INVESTIGATION OF 60 GHZ RADIO FRONT-ENDS IN NANOMETER CMOS

by

Sha Tao

M. Sc., KTH/ICT, 2009

A THESIS REPORT

submitted in partial fulfillment of the requirements for the degree

MASTER OF SCIENCE

SCHOOL OF INFORMATION AND COMMUNICATION TECHNOLOGY (ICT)
ROYAL INSTITUTE OF TECHNOLOGY (KTH)
STOCKHOLM, SWEDEN

October 2009
Abstract

In the past few years, silicon mm-wave, especially 60GHz CMOS design has experienced a transition from an obscure topic to a research hot spot. The focus of this thesis is the design of a 60GHz receiver front-end integrated circuit, together with device modeling solutions, using 65nm CMOS technology.

A 60GHz to 5GHz heterodyne receiver topology is initially architected to exploit its possible compatibility with the 5GHz legacy WLAN system. In order to implement this frontend, an EM simulation based device modeling methodology together with the corresponding design flow has been proposed, which is tailored for the specific 65nm CMOS design kits and the available simulation tool. Based on thorough analysis of the process feature, efforts on device modeling for 60GHz operation have been taken. For active device, an EM model, using exiting transistor compact model as core, is developed for the NFET valid in vicinity of 60GHz to account for parasitic elements due to wiring stacks. Solutions for implementing the passive components in the specific circuit blocks have been illustrated. In particular, constructing, optimizing and physically characterizing of spiral inductor operating around 60GHz frequency band has been demonstrated. After the modeling efforts, a single-stage cascode LNA and a single-gate transconductance-pumped mixer are individually designed in the IBM 65nm CMOS process, characterized by EM co-simulation, and then compared with the state-of-the-art. Finally, the LNA and mixer has been integrated, layout and simulated as a complete front-end. The frontend achieves a conversion gain of 11.9dB and an overall SSB noise figure of 8.2dB, with an input return loss of -13.7dB. It consumes 6.11mW power, and its layout occupies a die area of 0.33×0.44mm².
Acknowledgements

It was in this royal, old and modern institute that I turned from a girl into a young woman. I was taught vetenskap there, and konst; there, I experienced an amazing and peaceful time of my life…

Eventually, the two-year-period gets closure; it terminates with this thesis work. I would always feel fortunate enough to have decided to do my master thesis here, in RaMSiS. My gratitude to Prof. Mohammed Ismail, who initiated this 60GHz project so that I got this thesis topic, for his kindly suggestions and encouraging me to draft my first publication. Also, my gratitude to Dr. Ana Rusu, who has impressed me since the ASIC course, who has been an admirable instructor in several courses during my two years’ master study, for her welcoming me as a member in the group and help me to keep in the right track all the time. My best gratitude towards Saul Rodriguez, my mentor, who rescues me from deviated direction and taught me right from wrong, who shows me desirable attitude and manner for doing research; whose door has always been open to me when I feel depressed, anxious and desperate, whose gentleness, patience and sense of humor help me relax when facing difficulties; who is always caring about me in daily life and inviting me to interesting activities, who treats me just like a big brother. Last but not least, my appreciation to other group members - Hans Rabén, Julian Garcia, Samantha Yoder and Vasileios Manolopoulos – it was such a nice experience to work together and share ideas with you guys. Especially Hans, you did the similar topic as I do. During the time, we have had plenty of discussions about our works as well as other things of interest and we have been the defence opponent for each other. It was so nice to have you here working with me.

Upon graduation, I would like to mention the special ones in my life during the two years. All the SoC 2007 students, especially my boyfriend Mao Jia and my roommate Xu Ye, thanks for being with me here, in this fantastic city, Stockholm. My family and friends back in sweet home, Beijing, especially my mom and dad, though not physically being with me, but you are always my source of strength and spiritual pillar, all the time. I love all of you.
# Table of Contents

Abstract ........................................................................................................................................ iii

Acknowledgement .......................................................................................................................... iv

Table of Contents ............................................................................................................................. v

List of Figures ................................................................................................................................ vii

List of Tables .................................................................................................................................. ix

Chapter 1 Introduction ..................................................................................................................... 1

1.1 Background and Application ................................................................................................ 1

1.2 Problems and Challenges ...................................................................................................... 2

1.3 Scope and Objectives ............................................................................................................ 2

1.4 Thesis Report Outline ........................................................................................................... 3

Chapter 2 Receiver Architecture and Process Technology .............................................................. 4

2.1 Front End Architecture ......................................................................................................... 4

2.1.1 Existing 60GHz Receiver Architectures ....................................................................... 4

2.1.2 The Proposed Front-end Topology ............................................................................... 4

2.2 IBM 65nm CMOS Technology ............................................................................................ 6

2.2.1 Back-End-Of-Line ....................................................................................................... 6

2.2.2 Active Components ...................................................................................................... 7

2.2.3 Passive Components ..................................................................................................... 8

Chapter 3 Design Flow and Simulation Approach ....................................................................... 11

3.1 Simulation Methodology .............................................................................................. 11

3.1.1 Electromagnetic (EM) Simulation .............................................................................. 11

3.1.2 The Overall Simulation Approach .............................................................................. 12

3.2 Design Flow Description .................................................................................................... 13

