Extension of 0.18μm Standard CMOS Technology Operating Range to the Microwave and Millimetre-Wave Regime

A Thesis Submitted To
School Of Engineering
Electronics and Nanoscale Engineering Research Division
University Of Glasgow
In Fulfilment Of The Requirements
For The Degree Of
Doctor Of Philosophy

By
Salah-Aldeen Sharabi, *B.Eng. (Hons)*
December 2014

Abstract

There is an increasing interest in building millimetre-wave circuits on standard digital complementary metal oxide semiconductor (CMOS) technology for applications such as wireless local area networks (WLAN), automotive radar and remote sensing. This stems from the existing low cost, well-developed, high yield infrastructure for mass production. The overall aim of this thesis is to extend the operating range of 0.18 μm standard logic CMOS technology to millimetre-wave regime. To this end, microwave and millimetre-wave design, optimisation and modelling methodologies for active and passive devices and low noise circuit implementation are described. As part of the evaluation, new systematic and modular ways of making high performance passive and active devices such as spiral inductors, slow-wave coplanar waveguide (CPW) transmission lines, comb capacitors and NMOS transistors are proposed, designed, simulated, fabricated, modelled and analysed. Small-signal and noise de-embedding techniques are developed and verified up to 110 GHz, providing an increased accuracy in the device model, leading to a robust design at millimetre-wave frequencies. Reduced substrate losses resulting in increased quality factor are presented for optimised spiral inductor designs, featuring patterned floating shield (PFS), enabling improved matching network and a reduced chip area. Based on the proposed shielded slow-wave CPW, both the line attenuation and structure length are decreased, resulting in a more compact and simplified circuit design. An optimised transistor design, aimed at reducing the layout parasitic effects, was realised. The optimisation led to a significant improvement in the gain and noise performance of the transistor, extending its operation beyond the cut-off frequency (f_t). By combining all the optimised components, low noise amplifiers (LNAs) operating at 25 GHz and 40 GHz were implemented and compared. These LNAs demonstrate state-of-the-art performance, with the 40 GHz LNA exhibiting the highest gain and lowest noise performance of any LNA reported using 0.18 μm CMOS technology. On the other hand, the 25 GHz LNA showed a comparable performance to other reported results in literature using several topologies implemented in CMOS technology. These findings will provide a framework for expansion to smaller CMOS technology nodes with the view of extending to sub millimetre-wave frequencies.

"Character cannot be developed in ease and quiet. Only through experience of trial and suffering can the soul be strengthened, ambition inspired, and success achieved."

Helen Keller

Acknowledgements

This thesis is a milestone of more than 5 years of work. My experience at the University of Glasgow has been nothing short of amazing. This thesis presents the lessons learnt and also a result of many experiences I have encountered from many remarkable individuals whom I also wish to acknowledge. Though the list of people to thank is long, making this list is not the hard part. The hard part is finding the words that convey the sincerity and magnitude of my gratitude. It is with all this help that I stand where I do.

First and foremost, I would like to thank Professor Lee Cronin, who has been supportive since the day I began working for him. I would like to thank my project supervisor Dr. Edward Wasige for providing me with the opportunity to work on this project and for all the help provided during the course of my PhD, thanks for putting up with me. I would also like to thank my second supervisor Professor David Cumming.

Many thanks to my PhD examiners, Dr. Chris Oxley and Dr. Khaled Elgaid, for your helpful feedback during the my Viva. Also thanks are due to the Engineering, Physics and Science Research Council, EPSRC, and Texas Instruments for the fabrication support they provided during the course of this study.

Special thanks to the sales engineers from Castle microwave as well as Murray microwave and Focus microwave engineers for sharing their vast knowledge about noise measurements and techniques and for all their help with equipments and measurement setup ... I am honestly grateful.

Thanks to Ian McGregor and Huracio Cantu for all the technical discussions and help. Also many thanks to Christopher Boyle for his fantastic patience while proof reading this thesis.

The largest part of my life in Glasgow was dominated by the students and people that I worked with and learnt from. Thanks to those, and in no specific order: Kirsty Walls, Sankar Peralagu, Kamil Rudnicki, Steve Russell, Mohammed Al-Rawhani, Chris Martin, James Grant, Piotr Stolarz, Douglas McFarlane, Giangiacomo Cramarossa, Vasileios Papageorgiou, Mohsen Haji, Olysia Ignatova, Vincenzo Pusino, Lourdes Ferre Llin, Afesomeh, Jue Wang, Sanna Taking, Sonia Holik, Irene Suarez, Sergio Martin and Mike Lee.

I would like to thank my amazing family. Their support and patience have been unconditional all these years. Special thanks to my parents Nasib and Faiza, my fabulous grandmother and my incredibly patient uncles Yasser and Abed. Their unwavering love and care, and support provide me with the energy I need to pursuit my goals in life.

Finally, I would like to dedicate this thesis to the memory of a great man, Trevor Colin Fletcher. Your supportive and inspiring actions and comments helped me and encouraged me in the worst of times. We wish you were here to see this. Thank you for everything.

Contents

List of Figures

List of Tables

List of Acronyms

Acronym	Description
AC	alternating current
ADC	analog to digital converter
ADS	advanced design system
BiCMOS	bipolar complementary metal oxide semiconductor
CAD	computer-aided design
CG	common gate
CMOS	complementary metal oxide semiconductor
CPW	coplanar waveguide
CS	common source
DC	direct current
DUT	device under test
EM	electromagnetic
EPI	epitaxial
ESD	electrical static discharge
FET	field-effect-transistor
FOM	figure of merit
GCPW	grounded coplanar waveguide
GSG	ground-signal-ground
HBT	heterojunction bipolar transistor
HEMT	high electron mobility transistor
IC	integrated circuit
IF	intermediate frequency
ISS	impedance standard substrate
LNA	low noise amplifier
LO	local oscillator
LRM	line-reflect-match
LRRM	line-reflect-reflect-match

MEMS	micro electromechanical switch
MIM	metal insulator metal
MOM	metal oxide metal
MOS	metal oxide semiconductor
MOSFET	metal-oxide-semiconductor field-effect-transistor
MSG	maximum stable gain
NF	noise figure
NQS	non quasi-static
non-EPI	non-epitaxial
PFS	patterned floating shield
PGS	patterned ground shield
pHEMT	pseudomorphic high electron mobility transistor
PSD	power spectral density
RF	radio frequency
RFIC	radio frequency integrated circuit
RMS	root mean square
S-CPW	slow-wave coplanar waveguide
SMU	source/monitor unit
SoC	system on chip
SOI	silicon on insulator
SOLR	short-open-load-reciprocal
SOLT	short-open-load-thru
SNR	signal-to-noise ratio
STI	shallow trench isolation
TI	texas instruments
TLINP	ideal physical transmission line component
TMA	trimethyl-aluminum
TRL	thru-reflect-line
VCO	voltage controlled oscillator
VLSI	very large scale integration
VNA	vector network analyzer
WLAN	wireless local area networks

List of Symbols

Symbol	Description
A	area, attenuation, ABCD matrix element
A_v	voltage gain
B	susceptance, bandwidth
BeCu	Beryllium copper
C	capacitance
C_{ox}	oxide capacitance
C_p	parasitic capacitance
d_{in}	inner diameter of a spiral
d_{in}	outer diameter of a spiral
f	Frequency
f_{max}	frequency at unity power gain
f_t	frequency at unity current gain
F_{min}	minimum noise figure
G	Conductance, gain
G_A	available power gain
g_d	output conductance
g_m	transconductance
G_P	power gain
G_T	transducer gain
h	substrate thickness
I_{ds}	drain current
I_d - V_d	drain current vs drain voltage
I_d - V_g	drain current vs gate voltage
I_g	gate current
I_g - V_g	gate current vs gate voltage
i_d	drain noise current
i_g	gate noise current

k_B	Boltzmann constant
K	stability factor
L	inductance
l	Lengths of transmission lines
L_g	gate length
N	number of fingers, number of turns, Noise power
NF	Noise figure
NF_{min}	Minimum noise figure
P	power
P_{1dB}	power at 1-dB gain compression
P_{dBm}	power in dBm
P_{dc}	DC power
P_{in}	power input
P_R	received power
P_G	generated power
Q	quality factor
q	electronic charge
R	resistance
R_c	contact resistance
R_d	drain resistance
R_{ds}	output resistance
R_g	gate-resistance
R_n	equivalent noise resistance
R_{poly}	polysilicon resistance
R_s	source resistance
R_{sh}	sheet resistance
S	spacing
$S_{i,t}$	current thermal power spectra density
$S_{v,t}$	voltage thermal power spectra density
T	temperature
t	thickness
tan δ, tanD	Loss tangent of dielectric material
U	Mason's uniliteral gain
V_b	bulk voltage
V_{ds}	drain voltage
V_{gs}	gate voltage
V_s	source voltage
V_{th}	threshold voltage
W	Tungsten

W	width
W_e	electric energy
W_g	gate width
W_m	magnetic energy
Y	Admittances
Y_0	characteristic admittance
Y_c	correlation admittance
Y_{in}	Input admittance
Y_{opt}	optimum noise admittance
Y_s	source admittance
Z	Impedance
Z_0	Characteristic impedance
Z_{in}	Input impedance
Z_L	Load impedance
Z_S	Source impedance
α	attenuation constant
β	phase constant
γ	propagation constant
Γ_L	load reflection coefficient
Γ_S	source reflection coefficient
Γ_{opt}	optimum noise reflection coefficient
δ	skin depth
ϵ_0	free-space permittivity
ϵ_r	permittivity/dielectric constant
ϵ_{re}	Effective permittivity/dielectric constant
θ	Angle, electrical length
λ	Wavelength
μ	Permeability
μ_r	relative permeability
μ_0	free-space permeability
ρ	resistivity
τ	transit time
υ_p	phase velocity
ω	angular frequency

Publications

1. **S. Sharabi**, I. McGregor and E. Wasige, "Extension of 0.18μm CMOS Operating Range to mm-wave Regime via Layout Optimisation", *submitted to IEEE Electron Device Letters.*

2. **S. Sharabi** and E. Wasige, "Optimised Slow-Wave Transmission Line with Improved Quality Factor Implemented on Low Resistivity Silicon CMOS Process", *submitted to IEEE Microwave and Wireless Components Letters.*

3. S. A. O. Russell, **S. Sharabi**, A. Tallaire and D. A. J. Moran, "Hydrogen-Terminated Diamond Field Effect Transistors with Cut-Off Frequency of 53 GHz", IEEE Electron Device Letters vol. 33, no. 10, pp. 1471-1473, October. 2012.

4. S. A. O. Russell, **S. Sharabi**, A. Tallaire and D. A. J. Moran, "Diamond Field Effect Transistors", ARMMS RF & Microwave Society April 2012 Meeting, Oxfordshire, UK. 23rd-24th April. 2012.

5. S. A. O. Russell, **S. Sharabi**, A. Tallaire, H. McLelland and D. A. J. Moran, "Scaling of Diamond FET RF Performance, Including a Cut-Off Frequency of 53 GHz", 11th NPL Millimetre-wave Users Group Meeting, Luton, UK. 27th March. 2012.

6. S. A. O. Russell, **S. Sharabi**, D. Macfarlane, R. Caterino, J. Garrido and D. A. J. Moran, "A Comparison Between De-Embedding Strategies for the Extraction of the RF Performance of Hydrogen-Terminated Diamond FETs", Hasselt Diamond Workshop SBDD XVII, Hasselt, Belgium. 14th-16th March. 2012.

7. Moran, D. A J; Russell, S.A.O.; **Sharabi, S.**; Tallaire, A., "High frequency hydrogen-terminated diamond field effect transistor technology", 12th IEEE Nanotechnology Conference (IEEE-NANO), pp. 1-5, August. 2012.

8. S. Taking, A. Z. Khokhar, D. MacFarlane, **S. Sharabi**, A. M. Dabiran, E. Wasige, "New process for low sheet and ohmic contact resistance of AIN/GaN MOS-HEMTs", Microwave Integrated Circuits Conference (EuMIC), pp.306-309, 27-28 Sept. 2010.

Chapter 1

Introduction and Motivation

"Stay connected" is the motto of today's information technology age. Information exchange, in the form of data, speech, text, images, video and audio, is increasingly common in all parts of our daily life. Various communication technologies such as telephony, optical and wireless networks, internet and satellite have revolutionised the way we live. Such technologies have increased interaction and enabled easier connection, bringing people together and aiding the sharing of ideas, hopes, and fears.

The wireless technology market today features important existing and emerging applications such as mobile telephony, WLAN, radar and remote sensing. All these applications require faster speeds and higher data rates. This ever-increasing demand for a faster, larger bandwidth, wireless spectrum has resulted in a growing need for researchers to exploit higher frequencies, such as the upper microwave and millimetre-wave bands[1], to expand the available spectrum to meet these demands.

The most widely known development in the application of millimetre-wave technology occurred around the time of the second World War when the Massachusetts Institute of Technology (MIT) Radiation Laboratory developed millimetre-wave radar technology. Bell Telephone Laboratories had also explored millimetre-wave technology for conventional radio communication. Atmospheric attenuation due to oxygen, water vapour and rainfall, at these frequencies, was a serious disadvantage for long distance radio transmission. In 1949 R. Beinger conducted an early measurement clarifying the oxygen absorption near 60 GHz (5 mm wavelength) [2]. Since then, more suitable applications

[1] Millimetrewave frequencies range between 30GHz and 300GHz.

for microwave and millimetre-wave frequencies, such as automotive radar, medical imaging and optoelectronics have been identified. Figure 1.1 summarises a range of some of the applications in the millimetre-wave region today. These applications have been traditionally implemented using expensive compound semiconductor technologies such as Gallium Arsenide (GaAs) and Indium Phosphide (InP).

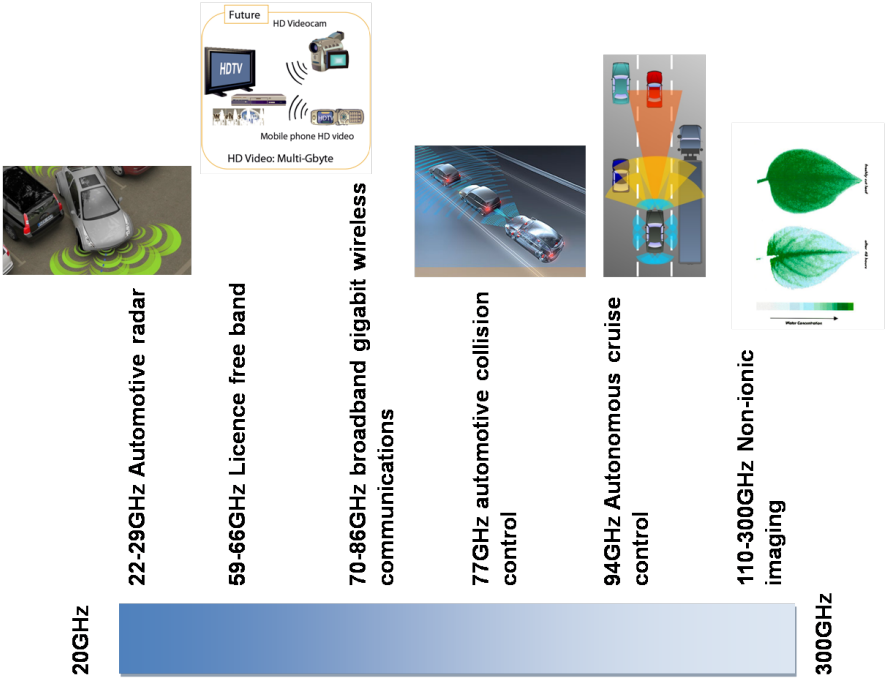


Figure 1.1: Range of some of the applications in the millimetre-wave region.

The main parameters that influence the choice of device technology for any microwave and millimetre-wave system are cost, performance, size, power consumption and time to market, in addition to high-density integration capability. The latter is very desirable since it enables smaller and faster systems that are cost effective and efficient. Modern CMOS technologies have emerged as a strong candidate for many commercial millimetre-wave applications. Silicon technology has been advancing since the invention of the silicon transistor in 1954, shortly after the invention of the silicon germanium transistor (SiGe) in 1947 at Bell Labs. The continued development of the transistor and semiconductor technology in the 1960's opened a new realm of opportunities for signal processing. Signals could be generated using solid-state circuits through the development of on-chip oscillators and phase-locked loops. Extensive analog functions can now be realised on-chip, as well as high speed digital functions. The first commercial metal oxide semiconductor (MOS) transistors were built with minimum feature size of

approximately 25 μm with low cut-off frequency (f_t). Using such a transistor as an amplification stage, at high GHz frequencies, was impossible as the operation frequency can only be a fraction of f_t. Recent developments in radio frequency (RF) CMOS technology have improved the technology node to 45 nm. Such sub-100 nm technology allows f_t of 150 GHz and beyond. This enables the design of microwave and millimetre-wave CMOS circuits.

1.1 Why Silicon CMOS?

There are three major semiconductor technologies for integrated circuit fabrication: silicon (Si), germanium (Ge) and III-V technologies such as GaAs and InP. From these three, silicon has been the mainstream semiconductor substrate material for integrated circuit (IC) circuits, mainly due to the following [3, 4]

- Silicon can be grown in very large scale single crystals (200 mm − 300 mm in silicon compared to 100 mm in GaAs) yielding many low-cost integrated circuits per wafer.
- Silicon has a good thermal conductivity (1.31 W/cm.°C compared to 0.55 W/cm.°C for GaAs and 0.68 W/cm.°C for InP) allowing efficient removal of dissipated heat.
- Silicon allows a controllable high dynamic range of doping concentration for both n-type and p-type impurities.
- Silicon has a good mechanical strength, facilitating ease of handling and fabrication.
- It is easy to make very low resistance ohmic contacts to silicon, thus minimising device parasitics.
- Silicon is abundant on earth in the form of sand and can be easily purified.

The continuous down scaling of the CMOS metal-oxide-semiconductor field-effect-transistor (MOSFET) has helped dominate the integrated circuit industry. The circuit performance has improved due to the lower power dissipation, faster speed and the reduced circuit size. However, when it comes to high frequency integrated circuit design silicon is not the prime candidate. This field has been dominated by III-V technologies for many

years. Silicon MOSFET performance is still substandard in comparison to III-V technologies (GaAs pseudomorphic high electron mobility transistor (pHEMT), InP high electron mobility transistor (HEMT), InP heterojunction bipolar transistor (HBT)) due to the following disadvantages:

1. **Lower carrier mobility**: Silicon carrier mobility is much lower than that of a III-V material, which could potentially introduce limitations to the performance of the device. While the f_t of today's CMOS technology is well under 200 GHz, InP devices with an f_t of approximately 400 GHz and an $f_{\max}$ of over 1 THz have been reported [5, 6].

2. **High sheet resistance of the polysilicon gates**: The gate material used for CMOS devices is polysilicon, which has a much higher sheet resistance ($\sim 10\,\Omega$/cm) than the metal used for the gates of GaAs field-effect-transistors (FETs). A higher gate resistance can reduce the transistor power gain and increase the noise figure. Fortunately, simple layout techniques can be used to minimise the effects of the polysilicon gate as will be described in this work.

3. **Low resistivity substrate**: The substrate resistivity of III-V material (GaAs substrate resistivity $\sim 10^7 - 10^9\,\Omega - \text{cm}$) [7] is significantly higher than silicon ($10 - 15\,\Omega - \text{cm}$). Signals that couple to the low resistivity silicon substrate suffer substantial losses, especially at microwave and millimetre-wave frequencies. This makes designing silicon integrated passive components extremely challenging.

4. **Complex device modelling**: GaAs FET can effectively be treated as a three terminal device. The existence of the bulk terminal and the body-effect complicate matters for silicon MOSFET modelling and design.

Despite all the advantages of III-V technology, the interest in silicon for commercial microwave and millimetre-wave systems is increasing. Commercial applications outlined in the previous section are cost sensitive and need to benefit from the economy of scale. The worldwide manufacturing capacity of silicon technologies, and its superior yield compared to its III-V rivals, can reduce the costs as required for commercial applications. In recent years advances in silicon technology, mainly driven by digital applications, has provided the opportunity for silicon devices to perform well in the high GHz frequency bands. In particular, the performance of CMOS technology, quantified

by the cut-off frequency (f_t), maximum oscillation frequency (f_{max}) and the minimum noise figure (NF_{min}), has significantly improved with geometry scaling and other relevant technology enhancements. At the 45 nm technology node[2], CMOS preforms with a $f_t = 280$ GHz, f_{max}=550 GHz and a NF_{min} of about 1 dB @ 30 GHz making it a good candidate for millimetre-wave systems. Apart from cost benefits, silicon presents a significant advantage over III-Vs in terms of yield and level of integration. This enables the implementation of entire transceiver topologies on a single die. These topologies are key to many commercial millimetre-wave systems such as car radar systems.

CMOS has a cost advantage over other silicon technologies such as bipolar complementary metal oxide semiconductor (BiCMOS) and silicon on insulator (SOI) technologies. CMOS offers the same level of performance as BiCMOS with half the geometry length. The routing metal density for BiCMOS, compared to an equivalent performance as CMOS chip, is however considerably lower. Even if the cost per area for the two technologies were equal, this difference in the metal density will directly increase the area of the digital portion of the system, resulting in a significant increase in the cost [8].

Silicon's real strength lies in its integration capability. In fact, BiCMOS, silicon germanium (SiGe), and possibly CMOS are the only feasible system on chip (SoC) solutions for communication systems. GaAs and InP lag far behind silicon technologies in device density and yield. Very large scale integration (VLSI) on III-V semiconductors is still far from reality. In Table 1.1, a comparison of the characteristics between silicon and

[2]Technology node for silicon is defined by the minimum gate channel length.

Table 1.1: Performance characteristics of silicon and III-V technologies.

Characteristic	Silicon			III-V		
	CMOS 0.18 μm	SiGe BiCMOS 0.18 μm	SOI CMOS 0.18 μm	GaAs HBT Power device	InP HBT 1 μm	InP HEMT 0.15 μm
f_t (GHz)	45	120	60	25	160	130
f_{max} (GHz)	60	100	75	50	200	185
NF_{min} (dB)	0.7@2 GHz	0.4@2 GHz	<1@2 GHz	N/A	N/A	<1@2 GHz
Substrate resistance (Ω-cm)	10	10	10-2000	1M	1M	1M
Cost	Low	Medium	Medium	Medium	High	High
Integration capability (SoC)	Good	Good	Good	Difficult		

III-V technologies in terms of figures of merits, substrate resistance, integration capability and cost, are illustrated. Finally, silicon technologies have a large cost advantage. Lower substrate cost, higher yield, and much greater infrastructure investments are all responsible for the lower cost per chip area of silicon. More importantly, silicon SoC capability can dramatically reduce the total system cost, including design and fabrication, by eliminating external components and reducing the board area. In view of this silicon is the technology adopted in this work

1.2 CMOS Millimetre-wave Challenges

Despite the advantages, CMOS also has some serious challenges. These can be summarised as follows:

- Traditionally radio frequency integrated circuit (RFIC) have been implemented in standard computer-aided design (CAD) packages using circuit models available for basic active and passive components. Every silicon foundry develops its own compact models which enable accurate simulations. However, these models are generally basic models for digital logic, and even if the foundry design kit includes an RF library, it is normally intended for low RF frequencies and can not guarantee accuracy at higher frequencies [9].

- The lack of accurate circuit models for passive and active components presents one of the most challenging problems for silicon RFIC designers. As the frequency of operation increases, the physical modelling of passive structures becomes more difficult due to the complexity of radio frequency phenomena like the eddy current effects and the substrate losses in the CMOS process.

- Furthermore, the active device's maximum gain decreases when operating near the technology cut-off frequencies (f_t, f_{max}). At this point the substantial passive losses, and the low active devices gain, can reduce the chances of the circuit meeting the aimed specifications. Hence, RFIC designers must make an extensive effort to model and optimise passive and active devices at microwave and millimetre-wave frequencies, to ensure perfectly matched measured and modelled performance, therefore ensuring successful ICs.

1.3 Design Methods and Approach

As mentioned earlier, millimetre-wave circuits have, traditionally, been implemented using III-V technologies (GaAs, InP,..etc). For these technologies, active and passive devices models were developed by the foundries and made available for standard CAD packages. This enabled accurate design and simulations far beyond 100 GHz [10]. On the other hand, CMOS active and passive device models, operating in the upper microwave and millimetre-wave frequencies, are not readily available. In fact, foundry provided, CMOS models of active devices are limited to low RF frequencies[9]. Furthermore, models of passive devices such as (transmission lines, spiral inductors, capacitors or resistors) are not suitable for designing circuits operating at high frequencies because of the lossy nature of the silicon substrate, the thin dielectric layers residing above the substrate, and the metal density requirements of CMOS technologies. Therefore, a systematic approach is needed for developing models that take into account the extrinsic layout parasitics, distributed effects and the challenges of CMOS. In view of this, custom models, that take into account the extrinsic layout parasitics and are accurate for the operation regimes, will be developed. Additionally, custom designed transmission lines and inductors, that exhibit lower losses at higher frequencies, along with interdigitated (comb) capacitors, RF and direct current (DC) probe pads will be designed, fabricated, measured and modelled in this work. All these components form a high frequency library that can be used to design circuits operating at microwave and millimetre-wave frequencies with an improved degree of accuracy. Figure 1.2 describes a systematic project design flow chart that shall be used to obtain a successful circuit design cycle in CMOS.

The flow chart includes test structures on chip, measurements, electromagnetic simulations and parasitic extractions to capture the high frequency behaviour of transistors, transmission lines, spiral inductors and capacitors. This is followed by the circuit design of test LNA amplifiers. For verification and characterisation of the active and passive components, two test chips[3], illustrated in Figure 1.3, containing test structures were fabricated and measured. The test chips comprised of a variety of passive and active structures of different sizes and type such as RF NMOS transistors (conventional and

[3]Two test runs, 40 chips per run, 30 chips were measured and characterised from the first run and 28 chips were measured and characterised from the second run.

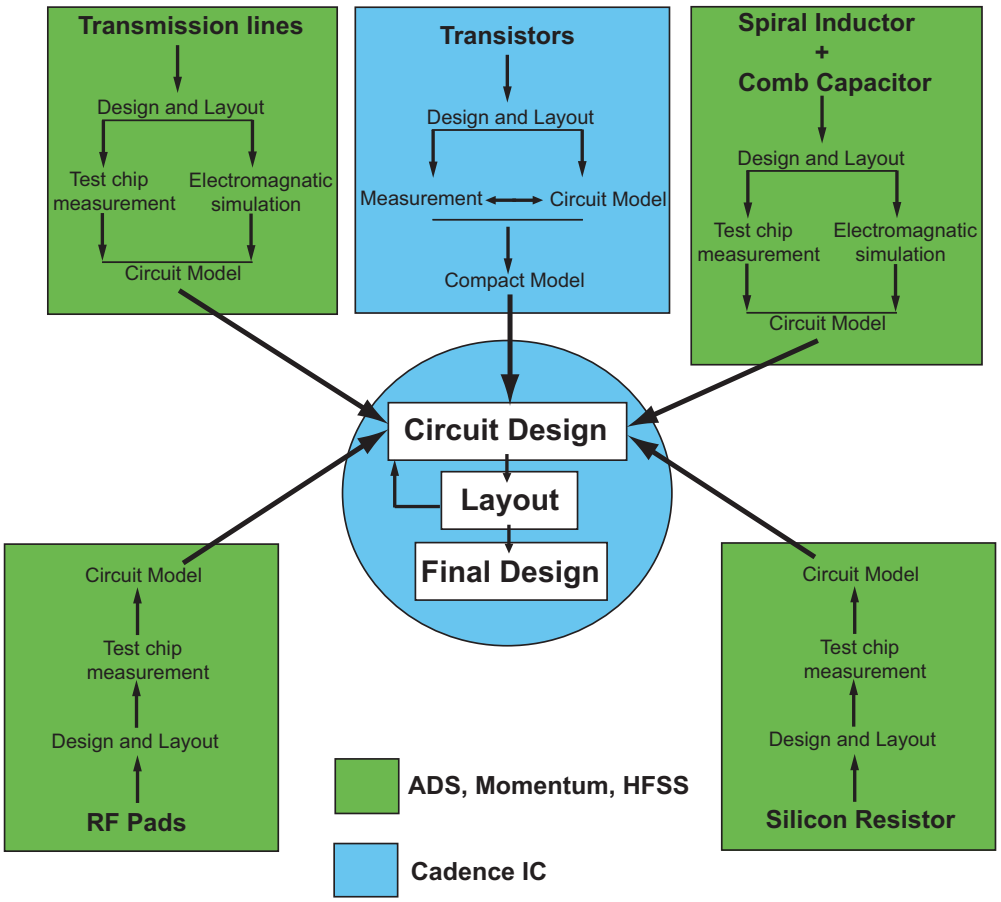


Figure 1.2: Design flow for a successful circuit design cycle in CMOS.

modified layout), Comb capacitors, spiral inductors, resistors, transmission lines, in addition to calibration and de-embedding structures. The following is a detailed list of fabricated structures:

1. Transistors
   (a) RF NMOS, Common source configuration, Gate width = **2** μm x 40 fingers.
   (b) RF NMOS, Common source configuration, Gate width = **4** μm x 40 fingers.
   (c) RF NMOS, Common source configuration, Gate width = **6** μm x 40 fingers.
   (d) RF NMOS, Common source configuration, Gate width = **8** μm x 40 fingers.
   (e) RF NMOS, Common source configuration, Gate width = 2 μm x **5 fingers**.
   (f) RF NMOS, Common source configuration, Gate width = 2 μm x **10 fingers**.
   (g) RF NMOS, Common source configuration, Gate width = 2 μm x **20 fingers**.
   (h) RF NMOS, Common source configuration, Gate width = 2 μm x **40 fingers**.
   (i) RF NMOS, Common source configuration, Gate width = 2 μm x **80 fingers**.

2. Transmission lines (Microstrip and CPW): unshielded, shielded (ground and floating) and Slow wave.

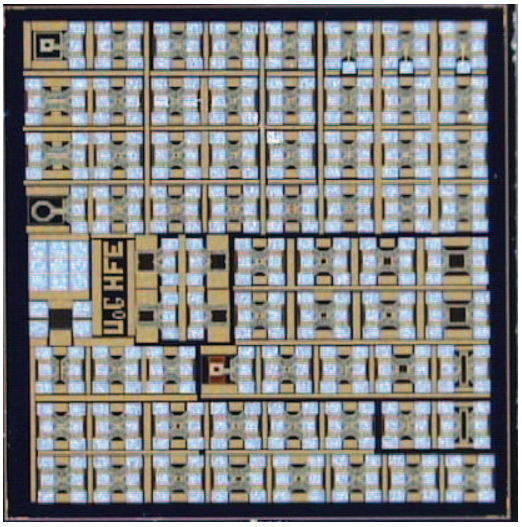
3. Multi-finger comb capacitors: single ended (supply-ground bypass capacitors) and two port (AC coupling capacitors).

4. Spiral inductors: shielded (ground and floating) and unshielded spirals.

5. 50 Ω poly-resistors.

6. On-wafer calibration and de-embedding structures.

(a) Test chip 1



(b) Test chip 2

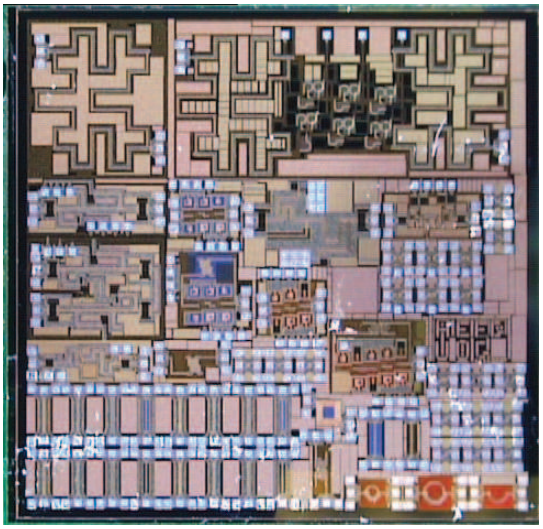


Figure 1.3: Test chips containing test structures such as: NMOS transistors, multi-finger comb capacitors, spiral inductors, 50 Ω resistors, transmission lines, in addition to calibration and de-embedding structures

Active and passive device model libraries were created using the measured and extracted data of the test structures. This enabled the design of CMOS circuits operating beyond the technology capabilities as will be shown in this work.

1.4 Thesis Aims and Outline

The work presented in this thesis has two main aims, the first is to optimise the performance of CMOS for millimetre-wave technology. To achieve this goal, millimetre-wave systems have been considered from both the device and circuit prospective. From the device point of view, a systematic way of making high performance active devices is

proposed and various device parameters such as power gain and noise performance are analysed and/or optimised. An extensive investigation of accurate modelling methods for transistors beyond 20 GHz is also performed. In relation to the circuit design side, several high performance, low power, low noise amplifiers were designed and implemented based on a comparative study highlighting several design topologies and techniques.

The second aim of this research was to investigate the limits of CMOS technology. More specifically, it attempts to answer the question: at any given technology node, what is the maximum frequency that one can design active circuits with acceptable performance and power dissipation? To this end, some techniques to maximise the $f_{\max}$ of a CMOS transistor are proposed and a low noise amplifier circuit was designed and implemented using these devices.

The organisation of this thesis is as follows: Chapter 2 introduces and discusses microwave and millimetre-wave metrology methods. Chapter 3 details the modelling and optimisation procedures of CMOS passive devices. Chapter 4 considers the thermal noise in microwave and millimetre-wave devices and circuits, detailing noise and small-signal modelling procedures of CMOS devices up to 50 GHz and 110 GHz respectively. Chapter 5 discusses the optimisation of the physical structure of the device in order to maximise performance. Chapter 6 outlines the design and implementation of several low noise amplifiers. Chapter 7 concludes the thesis and proposes some future research directions.

Chapter 2

High Frequency Characterisation and Metrology

Accurate microwave and millimetre-wave circuit design requires precise active and passive device characterisation. Appropriate procedures and methods of characterising device parameters, in addition to reliable measurements, are vital in order to determine the device functionality. Since the designs involve high frequency circuits, analysing the spectrum and power performance is also a necessity.

This chapter begins with an overview of the measurement equipment and setup, followed by a section explaining the design of CMOS device under test (DUT). This describes the challenges involved in measuring and de-embedding test fixtures used to characterise the small-signal and noise parameters of devices fabricated on silicon. Several conventional on-wafer small-signal de-embedding techniques are discussed and their shortcomings are identified, leading to proposed improved methodologies to overcome these shortcomings. The validity of the proposed techniques are then evaluated and compared against the conventional techniques. These de-embedding methods are presented and discussed in order to elucidate the robustness of these techniques for this work. Finally, the noise pad de-embedding methodology used for device noise characterisation is discussed.

2.1 Measurement Equipment

To verify the high frequency performance and the characteristics of circuits and devices fabricated on CMOS chips, precise measurement equipment is needed. The type of measurement varies depending on the circuit or device type, e.g. small-signal, power and noise measurements. The measurements are performed on-wafer using a probe station. The vector network analyzer (VNA) is one of the most important pieces of equipment and it is primarily used for low power scattering parameters (S-parameters) measurements. Every structure or circuit designed on chip requires terminal pads. Therefore, the probe station has access to these circuits by simply lowering the probes onto the terminal pads. The probes are a transition from coaxial to CPW and normally the tips are made of Tungsten (W) or Beryllium copper (BeCu). Figure 2.1 depicts an on-wafer measurement setup illustrating the CPW probes.

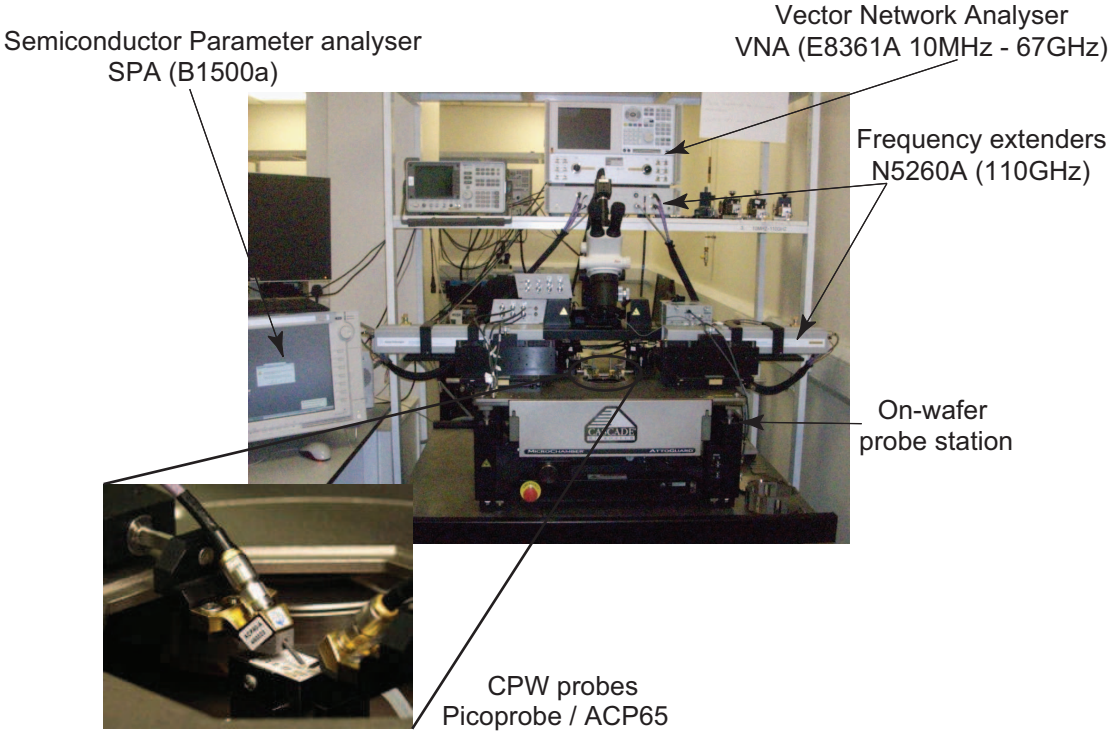


Figure 2.1: On-wafer probe station setup.

The VNA can be calibrated using one of the following established techniques: short-open-load-thru (SOLT), line-reflect-match (LRM) and thru-reflect-line (TRL). Calibration is required to define the reference point where the measurement data should be recorded. A detailed explanation is provided in Section 2.1.2.

Power measurements and characterisation can be cumbersome as it requires a frequency sweep while simultaneously varying the power levels. The measurement setup primarily consists of a frequency synthesiser (Figure 2.2(a)), a spectrum analyser (Figure 2.2(b)) and a power meter with a power sensor. The power sensor converts the RF and microwave signals to analog voltages, the latter are interpreted by the power meter into a displayed readable value that represents the power of the signal. The accuracy of the measurements can be affected by cables and probe losses. Hence, accurately calibrated cable lengths are vital in this type of measurement.

Noise measurements can be performed using a dedicated noise figure analyser and a calibrated noise source. However, in the absence of this piece of equipment a spectrum analyser could be utilised.

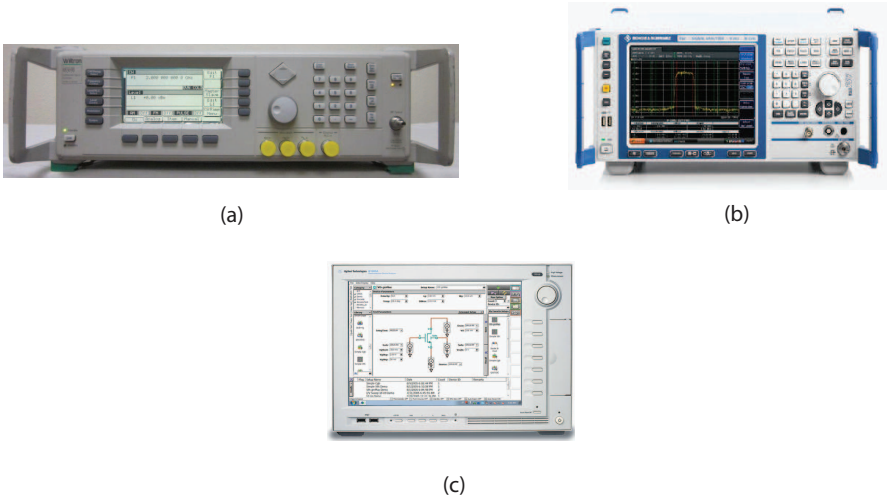


Figure 2.2: Characterisation and metrology equipment: (a) A frequency synthesiser (up to 60 GHz), (b) a spectrum analyser and (c) a semiconductor parameter analyser.

2.1.1 Spectrum Analyser

A spectrum analyser is an instrument that measures the power spectra of known and unknown signals. Spectrum analysis makes detection of dominant frequency, harmonics, distortion, bandwidth, power and other spectral components, much easier in comparison to time domain waveforms analysis [11].

Figure 2.3 depicts a simplified block diagram of a spectrum analyser. The input signal is fed into a mixer through an attenuator block, which controls the sensitivity and the level of the input signal to avoid mixer overload and distortion. If the input signal

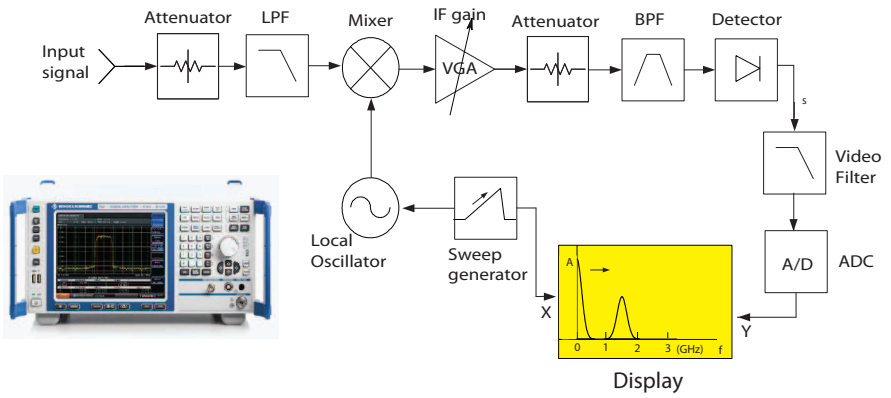


Figure 2.3: Simplified diagram of a spectrum analyser.

is weak, an external low noise amplifier can be used to boost it. The input signal is mixed with a swept local oscillator (LO) signal. The low pass filter block rejects the upper image frequency and removes any unwanted intermediate frequency (IF) signals. Spectrum analysers typically have a narrow IF bandwidth ($\sim 10\,\text{MHz}$) and are suited to narrow band signals. The IF is further filtered by a resolution bandwidth filter in order to remove any intrinsic noise embedded in the signal, which is then amplified or compressed before being detected by a power detector. Finally, the measured power or amplitude of the signal of interest is transferred to the display through an analog to digital converter (ADC).

State of the art spectrum analysers have wide operating bands, for example 3 Hz to 40 GHz for the FSV40 Rohde&Schwarz and $3\,\text{Hz} - 50\,\text{GHz}$ for the E4448A from Agilent Technologies. The frequency spectrum can be extended beyond the maximum internal frequency capability with an external waveguide subharmonic mixer, such as V-band ($50\,\text{GHz} - 75\,\text{GHz}$) or W-band ($75\,\text{GHz} - 110\,\text{GHz}$) mixers. Spectrum analysers can perform multiple measurements, for example evaluation of a signal spectrum and identification of unknown spectrum components, channel power or leakage, noise figure, phase noise and intermodulation measurements.

2.1.2 Vector Network Analyser

A VNA measures S-parameter vectors; the magnitude and phase, of all four S-parameters of a two-port networks. Figure 2.4 illustrates a block diagram of a VNA system. It can be seen that the VNA has two internal signal sources, RF and LO, that can sweep over the measurement frequencies of interest. Separate external sources may be needed for

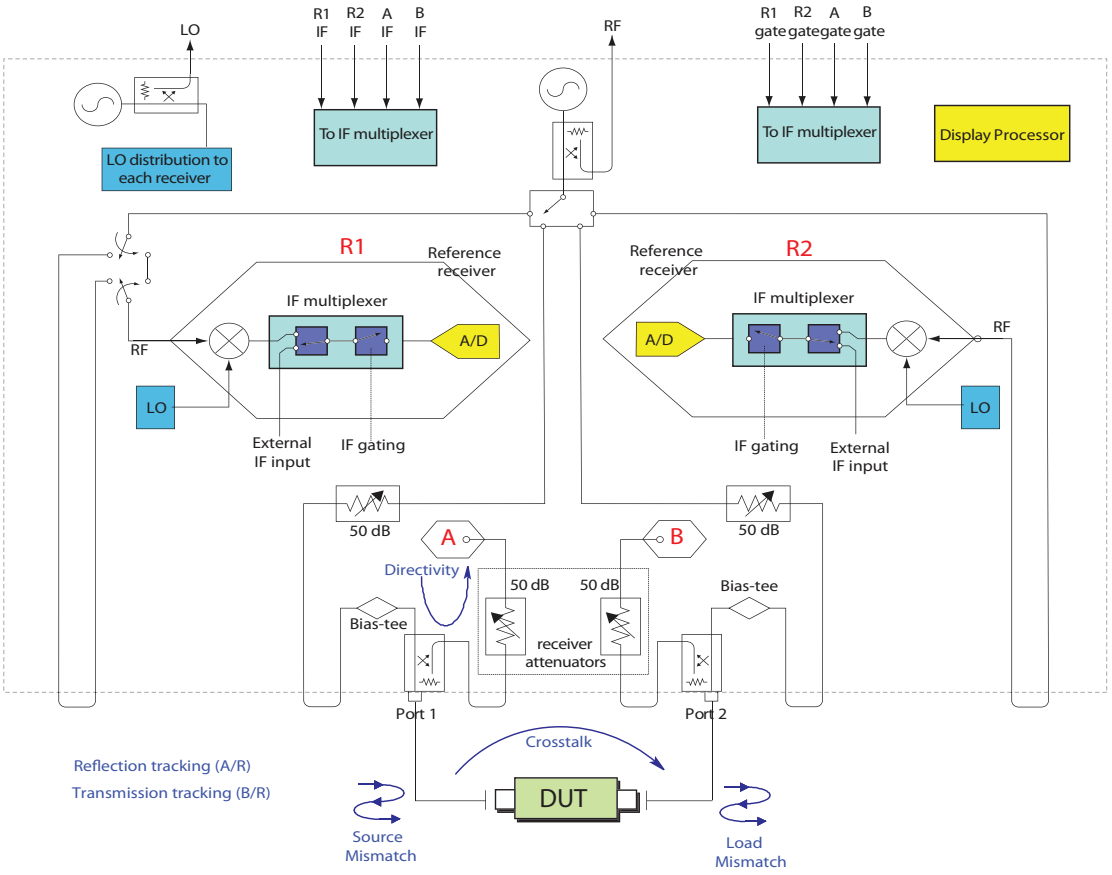


Figure 2.4: Vector network analyser test set block diagram with an illustration of one way, 6 systematic error terms.

operation frequencies beyond the VNA. The RF and LO are commonly set to be equal or slightly different (about 10 kHz difference), for measuring a linear network. This is governed by the setup of the IF bandpass filter or the IF bandwidth defined by the user, and is in the range of several tens of Hz to several hundreds of kHz depending on the application [11]. An actual system will have two or more stages of down-conversion mixing for IF to achieve good accuracy. Each test port has a directional coupler for separating the RF signals sent out to the DUT and the reflected signals from the DUT. The directivity of the coupler limits the minimum detected power of the reflected signals and therefore the return loss of the DUT.

Taking a two-port network measurement as a practical example, when a one-port reflection measurement (port 1) is in operation, the VNA terminates port 2 with a broadband matched load and sets its switch to the forward mode. The RF signal is split into two halves in terms of power. One goes into the mixer where it is mixed down by the LO to the IF as a reference, the other goes into the DUT through the directional coupler.

The reflected signal returns via the directional coupler to another mixer and is mixed with the LO to generate a test IF. The measured reflection coefficient, S_{11}, of the DUT is derived from the measured phase and magnitude difference of the test and reference IF signals. Similarly, the reflection coefficient of port 2, S_{22}, can be measured.

When measuring the transmission coefficients of a two-port network, the RF signal source still sends a signal to the DUT via the directional coupler while it is in forward mode, the signal will pass through the DUT and is separated by the directional coupler at test port 2 of the VNA. It is then down-converted to an IF signal as a test result. The ratio of the power and phase of the measured signals gives the forward transmission coefficient, or S_{21}, of the DUT. The reverse mode leads to the derivation of S_{12}. The final S-parameters of the DUT are then constructed by combining the two measured individual one-port reflection coefficients, S_{11} and S_{22}, and the two transmission coefficients, S_{21} and S_{12}.

Some VNAs can also provide additional functionality, such as power sweep and noise measurements. The output power of a VNA can be varied within a certain range, at a fixed frequency. This enables power characterisation and evaluation of a nonlinear network, for example the 1 dB compression point of amplifiers.

2.1.2.1 System Error and Calibration

There are three basic sources of measurement error associated with the VNA:

- **Systematic errors**: These are due to imperfections in the analyser and test setup. They are repeatable (therefore predictable), and assumed to be time invariant. Systematic errors are characterised during the calibration process and mathematically removed during measurements.
- **Random errors**: These are unpredictable since they vary with time in a random fashion. They therefore cannot be removed by calibration. The main contributors to random error are instrument noise.
- **Drift error**: Is primarily caused by temperature variation and it can be removed by further calibration. The time frame over which a calibration remains accurate is dependent on the rate of drift that the test system undergoes in the user's test

environment. Providing a stable ambient temperature helps minimise the drift error.

Figure 2.5 depicts the major systematic errors associated with network measurements. The errors relating to signal leakage are directivity and crosstalk (leakage). Errors related to signal reflections are source and load match. The final class of errors are related to the frequency response of the receivers, and are called reflection and transmission tracking. The full two-port error model includes all six of these terms for the forward direction and the same six (with different data) in the reverse direction, for a total of twelve error terms. This is why two-port calibration is often referred to as twelve-term error correction [12].

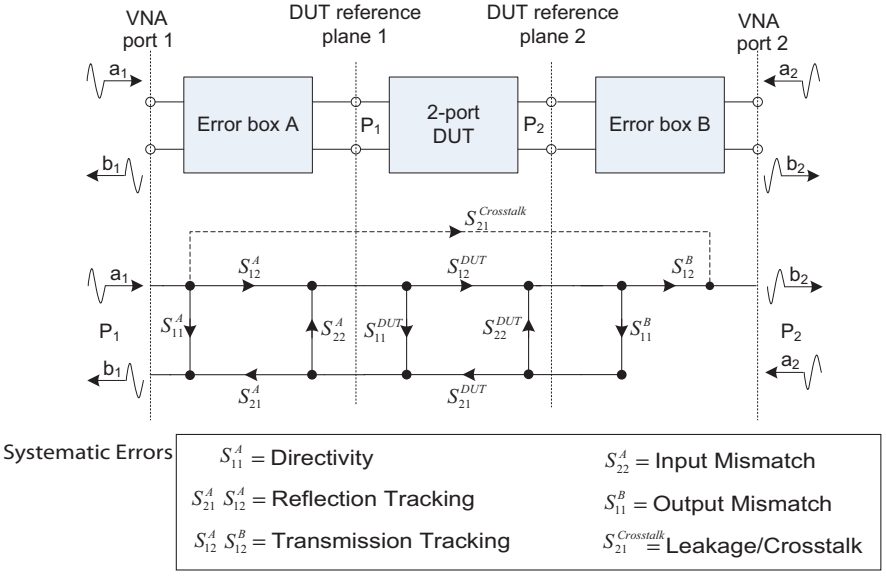


Figure 2.5: Block diagram of the system errors and a forward model of the 12-term error model for a two-port vector network analyser, and its signal flow graph representation with systematic errors.

2.1.2.2 Calibration Options

Figure 2.6 illustrates the basic calibration techniques available for network analysis measurements. Although the VNA can perform S-parameter measurement without any calibration, the measurement would be very inaccurate as any errors present are not corrected. Response calibration is very simple, it only requires a thru standard with

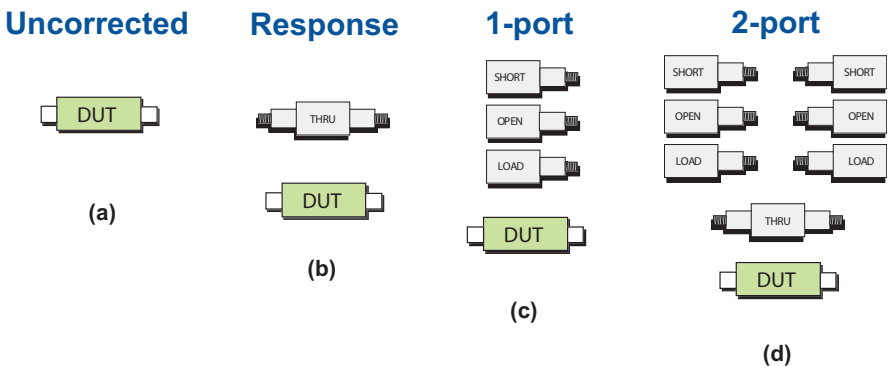


Figure 2.6: Calibration standards: (a) Uncorrected, (b) response, (c) 1-port, and (d) 2-port.

a known delay. This calibration option removes the frequency response errors (reflection and transmission errors)[1], and is often employed in low frequency measurements when accuracy is not important. The full 2-port calibration option is the most accurate method of all. It removes all the systematic errors (12 errors); it requires an open, short, broadband load standards and a thru standard with a known delay time. These are often referred to as mechanical calibration standards.

2.1.2.3 On-Wafer Calibration

Accurate well known calibration techniques for on-wafer measurements, that define the measurement reference plane at the probe tip, are widely available. On-wafer calibration standards are most often fabricated either on an alumina substrate, known as impedance standard substrate (ISS), or on the wafer containing the DUT. The calibration standards are: thin-film resistors (load), short-circuit connections (short) and 50 Ω transmission lines (thru). Figure 2.7 illustrates an ISS, provided by Cascade Microtech ®, fabricated on alumina. The inset shows the load, short, thru and alignment standards. The open standard is also available on substrate. However, it is normally performed by raising the probes in the air above the wafer.

The correct probe placement is very important to ensure accurate and repeatable calibration for on-wafer millimetre-wave measurements. Inconsistent probe placement and overdrive, can change the pad inductance and contact resistance resulting in errors [13].

[1]Response is a vector magnitude and phase normalisation of a transmission or reflection measurements.

Figure 2.8 shows the correct probe tips alignment and placement using the alignment markers on the ISS.

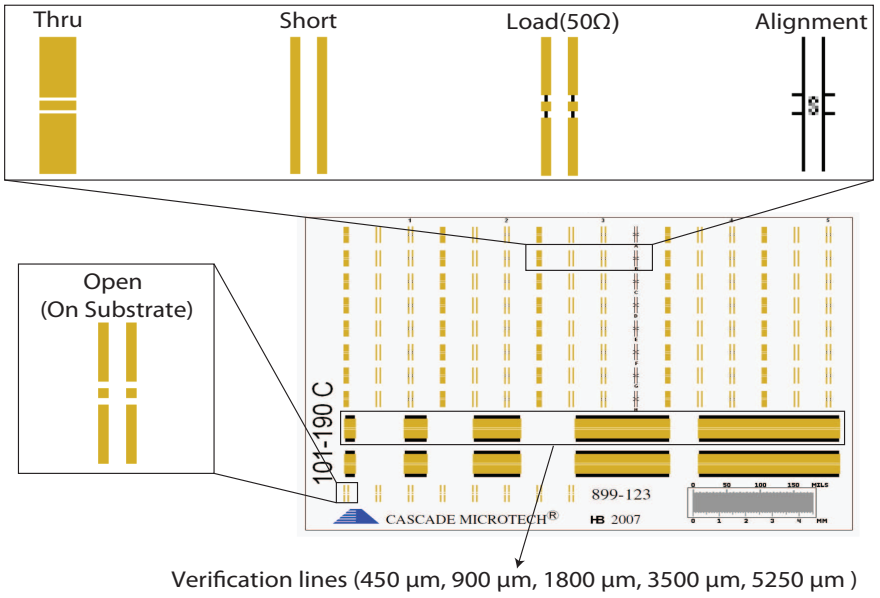


Figure 2.7: The impedance standard substrate (ISS) 101-190 fabricated by Cascade Microtech featuring the short, 50 Ω load, thru, on substrate open standards, alignments markers and verification lines of multiple length.

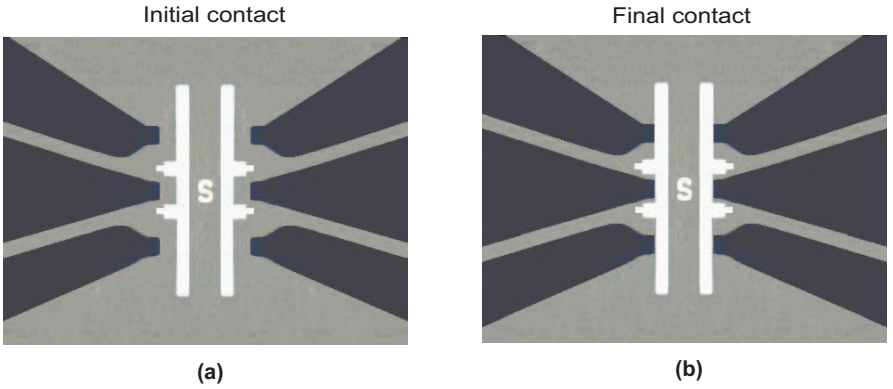


Figure 2.8: Correct alignment and placement of probe tips using the ISS Alignment markers: (a) Initial contact with ISS and (b) probe skating forward ($50\,\mu\text{m} - 250\,\mu\text{m}$).

On wafer calibration can be performed using the following VNA calibration techniques:

- **SOLT (short-open-load-thru)**:

  The most common and the simplest method amongst all the mentioned calibration methods. It requires a short, open, 50 Ω load (often laser-trimmed) and a 50 Ω transmission line (thru) with a known electrical length, as shown in Figure 2.9(a). This technique is suitable up to millimetre-wave frequencies. The open, short and load measurements will correct the forward and reverse directivity, reflection

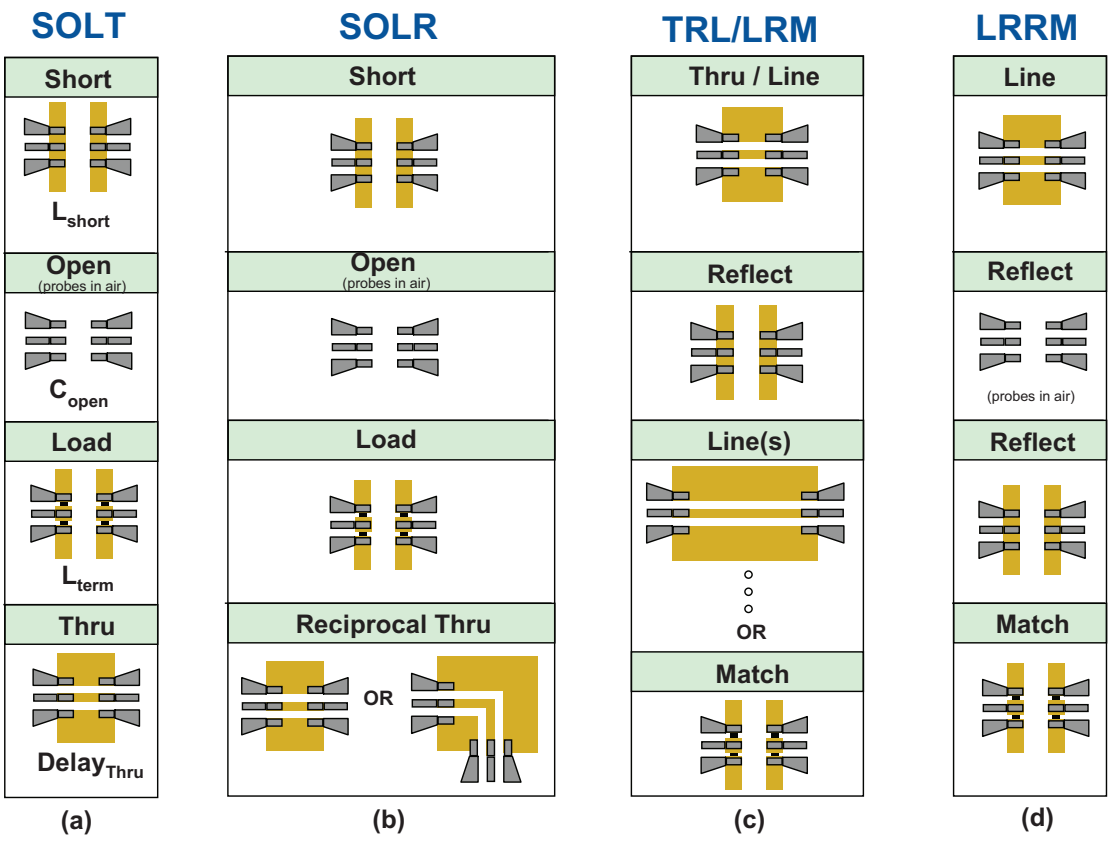


Figure 2.9: 2-port calibration techniques: (a) SOLT (short-open-load-thru), (b) SOLR (short-open-load-reciprocal), (c) TRL (thru-reflect-line) and LRM (line-reflect-match) and (d) LRRM (line-reflect-reflect-match).

tracking, isolation and port matching errors. On the other hand, the thru will correct the transmission reflection [14]. The SOLT standards are reasonably well modelled with simple lumped elements; open-circuit capacitance (C_{open})[2], load inductance (L_{load}), short-circuit inductance (L_{short}) and thru delay (normally 1 ps for 101-190 ISS).

