ANALYSIS, DESIGN, AND OPTIMIZATION OF INTEGRATED HIGH-CURRENT DEVICES IN HIGH-SPEED AND RADIO-FREQUENCY SEMICONDUCTOR SYSTEMS

A DISSERTATION

SUBMITTED TO THE DEPARTMENT OF ELECTRICAL ENGINEERING

AND THE COMMITTEE ON GRADUATE STUDIES

OF STANFORD UNIVERSITY

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

Choshu Ito

May 2003

I certify that I have read this dissertation and that in my opinion it is fully adequate, in scope and quality, as dissertation for the degree of Doctor of Philosophy.

_____

Robert W. Dutton (Principal Advisor)

I certify that I have read this dissertation and that in my opinion it is fully adequate, in scope and quality, as dissertation for the degree of Doctor of Philosophy.

_____

Kenneth E. Goodson

I certify that I have read this dissertation and that in my opinion it is fully adequate, in scope and quality, as dissertation for the degree of Doctor of Philosophy.

_____

Umran S. Inan

Approved for the University Committee on Graduate Studies.

Abstract

The integration of devices that must either sink or source high currents pose design problems that are often not well understood. The analysis and design of such devices and their surrounding impedances, operating at high frequencies, is the focus of this thesis.

Integration of large electrostatic discharge (ESD) protection devices with radio frequency (RF)/high-speed circuits requires sufficient excess current handling capabilities such that large capacitances are present in the signal path of high frequency signals. These capacitances may cause undesirable signal reflections and inefficient power transfer between the pad and the core system. This work quantifies these undesirable effects, and through the use of Smith Charts and transmission matrices, demonstrates a methodology for optimized, distributed ESD protection design which may be integrated with RF/high-speed circuits.

Requirements for available bandwidth in RF power amplifiers necessitate operation at high frequencies. The power delivery requirements for communication with the base station further require operation in the large-signal regime. However, when identical power amplifier device unit cells are replicated and integrated for increased output power, the performance of the resultant system does not scale proportionally with increasing device channel width. Various measurement and simulation results for both electrical and thermal characteristics are obtained to identify the cause of the scaling performance limits. It is demonstrated that electromagnetic coupling between the unit cells, which may be described in terms of mutual inductance, is most likely the dominant

cause of this effect, and measures to cancel these coupling effects are presented. Of these measures, the most promising consists of a system with baluns to split the signal into in-phase and out-of-phase components, then alternating these signals in adjacent unit cells, thereby generating negative mutual inductance. Simulation results indicate that using this technique results in significant performance improvement.

Acknowledgments

I would like to acknowledge all the people that have made my experience at Stanford possible.

I would first like to express my gratitude to Professor Robert Dutton for his constant support throughout my studies at Stanford. His guidance, going back to my undergraduate days, has been invaluable. I would like to thank Professor Kenneth Goodson for his support with regard to the ESD project, and Professor Umran Inan for agreeing to be on my reading committee at the last minute. I would also like to thank Professor Nick Bambos for chairing my oral defense.

I would also like to acknowledge Hitachi Ltd., for their support throughout my graduate studies through the CIS FMA program. In particular, I would like to thank Dr. Katsuhiro Shimohigashi, Dr. Isao Yoshida, Dr. Masao Hotta, Dr. Koichi Seki, Mr. Toru Fujioka, and the members of the High-Frequency Group at Hitachi, for all the support and discussions regarding the power amplifier project.

Also, I am grateful to Dr. Kaustav Banerjee for his guidance regarding ESD protection systems.

I would like to thank the staff and students of the Dutton group, past and present, for all their assistance and friendship. Special thanks to Fely Barrera, Miho Nishi, Maria Perea, and Dan Yergeau, for all their help in making things run smoothly, and to my officemates, Edward, Jaejune, Xin Yi, Nathan, Yi-Chang, Jae Wook, Hai, Jung-Hoon, and Robert, for all the discussions about anything and everything.

I would like to acknowledge Dr. Richard Dasher, Carmen Miraflor, Maureen Rochford, and all the staff of the Center for Integrated Systems, who have been especially helpful in my maintaining a good relationship with Hitachi.

I would also like to express my gratitude to all my friends that have supported and encouraged me throughout the years.

Lastly, I would like to thank my parents, for making all this possible.

List of Tables

List of Figures

1 Introduction

1.1 Motivation

The drive for smaller, and thereby faster, integrated circuits (ICs) has defined the progress of the semiconductor industry for the last half-century. Much of the progress of the semiconductor industry has been driven by the desire for more speed. Devices that are designed to operate at higher frequencies allow processing of more data per second. Financially, the utility and novelty of the fastest device command a premium. In wireless communications, operation at higher (RF) frequencies is necessary for larger bandwidths and for avoiding the lower frequency bands that are already in use. However, designing high-speed systems has its problems, as parasitics exert a larger influence on performance. Measurement becomes difficult due to parasitics presented by the measurement instruments affecting the results. Electromagnetic effects such as skin effect and radiation, while negligible at lower frequencies, manifest themselves at high frequencies. High-frequency operation also consumes more dynamic power, which is a

problem in itself for mobile applications, but in addition, the larger power consumption is coupled with increased heat dissipation, giving rise to potential thermal problems.

Integrating high-current devices into these high-frequency systems exacerbates these issues. High-current devices tend to be physically large, and thus present small impedances. The small impedances of these high-current devices are troublesome, because parasitics have a greater effect, and efficient power transfer becomes more difficult as the impedance transformation ratio increases. Also, high-current devices can consume a large amount of power and generate a large amount of heat. The effect of heat on surrounding circuitry is becoming increasingly important as dimensions shrink further.

In looking at the system design process, while the design processes for the digital logic and analog cores have matured, the design of high-current systems still utilizes trial-and-error. Analyzing the operation of these large devices is difficult, as they tend to operate in the non-linear, large-signal regime. This prevents the simple, linearized equations used in small-signal analysis from yielding meaningful results. Circuit models are generally inadequate for high-frequency, large-signal simulations—commonly available models may only be used for low-frequency large-signal, or high-frequency small-signal analyses. Device simulators that can simultaneously analyze the electrical, thermal, and electromagnetic characteristics do not exist, and even if they did, the amount of computation required would be prohibitive for these large devices in question.

Given these issues that exist in designing high-frequency systems that require high-current devices, this work examines the design of two of these systems. The first is the ESD protection circuit, which must conduct a large amount of current in a very short time to prevent the core circuitry from being damaged. The second is the RF power

amplifier for use in wireless communications. The power amplifier must be able to drive enough current onto an antenna such that the radiated signal may be received remotely.

1.2 Organization

Chapter 2 discusses the analysis and design of electrostatic discharge (ESD) protection systems for high-speed mixed-signal and RF systems. ESD is one of the most serious concerns in IC manufacturing. In fact, it is known to be the most common cause of all IC failures [1]. ESD is the transient discharge of static charge arising from human handling or contact with machines or other charged structures. The discharge typically occurs in a very short duration on the order of 10 ns to 100 ns, with currents ranging from 1 A to 10 A. Traditionally, protection circuits are designed and employed near the I/O pins to alleviate the impact of ESD events and to improve manufacturing yields. However, as the demand for wireless (RF) and high-speed mixed-signal systems continues to increase rapidly, integrating sufficient ESD protection with these systems poses a major design and reliability challenge. This is due to the fact that in applying ESD protection to these systems, the protection system must be transparent—the protection circuit must not affect desired signal propagation under normal operating conditions [2, 3]. ESD protection systems, which are most often found adjacent to I/O pins, present large capacitances under normal (non-ESD) conditions. In high-speed systems, large capacitances are anathema, as the signals short to ground. While Kleveland showed in [4] that a distributed protection system offers sufficient ESD protection, its effect on the high-speed signal under normal operation was not addressed. This chapter examines the effects of integrating ESD protection systems with high-speed

systems—in particular, the effects of integration on the high-speed signals under normal operation. Furthermore, a methodology by which ESD protection may be designed around high-speed systems is demonstrated.

Chapter 3 characterizes the RF power amplifier for wireless communications. Wireless communications, and the RF power amplifier in particular, are of interest because in recent years, the demand for wireless communications systems has been experiencing explosive growth. In the wireless voice space (cellular phones), an estimated one billion people subscribed to mobile phone service as of 2001, with a subscriber growth rate of 20% per annum [5]. In the wireless data space, the adoption of the IEEE 802.11 standard (Wi-Fi, Bluetooth) in 1997 was followed by mass production of computer peripherals that conform to this standard. Equipment compatible with the most common Wi-Fi standard (802.11b) started appearing on the market in 1999, and by 2002, the market for 802.11b equipment was expected to grow to $884 million [6]. This proliferation of wireless communications systems worldwide has increased the demand for better performance in these systems. Within the wireless system, the power amplifier subsystem is a very power-hungry circuit, as it needs to drive the modulated signal onto the transmission antenna with enough power that the receiving system can extract it out of the background noise. Because the power amplifier consumes much of the system power budget, improving its efficiency should significantly improve the overall system efficiency. However, measurements have shown that as more device fingers are placed on the same die to increase the available output power, the maximum output power does not increase proportionally against total device channel width. The ability to cure this phenomenon should yield a significant cut in the power consumption, allowing better

performance and longer battery life. This chapter presents the basic measurement results of an RF power amplifier, and hypotheses regarding the cause of this phenomenon.

Chapter 4 examines in detail the effects of mutual inductance as a possible cause for the performance degradation introduced in Chapter 3. Analysis of the power amplifier system using simulators and measurements demonstrate mutual inductance to be the dominant cause of lower maximum output power density with increased device width. Measures to counter this trend are also introduced and discussed.

Chapter 5 concludes the dissertation, suggesting possible future avenues of research in the topics discussed.

1.3 Contributions

The contributions of this work must be placed in context of the issues presented above. ESD protection for RF systems is a nascent field, where issues arising from the overlap of the two disciplines must be addressed, and this work provides insight for ESD protection designers into how to apply their knowledge to RF systems. With respect to the RF power amplifiers, progress in this area has mainly come through miniaturization and adoption of novel semiconductor material. This work focuses on neither of these, but rather on an area that has been overlooked. The insight that this work provides into the effects of device width on the performance of RF power amplifiers should prove valuable regardless of the semiconductor technology in use. Thus the contributions of this work may be summarized as follows:

1) Quantification of the impact of ESD protection systems on the performance of RF systems under normal operation.

2) Design methodologies for broadband, high-frequency, distributed ESD protection using Smith Charts or transmission matrices to minimize the effects of ESD protection on normal RF operation.

3) Quantification of changes in RF power amplifier characteristics due to changes in device channel width.

4) Identification of mutual inductance as the cause of observed performance degradation in RF power amplifiers with increased device channel width.

5) Novel RF power amplifier circuit topology using baluns to cancel the width effect.

2 RF ESD Protection System

As shown in Figure 2.0.1, the on-chip electrostatic discharge (ESD) protection circuit is placed between the signal pin and the core circuit. The protection circuit may be composed of various devices, such as diodes, transistors, or silicon controlled

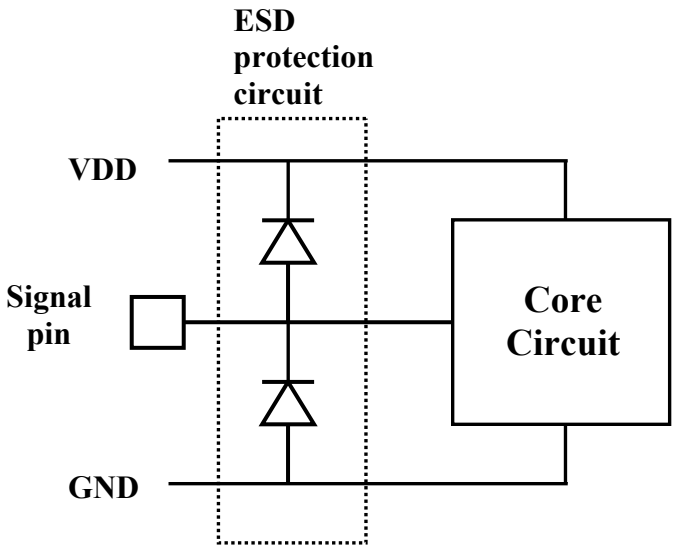


Figure 2.0.1. A simple ESD protection system example. The diodes shunt excess current applied to the signal pin towards VDD or GND to protect the core circuit.

rectifiers (SCRs), but in all cases, these devices shunt ESD current coming from the signal pin to the power/ground supply rails and away from the core circuits [1]. However, under normal operating conditions, these protection devices present capacitances and resistances to the signal path, and at sufficiently high frequencies, the capacitances look like short circuits to ground. Thus a poorly designed protection system can generate impedance mismatches, causing reflections of signals, corruption of signal integrity, and inefficient power transfer between the signal pin and the core circuit. In addition to these problems, both the incident and reflected signals can interfere with signals on adjacent wires through crosstalk. Also, while the operating frequency continues to rise, the size of the protection circuits and their associated capacitances are not decreasing as rapidly, resulting in increasingly inefficient power transfer. While narrowband RF operation with ESD protection is feasible by use of simple matching networks to tune out the ESD parasitic capacitance, broadband RF system protection poses a greater challenge [7, 8]. The simple approach of minimizing capacitance while maintaining high protection levels is becoming increasingly infeasible as the operating frequency rises beyond a few GHz [9]; alternate protection schemes such as the distributed transmission line ESD protection system may be necessary [4, 10]. Recent work has focused on comparison of ESD protection strategies for RF applications at 2 GHz [11]. However there is little published information that provides performance analysis of RF circuits with various ESD protection design options, particularly for the distributed protection scheme, which is attractive for operation in the multi-GHz regime. In this chapter, two design methodologies to quantify the impact of the parasitic capacitance and resistance associated with various distributed ESD protection circuit

designs and to optimize the number and length of coplanar waveguides (CPW) (used as transmission lines [12]) separating the distributed ESD elements, are shown [13-15]. Also, it is demonstrated that a 4-stage distributed ESD protection with coplanar waveguides can be employed to provide excellent RF performance for frequencies as high as 10 GHz.

2.1 Distributed ESD Protection System

As the operation frequency of systems increases, the parasitics associated with ESD protection systems become more significant in limiting bandwidth and power transfer. One method of minimizing the effects of these parasitics is to utilize a distributed protection scheme. A distributed ESD protection system is a modification of the distributed amplifier proposed by Ginzton [16], where ESD protection devices replace the amplifier elements. As proposed in [4], by incorporating the parasitic capacitance of the device into a discrete transmission line structure, the loading of the system by the ESD devices may be prevented.

For a constant total capacitance C, increasing the number of sections of the artificial transmission line increases the bandwidth of the transmission line. In [17], the iterative structure shown in Figure 2.1.1 is analyzed. In the π-section shown in Figure 2.1.1a, the capacitance C models the ESD protection device. For the input impedance Z_i to equal Z_o, which is the load impedance of the system and also the characteristic impedance of the transmission line, the inductance L must be chosen as shown in equation (2.1), where ω_o is the operating frequency.

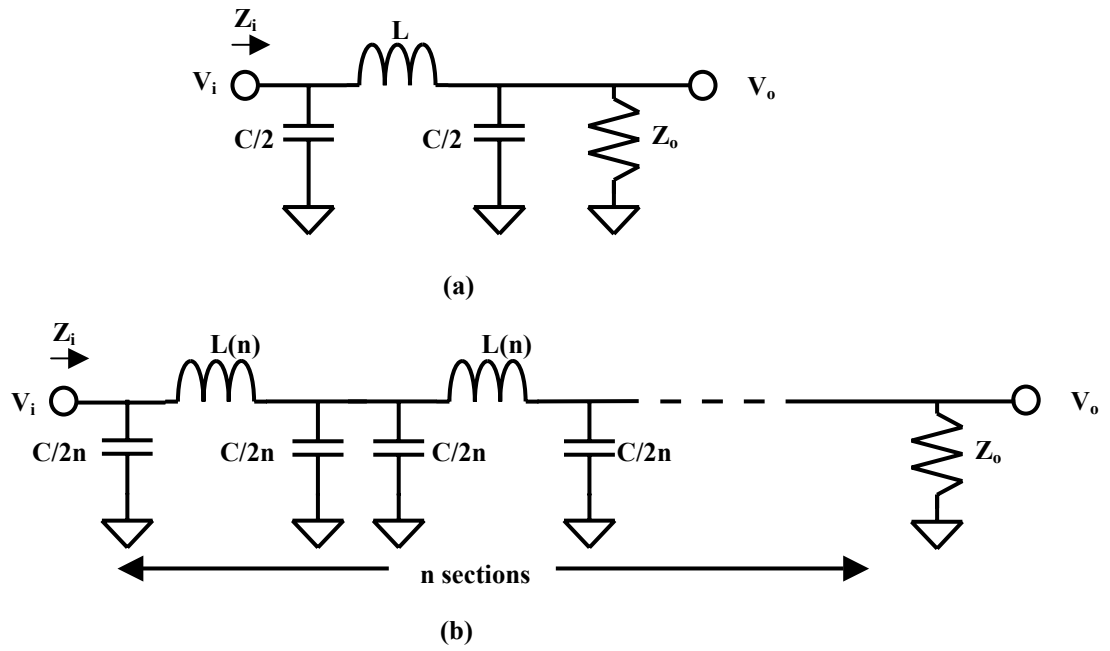


Figure 2.1.1. (a) LC single π-section forming an artificial transmission line. L is chosen such that Z_i equals Z_o at the frequency of interest. (b) n-section transmission line, with the same total capacitance as in (a). L is a non-linear function of *n* in this case.

$$L = \frac{4Z_o^2 C}{4 + (\omega_o Z_o C)^2} \tag{2.1}$$

In this case, the voltage gain is as shown in equation (2.2).

$$A_V = \frac{V_o}{V_i} = \left(1 - \frac{\omega^2 LC}{2} + j\omega\sqrt{LC}\sqrt{1 - \frac{\omega^2 LC}{4}}\right)^{-1} = \left(\frac{j\omega\sqrt{LC}}{2} + \sqrt{1 - \frac{\omega^2 LC}{4}}\right)^{-2} = \left(\sqrt{1 - \frac{\omega^2 LC}{4}} - \frac{j\omega\sqrt{LC}}{2}\right)^2 \tag{2.2}$$

The cutoff frequency for this system is ω_c such that the discriminant $(1-\omega^2 LC/4)$ equals zero, and is as shown in equation (2.3).

$$\omega_c = \frac{2}{\sqrt{LC}} \tag{2.3}$$

If there are *n* cascaded π-sections as shown in Figure 2.1.1b, the overall voltage gain would be as shown in equation (2.4), where L is a function of *n* as shown in equation (2.5).

$$A_v = \frac{V_o}{V_i} = \left(\sqrt{1 - \frac{\omega^2 L(n) C}{4n}} - \frac{j\omega \sqrt{L(n)\frac{C}{n}}}{2} \right)^{2n} \tag{2.4}$$

$$L(n) = \frac{4nCZ_o^2}{4n^2 + (\omega_o C Z_o)^2} \tag{2.5}$$

Substituting equation (2.5) into the discriminant from equation (2.4) and equating that to zero yields the cutoff frequency shown in equation (2.6).

$$\omega_c = \sqrt{\frac{4n^2 + \omega_o^2 Z_o^2 C^2}{Z_o^2 C^2}} \tag{2.6}$$

As can be seen, with increasing *n*, the cutoff frequency of the transmission line increases. Thus for a known capacitance, as determined from the ESD requirement, dividing the capacitive loading into sections allows operation at a higher frequency under normal conditions.

It is instructive to note that with increasing number of sections, the capacitance associated with each section, (C/2n), decreases, which will require smaller ESD protection devices between the transmission line sections. As the size of these devices decreases, there is some concern regarding their effectiveness. However, in [18], it was shown that smaller NMOS ESD protection devices are more robust since they conduct current more uniformly, while as the protection devices become large, current flows through only a small section of the device, giving rise to smaller failure currents per unit transistor width. Additionally, in [4], ESD measurements of a distributed protection system was performed, whereby it was shown that each smaller element of the distributed protection scheme does indeed turn on under human body model (HBM) and charged device model (CDM) stress. Although the above mentioned results lend strong support to

the distributed ESD protection scheme, the impact of this protection scheme on the RF signal under normal operation has not been quantified. The following sections show the effect of such ESD protection on high-speed signals and propose a methodology for designing optimal distributed ESD protection systems to minimize their effects on the integrity of high-speed signals.

2.2 S-Parameter Performance Analysis

In RF systems, the s-parameter matrix is often used to represent the characteristics of the network. As the system operating frequency increases such that the wavelength becomes comparable to the device dimensions, the wave-nature of signal propagation cannot be ignored. Also, generating pure open and short loads to calibrate the network measurement equipment becomes increasingly difficult at higher frequencies, as small parasitics greatly affect the impedances. The s-parameter matrix, which consists of ratios of outgoing to incoming signals measured with resistive terminations at each network port, satisfies the requirements for accurate high-frequency characterization [19]. Most often, RF networks are standardized to 50 Ω input and output impedances for maximum power transfer. S-parameter measurements are then taken as the reflection at, and transmission between, ports with 50 Ω termination at each port.

Figure 2.2.1a shows the measurements required for a general 2-port system. When a voltage is applied from a 50 Ω source (V_{src}), the ratio of the reflected signal to the incident signal is the input reflection s-parameter, s_{11}, which is also equal to the reflection coefficient Γ as shown in equation (2.7).

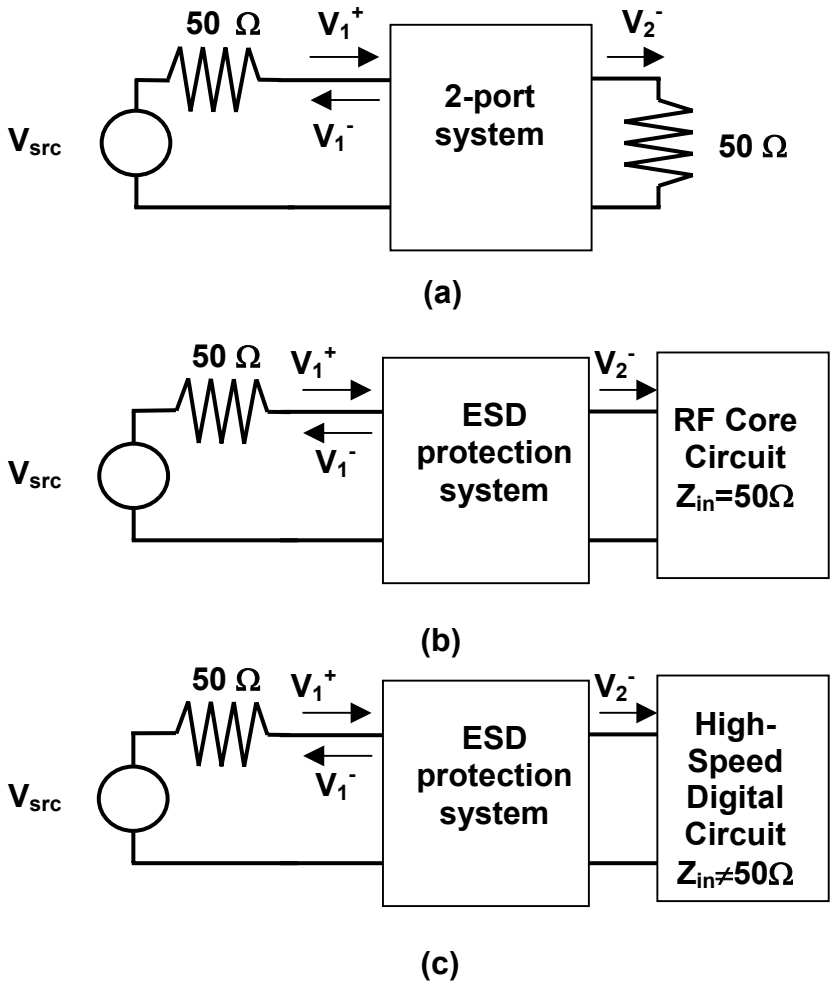


Figure 2.2.1. (a) General 50 Ω 2-port system showing the incident, reflected and transmitted voltages. The 50 Ω voltage source (V_{src}) drives the system, and after passing through the 2-port system, power is transferred to the 50 Ω load at the right. (b) 2-port model of an RF system with an ESD protection system. Note the analogy between (a) and (b). (c) With a high-speed digital circuit as the load, the impedance of the load is no longer at 50 Ω, and the situation differs from that of (a) and (b).

$$s_{11} = \frac{V_1^-}{V_1^+} = \Gamma \tag{2.7}$$

$$s_{21} = \frac{V_2^-}{V_1^+} \tag{2.8}$$

The forward transmission s-parameter, s_{21}, is the ratio of the outgoing signal at port 2 to the incident signal at port 1 as shown in equation (2.8). These two parameters characterize how much signal is reflected by the system, and how much signal is

transmitted to the 50 Ω load. Often, s_{11} is plotted on a Smith Chart, which is a diagram that allows representation of reflections and impedances [20]. Looking at the Smith Chart as a polar plot corresponds to looking at reflections, while the various arcs of the Smith Chart represent lines of constant resistance and reactance. Thus the magnitude of impedance mismatch and signal reflection may both be gleaned from Smith Chart plots.

A generalized RF system with an ESD protection network is shown in Figure 2.2.1b. S-parameters may be used to quantify the impact of the ESD protection system on the signal transmission between the input and the core circuit. The s_{11} of the ESD protection system shows how much signal is reflected back to the input, and s_{21} shows how much of the signal applied at the input reaches the core circuit. Thus the goal of the ESD protection circuit is to protect the core circuit while minimizing s_{11} and maximizing s_{21} of the system.

Since both the input and output impedances are standardized to 50 Ω in RF circuits, an analogous situation exists for ESD protection at the RF outputs. However, digital systems generally do not have standardized input and output impedances as shown in Figure 2.2.1c, therefore modifications are necessary in analyzing high-speed digital circuits, and this will be discussed in Section 2.4.

Starting with a standard 50 Ω system as is commonly found in RF systems, four different implementations of ESD protection are investigated, as shown in Figure 2.2.2. A 50 Ω signal source drives the input to the protection circuit, and the output of the protection circuit is connected to the system to be protected, as modeled by a 50 Ω load (R_{load}). In each circuit, the protection device is modeled as a capacitance and input resistance, and interconnects between the pin and ESD circuit or between distributed

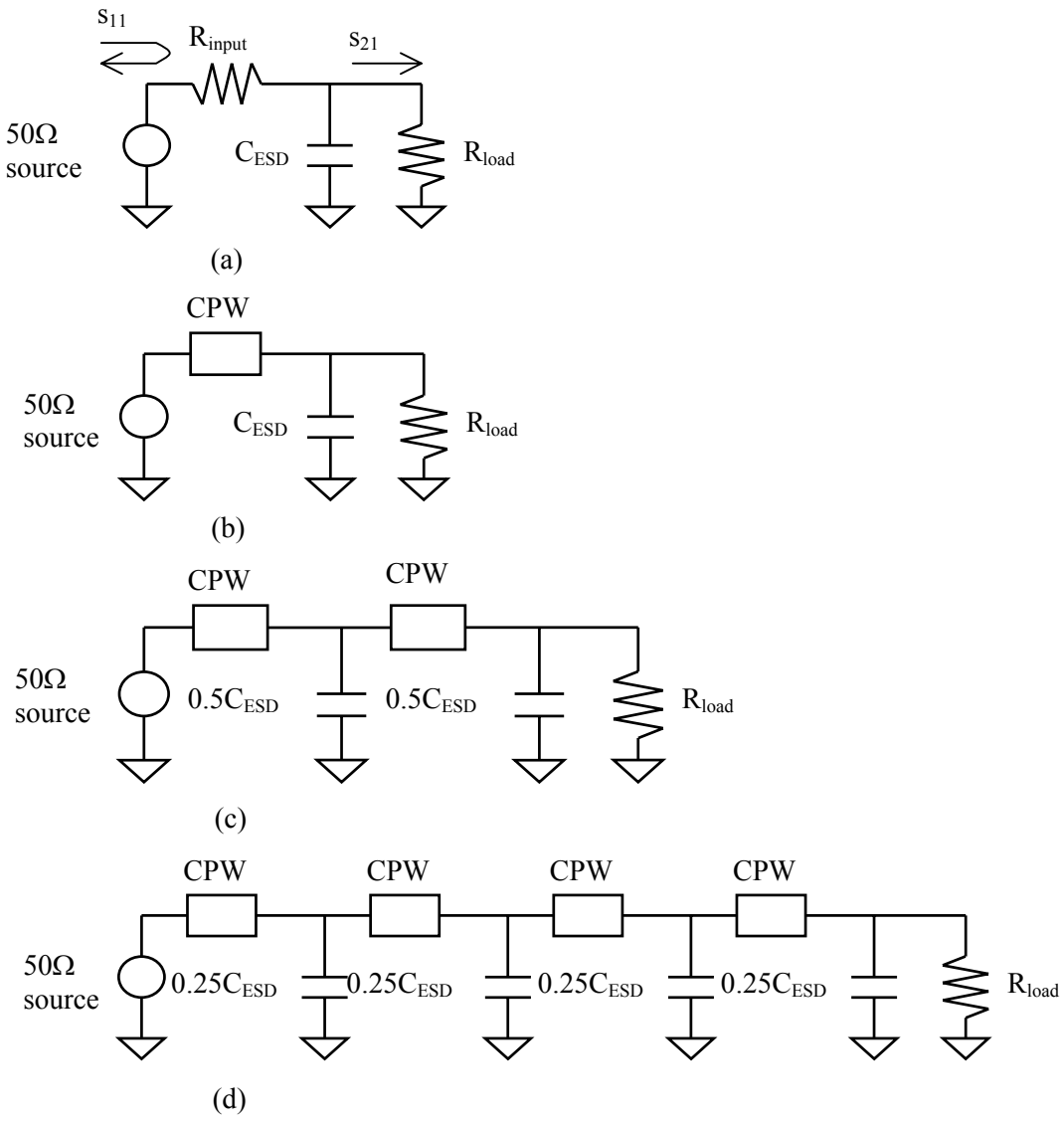


Figure 2.2.2. (a) Equivalent circuit with a general ESD protection device between the source and the load. (b) A length of CPW is added to the previous circuit, to improve impedance matching. (c) Two-section transmission line structure formed by CPW and the protection devices. (d) Four-section transmission line structure.

ESD elements are modeled by a resistance or a coplanar waveguide (CPW). Initially, the capacitance is assumed to be 200 fF, a value sufficient to provide a 2 kV ESD protection level [21]. Figure 2.2.2a, which represents the most general ESD protection, consists of the source, load, a resistor representing interconnect and device loss (R_{input}), and the

protection device and pad parasitic capacitance (C_{ESD}). Figure 2.2.2b introduces a CPW between the source and the protection device to provide a better impedance match. The required CPW length is calculated using Smith Charts and impedance transformations to minimize reflections. This methodology is discussed in detail in Section 2.5. Figures 2.2.2c and 2.2.2d show series circuits with smaller sections for a better broadband match as the circuit approaches an ideal transmission line for infinitesimally small sections. For the purpose of this study, transmission lines with a maximum of four sections were examined. Although more sections may yield better performance, any further gain would be marginal. Also, the added complexity of having more sections may be undesirable, and depending on the layout topology, it may be unreasonable to further divide the ESD device and pad capacitances into smaller elements. S-parameter simulations over the frequency range 0-10 GHz were performed on these circuits using the microwave circuit simulator ADS [22], to generate the reflection parameter s_{11}, and the transmission parameter s_{21}. As shown in Figure 2.2.2a, s_{11} corresponds to the amount of signal that is reflected at the input, and s_{21} corresponds to the amount of input signal that reaches the load. The objective of the system designer is to minimize s_{11} and maximize s_{21}. This study uses $|s_{11}|^2$ and $|s_{21}|^2$ as performance metrics since these coefficients are then directly proportional to power.

Next, the simulations are repeated for a set of capacitances, thus modeling different protection devices with different protection levels. The data from these simulations can be used to provide designers with insight into how much complexity is required in the protection system to obtain the desired ESD protection along with sufficient high-speed performance at the operating frequency of interest.

2.3 Results & Discussion

Figure 2.3.1 shows the results from the simulation of the simplest case from Figure 2.2.2a, with the input resistance set to zero and C_{ESD}=200 fF. Since the whole system is lossless, all the power loss is due to signal reflection caused by impedance mismatch. While most of the power reaches the load at low frequencies, the capacitance loads the circuit at higher frequencies. Using a larger protection device in the same circuit gives s_{11} like that in Figure 2.3.2. On this Smith Chart, it can be observed that the magnitude of the reflection increases with increased capacitance, thus less power is delivered to the load. Also, with increasing input resistance, the loss increases, as shown in Figure 2.3.3.

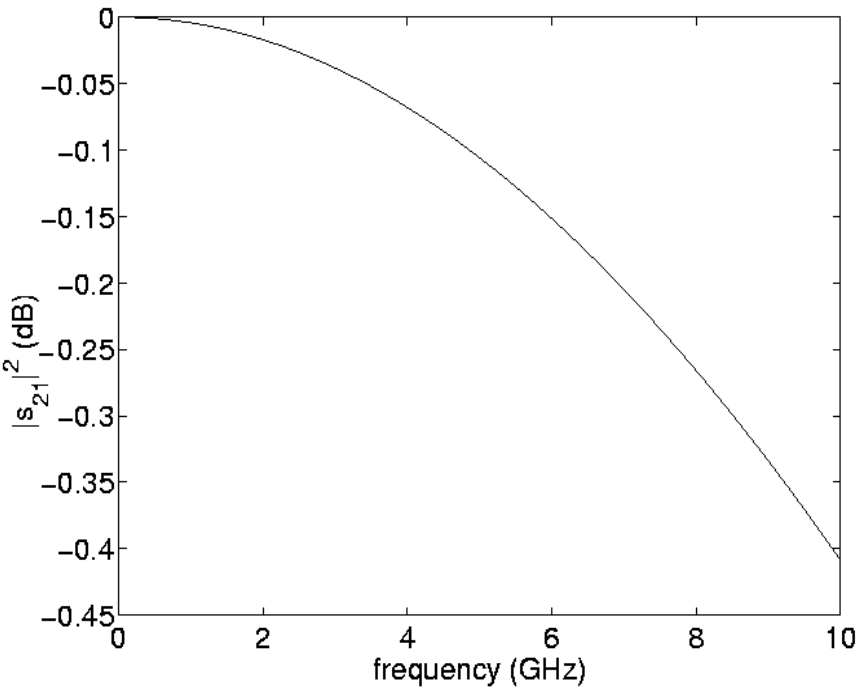


Figure 2.3.1. Plot of signal power loss vs. frequency for the circuit in Figure 2.2.2a, with an ESD device capacitance of 200 fF and R_{input} of zero.

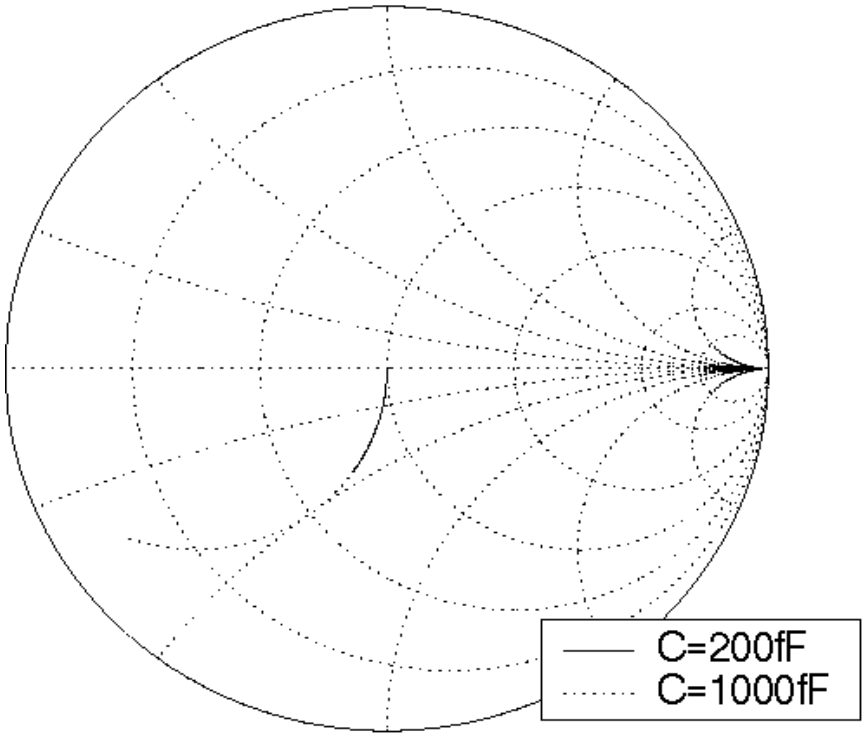


Figure 2.3.2. Smith Chart representation of the effect of C_{ESD} on s_{11} for Figure 2.2.2a, with zero series resistance (R_{input}=0), at f=10 GHz.

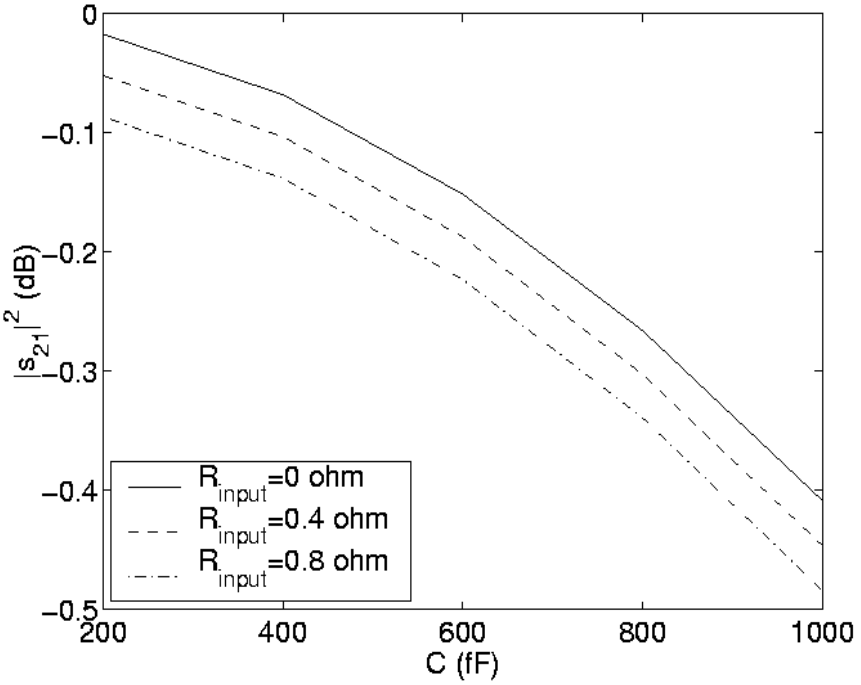


Figure 2.3.3. Performance vs. capacitance with varying R_{input} for Figure 2.2.2a at f=2GHz. Note the increase in loss with increased resistance.

The effect of employing CPWs in the ESD protection is also examined. Figure 2.3.4 corresponds to the circuit in Figure 2.2.2b, where a CPW with a characteristic impedance of 100 Ω as determined by appropriate physical parameters of the transmission line such as width, thickness, spacing, dielectric constant, and loss tangent, is added to provide some impedance match. In the case of Figure 2.3.4, the loss is due to the CPW loss and the mismatch loss. With a CPW length of 1.3 mm, there is a 0.25 dB loss even at low frequencies. As shown in Figure 2.3.5, this CPW has a loss of 0.18 dB/mm at low frequencies, with the loss becoming worse with increased frequency, to 0.39 dB/mm at 10 GHz. At higher frequencies, this CPW loss worsens, while the mismatch loss also becomes larger. Comparatively, this result is worse than that of Figure 2.3.1, but since Figure 2.3.1 shows an ideal case where there is no resistive loss at all, this is to be expected. Note that if the CPW were lossless, the results with the CPW

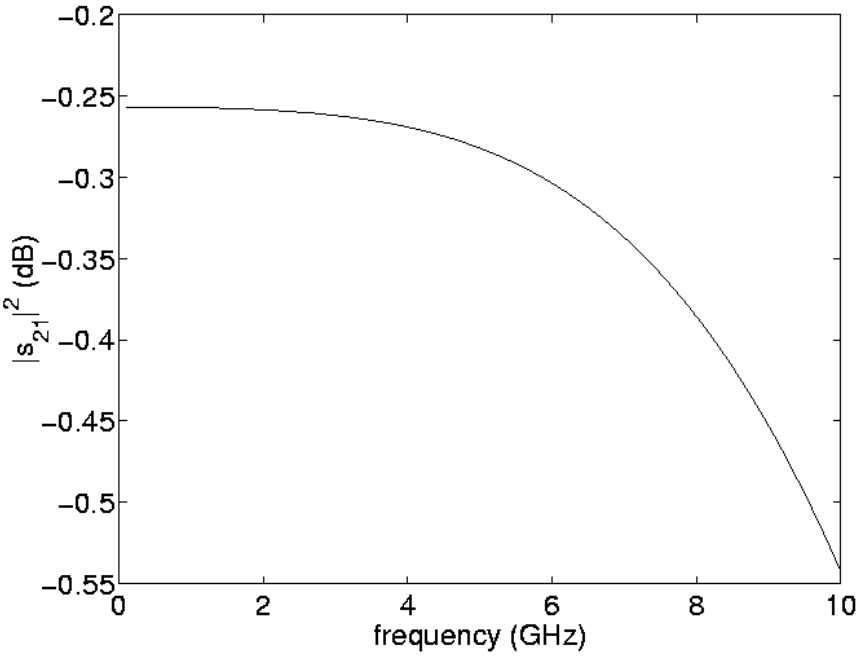


Figure 2.3.4. Plot of signal power loss vs. frequency for Figure 2.2.2b, with C_{ESD}=200 fF.

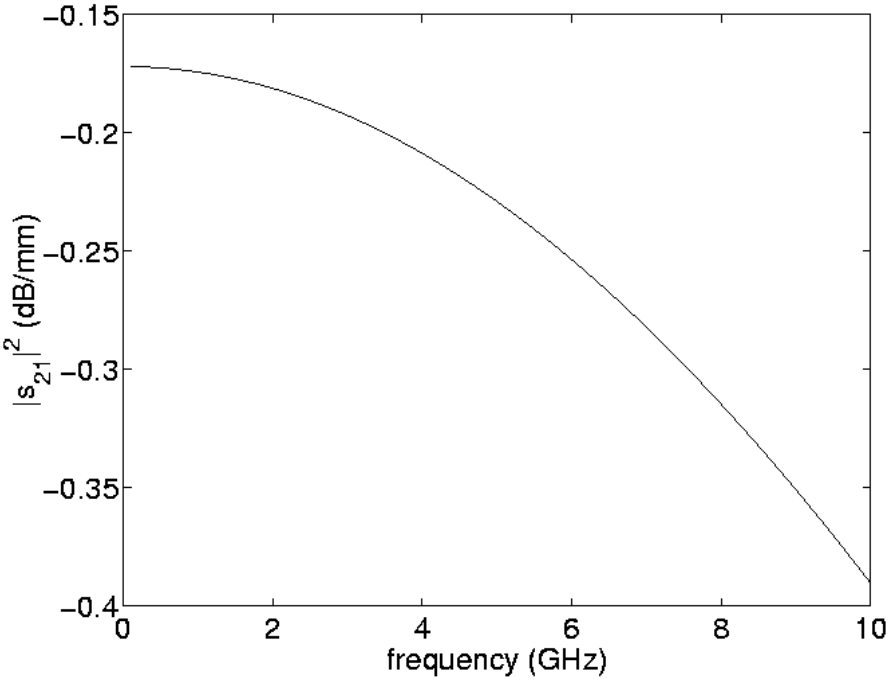


Figure 2.3.5. CPW loss per mm vs. frequency, showing the inherent loss found in the CPW due to resistive and dielectric loss.