3.2.1 The Systematic Procedure Using “RFIC Dynamic Link” .......................................... 13

3.2.2 The Simplified Design Flow in Practice ..................................................................... 14

Chapter 4 Device Modeling ........................................................................................................... 15

4.1 Modeling Strategies ............................................................................................................ 15

4.1.1 The Main Stream Modeling Methods ......................................................................... 15

4.1.2 The Proposed Modeling Method ................................................................................ 17

4.2 Active Device Modeling ..................................................................................................... 21

4.2.1 Review of PDK Offering ............................................................................................ 21
7.2 Future Work........................................................................................................................................66

Bibliography ...........................................................................................................................................67

Appendix A BEOL Set-up for EM Simulator .......................................................................................73
   A.1 Substrate Stack Setup for ADS Momentum/EMDS .................................................................73
   A.2 Layer Metallization for ADS Momentum/EMDS .................................................................74
   A.3 IBM 65nm BEOL Stacks Set-up File ..................................................................................75

Appendix B Schematics and Test-benches .........................................................................................76
   B.1 B.1 Rx Frontend Test-bench .........................................................................................76
List of Figures

Figure 1.1. 5/60 GHz system scenario in home environment .......................................................... 1
Figure 2.1. 60 GHz receiver block diagram .................................................................................. 5
Figure 2.2. BEOL of IBM 65nm CMOS technology ..................................................................... 6
Figure 2.3. Cross-section of the triple-well NFET ....................................................................... 8
Figure 2.4. MIM capacitor cross-section view ............................................................................... 10
Figure 3.1. Overall simulation approach ...................................................................................... 12
Figure 3.2. “RFIC Dynamic Link” design flow chart ................................................................. 13
Figure 3.3. The practical design flow in use .................................................................................. 14
Figure 4.1. Plots of L, Q and S11 for a PDK spiral inductor ....................................................... 19
Figure 4.2. Single level, parallel, and series stacked spirals ......................................................... 19
Figure 4.3. Optimization procedure for inductors ....................................................................... 20
Figure 4.4. Sub-circuit schematic of RF FET model ................................................................... 21
Figure 4.5. A NFET layout with wiring stacks ............................................................................ 22
Figure 4.6. 3D view of the V1-M5 stack in transistor layout ....................................................... 23
Figure 4.7. 3D view of the JT-LD stack in transistor layout ......................................................... 23
Figure 4.8. 3D view demonstration of the “via simplification” .................................................... 24
Figure 4.9. Comparison result as validation for the “via simplification” ........................................ 24
Figure 4.10. Schematic of capacitance extraction for RF NFET model ....................................... 25
Figure 4.11. The RF NFET model capacitive effects ................................................................. 25
Figure 4.12. 3D view of wiring stacks for the NFET ................................................................. 26
Figure 4.13. Capacitances extracted from the NFET wiring stacks ........................................... 26
Figure 4.14. Capacitive components for the new NFET model ................................................. 27
Figure 4.15. Comparison of S-parameters for a PDK inductor ................................................... 29
Figure 4.16. 3D view of the single-level spiral ........................................................................... 29
Figure 4.17. The two-port configuration to extract inductance ................................................... 30
Figure 4.18. S-parameters of the spiral ....................................................................................... 31
Figure 4.19. Inductance and quality value for the spiral ............................................................. 31
Figure 4.20. Cross-section of an inductor and its equivalent circuit .......................................... 32
Figure 4.21. Comparison of S-parameters for the original spiral with the pi model ................. 33
Figure 4.22. A more accurate lumped model ............................................................................. 34
Figure 4.23. Comparison of S-parameters for the spiral ............................................................ 34
Figure 5.1. Design flow of frontend circuit blocks ................................................................. 38
Figure 5.2. Transconductance versus drain and source bias voltages .............................. 40
Figure 5.3. Current voltage characteristics of the NFET .................................................. 40
Figure 5.4. Simulated $\text{NF}_\text{min}$ as a function of $W_g$ and $N_f$ at 60 GHz .................. 41
Figure 5.5. Current gain of the 65nm CMOS RF NFET ...................................................... 42
Figure 5.6. Power gain of the 65nm CMOS RF NFET ......................................................... 43
Figure 5.7. Circuit schematic of a single-stage cascode LNA .............................................. 44
Figure 5.8. Return losses at input and output ports .............................................................. 46
Figure 5.9. S21 and noise figure .......................................................................................... 47
Figure 5.10. Stability factor and maximum available gain .................................................. 47
Figure 5.11. Estimated input 1-dB compression point ....................................................... 47
Figure 5.12. Capacitive components for the new NFET model ......................................... 48
Figure 5.13. Extrapolation of the third-order intercept point ............................................ 48
Figure 5.14. Layout of the LNA including pads and power rails ....................................... 49
Figure 5.15. Simplified schematic of $g_m$ mixers .............................................................. 50
Figure 5.16. S-parameters of the spiral ............................................................................... 51
Figure 5.17. Current-voltage characterization of NFET ..................................................... 52
Figure 5.18. NFET characterization of drain current vs. gate-source voltage .................. 52
Figure 5.19. NFET characterization of $g_m$ versus drain and source bias voltages ........ 53
Figure 5.20. Conversion-gain sweeping as function of $V_{gs}$ and $W_g$ ............................... 53
Figure 5.21. Circuit schematic of a transconductance mixer ............................................. 54
Figure 5.22. Impedance transformation for RF, IF and LO ports ...................................... 55
Figure 5.23. Transient amplitude at RF input and IF output ............................................ 55
Figure 5.24. Return loss for RF, LO and IF ports ............................................................... 56
Figure 5.25. Conversion gain vs. LO power ...................................................................... 56
Figure 5.26. SSB and DSB NF vs. the LO power ............................................................... 57
Figure 5.27. RF input power versus IF output power to capture ICP ................................. 57
Figure 5.28. RF input power versus IF output power to capture IIP3 ................................ 57
Figure 5.29. Layout of the mixer including pads and power rails .................................... 58
Figure 5.30. Mixer EM Co-simulation Results ................................................................. 59
Figure 6.1. Schematic of the 60 GHz front-end ................................................................. 61
Figure 6.2. Front-end simulation results ........................................................................... 62
Figure 6.3. Layout of the integrated receiver frontend ...................................................... 64
List of Tables