- **SOLR (short-open-load-reciprocal)**:

  This is almost identical to SOLT calibration with only one difference. This technique, shown in Figure 2.9(b), assumes the thru standard is reciprocal and less than 180° [15, 16]. The advantage of this calibration method is that the thru standard does not require detailed information as long as it is reciprocal. This is particularly useful when two devices are far apart [17] and when calibrating a VNA with orthogonal probe positions [12].

- **TRL (thru-reflect-line(s))**:

[2]The open circuit may exhibit a negative capacitance since the probe lifted in the air has less tip loading than when it is in contact with a wafer.

Features much simpler standards compared to SOLT and short-open-load-reciprocal (SOLR). As shown in Figure 2.9(c), it only requires a thru standard, a short or an open standard, and one or more transmission line standards with electrical length shorter than 180° (lines should be between 20° and 160°).

- **LRM (line-reflect-match) and LRRM (line-reflect-reflect-match)**:

  These are shown in Figure 2.9(c) and (d). On-wafer LRM method requires a known electrical length line, two reflects standards either open or short, and two imperfectly matched loads [18]. The advantages of the LRM method is that it does not require prior details about the reflect standards. line-reflect-reflect-match (LRRM) is an improved version of LRM [19, 20]. It requires two additional undefined shorts and only one matched load should be measured at either port. This calibration method minimises the inaccuracies caused by probing misalignments with short and matched load standards [21]. However, it is unable to define the reference impedance.

2.2 Measurement Setups

In this section, the setup required for active and passive devices measurements including: DC, small-signal RF and noise measurements are described.

2.2.1 DC Measurements as a Prerequisite for RF Measurements

Active devices characterisation always begins with a DC performance check. The instrument used to perform the DC measurements is the B1500a semiconductor parametric analyser with 4 source/monitor units (SMUs), shown in Figure 2.1(c). The I-V characteristics of the transistors are measured as a function of gate source voltage (V_{gs}) and drain source voltage (V_{ds}). Each voltage sweep is asserted through a designated SMU. Figure 2.10(a) shows the I_d-V_d characteristics for six different V_{gs} values. It can be observed that in saturation a very large output resistance (R_{ds}) would be expected because of a very small increase in I_{ds} over a wide range of V_{ds}. Figure 2.10(b) shows the I_d-V_g characteristics of a transistor in the linear region with $V_{ds} = 0.05\,\mathrm{V}$. The gate-voltage axis intercept of the linear extrapolation of the I_dV_g curve at the point

of maximum transconductance, g_m, provides the extrapolated value of the threshold voltage, V_{th} [22].

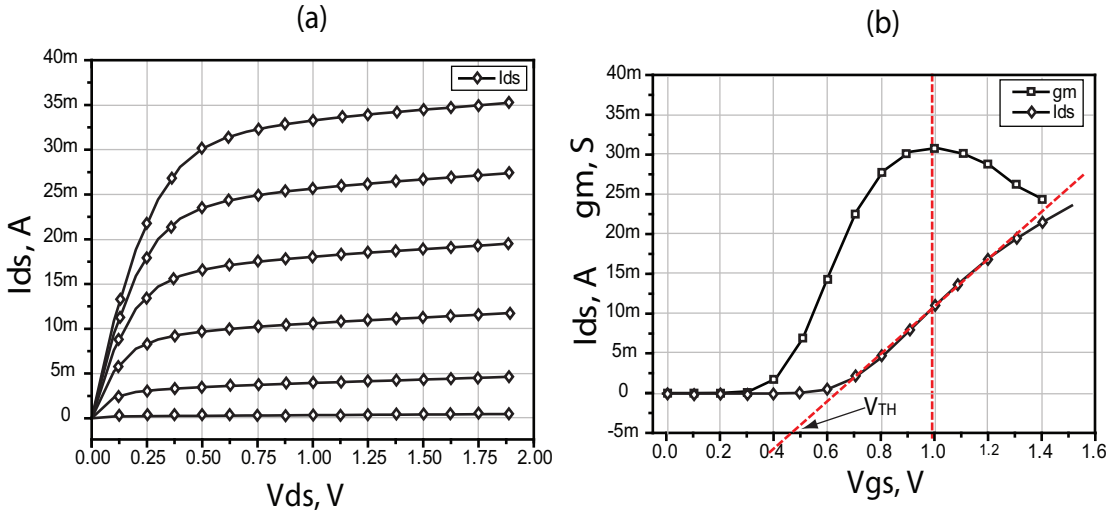


Figure 2.10: NMOS transistor DC measurements, device dimensions 80 μm (2 μm per finger): (a) I_d-V_d measurement plot and (b) I_d-V_g measurement plot with V_{th} extraction.

2.2.2 Small Signal High-Frequency Measurements

On-wafer S-parameter measurements up to 110 GHz were performed using a Cascade Microtech semi-auto probe station, ACP (65 GHz) or Microprobe (110 GHz), ground-signal-ground (GSG) coplanar waveguide probes, and an Agilent E8361A PNA with N5260A frequency extender. Figure 2.11 shows the small-signal measurement setup for 65 GHz and 110 GHz. The effects of the pads were de-embedded when measuring individual devices (transistors, transmission lines, inductors, capacitors etc). Pad removal is not necessary for the designed circuits such as amplifiers and oscillators as the pads are included in the design.

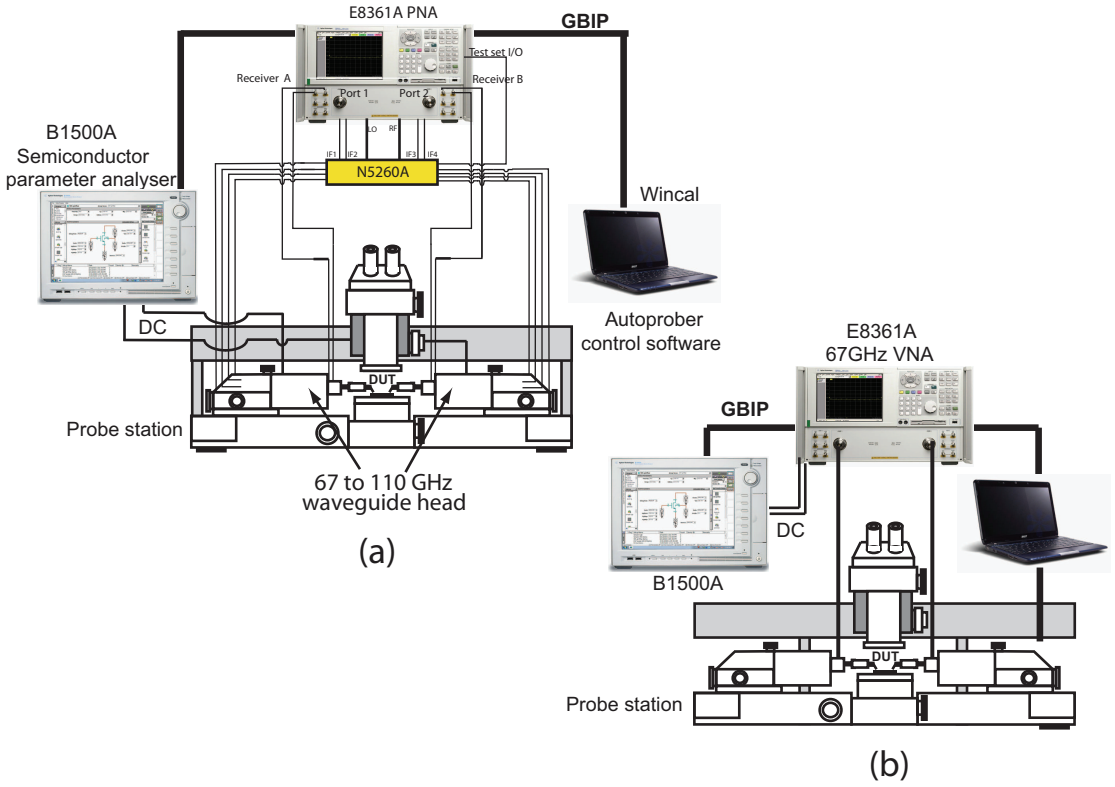


Figure 2.11: Small signal RF measurement setup: (a) 110 GHz measurement setup and (b) 65 GHz measurement setup.

2.2.3 High-Frequency Noise Measurements

The noise measurement setup varies depending on the purpose of the measurement, for example, the setup used to extract the noise parameters of a transistor is slightly different to that used to obtain the noise figure and gain of a low noise amplifier or any other circuit block. The difference is in the complexity and the time required to perform the measurements. The setups shown in Figure 2.12(a) and (b) depict the Y-factor and the gain methods respectively. These setups perform identical noise figure measurements. Figure 2.12(a) depicts the Y-factor setup [23, 24]. This setup requires a solid state noise source, that can operate in the desired frequency band (Agilent 346C Noise Source, 10 MHz to 26.5 GHz). However in the absence of such a source, or the need for higher frequencies, a frequency synthesiser with a variable output power could be used. For example the 60 GHz Wiltron 68187B synthesiser, with variable output power from −20 dBm to +20 dBm. This is illustrated in Figure 2.12(b) and referred to as the gain method [25]. The output of the frequency synthesiser is attenuated externally to achieve the low power needed for the noise measurement (≤-70dBm). This power is

calibrated using a HP8487A 50 GHz power sensor with a HP837B power meter, before performing any measurement.

The setup shown in Figure 2.12(c) is commonly known as source pull and it is based on a multiple source impedance technique proposed by Uhlir and Adamian to extract the noise parameters (NF_{min}, Γ_{opt}, R_n) [26, 27]. When characterising the noise performance

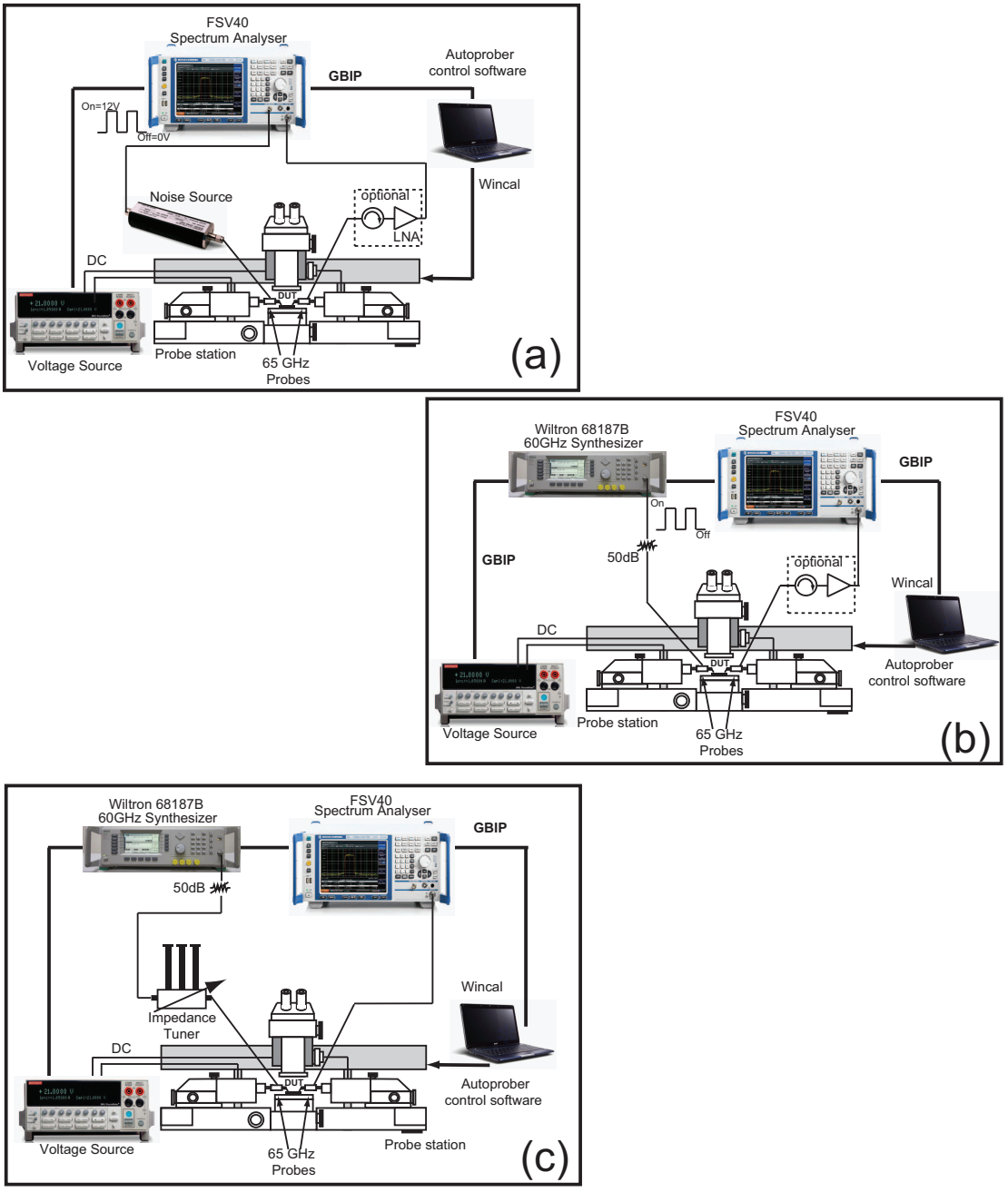


Figure 2.12: Noise measurement setup for various application: a) Y-factor method - solid state noise source (up to 26 GHz) based noise figure measurement setup, b) gain method - frequency synthesiser based noise figure measurement setup and c) device noise characterisation measurement setup.

of a transistor, it is highly unlikely that the minimum noise figure occurs at 50 Ω input impedance, but instead at some other, complex impedance. The minimum noise figure occurs at some specific impedance referred to as Γ_{opt}. A fourth value, noise resistance R_n, represents the rate of change of the level of noise when varying the source impedance presented to the DUT. Lower R_n means a more desirable, shallower slope where the noise figure does not rapidly increase as impedance moves away from Γ_{opt}. The noise source pull technique involves varying the source impedance presented to a transistor (DUT) and measuring the noise figure. Based on Equation 2.1 the noise parameters can be extracted from the measurement using the least-square fitting algorithm.

$$F = NF_{min} + \frac{4\,R_n}{Z_0}\frac{|\Gamma_{opt} - \Gamma_s|^2}{|1 + \Gamma_{opt}|^2\,(1 + |\Gamma_s|^2)} \tag{2.1}$$

where F is the noise figure, NF_{min} is the minimum noise figure, R_n is the noise resistance, Z_0 is the characteristic impedance, Γ_{opt} and Γ_s are the optimum and the source reflection coefficients respectively.

The source pull measurement system normally consists of the following: a noise figure analyser or a spectrum analyser to measure noise power (such as an Agilent N8975A, FSV Rhode&Schwartz, or any other spectrum analyser); a source tuner to vary the impedance presented to the DUT (can be either manual or automatic); a noise source or a frequency synthesiser, with an attenuated output, to represent hot and cold states to the DUT; a vector network analyser (VNA) to measure S-Parameters (such as an Agilent PNA); and miscellaneous components such as adapters, cables etc... The measurement is automated and the data is recorded using Labview scripts run from a PC. The source pull measurements in this study were kindly performed by Murray microwave engineers, using a source pull setup, shown in Figure 2.13, for 8-50 GHz noise parameter measurements using a Maury MT7553B01 noise receiver module and a Maury MT984AU01 automated tuner with the Agilent PNA-X network analyser.

Figure 2.13: Murray microwave source pull automated setup for 8-50 GHz noise parameter measurements.

2.3 Device Under Test

Unfortunately, when performing on-wafer measurements, it is not possible to directly probe and measure the intrinsic device. Instead the DUT is embedded in a test fixture consisting of probe pads and interconnects to the DUT. As the frequency measurement increases, the fixture de-embedding becomes far more important and challenging, particularly since the pads introduce parasitic effects, which can have an impact on the measured parameters of a device. These parasitics are usually much larger than the device used for modelling. For this reason understanding and designing the probe pads and interconnects, coupled with accurate de-embedding, are key to achieving better and more accurate device models.

CMOS devices with aluminium pads oxidise immediately once exposed to air, a thin layer of aluminium oxide is formed on top of the aluminium surface resulting in large contact resistance. Conventional RF probes use tips made of tungsten (W) that can break through the oxide layer formed on top of the aluminium pads due to its firmness. Unfortunately, tungsten itself oxidises and the contact resistance increases over time leading to poor measurement repeatability. Fortunately, nickel-alloy tips have been developed

that provide very low and stable contact resistance (less than 0.1 Ω) when probing aluminium pads [28]. This is critical since the contact resistance must be repeatable for accurate de-embedding.

To respect the confidentiality imposed by the non-disclosure agreement with Texas Instruments (TI), the following information shall remain general. More detailed information can be obtained directly from the foundry. TI 0.18 μm standard logic CMOS technology features both n-channel and p-channel type transistors with nominal operating voltages of 1.8 V and 5 V respectively. The latter exhibits thicker gate oxides. The CMOS process uses shallow trench isolation (STI) to allow a higher density of integration and is based on a p-type low resistivity substrate (0.015 Ω/cm), with an optional 5.5 μm thick 30 Ω/cm epitaxial (EPI) layer. Further frontend features include deep n-wells. The metal back-end can provide up to six metal layers (M1-M6) for interconnects. Figure 2.14 shows a detailed illustration of the technology layer structure, including metal, oxide, substrate and passivation layers with their corresponding parameters.

Layer	Thickness	ε_r	Resistivity
Passivation	3.25 µm	5.75	
Metal 6	1 µm / 2.3 µm		0.078 Ω/□
IMD5	1.38 µm	4	
Metal 5	0.53 µm		0.078 Ω/□
IMD4	1.38µm	4	
Metal 4	0.53 µm		0.078 Ω/□
IMD3	1.38µm	4	
Metal 3	0.53 µm		0.078 Ω/□
IMD2	1.38µm	4	
Metal 2	0.53 µm		0.078 Ω/□
IMD1	1.38 µm	4	
Metal 1	0.53 µm		0.078 Ω/□
ILD	0.75µm	4	
Polysilicon	0.2 µm		7.9 Ω/□
	w/o silicide		306 Ω/□
FOX	0.35 µm	3.9	
Substrate	725 µm	11.9	

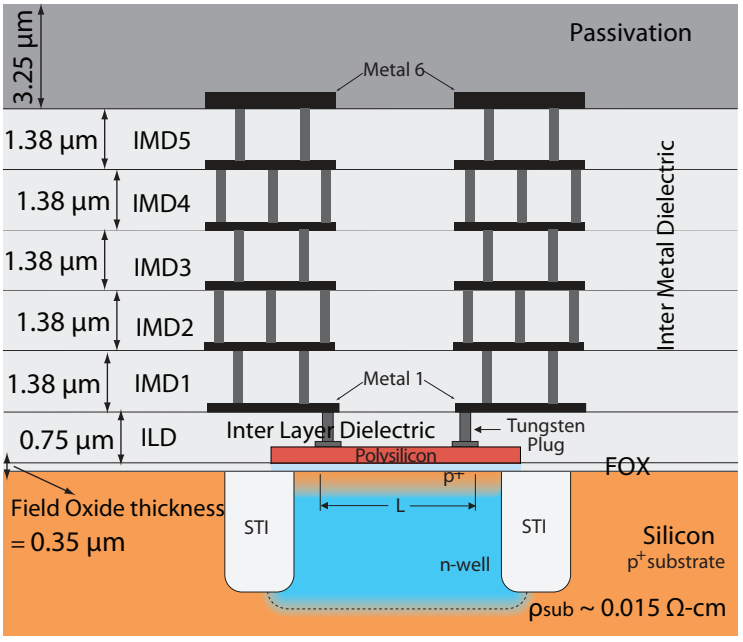


Figure 2.14: Detailed layer structure of TI CMOS technology process.

Electromagnetic simulators such as Agilent Momentum require substrate definition in order to simulate the designed structures correctly. The substrate definition includes the number of layers, position of each layer and composition of each layer. Silicon substrate permittivity and conductivity values are also needed. For the metal layer the required parameters are conductivity and metal thickness. The metal layers are interspaced by

SiO_2 dielectric layers. The backplating of the silicon substrate is represented by the closed boundary (ground) and at the top of the silicon substrate the free space is set as an open boundary (air).

2.3.1 Pad Design and Modelling

In order to perform any type of on-wafer measurements on the DUT, RF signals and bias supply voltages must be provided through appropriate pads. For the RF signals, coplanar waveguide (G-S-G) pads with a pitch[3] of 100 μm are used, as shown in Figure 2.15. The design of the probe pads on lossy silicon substrate involves a trade-off between resistive losses, parasitic capacitance and port-to-port isolation [13].

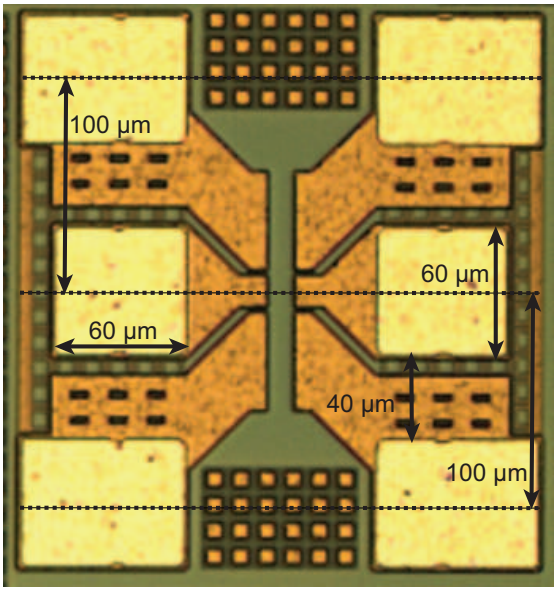


Figure 2.15: The layout of shielded RF GSG pads.

The pad capacitance should be small, otherwise it could potentially RF short-circuit high frequency signals to ground. The signal pad should therefore be implemented on the top most metal layers and it should occupy a small area [29]. The minimum pad size is limited by the smallest area on which a reliable probe contact is possible and the foundry design rules. The ground pads are realised by a stack of all metal layers from the top most metal layer (M6) down to the ground plane, which is a metal 1 sheet that covers the substrate. The size of the ground pads should be large to realise a large contact area and to accommodate any probe pitch. The passivation layer is removed over the entire pad surface. Dummy fillings of all the unused metal layers between the

[3]The distance between the centre points of two adjacent contacts.

pad and the ground plane are placed under the signal pads in order to comply with the metal density design rule.

The bias pads are similar to the ground pads in the G-S-G configuration mentioned above. The difference is that the metal stack of the bias pads descends down to the layer in which the bias signal is routed. The layers below the pad, except the ground plane, are removed in order to ensure substrate shielding.

Connecting the grounds on-chip through a ground plane ensures balanced ground currents and a very low-impedance ground return path. Imbalanced ground currents lead to parasitics that are different from those de-embedded during calibration. No electrical static discharge (ESD) protection is integrated in the pads as the realised circuits are intended for research purposes only. However, in a commercial design, ESD protection circuits are extremely crucial.

The pad model is extracted from an open and short circuited pad structure using simultaneous Y and Z-parameter fitting. As shown in Figure 2.16 the RF pad model is formed by two branches, one in series and another in parallel. The series part is composed of the pad inductance and resistance in series with a transmission line that represents the connection leading to the device. The parallel branch consists of the oxide capacitance and the ground shield resistance.

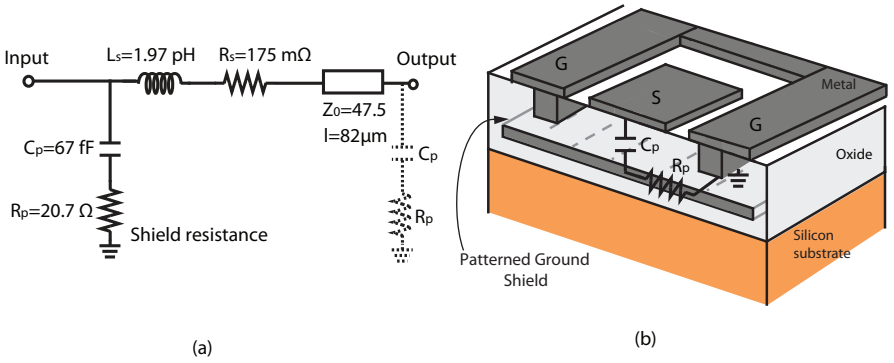


Figure 2.16: CMOS millimetre-wave G-S-G test pads: (a) Equivalent circuit representation of the test pad including inductances, capacitances and a transmission line and (b) 3D representation of one test pad illustrating the ground shield layer with the associated parasitics.

The lumped component values in the pad model are tuned to match the obtained pad measured network parameters. Figure 2.17 shows the measured Y and Z-parameters compared to the modelled parameters. Extracting the parallel capacitance and resistance, as well as the series inductance, is simple. However, the series resistor exhibits

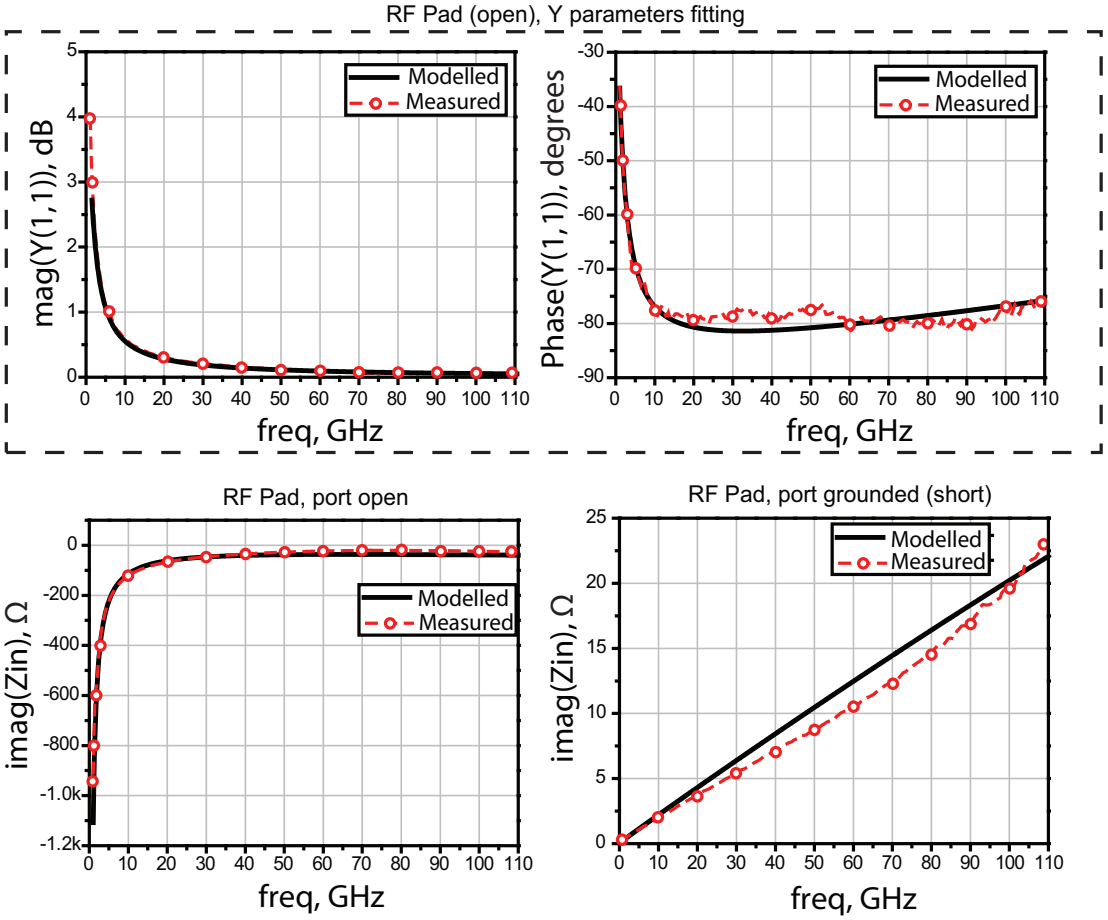


Figure 2.17: Modelled (solid lines) and measured (markers) Y and Z-parameters fitting for a G-S-G RF pad.

small values (a few hundred mΩ). This resistor contributes to the losses in the signal and affects the quality factor of the passive structure. For this reason it should be modelled accurately. Large variation in the value of this resistor has little effect on the S-parameters. Therefore, when optimising the pads or any other passive device model, S-parameter fitting should be avoided as it overshadows any series resistance in the model with a 50 Ω series resistor added by the simulation to each port. Instead Y and Z-fitting are used to ensure accurate model parameter values.

The transmission line in the pad model is represented by an Agilent advanced design system (ADS) transmission line model, CPW line or an ideal physical transmission line component (TLINP). These models are obtained from a measured transmission line fabricated on the same wafer. The TLINP line model requires the following parameters: characteristic impedance (Z_0), effective dielectric constant (k_{eff}), attenuation (α) in dB/m and the dielectric loss tangent (TanD). The TLINP model shall be discussed in further detail later on in this thesis.

2.4 Conventional On-Wafer Small Signal De-embedding

De-embedding is a term used to define the process of eliminating the influence of the transition region between the probe, probe pads and the DUT. The de-embedding structures should be implemented on wafer together with the DUT to ensure accurate results. In this section, conventional on-wafer de-embedding techniques are listed and explained along with their pros and cons.

2.4.1 Open De-embedding

Open de-embedding is the most commonly used approach due to its simplicity. It only requires an open dummy structure as shown in Figure 2.18. This technique assumes that the parasitics leading to the DUT are all in parallel with the DUT [30]. This assumption reduces the accuracy of the technique as it neglects the physical inductive and resistive nature of the pads. To clarify further, this de-embedding technique only removes the shunt parasitics from the measured device using an open dummy structure. Practically, this procedure is accurate up to 10 GHz for small on-chip structures.

The open de-embedding procedure can be performed using the following steps:

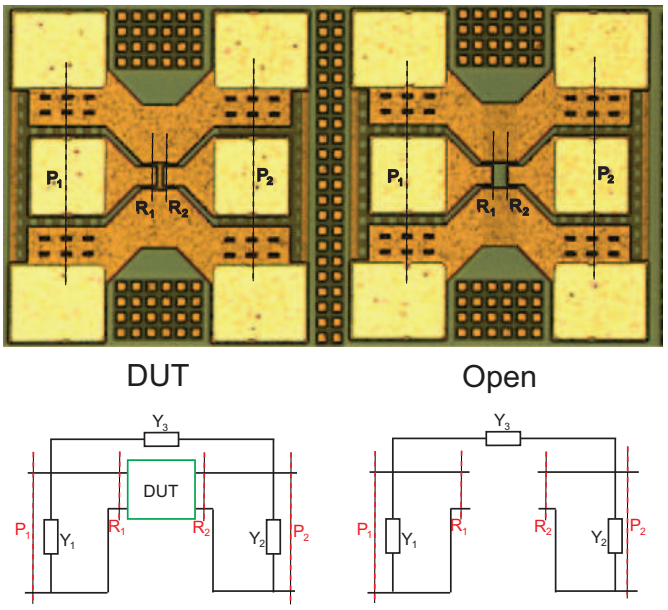


Figure 2.18: Open de-embedding: The DUT with the measurement pads and the open dummy structure with the corresponding equivalent circuits.

1. Measure the scattering parameters of the DUT (S^{Meas}_{DUT}) and the open structures (S^{Meas}_{Open}).

2. Convert the measured scattering parameters to Y-parameter matrices (Y^{Meas}_{DUT}, Y^{Meas}_{Open}).

$$\left[S^{Meas}_{DUT}\right] \longrightarrow \left[Y^{Meas}_{DUT}\right] \tag{2.2}$$

$$\left[S^{Meas}_{Open}\right] \longrightarrow \left[Y^{Meas}_{Open}\right] \tag{2.3}$$

3. Calculate the S-parameters (S_{DUT}) of the intrinsic DUT structure by subtracting the open structure from the DUT as follows:

$$\left[Y_{DUT}\right] = \left[Y^{Meas}_{DUT}\right] - \left[Y^{Meas}_{Open}\right], \tag{2.4}$$

$$\left[Y_{DUT}\right] \longrightarrow \left[S_{DUT}\right]. \tag{2.5}$$

2.4.2 Open-Short De-embedding

This technique is the most common approach used for de-embedding structures for up to very high frequencies. Open and short dummy structures are required to perform this technique as shown in Figure 2.19. The open dummy structure removes the shunt

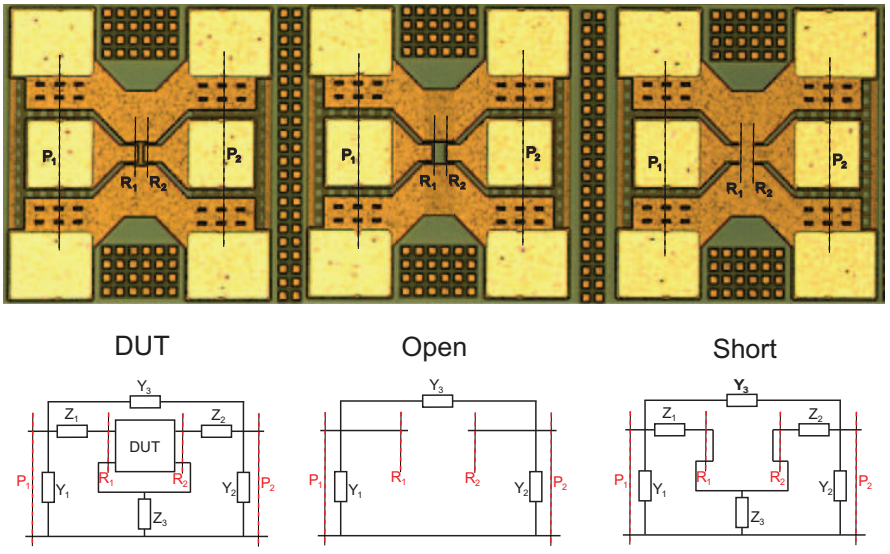


Figure 2.19: Open-short de-embedding: The DUT with the measurement pads and the open and short dummy structures with their equivalent circuits.

parasitic elements, whereas the short dummy structure removes the series elements, providing for far more accurate device de-embedding results.

The open-short de-embedding procedure involves the following steps:

1. Measure the scattering parameters of the DUT ($\mathrm{S}^{\mathrm{Meas}}_{\mathrm{DUT}}$), open ($\mathrm{S}^{\mathrm{Meas}}_{\mathrm{Open}}$) and the short ($\mathrm{S}^{\mathrm{Meas}}_{\mathrm{Short}}$) structures.

2. Convert the measured scattering parameters to Y-parameters matrices ($\mathrm{Y}^{\mathrm{Meas}}_{\mathrm{DUT}}$, $\mathrm{Y}^{\mathrm{Meas}}_{\mathrm{Open}}$ and $\mathrm{Y}^{\mathrm{Meas}}_{\mathrm{Short}}$).

$$\left[S^{\mathrm{Meas}}_{\mathrm{DUT}}\right] \longrightarrow \left[Y^{\mathrm{Meas}}_{\mathrm{DUT}}\right] \tag{2.6}$$

$$\left[S^{\mathrm{Meas}}_{\mathrm{Open}}\right] \longrightarrow \left[Y^{\mathrm{Meas}}_{\mathrm{Open}}\right] \tag{2.7}$$

$$\left[S^{\mathrm{Meas}}_{\mathrm{Short}}\right] \longrightarrow \left[Y^{\mathrm{Meas}}_{\mathrm{Short}}\right] \tag{2.8}$$

3. Subtract the open structure from the DUT then convert the results to Z-parameters using

$$\left[Y_{\mathrm{DUT/Open}}\right] = \left[Y^{\mathrm{Meas}}_{\mathrm{DUT}}\right] - \left[Y^{\mathrm{Meas}}_{\mathrm{Open}}\right], \tag{2.9}$$

$$\left[Y_{\mathrm{DUT/Open}}\right] \longrightarrow \left[Z^{\mathrm{Meas}}_{\mathrm{DUT/Open}}\right]. \tag{2.10}$$

4. Subtract the open structure from the short then convert the results to Z-parameters using

$$\left[Y_{\mathrm{Short/Open}}\right] = \left[Y^{\mathrm{Meas}}_{\mathrm{Short}}\right] - \left[Y^{\mathrm{Meas}}_{\mathrm{Open}}\right], \tag{2.11}$$

$$\left[Y_{\mathrm{Short/Open}}\right] \longrightarrow \left[Z^{\mathrm{Meas}}_{\mathrm{Short/Open}}\right]. \tag{2.12}$$

5. Calculate the S-parameters ($\mathrm{S}_{\mathrm{DUT}}$) of the intrinsic DUT structure by subtracting the results of step 3 and 4 as follows:

$$\left[Z_{\mathrm{DUT}}\right] = \left[Z^{\mathrm{Meas}}_{\mathrm{DUT/Open}}\right] - \left[Z^{\mathrm{Meas}}_{\mathrm{Short/Open}}\right], \tag{2.13}$$

$$\left[Z_{\mathrm{DUT}}\right] \longrightarrow \left[S_{\mathrm{DUT}}\right]. \tag{2.14}$$

At millimetre-wave frequencies, the accuracy of the open and open-short methods are questionable as the open and short dummy structures have some parasitics that are not present when the DUT is connected. The short dummy structure has an extra parasitic inductance to ground and the open dummy structure has an extra parasitic fringing capacitance. This may lead to over de-embedding and in some cases optimistic or degraded device performance [31]. To overcome this problem, 3-step, 4-step or even 5-step de-embedding techniques have been proposed [32–34]. These methods involve complex analysis and many dummy structures, this could lead to measurement errors due to probe misplacement, not to mention the extra area these dummy structures occupy, which for a compact CMOS chip is extremely vital. In this work a simple 2-step method, with much less area consumption, was developed to overcome these issues. Two new simple de-embedding methods are introduced and explained. These methods are validated through a comparative study with measured and de-embedded parameters extracted from various passive and active devices.

2.5 Proposed On-Wafer Small Signal De-embedding

2.5.1 Modelled Pads De-embedding

All the designed on-chip components use the same probe pads. This technique can be used efficiently as it relies on the modelled pads described in Section 2.3.1. The pad network parameters are converted to ABCD parameters and used as a chain network composed of a left pad, DUT, right pad, as shown in Figure 2.20. This technique provides higher accuracy at millimetre-wave frequencies since it models the non-ideality caused by the taper, connecting the probe pads to the interconnect transmission line that leads to the DUT.

This de-embedding procedure involves the following steps:

1. Measure the scattered parameters of the DUT (S^{Meas}_{DUT}) then convert to ABCD parameters (A^{Meas}_{DUT}).

$$\left[S^{\text{Meas}}_{\text{DUT}}\right] \longrightarrow \left[A^{\text{Meas}}_{\text{DUT}}\right] \tag{2.15}$$

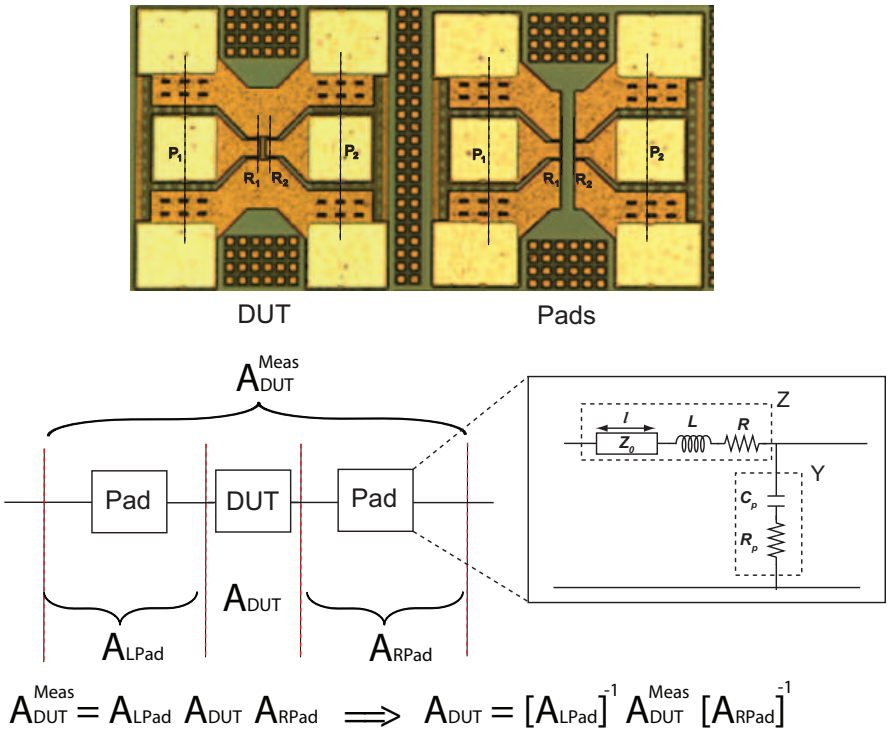


Figure 2.20: Modelled pad de-embedding technique.

2. Convert the scattered parameters of the modelled left and right pads to ABCD parameters matrices A_{LPad}, A_{RPad} respectively.

$$\left[S_{\mathrm{LPad}}\right] \longrightarrow \left[A_{\mathrm{LPad}}\right] \tag{2.16}$$

$$\left[S_{\mathrm{RPad}}\right] \longrightarrow \left[L_{\mathrm{RPad}}\right] \tag{2.17}$$

3. Calculate the S-parameters (S_{DUT}) of the intrinsic DUT structure using

$$\left[A_{\mathrm{DUT}}^{\mathrm{Meas}}\right] = \left[A_{\mathrm{LPad}}\right]\left[A_{\mathrm{DUT}}\right]\left[A_{\mathrm{RPad}}\right], \tag{2.18}$$

$$\left[A_{\mathrm{DUT}}\right] = \left[A_{\mathrm{LPad}}\right]^{-1}\left[A_{\mathrm{DUT}}^{\mathrm{Meas}}\right]\left[A_{\mathrm{RPad}}\right]^{-1}, \tag{2.19}$$

$$\left[A_{\mathrm{DUT}}\right] \longrightarrow \left[S_{\mathrm{DUT}}\right]. \tag{2.20}$$

2.5.2 Pad-Short De-embedding

This technique is proposed to improve the accuracy of the de-embedding process up to millimetre-wave frequencies. As shown in Figure 2.21, a zero length thru dummy structure and a short dummy structure are used to extract the pads parasitics. This

technique provides accurate de-embedding for both symmetric and asymmetric devices. The short structure compensates for the inductance to ground parasitic elements in the active device.

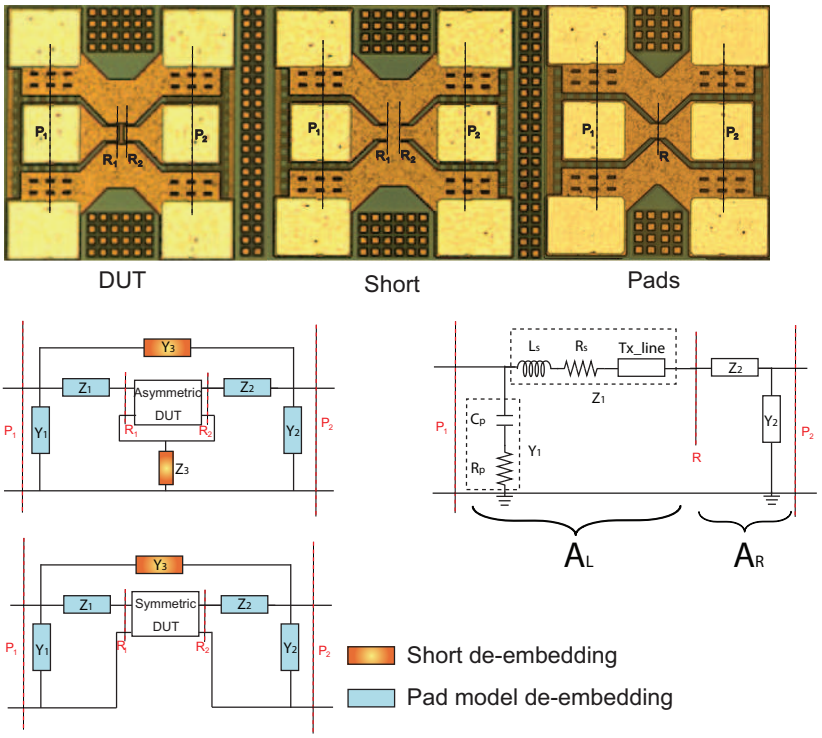


Figure 2.21: Modelled pad-short de-embedding technique.

The proposed de-embedding procedure involves the following steps:

1. Measure the scattered parameters of the DUT (S_{DUT}^{Meas}), the 0 length thru (S_{thru}) and the short (S_{Short}) structures.

2. Convert the measured scattered parameters to ABCD parameter matrices (A_{DUT}^{Meas}, A_{thru}, A_{Short}).

3. Calculate the ABCD parameters of the left and right side pads, [A_{LPad}] and [A_{RPad}] respectively, using

$$\left[A_{\mathrm{thru}}\right] = \left[A_{\mathrm{LPad}}\right]\left[A_{\mathrm{RPad}}\right], \tag{2.21}$$

$$\left[A_{\mathrm{LPad}}\right]\left[A_{\mathrm{RPad}}\right] = \begin{bmatrix} 1+\frac{y_1}{y_2} & \frac{1}{y_2} \\ y_1 & 1 \end{bmatrix}\begin{bmatrix} 1 & \frac{1}{y_2} \\ y_1 & 1+\frac{y_1}{y_2} \end{bmatrix}, \tag{2.22}$$

$$\left[A_{\mathrm{thru}}\right] = \begin{bmatrix} A_{\mathrm{thru}} & B_{\mathrm{thru}} \\ C_{\mathrm{thru}} & D_{\mathrm{thru}} \end{bmatrix} = \begin{bmatrix} 1+\frac{2y_1}{y_2} & \frac{2}{y_2}\left(1+\frac{y_1}{y_2}\right) \\ 2y_1 & 1+\frac{2y_1}{y_2} \end{bmatrix}, \tag{2.23}$$

$$\left[A_{\mathrm{LPad}}\right] = \begin{bmatrix} \frac{1+A_{\mathrm{thru}}}{2} & \frac{D_{\mathrm{thru}}-1}{C_{\mathrm{thru}}} \\ \frac{C_{\mathrm{thru}}}{2} & 1 \end{bmatrix}, \tag{2.24}$$

$$\left[A_{\mathrm{RPad}}\right] = \begin{bmatrix} 1 & \frac{D_{\mathrm{thru}}-1}{C_{\mathrm{thru}}} \\ \frac{C_{\mathrm{thru}}}{2} & \frac{1+A_{\mathrm{thru}}}{2} \end{bmatrix}. \tag{2.25}$$

4. Subtract the pads from the DUT using the short pads and convert to Z-parameters using

$$\left[A_{\mathrm{Short/Pad}}\right] = \left[A_{\mathrm{LPad}}\right]^{-1}\left[A_{\mathrm{Short}}\right]\left[A_{\mathrm{RPad}}\right]^{-1} \tag{2.26}$$

$$\left[A_{\mathrm{Short/Pad}}\right] \longrightarrow \left[Z_{\mathrm{Short/Pad}}\right] \tag{2.27}$$

5. Subtract the pads from the DUT and convert to Z-parameters using

$$\left[A_{\mathrm{DUT/Pad}}\right] = \left[A_{\mathrm{LPad}}\right]^{-1}\left[A_{\mathrm{DUT}}^{\mathrm{Meas}}\right]\left[A_{\mathrm{RPad}}\right]^{-1} \tag{2.28}$$

$$\left[A_{\mathrm{DUT/Pad}}\right] \longrightarrow \left[Z_{\mathrm{DUT/Pad}}\right] \tag{2.29}$$

6. Calculate the S-parameters (S_{DUT}) of the intrinsic DUT structure using

$$\left[Z_{\mathrm{DUT}}\right] = \left[Z_{\mathrm{DUT/Pad}}\right] - \left[Z_{\mathrm{Short/Pad}}\right] \tag{2.30}$$

$$\left[Z_{\mathrm{DUT}}\right] \longrightarrow \left[S_{\mathrm{DUT}}\right] \tag{2.31}$$

2.5.3 Experimental Results - Evaluation and Comparison

In order to evaluate the performance of the previously mentioned two methods (modelled pad and pad-short), comparisons were made with the conventional techniques (open

and open-short). The aim of the evaluation is to demonstrate that the proposed de-embedding techniques can be applied to symmetric (reciprocal) and asymmetric (non-reciprocal) structures. The first and most straight forward evaluation process involves de-embedding a symmetric structure such as a transmission line. The second evaluation process involves de-embedding an asymmetric structure for example a transistor.

1. **Symmetric structures de-embedding evaluation**

   The DUT used for the symmetric structure evaluation is a $300\,\mu$m long CPW line (coplanar waveguide), constructed to provide a characteristic impedance $Z_0 = 50\,\Omega$, using the top most metal layer (metal 6 = $1\,\mu$m thick) and separated from the meshed ground plane (metal 1) by a $6.54\,\mu$m layer of oxide with a dielectric constant of $\epsilon_r = 4$. The signal line width is $15\,\mu$m and the gap spacing between the signal and ground lines on each side is also $15\,\mu$m. It can be seen from the results presented in Figure 2.22 that the de-embedded extracted line parameters from the open method are higher than those extracted from the other methods. This is mainly due to the fact that open de-embedding does not consider the series parasitics and only accounts for the parallel ones. It therefore results in a larger error when compared to the other methods, especially at very high frequencies. On the other hand, the pad-short, modelled pad and open-short methods exhibit very little difference across the entire measurement spectrum. This is a good indicator that the proposed methods are reliable and can be used to de-embed the pad parasitic effects from any reciprocal structure up to 110 GHz.

2. **Asymmetric structures de-embedding evaluation**

   The asymmetric evaluation was conducted on an n-MOSFET, with a total gate width of $80\,\mu$m divided into 40 fingers, $2\,\mu$m wide each, with a channel length of $0.18\,\mu$m. The device was connected in a common source configuration and biased at a gate voltage of $V_g = 0.9\,\mathrm{V}$ and a drain voltage of $V_d = 1.3\,\mathrm{V}$.

   Figure 2.23 shows the maximum stable gain (MSG), the unity current gain ($\mathrm{f}_T = |\mathrm{h}_{21}|$) and the open-circuit voltage gain (Av_0). It can be seen that the most sensitive parameter to de-embedding variation, especially at very high frequencies, is Av_0. Therefore by comparing the results from open and open-short de-embedding with those obtained from the pad-short and modelled pad methods, it can clearly be seen that for the frequencies between $1\,\mathrm{GHz} - 40\,\mathrm{GHz}$ all the

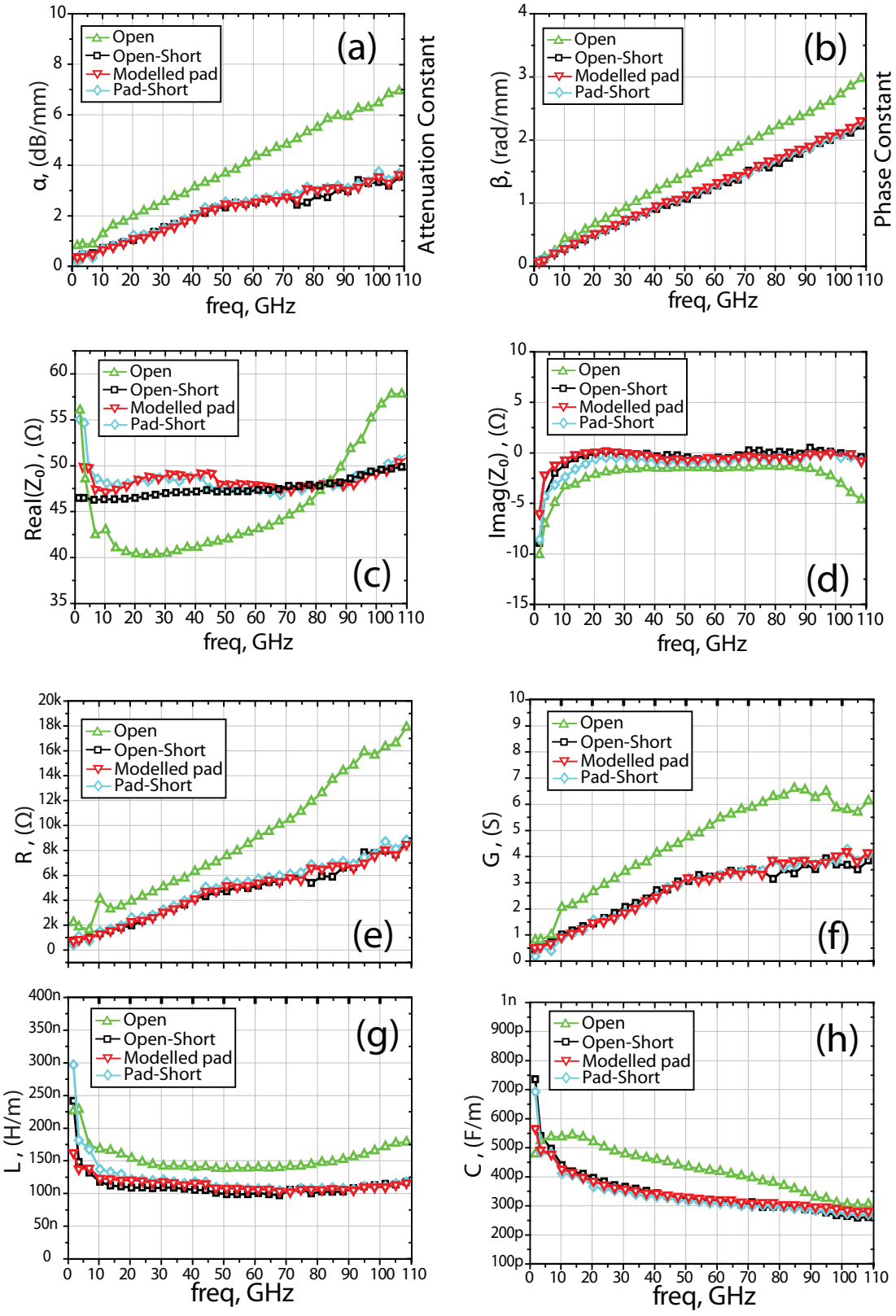


Figure 2.22: Plots comparing open, open-short, modelled pad and pad-short de-embedding methods evaluated on a CPW line: (a) The attenuation constant, (b) the phase constant, (c) the real part of the characteristic impedance of the line, (d) the imaginary part of the characteristic impedance of the line, (e) the extracted resistance, (f) the extracted conductance, (g) the extracted inductance per unit length and (h) the extracted capacitance per unit length.

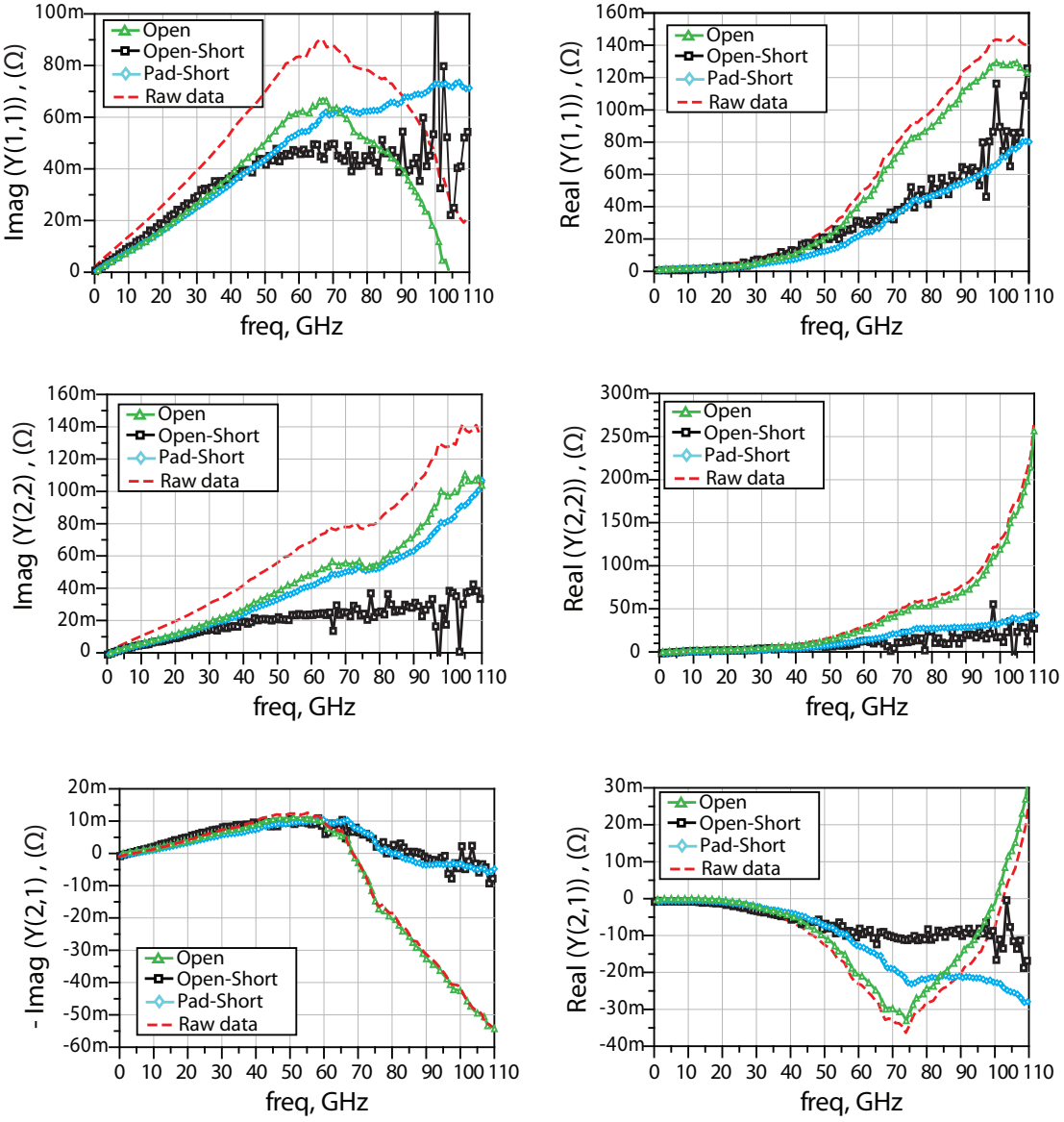


Figure 2.23: The evaluation and comparison of different de-embedding methods: Device capacitances C_{gs}, C_{gd} and C_{ds}.

methods show similar results. The open method starts to deviate at frequencies above 40 GHz, whereas the other de-embedding methods maintain a good agreement up to 75 GHz. Again, the difference occurring in the open method is mainly due to the fact that the series parasitics are totally neglected in the de-embedding process, this leads to under de-embedding errors. As for the other de-embedding methods, the difference starts to appear at frequencies above 75 GHz. At such high frequencies the series source to ground impedance, which serves as a negative feedback to the MOSFET, is normally neglected in the open and the modelled pad methods. This causes the difference in the results when compared to open-short

and pad-short as both of those methods consider the source impedance in the de-embedding process. Although open-short and pad-short take into account every parasitic element, the advantage that pad-short has over open-short is that, as discussed earlier in this chapter, the dummy structure of the latter suffers the extra fringing and series parasitics that do not exist when the device is attached. In contrast, the pad-short method relies on a thru structure to extract the pad parasitics and the short structure accounts for the series source to ground element. The results mentioned above demonstrate a good indicator that the pad-short method is reliable and can be used to de-embed the pad parasitic effects from any reciprocal or non-reciprocal structure up to 110 GHz.

