Bjorck, *Numerical Methods*, Translated by N. Anderson, Englewood Cliffs, N.J., Prentice-Hall, 1974.

[20] K. Nabors and J. White, *FastCap: A Multipole Accelerated 3-D Capacitance Extraction Program, IEEE Transactions on Computer-Aided Design*, Vol.10, No. 11, pp. 1447-1459, November 1991.

[21] 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.

[22] N. M. Nguyen, *Monolithic Microwave Oscillators and Amplifiers*, Ph.D. dissertation, University of California at Berkeley, 1991.

[23] N. Camilleri, J. Kirchgessner, J. Costa, D. Ngo and D. Lovelace, *Bonding Pad Models for Silicon VLSI Technologies and Their Effects on the Noise Figure of RF NPNs*, IEEE 1994 Microwave and Millimeter-Wave Monolithic Circuits Symposium, San Diego, CA: digest of papers, pp. 225-228.

[24] R. G. Meyer and W. D. Mack, *A 1-GHz BiCMOS RF Front-end IC*, IEEE Journal of Solid State Circuits, Vol. 29, No. 3, pp. 350-355, March 1994.

[25] T. C. Weigandt, B. Kim and P. R. Gray, *Analysis of Timing Jitter in CMOS Ring Oscillators,* Proceedings of the IEEE International Symposium on Circuits and Systems, London, UK, 30 May-2 June 1994.

[26] K. Sakui, S. S. Wong and B. A. Wooley, *The Effects of Hot Carriers Generation on the Operation of Neighboring Devices and Circuits*, Proceedings of IEEE VLSI Technology Symposium, Kyoto, Japan, pp. 11-12, 17-19 May 1993.

[27] R. Gharpurey and R. G. Meyer, *Analysis and Simulation of Substrate Coupling in Integrated Circuits,* to be published in The International Journal of Circuit Theory and Applications.

[28] 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.

Appendix A3.1: The Electrostatic Green Function in the Substrate Medium

Consider a point charge placed at point (x', y', z') in the upper most layer of the multilayered dielectric as shown in Fig. A1. The potential due to the point charge will be the solution of the Poisson equation, that is

$$\nabla^2 G(x, y, z, x', y', z') = -\frac{\delta(x-x')\,\delta(y-y')\,\delta(z-z')}{\varepsilon_N} \tag{A1}$$

Let G = X(x,x')Y(y, y')Z(z, z'). Substituting in the above equation, we have

$$YZ'\frac{d^2X}{dx^2} + Z'X\frac{d^2Y}{dy^2} + XY\frac{d^2}{dz^2}Z' = -\frac{\delta(x-x')\,\delta(y-y')\,\delta(z-z')}{\varepsilon_N} \tag{A2}$$

Assume $X = \cos\left(\frac{m_1\pi x}{a}\right); Y = \cos\left(\frac{n_1\pi y}{b}\right)$ where $m_1 \in [0, \infty), n_1 \in [0, \infty)$, to satisfy the boundary condition of zero normal E-field on the side walls.

We have from (A2)

$$\sum_{m_1=0}^{\infty}\sum_{n_1=0}^{\infty}\cos\left(\frac{m_1\pi x}{a}\right)\cos\left(\frac{n_1\pi y}{b}\right)\times\left\{\frac{d^2}{dz^2}Z' - \left(\left(\frac{m_1\pi}{a}\right)^2 + \left(\frac{n_1\pi}{b}\right)^2\right)Z'\right\} = -\frac{\delta(x-x')\,\delta(y-y')\,\delta(z-z')}{\varepsilon_N} \tag{A3}$$

Multiplying both sides by $\cos\left(\frac{m\pi x}{a}\right)\cos\left(\frac{n\pi y}{b}\right)$ and integrating over x and y from 0 to a and 0 to b