Table 4.1. Formulas for RLC parameters ................................................................. 33
Table 5.1. Performance summary of LNA with ideal passives................................ 45
Table 5.2. Performance comparison of CMOS LNAs around 60 GHz ..................... 50
Table 5.3. Performance comparison of CMOS downconversion mixers at 60 GHz .......... 60
Table 6.1. Performance summary of the 60 GHz front-end...................................... 65
Chapter 1

Introduction

1.1 Background and Application

The ever-increasing demand for higher wireless transfer capacity tends to saturate the low gigahertz bands of current in use wireless standards such as WiFi (Wireless Fidelity), UWB (Ultra-Wideband) and Bluetooth etc. One commonly accepted solution to this over congestion problem is to resort to the 60GHz band where bandwidth is abundantly available [1] [2].

The IEEE 802.15.3 group has investigated 7 GHz band spectrum around 60 GHz as an alternate physical layer to enable very high-data-rate applications such as high-speed internet access, streaming content downloads, and wireless data bus for cable replacement [3].

With the availability of this unlicensed wide band, there is growing interest in using this resource for new consumer applications, which require very high-data-rate and short-range wireless transmission, especially dense wireless local communications [1] [3].

Figure 1.1 [4] is an example of exploiting the 60 GHz band for WPAN (Wireless Personal Area Networks) application. This 60 GHz system operates in combination with 5 GHz system to achieve interoperability with legacy WLAN (Wireless Local Area Network) systems. Moreover, it can be used as a fallback option in case the received power at 60 GHz becomes insufficient.

Figure 1.1: 5/60 GHz system scenario in home environment [4]
1.2 Problems and Challenges

As stated above, the wireless system, operating at 60GHz is indeed a compelling candidate to deliver multi-gigabit speed in short-range. However, there is no free lunch in the world.

Wireless communication at 60 GHz suffers from several challenges, which include limited range due to large attenuation, substantial drop of receiver power due to antenna obstruction, several Doppler effects, and small channel dispersion [5].

Moreover, when it comes to the inexpensive and high-level integrated CMOS process, especially compared with its III-V semiconductor technologies counterparts, further issues would come along with [2]: CMOS has greater process variability, lower carrier mobility constants and smaller device breakdown voltages.

During the whole design procedure, three fundamental differences between 60 GHz design and low gigahertz design should always be account for [6]: 1) using transistors operating much closer to their cutoff frequencies, 2) operating with small wavelength signals, and 3) designing with parasitic elements that represent larger portion of impedance or admittance.

1.3 Scope and Objectives

This thesis work is one part of the research project MoDeM60 (Modeling and Design of Manufacturable 60 GHz RF CMOS Radios) currently conducted in RaMSiS group at ICT/KTH.

The ultimate purpose of this MoDeM60 project is to build a manufacturable 60 GHz radio for high volume low cost applications. Two main tasks are entailed in project [5]: firstly designing the RF front-ends to meet typical performance requirements and secondly introducing digital self-calibration techniques to overcome performance degradation.

The main objectives of this thesis are: 1) survey study of state-of-the-art 60 GHz radio design, 2) modeling of the active and passive devices to support 60 GHz, 3) set-up all design environment for EDA tools and 4) design of the blocks (LNA and mixer).
1.4 Thesis Report Outline

This master thesis final report consists of seven chapters.

Chapter 2 firstly presents a comparison of existing 60 GHz receiver collections and it followed by the proposed radio front-end topology. After that, descriptions of the relevant process features of the technology in use are given.

Chapter 3 addresses the general methodology adopted in this thesis work. An EM simulation approach is illustrated. A corresponding design flows is also briefly introduced.

Chapter 4 is one of the key sections in this thesis. It focuses on the device modeling issues required by the implementation of the front-end circuit-blocks. Beginning with the discussion of strategies and methods, modeling and implementation details for both active and passive components are then illustrated separately in terms of these aspects – the available models from design kits, the modeling strategies in accordance, the operation procedures and the results.

Chapter 5, another major section, demonstrates the design and implementation of the Rx sub-blocks. Initially, it addresses state-of-the-art and performance metrics of the 60 GHz low noise amplifier and down-conversion mixer designs. The chapter continues with elaborations on design, implementation, and characterization of a single-stage cascode LNA and a transconductance-pumped mixer in IBM 65nm CMOS.