2.6 On-Wafer Noise De-embedding

Noise de-embedding is very important when performing noise measurements mainly due to the probe pads added noise effects. The probe pads reactive network will transform the noise parameters of the DUT, in addition to the losses, thermal noise contribution to the overall noise measurements. Therefore, these effects have to be de-embedded in order to ensure accurate noise parameter extraction and device noise characterisation. Technically, any small-signal de-embedding method can be used for noise de-embedding with slight modification.

Noise de-embedding requires knowledge of two port noisy networks representation, correlation matrices and conversions as well as an understanding of two-port noisy network theory.

2.6.1 Conventional Noise De-embedding Technique

The simplest noise de-embedding technique is based on the open-short de-embedding technique described in Section 2.4.2. This approach requires open and short dummy structures and involves the following steps:

1. Convert the measured S- parameters of the DUT, and the open structure (S_{DUT}, S_{Open}) to Y-parameters (Y_{DUT}, Y_{Open}).

2. Determine the correlation matrix $C_{Y_{DUT}}$ using

$$\left[C_{Y_{DUT}}\right] = 2KT\Re\left[Y_{DUT}\right] \tag{2.32}$$

where $\Re[\ ]$ denotes the real part of a complex number.

2. Perform the open-short de-embedding technique as follows:

$$\left[Y_{DUT/Open}\right] = \left[Y_{DUT}\right] - \left[Y_{Open}\right] \tag{2.33}$$

$$\left[C_{Y_{DUT/Open}}\right] = \left[C_{Y_{DUT}}\right] - \left[C_{Y_{Open}}\right] \tag{2.34}$$

$$\left[Y_{Short/Open}\right] = \left[Y_{Short}\right] - \left[Y_{Open}\right] \tag{2.35}$$

$$\left[C_{Y_{Short/Open}}\right] = \left[C_{Y_{Short}}\right] - \left[C_{Y_{Open}}\right] \tag{2.36}$$

3. Convert the measured S-parameters of the DUT, and the short structure (S_{DUT}, S_{Short}) to Z-parameters (Z_{DUT}, Z_{Short}).

4. Convert $C_{Y_{DUT/Open}}$ and $C_{Y_{Short/Open}}$ to their Z-parameter correlation matrices $C_{Z_{DUT/Open}}$ and $C_{Z_{Short/Open}}$ using

$$\left[C_{Z_{DUT/Open}}\right] = \left[Z_{DUT/Open}\right]\left[C_{Y_{DUT/Open}}\right]\left[Z_{DUT/Open}\right]^{\dagger} \tag{2.37}$$

$$\left[C_{Z_{Short/Open}}\right] = \left[Z_{Short/Open}\right]\left[C_{Y_{Short/Open}}\right]\left[Z_{Short/Open}\right]^{\dagger} \tag{2.38}$$

where $[T]^{\dagger}$ is the transpose conjugate of [T].

5. Determine the correlation matrix C_Z using

$$\left[C_Z\right] = \left[C_{Z_{DUT/Open}}\right] - \left[C_{Z_{Short/Open}}\right] \tag{2.39}$$

6. Transform C_Z to C_A with

$$\left[C_A\right] = \left[T_A\right]\left[C_Z\right]\left[T_A\right]^{\dagger} \tag{2.40}$$

$$\left[T_A\right] = \begin{bmatrix} 1 & -A \\ 0 & -C \end{bmatrix} \tag{2.41}$$

7. Calculate the noise parameters NF_{min}, R_n and Y_{opt} using

$$\left[C_A\right] = \begin{bmatrix} R_n & \frac{NF_{min}-1}{2} - R_n Y_{opt}^* \\ \frac{NF_{min}-1}{2} - R_n Y_{opt} & R_n |Y_{opt}|^2 \end{bmatrix} \tag{2.42}$$

$$R_n = \frac{C_{A_{11}}}{2\,k_B\,T} \tag{2.43}$$

$$Y_{opt} = \frac{\sqrt{C_{A_{11}} C_{A_{22}} - (\Im(C_{A_{12}}))^2} + j\,\Im(C_{A_{12}})}{C_{A_{11}}} \tag{2.44}$$

$$NF_{min} = 1 + \frac{1}{k_B\,T}\left(\Re(C_{A_{12}}) + \sqrt{C_{A_{11}} C_{A_{22}} - (\Im(C_{A_{12}}))^2}\right) \tag{2.45}$$

$$NF_{min} = 1 + 2(C_{A_{12}} + C_{A_{11}}\,Y_{opt}^*) \tag{2.46}$$

8. Extract NF_{min}^{DUT}, R_n^{DUT} and Y_{opt}^{DUT} from C_A^{DUT} and compare to verify the de-embedding procedure.

This method is convenient for calculating the noise parameters of intrinsic devices. However, just like small-signal de-embedding, this technique could cause over or under de-embedding due to the extra parasitics introduced by the open and short dummy structures. To avoid this problem, a new method based on the technique described in Section 2.5.2 is proposed.

2.6.2 Proposed Noise De-embedding Technique

In the proposed method, the pads and the intrinsic device with some series and parallel parasitic elements are combined to form the DUT as shown in Figure 2.24.

The noise de-embedding is performed as follows:

1. Measure the scattered parameters of the DUT (S_{DUT}^{Meas}), the 0 length thru (S_{thru}) and the short ($\mathrm{S}_{Short}^{Meas}$) structures and convert it into Y-parameters (Y_{DUT}^{Meas}, Y_{thru}, $\mathrm{Y}_{Short}^{Meas}$) and ABCD parameters ($\mathrm{A}_{DUT}^{Meas}$, A_{thru}, $\mathrm{A}_{Short}^{Meas}$) .

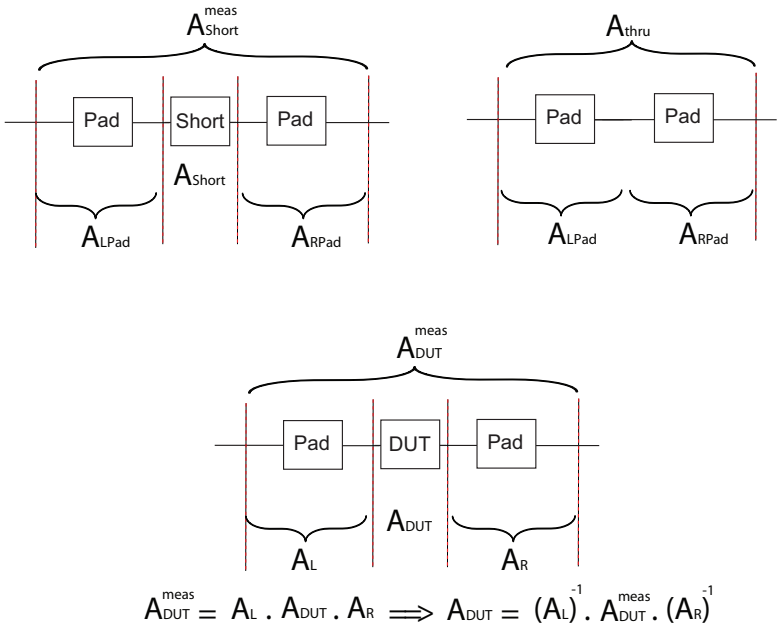


Figure 2.24: Noise de-embedding.

2. Calculate the ABCD parameters of the left and right pads [A$_{LPad}$] and [A$_{RPad}$], respectively using

$$\left[A_{thru}\right] = \left[A_{LPad}\right]\left[A_{RPad}\right] \tag{2.47}$$

$$\left[A_{LPad}\right] = \begin{bmatrix} \frac{1+A_{thru}}{2} & \frac{D_{thru}-1}{C_{thru}} \\ \frac{C_{thru}}{2} & 1 \end{bmatrix} \tag{2.48}$$

$$\left[A_{RPad}\right] = \begin{bmatrix} 1 & \frac{D_{thru}-1}{C_{thru}} \\ \frac{C_{thru}}{2} & \frac{1+A_{thru}}{2} \end{bmatrix} \tag{2.49}$$

3. The measured short structure can be represented as a chain of three cascaded noisy networks (A$_{Lpad}$, A$_{Short}$ and A$_{Rpad}$) and the measured DUT can be represented in a similar fashion. Therefore A$_{Short}$ and A$_{DUT}$ can be determined as follows:

$$\left[A^{Meas}_{Short}\right] = \left[A_{LPad}\right]\left[A_{Short}\right]\left[A_{RPad}\right] \tag{2.50}$$

$$\left[A^{Meas}_{DUT}\right] = \left[A_{LPad}\right]\left[A_{DUT}\right]\left[A_{RPad}\right] \tag{2.51}$$

$$\left[A_{Short}\right] = \left[A_{LPad}\right]^{-1}\left[A^{Meas}_{Short}\right]\left[A_{RPad}\right]^{-1} \tag{2.52}$$

$$\left[A_{DUT}\right] = \left[A_{LPad}\right]^{-1}\left[A^{Meas}_{DUT}\right]\left[A_{RPad}\right]^{-1} \tag{2.53}$$

4. Calculate Z$_{DUT}$ and Z$_{Short}$ from A$_{DUT}$ and A$_{Short}$.

5. Subtract Z_{Short} from Z_{DUT} and convert the result to ABCD matrix $A_{DUT/Short}$.

$$\left[Z_{DUT/Short}\right] = \left[Z_{Short}\right] - \left[Z_{DUT}\right] \tag{2.54}$$

6. Determine the correlation matrix $C_{A_{DUT/Short}}$ using

$$\left[C_{Z_{DUT/Short}}\right] = 2KT\Re\left[Z_{DUT/Short}\right] \tag{2.55}$$

$$\left[C_{A_{DUT/Short}}\right] = \left[T_A\right]\left[C_{Z_{DUT/Short}}\right]\left[T_A\right]^{\dagger} \tag{2.56}$$

$$\left[T_A\right] = \begin{bmatrix} 1 & -A \\ 0 & -C \end{bmatrix} \tag{2.57}$$

7. Calculate the noise parameters NF_{min}, R_n and Y_{opt}.

$$\left[C_A\right] = \begin{bmatrix} R_n & \frac{NF_{min}-1}{2} - R_n Y_{opt}^* \\ \frac{NF_{min}-1}{2} - R_n Y_{opt} & R_n |Y_{opt}|^2 \end{bmatrix} \tag{2.58}$$

$$R_n = \frac{C_{A_{11}}}{2\,k_B\,T} \tag{2.59}$$

$$Y_{opt} = \frac{\sqrt{C_{A_{11}}C_{A_{22}} - (\Im(C_{A_{12}}))^2} + j\,\Im(C_{A_{12}})}{C_{A_{11}}} \tag{2.60}$$

$$NF_{min} = 1 + \frac{1}{k_B\,T}\left(\Re(C_{A_{12}}) + \sqrt{C_{A_{11}}C_{A_{22}} - (\Im(C_{A_{12}}))^2}\right) \tag{2.61}$$

$$NF_{min} = 1 + 2(C_{A_{12}} + C_{A_{11}}\,Y_{opt}^*) \tag{2.62}$$

8. Extract NF_{min}^{DUT}, R_n^{DUT} and Y_{opt}^{DUT} from C_A^{DUT} and compare to verify the de-embedding procedure.

2.6.3 Results - Evaluation and Comparison

The proposed noise de-embedding method was evaluated on an NMOS transistor, with a total gate width of 80 μm divided into 40 fingers, 2 μm wide each, with a channel length of 0.18 μm. The device was connected in a common source configuration and biased at a gate voltage of V_g=0.9V and a drain voltage V_d=1.3V. Source pull noise measurement was performed on the device. Open-short and pad-short methods were used to de-embed the measured noise parameters.

The measurement results shown in Figure 2.25 depict a negligible difference between the two methods for all the noise parameters up to 20 GHz. This result shows that the pad-short method agrees with the published open-short de-embedding method at lower frequencies. However, as the frequency increases the difference between pad-short and open-short results increases. This is due to the extra fringing parasitic capacitance that the open dummy structures exhibits at high frequencies. The pad-short method uses a zero-length thru dummy structure rather than an open structure, this reduces the parasitic effects at higher frequency and makes the measurement result much smoother. Although pad-short results deviate from those of the open-short method, the difference at 50 GHz is still small (0.1%). This further indicates that the pad-short noise de-embedding method is reliable for the subtraction of both the pad parasitics up to 60 GHz.

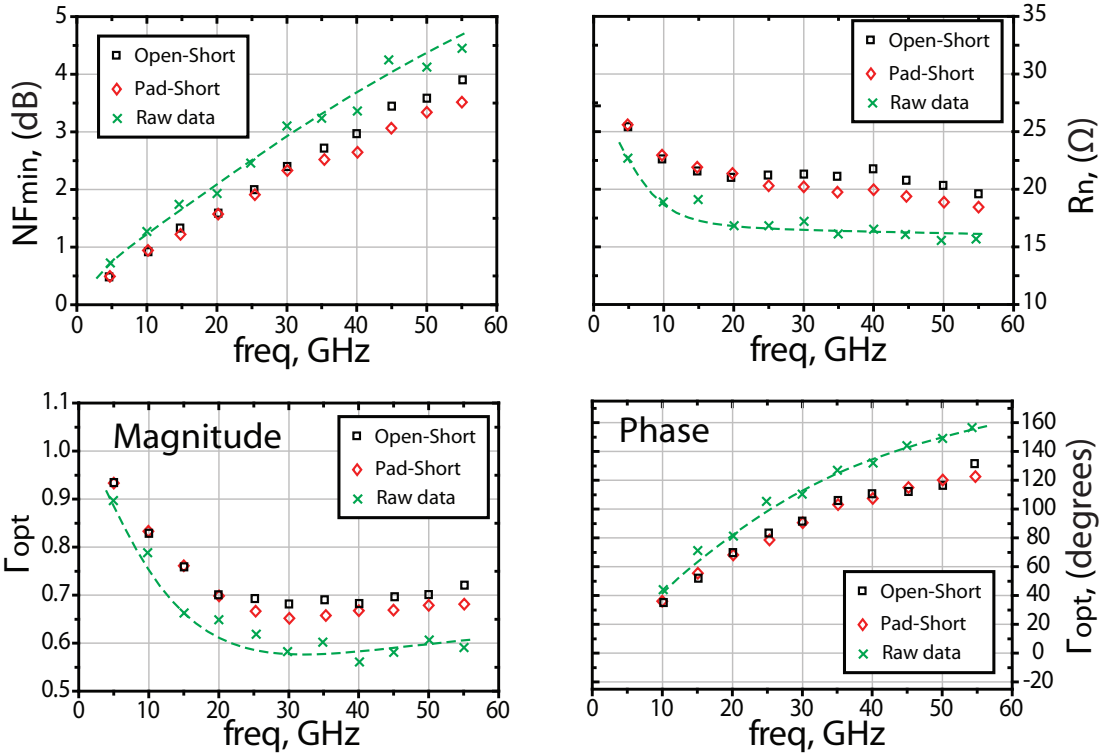


Figure 2.25: Plots comparing raw data, open-short and pad-short de-embedding methods performed on a NMOS transistor: a) The minimum noise figure (NF_{min}), b) the equivalent noise resistance (R_n), c) the magnitude of Γ_{opt} and d) the phase of Γ_{opt}.

2.7 Summary

This chapter began by illustrating the various important DC and RF measurement equipments. DC as a prerequisite for RF measurement, small-signal and noise measurement setups were illustrated and explained. Then the design of CMOS DUT was explained, challenges involving measuring and de-embedding test fixtures were discussed and addressed. Several on-wafer, conventional and proposed, small-signal and noise de-embedding techniques were presented. None of these methods is capable of completely removing the probe pad effects to very high frequencies. Each de-embedding technique introduces some error. This error can not be exactly determined by verification processes otherwise these processes could be used to de-embed the DUT perfectly and with great accuracy.

Chapter 3

Integrated CMOS Passive Components

3.1 Introduction

Traditionally, microwave and millimetre-wave circuits have been implemented using compound semiconductors such as GaAs and InP. In the past few decades however an increasing interest in high frequency silicon technology has been witnessed mainly due to its low cost, high integration capabilities, high dielectric constant, good thermal conductivity, minimal variation with temperature and frequency, and multi-interconnect metal layers (6 to 9 metal layers). The latter can be of a great help when it comes to building integrated passive devices such as transmission lines, inductors and capacitors. In spite of all the aforementioned benefits of silicon technology in microwave design, one major disadvantage is the lower resistivity of the silicon substrate making it prone to substrate losses in contrast to other technologies such as GaAs. This stems from the parasitic electromagnetic (EM) signal propagation through the silicon substrate [35]. It is therefore vital to design passive structures that can overcome this disadvantage giving an improved performance.

In this chapter, the different CMOS passive structures are first introduced, with the associated challenges highlighted. Following this techniques to overcome the shortcomings are presented and discussed. Finally a physical model of each of the above structures are implemented and verified.

3.2 On-chip Spiral Inductor

On-chip spiral inductors are the largest silicon area consuming devices for RFICs and play a very important role in any wireless system performance. Therefore, characterising and modelling them accurately are vital for any RFIC designer.

3.2.1 Structure

Typically, spiral inductors are designed by placing a metal track in a planar spiral configuration. The terminals of the inductor are the two ends of the spiral. The most common spiral configurations are circular, octagonal, hexagonal and square as shown in Figure 3.1. Spiral inductors are usually characterised by the line width 'W', the spacing between tracks 'S', the diameters 'd_{in}' and 'd_{out}', and the number of turns 'N'. Several topologies are available for implementing an on-chip inductor, of which the spiral inductor gives an optimal performance.

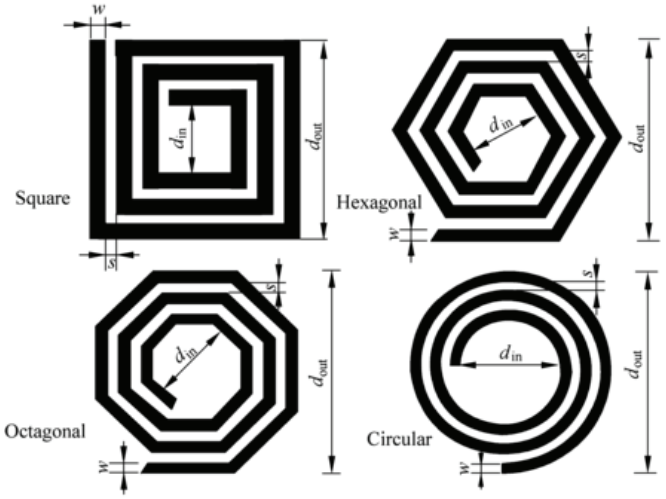


Figure 3.1: Planar spiral inductors with different shapes.

Figure 3.2 (a) shows a single-ended inductor, often used in matching networks and bias circuits, e.g. noise and power matching networks of an amplifier. This inductor is normally implemented using a track on the top metal layer with an underpass made on lower metal layers as shown in Figure 3.2 (a). Figure 3.2 (b) and (c) show symmetric spiral inductors with and without a centre tap, which connects to the centre node. These types of inductors are often used in differential circuits such as mixers and voltage controlled oscillators.

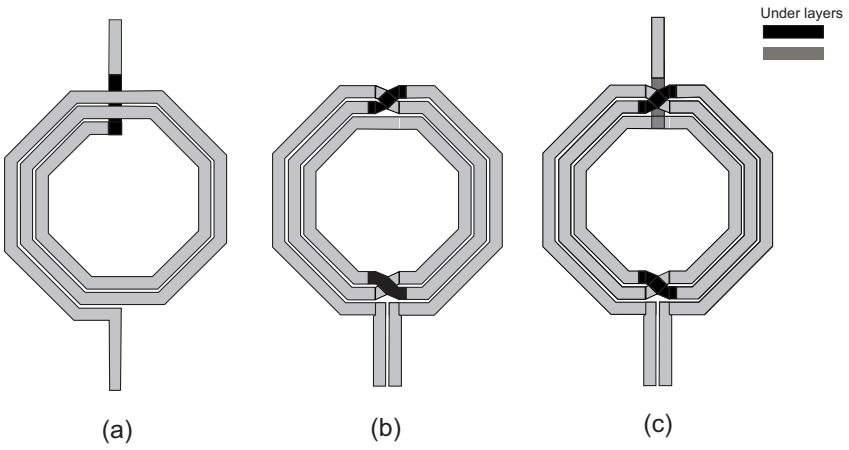


Figure 3.2: On-chip spiral inductors: (a) single-ended, (b) symmetric and (c) symmetric with a centre tap.

Spiral inductors are normally analysed using one of the following methods:

• **The numerical method:** This approach is based on finite element solutions of Maxwell's equations [36, 37], iterative convergence of the geometric mean distance and the arithmetic mean distance of each conductor segment with other neighbouring tracks in the structure. Despite the accuracy, this method requires expensive software and enormous computing capabilities.

• **The physical electrical equivalent circuit model (Compact model):** This method allows for fairly accurate circuit simulation and faster model development, as well as inductor optimisation. In addition, this method captures the principal phenomena that affect the performance of the spiral at high frequencies.

3.2.2 Performance Estimation

1. **Quality factor (Q)**

   The Q is a critical figure of merit to assess the performance of spiral inductors. It represents the efficiency of an inductor to store magnetic energy despite parasitic effects [38]. To a certain extent, Q can be interpreted as the difference between the average stored magnetic energy and average stored electric energy, divided by the total energy dissipated in a signal cycle.

$$Q \equiv 2\,\omega_0 \frac{\text{net energy stored}}{\text{average power loss}} = 2\,\omega_0 \frac{W_m - W_e}{P_R + P_G} \tag{3.1}$$

Q is also often defined as the ratio of the imaginary to the real part of the impedance of the spiral with one of its terminals connected to ground. The simplest and the most widely used definition of the quality factor of a spiral inductor is

$$Q = \frac{\Im(Z_{11})}{\Re(Z_{11})} = \frac{\omega L_s}{R_s} \tag{3.2}$$

where ωL_s and R_s represent the imaginary and the real parts of the complex impedance Z_{11} respectively.

2. **Self-Resonance Frequency**

   Since the spiral inductor exhibits both inductive and capacitive behaviours, the parasitic capacitances will resonate with the inductance at a certain frequency, referred to as self resonance frequency. Generally a spiral inductor that is closer to the substrate or larger in size has a higher total parasitic capacitance and a lower self-resonant frequency.

3. **Inductance and resistance**

   Inductance and resistance are two important parameters to estimate the performance of a spiral inductor. The characteristics of a spiral inductor are shown in Figure 3.3; inductance, resistance and quality factor. Note that at the resonance frequency, where the inductance crosses zero and turns capacitive, is the same point where the quality factor equals zero.

3.2.3 Physical Phenomena

The performance of silicon on-chip spiral inductors has improved with technology advancements, such as the use of copper metallisation, increase the metal thickness (increase the number of strapped metal layers), and the use of lower permittivity dielectrics. Despite these enhancements, a variety of electromagnetic and physical phenomena degrade the performance of on-chip spiral inductors at high frequencies as outlined below.

1. **Skin and Proximity Effect**

   The skin effect is a well-known phenomenon defined as the tendency of current to flow closer to the surface “skin” of a conductive material as the frequency increases

[39]. The depth of the current flow, skin depth (δ), is determined by

$$\delta = \sqrt{\frac{2\rho}{\omega\,\mu_0\,\mu_r}} \tag{3.3}$$

where ρ is the resistivity of the conductive material in Ω.cm, μ_r is the relative permeability of the material, μ_0 is the permeability of free space and $\omega = 2\pi f$ represents the angular frequency. Due to current crowding at the surface and corners of the conductor, the effective cross-sectional area of current flow shrinks and the resistance increases. Inductance is also affected by the skin effect. At low frequencies, the magnetic flux of the EM waves is contained within the conductor as well as outside it. As the frequency increases, the current flows near the surface thereby reducing the field lines inside the conductor, causing substantial EM field attenuation.

Proximity effect, another current crowding phenomenon, is the result of currents flowing in nearby conductors. Proximity effect causes the resistance to increase and the inductance to decrease in a manner similar to the skin effect. At higher frequencies the conductor becomes much thinner due to the skin effects. This enables designers to use thinner metal layers with little impact on the circuit loss.

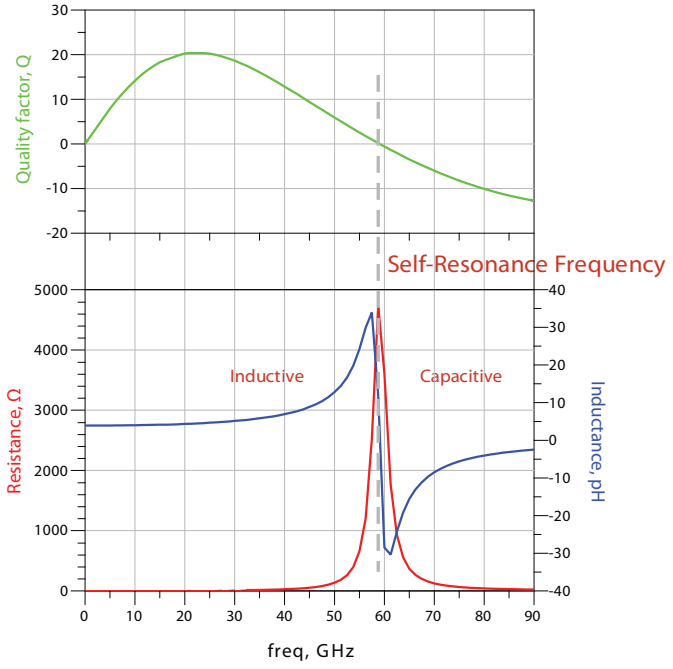


Figure 3.3: On-chip spiral inductors characteristics inductance, resistance and quality factor versus frequency. The marker indicates the self resonance frequency.

Although the resistance increases with frequency, the losses per wavelength can actually decrease.

The skin effect, or current constriction, is non-uniform as a function of the location in the spiral due to the non-uniformity of the magnetic field. At low frequencies the current is nearly uniform whereas at high frequencies non-uniform current flows due to proximity effects. The magnetic field is strongest in the centre of the spiral [40] and thus the time varying magnetic field produces eddy currents of greatest strength in the volume of conductors near the centre of the device. Since at high frequencies current constriction limits the current to the outer edges of the conductors, conductor width does not have as strong influence on minimising metal losses as at low frequencies. For this reason [40] suggests removing the inner turns to produce a 'hollow' spiral. Another approach is to decrease the width of the inner turns and to effectively move these turns closer to the outer edge. This approach contrasts with the approach suggested by [41] where the sum of the metal pitch and spacing, W+S, is kept constant.

2. **Electric Field Induced Substrate Loss**

   Silicon resistivity varies from insulating ($\rho \sim 10\,\mathrm{k\Omega.cm}$) for lightly doped silicon to conductive ($\rho \sim 10 - 0.1\,\Omega\mathrm{.cm}$) for heavily doped silicon. Typically, spiral inductors are fabricated on top of a doped silicon substrate and the bottom of the substrate is connected to ground. Consequently, a time-varying electric field originates from the inductor and mostly terminates at ground. This field penetrates the substrate and introduces loss. The amount of loss depends on the substrate doping, the inductor geometry, the height of the inductor above the substrate and the frequency of operation. Additionally, such an inductor couples noise capacitively into the substrate and is susceptible to noise from the substrate. A patterned ground shield (PGS) was introduced by [38] to address this problem. The PGS terminates the electric field and prevents it from reaching the substrate. This eliminates the electric field induced substrate loss but increases the parasitic capacitance, C_p, of the inductor. The reason behind the patterned rather than solid shield is to avoid eddy current flowing through it, which leads to performance degradation. In the presence of a well-designed PGS, the substrate loss factor can be minimised, leading to an improved quality factor and therefore an improved performance.

3. **Magnetic Field Induced Substrate Loss**

The magnetic fields interaction with the silicon substrate is key to determining the performance of passive RFIC structures. The time-varying magnetic field induces eddy currents in the substrate which cause loss. This loss manifests itself as a decrease in the quality factor of the inductors and increased attenuation in transmission lines. The induced eddy currents in the substrate create their own magnetic field, which opposes that of the spiral inductor, causing a decrease in the net inductance. Although patterned ground shields can prevent interaction of the electric field with the silicon substrate, they do not prevent magnetic field effects. Attempts have been made to eliminate substrate losses, caused by magnetic fields, by techniques such as fabricating suspended inductors by etching wells in the silicon underneath [42, 43], using micro electromechanical system (MEMS) to fabricate inductors far from the silicon surface [44] and using high-resistivity ($> 10\,\mathrm{k\Omega.cm}$) substrates [45].

3.2.4 High Frequency Equivalent Circuit

High frequency equivalent circuit representation of spiral inductors has been researched extensively [1, 46–54]. Figure 3.4(a) shows the equivalent circuit model of a spiral inductor on a doped silicon substrate, including the effects of various phenomena at high frequency. The DC values of the elements in the model are relatively simple to calculate. However the estimation of high frequency values can be very challenging. In the physical model shown in Figure 3.4(b) L_s and R_s are the wiring inductances and resistance respectively. L_s can be computed using Grover [55] and Greenhouse [56] methods. R_s is calculated by using conductor dimensions and the technology parameters such as metal sheet resistance ρ. L_{skin} and R_{skin} describe the skin and proximity effects. L_{skin}, R_{skin}, L_s and R_s can be merged into R_{eff} and L_{eff} for practical implementation reasons. C_s is the capacitance between the terminals of the inductor. C_{ox} is the oxide capacitance between the substrate and the wiring. Substrate resistance R_{si} is predominantly determined by the majority carrier concentration as determined by doping concentrations and the area the inductor occupies [57]. The substrate capacitance C_{si} is the self-capacitance and is attributed to the high frequency effects occurring in the substrate [38, 58]. Simple formulae for calculating the geometry dependent values for the equivalent circuit elements is provided in [59]. R_{sub} and L_{sub} indicate the eddy current effect due to the low

resistivity silicon substrate. This effect can be minimised by adding a substrate shield as shall be discussed later in this section.

3.2.5 Q Factor Improvement Methods

Poor quality factor leads to circuit inefficiencies and can affect the overall performance of the circuit. For example, if an inductor with a very poor Q was used in a matching network for a low noise amplifier, it may degrade the noise figure significantly and increase distortion in the circuit. An accurate calculation of a silicon on-chip inductor quality factor operating at high frequencies can be derived from the simple pi-model illustrated in Figure 3.4(b) and (c) as follows [60]:

$$Q = \underbrace{\frac{\omega L_s}{R_s}}_{\textbf{Ideal Q}} \underbrace{\frac{R_p}{R_p + \left[\left(\frac{\omega L_s}{R_s}\right)^2 + 1\right] R_s}}_{\textbf{Substrate Loss Factor}} \underbrace{\left[1 - \frac{R_s^2}{L_s}(C_s + C_p) - \omega^2 L_s (C_s + C_p)\right]}_{\textbf{Self Resonant Factor}} \tag{3.4}$$

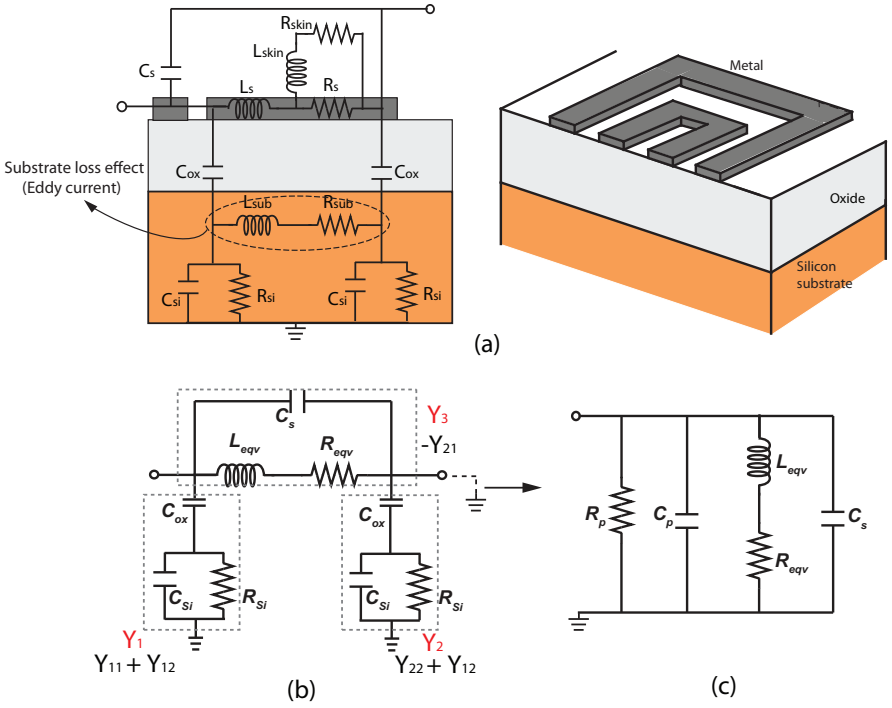


Figure 3.4: Cross section of an on-chip spiral inductor with associated parasitics at millimetre-wave frequencies: (a) Equivalent circuit model including effects of various phenomena, (b) simplified equivalent circuit pi-model and (c) single ended circuit model for quality factor calculation.

where

$$R_p = \frac{1}{\omega^2 C_{ox}^2 R_{si}} + \frac{R_{si}(C_{ox} + C_{si})^2}{C_{ox}^2} \tag{3.5}$$

$$C_p = C_{ox}\frac{1 + \omega^2(C_{ox} + C_{si})C_{si}R_{si}^2}{1 + \omega^2(C_{ox} + C_{si})^2 R_{si}^2} \tag{3.6}$$

The first term of the Equation 3.4 represents the ideal quality factor, while the second term is the substrate loss factor which accounts for the energy that is dissipated in the substrate. The third term is the self-resonance factor. It can be observed from Equation 3.4 that in order to improve the quality factor of an on-chip inductor the designer can:

1. **Reduce the series resistance (R_s):** By using wider metal tracks or using a metal with greater conductivity (e.g., copper, gold, silver) to reduce the series resistance improves inductor quality factor (Q). This method is the least effective since at high frequencies the skin effects force the signal to flow in a very thin area on the outer crust of the metal tracks. Therefore increasing the thickness of the tracks will have a minimal effect on the resistance.

2. **Reduce parasitic capacitance (C_p):** Placing the inductor as far away from the substrate as possible reduces C_p and improves the self-resonance factor.

3. **Eliminate the substrate loss factor**: By making the loss factor equal to unity the substrate loss is eliminated. This is achieved when R_p is ∞. This is possible when the silicon substrate is open. An insulating substrate [61] or a totaly etched substrate [62] or doping the substrate very lightly or very heavily[1] can make an open substrate. On the other hand, short substrate can be achieved by using patterned ground shielding layer between the inductor and the substrate [38, 63].

3.2.6 Implemented Design Methodology

It should be recognised that all the above-mentioned methods of improving the quality factor are limited by technological constraints. In view of this, seeking a method

[1]Doping the substrate very lightly or very heavily results in a very large R_p. This decreases the electrically induced substrate losses. Other high-frequency effects, however, induce loss in heavily doped substrates.

that abides by the technology rules and keeps within the technology constraints is very important. In this work, an optimised spiral inductor with a patterned shield layer is introduced. Two types of shield configurations in two forms (floating and grounded) are presented, compared and discussed.

3.2.6.1 Patterned Shielded Structure

In an attempt to reduce the substrate losses and improve the quality factor of silicon spiral inductors, researchers proposed creating a shield, made of lower metal layers or polysilicon, directly underneath the passive device. The shield can be grounded or floating. Figure 3.5 illustrates some of the patterned shields proposed by researchers. In order to comply with the design rules associated with CMOS technology, the shield layer can not be a solid sheet of metal. A shield should have a slotted pattern, positioned orthogonal to the transmission lines or spiral tracks, to reduce the effect of negative mutual coupling according to Lenzs Law[2] [64]. The slots act as an open circuit to the loop current, but must be sufficiently narrow to prevent leakage of the vertical electric field. Although a patterned shield seems like a very effective way to improve the quality factor, it can present several unwanted problems such as the significant reduction of self-resonant frequency caused by the increased parasitic capacitance (C_p). Therefore, the patterned shield design, layout, layer choice and shape, should be all considered carefully to take into account all the possible effects and drawbacks that may degrade the performance of passive devices. The shielded structures implemented in this work are shown in Figure 3.6. The patterned structure is a star-shape grounded and floating shield, illustrated in Figure 3.6 (a) and (b) respectively. The shield is constructed using strips of diffusion, silicided polysilicon and metal 1, as illustrated in Figure 3.6(c). These layers, based on their distance away from the spiral inductor tracks (implemented on metal 6), were chosen to minimise the parasitic capacitance. The width and the spacing between the shield strips are set to the minimum value allowed by the technology design rules, which is 0.45 μm wide strips spaced by 0.45 μm. The patterned shield, presented in Figure 3.6(a), is composed of 4 segments. Each segment consists of strips positioned transversely to the inductor tracks above it. Metal 1 and the diffusion strips are aligned above each other. Polysilicon strips fill in the gaps between

[2] An induced electromagnetic force (EMF) always gives rise to a current whose magnetic field opposes the original change in magnetic flux

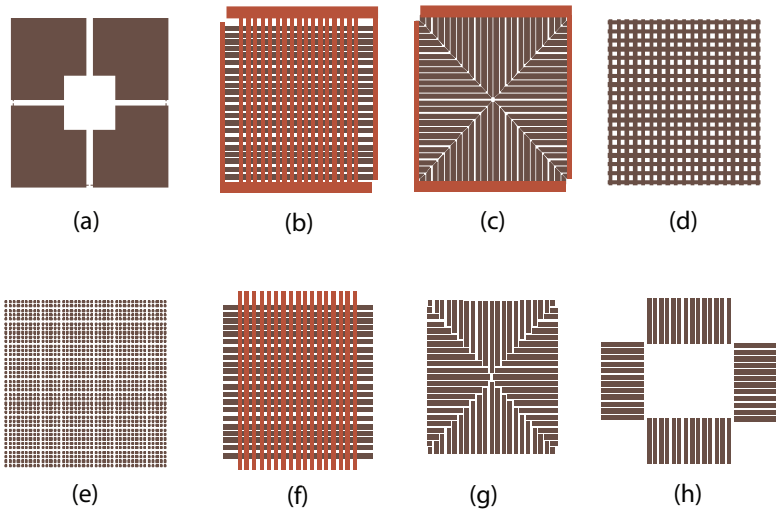


Figure 3.5: Patterned shield layers design: Patterned ground shield (top row) and patterned floating shield (bottom row).

the latter strips as shown by the 3D illustration in Figure 3.6. The grounded shield is formed by adding a star-shape strapping, as shown in Figure 3.6(a), that connects all the shield strips to the ground contact connected to the centre of the shield. The ground contact is strapped with the top layer metal using adequate contacts to provide a low-impedance path to ground. The presence of the patterned shield affects the high frequency equivalent circuit model of the spiral inductor. Figure 3.7 illustrates the equivalent circuit model for a spiral inductor with the presence of a patterned shield,

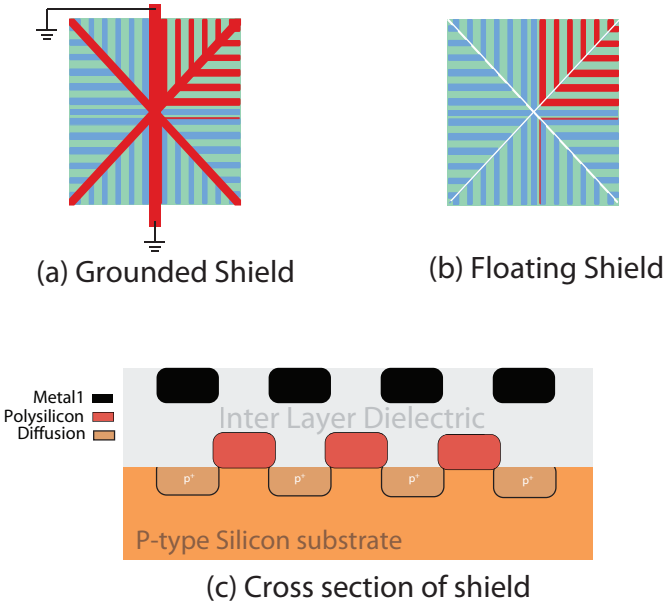


Figure 3.6: Designed patterned shield for silicon on-chip spiral inductors illustrating the star-shape configuration in two different forms (floating and grounded): (a) Star-shape patterned grounded shield; (b) star-shape patterned floating shield and (c) illustrates the side view of a cross section of the shield structure.

in both configurations grounded and floating, with circuit parameters to model the behaviour of the relevant high frequency effects. It can be seen that the spiral inductor with a patterned shield exhibits a parasitic capacitance C_p and a parasitic resistance R_{shld}. The parasitic capacitance is the capacitance between the spiral tracks and the shield. The value depends on the geometric parameters of the spiral and the distance separating the shield from the tracks. The parasitic resistance originates from the finite distance that the distributed capacitive current must flow through before it reaches the ground contact. The shorter the path the smaller the resistance. This was the concept behind the star-shape ground shield. The factor of 0.5 appears in the equivalent circuit of the spiral inductor because half the capacitance and the resistance are distributed in each leg of the equivalent circuit.

3.2.7 Experimental Results and Discussion

A comparative study was performed on three identical spiral inductors with different shielding topologies; unshielded, floating patterned shield and grounded patterned shield.

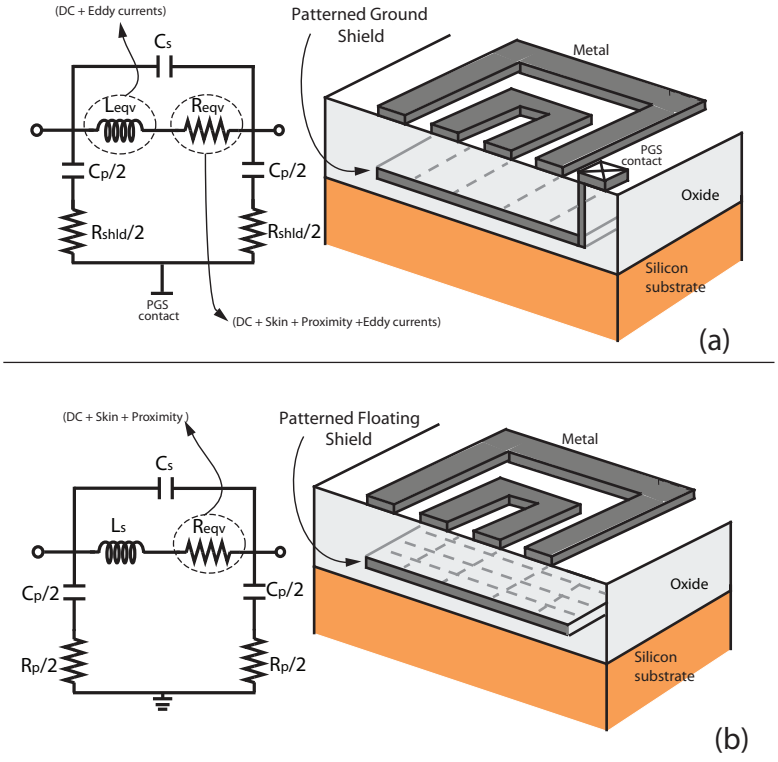


Figure 3.7: Equivalent circuit model for a spiral inductor with both patterned ground and floating shield, showing the relevant high frequency effects such as skin/proximity effects, eddy current and shield parasitics effects.

The spiral inductors were fabricated using a 1 poly, 6 metal layer 0.18 μm bulk CMOS process. The structures are two-port, formed of 2 turns and designed to have a 300pH dc inductance. These structures are constructed with the upper metal layer and with a conductor width of 9 μm, 2 μm spacing between the inductor tracks, a diameter of 50 μm and a total length of 450 μm, surrounded by a 25 μm wide ground guard ring formed of a metal stack (metal 2 up to metal 6).

In order to characterise the spiral inductors, a symmetric equivalent circuit lumped elements models were extracted from the measured data. The model can be extracted using one of two methods; Figure 3.8(a) depicts the equivalent circuit model extraction method of a grounded two-port spiral using the complex propagation constant γ and characteristic impedance Z_0 [1]. In contrast, Figure 3.8(b) depicts a symmetric two-port spiral inductor extraction method using Z and Y-parameters fitting method. The extracted model parameters of the spiral inductors discussed in this study are listed in Table 3.1.

Figure 3.9 compares the extracted inductance (a) and quality factor (b) of the unshielded and shielded (grounded and floating) spiral inductors discussed earlier in this study. It is evident from Figure 3.9 and Table 3.1 that shielding improves the quality factor

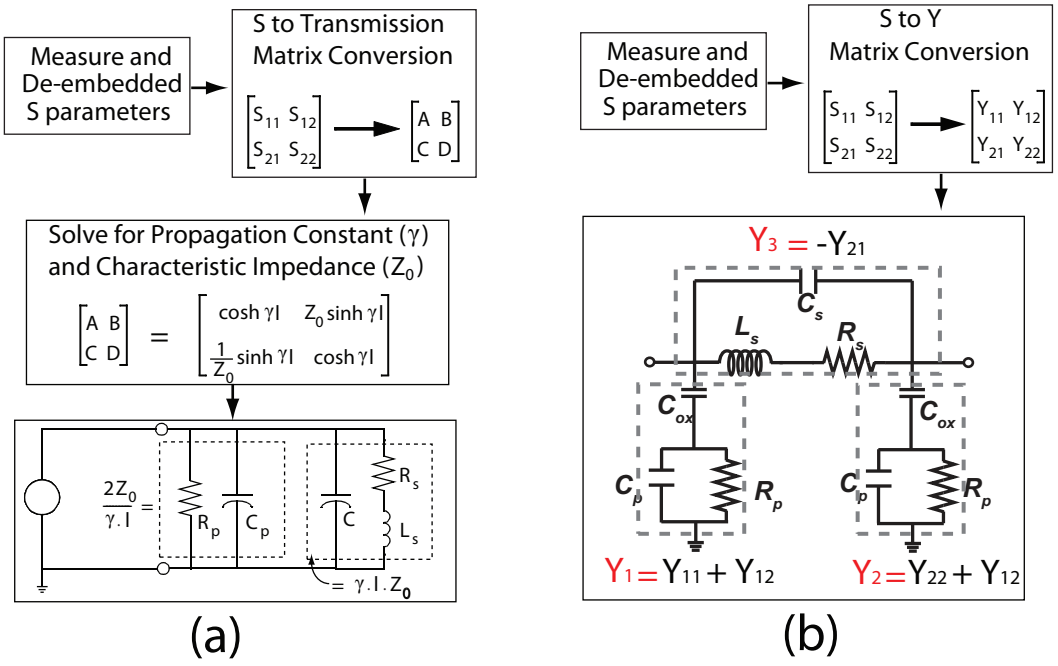


Figure 3.8: Testing and parameter extraction procedure of a silicon spiral inductor: (a) Extraction using the complex propagation constant γ and characteristic impedance Z_0 of the inductor proposed by [1] and (b) Y-parameter fitting extraction method.

Table 3.1: Extracted model parameters and the performance metrics of the spiral inductors listed in this study: Unshielded, Patterned grounded shield (PGS) and Patterned floating shield (PFS).

Spiral	**Model parameters**							
Inductor	L_s (pH)	R_s (mΩ)	C_{ox} (fF)	C_p (fF)	R_p (Ω)	C_s (fF)	$f_{resonance}$	Q
Unshielded	280	1260.37	215	0.1	580.4	8	31 GHz	≈8
PGS	275	1100	2.79	0.1	1000	5	44.8 GHz	≈18
PFS	260	1800	4.8	0.1	130	5	72 GHz	≈23

significantly (by a factor of 3) with a small decrease in the inductance value (≈5pH). The resonance frequency of the shielded structures is increased by a factor of 2.2 in comparison to the unshielded structure. While this highlights the merits of shielding the spiral inductor as part of millimetre-wave circuit design on silicon, the question remains as to which topology, floating or grounded, is more advantageous.

The spacing between the spiral tracks and the ground guard ring is 30 μm and the oxide thickness between the inductor tracks and the shielding layer is 5.4 μm. When the spiral shield is grounded the electric field from the spiral tracks couple to the ground shield through the 5.4 μm thick oxide, the oxide thickness dictates the spiral line capacitance. On the other hand, the floating shield channels the coupled electric field from the spiral tracks to the ground guard ring through the oxide. This halves the line capacitance since the effective oxide thickness is now doubled (10.8 μm). These effects can be seen in Figure 3.9 (a) where the resonant frequency of the grounded shield is lower than that of the floating shield by a factor of 1.75. Furthermore, the floating shield inductor and the patterned ground shield comparison reveals an improvement in the quality factor (21%) and an increase in the inductance (approximately 15pH). This analysis confirms the benefits of using spiral inductors, with a PFS, when designing circuits operating at millimetre-wave frequencies.

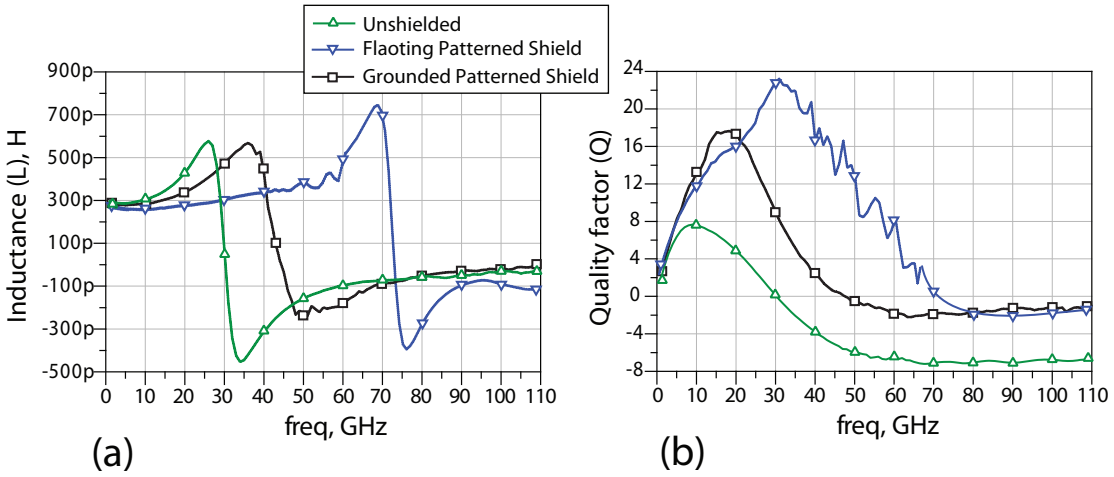


Figure 3.9: CMOS spiral inductors (unshielded, grounded and floating) electrical parameters comparison: (a) Inductance value and (b) quality factor.

3.3 CMOS Transmission Lines

3.3.1 Background

Transmission lines are very important in RFIC circuit design. As the frequency of operation increases the wavelength decreases, leading to shorter electrical length which can be easily implemented on chip. A transmission line is a distributed structure whose voltages and currents can vary in magnitude and phase over its length. It can be characterised by its equivalent distributed circuit model shown in Figure 3.10 [65].

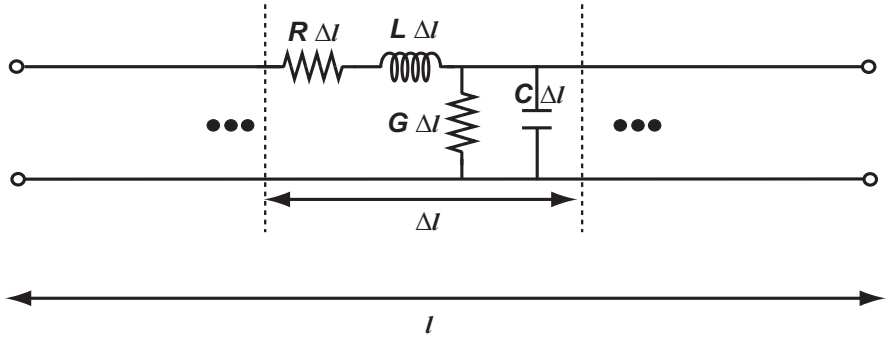


Figure 3.10: Transmission lines distributed model.

The short piece of the transmission line, of length $\triangle l$ is modelled as a frequency dependent lumped element circuit where R and L are the series resistance and inductance per unit length. G and C are the shunt conductance and capacitance per unit length. The R, L, G and C parameters that characterise a transmission line can be related to

its characteristic impedance (Z_0) and its complex propagation constant (γ) by [65–67]

$$Z_0 = \sqrt{\frac{R + j\omega_0 L}{G + j\omega_0 C}} \tag{3.7}$$

$$\gamma = \sqrt{(R + j\omega_0 L)(G + j\omega_0 C)} = \alpha + j\beta \tag{3.8}$$

$$\frac{\gamma}{Z_0} \equiv j\omega_0 C + G \tag{3.9}$$

$$\gamma\, Z_0 \equiv j\omega_0 L + R \tag{3.10}$$

where α and β are the attenuation and phase constants and that can be approximated as follows:

$$\alpha \approx \frac{R}{2Z_0} + \frac{GZ_0}{2} \tag{3.11}$$

$$\beta = \frac{2\pi}{\lambda_g} = \omega_0\sqrt{LC} \tag{3.12}$$

When a transmission line is used as a resonator in a matching network, the relevant definition of the quality factor (Q) of the line is [68]

$$Q_L = \frac{\omega_0 L}{R} \tag{3.13}$$

$$Q_C = \frac{\omega_0 C}{G} \tag{3.14}$$

$$\frac{1}{Q_{resonator}} \approx \frac{1}{Q_L} + \frac{1}{Q_C} \tag{3.15}$$

The relation between Q, α and β of a transmission line is given as [69]

$$Q = \frac{\beta}{2\alpha} \tag{3.16}$$

Among the various types of transmission lines, the microstrip line and the coplanar waveguide, illustrated in Figure 3.11, are most suited for silicon integrated circuits.

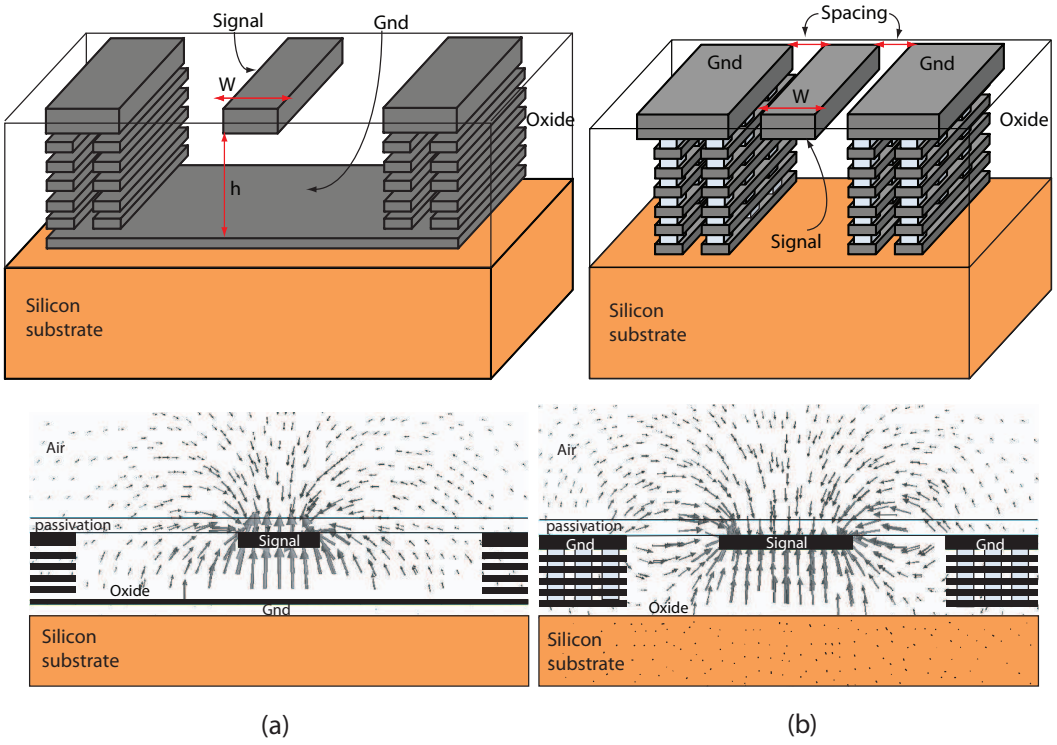


Figure 3.11: RFIC transmission lines types: (a) Simplified 3D and cross-section of the microstrip line with sidewalls and (b) simplified 3D and cross-section of the conventional coplanar waveguide (CPW).

3.3.1.1 Modelling for CAD

Several CAD packages were used in this work; Agilent ADS suite, AWR Microwave office, Ansoft HFSS and Cadence IC design tools. The majority of commercial microwave circuit simulators provide a length scalable model that captures the physical behaviour of the transmission line with good accuracy up to very high frequencies, often called "TLINP". This is simply obtained by optimising the model parameters to fit the measured S-parameters of a transmission line with a specific geometry. Once the model is achieved, and due to its length scalability, it can be used to model transmission lines of the same geometry of any length, by changing the length parameter in the model. This provides a significant advantage when it comes to designing circuits that utilise transmission lines, since all the line lengths can be tuned or optimised (using the optimisation suite) in order to obtain specific requirement, for example to realise a given impedance in a matching network. The TLINP model contains the following parameters; the characteristic impedance Z_0 (Ω), the physical length L (m), the effective dielectric constant ϵ_{re}, the attenuation A (dB/m), the frequency for scaling attenuation F (GHz) and finally the

dielectric loss tangent TanD. The main issue with TLINP models is that they can only be used for transmission lines with specific geometries that have been fabricated and measured. For geometry independent models the full-wave electromagnetic field simulator HFSS can be used. The only disadvantage associated with 3D full field analysis simulators such as HFSS is that transmission line construction and simulation can take hours or days to complete, depending on the complexity of the structure. Generally full field analysis possesses many advantages over equivalent circuit model based simulators, one of which is much higher accuracy. Once the simulation results of the transmission lines are obtained, a TLINP model is created for it to ease the circuit design process and reduce the design time.

3.3.1.2 Microstrip vs. Coplanar Transmission line

Microstrip lines are implemented on silicon by using two metal layers separated by an oxide layer grown on top of the silicon substrate. The top most layer, which offers the lowest sheet resistance and highest distance from substrate, is the signal line and the bottom layer is the ground plane. In standard CMOS, the distance between the back end of line (BEOL) metal layers cannot be altered, hence the distance between the signal line and the bottom ground plane (h) is fixed ($\sim 4 - 6\,\mu$m). The close proximity of the signal line and the ground plane leads to a large capacitance and small inductance per unit length. According to Equation 3.13 and Equation 3.14, this results in a low inductive and a large capacitive quality factor. In an attempt to improve the line inductance, microstrip lines provide another geometry design variable, the signal line width (W). This can be decreased leading to an increased inductance. However, reducing the line width will increase the line resistive losses resulting in quality factor degradation. A serious disadvantage of microstrip transmission lines in CMOS is the limited design freedom. Since (h) is fixed, the only variable is (W) and once the desired characteristic impedance has been chosen, there is only one signal width which will meet this requirement, and therefore no possibility of optimisation. The main advantage of this type of transmission line is the complete shielding from substrate losses. The ground layer forms a metal shield, stopping the electric field from penetrating into the silicon substrate ($\epsilon_r = 11.9$). Hence the shunt loss G is in the oxide layer ($\epsilon_r \approx 3.9$).