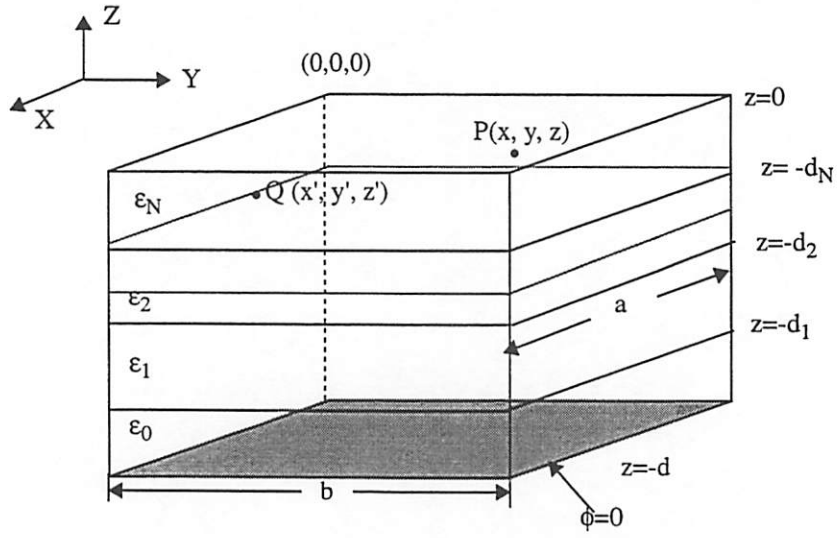


Figure A1: Multilayered-Substrate Model

respectively, we get

$$\frac{ab}{4}\left(\frac{d^2}{dz^2}Z' - \left(\left(\frac{m\pi}{a}\right)^2 + \left(\frac{n\pi}{b}\right)^2\right)Z'\right) = -\frac{\delta(z-z')}{\varepsilon_N}\cos\left(\frac{m\pi x'}{a}\right)\cos\left(\frac{n\pi y'}{b}\right) \tag{A4}$$

Let $Z'(z, z') = Z(z, z')\cos\left(\frac{m\pi x'}{a}\right)\cos\left(\frac{n\pi y'}{b}\right)$

Substituting the above expression in (A4), we have

$$\frac{ab}{4}\left(\frac{d^2Z}{dz^2} - \gamma_{mn}^2 Z\right) = -\frac{\delta(z-z')}{\varepsilon_N} \tag{A5}$$

where $\gamma_{mn} = \sqrt{\left(\frac{m\pi}{a}\right)^2 + \left(\frac{n\pi}{b}\right)^2}$. For $z \neq z'$ the RHS of (A5) is zero. A general solution to (A5) would be of the form $Z = \beta\sinh(\gamma_{mn}(d+z)) + \Gamma\cosh(\gamma_{mn}(d+z))$.

We assume that the source point and the observation point are in the uppermost layer that is $-d_1 \leq (z, z') \leq 0$. We first consider the solution for Z in the region $z < z'$, which is referred to as Z^1. The solution for Z in the bottom dielectric layer (ε_0) must be of the form

$$Z^1 = \beta_0\sinh(\gamma_{mn}(d+z)) + \Gamma_0\cosh(\gamma_{mn}(d+z)) \tag{A6}$$

where β_0 and Γ_0 are constants.

Since the backplane has been assumed to be at zero potential, we have $\Gamma_0 = 0$. Thus

$$Z^1 = \beta_0\sinh(\gamma_{mn}(d+z)) \tag{A7}$$

At the interface between dielectric layer ε_0 and ε_1, two boundary conditions need to be satisfied i.e. the potential and the normal component of the electric displacement must be continuous. This condition will be valid in general at any interface between dielectrics. Specifically, the set of equations relating (β_k, Γ_k) to (β_{k-1},Γ_{k-1}) can be represented in the matrix form as in (61)

$$\begin{bmatrix}\beta_k\\ \Gamma_k\end{bmatrix} = \begin{bmatrix}\frac{\varepsilon_{k-1}}{\varepsilon_k}\cosh^2(\theta_k) - \sinh^2(\theta_k) & , \left(\frac{\varepsilon_{k-1}}{\varepsilon_k} - 1\right)\cosh(\theta_k)\sinh(\theta_k)\\ \left(1 - \frac{\varepsilon_{k-1}}{\varepsilon_k}\right)\cosh(\theta_k)\sinh(\theta_k) & , \cosh^2(\theta_k) - \frac{\varepsilon_{k-1}}{\varepsilon_k}\sinh^2(\theta_k)\end{bmatrix}\begin{bmatrix}\beta_{k-1}\\ \Gamma_{k-1}\end{bmatrix} \tag{61}$$

where $\theta_k = \gamma_{mn}(d - d_k)$, $1 \le k \le N$, $\beta_0 = 1.0$ and $\Gamma_0 = 0$