Chapter 6 deals with the integration of the front-end sub-blocks, the LNA and mixer in this case. Re-design considerations for those blocks are described and then the whole Rx circuit schematic is given. The layout and simulation results of the integrated front-end are presented afterwards.

Chapter 7 provides a summary of this thesis assignment together with the remaining tasks and works to be further carried out in the future.
Chapter 2 Receiver Architecture and Process Technology

2.1 Front End Architecture

2.1.1 Existing 60GHz Receiver Architectures

With respect to the choice of the 60 GHz front-end radio architecture, there are, in general, two options: the homodyne (i.e., “direct conversion” or “zero IF”) and the heterodyne (including low IF, half-RF etc.). It would be hasty to tell which one works better at 60 GHz, before bringing them into comparison.

The advantage of homodyne is that it is uniquely well suited for monolithic integration, since it does not require image filtering, and it is intrinsically simple. However, this architecture has the disadvantages of intrinsic sensitivity to DC offset and LO leakage back to the antenna [1]. Several homodyne examples for 60 GHz application can be seen in reported works [7] [8] [9] [10] [11].

Compared with the direct-conversion type, the heterodyne (i.e., “super heterodyne”) receiver although suffers from the problem of image and half IF, proves to help with solving the LO problems of mm-wave transceiver designs. Plenty previous works on heterodyne type 60 GHz receivers have been reported in [10] [11] [12] [13] [14] [15].

2.1.2 The Proposed Front-end Topology

Despite of its intrinsic simplicity, the fact that homodyne architecture has not yet been massively adopted is not for no reason. The identical RF and LO frequency facilitates the LO leakage to mixer input, which would lead to self-mixing and consequently the DC offset at mixer output. This DC offset problem would significantly degrade overall performance of the receiver.

As regards the super-heterodyne topology, which was invented by Armstrong early in 1917, it is still of widespread use in modern wireless radio systems. In 60 GHz operation frequency range, converting the RF signal to an intermediate frequency will not only solve the DC offset of direct conversion but also alleviate the difficulty of generating local oscillation frequency to some extent. Moreover, for system application point of view, the 60-GHz system has the potential to combine the WLAN 5 GHz radio systems, so that it would be desirable to take an IF at 5 GHz
from this 60 GHz receiver. The idea is to make the existing IEEE 802.11 5-GHz RF serve as IF in 60-GHz system to construct a new dual mode operation.

Hence, a heterodyne 60 GHz Rx front-end is proposed to exploit its possible compatibility and interoperability with the 5 GHz WLAN system. The simplified block diagram of this 60 GHz receiver is depicted in Figure 2.1.

![Figure 2.1: 60 GHz receiver block diagram](image)

The 60 GHz front-end consists of a single-ended LNA and a transconductance which down-convert the signal to a 5 GHz intermediate frequency. The mixer is pumped by an external generated 55 GHz LO signal and it is followed by an IF buffer (which can be a simple cascade amplifier) to drive the 50 Ω load. In this particular application, neither IF amplifier nor IF I/Q mixers are required as succeeding stages for the reason that this IF signal is supposed to cooperate with the other 5 GHz applications instead of to be processed in baseband.

One of the merits of the proposed architecture is that it avoids the generation of differential LO signals for the mixer LO input, which is singly balanced. It exempts the difficulty of providing quadrature LOs as well. It is of relative simplicity and mitigates self-mixing of direct conversion. Furthermore, it fits the application of the 5/60 GHz dual-mode operating system illustrated in [4].

In the front-end architecture, both RF amplification stage (LNA) and down-conversion stage (mixer) are the most critical factors to affect the Rx noise and linearity performances and consequently the dynamic range of the whole system. However, they are also most challenging to implement. In this thesis work, those two blocks are designed and implemented in 65nm CMOS.
2.2 IBM 65nm CMOS Technology

The front-end circuits have been implemented in the Low Power and RF edition of IBM 65nm CMOS semiconductor process - CMS10LP/RFe [16]. It features a supply voltage of 1.2 V (thion oxide), four to nine copper metal levels (according to different wiring options), twin-well or triple-well CMOS technology on p-substrate, and minimum drawn gate length of 0.060μm etc.

According to state-of-the-art 60GHz designs, the 65nm technologies generally offer better performance than the earlier process nodes, such as 90nm and 130nm technologies. This design demands involving significant device modeling works, thus, it is of great importance to get sufficient knowledge of the process and exploit its features during the whole design course.

2.2.1 Back-End-Of-Line

In 60-GHz radio design (like all the other millimeter wave designs), information coming only from the active device is not sufficient. For frequencies above 10 GHz, each micrometer of back-end strip has a significant influence on the electrical behavior of the circuit [6]. Therefore, the BEOL (Back End of Line) information of technology should be well investigated before device modeling and components implementation for the circuit.

Figure 2.2 shows the BEOL of one particular stack option (8-Metal Analog Metal Stack) of the IBM CMOS 65nm technology used in this thesis work [17].

![Figure 2.2: BEOL of IBM 65nm CMOS technology [17]](attachment:image.png)
From the design manual [16], the detailed electrical and geometrical information of the metal stripes and dielectric layers can be obtained. Hence, the characteristics of this 65nm technology comparing with the previous CMOS nodes could be observed (also referring to the book chapter of [6]). The most significant one is the shrink of the vertical thickness of the metal and dielectric layers, which would certainly lead to increased integration density and increased influence of the substrate losses on the propagation constant. Accordingly, the dielectric oxide permittivity gets the trends of decreasing in order to limit the coupling effects between two conducting layers. The small metal line pitch together with the thin dielectric layers would also induce large substrate loss for the inductances. To cope with the loss, thick metal (e.g. 12x) and dielectric layers are used in the top levels.