On the other hand, coplanar transmission lines are implemented entirely in the top most metal layer and their characteristic impedance (Z_0), distributed inductance and capacitance are mostly determined by the width (W) of the signal line and the gap spacing (S) between the signal line and the ground planes. This provides an additional degree of freedom in their design as the signal line width can be optimised for minimum resistive lossess and the gap spacing can be optimised for the inductive and capacitive quality factor tradeoff. These parameters can be used to realise different characteristic impedances, however this could lead to an unwanted slotline mode. This can be eliminated by connecting the ground planes of each side of the signal line together with underpasses implemented on lower metal layers [70]. In this way the ground planes of the CPW remain at the same potential along the entire length the transmission line.

When implementing a transmission line in CMOS the metal density requirement has to be fulfilled, which means that there has to be adequate metal at all metal levels. This is achieved by connecting all the other metal layers together with vias to form the ground plane for the CPW. In the microstrip case, adding ground planes similar to a CPW line fulfills the metal density requirements. Due to the metal density requirements the wide ground plane on the lower metal level must have longitudinal slots, which do not interfere with the longitudinal ground currents of the microstrip line. As the distance between the signal line and ground plane (h) is rather low when the side ground planes are located far from the centre conductor, the signal propagates mostly in microstrip mode.

To achieve the optimum performance transmission lines in a CMOS process one can use the design flexibility of the CPW line and the shielding capability of the microstrip line. The use of grounded, shielded and slow-wave CPW lines can improve the performance, by reducing losses, when compared to a conventional coplanar waveguide [71].

3.3.1.3 Slow-wave Coplanar Waveguide: Operation and Concept

Previous studies have explored the possibility of reducing the silicon substrate losses by isolating the transmission line from the substrate using a solid sheet metal shield. Such a configuration is referred to as grounded coplanar waveguide (CPWG). Nevertheless losses caused by the eddy current, formed in the solid metal sheet, makes this technique very impractical. Inserting a solid metal shield on top of the silicon substrate significantly

reduces h, leading to a significant reduction in the quality factor and the characteristic impedance, hence making it more difficult to achieve 50 Ω characteristic impedance.

To overcome the limitations of CPW and CPWG, new topologies of coplanar transmission lines with improved performances have been investigated. The first CPW structure, on GaAs substrate, with a floating strips was introduced by [72]. An almost identical structure to the latter but with the floating strips above the CPW line was demonstrated by [73]. This structure was called a "cross-tie overlay CPW". These transmission lines displayed a slow-wave propagation phenomenon but suffered from large attenuation of 0.08 to 0.2 Np/mm. An attempt to implement a CPW with floating strips on a BiCMOS technology (10 Ω.cm silicon substrate) can be found in [74]. The results showed that from 15 to 40 GHz, a 50 Ω slow-wave CPW with floating shield achieves a quality factor ranging between 25 to 35.

The phase velocity of a CPW line is defined as

$$\upsilon_p = \frac{\omega}{\beta} = \frac{c}{\sqrt{\mu_r \epsilon_{re}}} = \frac{1}{\sqrt{L_l\, C_l}} \tag{3.17}$$

$$\epsilon_{re} = c^2 \left(\frac{\beta}{\omega}\right)^2 \tag{3.18}$$

where μ_r is the relative permeability and is equal to 1 in most mediums of interest. The speed of light in free space (c) is approximately $3{\times}10^8$m/s and ϵ_{re} is the effective relative permittivity (effective dielectric constant).

The phase constant can be calculated using

$$\beta = \omega \sqrt{L_l\, C_l} \tag{3.19}$$

where L_l and C_l are the inductance and capacitance per unit length.

Substituting Equation 3.19 in Equation 3.18 gives

$$\epsilon_{re} = c^2\, L_l\, C_l \tag{3.20}$$

According to Equation 3.17, a slow-wave phenomena in a transmission line can be achieved by either increasing the effective relative permittivity (ϵ_{re}), increasing the inductance of the line or increasing the capacitance of the line. Increasing the inductance

of the line degrades the quality factor as the increase in the conductor length creates more losses causing an increase in the signal attenuation. On the other hand, increasing the capacitance and ϵ_{re} can potentially improve the quality factor. Equation 3.20 denotes that any increase in the capacitance of the line would lead to a higher ϵ_{re}. By introducing a shield layer underneath the conventional CPW line, this would increase the capacitive effects between the line and the shield. This manifests as an increase in ϵ_{re} which leads to an increase in the phase constant (β), hence a decrease in the phase velocity υ_p thereby creating a slow-wave phenomena in the transmission line. The shield used in a slow-wave transmission line is constructed using equally spaced floating metal strips placed underneath the line itself. The choice of shield metal layer affects the capacitance and ϵ_{re}. The presence of the patterned strips shield, grounded or floating, maintains the line inductance of a conventional line, this is because the shield strips allow the magnetic field to flow through the substrate. As for the electric field, the shield gaps are small enough to stop it from leaking through to the substrate, therefore the field is focussed between the line and the floating shield.

Slow-wave transmission lines exhibit shorter physical lengths than conventional lines for a given electrical length, this is due to the increase in β which leads to a shorter wavelength λ, hence a decrease in the physical length according to

$$\theta = \beta l \tag{3.21}$$

Since the quality factor of the line is directly proportional to β, as shown in Equation 3.16, this indicates that a slow-wave transition line improves significantly in comparison to the conventional topology.

3.3.2 Experimental Results and Discussion

In this work, five CPW lines were fabricated, measured and compared; one conventional CPW, two slow-wave floating strips shield (S-CPW) lines and two slow-wave grounded strips shield (SG-CPW) lines. These structures are shown in Figure 3.14. The lines have a signal/ground line width of 15 μm/45 μm, with a 5 μm space between signal and ground lines. For all the transmission lines the signal line was constructed using the top most metal layer (metal 6). The ground planes were formed of a stack of all available

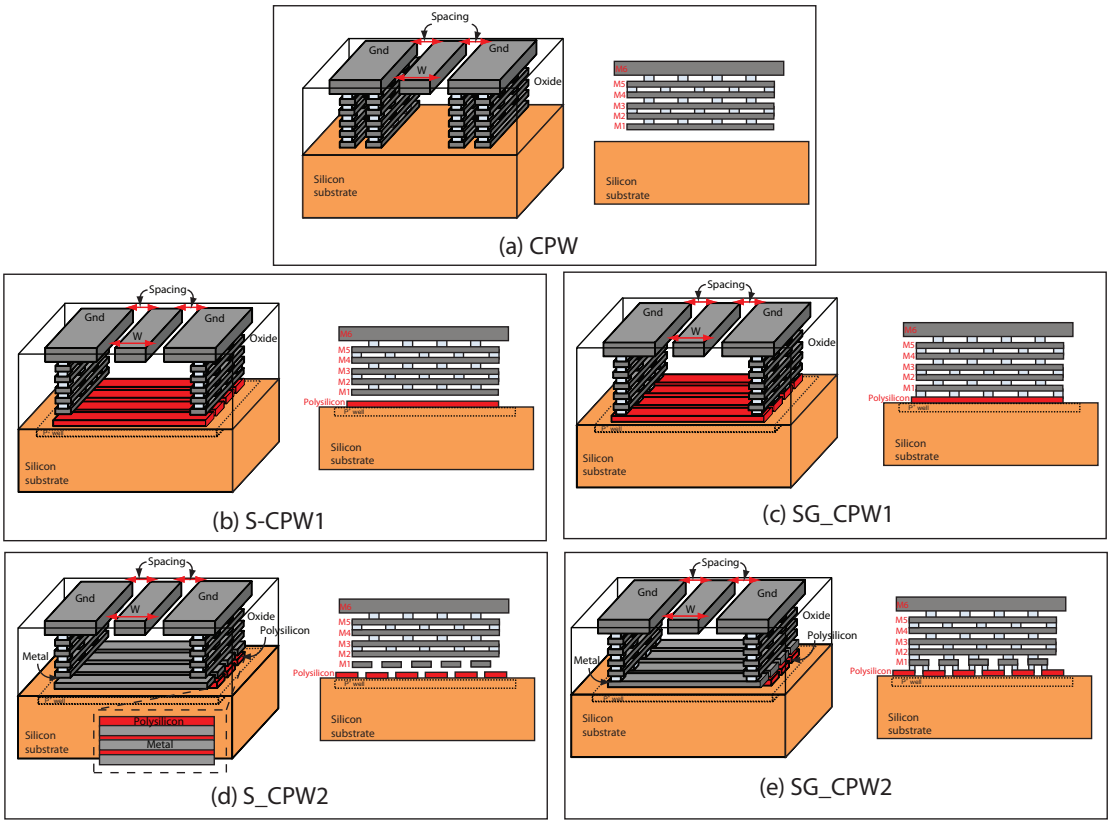


Figure 3.12: Designed CMOS coplanar waveguide (CPW) lines: (a) A conventional CPW line, (b) a slow-wave, polysilicon floating shield, CPW line, (c) a slow-wave, polysilicon grounded shield, CPW line, (d) a slow-wave, stacked floating shield, CPW line and (f) a slow-wave, stacked grounded shield, CPW line.

metal layers (metal 6 down to metal 1). The floating and grounded CPW transmission lines were designed with periodically slot-type shields, formed from strips of polysilicon (the lowest thickness layer), to reduce the losses due to eddy current. The shield strips are located directly beneath the CPW structure and oriented transversely to it. The width of the strips and the spacing between them were chosen to be 0.5 μm, this is the minimum allowed by technology design rules. The performance of the transmission lines is characterised by the following parameters: attenuation (α) in dB/mm, phase constant (β) in rad/mm, effective dielectric constant (ϵ_{re}), and the quality factor (Q). For each transmission line a line of 500 μm in length was measured, de-embedded and the previously mentioned parameters were extracted as follows: The measured S-parameters are converted to ABCD parameters, then the characteristic impedance (Z_0), and the

propagation constant (γ) of each line are calculated using

$$[S_{TL}] \rightarrow [ABCD_{TL}] = \begin{bmatrix} \cosh(\gamma\, l) & Z_0\, \sinh(\gamma\, l) \\ \frac{\sinh(\gamma\, l)}{Z_0} & \cosh(\gamma\, l) \end{bmatrix} \tag{3.22}$$

$$Z_0 = \sqrt{\left(\frac{B_{TL}}{C_{TL}}\right)} \tag{3.23}$$

$$\gamma = \frac{\text{acosh}(A_{TL})}{l} \tag{3.24}$$

$$\alpha = re(\gamma)(nepper/m) = re(\gamma) \times 8.685\,(dB/m) \tag{3.25}$$

$$\beta = imag(\gamma)\,(rad/m) \tag{3.26}$$

$$\epsilon_{eff} = c^2 \left(\frac{\beta}{\omega}\right)^2 \tag{3.27}$$

The equivalent circuit lumped parameters are then extracted using

$$R = re(Z_0\,\gamma)\,(\Omega/m) \tag{3.28}$$

$$L = \frac{imag(Z_0\,\gamma)}{\omega}\,(H/m) \tag{3.29}$$

$$C = \frac{imag\left(\frac{\gamma}{Z_0}\right)}{\omega}\,(F/m) \tag{3.30}$$

$$G = re\left(\frac{\gamma}{Z_0}\right)(S/m) \tag{3.31}$$

The loss tangent is then calculated as follows:

$$Tan\delta = \frac{G}{C\,\omega} \tag{3.32}$$

3.3.2.1 Conventional CPW Versus Slow-wave (S-CPW)

Figure 3.13 compares the attenuation per unit length (in dB/mm) of the conventional CPW (CPW) and the slow-wave CPWs (S-CPW1 and S-CPW2). It is evident that the slow-wave lines (S-CPW) have lower loss per unit length when compared to the conventional CPW, as was expected. As can be seen in Figure 3.13, the conventional

CPW exhibits an attenuation of 0.98 dB/mm compared to 0.36 dB/mm at 20 GHz, and 1.98 dB/mm compared 0.59 dB/mm at 40 GHz.

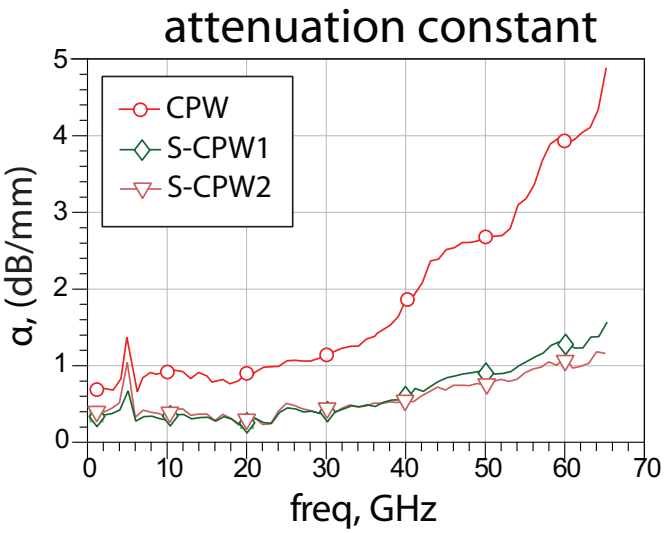


Figure 3.13: Attenuation per unit length of the conventional CPW (CPW) and the slow-wave CPW (S-CPW1 and S-CPW2) on silicon.

Figure 3.14(a) and (b) illustrate the relative dielectric constant and the phase constant of the conventional CPW and the slow-wave CPWs. The effective dielectric constant of the conventional CPW is approximately 3. This is to be expected as the surrounding material is silicon oxide and air. This demonstrates that the effective dielectric and the wavelength of a conventional CPW can never be altered as they are determined by the surrounding materials. Whereas, those of a slow-wave CPW can be adjusted by changing the structure physical layout as discussed earlier in this section. It is evident from Figure 3.14(a) that the dielectric constant of S-CPW1 and S-CPW2, 13.6 and 13.4 respectively, are roughly 5 times higher than that of the conventional CPW line implemented on silicon. It can be seen from Equation 3.27 that ϵ_{er} is inversely proportional to the wavelength and directly proportional to the phase constant. The slow-wave CPW therefore exhibits double the phase constant in comparison to conventional CPW. This renders the slow-wave CPW to be half the wavelength of its counter part the conventional CPW. This is particularly useful when a more compact transmission line implementation is required for impedance matching networks and phase shift applications.

Figure 3.15 depicts the quality factor of a CPW, S-CPW1 and S-CPW2. The slow-wave CPW line exhibits an approximately 50% higher quality factor, at the peak frequency,

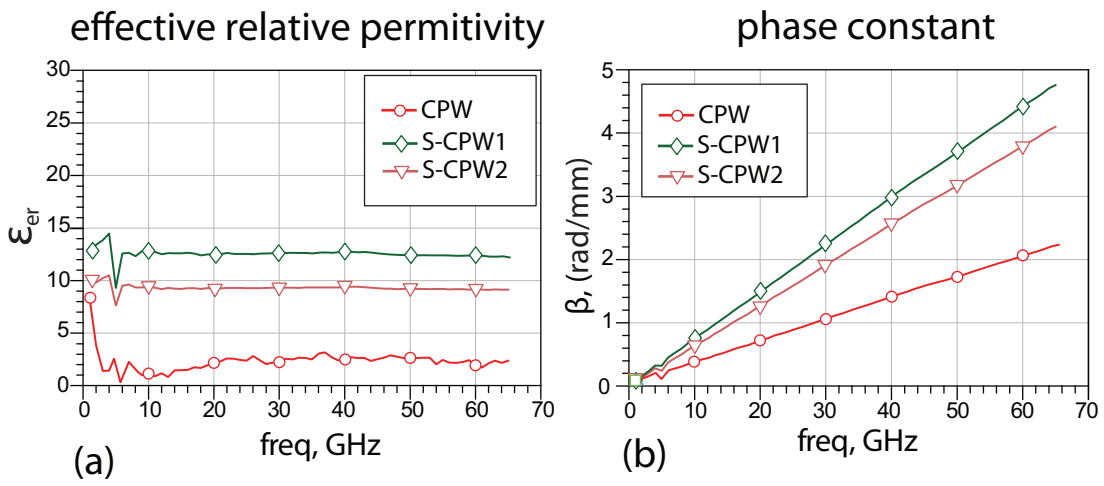


Figure 3.14: (a) Relative dielectric constants and (b) phase constant per unit length of the conventional CPW (CPW) and the slow-wave CPW (S-CPW1 and S-CPW2) on silicon.

than the conventional CPW. S-CPW presents a quality factor of 22 in comparison to 11 for the conventional CPW. This is expected as the shielding layer under the slow-wave line reduces the electromagnetic field penetrating the lossy substrate. Due to the resultant reduction in line attenuation, a high quality factor is realised for this structure. Contrary to this, the unshielded conventional CPW allows the electromagnetic field to penetrate the lossy silicon substrate causing high losses which in turn leads to low quality factor.

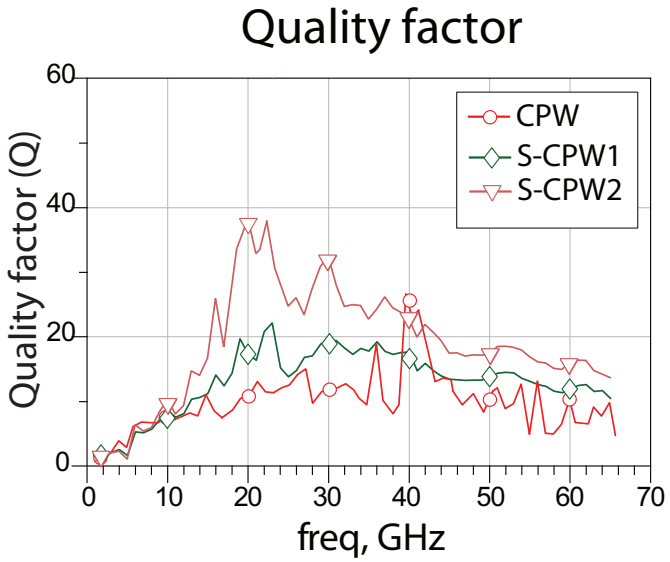


Figure 3.15: Quality factors of the conventional CPW (CPW) and the slow-wave CPW (S-CPW1 and S-CPW2) on silicon.

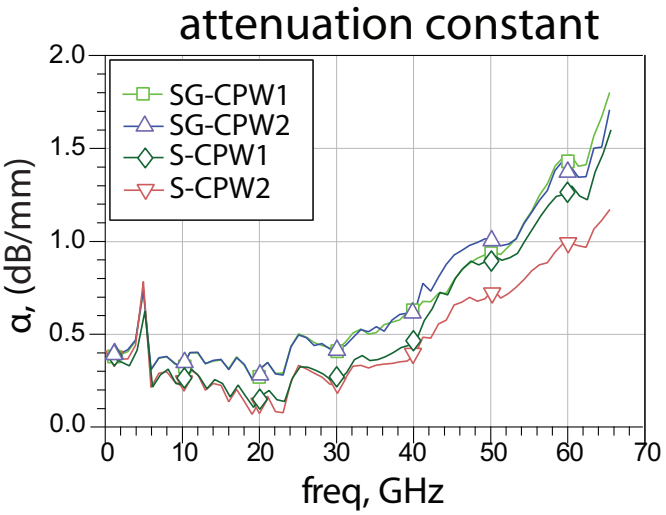


Figure 3.16: Attenuation per unit length of the slow-wave CPW (S-CPW1 and S-CPW2) and grounded slow-wave CPW (SG-CPW1 and SG-CPW2) on silicon.

3.3.2.2 Slow-wave CPW: Floating Shield (S-CPW) Versus Grounded Shield (SG-CPW)

In a slow-wave CPW the electric field couples from the signal track through the oxide thickness, down to the shield and back to the ground planes on both sides of the line. In the case of floating shield, since the shield has no potential, the floating shield forms a channel for the electric field. This results in coupling from the signal track, through the oxide thickness and shield, to reach the ground planes on each side of the signal line. This effectively doubles the oxide thickness, resulting in a smaller line capacitance, higher effective dielectric constant, higher phase constant and therefore shorter wavelength. In contrast, when the shield is grounded the electric field couples through the oxide thickness to the ground shield directly. Consequently the oxide thickness is halved and the line capacitance is increased.

Figure 3.16 compares the attenuation per unit length (in dB/mm) of the floating shield slow-wave CPW (S-CPW1 and S-CPW2) and ground shield slow-wave CPW (SG-CPW1 and SG-CPW2) structures. It can be seen that the ground shield slow-wave structures demonstrate similar attenuation per unit length across the entire measurement spectrum (0.29, 0.3 dB/mm at 20 GHz and 0.49, 0.5 dB/mm at 40 GHz). While the floating shield slow-wave (S-CPW1 and S-CPW2) demonstrate comparable attenuation per unit length up to 30 GHz (0.18, 0.15 dB/mm at 20 GHz), at higher frequencies, S-CPW1 attenuation starts to increase. This is due to the increased losses in the polysilicon shield and the

losses caused by the magnetic field penetration through the lossy silicon substrate. On the other hand, S-CPW2 maintains a low attenuation constant as it is designed to reduce the losses associated with polysilicon by using a second shield layer of metal 1 (lower sheet resistance). Additionally the configuration provides complete isolation from the silicon substrate, therefore reducing the magnetic penetration significantly. Furthermore the attenuation of the floating and grounded slow-wave lines is similar below 10 GHz. However, the attenuation of the ground shield SG-CPW1 and SG-CPW2 is 0.2 dB/mm higher, from 20 GHz up to 40 GHz, and at 60 GHz is 0.48 dB/mm higher than the floating shield S-CPW1 and S-CPW2. It is evident that floating shield slow-wave lines exhibit lower attenuation constant when compared to ground shield slow-wave. S-CPW2 demonstrates the lowest attenuation, approximately 30% lower attenuation per unit length, compared to all the other slow-wave structures.

The results in Figure 3.17 show that the quality factor (Q) of S-CPW2 with a floating shield is higher than the other slow-wave structures by a factor of 1.5. Finally a comparison of all the presented transmission lines is listed in Table 3.2. This provides a summary of the transmission line performance at 40 GHz, and illustrates the quality factor value at the peak frequency (20 GHz). These values are used to create the transmission lines CAD model discussed in the next section.

From the presented results, it can be seen that the floating shield slow-wave CPW can be used to achieve high quality factor at high frequencies. Slow-wave CPW structures exhibit shorter wavelength as a result of the high phase constant. This enables compact designs and reduces the required chip area.

Table 3.2: Comparison of model parameters of the transmission lines used in this study at 40 GHz.

Transmission line	**Parameters**				
	Z_0 (Ω)	ϵ_{re}	α (dB/mm)	β (rad/mm)	Q
CPW	46.7	2.6	1.97	1.43	11@20G
SG-CPW1	54	13.6	0.81	3.24	20@20G
SG-CPW2	53.4	13.4	0.82	3.16	22@20G
S-CPW1	49.2	12.4	0.84	2.98	18@20G
S-CPW2	56.4	9.3	0.5	2.52	37.8@20G

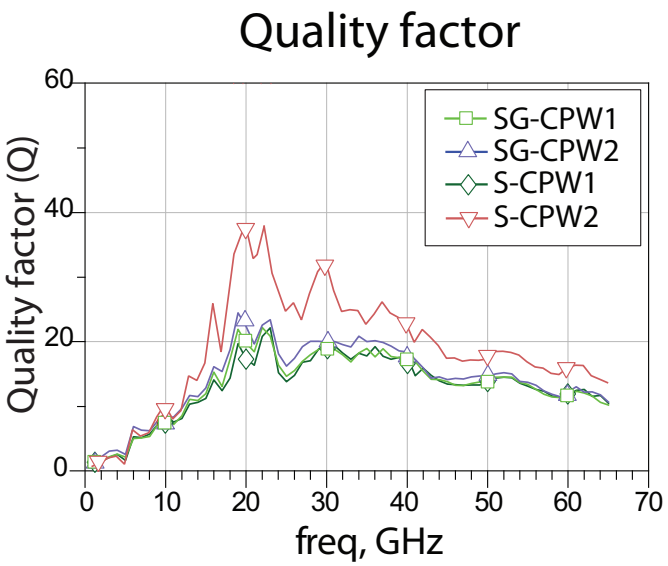


Figure 3.17: Quality factors of the slow-wave CPW (S-CPW1 and S-CPW2) and grounded slow-wave CPW (SG-CPW1 and SG-CPW2) on silicon.

3.4 On-Chip Capacitors

3.4.1 Introduction

The capacitor is a device that stores electrical energy. In microwave and millimetre-wave application capacitors are used in a DC block (alternating current (AC) coupling), supply-ground bypass and tuned resonators. A capacitor comprises a dielectric layer placed between two parallel metal electrodes. Classically the value of the capacitor is obtained by

$$C = \frac{\epsilon_0 \epsilon_r A}{t} \tag{3.33}$$

where $\epsilon_0\epsilon_r$ is the permittivity, t is the thickness of the dielectric material and A is the area of the metal electrodes. The main parameters that are used to characterise the quality of a capacitor are the quality factor Q, the resonant frequency and the capacity per unit surface area. This last parameter is of great importance for silicon integration as the cost of an integrated circuit is directly dependent on its surface area.

In CMOS technology, the main types of capacitors are:

I. MOS Capacitor

The MOS capacitor's structure is shown in Figure 3.18(a). The insulating dielectric (Oxide) is sandwiched between the top "metal" plate (heavily doped N^+ polysilicon)

and the semiconductor layer (p-type silicon). This type of capacitor is influenced by the technology scaling and provides a low capacitance per unit area and therefore are very impractical when a large capacitance is required.

This type of capacitor is very impractical when used in microwave circuits mainly due to the following reasons:

1. Affected by technology scaling, according to Equation 3.33 the insulator thickness is inversely proportional to the capacitance.
2. Provides low capacitance per unit area, which can be problematic and area consuming when a large capacitance is required (supply-ground bypass capacitor).
3. The lossy nature of this capacitor, due to the close proximity to the silicon substrate, makes it very impractical when used in matching circuits in the signal path.

II. metal insulator metal (MIM) Capacitor

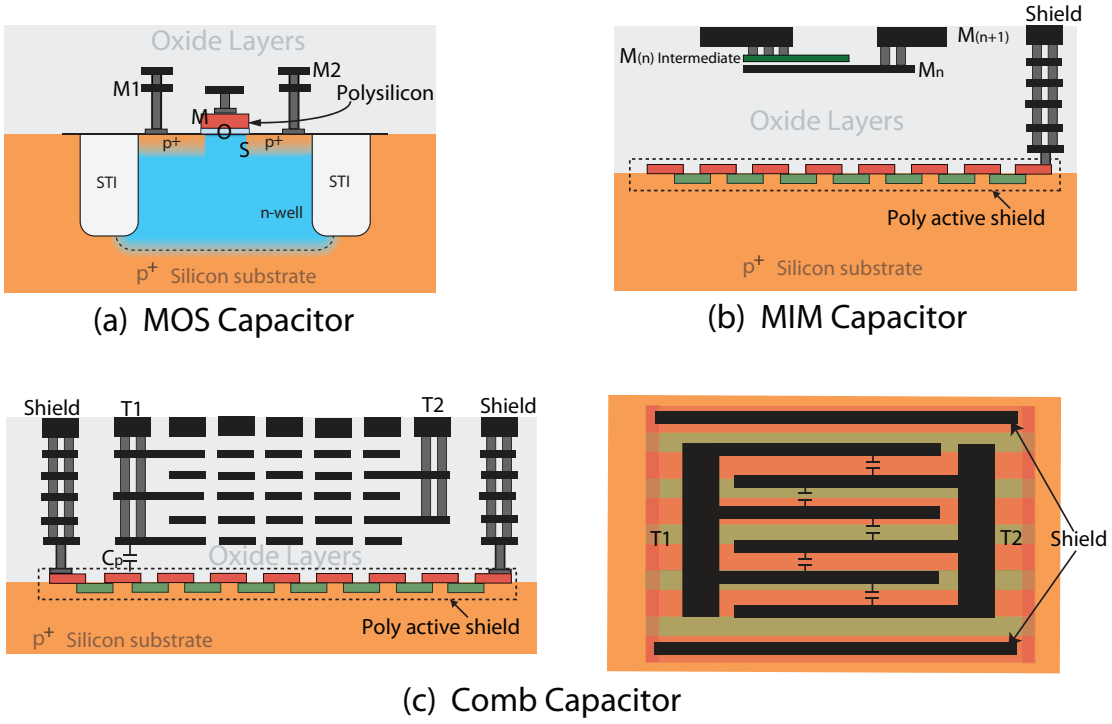


Figure 3.18: Capacitors in CMOS technology: (a) Metal oxide semiconductor (MOS) capacitor, (b) metal insulator metal (MIM) capacitor with shield and (c) comb capacitor with active poly shield.

A planar capacitor, as shown in Figure 3.18(b), is constructed from a thin insulation film between two metal layers [75, 76]. Normally implemented in the upper metal layers, this makes it insensitive to the silicon substrate as the electric field is largely enclosed between the metal layers that make up the capacitor. Unlike the MOS capacitor, MIM capacitors are not affected by technology scaling because the insulator thickness in this case does not scale according to process technology. This type of capacitor is very simple to design. However it requires intermediate metal levels and this option comes at an extra cost and may not be available in all CMOS foundries.

III. Comb (Interdigitated) Capacitor

Comb capacitors, shown in Figure 3.18(c), are often referred to as metal oxide metal (MOM) capacitors or interdigitated capacitors and are one of the most common type of capacitor implemented on silicon. It can be used in the pure digital CMOS process technology at no extra cost. In addition, large capacitance per unit area can be easily obtained. For this reason, this type of capacitor is employed in this work.

3.4.2 Designed Comb Capacitor

The comb capacitor can be formed by using metals in the same layer, or multiple layers, as interdigitated finger structures [77–79], where each other finger is connected to either port of the structure. In order to increase the capacitance and improve the quality factor (by decreasing the resistance) the capacitor was implemented using a stack of all available metal layers connected in parallel. Vertical fingers are formed of the odd metal layers while the horizontal fingers are formed of the even layers, as shown in Figure 3.19. The metal layers in each finger are connected together with the maximum number of vias allowed by the layout design rules in order to use the advantage of the via to via capacitance. The ports are composed of a wide stack of all metal layers. A shielding layer of active-poly fingers is used under the capacitor to reduce the substrate losses.

3.4.3 Modelling and Performance Results

The comb capacitors in this work were modelled using a symmetric 2-port model, and a 1-port model for the supply bypass capacitors as shown in Figure 3.20. C, L and R

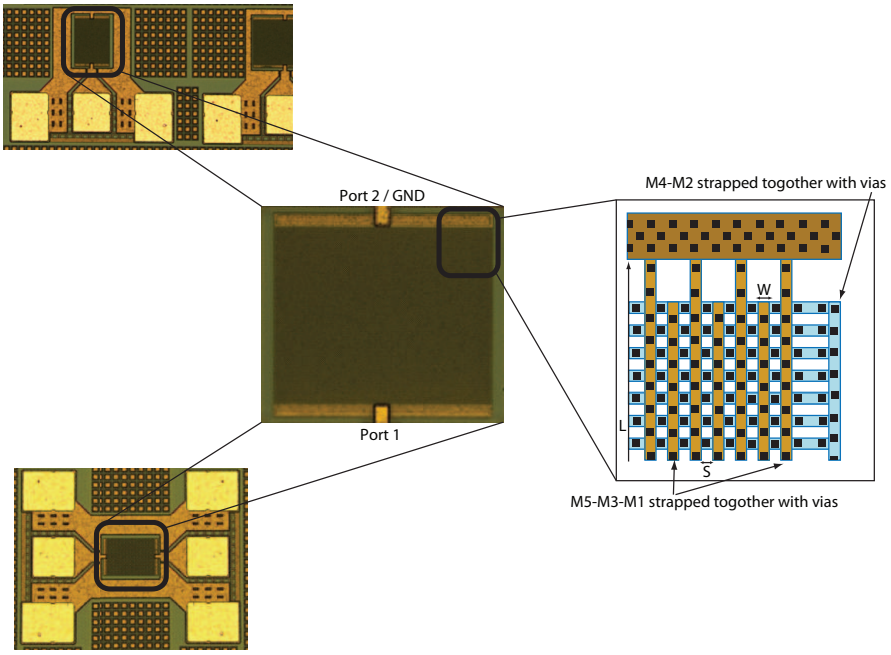


Figure 3.19: Top level view of a multi finger comb capacitor layout: DC bypass 1-port capacitor (top) and AC coupling and matching network 2-port capacitor (bottom).

represent the fingers and port networks capacitance, inductance, and resistance respectively. C_{ox} and R_{ox} model the oxide capacitance and resistance respectively, while C_{sub} along with R_{sub} model the substrate capacitance and resistance. In order to model the overall structure the circuit components values are optimised using Z and Y-parameters fitting to match the modelled network parameters to the measured ones.

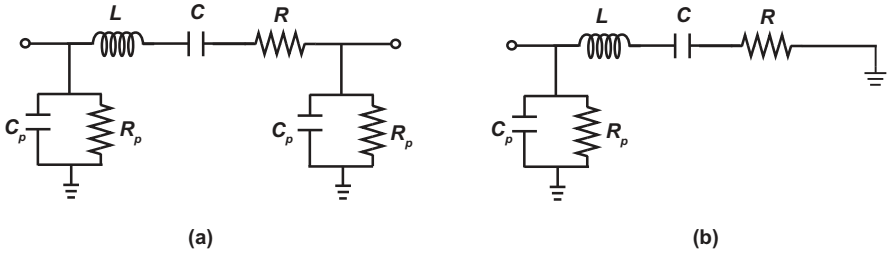


Figure 3.20: Comb capacitor model: (a) 2-port for matching networks and AC coupling and (b) 1-port supply bypass.

Several capacitors of different values were fabricated, measured and modelled. Table 3.3 shows the modelled circuit components and resonant frequency for those capacitors.

In conventional microwave and millimetre-wave circuits, transmission lines are used to provide the capacitance in matching networks. However, lumped capacitors can be used if their behaviour is modelled and characterised up to the frequencies of interest. It can be seen from Figure 3.21, which depicts the extracted capacitance of several lumped

Table 3.3: Extracted model parameters and the performance metrics of several multi-finger comb capacitors listed in this study; Two port(AC coupling caps) and one port (supply-ground bypass caps) with different dimensions N_vxN_h with different strip width(W) and spacing(S)

Dimensions				Model parameters					
			N_vxN_h	C (F)	L (pH)	R (Ω)	C_p (F)	R_p (Ω)	$f_{resonance}$
Two-port	W=0.28 μm	S=0.28 μm	18x18	71f	24	180m	26f	4.5K	≪ 100 GHz
			24x24	158.5f	26	264m	30f	4.5K	78.4 GHz
			120x20	661f	32.4	720m	27f	4.5K	34.5 GHz
			120x120	3.9p	42	1.46	53f	4.5K	34.8 GHz
One-port	W=0.28 μm	S=0.35 μm	18x18	87.4f	33	200m	26f	1.6K	93.6 GHz
			24x24	154f	35.4	261m	26f	1.6K	68 GHz
			120x20	641f	32.3	780m	28f	1.5K	35 GHz
			20x120	639f	33	780m	29f	1.5K	34.8 GHz
			120x120	3.78p	18	210m	52f	1.1K	19.2 GHz
			150x150	6.1p	13	177m	53f	1.1K	17.8 GHz
			200x200	10.5p	10.7	180m	62f	1.1K	14.9 GHz
			288x288	23.6p	12.4	175m	68f	1.1K	9.17 GHz

comb capacitors employed in matching networks, that the capacitance varies as the frequency increases. However for frequencies up to 60 GHz these variations do not have a critical effect on the matching circuit.

For both AC coupling and supply bypass the capacitor should ideally have infinite

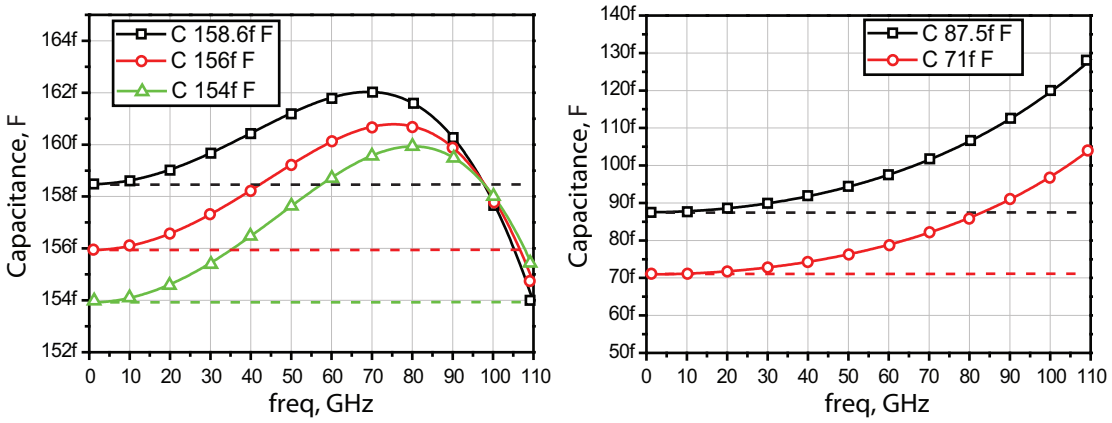


Figure 3.21: The extracted capacitance from two port lumped comb capacitors designed for matching networks (71fF, 87.5fF, 154fF, 156fF and 158.6fF).

impedance at DC and zero impedance at the frequency of interest. Therefore, large capacitors with a suitable self resonance frequency are required. Even if the capacitor self-resonates at a frequency below the operating frequency, as long as its impedance at the desired frequency is low enough, even if inductive, it can still be used as an effective structure for both supply bypass and AC coupling. Figure 3.22 and Figure 3.23 illustrate the measured and modelled, real and imaginary parts of Z_{11} and Y_{11} for a 6.1pF supply bypass capacitor. It can be seen that the capacitor exhibits low impedance at 25, 40 and 60 GHz, and hence a good choice for AC coupling and supply bypass capacitors for circuits designed at these frequencies.

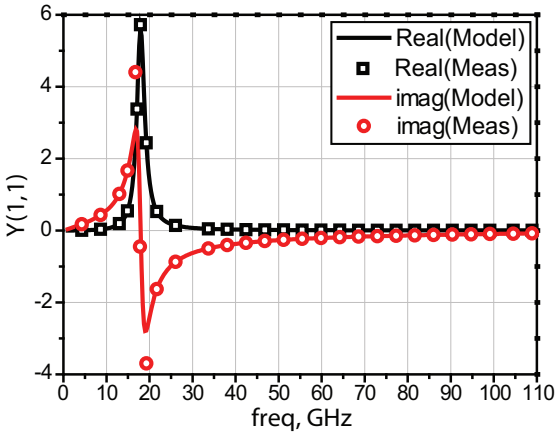


Figure 3.22: Y-parameters: Measured versus modelled for 6.1pF supply bypass comb capacitor.

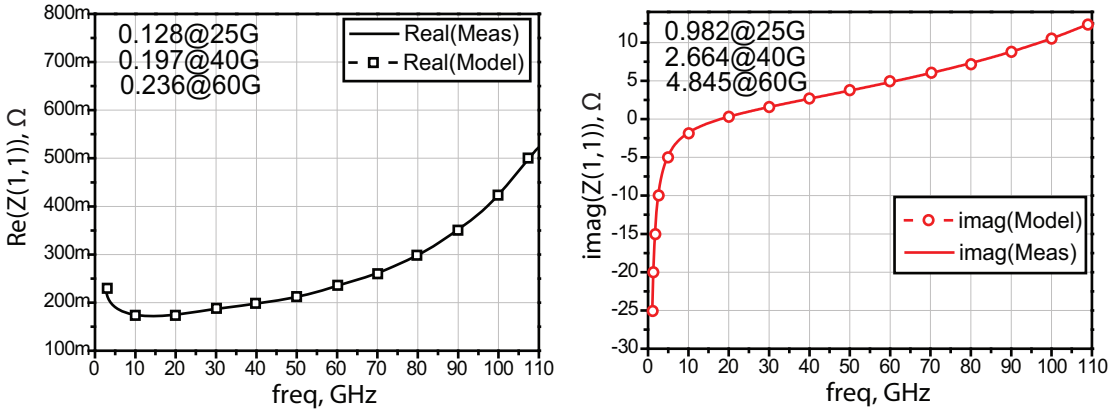


Figure 3.23: Z-parameters: Measured versus modelled for 6.1pF supply bypass comb capacitor.

3.5 On-Chip Resistors

3.5.1 Introduction

Resistors are important components in some RF circuits and are commonly used for current limiting, voltage division, signal isolation and load termination. Although resistors are linear components, their electrical behaviour in reality can be more complex. Any small deviations from linearity in resistors can lead to unacceptable levels of distortion, and parasitic capacitance can limit the frequency range over which a resistor can be used. For these reasons accurate modelling of resistor nonlinearities and parasitics is required for the design of RF circuits [80].

In CMOS technology several types of resistors are available:

I. Diffused and/or implanted resistor

A diffused resistor is formed using drain/source diffusion as shown in Figure 3.24(a). The sheet resistance of such resistors in a salicide process is in the range of 5 to 15 Ω/□ (Ohms per square). Diffusion resistors have a large voltage dependent parasitic capacitance associated with them; hence this type is rarely used in RF design [80].

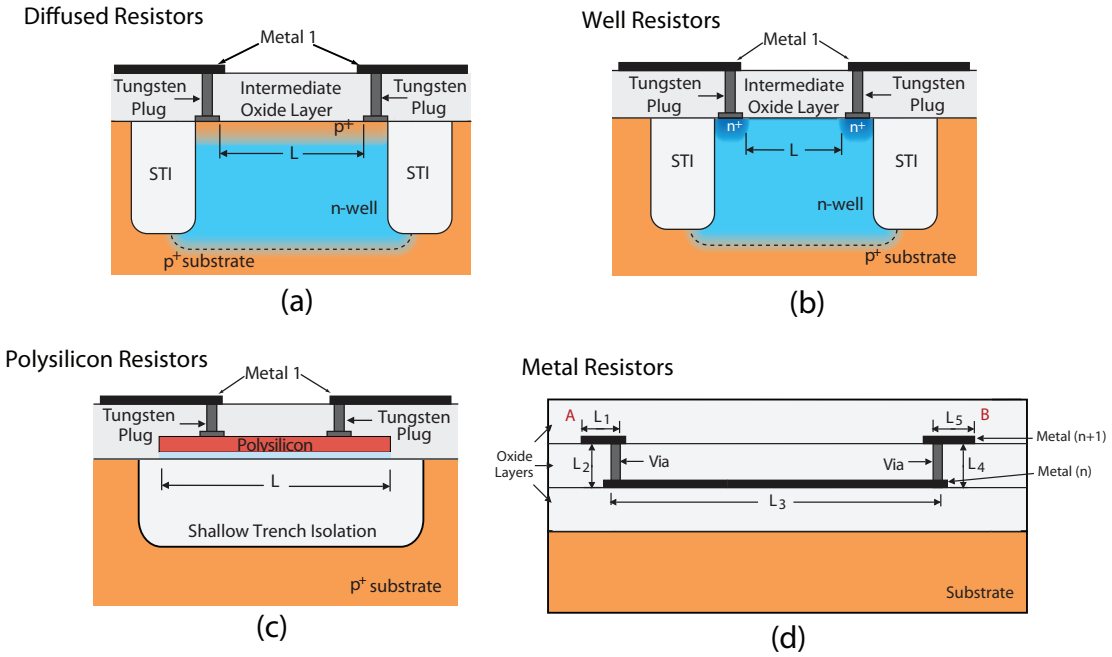


Figure 3.24: On-chip resistors in CMOS technology: (a) Diffused resistor, (b) well resistor, (c) metal resistor and (d) polysilicon resistor.

II. Well resistor (n-well or p-well)

Well resistors shown in Figure 3.24(b) are made up of a strip of n-well or p-well contacted at both ends with n+ drain/source diffusion. This type of resistor has a resistance of 1 to 10 $k\Omega/\square$. Well resistors are routinely used when a large resistance value is required, such as protection resistors and pull-up/down resistors. However, noise coupling through the substrate can be a concern when used in a low noise circuits.

III. Metal resistor

Metal resistors shown in Figure 3.24(c) give very low values of resistance ($0.078\Omega/\square$), therefore they are not commonly used.

IV. Polysilicon resistor

A polysilicon resistor is shown in Figure 3.24(d). This resistor is surrounded by a thick oxide and has a sheet resistance in the range of 30 to 200 $\Omega/\square$ depending upon doping levels. For a polysilicide process, the effective resistance of the polysilicon is about 10 $\Omega/\square$. Poly resistors in modern technologies can be designed to have small voltage and temperature coefficients. Moreover they offer much lower parasitic capacitance than metal, diffused and well resistors, and so are preferred for RF applications. In addition an accurate and physically based model for polysilicon resistors is available [80, 81].

3.5.2 Resistor Layout

Figure 3.25 shows the layout of an integrated resistor consisting of a rectangle of resistive material with contacts at either end. The low resistance of a contact effectively shorts out the material underneath it. The current flows in this resistor from the inner edge of one contact to the inner edge of the other.

In an integrated resistor, photolithography and etching can cause oxide openings to vary slightly, thus leading to resistance variations. Also nonuniform current flow near the contacts can increase its resistance. Factors relating to high frequency effects such as electromagnetic coupling, proximity effects, skin effect, radiation and substrate currents (ohmic, eddy and displacement currents) should be taken into account in order to accurately predict the value of a resistor. The resistor equation can be written as follows

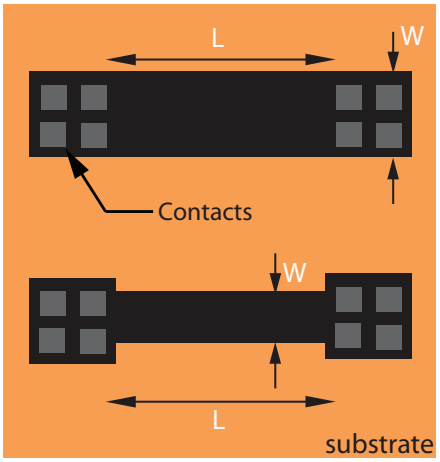


Figure 3.25: The layout of a typical integrated resistor consisting of the resistor body and two resistor heads.

[82]:

$$R = R_{sh}\frac{L}{W} + 2\left(R_c + R_{sh}\left(\frac{W_0}{W_c} + \triangle_{sq}\right)\right) \tag{3.34}$$

where

$$R_c = \frac{\sqrt{R_{sh}\,\rho_c}}{W_c}\coth\left(L_c\sqrt{\frac{R_c}{\rho_c}}\right) \tag{3.35}$$

$$\triangle_{sq} = \frac{1}{\pi}\left[\frac{1}{k}\ln\left(\frac{k+1}{k-1}\right) + \ln\left(\frac{k^2-1}{k^2}\right)\right] \tag{3.36}$$

$$k = \frac{W}{W - W_c} \tag{3.37}$$

Here L and W are the resistor effective length and width respectively, R_{sh} is the sheet resistance of the resistive material, R_c is the contact resistance, L_c and W_c are the length and the width of the contacts, ρ_c is the specific contact resistance and $\triangle_{sq}$ is the adjustment parameter for current crowding or spreading (nonuniform current flow).

Equation 3.34 shows that the length of the resistor is directly proportional to the resistor value. Hence, larger resistors would require larger area. Different layout techniques can be applied to create larger more compact resistors as shown in Figure 3.26. Dogbone resistors (Figure 3.26(a)) do not pack as densely as strip (Figure 3.26(b)) or serpentine resistors (Figure 3.26(c)). However, the errors caused by nonuniform current flow are smaller for dogbones than for either strip or serpentine resistors, but this does not actually improve the accuracy of the resistor [82]. To improve the accuracy of the

resistor and reduce process variations and mismatch, dummy structures should be used on each side of the designed resistor.

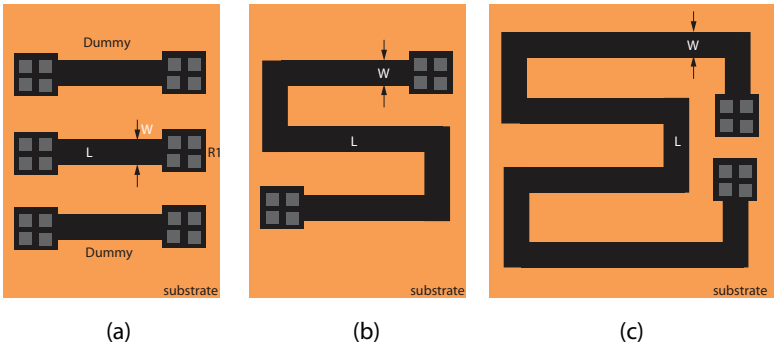


Figure 3.26: On-chip resistor layout examples: (a) Dogbone structure, (b) strip resistor and (c) serpentine resistor.

3.5.3 Modelling and Performance Results

In this work, resistors were not used directly in the signal path, but only for bias circuitry and load termination for passive couplers. The resistors are polysilicon and n-well resistors. Two types of poly resistors are provided in the design kit: N+ poly resistor (rnhpoly) and P+ poly resistor (rphpoly). The resistance value varies from 1 Ω/□ to 100 Ω/□ with silicide layer ($TiSi_2$ or $CoSi_2$ or PtSi) and from 100 Ω/□ to 2K Ω/□ for non-silicide resistor. The non-silicide resistors are fabricated by blocking the silicide deposition on top of the poly layer, thus achieving high resistivity [82, 83]. For large resistance values, N-well resistors were used (100 Ω/□ to 1M Ω/□), however the accuracy is not guaranteed. In order to get optimum performance up to millimetre-wave frequencies the width of the resistor has to be minimised, otherwise parasitics become dominant.

Unfortunately, the models provided by the design kit are insufficient for high frequency and new, more accurate models had to be created in order to account for the high frequency parasitics. The high frequency equivalent circuit model is illustrated in Figure 3.27.

In order to model the overall resistor structure, the circuit model components values are optimised using Z-parameter fitting. This will match the modelled network parameters to the measured parameters. The mean error value between the modelled and measured parameters is 3.4%. The circuit components values of both 50Ω polysilicon resistor and

the 2.4$k\Omega$ N-well resistor are listed Table 3.4. Figure 3.28 depicts the Z-parameters of the measured and modelled 50Ω polysilicon resistor and Figure 3.29 depicts the Z-parameters of the measured and modelled 2.4$k\Omega$ polysilicon resistor.

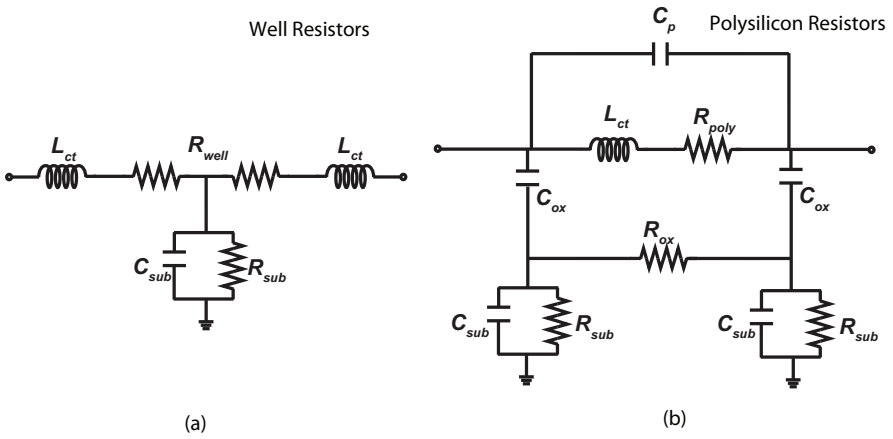


Figure 3.27: On-chip resistors equivalent circuit model: (a) Well resistor and (b) polysilicon resistor.

Table 3.4: The extracted parameters values for high frequency on-chip poly and well resistor models.

Polysilicon 50 Ω **resistor**						
R_{poly}	L_{ct}	C_p	C_{ox}	R_{ox}	C_{sub}	R_{sub}
50 Ω	16pH	20fF	19fF	45 kΩ	26fF	47 Ω
N-Well 2.4 kΩ **resistor**						
	R_{well}	L_{ct}	C_{sub}	R_{sub}		
	2.4 kΩ	13pH	26fF	40 Ω		

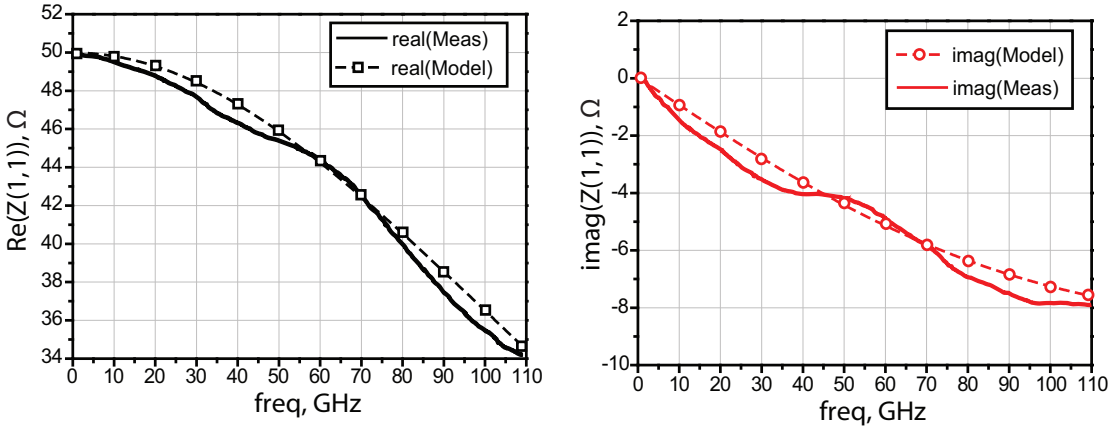


Figure 3.28: Z-parameters: Measurement versus model for the 50 Ω polysilicon resistor.

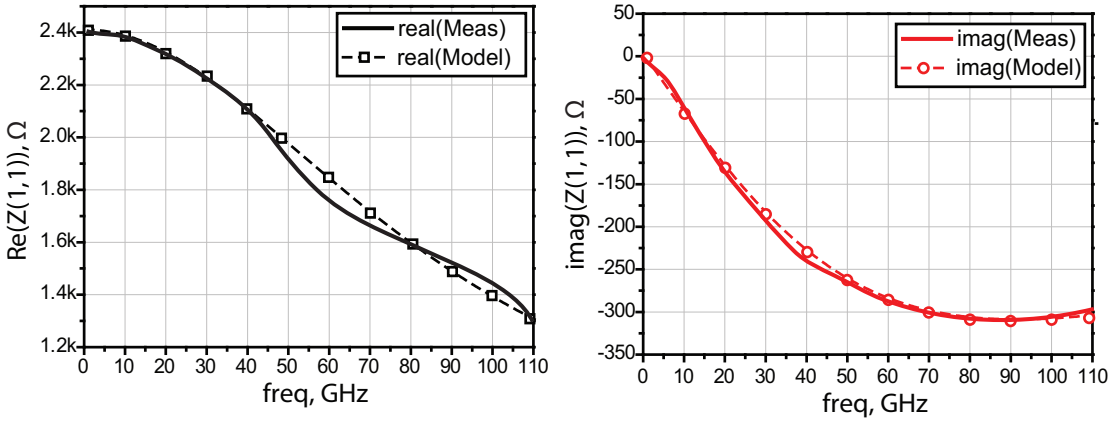


Figure 3.29: Z-parameters: Measurement versus model for the 2.4 kΩ N-well resistor.

3.6 Summary

This chapter discussed the design and optimisation of passive structures implemented on a lossy silicon substrate. The first passive structure discussed in this chapter was the spiral inductor. A brief background about the performance metrics and physical phenomena was introduced followed by an equivalent circuit model of a spiral inductor on silicon. The main challenges that degrade the performance were illustrated. Quality factor improvement using shielding was discussed. A grounded and floating patterned shield was proposed, followed by a case study comparison between the two shielding methods. The second passive structure discussed in this chapter was a transmission line on lossy silicon. These were briefly introduced and compared. Slow-wave CPW operation was described and explained. Followed by a comparative study between a conventional CPW, slow-wave CPW and grounded slow-wave CPW line was conducted. The study led to the conclusion that a slow-wave CPW provides the lowest attenuation and best quality factor when compared to the other lines listed in the study. Another passive structure listed in this chapter was a capacitor; several different types of CMOS capacitors were discussed. The comb capacitor was designed and modelled. The final section detailed the results , including all the necessary details to design poly resistors and methods to model them up to high frequencies. These passive structure will be used to design low noise amplifiers operating at millimetre-wave frequencies as shown later in this thesis.

Chapter 4

Millimetre-wave MOSFET Modelling: Small-Signal and Noise

4.1 Introduction

Generally circuit designers are used to assuming transistor models are provided by the foundry design kit. Hence the design process consists of placing the transistor in the schematic window, building the circuit around it, setting up the desired type of simulation and beginning the simulation process. They may perform their simulations in a number of different process corners but this is as much as they should worry about the device modelling concept. In the case of microwave and millimetre-wave circuit design, if this approach is employed the fabricated circuit may suffer huge variation between the model and measurement, and in most cases it may fail to function. In order to avoid these problems the transistors need to be accurately and extensively modelled.

Millimetre-wave transistor models are normally extracted from device measurement data. Models of this type can be categorised as;

- **Polynomial models:** These are often referred as models directly described from the S-parameter measurements. A block of 2-port S-parameters measured at specific bias points over a frequency range, used in circuit simulations as a black box.

- **Physical small-signal electrical equivalent circuit models:** As the name depicts, the device is represented by a circuit model based on the transistor physics. Each component in this circuit has a known origin. Extracting the circuit model components from the transistor measured S-parameters can be challenging.

Knowledge of the small-signal equivalent circuit model is key to developing circuit design methodologies or device optimisation. By understanding the influence of each circuit component on the device performance, designers can alter and optimise the design to suit a specific application. In this chapter, an overview of CMOS MOSFETs and fundamentals of high frequency noise is first presented. This provides the foundation for the small-signal and noise device modelling discussions that follow. Small-signal and noise modelling methodologies, as relevant to this work, are presented and verified thereafter.

4.2 Background Theory

4.2.1 CMOS MOSFETs

The basic structure of a CMOS transistor, as shown in Figure 4.1, comprises of a silicon substrate, heavily doped source and drain regions, an insulating layer (gate oxide) and a polysilicon gate. The insulating layer, silicon dioxide (SiO_2) or a high-k dielectric material, separates the gate from the conductive channel. The device is operated by two bias voltages; gate-source bias (V_{gs}) and a drain-source bias (V_{ds}). The former controls the amount of charge present in the channel, while the latter provides the

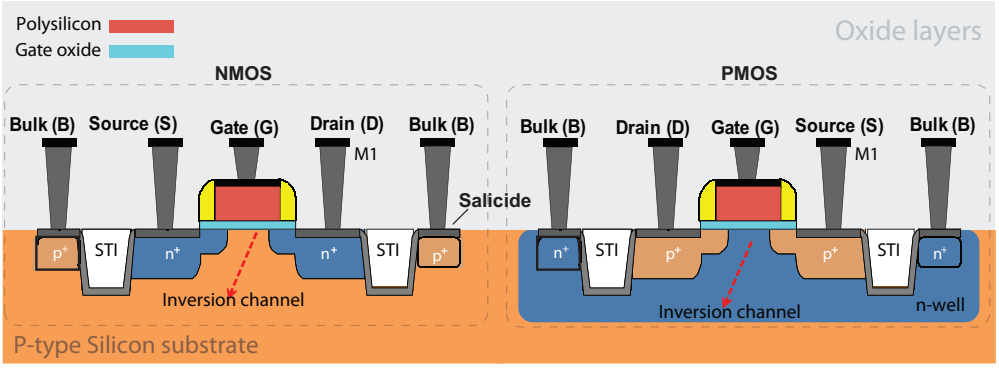


Figure 4.1: A cross section of a planar n-type MOSFET (NMOS) and p-type MOSFET (PMOS) structures in CMOS technology.

means of charge flow through the channel. The charge enters and exits the channel via the heavily doped source and drain ohmic contacts, n^+ for NMOS and p^+ for PMOS. This charge flow is referred to as the drain to source current (I_{ds}) [84]. The MOSFET can modulate I_{ds} by controlling the voltage applied across the gate and source terminals. Electrically this behaviour can be represented by a current controlled source defined by the transconductance (g_m).