In the uppermost layer we have

$$Z^l = \beta_N \sinh(\gamma_{mn}(d+z)) + \Gamma_N \cosh(\gamma_{mn}(d+z)) \tag{A8}$$

where $-d_1 \le z \le z' \le 0$. The field must satisfy the zero normal E-field condition at (z=0). Thus in the region $z' \le z \le 0$, the potential, denoted by Z^u, must be of the form

$$Z^u = \Gamma^l{}_N \cosh(\gamma_{mn} z) \tag{A9}$$

Applying symmetry on the function Z, i.e. requiring that Z must be symmetric under the interchange of the source and the observation points, we see that Z^u and Z^l must be of the form

$$Z^u = C(\beta_N \sinh(\gamma_{mn}(d+z')) + \Gamma_N \cosh(\gamma_{mn}(d+z'))) \cosh(\gamma_{mn} z) \tag{A10a}$$

$$Z^l = C(\beta_N \sinh(\gamma_{mn}(d+z)) + \Gamma_N \cosh(\gamma_{mn}(d+z))) \cosh(\gamma_{mn} z') \tag{A10b}$$

From (A10a) and (A10b) it can be observed that if z and z' are interchanged, then Z^u goes into Z^l and vice versa. (A10a) and (A10b) also satisfy the requirement that the potential must be continuous at z = z'.

To determine the constant C, (A5) is integrated over the discontinuity at z=z' which yields

$$\left.\frac{dZ}{dz}\right|_{z'-\delta}^{z'+\delta} = -\frac{4}{ab\varepsilon_N} \tag{A11}$$

At $z = z'+\delta$, Z is given by Z^u while at $z = z'-\delta$, Z is given by Z^l. Substituting (A10a) and (A10b) in (A11) and solving for C, we get

$$C = \frac{4}{ab\varepsilon_N \gamma_{mn}(\beta_N \cosh(\gamma_{mn} d) + \Gamma_N \sinh(\gamma_{mn} d))} \tag{A12}$$

Thus for m>0 and n>0 we have

$$G = \sum_{m=1}^{\infty}\sum_{n=1}^{\infty}\left\{\frac{4\left(\beta_N \sinh\left(\gamma_{mn}(d+z_l)\right)+\Gamma_N\cosh\left(\gamma_{mn}(d+z_l)\right)\right)\cosh\left(\gamma_{mn}z_u\right)}{ab\varepsilon_N\gamma_{mn}\left(\beta_N\cosh\left(\gamma_{mn}d\right)+\Gamma_N\sinh\left(\gamma_{mn}d\right)\right)}\right\}\times$$

$$\cos\left(\frac{m\pi x}{a}\right)\cos\left(\frac{m\pi x'}{a}\right)\cos\left(\frac{n\pi y}{b}\right)\cos\left(\frac{n\pi y'}{b}\right) \qquad \text{(A13)}$$

where z_u=max(z, z') and z_l =min(z, z').

For the case when (m=0, n>0) or (m>0, n=0), the form of (A13) remains the same except that the multiplicative constant 4 is replaced by 2. For the case (m=n=0) we see from (A5) that the general solution for Z must be of the form Z = Az+B. Following a similar argument as presented in the derivation of (A13), Z^l in the ε_0 layer is given by

$$Z^l = a(z')(z+d) \qquad \text{(A14)}$$

The function Z^l in any layer 'k' will be given by

$$Z^l = a(z')(\beta_k z+\Gamma_k) \qquad \text{(A15)}$$

Equating the potential and the field at every interface, we have the following matrix relation between (β_k, Γ_k) and (β_{k-1}, Γ_{k-1}).

$$\begin{bmatrix}\beta_k\\ \Gamma_k\end{bmatrix} = \begin{bmatrix}\frac{\varepsilon_{k-1}}{\varepsilon_k} & 0\\ \left(\frac{\varepsilon_{k-1}}{\varepsilon_k}-1\right)d_k & 1\end{bmatrix}\begin{bmatrix}\beta_{k-1}\\ \Gamma_{k-1}\end{bmatrix} \qquad \text{(A16)}$$

where β_0=1.0 and Γ_0=d,

The solution for Z^u must be of the form $Z^u = b(z')$. Applying the symmetry argument we have

$$Z^l = C(\beta_N z+\Gamma_N) \qquad \text{(A17a)}$$

$$Z^u = C(\beta_N z'+\Gamma_N) \qquad \text{(A17b)}$$

Integrating (A5) over the discontinuity at z = z'' we get

$$C = \frac{1}{ab\varepsilon_N\beta_N} \qquad \text{(A18)}$$

Thus, for m=n=0 we have

$$G_0 = \frac{1}{ab\varepsilon_N\beta_N}(\beta_N z_1 + \Gamma_N) \tag{A19}$$

The final expression for G is then given by (58).

The Green function for a multilayered substrate with a ground-plane has been derived here. The case of a two-layer substrate without a backplane has been analyzed in [A1].

Appendix A3.2: Derivation of the Coefficients of Potential

Equation (65) gives an expression for the coefficient-of-potential between two contacts i and j. This expression uses the definition in (64) for planar surface contacts. Equation (65) can be derived as a special case of three-dimensional contacts with a finite thickness.

Consider two contacts 'i' and 'j' with X, Y and Z coordinates $[(a_1,a_2), (b_1,b_2), (-c_2,0)]$ and $[(a_3,a_4), (b_3,b_4), (-c_4,0)]$ respectively, as shown below.

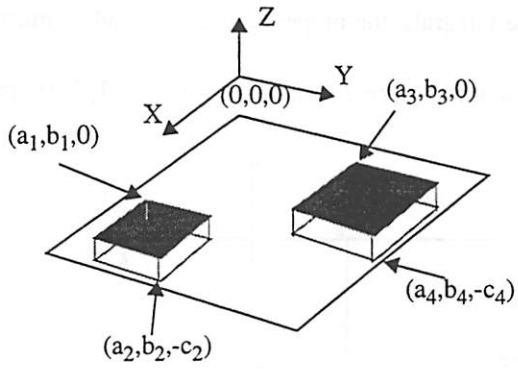


Figure A2a: Contact Layout

Using the volume definition for 'p_{ij}' from (49) we have

$$p_{ij} = \frac{1}{V_i V_j} \int_{-c_4}^{0} \int_{b_3}^{b_4} \int_{a_3}^{a_4} \int_{-c_2}^{0} \int_{b_1}^{b_2} \int_{a_1}^{a_2} G(x, y, z, x', y', z')\, dx\,dy\,dz\,dx'\,dy'\,dz' \tag{A20}$$

where $V_{(i,j)} = (a_{(2,4)}-a_{(1,3)})(b_{(2,4)}-b_{(1,3)})(c_{(2,4)})$ and G is given by (58). The order of integration can be changed as G is separated into X, Y and Z-dependent functions. If the integrations over X and Y coordinates are performed first we have, excluding the m=n=0 case,

$$p_{ij} = \frac{a^2 b^2}{\pi^4} \cdot \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} \left\{ \left(\int_{-c_4}^{0} \int_{-c_2}^{0} Z(z, z')\, dz\,dz' \right) \left(\sin\left(m\pi\frac{a_2}{a}\right) - \sin\left(m\pi\frac{a_1}{a}\right) \right) \left(\sin\left(m\pi\frac{a_4}{a}\right) - \sin\left(m\pi\frac{a_3}{a}\right) \right) \right.$$

$$\left. \frac{\left(\sin\left(n\pi\frac{b_2}{b}\right) - \sin\left(n\pi\frac{b_1}{b}\right) \right) \left(\sin\left(n\pi\frac{b_4}{b}\right) - \sin\left(n\pi\frac{b_3}{b}\right) \right)}{m^2 n^2 (a_2 - a_1)(b_2 - b_1)(c_2)(a_4 - a_3)(b_4 - b_3)(c_4)} \right\} \tag{A21}$$

where Z(z, z') is given by the term in brackets in (A13). Z(z, z') can be rewritten in a more convenient form as

$$Z(z,z') = \left(\frac{4(\beta_N \tanh(\gamma_{mn}d) + \Gamma_N)}{ab\varepsilon_N\gamma_{mn}(\beta_N + \Gamma_N \tanh(\gamma_{mn}d))} \cosh(\gamma_{mn}z_l) + \frac{4}{ab\varepsilon_N\gamma_{mn}} \sinh(\gamma_{mn}z_l) \right) \cosh(\gamma_{mn}z_u) \qquad (A22)$$

where z_u=max(z, z') and z_l =min(z, z'). Two integrals need to be computed in (A21). These are called I_1 and I_2 and are defined below.

$$I_1 = \int_{-c_4}^{0}\int_{-c_2}^{0} \cosh(\gamma_{mn}z_l)\cosh(\gamma_{mn}z_u)\,dzdz'; I_2 = \int_{-c_4}^{0}\int_{-c_2}^{0} \sinh(\gamma_{mn}z_l)\cosh(\gamma_{mn}z_u)\,dzdz' \qquad (A23)$$

In order to compute the above surface integrals, the proper values of z_l and z_u must be used as shown in Fig. A2b below. This choice is required in the expression for I_2 alone, since I_1 is symmetric in z_l and z_u.

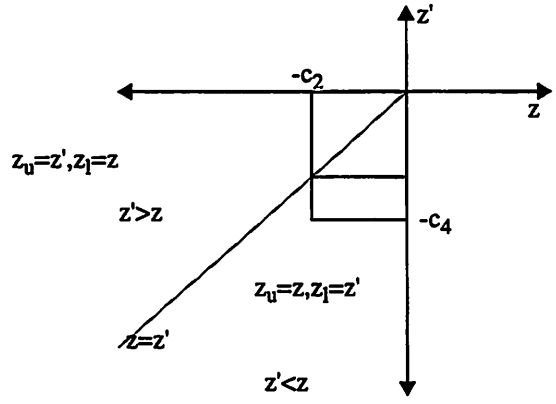


Figure A2b: Plane of Integration

We have

$$I_1 = \frac{\sinh(\gamma_{mn}c_4)\sinh(\gamma_{mn}c_2)}{\gamma_{mn}^2} \qquad (A24)$$

The second integral can be written as

$$I_2 = \int_{-c_2}^{0} \cosh(\gamma_{mn}z) \int_{-c_2}^{z} \sinh(\gamma_{mn}z')\,dz'dz + \int_{-c_2}^{0} \cosh(\gamma_{mn}z') \int_{-c_2}^{z'} \sinh(\gamma_{mn}z)\,dzdz'$$
$$\int_{-c_2}^{0} \sinh(\gamma_{mn}z)\,dz \int_{-c_4}^{-c_2} \cosh(\gamma_{mn}z')\,dz' \qquad (A25)$$

which yields

$$I_2 = \frac{c_s}{\gamma_{mn}} - \frac{\sinh(\gamma_{mn}c_s)\cosh(\gamma_{mn}c_g)}{\gamma_{mn}^2} \tag{A26}$$

where $c_g=\max(c_2, c_4)$ and $c_s=\min(c_2, c_4)$.

For the case when c_2 and c_4 are small, we have

$$I_1 = c_2c_4; I_2 = -\frac{\gamma_{mn}}{2}c_s c_g^2 \tag{A27}$$

The 'p_{ij}' term excluding the (m=n=0) case is thus given by

$$p_{ij} = \left\{ \sum_{m=0}^{\infty}\sum_{n=0}^{\infty} C_{mn}\frac{a^2b^2}{m^2n^2\pi^4} \times \left(f_{mn} - \frac{c_s c_g^2}{2ab\varepsilon_N c_2 c_4}\right)\left(\sin\left(m\pi\frac{a_2}{a}\right) - \sin\left(m\pi\frac{a_1}{a}\right)\right)\left(\sin\left(m\pi\frac{a_4}{a}\right) - \sin\left(m\pi\frac{a_3}{a}\right)\right) \right.$$
$$\left. \frac{\left(\sin\left(n\pi\frac{b_2}{b}\right) - \sin\left(n\pi\frac{b_1}{b}\right)\right)\left(\sin\left(n\pi\frac{b_4}{b}\right) - \sin\left(n\pi\frac{b_3}{b}\right)\right)}{(a_2-a_1)(b_2-b_1)(a_4-a_3)(b_4-b_3)} \right\} \tag{A28}$$

For $c_g=0$, the above equation reduces to (65), except for the (m=n=0) term.

In order to compute the (m=n=0) term we use (A19), that is $G_0 = (\beta_N z_1 + \Gamma_N)/(ab\varepsilon_N\beta_N)$.

The contribution to 'p_{ij}' from this term, denoted by p_{ij}^{00} is given by

$$p_{ij}^{00} = \frac{1}{c_2c_4}\int_{-c_4}^{0}\int_{-c_2}^{0} Z(z, z')\,dzdz' \tag{A29}$$

Using Fig. A2b, we get

$$p_{ij}^{00} = \frac{1}{c_2c_4ab\varepsilon_N\beta_N}\left\{ \int_{-c_s}^{0}\int_{-c_s}^{z} (\beta_N z' + \Gamma_N)\,dz'dz + \int_{-c_s}^{0}\int_{-c_s}^{z'} (\beta_N z + \Gamma_N)\,dzdz' + c_s\int_{-c_g}^{-c_s} (\beta_N z' + \Gamma_N)\,dz' \right\} \tag{A30}$$

which yields

$$p_{ij}^{00} = \frac{1}{c_2c_4ab\varepsilon_N\beta_N}\left(-\beta_N\left(\frac{c_s c_g^2}{2} + \frac{c_s^3}{6}\right) + \Gamma_N(c_2c_4)\right) \tag{A31}$$

For $c_g=0$, $p_{ij}^{00} = \Gamma_N/(ab\varepsilon_N\beta_N)$ as in (65).

Three-Dimensional contacts:

Eq. (65) shows the p_{ij}-coefficients for the case of two-dimensional contacts. There are several techniques to extend the technique to three-dimensional contacts as well.

The most accurate technique is to divide the Z-direction in the regions of the contacts into distinct layers. A contact would thus consist of several two-dimensional sheets of charge as shown in Fig. A2c below. The DCT for each unique interaction term for different z- and z'-values can be computed as shown for two contacts in Fig. A2d. The number of DCTs that need to be computed grows as n(n+1)/2 for n vertical layers. If the technique is kept general enough for optimization, then it becomes necessary to store all the DCT series completely. This limits the technique to about three vertical layers[14]. However if the generality

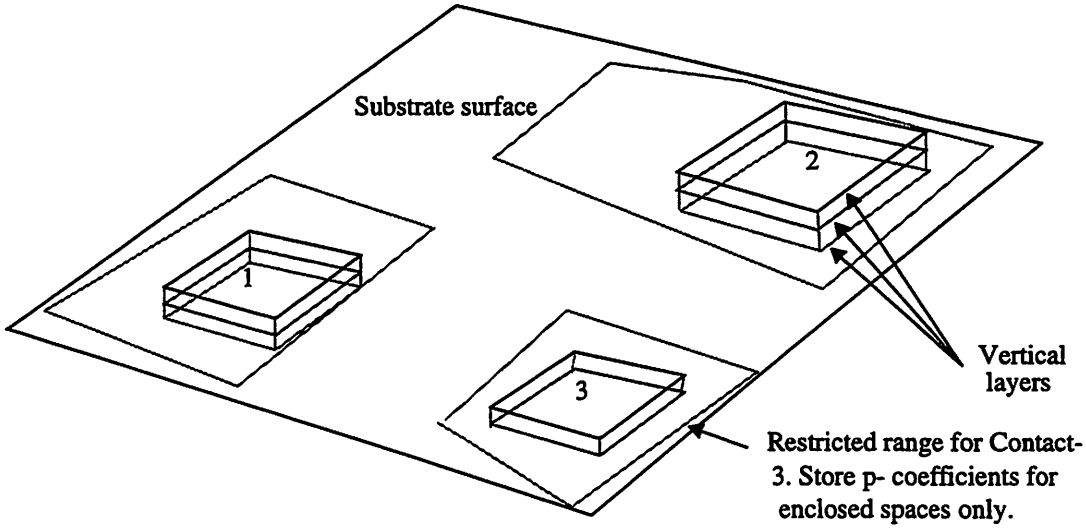


Figure A2c: Substrate Problem with Restricted Contact Movement

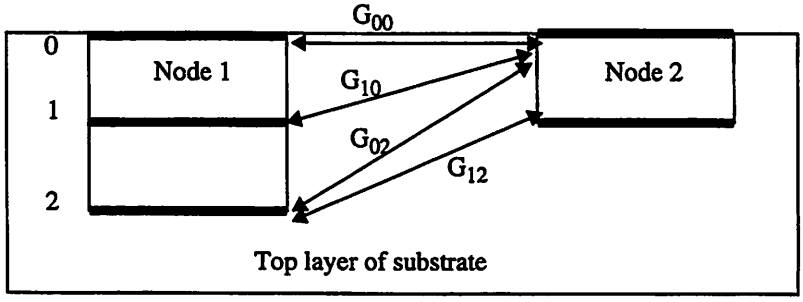


Figure A2d: Modeling of the Vertical Extension of a Contact

14. Three layers require storage for six DCT series.

of the technique is reduced somewhat, by restricting the field in which each contact can be moved, then only the corresponding terms of the DCT need be stored, and the technique can then be extended to a greater number of vertical layers (Fig. A2c). A problem with this technique is that the size of the [P] matrix can become very large. If each contact has n vertical subdivisions, then the size of the matrix increases by n^2.

A better technique for contacts with a small vertical extent is to use (A28). Here two DCT terms need to be calculated. The first series consists of terms of the form $\left(a^2b^2 \times f_{mn}\right) / \left(m^2n^2\pi^4\right)$ and the second one consists of terms of the form $\left(a^2b^2\right) / \left(2ab\varepsilon_N m^2n^2\pi^4\right)$. The individual p-coefficients are a sum of the surface DCT term as in (65) and a second term scaled by $-\left(c_s c_g^2\right) / (c_2 c_4)$. The second term represents the monopole contribution from the vertical charge distribution of the contact. In forming the [P] matrix using this algorithm, the depths of the contacts can be included by simply adding the second term to (65), with the appropriate scaling factor. The size of the matrix is unchanged. If all the contacts are of the same depth, then the monopole contribution can be computed exactly by computing the DCT of the complete expressions for I_1 and I_2 shown in (A24) and (A26).

Appendix A3.3: The Discrete Cosine Transform

The cosine transform of a two-dimensional series f_{mn} is defined as

$$D_{jk} = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f_{mn} \cos\left(\frac{\pi jm}{M}\right) \cos\left(\frac{\pi kn}{N}\right) \tag{A32}$$

This sum can be computed efficiently using several techniques as discussed in [A2]. An FFT based technique is discussed here.

The technique is based on reflecting the data f_{mn} about m=M and n=N such that

$$f_{m,\,(2N-n)} = f_{mn} \tag{A33a}$$

$$f_{(2M-m),\,n} = f_{mn} \tag{A33b}$$

$$f_{(2M-m),\,(2N-n)} = f_{mn} \tag{A33c}$$

Further by requiring that

$$f_{00} = f_{m0} = f_{0n} = f_{mN} = f_{Mn} \tag{A34}$$

Eq. (65) can be rewritten as

$$D_{jk} = \frac{1}{4}\left(\sum_{m=0}^{2M-1} \sum_{n=0}^{2N-1} f_{mn} \exp\left(i\frac{2\pi jm}{2M}\right) \exp\left(i\frac{2\pi kn}{2N}\right) \right) \tag{A35}$$

where $i = \sqrt{-1}$. The term in brackets can be recognized to be the $2M \times 2N$ two-dimensional FFT of the sequence f_{mn}.

More efficient techniques based on recursion can also be implemented. For an $M \times M$ order DCT computation this technique is $O(4M^2 \ln(4M^2))$. If f_{mn} are real data, the FFT can be reduced to order $2M^2 \ln(2M^2)$. A discussion of the two-dimensional FFT can be found in [A3].

References

[A1] 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 in Engineering Software, Vol. 20, No. 1, pp. 19-27, 1994.

[A2] K. R. Rao and P. Yip, *Discrete Cosine Transform: Algorithms, Advantages, Applications*, Boston: Academic Press, c1990.

[A3] A. V. Oppenheim and R. W. Schafer, *Discrete-Time Signal Processing*, Englewood Cliffs, N.J.: Prentice Hall, c1989.