In order to get accurate models for passive devices especially for the more complex inductances, EM (electro-magnetic) simulation is required to predict the effects of metal losses and parasitics. Therefore, the set-up for the technology BEOL in the EM simulator is the premise of modeling accuracy. Appendix A records the substrate-stack and layer-metallization set-up details for ADS momentum and EMDS.

2.2.2 Active Components

Transistors with several different electrical models are the only active components used in the circuits. Among the available Field-Effect Transistors (FET) provided by Physical-Design-Kits (PDK), two models are discussed in this section with special interests.

RF Transistor

Except the general purpose MOSFETs modeled by BSIM4.5, the PDK provides a cluster of RF FET models to support high frequency operation. They are built as sub-circuits, where the basic MOSFET model is wrapped by the parasitic gate resistance and wiring capacitances.

The intrinsic gate resistance, non-quasi-static (NQS) effect, and substrate resistance effect are enabled using the BSIM4 built-in gate resistance model and substrate resistance model whose parameters are adjusted to fit to the hardware data [20]. An issue needs special notice is that the accuracy is closely linked to the layout style, in other words, a PCELL (parameterized cell) compatible design will ensure the model accuracy. The preferred layout styles are [21]: gate fully strapped on one side or two sides through PC, CA and M1, double row CAs on gate stripes, single PC pitch, substrate RX ring fully strapped with CA and M1. Furthermore, the model-to-hardware correlations with two bias conditions and different device dimensions are given in the frequency range from 0.2-110GHz, which would certainly help with the modeling work a lot later.
**Triple-Well Transistor**

As an optional device, triple-well NFETs are available for the regular and RF mode NMOS transistors. Illustrated in Figure 2.3 [24], the triple-well devices provide FETs within a p-well that is isolated from the substrate. The isolation is accomplished by inserting a buried n-type layer, which is designated by the PI mask level between local p-well and p-substrate. A ring of n-well provides lateral isolation and connects to the PI region, which should be tied to a quiet power supply that is at a high potential to prevent forward biasing the PI/substrate or PI/p-well junctions.

![Cross-section of the triple-well NFET](image)

The obvious advancements of such a device model are the possibility to make multiple NFETs in a single well and providing noise/voltage bias isolation from the p-substrate, which would facilitate the circuit design.

### 2.2.3 Passive Components

The passive elements play a key role in RF and mm-wave CMOS design and often become circuit performance bottlenecks, which in most cases owe to the limited quality factor. Three reasons, generally, explain the limited Q-values of the passives [25]: substrate losses, metal losses and parasitic substrate capacitance. We can distinguish between distributed elements such as transmission lines, and lumped elements including inductors, capacitors and resistors [26].

**Inductors**

The design-kits contains three different inductor configurations: the symmetric parallel-stacked spiral inductor (symindp), the standard parallel-stacked spiral inductor (indp) and the standard series-stacked spiral inductor (inds). Each of these configurations can have either a low impedance Faraday shield groundplane or a high impedance P-substrate groundplane.
The configurations for the symindp consists of a parallel combination of the 1X, 2X, 4X, and 12X metal layers with the option of adding aluminum in parallel [21]. It is mainly used for differential circuits, which will not be adopted in our design, so that the rest two solutions would be laid more emphasis on.

The indp is offered for all BEOL stacks being supported. The vertical cross section of the parallel stacked inductor consists of a metal spiral at the top level of metal connected in parallel, through one or more bar vias, with identical spirals at the thick copper metal levels below it. The resultant low effective sheet resistance of the spiral helps to improve the peak Q value. As an alternative, the last aluminum level can be used in parallel with the copper levels, to reduce the spiral resistance [19].

The inds is offered for those BEOL stacks consisting of a thick copper inductor level and a thick aluminum level. The vertical cross section of the series stacked spiral inductor consists of a spiral at the thick aluminum level connected in series with a similarly wound spiral at the thick metal inductor level directly underneath it. This makes this inductor possible to achieve higher inductance per unit area than other inductor offerings [19].

**T-lines**

The IBM 65nm technology even provides several transmission line models, including single and coupled microstrip (singlewire&coupledwires), single and coupled CPW (coplanar waveguide) (singlecpw&coupledcpw), and single wire inductor line (rfline).

Take the singlecpw as an example, which is commonly used to model interconnects and critical wires. This device consists of a metal signal line between two coplanar ground wires above the silicon substrate. The model is implemented as a multi-segment RLC filter network [19]. The inductance and the resistance per unit length both depend on frequency due to skin and proximity effects. Dielectric losses in the oxide layer are negligible and are therefore neglected. The losses due to possible currents in the silicon substrate are incorporated into model. The model is valid in the bandwidth from DC until 100GHz [21].