4.2.1.1 CMOS MOSFET Operation Modes

MOSFET operation modes are controlled by voltage potentials between the gate and drain with respect to the source node (V_{gs} and V_{ds}). The gate node controls the MOSFET by setting the required electrical field to determine the charge concentration in the channel (bulk to oxide interface). Whereas the drain node controls the MOSFET operation regions and establishes the current flow between the drain and source, I_{ds}. The MOSFET operation modes as a function of gate voltage are divided into three regions: Accumulation (device off), depletion and inversion (device on). On the other hand, The MOSFET operation regions as a function of drain voltage are divided into three regions: cut off region, linear region and saturation region.

- **MOSFET Operation vs. Gate Voltage:**

  In a p-type substrate there is excess of positive majority carriers. An accumulation layer is created by applying a negative gate voltage. This attracts the positive majority carriers towards the oxide interface. On the other hand, any free negative minority carriers in the oxide interface are pushed deeper into the substrate away from the oxide interface. Hence the device is in *accumulation region* (off). Applying a small positive voltage to the gate, causes the positive majority carriers to repel away from the accumulation layer and deeper into the substrate leaving the negative charge behind. As the gate voltage increases the region near the bulk to oxide interface become more and more depleted from holes creating a *depletion region*. Further increase in the gate voltage attracts more and more free negative majority carriers, in the substrate, towards the bulk to oxide interface. Since the number of positive majority carriers in the p-type substrate increase due to repulsion from the bulk to oxide interface, negative minority carriers must be generated to maintain neutrality. Finally a continuous n-type channel region becomes present

at the bulk to oxide interface under gate, consisting of negative minority carriers that were just created. The semiconductor material near the bulk to oxide interface is said to be inverted thus the device is in *inversion region*. The gate voltage at which this channel is formed is called the threshold voltage, V_{th}. The inversion region is normally divided into two sub-regions weak and strong inversion, This refers to the regions before and after V_{th} respectively.

- **MOSFET Operation vs. Drain Voltage:**

  This dependency assumes that the device is in inversion region and the conducting channel is already formed. The drain bias is vital for establishing the current flow between drain and source, I_{ds}. The MOSFET is in the cut off region when V_{ds} = 0V, which ideally results in I_{ds} = 0A, as seen in Figure 4.2(a). This region becomes present if a small V_{ds} is applied, as shown in Figure 4.2(b). I_{ds} will then be about proportional to V_{ds}, hence a linear curve characteristic is obtained. By adjusting V_{ds} within this region, the behaviour of a voltage controlled resistor is obtained. As V_{ds} is further increased however, the size of the conducting channel starts to decrease at the drain end. At the so called pinch-off voltage Figure 4.2(c) the conducting channel reaches only just to about the drain diffusion. Drain current now starts to saturate, because only a depleted version of the original

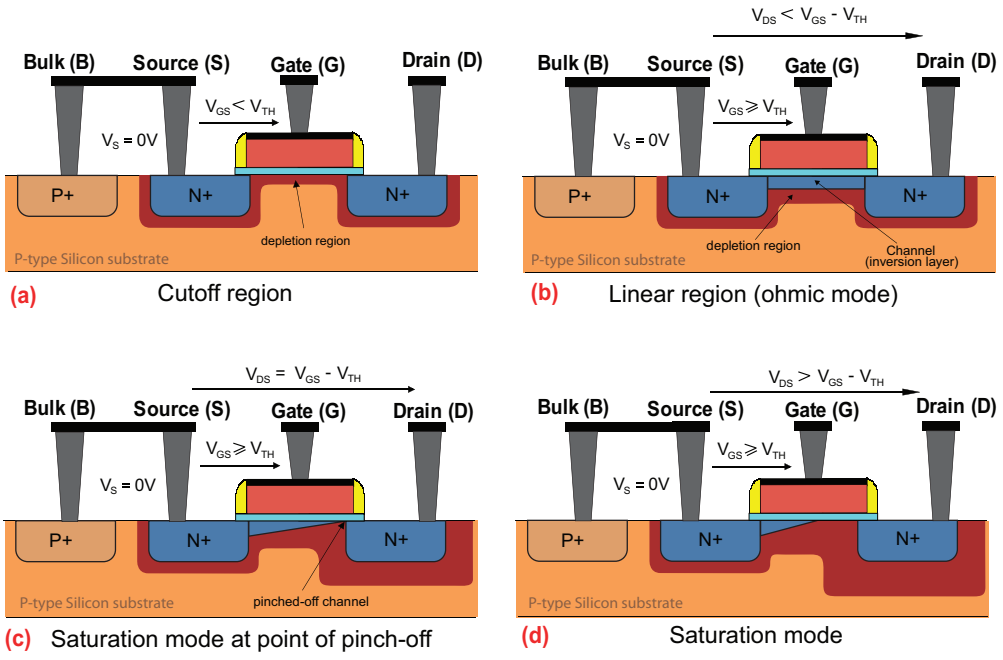


Figure 4.2: MOSFET operation modes with variable drain potential: (a) cut off region, (b) linear region, (c) pinch-off region and (d) saturation region.

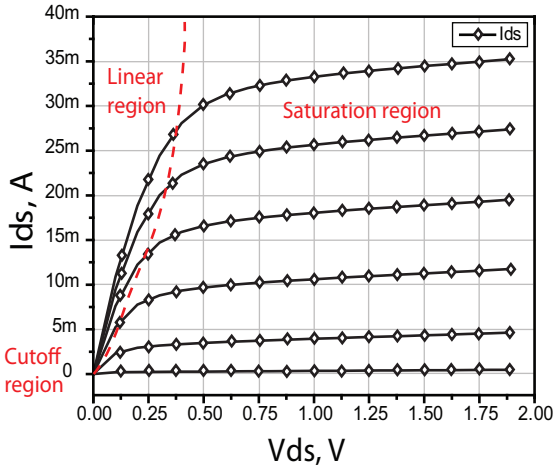


Figure 4.3: Current-voltage (ID-VD) characteristics of an n-type MOSFET. The plot illustrates the operation regions cutoff, linear and saturation. The dotted line seperates the linear region from saturation region.

channel is left on the oxide to substrate interface near the drain node. Since the MOSFET characteristics in linear region resemble a voltage controlled resistor, the output resistance is referred to as drain-to-source resistance, r_{ds}, or the drain-to-source conductance g_{ds} [1]. The saturation region ($(V_{gs} - V_{th}) < V_{ds}$) as V_{ds} is further increased after pinch-off, the depleted channel gradually replaces the original channel and so the pinch-off point moves closer and closer toward the source diffusion. As a result, the drain current becomes more and more saturated because of the decrease in effective channel length, Figure 4.2(d). As a result the velocity of the minority carriers saturates, and the ID-VD curve characteristic flattens as shown in Figure 4.3. A common exception from this characteristic comes as a consequence from channel-length modulation, which arise for short channel devices. That is, since the effective length of the channel is reduced with increasing V_{ds} the effective r_{ds} is also reduced. This effect results in a small proportional increase in I_{ds} even in this region. Channel-length modulation is expressed by the parameter λ, where $\lambda = 0$ indicates that the slope is zero in saturation region. The characteristics of the MOSFET in saturation region is very similar to that of a voltage-in-current-out amplifier, commonly named a transconductor. Because of this the transconductance gain of the MOSFET is a very interesting parameter for this specific region. This is modeled by the transistor transconductance, g_m[85].

[1] g_{ds} is the inverse of r_{ds}, ideally g_{ds} is infinite when the MOSFET is in cut off

Figure 4.4 depicts a simplified layout of an RF NMOS transistor implemented on a silicon substrate. The transistor comprises a multi-finger configuration with double a contacted gate polysilicon structure. Dummy gate poly is added to improve the gate structure formation. The metal 1 and metal 2 are used to connect the gate terminal. The source diffusions are connected using metal 1 to the bulk terminal (P-Well). The drain diffusion is connected to metal 1, 2 and 3, and taken to the opposite side of the gate terminal.

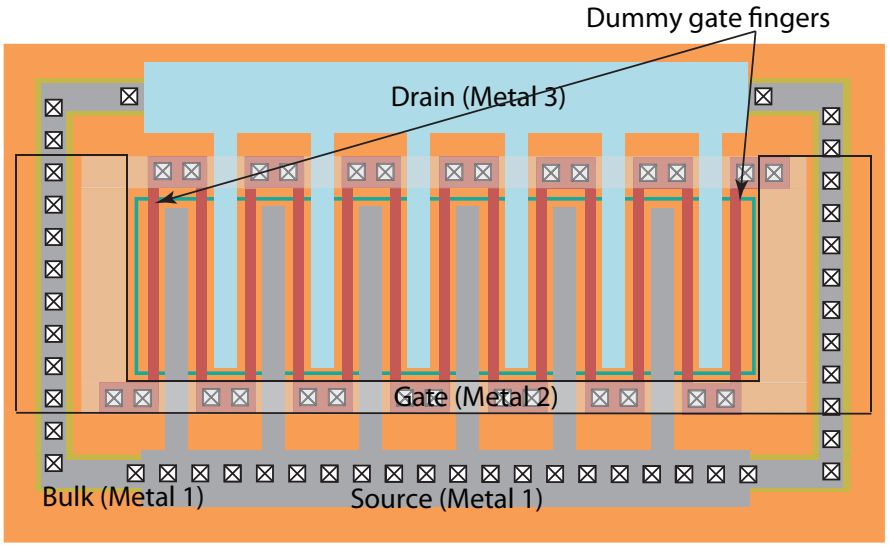


Figure 4.4: CMOS RF NMOS transistor layout featuring multi-finger double contacted gate CMOS technology.

4.2.1.2 High-Frequency Equivalent Circuit Representation

In DC operation, a MOSFET is controlled by four DC voltages; V_d, V_g, V_b and V_s. All defined with reference to ground. As mentioned earlier, in an NMOS transistor the channel of the device is formed by the charge (electrons) flowing from source to drain terminals in the inversion layer. When the variation of terminal voltages is sufficiently slow, the charges per unit area in the channel and the depletion region are unchanged and the device is assumed to be in quasi-static operation [84]. This assumption is only valid in DC and low frequency applications, for these conditions the quasi static model, shown in Figure 4.5(a), can be used with negligible errors. However, when operating at high frequencies the terminal voltages change more rapidly. As a result, the charges

flowing in the channel do not have sufficient time to follow the variation in terminal voltages and the quasi-static assumption becomes invalid[2].

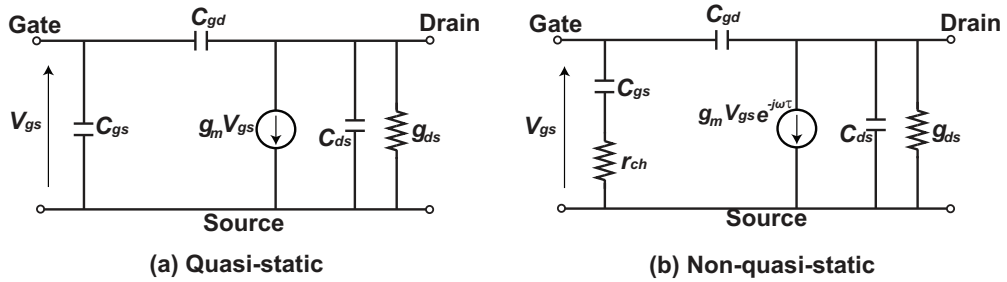


Figure 4.5: Intrinsic model of a MOSFET: (a) Quasi-static and (b) non-quasi-static.

There are several ways to account for the non-quasi-static (NQS) effects in a MOSFET model. The easiest approach is to model the NQS effects by using a single resistor connected to the gate of the device as shown in Figure 4.5(b). This resistor introduces a delay (τ) to the charge travelling in the channel [84].

The physical interpretation of the equivalent circuit of a MOSFET is illustrated in Figure 4.6. It can be seen that the MOSFET consists of intrinsic and extrinsic parts. The intrinsic part is an approximation of the device physics, and depends on the physical dimensions as well as the biasing conditions. It forms the core of the device and consists of a voltage dependent current source ($g_m.v_{gs}.e^{-j\omega\tau}$), capacitive elements such as the gate-drain capacitance (C_{gd}), the gate-source capacitance (C_{gs}) and the source-drain

[2]The upper frequency limit of the quasi-static model is proportional the cut-off frequency (f_T), which is approximately proportional to $1/L^2$, where L is the transistor channel length [86].

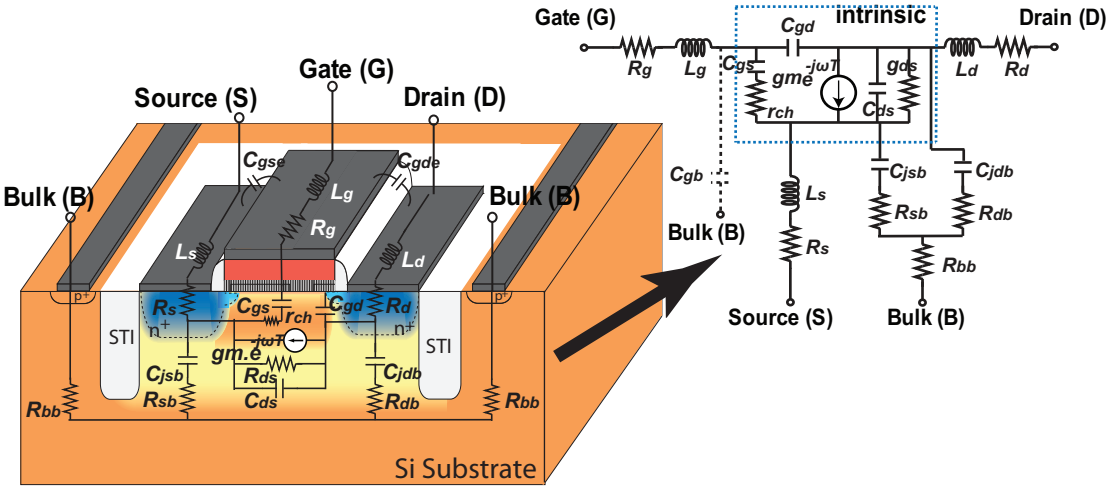


Figure 4.6: Physical interpretation of the equivalent circuit of a high frequency n-type MOSFET.

capacitance (C_{ds}). In addition to resistive elements which include the drain conductance (g_{ds}) and the channel resistance (r_{ch}).

The extrinsic part, on the other hand, is bias independent and includes all the parasitic elements from the nearest intrinsic device to the furthest, filling the gap between the intrinsic part of the device and the reference planes defined by de-embedded pads. The various inductances, capacitances, and resistances associated with the extrinsic par are as follows:

- **Extrinsic capacitances:** C_{gse}, C_{gde} and C_{dse} are capacitors located around the channel of the transistor. They are bias independent elements and proportional to the transistor dimensions. The extrinsic gate-source (C_{gse}) and gate-drain (C_{gde}) capacitances include overlap capacitances, located between the gate oxide and the diffusion of the source and the drain regions under the gate, and fringing capacitances from the gate sides to the source and drain terminals. Whereas, the capacitance (C_{dse}) represents the coupling between metal lines between the source and drain.

- **Extrinsic resistances:** Gate, source and drain resistors (R_g, R_s and R_d) are lumped resistors used in the device model to represent the distributed nature of the terminal resistances in multi-finger CMOS MOSFETs. The resistances R_s and R_d are proportional to the inverse of the transistor width [84]. They represent the metal losses and the contact resistances between the metal and the source and drain diffusion region. The gate resistance includes the resistance of the polysilicon gate fingers, which is proportional to the transistor width, and the resistance of some metallic contacts and lines, which connect the gate fingers together and to the reference plane.

- **Extrinsic inductances:** Gate, source and drain inductance (L_g, L_s and L_d). These are lumped inductors used in the model to represent the inductance of the terminals and they are usually in the order of a few pH in a sub-micron MOSFET.

- **Substrate capacitances:** C_{jdb} and C_{jsb}.

- **Substrate resistances:** R_{db}, R_{sb} and R_{bb}.

Conventional models may not include the substrate parasitics, but due to the resistive nature of the silicon substrate, the substrate network should be added. Substrate and

extrinsic parasitics may not have any effect when operating at DC and low frequencies. However, at millimetre-wave frequencies they influence the device performance significantly.

4.2.1.3 Figures of Merit

There are several performance figures of merit for a transistor, each defines the performance in a different way. The most popular figures of merit, as shown in Figure 4.7, are

1 **Cut-off frequency (f_t)**: Normally referred to as the transition frequency when the small-signal current gain reaches unity. This figure of merit is the most common performance metric, conventionally used as a speed benchmark for the active device in a given technology node. f_t can be defined as

$$f_t = \frac{i_{out}}{i_{in}} = \frac{1}{2\,\pi}\,\frac{g_m}{C_{gs} + C_{gd} + C_{gb}} \tag{4.1}$$

where C_{gs}, C_{gd}, C_{gb} are the gate-source, the gate-drain and the gate-bulk capacitances. g_m is the device transconductance. The cut-off frequency can be easily obtained from a measured device by taking the magnitude of h_{21}.

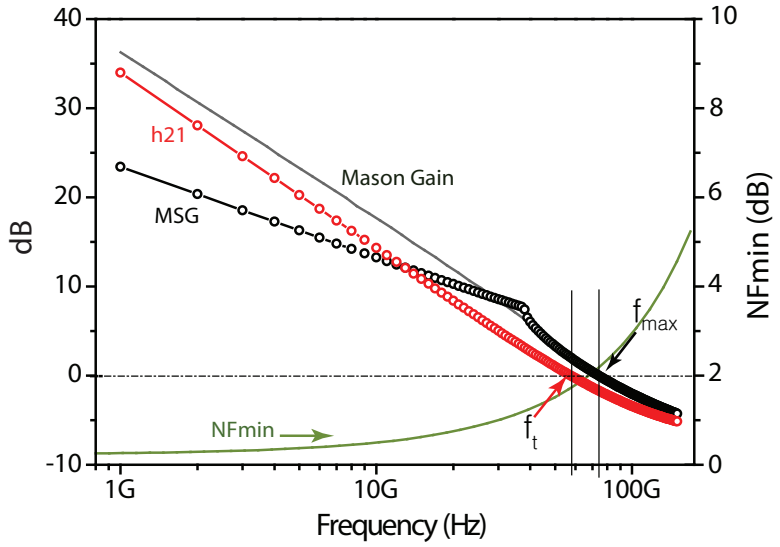


Figure 4.7: h_{21}, maximum stable gain (MSG), maximum unilateral gain (Mason Gain or U) and the minimum noise figure (NF_{min}) of a sample common source CMOS transistor. The f_t and f_{max} of the device are defined by the unity values of h_{21} and U curves respectively.

2 Maximum oscillation frequency (f_{max}): Defined as the transition frequency when the device changes from active to passive [87]. Therefore, it is a far more important figure of merit for the high frequency capabilities of a technology, than the cut-off frequency (f_t), as it dictates the operating frequency limitation. f_{max} can be calculated using

$$f_{max} = \frac{f_t}{2\sqrt{g_{ds}(R_g + R_s) + 2\pi f_t R_g C_{gd}}} \tag{4.2}$$

where R_g and R_s are the gate and the source resistances, C_{gd} is the gate drain capacitance.

3 Mason's Unilateral gain (U): Defined as the gain of the device when the device is unilateralised through some feedback mechanisms or using some circuit techniques [88]. Masons unilateral gain can be calculated using

$$U = \frac{\left|\frac{S_{21}}{S_{12}} - 1\right|^2}{2\left(k\left|\frac{S_{21}}{S_{12}}\right| - \Re\left(\frac{S_{21}}{S_{12}}\right)\right)} \tag{4.3}$$

where k is the stability factor and can be calculated as:

$$k = \frac{1 - |S_{11}|^2 - |S_{22}|^2 + \triangle^2}{2\,|S_{12}\,S_{21}|} \tag{4.4}$$

$$\triangle = S_{11}S_{22} - S_{12}S_{21} \tag{4.5}$$

The active device becomes passive when U = 1, this is the same frequency as f_{max}. In cases where f_{max} is beyond the frequency capability of the measurement equipment, it can be derived from the extrapolation of Masons unilateral gain assuming a 20 dB/decade slope.

4 Noise figure (NF): This figure of merit determines the noise sources inside the device as well as the source admittance that could achieve the minimum possible noise performance. The noise figure (NF) of a device achieves its minimum (NF_{min}) when the source admittance $Y_s = Y_{opt}$.

NF is calculated using

$$NF = NF_{min} + \frac{R_n}{G_s}|Y_s - Y_{opt}|^2 \tag{4.6}$$

NF_{min} can be approximated as [89]:

$$NF_{min} = 1 + K\frac{f}{f_t}\sqrt{g_m\left(R_g + R_s\right)} \tag{4.7}$$

where K is a constant representing the fitting coefficient which depends on the channel material.

It can be seen from Equations 4.1, 4.2 and 4.7 that f_t, f_{max} and NF_{min} depend on the parasitic capacitances and resistances of the device and can be improved by optimising the layout of the transistor. This will be discussed in detail in Chapter 5.

4.2.2 Fundamentals of High-Frequency Noise

There are two main types of noise sources in electronics. The first type is called artificial "external" noise source, this is normally a result of unintended coupling with other parts of the physical world such as power lines and electromagnetic waves. In principle this noise can be virtually eliminated by good shielding, isolation and careful layout design. The second type of noise source is classified under fundamental "internal" noise. It is an unpredictable microscopic event, associated with current conduction, inherent in a device or circuit, for example the "snowy picture" in analog TV sets or the hissing noise in audio systems. This type of noise can never be eliminated completely, but can be reduced.

If we bias a common-source transistor at a constant V_g greater than the transistor V_{th} with a fixed V_d, as shown in Figure 4.8, it is expected that the I_d is invariant as a function of time. However, if we monitor it with an oscilloscope, the actual current fluctuates randomly around the expected value I_d. These unwanted random fluctuations are normally referred to as "noise".

Noise is very unpredictable in the time domain due to its random nature. Therefore, it must be treated statistically by calculating the mean value $\overline{I_d(t)}$, the variance $\overline{I_d(t)^2}$ and the root mean square (RMS) value $\sqrt{\overline{I_d(t)^2}}$. In the frequency domain, noise is represented by its power spectral density (PSD), which is defined as the concentration of noise power at any given frequency, as illustrated in Figure 4.9. Notice at high frequencies, the PSD of the noise signal is frequency independent and it is known as

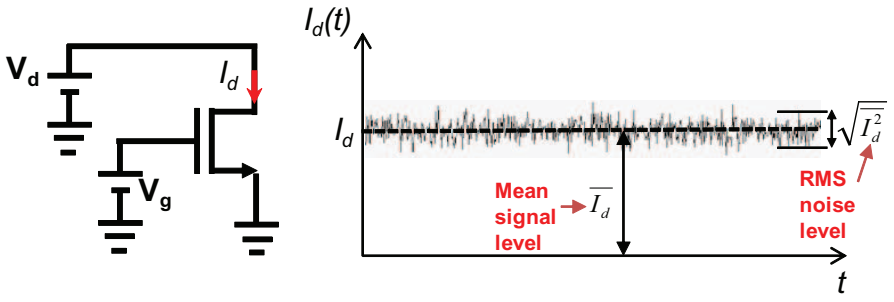


Figure 4.8: Drain current of a common-source transistor as a function of time at a fixed bias voltage.

"white noise". However, at lower frequency regions, the PSD of the noise signal is inversely proportional to the frequency, and it is often referred to as "pink noise" or $1/f$ noise.

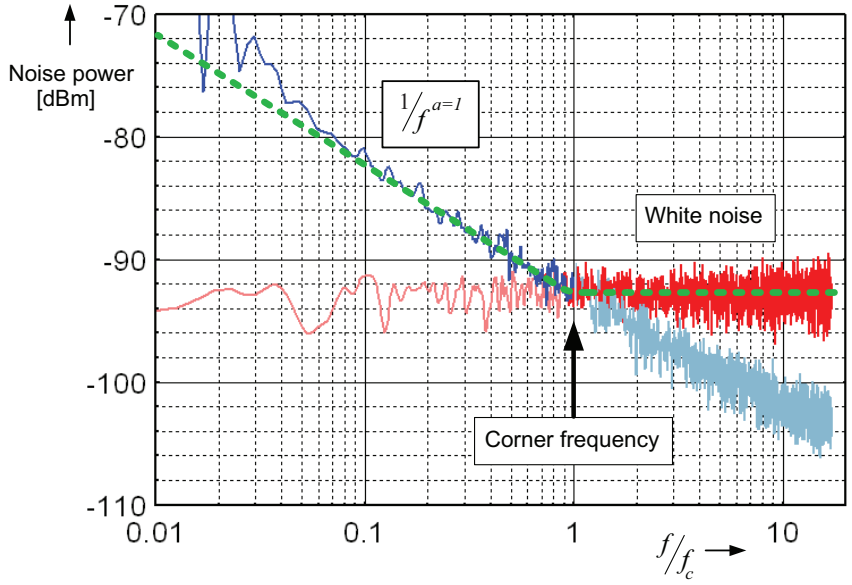


Figure 4.9: Power spectral density (PSD) of the drain current.

4.2.2.1 Noise in Semiconductor Devices

In a semiconductor device, several noise types are observed including thermal, shot, generation-recombination and flicker noise. These are described as follows:

1. **Thermal Noise (*Johnson/Nyquist Noise*)**: Thermal noise is one of the fundamental noise sources in any electronic circuit. It is caused by thermal excitation of charge carriers in a semiconductor. A thermally-excited charge carrier undergoes random changes in motion (Brownian motion), due to its collision with the

lattice of the semiconductor. As a result, it produces fluctuations in the output voltage/currents [90]. In 1927 Johnson discovered that the noise power spectral density of a semiconductor is independent of its material and the operation frequency, and determined only by the electrical resistance (R) and the temperature (T) under thermal equilibrium [90, 91]. The thermal power spectral density of the current and voltage fluctuations $S_{i,t}$, $S_{v,t}$, in A^2/Hz, V^2/Hz, can be represented by [92]

$$S_{i,t} = \frac{4k_B\,T\,\triangle f}{R} \tag{4.8}$$

$$S_{v,t} = 4k_B\,T\,\triangle f R \tag{4.9}$$

where k_B is Boltzmann's constant, $\triangle f$ is the effective noise frequency bandwidth and T is the absolute temperature in Kelvin. Ideally, the thermal noise power spectral density is constant throughout the whole frequency spectrum. This is usually referred to as "white noise". The thermal noise defines the fundamental limits in electronic circuits. For example, the noise floor of any electronic circuit operating at room temperature and 1Hz bandwidth is $-174\,\mathrm{dBm/Hz}$ (dBm is the decibel referenced to 1mW) due to the following:

$$\begin{aligned} P_t &= k_B\,T\,\triangle f && (4.10)\\ P_{dBm} &= 10\log(k_B\,T\,\triangle f\,\times\,1000) && (4.11)\\ &= 10\log(k_B\,T\,\times\,1000)\,+\,10\log(\triangle f) && (4.12)\\ &= -174\,+\,10\log(\triangle f) && (4.13) \end{aligned}$$

2. **Shot Noise**: Although the external DC current (I_{DC}) appears to be continuous and smooth in diodes, bipolar transistors and MOSFETs, it is in fact a result of random, independent pulses of current caused by the individual flow of carriers through a potential barrier. The fluctuation of the emission rate of the carriers is referred to as shot noise. Physically, shot noise is composed of thermal and generation-recombination components and dominates the noise characteristics of the device only when operating in the subthreshold region due to the current leakage (carrier transport). Shot noise is a white noise source and its power spectral density (A^2/Hz), is described as [92]

$$\overline{i_n^2} = 2q\,I_{DC}\;=\;S_{i,n}\,\triangle f \tag{4.14}$$

where q is the electronic charge ($1.66 \times 10^{-19}C$), $\triangle f$ is the frequency bandwidth and I_{DC} is the dc current flowing through the potential barrier.

3. **Generation-Recombination Noise (G-R Noise)**: Generation-recombination noise is caused by the fluctuation of the conductance because of the traps and recombination centres in semiconductors. The random trapping and de-trapping of carriers results in the fluctuation in the number of free carriers (N) per unit time, causing the conductance of the device to fluctuate. The power spectral density of G-R noise is given by [92]

$$S_N(f) = \overline{\triangle N^2}.\frac{4\tau}{1+(2\pi f\tau)^2} \tag{4.15}$$

where $\overline{\triangle N^2}$ is the variance of N , f is the frequency and τ is the lifetime of the carriers.

4. **Flicker Noise ($1/f$ Noise)**: Flicker noise is dominant at lower frequencies and is called $1/f$ noise because the spectrum varies as $1/f^{\alpha}$, where α is close to 1. Flicker noise originates from two effects; charge trapping at the Oxide (SiO_2) - Semiconductor (Si) interface and local fluctuations of carrier mobility. These two sources are described by the McWhorter model [93] and the Hooge model [94, 95]. Previous research has indicated both these models to be correlated at higher frequencies [96, 97].

4.2.2.2 Noise in High-Frequency MOSFETs

At high frequencies the main source of noise in a MOSFET is the channel thermal noise [92]. Shot noise is only dominant in the subthreshold region and flicker noise is negligible due to its $1/f$ characteristic.

MOSFET thermal noise originates from the noise source in the channel, as shown in Figure 4.10. The current fluctuation detected at the drain terminal of the MOSFET represents the channel thermal noise and is known as the drain current noise ($\overline{i_d^2}$). At high frequencies, channel thermal noise manifests itself in the gate current spectrum due to the capacitive coupling through the gate oxide. This noise current is known as induced gate noise ($\overline{i_g^2}$). The drain current noise and the induced gate noise are partially

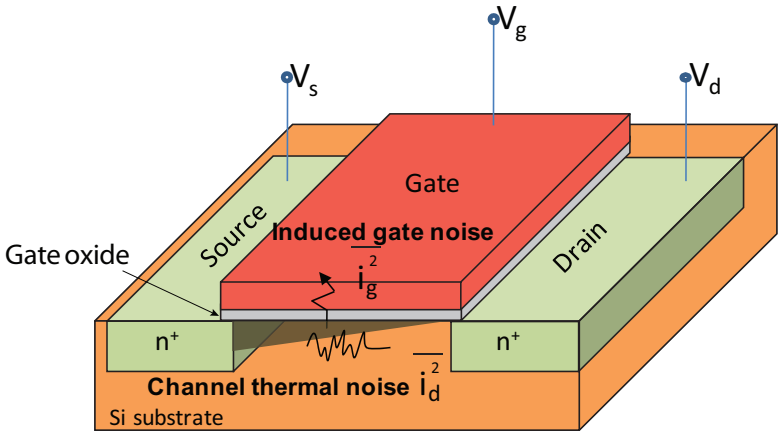


Figure 4.10: Thermal noise in MOSFETs: Gate noise current, drain noise currents and their correlation.

correlated since they originate from the same source, and their correlation is imaginary due to the capacitive nature.

Van der Ziel theory [92] expresses the drain noise current, induced gate noise current, and their correlation as

$$\overline{i_d^2} \doteq 4\,k_B\,T\,\triangle f\,\gamma\,g_{d0} = \overline{i_d i_d^*} \tag{4.16}$$

$$\overline{i_g^2} \doteq 4\,k_B\,T\,\triangle f\,\delta\,g_g = \overline{i_g i_g^*} \tag{4.17}$$

$$\overline{i_g i_d^*} \doteq c\sqrt{\overline{i_g^2}\,\overline{i_d^2}} \tag{4.18}$$

$$g_g \doteq \zeta\,\frac{\omega^2\,C_{gs}^2}{g_{d0}} \tag{4.19}$$

where γ, δ and ζ are bias dependent factors, and c is the cross correlation coefficient. g_{d0} and g_g are the drain output conductance under zero drain bias and the real part of the gate source admittance respectively.

For long channel MOSFETs operating in the saturation region, δ, ζ and c are reportedly 4/3, j0.395 and 1/5 [98] respectively. While γ satisfies the inequality $2/3 \leq \gamma \leq 1$. The value γ=1 is valid when the drain bias is zero, and 2/3 when the device is in the cut-off region [92]. A substantial increase in the values of γ and δ have been reported in short channel devices [99–102].

Figure 4.11 illustrates the most important noise sources in a MOSFET. The thermal noise generated by the distributed gate and substrate resistances, R_g and R_b, become very dominant at high frequencies. The thermal noise voltage generated by the gate and

source resistances is given by

$$\overline{v_{R_g}^2} \doteq 4\,k_B\,T\,\triangle f\,R_g \tag{4.20}$$

$$\overline{v_{R_b}^2} \doteq 4\,k_B\,T\,\triangle f\,R_b \tag{4.21}$$

The thermal noise generated by device terminal resistances, especially the gate and substrate resistances, can be reduced significantly by optimising the device layout. Multi-finger device layout with double-sided gate contacts reduces the gate resistance. Similarly, guard rings surrounding the device reduce the substrate resistance. Finally, placing an abundance of contacts and vias at the gate, drain and source terminals reduces contact resistances, which in turn minimises their thermal noise contribution.

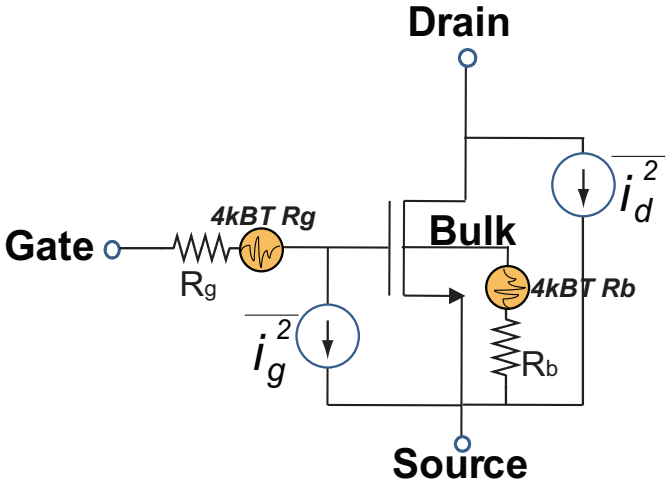


Figure 4.11: High frequency thermal noise sources in MOSFET: Gate noise current, drain noise currents, gate resistance and substrate resistance.

4.2.2.3 Noise Two-Port Theory

Noisy two-port network can be represented, using two-port theory, as a combination of a noiseless network and its corresponding noise sources [103]. This is illustrated in Figure 4.12. Depending on the type of noise source six representations can be derived, but the following three representations are particularly useful to this study: admittance for parallel connection, impedance for series connections and the chain for a cascade of stages [104].

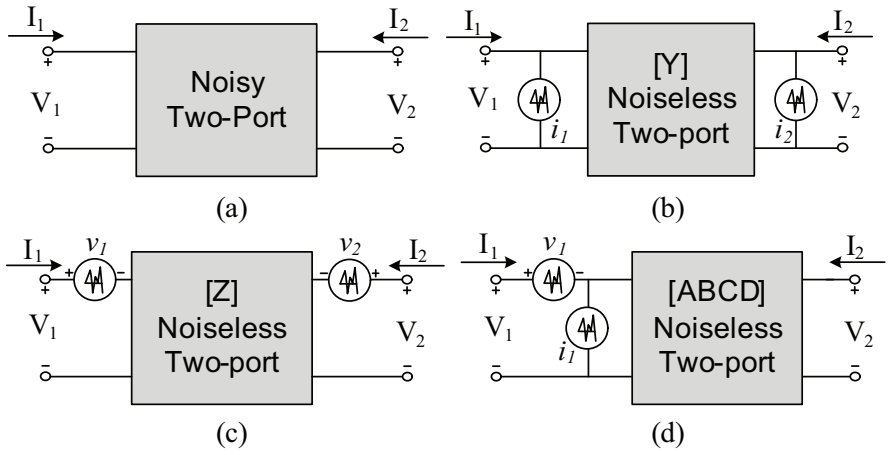


Figure 4.12: Noisy linear two-port networks: (a) General form, (b) admittance form, (c) impedance form and (d) chain matrix form.

1. **Admittance (Y) Noise Representation**

   This representation is useful when two networks $\overrightarrow{Y_1}$ and $\overrightarrow{Y_2}$ are connected in parallel:

$$\overrightarrow{C_Y} = \overrightarrow{C_{Y_1}} + \overrightarrow{C_{Y_2}} = \begin{bmatrix} C_{i_1 i_1^*} & C_{i_1 i_2^*} \\ C_{i_2 i_1^*} & C_{i_2 i_2^*} \end{bmatrix} \tag{4.22}$$

$$\overrightarrow{C_Y} = 2KT\Re\left[\overrightarrow{Y}\right] \tag{4.23}$$

   $\Re[\ ]$ denotes the real part of a complex number and $\overrightarrow{C_Y}$ is the correlation matrix of the admittance representation, in which $C_{i_m i_n^*}$ represents the self or cross-power spectral densities of the input and output current noise sources as illustrated in Figure 4.12(b).

2. **Impedance (Z) Noise Representation**

   This representation applies when two networks $\overrightarrow{Z_1}$ and $\overrightarrow{Z_2}$ are connected in series:

$$\overrightarrow{C_Z} = \overrightarrow{C_{Z_1}} + \overrightarrow{C_{Z_2}} = \begin{bmatrix} C_{v_1 v_1^*} & C_{v_1 v_2^*} \\ C_{v_2 v_1^*} & C_{v_2 v_2^*} \end{bmatrix} \tag{4.24}$$

$$\overrightarrow{C_Z} = 2KT\Re\left[\overrightarrow{Z}\right] \tag{4.25}$$

   where $\overrightarrow{C_Z}$ denotes the correlation matrix of the impedance representation, in which $C_{v_m v_n^*}$ represents the self or cross-power spectral densities of the input and output voltage noise sources as illustrated in Figure 4.12(c).

3. **Chain (ABCD) Noise Representation**

This representation is useful when two networks $\overrightarrow{A_1}$ and $\overrightarrow{A_2}$ are cascaded:

$$\overrightarrow{C_A} = \overrightarrow{A_1}\,\overrightarrow{C_{A_2}}\,\overrightarrow{A_1^\dagger} + \overrightarrow{C_{A_1}} = \begin{bmatrix} C_{v_n v_n^*} & C_{v_n i_n^*} \\ C_{i_n v_n^*} & C_{i_n i_n^*} \end{bmatrix} \tag{4.26}$$

$$\overrightarrow{C_A} = 2KT \begin{bmatrix} R_n & \frac{F_{min}-1}{2} - R_n Y_{opt}^* \\ \frac{F_{min}-1}{2} - R_n Y_{opt} & R_n \left|Y_{opt}\right|^2 \end{bmatrix} \tag{4.27}$$

where $\overrightarrow{C_A}$ denotes the correlation matrix of the ABCD representation, in which $C_{v_n i_n^*}$ represents the self or cross-power spectral densities of the input referred noise sources as illustrated in Figure 4.12(d). This representation is very important for this study as the correlation matrix can be directly obtained from the measured device noise characteristics.

- **Noise Representations Transformation**

Table 4.1: Transformation matrices to calculate the noise matrices of the other representations.

		Original Representation		
		C_Y	C_Z	C_A
Result Representation	Admittance C_Y'	$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$	$\begin{bmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{bmatrix}$	$\begin{bmatrix} -Y_{11} & 1 \\ -Y_{21} & 0 \end{bmatrix}$
	Impedance C_Z'	$\begin{bmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{bmatrix}$	$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & -Z_{11} \\ 0 & -Z_{21} \end{bmatrix}$
	ABCD C_A'	$\begin{bmatrix} 0 & A_{12} \\ 1 & A_{22} \end{bmatrix}$	$\begin{bmatrix} 1 & -A_{11} \\ 0 & -A_{21} \end{bmatrix}$	$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

Each representation can be transformed into another by the matrix operation:

$$\overrightarrow{C'} = \overrightarrow{T}\,\overrightarrow{C}\,\overrightarrow{T^\dagger} \tag{4.28}$$

where $\overrightarrow{C}$ and $\overrightarrow{C'}$ are the noise correlation matrices of the original and resulting representations respectively, $\overrightarrow{T}$ is the transformation matrix given in Table 4.1 and $\overrightarrow{T^\dagger}$ is the transpose conjugate of $\overrightarrow{T}$.

4.2.2.4 MOSFET Noise parameters

When a noisy network is connected to a noise source ($\overline{i_s^2}$) as illustrated in Figure 4.13, the noise figure of the network can be derived in terms of the R_n, G_u, G_s, G_c and B_c representation as follows:

$$F = \frac{\text{Total equivalent input noise power}}{\text{Noise power of source impedance}} \tag{4.29}$$

$$F = \frac{\overline{i_s^2} + \overline{\left|i_n + Y_s v_n\right|^2}}{\overline{i_s^2}} \tag{4.30}$$

$$= \frac{\overline{i_s^2} + \overline{\left|(i_c + i_u) + Y_s v_n\right|^2}}{\overline{i_s^2}} \tag{4.31}$$

$$= \frac{\overline{i_s^2} + \overline{\left|(Y_c v_n + i_u) + Y_s v_n\right|^2}}{\overline{i_s^2}} \tag{4.32}$$

$$= \frac{\overline{i_s^2} + \overline{\left|i_u + (Y_c + Y_s) v_n\right|^2}}{\overline{i_s^2}} \tag{4.33}$$

$$= 1 + \frac{\overline{i_u^2} + |Y_c + Y_s|^2 \overline{v_n^2}}{\overline{i_s^2}} \tag{4.34}$$

$$= 1 + \frac{\overline{i_u^2} + [(G_c + G_s)^2 + (B_c + B_s)^2] \overline{v_n^2}}{\overline{i_s^2}} \tag{4.35}$$

$$= 1 + \frac{G_u + [(G_c + G_s)^2 + (B_c + B_s)^2] R_n}{G_s} \tag{4.36}$$

$$\overline{i_s^2} = 4KT \triangle f \, G_s \tag{4.37}$$

$$\overline{i_u^2} = 4KT \triangle f \, G_u \tag{4.38}$$

$$\overline{v_n^2} = 4KT \triangle f \, R_n \tag{4.39}$$

where i_c is the component of i_n that is correlated with v_n, i_u is the component of i_n that is uncorrelated with v_n, and Y_c is the correlation admittance.

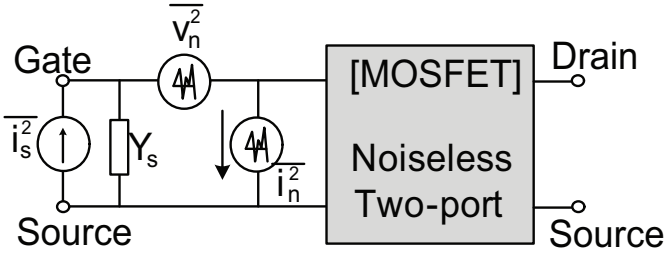


Figure 4.13: Circuit diagram of general two-port networks for noise figure calculation.

Since the derived expression of the noise figure is a function of the source admittance (Y_s), taking the derivative of Equation 4.36 with respect to Y_s yields the condition to achieve the minimum noise figure:

$$\frac{\partial F}{\partial G_s} = \frac{-G_u - [G_c^2 - G_s^2 + (B_c + B_s)^2]\, R_n}{G_s^2} \tag{4.40}$$

$$\frac{\partial F}{\partial B_s} = \frac{2(B_c + B_s)\, R_n}{G_s} \tag{4.41}$$

The source admittance which results in Equation 4.40 and Equation 4.41 equating to zero is determined by:

$$G_{opt} = \sqrt{\frac{G_u}{R_n} + G_c^2} \tag{4.42}$$

$$B_{opt} = -B_c \tag{4.43}$$

Introducing these two newly defined quantities removes G_u from Equation 4.36:

$$F = 1 + \frac{(G_{opt}^2 - G_{opt}^2)R_n}{G_s} + \frac{[(G_c + G_{opt}) + (G_s - G_{opt})]^2\, R_n}{G_s} + \frac{[(B_c + B_{opt}) + (B_s - B_{opt})]^2\, R_n}{G_s} \tag{4.44}$$

As $B_c + B_{opt} = 0$, rearranging Equation 4.44 yields another representation of the noise figure F_{min}, R_n, G_{opt} and B_{opt} as follows:

$$F = F_{min} + \frac{|Y_s - Y_{opt}|^2\, R_n}{G_s} \tag{4.45}$$

$$F_{min} = 1 + 2R_n(G_{opt} + G_c) \tag{4.46}$$

4.3 MOSFET Small-Signal Modelling

In this section, methods for extracting the small-signal equivalent circuit models are reviewed. Following this, the chosen extraction methodology is described. The proposed method is implemented and verified on a multi-fingered RF n-MOSFET with a W/L of 80 μm/0.18 μm formed from 40 gate fingers, each 2 μm wide.

4.3.1 Equivalent Circuit Model Extraction - Review

The extraction of the MOSFET small-signal equivalent circuit parameters is very important for circuit development and device performance optimisation. Several methods of extracting small-signal model parameters, from measured S-parameter data, have been reported [105–114]. The vast majority are based on the method presented by Dambrine [105]. The author explained a general method of parameter extraction, which is based on extracting the bias independent extrinsic components using analytical equations obtained from the device measured S-parameters, and then subtracting these extrinsic components from the measured S-parameters to obtain values for the intrinsic components. The work presented in [106] and [107] are similar. Both methods omit the extrinsic gate, drain and source inductances, as these models run up to 10 and 20 GHz respectively. The effect of the parasitic inductances is very small up to 10 GHz [108], therefore they can be omitted from the models. However, to achieve more accurate models and better admittance fitting up to 110 GHz the terminal inductances have to be included [114]. Similarly in [108], R_s and R_d were omitted for simplicity and ease of extraction. This practice could lead to significant variation between the measured and modelled S_{21} of the device. In [106] and [108] an extra transcapacitance (C_m) is introduced between the gate and the source. The transcapacitance is introduced to account for the capacitance between the gate and the drain, which is split between two non-reciprocal capacitances C_{gd} and C_{dg}, which are caused by the drain and gate biasing respectively where $C_m = C_{dg}$ - C_{gd}. According to [108] the transcapacitance takes care of the charging currents effects between the gate and drain and should be included in order to model Y_{21} and Y_{12} correctly. Additionally [108] and [106] provide a clear and much simpler method to extract the substrate parameters in comparison to all the other methods. The methods presented in [109–111] are complicated and involve a significant amount of mathematical manipulation and complex curve fitting. The extraction of the extrinsic resistances proposed in [107] was done by applying zero bias to the gate and drain. This eliminated the contribution of the voltage dependant parameters and hence reduced the equivalent circuit to R_g, R_d, R_s, C_{gs}, C_{gd} and C_{ds}. Analytical equations based on measured Y-parameters were then used to obtain values for the terminal resistances. The method presented in [106] is very similar to that in [107] except that the transistor was forward biased at the gate and zero biased at the drain. In such a case the device operates in the linear region and C_{gs} is approximately equal to C_{gd}, which makes

the effect of R_{ds} very small therefore it can then be neglected. The intrinsic behaviour of the transistor becomes fully symmetric with respect to the drain and the source terminals. Therefore, the effects of the transconductances and the intrinsic capacitances including the gate-bulk capacitance become very small and can be neglected [84]. This reduced the equivalent circuit to that presented in [107], but with the exception of C_{gd} which was neglected. The intrinsic extraction method in [107] was also adopted in [112].

4.3.2 Implemented Methodology for Parameters Extraction

In order to obtain the most suitable method for this work, an assessment process for every reported extraction method and its associated model was performed. The method presented in [106] was found to be the most suitable, since it is based on a model for a similar 0.18 μm RF Silicon MOSFET process. A few alterations were introduced, hence these adopted from [108] and [113]. For example the extrinsic terminal resistances were extracted prior to the intrinsic parameters, and the transcapacitance (C_m) was neglected as it has been demonstrated to have little impact on the results of S-parameter simulation [113].

4.3.2.1 Extrinsic Parameters Extraction

The extrinsic parameters are extracted as follows:

1. Bias the transistor at $V_{ds} = 0$V and a $V_{gs} = 1.5$V and measure the S-parameters.
2. Perform the de-embedding to remove the test fixture (probe pads) effects if needed, and extract the de-embedded S-parameters.
3. Convert the S-parameters to Z and Y-parameters.
4. Extract the terminal series resistances R_g, R_s and R_d using

$$R_s = \Re(Z_{12}) = \Re(Z_{21}) \tag{4.47}$$

$$R_d = \Re(Z_{22}) - R_s \tag{4.48}$$

$$R_g = \Re(Z_{11}) - R_s \tag{4.49}$$

5. Extract the terminal inductances (L_g, L_s and L_d) using

$$L_s = \frac{\Im(Z_{12})}{\omega} \tag{4.50}$$

$$L_d = \frac{\Im(Z_{22})}{\omega} - L_s \tag{4.51}$$

$$L_g = \frac{\Im(Z_{11})}{\omega} - L_s \tag{4.52}$$

where $\omega = 2\pi f$ and $\Re(\ )$ and $\Im(\ )$ are the real and imaginary parts of a complex number.

4.3.2.2 Intrinsic Parameters Extraction

The intrinsic parameters are extracted as follows:

1. Bias the transistor at the desired gate and drain voltages and measure the S-parameters.
2. Perform the de-embedding, if needed, and extract the de-embedded S-parameters.
3. Convert S-parameters to Z and Y-parameters.
4. Subtract the extrinsic elements from the measured network parameters mathematically or by using any circuit simulator. This can be achieved by connecting a negative resistance and inductance to the corresponding terminal as shown in Figure 4.14.
5. The equivalent circuit model of the intrinsic device can be represented as follows [108, 109]:

$$Y_i = \begin{bmatrix} \frac{r_{ch}C_{gs}^2\omega^2}{D_{gs}} + \frac{C_{gd}^2\omega^2}{D_{gd}} + j\omega\left(\frac{C_{gs}}{D_{gs}} + \frac{C_{gd}}{D_{gd}}\right) & -\left(\frac{C_{gd}^2\omega^2}{D_{gd}} + j\omega\frac{C_{gd}}{D_{gd}}\right) \\ \left(\frac{g_m|e^{-j\omega\tau}|}{1+j\omega C_{gs}r_{ch}} - \frac{C_{gd}^2\omega^2}{D_{gd}} - j\omega\frac{C_{gd}}{D_{gd}}\right) & g_{ds} + \omega^2C_{gd}^2 + \frac{\omega^2C_{jd}^2R_{bd}}{D_{jd}} + j\omega(C_{ds} + C_{gd} + \frac{C_{jd}}{D_{jd}}) \end{bmatrix} \tag{4.53}$$

where

$$D_{gs} = 1 + (\omega C_{gs}r_{ch})^2 \tag{4.54}$$

$$D_{gd} = 1 + (\omega C_{gd})^2 \tag{4.55}$$

$$D_{jd} = 1 + (\omega C_{jd}R_{bd})^2 \tag{4.56}$$

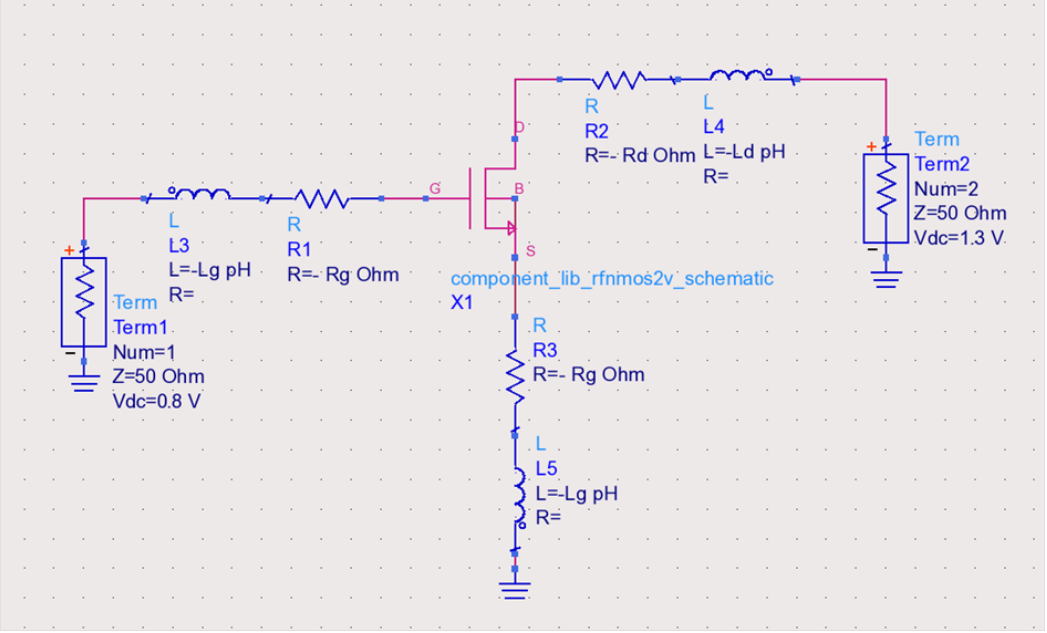


Figure 4.14: ADS schematic setup for extracting the intrinsic model parameters.

6. Extract g_m and g_{ds} from Y- parameters at $\omega = 0$ using

$$g_m = Re[Y_{21}]_{\omega=0} \tag{4.57}$$

$$g_{ds} = Re[Y_{22}]_{\omega=0} \tag{4.58}$$

7. Extract the intrinsic parameters using the following equations:

$$C_{gd} = -\frac{Im[Y_{12}]}{\omega} \tag{4.59}$$

$$C_{gs} = \frac{Im[Y_{11}] - \omega C_{gd}}{\omega}\left(1 + \frac{Re[Y_{11}]^2}{(Im[Y_{11}] - \omega C_{gd})^2}\right) \tag{4.60}$$

$$r_{ch} = \frac{Re[Y_{11}]}{(Im[Y_{11}] - \omega C_{gd})^2 + Re[Y_{11}]^2} \tag{4.61}$$

8. Calculate τ from

$$Y_{21} = \frac{g_m|e^{-j\omega\tau}|}{1 + j\omega C_{gs} r_{ch}} - \frac{C_{gd}^2\omega^2}{D_{gd}} - j\omega\frac{C_{gd}}{D_{gd}} \tag{4.62}$$

as follows:

$$Re[Y_{21} - Y_{12}] = \alpha(\cos\Theta - \beta\sin\Theta) \tag{4.63}$$

$$Im[Y_{21} - Y_{12}] = -\alpha(\beta\cos\Theta + \sin\Theta) \tag{4.64}$$

where

$$\Theta = \omega\tau \tag{4.65}$$

$$\beta = \omega C_{gs} r_{ch} \tag{4.66}$$

$$\alpha = \frac{g_m}{1 + \beta^2} \tag{4.67}$$

By rearranging Equation 4.63 and Equation 4.64, τ is obtained as follows:

$$\cos\Theta = \frac{Re[Y_{21} - Y_{12}]}{\alpha} + \beta \sin\Theta) \tag{4.68}$$

$$\sin\Theta = \frac{-(Im[Y_{21} - Y_{12}]) + \beta(Re[Y_{21} - Y_{12}])}{\alpha(1 + \beta)} \tag{4.69}$$

$$\tau = \frac{\arcsin(\frac{-(Im[Y_{21}-Y_{12}])+\beta(Re[Y_{21}-Y_{12}])}{\alpha(1+\beta)})}{\omega} \tag{4.70}$$

9. Extract the substrate parameters; this is a little more complex than the other parameters as it is not directly derived from the Y or Z-parameters. The admittance representation of the substrate network can be expressed as

$$Y_{sub} = Re[Y_{22}] - g_{ds} - \omega^2 C_{gd}^2 \tag{4.71}$$

$$\frac{\omega^2}{Y_{sub}} = \omega^2 R_{bd} + \frac{1}{C_{jd}^2 R_{bd}} \tag{4.72}$$

It is evident that Equation 4.72 is a straight line equation of the form $f(x) = mx+b$ where $f_x = \frac{\omega^2}{Y_{sub}}$, x= ω^2, m = R_{bd} and b = $\frac{1}{C_{jd}^2 R_{bd}}$. Therefore, R_{bd} can be extracted from the slope of the linear regression of $\frac{\omega^2}{\Re(Y_{sub})}$ versus ω^2. Once the value of R_{bd} is obtained, the substrate capacitance can be extracted using

$$C_{jd} = \frac{1}{\sqrt{\frac{\omega^2 R_{bd}}{Re[Y_{22}] - g_{ds} - \omega^2 C_{gd}^2} - \omega^2 R_{bd}^2}} \tag{4.73}$$

Finally C_{ds} can be extracted as follows:

$$C_{ds} = \frac{Im[Y_{22}]}{\omega} - \frac{C_{jd}}{D_{jd}} - C_{gd} \tag{4.74}$$

4.3.3 Experimental Results and Discussion

An experiment was conducted on a multi-fingered n-MOSFET with a W/L of 80 μm/0.18 μm formed from 40 gate fingers, each 2 μm wide, biased at I_{ds} =7.4mA (V_{gs} =0.8V and V_{ds} =1.3V). The device model parameters were extracted based on the aforementioned extraction technique for which the results are shown in Figure 4.15 and Figure 4.16 for the extrinsic and intrinsic parts respectively.

Figure 4.15(a) and (b) depict the extracted values of the extrinsic resistors and inductors with the calculated mean value of each parameter. It is evident that the extracted resistance values vary by no more than 11% beyond 2 GHz.

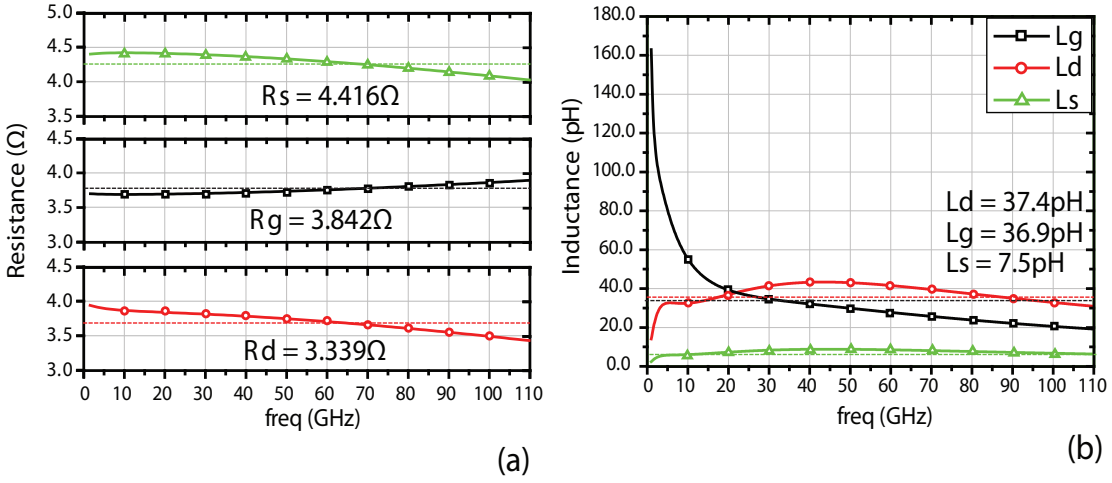


Figure 4.15: Plot of extracted extrinsic parameters: (a) Gate, drain and source resistance and (b) gate, drain and source inductances.

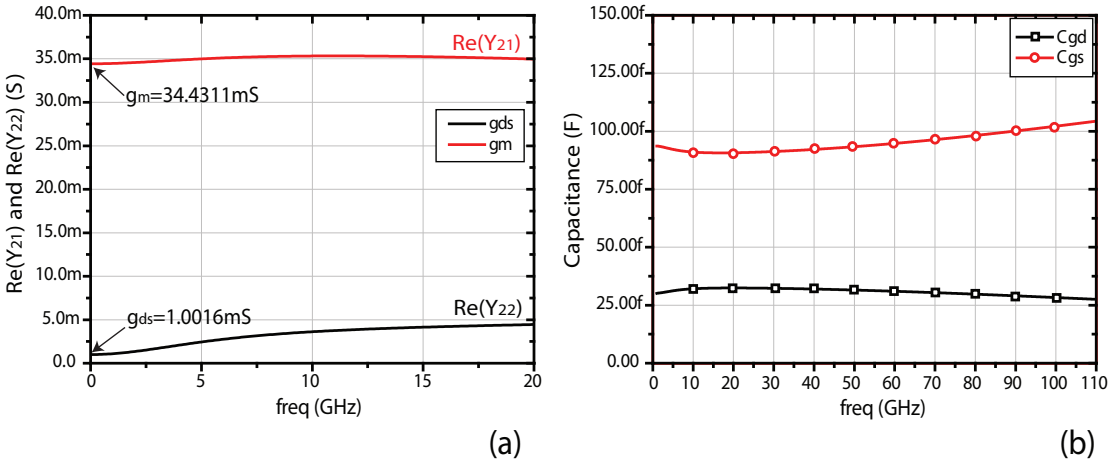


Figure 4.16: Plots of extracted (a) g_m and g_{ds} and (b) C_{gd} and C_{gs}.