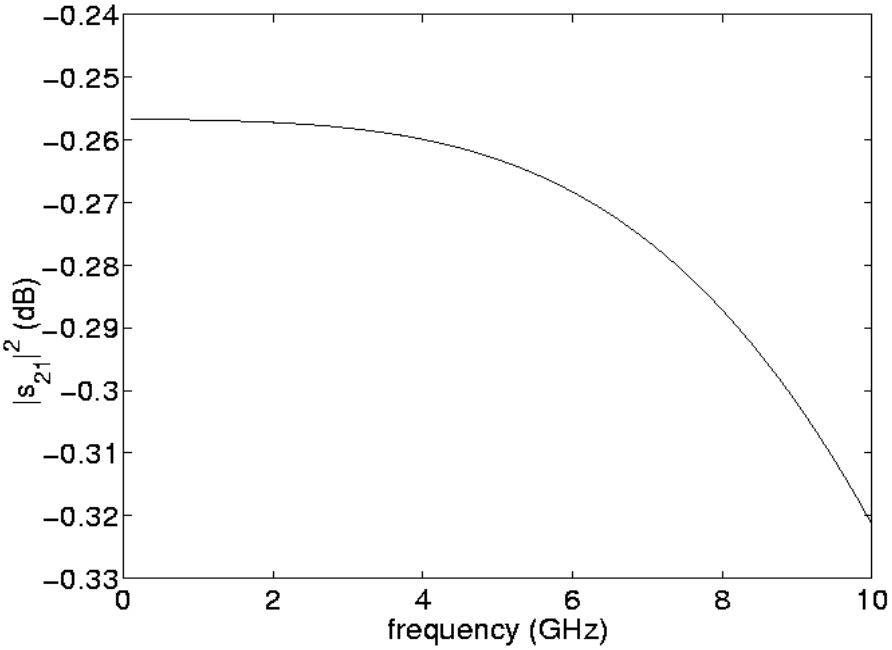


Figure 2.3.6. Plot of signal power loss vs. frequency for Figure 2.2.2c, with C_{ESD}=200 fF. Note the better loss characteristics compared to Figure 2.3.4.

would be better than that of Figure 2.3.1.

Figures 2.3.6 and 2.3.7 show the results from simulating the circuit in Figures 2.2.2c and 2.2.2d, respectively. The losses observed are less than those seen in both Figure 2.3.1 and Figure 2.3.4 at higher frequencies. At low frequencies, the CPW losses are again observed. Figure 2.3.7 shows loss characteristics that decrease by less than 0.02 dB between 0 and 10 GHz, with the maximum loss of 0.273 dB at 10 GHz, thus demonstrating good broadband characteristics.

Figure 2.3.8 shows the reflection parameter s_{11} corresponding to Figures 2.2.2b, 2.2.2c, and 2.2.2d. Note that the observed reflection becomes smaller with increased number of CPW sections.

The impact of the protection level is analyzed next by varying the size of the protection device (C_{ESD}). Figures 2.3.9-2.3.12 plot the power loss as represented by $|s_{21}|^2$

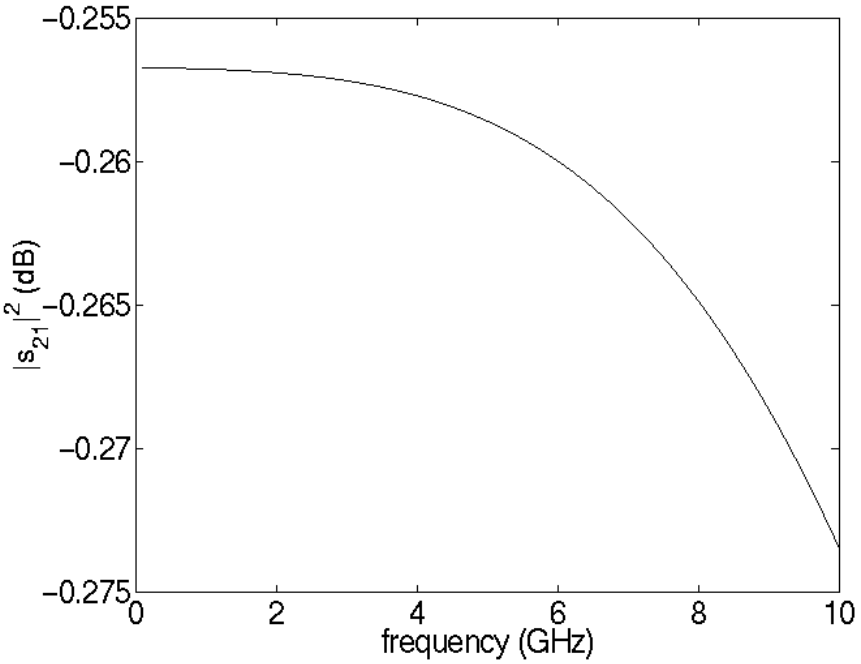


Figure 2.3.7. Plot of signal power loss vs. frequency for Figure 2.2.2d, with C_{ESD}=200 fF. The loss is now less than 0.275 dB at 10GHz.

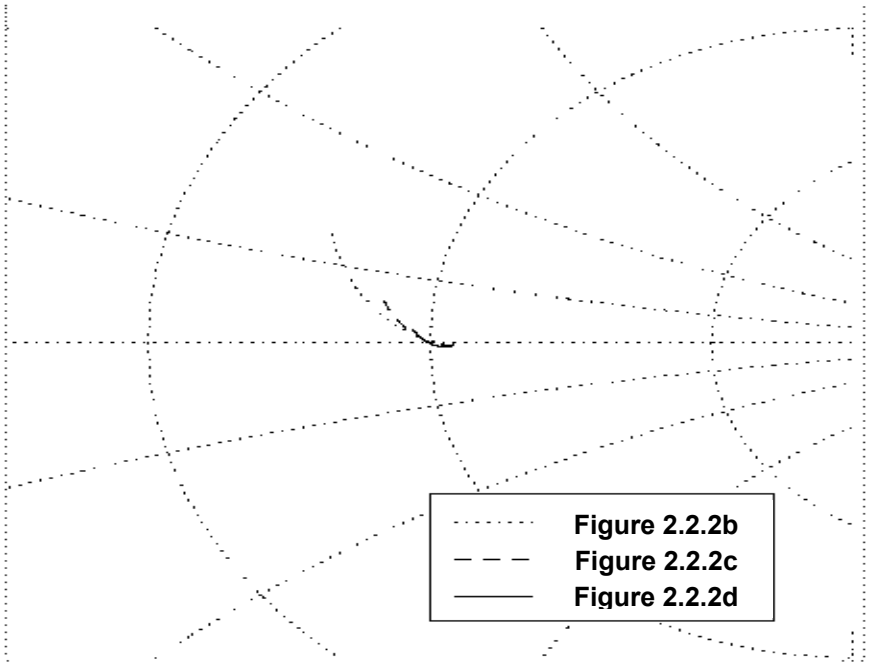


Figure 2.3.8. Smith Chart representation of s_{11} for Figures 2.2.2b, 2.2.2c, and 2.2.2d, over the frequency range 0-10 GHz, with C_{ESD}=200 fF. The reflection decreases with increasing number of sections.

in the protection systems in Figure 2.2.2, against the parasitic capacitance posed by the protection devices. It is clear that higher frequencies and larger capacitances generate larger losses. The ideal case (R_{input}=0) in Figure 2.3.9 shows that at low frequencies, the loss is minimal irrespective of the device capacitance. The single CPW case of Figure 2.3.10 shows resistive loss in the CPW, resulting in poor performance at all frequencies. However, a more distributed protection system can minimize the loss for a wider frequency range, as shown in Figures 2.3.11 and 2.3.12. Also, Figure 2.3.12, which represents the four-segment distributed protection, shows that the loss will only vary by 0.85 dB for all capacitances, thus showing immunity to variations in ESD device depletion capacitance due to changes in DC bias levels.

The above results indeed show that dividing the ESD protection device into a few smaller sections provides better broadband RF performance. But for high-speed digital

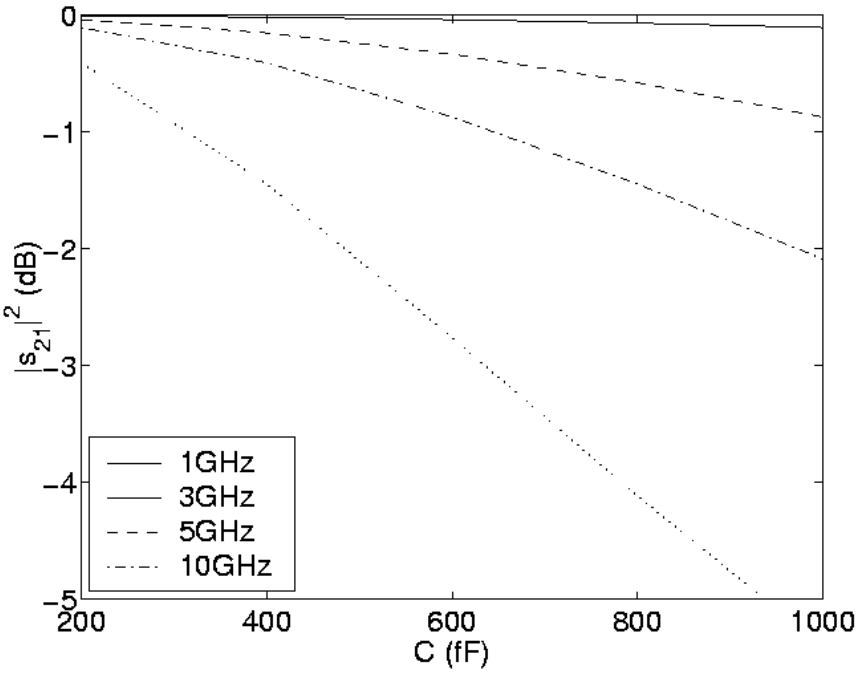


Figure 2.3.9. Signal power loss vs. parasitic capacitance, for a set of frequencies for the circuit in Figure 2.2.2a.

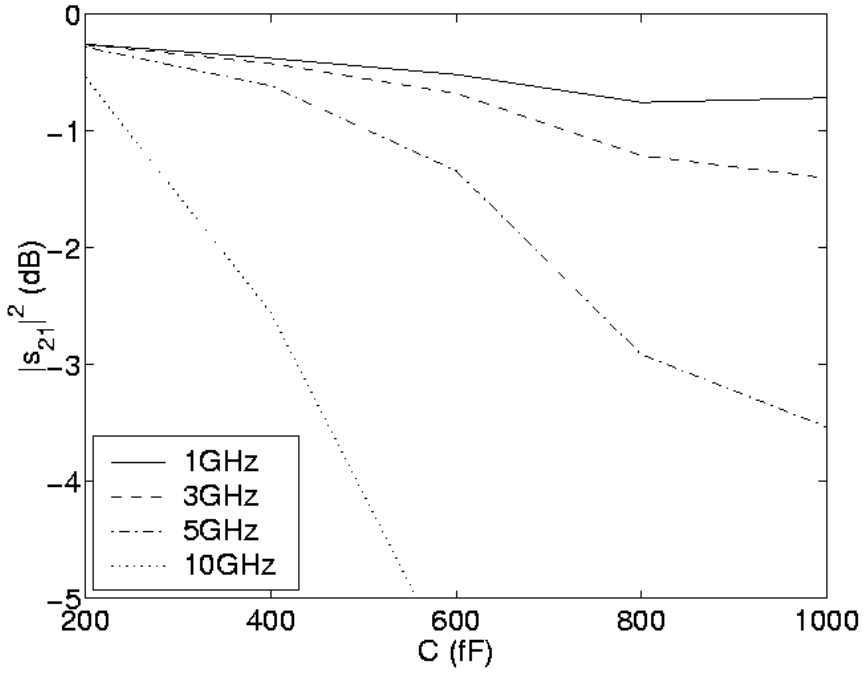


Figure 2.3.10. Performance vs. capacitance plot for the circuit in Figure 2.2.2b.

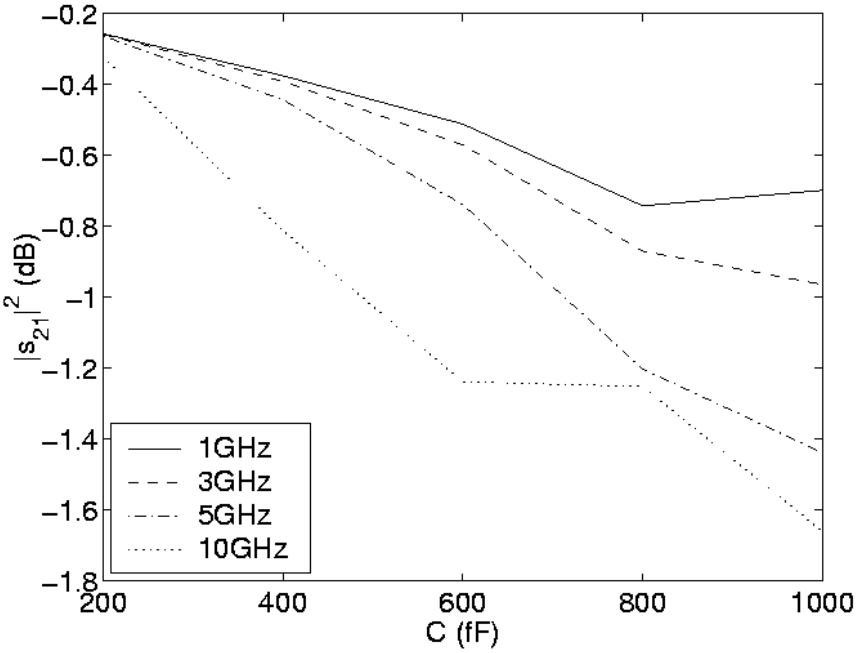


Figure 2.3.11. Performance vs. capacitance plot for the circuit in Figure 2.2.2c. Note that the curves start to converge for all frequencies.

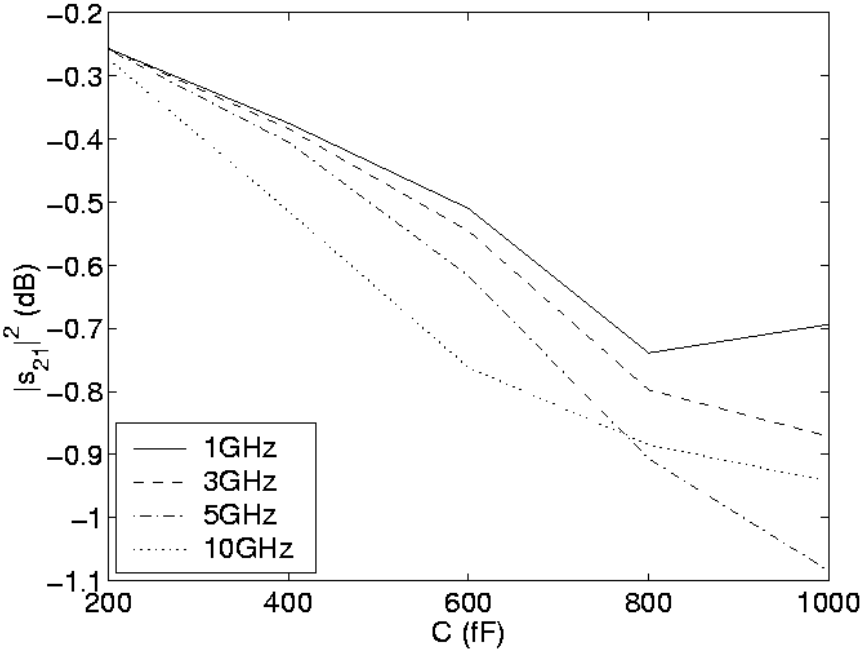


Figure 2.3.12. Performance vs. capacitance plot for the circuit in Figure 2.2.2d. Note that the variation of the loss over all measured frequencies and capacitances is only 0.85 dB.

systems, there are differences in the circuit characteristics that prevent the use of s-parameter analysis. These differences are addressed, and a solution towards designing effective distributed ESD systems, even in a high-speed digital environment, is proposed in the next section.

2.4 Modifications for Applications to High-Speed Digital Circuits

In the case of high-speed digital circuits, the impedance termination conditions are not as well defined as in the RF case. Input buffers present a capacitive load due to the gate capacitance of the devices, and output buffers present a capacitance due to the drain junction capacitance in parallel with the on-resistance (R_{on}) of the conducting device. These terminal impedances vary depending on the buffer sizes and voltage levels, and therefore transmission and reflection from a transmission line would also vary.

In order to facilitate the design of the distributed ESD protection system, we propose terminating these digital terminals resistively, as suggested in [23]. By applying a 50 Ω termination resistor for the input buffer as shown in Figure 2.4.1a, and for the output buffer as shown in Figure 2.4.1b, s-parameter analysis may then be performed on these digital circuits as well. The capacitances presented by the buffer transistors themselves may be absorbed into the ESD protection system capacitance for design purposes.

In addition to allowing s-parameter analysis, application of termination resistances prevents intersymbol interference (ISI) by not generating any reflections at

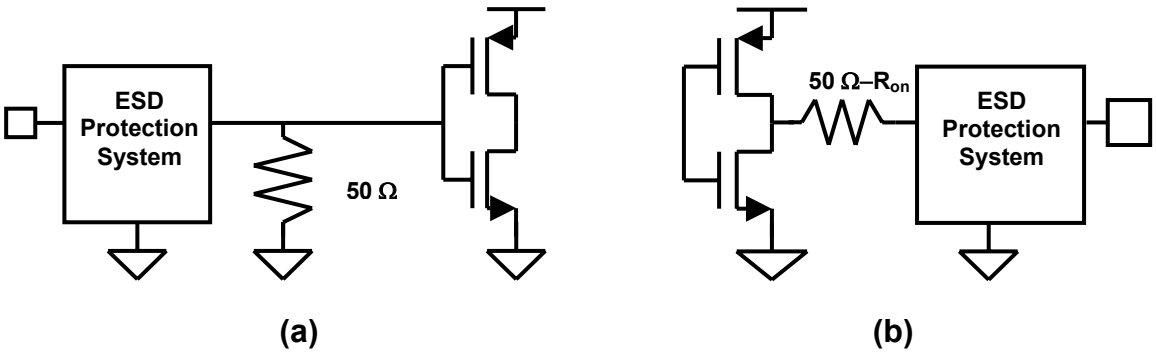


Figure 2.4.1. (a) Modified high-speed digital input buffer with a 50 Ω parallel resistive termination to prevent reflections. (b) Output buffer with a series resistive termination. Note that the series termination resistor would be less than 50 Ω depending on R_{on}, but since R_{on} should be small for a well-designed device, 50 Ω is a good approximation for the termination resistance.

the terminals. Also, crosstalk that generates reverse-traveling noise will no longer affect the forward-traveling signal because of proper termination at the terminals. The price that is paid for these benefits is additional die area, and perhaps an increase in power consumption.

With 50 Ω terminations both in RF and high-speed digital systems, effective signal transmission may be achieved by applying a distributed ESD protection system. Thus the next section will demonstrate methodologies for designing distributed ESD protection systems.

2.5 Design Methodologies for Optimized Distributed ESD Protection

Prior to determining the proper CPW characteristics, three parameters must be fixed. They are the maximum operating frequency f_{max}, the equivalent ESD capacitance C_{ESD}, and the CPW characteristic impedance Z_{CPW}. Selecting f_{max} should consist only of determining the maximum frequency specification for the core circuit. C_{ESD} should be

calculated after determining the proper ESD device size required for a particular protection level in a given technology. The equivalent capacitance may then be calculated from the device junction areas, or obtained through simulations.

The appropriate dimensions for the CPW as shown in Figure 2.5.1 must then be determined. The minimum width (*W*) must be such that given a line thickness (*t)*, which is technology dependent, the maximum possible ESD current can flow without causing open circuit failure of the line. An analytical model for the calculation of this width is provided in [24]. The distance between the substrate and the metal lines (*d*) is also technology dependent, but in general, the CPW should be placed as far away from the substrate as possible to reduce loss. The substrate resistivity also affects loss as shown in [25], where it is shown that the CPW on low-resistivity substrates suffers greater loss than corresponding CPW on high-resistivity substrates.

Once the minimum CPW signal line width is defined, the signal-to-ground spacings (*s*) of the CPW need to be determined from the desired characteristic impedance. In general, high characteristic impedance is desirable to minimize the required CPW length, but losses, which tend to be higher for higher impedance lines,

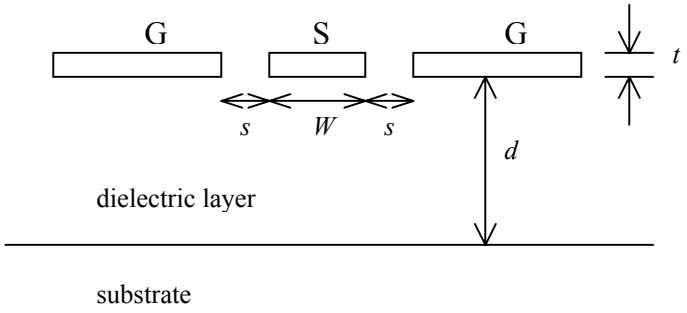


Figure 2.5.1. Cross-section of a coplanar waveguide (CPW). The lines are configured ground-signal-ground (G-S-G), with the key dimensions being the line width *W,* the spacing *s*, the CPW thickness *t*, and the distance of CPW from substrate *d*.

need to be minimized. The exact relationship between W, s, d, t, Z_{CPW}, substrate characteristics, and loss is a complex one that is beyond the scope of this study. In this work, a high characteristic impedance CPW line described in [26] is used. This CPW was fabricated on a low-resistivity (0.5 Ω-cm) substrate, with an aluminum (Al) metal layer 8 μm above the substrate—a distance typical for global interconnects. At present, VLSI metallization is realized with copper (Cu), which has higher electrical conductivity than Al [27], resulting in lower conductor loss. The CPW loss characteristics were measured in [26], and for this work, the ADS CPW model parameters were adjusted so that the loss characteristic of the model is similar to the experimentally measured loss. This loss characteristic is shown in Figure 2.3.5.

For the following example, f_{max} was set at 10 GHz, C_{ESD} was chosen to be 200 fF for a 2 kV protection level [21], and Z_{CPW} was chosen to be 100 Ω, a high-impedance, low-loss line according to [26]. For simulation in ADS, the transmission line has a thickness of 2μm, W=5μm, and s=10μm. Along with calculating the CPW length, the number of distributed sections to be created is also determined. The equivalent circuit used to achieve this is shown in Figure 2.5.2. First, factors of the ESD capacitance are

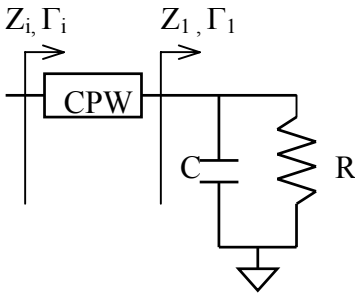


Figure 2.5.2. Circuit diagram for the example calculation in Section 2.5. R=50 Ω and f=10 GHz. The values of C and lengths of CPW vary. The impedances (Z) and reflection coefficients (Γ) are used to calculate the optimal C, CPW length, and number of distributed sections.

determined (in this case, 200 fF, 100 fF, 67 fF, 50 fF, 40 fF...) and are represented in the circuit as C. Then Z_1 is determined for each capacitance at f_{max}, and R is the 50 Ω load resistance. These impedances correspond to points on arc A on the Smith Charts (Figures 2.5.3 & 2.5.4). From the impedance Z_1, the reflection coefficient Γ_1 is calculated using the formula given by equation (2.9).

$$\Gamma_1 = \frac{Z_1 - Z_{CPW}}{Z_1 + Z_{CPW}} \tag{2.9}$$

Gamma, like s_{11}, represents the coordinate on the Smith Chart planes in Figures 2.5.3 and 2.5.4, with the center as the origin, and the outside circle being unity. Converting Γ_1 into polar coordinates gives the magnitude ρ and phase ϕ. By adding some length of CPW,

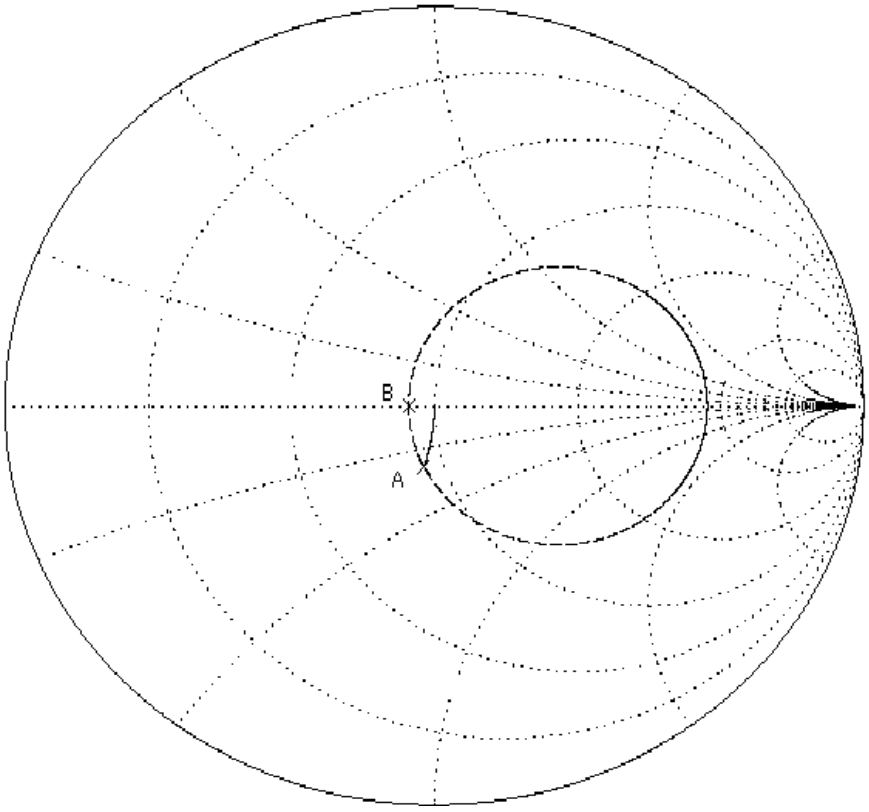


Figure 2.5.3. 50 Ω-normalized Smith Chart. Point A shows Γ_1 for 200 fF at 10 GHz. Point B shows Γ_i. The dotted circle shows the locus for different lengths of CPW.

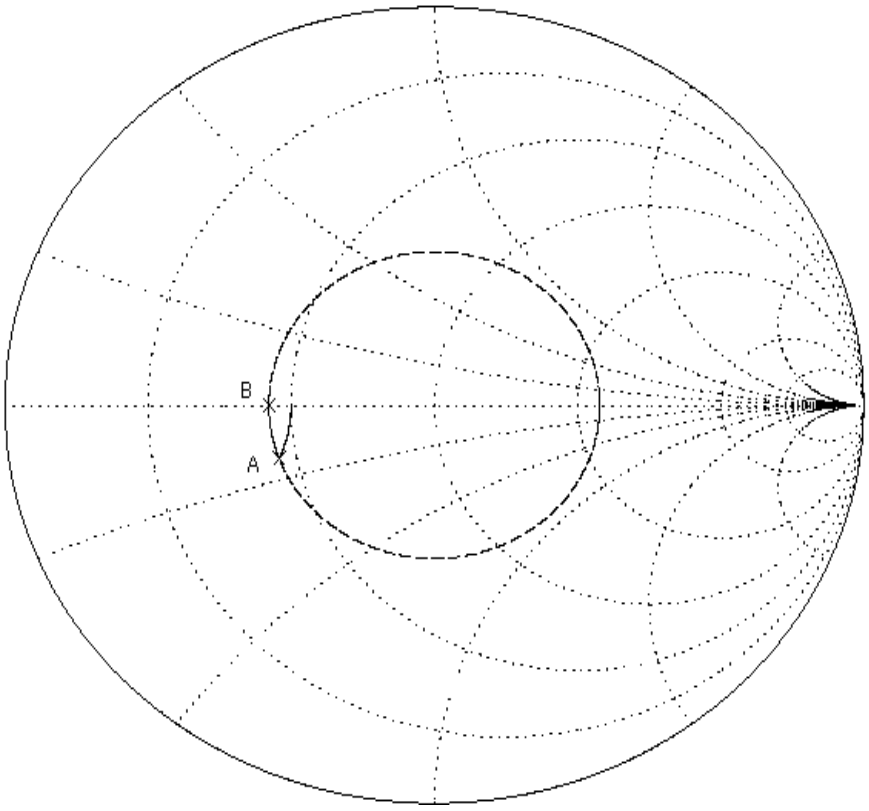


Figure 2.5.4. 100 Ω-normalized Smith Chart. Point A shows Γ_1 for 200 fF at 10 GHz. Point B shows Γ_i. The dotted circle shows the locus for different lengths of CPW.

we attempt to bring the phase to 180 degrees (point B). For a CPW with characteristic impedance of Z_{CPW}, the locus that results from adding CPW to Z_1 is a circle centered about the origin in Figure 2.5.4 (the dotted circle). Note that when this is viewed on a Smith Chart normalized to 50 Ω, the circle is not centered at the origin (Figure 2.5.3). Since a full circle around requires a CPW which is a half-wavelength ($\lambda/2$), the correct CPW length can be calculated from equation (2.10).

$$\text{CPW length} = \frac{|\phi - \pi|}{2\pi}\frac{\lambda}{2} \qquad (2.10)$$

Adding the correct CPW length should yield Γ_i with magnitude $\rho_i=\rho$ and ϕ_i=180 degrees. This can then be transformed into impedance with equation (2.11).

$$Z_i = \frac{(1+\Gamma_i)}{(1-\Gamma_i)} Z_{CPW} \tag{2.11}$$

As a measure of how close Z_i comes to the system impedance (Z_o) of 50 Ω, the standing wave ratio (SWR) can be calculated, as shown in equation (2.12).

$$SWR = \frac{1+\left|\frac{Z_i - Z_o}{Z_i + Z_o}\right|}{1-\left|\frac{Z_i - Z_o}{Z_i + Z_o}\right|} \tag{2.12}$$

A Z_i of 50 Ω results in an SWR of unity, and the closer the SWR is to unity, the better the match that is obtained. All the values calculated for this example are summarized in Table 2.1.

Now given the ESD capacitance and the maximum allowable SWR, the number of sections required to achieve those specifications can be readily calculated using the methodology above. Note that this is valid because R=50 Ω, and we try to bring Z_i to 50 Ω for each section. Thus the next section to be added can regard Z_i of the present section as just a 50 Ω load. However, if the capacitance of the section is large, the resulting SWR of the section will also be large, meaning that the Z_i of each section will not be equal to 50 Ω. In this case, error is introduced in the calculations, and this error is compounded when multiple sections, each with large SWR, are cascaded. While this

C (fF)	# sections	Z_1 (Ω)	Γ_1	ρ	ϕ (rad)	Arc length $\lvert\phi-\pi\rvert$ (rad)	Z_i (Ω)	SWR
200	1	35.85-22.52j	-0.43-0.24j	0.494	-2.64	0.502	33.9	1.48
100	2	45.51-14.30j	-0.36-0.13j	0.385	-2.79	0.355	44.4	1.13
67	3	47.88-10.08j	-0.35-0.092j	0.358	-2.88	0.259	47.3	1.06
50	4	48.80-7.66j	-0.34-0.069j	0.348	-2.94	0.200	48.4	1.03
40	5	49.22-6.19j	-0.338-0.056j	0.343	-2.98	0.163	49.0	1.02

Table 2.1. Summary of results from the example case. Note that with increased sections, Z_i and SWR approach 50 Ω and 1.00, respectively.

error may be somewhat mitigated by the series resistance found in the transmission line, if the SWR of a single section is large, then the system SWR may be larger than suggested by the simple calculations. But since this methodology has decreased accuracy only when the SWR is large, which is undesirable anyway, the calculation results should be acceptable as long as the SWR of each section is close to unity.

One last consideration is to determine whether this analysis, performed for a system operating at f_{max}, is valid when the system operates at a frequency lower than f_{max}. It can be shown that at lower frequencies, the capacitance has a larger impedance, thus the arc on the Smith Chart becomes shorter, and ρ decreases, resulting in a smaller SWR. Thus the f_{max} case is the worst case, and if the performance at this frequency is satisfactory, then the performance at any lower frequency will be at least as good as that seen at f_{max}.

While the above methodology relies on some geometrical concepts regarding the Smith Chart, there is an alternate, more algebraic methodology by which the same results may be obtained. This methodology uses ABCD-matrices, which, like the s-parameter matrix, is a matrix set that describes the characteristics of a 2-port network. The ABCD-matrix has the characteristics that when multiple 2-port networks are cascaded, the system response is characterized by the product of the ABCD-matrices of the component 2-port networks. This property of ABCD-matrices can also be used to calculate the correct CPW length. The general form of the ABCD-matrix for a transmission line with characteristic impedance Z_{o1} and length l, and for a shunt impedance Z, are shown in equations (2.13) and (2.14), respectively.

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} \cos\frac{2\pi}{\lambda}l & j\frac{Z_{o1}}{Z_o}\sin\frac{2\pi}{\lambda}l \\ j\frac{Z_o}{Z_{o1}}\sin\frac{2\pi}{\lambda}l & \cos\frac{2\pi}{\lambda}l \end{bmatrix} \tag{2.13}$$

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \frac{Z_o}{Z} & 1 \end{bmatrix} \tag{2.14}$$

In our example, the system impedance Z_o is 50 Ω, while Z_{o1} is 100 Ω. Looking again at Figure 2.5.2, the cascading of the CPW with the shunt impedance Z_l gives the ABCD-matrices for the system as shown in equation (2.15).

$$\begin{bmatrix} A_i & B_i \\ C_i & D_i \end{bmatrix} = \begin{bmatrix} \cos\frac{2\pi}{\lambda}l & j2\sin\frac{2\pi}{\lambda}l \\ j0.5\sin\frac{2\pi}{\lambda}l & \cos\frac{2\pi}{\lambda}l \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 50\left(\frac{1}{R} + j\omega C\right) & 1 \end{bmatrix} \tag{2.15}$$

The product of the ABCD-matrix for the CPW and for the parallel RC combination gives the system response looking into the input terminal, as denoted by the subscript *i*. From the product, the input impedance Z_i may be calculated by taking A_i/C_i of the matrix. By setting the imaginary component of this impedance to zero, since we desire the impedance to be a resistance close to 50 Ω, the CPW length may be calculated to be as shown in equation (2.16).

$$\frac{l}{\lambda} = \frac{1}{4\pi}\tan^{-1}\frac{2Z_o\omega C}{\frac{Z_{o1}}{Z_o}\left(\left(\frac{Z_o}{R}\right)^2 + (Z_o\omega C)^2\right) - \frac{Z_o}{Z_{o1}}} \tag{2.16}$$

Using this length, the input resistance, R_i, may be calculated as shown in equation (2.17).

$$R_i = \frac{Z_o^2 / R}{\left(\frac{Z_o}{R}\right)^2 \cos^2\left(\frac{2\pi}{\lambda}l\right) + \left(\frac{Z_o}{Z_{o1}}\sin\left(\frac{2\pi}{\lambda}l\right) + \omega C Z_o \cos\left(\frac{2\pi}{\lambda}l\right)\right)^2} \tag{2.17}$$

Repeating the calculations for the different capacitances, using ABCD-matrices yields results identical to those obtained using the Smith Chart method.

2.6 Conclusion

In conclusion, this chapter has presented quantitative methodologies to analyze the performance degradation of RF and high-frequency mixed-signal circuits due to the integration of distributed ESD protection devices at I/O pads. Detailed s-parameter analysis of high-frequency systems has been carried out to illustrate the impact of the parasitics associated with distributed ESD protection schemes. Circuit simulations using coplanar waveguides to absorb a 200 fF parasitic capacitance of the ESD protection devices have shown that a 4-stage distributed ESD protection system suffers only a 0.02 dB loss between 0-10 GHz, with a maximum loss of 0.273 dB at 10 GHz, showing excellent broadband characteristics. Two generalized design methodologies, using Smith Charts and ABCD-parameters, have been developed to optimize the number and length of coplanar waveguides separating the distributed ESD elements. For a parasitic capacitance of 200 fF, a 4-stage distributed ESD protection can be designed with 0.4 radians of 100 Ω transmission line to give an SWR of 1.03 at 10 GHz. Furthermore, termination schemes have been proposed to allow this analysis to be applicable to high-speed digital and mixed-signal systems. This work illustrates the utility of the distributed ESD protection structure for high-frequency applications and provides design methods, which can be useful to ESD and I/O designers.

3 RF Power Amplifier Scaling

The proliferation of wireless communications systems worldwide has increased the demand for better performance in these systems. Whether the transmission content is voice or data, most of the functions within the transmit/receive system remain the same. As shown in Figure 3.0.1, the transmit system consists of blocks which process, modulate, and amplify the input signal. The receive system consists of blocks which amplify, demodulate, and process the signal. Out of all these blocks, the power amplifier consumes much of the system power budget, and improving its efficiency should significantly improve the overall system efficiency.

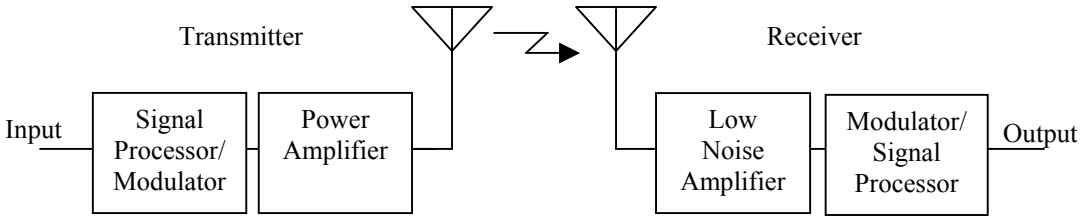


Figure 3.0.1. Block diagram of a simplified wireless communications system. Signal flows from the transmitter on the left side to the receiver on the right side.

Currently, the method by which more output power is derived from RF laterally-diffused (LD) MOSFET power amplifier devices consists of using devices with larger channel widths (W). However, experiments have shown that the RF performance per unit device width decreases as total W increases. Figure 3.0.2 shows two devices—Figure 3.0.2a, which has a channel width W, and Figure 3.0.2b, which has a channel width αW. When an input power P_{in} is applied to the device in Figure 3.0.2a, maximum output power Pout is obtained at a power-added efficiency η_1, where power added efficiency is defined as shown in equation (3.1).

$$\eta = \frac{P_{out} - P_{in}}{P_{dc}} \tag{3.1}$$

Ideally, when an input power αP_{in} is applied to the device in Figure 3.0.2b, both the maximum output power and the DC dissipation power should scale by α, and the efficiency should remain at η_1. However, in reality, the maximum output power (P_{out2}) of the device in Figure 3.0.2b is less than αP_{out}, and an input power greater than αP_{in} is

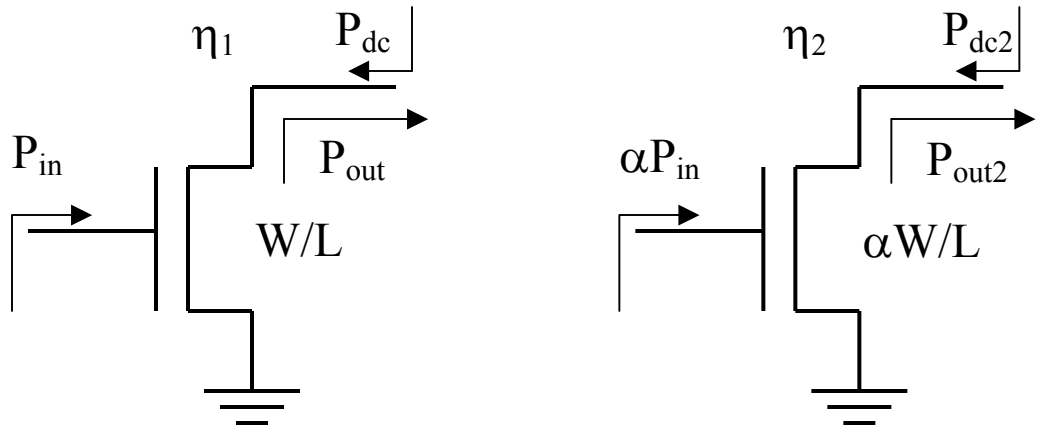


Figure 3.0.2. Schematics showing input powers, output powers, DC powers, and efficiencies for devices with different widths. When the width increases by a factor α, the new output power (P_{out2}) with new input power αP_{in}, is less than αP_{out}. This leads to $\eta_2 < \eta_1$.

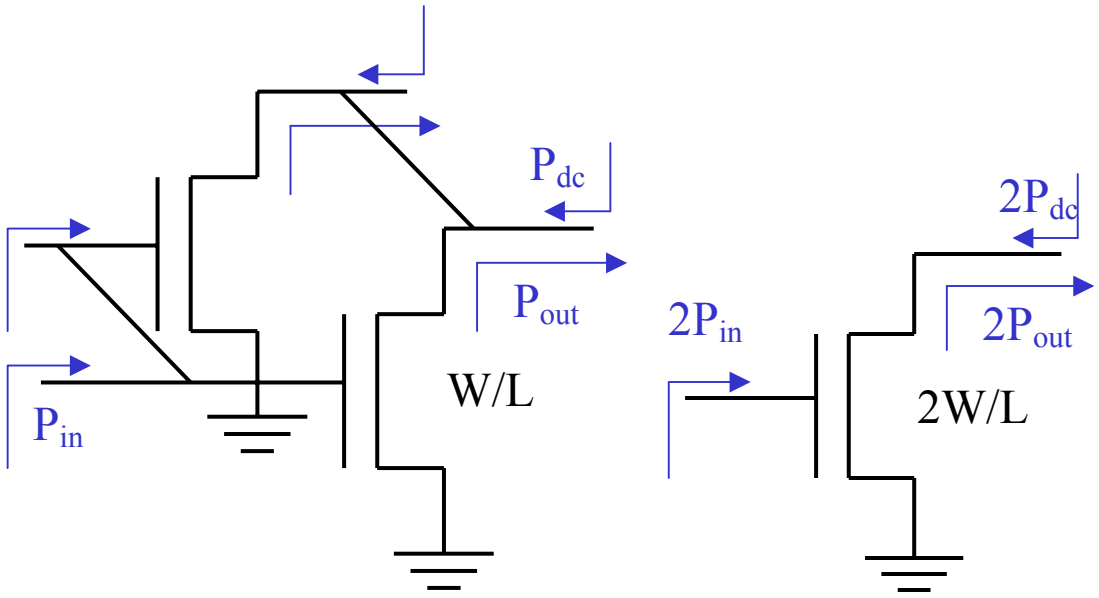


Figure 3.0.3. Schematics showing input powers, output powers, DC powers, and efficiencies for a device consisting of two unit cells. Note that the two schematics are equivalent.

required for an output power of αP_{out}, thereby lowering the efficiency.

In the device design process, obtaining a device with a larger W consists of placing identical unit cells in parallel. A case with two unit cells (α=2) is shown in Figure 3.0.3, where two identical devices are tied together in parallel. The device on the right side is the functional equivalent of the parallel unit cells on the left side. As labeled, each unit cell sees an input power P_{in}, a maximum output power P_{out}', and dissipates P_{dc}', where $P_{out}'<P_{out}$ from Figure 3.0.2, and P_{dc}' may not necessarily be the same as P_{dc}. Therefore, as a system, it takes in $2P_{in}$, outputs $2P_{out}'$, and dissipates $2P_{dc}'$.

Measurement results demonstrate this phenomenon, as can be seen in Figure 3.0.4, where output power is not linearly proportional with W [28]. Practically, higher efficiencies have been extracted from power amplifier systems through the use of power dividers and combiners as shown in Figure 3.0.5 [28], but these solutions require larger areas for the divider/combiner circuits, making them less than ideal.

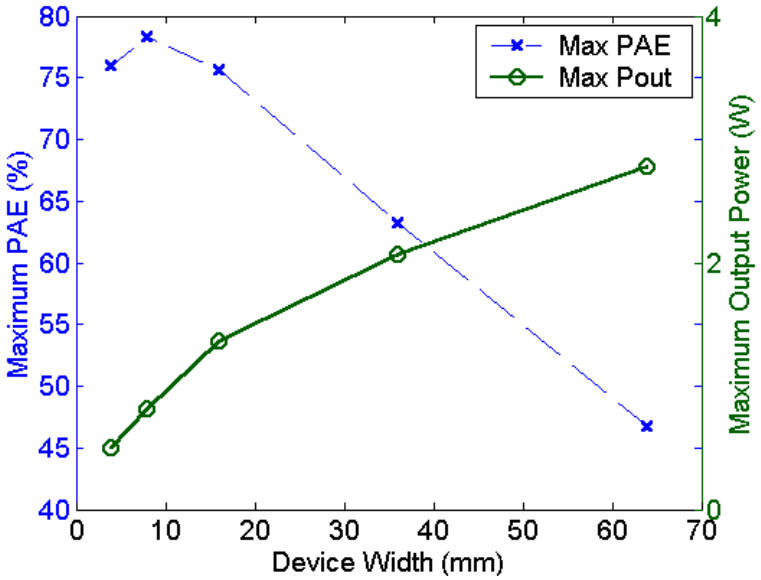


Figure 3.0.4. Measured performance against device width for L4B LDMOSFETs. The maximum power-added efficiency (PAE), measured under constant output power per width, decreases by about 0.6 %/mm in a near-linear fashion, and the maximum output power, measured under constant input power per width, does not increase linearly with device width.

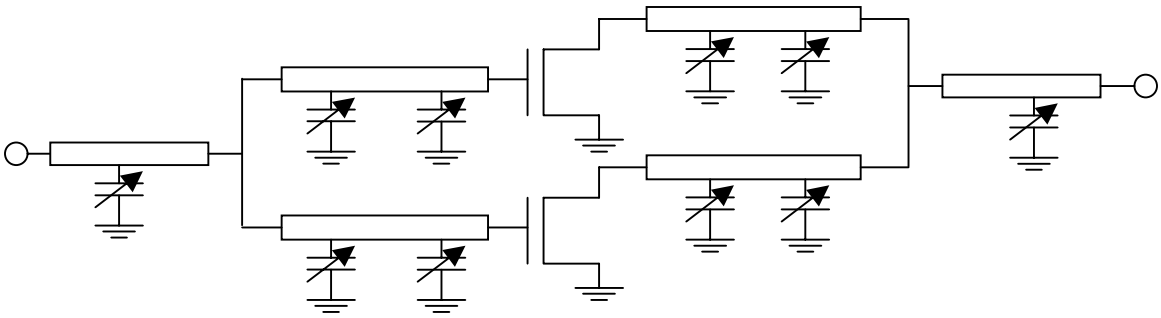


Figure 3.0.5. Circuit schematic showing power divider and combiner. This topology allows efficient device operation such that each half section operates optimally. However, the larger area requirement due to the transmission lines and tuning capacitors is undesirable.

This chapter examines the LDMOSFET in detail to identify the mechanism of the performance degradation that results from integrating identical unit cells onto one chip. First, Section 3.1 will examine the available measurement data. Hypotheses generated from the data will then be tested in Section 3.2 through experiments.

3.1 LDMOSFET Device

Shown in Figure 3.1.1 is the cross-section of an LDMOSFET device. The difference between this device and a regular MOSFET lies in the drain extension, a lightly-doped region that gives the LDMOSFET its higher breakdown voltage, a characteristic necessary for power amplifiers. As shown in Figure 3.1.2, the finger width (W_f) is the length of the active region in the z-direction. When there are multiple fingers, the number of gate fingers multiplied by W_f gives the total width. In power amplifier devices, a group of fingers connects to a set of terminal pads, and this collection is called a unit cell. Practically, in order to increase the width of the power amplifier device, unit cells are placed in parallel to obtain the desired total width. If each device finger is not

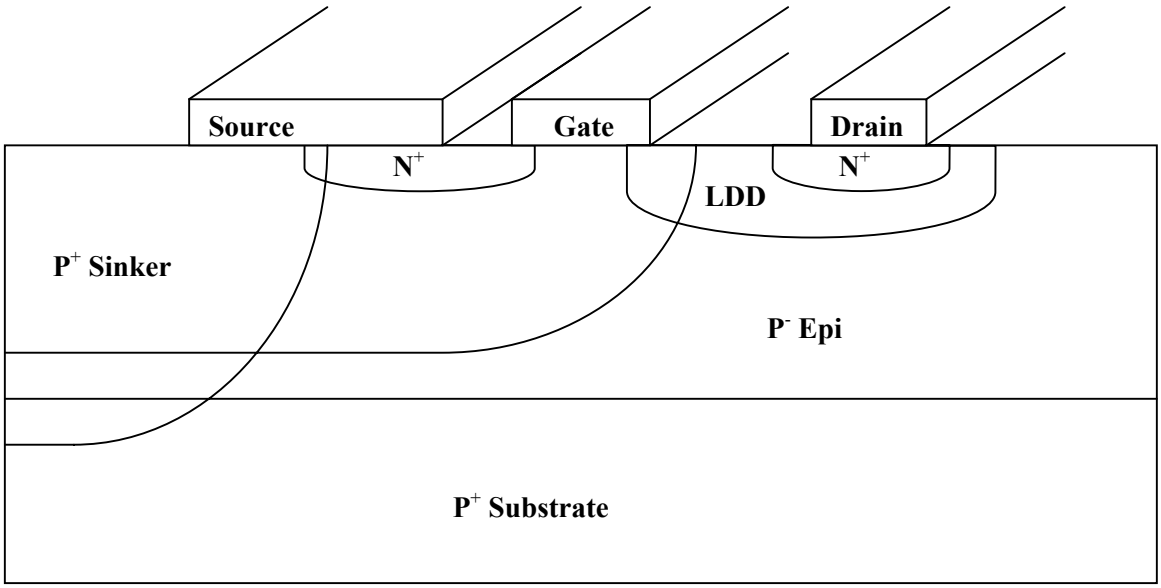


Figure 3.1.1. Cross-section of an RF LDMOSFET device. Compared to a standard MOSFET, the LDD region allows higher drain voltages, and the p^+ sinker allows the source contact to be on the backside of the chip.

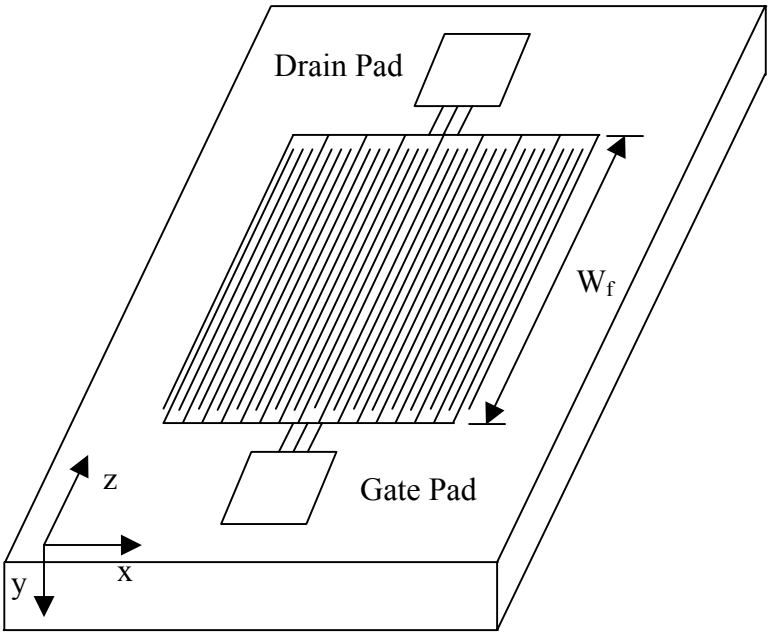


Figure 3.1.2. Top view of a multi-fingered LDMOSFET device. The alternating drain-gate-source fingers in parallel allows a large gate width to be packed densely. As shown, each finger has a width W_f, and in this case, with 16 gate fingers, the total device channel width is $16W_f$.

affected by the other parallel fingers, the total device should function as the sum of its parts. But measurements in the following sections show that the device does not function as well as the sum of its parts, and that this performance degradation is caused by an integration problem exacerbated by the high-frequency operation of the system.

3.1.1 DC Measurements

Pulsed current-voltage (IV) measurements have been made on L4B LDMOS power amplifier transistors provided by Hitachi Ltd. [29, 30] with device widths of 4mm, 8mm, and 16mm. The normalized results, shown in Figure 3.1.3, display three sets of curves measured with V_{gs} between 0 to 7V with 1V steps. For all widths, the drain currents per unit gate width show little variation, and no consistent trend with respect to changes in width. However, when the IV measurements are taken at different

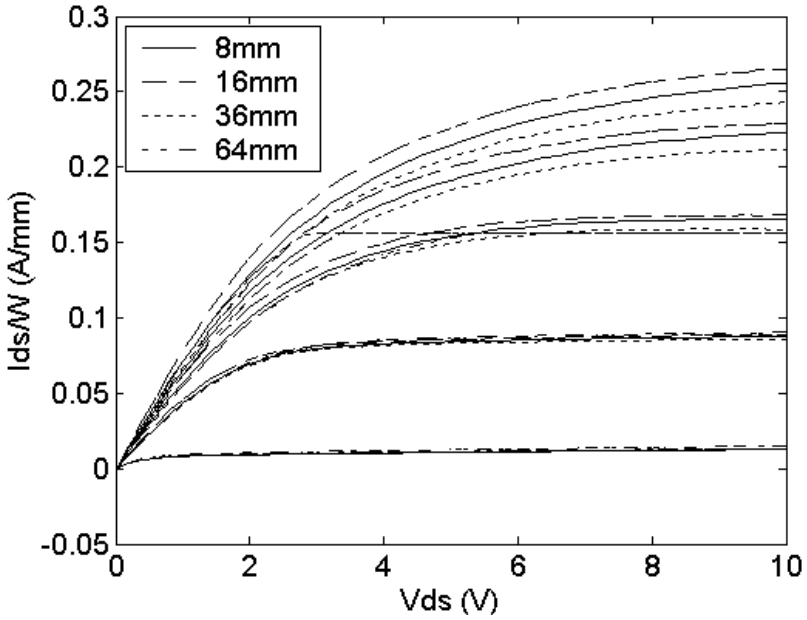


Figure 3.1.3. Width-normalized DC IV curves for devices with different widths, at Vg=1V to 5V, with 1V steps. For the most part, the currents are similar between all widths, and there is no consistent trend with respect to width. The deviation at larger gate biases are likely due to different self-heating in the devices. Also, the W=64mm curves at higher gate biases have hit the measurement device compliance to yield a horizontal line after the compliance point.

temperatures, the curves deviate as shown in Figure 3.1.4. Under normal RF operation, larger devices dissipate more power, and experience increased self-heating. Equation (3.2) shows the idealized drain current equation in saturation for an n-type MOSFET [31].

$$I_d = \mu_n C_{ox} \frac{W}{L} (V_{gs} - V_T)^2 \qquad (3.2)$$

The temperature sensitive parameters in equation (3.2) are the threshold voltage, V_T, and the electron mobility, μ_n. Equations (3.3) and (3.4) show the relationships between these two parameters and temperature [31].

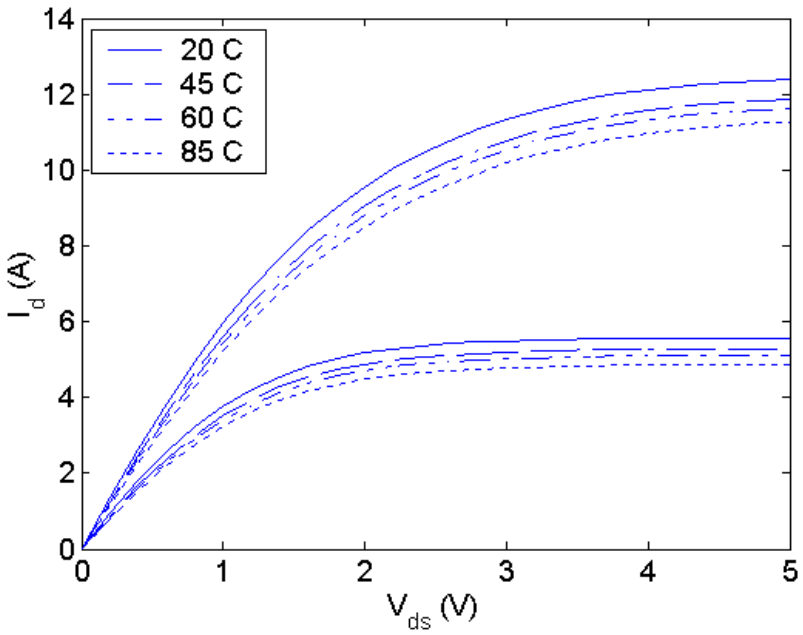


Figure 3.1.4. IV plots of a W=16mm device at different temperatures. The gate voltage is 2V for the lower set of curves, and 5V for the upper set of curves. Note that in both cases, the drain current for the warmer device is lower by around 10%.

$$\mu_n = 92\left(\frac{T}{300}\right)^{-0.57} + \frac{1268\left(T/300\right)^{-2.33}}{1+\left(\frac{N}{1.3\cdot 10^{17}}\left(\frac{T}{300}\right)^{-2.4}\right)^{0.91\left(T/300\right)^{-0.146}}} \tag{3.3}$$

$$V_T = 2\phi_F + \frac{K_s x_o}{K_o}\sqrt{\frac{4qN_A}{K_s\varepsilon_o}\phi_F} \tag{3.4}$$

where

$$\phi_F \approx \frac{kT}{q}\ln\left(\frac{N_A}{n_i}\right) \qquad \text{when } N_A >> n_i \tag{3.5}$$

$$n_i = 9.15\cdot 10^{19}\left(\frac{T}{300}\right)^2 e^{\frac{-0.5928}{kT}} \tag{3.6}$$

Figure 3.1.5 shows μ_n at different temperatures as calculated using the empirical relationship shown in equation (3.3), where *N* is the doping concentration. It can be seen

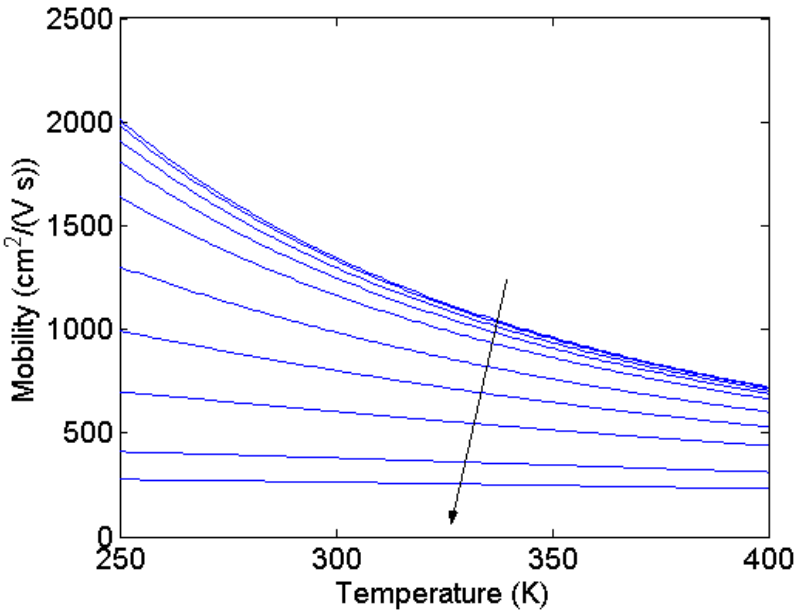


Figure 3.1.5. Plot of electron mobility against temperature for different doping concentrations. Generally, increased temperature decreases mobility. Also, as the arrow shows, increasing doping concentration decreases mobility, although at higher concentrations, mobility becomes less sensitive to temperature.

that for any given *N*, the mobility decreases with increased temperature. Figure 3.1.6 shows V_T at different temperatures for different doping concentrations, obtained from equations (3.4)-(3.6). Equation (3.6) is an empirically obtained relationship demonstrated in [32]. The plot shows that increasing temperature decreases V_T, regardless of the doping concentration. Since V_T and μ_n both decrease, they pull the drain current in opposite directions as temperature changes. That is, it can be seen from equation (3.2) that decreasing V_T increases drain current, and decreasing μ_n decreases drain current. When V_{gs} is large, changes in V_T affect the drain current less, while μ_n is always proportional to the drain current. Thus it is difficult to generalize which effect wins out, and both effects must be considered when heating occurs in the devices.

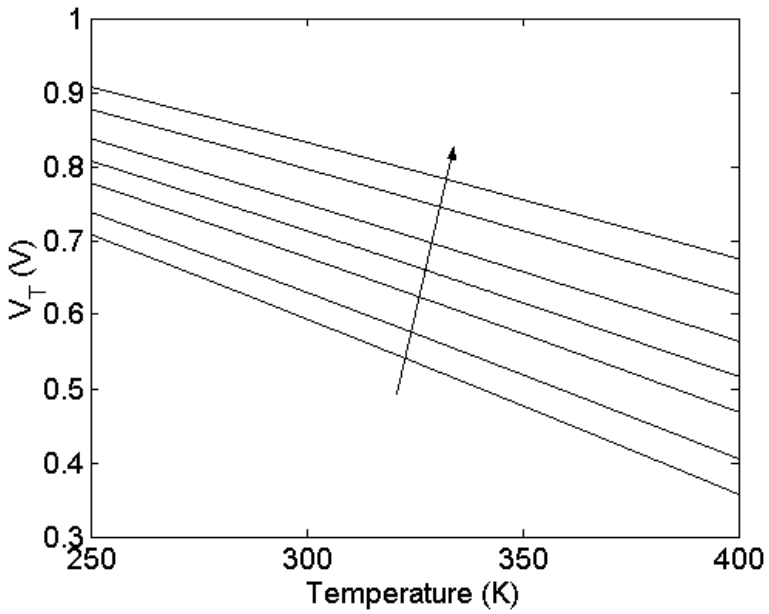


Figure 3.1.6. Plot of MOSFET threshold voltage against temperature for different doping concentrations. As temperature is increased, V_T decreases. While V_T increases for higher doping concentrations, the sensitivity of V_T to temperature does not change significantly.

3.1.2 CV Measurements

Low-frequency capacitance-voltage (CV) measurements have also been made on the transistors. Power amplifier capacitances are often characterized as C_{iss}, C_{oss}, and C_{rss}, and these measurements were taken on W=4mm, 8mm, 16mm, 36mm, and 64mm devices. The input short-circuit capacitance (C_{iss}), is measured across the gate and source terminals, with the drain terminal grounded, the output short-circuit capacitance (C_{oss}), is measured across the drain and source terminals, with the gate terminal grounded, and the reverse-transfer short-circuit capacitance (C_{rss}) is measured across the drain and gate terminals, with the source terminal grounded. So these three capacitances roughly correspond to C_{gs}, C_{ds}, and C_{gd}, respectively. The normalized measurement results, obtained by dividing the observed capacitance by the device width, are as shown in Figure 3.1.7-3.1.9, which show that the capacitances do not follow a consistent trend with

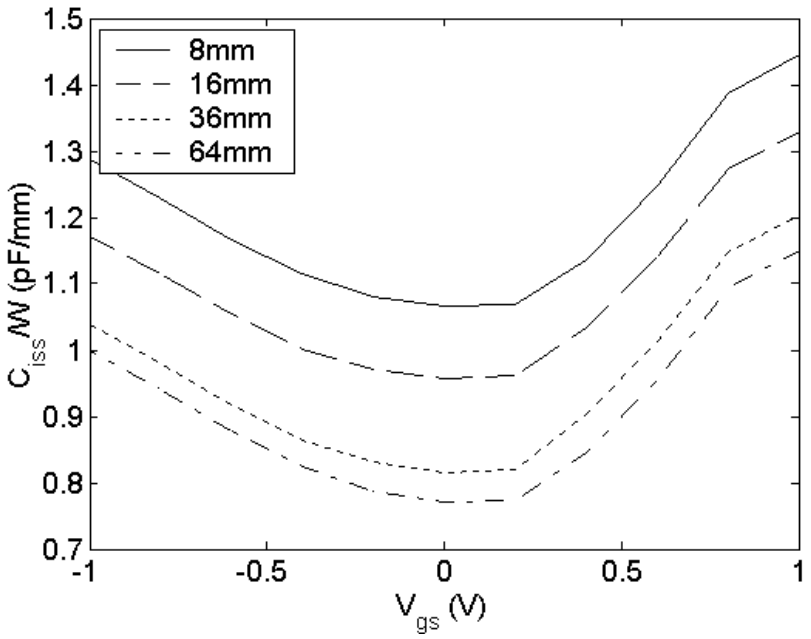


Figure 3.1.7. Width-normalized C_{iss} curves for devices with different widths. C_{iss} includes the gate pad capacitance and C_{gs}, and with constant package capacitance, constant gate protection diode capacitance, and fewer pads per device width for larger devices, the larger devices are expected to show smaller capacitances.

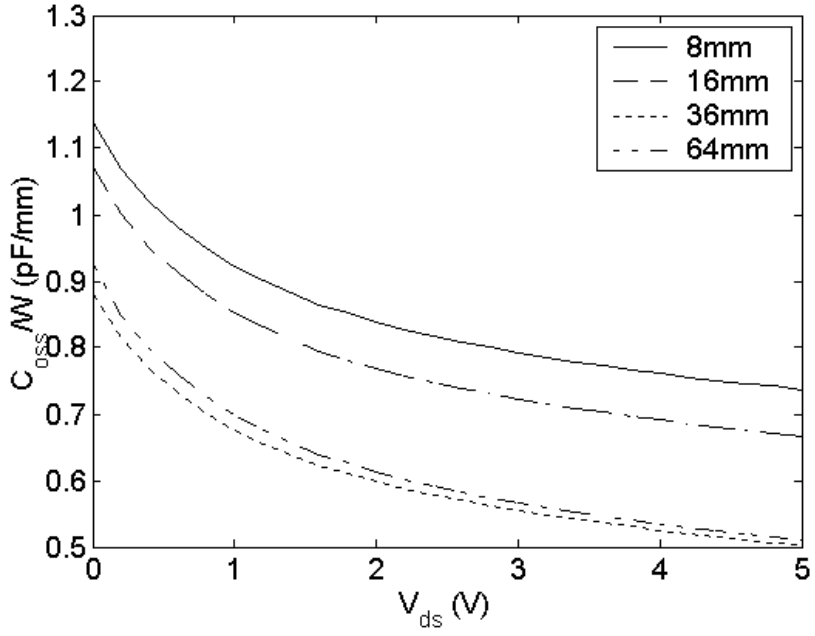


Figure 3.1.8. Width-normalized C_{oss} for different device widths. Like Figure 3.1.7, the larger devices are expected to exhibit smaller capacitances due to the pad density and package capacitance.

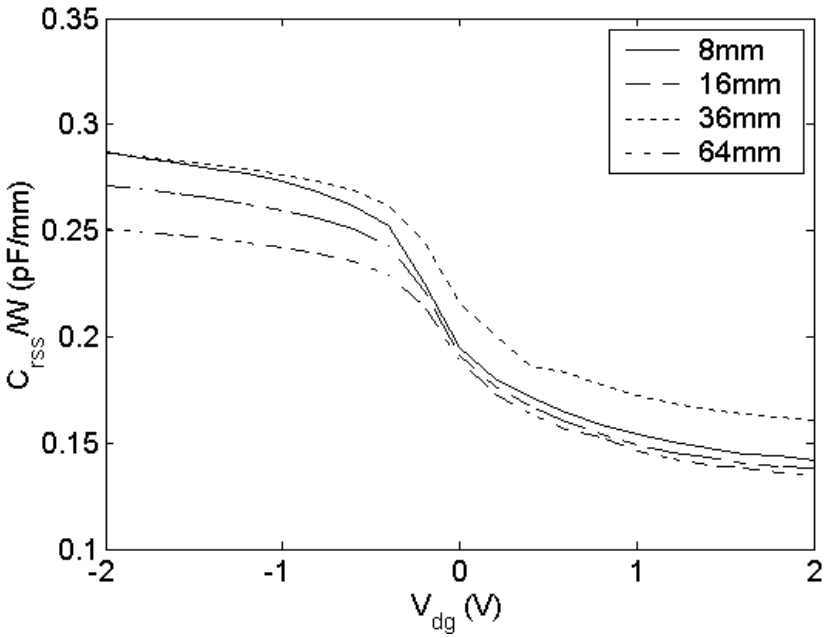


Figure 3.1.9. Width-normalized C_{rss} for devices with different widths. In this case, the capacitances show no trend with W, as parasitic capacitances should not show up in the measurements.

width. There is a small decrease in the normalized capacitance with increased device width, which may be attributed to: a smaller number of terminal pads per device width at larger widths, a constant package capacitance, and a constant gate protection diode capacitance. Figure 3.1.10 shows measurement results (provided by Hitachi Ltd.) for a different set of devices, but with the pad capacitance, package capacitance, and gate protection diode capacitance properly extracted and de-embedded. These results show constant internal device capacitance over device width. Since the internal capacitances stay constant, and the external capacitances may be countered with proper matching, these CV measurements do not provide any useful information regarding the cause of performance degradation with increased width.

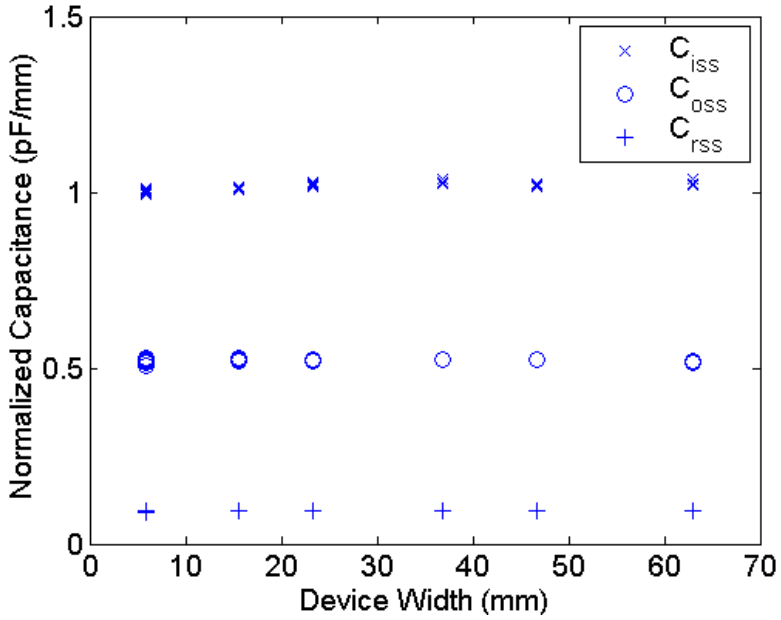


Figure 3.1.10. Normalized capacitance data after de-embedding the package capacitance, gate protection diode capacitance, and the pad capacitance. These capacitances show no change with width, as opposed to the curves in Figures 3.1.7-3.1.9.

3.1.3 Small-Signal RF Measurements

Small-signal RF measurements are made using the HP 8510 network analyzer [33]. Measurements were made on transistors with widths of 4mm, 8mm, 16mm, 36mm, and 64mm between 500MHz and 5.5GHz at a gate bias of 1.2V, and a drain bias of 3.5V. In order to properly compare the s-parameters, the data was normalized to a width of 1mm. The normalization procedure first consisted of transforming the s-parameter results into y-parameters according to the relationships shown in equation (3.7).

$$\begin{bmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{bmatrix} = \begin{bmatrix} \dfrac{(1-s_{11})(1+s_{22})+s_{12}s_{21}}{(1+s_{11})(1+s_{22})-s_{12}s_{21}} & \dfrac{-2s_{12}}{(1+s_{11})(1+s_{22})-s_{12}s_{21}} \\ \dfrac{-2s_{21}}{(1+s_{11})(1+s_{22})-s_{12}s_{21}} & \dfrac{(1+s_{11})(1-s_{22})+s_{12}s_{21}}{(1+s_{11})(1+s_{22})-s_{12}s_{21}} \end{bmatrix} \tag{3.7}$$

The y-parameters were then divided by width, as shown in equation (3.8).

$$\tilde{y} = \frac{y}{W} \tag{3.8}$$

Finally, the normalized y-parameters were transformed back to s-parameters using equation (3.9).

$$\begin{bmatrix} \tilde{s}_{11} & \tilde{s}_{12} \\ \tilde{s}_{21} & \tilde{s}_{22} \end{bmatrix} = \begin{bmatrix} \dfrac{(1-\tilde{y}_{11})(1+\tilde{y}_{22})+\tilde{y}_{12}\tilde{y}_{21}}{(1+\tilde{y}_{11})(1+\tilde{y}_{22})-\tilde{y}_{12}\tilde{y}_{21}} & \dfrac{-2\tilde{y}_{12}}{(1+\tilde{y}_{11})(1+\tilde{y}_{22})-\tilde{y}_{12}\tilde{y}_{21}} \\ \dfrac{-2\tilde{y}_{21}}{(1+\tilde{y}_{11})(1+\tilde{y}_{22})-\tilde{y}_{12}\tilde{y}_{21}} & \dfrac{(1+\tilde{y}_{11})(1-\tilde{y}_{22})+\tilde{y}_{12}\tilde{y}_{21}}{(1+\tilde{y}_{11})(1+\tilde{y}_{22})-\tilde{y}_{12}\tilde{y}_{21}} \end{bmatrix} \tag{3.9}$$

Figure 3.1.11 shows the normalized s-parameter results. If the characteristics of the devices scaled properly with width, the curves on each of the s-parameter plots should be identical after normalization. However, from the s_{11} plot, it can be seen that as width increases, the curves shift to the right and rotate clockwise on the Smith Chart. In terms of impedance, the rightward shift indicates an increase in the device input resistance, while the clockwise movement indicates an increase in the device input reactance. The s_{22} plot shows similar movements in the curves, indicating an increasing output resistance and reactance with increasing width. As for the s_{21} curves, plotted on a polar chart, the lengths of the curves are shown to decrease with increasing width, indicating the decrease of forward transmission. The s_{12} plot shows curves with larger magnitudes as width increases, indicating an increasing reverse transmission. While the low-frequency measurements from the previous sections showed proper scaling with width, these RF measurements clearly show a non-proportional scaling with width.

3.1.4 Maximum Available Gain, Maximum Stable Gain, and Stability

From the s-parameter data obtained in the previous section, the maximum possible performance of the amplifier may be quantified. The maximum power transfer

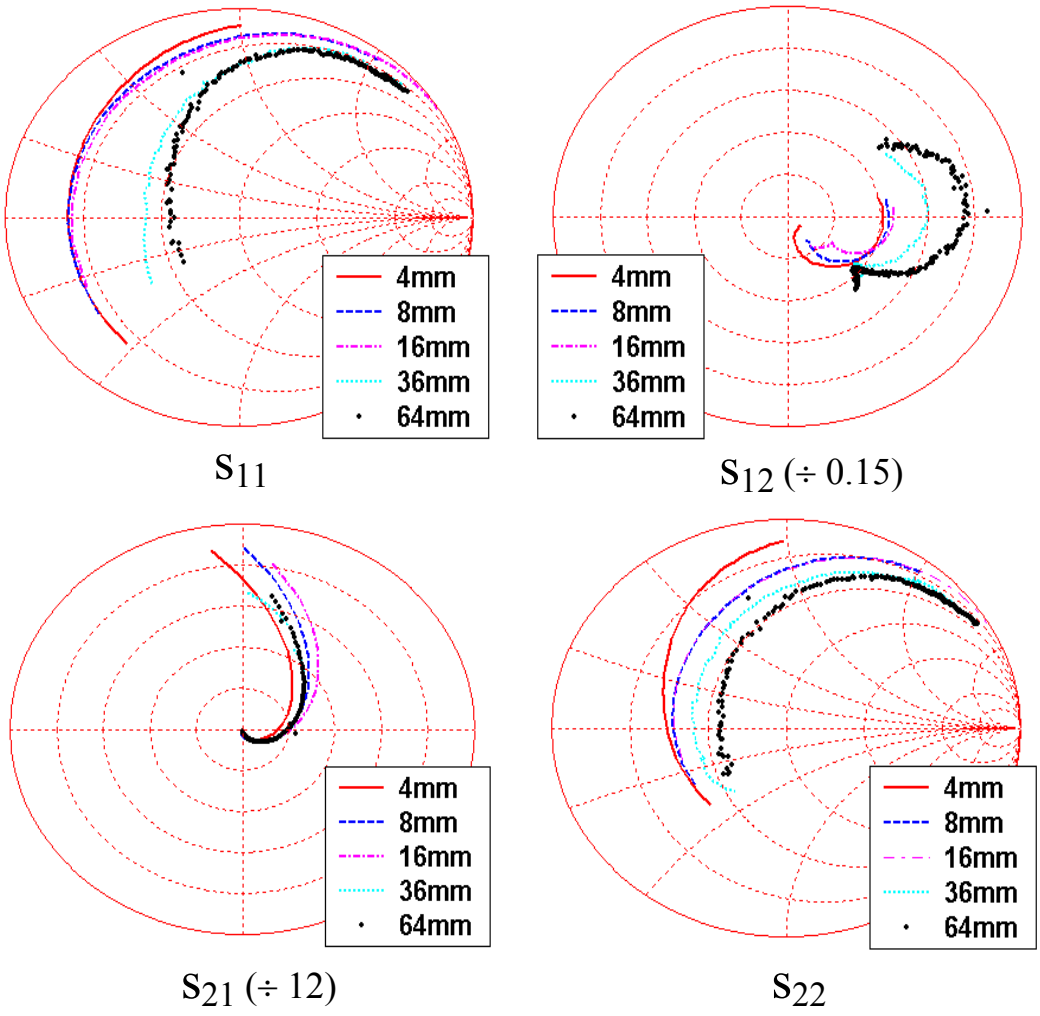


Figure 3.1.11. Measured small-signal s-parameter results for devices with different widths. The frequency is swept from 0.5-5.5GHz, while V_{gs}=1.2V, and V_{ds}=3.5V. It can be seen that the input resistance and output resistance increase with W, and the forward transmission decreases with W.

theorem states that given some complex load impedance, maximum power will be delivered to the load when the source impedance is the complex conjugate of the load impedance. In an amplifier system, power transfer occurs from the generator to the amplifier, and from the amplifier to the load. Ideally, maximum power transfer occurs in both these instances through simultaneous conjugate impedance matching at the input

and the output. The system gain that could be achieved by simultaneous conjugate matching is defined as the maximum available gain (MAG) [34-36].

In [35], the derivations for the optimal source and load impedances are presented in terms of z-parameters, as shown in Equations (3.10)-(3.11).

$$Z_s^* = z_{11} - \frac{z_{12}z_{21}}{z_{22} + Z_L} \tag{3.10}$$

$$Z_L^* = z_{22} - \frac{z_{12}z_{21}}{z_{11} + Z_s} \tag{3.11}$$

Further manipulations of the equations yield the real and imaginary parts of the optimal source and load impedances, as shown in Equations (3.12)-(3.15).

$$R_{Lo} = \frac{1}{2\,\mathrm{Re}(z_{11})}\sqrt{\left[2\,\mathrm{Re}(z_{11})\mathrm{Re}(z_{22}) - \mathrm{Re}(z_{12}z_{21})\right]^2 - \left|z_{12}z_{21}\right|^2} \tag{3.12}$$

$$X_{Lo} = \frac{\mathrm{Im}(z_{12}z_{21})}{2\,\mathrm{Re}(z_{11})} - \mathrm{Im}(z_{22}) \tag{3.13}$$

$$R_{so} = \frac{\mathrm{Re}(z_{11})}{\mathrm{Re}(z_{22})} R_{Lo} \tag{3.14}$$

$$X_{so} = \frac{\mathrm{Im}(z_{12}z_{21})}{2\,\mathrm{Re}(z_{22})} - \mathrm{Im}(z_{11}) \tag{3.15}$$

It can be seen that these equations are valid only when the quantity under the square root in Equation (3.12) is greater than or equal to zero, and in the case that this quantity is less than zero, oscillations may occur when simultaneous conjugate impedance matches are applied. Thus this quantity presents a measure of the stability of the system, and in [36], the stability factor (k) is obtained by setting the quantity under the square root to zero and rearranging the terms, as shown in Equation (3.16).

$$k = \frac{2\,\mathrm{Re}(z_{11})\mathrm{Re}(z_{22}) - \mathrm{Re}(z_{12}z_{21})}{|z_{12}z_{21}|} \tag{3.16}$$

The stability factor indicates the degree of stability of a system, with k>1 being the stable condition. When k is found to be less than one, the system is unstable under simultaneous conjugate impedance match, and the concept of maximum available gain is meaningless. In order to quantify the performance of a potentially unstable system, the maximum stable gain (MSG) is used. The maximum stable gain is thus defined as the gain of the system while setting the input and output impedance matches such that k=1 [35].

Another important result shown in [36] is that the parameters used in the stability factor calculation need not be z-parameters. It can also be y-, h-, or g-parameters, and this fact can be used to derive some intuitive understanding about stability in MOSFET amplifiers. For example, given a simple MOSFET transistor model as shown in Figure 3.1.12, the y-parameters for the transistor can be readily calculated to be as follows:

$$y_{11} = j\omega(C_{gs} + C_{gd}) \tag{3.17}$$

$$y_{12} = -j\omega C_{gd} \tag{3.18}$$

$$y_{21} = g_m - j\omega C_{gd} \tag{3.19}$$

$$y_{22} = j\omega(C_{ds} + C_{gd}) \tag{3.20}$$

Using these relationships and the invariance of Rollett's formula, it can be shown that the stability factor in terms of the model parameters is as shown in Equation (3.21).

$$k = \frac{1}{\sqrt{1 + \frac{g_m^2}{\omega^2 C_{gd}^2}}} \tag{3.21}$$

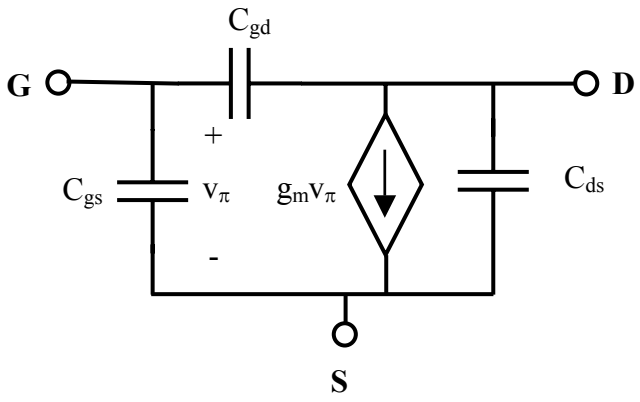


Figure 3.1.12. Simplified MOSFET model used to calculate the y-parameters of the transistor.

By analyzing the equation, it can be seen that in this case, the stability factor can never be greater than unity. However, once inductances and resistances are introduced to the model at the device terminals, the y-parameter expressions, and therefore the expressions for the stability factor and gain, become complicated, and understanding the effects of those factors is less obvious. But in general, introduction of lossy elements increases stability and decreases gain.

In terms of s-parameters, the following equations may be used to calculate the stability factor *k*, the maximum stable gain, and the maximum available gain [37-40]:

$$k = \frac{1 + \left|s_{11}s_{22} - s_{21}s_{12}\right| - \left|s_{11}\right|^2 - \left|s_{22}\right|^2}{2\left|s_{21}s_{12}\right|} \tag{3.22}$$

$$MSG = \left|\frac{s_{21}}{s_{12}}\right| \tag{3.23}$$

$$MAG = \left|\frac{s_{21}}{s_{12}}\right|\left(k - \sqrt{k^2 - 1}\right) \tag{3.24}$$

Since the calculations for MAG and MSG require all four s-parameters, the maximum gain plot is a good composite indicator of device characteristics. Like the individual s-

parameter curves, the maximum gain curves of devices with different widths would also overlap if the devices scaled properly with width. But as shown in Figure 3.1.13, the curves for the measured devices show greater decreases in gain at higher frequencies for the larger devices. In going from 4mm to 64mm, the gain at 500MHz decreases from 22.5dB to 20dB, and the corner frequency decreases from 3GHz to 600MHz.

It is important to note here that although the small-signal measurement results show some performance degradation with increased width, there is no guarantee that the large-signal performance degradation seen in Figure 3.0.4 is also due to the same cause.

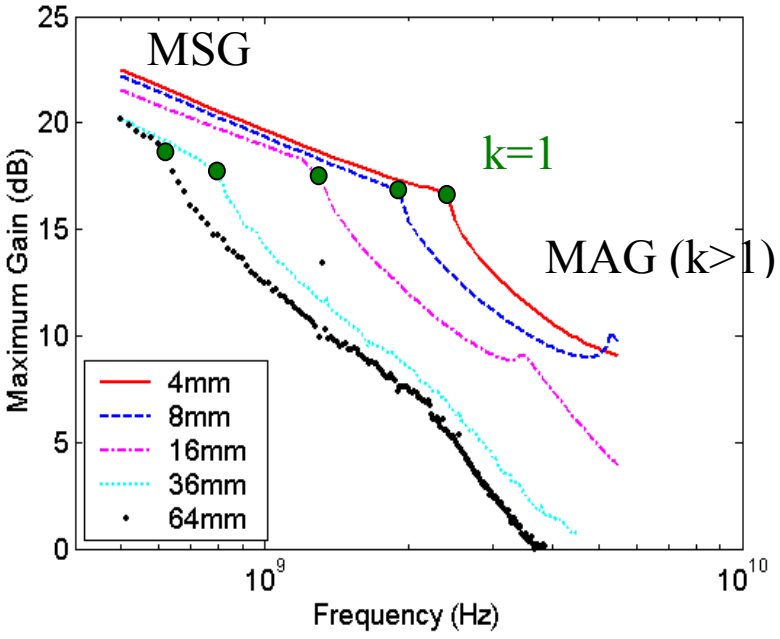


Figure 3.1.13. The maximum gain plots for different widths, calculated from the data shown in Figure 3.1.11. As W increases, the low-frequency gain decreases, and the cut-off frequency decreases.

3.1.5 Large-Signal RF Measurements

In order to characterize the maximum large-signal performance of a device, the measurements must be taken after optimizing the source and load impedances. Because of the non-linearities of the devices, the source and load impedances that allow delivery of maximum power are not necessarily the complex conjugates of the device input and output impedances, respectively. While complex conjugate impedance match allows maximum power transfer in the small-signal regime, the optimal impedance match is often determined empirically for large-signal operation, as factors such as bias conditions and input power also affect the optimal impedances. In addition, different optimal input impedances exist for different output impedances, and vice versa, necessitating the simultaneous determination of optimal terminal impedances. These measurements to determine the optimal input and output impedances are called source-pull and load-pull measurements, where source-pull measures the device characteristics for different source impedances given a constant load, and load-pull measures the device characteristics for different load impedances given a constant source impedance.

These measurements to determine device characteristics under optimal terminal impedances were taken on devices with W=4mm, 8mm, 16mm, 36mm, and 64mm using the Automated Tuner System from Maury Microwaves [41]. Figure 3.1.14 shows the measurement setup, which consists of a signal generator, power supplies, impedance tuners, a controller for the tuners, and a power meter. The gate bias was set to 1.2V, the drain bias to 3.5V, the operating frequency to 900MHz, and the input power to 7dBm/mm device width. After initial measurements, it was discovered that the optimal source impedances were relatively insensitive to changes in load impedance—this

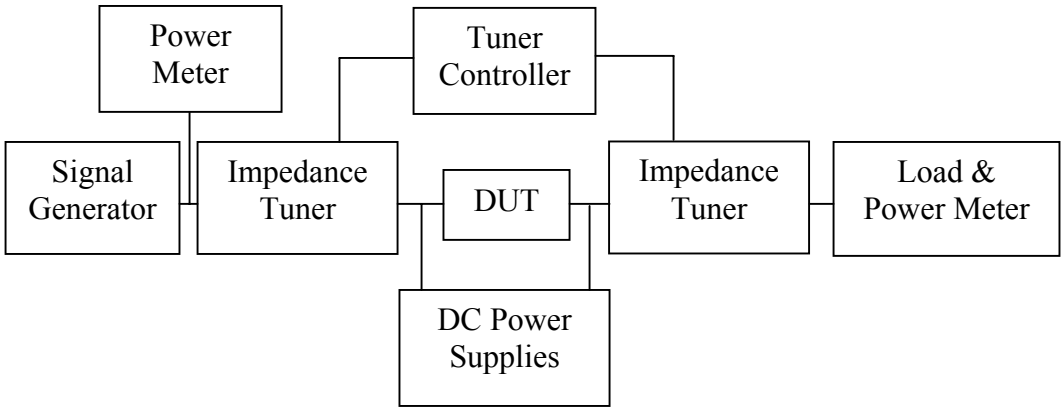


Figure 3.1.14. Load-pull measurement setup. The key components are the impedance tuners at the input and output. The tuners change the source and load impedances to determine the change in system performance at different impedances.

W (mm)	Z_o (Ω)	Z_o (Ω-mm)	PAE (%)	Output Power (mW)	Output Power (dBm)	Power Gain (dB)
4	12.5+11.3j	50+45j	76.0	124.6	20.96	7.96
8	8.13+3.75j	65+30j	78.3	202.5	23.06	7.06
16	5-1.25j	80-20j	75.6	340.3	25.32	6.32
36	2.5-5.56j	90-200j	63.2	524.3	27.20	4.65
64	2.34-5.47j	150-350j	46.7	688.0	28.38	3.38

Table 3.1.1. Load-pull measurement results for devices with different W. As W increases, the optimal load resistance $\mathrm{Re}\{Z_o\}$ decreases, but not proportionally with W, as the normalized Z_o shows. Also, the maximum PAE and the power gains at the optimal impedances decrease.

simplified the optimization process, as once the optimal source impedance was determined through a source-pull with a 50Ω load, that source impedance could then be used for the load-pull measurement. The measurement results are summarized in Table 3.1.1, where the maximum PAE, the optimal load impedance, and the output power at this impedance, are shown for each W. The maximum PAE contours, obtained through measuring the output powers at different load impedances, are shown in Figures 3.1.15 and 3.1.16. On the Smith Chart in Figure 3.1.15, the contours move in the negative reactance direction with increasing width, and at large widths, the contours shrink in size due to their location at high impedances. Plotting the same data on a width-normalized

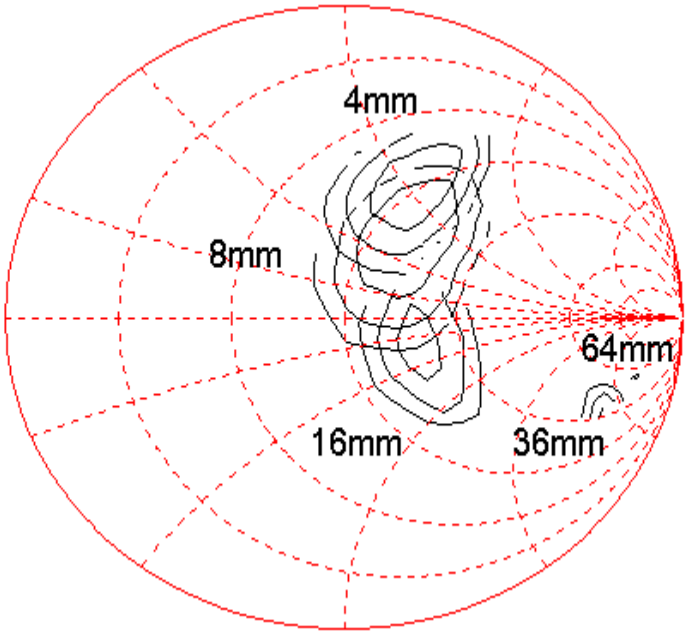


Figure 3.1.15. Measured load-pull contours for L4B MOSFETs on a width-normalized Smith Chart. As W increases, the contours get compressed in the high impedance region. The contours also move in the negative reactance and positive resistance direction.

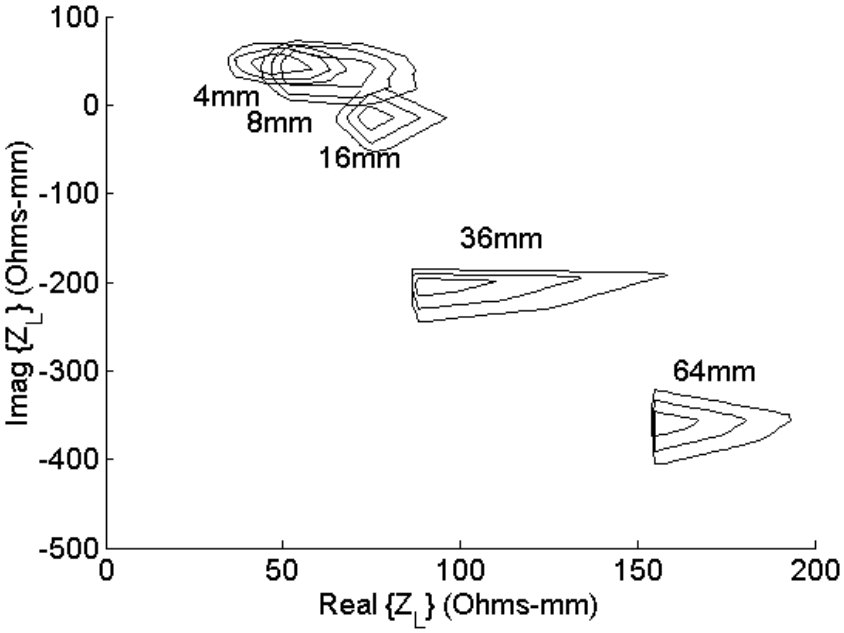


Figure 3.1.16. Width-normalized, measured load-pull contours for various device widths under maximum PAE tuning. It can be seen that the real part of the impedance increases with W, and the imaginary part decreases with W.

complex impedance plane gives Figure 3.1.16. As W increases, the contours move in the negative reactance direction and the positive resistance direction.

3.1.6 Large-Signal RF Simulations

Further characterization in the RF regime utilized the Root MOSFET models that were extracted using the HP85124A Pulsed Modeling System from Agilent Technologies [42-45]. Root models are derived from small-signal s-parameters for use in large-signal RF simulations. The pulsed system allows measurement of device s-parameters at higher gate and drain biases without fear of destroying the device, thus allowing accurate simulations over a larger domain. Also, the pulsed small-signal measurements prevent electrothermal self-heating effects from affecting the measurements, since the small signal amplitudes and the short pulse durations minimize power dissipation in the device. Models were generated for devices at three gate widths (4mm, 8mm, and 16mm), and these models were then used to perform load-pull analyses in the circuit simulator ADS from Agilent Technologies [45].

The input power density to the devices was set to 7dBm/mm device width, the gate bias to 1.2V, drain bias to 3.5V, and the frequency to 900MHz. After optimizing the impedance of the generator for maximum PAE, the results in Table 3.1.2 were obtained.

W (mm)	Max. PAE (%)	Output Power/Width @ Max. PAE (dBm/mm)	Power Gain @Max. PAE (dB)
4	70.4	18.9	11.9
8	70.0	18.9	11.9
16	80.4	18.7	11.7

Table 3.1.2. Electrical performance results of load-pull simulations for different devices when the impedances are matched for maximum PAE. Note the increase in PAE for the larger device while maintaining the same output power per width.

The table shows that as the width increases, the maximum PAE actually increases while the output power at these points scale with W, contrary to experimental observations. Figure 3.1.17 shows the width-normalized PAE contours on a Smith Chart, where each circle correspond to a constant PAE, in 1% steps from the maximum. The top set of circles belong to the W=4mm model, the middle to the W=8mm model, and the bottom set to the W=16mm model. It should be noted here that the contours consist of similarly sized circles, which move in the negative reactance direction while maintaining a similar resistance. Figure 3.1.18 shows the same contours on a complex impedance plane. It is more evident how the optimal load resistance stays constant while the optimal load reactance decreases with increasing W. Between the 4mm and 8mm contours, the movement is about j40 Ω, and between the 8mm and 16mm contours, the movement is

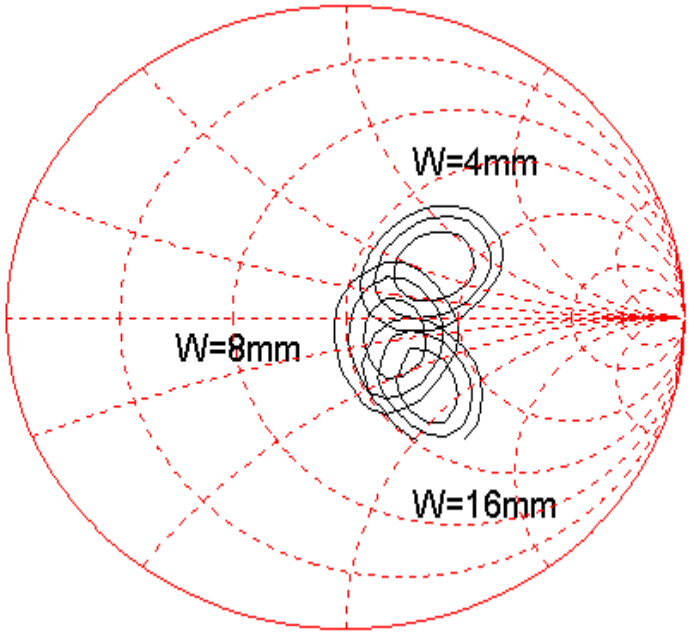


Figure 3.1.17. PAE contours (1 % per circle) drawn on a Smith Chart for W=4mm, 8mm, and 16mm devices, normalized to W=1mm. The contours are separated along the constant resistance circle by about j0.8*Z_o Ω while retaining similarly sized circles.

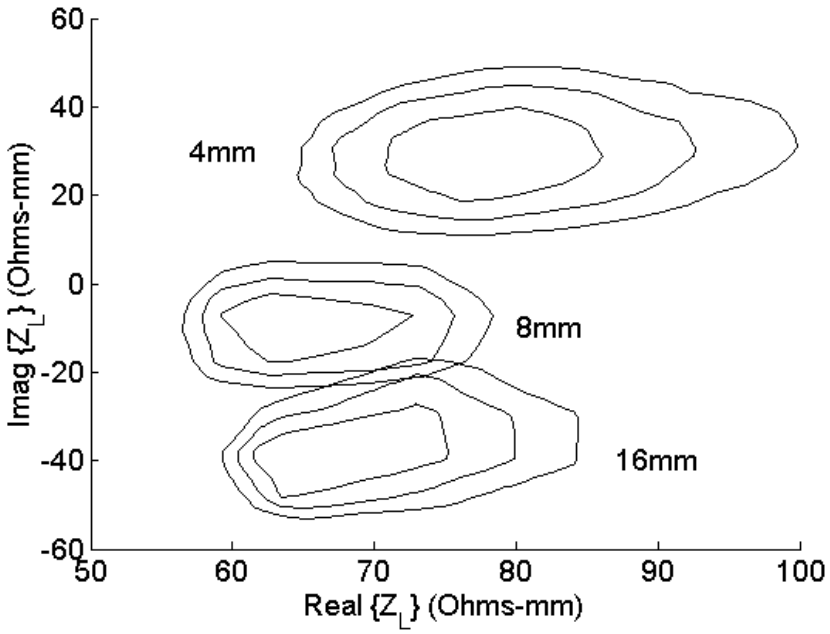


Figure 3.1.18. Width-normalized PAE load-pull contours of Figure 3.1.17 on a complex impedance plane for W=4mm, 8mm, and 16mm Root models, when impedances are matched for maximum PAE. As W is increased, the contours move in the negative reactance direction.

W (mm)	Max. Output Power/Width (dBm/mm)	Max. Power Gain (dB)	PAE @ Max. Output Power (%)
4	21.2	14.2	53.0
8	21.1	14.1	53.9
16	21.0	14.0	62.6

Table 3.1.3. Electrical performance results of load-pull simulations for different devices when the impedances are matched for maximum output power. Like Table 3.1.2, the PAE increases for the larger device while maintaining the same maximum output power per width.

about j30 Ω.

When the generator impedance was optimized for maximum output power, results as shown in Table 3.1.3, Figure 3.1.19, and Figure 3.1.20 were obtained. Table 3.1.3 shows that as the device width increases, the maximum output power per width stays constant with an improving PAE, which is again contrary to experimental observations. Figure 3.1.19 shows the width-normalized output power contours on a Smith Chart for the three models, with 0.1dB steps from the maximum. Again the contours belong to

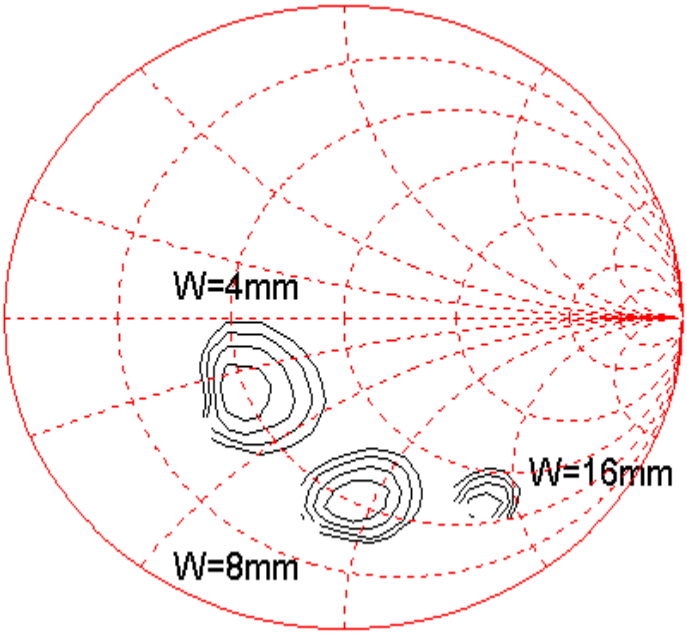


Figure 3.1.19. Output power contours (0.1 dB per circle) on a Smith Chart for W=4mm, 8mm, and 16mm devices. Again, the contours move by about j0.8*Z_o Ω each time. The contour for W=16mm look smaller because the impedance scale is compressed at the right side of the plot where higher impedances are represented.

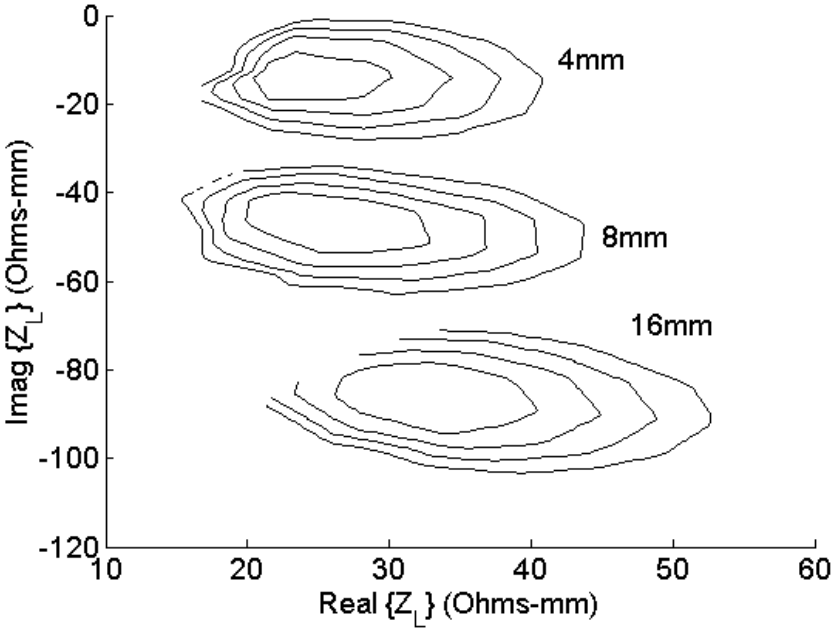


Figure 3.1.20. Width-normalized output power load-pull contours of Figure 3.1.19 on a complex impedance plane for W=4mm, 8mm, and 16mm Root models, when impedances are matched for maximum output power. Like Figure 3.1.18, as W is increased, the contours move in the negative reactance direction, with a slight increase in the resistance.

W=4mm, 8mm, and 16mm from the top set of circles to the bottom set. The incomplete circles for the W=16mm contours are due to the difficulty in simulating the devices loaded with very small impedances. The contours decrease in size on the Smith Chart plot, but this is due to the compression of the impedance scale as one moves towards the right hand side of the chart where higher impedances are represented. On the complex impedance plane, these contours look as shown in Figure 3.1.20. Like the PAE contours in Figures 3.1.17 and 3.1.18, these contours move in the negative reactance direction by j40Ω, then j30Ω with increasing width. There is a slight movement towards higher resistance as well, by 10Ω from W=4mm to W=16mm.

3.1.7 Section Summary

In summary, the above electrical measurements of the LDMOS power amplifier devices show that:

1) The width-normalized, isothermal DC currents stay constant with increased width.
2) When the temperature changes, the IV characteristics also change.
3) The width-normalized device capacitances stay constant with increased width.
4) The device input and output resistance and reactance increase with increased width.
5) The small-signal gain decreases with increased width.
6) The optimal load resistance increases and the optimal load reactance decreases with increased width.

7) The constant output power contours over the load impedance plane become larger, especially in the resistance direction, with increased width.

These results, combined with further experimentation, will aid in determining the cause behind the performance degradation observed with increased device width.

3.2 Identifying the Cause of Performance Degradation

Using the insight obtained through electrical measurements in the previous section, possible causes for the degradation were hypothesized, and experiments have been designed to test these hypotheses. The three hypotheses developed to explain the performance degradation of the larger devices are:

1) Electrothermal Effects
2) Losses in the Matching Networks
3) Internal Device Effects

In theory, all three of these hypotheses could yield lower performance with increased device width, as will be shown in the following sections. But a combination of measurements and simulations were performed such that their results either support or discredit each hypothesis.

3.2.1 Electrothermal Effects

Transistor characteristics change with temperature, as shown in the DC IV plots in Section 3.1.1. The LDMOS transistors, while delivering power to the antennae, also consume power, which is dissipated on the chip as heat. Because silicon is not a very good thermal conductor compared to the metal package, the temperature on-chip is

higher compared to the package and to the ambient. As the transistor channel width increases, the physical area of the chip does not necessarily increase proportionally, and the thermal resistance per channel width increases. The higher thermal resistance, under constant heat generation, would lead to higher temperatures. Semiconductor physics indicate that higher temperature yields higher carrier concentration, lower carrier mobility, and lower threshold voltage, thereby altering performance [31, 46, 47]. But in quantifying the electrical performance degradation due to thermal characteristics, the temperature rise was first quantified through the use of the finite-element simulator ANSYS. From the electrical efficiency data, the amount of heat generated by the device was computed. Calculations yielded higher heat generation for devices with larger widths due to the lower electrical efficiencies. This information, along with the geometry of the device, was entered into ANSYS to generate the thermal distribution map of the surface of the device. It was assumed that the heat generation occurs uniformly in the active region at the surface of the device, as suggested in [47]. These thermal simulations were performed for models of devices with W=16mm, 36mm, and 64mm. The geometries and resulting thermal maps are shown in Figures 3.2.1-3.2.3. In the figures, the box represents the silicon device, which was placed on a large copper block emulating the package. The temperature at the bottom of the block was held at 300K, and the heat was applied uniformly to the small white boxes on the chip surface. For the 16mm device, the maximum temperature was 306K, and the temperature at the edge of the active region was 303K. For the 36mm device, the maximum was 316K, and the edge temperature was 308K. For the 64mm device, the maximum was 321K, and the edge temperature was 311K. This information is summarized in Table 3.2.1. Thus it can

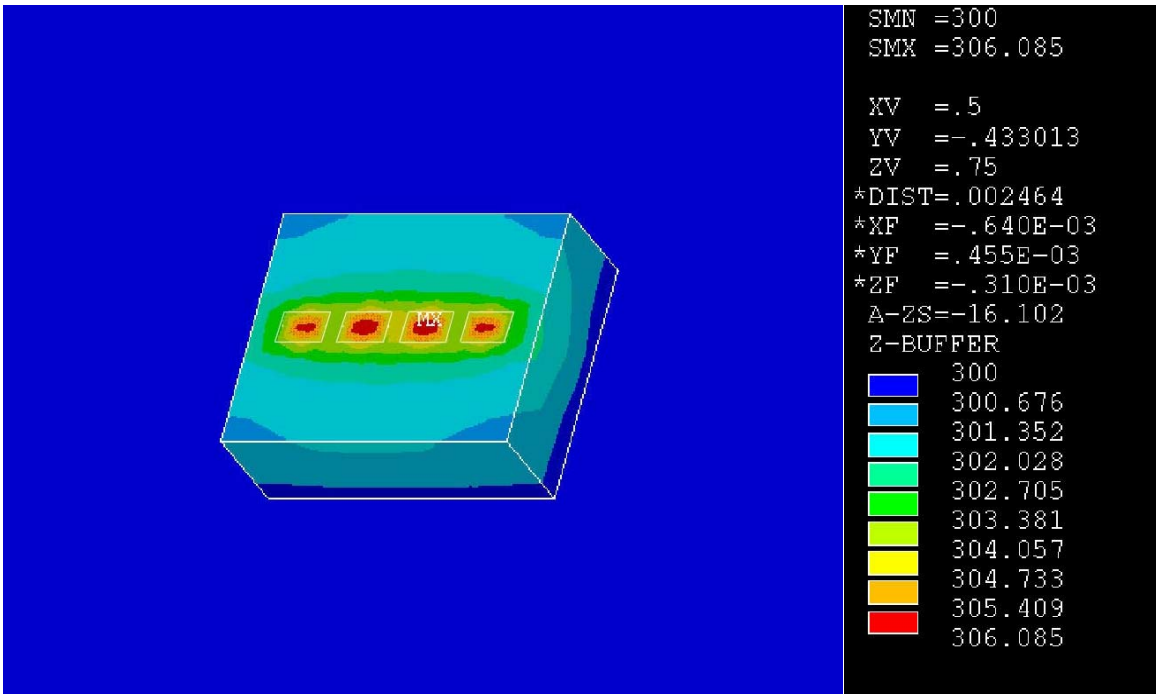


Figure 3.2.1. Thermal map of a W=16mm device. The 4 boxes show the active regions for the 4 unit cells which compose this device. Note the maximum temperature of 6 degrees above ambient, and the temperature variation in each box of about 3 degrees.

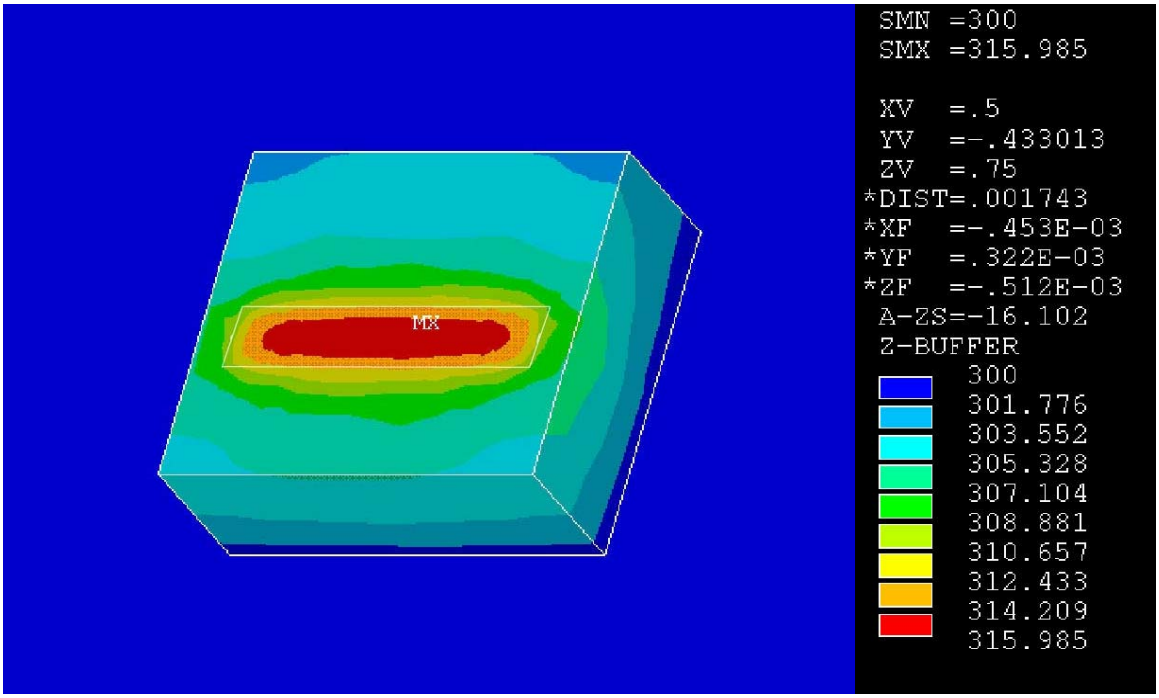


Figure 3.2.2. Thermal map of a W=36mm device. Again, the box shows the active region. This time, the maximum temperature is about 16 degrees above ambient, and although the central unit cells seem to have identical temperatures, the outer unit cells are about 7 degrees cooler.

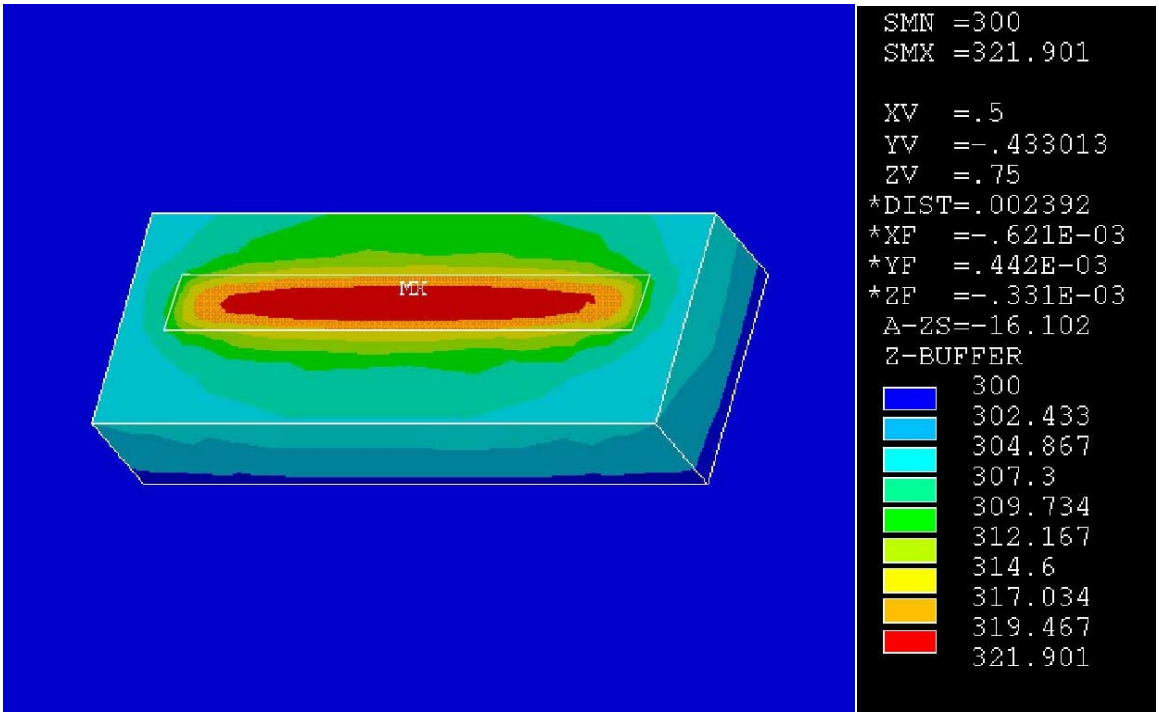


Figure 3.2.3. Thermal map of a W=64mm device. The box shows the active region. The maximum temperature is 22 degrees above ambient, and the outer unit cells are about 10 degrees cooler than the central unit cells.

W (mm)	T_{center} (K)	T_{edge} (K)	ΔT (K)
16	306	303	3
36	316	308	8
64	321	311	10

Table 3.2.1. Chip surface temperatures from ANSYS simulations. The larger devices have higher overall temperatures due to larger power dissipation, and have larger temperature differences between the central region and the edge of the chip.

be seen that the larger devices have higher peak temperatures, the centers of the devices are warmer than the edges, and the temperature difference between the center and the edge of the active region increases with device width.

To determine the effects of heat on operating conditions, DC IV traces were taken on a device at different temperatures as already shown in Figure 3.1.4. It can be seen that the curves compress in the vertical direction as drain currents decrease at higher temperatures. This compression in the drain currents indicates smaller AC current

swings, and thus smaller output powers. Also, when a load-pull analysis based on load lines as suggested by Cripps [48] is applied to the higher temperature curves, the load lines must be flatter due to the lower drain currents, and therefore the optimal load resistance seen at the drain terminal must be higher than that of the lower temperature IV curves, as illustrated in Figure 3.2.4. This agrees with the observed optimal load impedance from the load-pull measurements shown in Figures 3.1.15-3.1.18, and coincides with previous work performed on GaAsFETs [49]. However, this analysis does not quantify how this change in optimal load impedance affects performance in terms of output power.

Experimentally, electrical measurements of a 16mm device were taken at different temperatures by heating the chuck with a thermoelectrical heater. If higher temperatures are indeed responsible for the electrical performance degradation, taking

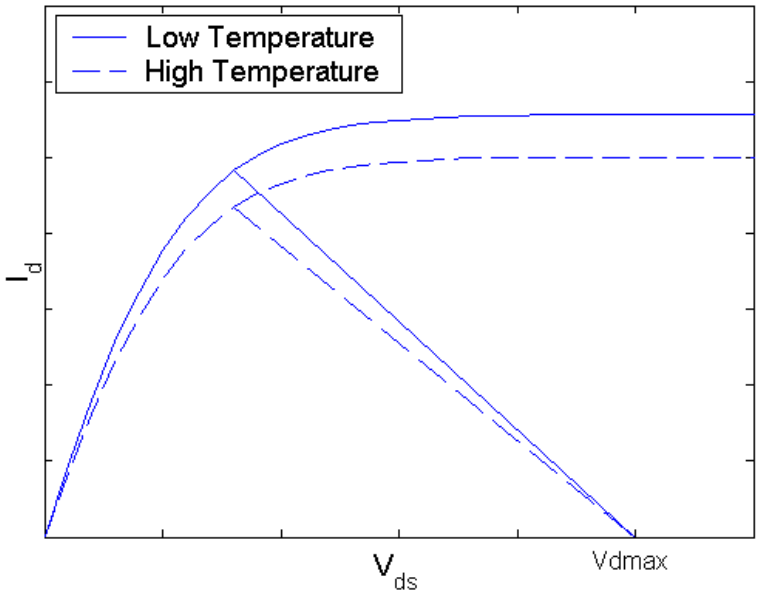


Figure 3.2.4. IV plot illustrating the effect of higher temperature on loadlines. When the IV curves become compressed vertically, the optimal loadline also flattens, resulting in an increase in the optimal load resistance.

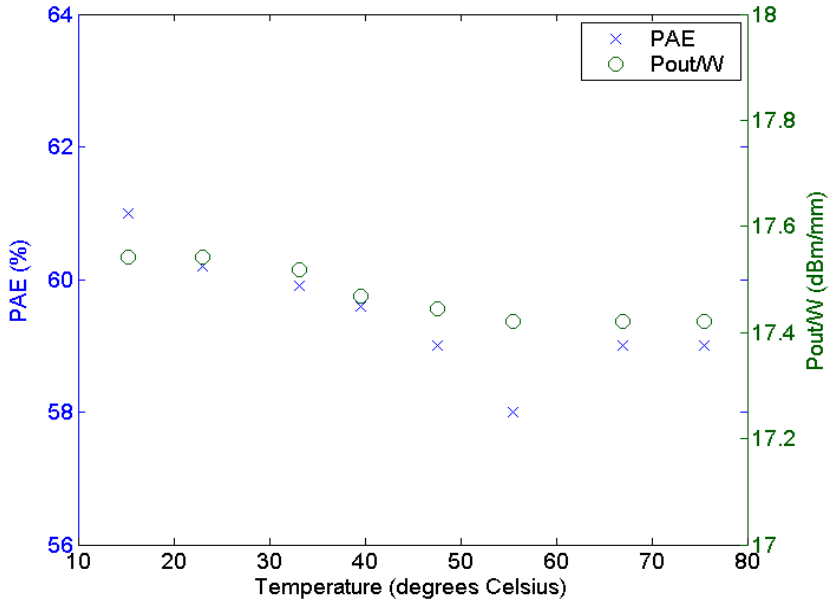


Figure 3.2.5. Electrothermal measurement results for W=16mm device. As the temperature of the device is increased, the PAE and output power per width do not change significantly.

measurements of a small device (in this case a W=16mm device) while heating it should result in lower performance. However, the results in Figure 3.2.5 show that even at temperatures significantly higher than those yielded by the simulator, the electrical efficiency and power output of the device do not decrease significantly.

Another aspect of the thermal effect is the non-uniform temperature at the surface of the device. In examining the active region of the device and its thermal contours, Figures 3.2.1-3.2.3 show that the central area of the device is warmer than the outer parts. This was also shown in [47]. The non-uniformity of the temperature distribution can cause variations in operating conditions between unit cells. It is known that the optimal load impedance for a unit cell is different at different temperatures [49]. If there are many unit cells, all at different temperatures, it is impossible to present a single load impedance that optimally matches to all of the unit cells. If this were the case, as device

width increases, the electrical performance would decrease due to the increase in the temperature gradient across the chip.

However, as the device width increases, the temperature gradient is largely limited to the edges of the chip. This would suggest that as the device width increases, the effect of non-uniform temperature distribution becomes less significant, as the proportion of the active region that has the same temperature increases with device width.

Through these experiments, it was concluded that although higher temperatures and larger temperature gradients are observed in the larger devices, the decreased electrical performance could not be attributed to the thermal effects. Within the temperature range in which these devices are expected to operate, measurements have shown that the performance does not decrease significantly. Also, the on-chip temperature gradient is such that a large portion of the central region of the larger devices has the same temperature, and significant gradients are observed only at the edges of the devices. This suggests near-uniform operating conditions throughout most of the device, and it is therefore unlikely that the thermal gradient at the edges cause significant performance degradation.

3.2.2 Matching Network

The second possible cause for the lower performance observed in larger devices is increasing power loss in the matching networks while taking power measurements. The results from the load-pull measurements are summarized in Table 3.1.1, and will be used as a starting point for this investigation.

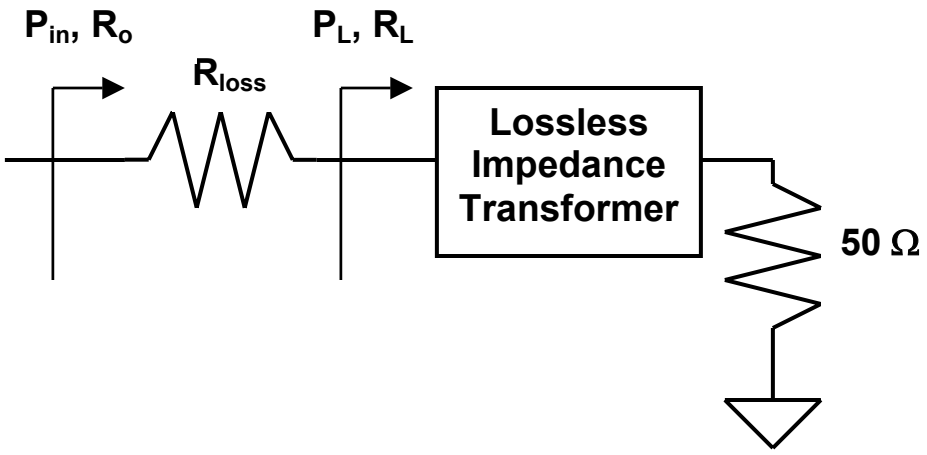


Figure 3.2.6. Simplified model of the output matching network to determine the width-dependent loss. The transistor sees the resistance R_o, and delivers P_i to the network. After losing some power in R_{loss}, the output power, P_L is delivered to the transformed load resistance R_L.

Assuming there is a lossy element in the output matching network, the simple circuit shown in Figure 3.2.6 is used to represent the matching network. The resistance looking into this circuit, R_o, corresponds to the optimal load resistance shown in Table 3.1.1. The loss element, R_{loss}, is in series with the transformed 50-ohm load resistance (R_L). The series combination of R_{loss} and R_L should yield the equivalent resistance R_o, as shown in equation (3.25).

$$R_o = R_{loss} + R_L \tag{3.25}$$

The observed output power is dissipated in R_L, and there is some loss associated with R_{loss}, so let the input power P_i be the sum of the output power and the loss. Iterative calculations show that when R_{loss} is constant and R_L approximately scales with W, in order for equation (3.25) to hold using the data in Table 3.1.1, R_{loss} is around 1.5Ω.

If we now look again at Table 3.1.1, at W=4mm, the observed output power is 21.0dBm (125mW) and R_o=12.5Ω. Then according to the equation (3.26) which shows

W (mm)	R_o (Ω)	R_{loss} (Ω)	R_L (Ω)	P_i (mW)	P_L (mW)
4	12.5	1.5	11	142	125
8	8.13	1.5	6.6	284	230
16	5	1.5	3.5	568	392
36	2.5	1.5	1.0	1278	486
64	2.34	1.5	0.84	2272	768

Table 3.2.2. Calculation results from using the circuit in Figure 3.2.6 to calculate the output power for different W. This assumes the transistor scales, and that there is a constant R_{loss} in the matching network. Thus even though R_L approximately scales with W, the power delivered to the load does not, as R_{loss} becomes more significant with respect to R_L at larger W.

the relationship between P_i and the power delivered to R_L, R_L=11Ω, P_i=142mW, and the lost power is 17mW.

$$P_L = P_i \frac{R_L}{R_L + R_{loss}} \tag{3.26}$$

If the transistor scales to W=8mm, P_i should be 284mW. But with an observed R_o=8.13Ω, and R_{loss} still 1.5Ω, the delivered power would be 232mW (23.6dBm), and the loss is 52mW. Note here that while the device size doubled, the power lost has more than doubled. If the calculations are repeated for W=16mm, 36mm, and 64mm, the results shown in Table 3.2.2 are obtained. Table 3.2.2 shows that if a lossy element (R_{loss}) exists, even if the power from the transistor (P_i) scales with W, the output power does not scale. Figure 3.2.7 shows these output powers plotted against W. Also shown in Figure 3.2.7 is the plot of measured output power against W from Table 3.1.1. A comparison of the two curves shows a marked resemblance, both in value and shape.

While this investigation has examined the output matching network, the same logic could be applied to the input matching network. It has been experimentally observed that measures to increase the input impedance of the device, which would decrease the fraction of lost power to delivered power, have improved the system

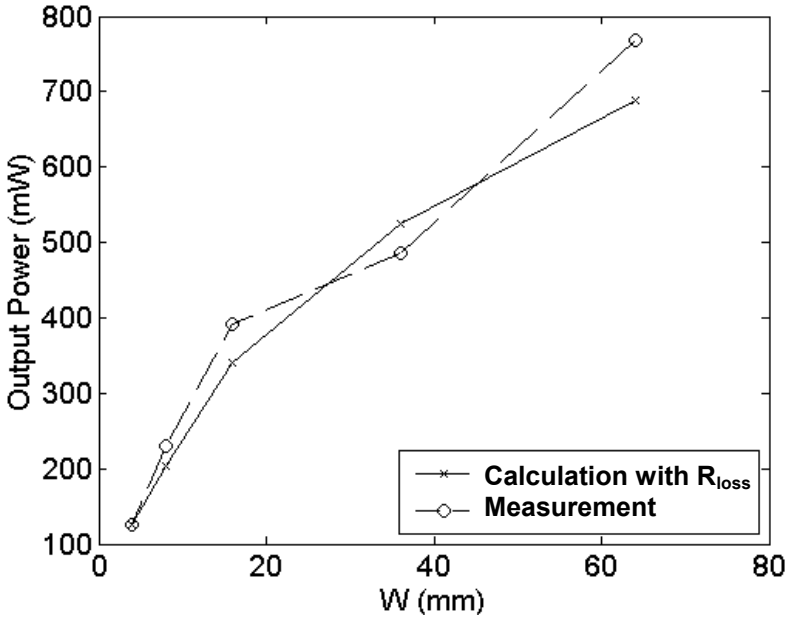


Figure 3.2.7. Measured output power as seen in Table 3.2.2, compared to the results of calculating loss using the circuit in Figure 3.2.6. With R_{loss} of 1.5 Ω, the output power curves look similar to each other.

performance [50]. Also, it may be argued that matching network loss is prevented through the use of power dividers/combiners as shown in [28] by successfully raising the impedance of the devices as seen by each matching network.

But it is unlikely that a series resistance as large as 1.5Ω exists in the matching network. The load-pull measurement results in Section 3.1.5 demonstrated the lower performance of larger devices after having the matching network de-embedded—the calibration of the impedance transformers are performed so that the reference plane of the measurements are at the device. Since it is extremely unlikely that a resistance as large as 1.5Ω is unaccounted for after the calibration of the matching networks, it is improbable that the losses in the matching networks dominate the performance degradation in larger devices.

3.2.3 Internal Device Effects

While the previous two hypotheses have taken a system level perspective and identified potential problems that may arise due to integration of identical devices, the possibility remains that the device to be integrated may not all be identical. There is a possibility that even on the same die, different unit cells have sufficiently different characteristics. Or even if the unit cells on the die are identical, the integration of multiple unit cells on a die, and simultaneous operation of the unit cells, cause different electrical characteristics in different unit cells. In either case, the parameter that is different between unit cells must be identified. To this end, a W=4mm Root model was simulated in ADS while altering its characteristics by adding components around the model. In altering the model, the parameters in a standard hybrid-pi model as shown in Figure 3.2.8 were used. By adding external components to this model, it would behave as though some of its internal parameters have changed. After altering each parameter, the small-signal s-parameter simulation results were compared to measured s-parameter results of devices with different W, in the hopes that similar trends may be identified between the simulated and measured results. In this way, the internal component that is not scaling with width could be identified.

S-parameter simulations were run after altering only one hybrid-pi model parameter (R, C, L, g_m) at a time, by four different values each, as shown in Table 3.2.3. The complete results are shown in Appendix A, but out of the entire matrix, the plots that provide insight are reproduced as Figures 3.2.9-3.2.18. In these figures, the arrows denote the change in device characteristics as the parameter value deviates farther away from the original value. When comparing these plots to the measured s-parameter and

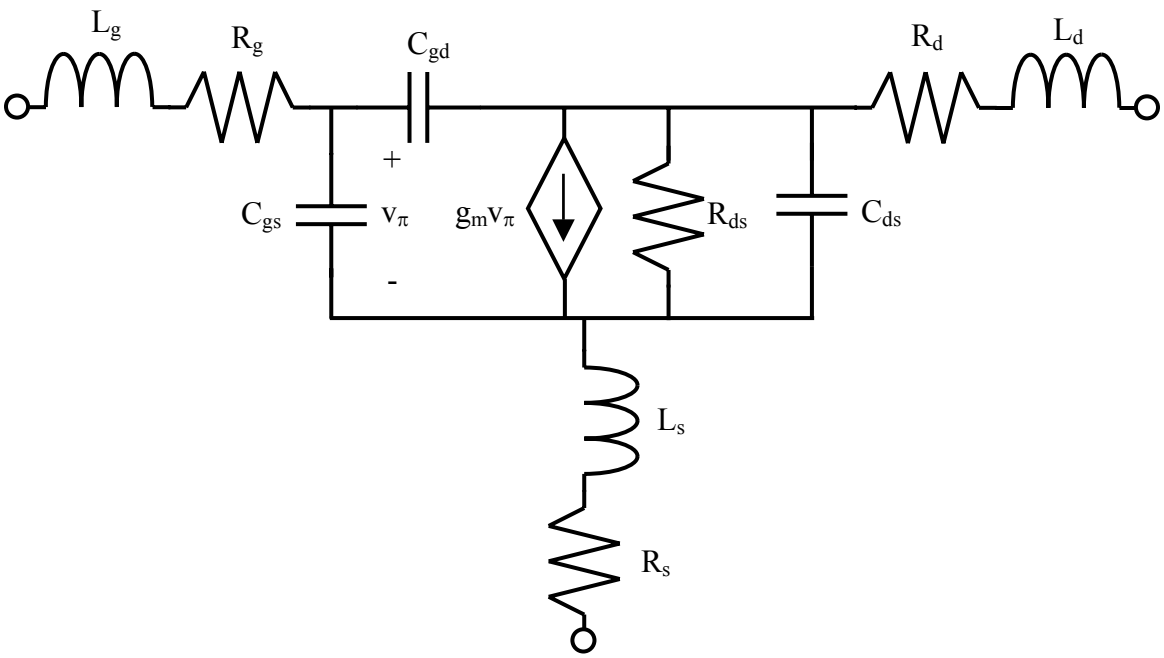


Figure 3.2.8. Hybrid-pi model for a MOSFET. The values of the parameters shown in this figure were altered to determine the cause of non-scaling performance with increased width.

Parameters	Default	Extra Parameter Values			
R_g	0	1Ω	2Ω	3Ω	4Ω
R_d	0	1Ω	2Ω	3Ω	4Ω
R_s	0	0.25Ω	0.5Ω	0.75Ω	1.0Ω
C_{gs}	0	1pF	2pF	3pF	4pF
C_{ds}	0	1pF	2pF	3pF	4pF
C_{gd}	0	50fF	100fF	150fF	200fF
L_g	0	1nH	2nH	3nH	4nH
L_d	0	1nH	2nH	3nH	4nH
L_s	0	200pH	400pH	600pH	800pH
g_m	0	-10mS	-20mS	-30mS	-40mS
R_{ds}	∞	1000Ω	500Ω	200Ω	100Ω

Table 3.2.3. Parameter values that were added to the W=4mm Root model in s-parameter simulations of Section 3.2.3. Four extra simulations per parameter were performed to observe how the s-parameters change with a change in each parameter, as shown in Figures 3.2.9-3.2.30.

maximum gain plots of Figures 3.1.11 and 3.1.13, it can be seen that the clockwise movement of s_{11} and s_{22} with increased W may be accounted for by increased L_g (Figure 3.2.9) and L_d (Figure 3.2.11), respectively. But a change in L_g and L_d does not alter the maximum gain plot at all (Figures 3.2.10 and 3.2.12) because the reactance added to the terminals may be matched out. The change in the measured maximum gain, consisting of lower gain at low frequency and lower corner frequency with increased W, can be mimicked by increased L_s (Figure 3.2.14), decreased g_m (Figure 3.2.16), and/or increased C_{gd} (Figure 3.2.18).

However, out of L_s, C_{gd}, and g_m, no decrease in g_m is detected for increased W from the DC IV curves of Figure 3.1.3. Similarly, no increase in C_{gd} is detected in the CV plot of Figure 3.1.10. Which leaves L_s as the only possible suspect parameter that remains to be investigated.

Putting the results together, it would seem that L_g, L_d, and L_s all increase with increasing W. It does not seem coincidental that the three parameters remaining are all inductances. Now, a plausible mechanism by which the inductors increase is required to support this hypothesis. In revisiting the physical structure of the devices, it can be seen that when the device width increases, there are more unit cells. The number of parallel fingers increases, and the number of bondwires increases. It is quite possible that the parallel current-carrying structures, when operated at RF, may interact with each other. This may be described in circuit terms as mutual inductance, and will be further investigated in the following chapter.

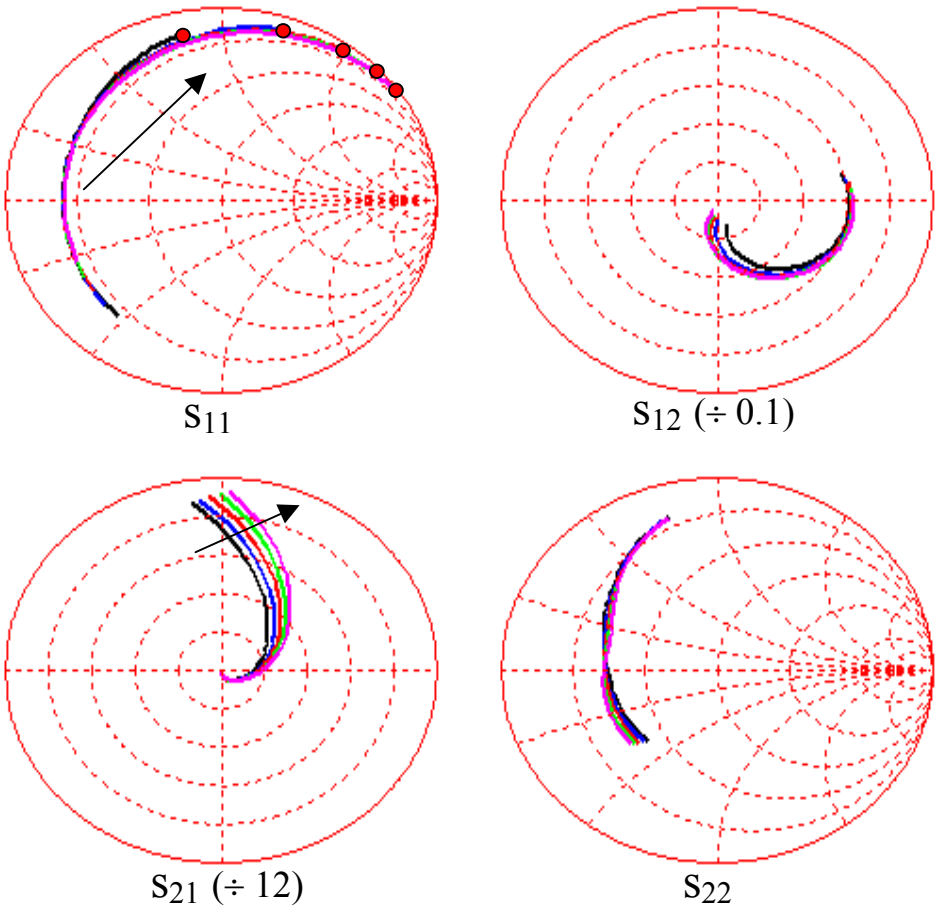


Figure 3.2.9. Small-signal s-parameters of the W=4mm Root model with extra gate inductance (L_g). An increase in forward transmission is observed in the s_{21} plot and a large increase in the input reactance is observed in the s_{11} plot as L_g is increased.

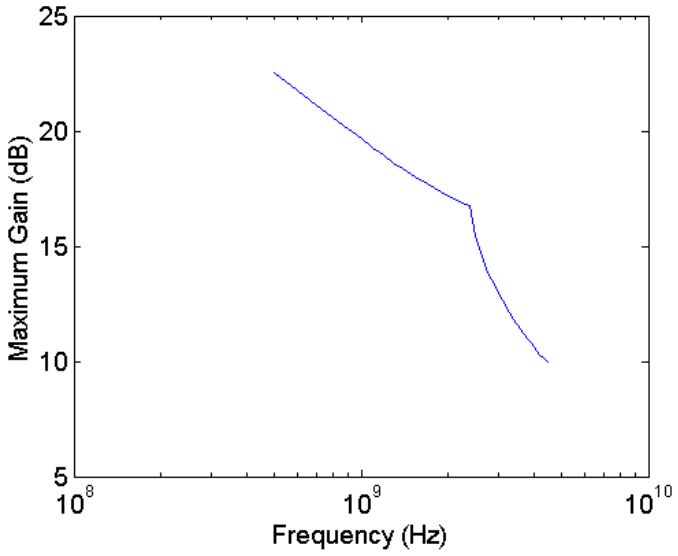


Figure 3.2.10. Maximum gain plots calculated from the data shown in Figure 3.2.9 (extra L_g). Because a change in L_g may be countered by a reactive matching circuit, no change is observed in the maximum gain plots.

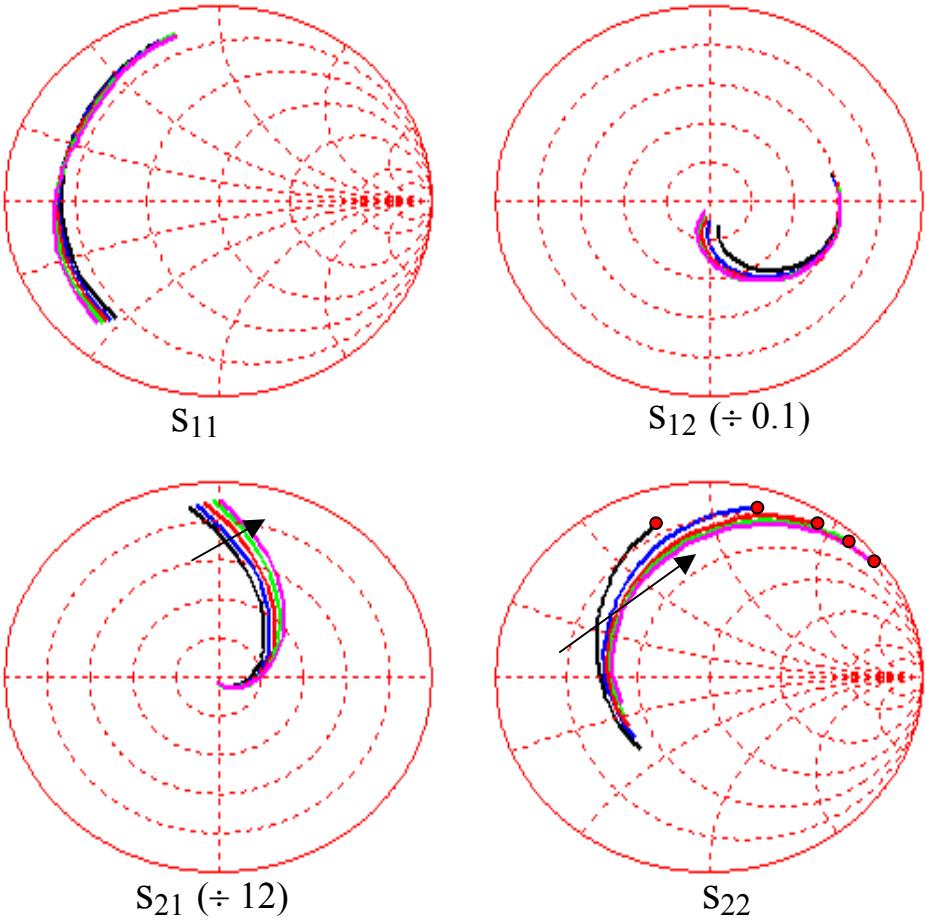


Figure 3.2.11. Small-signal s-parameters of the W=4mm Root model with extra drain inductance (L_d). An increase in forward transmission is observed in the s_{21} plot and a large increase in the output reactance is observed in the s_{22} plot, as more L_d is added.

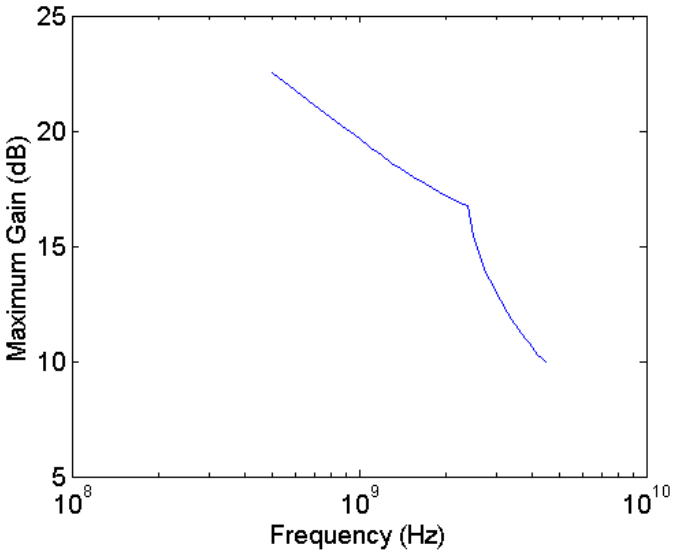


Figure 3.2.12. Maximum gain plots calculated from the data shown in Figure 3.2.11 (extra L_d). Because a change in L_d may be countered by a reactive matching circuit, no change is observed in the maximum gain plots.

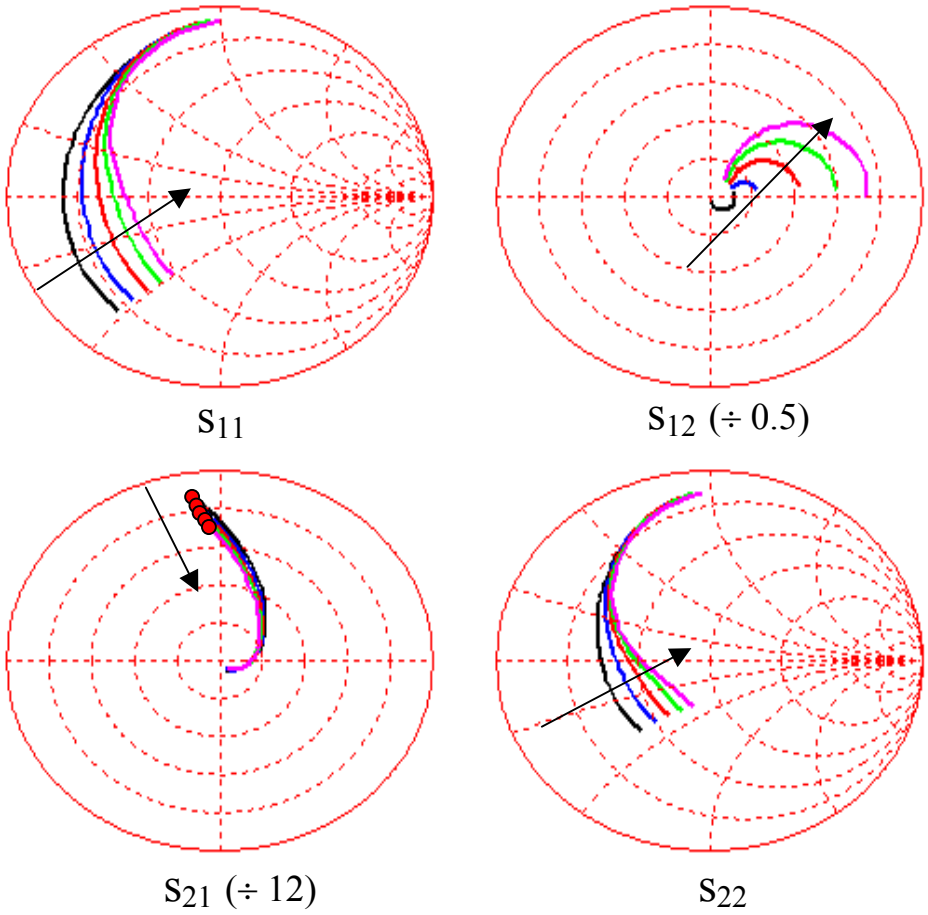


Figure 3.2.13. Small-signal s-parameters of the W=4mm Root model with extra source inductance (L_s). A decrease in forward transmission is observed in the s_{21} plot, a large increase in reverse transmission is observed in the s_{12} plot, and increases in the input and output resistances is observed in the s_{11} and s_{22} plot, respectively, as more L_s is added.

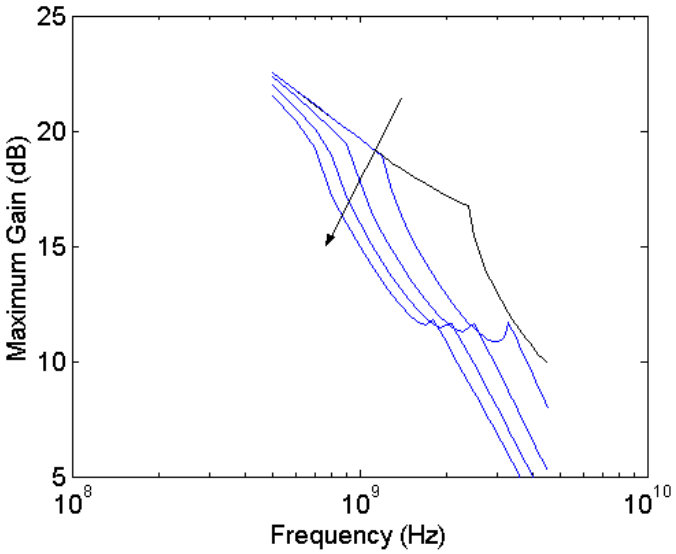


Figure 3.2.14. Maximum gain plots calculated from the data shown in Figure 3.2.13 (extra L_s). The low-frequency gain decreases, and the cut-off frequency moves drastically lower as L_s increases.

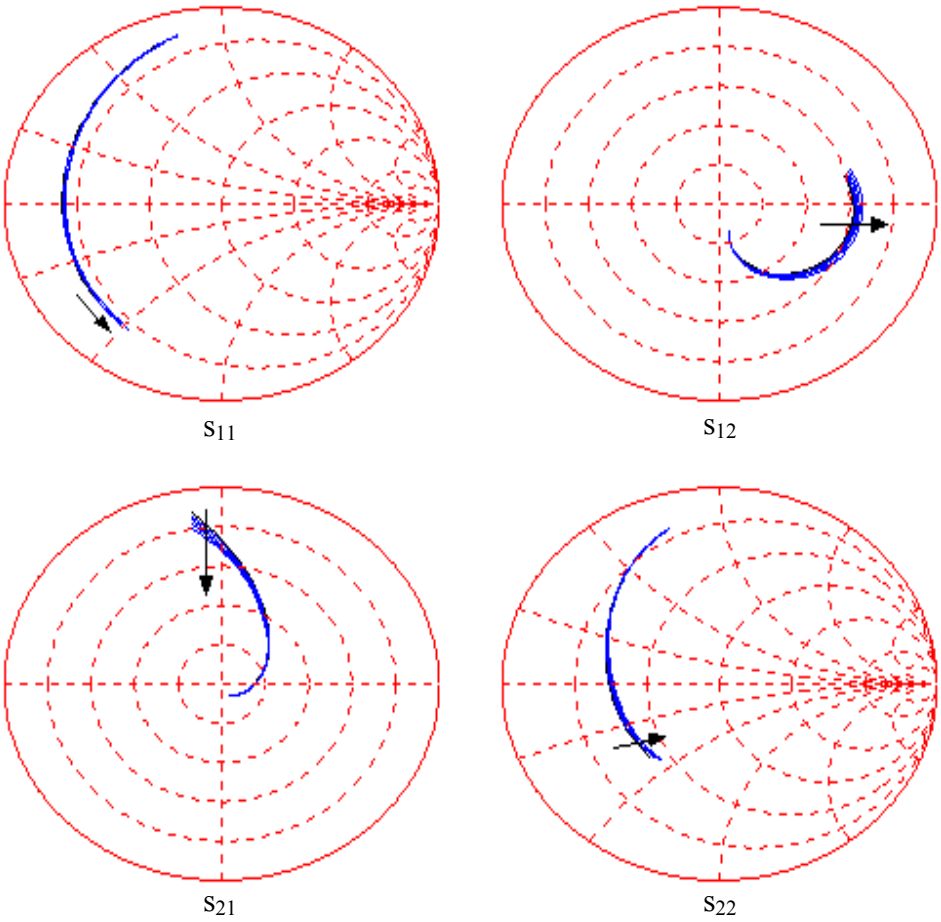


Figure 3.2.15. Small-signal s-parameters of the W=4mm Root model with decreasing transconductance (g_m). A decrease in forward transmission in s_{21}, an increase in reverse transmission in s_{12}, a decrease in the input reactance in the s_{11}, and an increase in the output resistance in s_{22} are observed as more g_m is subtracted.

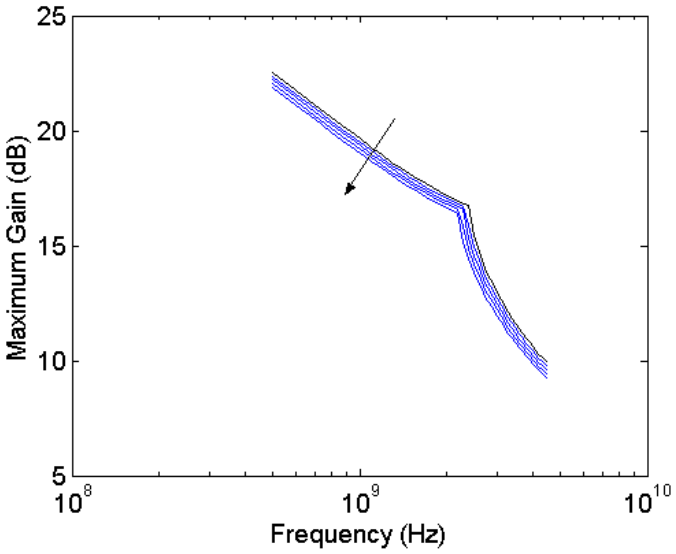


Figure 3.2.16. Maximum gain plots calculated from the data shown in Figure 3.2.15 (subtracting g_m). The curves shift downward slightly as g_m is decreased.

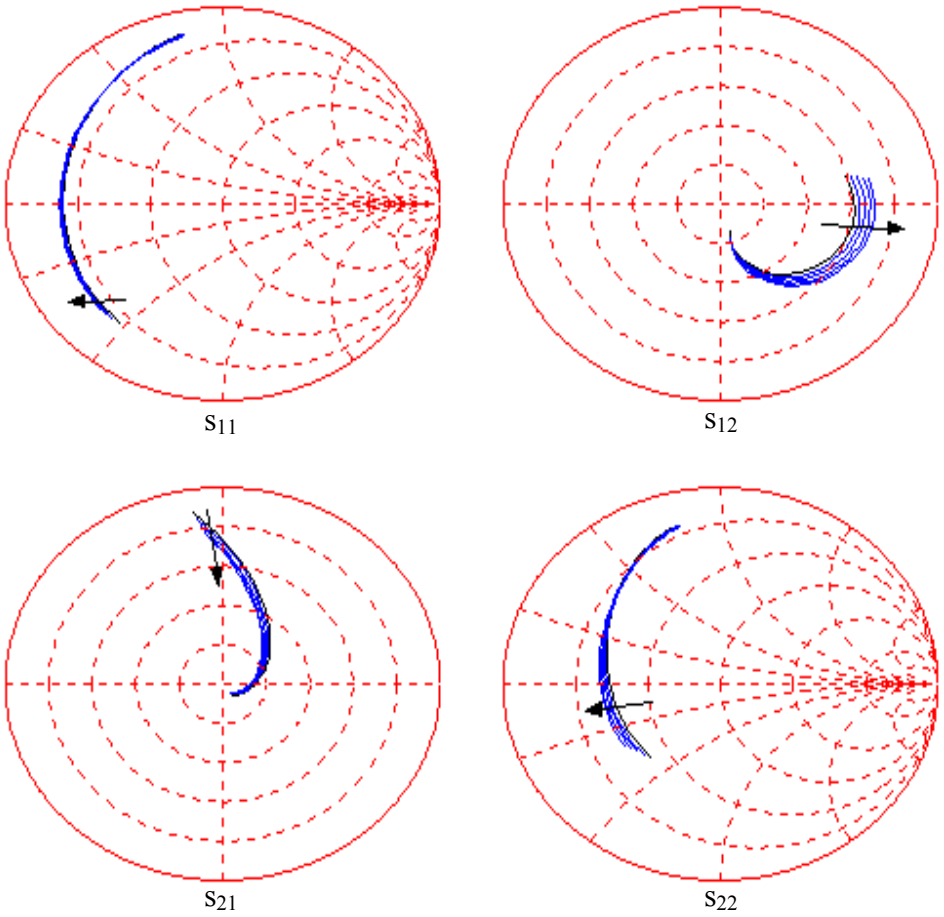


Figure 3.2.17. Small-signal s-parameters of the W=4mm Root model with extra gate-drain capacitance (C_{gd}). A decrease in forward transmission is observed in s_{21}, an increase in reverse transmission is observed in the s_{12} plot, and small decreases in the input and output resistances are observed in the s_{11} and s_{22} plots, as more C_{gd} is added.

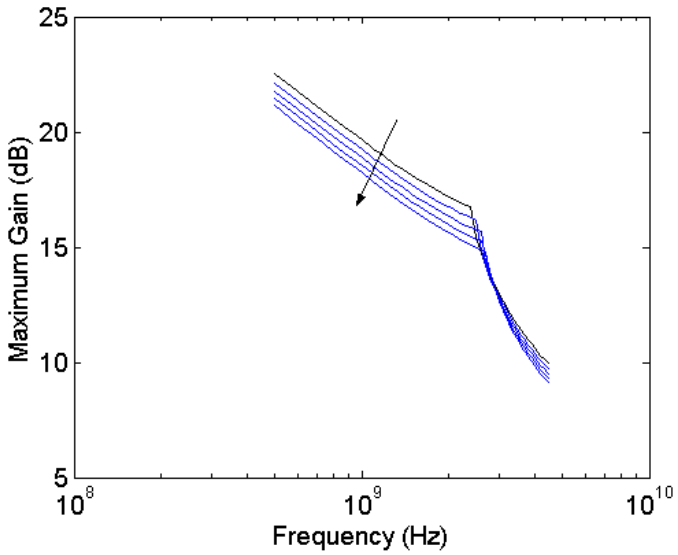
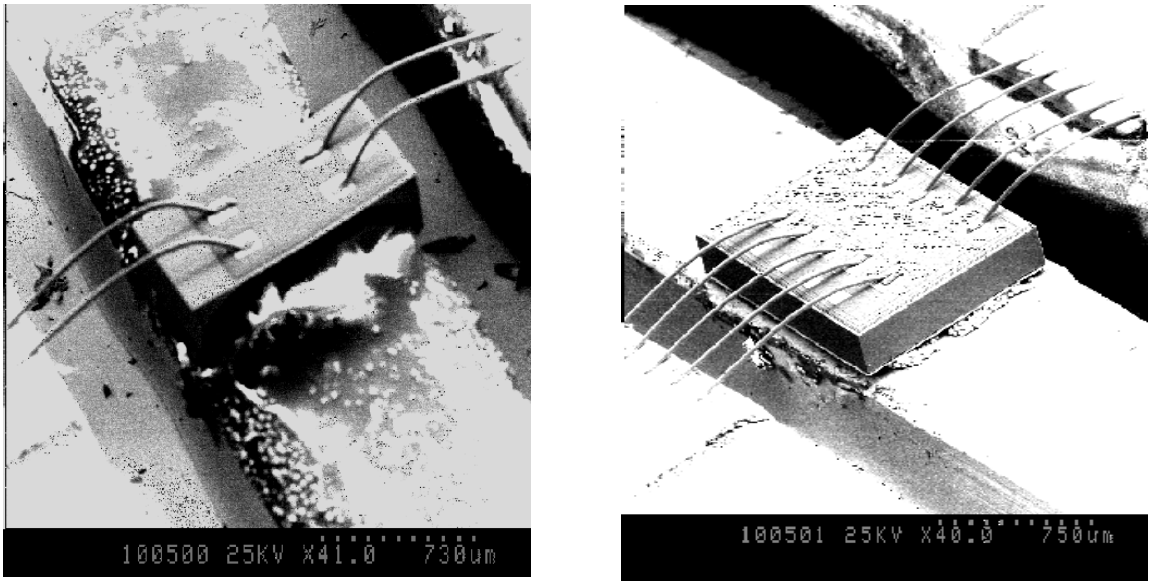


Figure 3.2.18. Maximum gain plots calculated from the data shown in Figure 3.2.17 (extra C_{gd}). The low-frequency gain decreases, and the cut-off frequency moves slightly higher as C_{gd} increases.

4 Effects of Mutual Inductance

While the previous chapter has identified a plausible cause for performance degradation in scaled power LDMOS transistors resulting from mutual inductance, this chapter focuses on examining mutual inductance in more detail. Moreover, some possible countermeasures to recover the lost performance will be considered.

In looking at the SEM photographs of Figure 4.0.1, an increase in device width consists of adding duplicate cells in parallel, thereby increasing the total number of device fingers. The microphotographs show a W=8mm device (Figure 4.0.1a), and a W=36mm device (Figure 4.0.1b), where in both cases, the terminal on the left is the gate, the terminal on the right is the drain, and the source terminal is on the bottom of the device. Figure 4.0.2 shows an illustrative diagram of the two-bondwire device (Figure 4.0.1a), including the orientation axes and the device fingers. Each finger is approximately 250 μm long, and increasing the total W consists of placing more fingers in parallel. If the parallel fingers have no interaction between them, all the fingers should



(a) (b)

Figure 4.0.1. (a) SEM photo of a device with W=8mm. The bondwire geometry is used to extract the inductance values. (b) SEM photo of a W=36mm device. The 5-wire structure should give more total mutual inductance.

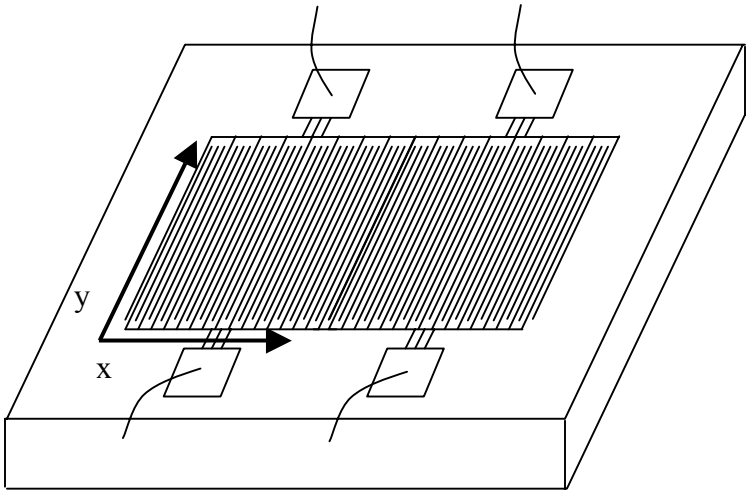


Figure 4.0.2. Drawing showing the finger structure of the two-unit cell device shown in Figure 4.0.1a. The interdigitated structure allows for a high ratio of total device width to die area.

act identically, and the system should scale proportionally with W. But interaction in the form of inductive coupling must be considered in both the bondwires and the fingers. As alternating current runs through the gate and drain bondwires, and through the gate, drain,

and source fingers, magnetic fields are generated, affecting current flow. Self-inductance, or the added impedance due to this effect, may be described as shown in equation (4.1), where l is the length of the conductor, and r is the radius of the cross-section of the conductor [51].

$$L = \frac{\mu_o l}{2\pi}\left(\ln\left(\frac{2l}{r}\right) - 0.75\right) \tag{4.1}$$

When the generated magnetic field is coupled in some other conductor, it is quantified as mutual inductance, and for parallel wires, the mutual inductance is a function of the conductor length l and separation d, as shown in equation (4.2) [51].

$$M = \frac{\mu_o l}{2\pi}\left(\ln\left(\frac{2l}{d}\right) - 1 + \frac{d}{l}\right) \tag{4.2}$$

4.1 Mutual Inductance Extraction

4.1.1 Bondwires

In order to estimate the mutual inductance, the bondwire structures of Figure 4.0.1 were entered into the bondwire geometry extracting software described in [52], and the field solver FASTHENRY [53] was used to estimate the mutual inductance between bondwires.

The following matrix equations show the inductance matrices for both the 2-wire and the 5-wire cases as generated by FASTHENRY.

$$\begin{bmatrix} L_{self1} & M_{12} \\ M_{21} & L_{self2} \end{bmatrix} = \begin{bmatrix} 1.21 & 0.45 \\ 0.45 & 1.19 \end{bmatrix}(nH) \tag{4.3}$$

$$L = \begin{bmatrix} 0.932 & 0.357 & 0.237 & 0.176 & 0.142 \\ 0.357 & 0.949 & 0.360 & 0.239 & 0.182 \\ 0.237 & 0.360 & 0.959 & 0.371 & 0.249 \\ 0.175 & 0.239 & 0.372 & 0.917 & 0.361 \\ 0.141 & 0.182 & 0.269 & 0.361 & 0.947 \end{bmatrix} (nH) \tag{4.4}$$

As can be seen in the left side of equation (4.3), the major diagonal represents the self-inductance, and the mutual inductance appearing off of the major diagonal decreases as the separation between the wires increase. The first matrix shows that both the self and mutual inductances of the two wires are very similar. The symmetry across the diagonal is expected, as the mutual effects should be reciprocal between wires. The second matrix equation, equation (4.4), also shows the symmetry across the diagonal, and when the individual inductances are summed up, the center wire sees about 15 % more inductance than the end wires. The matrices also show that the 2-wire case has higher self-inductance and mutual inductance from the adjacent wire than the 5-wire case. This is most likely due to the longer bondwires for the 2-wire case, as evidenced by the greater height of the 2-wire case in Figure 4.0.1.

In the 5-wire case, the difference in total inductance seen by the different wires generates a phase difference at the device end of the bondwires. But since the gate and drain on the device itself is connected, there is current flow across the device in the X-direction, which results in loss that is not present in the small devices with one or two bondwires.

4.1.2 Fingers

While the analysis of bondwire effects is relatively straightforward, the analysis of the influence of device fingers is not so simple. The complexity of analyzing fingers

stems mainly from the distributed nature of the device. As shown in Figure 4.0.2, the fingers run in the Y-direction, and a channel exists all along this axis. From a circuit perspective, one set of gate, drain, and source fingers would look as shown in Figure 4.1.1. Note that only a fraction of the current that enters or leaves a finger runs through the whole finger. So as a first-order estimation of the finger inductance, an approach similar to the bondwire case was applied.

The design of the device was such that the length of each finger was 250μm, and the pitch between each drain finger was 12μm. FASTHENRY simulation was performed on a 20-finger system, and the resulting inductance values are shown in Figure 4.1.2. The self-inductance of each finger was found to be 181pH, and a decrease in mutual inductance with distance was observed. Due to the distributed nature of the fingers, the numbers presented by FASTHENRY place an upper bound on what the finger inductance may be, as only a small fraction of the current travels through the whole finger.

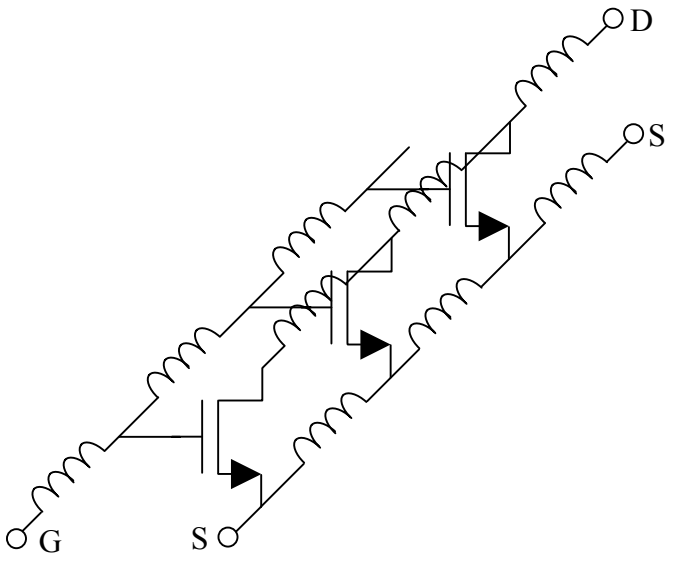


Figure 4.1.1. Circuit drawing representing the distributed nature of each finger of the active device. The current that flows into/out of any of the terminals does not flow through all the inductances, making estimation of finger inductance difficult.

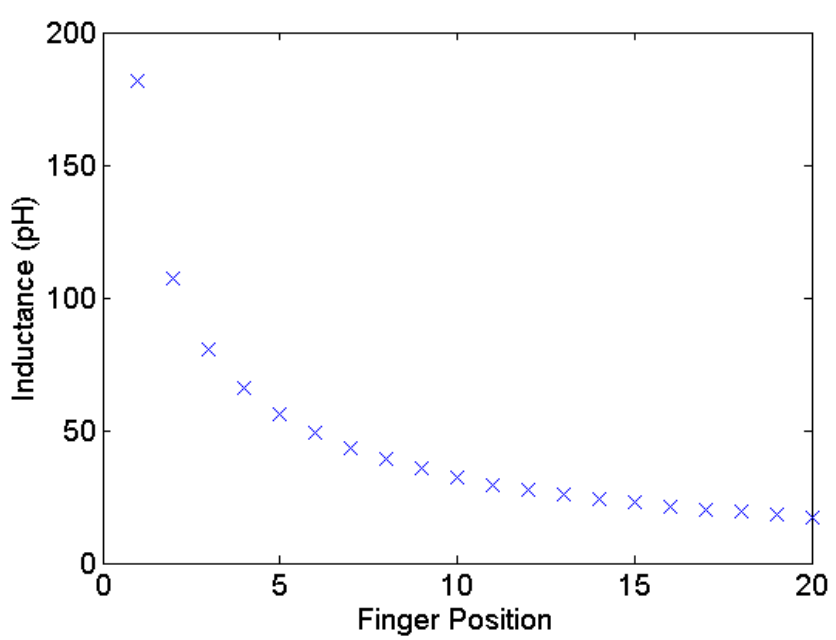


Figure 4.1.2. Mutual inductance values for a 250mm long finger, with 100um spacing between fingers. As the separation distance increases, the mutual inductance decreases.

Since the source finger structure has similar length and pitch as the drain fingers, the inductance matrix for the source finger would most likely be similar to that of the drain fingers. But the gate finger structure has twice as many fingers, at half the pitch, so while not quite precise, it is expected that the gate inductance matrix be twice the size, with twice the inductance values found in the drain matrix.

Also, since the gate, drain, and source fingers all run in parallel, the mutual coupling between gate/drain, gate/source, and drain/source must also be taken into account. Again, without knowing the exact current paths through the fingers, a precise FASTHENRY simulation only gives a precise solution to an inaccurate setup, and so its result is good only as an estimation of what the inductance matrix might be.

In order to validate the existence of mutual coupling between the bondwires and fingers, the estimated inductance values were used then to estimate the mutual inductance

observed at the unit cell level, and the circuit simulator ADS was used, as shown in the following section, to determine whether a model of a small device with extra mutual inductance mimics a large device.

4.2 Validation

4.2.1 Small-Signal Validation

With the mutual inductance of the bondwires and the fingers estimated, they were applied to the Root transistor model of a W=4mm device in ADS to see whether the effects of mutual inductance reproduces the s-parameters of larger devices. As shown in Figure 4.2.1, the gate-to-gate (M_{gx}), gate-to-drain (M_{gdx}), gate-to-source (M_{gsx}), drain-to-drain (M_{dx}), drain-to-source (M_{dsx}), and source-to-source (M_{sx}) finger mutual inductances as well as the bondwire mutual inductances at the gate (M_{gbx}) and drain (M_{dbx}) were added to the model. The subscript *x* after each mutual inductance label signifies the distance between the unit cells in question. Mutual inductance between adjacent unit cells are represented by x=1, between unit cells sandwiching another are represented by x=2, and so forth. Because the bondwire lengths, bondwire spacings, and unit cell spacings differ with total W, the inductances extracted in the previous section were not expected to be exact, and thus were used only as initial estimates. From these initial estimates, the inductance values were varied such that the simulation results fit the measurement results.

The first simulation compared measured data from a W=8mm device to a model consisting of two W=4mm Root models in parallel. The models were connected as shown in Figure 4.2.1, mutual inductance values as shown in Table 4.2.1 were assigned,

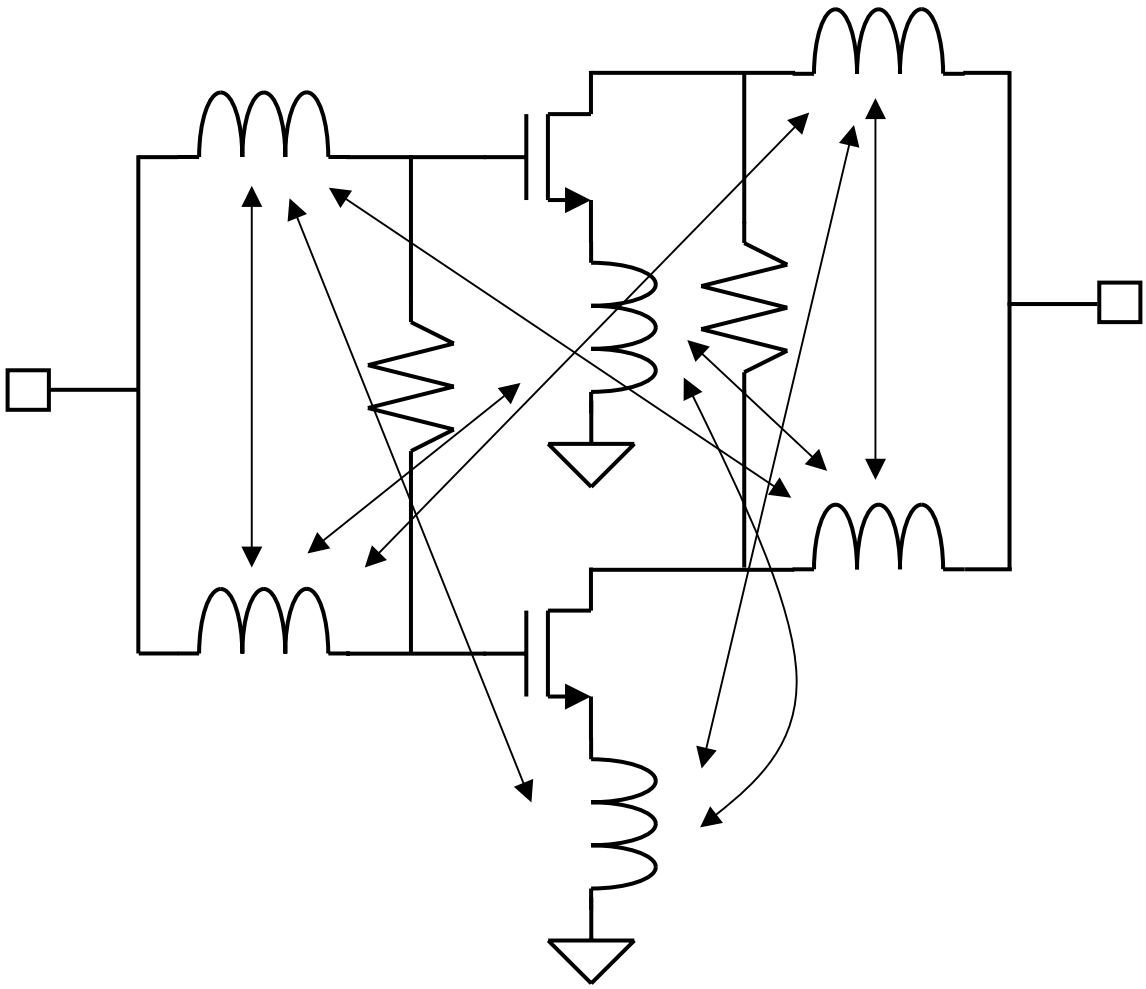


Figure 4.2.1. Circuit diagram showing the mutual inductance coupling between unit cells. Due to the configuration as shown in Figure 4.1.1, current flowing through parallel fingers affect each other, generating a complex web of interaction between the conductors at the gate, drain, and source terminals.

x	Mg (pH)	Md (pH)	Ms (pH)	Mgs (pH)	Mgd (pH)	Mds (pH)	Mgb (nH)	Mdb (nH)
1	100	50	50	100	100	50	1.3	1.3

Table 4.2.1. Mutual inductance between 2 unit cells to model the W=8mm device with two W=4mm devices.

and small-signal s-parameter simulation was performed on the model with Vgs=1.2V, Vds=3.5V, and f=0.5-3.0GHz. The comparison results are shown in Figure 4.2.2, where it can be seen that the model with the extra inductance approximates the measured data better than that of the W=4mm model by itself. Also, the maximum gain curves

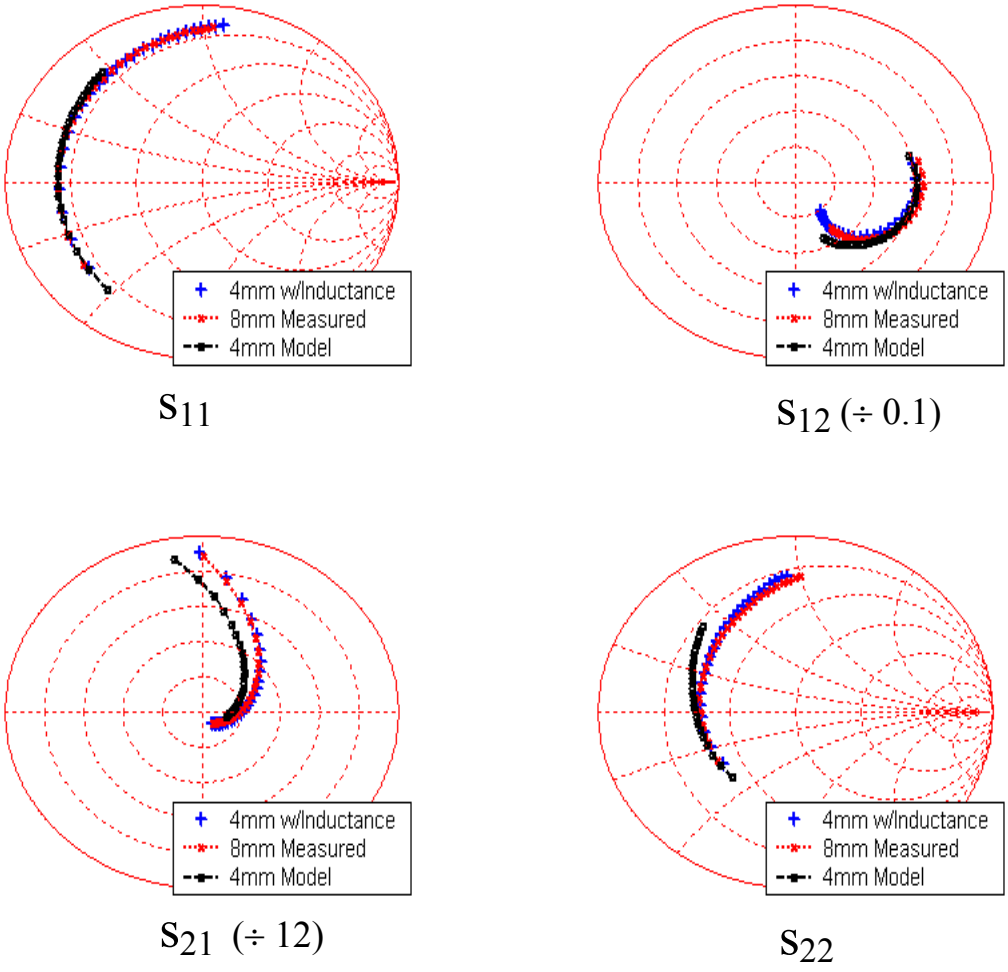


Figure 4.2.2. S-parameter results of a simulated 4mm Root model, a simulated 4mm Root model with extra mutual inductance, and a measured 8mm device. By adding the extra mutual inductance components to the 4mm model, the characteristics of the 8mm device may be reproduced.

calculated from the s-parameters are shown in Figure 4.2.3, where a good match between the measured data and the W=4mm model with inductance can be seen.

Using the same methodology as for the previous simulations, the measured results from a W=16mm device were compared to a model consisting of four W=4mm models. There were more mutual inductance values to consider, as shown in Table 4.2.2, and the simulation was performed at the same biases and frequency range. The results are shown in Figure 4.2.4, where the model with mutual inductance approximates the measured data

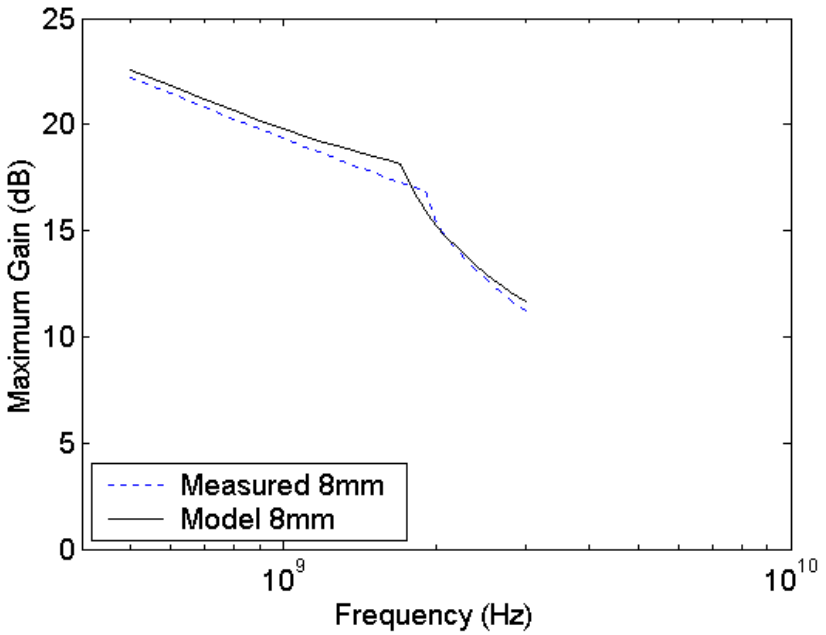


Figure 4.2.3. Maximum small-signal gain simulation results of a W=4mm Root model with mutual inductance compared to measured maximum small-signal gain of a W=8mm device, as calculated from the s-parameter data shown in Figure 4.2.2. The curves show very good agreement.

x	Mg (pH)	Md (pH)	Ms (pH)	Mgs (pH)	Mgd (pH)	Mds (pH)	Mgb (nH)	Mdb (nH)
1	50	25	25	50	50	25	1.3	1.3
2	40	20	20	40	40	20	0.91	0.91
3	30	15	15	30	30	15	0.65	0.65

Table 4.2.2. Mutual inductance between unit cells to model the 4-unit cell, W=16mm device. The on-chip mutual inductance values at x=1 are half that of Table 4.2.1 to account for the shorter finger lengths of the W=16mm device, and the values decrease as the distance between unit cells increases.

better than the W=4mm model, although a significant discrepancy exists between the model with mutual inductance and the measured data. This is most likely due to the fact that while the 4mm Root model consisted of sixteen 250μm-long fingers, the actual W=16mm device consisted of 128 125μm-long fingers. With the shorter fingers, the actual device can be expected to have smaller gate and drain resistances, which is observed in Figure 4.2.4. The maximum gain curves are shown in Figure 4.2.5, where these curves match only qualitatively, due to the reasoning given above.

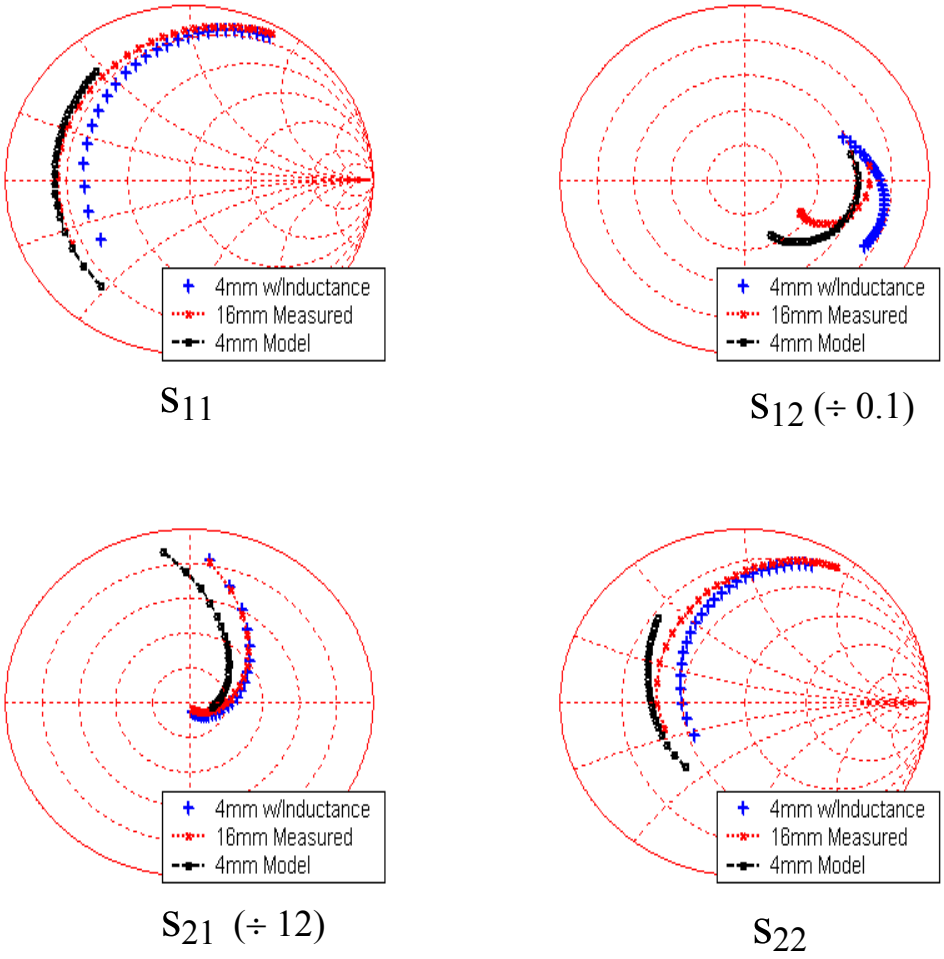


Figure 4.2.4. S-parameter results of a simulated 4mm Root model, a simulated 4mm Root model with extra mutual inductance, and a measured16mm device. By adding the extra mutual inductance components to the 4mm model, the characteristics of the 16mm device may be reproduced. Overestimation of the input and output resistances may be seen in the s_{11} and s_{22} plots, likely caused by the shorter fingers of the measured 16mm device.

The comparison becomes less straightforward with the 36mm device, as the size of the unit cell for the actual W=36mm device is not 4mm. The actual device has five unit cells, meaning each unit cell consists of W=7.2mm. For the purpose of this comparison, instead of having nine 4mm Root models in parallel to total 36mm, only five were used to match the five unit cells in the actual device. Using the mutual inductance values as shown in Table 4.2.3, the simulation was run under the same bias and frequency

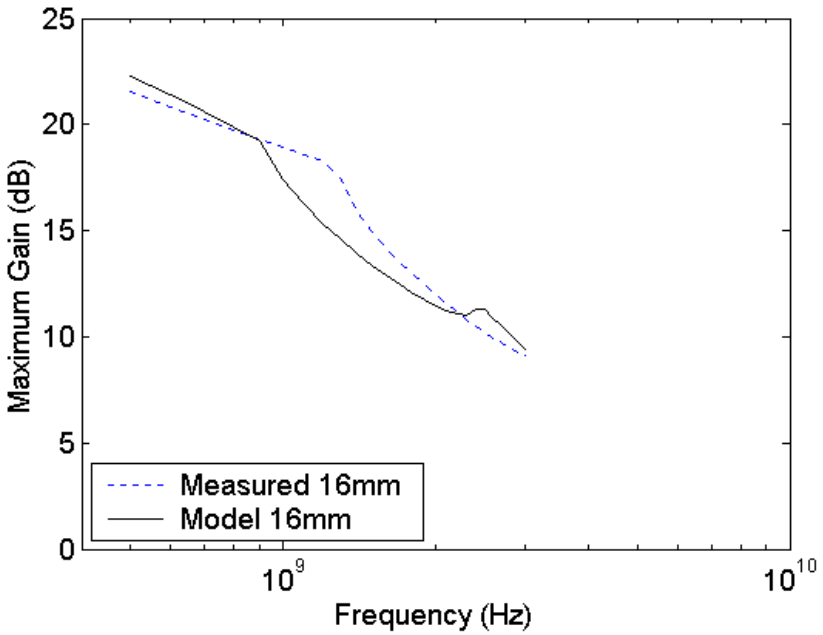


Figure 4.2.5. Maximum small-signal gain simulation results of a W=4mm Root model with mutual inductance compared to measured maximum small-signal gain of a W=16mm device, as calculated from the s-parameter data shown in Figure 4.2.4. The corner frequency is different, mainly due to the mismatch in the s_{21} data.

x	Mg (pH)	Md (pH)	Ms (pH)	Mgs (pH)	Mgd (pH)	Mds (pH)	Mgb (nH)	Mdb (nH)
1	100	50	50	100	100	50	1.3	1.3
2	80	40	40	80	80	40	0.91	0.91
3	60	30	30	60	60	30	0.65	0.65
4	40	20	20	40	40	20	0.39	0.39

Table 4.2.3. Mutual inductance between unit cells to model the 5-unit cell, W=36mm device. The inductance values at x=1 are identical to that of Table 4.2.1, and the values decrease with distance proportionally, compared to Table 4.2.2.

conditions, and the s-parameter results are shown in Figure 4.2.6. While the actual device data is significantly different from that of the original W=4mm model, the model with extra inductance matches the actual device reasonably well. The maximum gain curves of Figure 4.2.7 also show a reasonable resemblance between the measured and simulated data.

The same problem mentioned above is encountered when comparing the models to the W=64mm device. The actual device consists of eight 8mm unit cells. Thus again,

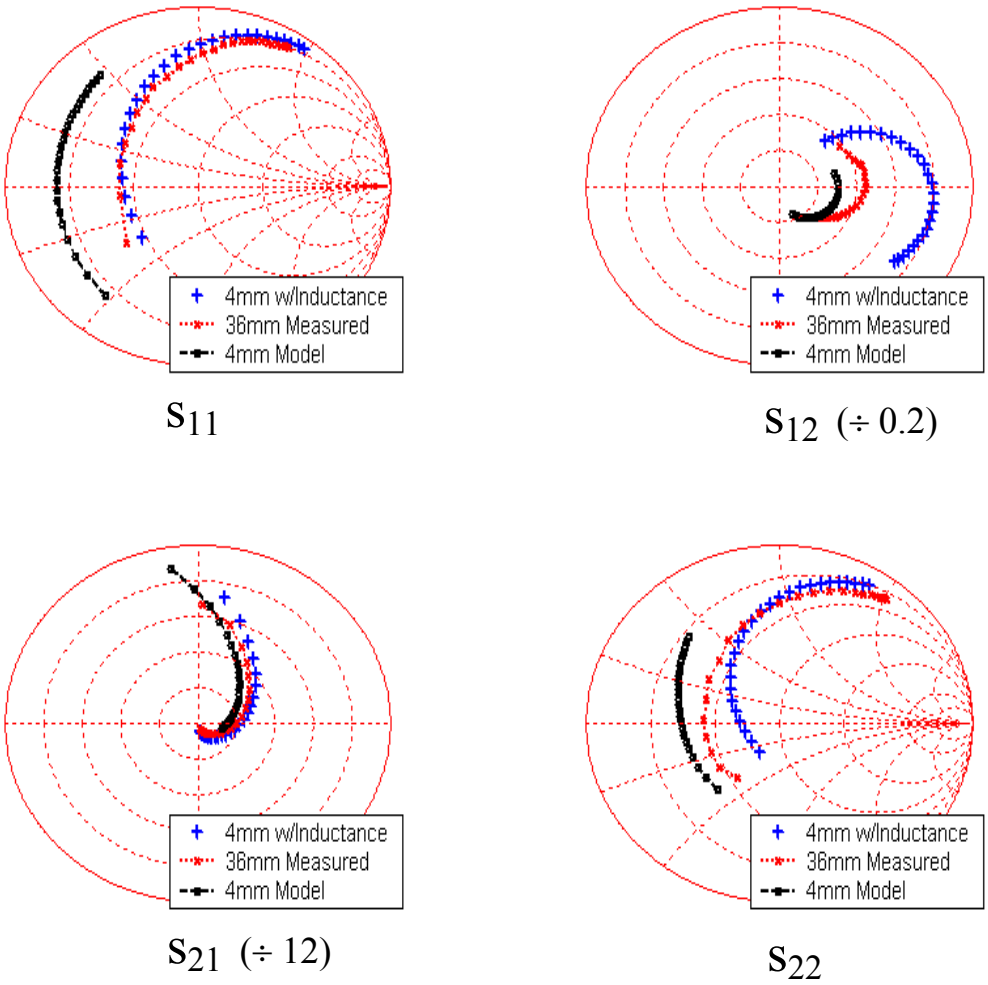


Figure 4.2.6. S-parameter results of a simulated 4mm Root model, a simulated 4mm Root model with extra mutual inductance, and a measured 36mm device. Again, the mutual inductance allows better modeling of the larger device. The remaining error is most likely due to the discrepancies in the unit cell performances (4mm unit cell modeled vs. 7.2mm unit cell measured).

eight 4mm Root models were placed in parallel for the comparison. The mutual inductance values as shown in Table 4.2.4 were applied, and the simulation was performed with the same bias and frequency conditions. The resulting s-parameters are shown in Figure 4.2.8, and the maximum gain curves in Figure 4.2.9. Again, the model with mutual inductance matches the measured data with reasonable accuracy.

The overall trend as W increases can be seen in the maximum gain curves in Figures 4.2.10 and 4.2.11. Figure 4.2.10 shows the maximum gain for different W taken

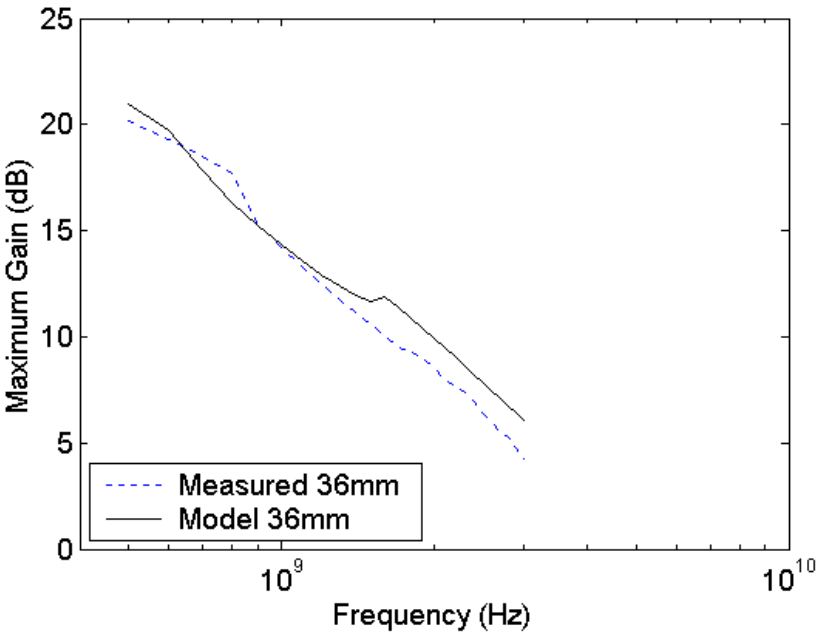


Figure 4.2.7. Maximum small-signal gain simulation results of a W=4mm Root model with mutual inductance compared to measured maximum small-signal gain of a W=36mm device, as calculated from the s-parameter data shown in Figure 4.2.6. The curves show reasonable agreement.

x	Mg (pH)	Md (pH)	Ms (pH)	Mgs (pH)	Mgd (pH)	Mds (pH)	Mgb (nH)	Mdb (nH)
1	100	50	50	100	100	50	1.30	1.30
2	80	40	40	80	80	40	0.91	0.91
3	60	30	30	60	60	30	0.65	0.65
4	40	20	20	40	40	20	0.39	0.39
5	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0

Table 4.2.4. Mutual inductance between unit cells in modeling the 8-unit cell, W=64mm device. Mutual inductance between unit cells farther than 4 unit cells are ignored, making the contents of this table similar to that of Table 4.2.3. But due to the difference in the device geometry, the electrical characteristics of the W=36mm and the W=64mm are different.

from measured devices. It can be seen that as W increases, the gain at lower frequency decreases, and the knee frequency at which the gain starts to drop steeply, decreases. Figure 4.2.11 shows the maximum gain curves from the various simulated W=4mm

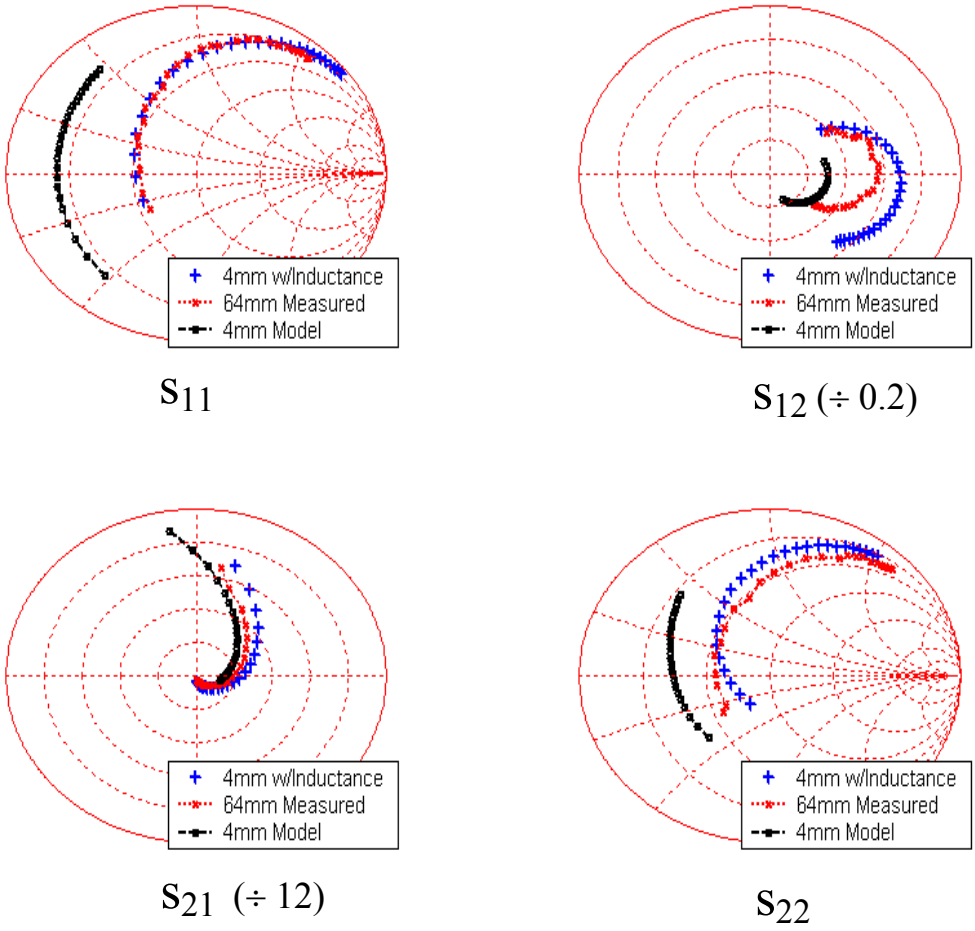


Figure 4.2.8. S-parameter results of a simulated 4mm Root model, a simulated 4mm Root model with extra mutual inductance, and a measured 64mm device. Again, the mutual inductance allows better modeling of the larger device. Like the 36mm comparison, the remaining error is most likely due to the discrepancies in the unit cell performances (4mm unit cell modeled vs. 8mm unit cell measured).

devices with mutual inductance. In comparing the two figures, it can be seen that the simulated results also show the trend with respect to gain and knee frequency.

While the s-parameter data comparing the measured and simulated systems resemble each other reasonably, errors between the two datasets were observed, especially for larger values of W. The error between the 4-mm with mutual inductance model and the large device measurement results may be attributed to several factors. First, the distributed nature of the unit cells poses problems. Namely, the magnetic flux generated is not uniform throughout the unit cell due to the different amounts of current

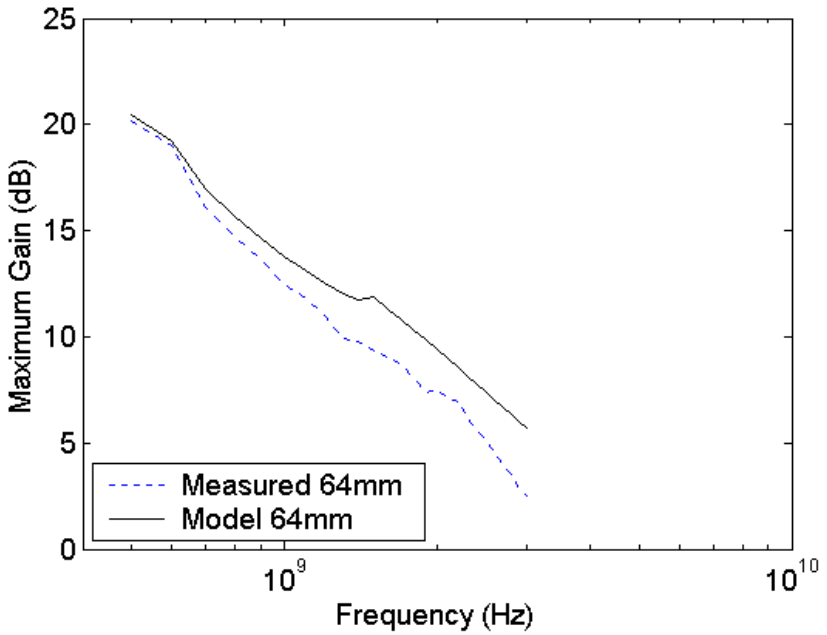


Figure 4.2.9. Maximum small-signal gain simulation results of a W=4mm Root model with mutual inductance compared to measured maximum small-signal gain of a W=64mm device, as calculated from the s-parameter data shown in Figure 4.2.8. The curves agree overall, but show better matching at lower frequencies.

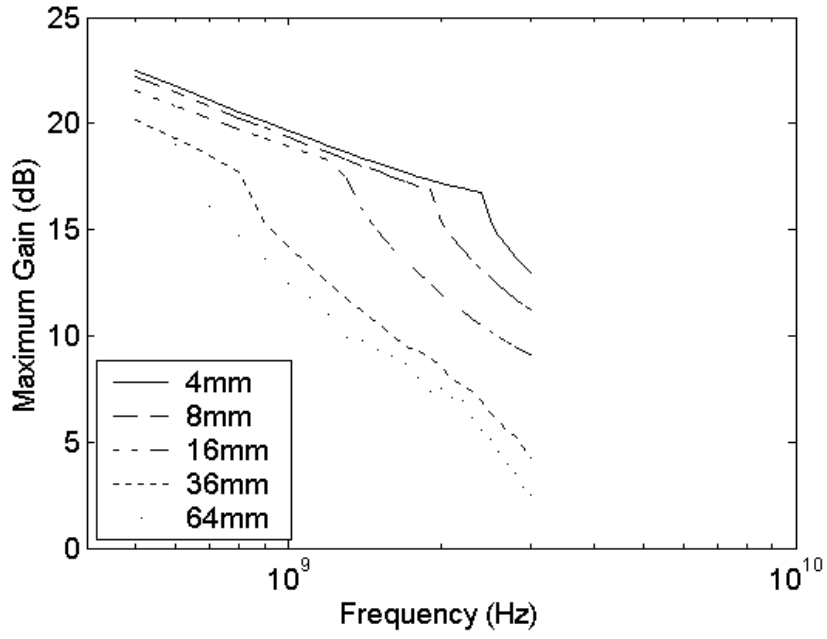


Figure 4.2.10. Small-signal maximum gain calculated from s-parameter measurements. As the width increases, the low-frequency gain decreases and the corner frequency decreases.

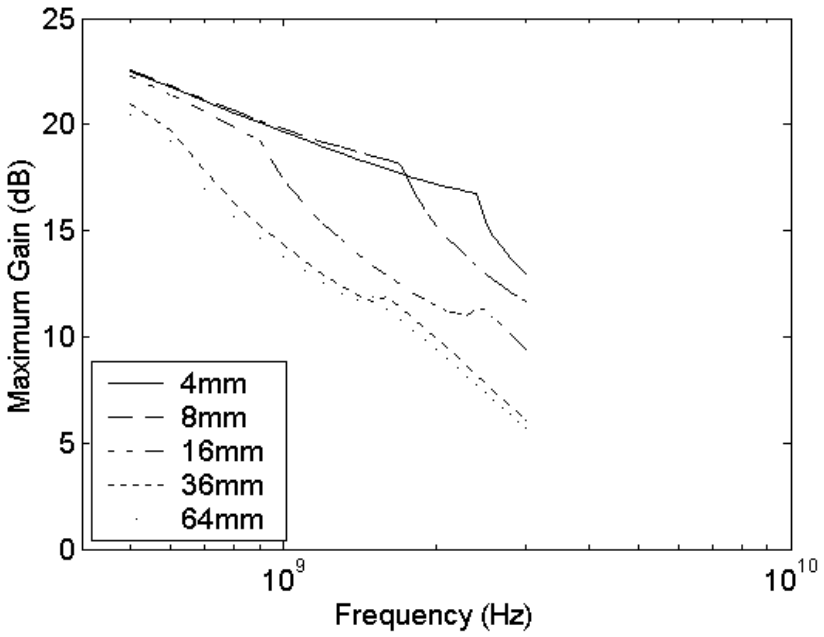


Figure 4.2.11. Small-signal maximum gain plots from s-parameter simulations of devices with mutual inductance. Like Figure 4.2.10, as the width increases, the low-frequency gain decreases, and the corner frequency decreases.

flowing through different device fingers. Depending on the location of the finger within the unit cell, its effect on other unit cells will be different, and the effect of other unit cells on that finger will be different. The approach taken, which consists of using a lumped mutual inductance term at the device terminals, while simple, cannot account for this distributed effect. In order to model the position-dependence of the magnetic flux with more precision, the variation of generated and observed magnetic flux in both the X and Y directions need to be modeled. In addition, applying this two-dimensional flux map to the circuit simulator requires further segmentation of the active device, ideally into small sections of a finger, as shown in Figure 4.1.1. This approach is currently impractical, as the component count increases dramatically, and a device model of the section of a finger, which is not currently available, is required.

The second reason for the difference between the measured results and the 4-mm model with inductance results is the fact that the unit cell size does not stay constant for the available test devices. While it is desirable to have constant-size unit cells for analysis, real devices are not designed with that in mind. Thus the real W=36mm device, for example, does not consist of nine, 4mm unit cells, but fewer, larger unit cells. An unavoidable discrepancy occurs due to using the 4mm Root model as the standard against which the measurement was compared. Due to the fact that large scaled devices composed of 4mm unit cells were not available, some error between the measured results and simulated results is expected.

4.2.2 Large-Signal Validation

The question that remains then is whether the change in small-signal s-parameters caused by mutual inductance translates into lower large-signal performance. In an attempt to answer this question, load-pull simulations in ADS were performed using the 4mm Root model in conjunction with the mutual inductance parameters extracted in the previous section.

Figures 4.2.12 and 4.2.13 show the load-pull simulation results for device models determined in the previous section, when the input power density was set to 7dBm/mm, f=900MHz, Vgs=1.2V, and Vds=3.5V. As mentioned in the previous section, the number of unit cells was kept the same between the actual devices and the models. Thus the 36mm and 64mm models had five and eight unit cells, or actual widths of 20mm and 32mm, respectively. For these cases, the input power was set to correspond to the actual widths. In order to take this effect into account, the comparison against width was not

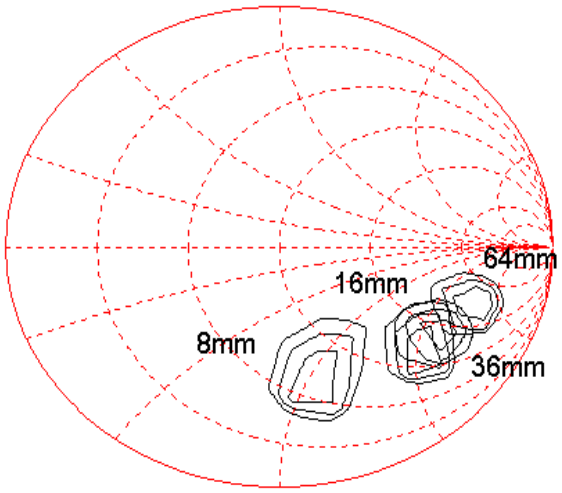


Figure 4.2.12. Normalized output power load-pull simulation results of Root model devices with mutual inductance parameters shown in Tables 4.2.1-4.2.4. The movement of contours in the negative reactance and positive resistance, as well as the shrinking of the contours, correlates well with the measured load-pull results of Figure 3.1.15.

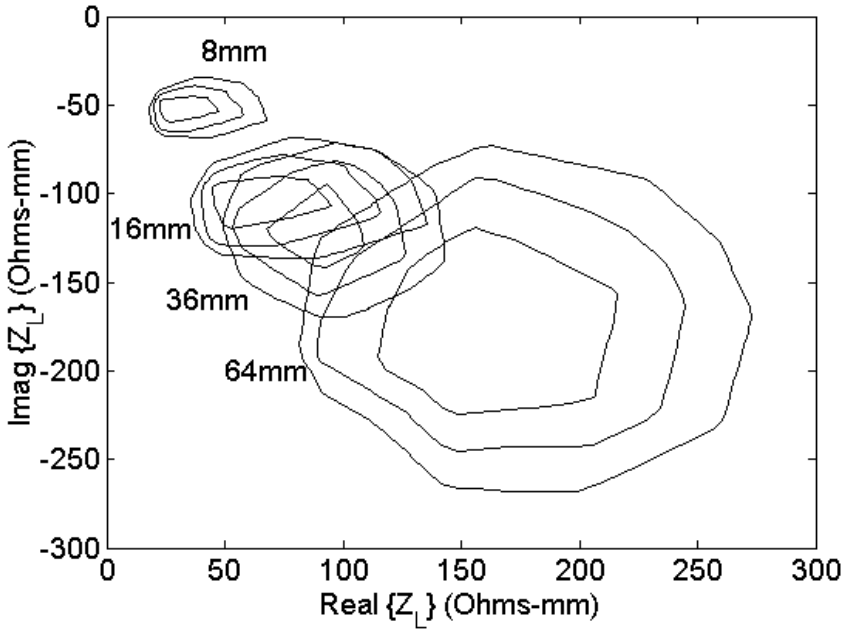


Figure 4.2.13. The results of Figure 4.2.12 on a Cartesian impedance plane. The movement of the contours with respect to impedance is more obvious, and in this plane, the contours increase in size with increased width.

evaluated with maximum output power, but rather maximum power gain. On the Smith chart of Figure 4.2.12, it can be seen that as width increases, the maximum power gain contours shrink in size as they move in the positive resistance and negative reactance

W (mm)	GpMeas (dB)	GpSim (dB)
4	11.6	14.18
8	10.5	14.16
16	10.4	12.62
36	8.79	11.27
64	8.80	8.56

Table 4.2.5. Maximum output power gain for the load-pull simulations using the devices with mutual inductance shown in Tables 4.2.1-4.2.4, compared to measurement. The simulated results show a greater decrease in power gain with increased width.

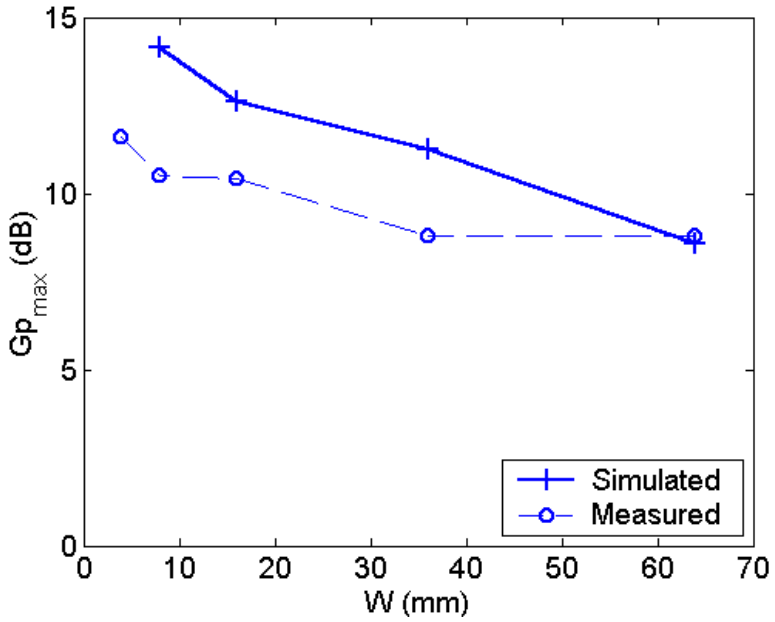


Figure 4.2.14. Comparison of maximum power gains between simulations with mutual inductance and measurements. Both curves show a decrease in the maximum power gain as width increases.

directions. When this data is converted to a width-normalized Cartesian impedance plane of Figure 4.2.13, the change in impedance becomes more obvious, and it can be seen that the contours increase in size. These trends compare favorably with measurement data shown in Figure 3.1.15 and 3.1.16, in that the contour size trends and impedance trends are the same.

Table 4.2.5 shows the maximum simulated output power gain for each model. Figure 4.2.14 shows the plot of maximum power gain against width for both the

measured and simulated devices. It can be seen that the two traces on the plot decrease in a similar manner as the width increases.

From the observed decrease in maximum power gain with increased width for the models, it has been shown that by taking into account the mutual inductance between unit cells, the large-signal characteristics of the models approximate those of real devices.

4.3 Preventing Performance Degradation

In order to prevent performance degradation observed in previous sections, the effect of mutual inductance at the bondwires and fingers must be minimized. In minimizing the effects of mutual inductance, the mutual components could be reduced, or in some cases, equalized across unit cells. There are a number of methods by which the mutual inductance effects may be reduced, each with a different degree of practicality. The following sections will discuss some of these possible measures, including their effectiveness and practicality.

4.3.1 Physical Separation

A simple solution towards reducing the effect of mutual inductance would be to physically separate the unit cells as shown in Figure 4.3.1, since mutual inductance decreases with distance as shown by equation (4.2). This would lead to greater separation between bondwires as well, leading to smaller bondwire mutual inductance. Figure 4.3.2 shows the relationship between mutual inductance and the separation distance. The curve shows that while increasing the separation distance does indeed decrease mutual inductance, even a large separation of 200μm still yields a mutual

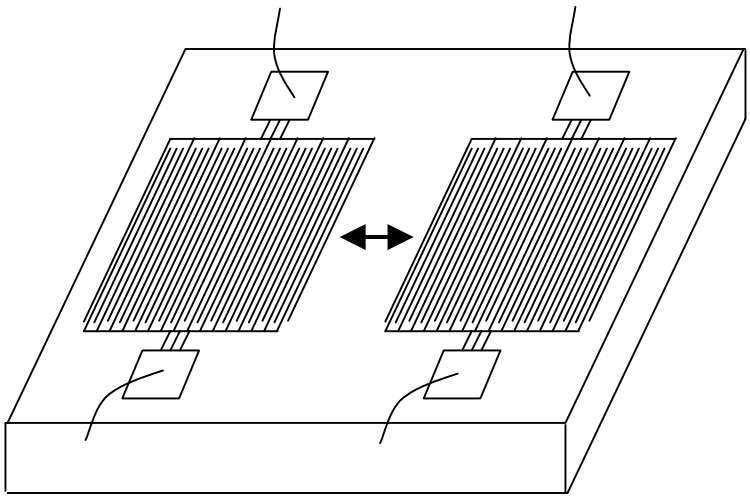


Figure 4.3.1. Drawing indicating the separation of unit cells in order to reduce the mutual inductance between the parallel bondwires. Notice that die area is wasted by increasing the separation distance if the center is left empty. Or if the center is filled with additional fingers, there is a greater variance in the signal path lengths between the fingers close to the bondpads and far from the bondpads, leading to phase shifts.