**Capacitors**

CMOS10LP/RFe provides two capacitor options: the commonly used high-precision metal-insulator-metal capacitor (mimcap) and the vertical natural capacitor (vncap) that is constructed from fingers of metal wires and vias.
The provided MIM capacitor depicted in Figure 2.4 [19] is formed by adding two masks, QT and HT, between the last copper metal and terminal aluminum layer LB or LD. These layers are QT (bottom plate) and HT (top plate). Both plates of the capacitor are connected to the terminal aluminum metal level through aluminum VV vias. In [21], the model-to-hardware correlation for capacitance and Q value with varied areas are given in the frequency range from DC to 50GHz.

![MIM capacitor cross-section view](image)

Figure 2.4: MIM capacitor cross-section view [19]

The vertical natural capacitor is formed by creating a set of interdigitated fingers on any number of contiguous lower metal levels. Fingers on separate metal levels are connected by vias. It has been specially noted in model guide [21] that a square vncap often provides the best trade-off between high capacitance values and high Q. As the length of the fingers increases, the series resistance also increases and leads to a decrease in the device Q. This will also result in a higher value of inductance and, thus, a lower self-resonant frequency.

**Resistors**

The design-kits supports four kinds of resistors: the OP N+ diffusion resistor (opndres), the OP P+ Poly resistor (opppres), the N-Well resistor (nwres) and the RP P+ resistor (oprppres) with resistive values of 635ohm/sq, 564ohm/sq, 115ohm/sq and 370ohm/sq respectively. Generally, since resistor values can differ slightly from different wafer locations, it is good practice to use identical resistor values and types for implementation of e.g. transistor biasing in circuit [25].
Chapter 3  Design Flow and Simulation Approach

3.1 Simulation Methodology

The simulation method of the 60 GHz front-end circuit would certainly give significant difference to the traditional low gigahertz RF design flow. It requires extra mm-wave device modeling efforts before the circuit design and asks for far more sophisticated parasitic extraction operation after the layout. Those two demands inevitably involve the electromagnetic simulation.

3.1.1 Electromagnetic (EM) Simulation

There are several technical approaches to EM simulation, among which two methods are mostly adopted: Method-Of-Moments (MoM) and Finite-Element-Method (FEM). Each one of them is aligned with specific applications and instantiated by some commercial EM simulation tools accordingly.

In particular, MoM involves careful evaluation of Green’s functions. Maxwell’s equations are then transformed into integral equations, which yield the coupling matrix equation of the structure. While FEM is based on volumetric meshing, in which the full problem space is divided into thousands of smaller regions and represents the field in each sub-region with a local function [40].

In the state-of-the-art commercial EM software, 2.5D ADS Momentum is the representative of MoM while FEM has the 3D Ansoft HFSS as the counterpart. In spite of the advertising claims from those companies, there is no perfect numerical technique that is most efficient and accurate in every possible situation. Hence, the choice of the approach consequently the simulation tool should depend on the different modeling and characterization practice.

For modeling on-chip inductances, the method-of-moments based simulation tools offer significant advantages over the other techniques. More particularly in terms of simulation time and computer requirements, since only field quantities on metal surfaces are introduced as unknowns in the MoM formulation. Therefore, in this thesis work, Momentum is used to characterize the inductors and transmission lines whose accuracy has been proved by many former designs [41].

For extracting transistors wiring models, the finite-element-method would be more preferable with respect to the general 3D nature of the structure. However, when using the stand-alone FEM field solver such as HFSS in a circuit design, the typical practice is to draw structures, run EM
simulations with HFSS, and then bring the S-parameter data back to circuit design environment. Hence, if there is a way to integrate 3D EM solver into the circuit design environment, design time would be reduced in terms of many eliminations on design steps, such as layout data conversion, import and export process etc. The 3D FEM field solver EMDS that is integrated in ADS circuit design environment is chosen to model the transistor wiring parasitic in this work.

3.1.2 The Overall Simulation Approach

With respect to the simulation after layout, the traditional parasitic extraction method (PEX) is inaccurate at such a high frequency. It is impractical to put the whole layout into EM simulator. A new approach is developed here to cope with the co-simulation problem, in order to achieve the balance between better accuracy of parasitic prediction and less consumed in time in simulation.

The simulation approach mainly consists of three separated parts (depicted in Figure 3.1). One is Momentum simulation (a), which includes inductive elements such as spiral inductances, waveguides and critical inter-connections. This simulation results in S-parameter models, which are used in the later co-simulation stage. Another is EMDS simulation (b), aiming to extract S-parameter model or lumped model for wirings of the transistor from the second metal level to the top metal. This wiring-structure model is then combined with the original model of the MOS transistor. The other components are implemented directly using the models from design kits (c). Finally, in ADS, a co-simulation is performed using the models from the above three sources.

![Figure 3.1: Overall simulation approach](image)
3.2 Design Flow Description

Same as the situation in simulation approach, this 60 GHz radio design should have a unique design flow instead of the traditional RFIC design flow. In this section, a systematic design procedure using the “RFIC Dynamic Link” is firstly proposed. It then followed by an improved and simplified design flow, which would be utilized in design practice.

3.2.1 The Systematic Procedure Using “RFIC Dynamic Link”

Based on the fact that the selected technology only provide PDK for Cadence meanwhile it is indispensible for mm-wave design to involve components modeling and electromagnetic test which commonly done by ADS, the design flow in this case is then chosen as a combination of these two tools to take advantage of the strengths and capability of both design environment.

Because of the desire to use multiple tools, Agilent Technologies has developed the “RFIC Dynamic Link”, which enables both top-down and bottom up design and simulation in ADS using IC designs from the Cadence database. There is a design flow chart provided by Agilent for reference as shown in Figure 3.2 [42].