Figure 4.16(a) depicts the extracted transconductance (g_m) and the output conductance (g_{ds}) from the real part of Y_{21} and Y_{22} respectively at ω=0. It can be seen

from Figure 4.16(b), the extracted capacitances are almost constant over the frequency spectrum. Similarly the extracted resistances shown in Figure 4.15(a) remain almost constant throughout the entire frequency spectrum. This demonstrates the validity of the chosen parameter extraction method and implies that the pad de-embedding is done correctly.

As stated previously Equation 4.72 is a straight line equation, m = R_{bd} and b = $\frac{1}{C_{jd}^2 R_{bd}}$. Therefore, R_{bd} can be extracted from the slope of the linear regression of $\frac{\omega^2}{\Re(Y_{sub})}$ versus ω^2 as shown in from Figure 4.17(a). The extracted substrate capacitance is illustrated in Figure 4.17(b).

The model parameters extracted using the discussed method provide accurate values up to 40 GHz. In order to achieve higher accuracy and better fitting between the simulated and measured responses across the whole frequency spectrum, further tuning or optimisation of all the model component values should be performed. Table 4.2 shows the extracted parameter values and the optimised parameter values for the transistor used in this study. It can be observed that the optimised and extracted values are similar. This further demonstrates the validity of the proposed method to provide a reliable model extraction up to 110 GHz.

The proposed method is also compared to a modified BSIM3v3 model. The BSIM3 model is an industry standard MOSFET model for deep sub-micrometre applications [115]. Recent studies have shown that the BSIM3 model needs some improvements to be suitable for CMOS circuit design at RF and microwave frequencies [116]. The BSIM3v3

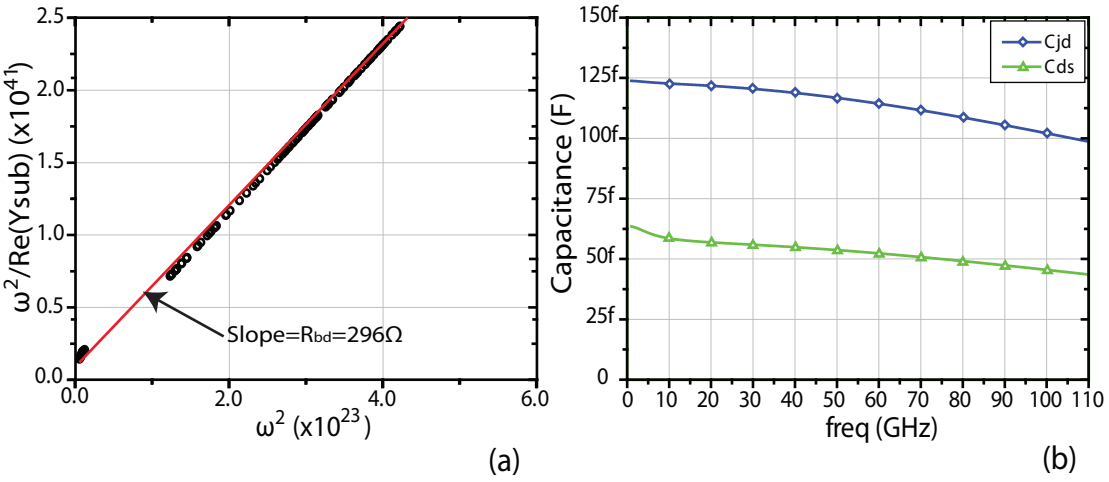


Figure 4.17: Plots of extracted (a) substrate resistance (R_{bd}) and (b) substrate capacitances (C_{ds} and C_{jd}).

Table 4.2: Extracted and optimised small-signal model parameter values.

Model Parameter	Extracted Value	Optimised Value
g_m, mS	34.43	34.43
g_{ds}, mS	1.0016	1.1
C_{gd}, fF	30.7	30.7
C_{gs}, fF	95.35	95
C_{ds}, fF	52.58	53
R_g, Ω	3.84	5.1
R_d, Ω	3.39	3.4
R_s, Ω	4.42	4.2
L_g, pH	36.9	37
L_d, pH	37.4	37
L_s, pH	7.5	10.8
$R_{bd} = R_{bs}$, Ω	592	600
$C_{jbd} = C_{jbs}$, fF	114	115
τ, ps	1.37	1.37
r_{ch}, Ω	1.92	1.92

model, obtained from the foundry design kit, was adapted in this study and is used as a core model as shown in Figure 4.18. Additional terminal resistance and inductance at the gate, drain, source and substrate terminals were introduced. Capacitances and

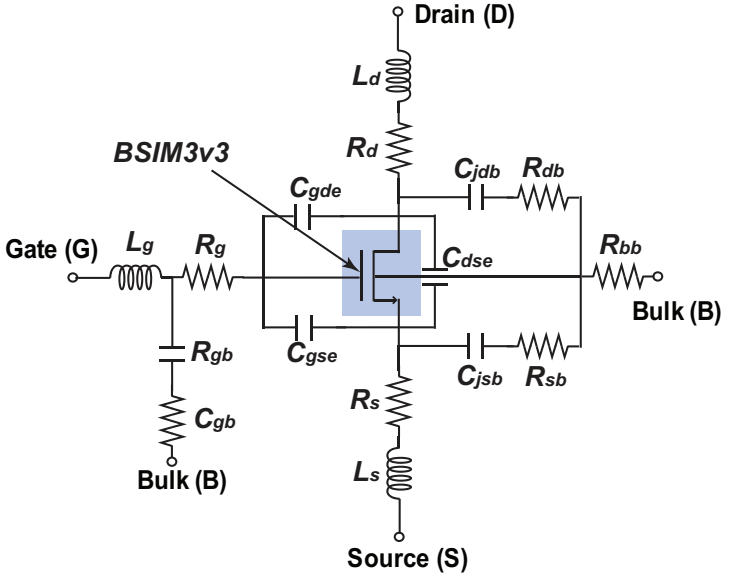


Figure 4.18: High frequency equivalent circuit model based on BSIM3v3 core.

resistances are added to model the source and drain junction with the substrate. Capacitances also are added between the gate-source and gate-drain terminals to compensate for any inaccuracy in the internal capacitance model.

Figure 4.19 shows the the Y-parameters of the measured and simulated model results, biased at $V_{gs} = 0.8$V and $V_{ds} = 1.3$, obtained using the proposed equivalent circuit

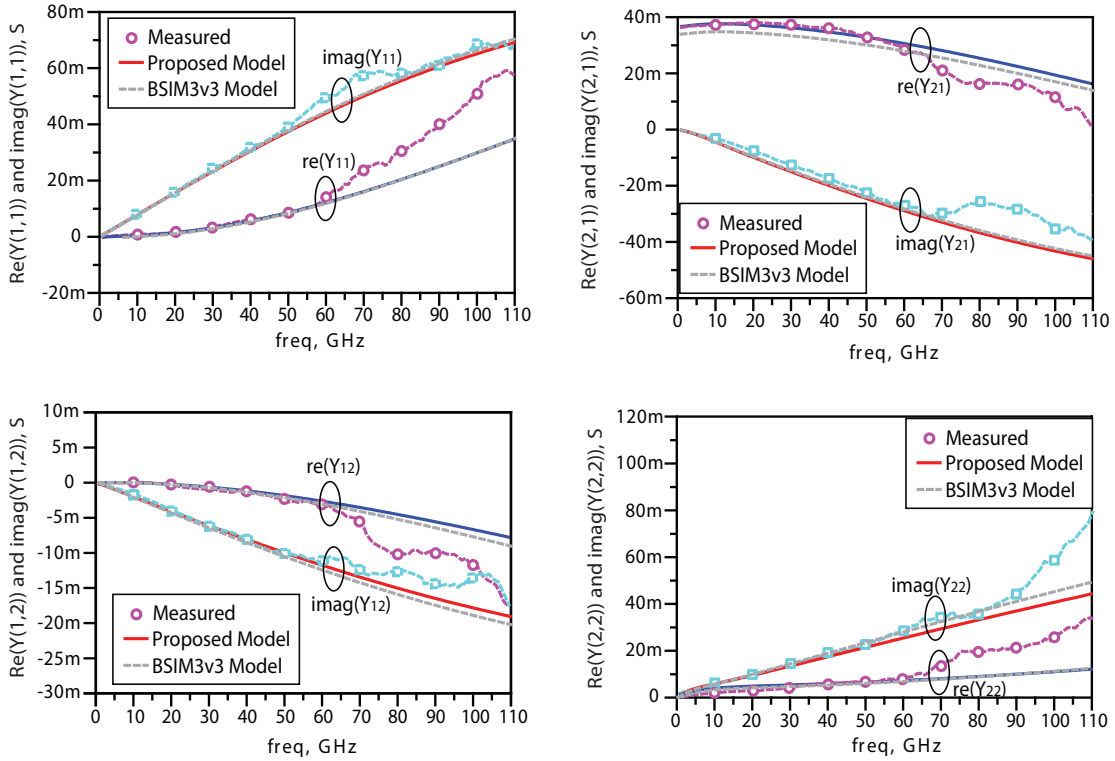


Figure 4.19: Y-parameters of measurement (symbol), proposed model (solid line) and BSIM3v3 model (dotted line) for the 80 μm, 40 finger n-MOSFET biased at $V_{gs} = 0.8V$ and $V_{ds} = 1.3V$.

model. It can been seen that the proposed model is as good as the BSIM3v3 model, and most importantly the simulated model results agree with the measured device data up to 60 GHz. The total error [117] between modelled and measured Y-parameters is only 0.1% up to 60 GHz. As the frequency increases, the fitting of the output admittance Y_{22} becomes increasingly cumbersome due to the increase in substrate coupling effects through the source and drain junction capacitance and substrate resistance [108]. This causes the discrepancy in the output admittance Y_{22} response at very high frequencies. Similarly C_{gs} significantly affects the input admittance Y_{11} at high frequency [108] and C_{gd} heavily dominates the forward and reverse admittance Y_{12} and Y_{21}.

4.4 MOSFET Noise Modelling

The aim of this section is to determine the MOSFET noise parameters, which when used in conjunction with the small-signal device model developed earlier in this work, creates a complete MOSFET model that enables the design of low noise circuits, such as low noise amplifiers.

Based on the two-port noise theory, presented in Section 4.2.2.3, a method for extracting MOSFET's noise parameters such as channel thermal noise, induced gate noise and their cross-correlation from the measured S-parameters and RF noise parameters, is discussed. The method is then tested and verified on a multi-fingered RF n-MOSFET with a W/L of 80 μm/0.18 μm formed from 40 gate fingers, each 2 μm wide.

4.4.1 Noise Parameters Extraction Method

The noise model of a CMOS transistor is shown in Figure 4.20. It is simply the small-signal equivalent model of transistor with intrinsic device noise sources introduced at the gate and drain nodes. The mean-square noise currents due to the drain and gate noise are modelled with noise generators i_{nd} and i_{ng} respectively. The MOSFET noise parameters NF_{min},R_n and Y_{opt} can be obtained from the measured noise parameters using the noise correlation matrix representation of a two-port device. The noise correlation matrix of a two-port device is given by

$$\overrightarrow{C_A} = \begin{bmatrix} C_{v_n v_n^*} & C_{v_n i_n^*} \\ C_{i_n v_n^*} & C_{i_n i_n^*} \end{bmatrix} \tag{4.75}$$

$$= 2KT \begin{bmatrix} R_n & \frac{F_{min}-1}{2} - R_n Y_{opt}^* \\ \frac{F_{min}-1}{2} - R_n Y_{opt} & R_n \left|Y_{opt}\right|^2 \end{bmatrix} \tag{4.76}$$

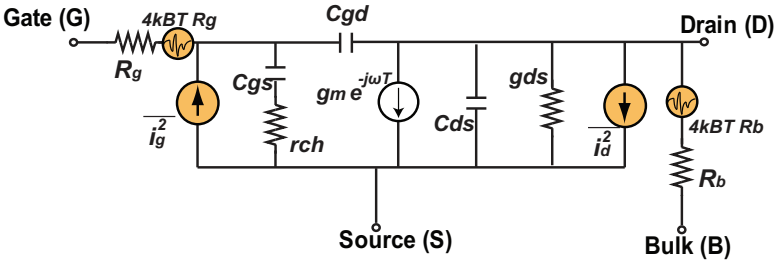


Figure 4.20: High frequency noise small-signal equivalent model of a MOSFET.

From Equation 4.75 and Equation 4.76, the device noise parameters can be written in terms of the measured noise parameters as follows:

$$F_{min} = 1 + \frac{\overline{v_n i_n^*} + \overline{v_n^2} Y_{opt}^*}{2KT\triangle f} \tag{4.77}$$

$$R_n = \frac{\overline{v_n^2}}{4KT\triangle f} \tag{4.78}$$

$$Y_{opt} = \sqrt{\frac{\overline{i_n^2}}{\overline{v_n^2}} - \left(\Im\left[\frac{\overline{v_n i_n^*}}{\overline{v_n^2}}\right]\right)^2} + j\Im\left[\frac{\overline{v_n i_n^*}}{\overline{v_n^2}}\right] \tag{4.79}$$

Direct transformation using Table 4.1 enables the representation of these parameters in terms of channel thermal noise ($\overline{i_d^2}$) and induced gate noise ($\overline{i_g^2}$) as follows:

$$\overline{v_n^2} = |A_{12}|^2 \overline{i_d^2} \tag{4.80}$$

$$\overline{v_n i_n^*} = A_{12} \overline{i_d i_g^*} + A_{12} A_{22}^* \overline{i_d^2} \tag{4.81}$$

$$\overline{i_n^2} = \overline{i_g^2} + A_{22} \overline{i_d i_g^*} + A_{22}^* \overline{i_g i_d^*} + |A_{22}|^2 \overline{i_d^2} \tag{4.82}$$

Substituting these expressions into Equation 4.77, Equation 4.78 and Equation 4.79 yields

$$F_{min} = 1 + \frac{|A_{12}|^2 \overline{i_d^2}}{2KT\triangle f}\left[\sqrt{\frac{\overline{i_g^2} - |\overline{i_g i_d^*}|^2/\overline{i_d^2}}{|A_{12}|^2 \overline{i_d^2}} + \Re\left(\frac{A_{22}}{A_{12}} + \frac{\overline{i_g i_d^*}}{A_{12}\overline{i_d^2}}\right)^2} + \Re\left(\frac{A_{22}}{A_{12}} + \frac{\overline{i_g i_d^*}}{A_{12}\overline{i_d^2}}\right)\right] \tag{4.83}$$

$$R_n = |A_{12}|^2 \frac{\overline{i_d^2}}{4KT\triangle f} \tag{4.84}$$

$$Y_{opt} = \sqrt{\frac{\overline{i_g^2} - |\overline{i_g i_d^*}|^2/\overline{i_d^2}}{|A_{12}|^2 \overline{i_d^2}} + \Re\left(\frac{A_{22}}{A_{12}} + \frac{\overline{i_g i_d^*}}{A_{12}\overline{i_d^2}}\right)^2} - j\Im\left(\frac{A_{22}}{A_{12}} + \frac{\overline{i_g i_d^*}}{A_{12}\overline{i_d^2}}\right)^2 \tag{4.85}$$

The network parameters can be approximated as

$$A_{12} \approx -\frac{1}{g_m} \tag{4.86}$$

$$A_{22} \approx -\frac{j\omega C_{gs}}{g_m} \tag{4.87}$$

Combining these equations with the drain and gate current noise power spectrum gives

$$F_{min} \approx 1 + \frac{\omega}{\omega_T}\sqrt{\gamma\,\delta\,\zeta\,(1 - |c|^2)} \tag{4.88}$$

$$R_n \approx \frac{\gamma\, g_{d0}}{g_m^2} \tag{4.89}$$

$$G_{opt} \approx \frac{g_m \omega C_{gs}}{g_{d0}}\sqrt{\frac{\delta\,\zeta\,(1 - |c|^2)}{\gamma}} \tag{4.90}$$

$$G_{opt} \approx -\omega C_{gs}\left(1 - c\frac{g_m}{g_{d0}}\sqrt{\frac{\delta\,\zeta}{\gamma}}\right) \tag{4.91}$$

Inversely, $\overline{i_g^2}$, $\overline{i_g i_d^*}$, $\overline{i_d^2}$ can be expressed in terms of F_{min}, R_n, G_{opt}, B_{opt} as

$$\overline{i_g^2} = 4KT\triangle f\, R_n\big(|Y_{opt}|^2 + |Y_{11}|^2 - 2\Re[Y_{11}Y_c^*]\big) \tag{4.92}$$

$$\overline{i_g i_d^*} = 4KT\triangle f\, R_n(Y_{opt} - Y_c)Y_{21}^* \tag{4.93}$$

$$\overline{i_d^2} = 4KT\triangle f\, R_n|Y_{21}|^2 \tag{4.94}$$

$$Y_c = \frac{F_{min} - 1}{2R_n} - Y_{opt} \tag{4.95}$$

4.4.2 Experimental Results and Discussion

An experiment based on a multi-fingered n-MOSFET with a W/L of 80 μm/0.18 μm formed from 40 gate fingers, each 2 μm wide, biased at I_{ds} =7.4mA (V_{gs} =0.8V and V_{ds} =1.3V) was undertaken. The intrinsic device noise sources were extracted from the measured noise parameters and the elements of the high frequency small-signal model, using the procedure described in the previous section.

Figure 4.21 shows the extracted drain noise ($\overline{i_d^2}$), induced gate noise ($\overline{i_g^2}$), and their correlation ($\overline{i_g i_d^*}$) versus frequency characteristics. It can be seen that $\overline{i_d^2}$ is independent of frequency, whereas $\overline{i_g^2}$ increases in proportion to f^2. The noise correlation $\overline{i_g i_d^*}$ increases in proportion to frequency, which agrees with the theoretical study. The gate shot noise is estimated using Equation 4.14. For I_g = 0.3 μA, the shot noise is estimated to be about 9.96 x 10^{-26} A^2/Hz at the frequency and bias condition of interest. This shows that the gate shot noise is insignificant at high frequencies.

The correlation between channel thermal noise and induced gate noise as a function of frequency is shown in Figure 4.21(b). The imaginary part of the correlation term is

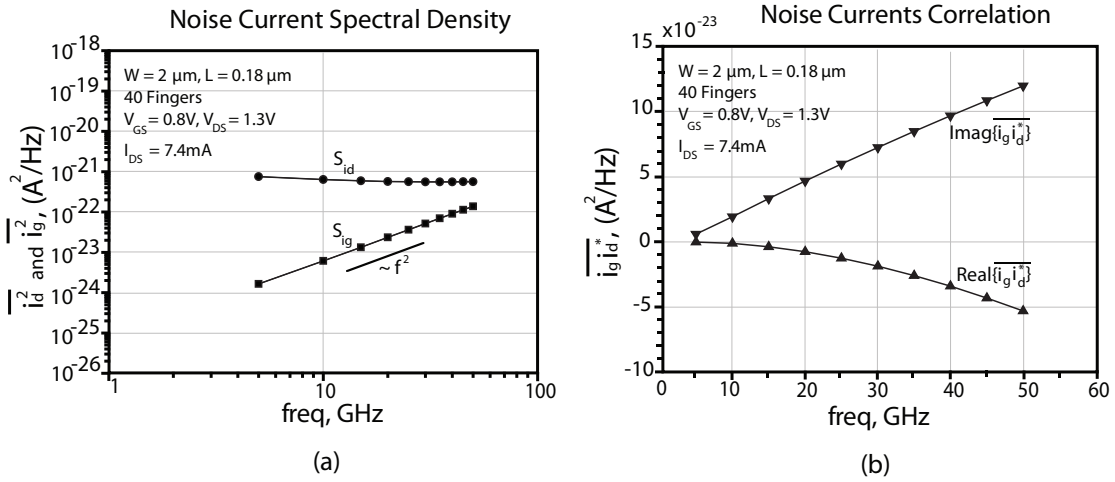


Figure 4.21: (a) Drain thermal noise ($\overline{i_d^2}$), gate current noise ($\overline{i_g^2}$) and (b) their correlation ($\overline{i_g i_d^*}$)(b) plotted versus frequency, for a 40 finger NMOS with a W/L of 80/0.18 biased at I_{ds} =7.4mA (V_{gs} =0.8V and V_{ds} =1.3V).

proportional to the frequency, whereas the real part of the correlation term is almost zero up to about 10 GHz. This is mainly due to the capacitive coupling of the gate-induced noise through C_{gs}.

The dependence of the channel thermal noise and induced gate noise on the drain current, at a fixed drain voltage and varying gate voltage, is shown in Figure 4.22. The channel thermal noise shows a strong bias dependence and increases as V_{gs} increases. The induced gate noise shows a weaker bias dependence and increases slightly with increase in V_{gs}. This is due to the increase in the effective channel resistance due to velocity saturation at high gate bias conditions. The bias dependence of noise correlation between the channel thermal noise and induced gate noise is plotted in Figure 4.22(b). The imaginary part of the correlation term increases with increasing gate bias, whereas the real term exhibits a weak dependence on the gate bias.

The drain and gate noise factors, γ and δ, are evaluated from $\overline{i_d^2}$ and $\overline{i_g^2}$, respectively. Figure 4.23 shows the bias dependence of γ and δ. It is observed that both factors increase with increasing V_{gs}. There is a moderate enhancement of γ compared to the theoretical long channel value of 2/3. This enhancement in short-channel MOSFETs is attributed to channel length modulation and carrier heating. The maximum value of γ in the saturation region is around 2, which is in agreement with the values reported for 0.18 μm CMOS technology [118–120]. It is observed that δ is close to the long-channel value of 4/3 for small gate bias and becomes as large as 8 at higher V_{gs}.

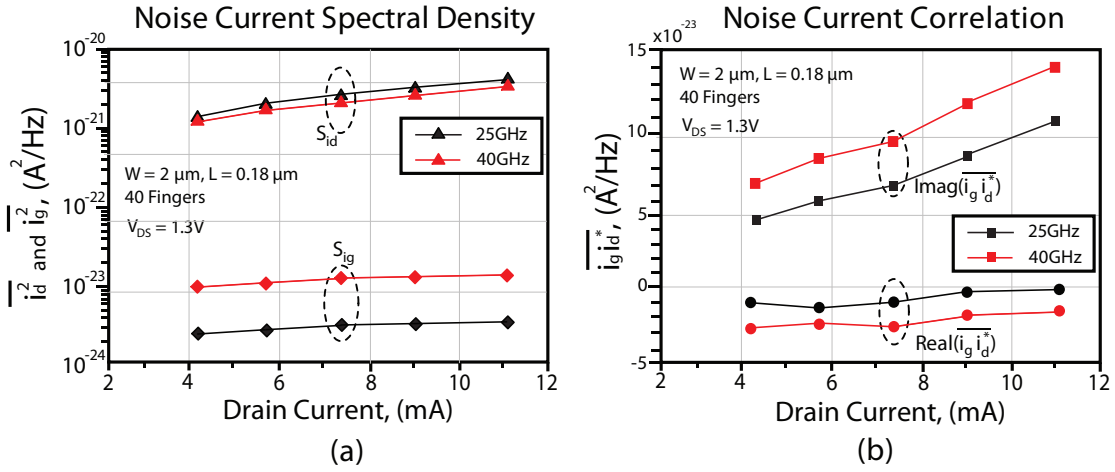


Figure 4.22: (a) Drain thermal noise ($\overline{i_d^2}$), gate current noise ($\overline{i_g^2}$) and (b) their correlation ($\overline{i_g i_d^*}$) plotted versus drain current at 25 GHz (black) and 40 GHz (red), for a 40 finger NMOS with a W/L of 80/0.18 biased at V_{ds} =1.3V and varying V_{gs}.

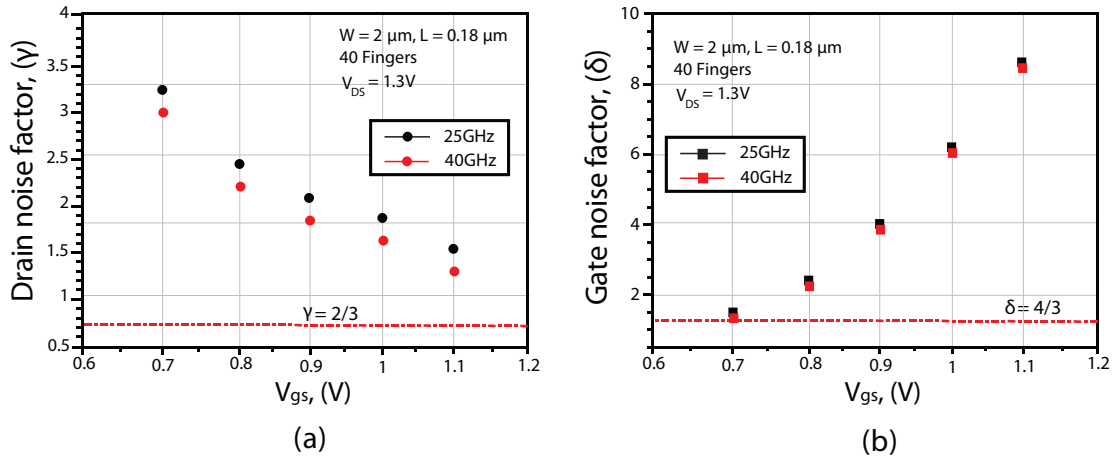


Figure 4.23: Drain and gate noise factors, γ and δ, bias dependency illustrated for a 40 finger NMOS with a W/L of 80/0.18 biased at V_{ds} =1.3V and varying V_{gs}

The increase in δ at high gate bias is due to the increase in effective channel resistance because of velocity saturation. However, it is observed that the contribution of the induced gate noise to the total device noise is relatively small for these highly scaled MOSFETs.

The extracted noise currents are fed into the equivalent circuit model to obtain the simulated high frequency noise parameters, NF_{min}, R_n, and Γ_{opt}. Figure 4.24(a) shows the measured and simulated NF_{min} and R_n as a function of frequency. It can be seen that the 0.18 μm NMOS transistor exhibits a NF_{min} of 1.5 dB at 25 GHz, which is very useful for K-band applications. Figure 4.24(b) depicts the optimum source reflection coefficient versus frequency.

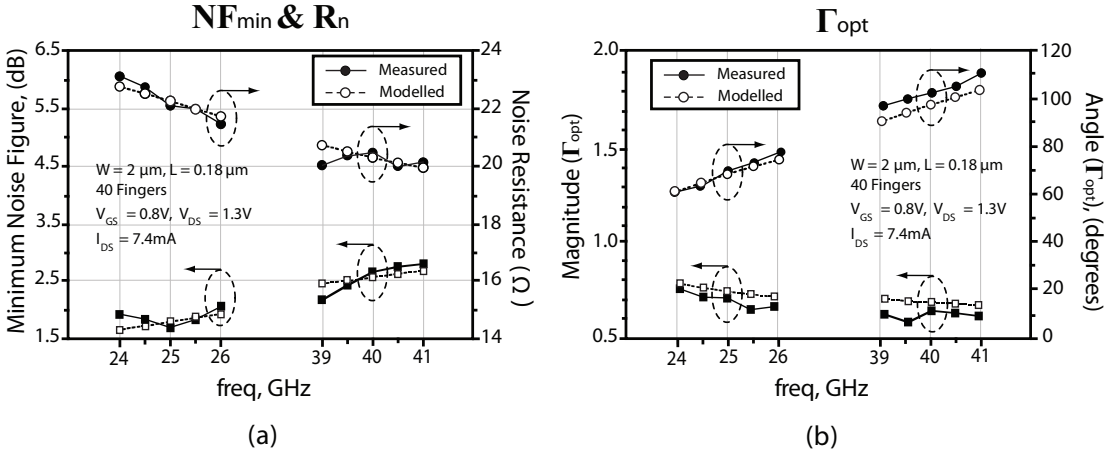


Figure 4.24: Measured and modelled noise parameters versus frequency: (a) Minimum noise figure (NF_{min}) and noise resistance (R_n) and (b) noise reflection coefficient Γ_{opt}.

Figure 4.25 compares the results of the combined model against the experimental measurements at different bias currents. The comparison is performed at two different frequencies of 25 GHz (microwave) and 40 GHz (millimetre-wave). These frequencies were chosen for validation purposes as the designed low noise amplifiers of this work operate at these frequencies. In general, there is a good agreement between the measured and modelled parameters of the device. The maximum deviation at both frequencies for all the noise parameters is less that 1%. This confirms the validity of the combined small-signal and noise circuit model.

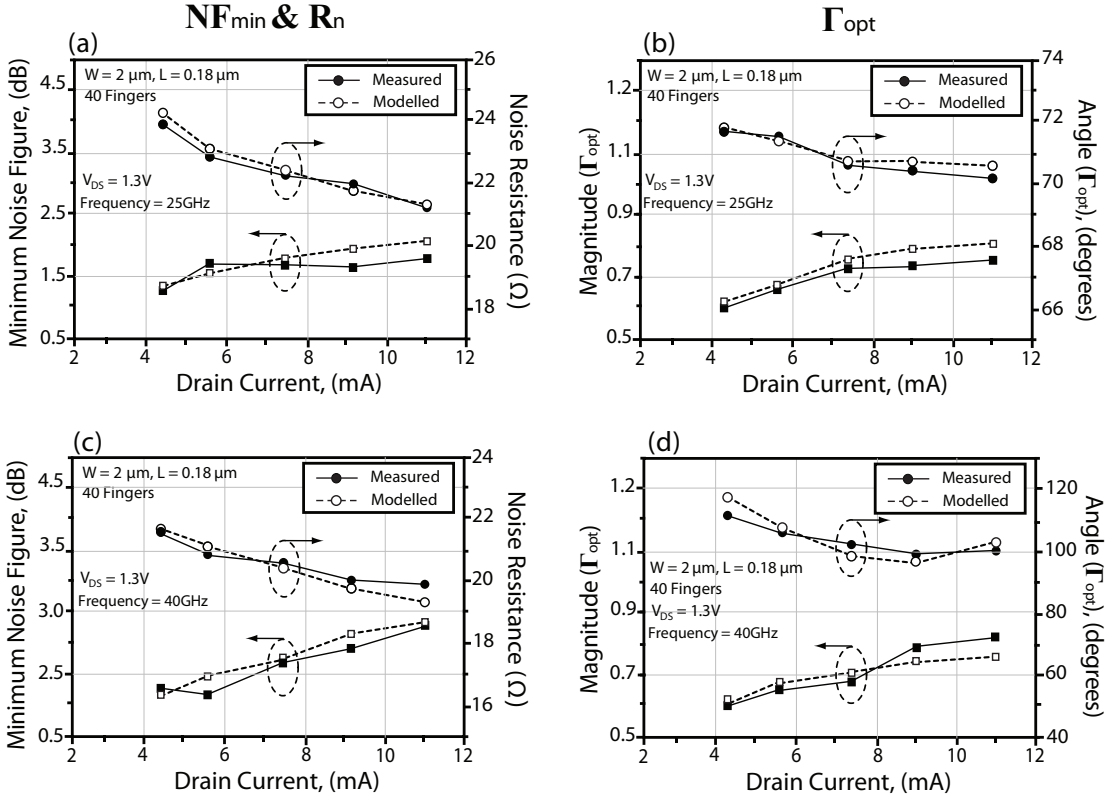


Figure 4.25: Measured and modelled noise parameters versus drain current at 25 GHz and 40 GHz: (a) Minimum noise figure (NF_{min}) and noise resistance (R_n) at 25 GHz, (b) noise reflection coefficient (Γ_{opt}) at 25 GHz, (c) minimum noise figure (NF_{min}) and noise resistance (R_n) at 40 GHz and (d) noise reflection coefficient (Γ_{opt}) at 40 GHz.

4.5 Summary

Designing reliable and robust CMOS low noise circuits hinges on the provision of accurate device models. In order to develop such models, a good understanding of the basics of CMOS MOSFETs and noise fundamentals are vital and therefore were introduced at the beginning of this chapter. Following on from this, the small-signal models reported in literature were reviewed and assessed to arrive at the most suited methodology for this work. A description of the method applied to extract the device noise parameters and thermal currents from the measured data followed. Each model was tested and verified on a multi-fingered n-MOSFET. The results show a good agreement between the measured and modelled responses with an error of under 1%. This gives confidence in adopting these modelling methodologies in the design of low noise circuits, an example of which will be illustrated in Chapter 6.

Chapter 5

Millimetre-wave Transistor Optimisation

5.1 Introduction

Traditionally, transistor design is in the domain of the process engineers. It is common practice for circuit designers to employ transistor design, without any alterations, in their circuits. This is applicable for circuits designed to operate at low frequencies. However at millimetre-wave frequencies device layout parasitic effects become very dominant and can significantly influence the performance of the device. Therefore millimetre-wave circuit designers are advised to alter the device layout with the aim to improve the device performance metrics to serve a specific application [121]. In this chapter CMOS transistor optimisation reported in literature are introduced, following this a design methodology for noise and power performance improvement is discussed. Finally a few examples of device optimisation including a novel "stacked" device are presented and compared.

5.2 Millimetre-wave Device Optimisation - An Overview

The device layout, geometry and wiring play a vital role in increasing or decreasing the device capacitive and resistive parasitics, leading to an enhanced or degraded performance. Methods to enhance the device performance have been discussed extensively in

literature. The impact of finger length on the device high frequency characteristics, as the device scales down, was presented by Morifuji and Woerlee [122, 123]. A TCAD simulation based MOSFET layout optimisation method was proposed by Tatsumi and Nakamura [124, 125]. Cao and Razavi [126, 127] highlighted the effects of the device parasitic capacitances and optimised the core transistor structure to reduce the effects. Chan and Jhon [128, 129] focused on the impact of wiring effect to reduce the device capacitive and resistive parasitics, leading to an improved f_{max} and f_t. Several gate layout structures using different gate connection, including waffle, meander and comb structures, mainly aimed at reducing the gate resistance and thereby improving f_{max}, were suggested by Wu and Kim [130, 131]. Similarly, Nicolson and Voinigescu [132] investigated and optimised the MOSFET gate layout structure and demonstrated a 77 GHz and a 99 GHz LNA using the optimised device. Another optimised layout design, referred to as a round table structure, was proposed by Heydari [133] to realise a power amplifier operating at 60 GHz.

Despite the impressive performance improvements, the vast majority of the aforementioned work have only focussed on the optimisation of f_{max} and f_t. The reports that discussed the device noise performance mainly referred to it as a by-product of the optimisation procedure and not as an aim. Designing low noise circuits, especially low noise amplifiers, requires equal attention to both noise and power optimisation. Therefore, in this study the focus was on reporting the layout and sizing effects on the device noise in addition to the power performance. A new modular layout approach is proposed with the aim to improve the RF performance of CMOS MOSFET and enable device operation beyond the cut-off region.

5.3 Device Optimisation Methodology

To ascertain the effect of sizing and layout on device performance a device parameter sensitivity analysis was performed, as shown in Figure 5.1. Normalised sensitivity calculation was used, where one component at a time was varied and the outcome presented as a percentage change in response versus a 1% change in the component value.

For this work the study was conducted on a 40 finger NMOS transistor with a W/L of 80/0.18, biased at I_{ds} = 7.4mA (V_{gs}=0.8V and V_{ds} =1.3V). A physical small-signal

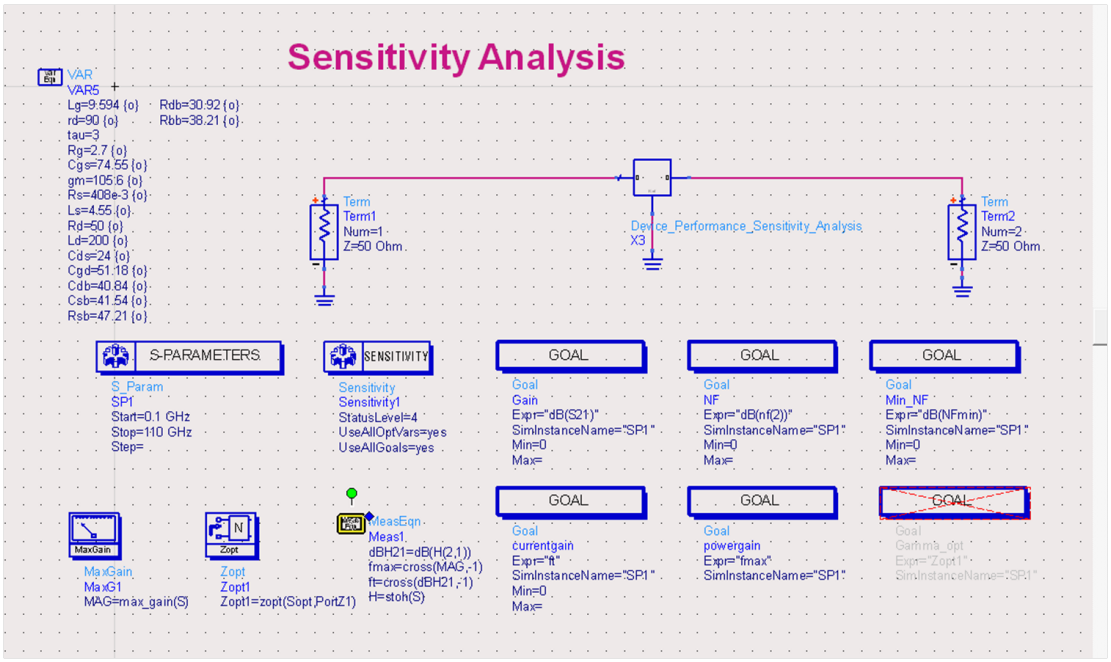


Figure 5.1: The device sensitivity analysis setup implemented in Agilent ADS.

model was extracted, as shown in Section 4.3.2, and a normalised parameter sensitivity analysis was performed. This used Agilent ADS to determine the effect of each parasitic on different device performance metrics.

Figure 5.2 and Figure 5.3 illustrate the results obtained from the performed sensitivity analysis. It can be seen from Figure 5.2 that the gate capacitances (C_{gd} and C_{gs}) have the most significant contributions in degrading the device power metrics. They show the largest effects on the cut-off frequency (f_t), the maximum oscillation frequency (f_{max}) and the power gain of the device. The second largest contributor is the gate resistance, followed by the drain resistance. These results agree with the reported literature and can be predicted from the device figure of merits mathematical estimation. The drain inductance and the substrate network also have a moderate influence on the aforementioned figures of merit and therefore cannot be neglected. These results suggest that reduction in R_g, R_d, C_{gd} and C_{gs} as well as the drain inductance (L_d) and the substrate network (R_{sub} and C_{sub}) could substantially improve the power performance of the device and increase the cut-off frequencies, and thus the operating region of the device.

It is evident from Figure 5.3 that the gate capacitances have the largest bearing on the device noise degradation. The gate source capacitance degrades the noise parameters, whereas the gate drain capacitance has the opposite effect. The terminal resistances

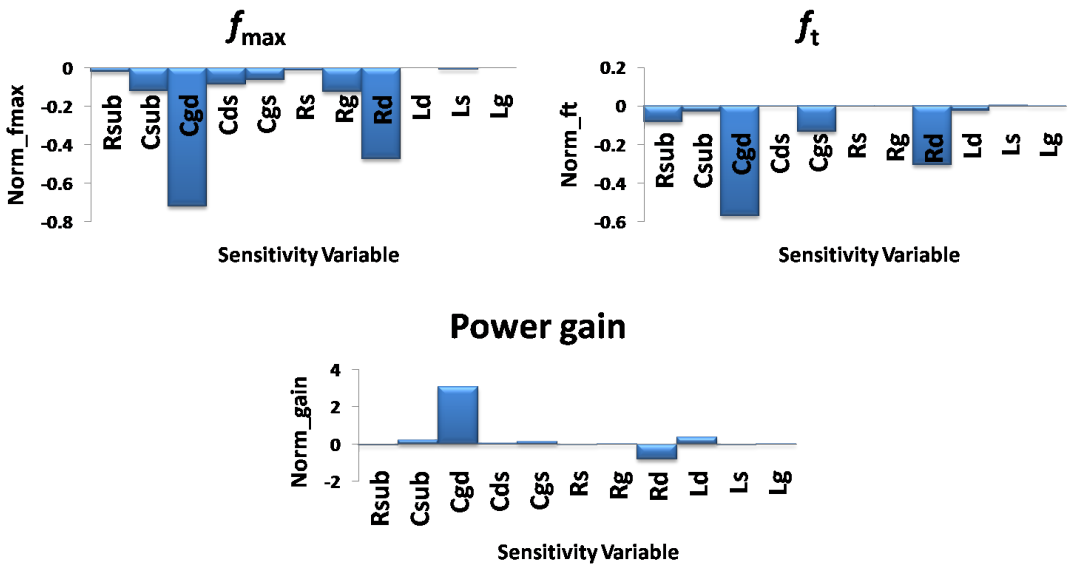


Figure 5.2: Power sensitivity analysis of the device to various parasitic components of the device model.

have a considerable influence on the device noise degradation. Variation in the gate resistance has a significant effect on the device noise parameters. The next most significant component is the source resistance. The drain and substrate resistances demonstrate much lower impact on the device noise parameters when compared to gate and source resistances. These findings agree with the mathematical prediction of the device noise figure of merit. However the mathematical estimation doesn't quantify the magnitude of the effect each of the components has on the noise parameters. Furthermore the source inductance, which demonstrates a significant sensitivity to the noise parameters, does not feature in the mathematical equation of the device noise parameters.

All the terminal resistances and gate capacitances are controllable and therefore, for low noise application, the device can be tailored, by reasonable device layout, to reduce

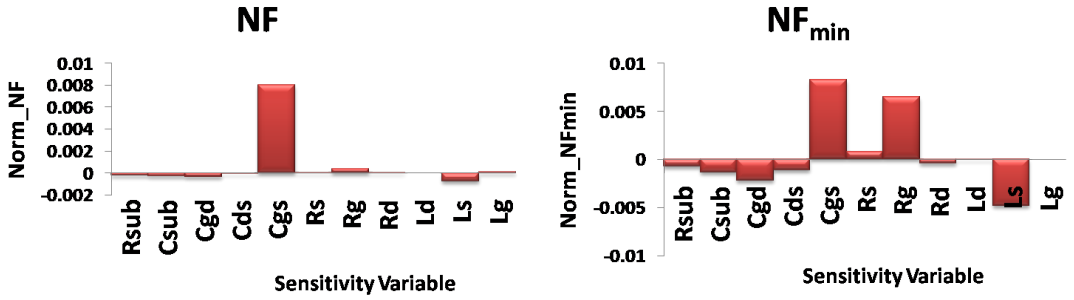


Figure 5.3: Noise sensitivity analysis of the device to various parasitic components of the device model.

C_{gs}, R_g, R_s and R_d and increase C_{gd} and L_s. These modifications would improve the device noise performance significantly. From the analysis above, it can be concluded that C_{gd}, R_g and R_d provide the largest impact on the device f_{max}. On the other hand, C_{gs}, C_{gd} and R_g have the largest effect on the device minimum and overall noise figure. Thus, in order to improve the device power and noise performance, the device layout should be altered in a such a way to give first priority to minimising these device components. Special attention should be paid to C_{gd} as it demonstrates opposite effects on the power and noise parameters; C_{gd} should be optimised to satisfy a good trade off between power and noise performance. A detailed analysis of parasitic capacitances and resistances, together with methods to minimise their effects, are described in the next section.

5.3.1 Parasitic Resistance Optimisation Methods

As detailed earlier, among all the parasitic resistances, the gate resistance demonstrates the largest impact on the device noise and power performance in the microwave and millimetre-wave frequency range. At these frequencies, a large gate resistance tends to increase the thermal noise and decrease the maximum available gain. The gate resistance depends on the gate material and the physical structure of the device. In a deep sub-micron salicided CMOS process, metal vapour is deposited on top of the gate polysilicon, this is known as salicide and reduces the gate sheet resistance to less than $10\,\Omega/\square$. At high frequencies each individual gate finger can be viewed as a distributed R-C transmission line, where the gate resistance can be approximated by [134, 135]

$$R_g = \frac{R_\square\, W}{3\, n^2\, L} \tag{5.1}$$

where $R_\square$ is the gate polysilicon sheet resistance, W is the gate finger width, L is the channel length and n is a constant that depends on the gate number of contacts (n=1 if gate is connected from one side and n=2 if both sides of gate are connected). Equation 5.1 suggests that as the sheet resistance is fixed and the channel length is governed by the technology, using multiple shorter gate fingers is one way to reduce the gate resistance. Similarly, connecting the gate fingers from both sides further decreases the gate resistance.

The source and drain resistances according to [84] can be presented as

$$R_d = R_{d0} + \frac{r_{dw}}{N_f W} \tag{5.2}$$

$$R_s = R_{s0} + \frac{r_{sw}}{N_f W} \tag{5.3}$$

where r_{sw} and r_{dw} are the drain and source resistances per unit width, while R_{d0} and R_{s0} represent the part of the series resistances that are independent of width. It is evident from Equation 5.2 and Equation 5.3 that increasing the number of fingers and connection vias reduce the value of R_d and R_s.

The impact of substrate resistance on the device performance is significant and cannot be neglected at microwave and millimetre-wave frequencies. In RF CMOS IC design, the non-epitaxial (non-EPI) process is preferred to the EPI process because of the higher substrate resistance. In a non-EPI CMOS process, a high substrate resistance ($6\,\Omega - \text{cm}$ to $10\,\Omega - \text{cm}$) helps to reduce the parasitic capacitance to the substrate and the unwanted signal coupling between the devices. Substrate resistive and capacitive coupling strongly depend on the device geometry or the distance separating the source and drain junctions from the bulk connection. It is always preferable to have the shortest path between the intrinsic device and the substrate connection with minimum substrate resistance. Thus a device is usually wrapped with a ring of substrate contact, often referred to as guard ring, with as many substrate contacts as possible.

5.3.2 Parasitic Capacitance Optimisation Methods

As demonstrated earlier, C_{gd} and C_{gs} have the highest impact on the power and noise parameters of the device. In order to interface the RF device with other components in the circuit such as spiral inductors, transmission lines and capacitors, the RF device requires multi-layer interconnects, to reach the top metal layer where these components are implemented. The sidewall overlap of the previously mentioned interconnects results in increased gate parasitic capacitances. Thus gate parasitic capacitances depend heavily on the device physical layout and can be controlled by changing the drain, source and gate interconnects during transistor design.

Figure 5.4 illustrates two fabricated devices and their cross section in three different orientations. Device 1 is a standard RF layout used in most RF transistors, whereas

device 2 is an optimised layout aimed at reducing the gate parasitic capacitances. Both devices feature a a multi finger, meandered type gate poly to reduce the gate resistance. It can be seen from the cross section of device 1 (across A-A') that the drain node is surrounded by the gate interconnect layer. This leads to a large unwanted gate drain capacitance. This issue has been circumvented in device 2 for which the gate interconnect ends at metal 2, which results in a much lower capacitive coupling between the gate and drain interconnects. This can be understood by comparing the cross section of devices along A-A' in Figure 5.4. The optimised layout, with reduced gate interconnect layers, also enabled a reduction of the gate source capacitance, leading to smaller direct coupling paths between the gate and source. This can be understood by comparing the cross section of devices along B-B' in Figure 5.4. A detailed discussion of the layout effects on the device f_{max}, f_t and NF_{min}, together with the measurement results and the extracted capacitances and resistances from the small-signal model, are provided later in this chapter.

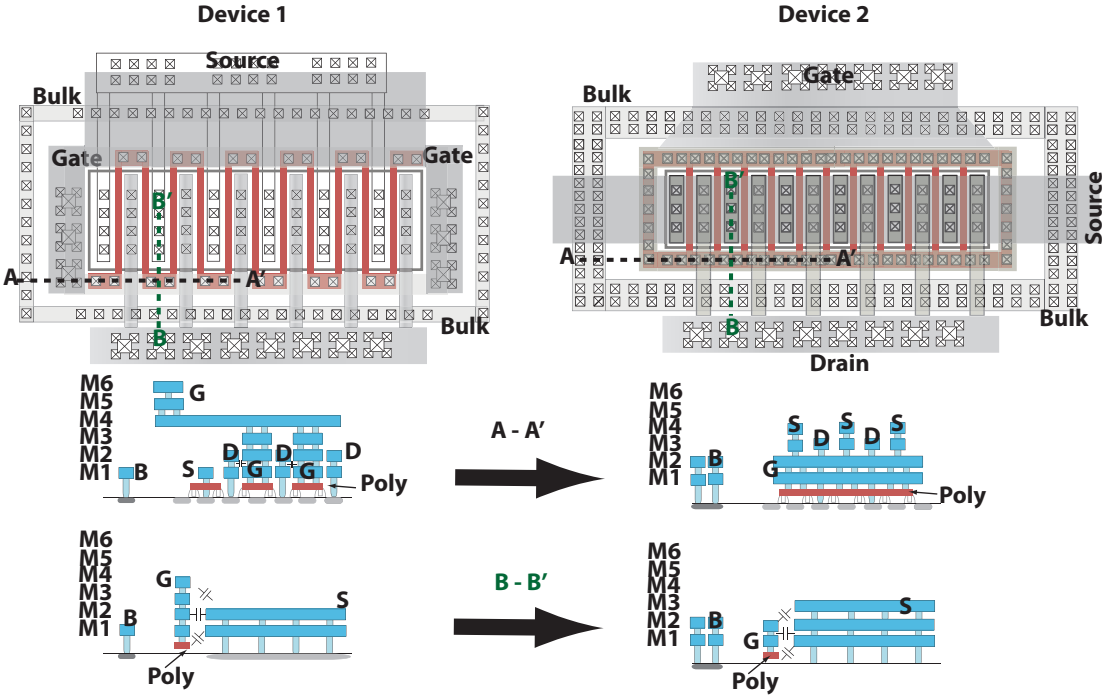


Figure 5.4: Conventional layout (device 1) and optimised layout (device 2). The step by step cross section of A-A' and B-B' demonstrate the reduction of parasitic capacitances (C_{gd} and C_{gs}) in the optimised device layout.

5.3.3 Device Sizing Optimisation

In a multi-finger CMOS transistor, varying the device gate width and the number of fingers results in an increase or a decrease in the gate parasitic resistances and capacitances. This directly affects the maximum stable gain, f_{max}, f_t and NF_{min}. To determine these effects, several NMOS transistors with gate finger widths (8, 6, 4 and 2 μm) and numbers (80, 40, 20, 10 and 5 fingers) were fabricated and measured.

5.3.3.1 Gate Width Optimisation

This section discusses the effects of device gate finger width on device noise and power performance. Figure 5.5 depicts the measured device figures of merit f_{max}, f_t and NF_{min} versus gate finger width. It is evident from Figure 5.5(a) that increasing the transistor finger width results in decreasing f_{max}. This is expected as the device gate and drain resistances decrease resulting in a higher device f_{max}. It is further noted that increasing the finger width causes a gradual decrease in f_{max} with the highest gradient occurring for values below 4 μm.

The transistor f_t is mainly affected by the variation in the gate capacitive parasitics. The gate resistances have negligible influence on f_t, whereas the gate capacitances have a significant influence. In fact f_t is completely dependent on the gate capacitances. Therefore, it is expected that increasing the finger width, which results in higher gate capacitances, leads to a decreased f_t. However, as shown in Figure 5.5(b), devices with different gate finger width exhibit very little variation in f_t. This is due to the fact that the gate parasitic capacitance is mostly dominated by the device multi layer interconnects and as all the devices in this study have identical layouts, the interconnect capacitances dominate the internal capacitances and mask their contribution.

The effects of gate finger width on the device minimum noise figure is illustrated in Figure 5.5(c) and (d). It can be observed that the device NF_{min} increases with the increase of the finger width, regardless of the frequency of operation. This is attributed to the increase in the parasitic resistance, which leads to increased thermal noise and therefore a degraded minimum noise figure. The lowest NF_{min} is achieved by using the lowest finger width. In view of this, the optimum gate finger width is chosen as 2 μm,

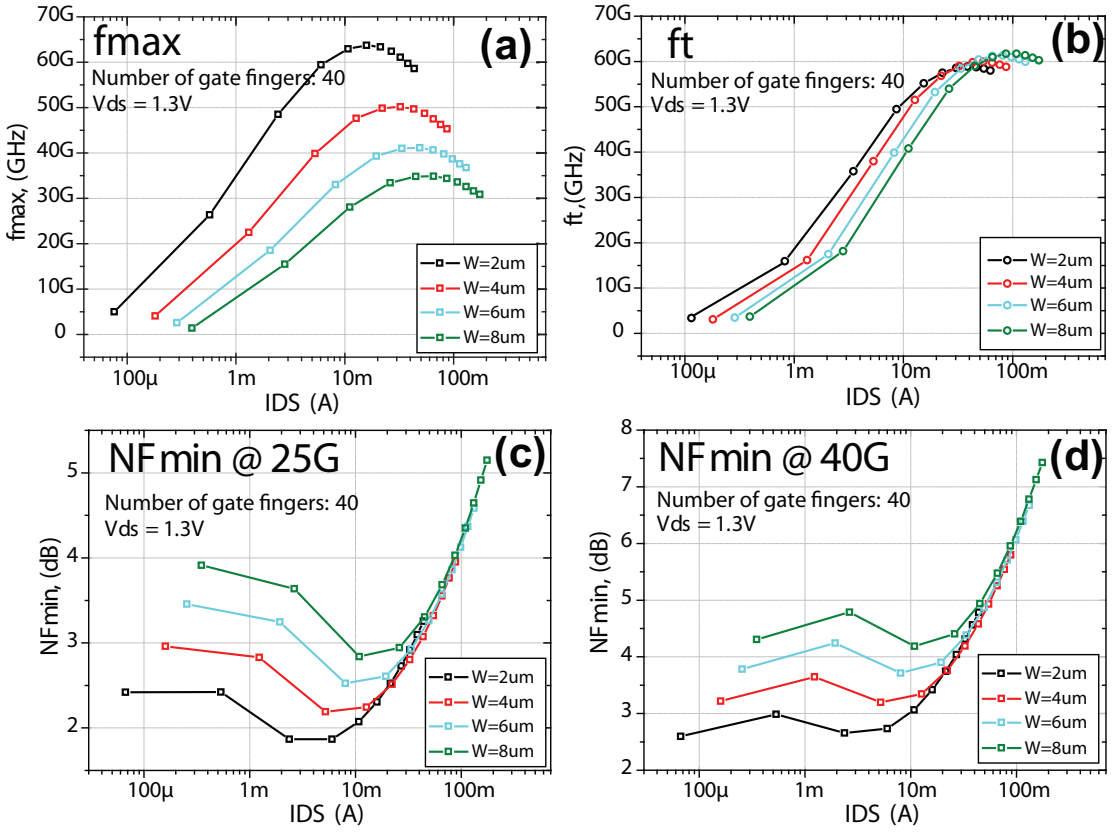


Figure 5.5: Transistor figures of merit versus gate finger width at multiple bias points: (a) f_{max}, (b) f_t, (c) NF_{min} at 25 GHz and (d) NF_{min} at 40 GHz.

as it provides the device with a high f_{max} and a low NF_{min}, which is ideal for low noise circuits and application.

5.3.3.2 Number of Fingers Optimisation

Once the optimum finger width is chosen, the number of fingers should be increased in order to increase the effective device size. Similar to the experiment conducted for the gate finger width, the device f_{max}, f_t and NF_{min} are compared versus the number of gate fingers for devices with a fixed gate width of 2 μm and over multiple bias points. These results are given in Figure 5.6. It is evident from Figure 5.6(a) that increasing the number of gate fingers decreases f_{max} and increases NF_{min}, as shown in Figure 5.6(c) and (d). This is due to the increase in the physical layout of the device, which leads to an increase in the gate, drain and source networks and therefore increase the gate and drain resistive losses leading to a degradation in the device figures of merit. The device f_t decreases as the number of fingers (size) increases, this is due to the increase in

the parasitic gate capacitances. This increase is caused by the increase in the effective coupling area between the device interconnects.

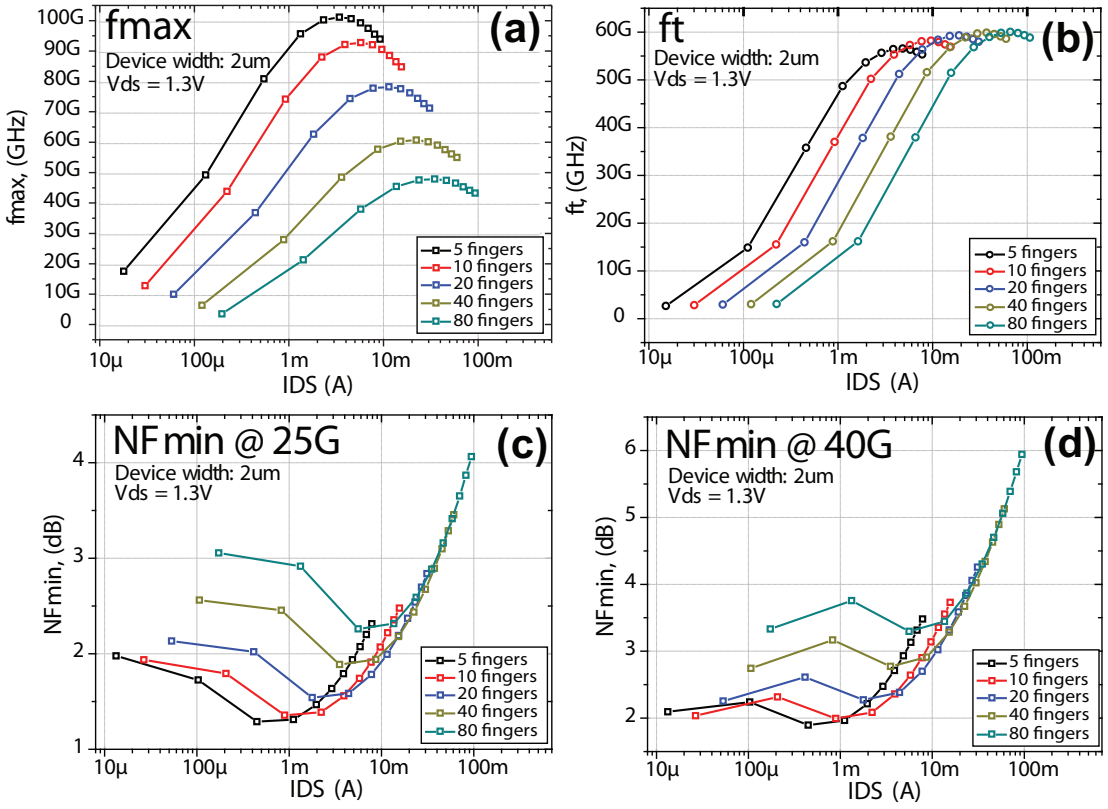


Figure 5.6: Transistor figures of merit versus number of gate fingers at multiple bias points: (a) f_{max}, (b) f_t, (c) NF_{min} at 25 GHz and (d) NF_{min} at 40 GHz.

5.4 Stacked Device Structure

Based on the optimisation analysis performed above, the device performance can be optimised by employing a multi finger device layout with modified gate, drain and source contacts, increased number of vias, and a wide guard ring with multiple rows of substrate contacts to reduce the substrate effects. These modifications would potentially minimise the parasitic resistances and boost the power and noise performance of the device. Based on the earlier findings, a new device structure can be proposed. The idea behind the optimised device is to reduce the parasitics that have the highest impact on f_{max} and NF_{min}. This is achieved by using a modular method in the device design. The proposed device consists of multiples of 10 μm wide devices, formed of 5×2 μm wide circular type poly gate fingers, with double contacts to reduce the poly gate resistance. The drain, source and circular gate metal are formed on the M2 and M3 metal layers. The core

device dimension were chosen based on the results discussed earlier in this chapter. The core devices are then connected, in a meandered fashion, to form the final device. The number of core devices needed depends on the desired overall device size. For example, if the designer requires a W=80 μm device, then the stacked structure will be formed of 8 core transistors connected in a meander fashion as shown in Figure 5.7. The core devices are connected using a global double-contact multi-path connections formed on M4, M5 and M6, between the gates, sources and drains of the core units. This configuration will enable connection with other RF devices such as active devices, spiral inductors, transmission lines or capacitors.

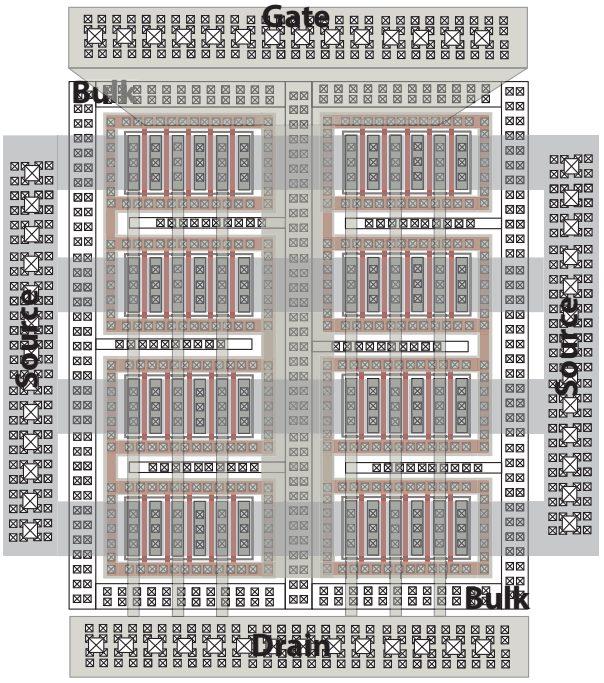


Figure 5.7: Layout of the proposed device structure; final device size is 80 μm. The core devices are formed of 5×2 μm devices.

5.5 Experimental Results and Discussions

In order to assess the proposed optimised device a comparative study was conducted. In this study, three NMOS transistors fabricated with a gate length of 0.18 μm and a total width of 80 μm, on a low resistivity silicon substrate using 0.18 μm digital logic CMOS process with one poly and six metal layers (1P6M). Figure 5.8 illustrates the devices included in this study, device 1 and device 2 have been discussed earlier in this

chapter. Device 1 is a common layout widely used in RF transistor design, a meandered poly-gate is employed to reduce the gate resistance. Device 2 is an optimised layout that is aimed to improve the parasitic resistances and capacitances. The gate resistance is reduced by using a double contact poly-gate, connected to a circular gate metal using multiple vias to reduce the gate resistance even further. Finally device 3 is the proposed stacked device, described in the previous section, aimed at improving the device noise and power performance.

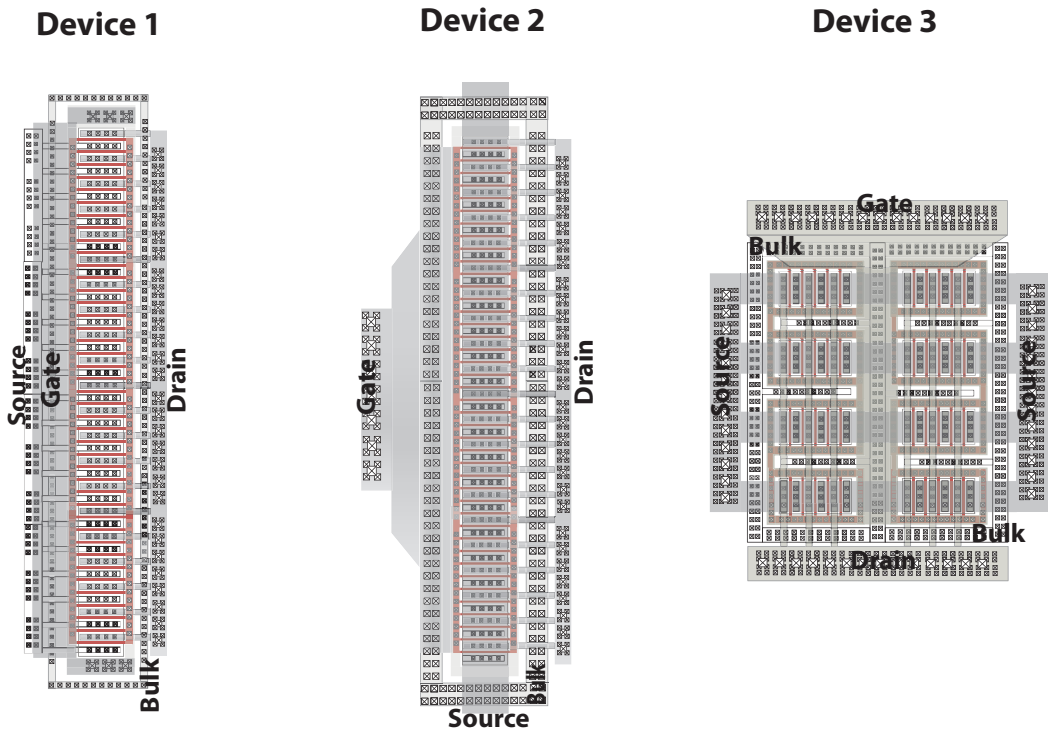


Figure 5.8: Layout of the proposed device structure; final device size is 80 μm. The core devices are formed of 5×2 μm devices.

All devices were biased at V_{gs}=0.8V and V_{ds}=1.3V, and measured up to 110 GHz. The probe pads were de-embedded from each device and the small-signal model parameters were extracted (averaged between 10 GHz and 65 GHz) for each device. The extracted values of C_{gs}, C_{gd}, R_g and R_d are listed in Table 5.1. It can be seen from the table that the total gate capacitance of device 3 is reduced by approximately 5 when compared to the other two devices. This is due to the reduced sidewall capacitive coupling from the top metal interconnects. The gate resistance is halved, from 5.1 Ω down to 2.3 Ω. The drain resistance has been reduced by 23% when compared to device 1. The drain current of device 3 is approximately 6% higher than the other two devices owing to the reduced drain resistance. Furthermore the extracted results illustrated in Figure 5.9 depict the

extracted maximum stable gain, which represents f_{max}, and the magnitude of h_{21}, which represents f_t, of devices 1, 2 and 3. The measurements suggest a significant performance improvement in the optimised device (device 3) compared to the other devices (devices 1 and 2). Device 3 features an f_{max} of approximately 50% increase when compared to the conventional device (device 1) and a 34% increase when compared to device 2. f_t remains almost constant in all devices. Device 3 demonstrates a 35% improvement in NF_{min}, in comparison to devices 1 and 2 as shown in Figure 5.10.

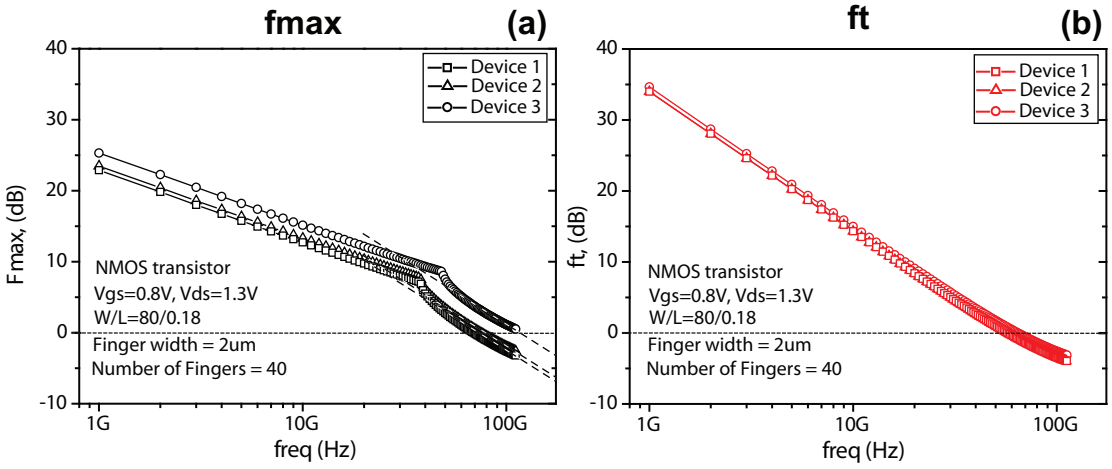


Figure 5.9: f_{max}, represented by the maximum stable gain, and f_t, represented by the magnitude of h_{21}, of the devices included in this study (device 1, 2 and 3).

Table 5.1: The comparative study results of 3 NMOS devices with different layout methodologies, including the extracted parasitic capacitances, gate and drain resistances, DC current, transcunductance and RF figures of merit (at V_{gs}=0.8V and V_{ds}=1.3V).

	Device 1	**Device 2**	**Device3**
C_{gs}, (fF)	95.3	102.3	84
C_{gd}, (fF)	30.7	32	30
C total, (fF)	126	124.4	114
I_{ds}, (mA)	7.4	7.4	7.8
g_m, (mS)	34.4	32	34.2
R_g, (Ω)	5.1	3.7	2.3
R_d, (Ω)	3.4	3.1	2.6
f_t, (GHz)	50	54	60
f_{max}, (GHz)	66	79	118
NF_{min}@ 25 GHz,(dB)	2.0	1.8	1.3
NF_{min}@ 40 GHz,(dB)	3.1	2.8	2.2

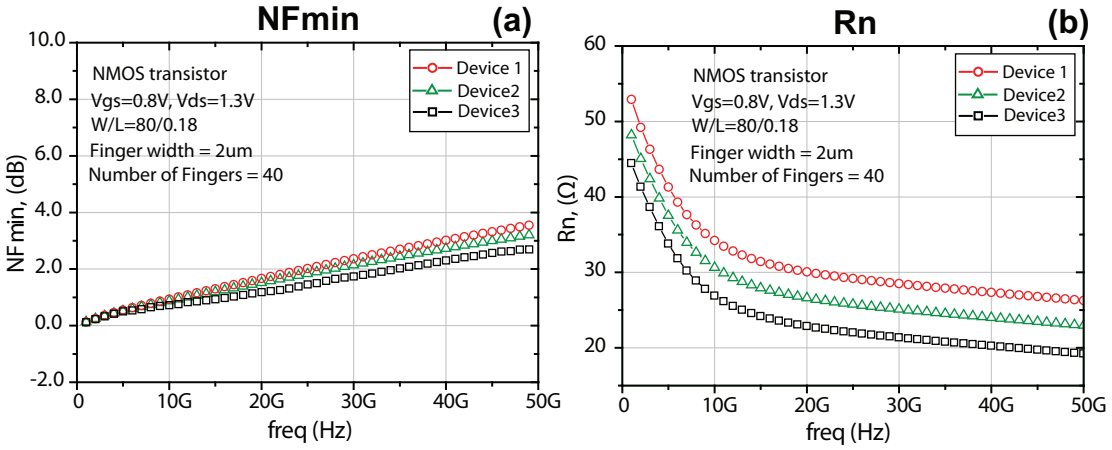


Figure 5.10: NF_{min} and R_n of the devices included in this study (device 1, 2 and 3).

This experiment demonstrates that device optimisation could potentially double the device power performance and reduce the minimum noise figure significantly. This enables the design of low noise amplifiers in the millimetre-wave frequency regime. This will be further demonstrated in the next chapter.

5.6 Summary

Transistors operating in the millimetre-wave regime are assessed based on various figures of merit such as f_t, f_{max} and NF_{min}. Most of these device parameters, especially f_{max} and NF_{min}, are strongly dependent on the device layout and the physical structure. The performance of the transistor can be optimised by a study of the sensitivity of the layout parasitics that can degrade those figures of merit and the layout is improved based on these findings. Taking into account all the vias and metal layers in the device under test, the optimised stacked layout specifically focuses on optimising the effects of the terminal parasitic capacitances and resistances. By changing the device size and the source, drain and gate interconnects, the capacitive and resistive parasitics can be optimised effectively, leading to optimum device that provides an improved low noise and high power capabilities, satisfying the trade off between noise and power performance. The stacked device designed as a result of this study is an engineered structure that shows an improved performance with an extrapolated f_{max} of 120 GHz and a minimum noise figure of 1.3 dB in the millimetre-wave regime.

Chapter 6

Low Noise Amplifier

6.1 LNA Fundamentals

A low noise amplifier is usually the first gain stage in any wireless receiver. The incoming wireless signal from the antenna, which is typically very weak, is fed to the input of a LNA. In order for the following mixer to process the incoming signal the LNA needs to amplify it. Thus, the LNA needs to have a certain power gain. The noise generated by the LNA is directly added to the signal in the amplifying process and reduces the signal-to-noise ratio (SNR). In contrast, the noise contribution from the following stages of the receiver is attenuated by the LNA gain. To satisfy the system noise requirement, the noise contribution from the LNA should not be large. The different metrics and topologies of the LNA are discussed in the following sections with the aim to reduce the LNA noise.

A LNA, as the name suggests, is an amplifier where particular emphasis has been put on its noise characteristics. In a system composed of a series of cascade stages, as shown in Figure 6.1, the total noise factor can be calculated using Friis's formula:

$$F_{tot} = F_1 + \frac{F_2 - 1}{G_1} + \frac{F_3 - 1}{G_1 G_2} + \frac{F_4 - 1}{G_1 G_2 G_3} + \cdots \tag{6.1}$$

It can been seen that the system's first stage (LNA) noise factor (F_1) and gain (G_1) dominate the overall noise factor (F_{tot}). Therefore, the low noise amplifier should be

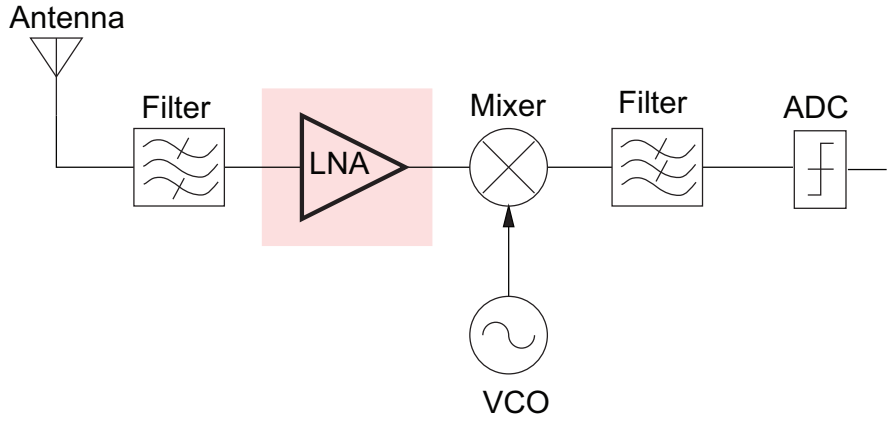


Figure 6.1: Typical architecture of a radio receiver.

designed to provide sufficiently high signal gain, to overcome the noise of the succeeding stages while maintaining a low noise figure.

The gain of a device is its ability to amplify the power or the amplitude of the input signal. It is defined as the ratio of the output to the input signal and is often referred to in terms of decibels

$$A_V = 20 \log\left(\frac{V_{out}}{V_{in}}\right) \tag{6.2}$$

Figure 6.2 shows a single-stage RF amplifier block diagram. Such an amplifier is normally characterised using S-parameters and terminated by an arbitrary source and load impedance Z_S and Z_L.

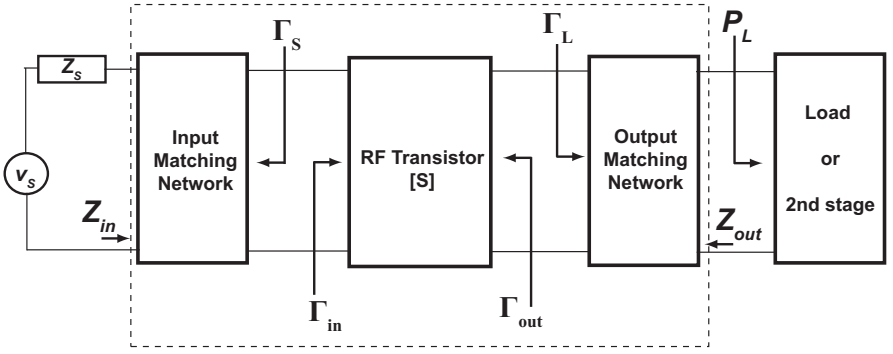


Figure 6.2: The block diagram of a typical RF amplifier including matching networks.

The input and output reflection coefficients, Γ_{IN} and Γ_{OUT}, can be calculated as follows:

$$\Gamma_{in} = S_{11} + \frac{S_{12}S_{21}\Gamma_L}{1 - S_{22}\Gamma_L} \tag{6.3}$$

$$\Gamma_{out} = S_{22} + \frac{S_{12}S_{21}\Gamma_S}{1 - S_{11}\Gamma_S} \tag{6.4}$$

$$\Gamma_S = \frac{Z_S - Z_0}{Z_S + Z_0} \tag{6.5}$$

$$\Gamma_L = \frac{Z_L - Z_0}{Z_L + Z_0} \tag{6.6}$$

$$\tag{6.7}$$

where Γ_S and Γ_L are the source and load reflection coefficients.

Power gain is generally defined as the ratio of the power delivered to the load and the power delivered by the source. However, as simple as it seems, this definition is not entirely relevant and is difficult to quantify as the source impedance is difficult to specify. For this reason a number of specific, and therefore more useful, definitions have evolved. Most notable are perhaps the transducer gain (G_T) which is the ratio of average power delivered to the load and the maximum available average power from the source:

$$G_T = \frac{1 - |\Gamma_s|^2}{|1 - \Gamma_{IN}\Gamma_S|^2}|S_{21}|^2\frac{1 - |\Gamma_L|^2}{|1 - S_{22}\Gamma_L|^2} \tag{6.8}$$

Power gain (G_P) is defined as the ratio of power delivered at the load and the power available from the source:

$$G_P = \frac{1}{1 - |\Gamma_{IN}|^2}|S_{21}|^2\frac{1 - |\Gamma_L|^2}{|1 - S_{22}\Gamma_L|^2} \tag{6.9}$$

Available power gain (G_A) is defined as the ratio of maximum available average power at the load and the maximum available average power from the source:

$$G_A = \frac{1 - |\Gamma_s|^2}{|1 - S_{11}\Gamma_S|^2}|S_{21}|^2\frac{1}{1 - |\Gamma_{OUT}|^2} \tag{6.10}$$

Maximum power is only obtained when an amplifier has complex conjugate terminations, i.e. $\Gamma_{in} = \Gamma_S^*$ and $\Gamma_{out} = \Gamma_L^*$.

The stability of an amplifier is a critical consideration. An amplifier could easily become a very poor oscillator if it was not designed carefully. Oscillation is potentially possible if either the input or the output have a negative resistance. This translates to having both Γ_{in} and Γ_{out} larger than one [136]. Considering a matched two-port network, as shown in Figure 6.2, it is clear that the stability of the circuit depends on the value of Z_S and Z_L as these values are added to the input and output impedance of the network. If a two port network is stable for only some values of the source and load impedances, its stability is conditional. Likewise, unconditional stability happens when the circuit is stable for all positive real values of Z_S and Z_L. In this case, the following conditions should be met:

$$|\Gamma_S| < 1 \qquad |\Gamma_L| < 1 \tag{6.11}$$

$$|\Gamma_{IN}| < \left| S_{11} + \frac{S_{21}S_{12}\Gamma_L}{1 - S_{22}\Gamma_L} \right| < 1 \tag{6.12}$$

$$|\Gamma_{OUT}| < \left| S_{22} + \frac{S_{21}S_{12}\Gamma_S}{1 - S_{11}\Gamma_S} \right| < 1 \tag{6.13}$$

As the input impedance of the network is a function of the load impedance, the input stability depends on the load. This determines the values of Z_L that make the input unconditionally stable.

On the Smith chart, the boundary between these two regions is a circle and is called the stability circle [136]. The magnitude of S_{11} determines whether the stability happens for load values inside or outside the stability circle. If $S_{11} > 1$, which is the case for most practical microwave CMOS amplifiers, the area outside the circle ensures stability as shown on Figure 6.3. The same argument holds for source stability circles. To make a network stable for any value of the source and load impedance, both these stability circles need to fall outside the unit circle on the Smith Chart, assuming positive $|S_{11}|$ and $|S_{22}|$. Solving Equation 6.11 gives a necessary and sufficient condition for stability:

$$k = \frac{1 - |S_{11}|^2 - |S_{22}|^2 + |\triangle|^2}{2|S_{21}S_{12}|} < 1 \tag{6.14}$$

$$|\triangle| = S_{11}S_{22} - S_{12}S_{21} < 1 \tag{6.15}$$

where k is called the stability factor and is a function of frequency. As the frequency increases, the gain of the device decreases, making the device more and more stable.

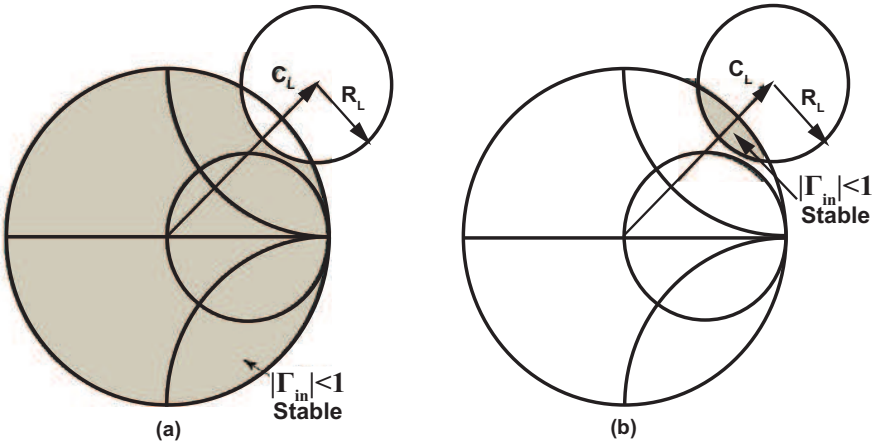


Figure 6.3: Stable region on the Smith Chart for (a) $|S_{11}| < 1$ and (b) $|S_{11}| > 1$.

One simple way to stabilise an active device is to add a series resistance or a shunt conductance to the unstable port. In practice, due to to the coupling between the input and output of the amplifier, it is usually sufficient to stabilise just one of the ports. However, the penalty is reduced power gain and a larger noise figure.

6.2 Single Ended Low Noise Amplifier Architectures and Design

Impedance matching is very important in low noise amplifier design. In most cases the source impedance of the LNA, following a typical antenna, is 50 Ω. As the input impedance of the transistor is almost purely capacitive, providing a good match to the input source can be very challenging. There are several popular LNA circuit architectures, implementable in a CMOS integrated circuit, distinguished from each other by the transistor topology and the way input impedance matching is achieved. These are shown in Figure 6.4.

6.2.1 Common Drain

A simplified common drain amplifier is shown in Figure 6.4(a). This configuration is generally used as a buffer (unity voltage gain) with low output impedance and high input impedance. The input signal is applied to the gate and the output is taken from the source. The output current (at the source) is very large compared to the input current,

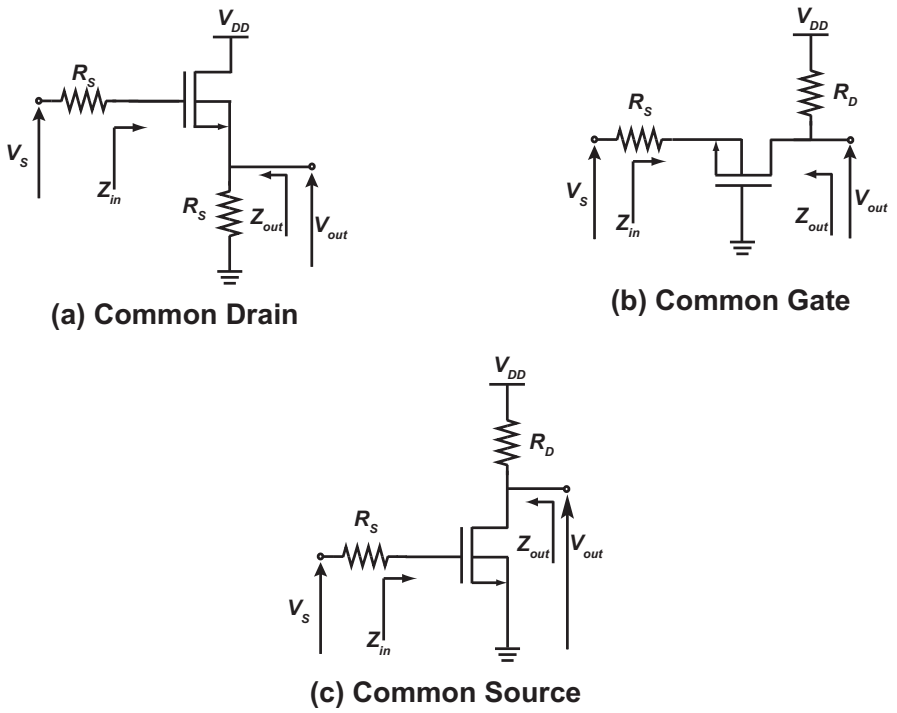


Figure 6.4: Amplifier topologies: (a) Common drain, (b) common gate and (c) common source.

therefore this amplifier can achieve high power gain, sufficient to reduce the significance of the noise figure of subsequent stages in the receiver system. However, the large input voltage swing, required to turn on the transistor, is a major drawback of this topology. As the incoming signal in any LNA is usually very weak, the requirement of a high voltage swing at the input is not an option. Therefore, the common-drain configuration is not suitable for the first stage of an LNA design.

6.2.2 Common Gate

A simplified common gate amplifier is shown in Figure 6.4(b). The input signal is applied to the source and the output is taken from the drain. Using this configuration the input impedance matching can cover a wide frequency band. Hence the name wide-band topology. The real part of the input impedance is just $1/g_m$. Therefore, $50\,\Omega$ impedance matching is easily obtained by carefully choosing the size of the transistor and the biasing condition. This however creates a drawback for this configuration, in that g_m is fixed once the source resistance R_s is known.

6.2.3 Common Source

A simplified common source configuration is shown in Figure 6.4(c). The input signal is applied to the gate and the output is taken from the drain. This configuration has a high voltage gain and high input impedance. Input impedance matching is established in several ways. The most straight forward approach to achieve $50\,\Omega$, as shown in Figure 6.5(a), by directly placing a $50\,\Omega$ resistor in parallel with the gate of the transistor. In [137] it was found that this matching method could double the noise figure. This is because the resistor adds its own thermal noise to the circuit and attenuates the incoming signal, hence this method is impractical for low noise applications. Another approach used to obtain good input matching is shown in Figure 6.5(b). Here a shunt feedback is used to set the input impedances of the LNA. Unlike the resistive termination case, it does not attenuate the input signal before reaching the gate of the transistor, so the noise figure can be improved. However, the feedback resistor continues to generate thermal noise of its own and could contribute significantly to the overall noise figure. This approach was implemented in [138, 139]. The comparison in [137] showed that amplifiers using this technique are usually broadband amplifiers and their power consumption is

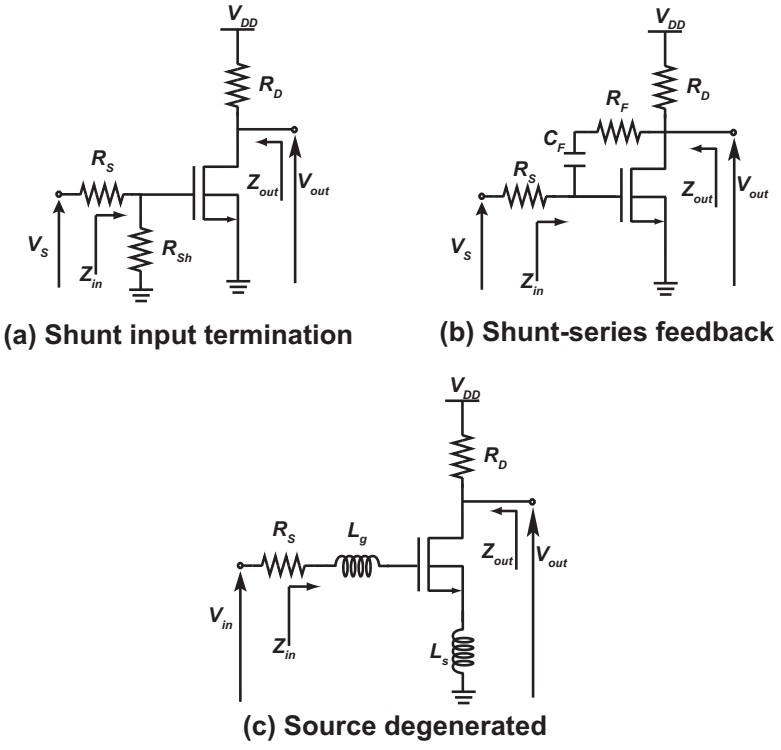


Figure 6.5: Common source amplifier input matching topologies: (a) Resistive termination, (b) shunt and series feedback and (c) inductive degenerated.

much higher than other amplifiers with similar performance. Figure 6.5(c) is a probably the most popular narrow-band[1] CMOS LNA. It uses inductive source degeneration to achieve input impedance matching (for example see [140]). The input impedance for this configuration (Z_{in}) is:

$$Z_{in} = j\,\omega\,(L_s\,+\,L_g) + \frac{1}{j\,\omega\,C_{gs}}\,+\,R_g\,+\frac{g_m}{C_{gs}}L_s \tag{6.16}$$

where C_{gs} is the gate to source capacitance of the transistor, L_g and L_s are the gate and source inductance respectively, and R_g is the gate resistance. Around the operation frequency $\omega_0 = 1/\sqrt{C_{gs}(L_g + L_s)}$, the input impedance only presents a real part $Z_{in} = (g_m/C_{gs})L_s$. Detailed analysis can be found in [141]. The noise figure of this LNA can be around 2 dB or even lower at lower frequencies [137].

6.2.4 Inductive Degenerative Common Source Versus Common Gate

The main difference between a common gate and common source LNA is the input impedance Z_{in}. Z_{in} of a common gate has a real part $R_{in} = 1/g_m$, whereas Z_{in} of a common source is mostly imaginary. Ideally the LNA should achieve a low noise figure, and also obtain a good input matching to the source impedance to avoid any loss of incoming signal power. These requirements lead to a major difference in common source and common gate LNA design and are discussed in this section.

The work presented in [142] shows that the lowest noise figure for any given MOSFET can be obtained by presenting an optimum signal source admittance $Y_{opt} = G_{opt} + jB_{opt}$. This work also shows that the optimum susceptance B_{opt} corresponds to a negative capacitor, which is commonly approximated by an inductor at a single frequency. This narrow band approximation limits the bandwidth for which the LNA noise figure can approach the minimum noise figure. As the signal source admittance $Y_s = G_s + jB_s$ deviates from Y_{opt}, the noise figure degrades according to

$$F\,=\,F_{min}\,+\,\frac{R_n}{G_s}[(G_s - G_{opt})^2\,+\,(B_s - B_{opt})^2] \tag{6.17}$$

[1] It is narrow-band because the impedance matching is only established within a very narrow frequency range due to the resonant nature of the reactive matching network.

A LNA designed according to this technique will achieve the lowest noise figure, F_{min}. However the input impedance match is neglected and the output power is compromised. To address this issue a real part in the input impedance has to be generated, this is achieved by employing an inductive source degeneration LNA. Several studies [137, 143–147] have shown that this topology can achieve good performance at low GHz bands. The feedback effect of the source inductor L_s results in a real input impedance component $\frac{g_m L_s}{C_{gs}}$. This resistance is noiseless and very convenient when an input power match is required. In addition to altering the transistor input impedance, the source inductor affects the optimum signal source admittance, Y_{opt}, that minimises the noise figure [148].

A simple design procedure to achieve noise and power matching simultaneously, was proposed in [149]. This design flow is summarised as follows:

1. First the drain current density associated with the lowest F_{min} is determined.
2. The transistor is scaled by varying the number of gate fingers while keeping constant finger unit width and current density until R_o reaches R_s.
3. The source degeneration inductance L_s is added to bring the real part of the amplifiers input impedance Z_{in} to R_s.
4. The gate inductance L_g is added to neutralise the imaginary part of Z_{in} at the operating frequency.

By this means, Z_{in} and R_o of the amplifier are matched to R_s. However, $F = F_{min}$ is only obtained here when X_o is zero, which is not true due to the correlation between the drain and gate noise current, i_{nd} and i_{ng}. Although common source degeneration techniques can bring Z_o closer to Z_s while maintaining a low F_{min}, power and noise matching cannot be perfectly achieved at the same time. A tradeoff between the noise figure and the signal power transfer is inevitable in common source LNA design. For an inductively degenerated common source, assuming i_{nd} is the dominant internal noise source, the noise factor of the stage can be approximated as follows [150]:

$$F \approx 1 + \gamma\, g_{d0} R_s \Big(\frac{\omega_0}{\omega_t}\Big)^2 \tag{6.18}$$

It is evident from Equation 6.18 that the noise figure is dependant on ω_0/ω_t. This indicates that an inductively degenerated common source LNA is well suited for applications

operating at frequencies well below the cut-off frequency of the device. However, the performance of this topology degrades substantially when ω_0 becomes comparable to ω_t. In contrast for common gate LNA the input impedance, given by $1/g_m$, is used to match the input to R_s. The source inductance L_s is used to resonate out the capacitance seen at the source at the operating frequency. Obviously g_m is fixed to $1/R_s$ for the purpose of power matching. Scaling the transistor size with constant g_m will either result in a low w_t and thus a high F_{min}, or a Z_0 far away from R_0. Therefore, compared to a common source with an inductive degeneration technique, the common gate stage lacks the flexibility of adjusting transistor size to bring Z_0 closer to R_0 while maintaining a low F_{min}. In other words, the goals of power matching and a low noise figure strongly conflict with each other in a common gate stage. At perfect power matching, and taking i_{nd} into account only, the noise factor for this topology is represented by

$$F \geq 1 + \frac{\gamma}{\alpha} \tag{6.19}$$

where $a = g_m/g_{d0}$. In short channel devices, due to short channel effects $a \leq 1$ and excess thermal noise from hot electrons $\gamma \geq 2/3$ (for a long channel device $a = 1$ and $\gamma = 2/3$). This yields about 2.2 dB and 4.8 dB for a long and short channel device respectively, this will most likely be higher in practice. It can be seen from Equation 6.19, that the noise figure is independent of frequency, indicating the performance of the common gate stage degrades more gradually with the increase of the operating frequency. However the achievable noise figure at the power matching condition is far above F_{min}, this disqualifies the common gate as an optimal design.

Table 6.1 illustrates a comparison between common gate and inductively degenerated common source amplifier topologies indicating the input impedance, noise figure and gain. It can be seen from this table that the gain of a single stage common source amplifier is $A_v = -g_m Z_{load}$, which is the same as for a common gate stage but with the 180° phase shift.

Table 6.1: Noise figure, gain and input impedance comparison of a common gate and inductively degenerated common source CMOS LNA architectures.

Topology	**Input Impedance (Z_{in})**	**Noise factor F**	**Gain**
Common gate	$1/g_m$	$1+\frac{\gamma}{\alpha}$	$g_m Z_L$
Inductive degeneration	$j\omega(L_g+L_s)+\frac{1}{j\omega C_{gs}}+\frac{g_m L_s}{C_{gs}}$	$1+\gamma\, g_{d0} R_s(\frac{\omega_0}{\omega_t})^2$	$-g_m Z_L$

6.2.5 The Common Source Common Gate Cascode

To exploit the advantages of both the common source and common gate topologies, a cascode design with a common source first stage and a common gate second stage shown in Figure 6.6 is studied. This topology has been, and still is by far, the most common LNA topology used [151–153]. In a conventional common source stage, C_{gd} provides a feedback path between the input and output which degrades reverse isolation and therefore affects the amplifier stability and reduces the bandwidth. In a common gate amplifier, this problem does not exist as C_{gd} is connected to ground. In a cascode configuration, C_{gd} of transistor Q1 is no longer a feedback path as it is not located between

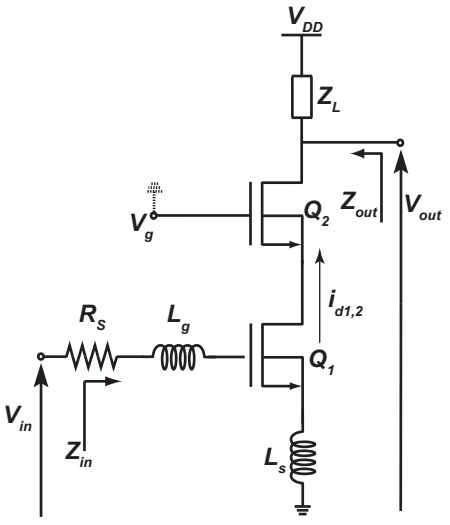


Figure 6.6: The common source common gate cascode LNA topology (with inductive degeneration).

the amplifier input and output anymore. Therefore reverse isolation and stability can be improved[2].

Simple mathematical analysis of the circuit in Figure 6.6 reveals that the gain of the amplifier is the gain of the common source stage only. The common gate increases the output impedance by $g_{m2}.r_{o2}$ where g_{m2} and r_{o2} are the transconducatnce and the drain-source resistance of Q2. The advantage of this topology is that it allows independent optimisation of noise by optimising the common source and common gate transistors respectively. This topology has a few disadvantages, one of which is that the transconductance of the cascode amplifier is limited by the transconductance of the common source stage only [154]. This can be solved by optimising the common source transistor noise and power performance. Another disadvantage is that at high frequencies the parasitics at the intermediate node between the two cascode transistors can significantly degrade noise and gain [155]. This issue can be rectified by adding a parallel inductor to the source of transistor Q2, forming a parallel resonator with the parasitic capacitance at the intermediate node [156].

6.3 State of the Art Microwave and Millimetre-wave LNA

This section presents the performance results of some recently published CMOS narrow-band LNA designs operating in the microwave and millimetre-wave frequencies. The common source, common gate and cascode architectures are hereafter called the CS, CG, and CS.CG respectively. The performance results are listed in Table 6.2. The top section of the table presents LNAs designed for microwave frequencies, whereas the lower section presents those designed for millimetre-wave frequencies.

It is evident from the bottom section of Table 6.2 that CMOS millimetre-wave LNAs have been developed mostly in 90nm technology. It is only at this technology node that the noise and power capabilities of the active device become competitive, without any modifications, which enables the design of building blocks operating in the millimetre-wave frequency bands.

[2]Eliminating the large Miller capacitance that appears at the gate of the common source transistor due to the parasitic capacitance between the gate and drain, C_{gd}.

Table 6.2: State-of-the-art CMOS microwave and millimetre-wave LNAs.

Ref	Technology (nm)	Topology	Frequency (GHz)	Gain (dB)	NF (dB)	P_{1dB} (dBm)	Supply (V)	P_{dc} (mW)	Size (mm^2)
[157]	180	CS + CS	24	13.1	3.9	–	1	14	0.34
[158]	180	CS + CS	24	12.8	3.3	–	1	8	1.9
[159]	180	CS	24.5	12	6.4	–	0.8	8	0.26
[160]	180	CS	24	14	4.9	-20	1	7.9	0.23
[161]	180	CS + CS.CG	24	9.3	5	-14	1.8	27	0.39
[162]	180	CG + CS + CS	21.8	15	6	-23	1.5	24	0.05
[163]	180	CS + CS + CS	23.7	12.8	5.6	-11.1	1.8	54	0.73
[164]	180	CS.CG	24	11.5	5.7	-20	1.4	3	0.31
[165]	130	CS.CG	26.2	8.4	4.8	–	1	0.8	0.16
[166]	90	CS + CS.CG	28.5	20	2.9	-17	1	16.2	0.67
[167]	90	CS	24	7.5	3.2	–	1	10.6	–
[168]	90	CS.CG + CS.CG	31-34	18.6	3	–	1.2	10	0.856
[169]	90	CS	20	8.4	3.1	–	1.4	14	0.56
[170]	90	CS + CS	20	8	5.3	–	0.66	11	0.84
[171]	45	CS.CG	23	7.1	4	-9.5	1	3.6	0.81
[172]	180	3 x CS.CG	40	7	–	5	3	300	2.04
[173]	130	CS + CS + CS	43	20	6.3	4	1.5	36	0.525
[174]	130	3 x CS.CG	40	19	–	-0.9	1.5	36	1.43
[175]	130	3 x CS.CG	55	24.7	8	2	2.4	54	0.48
[176]	90 SOI	CS.CG	35	11.9	3.6	4	2.4	40.8	0.18
[176]	90 SOI	CS.CG	40	9.5	4	4	2.4	40.8	0.18
[168]	90	CS.CG + CS.CG	31-34	18.6	3	N/A	1.2	10	0.856
[177]	90	CS + CS + CS	58	14.6	<5.5	-0.5	1.5	24	0.56
[9]	90	CS + CS + CS	60	12.2	6.5	4	1	10	0.48
[178]	90	CS	30	7.3	–	-5.75	1.5	10.5	–

6.4 Millimetre-wave LNA Design Methodology

Traditionally LNA design used a combination of passive elements to transform the signal source impedance to the optimum noise impedance of a transistor, biased at optimum NF_{min} current density [136]. While this approach improves the noise, the gain is compromised due to input reflection coefficient mismatch. The conventional method for simultaneous noise and input impedance matching was first presented in [179] and is widely used. Several design approaches for high performance and low power LNAs

followed [154, 180–182]. Unfortunately these methods were tailored for low frequency applications, hence these methods suffer more disadvantages and limitations as the frequency of operation increases [183]. For example, the pad capacitance which appears in parallel with the signal source impedance becomes increasingly significant at higher frequencies and cannot be neglected. Therefore designers have to include the pads in the input matching network. Furthermore the number of stages in the amplifier have to be increased due to the limited device gain at millimetre-wave frequencies. This results in a multi-stage amplifier of large devices, which means higher power consumption and noise performance degradation. In this section a simple design methodology for microwave and millimetre-wave LNA, that avoids the limitations caused by the conventional methods mentioned above shall be presented. The design flow is similar to the conventional inductively degenerated common source LNA design method described in Section 6.2.3. It shares the same procedure to design the input network, the transistor M1, and the LNA load network. The only difference lies in the cascode stage with cross coupling capacitor C_c and the added inductor L_{add}. The main target of this design is to verify the proposed technique in the cascode stage and the comparison between the conventional CS-LNA and the proposed CS-LNA is focused in the cascode stage. For this purpose, the conventional CS-LNA and the proposed CS-LNA share the same input stage. The performance of the input stage does not limit the validation of the proposed technique.

The design flow of the millimetre-wave low noise amplifier is outlined in the following steps:

1. Determine the device current density that will provide the lowest minimum noise figure. This can be achieved using S-parameters simulation and the noise characteristics of the active devices.

2. The minimum, maximum and variation of the device gain and noise parameters are analysed using different bias points. The optimum bias point is obtained leading to a low noise figure and high gain, while conserving the power.

3. Check that the circuit meets the requisite conditions for unconditional stability (K$>$ 1 and $|\triangle| < 1$) across the entire desired bandwidth..

4. Select the size of the source degeneration inductor (L_s) such that the stability conditions are met at the desired frequency. The use of inductive degeneration in

this method will also serve its original purpose in increasing the real part of the input impedance. Thus, as this inductor increases, the impedance will get closer to 50 Ω.

5. The constant noise figure circles are drawn on the Smith chart with the constant gain circles. Since the maximum gain and minimum noise figure can not be satisfied simultaneously, the trade-off must be made with respect to the specification.

6. Design the matching network of input port, interstage and output port.

The technique above has several advantages over other matching techniques. It is simple and requires only one additional matching component in series with the input of the transistor. It produces a relatively broadband match and it achieves simultaneous noise and power matching of the transistor. This makes it a preferred method for matching. There are several instances where this method cannot be applied without modification. For some designs, power constraints may not accommodate the necessary current to achieve the best noise performance. In this case, the design could proceed as before, except with a less than optimum current density. In other cases, the design may not provide the necessary gain required for some applications. In such case more current may be needed or some other matching technique that does not require degeneration may have to be employed. Also linearity constraints may demand a larger amount of degeneration than this method would produce.

6.5 Experimental Results

In this work low noise microwave and millimetre-wave amplifiers were designed and fabricated using the optimised stacked common source device and conventional RF device, discussed in the previous chapter. The low noise performance and the high available gain of the stacked device provides the means to reduce the number of stages of the amplifier, and therefore reduce the power consumption of the LNA. Furthermore, in order to simplify the input and output matching networks, a high pass combination of shielded and unshielded inductors, transmission lines and capacitor are adopted. In addition to matching, the inductors and capacitors also serve as an RF choke and a DC block respectively. In all the designs, the input and output access CPW pads were used as part

of the matching networks. This saves the de-embedding after measurements and most importantly accounts for the pad capacitance, which could alter the matching network response as the frequency of operation increases. All the amplifiers were fabricated in a standard digital 0.18 μm CMOS process provided by Texas Instruments. Slotted metal sheets cover a large area of the chip surface, this is a compulsory design rule that ensures minimum and maximum density of metals per any unit area. Measurements were taken directly using the on-wafer setup as illustrated in Chapter 2. Noise measurements were performed using the gain method (as described in Chapter 2), where the input of the amplifier is terminated with 50 Ω and the output noise power is measured using a spectrum analyser.

6.5.1 25GHz LNA: Shielded Versus Unshielded Matching Networks

The aim of this study is to illustrate the effects of shielding the matching networks on the power performance of the LNA. For this purpose two cascade source degenerated LNAs were designed at 25 GHz, the core device is a 40 fingers common source optimised stacked device with a W/L of 80/0.18. The chip photographs of the fabricated LNAs are shown in Figure 6.7 (a) and (b).

The amplifiers were powered from a 1.3V power source. The gate voltage is 0.8V, the drain voltage is 1.3V. Each stage draws a current of 7.3mA. The total current is divided equally between both stages. The measured S-parameters of both LNAs are shown in Figure 6.8. The shielded LNA (Figure 6.8(a)) exhibits a power gain of 13.64 dB at 25.4 GHz, a good reverse isolation, S_{12}, of about -29 dB and a relatively wide bandwidth of approximately 7 GHz (from approximately 23 GHz to 30 GHz). In comparison, the unshielded LNA (Figure 6.8(b)) exhibits a peak power gain of 6.48 dB at 25.4 GHz, a reverse isolation, S_{12}, of about -25 dB and a much narrower bandwidth of approximately 4 GHz (from approximately 24 GHz to 28 GHz).

It is evident from Figure 6.8 that the input and output return losses using the shielded LNA, are improved significantly when compared to the unshielded LNA, this was as expected. S_{11} improves from -8 dB to -13 dB and S_{22} improves from -4.8 dB to -25.4 dB for the unshielded and shielded LNAs respectively. Even though the amplifiers are not perfectly matched at the desired frequency, this is mainly due to the inaccuracies in the

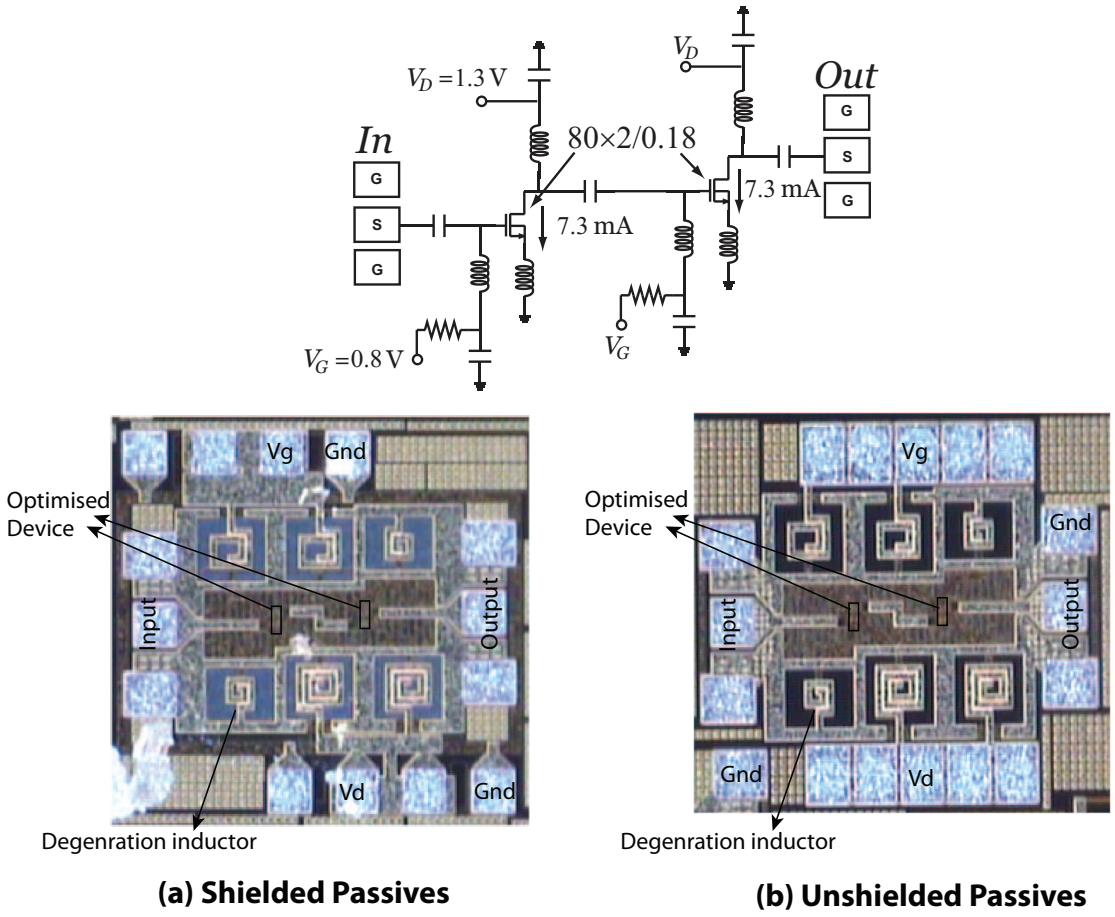


Figure 6.7: Chip photograph of two 25 GHz cascade source degenerated LNAs implemented in 0.18 μm CMOS process: (a) shielded and (b) unshielded inductors in the matching networks.

device model at the time of the design, the power gain performance has improved by a factor 2 when compared at the peak frequency.

The noise figure of the amplifiers is measured at a few points across the bandwidth. Figure 6.9(a) depicts the measured gain and noise figure for the shielded LNA. It can be seen that the noise figure is 3.38 dB at 25.4 GHz, which is the frequency where the peak power gain is achieved. Furthermore, the NF is lower than 4.8 dB over the entire amplifier bandwidth. Similarly Figure 6.9(b) depicts the measured gain and noise figure for the unshielded LNA. The noise figure is 3.43 dB at 25.4 GHz and the NF over the entire bandwidth is lower than 4.5 dB. The noise figure is comparable in both LNAs and this is expected as the noise contribution is mainly governed by the device noise. As the substrate noise consists of the substrate resistance and the coupling capacitance, which forms a low pass filter, this renders the substrate noise of no importance at high frequencies.

The shielded LNA in this work presents comparable, and in some metrics better performance when compared to the reported literature. This is shown in Table 6.3, this summarises the performance of recently reported amplifiers operating around 25 GHz.

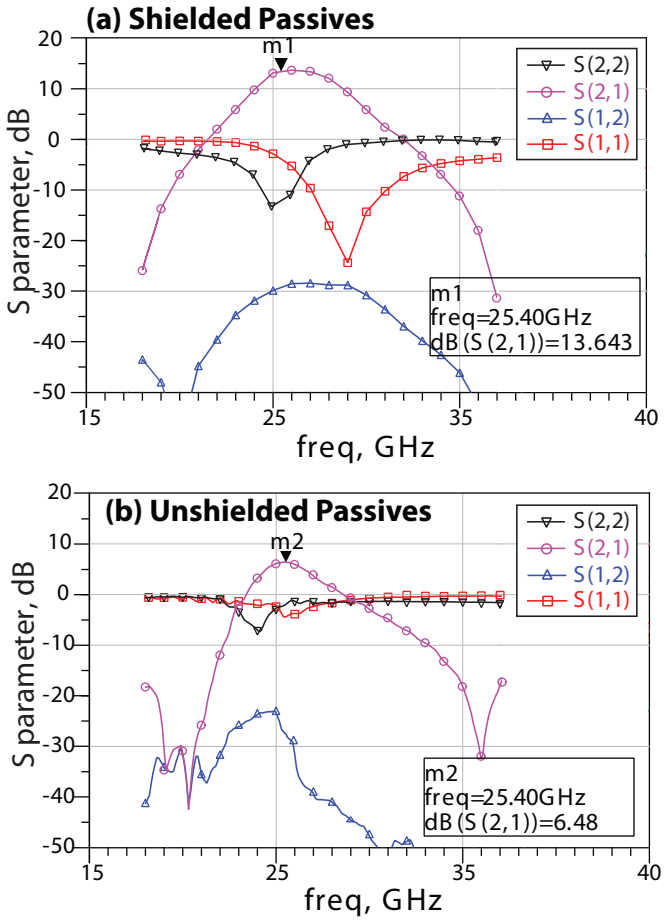


Figure 6.8: Measured S-parameters of the represented 25 GHz LNAs: (a) shielded and (b) unshielded inductors matching network.

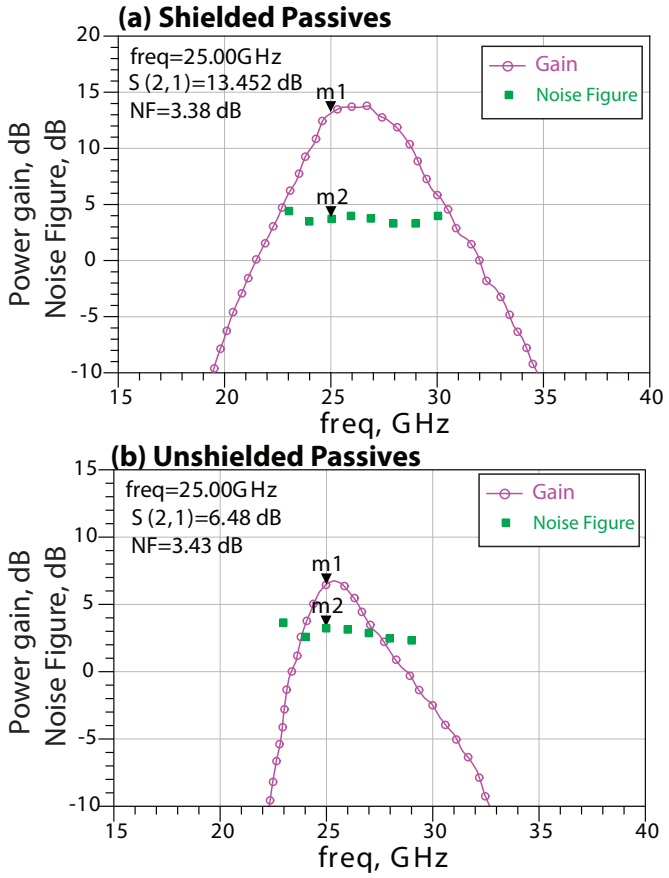


Figure 6.9: Measured gain and noise figure of the represented 25 GHz LNAs: (a) shielded and (b) unshielded inductors matching network.

Table 6.3: Performance comparison of recently reported CMOS LNAs operating around 25 GHz.

Ref	Technology (nm)	Topology	Frequency (GHz)	Gain (dB)	NF (dB)	Supply (V)	P_{dc} (mW)	Size (mm^2)
[157]	180	CS + CS	24	13.1	3.9	1	14	0.34
[158]	180	CS + CS	24	12.8	3.3	1	8	1.9
[159]	180	CS	24.5	12	6.4	0.8	8	0.26
[160]	180	CS	24	14	4.9	1	7.9	0.23
[161]	180	CS + CS.CG	24	9.3	5	1.8	27	0.39
[162]	180	CG + CS + CS	21.8	15	6	1.5	24	0.05
[163]	180	CS + CS + CS	23.7	12.8	5.6	1.8	54	0.73
[164]	180	CS.CG	24	11.5	5.7	1.4	3	0.31
This work	180	CS + CS	25.4	13.54	3.4	1.3	18.8	0.33

6.5.2 40GHz LNA: Optimised Versus Conventional Device

This aim of this study is to demonstrate the capability of the optimised device when used at frequencies close to the technology cut-off frequency. In this experiment two single stage source degenerated LNAs were designed at 40 GHz. The core devices are 40 finger common source with a W/L of 80/0.18. The first LNA (LNA1) used an optimised stacked device, whereas the second LNA (LNA2) used a conventional RF device. Both LNAs employed floating shield slow-wave CPW lines in the design of the input and output matching networks. The chip photographs of the fabricated LNAs are shown in Figure 6.10(a) and (b).

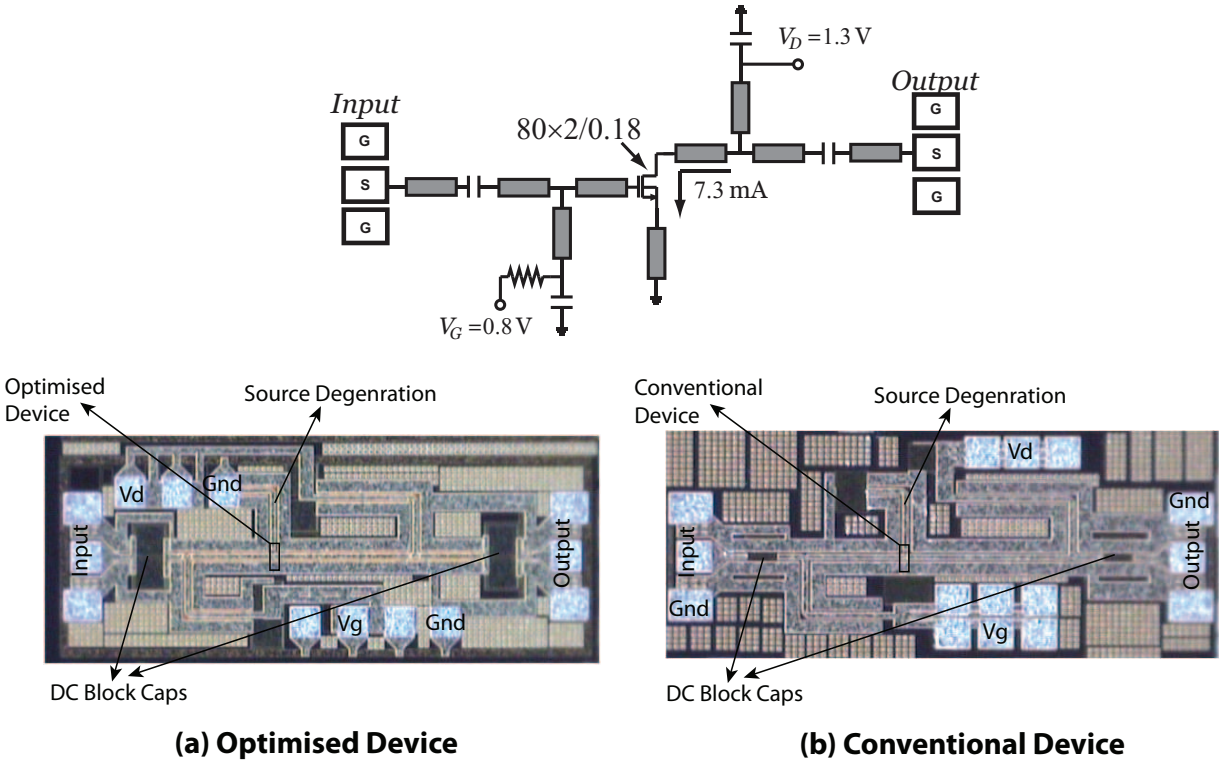


Figure 6.10: A chip photograph of two 40 GHz source degenerated LNAs implemented using 0.180 μm CMOS process. The matching networks are constructed using floating shield slow-wave CPW lines: (a) LNA1 utilising an optimised stacked device and (b) LNA2 utilising a conventional RF device.

The amplifiers are powered from a 1.3 V power source. The gate voltage is 0.8 V, the drain voltage is 1.3 V and draws 7.3 mA. This interprets to a power consumption of 9.49 mW. The measured S-parameters of both LNAs are shown in Figure 6.11. LNA1 (Figure 6.11(a)) exhibits a peak power gain of 11.16 dB at 38.4 GHz and a power gain of 10.42 dB at 40 GHz, a reverse isolation, S_{12}, less than -30 dB and a wide bandwidth of approximately 8 GHz (from approximately 34 GHz to 46 GHz). In contrast LNA2

(Figure 6.11(b)) exhibits a peak power gain of 3.88 dB at 40 GHz, a reverse isolation, S_{12}, of about −28 dB and a significantly narrower bandwidth of approximately 2 GHz (from approximately 39 GHz to 41 GHz).

It can be seen from Figure 6.11 that the optimised device, used in LNA1, has almost tripled the amplifier gain at 40 GHz when compared to the conventional device used in LNA2. The power gain is improved from 3.9 dB up to 10.4 dB at 40 GHz, even the noise figure improves significantly at this frequency, as shown in Figure 6.12. In fact the noise figure is almost halved; LNA2 exhibits a noise figure of 7.34 dB while LNA1 is 5.24 dB at 40 GHz. Furthermore, the noise figure of LNA1 is lower than 5.5 dB, over the amplifier bandwidth, while LNA2 displays a noise figure in the region of 7 dB over the entire bandwidth.

Table 6.4 presents and compares the performance of the CMOS millimetre-wave low noise amplifiers fabricated and measured in this work compared to results previously

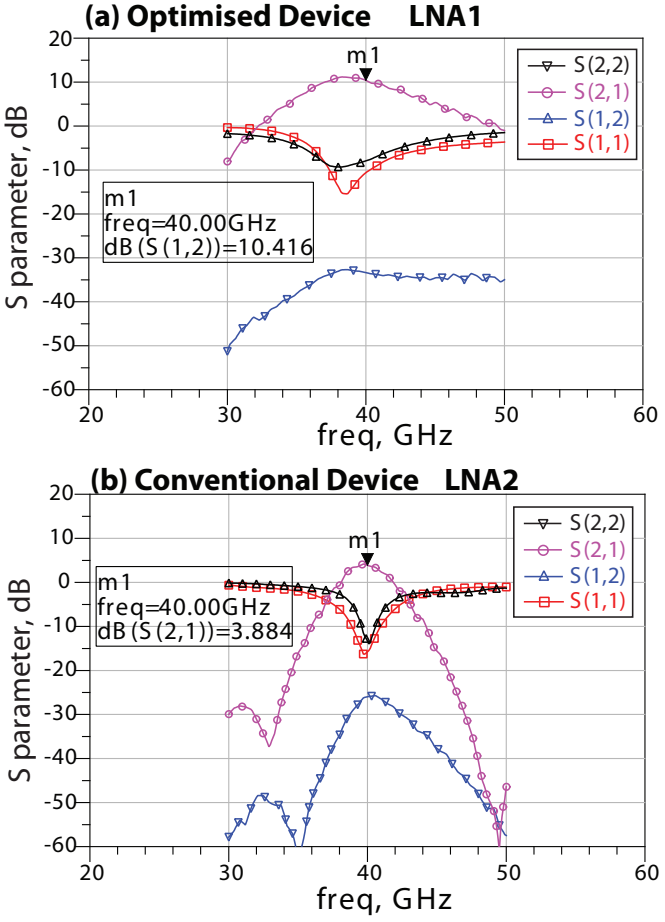


Figure 6.11: Measured S-parameters of the represented 40 GHz LNAs: (a) Optimised stacked device and (b) conventional RF device.

reported in literature. The work presented in [172] is based on the same technology node as this study. However to achieve a reasonably acceptable gain (7 dB) and a good noise figure (5 dB), the authors used a three stage common source common gate cascode topology. This is mainly due to the weak power gain that the core device can deliver. The optimised device presented in this work, with its improved power and noise capabilities, resolves this issue and eliminates the need for multiple stages yielding comparable power gain and noise figure from a single stage with lower power consumption and smaller chip area.

The work in [173] used a 130nm technology node, which naturally renders the devices to have higher power and lower noise performance in comparison to the 180nm node. However, the design in this study is still comparable with the referenced work, even though the latter used a three stage common source configuration. The work presented in [178] is probably most relevant with which to compare, as both LNAs are based on a single stage common source configuration. Nevertheless the amplifier in [178] is

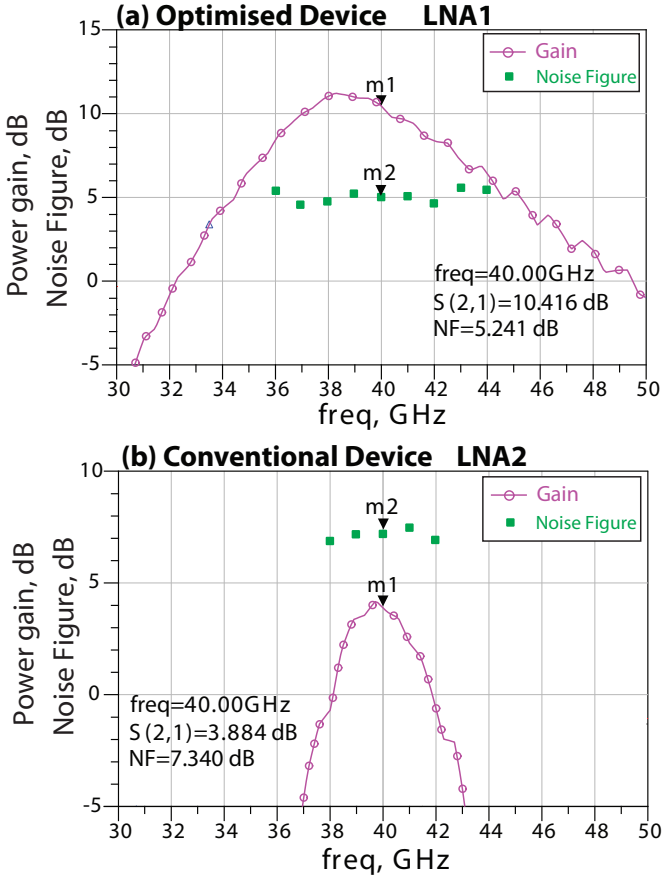


Figure 6.12: Measured gain and noise figure of the represented 40 GHz LNAs: (a) Optimised stacked device and (b) conventional RF device.

Table 6.4: Performance comparison of reported millimetre-wave CMOS LNAs.

Ref	Technology (nm)	Topology	Frequency (GHz)	Gain (dB)	NF (dB)	Supply (V)	P_{dc} (mW)	Size (mm^2)
This work	180	CS	40	10.4	5.24	1.3	18.8	0.38
[172]	180	3 x CS.CG	40	7	5	3	30	2.04
[173]	130	CS + CS + CS	43	20	4	1.5	36	0.525
[178]	90	CS	40	6.1	–	1.5	19	–

implemented in a more advanced CMOS technology (90nm) in comparison to 180nm CMOS used in this study. It can be seen that the gain in the reported study is less than that obtained in this study. However the power consumption is less in the published LNA, this is due to the difference in the technology node.

6.6 Summary

This chapter discussed different single-ended LNA topologies and highlighted the advantages and disadvantages of each topology. Based on the findings of this discussion, the inductor degenerated common source topology was identified to demonstrate the performance that fulfills the gain, noise figure and impedance matching requirements with minimum power consumption for narrow band designs. This was followed by a presentation and comparison of the state-of-the-art microwave and millimetre-wave CMOS LNAs implemented in technology nodes from 180 nm down to 45 nm. Following on from this, a modified inductive degenerative LNA design methodology was illustrated. A comparison was then made against the conventional design method. LNAs, using optimised stacked transistors, operating at microwave (25 GHz) and millimetre-wave (40 GHz) frequencies were designed, fabricated, measured and analysed. At 25 GHz, the shielded matching network showed a factor of 2 improvement in gain over the unshielded matching network. Moreover, the shielded LNA demonstrated comparable noise and gain performance to other reported state-of-the-art LNAs for the same technology node. For the 40 GHz LNA, a significant gain improvement and a noise reduction was measured from the the stacked (optimised) as opposed to the conventional transistor layout. The optimised transistor layout coupled with the shielded matching networks provided higher gain

and lower noise figure on a more compact chip area, compared with published results reported on several CMOS technology nodes.

Chapter 7

Conclusion and Future Work

In this thesis a study was conducted to demonstrate the feasibility of designing millimetre-wave modules in a low cost, standard 0.18 μm digital CMOS technology. To achieve this the curent limits of CMOS technology had to be investigated and understood. More specifically, the limitations associated with the lossy silicon substrate at millimetre-wave frequencies and the performance degradation caused by dominating parasitic effects in active devices at those frequencies. With knowledge of the technological challenges, new systematic methods of making high performance passive and active devices were proposed, designed, simulated, fabricated, modelled and analysed. The following achievements of this study were:

- Two new de-embedding small-signal and noise techniques were proposed and evaluated up to frequencies of 110 GHz. This provides more accurate device de-embedding in comparison to currently implemented methods. This results in a more accurate device model and therefore more robust device design at millimetre-wave frequencies.

- An evaluation of shielded (grounded and floating) and unshielded CMOS spiral inductors was performed. These were fabricated, measured and modelled. As a result of this work the significant improvement achieved in the use of shielded spiral inductors, especially floating shield, was identified for millimetre-wave devices. This work will contribute to the further reduction in the circuit dimensions of millimetre-wave building blocks.

- Shielded and unshielded slow-wave CPW transmission lines were investigated resulting in a new shielding methodology being proposed, fabricated and analysed. This study achieved a reduction in the line attenuation and decreased the structure length, this results in more compact and less cumbersome circuit design.

- Multi-layered comb capacitors were designed, fabricated, measured and modelled. This work will contribute to the reduction in circuit dimensions and improved modelling capabilities up to millimetre-wave frequencies.

- The impact of transistor layout parasitics on the device power and noise performance was investigated. A new transistor layout was proposed based on those findings. This device achieved an improvement in gain and noise performance compared to conventional RF device layout. The use of this device will make the operation of circuits beyond the technology cut-off frequency possible.

- Combining the previously detailed devices, two low noise amplifiers at 25 GHz and 40 GHz were designed, implemented and measured. The proposed 25 GHz LNA demonstrated the highest gain and lowest noise performance of any LNA operating at this frequency reported to date.

To begin this study, the metrology at high frequencies was explored to identify the required metrology equipment and methods to achieve accurate measurement of devices up to millimetre-wave frequencies. This is important in order to achieve accurate models of all the developed devices employed in the design blocks. Initially the conventional methods for de-embedding the DUT test pads were investigated, and shortcomings of these de-embedding methods were identified. With this knowledge two new methods were proposed to overcome these shortcomings. The methods were evaluated and proved to be valid for up to millimetre wave frequencies. Symmetric and asymmetric structures were fabricated and the measurement results were de-embedded using the conventional and proposed methods. The evaluation of these structures showed the proposed methods suitably compensated for the discrepancies existing in the conventional methods. This results in the ability to measure devices up to millimetre-wave frequencies with an increased accuracy compared to previous methods.

Circuits operating in the millimetre-wave region consist of a combination of components, each component contributes to the overall performance of the circuit. Therefore a clear

understanding and accurate models of each component is key for a successful design. In order to obtain those accurate models up to millimetre-wave frequencies, this work began with the design consideration of each component, such as spiral inductors, transmission lines, capacitors, resistors and transistors operating in this frequency band. All of these components were designed and implemented in a standard CMOS process. The first passive structure studied was spiral inductors. Several shielded and unshielded spirals were fabricated, measured, modelled and analysed. A comparative study was conducted on a floating patterned shield inductor, a grounded patterned shield inductor and an unshielded inductor. The comparison reveals that shielding improves the quality factor of the spiral. Furthermore floating patterned shield increases the quality factor by 21%. This identified that the shielded inductors exhibit a higher quality factor compared to the unshielded at these frequencies.

The second passive device to be explored were transmission lines on lossy silicon. Different transmission line layouts were compared. Multiple types of shielded slow-wave CPW transmission lines were designed, fabricated, measured, modelled and analysed. A novel shielding methodology was proposed and compared to conventional slow-wave CPW. The results of this comparison showed that the proposed floating shield slow-wave CPW provides the lowest attenuation and best quality factor, by a factor of 1.5, when compared to the other structures. In addition to this work compact comb capacitors and RF resistors were designed, fabricated, measured and modelled up to 110 GHz. The capacitors were designed for matching networks as well as AC coupling and supply bypass for circuits designed at these frequencies. The development of these passive structures for operation at millimetre-wave was essential to the design of low noise amplifiers operating at millimetre-wave frequencies, the final aim of this work.

The final component explored was the transistor. A sensitivity analysis was performed on multiple NMOS transistors in order to ascertain the impact of layout parasitics on the device performance. Based on the findings of the study an optimised transistor layout, that decreases the effects of the parasitics responsible for device performance degradation, was proposed, fabricated, measured and analysed. A comparative study was conducted to prove the performance improvement. The proposed transistor layout featured an approximately 50% increase in f_{max} and a 34% improvement in the minimum noise figure when compared to the conventional RF device layout. A small-signal and noise model for conventional and optimised stacked NMOS transistor was extracted from

measured data. The modelling methods were verified and the results showed an error less of than 1% between the measured and modelled responses. These results provided an assurance that the modelling methodologies employed in this work are valid and the resulting models can be used in the design of millimetre-wave circuits with confidence.

All the aforementioned components were integrated to realise low noise amplifiers, with the purpose of optimising the design at 25 GHz and 40 GHz. In relation to this, two identical LNAs, with two different matching networks (shielded and unshielded), operating at 25 GHz were designed, fabricated, measured and analysed. Although the results indicated a comparable noise figure between the LNAs, a factor of two improvement in the power gain was obtained from the LNA with the shielded matching network. A benchmarking study highlighted the shielded LNA to have the highest performance in comparison to the state-of the-art LNAs implemented using the same technology. The amplifiers at 40 GHz were designed, fabricated, measured and analysed, based on two different transistor layouts; conventional and stacked. The latter being proposed for performance optimisation. The study showed that the gain tripled for the stacked layout, with a factor of two noise figure reduction compared to the conventional layout. Benchmarking this LNA to that reported in literature for a similar technology node revealed a 30% enhancement in the gain. Moreover, the high gain capability of the stacked transistor layout eliminated the need for multiple stage LNA, resulting in a more compact design and therefore smaller chip area.

This work has demonstrated the extension of 0.18 μm CMOS technology into the millimetre-wave regime, thereby extending the technology beyond its current capabilities. This enables millimetre-wave technology to be implemented in the same processes as other digital circuitry enabling integrated solutions and reducing the production costs. This could be extremely beneficial for commercial application (medical, automotive radar, entertainment and broadcasting industries) where cost, reliability and compact integration is of a great importance.

7.1 Future Work

The focus of this work has been extending the performance of 0.18 μm CMOS technology into the millimetre-wave regime. The natural progression from the work presented here

would be towards the implementation of low cost, compact, CMOS millimetre-wave transceivers implemented on a single chip. In order to achieve this, the following points have to be investigated and achieved:

- **Transistor enhancement**: The transistor layout optimisation can be investigated further by employing other layout methodologies to reduce the transistor parasitics further and reduce transistor overall area. For example using hexagonal shaped devices improves stacking and reduce size. Moreover, an investigation into other transistor topologies, such as differential and cascode common source-common gate could be conducted. This may help expand the device capability and create a few alternative routes of investigation for circuit topologies.

- **Scalable non-linear millimetre-wave transistor models**: Reliable and accurate scalable device models are key to a successful and robust circuit designs, presently the models provided by CMOS foundries are insufficient when used in the millimetre-wave regime. Accurate scalable, non-linear transistor models are required to enable the design of reliable millimetre-wave non-linear modules such as power amplifiers, mixers and oscillators.

- **ESD protection**: ESD is one of the most common challenges for microwave and millimetre-wave systems implemented in CMOS technology. The thin gate oxide and low gate oxide breakdown, make the transistor more vulnerable to ESD damage. The low noise amplifier (LNA) is usually the first stage in any wireless transceiver. ESD is normally desirable at this stage to protect the system. However, and as we have seen in this work, in this frequency range the circuit characteristics are very sensitive to the parasitic components introduced by the ESD blocks. The noise introduced by the ESD block which could degrade the noise performance of the LNA and consequently the overall noise performance of the transceiver. Therefore, ESD circuits should be added to the design. But first different ESD circuits topologies should be studied and analysed looking into the advantages and disadvantages of each topology. Then implement the best ESD circuit in millimetre-wave circuits to achieve a high performance LNA with sufficient robustness.

- **CMOS millimetre-wave Power amplifiers**: Despite the effort to improve the amplifier noise and gain at millimetre-wave frequencies. The difficulty of generating

sufficiently high output power at the transmitter end is still a challenge due to the low break down voltage in CMOS technology. Power combining techniques, such as multilevel and spatial combining, could be the solution to provide high power at low supply voltages.

- **CMOS millimetre-wave oscillators and mixers**: Provided that a non-linear transistor model is available, different CMOS oscillator and mixer topologies can be investigated and implemented.

- **Fully-integrated millimetre-wave CMOS transceivers**: Once all of the individual circuit blocks have been designed, the natural next step is to integrate all stages into a single-chip to achieve a fully integrated millimetre-wave transceiver. This would enable compact, low cost, manufacturable millimetre-wave modules to be realised for high frequency applications.

Bibliography

[1] C. Yue, C. Ryu, J. Lau, T. Lee, and S. Wong, "A physical model for planar spiral inductors on silicon," in *Electron Devices Meeting, 1996., IEEE International*, 1996, pp. 155–158.

[2] R. Beringer, "The absorption of one-half centimeter electromagnetic waves in oxygen," *Phys. Rev.*, vol. 70, pp. 53–57, 1946.

[3] S. Sze, *High-speed semiconductor devices.* Wiley, 1990.

[4] R. S. Muller and T. Kamins, *Device electronics for integrated circuits.* Wiley, 1977.

[5] R. Lai, X. Mei, W. Deal, W. Yoshida, Y. Kim, P. Liu, J. Lee, J. Uyeda, V. Radisic, M. Lange, T. Gaier, L. Samoska, and A. Fung, "Sub 50 nm InP HEMT device with fmax greater than 1 THz," in *Electron Devices Meeting, 2007., IEEE International*, Dec. 2007, pp. 609–611.

[6] V. Jain, J. Rode, H.-W. Chiang, A. Baraskar, E. Lobisser, B. Thibeault, M. Rodwell, M. Urteaga, D. Loubychev, A. Snyder, Y. Wu, J. Fastenau, and W. Liu, "1.0 THz fmax InP DHBTs in a refractory emitter and self-aligned base process for reduced base access resistance," in *Device Research Conference (DRC), 69th Annual*, Jun. 2011, pp. 271 –272.

[7] I. Bahl and P. Bhartia, *Microwave solid state circuit design.* Wiley-Interscience, 2003.

[8] A. Niknejad and H. Hashemi, "Introduction to mm-wave silicon devices, circuits, and systems," in *mm-Wave Silicon Technology*, ser. Series on Integrated Circuits and Systems, A. Niknejad and H. Hashemi, Eds. Springer US, 2008, pp. 1–24.

[9] B. Heydari, M. Bohsali, E. Adabi, and A. Niknejad, "Millimeter-wave devices and circuit blocks up to 104 GHz in 90 nm CMOS," *Solid-State Circuits, IEEE Journal of*, vol. 42, no. 12, pp. 2893–2903, 2007.

[10] M. Kantanen, M. Varonen, M. Karkkainen, T. Karttaavi, R. Weber, P. Leuther, M. Seelmann-Eggebert, T. Narhi, and K. A. I. Halonen, "Coplanar 155 GHz MHEMT MMIC low noise amplifiers," in *Microwave Conference Proceedings, 2006. APMC 2006. Asia-Pacific Conference Proceedings*, Dec. 2006, pp. 173–176.

[11] R. Witte, *Spectrum and network measurements.* Noble Publishing Corporation, 1993.

[12] H. Van Hamme and M. Vanden Bossche, "Flexible vector network analyzer calibration with accuracy bounds using an 8-term or a 16-term error correction model," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 42, no. 6, pp. 976–987, 1994.

[13] S. Wartenberg, "Selected topics in RF coplanar probing," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 51, no. 4, pp. 1413–1421, 2003.

[14] S. Rehnmark, "On the calibration process of automatic network analyzer systems (short papers)," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 22, no. 4, pp. 457–458, 1974.

[15] S. Basu and L. Hayden, "An SOLR calibration for accurate measurement of orthogonal on-wafer DUTs," in *Microwave Symposium Digest, 1997., IEEE MTT-S International*, vol. 3, Jun. 1997, pp. 1335–1338.

[16] U. Stumper, "Influence of nonideal calibration items on S-parameter uncertainties applying the SOLR calibration method," *Instrumentation and Measurement, IEEE Transactions on*, vol. 58, no. 4, pp. 1158–1163, 2009.

[17] A. Ferrero and U. Pisani, "Two-port network analyzer calibration using an unknown 'thru'," *Microwave and Guided Wave Letters, IEEE*, vol. 2, no. 12, pp. 505–507, 1992.

[18] H.-J. Eul and B. Schiek, "Thru-match-reflect: One result of a rigorous theory for de-embedding and network analyzer calibration," in *Microwave Conference, 18th European*, Sep. 1988, pp. 909–914.

[19] F. Purroy and L. Pradell, "New theoretical analysis of the LRRM calibration technique for vector network analyzers," *Instrumentation and Measurement, IEEE Transactions on*, vol. 50, no. 5, pp. 1307–1314, 2001.

[20] A. Davidson, K. Jones, and E. Strid, "LRM and LRRM calibrations with automatic determination of load inductance," in *ARFTG Conference Digest-Fall, 36th*, vol. 18, Nov. 1990, pp. 57–63.

[21] J. Ou and M. Caggiano, "Determine two-port S-parameters from one-port measurements using calibration substrate standards," in *Electronic Components and Technology Conference, 2005. Proceedings. 55th*, May. 2005, pp. 1765–1768.

[22] J. Liou, A. Ortiz-Conde, and F. Garcia-Sanchez, *Analysis and design of MOSFETs: Modeling, simulation, and parameter extraction*, ser. Software Engineering; 4. Springer US, 1998.

[23] A. Technology, "Noise figure measurement accuracy - the Y-factor method, application note 57-2."

[24] R. Leffel, "The Y factor technique for noise figure measurements, application note."

[25] I. Maxim Integrated Products, "Three methods of noise figure measurement, tutorial 2875," Nov 2003.

[26] V. Adamian and J. Uhlir, A., "Simplified noise evaluation of microwave receivers," *Instrumentation and Measurement, IEEE Transactions on*, vol. 33, no. 2, pp. 136–140, 1984.

[27] V. Adamian and A. Uhlir, "A novel procedure for receiver noise characterization," *Instrumentation and Measurement, IEEE Transactions on*, vol. 22, no. 2, pp. 181–182, 1973.

[28] J. Browne, "On-wafer probes test Si ICs to 110 GHz," *Microwave and RF*, vol. 42, pp. 84–90, 2003.

[29] D. Williams, A. Byers, V. Tyree, D. Walker, J. Ou, X. Jin, M. Piket-May, and C. Hu, "Contact-pad design for high-frequency silicon measurements," in *Electrical Performance of Electronic Packaging, 2000, IEEE Conference on.*, Oct. 2000, pp. 131–134.

[30] T. Kolding, "On-wafer calibration techniques for giga-hertz CMOS measurements," in *Microelectronic Test Structures, 1999. ICMTS 1999. Proceedings of the 1999 International Conference on*, Mar. 1999, pp. 105–110.

[31] R. Gillon, W. Van De Sype, D. Vanhoenaker, and L. Martens, "Comparing high-frequency de-embedding strategies: immittance correction and in-situ calibration," in *Microelectronic Test Structures, 2000. ICMTS 2000. Proceedings of the 2000 International Conference on*, Mar. 2000, pp. 241–245.

[32] E. Vandamme, D. Schreurs, and C. Van Dinther, "Improved three-step de-embedding method to accurately account for the influence of pad parasitics in silicon on-wafer RF test-structures," *Electron Devices, IEEE Transactions on*, vol. 48, no. 4, pp. 737–742, 2001.

[33] T. Kolding, "A four-step method for de-embedding gigahertz on-wafer CMOS measurements," *Electron Devices, IEEE Transactions on*, vol. 47, no. 4, pp. 734–740, 2000.

[34] I. M. Kang, S.-J. Jung, T.-H. Choi, J.-H. Jung, C. Chung, H.-S. Kim, H. Oh, H.-W. Lee, G. Jo, Y.-K. Kim, H.-G. Kim, and K.-M. Choi, "Five-step (pad-pad short-pad open-short open) de-embedding method and its verification," *Electron Device Letters, IEEE*, vol. 30, no. 4, pp. 398–400, 2009.

[35] R. Amaya, P. Popplewell, M. Cloutier, and C. Plett, "Analysis and measurements of EM and substrate coupling effects in common RF integrated circuits," in *Custom Integrated Circuits Conference, 2004. Proceedings of the IEEE 2004*, Oct. 2004, pp. 363 – 366.

[36] *Maxwell 2D Parameter Extractor Users Reference*, Ansoft Corporation, 2001.

[37] A. E. Ruehli, "Equivalent circuit models for three-dimensional multiconductor systems," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 22, no. 3, pp. 216–221, 1974.

[38] C. Yue and S. Wong, "On-chip spiral inductors with patterned ground shields for Si-based RF ICs," *Solid-State Circuits, IEEE Journal of*, vol. 33, no. 5, pp. 743–752, 1998.

[39] S. Wentworth, *Fundamentals of electromagnetics with engineering applications*. Wiley, 2006.

[40] J. Craninckx and M. S. J. Steyaert, "A 1.8-GHz low-phase-noise CMOS VCO using optimized hollow spiral inductors," *Solid-State Circuits, IEEE Journal of*, vol. 32, no. 5, pp. 736–744, 1997.

[41] J. Long, M. Copeland, S. Kovacic, D. Malhi, and D. Harame, "RF analog and digital circuits in SiGe technology," in *Solid-State Circuits Conference, 1996. Digest of Technical Papers. 42nd ISSCC., 1996 IEEE International*, Feb. 1996, pp. 82–83.

[42] H. Jiang, Y. Wang, J.-L. A. Yeh, and N. C. Tien, "On-chip spiral inductors suspended over deep copper-lined cavities," *IEEE Transactions on Microwave Theory Techniques*, vol. 48, pp. 2415–2423, 2000.

[43] J.-B. Yoon, Y.-S. Choi, B.-I. Kim, Y. Eo, and E. Yoon, "CMOS-compatible surface-micromachined suspended-spiral inductors for multi-GHz silicon RF ICs," *Electron Device Letters, IEEE*, vol. 23, no. 10, pp. 591–593, 2002.

[44] C. H. Ahn and M. G. Allen, "Micromachined planar inductors on silicon wafers for MEMS applications," *Industrial Electronics, IEEE Transactions on*, vol. 45, no. 6, pp. 866–876, 1998.

[45] K. Benaissa, J.-Y. Yang, D. Crenshaw, B. Williams, S. Sridhar, J. Ai, G. Boselli, S. Zhao, S. Tang, S. Ashburn *et al.*, "RF CMOS on high-resistivity substrates for system-on-chip applications," *Electron Devices, IEEE Transactions on*, vol. 50, no. 3, pp. 567–576, 2003.

[46] Y. Cao, R. Groves, N. Zamdmer, J.-O. Plouchart, R. Wachnik, X. Huang, T.-J. King, and C. Hu, "Frequency-independent equivalent circuit model for on-chip spiral inductors," in *Custom Integrated Circuits Conference, 2002. Proceedings of the IEEE 2002*, 2002, pp. 217–220.

[47] S. Kim and D. Neikirk, "Compact equivalent circuit model for the skin effect," in *Microwave Symposium Digest, 1996., IEEE MTT-S International*, vol. 3, 1996, pp. 1815–1818.

[48] D. Melendy, P. Francis, C. Pichler, K. Hwang, G. Srinivasan, and A. Weisshaar, "A new wideband compact model for spiral inductors in RFICs," *Electron Device Letters, IEEE*, vol. 23, no. 5, pp. 273–275, 2002.

[49] R. Bunch, D. Sanderson, and S. Raman, "Quality factor and inductance in differential IC implementations," *Microwave Magazine, IEEE*, vol. 3, no. 2, pp. 82–92, 2002.

[50] E. Ragonese, T. Biondi, A. Scuderi, and G. Palmisano, "A lumped scalable physics-based model for silicon spiral inductors," in *Electron Devices for Microwave and Optoelectronic Applications, 2002. EDMO 2002. The 10th IEEE International Symposium on*, Nov. 2002, pp. 119–124.

[51] M. Yang, T. Yeh, W. C. Lin, H. Hsu, P. Ho, Y. Wang, Y. Chia, and D. Tang, "Characterization and model of high quality factor and broadband integrated inductor on Si-substrate," in *Microwave Symposium Digest, 2003 IEEE MTT-S International*, vol. 2, Jun. 2003, pp. 1283–1286.

[52] M. Fujishima and J. Kino, "Accurate subcircuit model of an on-chip inductor with a new substrate network," in *VLSI Circuits, 2004. Digest of Technical Papers. 2004 Symposium on*, Jun. 2004, pp. 376–379.

[53] S. Mohan, M. del Mar Hershenson, S. Boyd, and T. Lee, "Simple accurate expressions for planar spiral inductances," *Solid-State Circuits, IEEE Journal of*, vol. 34, no. 10, pp. 1419–1424, 1999.

[54] X. Shi, J.-G. Ma, K.-S. Yeo, A. Do, and E. Li, "Equivalent circuit model of on-wafer CMOS interconnects for RFICs," *Very Large Scale Integration (VLSI) Systems, IEEE Transactions on*, vol. 13, no. 9, pp. 1060–1071, 2005.

[55] F. Grover, *Inductance calculations: working formulas and tables*, ser. Dover phoenix editions. D. Van Nostrand Company, Incorporated, 1946.

[56] H. Greenhouse, "Design of planar rectangular microelectronic inductors," *Parts, Hybrids, and Packaging, IEEE Transactions on*, vol. 10, no. 2, pp. 101–109, 1974.

[57] J. Burghartz, D. Edelstein, M. Soytter, H. Ainspan, and K. Jenkins, "RF circuit design aspects of spiral inductors on silicon," in *Solid-State Circuits Conference, 1998. Digest of Technical Papers. 1998 IEEE International*, Feb. 1998, pp. 246–247.

[58] C. Yue and S. Wong, "Physical modeling of spiral inductors on silicon," *Electron Devices, IEEE Transactions on*, vol. 47, no. 3, pp. 560–568, 2000.

[59] R. Adler, *Electromagnetic energy transmission and radiation.* John Wiley, 1960.

[60] O. Murphy, K. McCarthy, C. Delabie, A. Murphy, and P. Murphy, "Design of multiple-metal stacked inductors incorporating an extended physical model," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 53, no. 6, pp. 2063–2072, 2005.

[61] W. Kuhn, X. He, and M. Mojarradi, "Modeling spiral inductors in SOS processes," *Electron Devices, IEEE Transactions on*, vol. 51, no. 5, pp. 677–683, 2004.

[62] M.-C. Hsieh, Y.-K. Fang, C.-H. Chen, S.-M. Chen, and W.-K. Yeh, "Design and fabrication of deep submicron CMOS technology compatible suspended high-Q spiral inductors," *Electron Devices, IEEE Transactions on*, vol. 51, no. 3, pp. 324–331, 2004.

[63] S.-M. Yim, T. Chen, and K. O, "The effects of a ground shield on the characteristics and performance of spiral inductors," *Solid-State Circuits, IEEE Journal of*, vol. 37, no. 2, pp. 237–244, 2002.

[64] R. Schmitt, *Electromagnetics explained [electronic resource]: a handbook for wireless/RF, EMC, and high-speed electronics*, ser. Edn Series for Design Engineers. Newnes, 2002.

[65] G. González, *Microwave transistor amplifiers: analysis and design.* Prentice Hall, 1997.

[66] D. Williams and R. B. Marks, "Transmission line capacitance measurement," *Microwave and Guided Wave Letters, IEEE*, vol. 1, no. 9, pp. 243–245, 1991.

[67] J. Brews, "Transmission line models for lossy waveguide interconnections in VLSI," *Electron Devices, IEEE Transactions on*, vol. 33, no. 9, pp. 1356–1365, 1986.

[68] A. Niknejad and H. Hashemi, *mm-wave silicon technology: 60GHz and beyond*, ser. Series on Integrated Circuits and Systems. Springer Science+Business Media, 2008.

[69] H. Hsu, "On the general relation between α and Q (correspondence)," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 11, no. 4, pp. 258–258, 1963.

[70] F. Ellinger, "26-42 GHz SOI CMOS low noise amplifier," *Solid-State Circuits, IEEE Journal of*, vol. 39, no. 3, pp. 522–528, 2004.

[71] M. Varonen, M. Karkkainen, M. Kantanen, and K. A. I. Halonen, "Millimeter-wave integrated circuits in 65-nm CMOS," *Solid-State Circuits, IEEE Journal of*, vol. 43, no. 9, pp. 1991–2002, 2008.

[72] H. Hasegawa and H. Okizaki, "M.I.S. and Schottky slow-wave coplanar striplines on GaAs substrates," *Electronics Letters*, vol. 13, pp. 663–664, 1977.

[73] W. Te-Hu and T. Itoh, "Compact grating structure for application to filters and resonators in monolithic microwave integrated circuits," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 35, no. 12, pp. 1176–1182, 1987.

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

[75] A. Mellberg and J. Stenarson, "An evaluation of three simple scalable MIM capacitor models," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 54, no. 1, pp. 169–172, 2006.

[76] C.-H. Ng, C.-S. Ho, S.-F. S. Chu, and S.-C. Sun, "MIM capacitor integration for mixed-signal/RF applications," *Electron Devices, IEEE Transactions on*, vol. 52, no. 7, pp. 1399–1409, 2005.

[77] K. Subramaniam, A. Kordesch, and M. Esa, "Design and modeling of metal finger capacitors for RF applications," in *Applied Electromagnetics, 2005. APACE 2005. Asia-Pacific Conference on*, Dec. 2005, pp. 293–296.

[78] T. Sowlati, V. Vathulya, and D. Leenaerts, "," in *Low Power Electronics and Design, International Symposium on, 2001.*, 2001, pp. 243–246.

[79] A. Imamura, M. Fujishima, and K. Hoh, "Bending-comb capacitor with a small parasitic inductance," in *VLSI Circuits Digest of Technical Papers, 2002. Symposium on*, Jun. 2002, pp. 22–25.

[80] Z. Yu and C. McAndrew, "RF CMOS is more than CMOS: Modeling of RF passive components," in *Custom Integrated Circuits Conference, 2009. CICC '09. IEEE*, Sep. 2009, pp. 407–414.

[81] R. V. H. Booth and C. McAndrew, "A 3-terminal model for diffused and ion-implanted resistors," *Electron Devices, IEEE Transactions on*, vol. 44, no. 5, pp. 809–814, 1997.

[82] A. Hastings, *The art of analog layout*, 2nd ed. Pearson Prentice Hall, 2006, ch. 5.

[83] B. Razavi, *Design of analog CMOS integrated circuits*, ser. McGraw-Hill higher education. McGraw-Hill Education (India) Pvt Limited, 2002.

[84] T. Ytterdal, Y. Cheng, and T. Fjeldly, *Device Modeling for analog and RF CMOS circuit design*. Wiley, 2003.

[85] H. Lund, "IV and CV characterization of 90nm CMOS transistors," 2006.

[86] Y. Tsividis, *Operation and modeling of the MOS transistor*, ser. McGraw-Hill series in electrical engineering: VLSI. McGraw-Hill, 1987.

[87] R. Spence, *Linear active networks*. Wiley-Interscience, 1970.

[88] S. Mason, "Power gain in feedback amplifier," *Circuit Theory, Transactions of the IRE Professional Group on*, vol. CT-1, no. 2, pp. 20–25, 1954.

[89] H. Fukui, "Optimal noise figure of microwave GaAs MESFET's," *Electron Devices, IEEE Transactions on*, vol. 26, no. 7, pp. 1032–1037, 1979.

[90] H. Nyquist, "Thermal agitation of electric charge in conductors," *Phys. Rev.*, vol. 32, pp. 110–113, 1928.

[91] J. B. Johnson, "Thermal agitation of electricity in conductors," *Phys. Rev.*, vol. 32, pp. 97–109, 1928.

[92] A. Van Der Ziel, *Noise in solid state devices and circuits*. Wiley, 1986.

[93] A. L. McWorther, *Semiconductor surface physics*. University of Pennsylvania Press, Philadelphia, 1957.

[94] A. van der Ziel, "Unified presentation of 1/f noise in electron devices: fundamental 1/f noise sources," *Proceedings of the IEEE*, vol. 76, no. 3, pp. 233–258, 1988.

[95] F. Hooge, "1/f noise," *Physica B+C*, vol. 83, no. 1, pp. 14 – 23, 1976.

[96] S. Christensson, I. Lundstrm, and C. Svensson, "Low frequency noise in MOS transistors – I Theory," *Solid-State Electronics*, vol. 11, no. 9, pp. 797–812, 1968.

[97] S. Christensson and I. Lundstrm, "Low frequency noise in MOS transistors – II Experiments," *Solid-State Electronics*, vol. 11, no. 9, pp. 813–820, 1968.

[98] A. Van der Ziel, *Solid state physical electronics*, ser. Solid state physical electronics series. Prentice Hall, 1976, ch. 18.

[99] A. Abidi, "High-frequency noise measurements on FET's with small dimensions," *Electron Devices, IEEE Transactions on*, vol. 33, no. 11, pp. 1801–1805, 1986.

[100] R. Jindal, "Hot-electron effects on channel thermal noise in fine-line NMOS field-effect transistors," *Electron Devices, IEEE Transactions on*, vol. 33, no. 9, pp. 1395–1397, 1986.

[101] S. Tedja, J. Van der Spiegel, and H. Williams, "Analytical and experimental studies of thermal noise in MOSFET's," *Electron Devices, IEEE Transactions on*, vol. 41, no. 11, pp. 2069–2075, 1994.

[102] A. Scholten, H. J. Tromp, L. Tiemeijer, R. van Langevelde, R. Havens, P. De Vreede, R. F. M. Roes, P. Woerlee, A. Montree, and D. B. M. Klaassen, "Accurate thermal noise model for deep-submicron CMOS," in *Electron Devices Meeting, 1999., IEEE International*, Dec. 1999, pp. 155–158.

[103] H. Haus and R. Adler, *Circuit theory of linear noisy networks*. BiblioBazaar, 2011.

[104] H. Hillbrand and P. Russer, "An efficient method for computer aided noise analysis of linear amplifier networks," *Circuits and Systems, IEEE Transactions on*, vol. 23, no. 4, pp. 235–238, 1976.

[105] G. Dambrine, A. Cappy, F. Heliodore, and E. Playez, "A new method for determining the FET small-signal equivalent circuit," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 36, no. 7, pp. 1151–1159, 1988.

[106] A. Tong, K. S. Yeo, L. Jia, C. Q. Geng, J.-G. Ma, and M. Do, "Simple and accurate extraction methodology for RF MOSFET valid up to 20 GHz," *Circuits, Devices and Systems, IEE Proceedings* -, vol. 151, no. 6, pp. 587–592, 2004.

[107] D. Lovelace, J. Costa, and N. Camilleri, "Extracting small-signal model parameters of silicon MOSFET transistors," in *Microwave Symposium Digest, 1994., IEEE MTT-S International*, May. 1994, pp. 865–868.

[108] I. Kwon, M. Je, K. Lee, and H. Shin, "A simple and analytical parameter-extraction method of a microwave MOSFET," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 50, no. 6, pp. 1503–1509, 2002.

[109] S. Lee, H.-K. Yu, C.-S. Kim, J.-G. Koo, and K.-S. Nam, "A novel approach to extracting small-signal model parameters of silicon MOSFET's," *Microwave and Guided Wave Letters, IEEE*, vol. 7, no. 3, pp. 75–77, 1997.

[110] S. Lee and H.-K. Yu, "A semianalytical parameter extraction of a SPICE BSIM3v3 for RF MOSFET's using S-parameters," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 48, no. 3, pp. 412–416, 2000.

[111] S. Lee, C.-S. Kim, and H.-K. Yu, "A small-signal RF model and its parameter extraction for substrate effects in RF MOSFETs," *Electron Devices, IEEE Transactions on*, vol. 48, no. 7, pp. 1374–1379, 2001.

[112] S.-M. Jen, C. Enz, D. Pehlke, M. Schroter, and B. Sheu, "Accurate modeling and parameter extraction for MOS transistors valid up to 10 GHz," *Electron Devices, IEEE Transactions on*, vol. 46, no. 11, pp. 2217–2227, 1999.

[113] A. Abuelmaatti and I. Thayne, "A new RF SiCMOS SDD model for quantifying individual contribution to distortion from transistor's nonlinear parameters," in *Electronics, Circuits and Systems, 2006. ICECS '06. 13th IEEE International Conference on*, Dec 2006, pp. 5–8.

[114] M.-S. Yang and S.-G. Lee, "Linearity improvement technique for the common-source transconductance stage," in *Midwest International Conference on Circuit And System, 2003. Proceedings. MWSCAS '03.*, 2003.

[115] W. Liu, X. Jin, J. Chen, M.-C. Jeng, Z. Liu, Y. Cheng, K. Chen, M. Chan, K. Hui, J. Huang, R. Tu, P. Ko, and C. Hu, "BSIM3v3.2.2 MOSFET model users' manual," EECS Department, University of California, Berkeley, Tech. Rep. UCB/ERL M99/18, 1999.

[116] W. Liu, R. Gharpurey, M.-C. Chang, U. Erdogan, R. Aggarwal, and J.-P. Mattia, "RF MOSFET modeling accounting for distributed substrate and channel resistances with emphasis on the BSIM3v3 SPICE model," in *Electron Devices Meeting, 1997., IEEE International*, Dec 1997, pp. 309–312.

[117] C. Biber, M. Schmatz, T. Morf, U. Lott, and W. Bachtold, "A nonlinear microwave MOSFET model for SPICE simulators," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 46, no. 5, pp. 604–610, 1998.

[118] A. Scholten, L. Tiemeijer, R. van Langevelde, R. Havens, A. Zegers-van Duijnhoven, and V. Venezia, "Noise modeling for RF CMOS circuit simulation," *Electron Devices, IEEE Transactions on*, vol. 50, no. 3, pp. 618–632, 2003.

[119] K. Han, J. Gil, S.-S. Song, J. Han, H. Shin, C.-K. Kim, and K. Lee, "Complete high-frequency thermal noise modeling of short-channel MOSFETs and design of 5.2-GHz low noise amplifier," *Solid-State Circuits, IEEE Journal of*, vol. 40, no. 3, pp. 726–735, 2005.

[120] A. Roy and C. Enz, "Compact modeling of thermal noise in the MOS transistor," *Electron Devices, IEEE Transactions on*, vol. 52, no. 4, pp. 611–614, 2005.

[121] C.-S. Kim, H.-K. Yu, H. Cho, S. Lee, and K. S. Nam, "CMOS layout and bias optimization for RFIC design applications," in *Microwave Symposium Digest, 1997., IEEE MTT-S International*, vol. 2, June 1997, pp. 945–948 vol.2.

[122] E. Morifuji, H. Momose, T. Ohguro, T. Yoshitomi, H. Kimijima, F. Matsuoka, M. Kinugawa, Y. Katsumata, and H. Iwai, "Future perspective and scaling down roadmap for RF CMOS," in *VLSI Circuits, 1999. Digest of Technical Papers. 1999 Symposium on*, Jun. 1999, pp. 165–166.

[123] P. H. Woerlee, M. Knitel, R. van Langevelde, D. Klaassen, L. Tiemeijer, A. Scholten, and A. Zegers-van Duijnhoven, "RF-CMOS performance trends," *Electron Devices, IEEE Transactions on*, vol. 48, no. 8, pp. 1776–1782, 2001.

[124] T. Tatsumi, "Geometry Optimization of Sub-100nm Node RF CMOS Utilizing Three Dimensional TCAD Simulation," in *Solid-State Device Research Conference, ESSDERC 2006. Proceeding of the 36th European*, Sep. 2006, pp. 319–322.

[125] A. Nakamura, N. Yoshikawa, T. Miyazako, T. Oishi, H. Ammo, and K. Takeshita, "Layout optimization of RF CMOS in the 90nm generation by a physics-based model including the multi-finger wiring effect," in *Radio Frequency Integrated Circuits (RFIC) Symposium, 2006 IEEE*, Jun. 2006, pp. 1–4.

[126] C. Cao and K. O, "A 90-GHz voltage-controlled oscillator with a 2.2-GHz tuning range in a 130-nm CMOS technology," in *VLSI Circuits, 2005. Digest of Technical Papers. 2005 Symposium on*, Jun. 2005, pp. 242–243.

[127] C. Liang and B. Razavi, "Systematic transistor and inductor modeling for millimeter-wave design," *Solid-State Circuits, IEEE Journal of*, vol. 44, no. 2, pp. 450–457, 2009.

[128] C.-Y. Chan, S.-C. Chen, M.-H. Tsai, and S. Hsu, "Wiring effect optimization in 65-nm low-power NMOS," *Electron Device Letters, IEEE*, vol. 29, no. 11, pp. 1245–1248, 2008.

[129] H.-S. Jhon, J.-H. Lee, J. Lee, B. Oh, I. Song, Y. Yun, B.-G. Park, J.-D. Lee, and H. Shin, "f_{max} improvement by controlling extrinsic parasitics in circuit-level MOS transistor," *Electron Device Letters, IEEE*, vol. 30, no. 12, pp. 1323–1325, 2009.

[130] W. Wu, S. Lam, and M. Chan, "Effects of layout methods of RFCMOS on noise performance," *Electron Devices, IEEE Transactions on*, vol. 52, no. 12, pp. 2753–2759, 2005.

[131] C.-S. Kim, J.-W. Park, H.-K. Yu, and H. Cho, "Gate layout and bonding pad structure of a RF n-MOSFET for low noise performance," *Electron Device Letters, IEEE*, vol. 21, no. 12, pp. 607–609, 2000.

[132] S. T. Nicolson and S. Voinigescu, "Methodology for simultaneous noise and impedance matching in W-band LNAs," in *Compound Semiconductor Integrated Circuit Symposium, 2006. CSIC 2006. IEEE*, 2006, pp. 279–282.

[133] B. Heydari, M. Bohsali, E. Adabi, and A. Niknejad, "Low-power mm-wave components up to 104GHz in 90nm CMOS," in *Solid-State Circuits Conference, 2007. ISSCC 2007. Digest of Technical Papers. IEEE International*, Feb. 2007, pp. 200–597.

[134] C. Enz, "An MOS transistor model for RF IC design valid in all regions of operation," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 50, no. 1, pp. 342–359, Jan 2002.

[135] B. Razavi, R.-H. Yan, and K. Lee, "Impact of distributed gate resistance on the performance of MOS devices," *Circuits and Systems I: Fundamental Theory and Applications, IEEE Transactions on*, vol. 41, no. 11, pp. 750–754, 1994.

[136] G. González, *Microwave transistor amplifiers: analysis and design.* Prentice Hall, 1997.

[137] D. Shaeffer and T. Lee, "A 1.5-V, 1.5-GHz CMOS low noise amplifier," *Solid-State Circuits, IEEE Journal of*, vol. 32, no. 5, pp. 745–759, 1997.

[138] R. Benton, M. Nijjar, C. Woo, A. Podell, G. Horvath, E. Wilson, and S. Mitchell, "GaAs MMICs for an integrated GPS front-end," in *Gallium Arsenide Integrated Circuit (GaAs IC) Symposium, 1992. Technical Digest 1992., 14th Annual IEEE*, Oct. 1992, pp. 123–126.

[139] N.-H. Sheng, W.-J. Ho, N.-L. Wang, R. Pierson, P. Asbeck, and W. L. Edwards, "A 30 GHz bandwidth AlGaAs-GaAs HBT direct-coupled feedback amplifier," *Microwave and Guided Wave Letters, IEEE*, vol. 1, no. 8, pp. 208–210, 1991.

[140] H.-W. Chiu, S.-S. Lu, and Y.-S. Lin, "A 2.17-dB NF 5-GHz-band monolithic CMOS LNA with 10-mW DC power consumption," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 53, no. 3, pp. 813–824, 2005.

[141] C. Zhang, D. Huang, and D. Lou, "Optimization of cascode CMOS low noise amplifier using inter-stage matching network," in *Electron Devices and Solid-State Circuits, 2003 IEEE Conference on*, Dec. 2003, pp. 465–468.

[142] H. Haus, W. R. Atkinson, G. M. Branch, W. Davenport, W. H. Fonger, W. A. Harris, S. W. Harrison, W. McLeod, E. K. Stodola, and T. E. Talpey, "Representation of noise in linear twoports," *Proceedings of the IRE*, vol. 48, no. 1, pp. 69–74, 1960.

[143] J. C. Rudell, J. jiunn Ou, T. B. Cho, G. Chien, F. Brianti, J. A. Weldon, P. R. Gray, J. Rudell, C. Sitea, J. Ou, J. Cho, T. Chien, G. Brianti, and F. Weldon, "A 1.9-GHz wide-band IF double conversion CMOS receiver for cordless telephone applications," *IEEE Journal of Solid-State Circuits*, vol. 32, pp. 2071–2088, 1997.

[144] S. Wu and B. Razavi, "A 900-MHz/1.8-GHz CMOS receiver for dual-band applications," *Solid-State Circuits, IEEE Journal of*, vol. 33, no. 12, pp. 2178–2185, 1998.

[145] H. Darabi and A. Abidi, "A 4.5-mW 900-MHz CMOS receiver for wireless paging," *Solid-State Circuits, IEEE Journal of*, vol. 35, no. 8, pp. 1085–1096, 2000.

[146] Y.-C. Ho, M. Biyani, J. Colvin, C. Smithhisler, and K. O, "3 V low noise amplifier implemented using a 0.8 μm CMOS process with three metal layers for 900 MHz operation," *Electronics Letters*, vol. 32, no. 13, pp. 1191–1193, 1996.

[147] P. Leroux, J. Janssens, and M. Steyaert, "A 0.8-dB NF ESD-protected 9-mW CMOS LNA operating at 1.23 GHz [for GPS receiver]," *Solid-State Circuits, IEEE Journal of*, vol. 37, no. 6, pp. 760–765, 2002.

[148] W. Guo and D. Huang, "The noise and linearity optimization for a 1.9-GHz CMOS low noise amplifier," in *ASIC, 2002. Proceedings. 2002 IEEE Asia-Pacific Conference on*, 2002, pp. 253–257.

[149] T. Manku, "Microwave CMOS-device physics and design," *Solid-State Circuits, IEEE Journal of*, vol. 34, no. 3, pp. 277–285, 1999.

[150] D. K. Shaeffer, "The design and implementation of low-power CMOS radio receivers," Ph.D. dissertation, Stanford University, 1998.

[151] Y.-J. Lin, S. Hsu, J.-D. Jin, and C. Chan, "A 3.1–10.6 GHz ultra-wideband CMOS low noise amplifier with current-reused technique," *Microwave and Wireless Components Letters, IEEE*, vol. 17, no. 3, pp. 232–234, 2007.

[152] W.-N. Chen, K.-H. Cheng, and T. Duan-Chang, "A 10GHz with low power and low noise figure amplifier Using 0.18mum CMOS technology," in *Solid-State and Integrated Circuit Technology, 2006. ICSICT '06. 8th International Conference on*, Oct. 2006, pp. 1556–1558.

[153] H. Fouad, K. Sharaf, E. El-Diwany, and H. EL-Hennawy, "An RF CMOS cascode LNA with current reuse and inductive source degeneration," in *Circuits and Systems, 2001. MWSCAS 2001. Proceedings of the 44th IEEE 2001 Midwest Symposium on*, vol. 2, Aug. 2001, pp. 824–828 vol.2.

[154] M. Ben Amor, A. Fakhfakh, H. Mnif, and M. Loulou, "Dual band CMOS LNA design with current reuse topology," in *Design and Test of Integrated Systems in Nanoscale Technology, 2006. DTIS 2006. International Conference on*, Sep. 2006, pp. 57–61.

[155] C.-Y. Cha and S.-G. Lee, "A low power, high gain LNA topology," in *Microwave and Millimeter Wave Technology, 2000, 2nd International Conference on. ICMMT 2000*, Sep. 2000, pp. 420–423.

[156] B. Razavi, "A millimeter-wave cmos heterodyne receiver with on-chip lo and divider," *Solid-State Circuits, IEEE Journal of*, vol. 43, no. 2, pp. 477–485, 2008.

[157] S.-C. Shin, M.-D. Tsai, R.-C. Liu, K.-Y. Lin, and H. Wang, "A 24-GHz 3.9-dB NF low-noise amplifier using 0.18 mu;m CMOS technology," *Microwave and Wireless Components Letters, IEEE*, vol. 15, no. 7, pp. 448–450, 2005.

[158] A. Sayag, S. Levin, D. Regev, D. Zfira, S. Shapira, D. Goren, and D. Ritter, "A 25 GHz 3.3 dB NF low noise amplifier based upon slow wave transmission lines and the 0.18 μm CMOS technology," in *Radio Frequency Integrated Circuits Symposium, 2008. RFIC 2008. IEEE*, Jun. 2008, pp. 373–376.

[159] H.-S. Jhon, I. Song, J. Jeon, H. Jung, M. Koo, B.-G. Park, J. Lee, and H. Shin, "8mW 17/24 GHz dual-band CMOS low-noise amplifier for ISM-band application," *Electronics Letters*, vol. 44, no. 23, pp. 1353–1354, 2008.

[160] C.-Y. Chu, C.-C. Wei, H.-C. Hsu, S.-H. Feng, and W.-S. Feng, "A 24GHz low-power CMOS receiver design," in *Circuits and Systems, 2008. ISCAS 2008. IEEE International Symposium on*, May. 2008, pp. 980–983.

[161] H.-Y. Yang, Y.-S. Lin, and C.-C. Chen, "0.18 μm 21-27 GHz CMOS UWB LNA with 9.3 ± 1.3 dB gain and 103.9 ± 8.1 ps group delay," *Electronics Letters*, vol. 44, no. 17, pp. 1014–1016, 2008.

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

[163] K.-W. Yu, Y.-L. LU, D.-C. Chang, V. Liang, and M. Chang, "K-band low-noise amplifiers using 0.18 μm CMOS technology," *Microwave and Wireless Components Letters, IEEE*, vol. 14, no. 3, pp. 106–108, 2004.

[164] C.-C. Kuo and H. Wang, "A 24-GHz low power low noise amplifier using current reuse and body forward bias techniques in 0.18-μm CMOS technology," in *Microwave Conference Proceedings, 2010. APMC 2010. Asia-Pacific Conference Proceedings*, Dec. 2010, pp. 1509–1512.

[165] Y. Su and K. O, "An 800-μW 26-GHz CMOS tuned amplifier," in *Radio Frequency Integrated Circuits (RFIC) Symposium, 2006 IEEE*, Jun. 2006, pp. 1–4.

[166] E. Adabi, B. Heydari, M. Bohsali, and A. Niknejad, "30 GHz CMOS low noise amplifier," in *Radio Frequency Integrated Circuits (RFIC) Symposium, 2007 IEEE*, Jun. 2007, pp. 625–628.

[167] O. Dupuis, X. Sun, G. Carchon, P. Soussan, M. Ferndahl, S. Decoutere, and W. De Raedt, "24 GHz LNA in 90nm RF-CMOS with high-Q above-IC inductors," in *Solid-State Circuits Conference, 2005. ESSCIRC 2005. Proceedings of the 31st European*, Sep. 2005, pp. 89–92.

[168] M. A. T. Sanduleanu, G. Zhang, and J. Long, "31-34GHz low noise amplifier with on-chip microstrip lines and inter-stage matching in 90-nm baseline CMOS," in *Radio Frequency Integrated Circuits (RFIC) Symposium, 2006 IEEE*, Jun. 2006, pp. 1–4.

[169] L. Aspemyr, H. Jacobsson, M. Bao, H. Sjoland, M. Ferndahl, and G. Carchon, "A 15 GHz and a 20 GHz low noise amplifier in 90 nm RF-CMOS," in *Silicon Monolithic Integrated Circuits in RF Systems, 2006. Digest of Papers. 2006 Topical Meeting on*, Jan. 2006, pp. 1–4.

[170] M. Bao, H. Jacobsson, L. Aspemyr, A. Mercha, and G. Carchon, "A 20 GHz sub-1V low noise amplifier and a resistive mixer in 90 nm CMOS technology," in *Microwave Conference Proceedings, 2005. APMC 2005. Asia-Pacific Conference Proceedings*, vol. 5, Dec. 2005, pp. 1–4.

[171] W.-C. Wang, Z.-D. Huang, G. Carchon, A. Mercha, S. Decoutere, W. De Raedt, and P. Chung-Yu Wu, "45-nm Planar bulk-CMOS 23-GHz LNAs with high-Q above-IC inductors," in *Circuits and Systems (ISCAS), Proceedings of 2010 IEEE International Symposium on*, May. 2010, pp. 741–744.

[172] H. Shigematsu, T. Hirose, F. Brewer, and M. Rodwell, "Millimeter-wave CMOS circuit design," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 53, no. 2, pp. 472–477, 2005.

[173] J.-H. Tsai, W.-C. Chen, T.-P. Wang, T.-W. Huang, and H. Wang, "A miniature Q-band low noise amplifier using 0.13-μm CMOS technology," *Microwave and Wireless Components Letters, IEEE*, vol. 16, no. 6, pp. 327–329, 2006.

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

[175] C.-M. Lo, C.-S. Lin, and H. Wang, "A Miniature V-band 3-Stage Cascode LNA in 0.13μm CMOS," in *Solid-State Circuits Conference, 2006. ISSCC 2006. Digest of Technical Papers. IEEE International*, Feb. 2006, pp. 1254–1263.

[176] F. Ellinger, "26-42 GHz SOI CMOS low noise amplifier," *Solid-State Circuits, IEEE Journal of*, vol. 39, no. 3, pp. 522–528, 2004.

[177] T. Yao, M. Gordon, K. K. W. Tang, K. H. K. Yau, M.-T. Yang, P. Schvan, and S. Voinigescu, "Algorithmic design of CMOS LNAs and PAs for 60-GHz radio," *Solid-State Circuits, IEEE Journal of*, vol. 42, no. 5, pp. 1044–1057, 2007.

[178] M. Masud, H. Zirath, M. Ferndahl, and H. Vickes, "90 nm CMOS MMIC amplifier," in *Radio Frequency Integrated Circuits (RFIC) Symposium, 2004. Digest of Papers. 2004 IEEE*, Jun. 2004, pp. 201–204.

[179] S. Voinigescu, M. Maliepaard, J. Showell, G. Babcock, D. Marchesan, M. Schroter, P. Schvan, and D. Harame, "A scalable high-frequency noise model for bipolar transistors with application to optimal transistor sizing for low-noise amplifier design," *Solid-State Circuits, IEEE Journal of*, vol. 32, no. 9, pp. 1430–1439, 1997.

[180] S.-H. Lee, Y.-Z. Juang, C.-F. Chiu, and H.-K. Chiou, "A novel low noise design method for CMOS L-degeneration cascoded LNA," in *Circuits and Systems, 2004. Proceedings. The 2004 IEEE Asia-Pacific Conference on*, vol. 1, Dec. 2004, pp. 273–276.

[181] M. Egels, J. Gaubert, P. Pannier, and S. Bourdel, "Design method for fully integrated CMOS RF LNA," *Electronics Letters*, vol. 40, no. 24, pp. 1513–1514, 2004.

[182] R. Rafla and M. El-Gamal, "Design of a 1.5 V CMOS integrated 3 GHz LNA," in *Circuits and Systems, 1999. ISCAS '99. Proceedings of the 1999 IEEE International Symposium on*, vol. 2, Jul. 1999, pp. 440–443.

[183] S. Nicolson and S. Voinigescu, "Methodology for simultaneous noise and impedance matching in W-band LNAs," in *Compound Semiconductor Integrated Circuit Symposium, 2006. CSIC 2006. IEEE*, Nov. 2006, pp. 279–282.