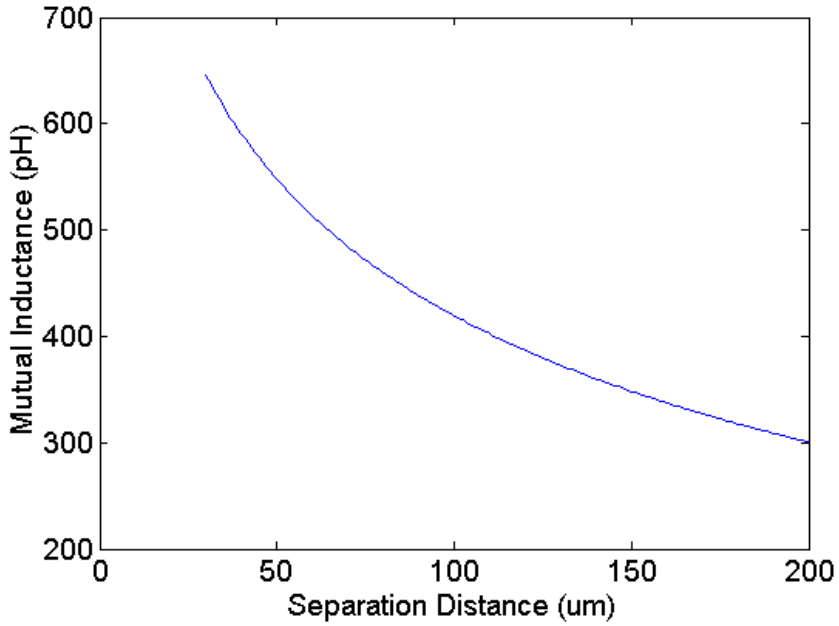


Figure 4.3.2. Plot of mutual inductance against distance, as calculated using equation (4.2). The mutual inductance decreases with distance, but not sharply enough for the separation of unit cells to be effective towards minimizing mutual inductance.

inductance of 300pH. Too large a separation distance is also problematic, however, since the effective bondwire resistance of the system increases because fewer bondpads and bondwires can be placed with large separations, and the on-chip phase shift between

fingers close to the bondpads and fingers far from the bondpads increases. Thus although a large separation distance would mean lower mutual inductance, other factors make this solution most likely infeasible.

4.3.2 Changing Bondwire Lengths

Since multiple bondwires run in parallel to connect the external system to the power amplifier unit cells, mutual inductance affects the bondwire system. When there is a system of parallel, identical bondwires, it can be expected that the system is symmetrical about the central wire(s). While pairs of wires equidistant from this central line of symmetry may be expected to have identical inductive characteristics, each set of wires at different distances should have different inductances. This position dependence of mutual inductance leads to phase-shifting between unit cells, as the signals see different impedances through the bondwires. In order to prevent this phasing, ideally, the bondwires would be removed, by using technology that does not require bondwires--flip-chip technology, for example. However, if removing bondwires is not possible, and packaging constraints prevent minimizing bondwire lengths, an alternative would be to adjust the lengths of the bondwires to equalize the total inductance of each wire. Once equalized, the inductance, and the resulting reactance, could be matched by an external matching circuit. The following demonstrates the methodology by which inductance equalization may be achieved.

From equations (4.1) and (4.2), the self and mutual inductances for a system of n parallel bondwires may be calculated. When the bondwires are all 1000μm long Figure 4.3.3 shows the FASTHENRY simulation results for the three- and five-wire cases,

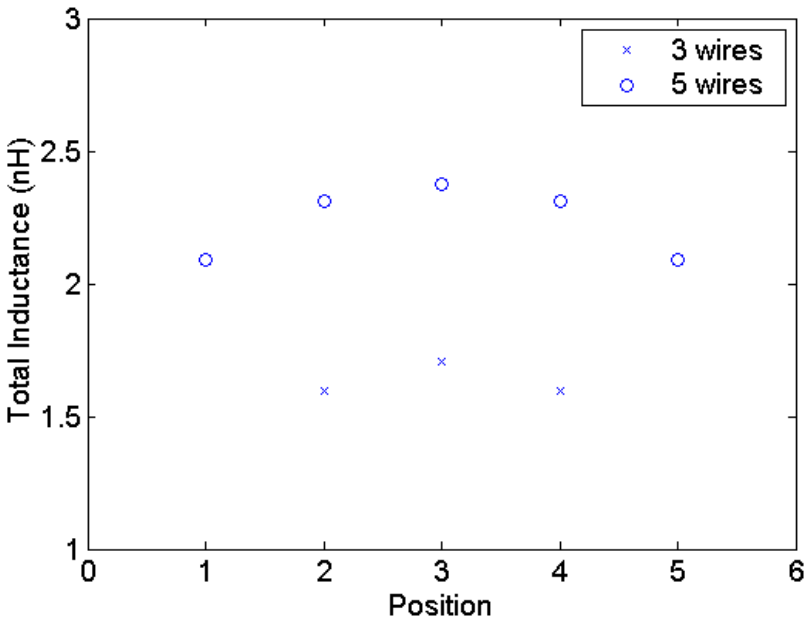


Figure 4.3.3. FASTHENRY simulation results for 3 and 5 parallel bondwires when all their lengths are *lmin* (1mm). As expected, the central wires see more total inductance than the outer wires.

where the inductance experienced by each wire is plotted against the position of the wire in the system. It can be seen that the central wires see significantly more inductance than the outer wires. This is due to the fact that the central wires have more wires that are close to it than the outer wires. In order to equalize the total inductance of each wire, equation (4.2) suggests that the central wires be shorter than the outer ones. This inductance equalization problem is a non-linear optimization problem, but the solution is not very difficult to obtain, as even the longest (outer) wire tends to be not too much longer than the shortest (center) wire.

In a three-wire case, the problem may be stated as shown in equations (4.5)-(4.6).

$$\begin{bmatrix} L_1 & M_{12} & M_{13} \\ M_{12} & L_2 & M_{23} \\ M_{13} & M_{23} & L_3 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} L_{T1} \\ L_{T2} \\ L_{T3} \end{bmatrix} \tag{4.5}$$

where $L_1 = L_3, M_{12} = M_{23}$ by symmetry. Given equation (4.5), the equality shown in equation (4.6) is desired for equality of total inductance in each bondwire.

$$L_{T1} = L_{T2} \tag{4.6}$$

Since by symmetry, $L_{T1}=L_{T3}$, this is the only equality condition that needs to be satisfied. Given a minimum bondwire length l_{min} as governed by the package geometry, assigning this length to the central wire, and assuming that the length parameter l used in the mutual inductance calculation is the length of the shorter of the two wires, the length of the outer wire may be calculated from these equations and equations (4.1) and (4.2). Table 4.3.1 shows the results of the calculations when l_{min} is chosen to be 1000 μm, r=25 μm, and d=80 μm.

In the more complex, five-wire case, the format of the problem remains the same, but there are more equations to be solved simultaneously. By using a computer program such as MATLAB, the optimal lengths of the bondwires may be calculated, and the results given the same l_{min}, r, and d, are tabulated in Table 4.3.2.

As a verification of this simple inductance equalization, the calculated bondwire lengths shown in Tables 4.3.1 and 4.3.2 were entered into FASTHENRY to determine

	Outer	Center	Outer
Length l (mm)	1.085	1.000	1.085

Table 4.3.1. Optimized bondwire lengths for the 3-wire case to equalize the total inductance in the 3 wires. Note that the outer wires need to be 8.5% longer than the central wire.

	Outer	------	Center	------	Outer
Length l (mm)	1.200	1.026	1.000	1.026	1.200

Table 4.3.2. Optimized bondwire lengths for the 5-wire case. The outer wires must be 20% longer compared to the central wire for the total inductance to be equalized.

whether the simulator returns equalized inductance values. As shown in Figure 4.3.4, FASTHENRY results show that the total bondwire inductances are equalized when the length values obtained from calculations are used.

Whereas the system of three identical wires yielded an inductance of the central wire that was 7.0% greater than the inductance of the outer wire (Figure 4.3.3), by altering the lengths of the wires, the difference in inductance was decreased to 0.3%. Similarly, for the five-wire case, the identical parallel wires showed a difference of 13.5%, but this variance was reduced to 2.3% by altering the bondwire lengths.

Using this methodology to equalize the bondwire inductances, the signal phases at the gate and drain pads on-chip may be equalized to prevent loss caused by the x-direction current flow mentioned in Section 4.1.1, given identical unit cells.

In looking at the on-chip device fingers, this method of equalizing mutual

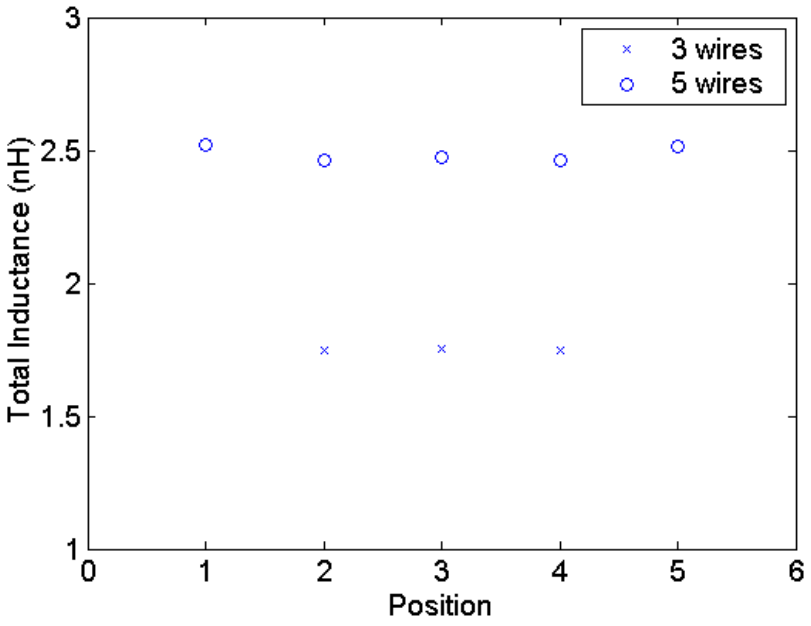


Figure 4.3.4. FASTHENRY simulation results of 3 and 5 parallel bondwires with optimized lengths of Tables 4.3.1 and 4.3.2. Note that the total inductance that each bondwire sees has been equalized.

inductance unfortunately cannot be applied very well. Again, the distributed nature of the fingers prevents accurate modeling of the finger inductance. Although the trend should be such that the fingers in the central unit cells should be shorter than the fingers in the outer unit cells for the same reasons given earlier in this section, without knowledge of the inductance in the fingers, changing the finger lengths would require trial-and-error. But due to the expense of multiple device fabrication runs, this trial-and-error approach would not be very cost-effective.

4.3.3 Alternating Inverted-Phase Signals

Without the ability to equalize on-chip inductances between unit cells via changing finger lengths, alternative solutions must be devised to equalize or reduce the on-chip mutual inductance.

One such solution would require the use of baluns at the gate and drain to split and combine signals that have opposite phases, as shown in Figure 4.3.5. By feeding inverted signals in alternate unit cells, it can be expected that the magnetic fields observed by each unit cell would be reduced, as adjacent unit cells would be generating magnetic fields with opposite polarities.

Starting with the mutual inductance values in Tables 4.2.1-4.2.4, and applying them to the circuit shown in Figure 4.3.5, simulations in ADS were performed. In addition to the circuit seen in Figure 4.3.5, larger devices were simulated by placing more 4mm unit cells in parallel, with alternating connections to the in-phase or inverted signals.

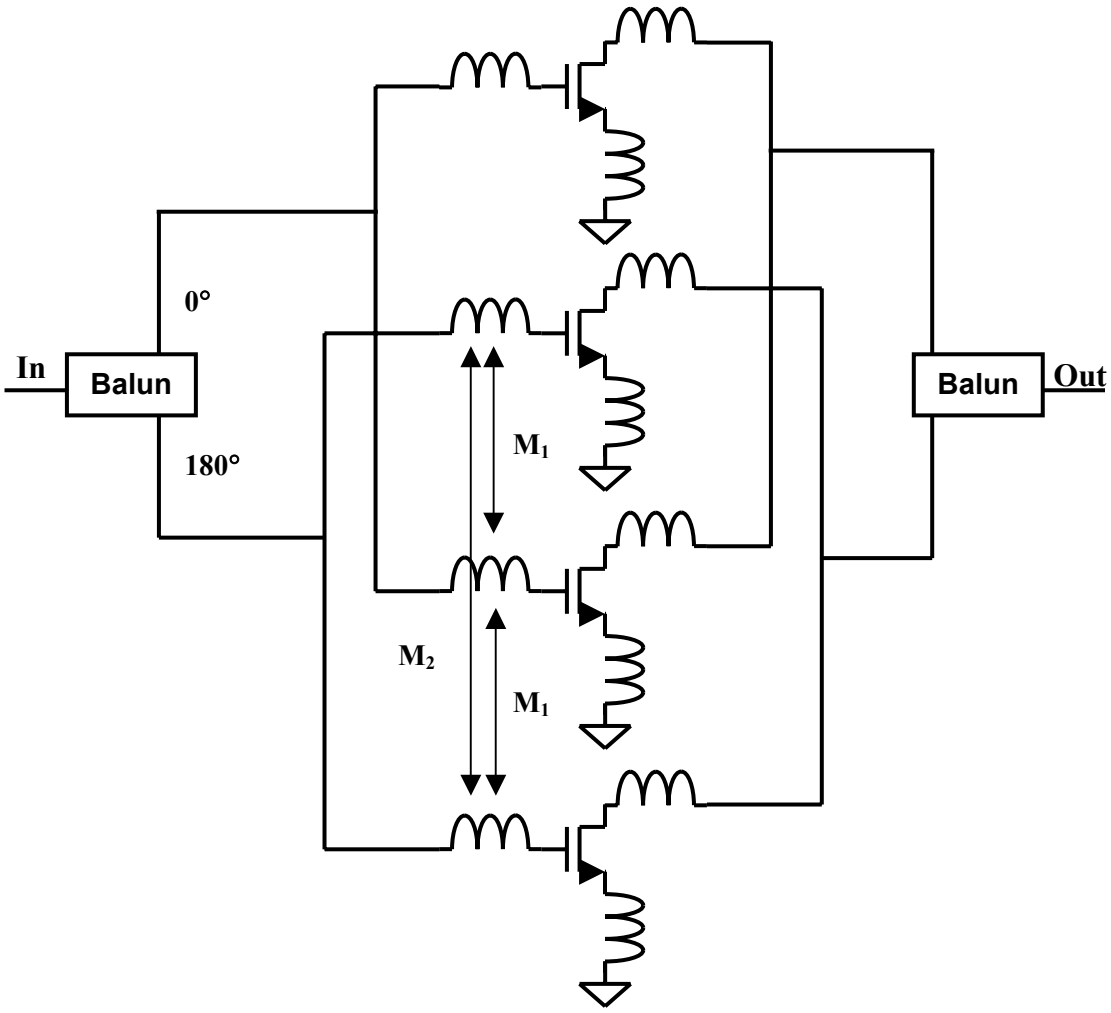


Figure 4.3.5. Power amplifier circuit diagram to counter the mutual inductance at the three terminals. Phase inversion is achieved in adjacent unit cells by using baluns. Phase inversion generates negative mutual inductance between adjacent unit cells ($M_1<0$), thus countering the source degeneration effect caused by positive mutual inductance observed in traditional power amplifier devices. Although $M_2>0$, since $|M_1|>|M_2|$ due to the larger distance for M_2, the net effect of all the mutual inductances should be negative. This canceling of mutual inductance should not only occur at the gate terminal as shown, but also at the drain and source terminals.

S-parameter simulation results of the 8mm, 16mm, 36mm, and 64mm device as modeled in Section 4.2.1 with baluns at the input and output are shown in Figure 4.3.6. Unlike Figure 3.1.11, where the width-dependence of the s-parameters was obvious, all the s-parameter curves in Figure 4.3.6 lie on top of one another, indicating that the mutual inductance effects have been neutralized by the baluns.

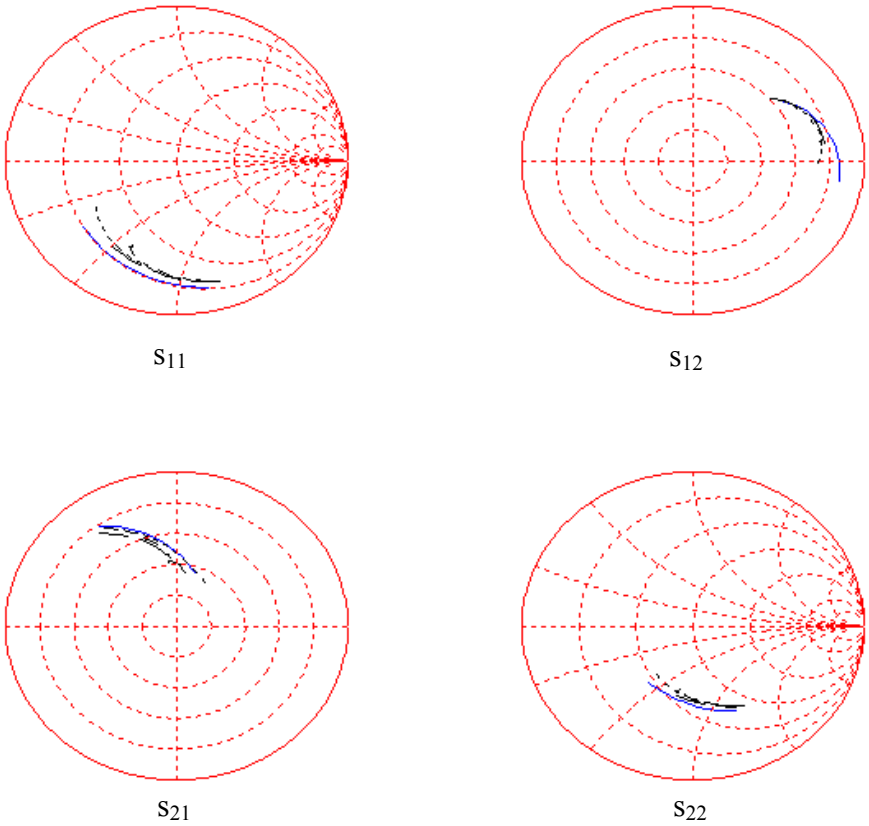


Figure 4.3.6. Small-signal s-parameter simulation results for different width devices with baluns. Note that irrespective of width, the curves are identical, showing that the addition of baluns counters the effect of mutual inductance.

Using the same devices with baluns, load-pull simulations were run in ADS. The results are shown in Smith Chart form and in normalized impedance form in Figures 4.3.7 and 4.3.8, respectively, and summarized in the form of a maximum output plot in Figure 4.3.9. Also plotted in Figure 4.3.9 is the maximum power gain curve of the devices without baluns as shown in Figure 4.2.14. Comparing the devices with and without baluns, it can be seen that the maximum power gain of devices without baluns decreases from 14dB to 8dB between 8mm and 64mm, while the maximum power gain for devices with baluns stays constant at around 14dB for all widths.

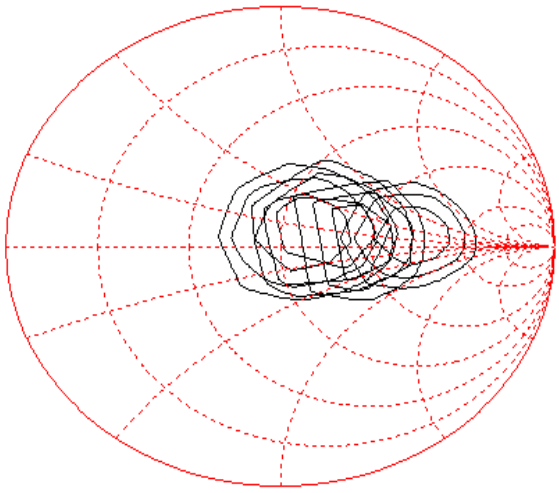


Figure 4.3.7. Maximum power gain contours for the load-pull simulations with baluns on devices with different widths. Each contour represents a 0.5dB step. The contours do not move in the reactance direction with increased width, but only increases in resistance.

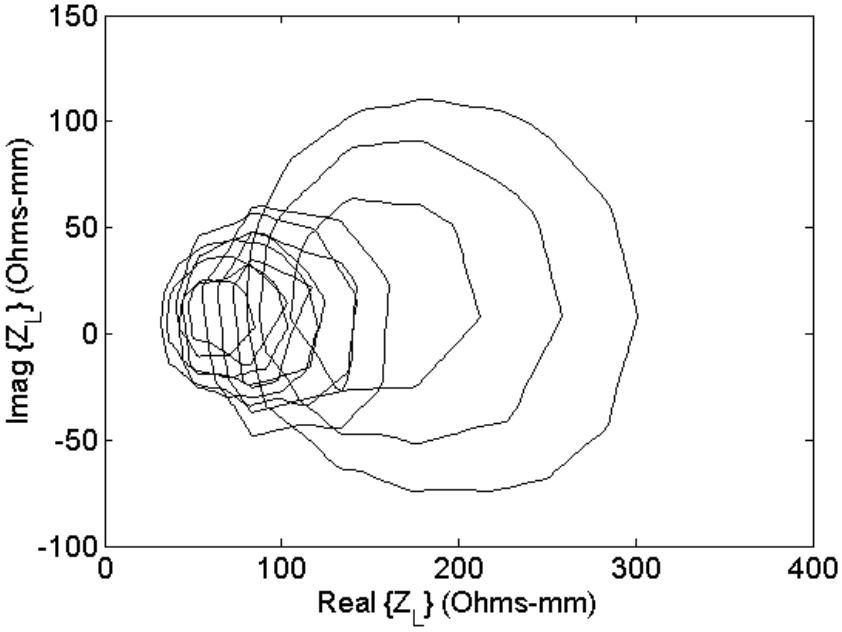


Figure 4.3.8. Maximum power gain contours from Figure 4.3.7 on a Cartesian impedance chart. The optimal reactance stays at zero for all widths, and the optimal resistance increases with width. Also, the contours increase in size as the device size increases.

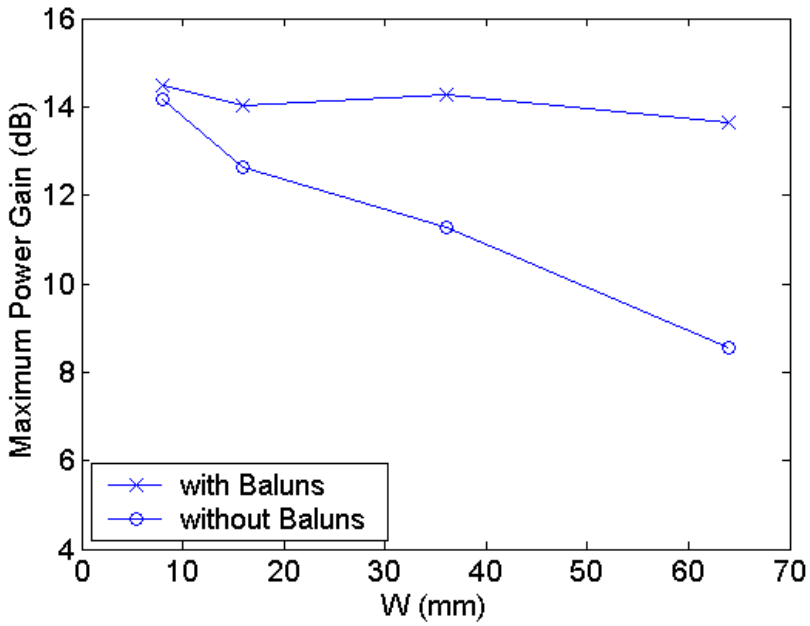


Figure 4.3.9. Maximum output power gain against device width for devices without baluns and devices with baluns. By adding baluns, the power gain lost due to mutual inductance is recovered.

4.4 Conclusion

The previous chapter demonstrated that inductance values were not scaling properly with width as more unit cells were integrated onto one die, and it was suggested that mutual inductance was the mechanism by which this non-scaling occurred. This chapter thus examined mutual inductance in great detail. Mutual inductance values were extracted, and these values were applied to circuit models to verify that mutual inductance causes the performance degradation observed in Chapter 3. Once it was established that mutual inductance causes the expected change in electrical parameters of the LDMOSFET, measures to minimize the effects of mutual inductance were developed and tested through simulations. It was shown that while trying to equalize the total inductance seen by each unit cell via changing conductor lengths works well with

bondwires, unfortunately, on-chip fingers cannot be treated in a similar way, due to the distributed, active nature of the LDMOSFET. A better approach would be to apply baluns at the input and output, with the phase-shifted signals connected to alternating unit cells on-chip, in order to cancel the largest mutual inductance component from each adjacent cell with an opposite-polarity field. Simulations showed that use of baluns makes the normalized small-signal s-parameters and large-signal maximum power gain independent of device width.

Thus a problem stemming from the integration of multiple LDMOSFET cells operating at high frequencies was characterized, root causes for loss in performance were identified, and solutions offered. The cause of the problem was limited to high frequencies, since the mutual inductance effect is largely negligible at low frequencies, making the analysis that much more difficult.

5 Conclusions

This dissertation discussed issues that arise from integration of high-current systems that operate at high frequencies. Two cases were studied—the first involved applying ESD protection to high-speed circuits, and the second involved placing more RF power amplifier cells in parallel. This chapter summarizes the findings and provides direction for future work in these areas.

5.1 Summary

In integrating ESD protection with high-speed circuitry, the concerns are two-fold: whether the ESD protection actually shunts enough current during ESD events, and whether the ESD protection still allows proper normal operation. The first concern is not the focus of this dissertation, and is addressed in other literature [4]. Chapter 2 addressed the second concern, by quantifying the effect of ESD protection on normal, high-speed operation through the use of s-parameters. In protecting high-speed systems, the

dominant factor is the capacitance of the ESD protection system that appears on the signal path. This capacitance shorts AC signals to ground, and presents an obstacle to efficient power transfer. But it was found that applying a distributed protection scheme prevents the degradation in power transfer that would be observed with a more coarsely lumped protection scheme. Also, a methodology by which the distributed ESD protection could be designed was presented. Given the requisite ESD protection device, its total parasitic capacitance may be extracted, from which a distributed ESD protection system can be designed to minimize loading of the normal operation signal. Thus an approach to solving the problem of integrating ESD protection with high frequency circuits was demonstrated.

In Chapters 3 and 4, the integration of RF power amplifiers was addressed. Measurements showed that performance per device channel width decreased with increased width. Low-frequency characteristics, as well as thermal characteristics showed little width-dependence, which left high-frequency characteristics as the most likely cause of this phenomenon. In Chapter 4, it was demonstrated that electromagnetic coupling between unit cells was the most likely cause of performance degradation. Evidence consisting of measurement and simulation results was presented to support the case. Also, methods by which the performance degradation may be prevented, were presented. In particular, equalization of mutual inductance by altering conductor lengths and spacings, as well as cancellation of the magnetic field through the use of baluns, was discussed. Simulations testing these measures were performed, and the results showed the recovery of performance that otherwise would have been lost. Thus the mechanism

by which performance decreases with integration of RF power amplifier unit cells was identified, and methods to prevent this phenomenon were suggested.

5.2 Future Work

There are a few related issues outside the scope of this work that may be of interest in the future. First and foremost is the development of a fast simulator capable of unified electromagnetics, circuits, and semiconductor device physics simulations. The current state of simulators requires a piece-meal approach--the data from a device simulator and an electromagnetic simulator must be extracted into a model that the circuit simulator can use. While this approach provides results, there is little ease-of-use. With further miniaturization and integration, circuit simulations that require analysis of electromagnetics and semiconductor device physics should become more common, and a faster, easier-to-use, unified simulator would be needed to allow more engineers to design these systems. Once such a simulator becomes commonly available, resources may be utilized more efficiently, with specialists free to tackle the more difficult problems.

While working on analyzing ESD protection systems for high-speed circuits, an interesting extension was suggested by a reviewer of the work. The suggestion was to quantify the effects of ESD protection on noise behavior. In a wireless receiver system, the first stage is usually the low-noise amplifier (LNA). ESD protection for this amplifier would be desirable, but the noise added by the protection system must be minimized, in addition to minimizing the high-speed loading problem. Analysis of the noise impact of ESD protection on RF systems would therefore be useful in decreasing defective RF front-end systems.

Regarding the RF power amplifier system, the methods of improving performance (preventing performance degradation) have yet to be implemented. While the simulations show good results, there may be unexpected obstacles in extracting better performance out of larger power amplifier devices. Also, the power amplifiers investigated in this work delivered output powers of a few watts. For base-station class amplifiers, delivering more than 100W, the same phenomenon is observed, but it is unknown whether the discoveries of this work also directly apply to larger amplifiers. The thermal effects are expected to be more prominent, and the sensitivity to losses in the matching networks could be greater due to the lower device impedances.

Appendix A
S-Parameter Simulation Results

In Chapter 3, s-parameter simulations while altering various hybrid-pi model parameters were discussed in order to determine which model parameter is responsible for the width-related changes in s-parameters. The simulation results that demonstrated the most promise of identifying the cause for the width effect were shown in Chapter 3, but this appendix shows the complete results from these simulations. For the most part, these results show that changes in these model parameters cause the s-parameters to move in a fashion that is not observed in measurement, thus eliminating these parameters from further consideration.

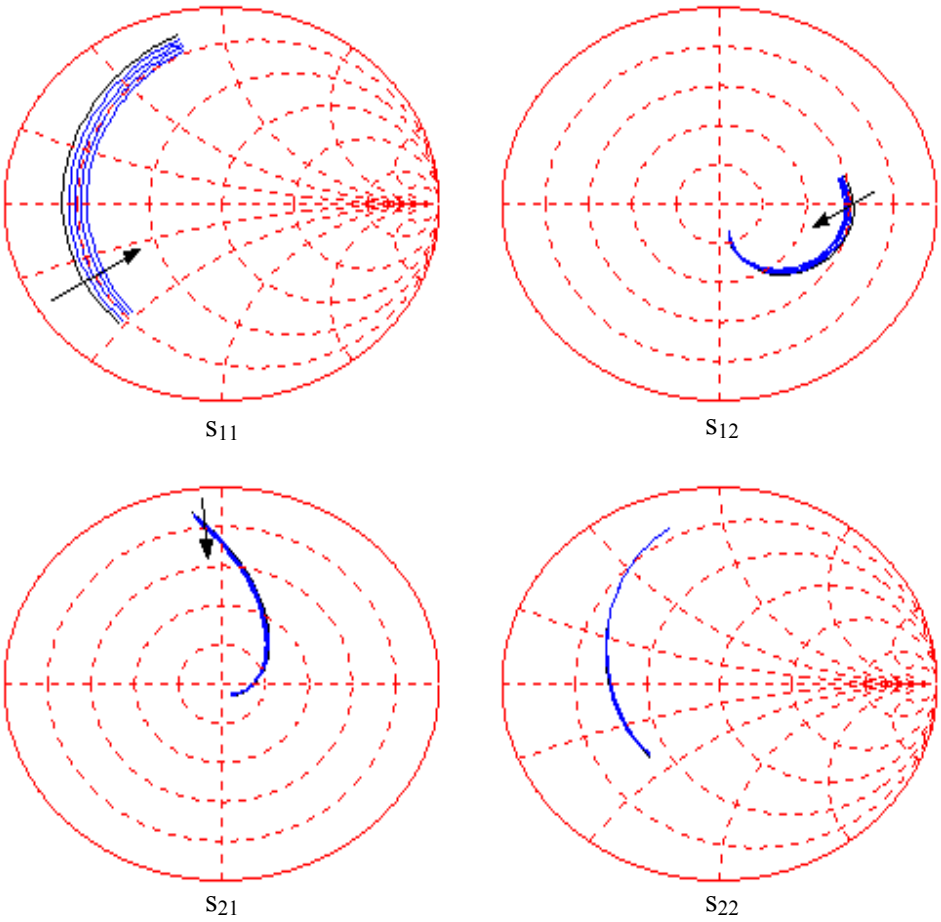


Figure A.1. Small-signal s-parameters of the W=4mm Root model with extra gate resistance (R_g). An increase in input resistance may be observed in the s_{11} plot as more R_g is added.

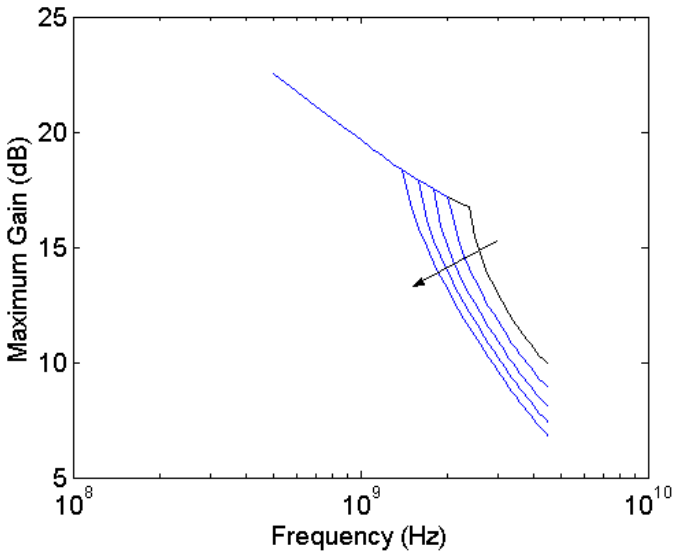


Figure A.2. Maximum gain plots calculated from the data shown in Figure A.1 (extra R_g). The low-frequency gain shows no change, but the cut-off frequency decreases with increased R_g.

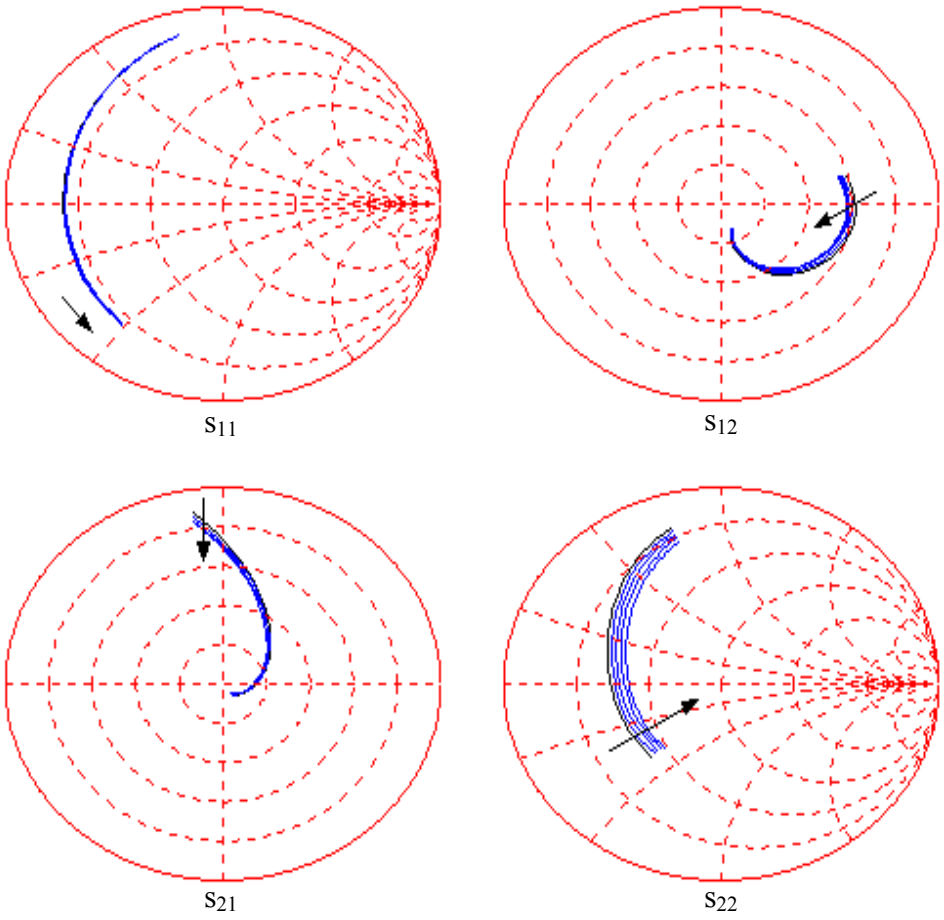


Figure A.3. Small-signal s-parameters of the W=4mm Root model with extra drain resistance (R_d). An increase in output resistance may be observed in the s_{22} plot as more R_d is added.

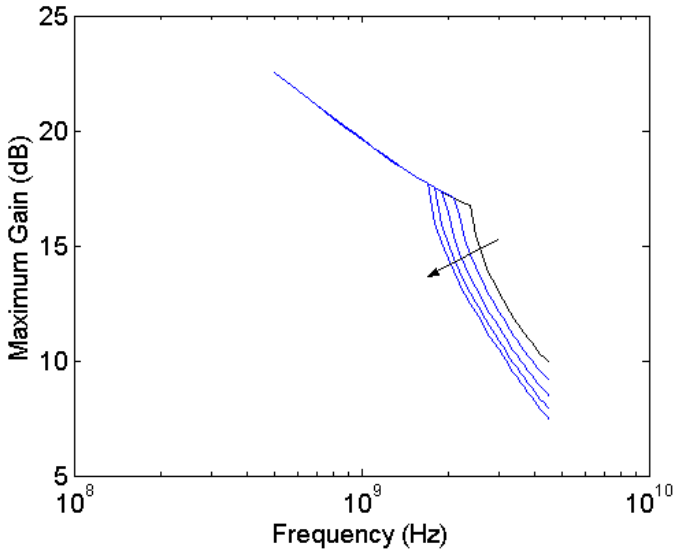


Figure A.4. Maximum gain plots calculated from the data shown in Figure A.3 (extra R_d). The low-frequency gain shows no change, but the cut-off frequency decreases with increased R_d.

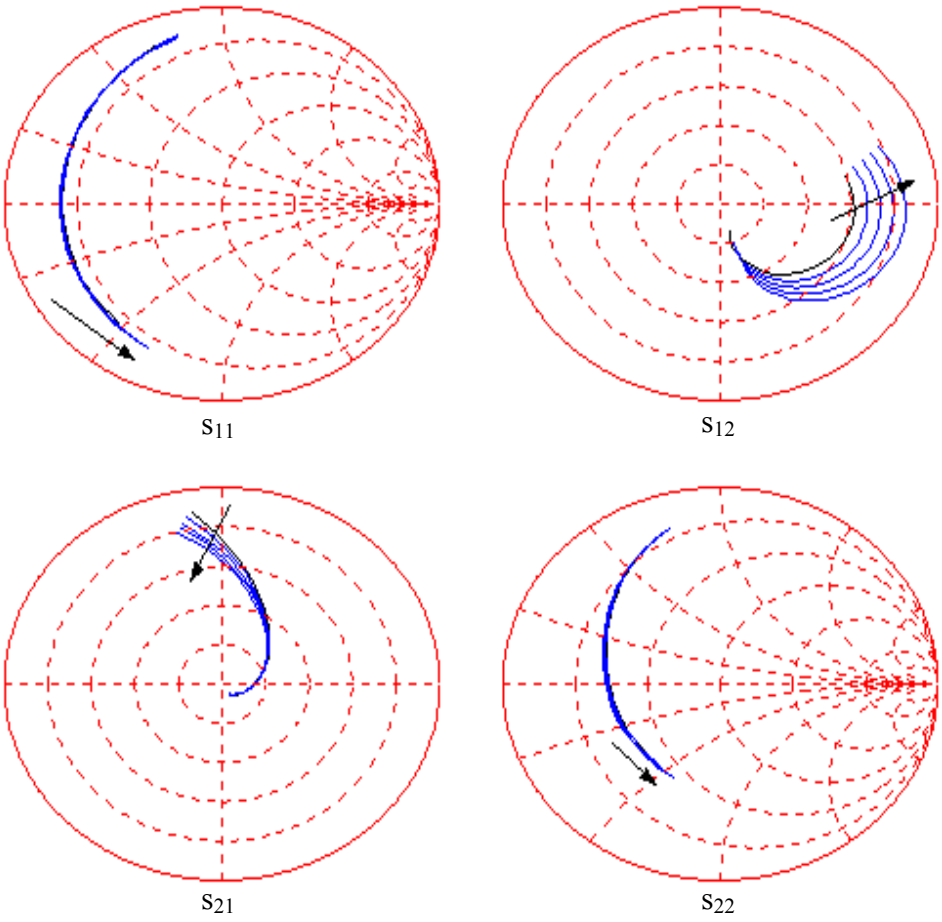


Figure A.5. Small-signal s-parameters of the W=4mm Root model with extra source resistance (R_s). A decrease in forward transmission is observed in s_{21}, an increase in reverse transmission is observed in the s_{12} plot, and negative input and output reactance is observed in the s_{11} and s_{22} plots, as more R_s is added.

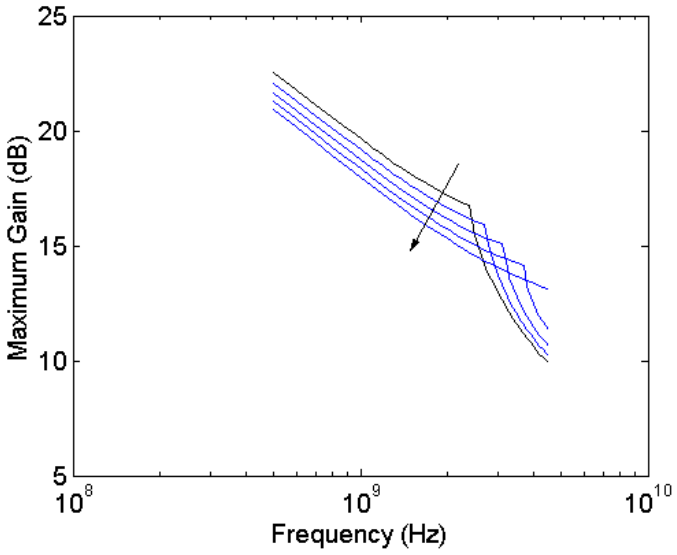


Figure A.6. Maximum gain plots calculated from the data shown in Figure A.5 (extra R_s). The low-frequency gain decreases, while the cut-off frequency increases with increased R_s.

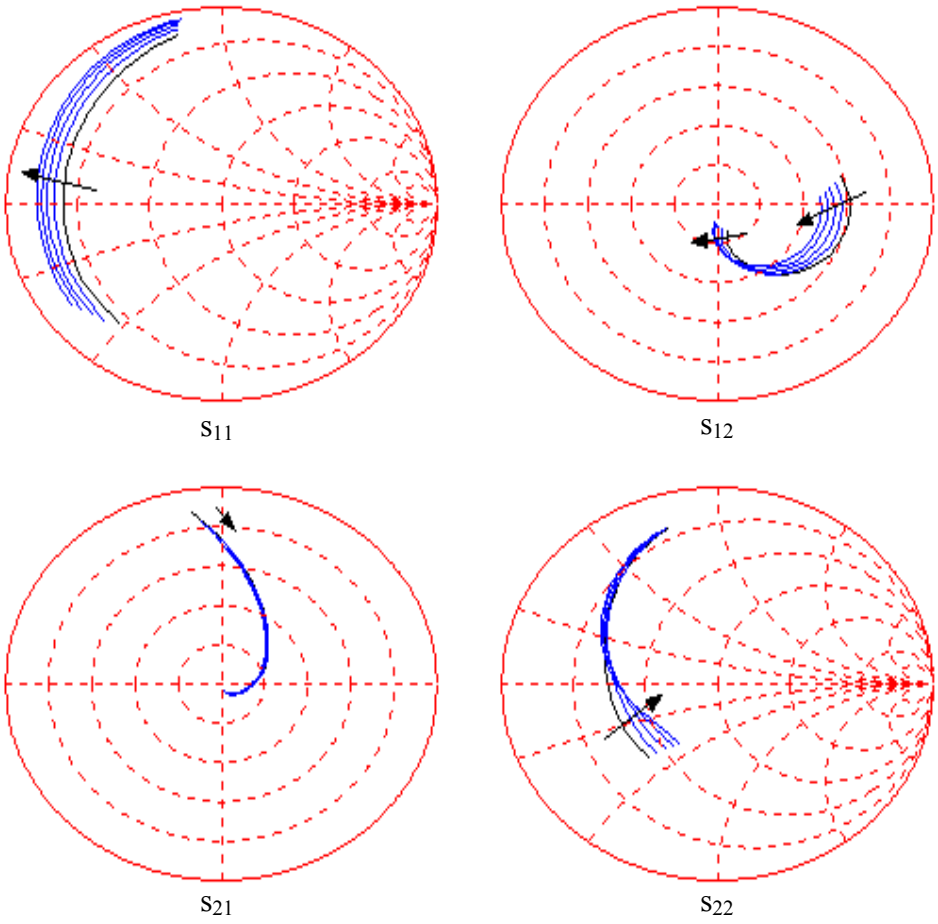


Figure A.7. Small-signal s-parameters of the W=4mm Root model with extra gate-source capacitance (C_{gs}). A decrease in forward transmission is observed in s_{21}, a decrease in input resistance is observed in s_{11}, and an increase in output resistance is observed in the s_{22} plot, as C_{gs} is increased

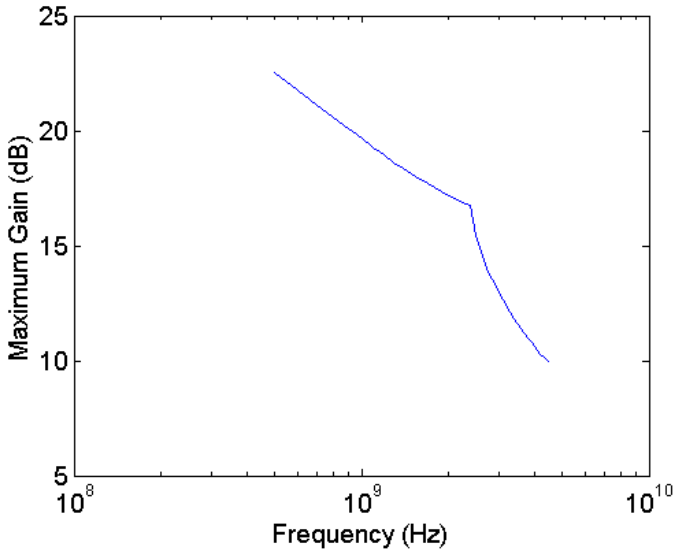


Figure A.8. Maximum gain plots calculated from the data shown in Figure A.7 (extra C_{gs}). Because a change in C_{gs} may be countered by a reactive matching circuit, no change is observed in the maximum gain plots.

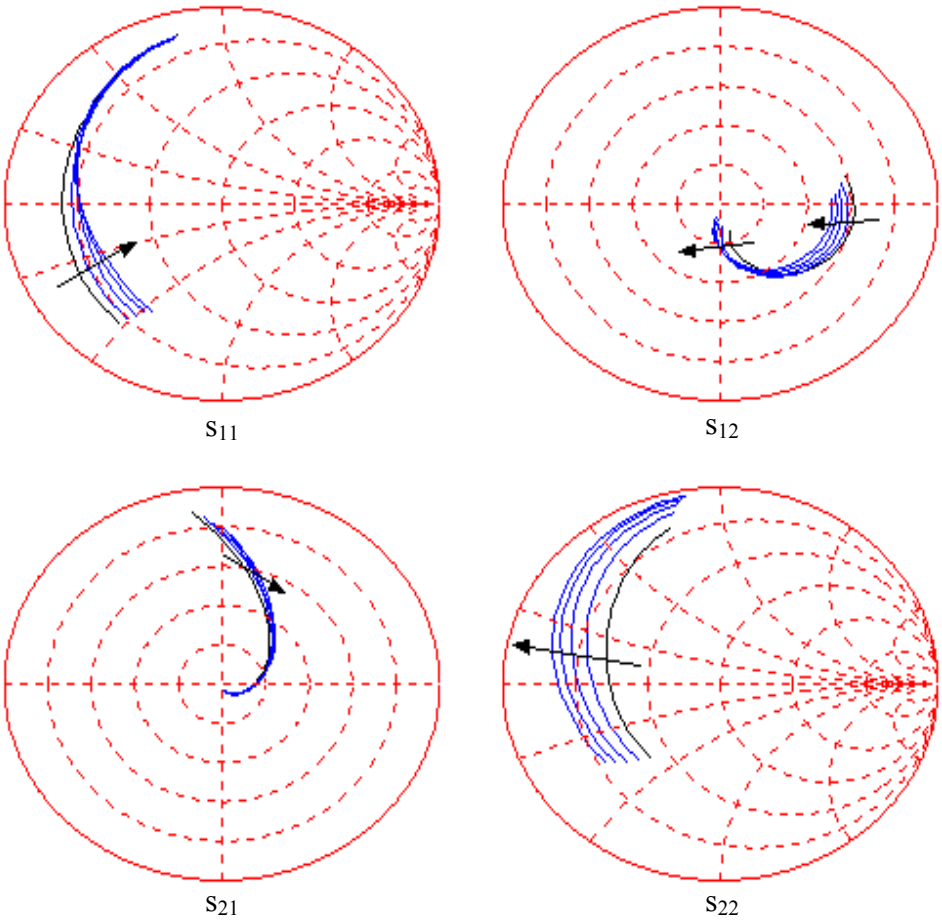


Figure A.9. Small-signal s-parameters of the W=4mm Root model with extra drain-source capacitance (C_{ds}). A decrease in forward transmission in s_{21}, a shift in reverse transmission in the s_{12} plot, an increase in the input resistance from s_{11}, and a decrease in the output resistance from s_{22} are observed as C_{ds} is increased.

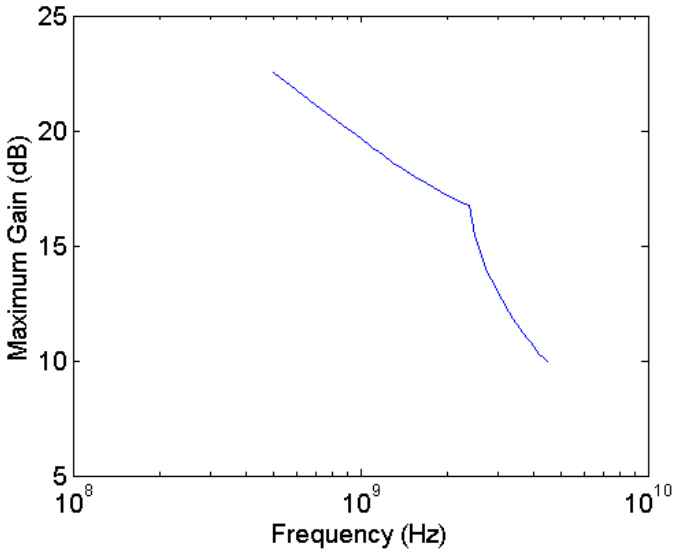


Figure A.10. Maximum gain plots calculated from the data shown in Figure A.9 (extra C_{ds}). Like Figure A.8, because a change in C_{ds} may be countered by a reactive matching circuit, no change is observed in the maximum gain plots.

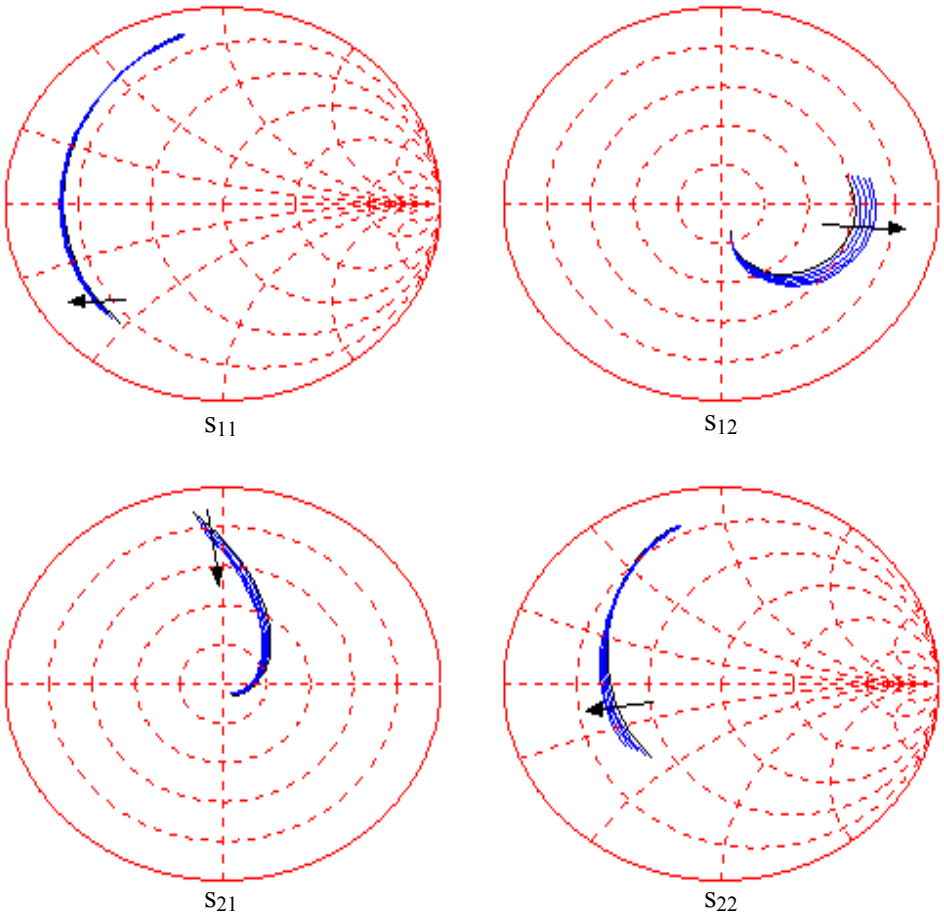


Figure A.11. Small-signal s-parameters of the W=4mm Root model with extra gate-drain capacitance (C_{gd}). A decrease in forward transmission is observed in s_{21}, an increase in reverse transmission is observed in the s_{12} plot, and small decreases in the input and output resistances are observed in the s_{11} and s_{22} plots, as more C_{gd} is added.

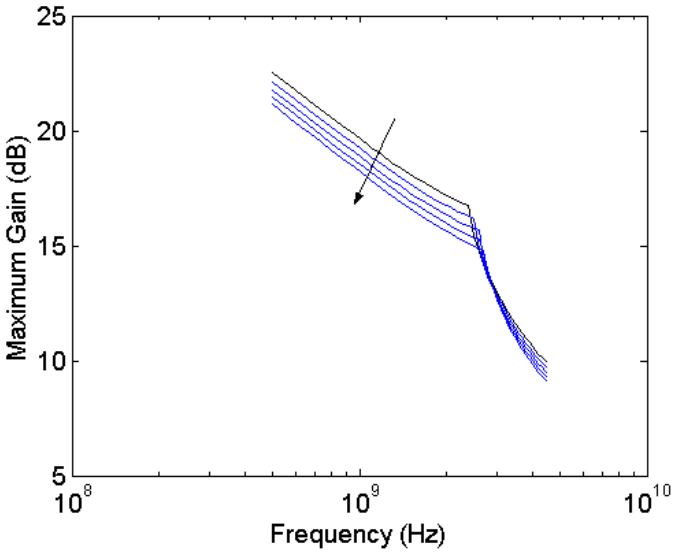


Figure A.12. Maximum gain plots calculated from the data shown in Figure A.11 (extra C_{gd}). The low-frequency gain decreases, and the cut-off frequency moves slightly higher as C_{gd} increases.

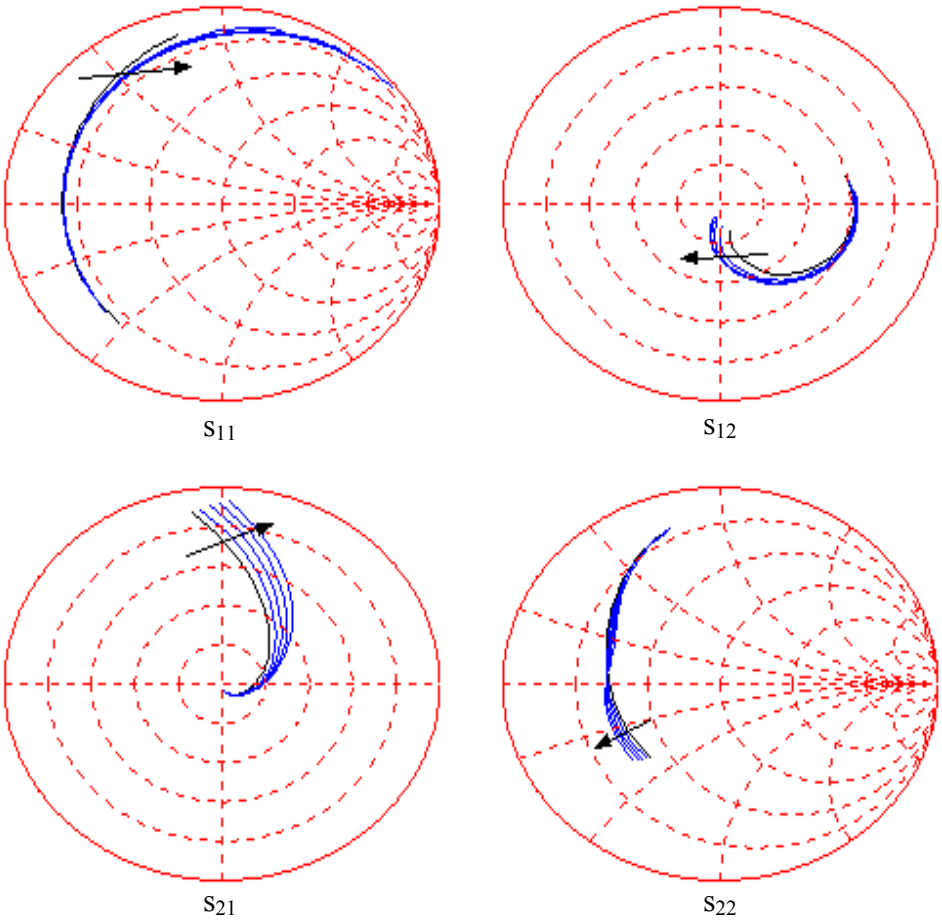


Figure A.13. Small-signal s-parameters of the W=4mm Root model with extra gate inductance (L_g). An increase in forward transmission is observed in the s_{21} plot and a large increase in the input reactance is observed in the s_{11} plot as L_g is increased.

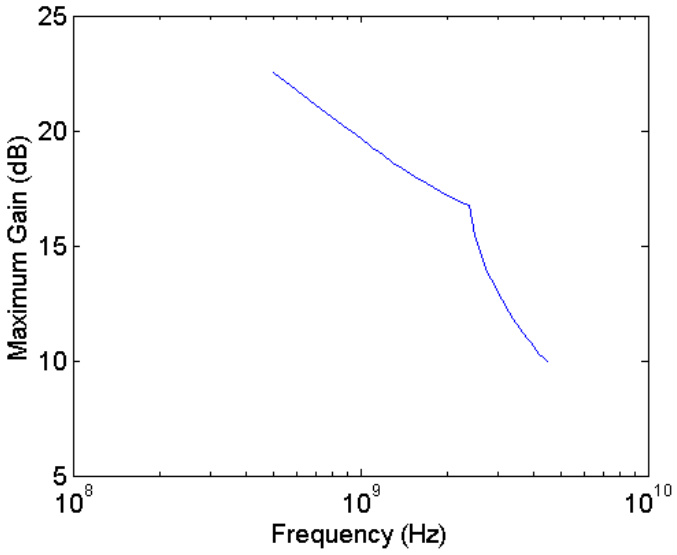


Figure A.14. Maximum gain plots calculated from the data shown in Figure A.13 (extra L_g). Because a change in L_g may be countered by a reactive matching circuit, no change is observed in the maximum gain plots.

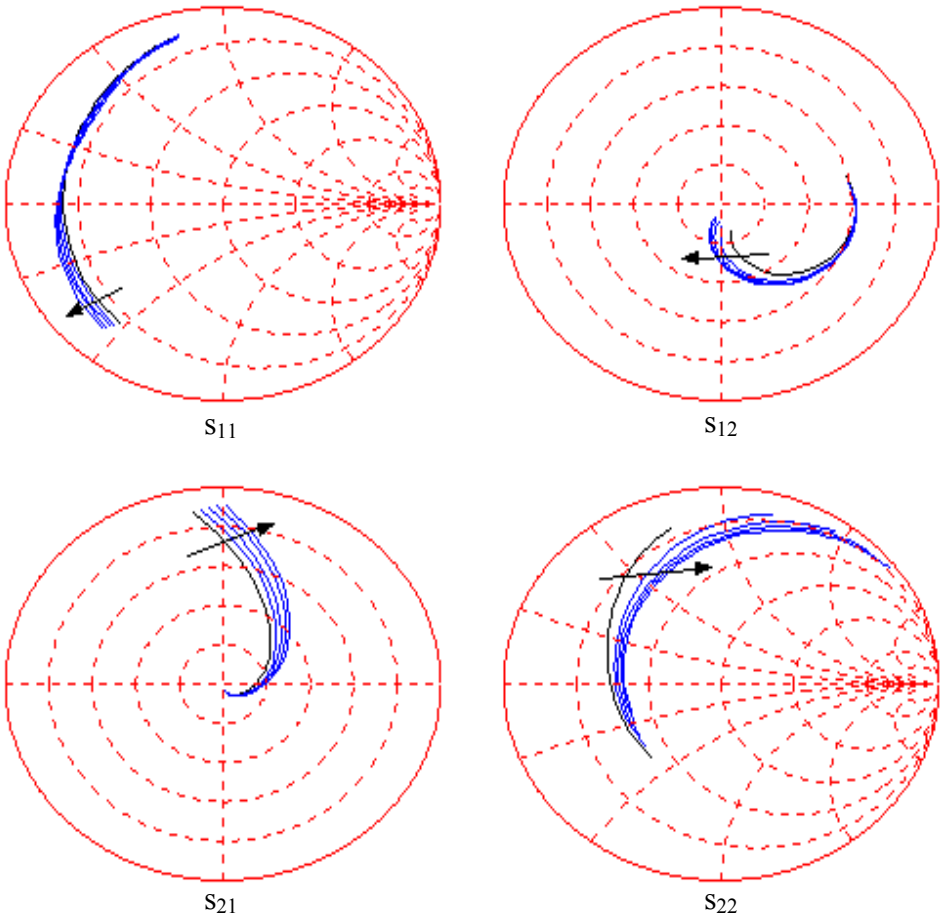


Figure A.15. Small-signal s-parameters of the W=4mm Root model with extra drain inductance (L_d). An increase in forward transmission is observed in the s_{21} plot and a large increase in the output reactance is observed in the s_{22} plot, as more L_d is added.

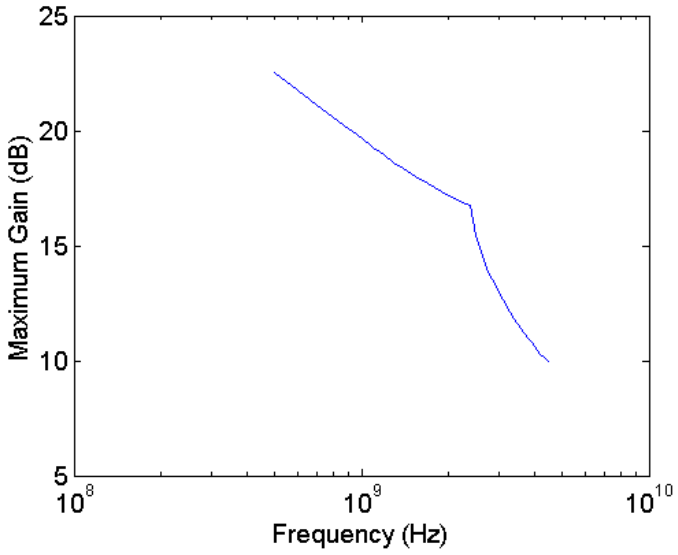


Figure A.16. Maximum gain plots calculated from the data shown in Figure A.15 (extra L_d). Because a change in L_d may be countered by a reactive matching circuit, no change is observed in the maximum gain plots.

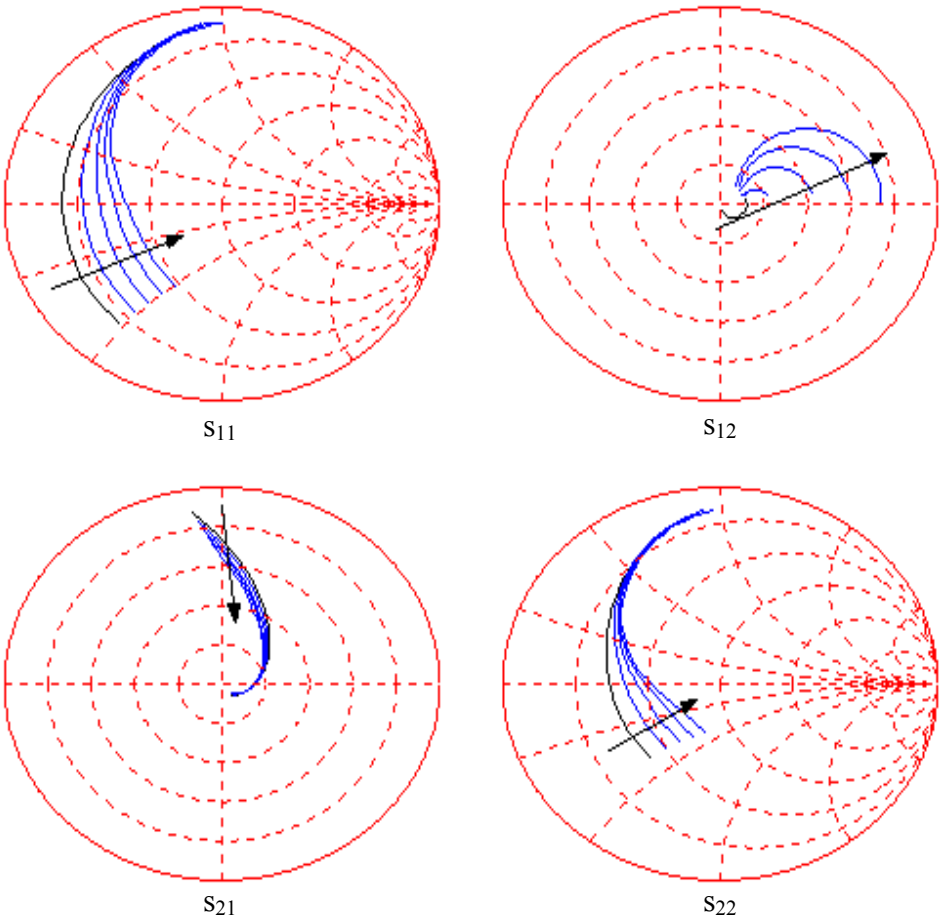


Figure A.17. Small-signal s-parameters of the W=4mm Root model with extra source inductance (L_s). A decrease in forward transmission is observed in the s_{21} plot, a large increase in reverse transmission is observed in the s_{12} plot, and increases in the input and output resistances is observed in the s_{11} and s_{22} plot, respectively, as more L_s is added.

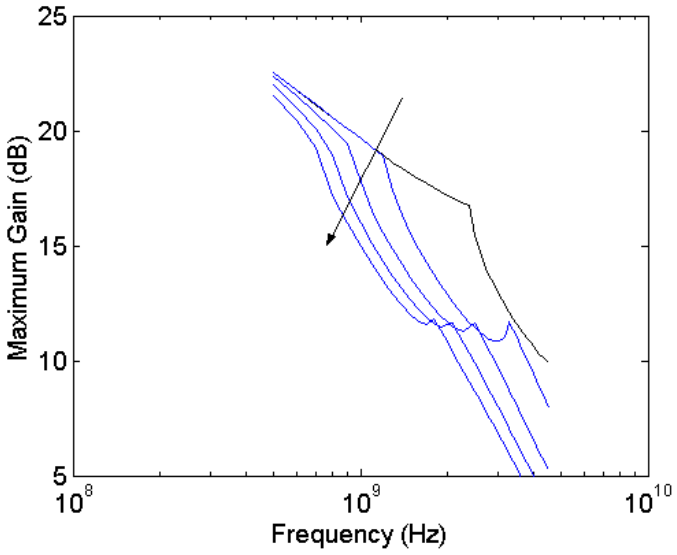


Figure A.18. Maximum gain plots calculated from the data shown in Figure A.17 (extra L_s). The low-frequency gain decreases, and the cut-off frequency moves drastically lower as L_s increases.

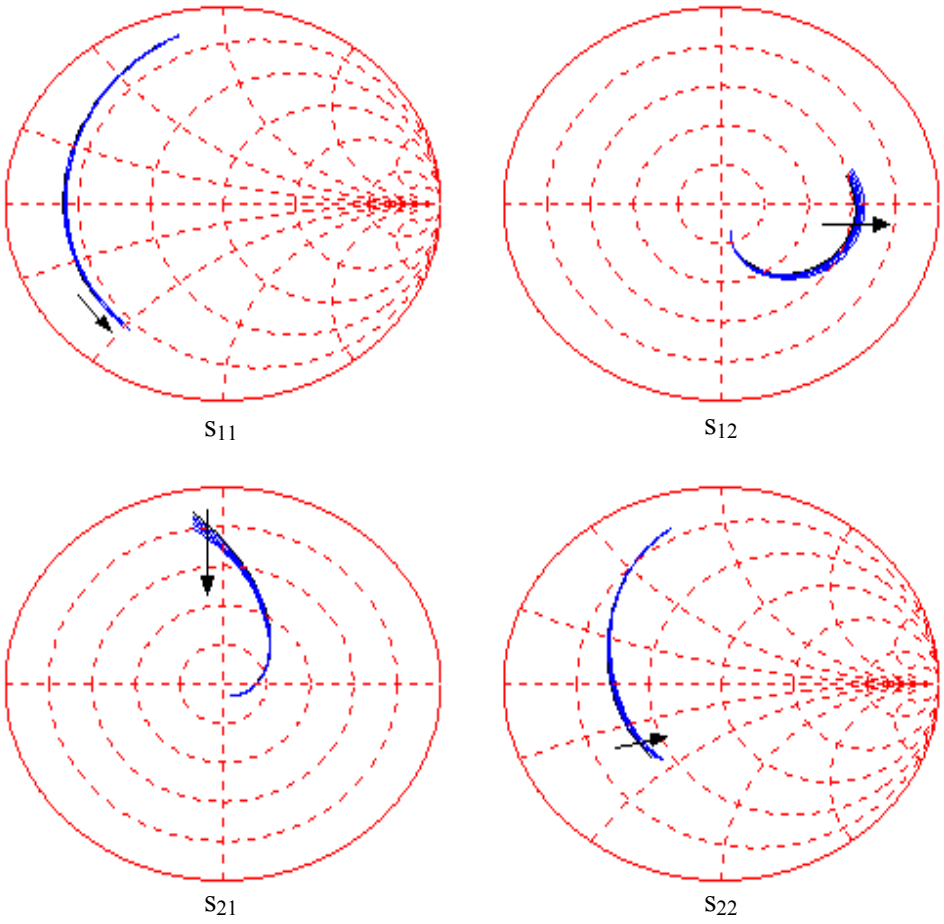


Figure A.19. Small-signal s-parameters of the W=4mm Root model with decreasing transconductance (g_m). A decrease in forward transmission in s_{21}, an increase in reverse transmission in s_{12}, a decrease in the input reactance in the s_{11}, and an increase in the output resistance in s_{22} are observed as more g_m is subtracted.

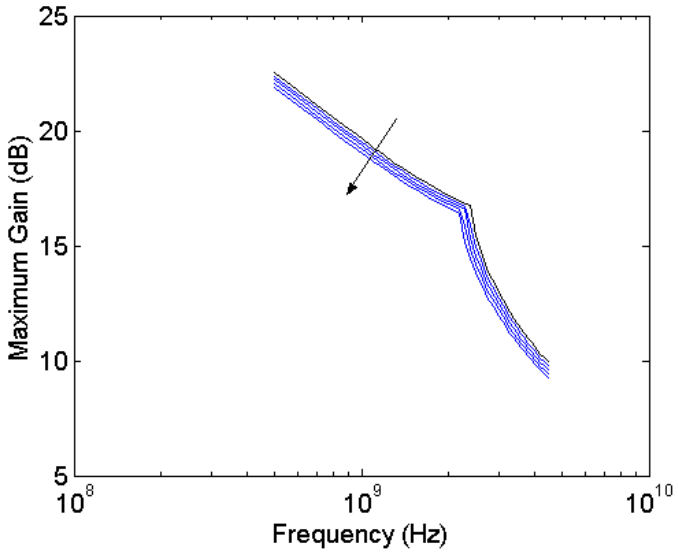


Figure A.20. Maximum gain plots calculated from the data shown in Figure A.19 (subtracting g_m). The curves shift downward slightly as g_m is decreased.

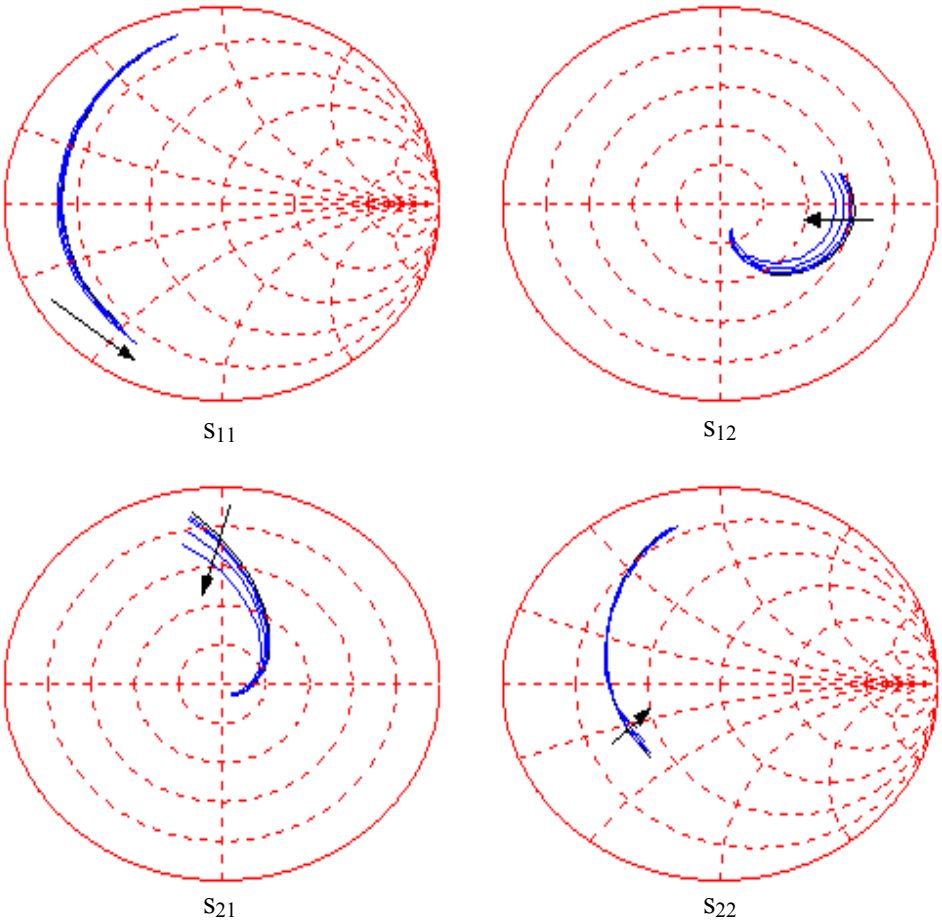


Figure A.21. Small-signal s-parameters of the W=4mm Root model with decreasing output resistance (R_{ds}). A decrease in forward transmission in s_{21}, a decrease in reverse transmission in s_{12}, a decrease in the input reactance in the s_{11}, and an increase in the output resistance in s_{22} are observed as smaller R_{ds} is placed in parallel with the existing R_{ds}.

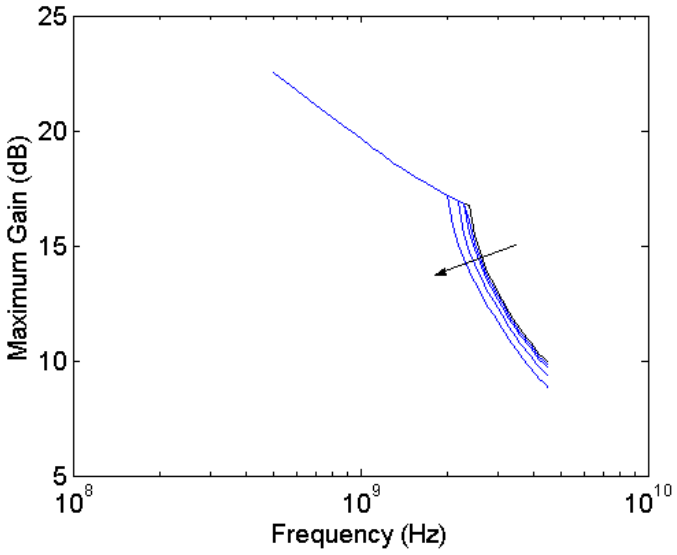


Figure A.22. Maximum gain plots calculated from the data shown in Figure A.21 (extra R_{ds}). The low-frequency gain shows no change, but the cut-off frequency decreases with decreased R_{ds}.

Bibliography

[1] A. Amerasekera and C. Duvvury, *ESD in Silicon Integrated Circuits*, Wiley, NY, 1995.

[2] A. Amerasekera, "RF Protection Circuit Design Approaches," Tutorial Notes, *EOS/ESD Symposium*, 1999.

[3] S. H. Voldman, "The State of the Art of Electrostatic Discharge Protection: Physics, Technology, Circuits, Design, Simulation, and Scaling," *IEEE J. of Solid-State Circuits,* vol. 34, no. 9, pp. 1272-1282, 1999.

[4] B. Kleveland, T. J. Maloney, I. Morgan, L. Madden, T. H. Lee, and S. S. Wong, "Distributed ESD Protection for High-Speed Integrated Circuits," *IEEE Electron Device Lett.,* vol. 21, no. 8, pp. 390-392, 2000.

[5] Bensinger, A. "No Clear Signal for Wireless," *Businessweek*, 30 Aug., 2002. <http://www.businessweek.com/investor/content/aug2002/pi20020830_2772.htm>.

[6] Gurley, W. J., "The next big thing? Try 802.11b," *News.com*, 19 Feb., 2001. <http://news.com.com/2010-1072-281406.html>.

[7] P. Leroux, and M. Steyaert, "High-performance 5.2GHz with on-chip inductor to provide ESD protection," *Electronics Letters*, vol. 37, no. 7, pp. 467-469, 2001.

[8] P. Leroux, J. Janssens, and M. Steyaert, "A 0.8dB NF ESD-protected 9mW CMOS LNA," *Tech. Digest of ISSCC,* 2001, pp. 410-411.

[9] K. Bock, "ESD Issues in Compound Semiconductor High-Frequency Devices and Circuits," *Proc. EOS/ESD Symp.,* 1997, pp. 1-12.

[10] B. Kleveland, and T. Lee, US Patent # 5,929,969, Oct. 1999.

[11] C. Richier, P. Salome, G. Mabboux, I. Zaza, A. Juge, and P. Mortini, "Investigation on Different ESD Protection Strategies Devoted to 3.3V RF Applications (2GHz) in a 0.18um CMOS Process," *Proc. EOS/ESD Symp.*, 2000, pp. 251-259.

[12] C. P. Wen, "Coplanar Waveguide: A Surface Strip Transmission Line Suitable for Nonreciprocal Gyromagnetic Device Applications," *IEEE Trans. Microwave Theory & Techniques*, vol. MTT-17, no. 12, pp. 1087-1090, 1969.

[13] C. Ito, K. Banerjee, and R. W. Dutton, "Analysis and Design of ESD Protection Circuits for High-Frequency/RF Applications," *IEEE Int. Symp. on Quality Electronic Design*, 2001, pp. 117-122.

[14] C. Ito, K. Banerjee, and R. W. Dutton, "Analysis and Optimization of Distributed ESD Protection Circuits for High-Speed Mixed-Signal and RF Applications," *Proc. EOS/ESD Symp.*, 2001, pp. 355-363.

[15] C. Ito, K. Banerjee, and R. W. Dutton, "Analysis and Design of Distributed ESD Protection Circuits for High-Speed Mixed-Signal and RFICs," *IEEE Trans. Electron Devices*, vol. 49, no. 8, pp. 1444-1454, 2002.

[16] E. L. Ginzton, W. R. Hewlett, J. H. Jasberg, and J. D. Noe, "Distributed Amplification," *Proc. IRE*, pp. 956-969, 1948.

[17] P. C. Magnusson, G. C. Alexander, and V. K. Tripathi, *Transmission Lines and Wave Propagation 3rd Ed.*, CRC Press, Boca Raton, FL, 1992.

[18] K.-H. Oh, C. Duvvury, C. Salling, K. Banerjee, and R. W. Dutton, "Non-uniform Bipolar Conduction in Single Finger NMOS Transistors and Implications for Deep Submicron ESD Design," *Int. Reliability Physics Symp. (IRPS)*, 2001, pp. 226-234.

[19] R. Anderson, L. Smith, and J. Gruszynski, "S-Parameter Techniques," Hewlett Packard Application Note 95-1, 1996.

[20] S. Ramo, J. R. Whinnery, and T. Van Duzer, *Fields and Waves in Communication Electronics*, *Ch. 5 Transmission Lines,* John Wiley, & Sons, NY, 1994.

[21] A. Amerasekera, "RF Protection Circuit Design Approaches," Tutorial Notes, *EOS/ESD Symposium*, 1999.

[22] Advanced Design System, Agilent Technologies v.1.3, 2000.

[23] W. J. Dally and J. W. Poulton, *Digital Systems Engineering,* Cambridge University Press, Cambridge, UK, 1998.

[24] K. Banerjee, A. Amerasekera, N. Cheung, and C. Hu, "High-Current Failure Model for VLSI Interconnects Under Short-Pulse Stress Conditions," *IEEE Electron Device Lett.*, vol. 18, no. 9, pp. 405-407, 1997.

[25] G. Carchon and B. Nauwelaers, "Accurate transmission line characterisation on high and low-resistivity substrates," *Proc. Microwaves, Antennas and Propagation*, 2001, pp. 285-290.

[26] B. Kleveland, T. H. Lee, and S. S. Wong, "50GHz Interconnect Design in Standard Silicon Technology," *Proc. MTT-S*, 1998, pp. 1913-1916.

[27] *The International Technology Roadmap for Semiconductors (ITRS),* 2001, <http://public.itrs.net/Files/2001ITRS/Home.htm>.

[28] I. Yoshida, M. Katsueda, M. Morikawa, Y. Matsunaga, T. Fujioka, M. Hotta, Y. Nunogawa, K. Kobayashi, S. Shimizu, and M. Nagata, "A 3.6V 4W 0.2cc Si Power-MOS Amplifier Module for GSM Handset Phones," *Int. Solid-State Circuits Conf.*, 1998, pp. 3.4.1-3.4.8.

[29] I. Yoshida, M. Katsueda, Y. Maruyama, and I. Kohjiro, "A Highly Efficient 1.9-GHz Si High-Power MOS Amplifier," *IEEE Trans. Electron Devices*, vol. 45, no. 4, pp. 953-956, 1998.

[30] I. Yoshida, "2-GHz Si power MOSFET technology," *Int. Electron Devices Meeting Technical Digest*, 1997, pp. 51 –54.

[31] R. F. Pierret, *Semiconductor Device Fundamentals*, Addison-Wesley, Reading, MA, 1996.

[32] A. B. Sproul and M. A. Green, "Improved value for the silicon intrinsic carrier concentration from 275-375K," *J. Applied Physics*, no. 70, pp. 846-854, 1991.

[33] Agilent HP8510 Network Analyzer

[34] "IRE Standards on Electron Tubes," *Proc. IRE*, vol. 45, pp. 983-1010, 1957.

[35] M. A. Karp, "Power Gain and Stability," *IRE Trans. Circuit Theory*, vol. CT-4, pp.339-340, 1957.

[36] J. M. Rollett, "Stability and Power-Gain Invariants of Linear Twoports," *IRE Trans. Circuit Theory*, vol. 9, no.1, pp.29-32, 1962.

[37] G. Vendelin, A. M. Pavio, and U. L. Rohde, *Microwave Circuit Design Using Linear and Nonlinear Techniques*, Wiley, New York, 1990.

[38] R. S. Pengelly, *Microwave Field-Effect Transistors--Theory, Design, and Applications, 2nd Ed.* Research Studies Press, Letchworth, England, 1986.

[39] R.Mavaddat, Network Scattering Parameters, World Scientific Publishing, Singapore, 1996.

[40] Agilent Application Note AN154, S-parameter Design, Agilent Technologies, 2000.

[41] Maury Microwaves Automatic Tuner System

[42] D. E. Root and S. Fan, "Experimental Evaluation of Large-Signal Modeling Assumptions Based on Vector Analysis of Bias-Dependent S-Parameter Data from MESFETs and HEMTs," *IEEE MTT-S Digest*, 1992, pp. 255-258.

[43] D. E. Root, "Introduction to large-signal models - ISCAS2001 Tutorial/Short course & special session on high speed devices & modeling," *Int. Symp. on Circuits and Systems Tutorial Guide,* 2001, pp. 2.3_1 -2.3_7.

[44] D. E. Root, "Advanced nonlinear modeling - ISCAS2001 Tutorial/short course & special session on high speed devices & modeling," *Int. Symp. on Circuits and Systems Tutorial Guide,* 2001, pp. 2.7_1 -2.7_8.

[45] Advanced Design System, Agilent Technologies v. 1.5, 2001.

[46] G. M. Dolny, G. E. Nostrand, K. E. Hill, "The Effect of Temperature on Lateral DMOS Transistors in a Power IC Technology." *IEEE Trans. Electron. Devices* vol. 39, no. 4, pp. 990-995, 1992.

[47] P. Khandelwal, M. Trivedi, K. Shenai, and S. K. Leong, "Thermal and Package Performance Limitations in LDMOSFET's for RFIC Applications," *IEEE Trans. Microwave Theory & Techniques*, vol. 47, no. 5, pp. 575-585, 1999.

[48] S. C. Cripps, *RF Power Amplifiers for Wireless Communications*, Artech House, Norwood, MA, 1999.

[49] E. Gebara, J. Laskar, M. Harris, and T. Kikel, "Development of Cryogenic Load-Pull Analysis: Power Amplifier Technology Performance Trends," *IEEE MTT-S Digest*, 1998, pp. 1663-1666.

[50] E. Yanokura, T. Seki, I. Takei, Y. Maruyama, Y. Fujita, M. Katsueda, I. Yoshida, M. Ohnishi, and K. Sekine, "A 1.5 GHz, 35-W Si-MOSFET with an Internal Matching Circuit," *Int. Symp. on Power Semiconductor Devices & ICs*, 1994, pp. 425-429.

[51] F. W. Grover, *Inductance Calculations: Working Formulas and Tables*, Dover Publications, Inc., NY, 1946.

[52] X. Qi, C. P. Yue, T. Arnborg, H. Soh, H. Sakai, Z. Yu, R. W. Dutton, "A Fast 3-D Modeling Approach to Electrical Parameters Extraction of Bonding Wires for RF," *IEEE Trans. on Advanced Packaging*, vol. 23, no. 3, pp. 480-488, 2000.

[53] M. Kamon, M. J. Tsuk, and J. White, "FASTHENRY: A Multipole Accelerated 3-D Inductance Extraction Program," *IEEE Trans. MTT* vol. 42, no. 9, pp. 1750-1758, 1994.