![Figure 3.2: “RFIC Dynamic Link” design flow chart [42]](image)

Based on the provided “RFIC Dynamic Link” chart, a formal design procedure is then proposed for the 60GHz Rx design as the following steps with some iteration.

1. To check the validity of the substrate stack set-up, import the layouts of the selected components from Cadence Virtuoso Physical Design environment to ADS momentum.
2. In Momentum and EMDS, model the devices that are going to be used in the circuit design.

3. Use Cadence Virtuoso to capture the spectre model, which can be directly used, and the circuit schematic is then instantiated as a cell view in ADS through the Dynamic Link.

4. Use the Cadence sub-circuits together with the modeled devices to complete the circuit and perform the required simulation and optimization in ADS.

5. After the pre-simulation, layout is then carried out in Cadence Virtuoso. One part generated from the Virtuoso schematics and the other part imported from ADS momentum layout.

6. Back in momentum, selected portions of the layout are checked with EM simulation.

7. Next, the physical verification is carried out in Cadence Assura, including DRC, LVS and PE;

8. Finally, the post-layout simulation would be done in ADS considering the extracted parameters. If result meets the specs, the GDSII file export from the Cadence is ready for tape out.

### 3.2.2 The Simplified Design Flow in Practice

It happens that the ADS’s built-in netlist-translation feather enables the indirect use of spectre models from PDK. Consequently, the above design steps can be modified into a simplified design flow (illustrated in Figure 3.3), which would keep the design phases mainly in ADS.

![Diagram](image-url)

**Notes:**

1. Require newly generated models for the circuit reality and passive achievability;
2. Adjust the circuit components in terms of better floor-plan;
3. Modify the design account for the co-simulation performance.

Figure 3.3: The practical design flow in use
Chapter 4  Device Modeling

As the performance margin of CMOS transistors shrinking dramatically at 60 GHz, accurate device modeling becomes indispensible and is normally deemed as the premise of design success. In this chapter, the unique methods and strategies as well as the derived models and simulation results for both active and passive components are described respectively.

4.1 Modeling Strategies

4.1.1 The Main Stream Modeling Methods

Historically, especially in the course of the last decade, many modeling attempts of transistors and passive structures for the 60 GHz application have been reported in the aim of providing better accuracy in the circuit design. Those previous works can serve as source of evidence and inspiration for the proposed modeling methodology.

Transistor Modeling

According to the state of the art, the active device modeling methods could be classified as four types. 1) The conventional RF CMOS transistor modeling method, which has been verified only in low-gigahertz frequencies, are based on a BSIM model with external parasitic to model the substrate and gate resistance [27][28]. 2) The traditional microwave transistor modeling approach is measurement-based and uses tested S-parameter data from fabricated devices [29][30]. 3) A modeling method for the 60 GHz CMOS using the circuit model for fixed device layouts, which sufficiently models small- and large-signal transistor performance up to 65 GHz is commonly adopted in the recent years [8][12][31][32][33][34]. 4) A simulation-based, systematic transistor modeling methodology, which can provide high accuracy and superior flexibility in mm-wave CMOS circuits, is newly proposed [35].

Apparently, the conventional RF CMOS transistor modeling method will not be valid anymore at the 60 GHz Radio design for that it ignores some critical parasitic effects at mm-wave frequencies, such as the delay effect modeled by the inductors. The traditional microwave transistor modeling approach would also be an improper way due to the facts that it lacks of the possibility to extrapolate to frequencies beyond the measurement capabilities of the test equipment and an accurate large-signal model for the design of nonlinear circuit blocks [3].
Hence, the methods, which are reasonable to use in the current 60 GHz CMOS design, are the remaining two mentioned in the last paragraph.

**Passive Elements**

A direct benefit of designing at 60 GHz is the reduced geometry of on-chip passives. However, some issues such as the degraded Q value inevitably emerge. In this section, two main passive components – capacitive element and inductive element are discussed.

In the state of the art, on-chip capacitances have two kinds of implementations: one is metal-insular-metal capacitor (MIM); the other is finger capacitor (MOM). Generally, MIM has good self-resonance frequency and low parasitic capacitance to substrate, but a very poor quality factor although it is sufficiently high at low frequencies. MOM, on the other hand, exhibits a high quality with almost no impact on the circuit performance but have a higher parasitic capacitance to substrate [36]. Finger capacitors are commonly utilized in published works [31][36][38][39] to achieve reasonably high capacitance. They are still some other implementations choose to use the metal-insular-metal capacitors [30][31][37] to take the advantage of lower absolute variation.

The implementation of high-Q precise-valued inductances poses a major challenge for silicon implementations at 60 GHz [3]. For 60GHz circuit design, the passive component values are very small, requiring inductance values on order of 100pH [32]. Currently, there are two basic approaches: the transmission line and the lumped inductor. The transmission line approach is extensively adopted by the majorities while the spiral inductor is selected in some certain applications.

The lumped inductor approach has better inductance per area-ratio [38], consequently smaller dimensions that result in significant area savings. In addition, test structure measurements have indicated that the spiral inductor implementation systematically leads to higher Q while minimizing area [29]. However, the spiral inductor is more susceptible to coupling that may cause discrepancies between design and actual performance [36], which makes it more challenging to model accurately. In contrast to spiral inductors, transmission lines (T-lines) substantially confine the electric and magnetic fields so that is easier to model [9]. Another benefit of using T-lines is that the well-defined ground return path confines the fields and significantly reduces coupling to adjacent structures [32]. Nevertheless, the length of such lines leads to disproportionately tall layouts, making the routing of the signal and power lines difficult [9].
4.1.2 The Proposed Modeling Method

Based on survey of state of the art millimeter-wave device modeling, together with characteristics of the process and availability of the EDA tools, the approaches to model the active and passive elements in particular the transistor and inductor are generally schemed in this section.

Transistor Modeling Methodology

Since there is no fabricated device for measurement available at this stage, and since a more geometrically scalable and layout-optimized model would benefit the circuit performance, an electromagnetic simulation based, existing compact model reused transistor modeling approach is developed for this work.

As stated in last chapter, the 3D EM (three-dimension electro-magnetic) simulator EMDS, which has already been integrated into ADS, is used to carry out the full-wave electromagnetic simulation required by the demand of accurate prediction of parasitics in transistor mm-wave modeling. Compared with stand along EM solver, this EMDS/ADS integration definitely facilitates the Circuit and layout EM model co-simulation a lot.

In essence, the NFET transistor 60 GHz model is customized in a way that the provided RF model serves as the core and then its metallization to top metals and its interconnection to outside world are captured by a combination of selective EM simulation. The detail modeling strategy and processing steps will be explained later on.

The proposed transistor model, in general means, is supposed to achieve these objectives. It should not only accurately predict bias dependence of small-signal parameters at 60 GHz operation but also correctly describe the nonlinear behavior of the device to run the large-signal simulation properly for the mixer as well. The RF noise should also be well characterized, which is crucial for the low noise amplifier design. Additionally, some of the significant effects shown in mm-wave frequency range should be included, for instance, the non-quasi-static effect and the parasitic effect source from substrate network etc.

Considering the above requirements as well as the reality of the existing RF model from PDK (whose details illustrated in the next section), two principles would be adopted. 1) Fully exploit as much as possible the merits of provided compact model for its verified accuracy up to hundreds of gigahertz by the hardware testimony and to avoid repeated effort. 2) The EM model for transistor metallization parasitic should always be extracted for a certain and optimized layout for that extra effects introduced by transistor wiring stacks and connections play a dominate role at such high frequencies as the range among 60 GHz.
**Inductor Modeling Methodology**

Inductors are used in the mm-wave circuit design as bias feeding, impedance matching or tuning out the capacitive parasitics for the transistors. The inductance values on the order of 100pH can handle the operation frequency range around 60 GHz. In a broad sense, an integrated inductor can be in the form either the classic spiral inductor or a short section of the transmission line. A spiral can actually be seen as a differential or an inductive T-line with high characteristic impedance.

The value of the inductance of a shorted transmission line depends only on the length, $Z_0$, and propagation constant $\gamma = \alpha + j \beta$. So that if a transmission line is well characterized, then any arbitrary inductance can be synthesized by varying the length of the line [6]. In most of the cases, modeling method for the t-lines is based on the measurement data and full-wave EM simulation. The scalable electrical models are then characterized and optimized to fit most accurately at the interested frequencies. Normally, first order frequency-dependent-loss is adopted on the T-line model, which assumes no coupling to adjacent structures. This assumption is justified since well-defined ground return path helps confine the magnetic and electric fields, and the close proximity of the adjacent grounds to the signal line helps to minimize any second-order effects [43].

The IBM PDK already provide a collection of different purpose, well characterized transmission line models, which have been verified versus field solver data for various test case geometries through the wide frequency range from 0 to 100GHz. Therefore, it would be advisable to utilize those electrical models in circuit design with some extra selection and characterization efforts.

As regard to the spiral inductor models, in particular the available structures, it would be far more problematic to utilize those spiral physical models in 60 GHz operation. EM simulations are carefully carried out on both the series and parallel-stacked spiral physical layouts given by the PDK, whose geometries are chosen as to achieve largest possible self-resonate frequency. Unfortunately, however, it turns out to be insufficient for operations even above 50 GHz. To illustrate, the inductance and quality value together with the S11 determined from the EM simulation for a PDK inds (series-connected type inductor) are plotted in Figure 4.1 from 1 to 60 GHz. Note that, at ca. 51 GHz, the inductance value becomes negative as well as the Q factor. This indicates that above 51 GHz, the capacitive effects begin to dominate the behavior of this inductor, which can also be intuitively observed by the s11 on the smith-chart.
By rule of thumb, the single layer inductor would give a much higher resonance frequency compared with the stacked ones in the sacrifice of quality factor and area. A good proof is given by [44], shown in Figure 4.2. Hence, the only option to make the inductor reasonably work at 60 GHz for this PDK is to build up a single level spiral.

Figure 4.2: Single level, parallel, and series stacked spirals [44]
Accurate characterization of the electrical behavior of this self-built spiral is therefore of great importance for the later circuit design. Traditionally, spirals on Silicon have been characterized by measurements, where a test wafer with a large number of spirals is designed, fabricated and measured, for instance the method described in [45]. Comparatively, a characterization based on the EM simulation avoids the need for a specific test wafer dedicated to the spiral and allows a predictive design. