K-Band Low-Noise Amplifier Design in CMOS Technology

by

Dustin Dunwell

A thesis submitted to the
Department of Electrical and Computer Engineering
in conformity with the requirements for
the degree of Master of Science (Engineering)

Queen's University
Kingston, Ontario, Canada
August 2006

Abstract

This thesis investigates some the challenges associated with high-speed circuit design in silicon CMOS. It begins with the use of on-chip inductor structures, which consume a large chip area, are difficult to accurately model and create significant sources of loss and noise at high-frequencies. By modifying a standard microstrip transmission line structure to include a series-stub section, a novel inductor structure is presented with dimensions that are easily implementable on-chip. Measured results show that this structure is able to provide Q factor improvements of up to 20% over spiral inductors at frequencies above 25 GHz.

Secondly, the use of accumulation-mode MOS (AMOS) varactors is investigated at frequencies above 20 GHz and a new model for their behaviour is presented. This model relies on physical equations for each of its components, making it easily adaptable to any technology node or layout dimensions. Comparison to measured results shows good agreement even for capacitances on the order of tens of femtofarads.

Finally, two low-noise amplifiers (LNAs) are presented in order to demonstrate the usefulness and functionality of the above concepts. The first amplifier makes use of the SSTL inductor structures to implement a two-stage cascode device for operation at 23 GHz. This cascode configuration improves reverse isolation helping to stabilize the amplifier and simplify matching network design. Although cascode topologies

incur a penalty to the noise figure, this penalty is offset by the improved modeling and Q factor of the SSTL inductors, allowing the amplifier to achieve a measured NF of 5.9 dB and a gain of 13.5 dB at 23 GHz.

The second amplifier uses the varactor models to implement a capacitive neutralization scheme to improve the reverse isolation of a differential LNA. This topology has not been reported in a high-frequency amplifier to date as very few publications exist examining differential LNA performance at frequencies above 20 GHz. Measured results show that while the gain of 5.2 dB of this amplifier falls short of that predicted by simulation results, it is still able to produce a NF of only 4.5 dB at approximately 23 GHz.

Acknowledgments

This section is to acknowledge and thank the many parties that generously assisted in the completion of this thesis.

Firstly, the interest and enthusiasm shown by Dr. Brian Frank in his teaching has played a large part in sparking my own interest in very high speed silicon circuit design. His continuing guidance and support, not to mention his problem solving skills, have been instrumental in the successful completion of my graduate studies in the area.

Secondly, thanks is owed to the other graduate students in the Very High Speed Circuits research lab at Queen's University who were never too busy to lend a helping hand and whose company during long hours at the lab was a comfort during the most stressful times.

Financial support for much of this work was sponsored by the Natural Sciences and Research Council (NSERC) as well as Queen's University.

I would also like to thank my family whose tolerance in my continued absence and busy schedule has not gone unnoticed. Finally, I would like to thank my girlfriend and best friend for her understanding and perseverance during the past two years. Kelly: your pride and devotion means the world to me. You have never and will never come second.

Acronyms, Abbreviations, and Symbols

Abbreviation	Definition
γ	channel thermal noise coefficient
α	noise parameter
δ	gate noise coefficient
ϵ_{ox}	oxide permittivity
ϵ_{si}	silicon permittivity
μ_n	electron mobility
ω	frequency in rad/s
ADS	Advanced Design System (from Agilent)
AMOS	Accumulation-mode metal oxide semiconductor
c	correlation coefficient
CAD	Computer aided design
CG	Common gate
CS	Common source
CMOS	Complimentary metal oxide semiconductor

CNM	Classical noise matching
CPW	Coplanar waveguide
CTM	Capacitor top metal
DC	Direct current
DSP	Digital signal processing
EM	Electromagnetic
f	frequency in Hertz
f_{MAX}	Maximum frequency of oscillation
f_T	Unity current gain frequency
FET	Field effect transistor
GaAs	Gallium arsenide
HFNM	High frequency noise matching
InP	Indium phosphide
k	Boltzmann's constant
LNA	Low noise amplifier
MAG	Maximum available gain
MIM	Metal-insulator-metal
MOSFET	Metal oxide semiconductor field effect transistor
MOS	Metal oxide semiconductor
MPG	Metal 1 patterned ground
NF	Noise figure
N_{well}	n well carrier concentration
N_{sub}	p substrate carrier concentration
N_i	Intrinsic silicon carrier concentration

NSERC	Natural Sciences and Research Council
PCNO	Power constrained noise optimization
PPG	Polysilicon patterned ground
Q	Quality factor
q	Electron charge
RF	Radio Frequency
SiGe	Silicon germanium
SOI	Silicon on insulator
SSTL	Series stub transmission line
T	Temperature in Kelvin
VCO	Voltage controlled oscillator
VNA	Vector network analyzer

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Wireless Communication

The creation of the wireless communications industry, sparked by the first successful demonstration of the cellular phone in 1973, has grown dramatically and today consists of many companies, employing approximately a quarter of a million people, grossing over $100 million per year and servicing over 200 million subscribers [1] in the United States alone. The extremely competitive nature of this market has led to a surge in technological development as companies attempt to provide faster, cheaper and more versatile wireless communication options not only for voice transmission, but for data as well.

One consequence of this technological push is that digital communications, which offer improved noise immunity and more efficient spectrum usage, are becoming more prevalent in consumer applications. The digital signal processing (DSP) required in such high-speed communication systems is almost always implemented on silicon

Complimentary Metal Oxide Semiconductor (CMOS) technology since its low substrate resistivity helps to prevent latchup in the transistors used. As a result, CMOS fabrication technology has received a great deal of attention in recent years and has now become a very reliable and relatively inexpensive option for integrated circuit design.

A second consequence of the increased competition in the wireless market is the fact that increased data transmission rates often require increased bandwidth. This, coupled with the fact that emerging applications must use frequency bands that are not already in use, means that regulating agencies around the world are being forced to allocate frequency spectrum at ever increasing frequency ranges, which are not currently in use.

By combining these two consequences, it becomes apparent that there exists a great deal of demand for the integration of analog and digital circuits designed to operate at increasingly higher frequencies. Unfortunately, while CMOS technology lends itself well to digital applications, its use in radio frequency (RF) and millimeter wave applications poses several challenges to the circuit designer.

1.2 The RF CMOS Challenge

In the recent past, RF and millimeter-wave circuit design was limited to specialized, high-performance substrate materials such as silicon-germanium (SiGe), indium-phosphide (InP) or gallium-arsenide (GaAs). While these materials provide decidedly superior device physics in terms of substrate resistivity and carrier mobility, they are significantly more expensive and difficult to integrate with digital circuitry than silicon. This has caused silicon to maintain its appeal for RF and millimeter wave circuit

applications and, thanks to the steady and fast-paced evolution of silicon technology fueled by the staggering investment by the digital community, silicon CMOS has emerged as a viable and inexpensive option for high-speed circuit design [2].

Validating this statement is the fact that emerging sub 0.1 μm CMOS technologies are able to demonstrate unity current gain frequencies, f_T, and maximum oscillation frequencies, f_{MAX}, exceeding 150 GHz [3] while other technologies with gate lengths greater than 0.1 μm have displayed f_T and f_{MAX} values greater than 40 GHz. These parameters illustrate the potential usefulness of CMOS technology in high-speed circuit designs. Also helping this case is the fact that new CMOS technologies commonly feature an increased number of metal layers, distancing passive structures from the substrate and increasing their Q factors.

As an added advantage, silicon Metal Oxide Semiconductor Field Effect Transistors (MOSFETs) benefit from lower transconductance and capacitance per unit gate width and from the layout dependence of gate resistance to achieve lower noise figures than older III-V semiconductor technologies of comparable or higher f_T and f_{MAX} [3]. However, these benefits are extremely difficult to achieve in practice due to fact that the very high noise resistance, low optimum noise admittance and strong layout dependence make CMOS circuits more sensitive than III-V technologies to process variations, impedance mismatch, and model inaccuracies [3]. As a result, there is still a great deal of progress to be made in high-speed CMOS circuit design, especially those that focus on minimizing the noise produced.

1.3 Literature Review

Before outlining the contents and major contributions of this thesis, it is useful to examine the current state of the research presented in relevant publications. This section is organized into subsections — one for each major topic covered in the thesis — and presents not only the evolution of each topic and the current state of the art, but also outlines areas in each section that require further research in the hopes of improving aspects of their performance.

1.3.1 Inductors

The use of integrated circuits requires a two-dimensional inductor implementation and planar spiral inductors, as illustrated in Figure 1.1, have long served as an effective method of achieving this. However, despite their common use in integrated circuit design, accurate modeling of their inductive properties is a difficult problem for designers and research working toward this goal has been conducted since 1980 [4].

More recently, advances in the CAD tools used to help model the properties of integrated circuit components, such as three-dimensional numerical simulators, have enabled the accurate prediction and Q factor optimization of spiral inductors for frequencies up to 5 GHz [5]. As CMOS operating frequencies continue to increase, however, spiral inductor modeling techniques have struggled to keep up due to the complexity involved in modeling the substrate behaviour at these frequencies. Only very recently have spiral inductor extraction and modeling techniques been presented that can produce reliable results at frequencies up to 20 GHz [6].

Despite these advances in modeling and optimization, the high-frequency substrate

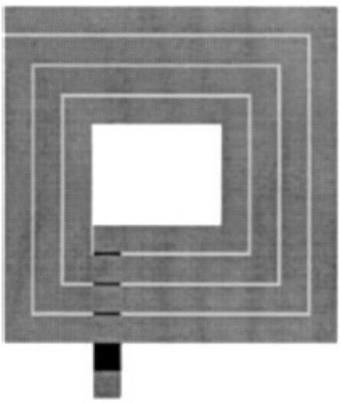

Figure 1.1: Layout of a standard 3.75 turn square spiral inductor.

coupling that inevitably occurs with the use of on-chip spiral inductors results in low Q factors in CMOS technology. Although the reported Q factors can reach values of nearly 20 in some cases, such as in [6] for 0.13 μm technology with a 3 μm thick top metal layer, this has only been achieved for frequencies below 10 GHz. As operating frequencies increase beyond 20 GHz, which is usually well beyond the maximum Q frequency, Q factors are typically reduced to less than 7. Despite the introduction of symmetrical geometries or patterned ground shielding [7], which can help mitigate these substrate effects, the Q factors that are achievable at frequencies above 20 GHz leave much to be desired.

As an alternative, the use of simple microstrip transmission lines as a means of implementing an on-chip inductance is receiving renewed interest. In fact, recent research has successfully reported:

- Accurate characterization of the high frequency behaviour of transmission lines [8].

- The use of non-standard structures, such as ground shielding beneath coplanar waveguide (CPW) transmission lines to improve performance [9].
- Implementation of microstrip transmission lines as inductors in a 60 GHz CMOS receiver [10].

Although promising at high-frequencies, the length of the transmission line required to implement suitable inductances at approximately 25 GHz in 0.18 μm CMOS technology can be prohibitive and this issue must be addressed before their use can become commonplace in integrated circuit design at these frequencies.

1.3.2 Varactors

Thanks to the addition of a capacitor top metal (CTM) layer to most recent CMOS technologies, which provides an extra metal layer placed extremely closely to a standard metal layer so as to increase the capacitance between these sections, metal-insulator-metal (MIM) capacitors can be easily implemented on-chip. However, in some applications it is advantageous or even necessary to use a variable capacitor, or varactor, to be able to tune the capacitance produced.

Traditionally, this type of capacitance has been most easily implemented using a reverse-biased $p - n$ junction embedded in the silicon substrate. In such a device, the depletion region formed at the $p - n$ junction varies with the voltage difference seen across the junction, allowing the resulting capacitance seen across the depletion region to be varied by changing the applied bias voltage. Although modeling the capacitance of such a device is relatively straight forward, it suffers from a higher power consumption and phase noise and lower tuning range than the metal oxide semiconductor (MOS) varactor proposed in [11]. This device relies on the voltage

sensitive capacitance seen between the gate and channel of a standard MOSFET to implement a variable capacitance.

This concept has since evolved to the Accumulation-mode MOS (AMOS) varactor, which is similar to an n-channel MOSFET, but differs in that it is fabricated in an n well as opposed to the normal p type substrate [12]. This eliminates the parasitic $p-n$ junction capacitances at the source and drain diffusion regions that would otherwise limit the tuning range. This AMOS varactor has since found considerable application in RF CMOS applications, such as voltage controlled oscillators (VCOs), operating at frequencies of up to 40 GHz [13].

Unfortunately, as was the case with spiral inductors, the modeling of these devices poses a significant challenge, especially at high frequencies. In [14] a physics based model for the AMOS varactor in 0.25 μm CMOS was presented, accurately predicting capacitances between 0.6 and 2.6 pF. Unfortunately, this model was created for a non-standard structure using shallow trench isolation regions to separate the channel from the n^+ diffusion regions and was only presented for frequencies up to 5 GHz. It also used different models for different modes of operation, making its implementation in circuit simulators challenging. In [15] an AMOS varactor model for 0.18 μm CMOS was created by extracting lumped element values from measured data. This model displayed good results at higher frequencies, but achieves self resonance at a frequency of 18 GHz. In addition, the curve fitting parameters used in much of the extraction makes adaptation of the model to other varactor dimensions or CMOS gate lengths extremely difficult.

Currently, no AMOS varactor models have been presented for frequencies above

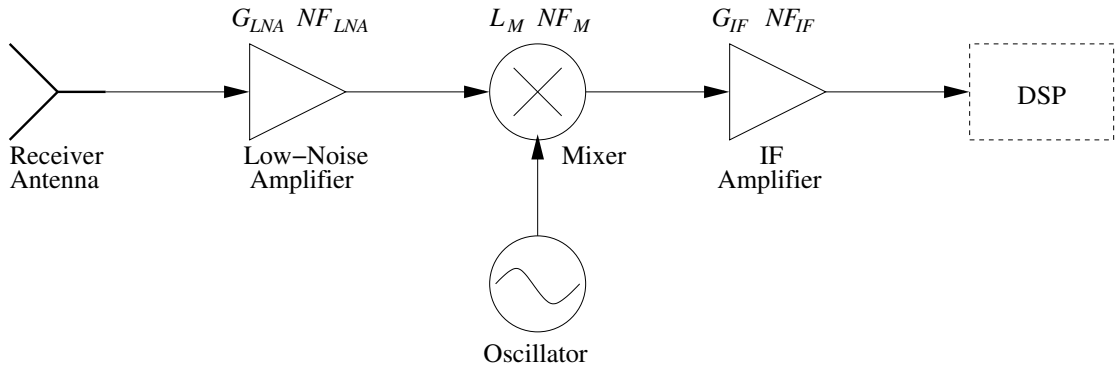

Figure 1.2: Typical receiver architecture showing relevant gains and noise figures.

20 GHz in a standard CMOS technology, making their implementation at these frequencies extremely difficult, especially for very small capacitance values. Further work would have to be completed in this area in order to allow such varactors to be implemented with a reasonable level of confidence.

1.3.3 Low-Noise Amplifiers

The low-noise amplifier (LNA) is an integral part of any wireless receiver. Equation 1.1 calculates the noise figure for a typical receiver architecture, NF_{rec}, as shown in Figure 1.2

$$NF_{rec} = NF_{LNA} + \frac{NF_M - 1}{G_{LNA}} + \frac{NF_{IF} - 1}{G_M G_{LNA}} \tag{1.1}$$

where the loss L or gain G and noise figure, NF, of each component is displayed in Figure 1.2. This illustrates the fact the noise and gain performance of the LNA has a direct and significant effect on the noise figure of the entire receiver. As such, LNA design has received a great deal of attention in recent publications and several successful designs have been reported at frequencies above 20 GHz in 0.18 μm CMOS technology.

These amplifier designs typically employ one of four common design techniques,

which are summarized in [16], to achieve minimum noise while also considering power constraints and gain. Of the successful designs reported for frequencies between 20 and 26 GHz, measured noise figures vary widely from values as low as 3.9 dB [17] to others as high as 8.9 dB [18]. Although the circuit topologies used for these designs vary widely, the most common implementations are single-ended amplifiers that focus on achieving the lowest possible noise while maximizing gain and minimizing power consumption. This is most often accomplished with the help of inductive degeneration in the source of the amplifier, as described in [16]. Unfortunately, [19] illustrates that the low Q factor of the on-chip inductors used in these topologies has the inherent effect of increasing the noise seen at the input of the amplifier.

Despite the seemingly abundant high-frequency LNA publications, one issue that has been commonly ignored is that of the high-frequency feedback path created by the gate to drain capacitance, C_{gd}, common to all CMOS transistors. With the exception of [20] and [21], which use cascode topologies in 0.12 μm and 90 nm CMOS technologies, respectively, little attention has traditionally been paid to the reverse isolation of the amplifier. This can be a concern for amplifier stability as well as matching network design, especially at frequencies above 20 GHz.

Another area that has seen little attention in high-frequency LNA publications is the use of differential designs. One design for use at 20 GHz is presented in [22], but displays large discrepancies between the modeled and measured results and reports a noise figure of 5.5 dB. Although differential topologies often require higher power consumption and chip space, their ability to reject common-mode disturbances means that their noise performance should compete with the best reported single-ended implementations.

1.4 Thesis Overview and Major Contributions

This thesis is organized topically by chapter, each chapter presenting a study of a specific component or entire circuit block designed for high-frequency operation in 0.18 μm CMOS. Chapters two and three present novel structures or improved models for common passive devices, while chapters four and five examine specific circuits that utilize the structures and models presented in chapters two and three.

Chapter two highlights the inaccuracies encountered in the modeling of spiral inductors at frequencies above 20 GHz. It then presents a novel series stub transmission line (SSTL) structure, which is based on microstrip transmission lines. This design simplifies modeling by limiting coupling to the substrate and addresses the issue of prohibitively long line lengths by using a three-dimensional stub section of transmission line, allowing for relatively large inductances to be easily implemented on-chip. As a result of the ease in modeling this structure, optimization was possible and both simulated and measured Q factors show an improvement over their spiral inductor counterparts.

Chapter three explains the functionality of recently developed AMOS varactors. Although these varactors are extremely useful since their variable capacitances can be implemented for a wide range of capacitance values, no circuit model has yet been published for standard CMOS technology at frequencies above 20 GHz. This chapter presents a model for these frequencies that relies heavily on semiconductor physics and validates it through comparison with measured data. The model is adaptable to any layout dimensions and any CMOS technology and can be easily implemented in circuit simulation.

Chapter four presents a two-stage, single-ended LNA design for operation at 23

GHz. This design employs a cascode topology in an effort to improve the reverse isolation of the amplifier, helping to improve stability and ease the design of the input and output matching networks. It also uses the SSTL inductor structures presented in Chapter 2 in an effort to show that their accurate inductance modeling and improved Q factor can be used to improve the noise performance of the LNA. After overcoming issues with instability created by parasitics associated with the testing equipment, the measured LNA shows good results with a noise figure of 5.9 dB and a gain of 13.5 dB at 23 GHz. These results are not the best yet reported, but rank favourably within the range of other published LNAs in this frequency range.

Chapter five considers a two-stage, differential LNA design for operation at 23 GHz. This design also addresses the issue of reverse isolation but uses capacitive neutralization, instead of a cascode topology, in an attempt to mitigate reverse feedthrough. Since the capacitances used to implement this neutralization must be precisely matched to the gate-to-drain capacitance of the transistors, AMOS varactors are used. This novel use of the AMOS varactors helps to further validate the varactor model presented in Chapter 3. The resulting amplifier displays an excellent measured noise performance, achieving a low NF of only 4.5 dB at 23.5 GHz, which is close to that predicted by EM simulation results. Although the differential LNA suffers from a similar gain shortage to that reported in [22], achieving a measured maximum value of 5.2 dB, this result can be duplicated by using measured transistor data in place of the high-frequency models used in original simulations. Despite this, the resulting LNA is effective at mitigating the noise contributions of subsequent circuit components, as illustrated by Equation 1.1 and is suitable for use in a differential CMOS front end.

Chapter six concludes this work, and summarizes the results contained herein. It also proposed areas suitable for further research.

Chapter 2

On-Chip Inductors

2.1 Introduction

One unavoidable issue associated with the evolution of RF CMOS technology is the fact that on-chip passive structures such as Metal-Insulator-Metal (MIM) capacitors and spiral inductors often account for the majority of the chip area consumed. Spiral inductor structures are of particular concern since they not only require a tremendous amount of chip space, but also contribute heavily to the loss and noise generated by the circuit at frequencies above 20 GHz. Not surprisingly, a great deal of research has been conducted in recent years in the hopes of either accurately modeling the high frequency behaviour of spiral inductors [5] or creating new spiral structures to help improve performance [7].

Although circuits using spiral inductors have been successfully designed at frequencies above 20 GHz, their modeling and optimization still remain a challenge and their poor Q factor performance can often severely limit the overall noise performance of the circuit. This chapter briefly summarizes the difficulties associated with

the high-frequency operation of spiral inductors and then introduces a novel Series Stub Transmission Line (SSTL) alternative, which can be used to improve on-chip inductor performance.

2.2 Spiral Inductors

Spiral inductors have traditionally offered the most compact, and therefore economical, means of producing an on-chip inductor in CMOS circuit design. Their use in the low gigahertz range has been well characterized and through the efforts of many recent publications, summarized in [23] and [24], their optimization and implementation have become fairly straight forward in this frequency range. However, as operating frequencies surpass 20 GHz, the behaviour of spiral inductors becomes more complex due to the fact that [10]:

1. Substrate eddy currents become more prevalent and limit the Q factor.

2. Magnetic coupling to the substrate begins to have a significant effect on the overall inductance value.

Since both of the above factors require a detailed knowledge of the substrate profile to be accurately modeled, optimization and implementation of spiral inductors at these frequencies becomes a non-trivial concern for the circuit designer. To illustrate this, Table 2.1 shows the measured inductance and Q factor of a single-turn, 100 μm diameter, square spiral inductor and compares these results to an electromagnetic (EM) simulation conducted in ADS Momentum and a lumped element model based simulation conducted in Asitic. The apparent differences highlight the non-trivial

	Inductance (pH)	Q Factor
Asitic	188	5.402
Momentum	362	12.358
Measured	351	7.429

Table 2.1: Modeling of a single-turn, 100 μm diameter, 15 μm trace width, square spiral inductor at 24 GHz.

nature of high-frequency spiral inductor modeling, resulting in an undesirable degree of uncertainty when using such structures in circuit designs.

It should be noted that there are several different methods commonly used to calculate the Q factor of a spiral inductor [25]. Since the inductors tested were simulated well below their self-resonance frequencies, the definition shown in equation 2.1 is valid and was therefore used for all Q factor calculations in this chapter unless otherwise stated.

$$Q = \left| \frac{\Im(Z_{11})}{\Re(Z_{11})} \right| \tag{2.1}$$

Both of the programs used to obtain the results in Table 2.1 provide accurate simulation results for frequencies below that at which the inductor achieves its maximum Q factor (usually about 10 GHz). Figure 2.1 makes this problem apparent, comparing the measured Q factor of two square spiral inductors with EM simulation results obtained using Agilent ADS Momentum software. Both spirals are 1.25 turns long and have an outer diameter of 100 μm but have different trace widths of 10 and 15 μm, showing that the high-frequency discrepancy exists for different inductance values.

In addition to the modeling discrepancies highlighted above, it is also apparent that the Q factor of these devices leaves much to be desired. With such low values, their use in high-frequency circuits will inevitably create sources of considerable noise

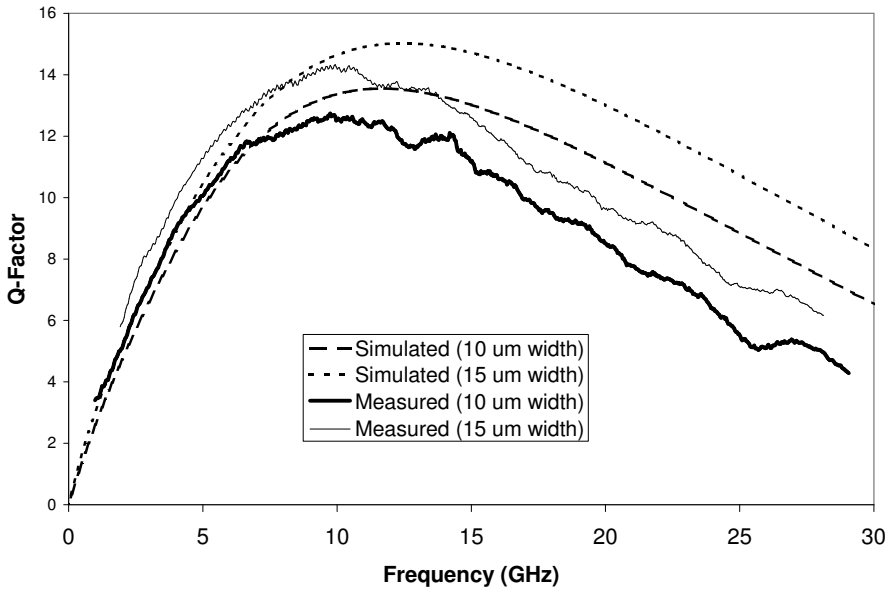

Figure 2.1: Circuit simulators accurately predict inductor behaviour below the maximum Q frequency but struggle at higher frequencies.

and loss, which is of particular concern in low-noise applications. Although some Q factor enhancement schemes, such as the patterned ground shield illustrated in Figure 2.2 [26], have been proposed in recent literature, their improvements are limited to frequencies below or close to the maximum Q frequency. This is illustrated by the results from [7] reproduced in Figure 2.3.

Other Q-enhancement techniques that have been proposed include symmetrical square or octagonal spirals [25]. Although these geometries show more promise at high frequencies than the patterned ground shield option, EM simulations show that they still provide little or no improvement in Q for frequencies above 15 GHz. Figure 2.4 compares the Q factor of 270 pH standard square, symmetrical square and octagonal spiral inductors and illustrates the marginal improvement achieved by using these non-standard geometries at frequencies above 15 GHz.

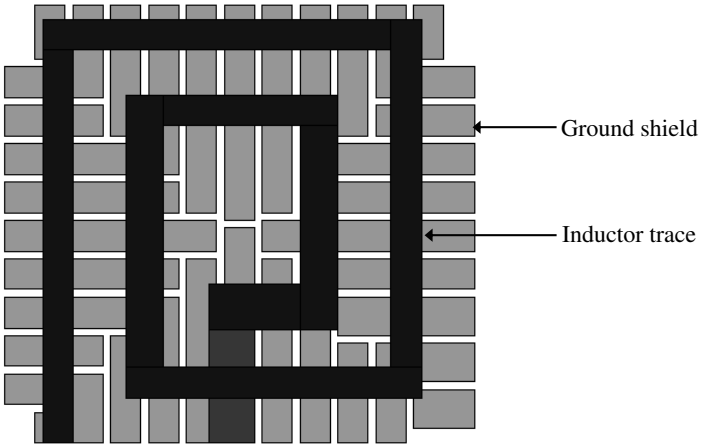

Figure 2.2: Layout of a patterned ground shield beneath a spiral inductor [26].

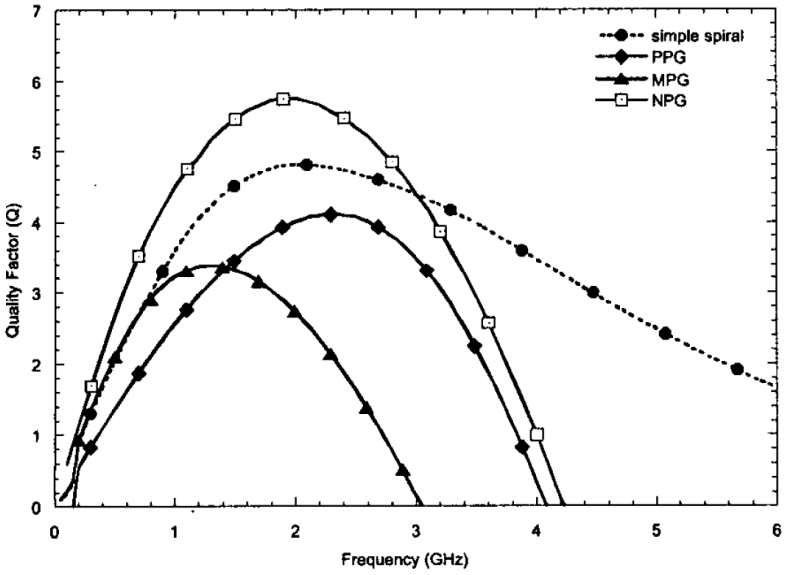

Figure 2.3: Patterned ground shields using polysilicon (PPG), metal 1 (MPG) and n^+ diffusion (NPG) offer no Q factor improvement at frequencies well above the maximum Q frequency [7].

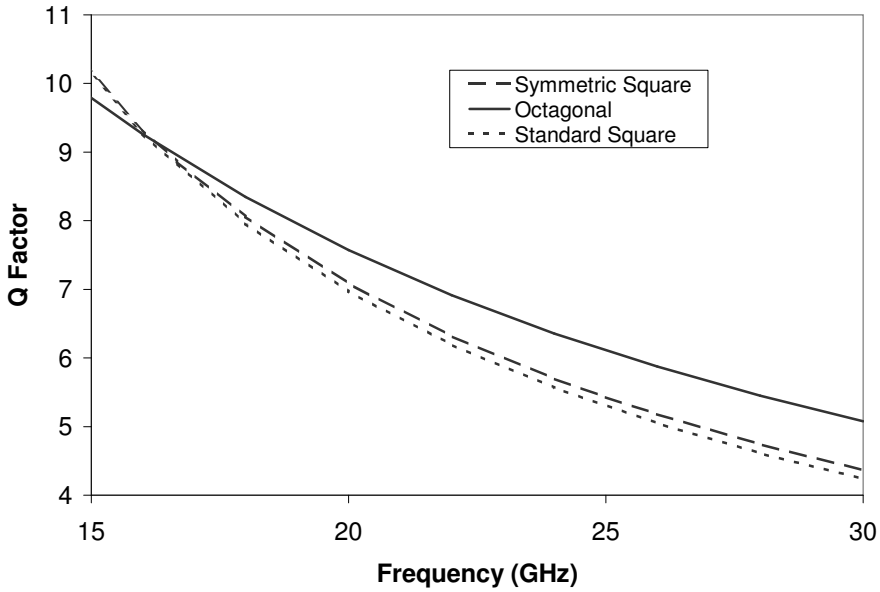

Figure 2.4: High frequency comparison of common spiral inductor geometries.

Clearly, the benefits of an inductor structure that offers either simplified and more accurate modeling or a Q factor improvement at high frequencies would be extremely attractive in RF and millimeter-wave circuit design. The series stub transmission line (SSTL) inductor structure presented in the remainder of this chapter presents advantages over spiral inductors in both of these areas, making its implementation in low-noise circuit applications extremely promising.

2.3 Series Stub Transmission Line Inductors

As an alternative to spiral inductors, transmission lines can also provide an equivalent on-chip inductance. Coplanar waveguides (CPWs), which consist of one signal line placed between two adjacent ground planes, can be designed for minimum loss by optimizing the signal line width and for any system impedance by choosing the

appropriate space between the signal and ground conductors. On the other hand, microstrip transmission lines are particularly attractive to circuit designers since these structures employ the use of a ground plane on a low metal layer, shielding the structure from the substrate, as illustrated in Figure 2.5.

This results in substantially confined electric and magnetic fields and hence simplifies modeling considerably [10]. The downside, which has limited the usefulness of microstrip transmission lines as inductors, is the fact that the line lengths required to achieve the necessary inductance values can be prohibitively long. However, as operating frequencies continue to increase, the reactive elements needed for matching networks and resonators becomes increasingly small, typically requiring inductance values of less than 400 pH.

This means that transmission line inductors are now nearly short enough to be implemented on-chip and has lead to a recently renewed interest in their use as an alternative to spiral inductors [8]. The inductive properties of microstrip transmission lines have been explored for many years and several simplified equations, summarized in [27], have been presented that can be used to calculate their effective inductance. Unfortunately, both calculations and EM simulations reveal that in 0.18 μm CMOS a 300 pH inductance would require a length of microstrip line of 700 μm, which is prohibitive in many cases.

The addition of a series-stub section to a microstrip line, as illustrated in Figure 2.6.(a), can be used to reduce the length of $L1$ required to implement a given inductance. This idea can be extended to take advantage of the multiple metal layers present in CMOS technology to create a three-dimensional structure, thereby reducing the total chip area consumed by the transmission line, an advantage that can not

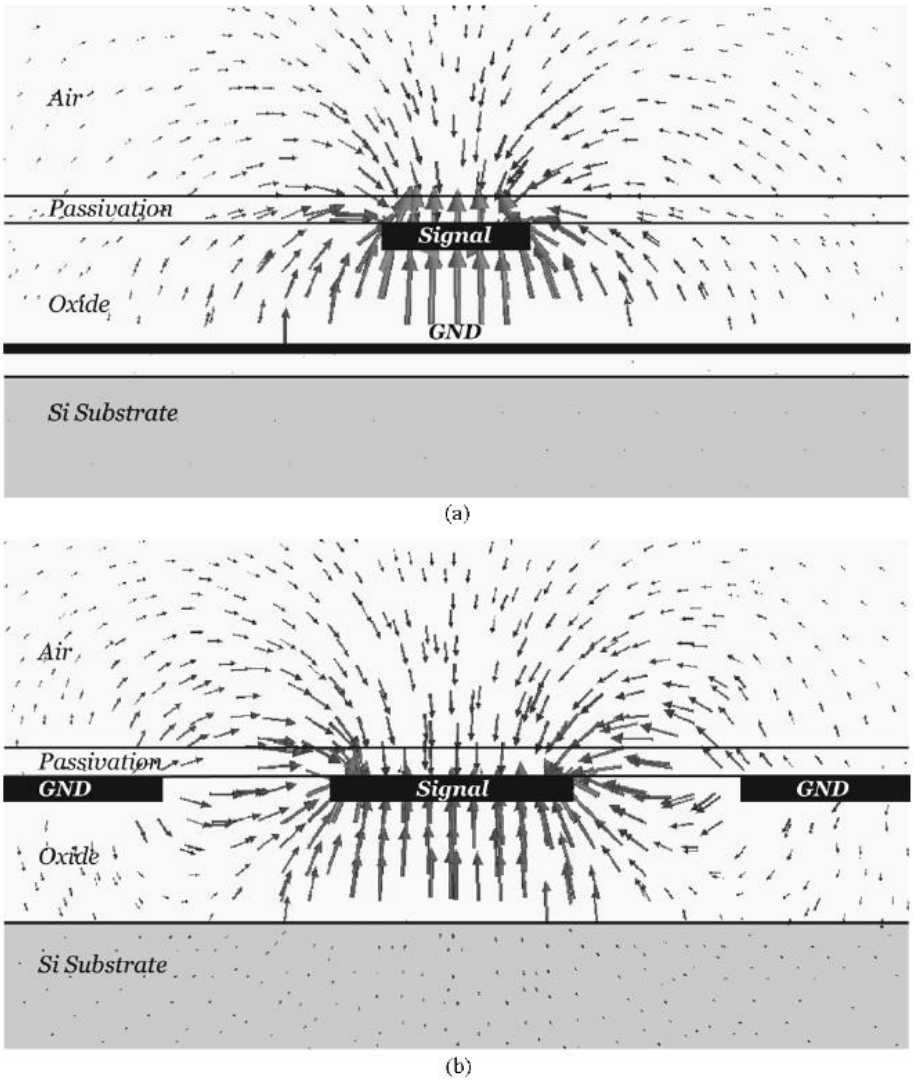

Figure 2.5: Electric field distributions from 3-D EM simulations of (a) microstrip and (b) CPW transmission lines [8].

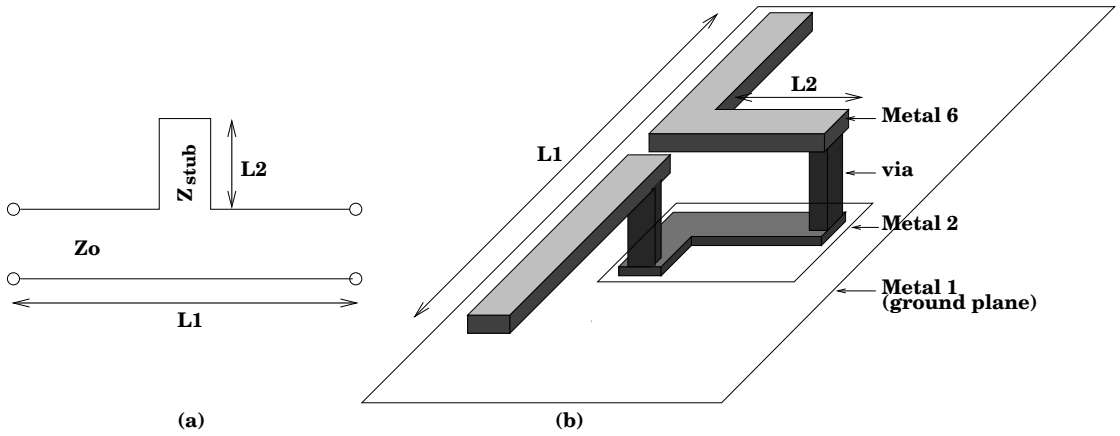

Figure 2.6: The (a) equivalent circuit of a transmission line with a series stub and (b) proposed CMOS implementation.

be realized with the use of other two-dimensional structures (such as a meander line). Figure 2.6.(b) illustrates the novel SSTL structure introduced in this section, which capitalizes on this idea, using several of the metal layers available in the six-layer, 0.18 μm CMOS process with thick top-metal used in this study. This reduces the required length of transmission line $L1$ by twice the length of the series stub (2 x $L2$), so that the 300 pH inductance discussed earlier could now be implemented by an SSTL inductor with $L1 = 500$ μm and stub length $L2 = 100$ μm, dimensions that are much more feasible for on-chip implementation. The flexibility of this structure also means that $L2$ can be increased further, thereby decreasing $L1$, until the overall dimensions of the inductor are suitable for the desired implementation.

While this SSTL geometry benefits the circuit designer by reducing the chip area consumed by the inductor, this gain comes at the cost of a small sacrifice in Q factor and inductance when compared to a simple microstrip line. In order to quantify this tradeoff, EM simulations were performed comparing the two structures and their results are presented in Figure 2.8, with illustrations of the microstrip and SSTL

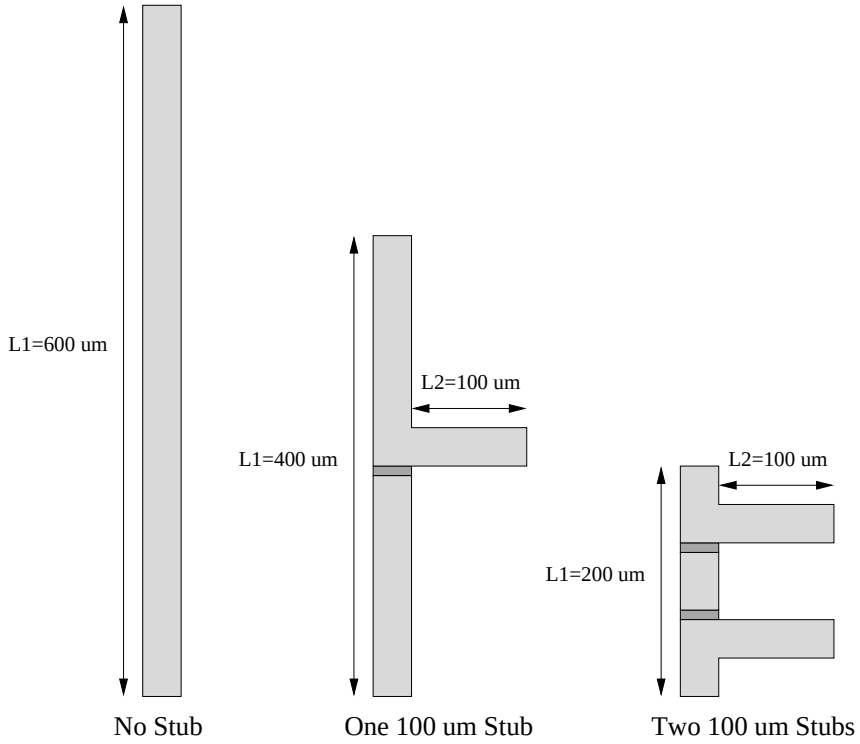

Figure 2.7: Geometries of the microstrip and SSTL structures used in EM simulation. All structures have the same equivalent line length of 600 μm.

geometries used shown in Figure 2.7.

By adding one or two stubs with $L2 = 100$ μm and reducing $L1$ by the corresponding 200 μm increments to keep the total line length constant, the effect of the addition of stubs on the inductance (Figure 2.8 (a)) and Q factor (Figure 2.8 (b)) can be easily seen. From these results it becomes clear that the addition of a single stub can reduce the chip area required by the inductor with only a small sacrifice in Q factor, while the addition of more than one stub reduces both inductance and Q factor considerably. In addition, the geometry of a multi-stub SSTL makes implementation in a circuit much more difficult. For these reasons, all further SSTL analysis is limited to inductors using only a single series stub.

It is difficult to directly compare the chip space consumed by an SSTL inductor to a

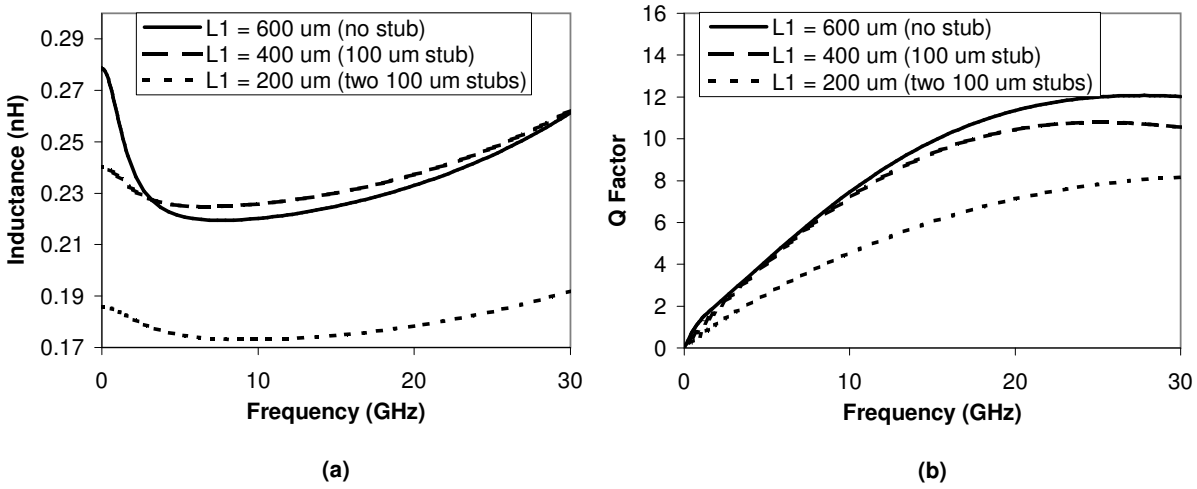

Figure 2.8: The effects on (a) inductance and (b) Q factor encountered when adding series stubs to a microstrip line.

spiral inductor of equivalent inductance due to the fact that the inductance of a spiral can easily be increased by adding turns. Although this does not increase the chip space consumed by the spiral, it instead has the detrimental effect of lowering both the Q factor and self-resonance frequency. Meanwhile, increasing SSTL inductance values must result in an increase in line length, and therefore chip space, but does not necessarily result in a reduction in Q factor or self-resonance, as will be seen later in the chapter. Overall, chip area consumed is usually on the same order of magnitude and even in cases where the SSTL inductor size exceeds that of the equivalent spiral, it may be easier to implement in the circuit layout due to its long and narrow geometry and its flexibility in terms of stub placement.

2.3.1 SSTL Optimization

With the benefits of the SSTL structure clearly outlined, an analysis of each of the design variables in the geometry becomes necessary so as to optimize the Q factor and

inductance. It should be noted that, in order to reduce the computation time required in performing numerous EM simulations for optimization, some small simplifications (such as placing the ground reference terminal directly on the metal 1 ground plane instead of placing vias to return the ground plane to the top metal layer) were made to the inductor layouts in this section. These modifications were consistent in all simulations performed, making the optimized results valid, but absolute inductance and Q factor values are slightly optimistic. After completing the optimization in this section, more detailed simulations providing more accurate inductance and Q factor results are provided in the following section.

The first question of optimization to be addressed is the choice of metal layers to be used in the SSTL structure. By removing a section of the Metal 1 ground plane directly beneath the series stub, the capacitance to the ground plane is minimized, thereby maximizing the SSTL inductance. EM simulations reveal that this not only results in a significant increase in inductance but also in an increase in Q factor of approximately 2. The downside to this approach is that it allows for a small amount of coupling to the substrate through this hole in the ground plane and raises the question as to whether the lower layer of the series-stub should be placed on a high metal layer so as to reduce this substrate coupling, or on a low metal layer so as to reduce the capacitance between this line and the upper layer of the series-stub.

EM simulations of an SSTL inductor with a main transmission line length of $L1 = 500$ μm and a stub length of $L2 = 100$ μm (and therefore a total length equivalent to a simple microstrip line of 700 μm) were conducted to test the effects of the choice of metal layer for the lower layer of the series-stub. These results, which are displayed in Figure 2.9, included the effects of all metal thicknesses and display a uniform

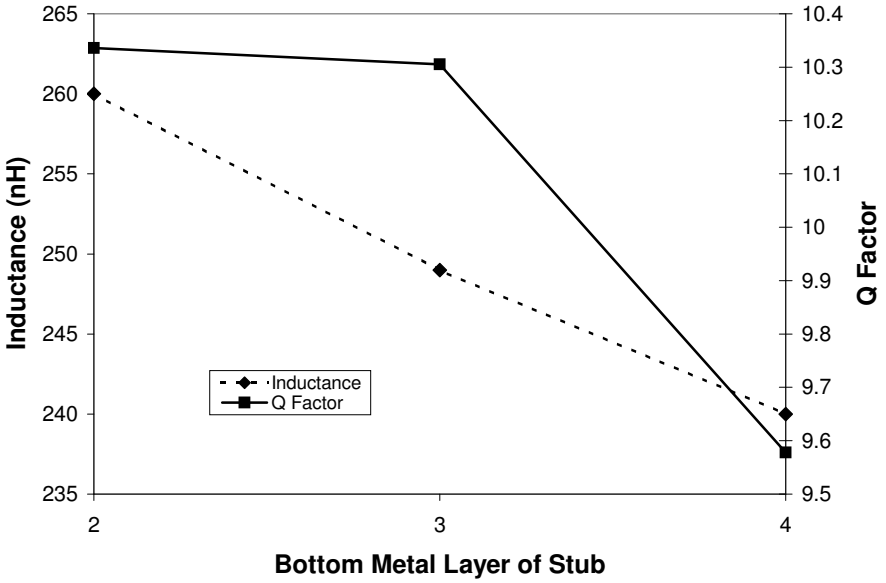

Figure 2.9: Choice of the series-stub bottom metal layer.

decrease in the total inductance of the SSTL inductor as the bottom metal layer is moved farther from the substrate. The decrease in Q factor that can also be seen is the result of the change in impedance of the series stub that results from differences in the capacitive coupling between the top and bottom layers of the series-stub. This changes the peak Q factor frequency of the SSTL which, at a frequency of 25 GHz, translates to a small drop in Q factor as the bottom layer of the stub is moved closer to the top metal layer. This justifies the choice of the use of metal 2 in the SSTL structure but illustrates that this optimization should be re-examined if the SSTL structure is to be used at frequencies at or above the peak Q frequency.

The next step in optimizing the SSTL structure is to determine the appropriate trace width. This involves a tradeoff between higher inductance (and therefore shorter required line length) for thin trace widths and higher Q factor for wide trace widths. Figure 2.10 illustrates the change in inductance (a) and Q factor (b) encountered as

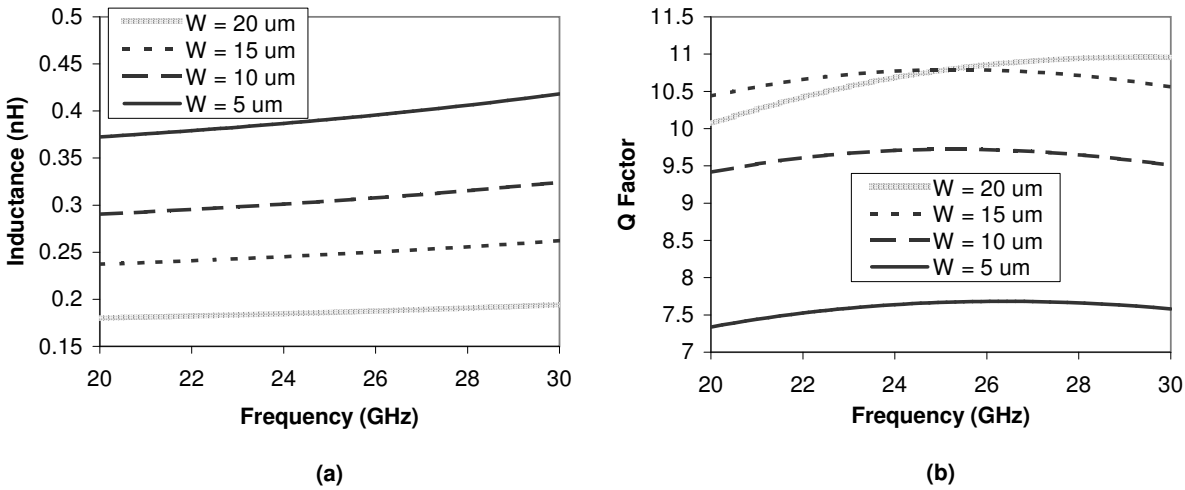

Figure 2.10: Optimization of the SSTL trace width.

the trace width of an L1 = 400 μm and L2 = 100 μm SSTL inductor is varied from 5 to 20 μm. Since the Q factor, which is of high importance, reaches a maximum at 24 GHz for a trace width of 15 μm and the inductor displays a good level of inductance for this same width, 15 μm was the final width chosen for further simulation and fabrication.

Finally, the last parameter of interest that can optimized in the SSTL inductor, the length of the stub itself, should be considered. Again, EM simulations of various stub lengths were performed and the results are displayed in Figure 2.11. In order to compare the inductance and Q factor fairly, the SSTL length $L1$ was reduced by 40 μm for every 20 μm increase in the stub length $L2$ in order to keep the total line length constant. Although we have established that the addition of the stub incurs a slight penalty in the Q factor, Figure 2.11 illustrates the fact that, at a frequency of 25 GHz, both the inductance (a) and Q factor (b) can be optimized by choosing a stub length of approximately 80 μm.

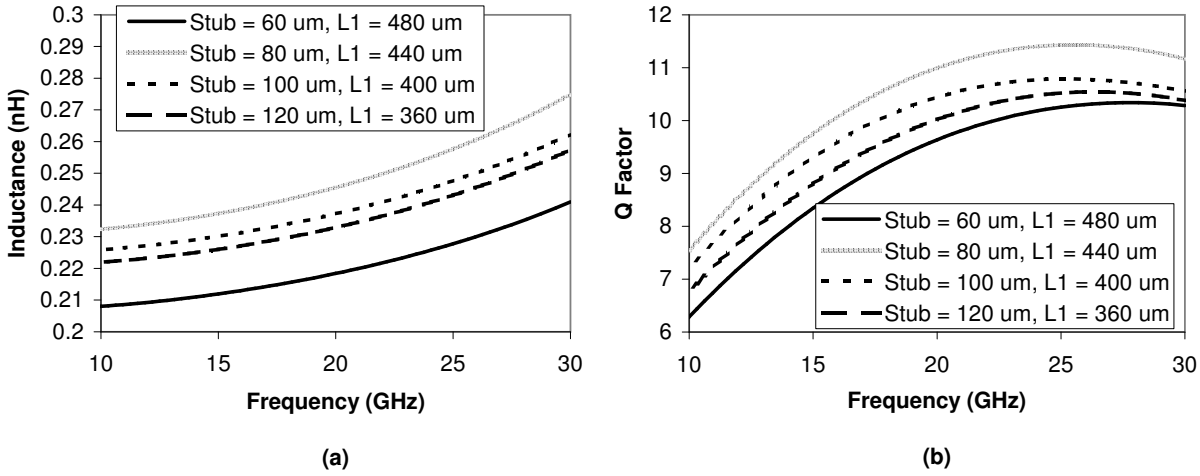

Figure 2.11: Optimization of the SSTL stub length.

2.3.2 Simulation and Measurement Results

With the SSTL inductor structure optimized, an in-depth analysis of the Q factor behaviour was then conducted. Since microstrip structures inherently shield the transmission line from the substrate, the only sources of resistive loss that remain come from a small amount of coupling to the substrate in the series-stub area, and the ohmic resistance of the line itself. The overall effect of this is that the Q factors of SSTL structures are only weakly dependent on the line length and hence the inductance value. This contrasts sharply with spiral inductor structures where longer line lengths (meaning either larger diameter or an increased number of turns) lead to increased substrate coupling and hence lower Q. Also, since the substrate coupling is limited to a very small area, one can intuitively see that the reactive portion of the impedance should increase more quickly than the resistive portion over a very wide frequency range. This dictates that the Q factor of the device (recall that for an inductor $Q = \frac{\omega L}{R}$) should also increase over the same frequency range.

This behaviour is verified by the EM simulation results in Figure 2.12, which

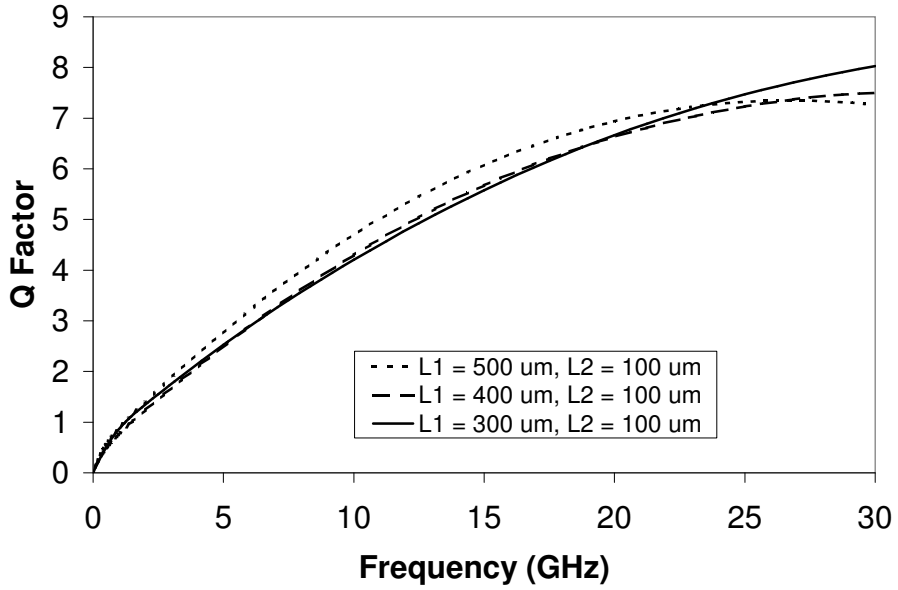

Figure 2.12: Simulated Q Factor of SSTL inductors.

display a monotonically increasing Q factor with a variation of less than 1 even as the operating frequency increases to 30 GHz and the equivalent line length is varied from 500 μm to 700 μm. Above 20 GHz the ohmic losses induced by the skin effect of the transmission line have a more significant effect for the longer line lengths, causing variations in the Q factor that are dependent on this length.

S-parameter measurements were completed on an Agilent 8510 vector network analyzer (VNA). The Q factor results are displayed in Figure 2.13 and show a good correlation between EM simulation and measured results for SSTL inductors with total equivalent line lengths of 500, 600 and 700 μm. Although there is some deviation at higher frequencies, these discrepancies between the simulated and measured Q are limited to a difference of less than 0.5, which is a significant improvement over the discrepancies seen in the spiral inductor results as shown in Figure 2.1.

The second parameter of importance, the inductance value itself, shows a very

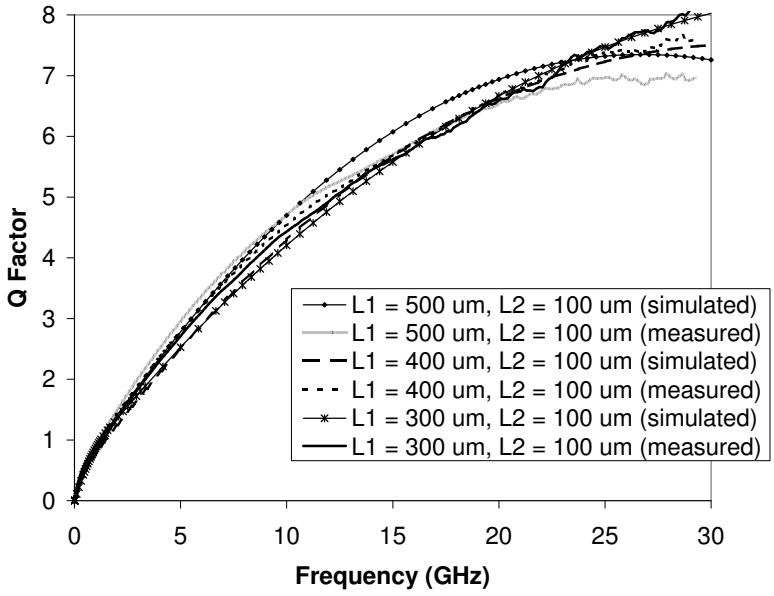

Figure 2.13: Simulated vs measured SSTL Q factor.

strong correlation between measured and EM simulation data over all frequencies of interest and for all line lengths, as can be seen in Figure 2.14. To quantify this correlation, discrepancies are limitted to less than 20 pH from 10 to 30 GHz, with this variation shrinking to much less than 10 pH at 25 GHz. This therefore validates the assertion that the SSTL inductor structure, with its ground plane acting as a shield between the microstrip line and the substrate, lends itself more effectively to modeling and optimization than spiral structures, which suffer from the high frequency modeling inaccuracies discussed earlier.

Finally, Figure 2.15 compares the measured Q factor of SSTL inductors to single-turn spiral inductors of approximately the same surface area. These results show a distinct high frequency advantage in favour of the SSTL structure, which continues to grow as the operating frequency increases to 30 GHz. This increase in Q factor can play an important role in reducing the noise generated by a circuit in noise-sensitive

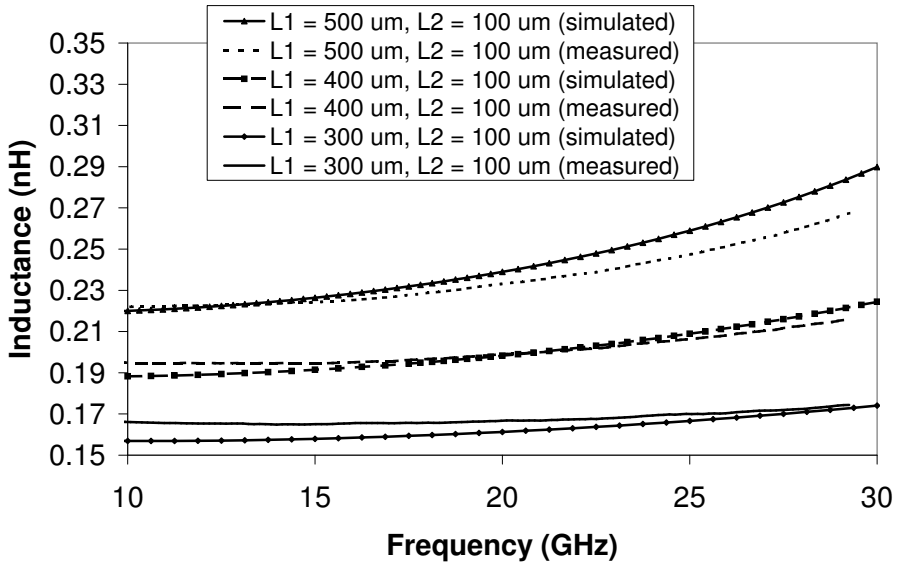

Figure 2.14: Simulated vs measured SSTL inductance.

applications, such as a low-noise amplifier (LNA) in the front end of a high-frequency CMOS receiver. As such, these results are of great interest to circuit designers working at these frequencies and have been published in [28].

2.4 Future Work

The idea of patterned ground shielding mentioned earlier was shown to be ineffective for increasing the Q factor of spiral inductors designed to operate well below the maximum Q frequency [7]. Very recently, however, this idea has been extended to other passive devices such as coplanar waveguide (CPW) transmission lines. In [9] the use of floating metal 1 strips beneath the signal path are shown to encourage coupling to the adjacent ground planes and minimize the penetration of electric field lines into the substrate.

While SSTL inductors consist mainly of a microstrip structure, with a solid ground

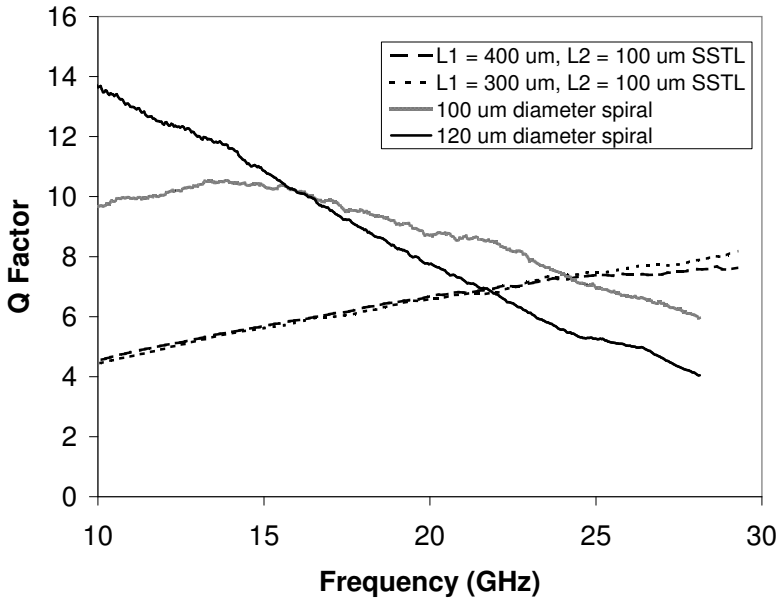

Figure 2.15: Q factors of SSTL and spiral inductors.

shield beneath the transmission line, the ground shield beneath the stub section has been removed for Q factor and inductance improvement, as discussed earlier. While the losses incurred from the resulting substrate coupling are shown to be smaller than would be encountered if the ground shield beneath the stub were not removed, they are still noticeable. As an alternative, it may be possible to mitigate this coupling by using the floating shield technique, discussed in [9], beneath the stub. This concept is illustrated by a top view of the shielded stub shown in Figure 2.16.

Preliminary EM simulations of this technique on an SSTL inductor with $L1 = 400$ μm and $L2 = 100$ μm, are shown in Figure 2.17 and confirm that this method is promising for increasing the Q factor. The simulation was completed for $W = 1.6$ μm wide and $L = 25$ μm long floating shield strips, separated by $S = 1.6$ μm, with their ends spaced $G = 0.3$ μm from the adjacent ground plane. Although the Q factor improvement shown in Figure 2.17 is fairly small, future work on the dimensions and

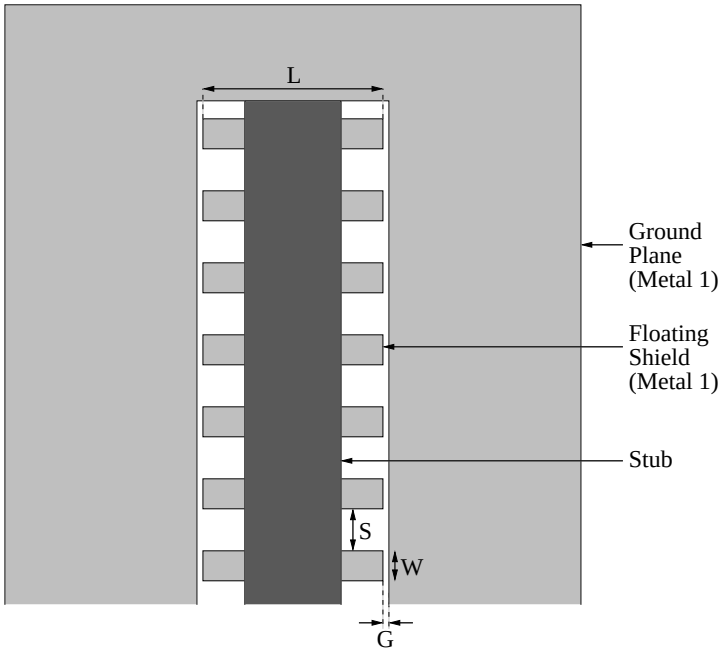

Figure 2.16: Floating ground shields reduce coupling from the series-stub to the substrate.

metal layers used for the floating shield could further improve the resulting Q factor.

2.5 Conclusions

Despite the efforts of recent literature, spiral inductors remain difficult to model and suffer from low Q factors in silicon CMOS at frequencies above 20 GHz. Fortunately, as operating frequencies continue to increase, the inductance values required in matching networks and resonators must decrease accordingly. This means that transmission line inductors, which were once prohibitively long, are finding renewed application in RF and millimeter-wave CMOS design. The SSTL inductor structure presented in this chapter uses a 3-dimensional series stub to minimize its required

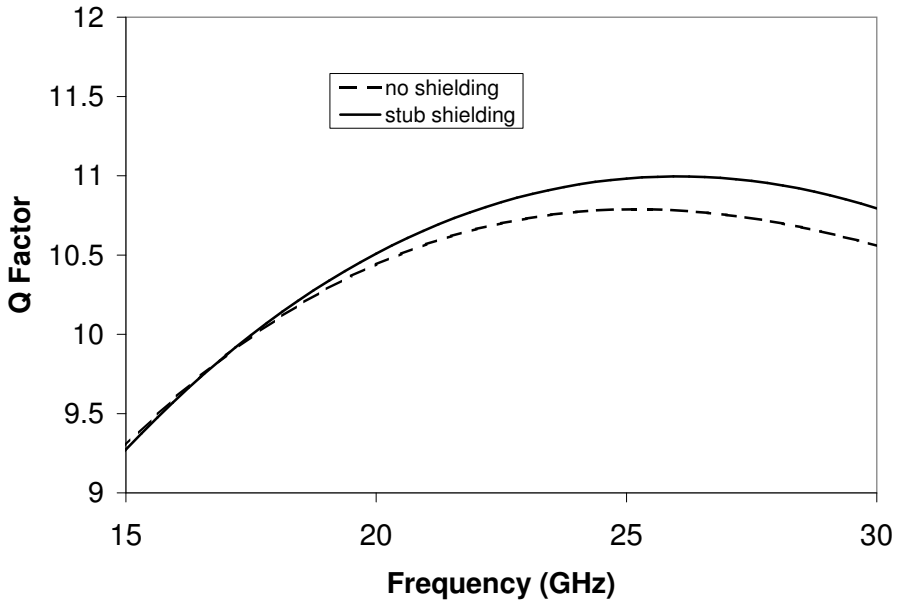

Figure 2.17: EM simulation results showing the Q factor improvement made possible through the use of floating ground structures beneath the stub of an SSTL inductor with $L1 = 400$ μm and $L2 = 100$ μm.

chip area and exhibits a distinct high-frequency Q factor advantage over comparably sized spiral inductors. In addition, the use of a metal ground plane on a low metal layer, which is inherent to microstrip lines shields the SSTL from the effects of the substrate, making simulation and optimization much simpler and more accurate than for spiral inductors.

Chapter 3

Varactors

3.1 NMOS Varactors

Variable capacitors, or varactors, have often been implemented in CMOS technology using a reverse-biased $p - n$ junction diode. However, the poor Q factor of such devices, even in the low gigahertz region of operation, leaves much to be desired [29]. A second option that is readily available in any CMOS process is the MOS transistor itself, whose capacitance value can be tuned by the changing the applied voltage. By tying the source, drain and bulk terminals together and varying the voltage applied between the resulting terminal and the gate of an NMOS device, the charge layer beneath the gate oxide in a standard NMOS transistor will vary accordingly. This creates a capacitance between the heavily doped n^+ polysilicon gate electrode and the drain/source/body connection, that can be easily tuned by varying a single DC voltage. This variation in capacitance, as illustrated in Figures 3.1 to 3.3, can be characterized by three modes of device operation [30]:

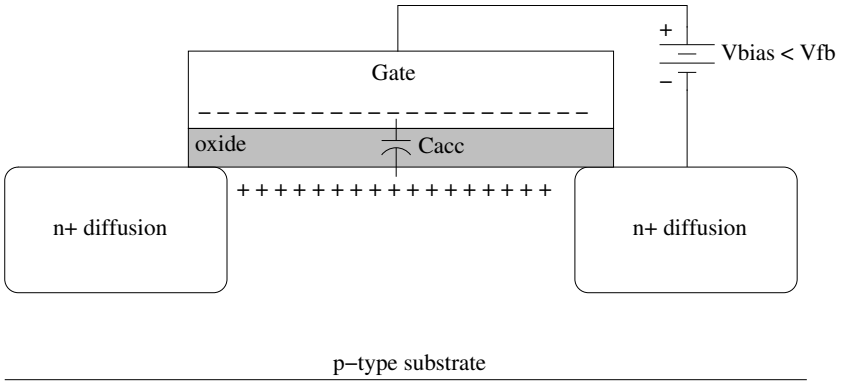

Figure 3.1: Mobile carriers form a positive charge layer below the gate oxide when the NMOS varactor is in accumulation mode.

1. *Accumulation* in which positively charged mobile carriers supplied by the body "accumulate" in the channel (Figure 3.1). This occurs when the applied bias voltage is less than the flat band voltage, V_{fb}, which is the voltage at which there is no charge buildup in the channel.

2. *Depletion* in which the channel is "depleted" of any mobile carriers, leaving only a charge separation region (Figure 3.2). This occurs when the V_{bias} is between V_{fb} and the threshold voltage, V_{th}.

3. *Inversion* in which negatively charged mobile carriers supplied from the n^+ diffusion regions aggregate in the channel, thereby "inverting" the conductivity type of the channel (Figure 3.3). This occurs when the control voltage is greater than V_{th}.

From the capacitances displayed in these figures, it is apparent that the capacitance seen between the varactor terminals is at its peak in the accumulation, C_{acc}, and inversion, C_{inv}, modes. It can be roughly calculated using the equation for the

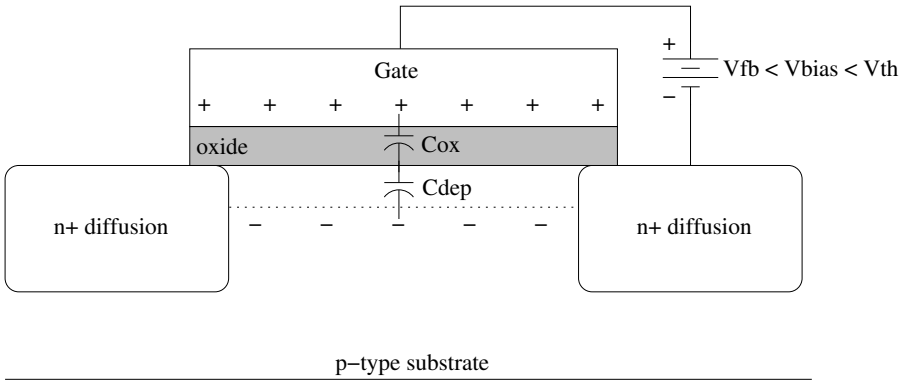

Figure 3.2: The absence of a charge layer below the gate oxide reduces the overall capacitance when the NMOS varactor is in depletion mode.

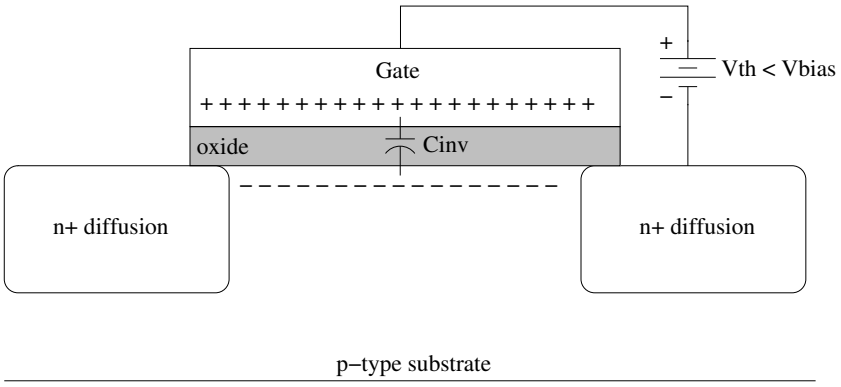

Figure 3.3: Mobile carriers form a negative charge layer below the gate oxide when the NMOS varactor is in inversion mode.

capacitance across the gate oxide

$$C_{acc} = C_{inv} = W_{eff} L_{eff} \frac{\epsilon_{ox}}{t_{ox}} n \tag{3.1}$$

where ϵ_{ox} is the oxide permittivity, t_{ox} is the oxide thickness and n is the number of gate fingers. L_{eff} is the effective channel length after it has been reduced by the lateral diffusion of the source and drain regions, as explained later, and is illustrated in Figure 3.7. Similarly, W_{eff} accounts for lateral diffusion of these regions along the width of the drain and source regions.

As can be seen in Figure 3.2, the creation of a depletion region beneath the gate oxide leads to the creation of an additional capacitance, C_{dep}, in series with the oxide capacitance. If we define the oxide capacitance per unit area as

$$C'_{ox} = \frac{\epsilon_{ox}}{t_{ox}} \tag{3.2}$$

and the depletion region capacitance per unit area as

$$C'_{dep} = \frac{\epsilon_{si}}{w_d} \tag{3.3}$$

where w_d is the depletion region width and ϵ_{si} is the permittivity of silicon, the capacitance of the varactor in depletion mode, C_{var}, can be calculated as [30]

$$C_{var} = W_{eff} L_{eff} \frac{C'_{ox} C'_{dep}}{C'_{ox} + C'_{dep}} n \tag{3.4}$$

Since the width of the depletion region varies according to the bias voltage applied between the varactor terminals, so too does the resulting C_{dep}. Equation 3.4 describes the effect of two series capacitances and as such, it is apparent that any value of C'_{dep} will result in a reduction of C_{var}, but also that this reduction will become negligible as C'_{dep} becomes much larger than C'_{ox}. The capacitance curve for all three regions of varactor operation will look similar to that of Figure 3.4 [29].

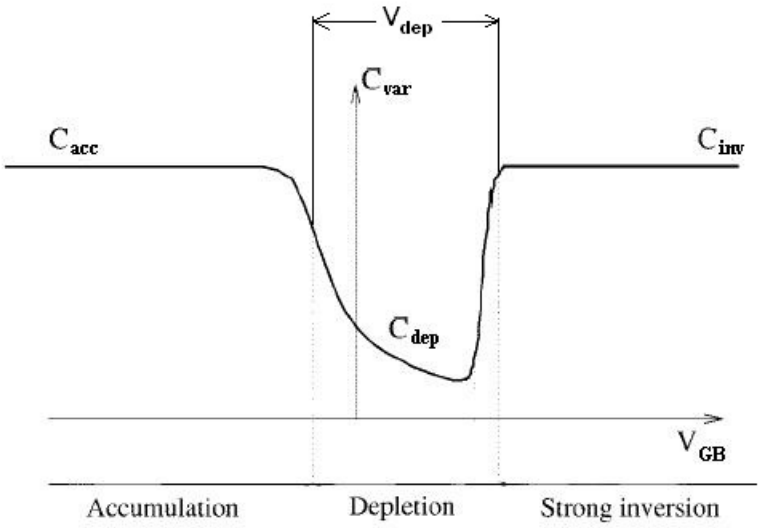

Figure 3.4: Capacitance curve of an NMOS varactor as bias voltage across the terminals is varied [29].

3.2 AMOS Varactors

As shown in Figure 3.4, the voltage range over which an NMOS varactor capacitance can be tuned is limited to the relatively narrow depletion voltage range, V_{dep}. As a result, a more attractive varactor option is to modify the NMOS device such that it is limited to operation in the accumulation and depletion regions only, allowing the circuit designer to tune the varactor between a high and low capacitance value over a large range of bias voltages. Such a varactor is known as an Accumulation-mode Metal Oxide Semiconductor (AMOS) varactor and can be created by placing the n^+ diffusion regions of an NMOS device in an n well region instead of the p type substrate. The resulting varactor is illustrated by a cross-section of the device in Figure 3.5 [29]. With the n^+ diffusion regions surrounded by an n well, there is no supply of positively charged mobile carriers into the channel and the device cannot

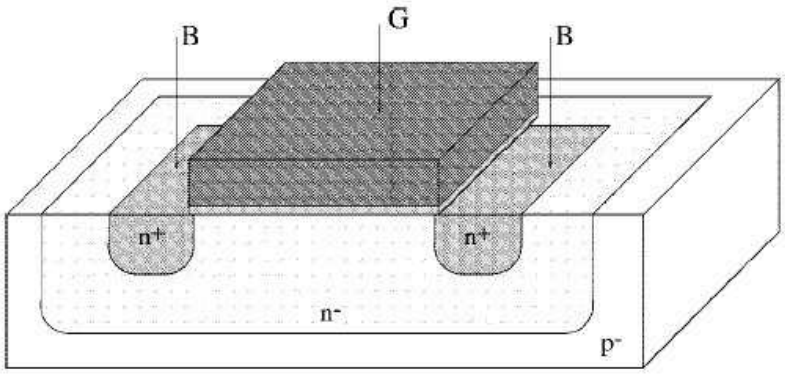

Figure 3.5: Cross-section of an AMOS varactor [29].

achieve inversion.

3.3 AMOS Varactor Modeling

As CMOS device dimensions continue to be scaled downwards into the nanometer regime, opening the door for high gigahertz operating frequencies, the inductance and capacitance of the passive components required in these circuits must shrink accordingly. In fact, for the circuit designs in 0.18 μm CMOS operating above 20 GHz presented in this research, it is not uncommon to require capacitances on the order of tens of femtofarads. As such, the modeling of any varactors to be used to implement such small capacitances must be extremely accurate in order to avoid multiple fabrication iterations. Recent publications have explored this issue and presented models which can accurately duplicate the measured behaviour of their respective varactors. However, most of the models published to date suffer from at least one of the following difficulties:

- They rely on components which are derived from extracted data, as can be seen in [15]. These components therefore often use non-physical, curve-fitting

parameters, making the adaptation of such models to other CMOS technologies or even different varactor dimensions extremely difficult.

- Those that are based on the physical AMOS structure and use analytical equations based on the surface potential, such as those in [14] and [12], often employ separate models for the various operating regimes. This makes their use extremely difficult in circuit simulation.

In this chapter, a new AMOS varactor model is presented which relies on equations based in semiconductor physics, minimizing the use of curve fitting factors. It also avoids the use of boundary conditions and different models for the various regions of operation, allowing for accurate and easy implementation in circuit simulation. Since the intended capacitance range of the varactors being modeled is on the order of tens of femtofarads, an in depth analysis of all the parasitics associated with the AMOS varactor, including the substrate in which is fabricated, is essential to producing an accurate model. Figure 3.6 illustrates the complexity of this problem and identifies the lumped elements that compose the varactor model presented in this chapter. Some layout components, such as via connections, are omitted from this diagram since they make negligible parasitic contributions to the device. Then, in Figure 3.7 the lumped elements have been replaced by an illustration of the important dimensions, which will be used in analytical equations to derive values for the lumped elements.

In order to simplify the reading of the remainder of this chapter, the constant values used for each of these parameters, as well as any other constants used in the equations to follow, are displayed in Appendix A. The sources from which each of these values were taken is also given so that the model can easily be reproduced or adapted to fit other CMOS technology nodes or varactor layouts.

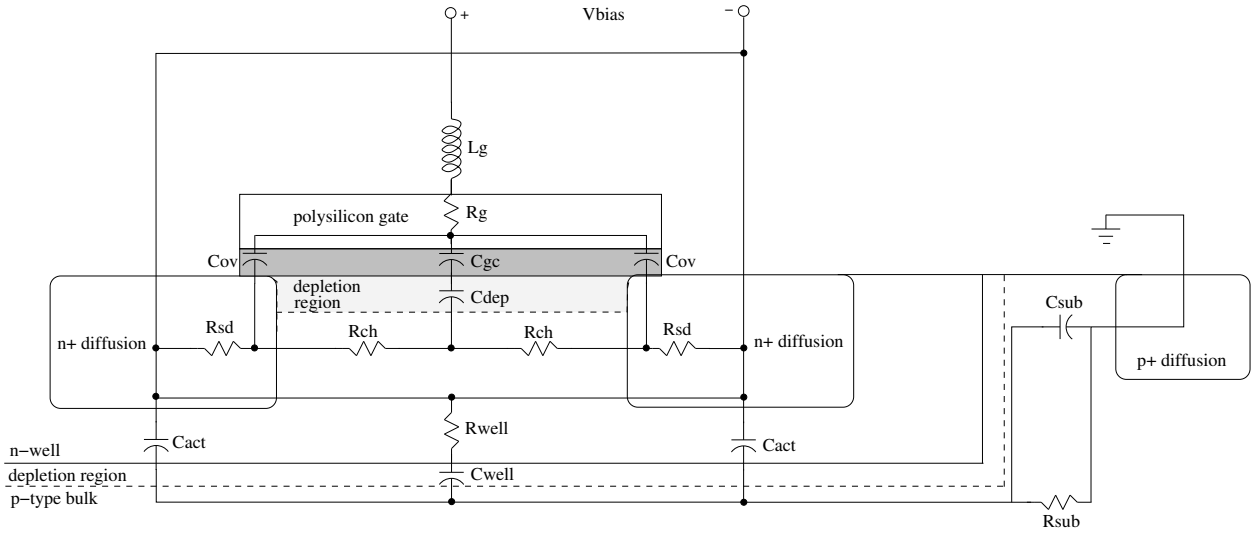

Figure 3.6: Cross-section of an AMOS varactor showing lumped element components to be considered in device modeling.

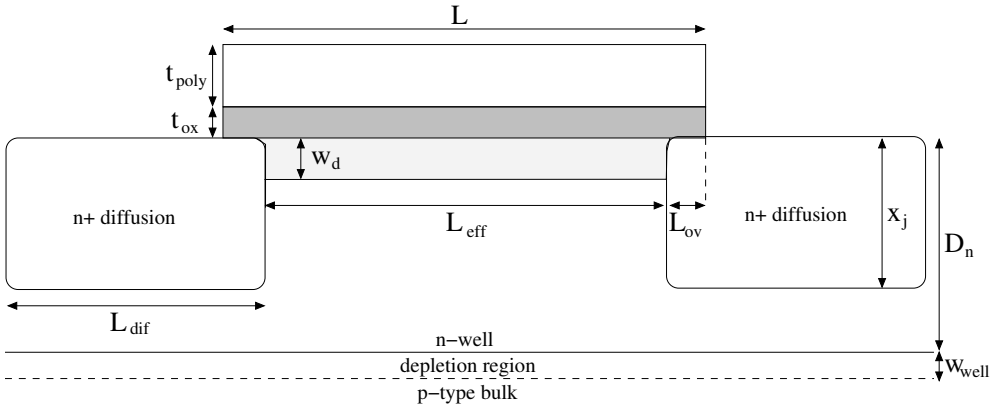

Figure 3.7: Cross-section of an AMOS varactor showing the important dimensions used to determine lumped element values.

3.3.1 Channel Capacitance Model

Since the principle of operation for the varactor depends on the capacitance created between the gate electrode and the body, an accurate physical model of this capacitance is essential. The gate to channel capacitance, C_{gc}, is essentially a parallel plate capacitance and can be simply calculated as

$$C_{gc} = C'_{ox} A_g n \tag{3.5}$$

where n is the number of gate fingers and A_g is the area in the channel beneath the gate electrode given by

$$A_g = L_{eff} W_{eff} \tag{3.6}$$

As explained earlier, when the voltage seen at the gate terminal drops below the voltage of the channel, a depletion region begins to form beneath the gate oxide. The width of this depletion region varies with the magnitude of the voltage difference creating an additional capacitance, C_{dep}, in series with C_{gc}, effectively reducing the total capacitance seen. Since this depletion region exists only when the varactor is in depletion mode, it is tempting to create two distinct operating cases for the varactor model in order to accurately describe this effect. However, by using semiconductor physics based equations it is possible to create a single, unified model for C_{gc}. Using this technique, the width of the depletion region can be calculated as [14]

$$w_d = \sqrt{\frac{2\epsilon_{si}}{qN_{well}}} \sqrt{V_{eff} - \frac{Q_{dep}}{C'_{ox} A_g}} \tag{3.7}$$

where q is the elementary charge and N_{well} is the carrier concentration of the n well region. Q_{dep} is the depletion region charge, which is calculated as

$$Q_{dep} = qN_{well} w_d A_g \tag{3.8}$$

Finally, V_{eff} represents the effective voltage seen across the variable capacitance and is used to ensure that Q_{dep} approaches zero with the device being driven into accumulation. It can be calculated by [14]

$$V_{eff} = \frac{1}{2}\left(\sqrt{V_{bias}^2 + \delta_d} - V_{bias}\right) \tag{3.9}$$

V_{bias} is the applied bias voltage and δ_d is used to determine the speed at which w_d (and hence C_{dep}) changes with the applied voltage and is usually chosen to have a value of approximately 0.01. By substituting equation 3.7 into equation 3.8 we obtain the following result for the depletion region charge

$$Q_{dep} = -qN_{well}\frac{\epsilon_{si}}{C'_{ox}}A_g \pm \sqrt{q^2N_{well}^2\frac{\epsilon_{si}^2}{(C'_{ox})^2}A_g^2 + 2qN_{well}\epsilon_{si}V_{eff}A_g^2} \tag{3.10}$$

By using this in equation 3.7 we obtain a continuous function for w_d at any applied bias voltage. Although this function never reaches zero regardless of the bias voltage applied, Figure 3.8 shows that w_d is very small for positive gate voltages. Since C_{dep} is simply calculated as

$$C_{dep} = \frac{\epsilon_{si}}{w_d}A_g n \tag{3.11}$$

this implies that C_{dep} actually increases as w_d decreases, as can also be seen in Figure 3.8. By examining equation 3.4, given previously for overall capacitance in depletion mode, it becomes apparent that as C_{dep} continues to increase, its effect on the overall capacitance decreases and eventually becomes negligible.

3.3.2 Overlap Capacitance Model

The second major contributor to the capacitance seen between the varactor terminals is the overlap capacitance, C_{ov}, created by the overlap of each gate finger over the

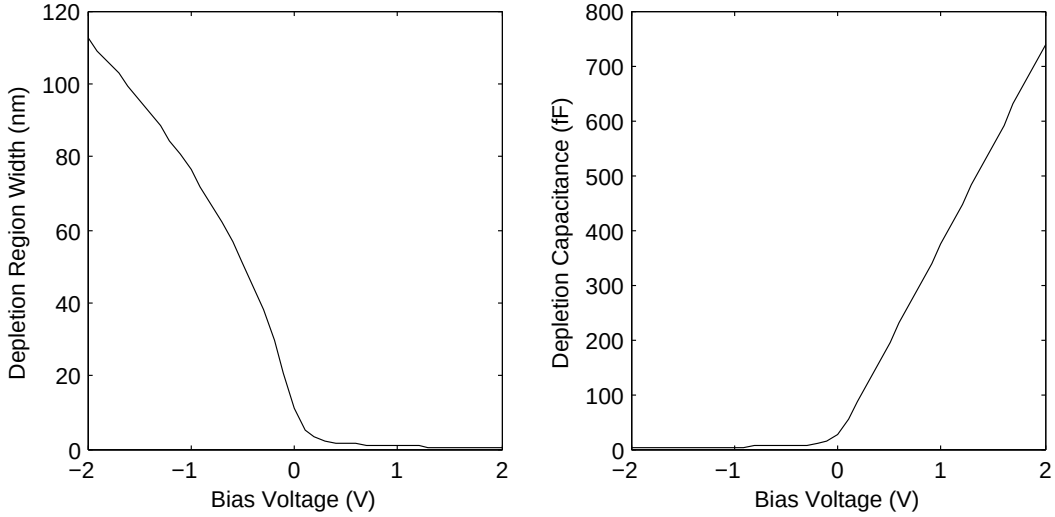

Figure 3.8: Simulated depletion region width and resulting capacitance.

source and drain regions. This overlap is caused by the lateral diffusion of the n^+ diffusion regions and is relatively independent of gate length. Unfortunately, the simple parallel plate capacitance model that is most often used to describe C_{ov} is insufficient to describe the magnitude of the actual measured capacitance due to the fact that the fringing capacitances also make a significant contribution [31]. Figure 3.9 illustrates the origins of these fringing capacitances.

While fringing capacitances are often accounted for by simply multiplying the parallel plate capacitance, C_{pp}, by a factor of about 1.3, it is possible to eliminate this guesswork and improve the accuracy of the capacitance model. To do this we begin by calculating the actual parallel plate capacitance of the overlap as

$$C_{pp} = C'_{ox} W_{eff} L_{ov} n \tag{3.12}$$

where L_{ov} is the length of the overlap region on one side of the channel, which can be taken from the fabrication process parameter. In the 0.18 μm CMOS process used

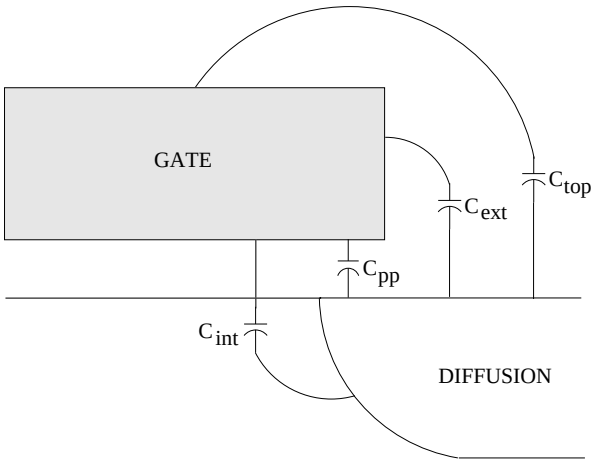

Figure 3.9: Fringing capacitances make significant contributions to the total overlap capacitance.

for varactor fabrication in this work, this length can be reliably approximated as

$$L_{ov} = \frac{1}{6}L \tag{3.13}$$

Using this information the effective gate width, W_{eff}, can also be calculated. If we assume that the lateral diffusion of the drain and source regions is equal along the length and width dimensions then W_{eff} can be approximated in 0.18 μm CMOS as

$$W_{eff} = \frac{1}{3}L + W_{lay} \tag{3.14}$$

where W_{lay} is the width of the diffusion region in the varactor layout.

Secondly, we examine the extrinsic fringing capacitance, C_{ext}, which is calculated using the following equation from [32]

$$C_{ext} = \frac{2}{\pi}\epsilon_{ox} \ln\left(\frac{t_{poly}}{t_{ox}}\right) W_{eff} 2n \tag{3.15}$$

where t_{poly} is the thickness of the polysilicon gate material.

Although the top side of the gate is separated from the diffusion region by at least $t_{poly} + t_{ox}$, it can also have a significant effect on the overlap capacitance, which can

be calculated in a similar fashion to C_{ext} above using the following equation from [33]

$$C_{top} = \epsilon_{ox} \ln\left(1 + \frac{L}{t_{poly} + t_{ox}}\right) W_{eff} 2n \tag{3.16}$$

Finally, the intrinsic fringing capacitance, C_{int}, must be determined. Since this capacitance occurs through the channel region, it can only exist in the absence of a charge layer. This means that the intrinsic fringing capacitance is negligible when the varactor is in accumulation mode and then increases as the depletion region grows, exposing more of the inner side wall of the n^+ diffusion region. As a result, C_{int} varies with the width of the depletion region and can be calculated using [32]

$$C_{int} = \frac{2}{\pi}\epsilon_{si} \ln\left(1 + \frac{w_d}{t_{ox}} \sin\left(\frac{2\epsilon_{ox}}{\pi\epsilon_{si}}\right)\right) W_{eff} 2n \tag{3.17}$$

The total overlap capacitance can now be calculated using equation 3.18. This overlap capacitance is significant as it can contribute more strongly to the overall capacitance seen across the varactor terminals than the channel capacitance itself, especially when the device is in depletion mode. In addition, it is worth noting that, in standard 0.18 μm CMOS, the fringing capacitance can account for anywhere from 30 to 50 percent of the total overlap capacitance. This emphasizes the fact that the use of a simple "fringe factor" in the overlap capacitance equation is inadequate for accurate modeling.

$$C_{ov} = C_{pp} + C_{ext} + C_{top} + C_{int} \tag{3.18}$$

At last, a model for the total capacitance seen between the varactor terminals can be produced by the simple addition of this overlap capacitance with the channel capacitance presented in the previous section. Figure 3.10 illustrates this capacitance (ignoring channel resistance and other parasitics) as a function of bias voltage applied to the gate terminal of a single finger varactor with a gate width of 2.5 μm. Although

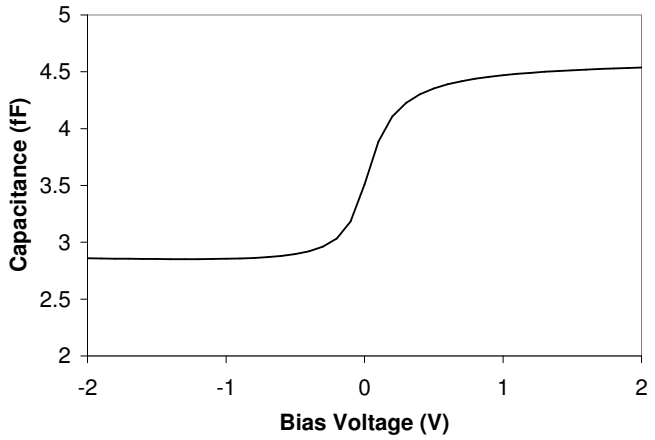

Figure 3.10: Total capacitance seen between the varactor terminals (including fringing effects) for a single finger varactor with a bias voltage applied to the gate terminal.

these simulations were conducted at a frequency of 25 GHz, the capacitance shows little variation with frequency and a similar curve is produced at lower frequencies as well.

3.3.3 Channel Resistance Model

A secondary effect of the creation of a depletion region beneath the varactor gate terminal is that the effective cross sectional area through which the signal can propagate through the channel is now reduced. This means that increasing the width of the depletion region should result in an increase in the channel resistance. By adapting the channel resistance equation given in [14] (which includes considerations for the shallow trench isolation regions used), the following equation can be used to model the channel resistance

$$R_{ch} = \frac{\frac{L_{eff}}{2}}{qN_{well}\mu A_{ch}n} \tag{3.19}$$

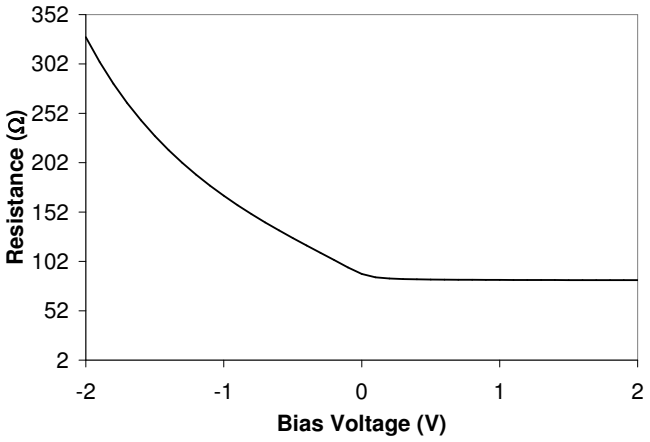

Figure 3.11: Channel resistance for a single finger varactor with a bias voltage applied to the gate terminal.

where L_{eff} is the effective channel length, simply calculated as

$$L_{eff} = L - 2L_{ov} \tag{3.20}$$

and where A_{ch} is the cross-sectional area of the channel given by

$$A_{ch} = W_{eff}\,(X_j - w_d) \tag{3.21}$$

and X_j is the depth of the n^+ diffusion region. It should be noted that a factor of $\frac{1}{2}$ has been added to L_{eff} since in this model C_{var} is considered to be a lumped capacitance connected to the center of the channel, leaving a distance of $\frac{L_{eff}}{2}$ between this point and either of the n^+ diffusion regions. As expected, the channel resistance curve shown in Figure 3.11 for a single finger varactor with $W = 2.5$ μm displays the same behaviour as the depletion region width.

3.3.4 Substrate Model

For an AMOS varactor designed in bulk CMOS, an accurate model of the well and substrate effects is vital in order to predict the device behaviour. While accurately

modeling all of the parasitics associated with the substrate is a time consuming and challenging task, the simplified substrate model presented in this section addresses the most significant elements, which is adequate to accurately predict varactor performance. As illustrated in Figure 3.6, the most significant elements of the substrate model are the n well resistance, R_{well}, the capacitance across the depletion region at the n well to p substrate interface, C_{well}, the capacitance between the n^+ diffusion regions and the p substrate, C_{act}, and the capacitance and resistance between the edge of the n well and the physical ground contact, R_{sub} and C_{sub} respectively.

While substrate lumped element models, ranging from a few elements in [15] to the eight-element configuration in [22], have been presented in recent literature, almost all of these models obtain component values using parameter extraction methods from measured data. The downside to this approach is that very little is learned concerning the origins of the component values, making it extremely difficult to use these extracted models to account for any layout effects. As a result, the application of any such model in varactor design will greatly reduce the level of confidence in the predicted varactor behaviour. As an alternative, the substrate model presented in this section applies semiconductor physics based models to each of the lumped elements and, although some simplifications must still be made, it eliminates much of the guesswork traditionally involved when using substrate models based on parameter extraction.

We begin by calculating the capacitance C_{act}, which can be considered to be a simple parallel plate capacitance between the charge layer formed at the n^+ diffusion to n well interface and the n well to p substrate interface, with the n well acting as the dielectric. The resulting equation for this calculation is therefore a function of

the total area beneath all n^+ diffusion regions, A_{act}, as

$$C_{act} = \frac{\epsilon_{si}}{D_n - X_j} A_{act} \tag{3.22}$$

Next, by applying an equation similar to equation 3.19, the vertical resistance from the n^+ diffusion regions, through the n well, to the p substrate can be obtained. By taking the total area of the n well region as the cross-sectional area, A_{well}, to be used in the resistance equation, we obtain the following equation

$$R_{well} = \frac{D_n - X_j}{qN_{well}\mu_n A_{well}} \tag{3.23}$$

where D_n is the depth of the n well and μ_n is the electron mobility.

A similar approach can also be applied to determine the substrate resistance and capacitance seen between the n well and the substrate ground contacts. The signal path for this element is extremely difficult to physically model due to the fact that the cross sectional area through which the signal flows in the substrate is ambiguous since, although the substrate has a very large depth in relation to the n well, it is unlikely that signal will utilize all of this area since it will involve higher resistance path.

In comparing possible modeling techniques to measured data, acceptable simplifications that can account for this issue were determined. In calculating the resistance it was determined that using the distance between the nearest ground contact and the edge of the n well, D_{gnd}, as illustrated in Figure 3.12, is an effective approach.

Secondly, using the cross-sectional area of one side of the p^+ ground contact to constrain the area through which the signal can flow from the n well charge layer produces accurate results for the substrate resistance. This area, A_{sub}, can be therefore calculated as follows

$$A_{sub} = X_p L_{con} \tag{3.24}$$

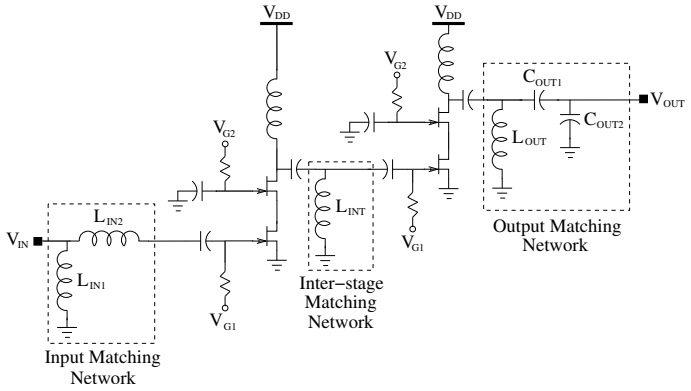

Figure 3.12: Illustration of the length used to calculate resistance between the n well and the ground contacts.

where X_p is the depth and L_{con} is the length of the p^+ ground contact implanted in the substrate. By using this area in place of A_{well} and using the distance between the n well and the ground contact, D_{gnd}, in place of D_n in equation 3.23, the resistance, R_{sub}, seen between the n well and ground contact can be calculated as

$$R_{sub} = \frac{D_{gnd}}{qN_{sub}\mu_n A_{sub}} \tag{3.25}$$

where N_{sub} is the substrate carrier concentration.

The area to be used in the substrate capacitance is slightly more involved since the effects of fringing, both through the substrate and through the oxide must be considered, as well as the capacitance between the n well and the sides and bottom of the p^+ ground contact. An acceptable simplification that can be made to account for these effects is to use the area of the side of the n well ($A_{well}^{side} = L_{well}D_n$) to calculate the substrate capacitance as

$$C_{sub} = \frac{\epsilon_{si}}{D_{gnd}} A_{well}^{side} \tag{3.26}$$

It should also be noted that if multiple ground connections are in place in the circuit layout then the resulting resistances and capacitances should be considered to be in parallel.

Finally, only the capacitance encountered between the n well and p substrate, C_{well}, remains to be calculated. The varactor configuration discussed thus far relies on a bias voltage applied to the gate terminal, meaning that the RF signal enters the varactor's n well through the n^+ diffusion regions. This in turn means that whatever DC bias voltage, V_{DC}, is in place in the circuit at the varactor's drain/source connection point is also passed to the n well. In most circuit applications this DC bias has a positive value, effectively reverse biasing the $p-n$ junction at the well to substrate interface and allowing C_{well} to be calculated using the following equation from [34]

$$C_{well} = \epsilon_{si} A_{well} \sqrt{\frac{q}{2\epsilon_{si}\left(\Phi_{Bi} + V_{DC}\right)} \left(\frac{N_{sub} N_{well}}{N_{well} + N_{sub}}\right)} \tag{3.27}$$

where Φ_{Bi} is the built in electrostatic potential barrier, which is defined as the difference between the fermi levels of the p and n materials, Φ_{Fp} and Φ_{Fn} respectively. These values are obtained using the following equations from [32]

$$\Phi_{Fp} = \frac{kT}{q} \ln\left(\frac{N_{sub}}{N_i}\right) \tag{3.28}$$

$$\Phi_{Fn} = -\frac{kT}{q} \ln\left(\frac{N_{well}}{N_i}\right) \tag{3.29}$$

$$\Phi_{Bi} = \Phi_{Fp} - \Phi_{Fn} \tag{3.30}$$

where T is the operating temperature and N_i is the intrinsic carrier concentration in undoped silicon.

With all four of the substrate model lumped elements defined, it is apparent that decisions made during layout such as the n well size, its distance from ground contacts, and the size and number of ground contacts can have a significant effect on the values of these components. It is therefore useful to analyze the effect that these choices can have on circuit performance. Figure 3.13 illustrates the simplified lumped

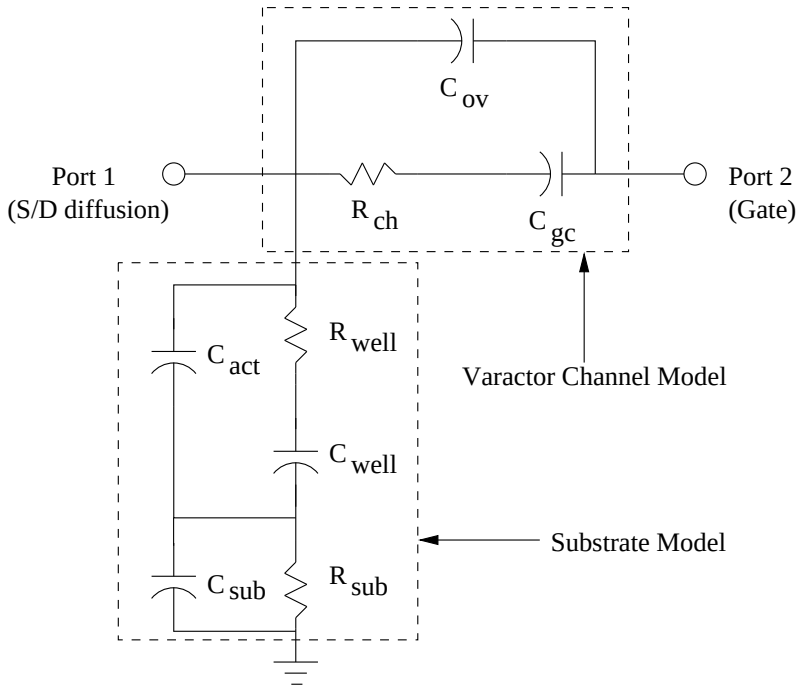

Figure 3.13: Simplified lumped element varactor model used to analyze the effect of substrate parasitics.

element model, with the components that do not have a significant effect on varactor behaviour removed, that was used to analyze the effects of the substrate parasitics on varactor behaviour.

Since the parasitic substrate model is composed of parallel and series RC circuit components in parallel with the varactor channel model components, it is difficult to intuitively visualize the effects of the individual substrate component values on the overall varactor Q factor. Instead ADS simulations analyzing the Q factor as a function of the substrate component values were performed. First, since the circuit designer has control over the distance between the ground contacts and the varactor n well, a degree of control over the values of R_{sub} and C_{sub} can be achieved. Since these components are located in the substrate signal path, increasing their impedance

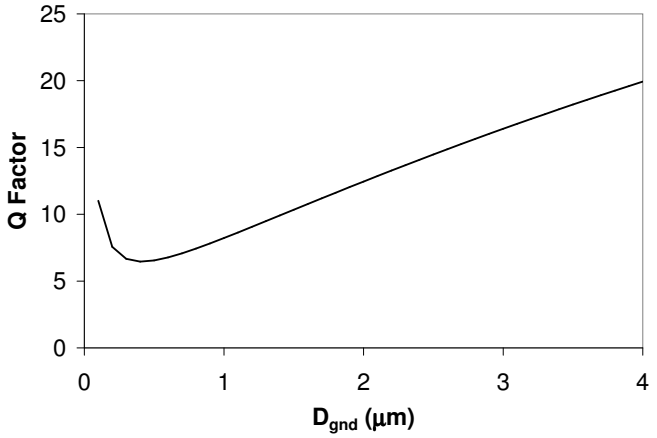

Figure 3.14: Effect of ground contact placement on the Q factor of a 9-finger varactor.

should reduce the signal flow through this parasitic path and increase the device Q factor as a result. To verify this, Figure 3.14 illustrates that, at 25 GHz, as D_{gnd} is increased, thereby increasing R_{sub} and decreasing C_{sub}, the Q factor of a 9-finger varactor (which will be discussed more later) tends to increase as well. This suggests that designers should arrange ground contacts far from the varactor wells so as to increase the impedance between these wells and the ground contacts in order to maximize Q factor.

The other substrate parameter over which the circuit designer has some freedom is the size of the n well in which the varactor is placed. While this well must be large enough to encompass the gate fingers and the n^+ diffusion regions, design rules typically allow for the well to be made larger than necessary. Although the value of C_{act} is dependent on fixed process parameters such as the depth of the n well and n^+ diffusion regions, increasing the area of the well has the effect of reducing R_{well} while at the same time increasing C_{well}. Simulation results show that the net result of this behaviour is actually an increase in the series resistance and a decrease in capacitance seen at the varactor input. This results in a decrease in Q factor, as is illustrated in

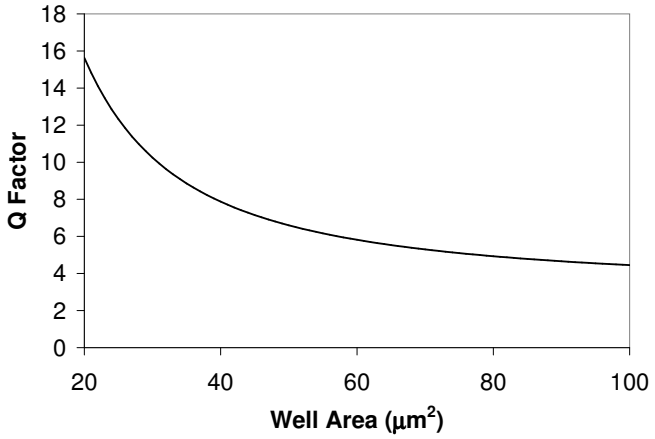

Figure 3.15: Effect of n well area on the Q factor of a 9-finger varactor.

Figure 3.15, leading to the conclusion that in order to maximize the Q factor of the varactor, the n well size should be minimized.

3.3.5 Additional Parasitics

The final pieces of the varactor model puzzle that remain to be calculated are the parasitic resistances and inductances associated with the varactor contacts, as identified in Figure 3.6. While these components are usually very small in comparison to the lumped elements discussed thus far, they can still have a noticeable effect on the performance of varactors designed for low capacitance values and should be included for best accuracy. First, the resistance inherent to the use of n^+ polysilicon as the gate material, R_g, has been well documented [14] and can be calculated as

$$R_g = \frac{W_{eff}}{L}\frac{R_{sq}^g}{12n} \tag{3.31}$$

where R_{sq}^g is the resistance per square of the gate material and the factor of 12 is used to account for the fact that the gate fingers are connected at each end.

At the opposite varactor terminal, the resistance through the n^+ diffusion region

must be considered. This resistance is calculated in a similar fashion to R_g but relies on the length of the n^+ diffusion region, L_{dif}, and the sheet resistance of the diffusion region, R^d_{sq}, as can be seen in the following equation from [35]

$$R_{sd} = \frac{\frac{L_{dif}}{2}}{2nW_{eff}} R^d_{sq} \tag{3.32}$$

The final parasitic component, which is not a physical part of the varactor but is nonetheless necessary in order to connect the varactor to the appropriate point in the circuit is the inductance of the connecting via, L_g. While this component depends largely on the metal layer and vias used to connect the varactor to the rest of the circuit and therefore varies a great deal, it is still useful to provide a general formula, which can be easily adapted for any layout dimensions. To obtain this formula we begin by treating the via as a dipole antenna and determining the magnitude of the non-radiative term of the magnetic field equation given in [36] as

$$\left|\vec{H}\right| = \frac{I_0 L_{via} \sin\theta}{4\pi r^2} \tag{3.33}$$

where L_{via} is the length of the via and I_0 is the current through the via. Since the inductance of the via is related to the integral over volume of the magnetic field as the field radius approaches infinity, given in [37] by

$$L = \frac{\mu}{I_0^2} \int\int\int \left|\vec{H}\right|^2 dV \tag{3.34}$$

we finally obtain the following result for the inductance of the via region

$$L_g = \frac{\mu_0 L_{via}}{6\pi r_{via}} \tag{3.35}$$

where r_{via} is the radius of the via or equivalent radius of an array of vias.

It is worth noting that several other via regions are also present in the varactor layout which are used to provide contacts to the gate fingers, diffusion regions and

ground contacts. As a result, if these sections are closely spaced, it may be possible for significant parasitic capacitances to be generated between these regions. However, in the layouts examined in this thesis all adjacent sections were separated by relatively large distances. This, combined with the very small surface area of these vias, meant that the resulting capacitances were negligible (typically much less than 1 fF) and were hence neglected from further analysis.

3.4 Measured Results

Using the modeling theory discussed in the previous section a complete model can now be simulated to accurately predict the capacitive and resistive behaviour of an n finger varactor. However, the difficulty associated with using a single varactor is that changing the DC control voltage applied to either of the varactor terminals will affect the DC bias voltage seen by the circuit at that point. One method that is commonly used to eliminate this problem is a series, back-to-back connection of two identical varactors, which isolates the varactor control voltage from the circuit connection points. This differential structure is illustrated by the lumped element representation in Figure 3.17, which shows a clear DC isolation between V_{bias} and circuit connection points Port 1 and Port 2. The two possible configurations for these differential AMOS varactors are the gate biased and diffusion biased topologies as shown in Figures 3.16 (a) and (b), respectively. In this section, the varactor structure shown in Figure 3.16 (a) is examined.

In order to verify the accuracy of the proposed model, a back-to-back varactor with the same configuration as that shown in Figure 3.17, was fabricated in 0.18 μm CMOS by labmate John Carr for publication in [38]. Each varactor in the design

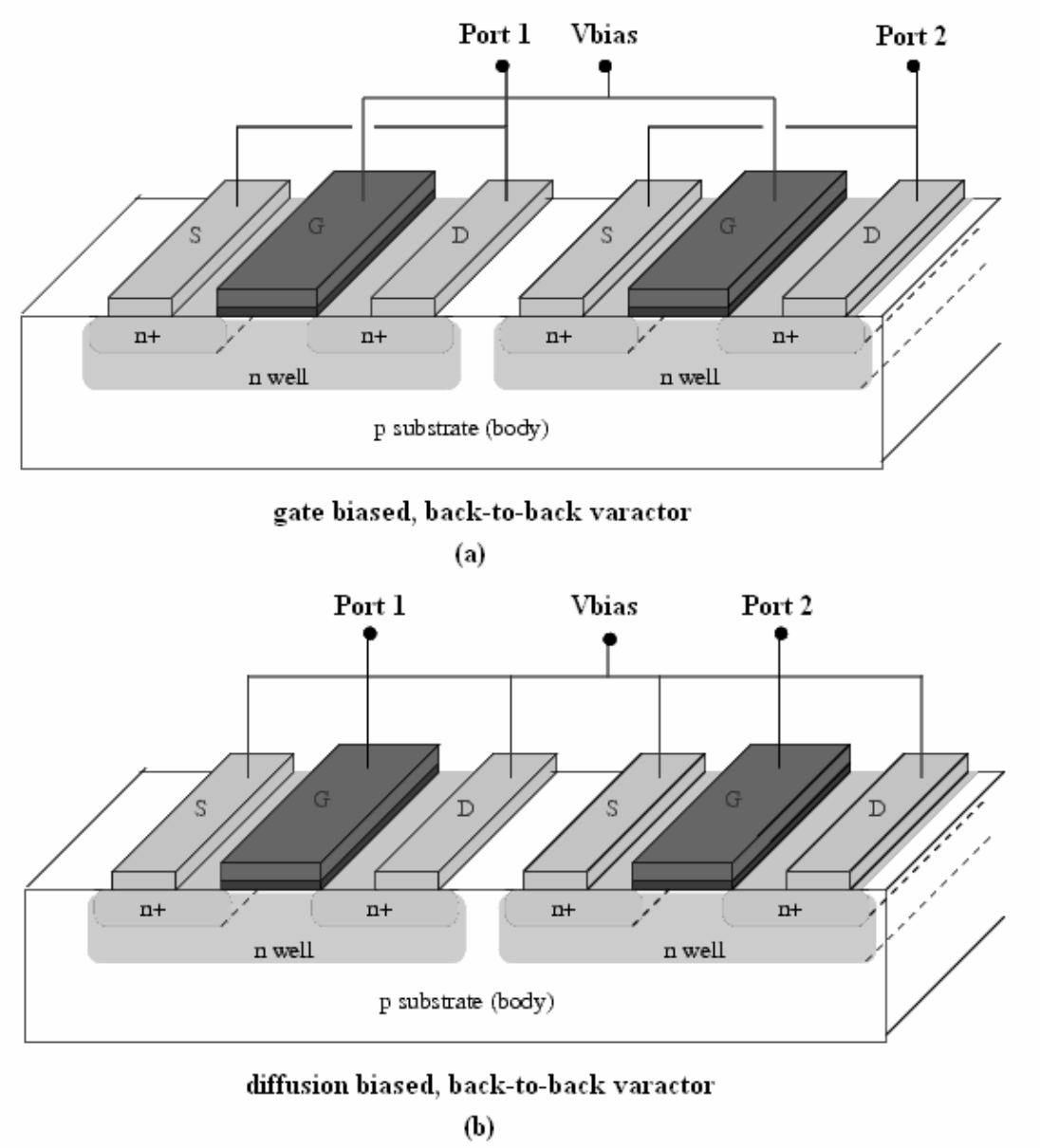

Figure 3.16: Two possible differential AMOS varactor configurations.

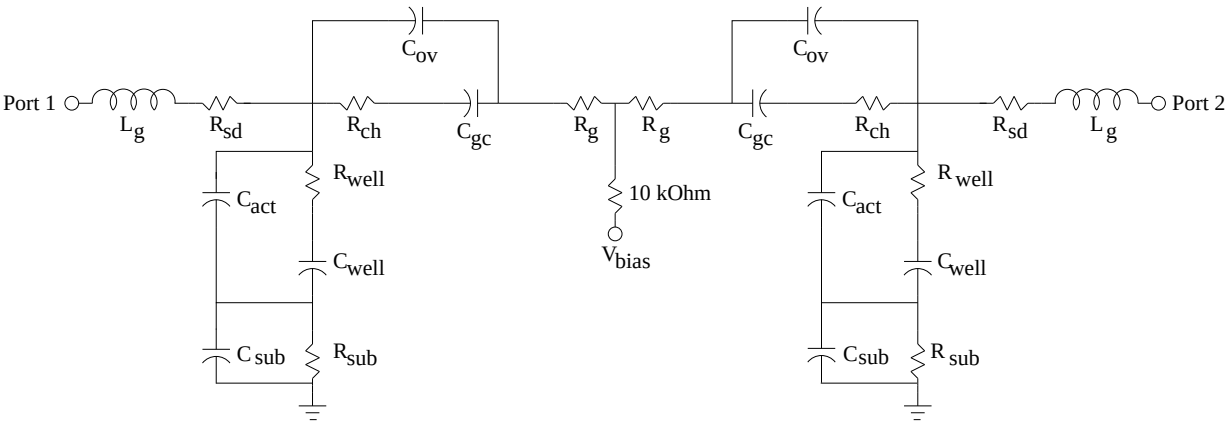

Figure 3.17: Back-to-back varactor model including all parasitic components.

consists of 9 gate fingers (and therefore 10 diffusion regions) and was designed to operate at 25 GHz in a circuit with a DC voltage of +2.8 V at both the input and output terminals. As a result, although the x-axis of all of the plots to follow show the application of a negative control voltage, these levels are actually in relation to the 2.8 V seen at the varactor ports 1 and 2 so the absolute control voltage sweep ranges from +0.8 to +4.8 V.

At the time of fabrication the information presented earlier for maximizing the Q factor was not yet available and as a result, the n well dimensions were not minimized and the ground contacts were not optimally arranged. However, simulation results indicate that this optimization would only increase the Q factor of the back-to-back structure by approximately 10 % and since the model used to account for the effects of these layout decisions has been well-defined in the previous sections, it is still possible to accurately predict the varactor behaviour.

To avoid any uncertainty introduced by attempting to deembed the parasitic effects of the fabricated test structure used, this test structure (consisting of the top metal ground plane, feed lines and probing pads) was simulated in ADS Momentum

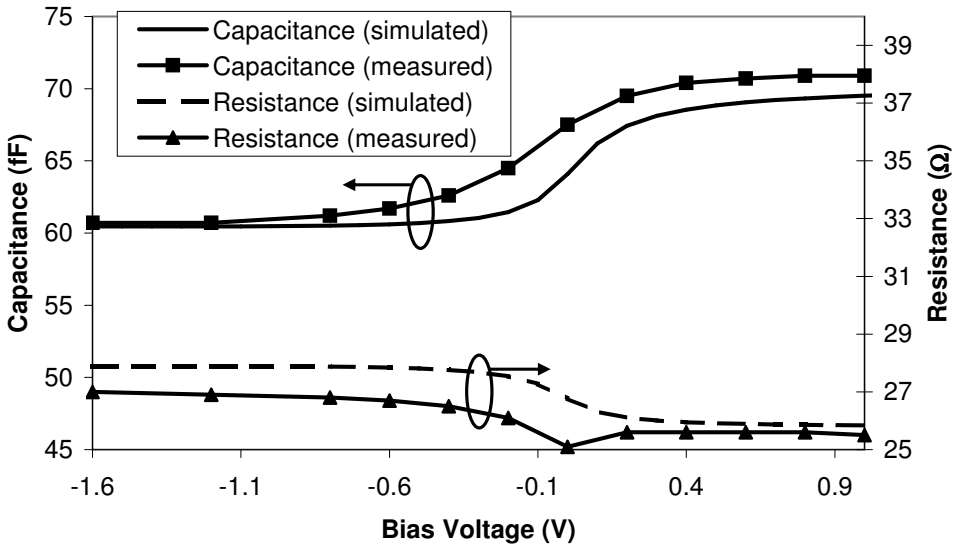

Figure 3.18: Simulated and measured data of two back-to-back, gate biased varactors.

and the lumped element varactor model was then used in conjunction with this to accurately duplicate the measured results. Figure 3.18 compares the simulated capacitance and resistance of this varactor structure with the measured data collected by labmate John Carr [38], showing good agreement between the two sets of curves. It should be noted that the resistance includes the 50 Ω terminations at ports 1 and 2. The n well area of the fabricated varactor was $A_{well} = 51.4\ \mu m^2$ and the substrate resistances, calculated using the method outlined in the previous section, were R_{sub}^{in} = 527 Ω and R_{sub}^{out} = 1074 kΩ for the input and output varactors, respectively. The difference between these two values is due to differences in the ground contacts placed around each n well region.

It is also useful to examine the behaviour of the varactor with the effects of the test structure removed, in order to facilitate implementation in other circuit designs and comparison with other published results. Simulations of the capacitance and resistance of a single varactor with the same number of fingers and layout geometry as

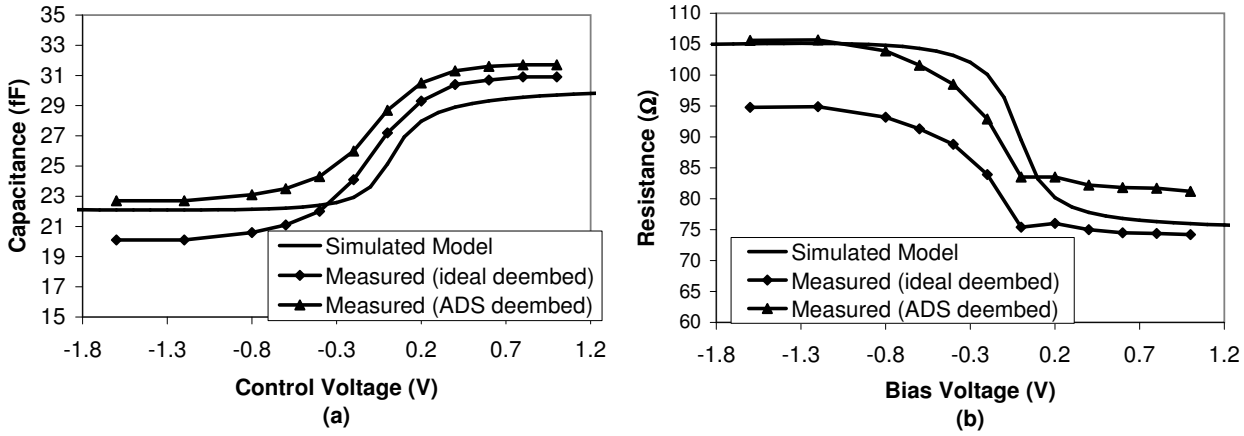

Figure 3.19: Deembedded data of the 9-finger, back-to-back, gate biased varactor.

presented above were conducted with ADS at a frequency of 25 GHz and are displayed in Figure 3.19, comparing these results with two different methods for deembedding the effects of the test structure. The first method models the test structure as a simple series connection of an ideal capacitance of $C_{pad} = 38$ fF and an ideal resistance of $R_{pad} = 36\ \Omega$ between the input and output terminals of the test structure and ground, resulting in an impedance of

$$Z_{pad} = \frac{1}{Y_{pad}} = R_{pad} + \frac{1}{j\omega C_{pad}} \tag{3.36}$$

Using this information, the deembedded results can be extracted from the measured data by subtracting Y_{pad} from both Y_{11} and Y_{22}. The second method uses EM simulation data of the test structure and the deembedding function in ADS to remove the effects of the test structure. As is clear in Figure 3.19, the presented model compares well with either deembedding method.

Unfortunately, no published data could be found for varactors in bulk 0.18 μm CMOS at frequencies above 20 GHz. This, combined with the fact that other publications rarely include all the layout dimensions necessary to recreate all of the model

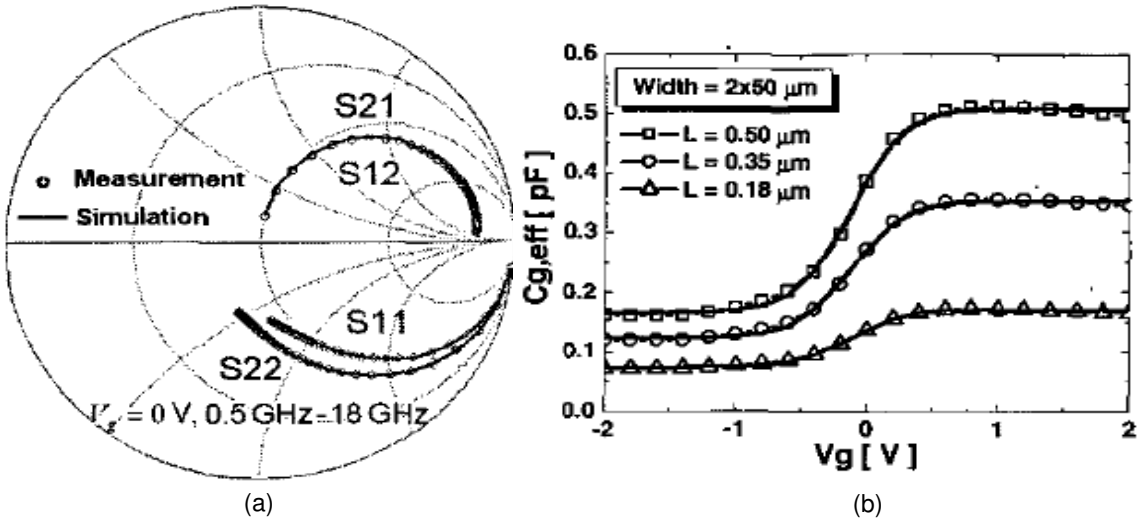

Figure 3.20: (©2003 IEEE) Simulated and measured (a) S-parameter data and (b) channel and overlap capacitance of a 50 finger, $L = 0.5$ μm, $W = 2$ μm varactor reproduced from [15].

components, makes direct comparison to other reported data difficult. Results, however, are encouraging since general curve shapes match those reported in [15], [14], [39] and [13], which use either SOI CMOS or bulk CMOS at significantly lower frequencies. This is illustrated by Figure 3.20 (a), which shows the measured and simulated S-parameters at frequencies between 0.5 and 18 GHz for a 50 finger varactor with gate dimensions $L = 0.5$ μm and $W = 2$ μm, obtained in [15]. This is complemented by Figure 3.20 (b), which shows the combined channel and overlap capacitance, $C_{g,eff}$, for three different 50 finger varactors with gate finger widths of $W = 2$ μm and lengths varying from 0.18 to 0.5 μm, each of which were fabricated in a 0.18 μm CMOS process.

Figure 3.21 then presents the same parameters obtained through simulation of the lumped element model presented in this chapter. Although the dimensions of the n well and substrate ground contacts necessary to calculate the substrate components

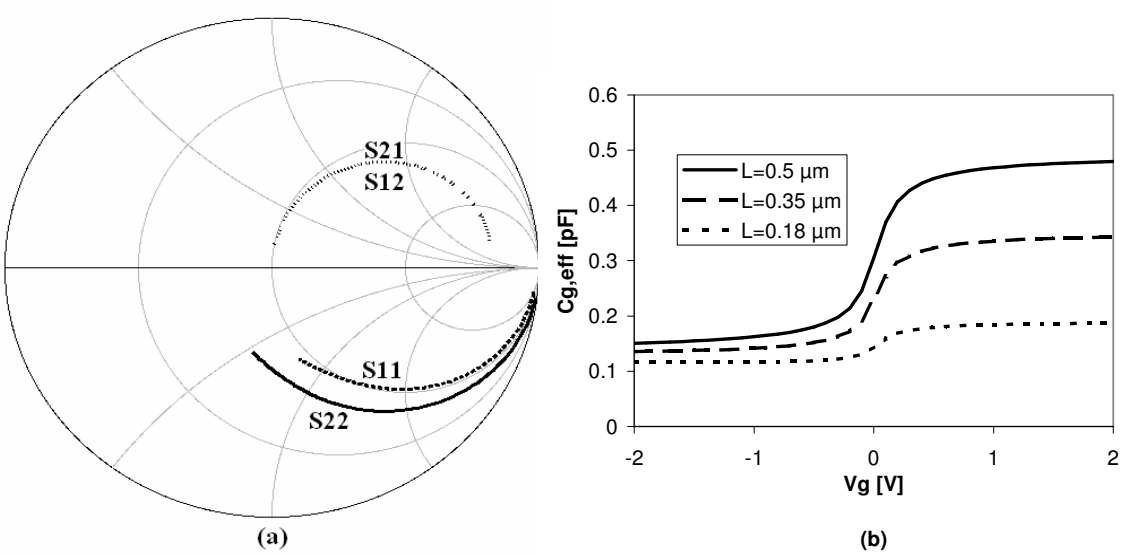

Figure 3.21: Simulated (a) S-parameters from 0.5 to 18 GHz and (b) channel and overlap capacitance of the varactor presented in [15] using the lumped element model presented in this section.

were not included in [15], the extracted lumped element values given for the substrate model in the publication were used instead. Despite the difference in operating frequency and possible variations in the CMOS processes used, in comparing these two figures it is clear that a good match is obtained, helping to validate the model presented in this chapter.

3.5 The Diffusion Biased Alternative

In the case of the varactor structure discussed in the previous section, with a control voltage applied to the gate terminals, the RF signal enters the varactor through the n^+ diffusion regions. When this occurs, the signal is immediately exposed to the substrate parasitics, impacting the total capacitance and resistance seen and reducing the Q factor of the device. One possible way to mitigate these effects is to instead

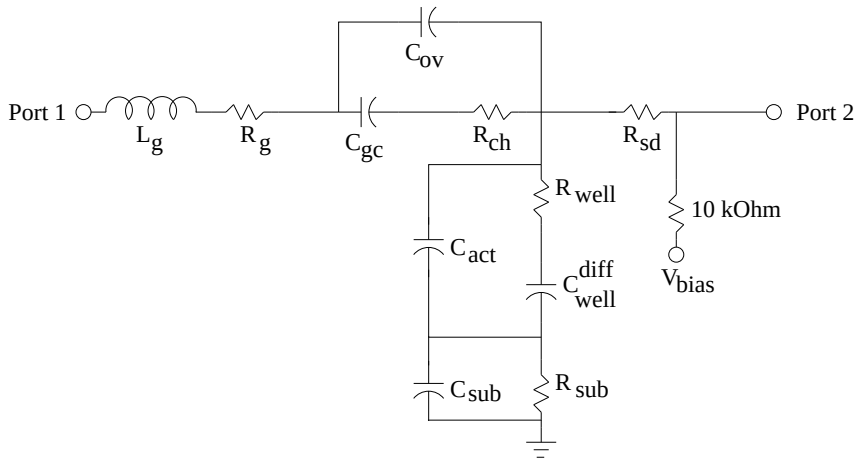

Figure 3.22: Lumped element model of a diffusion biased varactor.

apply the control voltage to diffusion regions and apply the RF signal to the gate terminal. The resulting lumped element model is shown in Figure 3.22.

With the varactor arranged in this way, the substrate parasitics are not encountered until after the RF signal has passed through the first AMOS varactor channel capacitance and resistance, reducing the overall effect of the substrate. This has the inherent effect of decreasing the total capacitance of the circuit since the substrate capacitance is no longer in parallel with the oxide capacitance, as discussed in [40], but can also increase the Q factor significantly. The equations used to calculate all of the varactor model components in the gate biased configuration discussed previously still apply to the diffusion biased case with the exception of C_{well}. Since the DC control voltage, V_{bias}, is now being applied to the n^+ diffusion regions (and hence the entire n well itself) the bias voltage seen across the $p-n$ junction between the n well and p substrate must vary with V_{bias}. This is reflected by redefining equation 3.27 to depend on the value of V_{bias} instead of the fixed V_{DC} at that point of the circuit as

follows

$$C_{well}^{diff} = \epsilon_{si} A_{well} \sqrt{\frac{q}{2\epsilon_{si}\left(\Phi_{Bi} + V_{bias}\right)} \left(\frac{N_{sub} N_{well}}{N_{well} + N_{sub}}\right)} \tag{3.37}$$

which remains valid as long as V_{bias} is positive, keeping the $p - n$ junction in reverse bias.

To illustrate the effects of connecting the varactor in this manner, Figure 3.23 compares the simulated (a) capacitance, (b) resistance and (c) Q factor of the back-to-back, 9-finger, gate biased varactor presented in the previous section, to an identical varactor with the bias voltage applied to the diffusion region instead. In addition to reversing the shape of the curves due to the fact that the bias voltage is now applied to the diffusion regions instead of the gate, the removal of the substrate parasitics from the RF input of the varactor results in a reduction of parasitic effects. This lowers the total capacitance and series resistance, and increases the Q factor of the device. Although the Q factor of both devices appears to be relatively low, this is due in part to the inclusion of the entire measurement test structure in the simulations and in part to the fact that the back-to-back structure effectively doubles the resistance and halves the capacitance of the varactor structure, creating a reduction by a factor of 4 in the overall Q factor.

This back-to-back, diffusion biased configuration of the varactor was also fabricated in bulk 0.18 μm CMOS using the same test structure as the gate biased case presented above. Figure 3.24 compares the measured results with that predicted by the varactor model developed thus far. As with the gate biased case, the lumped element model provides an accurate representation of the measured results and adds further evidence that this physical model can provide an accurate representation of

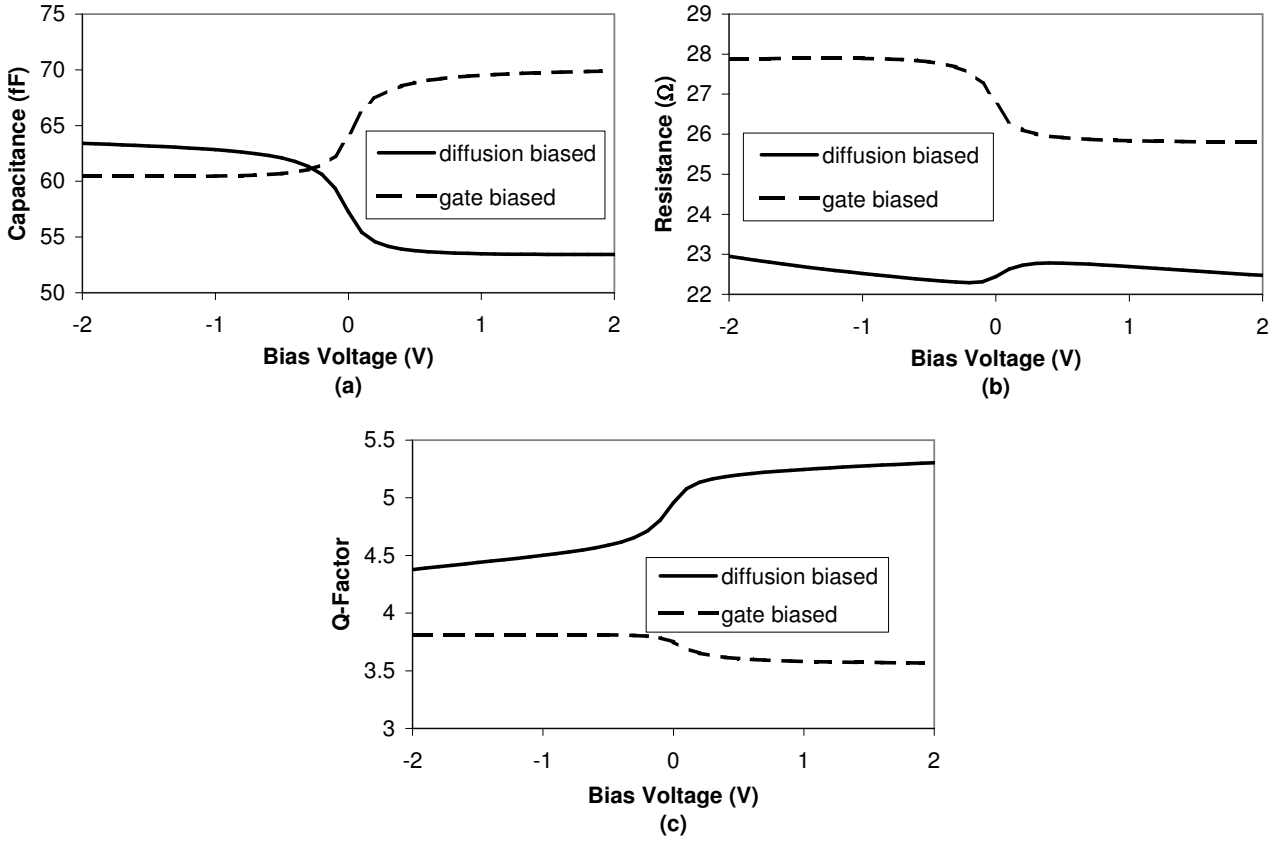

Figure 3.23: Simulated effect on (a) capacitance, (b) resistance and (c) Q factor when applying the varactor DC control voltage to the diffusion regions.

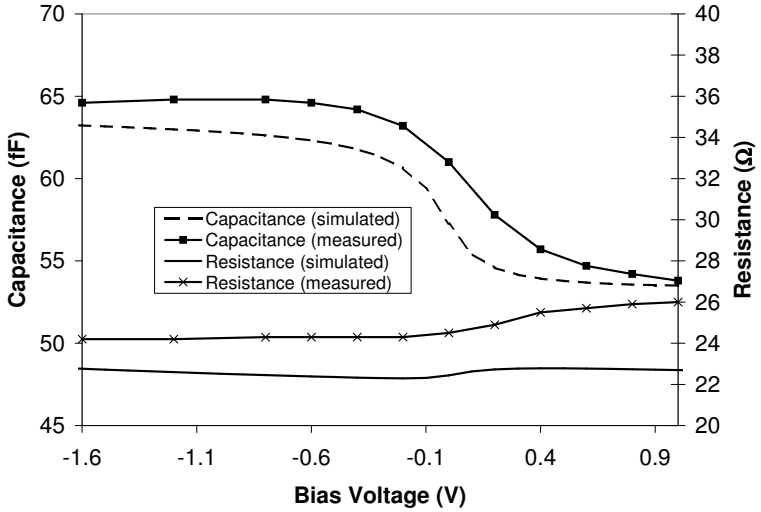

Figure 3.24: Measured and simulated data for a back-to-back diffusion biased varactor.

varactor behaviour, allowing for the varactor to be implemented in other circuit designs with confidence. This is extremely beneficial to circuit designs at frequencies above 20 GHz and as such, these results have been submitted for publication in January 2007 [41].

3.6 Conclusion

An equivalent lumped element model has been presented for an AMOS varactor, which is valid for frequencies above 20 GHz. This model is also valid in any mode of operation, making it easily implementable in circuit simulators and has the advantage of using semiconductor physics based equations to determine the values for each of its components. As a result, the proposed model can be applied to any bulk CMOS process and is flexible enough to account for individual layout geometries. This has the added advantage of allowing for optimization of the layout and shows that in

order to optimize the Q factor, the circuit designer should minimize the required n well size and increase the resistance seen between this well and the surrounding ground contacts wherever possible.

In order to validate this model, two 9-finger varactors were fabricated in 0.18 μm bulk CMOS. Both structures were arranged as two varactors connected in series so as to isolate the control voltage applied in between these two devices from the rest of the circuit. In one structure this bias voltage was applied to the gate terminals of the varactors and in the other to the diffusion regions. The proposed model shows excellent agreement with measured results for both of these configurations, allowing it to be used with confidence when implementing these varactors in other circuit designs.

Chapter 4

Single-Ended LNA Design

4.1 Introduction

The low-noise amplifier (LNA) is an integral part of any wireless receiver. As outlined in Chapter 1, its noise and gain performance play a major role in the overall sensitivity of the receiver, making its optimization and implementation of paramount importance in any receiver structure. As a result, LNAs have received their fair share of attention in recent publications, including those dealing with CMOS circuits designed to operate above 20 GHz.

Since silicon MOSFETs benefit from lower transconductance and capacitance per unit gate width than older III-V semiconductor technologies of comparable or higher f_T and f_{MAX}, it should in theory be possible to create lower noise figures in CMOS technology [3]. Unfortunately, due to the process variation, modeling inaccuracies and low-Q passive structures discussed previously, this is rarely achieved in practice and published results display LNA noise figures (NF) typically on the order of 4 to 9 dB [42], [43], [17]. Not only are these results above the minimum achievable noise

figures in 0.18 μm CMOS, but they also exhibit a larger degree of variation than can be explained by the different design techniques used. In fact, even single publications reporting multiple LNAs using the same design technique, such as [18], can show variations in NF greater than 1 dB. This underscores the difficulties encountered by circuit designers attempting to minimize the noise produced by high-speed CMOS circuits.

This chapter uses the novel SSTL inductor concept explained in Chapter 2 in a single-ended, cascode LNA design. With the improved modeling and Q factor of these inductors, it should be possible to improve the noise figure of the LNA and push noise performance closer to minimum achievable values. It also explains, in detail, the steps taken in the design and optimization of this device including the transistor size and matching network selection in an attempt to eliminate any guesswork typically encountered by circuit designers working with high-speed CMOS circuitry.

4.2 Transistor Size and Bias Condition

The first parameter to be determined in any LNA design is the size of the transistors to be used. This decision will have a direct impact on the minimum achievable noise figure, NF_{min}, of the circuit as defined by [44]

$$NF_{min} = 1 + \frac{2}{\sqrt{5}}\frac{\omega}{\omega_T}\sqrt{\gamma\delta\left(1 - |c|^2\right)} \tag{4.1}$$

where ω is the operating frequency in radians per second, ω_T is the unity current gain frequency in radians per second defined as the ratio of device transconductance, g_m, to the gate-source capacitance, C_{gs}. Terms γ, δ and c are noise parameters which can be taken to be constants whose values are displayed in Table 4.1. While these values

Noise Parameter	G_{opt}	α	γ	δ	c
Value	0.02	0.85	1.2	8.2	j0.2

Table 4.1: Noise parameters for 0.18 μm CMOS at 24 GHz [45].

are based on the results reported in [45] for frequencies up to only 6 GHz, they show very weak frequency dependence and should therefore be valid at 24 GHz. It should also be noted that parameters γ and δ are directly dependent on the bias conditions applied to the transistor and that G_{opt} and c, which are functions of γ and δ, can also show bias-dependent variations [44]. As a result, if the optimum bias conditions that are found using the technique presented later in this section differ from those used to find the optimum transistor size, it may be necessary to repeat the optimum transistor size calculations with updated values for the parameters in Table 4.1.

The optimum transistor size for minimum noise generation has been well documented by CMOS researchers and can be extracted in a number of different ways depending on the design method chosen and power limitations. Ignoring methods that use inductive degeneration due to the resulting noise degradation, as will be discussed later, the optimum transistor size can be obtained from the optimum gate-source capacitance, C_{gs}. If power consumption is not a concern, the optimum value for C_{gs} can be obtained using the classical noise matching (CNM) technique presented in [44] as

$$C_{gs}^{CNM} = \frac{G_{opt}}{2\pi f \alpha \sqrt{\frac{\delta}{5\gamma}\left(1 - |c|^2\right)}} \tag{4.2}$$

where f is the operating frequency and G_{opt} and α are noise parameters as defined in Table 4.1. Using this CNM method, and the noise parameter values displayed in Table 4.1, the optimum device width at 24 GHz is found to be 203 μm. This size is extremely large and therefore consumes too much power to be feasible for most

applications.

As an alternative, a lower optimum C_{gs} can be found using the power constrained noise optimization (PCNO) technique presented in [44] as

$$C_{gs}^{PCNO} = \frac{3}{2\omega R_s Q_{sP}} \tag{4.3}$$

Here $R_s = 50\ \Omega$ and Q_{sP} is a quality factor defined for a value of source conductance, G_s, that is different than the optimal noise value G_{opt} and is calculated as

$$Q_{sP} = |c| \sqrt{\frac{5\gamma}{\delta}} \left[1 + \sqrt{1 + \frac{3}{|c|^2} \left(1 + \frac{\delta}{5\gamma} \right)} \right] \tag{4.4}$$

Using the noise parameter values from Table 4.1 once again, the PCNO method yields an optimum device width at 24 GHz of 118 μm. Although this results yields a value for the minimum noise figure, NF_{min}, which is typically between 0.5 and 1 dB higher than the CNM case [44], this tradeoff is usually desirable for RF CMOS circuit designers.

Unfortunately, both noise models used in the above methods for determining optimum device size suffer from high frequency inaccuracies due to the fact that they neglect the gate-drain capacitance. Although this assumption is a good one for long channel devices operating well below f_T, it is not valid for deep submicron MOSFETs operating above 20 GHz. Instead, it is more appropriate to use the noise model derived more recently in [46], which uses the constant values given in Table 4.1 to ultimately achieve an expression for the optimum source resistance, R_{sopt}, given by

$$R_{sopt} = 0.2 R_g + \frac{0.25}{\omega^2 C_{gs}^2} - 0.1712 R_g^2 \omega^2 C_{gs}^2 \tag{4.5}$$

where R_g is the total gate resistance as discussed in Chapter 3. Since both R_g and C_{gs} are functions of the total device width, the optimum device width, W_{opt}, of the transistor can easily be determined by setting $R_{sopt} = 50\ \Omega$ and solving for W_{opt}.

Optimization Method	CNM	PCNO	HFNM
Optimum Width (μm)	203	118	80
Number of 2.5 μm Fingers	81	47	32

Table 4.2: Comparison of transistor size optimization methods for 0.18 μm CMOS at 24 GHz.

Using this technique, which we will call the high frequency noise matching (HFNM) technique, W_{opt} is found to be only 80 μm without power constraints applied. It is worth noting that this result is in stark contrast to the optimum device widths obtained previously and the results of all three methods, including the corresponding number of 2.5 μm wide fingers required to achieve each optimum width, are summarized in Table 4.2

The second parameter that can be considered before deciding on device topology is the bias condition to be applied to the transistors. While long channel devices display a monotonically decreasing noise figure with increased bias current (and hence power consumption), this is not the case for short channel devices. In this situation the introduction of velocity saturation means that an increase in the drain current when in saturation can increase the noise produced. As a result an optimum bias condition exists, which is identified in [46] to be achieved when the gate voltage, V_g, is set to approximately 1 V. While this result is accurate for the technology node in which it was reported, it will be different for each successive CMOS technology size, which inherently involves different threshold and saturation voltages.

As an alternative, it is possible to optimize the bias current drawn by the transistor as opposed to the gate voltage applied. As reported in [3], this parameter remains relatively constant regardless of the CMOS technology node being used, making it more versatile and easily applicable to new CMOS gate lengths. Using this technique,

it was determined that the optimum bias current of 0.25 $\frac{mA}{\mu m}$ could be obtained for an 80 μm width transistor by applying a gate voltage of approximately 0.9 V.

4.3 The Cascode Structure

With the transistor size and bias conditions set, the next step in the LNA design is to determine the topology to be used. Recent high-frequency CMOS publications have explored several single-ended options, the most common of which are illustrated in Figure 4.1. All of these topologies focus exclusively on minimizing the noise figure while maintaining power consumption levels low enough to be practical in most CMOS receiver applications.

Unfortunately all of the topologies illustrated in Figure 4.1 suffer from one of two shortcomings:

1. They do not address the issue of feedback created by C_{gd}. This creates coupling between the input and output, making matching network design extremely challenging and can also create instability.

2. They employ inductive feedback mechanisms to help mitigate the effects of C_{gd}, which hinder noise performance and are difficult to accurately implement, introducing another source of potential instability.

An alternative method that can be used to improve reverse isolation and address the high-frequency feedback issues associated with C_{gd} is the cascode topology, as displayed in Figure 4.2. The insertion of a Common Gate (CG) stage at the drain of the initial Common Source (CS) transistor eliminates the direct path from output to input via the C_{gd} of a single transistor and enhances reverse isolation considerably.

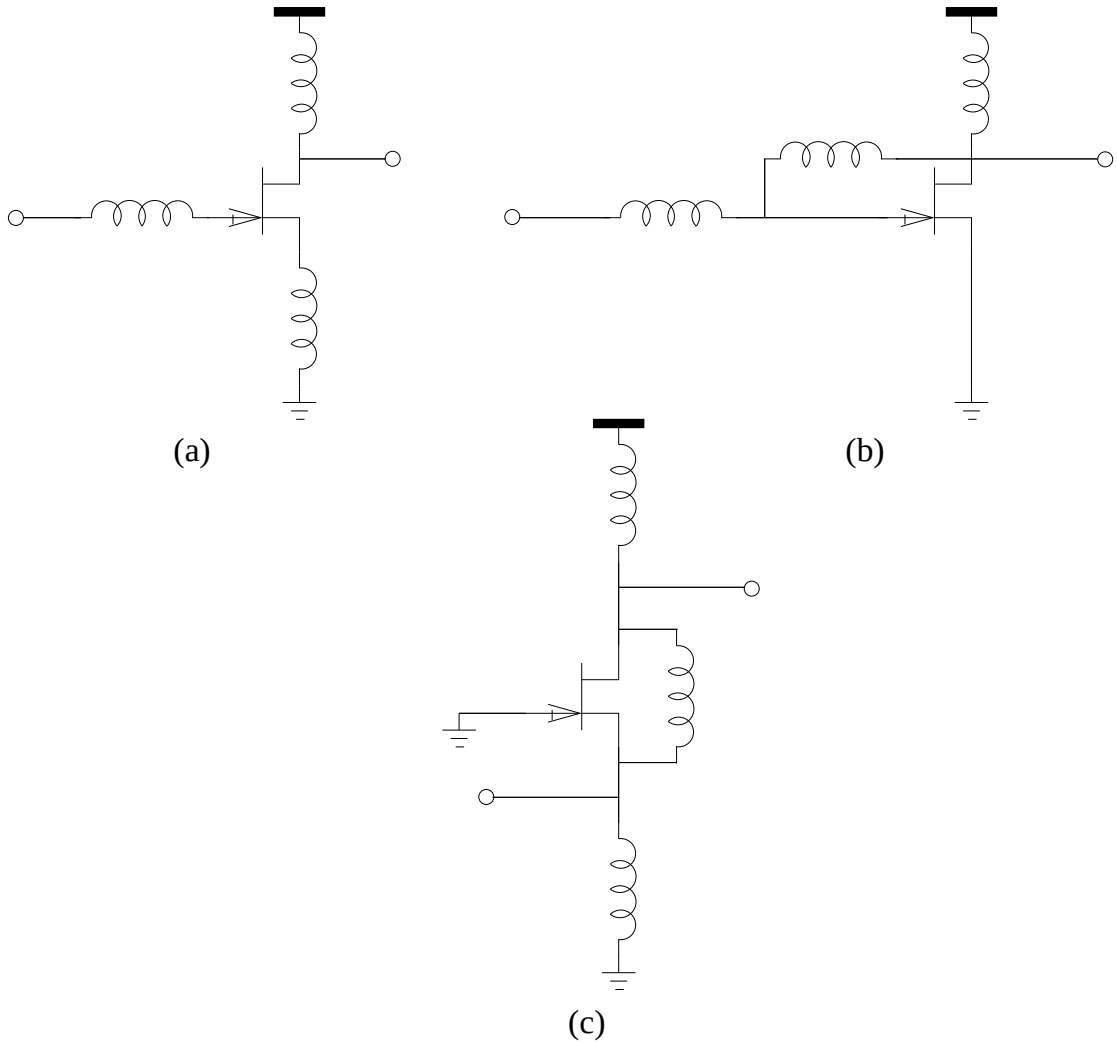

Figure 4.1: Common high-frequency CMOS LNA topologies (a) Common Source (CS) with degeneration, (b) CS with parallel feedback, (c) Common Gate (CG) with feedback.

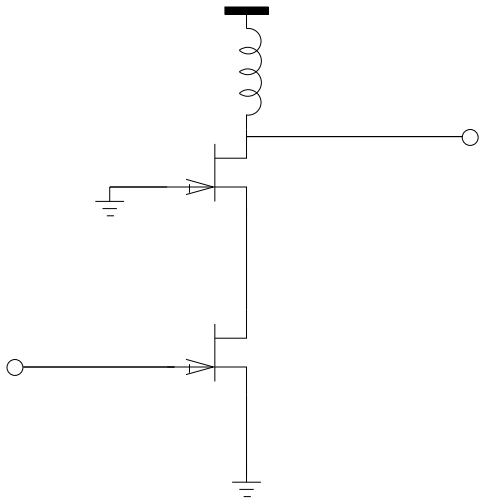

Figure 4.2: Cascode structure used to improve reverse isolation.

Although this configuration has seen extensive use at low frequencies, it has traditionally seen little use at frequencies above 20 GHz due to the fact that the addition of the CG stage typically increases noise figure by at least 0.5 dB. Also, the stacked transistor configuration requires a higher supply voltage rail and hence, increased power consumption. Although this problem can be overcome using the more recently proposed folded cascode structure [47], this technique increases the total bias current and hence power consumption of the circuit considerably. Despite these shortcomings, the large reduction in the feedback path and resulting increase in stability of the LNA obtained by using the cascode configuration should not be underestimated. This statement is backed by recent publications, such as [21], which reports a successful cascode LNA design in 90 nm silicon-on-insulator (SOI) technology with a noise figure of 3.6 dB and a gain of 11.9 dB at a frequency of 35 GHz.

Additional benefits of the cascode topology include increased gain from the stacked transistor configuration, which helps to minimize the noise contribution of successive amplifier stages, and the isolation of input and output matching sections, simplifying

the design of these networks considerably. These benefits, coupled with the increased Q factor obtained by using the SSTL inductors introduced in Chapter 2, make the cascode structure a suitable candidate for LNA design above 20 GHz, able to perform at noise levels comparable to the best previously published results for standard 0.18 μm CMOS.

4.4 Matching

The matching networks used at the input and output of the LNA, as well as between stages in a multi-stage amplifier are non-trivial matters that should be considered carefully. While the classic noise matching (CNM) technique requires that the designer sacrifice some of the amplifier gain in order to achieve the optimum noise match, other techniques have also been presented, and compared in [16], which use inductive degeneration in the source and gate of the transistor to help combine the optimum noise and impedance matching conditions. This technique is illustrated with the help of Figure 4.3.

To achieve the optimum noise matching condition, the impedance seen looking into the source as illustrated by Z_s in Figure 4.3 must be set equal to the optimum noise impedance, Z_{opt}^0, with the help of matching circuitry. The degeneration technique works using the principle that the addition of the source inductor, L_s, can change this optimum noise impedance by [16]

$$Z_{opt} = Z_{opt}^0 - j\omega L_s \tag{4.6}$$

where Z_{opt} is the new optimum noise impedance obtained after the addition of L_s. Since the input impedance of the amplifier is inherently capacitive, Z_{opt}^0 is typically

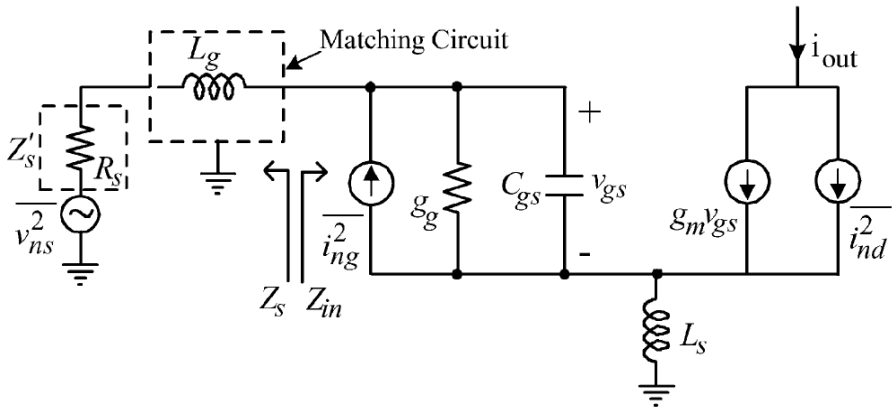

Figure 4.3: Basic matching technique used to combine optimum noise and impedance matching conditions [16].

inductive in nature. As a result, since simple $Z_s = 50\ \Omega$ sources are typically used in practice, the subtraction of $j\omega L_s$ in equation 4.6 helps to reduce the imaginary part of Z_{opt}^0 and hence the difference between Z_{opt} and Z_s.

This source inductance, along with the addition of the gate inductance, L_g, also has the added effect of changing the input impedance, Z_{in}, to be [16]

$$Z_{in} = j\omega L_g + \frac{L_s g_m}{C_{gs}} + j\omega L_s - \frac{j}{\omega C_{gs}} \tag{4.7}$$

From this equation we notice that the addition of L_s creates a real part to Z_{in}, which helps to decrease the discrepancy between the real parts of Z_{in} and Z_{opt} and the proper selection of L_s can also set the this real component to 50 Ω. In addition, L_s has the added benefit of reducing the imaginary component of Z_{in}, created by C_{gs}. L_g adds an extra degree of freedom and allows for this imaginary component to be eliminated completely.

Unfortunately, while degeneration techniques seem effective in theory, their implementation can be significantly more complicated, especially at frequencies above 20

GHz. This is due to the fact that the feedback path through C_{gd}, which was ignored in the above analysis, means that matching networks and additional stages connected to the output of the transistor will also have a significant effect on Z_{in}. Although this reverse isolation is improved by the use of the cascode structure, simulation results reveal that assuming the cascode stage to be unilateral will lead to a significant amount of mismatch when the input and output matching networks are designed.

Further complicating the use of degeneration inductors is the fact that, as discussed in Chapter 2, the low Q factor of on-chip inductors means that these elements will have significant series resistances associated with them. This results in an inherent increase in the noise figure of the device, which can be quantified using the following equation from [19]

$$NF = 1 + \left[\gamma g_{d0} R_s \left(\frac{\omega_o}{\omega_T}\right)^2 \left(1 + \frac{\delta\alpha^2}{5\gamma}\right) + 2\,|c| \left(\frac{\omega_o}{\omega_T}\right) \sqrt{\frac{\delta\gamma}{5}} + \frac{\delta}{5 g_{d0} R_s}\right] + \frac{R_{L_g} + R_{L_s}}{R_s} \tag{4.8}$$

where the term of interest is the third term on the right hand side, which shows an increase in NF for any value of gate inductor resistance R_{L_g} or source inductor resistance R_{L_s} even though this increase is lessened by the source resistance R_s.

As a result, it is clear that the use of series inductors in the gate or source regions should be avoided wherever possible in order to minimize NF. This can be accomplished by resorting to the CNM technique and although this method results in a reduced gain, it is possible to compensate for this with the addition of a second and, if needed, third amplification stage. As long as the gain of the first amplifier stage is sufficient, the addition of these stages will have a smaller effect on the overall NF than the use of series inductors in the matching networks.

4.5 Simulation Results

The modeling inaccuracies and non-ideal transistor behaviour discussed earlier, such as non-ideal reverse isolation and parasitic substrate coupling, highlight the extremely complex nature of high-frequency circuit design. As a result, accurately predicting circuit behaviour using hand calculations becomes an extremely difficult undertaking. Instead, the use of circuit simulation programs, which are becoming increasingly sophisticated and accurate, has become not only commonplace but a necessary design step before fabrication can occur.

In this work circuit simulation was used not only for verification and fine tuning of the circuit, but also to help select the lumped elements to be used in the matching networks. First, though it could not be eliminated all together, the use of series inductors was avoided wherever possible in the matching networks for the reasons discussed earlier. Then, with this in mind, the optimization tools in the Agilent ADS circuit simulator were employed to choose lumped element component values that would minimize the noise, maximize gain, and produce suitably low input and output return losses. The final circuit layout resulting from this optimization is displayed in Figure 4.4.

The optimized matching network components are highlighted by the dashed boxes in Figure 4.4 and their final values are given in Table 4.3. Of these matching network components, only L_{IN2} was implemented with a spiral inductor to simplify layout geometry, while the rest were implemented using SSTL inductors. Capacitors and inductors shown in the circuit schematic that are not enclosed by the matching network boxes are large valued lumped components, which are in place to maintain proper bias conditions without greatly affecting matching.

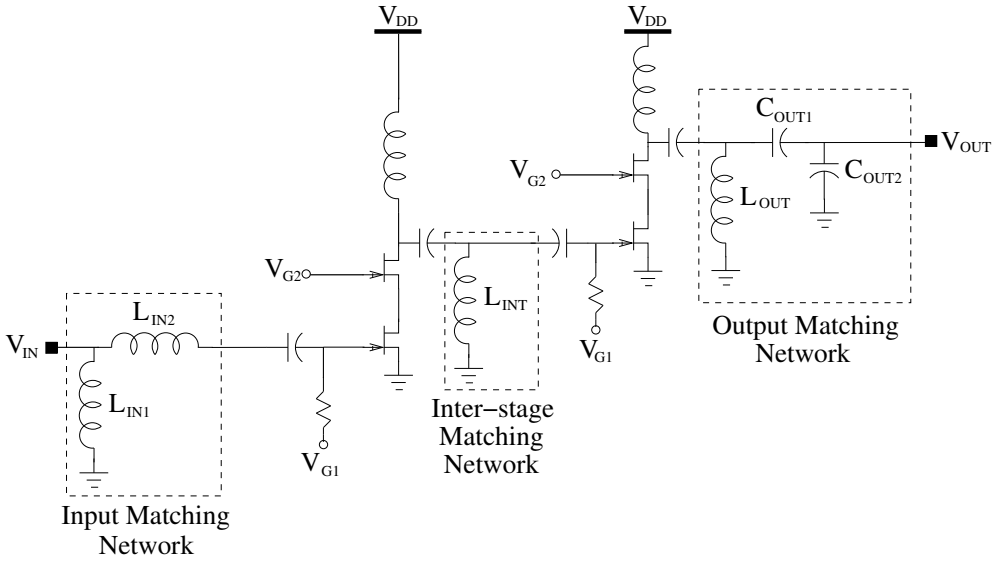

Figure 4.4: Final LNA circuit schematic produced using ADS optimization.

Component	L_{IN1}	L_{IN2}	L_{INT}	L_{OUT}	C_{OUT1}	C_{OUT2}
Value	290 pH	130 pH	220 pH	290 pH	186 fF	328 fF

Table 4.3: Matching network component values obtained using ADS optimization.

With the circuit design optimized for minimum noise and good input and output return loss, simulations show promising results for the cascode LNA. Even after replacing the inductors in the schematic with electromagnetic (EM) simulations of the individual spiral or SSTL inductors to accurately model the inductive behaviour and losses in these components, S-parameter simulations show excellent gain, reverse isolation, and input and output matching. These results are displayed in Figure 4.5.

Highlights of these results include a peak gain of over 17.5 dB and S_{11} and S_{22} values that are suitably low at the peak gain frequency and remain below zero at all frequencies. To confirm that the amplifier remains stable over these frequencies, Figure 4.6 displays the source and load stability factors μ and μ' (defined in equation 4.9 in the following section) of the amplifier. Although Figure 4.6 is used to highlight

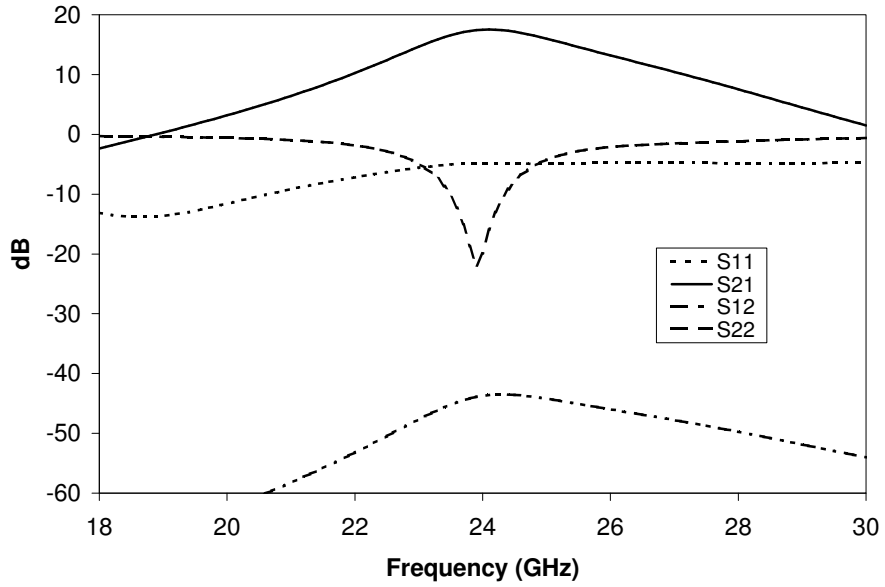

Figure 4.5: Simulated S-parameters of the complete circuit using EM simulation data for spiral and SSTL inductors.

frequencies between 18 and 30 GHz, it is important to note that both of these factors remain greater than 1 for all frequencies from DC to f_{max}, meaning the transistor is unconditionally stable. It should also be noted that although S_{11} achieves its minimum value of approximately -13 dB at a frequency of 18.7 GHz, well away from the peak gain frequency, this is not an accidental mismatch. Instead, as discussed earlier, the the LNA has been designed for the lowest possible noise figure meaning that the use of source degeneration to achieve simultaneous noise and impedance matching has been avoided and a slight impedance mismatch at the input has been implemented, sacrificing some gain for a lower noise figure.

Simulated results of the noise figure are shown in Figure 4.7. These results show the NF to be close to NF_{min} and below 5 dB at the 24 GHz operating frequency, placing it amongst the best reported results in recent publications. Also worth noting

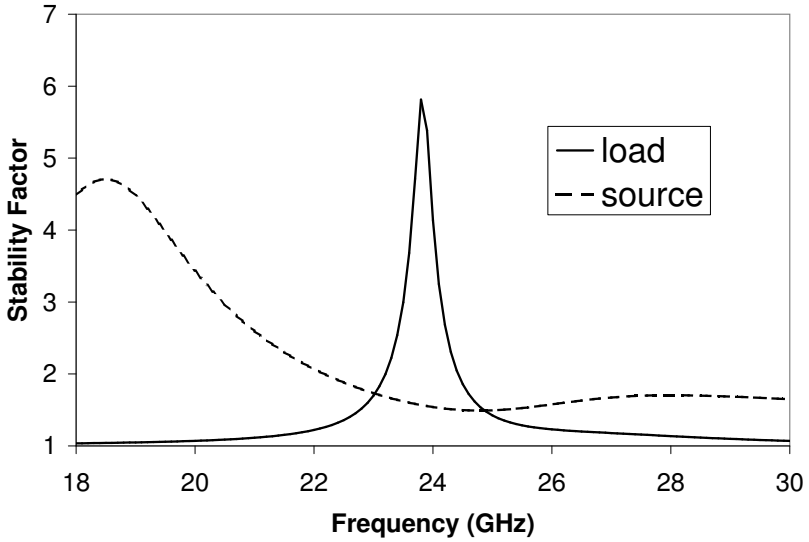

Figure 4.6: Simulated source and load stability factors show unconditional stability.

is the fact that the entire circuit consumes 58.1 mW (which is comparable to other published LNAs using a cascode topology) when the applied gate voltage is 0.8 V and that simulations show that increasing this bias voltage to the optimum 0.9 V obtained earlier increases power consumption to 85.5 mW but reduces NF by approximately 0.2 dB.

An additional degree of confidence in circuit performance can be gained by performing a full EM simulation of all passive elements in the circuit. Although the transistors can not be included in this simulation, the RF models provided by the Taiwan Semiconductor Manufacturing Company (TSMC) can be inserted into the entire passive network after S-parameter simulation has been completed, as illustrated in Figure 4.8.

This type of analysis simulates the coupling between on-chip components and is the most accurate means of predicting RF CMOS circuit behaviour. Unfortunately, it is extremely computationally intensive and while it could not be completed before

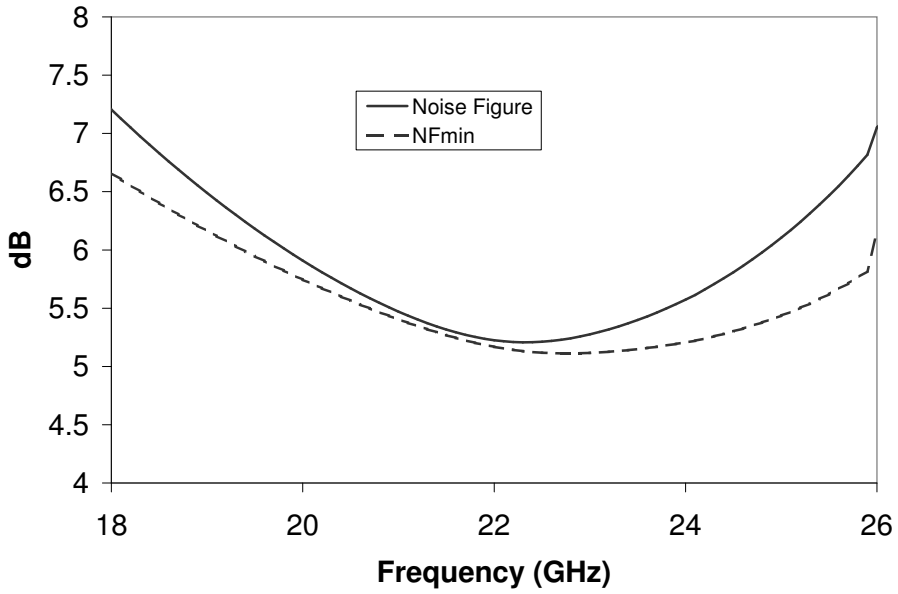

Figure 4.7: Simulated noise performance using EM simulation data for spiral and SSTL inductors.

submission for fabrication, it was later accomplished. EM simulated S-parameter and noise figure results are displayed in Figure 4.9 and Figure 4.10 respectively.

It is apparent that the S-parameter results of the full circuit EM simulation differ slightly from those obtained from the initial circuit schematic simulations described earlier. First, the operating frequency and the optimum input and output return loss frequencies have been shifted down by approximately 2 GHz. This is likely due to additional parasitic capacitances that were not included in the original simulations and to interactions between adjacent inductors in the matching and bias networks. Secondly, while the minimum noise figure value is relatively unchanged, it also shows the same frequency shift as the S-parameters and the noise performance of the device shows a sharper increase in both NF and NF_{min} as the frequency moves from the optimal operating condition. Despite these changes, the LNA still shows a good gain

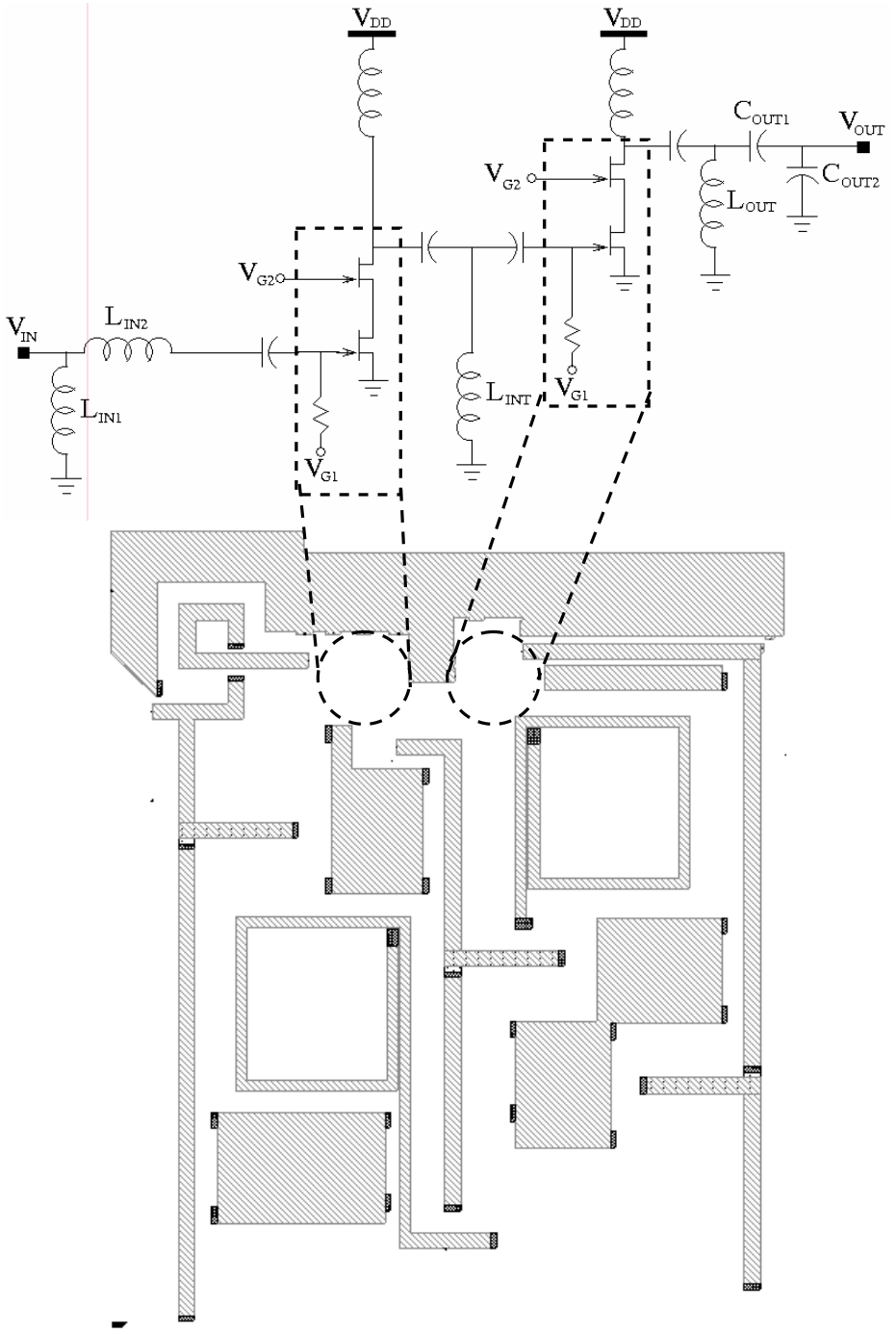

Figure 4.8: Insertion of RF transistor models into the full passive EM simulation.

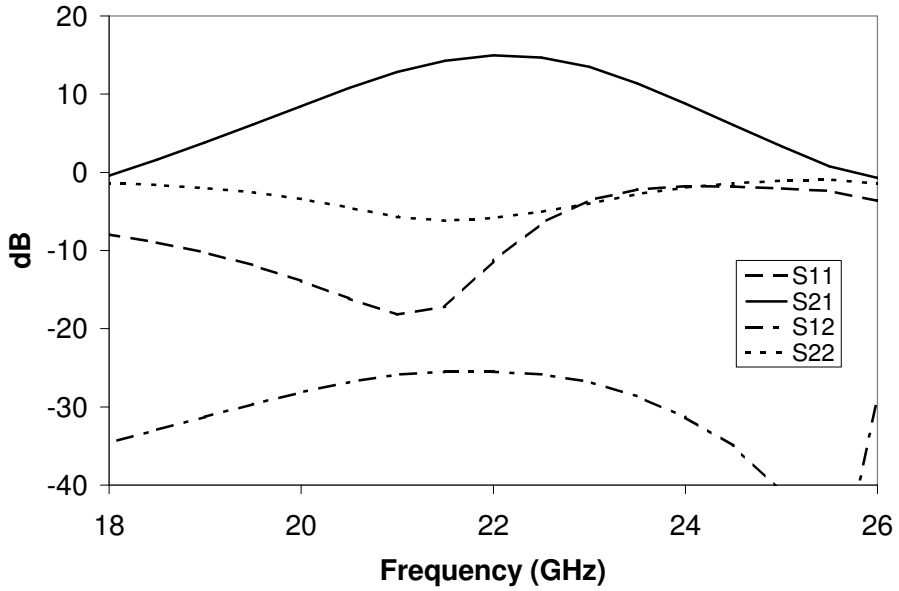

Figure 4.9: Full circuit EM S-parameter simulation results.

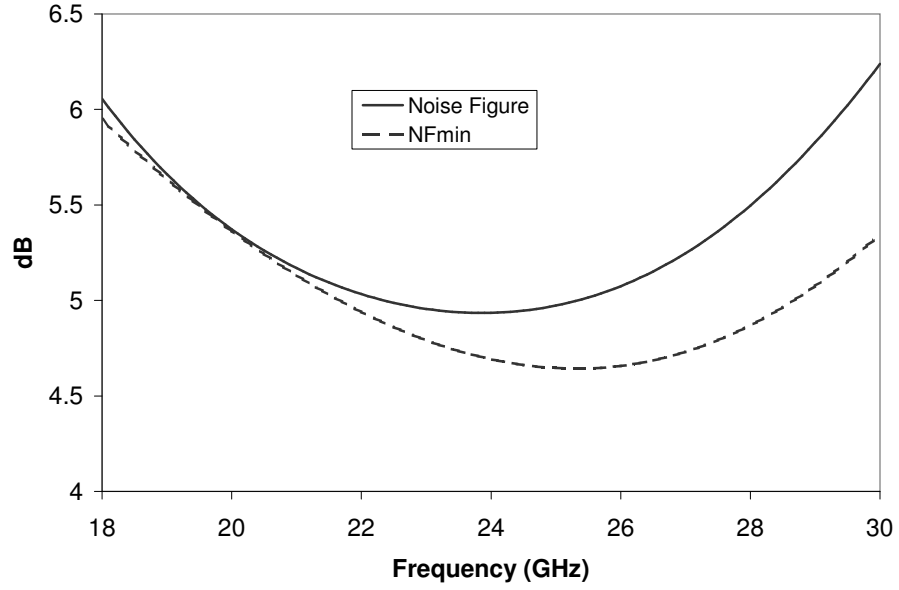

Figure 4.10: Noise figure results obtained from a full circuit EM simulation.

of almost 15 dB, acceptable input and output return loss, and a minimum noise figure of approximately 5.2 dB, supporting the earlier assertion that this design should rank amongst the best reported in other publications.

4.6 Measured Results

As with any integrated circuit design, the best way to demonstrate its functionality is with accurate measured results. As such, a prototype was fabricated in 0.18 μm CMOS with a thick top metal layer and a photograph of this chip is shown in Figure 4.11. Unfortunately, obtaining measured results is rarely a trivial matter, especially at very high frequencies in CMOS devices. This section outlines the difficulties encountered during measurement as well as the final results obtained and compares these results to simulation.

Stability is a major concern in any amplifier design since instability can create oscillation, making the amplifier useless. As such, an analysis of the amplifier stability is a wise course of action before fabricating any amplifier design. This can be done by checking to see if the amplifier is unconditionally stable for any source or load reflection coefficients, Γ_S or Γ_L respectively. For any two port device, this stability check can be performed by ensuring that the following equation for the stability factor, μ, is greater than 1 [48]:

$$\mu = \frac{1 - |S_{11}|^2}{|S_{22} - \Delta S_{11}^*| + |S_{12}S_{21}|} \tag{4.9}$$

where

$$\Delta = S_{11}S_{22} - S_{12}S_{21} \tag{4.10}$$

Fortunately, the losses from the low-Q passive structures employed in the amplifier

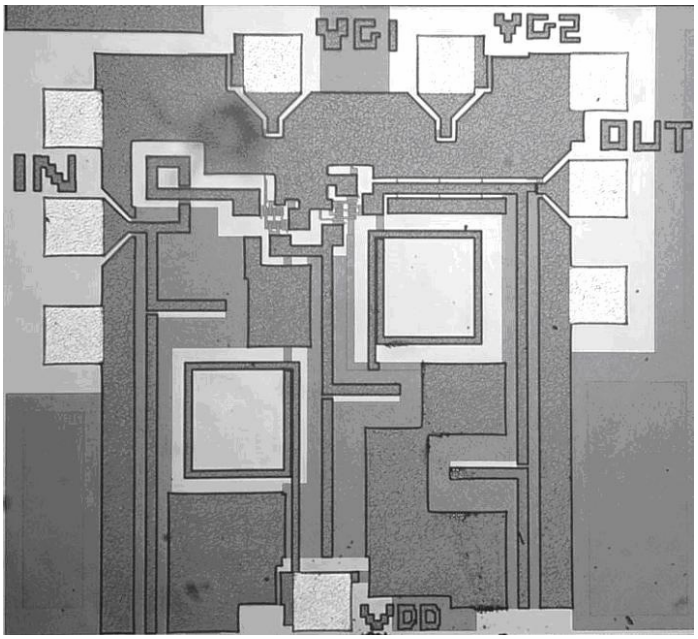

Figure 4.11: Die photograph of the fabricated single-ended LNA.

matching networks, as described in Chapter 2, coupled with the use of the cascode structure to improve reverse isolation should mean that stability is not a concern in the amplifier design presented in this chapter. This was verified by ADS simulations, displayed in the previous section, which show the amplifier to be unconditionally stable over all frequencies. The problem, however, is complicated by the fact that the operating conditions of the amplifier can be significantly altered during physical testing due to the parasitics of the cables used to supply DC power to the amplifier.

In an effort to mitigate these effects on-chip, 10 kΩ resistors were placed in the DC supply lines for all gate voltages (since no current is drawn at these points) and 3 pF decoupling capacitors were placed at each DC supply pad, including V_{DD}. Unfortunately, spectrum analysis measurements proved that these steps were not effective since low-frequency oscillations at approximately 16 MHz were created by the

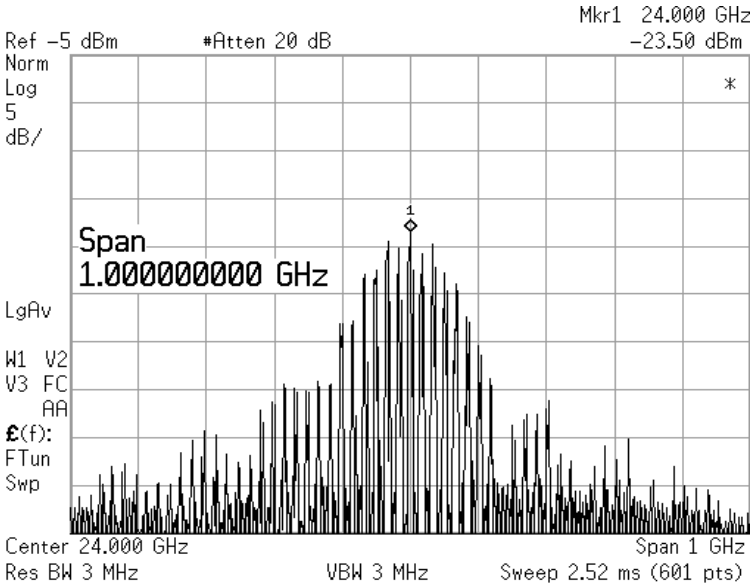

Figure 4.12: Spectrum analysis of the LNA output signal when low frequency oscillation is present.

parasitics on the V_{DD} supply line. These oscillations were subsequently up-converted to the operating band of the amplifier as is clear in the spectrum displayed in Figure 4.12, which is generated by the LNA when a 24 GHz signal with a strength of -10 dBm is applied to the amplifier.

After observing this behaviour, a more in-depth analysis of the low-frequency stability of the LNA was performed with ADS. Although, it is true that both the source and load stability factor remains greater than 1 at all frequencies, Figure 4.13 shows that the load stability factor gets very close to 1 at frequencies below 1 GHz. The measured results show that this factor, combined with the parasitics and noise generated by the DC connection to V_{DD} were enough to cause the amplifier to oscillate.

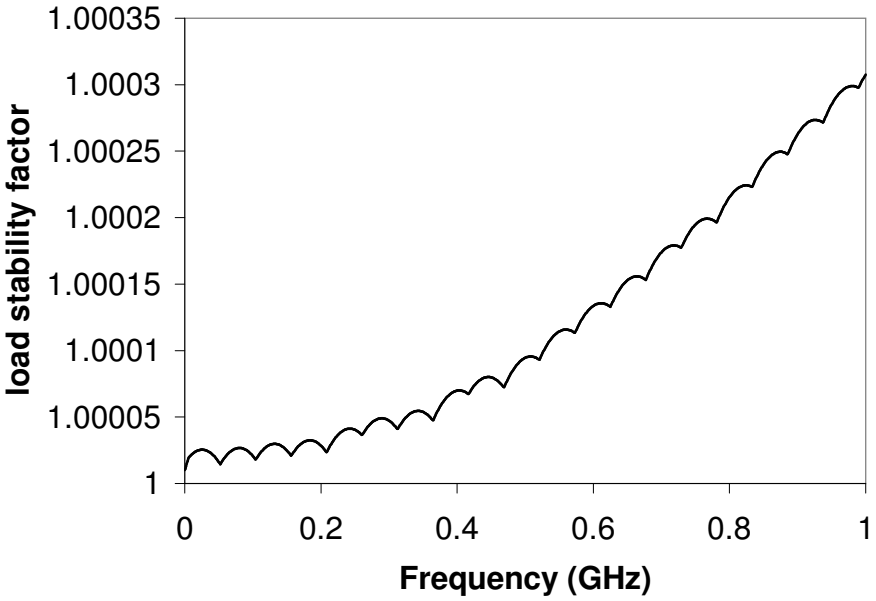

Figure 4.13: Low-frequency analysis of the load stability factor.

These oscillations have the detrimental effect of not only making the amplifier unusable but also, if the voltages incurred by the oscillation are large enough, damaging the circuit so that even if the oscillations can be removed, functionality will be lost. This problem was encountered several times during measurement when, if the chip was left to oscillate for long periods of time, a DC connection between V_{DD} and both V_{G1} and V_{G2} was created.

This problem was likely caused by a breakdown in the reverse biased well conditions seen at the boundary between the deep n well and p type substrate, as illustrated in Figure 4.14. Designed to have a high voltage difference of 3.6 V across this boundary, the existence of oscillation on the V_{DD} line can increase this difference beyond maximum tolerances, causing the $p - n$ junction to breakdown and allowing for the positive DC voltage in the n well to be transferred to the substrate. This results in DC current being drawn into the substrate and allows for some DC voltage to be

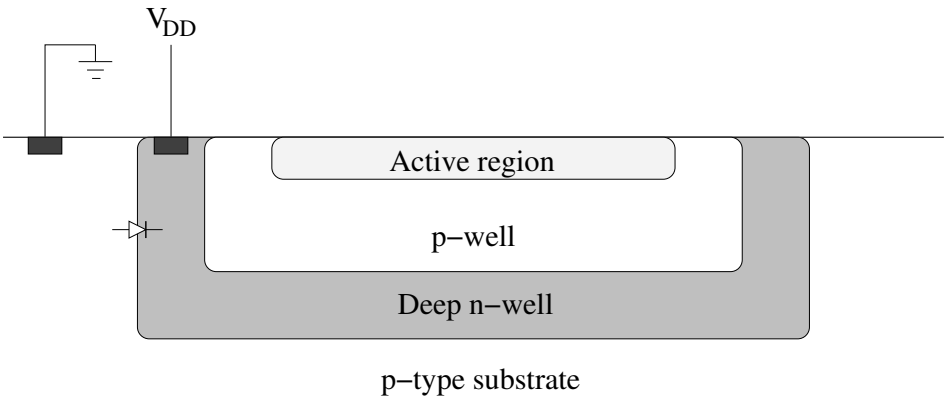

Figure 4.14: Well breakdown caused by oscillation in the amplifier.

transferred to the gates of each transistor via the (normally) reverse biased electrostatic discharge (ESD) protection diodes that were placed between each gate and the substrate in order to protect the circuit from excessive charge buildup on the gate nodes.

Fortunately, with the cause of the oscillations identified, it was possible to prevent further well breakdown on new test chips and eliminate the oscillations by removing the parasitics and noise generated on the V_{DD} supply line with the use of a combination of ferrite beads and simple RC filtering. A detailed description of this setup is given in Appendix B. With this oscillation prevention system in place, spectrum analysis measurements were again completed and these new results are displayed in Figure 4.15. With the oscillations removed and a -10 dBm signal applied, the amplified signal is free of any sideband components and the output power has increased significantly, showing enough positive gain in the LNA to compensate for approximately 9 dB of loss encountered in the cables and bias tees used in the measurement setup.

With the oscillation problem solved, accurate measurement of noise figure (NF) and gain of the LNA was possible. These measurements were completed on the

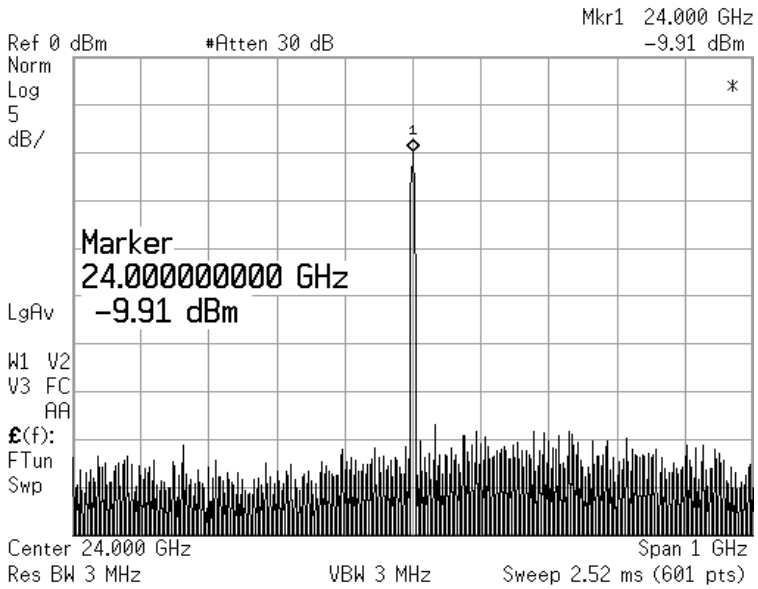

Figure 4.15: Spectrum analysis of the LNA output signal when oscillation has been eliminated.

spectrum analyzer with a noise source connected to the input of the LNA, and results are displayed in Figure 4.16. With the bias conditions set to $V_{DD} = 3.6$ V, $V_{G2} = 2.6$ V and $V_{G1} = 0.8$ V, resulting in a bias current of 15.4 mA, the measured NF of the amplifier reaches a minimum of 5.9 dB at approximately 23 GHz. The measured curve appears to follow the simulated trendline and while the NF is approximately 0.7 dB higher than simulations using ADS transistor models predict, this discrepancy is consistent with that reported in [46].

Somewhat surprising is the fact that the measured gain achieves its maximum value of 13.5 dB at a frequency of just over 23 GHz, approximately 1 GHz higher than that predicted by simulations of the entire LNA. This is likely caused by inaccuracies in the simulation due to two factors:

1. Simplifications in the layout (such as removal of the metal-insulator-metal (MIM)

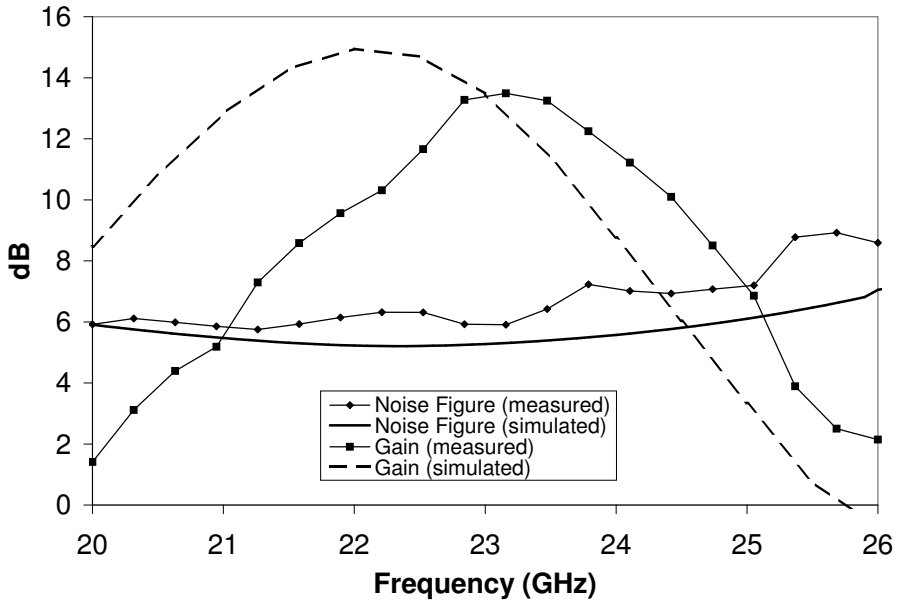

Figure 4.16: Noise figure and gain comparison between EM simulation and measurements taken using a spectrum analyzer.

capacitors) and substrate definition (such as ignoring the thickness of the metal 1 ground plane regions) used in the EM simulation were made in order to reduce the computing resources required.

2. Small inaccuracies in the high frequency transistor model used are amplified when two transistors are stacked in the cascode configuration. This problem is explored further in the following section.

This frequency shift between simulated and measured results was also verified by performing S-parameter measurements on the vector network analyzer. Results of these measurements are shown in Figure 4.17 and display a similar frequency shift, not only in gain but in input and output return loss as well, to that reported by the spectrum analyzer.

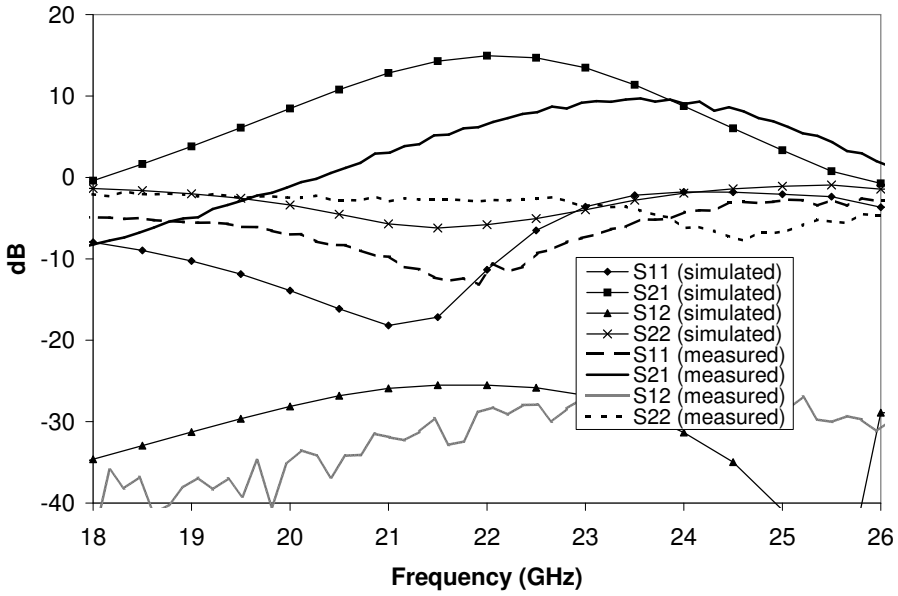

Figure 4.17: Comparison of measured and EM simulation S-parameter data.

Finally, measurements of the output power capabilities and linearity of the amplifier, which are important considerations in any practical application for the LNA, were performed with the spectrum analyzer. Figure 4.18 displays the results of these measurements taken at a frequency of 23 GHz. They highlight a good input referred 1 dB compression point of -7 dBm and a high IIP3 of +3 dBm, both of which should be high enough to be suitable for implementation in the front end of a CMOS receiver where operating power levels are typically significantly lower than this.

In order to give some context to the measured results reported in this section, Table 4.4 compares this work with other recent publications. This highlights the impressive performance of the cascode LNA presented in this chapter and the results have subsequently been published in [28].

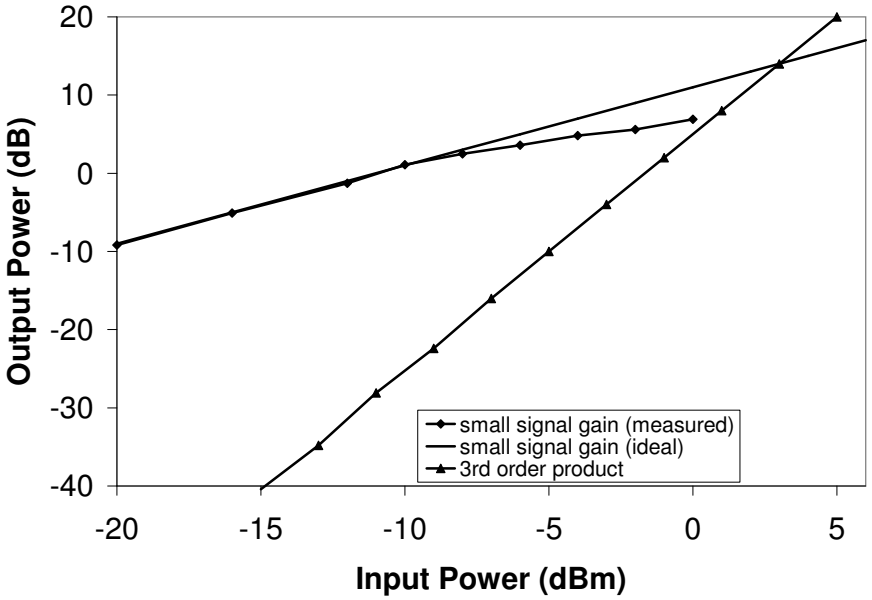

Figure 4.18: Output power and linearity measurements.

Ref.	Technology	f_T (GHz)	f_{op} (GHz)	S_{21} (dB)	NF (dB)	P_{1dB} (dBm)	$IIP3$ (dBm)	P_{DC} (mW)
[21]*	90 nm SOI	149	35	15	3.6	4	N/A	40.8
[20]*	0.12 μm	N/A	20.3	18.9	4.26	-16	N/A	57
[42]	0.18 μm	54	21.8	15	6	-8	N/A	24
[17]	0.18 μm	70	24	13.1	3.9	-12.2	0.54	14
[18]	0.18 μm	45	23.7	12.9	5.6	-11.1	2.04	54
This work*	0.18 μm	45	23.2	13.5	5.9	-7	3	55.4

* Designs using a cascode topology.

Table 4.4: Comparison of measured results in this work and other published results.

4.7 Future Work

The parasitics associated with a CMOS transistor can have a significant effect on the performance of the MOSFET, especially at high frequencies. Therefore, in order to facilitate accurate circuit design at these frequencies, it is necessary to create an accurate representation of these effects. The RF model supplied by TSMC for 0.18 μm CMOS accounts for these high-frequency effects in ADS simulations. A comparison of measured and simulated S_{11} and S_{22} from 0 to 30 GHz for a 30 finger cascode test structure with the body terminals grounded is illustrated in Figure 4.19. These results show that although this model provides a good representation for the cascode structure for lower frequencies, the curves diverge significantly at frequencies above 10 GHz. In order to provide the best possible comparison to the measurements discussed in the previous section, simulation results presented in this section use an EM simulation of the test structure and compare this data to the measured data without performing any deembedding.

To further highlight this discrepancy, Figure 4.20 displays the magnitudes of the measured and simulated S-parameters of a 30 finger cascode structure. From these results it is apparent that if EM simulations are to be produced to accurately predict measured performance of a cascode structure, further work on the high-frequency transistor model must be completed.

4.8 Conclusions

A two-stage, narrowband, single-ended LNA has been presented for use at 23 GHz. This amplifier is optimized for lowest noise in terms of transistor size, bias conditions

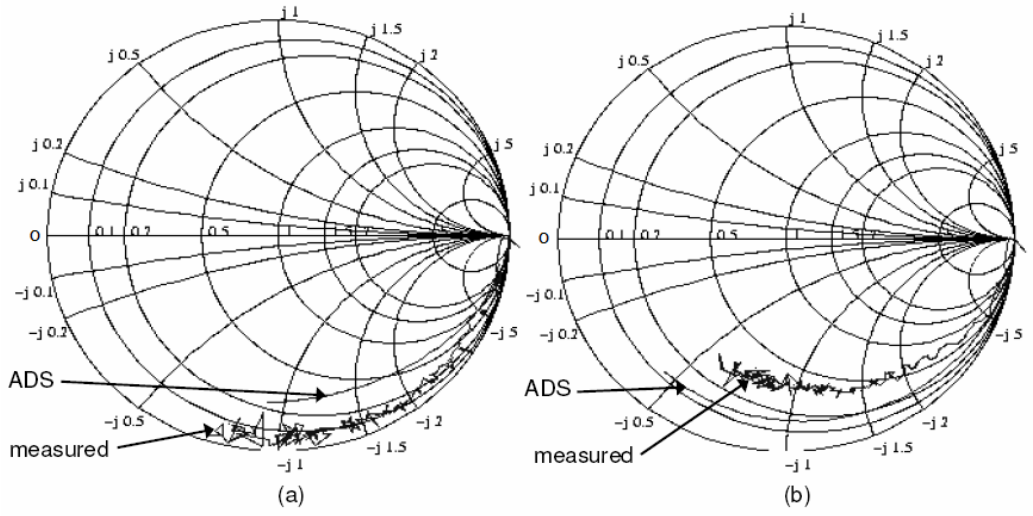

Figure 4.19: Comparison of measured and simulated (a) S_{22} and (b) S_{11} of a 30 finger cascode structure.

and matching networks and uses a cascode structure for both stages in an effort to improve reverse isolation. Although the cascode configuration suffers from a small, inherent penalty to the NF, the use of high-Q, easily modeled SSTL inductor structures in combination with the simplified design created by the improved reverse isolation mean that it is possible to create a cascode amplifier with gain and noise performance ranking amongst the best presented in current literature.

The final circuit was fabricated in 0.18 μm CMOS and, after eliminating oscillations caused by parasitics on the drain voltage supply line, measured results display a high gain of approximately 13.5 dB and good noise performance with an NF of approximately 5.9 dB at 23 GHz. The uncertainty in this measured result was calculated using [49] to be approximately $\pm$0.5 dB. These results fall within the range of other reported results for single-ended LNAs, but do not match the noise performance of the best published amplifiers in the same technology.

Comparison of simulated and measured results shows a discrepancy between the

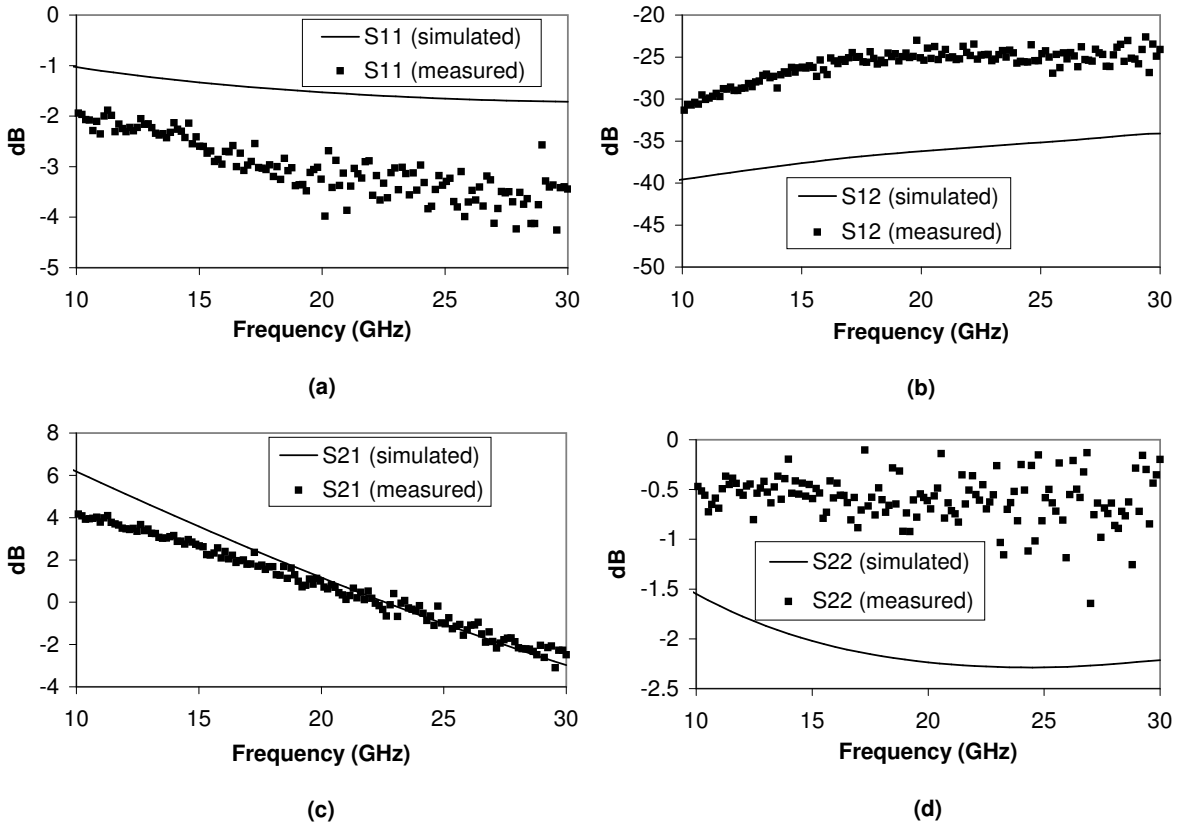

Figure 4.20: Comparison of measured and simulated S-parameter magnitudes of 30 finger a cascode structure.

two even in accurate, full-circuit EM simulation. This highlights inaccuracies in the high-frequency transistor models provided by TSMC, especially when these models are used in a cascode configuration. Further work done in an effort to improve the transistor modeling could help to fine-tune matching and improve the gain and noise performance of the LNA to allow it to surpass the performance of the best previously reported designs.

Chapter 5

Differential LNA Design

5.1 Introduction

The cascode LNA presented in the previous chapter presents a viable option for single-ended, noise sensitive amplification at 23 GHz. Differential LNAs can also be used to provide low-noise, high-frequency amplification and can be extremely useful in circuit design since certain structures, such as some mixers and A/D converters, are naturally balanced [50]. They also have the added advantage of rejecting common-mode disturbances, which can be particularly attractive for mixed-signal applications where both the supply and substrate voltages may contain significant amounts of noise [44].

This chapter presents the design and fabrication of a 2-stage, differential LNA. Although designed to operate at a frequency of 25 GHz, measured s-parameter results are presented showing peak noise and gain performance at 23 GHz. Although differential LNA design has received far less attention in recent research and extremely few published results are currently available for frequencies above 20 GHz, the simulated

and measured results presented in this chapter indicate that the differential LNA can perform at least as well as its single-ended counterparts.

While the cascode topology presented previously can also be implemented in a differential configuration, the presence of two signal paths that are 180^0 out of phase in differential circuit design presents another attractive option for reducing the effects of C_{gd} and improving reverse isolation. This technique, which is often referred to as neutralization, uses the varactor models presented in Chapter 3 to implement cross-coupled varactors, which are used to cancel signal flow through C_{gd}.

5.2 Neutralization

As silicon CMOS circuitry continues to scale downwards in size, leading to an increase in operating frequency, several obstacles in RF circuit design that could once be neglected are beginning to become more problematic. One such problem is the fact that as technology sizes decrease, the gate-drain capacitance, C_{gd}, of MOSFETs becomes comparable in magnitude to the gate-source capacitance, C_{gs}, meaning that feedback via C_{gd} can no longer be neglected in circuit analysis [51].

As presented in the previous chapter, one method commonly used to reduce this feedback in CMOS amplifiers has been the use of a common gate transistor in the output path in a cascode configuration. While this work has shown that the cascode topology is a viable option for LNAs operating above 20 GHz, it does have some shortcomings that can be significant depending on the intended application, as mentioned in the previous chapter:

1. A two-transistor stack is not optimal for operation at the lowest possible supply voltage, making it unsuitable for low voltage applications [51].

2. The output common gate transistor typically adds approximately 0.5 dB of noise to the amplifier, meaning that although it can achieve good noise performance, this performance can be improved if this stacked transistor can be removed.

Not surprisingly, recent research is exploring alternative methods of reducing the feedback introduced by C_{gd}. Specifically, the concepts of neutralization and unilateralization, first introduced in 1955 [52], are finding renewed application in modern technology. Neutralization is broadly defined in [52] as the process of balancing out an undesirable effect, while unilateralization refers to a specific type of neutralization that converts a bilateral four-terminal network to a unilateral one. This chapter uses extensive ADS simulation to examine the application of a capacitive neutralization scheme, proposed in [44], to improve the performance of a two-stage differential amplifier. Results show the neutralization scheme able to offer performance competitive to that of cascode topology while reducing supply voltage and even noise figure [51].

5.2.1 Theory

The gate-drain capacitance, C_{gd}, is an unavoidable parasitic in standard CMOS technologies largely caused by lateral diffusion of the drain dopant under the polysilicon gate material [51]. This capacitance provides a non-inverting feedback path, which degrades circuit performance by reducing amplifier forward gain, reverse isolation and device cutoff frequency, f_t, as is illustrated by equation 5.1

$$f_t = \frac{g_m}{2\pi(C_{gs} + C_{gd})} \tag{5.1}$$

where g_m is the transistor transconductance.

The gate to drain capacitance also has the added effect of increasing the input capacitance, C_{in}, by the Miller effect as can be seen by equation 5.2 below for a

MOSFET in common source configuration

$$C_{in} = C_{gs} + C_{gd}(1 + A_v) \tag{5.2}$$

where A_v is the forward gain of the transistor.

Neutralization can improve circuit performance by canceling signal flow through C_{gd}. This is accomplished by adding secondary signal paths around the amplifier so that the net signal flow through C_{gd} and the additional signal path is zero [51]. This addresses the issues of reduced forward gain and reverse isolation but not necessarily the issue of increased input capacitance.

Three common neutralization topologies are shown in Figure 5.1 below, each of them having their respective difficulties that must be addressed to make implementation feasible. Figure 5.1(b) uses an inductor L to resonate with C_{gd}, effectively canceling its effect. However, this technique is of limited usefulness due to the large inductance values required at high operating frequencies. Other obstacles associated with this topology include the fact that the low Q factor of a monolithic inductor will increase the amplifier noise figure as well as the fact that the large DC block capacitor, C_{BIG}, will increase the input capacitance, limiting the switching speed of the transistor [51].

Figure 5.1(c) uses a transformer to induce magnetic coupling between drain and source, feeding back a portion of the output signal, which can effectively cancel the feedback from output to input. This topology encounters a similar difficulty to that of figure 5.1(b) in that the monolithic transformer used has low Q at frequencies above 20 GHz in standard CMOS, degrading the noise performance of the amplifier. Also complicating matters is the fact that C_{gd} is difficult to accurately predict at frequencies above 20 GHz, being susceptible to process variations and changing with applied

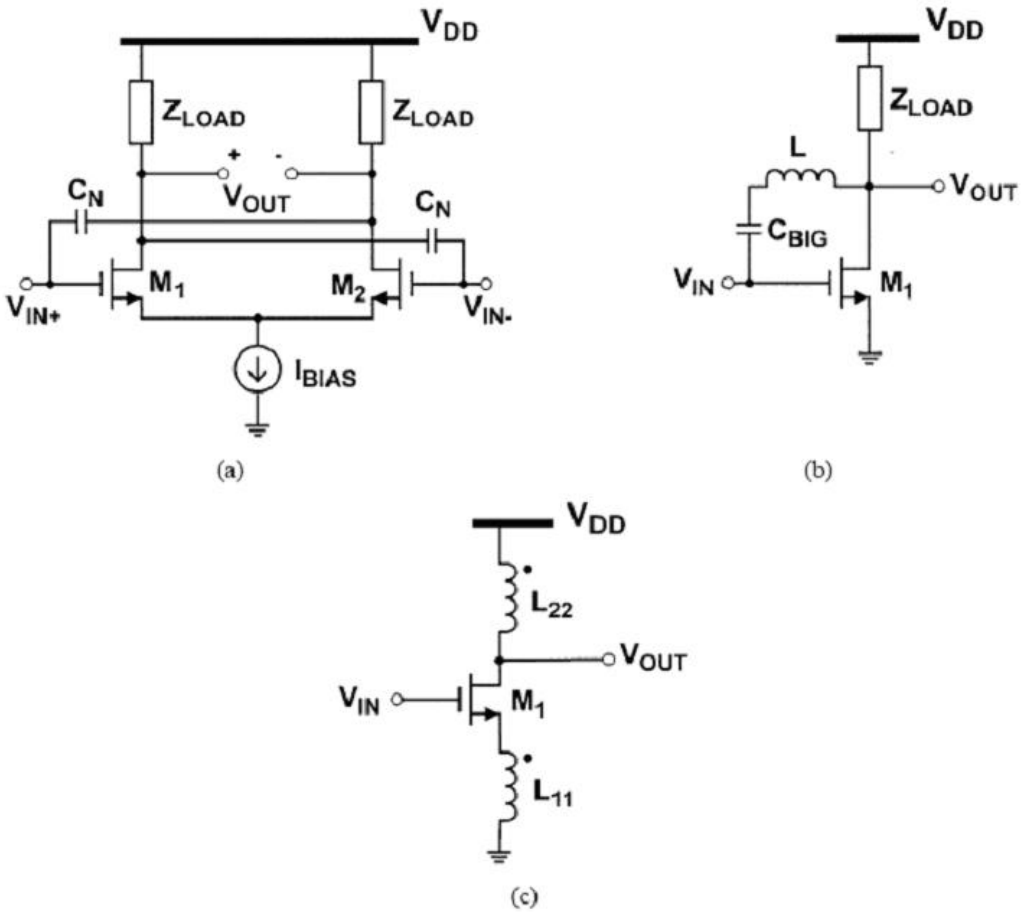

Figure 5.1: Common neutralization topologies using (a) neutralizing capacitors, (b) a resonating inductor and (c) transformer feedback [51].

bias voltage. As a result, it becomes exceedingly difficult to design a transformer with the appropriate turns ratio and self inductance, which in turn leads to a shift in operating frequency or even instability [46].

Figure 5.1(a) uses neutralizing capacitors C_N to cancel signal flow through C_{gd}. Since the drain voltages of the differential pair are 180^0 out of phase, the current through C_N will therefore be equal in magnitude and opposite in phase to the current flowing through C_{gd}. This relies on the precise matching of capacitances C_N and C_{gd}, which can be difficult due to the difficulties in accurately predicting C_{gd} as discussed above. Furthermore, signal flow through C_N is actually positive feedback, which can cause instability if C_N is inaccurately matched to C_{gd} [51]. However, unlike inductors and transformers, the design of variable capacitors using MOSFET devices is much more feasible. This makes it possible for the neutralization capacitances to be tuned to appropriate capacitance values post-fabrication, compensating for any transistor mismatch or process variations in C_{gd}. These factors indicate that the method of using neutralizing capacitors to compensate for C_{gd} is the most promising topology from Figure 5.1 for CMOS devices operating above 20 GHz and this topology was therefore chosen for analysis and fabrication.

5.2.2 Simulation

In order to validate the theory described above and quantify the improvements in amplifier performance made possible by neutralization, basic circuit simulations were performed in ADS. After selecting a transistor width of 80 μm (32 fingers x 2.5 μm per finger) for minimum noise, as was described in the previous chapter, simulations were performed to analyze the effect of adding neutralizing capacitors to a simple

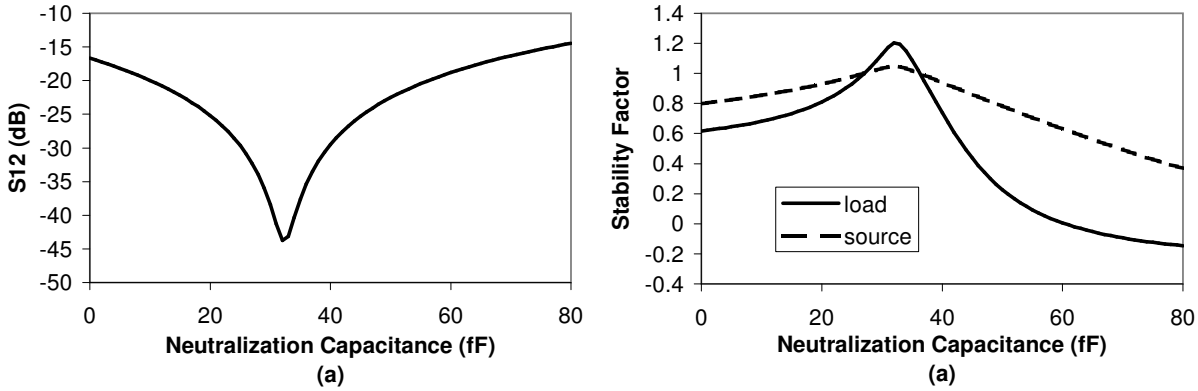

Figure 5.2: Improved (a) reverse isolation and (b) stability achieved by using neutralizing capacitances in a differential pair.

differential pair, before the addition of matching networks.

Figure 5.2 displays simulation results conducted at 25 GHz showing (a) the significant improvement in reverse isolation and the resulting increase in (b) stability achieved by placing $C_N = 32$ fF neutralizing capacitors in the differential pair. This figure also illustrates the sensitive nature of this neutralization. The sharp increase in S_{12} as C_N deviates from 32 fF in either direction, and the fact that the amplifier remains unconditionally stable for only a small range of values of C_N between 28 and 36 fF, indicate that any process variations in C_{gd} or even parasitic capacitances from the interconnects used in the transistor layout could cause instability in the amplifier. This supports the use of varactors to implement C_N to help reduce this possibility. It should be noted that although this stability analysis was conducted at a single frequency of 25 GHz and therefore does not indicate unconditional amplifier stability, simulation results indicating that unconditional stability is indeed achieved for all frequencies are presented later in the chapter.

This neutralization technique also has secondary effects on other amplifier parameters. The maximum available gain (MAG) of an unconditionally stable amplifier is

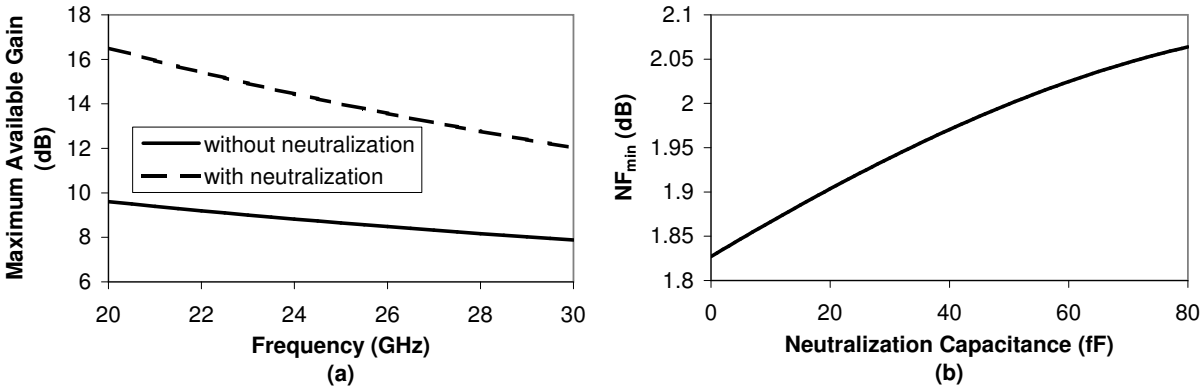

Figure 5.3: Effect of neutralization capacitors on (a) maximum available gain and (b) NF_{min}.

a function of S_{21} and S_{12} as follows

$$MAG \propto \frac{|S_{21}|}{|S_{12}|} \tag{5.3}$$

Since we have seen that the addition of an appropriately sized C_N can significantly decrease S_{12}, it stands to reason that this should have the secondary effect of increasing the MAG. Figure 5.3 (a) confirms this for frequencies between 20 and 30 GHz. Figure 5.3 (b) shows that the tradeoff for all of the advantages seen thus far is a monotonic increase in the minimum noise figure, NF_{min}, that is attainable at 25 GHz when using neutralizing capacitances. This increase is relatively small, however, and is outweighed by the advantages achieved by the addition of C_N in most applications.

5.3 Circuit Design

Using the same transistor size and bias conditions to produce optimal noise performance presented in the previous chapter, a 2-stage, differential low noise amplifier

was designed using the neutralization scheme discussed above. As with the single-ended LNA, this differential amplifier was designed using the classic noise matching technique for best noise, while sacrificing some gain. Also, similar to the LNA from the previous chapter, attempts were made to avoid the use of inductive degeneration in the source of the transistors used in this design.

The design method of the differential LNA presented in this chapter, however, differs significantly in the approach taken to the optimization of the matching networks. Instead of treating the bias networks as large valued inductive and capacitive elements designed only to provide proper bias conditions to the transistors in the circuit, their usefulness as part of the matching networks was identified and optimized. This idea is illustrated with the help of Figure 5.4, which shows the method used to supply a DC bias to the drains of the differential circuit. Here capacitors C_{IN} and C_{OUT} are used to isolate the bias voltages from the input and output and C_{SHUNT} serves the dual purpose of isolating noise from the DC voltage source and providing an RF ground from a matching perspective.

From this picture it is clear that if the drain bias inductor, L_{BIAS}, is properly tuned, it can serve as a shunt inductance in the output matching network. The same is true for the blocking capacitance C_{OUT}, which can provide the desired series capacitance to properly match the output of the differential circuit. This is extremely beneficial to the circuit designer due to the fact that it reduces the number of lumped elements required at the output of the amplifier thereby reducing the need for low-Q on-chip inductors, which can be detrimental to the noise performance. This is also beneficial in terms reducing of total fabrication space (and therefore cost) required for the chip, which can be especially important for differential designs since all lumped

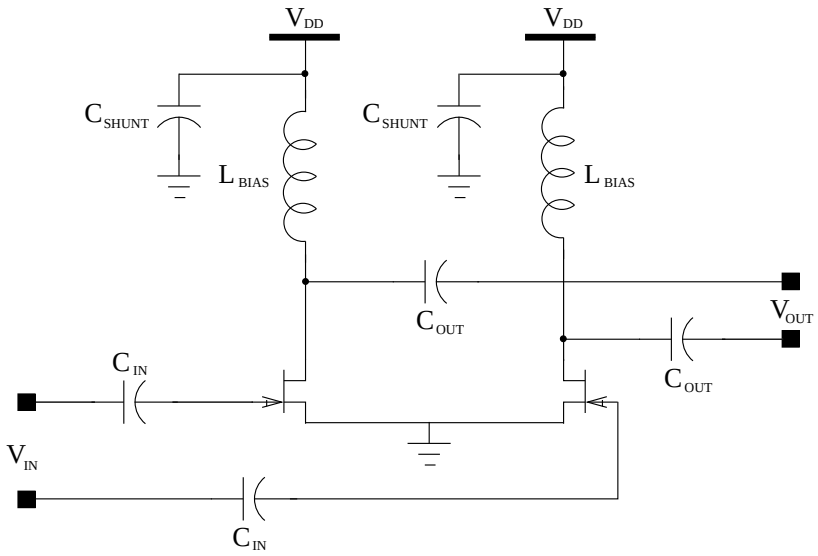

Figure 5.4: Components used to provide DC bias to the amplifier can be used for input and output matching as well.

elements used must be mirrored in each signal path.

Also worth noting is the fact that, although the addition of neutralizing capacitors improves the feedback from drain to gate, tuning the drain bias inductor also has an effect on the impedance seen at the input of the amplifier. This means that the creation of the output and input matching networks can not be considered to be separate entities. This makes the complexity of high-frequency LNA design in CMOS technology apparent, especially when multiple amplifier stages are used, requiring the creation of interstage matching networks as well. As a result, optimization tools within the ADS circuit simulation program were used extensively to help design the matching networks.

Figure 5.5 displays the schematic of the final circuit design including matching and drain bias networks. Since no DC current is required at the transistor gates, bias voltages were applied to these points using 10 kΩ resistors and were omitted from Figure 5.5 for simplicity. Of special interest is the output matching network, which

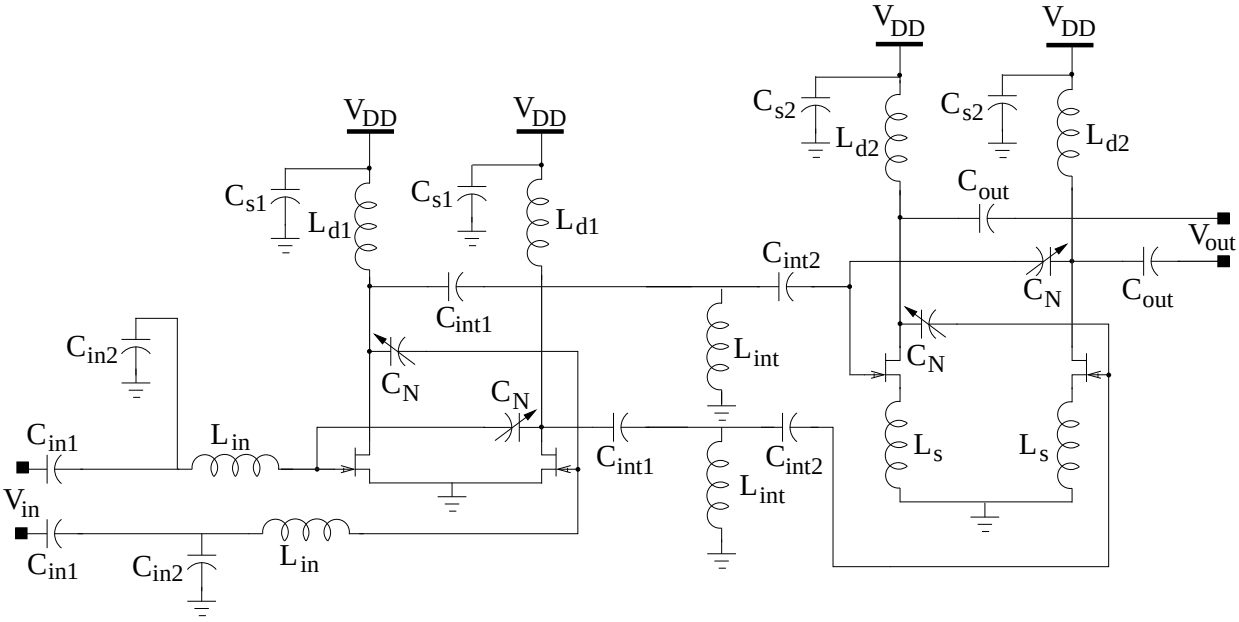

Figure 5.5: Final design of the differential LNA including matching and drain bias networks.

has been significantly simplified by careful tuning of the bias network required to supply V_{DD} to the transistor drains.

A second point of interest in Figure 5.5 is that inductive degeneration was used in the source of the second amplifier stage. This is due to the fact that, although these components do increase the noise figure, this increase is very small since it occurs in the second stage of the amplifier and its effect is therefore reduced by the gain of the first stage. This slight increase in noise figure was determined to be an acceptable tradeoff for the improved interstage matching that it helps to achieve, increasing the gain of the amplifier and placing its peak gain frequency at the desired operating frequency of 25 GHz.

Table 5.1 gives inductance and capacitance values for each of the lumped elements used in the final circuit design. Due to the nature of differential circuit design, which requires that all components be mirrored precisely in the circuit layout in order to

C_{in1}	C_{in2}	L_{in}	L_{d1}	C_{s1}	C_{int1}	L^{*}_{int}	C_{int2}	L^{**}_{s}	L_{d2}	C_{s2}	C_{out}
1.4 pF	180 fF	230 pH	290 pH	3 pF	1.4 pF	200 pH	1.4 pF	65 pH	140 pH	1 pF	125 fF

* Inductors implemented using the SSTL technique.
** Inductors implemented using simple microstrip lines.

Table 5.1: Lumped element component values used in the differential LNA design.

reject common-mode disturbances [44], the use of elongated SSTL structures was often difficult to layout efficiently. As a result, only the relatively small inductance for L_{int} was implemented using an SSTL structure, while the very small inductance required for L_s was implemented with a simple microstirp transmission line. All other inductors were implemented using octagonal spiral designs.

A final point of interest is that the sources of all the transistors in this circuit are tied directly to ground whereas traditional differential pairs often involve the use of current sources in these locations instead. These current sources make the drain currents in the differential pair independent of the common-mode voltage present at the gates of each transistor, thereby helping to stop the amplification of any common-mode signals present on the input line and improving the common-mode rejection ratio ($CMRR$) [47]. The downside to the use of these current sources is that they require an increase in the supply voltage rail to provide proper biasing, which increases the power consumed by the LNA. Since the differential circuit presented here already consumes more power than its single-ended counterpart, the decision was made to avoid the use of current sources, trading a penalty in $CMRR$ for decreased power consumption. The differential circuit presented in this chapter still achieves some advantages over its single ended counterpart since it is able to reject supply noise on V_{dd} and provide higher output voltage swings. In addition, the penalty incurred to the $CMRR$ should be small if the intended application has a well-defined common-mode DC level that willl not experience large disturbances [47].

5.4 Varactor Implementation

The varactor model presented in Chapter 3 was not finalized by the time the differential LNA was prepared for fabrication. As such, measured varactor data from the 9-finger AMOS varactors presented in [38] was used to provide a reasonable representation of the varactor behaviour. Although these varactors were not large enough to implement the required neutralizing capacitances, additional ideal capacitor components were connected in parallel with the data components, effectively increasing the capacitance to the required value. Using this information, along with an approximate capacitance per varactor finger value obtained from analyzing the measured 9-finger varactor data, it was determined that 16-finger varactors, connected in a back-to-back manner so as to isolate the applied control voltage from the circuit, were necessary to achieve the desired 32 fF capacitance for C_N.

In order to maximize the Q factor of these varactors, they were designed to have the control voltage applied to their diffusion regions, as presented in Chapter 3. Complicating things somewhat in the varactor implementation was the fact that when connecting one of these varactors between the drain of one transistor in a differential pair and the gate of the other, a different effective DC bias voltage is seen across each of the varactors. As such, the control voltage applied to the diffusion regions of the varactors will have a different effect on each of the varactors in the back-to-back configuration. Figure 5.6 illustrates this idea, showing that when the back-to-back varactor is placed between a drain voltage, V_d, of 1.8 V and a gate voltage, V_g, of 0.8 V, the application of a 1 V control voltage to the diffusion regions will create a different effective voltage between the gate and channel of each of the varactors.

While this difference in effective voltages should not reduce the total swing in

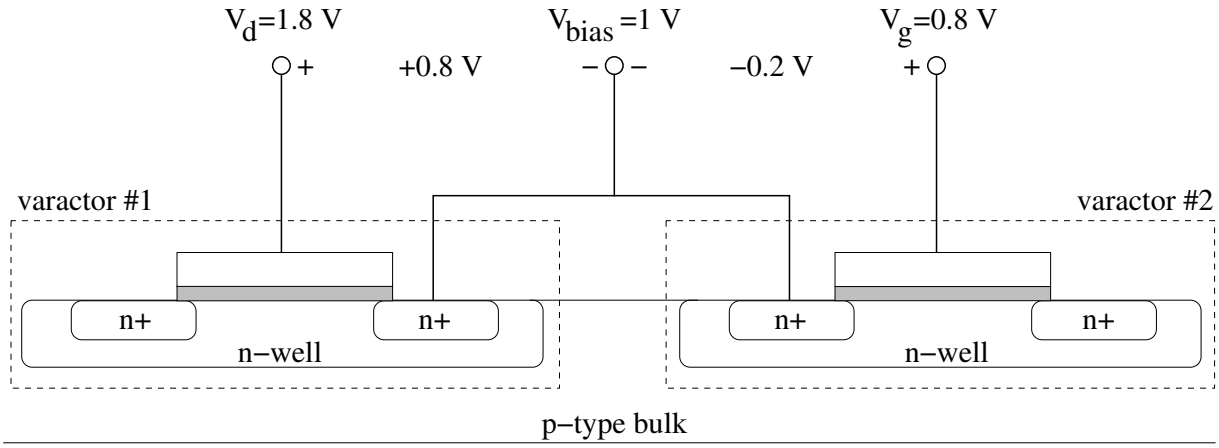

Figure 5.6: Varactor control voltage provides a different effective voltage to each varactor if DC bias at the circuit connection points is different.

capacitance that is achievable by the varactor structure, it does stretch the voltage range over which the varactor can be tuned. This can actually be helpful to the circuit designer since it means that the capacitance implemented by the varactor is less sensitive to slight variations in the control voltage, as can be the case when the varactors presented in Chapter 3 are biased near the middle of their capacitance range.

After the varactor model from Chapter 3 was finalized, simulations of the 16-finger varactor capacitances that were implemented in the differential LNA were completed and are shown in Figure 5.7. These results show the expected increase in the voltage range over which the varactor can be tuned and also shows that the use of 16-finger varactors is appropriate for implementing a 32 fF capacitance, indicating that a control voltage of approximately 0.9 V should achieve the desired results.

To verify this assertion, a test structure of the 16-finger varactor used in the differential LNA design was fabricated. Figure 5.8 illustrates the measured capacitance and resistance of this device, as well as the values predicted by ADS simulation of the

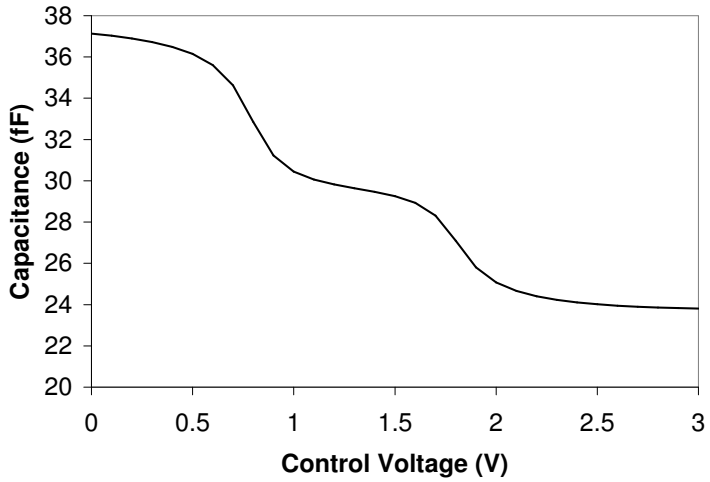

Figure 5.7: Capacitance curve for the back-to-back varactors used to implement C_N.

varactor models from Chapter 3 placed within an EM simulation of the entire test structure. The good agreement shown between these two sets of curves further validates the varactor model and provides added confidence that the 16-finger varactors are appropriately sized for the differential LNA.

5.5 Simulation Results

Since the varactor models had not yet been finalized, initial simulations were carried out using an ideal capacitance in parallel with measured varactor data from smaller varactors to achieve the neutralization. These simulations were carried out using isolated EM simulation data for each of the inductors used in the design in combination with the RF tranistor models for 0.18 μm CMOS supplied by TSMC. The results of the S-parameter simulations conducted in this manner are shown in Figure 5.9.

Highlights of these results are an excellent gain of 13.3 dB and a good output return loss reaching a minimum value of approximately -15 dB at 25 GHz. Input

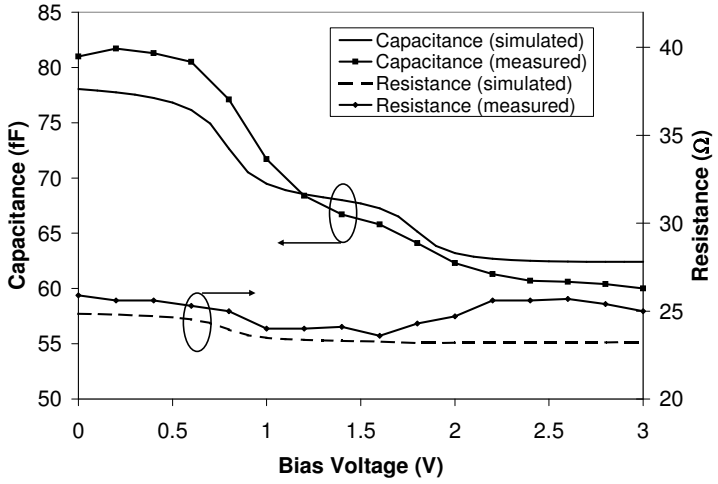

Figure 5.8: Measured capacitance and resistance of the back-to-back, 16-finger varactor used in the differential LNA.

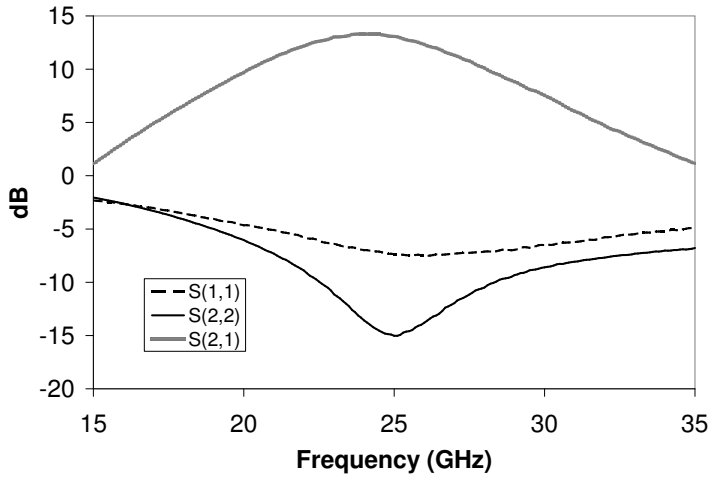

Figure 5.9: Simulated S-parameters using EM simulation data for individual inductor components.

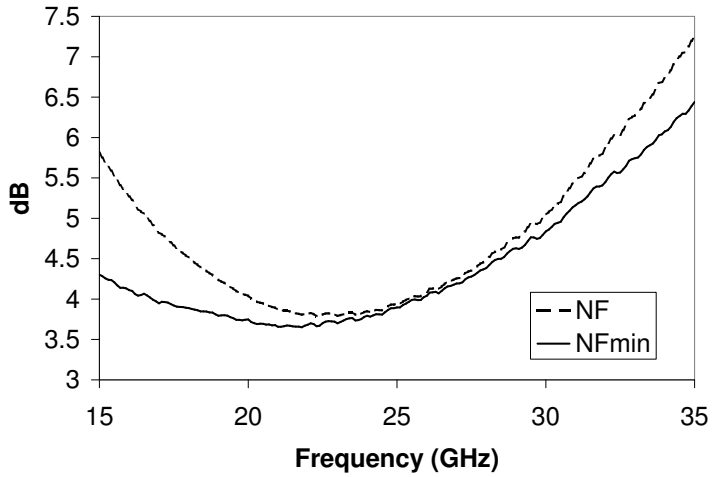

Figure 5.10: Simulated noise figure and minimum noise figure using EM simulation data for individual inductor components.

return loss was designed to be slightly higher in order to produce the minimum possible noise figure but still achieves an acceptably low value of less than -7 dB at 25 GHz. Although not shown in Figure 5.9, reverse isolation was excellent with the help of the neutralizing varactors, remaining below -50 dB for all frequencies. Figure 5.10 then shows the noise performance predicted by this simulation. From these results it is clear that the differential LNA shows excellent noise performance with a noise figure (extremely close to NF_{min}) of only 3.9 dB at 25 GHz.

To account for interactions between the passive components in the circuit design as well as the effects of the finite impedance seen by the ground plane, a full EM simulation of all passives in the differential LNA was also conducted. The S-parameter data of this simulation was then combined with the TSMC transistor models for the amplifier stages as illustrated in the previous chapter. This, along with the use of the high-frequency varactor models presented in Chapter 3, provides an extremely accurate representation of the final circuit behaviour. The final S-paramter results

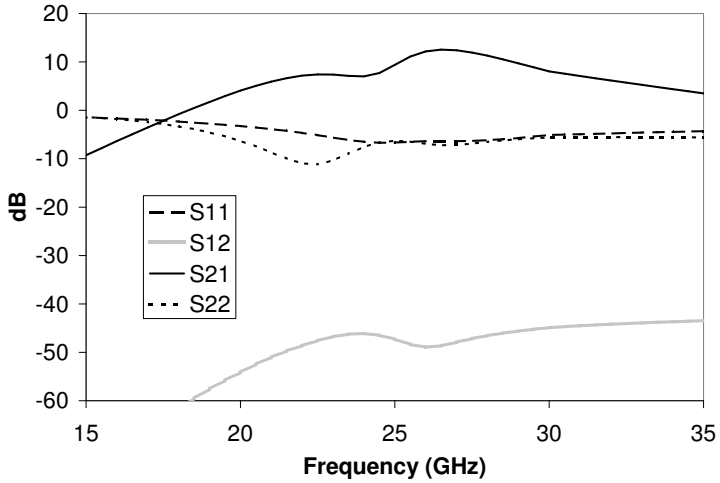

Figure 5.11: Simulated S-parameters obtained using a single EM simulation of all passives and the varactor model presented in Chapter 3.

of this simulation are shown in Figure 5.11 and illustrate a shift in the optimum output match frequency, resulting in a change in the S_{21} curve and an increase in S_{12}. Unfortunately, the computational resources required to perform such a simulation are extreme, requiring the use of advanced computing facilities and as such, this simulation could not completed before submission of the design for fabrication meaning that further circuit tuning to compensate for these effects was not possible. However, these simulation results show that the amplifier still achieves an excellent reverse isolation of less than -40 dB at all frequencies, and a good peak gain of 12.6 dB at 26.5 GHz. This simulation was completed with a control voltage of 1 V applied to the varactor models and shows that the circuit draws a total of 31.7 mA from a 1.8 V supply voltage when the gate voltage of each transistor is set to 0.8 V.

Another important factor to be considered is the stability of the amplifier. As outlined in the previous chapter, this can be determined for any two port device by

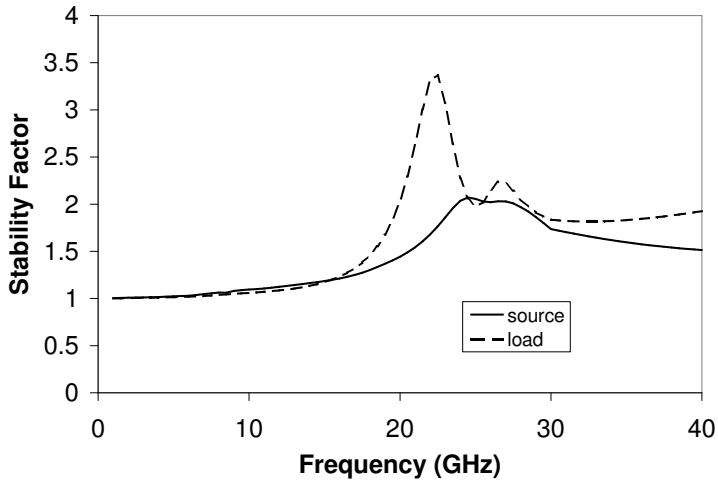

Figure 5.12: Simulated stability factors obtained using a single EM simulation of all passives and the varactor model presented in Chapter 3.

calculating the stability factor, μ, using the equation

$$\mu = \frac{1 - |S_{11}|^2}{|S_{22} - \Delta S_{11}^*| + |S_{12}S_{21}|} \tag{5.4}$$

where

$$\Delta = S_{11}S_{22} - S_{12}S_{21} \tag{5.5}$$

If μ is greater than 1, then the device is unconditionally stable at that frequency. Figure 5.12 illustrates the source and load stability factors produced by the EM simulation of the passive devices along with the transistor and varactor models, and shows that the device is unconditionally stable at all frequencies of interest.

Finally, the noise performance of the circuit predicted by this type of simulation is presented in Figure 5.13. Since the interactions of the passive components did not have a significant effect on the input match, as seen in Figure 5.11, the simulated noise figure still approaches NF_{min} in the frequency range of interest. However, this minimum value shows an increase of approximately 0.5 dB due largely to the effects of the non-ideal ground plane implementation used in the simulation.

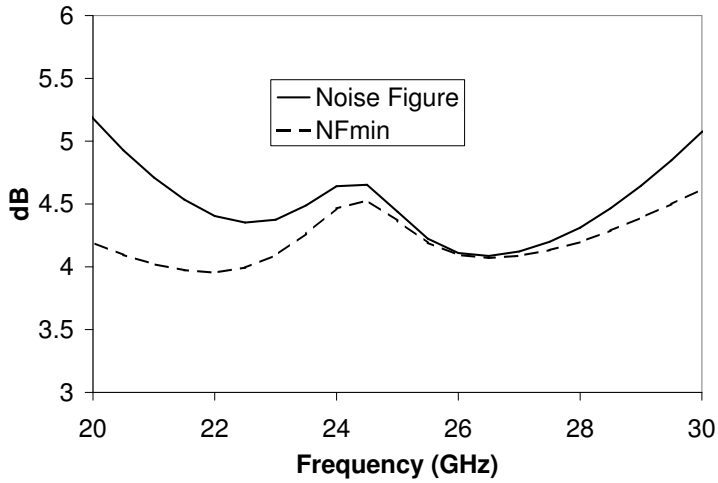

Figure 5.13: Simulated noise figure and minimum noise figure obtained using a single EM simulation of all passives and the varactor model presented in Chapter 3.

5.6 Measured Results

As with the single-ended LNA, a prototype of the differential LNA presented in this chapter was fabricated in a standard 0.18 μm CMOS process with a thick top metal layer. A photograph of the final product is shown in Figure 5.14. This section outlines the results obtained from the physical measurement of this prototype.

Since parasitics on the wires used to supply DC power to the prototype chip led to stability problems with the single-ended LNA in the previous chapter, this issue was immediately analyzed with the differential LNA. Once again, low-frequency oscillations at about 16 MHz were observed in the output spectrum of the amplifier. Fortunately, the same RC filtering and resistance in the V_{dd} supply line used to eliminate the oscillations of the single-ended LNA (described in detail in Appendix B) in combination with large 10 kΩ resistances in the gate and varactor tuning voltage lines proved effective in eliminating the oscillations of the differential LNA as well. Figure 5.15 shows the resulting output spectrum when a 26 GHz signal, with a strength of

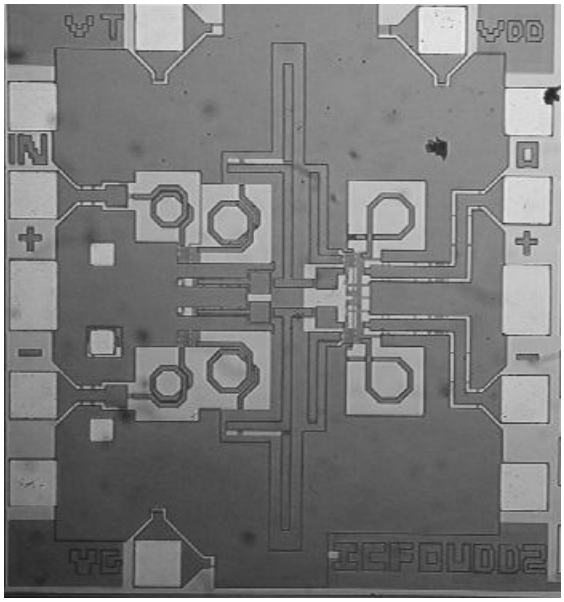

Figure 5.14: Die photograph of the prototype differential LNA fabricated in 0.18 μm CMOS.

-20 dBm, is applied to the input and the oscillation suppression filtering is applied.

With the oscillations removed, it was then possible to conduct S-parameter measurements with the vector network analyzer (VNA). Unfortunately, since the available VNA does not have the capability of performing differential S-parameter measurements directly, measurements were taken in a single-ended fashion and were combined mathematically to produce differential results. This process is illustrated with the help of Figure 5.16, which shows that by eliminating the use of 180^0 hybrid couplers, the (a) 2-port differential device can actually be measured as a (b) 4-port device instead. A more detailed analysis of the final measurement setup used to perform these measurements is presented in Appendix B.

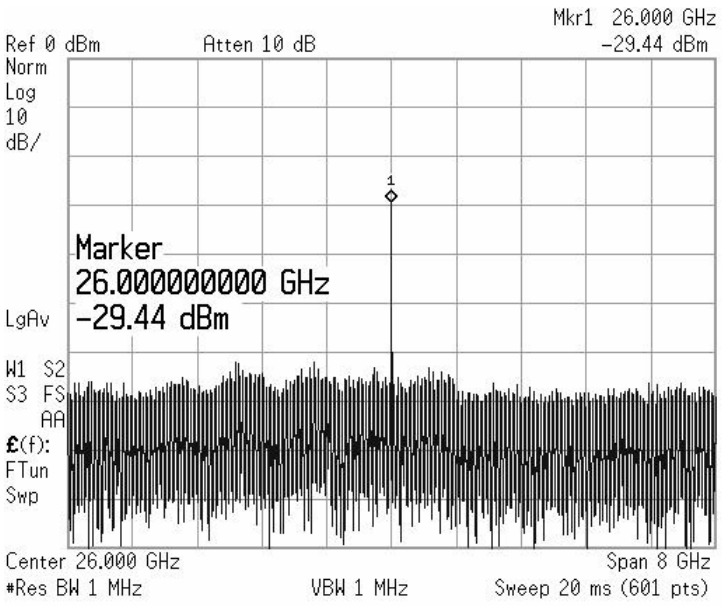

Figure 5.15: Output spectrum of the differential LNA when oscillation suppression is applied to V_{dd}.

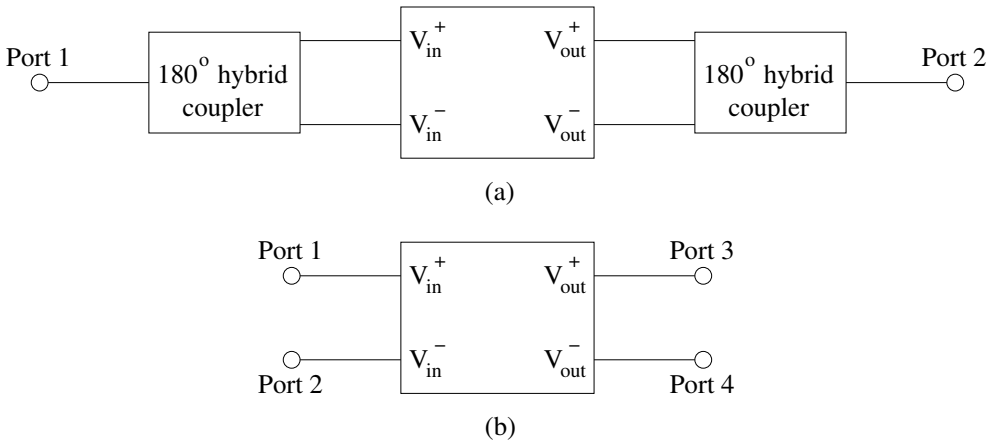

Figure 5.16: By removing the hybrid couplers from the test setup in (a) the differential LNA can be considered to be a (b) 4-port device.

The resulting 4-port measured S-parameters can then be combined using the following equations to provide both differential and common-mode results [50]

$$S_{d1d1} = \frac{1}{2}\left(S_{11} - S_{21} - S_{12} + S_{22}\right) \tag{5.6}$$

$$S_{d1d2} = \frac{1}{2}\left(S_{13} - S_{23} - S_{14} + S_{24}\right) \tag{5.7}$$

$$S_{d2d1} = \frac{1}{2}\left(S_{31} - S_{41} - S_{32} + S_{42}\right) \tag{5.8}$$

$$S_{d2d2} = \frac{1}{2}\left(S_{33} - S_{43} - S_{34} + S_{44}\right) \tag{5.9}$$

$$S_{c1c1} = \frac{1}{2}\left(S_{11} + S_{21} + S_{12} + S_{22}\right) \tag{5.10}$$

$$S_{c1c2} = \frac{1}{2}\left(S_{13} + S_{23} + S_{14} + S_{24}\right) \tag{5.11}$$

$$S_{c2c1} = \frac{1}{2}\left(S_{31} + S_{41} + S_{32} + S_{42}\right) \tag{5.12}$$

$$S_{c2c2} = \frac{1}{2}\left(S_{33} + S_{43} + S_{34} + S_{44}\right) \tag{5.13}$$

where subscript d's represent differential parameters and subscript c's represent common-mode parameters.

Single-ended measurements of the differential LNA were therefore taken with the amplifier biased at V_{dd} = 1.8 V and V_g = 0.8 V, drawing a total of 55.8 mW of DC power. Then, using the technique outlined above, differential S-parameters were constructed from the measured single-ended data and are shown in Figure 5.17. These results show that the varactor neutralization is as effective as the cascode structure at improving the reverse isolation, keeping S_{12} below -20 dB at approximately all frequencies. The input and output return loss values are reasonably low, with S_{22} equal to approximately -6.5 dB and S_{11} (which was designed to produce optimal noise performance instead of gain) equal to -3.7 dB at the maximum gain frequency. Finally, the measured gain displays maximum value of 5.2 dB at a frequency of 22.6 GHz.

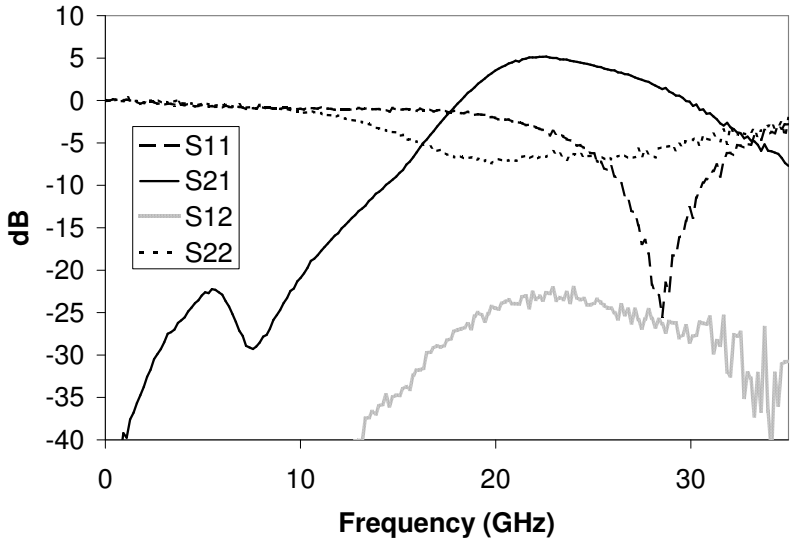

Figure 5.17: Differential S-parameters created using single-ended measurements and Equations 5.6 to 5.9.

This result is lower in both frequency and magnitude than predicted by simulations shown in the previous section. Although similar differences are reported between the simulated and measured S_{21} of the high-frequency differential LNA [22], little evidence is presented explaining the cause of this shift.

One factor contributing to this shift is a difference between the measured gain of an individual transistor and that provided by the high-frequency transistor models used in ADS. To verify this, a 32-finger common source transistor, identical to those used in the differential LNA, was fabricated as a test structure on the same chip. After deembedding the effects of the measurement structure with the help of EM simulation data collected in ADS, the measured transistor data was used in a simulation of the differential LNA in place of the TSMC transistor models. The results of this simulation provide a good match with the measured S-parameters, as shown in Figure 5.18. In particular, these results seem to explain the drop in the magnitude

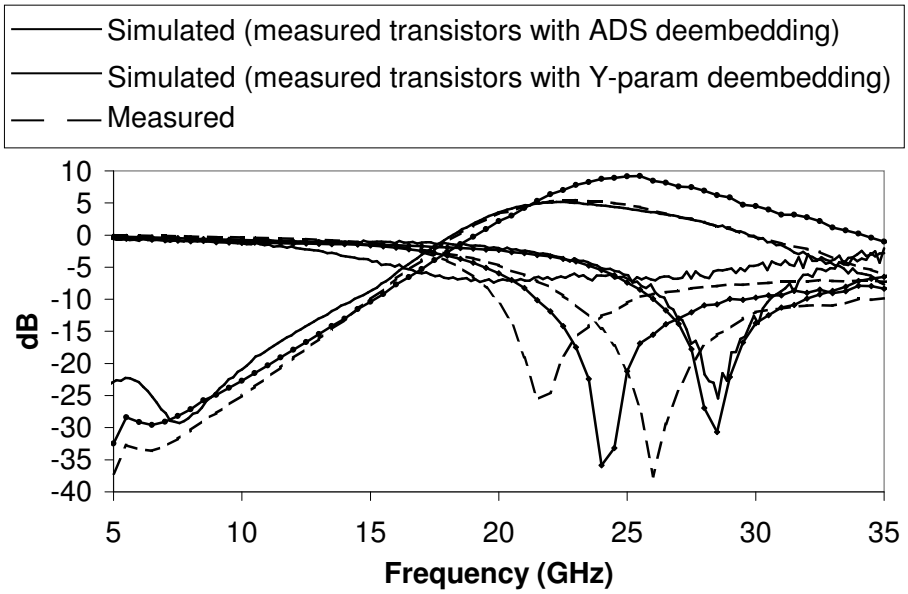

Figure 5.18: Using measured transistor data provides more accurate S-parameter results than using TSMC transistor models.

and frequency of the measured gain of the amplifier and indicate that further work must be done on the high-frequency transistor models to obtain an accurate prediction of amplifier gain. Figure 5.18 also displays the results obtained by deembedding the measured transistor data using measured data for a similar test structure in combination with the Y-parameter method outlined in [53]. While this method does not predict the drop in gain seen by measurement, it does provide an accurate prediction for the input return loss.

A second result of interest that can be obtained from these single-ended s-parameters is the common-mode rejection ration ($CMRR$), which is a ratio of the differential-mode gain, S_{d2d1}, to the common-mode gain, S_{c2c1} [54]. For the amplifier presented in this chapter the $CMRR$ is expected to be low as a result of the omission of current

sources in each differential pair in order to achieve a reduction in power consumption, which is likely to be more of a concern in many CMOS receiver applications. Since the differential-mode gain reported above achieves a peak value of 5.2 dB at 22.6 GHz and the common-mode gain is calculated using Equation 5.12 to be 0 dB at this frequency, the $CMRR$ is therefore also equal to 5.2 dB. While the $CMRR$ is rarely reported in differential amplifier publications, making comparison to other results difficult, it should be possible to improve the $CMRR$ with the use of current sources in each differential pair in the circuit, as described earlier in the chapter. This decision must be made based on the intended application of the differential LNA, as the inclusion of these current sources comes at a cost of a significant increase in power consumption.

Gain and noise figure measurements were then taken using the spectrum analyzer and a detailed look at the measurement setup used for these measurements is presented in Appendix B. Despite the low gain of the amplifier, it still exhibits good agreement between the measured and simulated noise figure, as illustrated by Figure 5.19, achieving a minimum value of only 4.5 dB at a frequency of 23.5 GHz. In addition, the gain measured by the spectrum analyzer compares well with the simulated results obtained using transistor data that has been deembedded using the Y-parameter method discussed above, achieving a maximum value of 7.7 dB at a frequency of 26.2 GHz.

Although this measurement was performed differentially and did not require the S-parameter reconstruction used with the VNA, it should be noted that due to the low gain of the amplifier and the relatively high levels of loss associated with the cables, bias tees and hybrid couplers used to connect the measurement equipment to

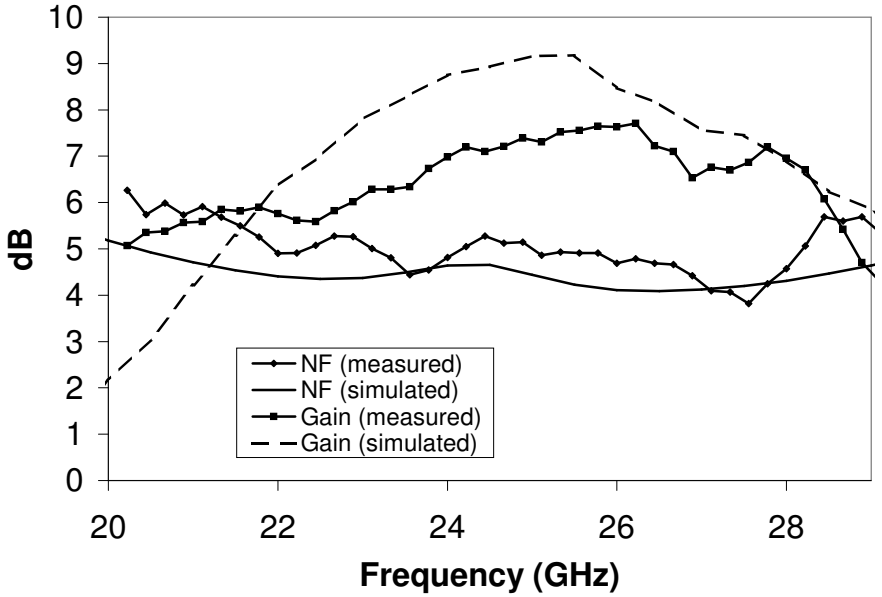

Figure 5.19: Measured noise figure and gain.

the differential LNA, uncertainty in the noise figure measurements was calculated to be a relatively high value of ± 0.7 dB using the calculation tool provided by [49]. This, in combination with non-ideal variations in the phase and magnitude of the signal splitting of the hybrid couplers used, leads to the variations in the measured noise figure and gain curves seen in Figure 5.19 and can account for the difference between the gain measured by the spectrum analyzer and that measured by the VNA. In fact, even though a minimum noise figure of value of 3.8 dB is actually achieved at 27.6 GHz, this value is not reported as the actual NF of the amplifier since it is measured well beyond the maximum rated operating frequency of 26 GHz of the hybrid couplers used, creating low levels of confidence in this result.

Finally, measurements of the output power capabilities of the differential LNA were performed with the spectrum analyzer. Figure 5.20 displays the results of these measurements, which were taken at a frequency of 26 GHz. They highlight a good

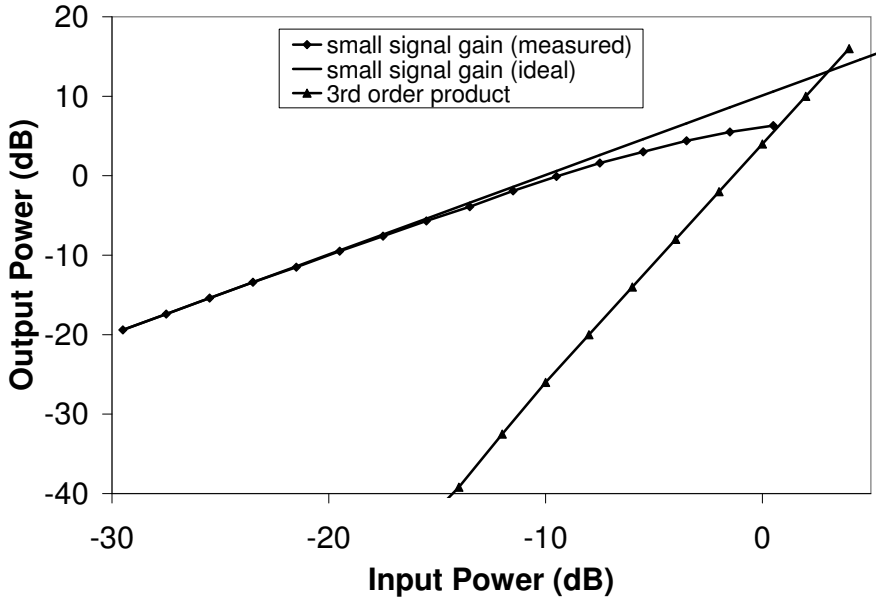

Figure 5.20: Measured output power and linearity of the differential LNA.

input referred 1 dB compression point of -7.5 dBm and a high IIP3 of +3 dBm, both of which are very similar to the single-ended LNA presented in the previous chapter and should be suitable for implementation in the front end of most CMOS receivers.

5.7 Conclusion

A differential, two-stage LNA has been presented for use in 0.18 μm CMOS at a frequency of 23 GHz. As with the single-ended LNA presented in the previous chapter, the transistor size, bias conditions and matching networks were all designed to produce minimum noise. The design also uses the AMOS varactors presented in Chapter 3 to improve the reverse isolation of the amplifier, leading to improved stability and simplifying the matching network design.

Although extremely useful, especially in applications using naturally balanced mixers or A/D converters, very few differential LNAs have been reported at frequencies above 20 GHz to date. Despite this, the differential LNA presented in this chapter exhibits a noise performance comparable to the best single-ended LNAs presented in the previous chapter, achieving a measured noise figure of only 4.5 dB at a frequency of 23.5 GHz. As such, the results have been submitted for publication in January 2007 [41].

Although, the amplifier gain falls short of that predicted by simulations using high-frequency transistor models, the measured LNA gain can be accurately reflected by simulations using measured transistor data and suggests that the models used in the original simulations overestimate the transistor gain at frequencies above 20 GHz. Despite this, the measured gain of 5.2 dB at 22.6 GHz is high enough to be effective at limiting the noise contributions from successive components in a CMOS receiver to the overall NF and could be used in a differential front end.

Chapter 6

Conclusions and Future Work

6.1 Conclusions

As silicon CMOS technology continues its progress in achieving shorter gate lengths, and hence higher f_T values, it is becoming more and more suitable for very high frequency applications. Although circuit design at frequencies above 20 GHz in 0.18 μm CMOS is challenging, the appeal of its low-cost and ease of integration with digital applications has meant that a great deal of research has recently been conducted in this area. This thesis has contributed to this body of knowledge by proposing solutions to some of the most common problems encountered by researchers in the implementation of passive devices at these frequencies, and by proving the functionality of these solutions by implementing them in LNA circuit designs, suitable for the front-end of a high-frequency CMOS receiver.

In the first half of this thesis, the research concerning these passive devices is presented. In Chapter 2, after highlighting the difficulties encountered with the use of spiral inductors, a novel SSTL inductor structure is presented. The advantages of

this structure over the more common spiral or microstrip transmission line inductor structures are threefold:

1. The SSTL structures are shielded from the substrate effects, simplifying the modeling of their behaviour and allowing for accurate prediction of their inductance and Q factor even as frequencies increase beyond 30 GHz.

2. Thanks to these simplified models, optimization of the SSTL structures is straight forward and their optimized Q factors show a distinct high-frequency advantage over spiral inductors.

3. The three-dimensional stub structure reduces the length and overall chip space that would otherwise be required if the inductance were implemented by a simple microstrip structure.

These advantages are confirmed by both simulated and measured results and should remain an attractive alternative to both spiral and simple microstrip inductor structures as researchers continue to design CMOS circuitry at ever higher frequencies.

In Chapter 3 a new model for the relatively recently proposed AMOS varactors is presented. Unlike any previously published model for this device, this model is valid for frequencies above 20 GHz and is based on physical equations that can be easily implemented in circuit simulators. The model is explained in great detail, giving it the advantage of being easily adaptable for any CMOS technology node and for any layout variations. The model also shows good correlation with measured results, giving a great deal of confidence to its ability to predict varactor behaviour in other circuit designs. This should enable the AMOS varactor's usefulness to expand from its traditional role in VCO tuning to other circuits requiring variable or very precise

capacitance values.

In the second half of this thesis, the information presented in Chapters 2 and 3 is put to practical use in two separate LNA designs. Chapter 4 presents a two-stage, single-ended LNA designed to operate at 23 GHz. This LNA makes extensive use of the SSTL inductor design in the input, output and interstage matching networks and uses a cascode configuration to help improve the reverse isolation. This helps to stabilize the amplifier and isolate the input and output matching networks of each transistor stage, but comes at the cost of an increase in noise figure. Although this should mean that the noise figure produced by the amplifier is higher than other reported LNAs not using the cascode configuration, the 5.9 dB measured noise figure ranks in the middle of other published results. This supports the use of the SSTL inductors in the design as their improved Q factor helps to offset the noise produced by the cascode configuration.

In Chapter 5 another low-noise amplifier is presented for operation at 23 GHz, this time in a differential configuration. Again, attention was paid to the improvement of reverse isolation and a neutralization scheme involving cross-coupled capacitances was used to achieve this goal. This design has yet to be reported at high-frequencies and simulations revealed that a slight mismatch in these neutralizing capacitances could lead to instability in the amplifier. As a result, AMOS varactors were implemented with the help of the high-frequency models presented in Chapter 3. These varactors, along with the use of some SSTL inductors, helped to produce an LNA with an excellent measured noise performance of only 4.5 dB at a frequency of 23.5 GHz. Although the measured gain falls short of that predicted by simulations, it still achieves a maximum value of only 5.2 dB and is useful for limiting overall receiver

noise figure in a differential CMOS receiver.

6.2 Future Work

The work presented in this thesis offers opportunity for several areas of further study. First of all, the novel SSTL inductor structures presented in Chapter 2 opens up a great deal of possible future study. The use of recently proposed floating ground structures [9] may be useful in shielding the SSTL from the substrate in the stub section where the ground plane has been removed, which should help to further increase the Q of the device. Furthermore, the inherent inductance of the microstrip structure, along with the capacitance generated between the each layer of the stub section could possibly be used as a high-Q resonator, extending the SSTL's usefulness to other circuit building blocks.

Secondly, the low-frequency instability encountered by the single-ended LNA presented in Chapter 4 was determined be caused by the parasitics and noise associated with the connections used to supply DC power to the circuit. Although the problem was eventually eliminated with the use of RC filtering on the supply line, this solution was only temporary and it would be useful to examine the problem further in the hopes for formalizing a method to prevent these oscillations, preferably with some kind of on-chip solution.

Finally, the discrepancies between the measured transistor performance and that predicted by the TSMC high-frequency models should be addressed. The results presented in Chapter 4 shows that these differences can be especially significant when arranging transistors in a cascode configuration and that the are particularly inaccurate in predicting the gain of the amplifiers. Further work done to modify the TSMC

model to achieve a better fit to the measured data would be useful in future designs using either cascode or differential topologies and could help to further reduce the noise figure of LNAs designed in this way.

Bibliography

[1] CTIA Research, "CTIA's wireless industry INDICES semi-annual data survey results," Cellular Telecommunications & Internet Association, Tech. Rep., January 2005.

[2] T. Lee, "CMOS RF: (still) no longer an oxymoron," in *IEEE Radio Frequency Integrated Circuits Symposium*, 1999, pp. 3–6.

[3] S. Voinigescu, T. Dickson, R. Beerkens, I. Khalid, and P. Westergaard, "A comparison of Si CMOS, SiGe BiCMOS, and InP HBT technologies for high-speed and millimeter-wave ICs," in *Topical Meeting on Silicon Monolithic Integrated Circuits in RF Systems*, 2004, pp. 111–114.

[4] R. Rodriguez, J. Dishman, F. Dickens, and E. Whelan, "Modeling of two-dimensional spiral inductors," *IEEE Transactions on Components, Hybrids, and Manufacturing Technology*, vol. 3, no. 4, pp. 535–541, December 1980.

[5] J. Long and M. Copeland, "The modeling, characterization, and design of monolithic inductors for silicon RF IC's," *IEEE J. Solid-State Circuits*, vol. 32, no. 3, pp. 357–369, March 1997.

[6] J. Guo and T. Tan, "A broadband and scalable model for on-chip inductors incorporating substrate and conductor loss effects," *IEEE Trans. Electron Devices*, vol. 53, no. 3, pp. 413–421, March 2006.

[7] Y. Chen, D. Bien, D. Heo, and J. Laskar, "Q-enhancement of spiral inductor with N+ diffusion patterned ground shields," in *IEEE MTT-S International Microwave Symposium Digest*, 2001, pp. 1289–1292.

[8] C. Doan, S. Emani, A. Niknejad, and R. Brodersen, "Millimeter-wave CMOS design," *IEEE J. Solid-State Circuits*, vol. 40, no. 1, pp. 144–155, January 2005.

[9] T. Cheung and J. Long, "Shielded passive devices for silicon-based monolithic microwave and millimeter-wave integrated circuits," *IEEE J. Solid-State Circuits*, vol. 41, no. 5, pp. 1183–1200, May 2006.

[10] B. Razavi, "A 60-GHz CMOS receiver front-end," *IEEE J. Solid-State Circuits*, vol. 41, no. 1, pp. 17–22, January 2006.

[11] J. Nadan, "Experimental capacitance-voltage characteristics of three-terminal digital MOS varactors," *Proceedings of the IEEE*, vol. 61, no. 10, pp. 1512–1513, October 1973.

[12] T. Soorapanth, C. Yue, D. Shaeffer, T. Lee, and S. Wong, "Analysis and optimization of accumulation-mode varactor for RF ICs," in *1998 Symposium on VLSI Circuits. Digest of Technical Papers*, 1998, pp. 32–33.

[13] N. Fong, G. Tarr, N. Zamdmer, J.-O. P., and C. Plett, "Accumulation MOS varactors for 4 to 40 GHz VCOs in SOI CMOS," in *IEEE International SOI Conference*, 2002, pp. 158–160.

[14] J. Maget, R. Kraus, and M. Tiebout, "A physical model of a CMOS varactor with high capacitance tuning range and its application to simulate a voltage controlled oscillator," in *International Semiconductor Device Research Symposium*, 2001, pp. 609–612.

[15] S. Song and H. Shin, "An RF model of the accumulation-mode MOS varactor valid in both accumulation and depletion regions," *IEEE Trans. Electron Devices*, vol. 50, no. 9, pp. 1997–1999, September 2003.

[16] T. Nguyen, C. Kim, G. Ihm, M. Yang, and S. Lee, "CMOS low-noise amplifier design optimization techniques," *IEEE Trans. Microwave Theory Tech.*, vol. 52, no. 5, pp. 1433–1442, May 2004.

[17] S. Shin, M. Tsai, R. Liu, K. Lin, and H. Wang, "A 24-GHz 3.9-db NF low-noise amplifier using 0.18 μm CMOS technology," *IEEE Microwave Wireless Compon. Lett.*, vol. 15, no. 7, pp. 448–450, July 2005.

[18] K. Yu, Y. Lu, D. Chang, V. Liang, and M. Chang, "K-band low-noise amplifiers using 0.18 μm CMOS technology," *IEEE Microwave Wireless Compon. Lett.*, vol. 14, no. 3, pp. 106–108, March 2004.

[19] K. Sun, Z. Tsai, K. Line, and H. Wang, "A noise optimization formulation for CMOS low-noise amplifiers with on-chip low-Q inductors," *IEEE Trans. Microwave Theory Tech.*, vol. 54, no. 4, pp. 1554–1560, April 2006.

[20] U. Basaran, R. Tao, L. Wu, and M. Berroth, "Low-noise amplifier in 0.12 μm standard CMOS technology for K-band," *IEE Electronics Letters*, vol. 41, no. 10, pp. 592–593, May 2005.

[21] F. Ellinger, "24-42 GHz SOI CMOS low noise amplifier," *IEEE J. Solid-State Circuits*, vol. 39, no. 3, pp. 522–528, March 2004.

[22] J. Guo and Y. Lin, "A lossy substrate model for sub-100 nm, super-100 GHz f_T RF CMOS noise extraction and modeling," in *IEEE Radio Frequency Integrated Circuits Symposium*, 2005, pp. 145–148.

[23] F. Huang, N. Jiang, and E. Bian, "Characteristic-function approach to parameter extraction for asymmetric equivalent circuit of on-chip spiral inductors," *IEEE Trans. Microwave Theory Tech.*, vol. 54, no. 1, pp. 115–119, January 2006.

[24] A. Telli, I. Demir, and M. Askar, "Practical performance of planar spiral inductors," in *IEEE International Conference on Electronics, Circuits and Systems*, 2004, pp. 487–490.

[25] A. Niknejad and R. Meyer, "Analysis, design, and optimization of spiral inductors and transformers for Si RF IC's," *IEEE J. Solid-State Circuits*, vol. 33, no. 10, pp. 1470–1481, October 1998.

[26] B. Frank, "Spiral inductor design," web site:

http://bmf.ece.queensu.ca/mediawiki/index.php/Spiral_Inductor_Design, May 2006.

[27] F. Leferink, "Inductance calculations; methods and equations," in *IEEE International Symposium on Electromagnetic Compatibility*, 1995, pp. 16–22.

[28] D. Dunwell and B. Frank, "24 GHz low-noise amplifiers using high Q series-stub transmission lines in 0.18 μm CMOS," in *IEEE Canadian Conference on Electrical and Computer Engineering*, 2006, pp. 1887–1890.

[29] P. Andreani and S. Mattisson, "On the use of MOS varactors in RF VCOs," *IEEE J. Solid-State Circuits*, vol. 35, no. 6, pp. 905–910, June 2000.

[30] Microelectronics Processing Technology, "Metal oxide semiconductor (MOS) capacitor," web site:

http://hackman.mit.edu/6152J/SP_2004/supplementals/sp_2005_6_152J_ST04_MOSCap.pdf, Spring 2004.

[31] R. Shrivastava and K. Fitzpatrick, "A simple model for the overlap capacitance of a VLSI MOS device," *IEEE Trans. Electron Devices*, vol. 29, no. 12, pp. 1870–1875, December 1982.

[32] Y. Tsividis, *Operation and Modeling of the MOS Transistor*. McGraw-Hill, 1999.

[33] C. Wang, "Identification and measurement of scaling-dependent parasitic capacitances of small-geometry MOSFET's," *IEEE Trans. Electron Devices*, vol. 43, no. 6, pp. 965–972, June 1996.

[34] B. Streetman, *Solid State Electronic Devices*. Prentice Hall, 1995.

[35] C. Enz, "An AMOS transistor model for RF IC design valid in all regions of operation," *IEEE Trans. Microwave Theory Tech.*, vol. 50, no. 1, pp. 342–359, January 2002.

[36] C. Balanis, *Antenna Theory: Analysis and Design*. Harper & Row, 1982.

[37] F. Ulaby, *Applied Electromagnetics*. Prentice Hall, 1999.

[38] J. Carr and B. Frank, "A 38 GHz accumulation MOS differentially tuned VCO design in 0.18 μm CMOS," in *Topical Meeting on Silicon Monolithic Integrated Circuits in RF Systems*, 2006, pp. 170–173.

[39] K. Molnar, "MOS varactor modeling with a subcircuit utilizing the BSIM3v3 model," *IEEE Trans. Electron Devices*, vol. 49, no. 7, pp. 1206–1211, July 2002.

[40] N. Fong, J.-O. Plouchart, N. Zamdmer, D. Liu, L. Wagner, C. Plett, and N. Tarr, "A 1-v 3.8-5.7 GHz wide-band VCO with differentially tuned accumulation MOS varactors for common-mode noise rejection in CMOS SOI technology," *IEEE Trans. Microwave Theory Tech.*, vol. 51, no. 8, pp. 1952–1959, August 2003.

[41] D. Dunwell and B. Frank, "Accumulation-mode MOS varactors for RF CMOS low-noise amplifiers," in *IEEE Topical Meeting on Silicon Monolithic Integrated Circuits in RF Systems*, 2007.

[42] X. Guan and A. Hajimiri, "A 24 GHz CMOS front-end," *IEEE J. Solid-State Circuits*, vol. 39, no. 2, pp. 368–373, February 2004.

[43] K. Yu, Y. Lu, D. Huang, D. Chang, V. Liang, and M. Chang, "24 GHz low-noise amplifier in 0.18 μm CMOS technology," *IEE Electronics Letters*, vol. 39, no. 22, pp. 1559–1560, October 2003.

[44] T. Lee, *The Design of CMOS Radio Frequency Integrated Circuits*. Cambridge University Press, 1998.

[45] C. Chen, M. J. Deen, Y. Cheng, and M. Matloubian, "Extraction of the induced gate noise, channel noise, and their correlation in submicron MOSFETs from

RF noise measurements," *IEEE Trans. Electron Devices*, vol. 48, no. 12, pp. 2884–2892, December 2001.

[46] M. Hossain, "A fully integrated K-band receiver front-end in CMOS," Master's thesis, Queen's University, 2004.

[47] B. Razavi, *Design of Analog CMOS Integrated Circuits*. McGraw-Hill, 2001.

[48] D. Pozar, *Microwave Engineering*. John Wiley & Sons, 1998.

[49] Agilent Technologies, "The noise figure uncertainty calculator," web site: http://contact.tm.agilent.com/Agilent/tmo/Format/K2Pages/NFUcalc.html, July 2006.

[50] Anritsu, "Application note: Three and four port S-parameter measurements," Anritsu, Tech. Rep., May 2002.

[51] D. Cassan and J. Long, "A 1-V transformer-feedback low-noise amplifier for 5-GHz wireless LAN in 0.18 μm CMOS," *IEEE J. Solid-State Circuits*, vol. 38, no. 3, pp. 427–435, March 2003.

[52] C. Cheng, "Neutralization and unilateralization," *IRE Transactions on Circuit Theory*, vol. 2, no. 2, pp. 138–145, June 1955.

[53] J. Hasch and E. Kasper, "S-parameter characterization of mm-wave IMPATT oscillators," in *Topical Meeting on Silicon Monolithic Integrated Circuits in RF Systems. Digest of Papers.*, 2006, pp. 213–217.

[54] Agilent Technologies, "Agilent balanced measurement example: Differential amplifiers," Agilent Technologies, Tech. Rep., September 2001.

[55] J. Su, H. Hsu, S. Wong, C. Chang, T. Huang, and J. Sun, "Improving the RF performance of 0.18 μm CMOS with deep n-well implantation," *IEEE Trans. Electron Devices*, vol. 22, no. 10, pp. 481–483, April 2001.

[56] Taiwan Semiconductor Manufacturing Company, "TSMC 0.18 μm mixed signal 1P6M salicide 1.8 V/3.3 V SPICE models," Proprietary Information, 2002.

[57] W. Liu, *MOSFET Models for SPICE Simulation Including BSIM3v3 and BSIM4*. John Wiley & Sons, 2001.

Appendix A

Varactor Model Parameter Values

The parameter values used in each of the varactor model lumped element calculations were amalgamated from a wide number of sources. In an effort to make the chapter easier to follow and to allow the presented model to be easily reproduced, the values used, as well as the sources from which they were taken, are listed in Table A.1 below. Those parameters without references are common constants that should be readily accessible from many sources.

Symbol	Description	Value	Reference
D_n	n well diffusion depth	0.8 μm	[55]
ϵ_0	Dielectric constant of free space	8.854x10$^{-12}\frac{F}{m}$	
ϵ_{ox}	Oxide dielectric constant	3.9 ϵ_0	[56]
ϵ_{si}	Silicon dielectric constant	11.8 ϵ_0	[32]
k	Boltzman's constant	1.381x10$^{-23}\frac{J}{K}$	
L	Channel length	0.18 μm	
L_{dif}	Diffusion region length	0.6 μm	*
L_{via}	Via length	7.8 μm	*
N_i	Silicon intrinsic carrier concentration	1.18x10^{16} m^{-3}	[32]
N_{sub}	Substrate doping density	6x10^{22} m^{-3}	[57]
N_{well}	n well doping density	1.7x10^{23} m^{-3}	[57]
Φ_P	Flat-band voltage constant	-0.55 V	[30]
q	Elementary charge	1.6x10^{-19} C	
R^d_{sq}	Diffusion region resistance	6.8$\pm$2.5 $\frac{\Omega}{sq}$	[56]
R^g_{sq}	Polysilicon resistance	7.4$\pm$2 $\frac{\Omega}{sq}$	[56]
r_{via}	Via radius	0.26 μm	*
T	Operating temperature	300 K	
t_{ox}	Oxide thickness	4.08 nm	[56]
t_{poly}	Polysilicon thickness	200 nm	[56]
μ_n	Surface electron mobility	0.067 $\frac{m^2}{Vs}$	[57]
μ_0	Permeability of free space	1.2566x10$^{-6}\frac{H}{m}$	
W_{lay}	Gate finger width in layout	2.57 μm	*
X_j	n^+ diffusion depth	0.15 μm	[57]

* Layout dependent parameters

Table A.1: Parameter values used in calculations of the varactor model lumped elements.

Appendix B

LNA Measurement Setup

The difficulties encountered during measurement of both the single-ended and differential LNAs due to the parasitics associated with DC supply connections are not uncommon when performing high-frequency testing. As such, a detailed look at the setups used to perform measurements while stopping oscillation is presented in this section since it may be useful in compensating for similar problems during future measurements.

B.1 Single-Ended LNA

A block diagram representation of the test setup used to perform noise and gain measurements of the single-ended LNA is presented in Figure B.1. In order to account for losses in the bias tees and cables, loss compensation was set on the spectrum analyzer to 4.5 dB before the device under test (DUT) and 4.5 dB after the DUT. These values were obtained earlier by taking measurements of each of these passive devices using the vector network analyzer (VNA). Although measuring these losses

by connecting the passive devices directly between the signal generator and spectrum analyzer yield higher values of approximately 5.7 dB before the DUT and 5.2 dB after the DUT, more confidence was placed in the absolute power levels recorded by the VNA and these earlier values were used instead. It should also be noted that the 70 Ω resistance placed in the DC bias path has an obvious effect on the amplifier bias conditions. To compensate for this, the supply voltage for V_{DD} was increased to 5.4 V as this value yielded the correct voltage at the circuit connection point and provided an appropriate bias current. The same bias network and oscillation suppression techniques were used to perform s-parameter measurements with the vector network analyzer (VNA) and as such the test setup used to perform these measurements is very similar to that of Figure B.1 and is not reproduced.

A more detailed look at the RC filter used to mitigate oscillations on the drain voltage supply line is presented in Figure B.2. Since this filter was designed to combat low-frequency oscillations, a simple soldered combination of discrete resistors and capacitors, when used in combination with the ferrite beads, was sufficient to remove the oscillations. Parallel or series combinations of resistors and capacitors were used in order to achieve the appropriate resistance and capacitance values from available components.

B.2 Differential LNA

Measurement of the noise figure and gain of the differential LNA was conducted with a similar setup to that used for the single-ended LNA. Differences include the use of 180^0 hybrid couplers to create a differential signal at the input and then recombine this signal at the output, and the addition of 10 kΩ resistors in the bias lines used to

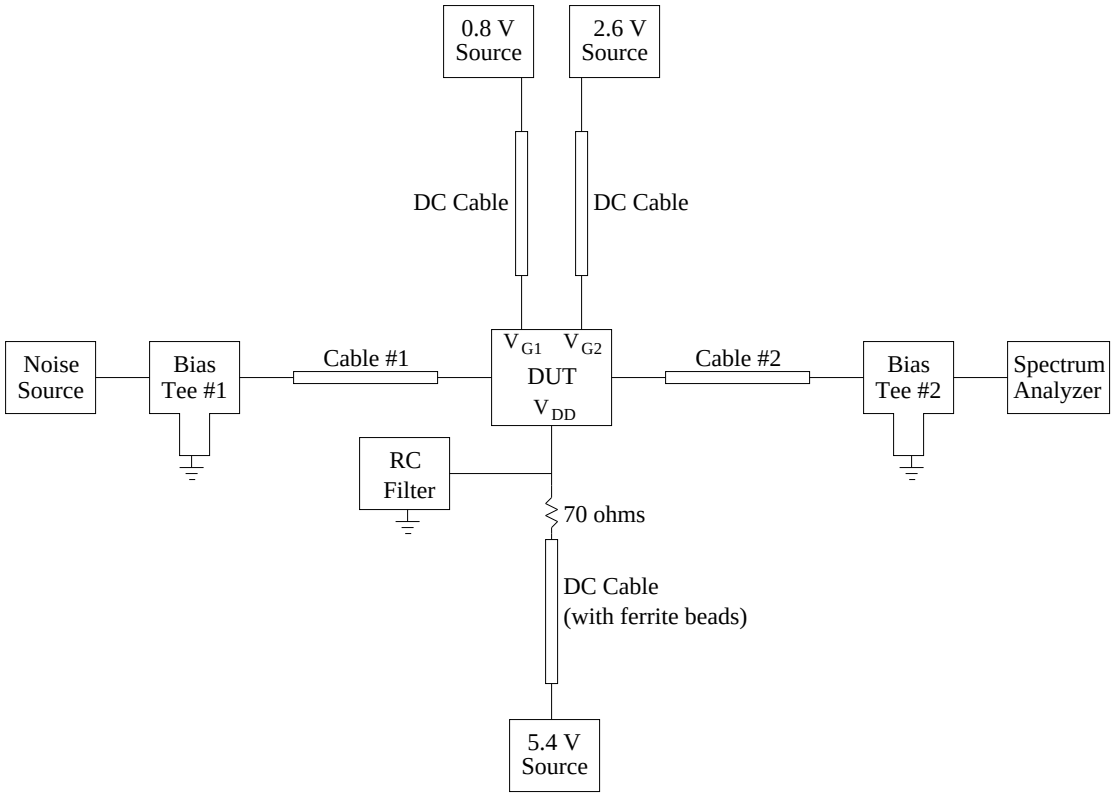

Figure B.1: Test setup used to eliminate oscillation while performing noise figure measurements on the single-ended LNA.

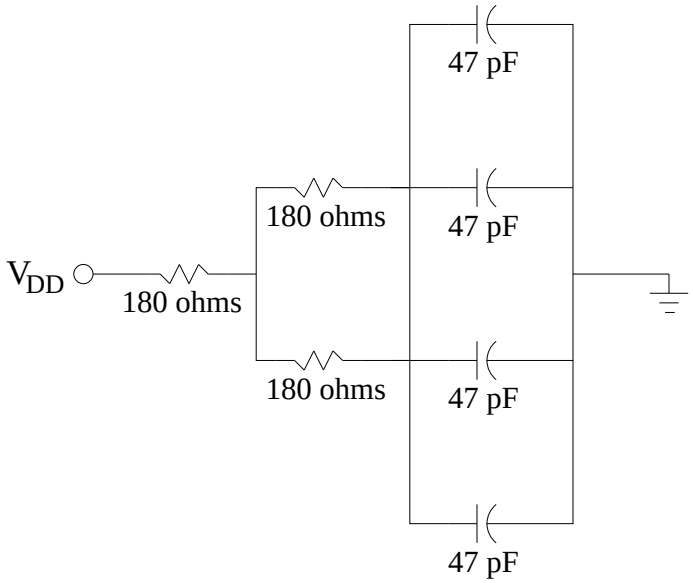

Figure B.2: Simple RC filter used to kill oscillation the drain voltage supply line.

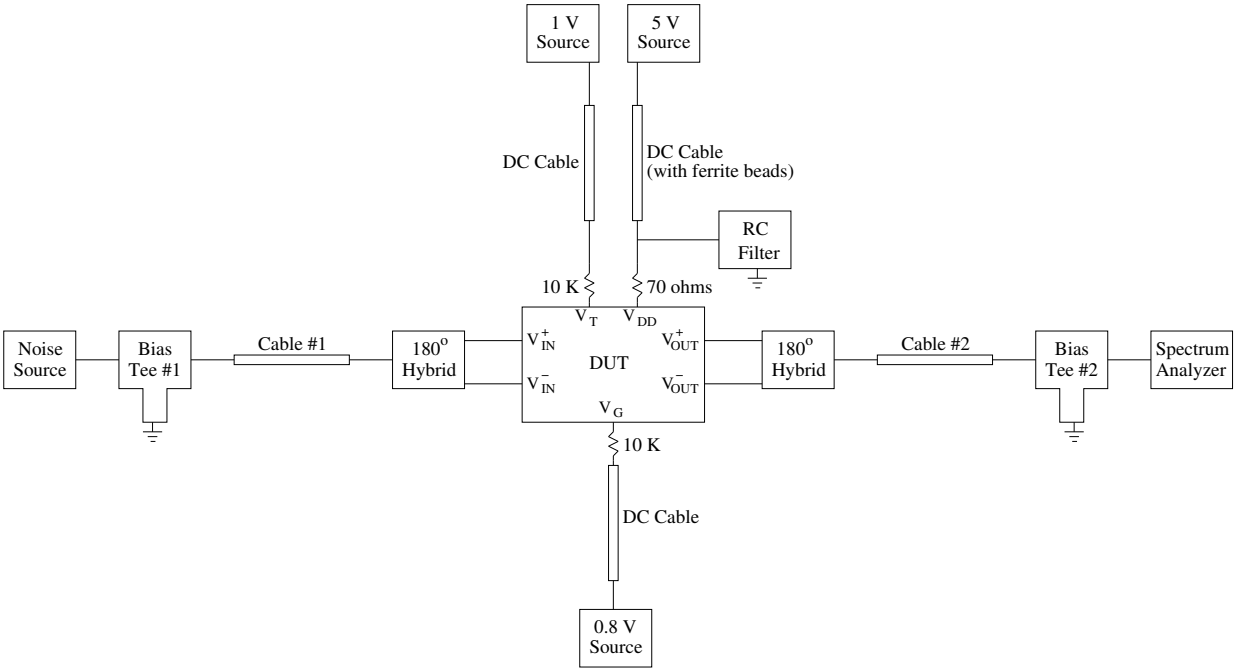

Figure B.3: Test setup used to eliminate oscillation while performing noise figure measurements on the differential LNA.

supply the varactor tuning voltage, V_T, and the gate bias voltage, V_G, to the DUT. These resistors help to further isolate the circuit from any parasitics on the DC bias lines, which do not supply current to the LNA. The resulting setup is displayed in Figure B.3.

Unfortunately, fully differential calibration and measurement was not possible with the available VNA. Instead, single-ended measurements were performed and then combined mathematically to produce differential s-parameters, as explained in Chapter 5. Figure B.4 illustrates this idea by showing the test setup used to perform the measurement of S_{31}. Also worth noting is the inclusion of a power splitter between the DUT and port 2 of the VNA. This is used to direct a portion of the output signal to the spectrum analyzer in order to monitor the output spectrum and ensure that oscillations are not present. However, although the effects of this setup can

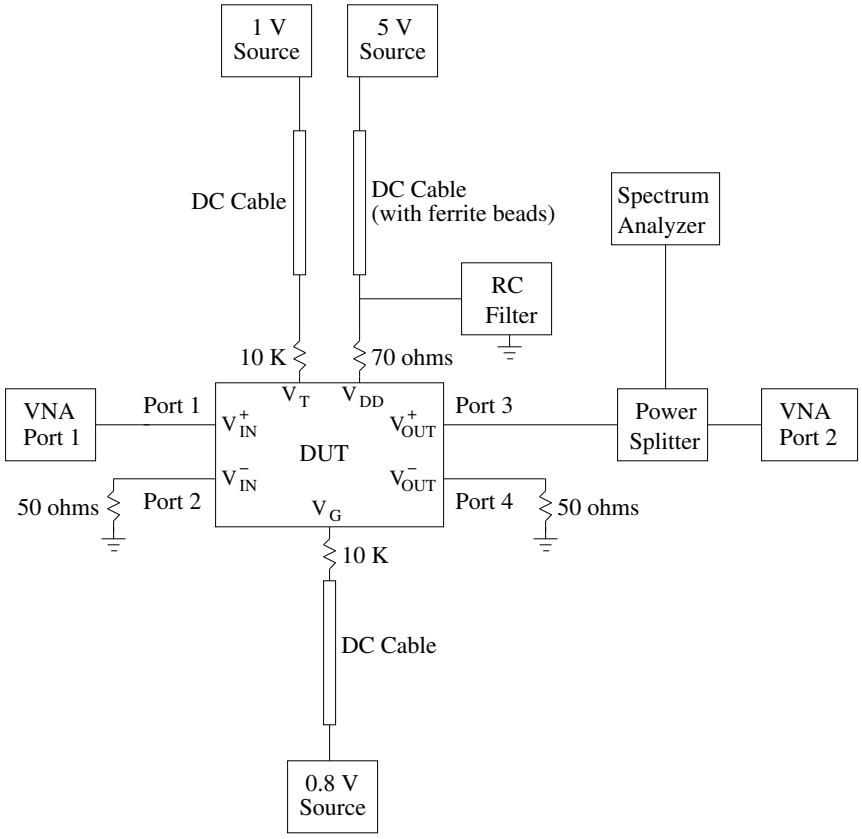

Figure B.4: Test setup used to eliminate oscillation while performing s-parameter measurements on the differential LNA.

be partially removed by calibration of the VNA, the measured s-parameters can be improved if this section is removed. As a result, after ensuring that oscillations were successfully suppressed, the power splitter and spectrum analyzer were removed from the test setup and the VNA was recalibrated to obtain valid and repeatable s-parameter measurements.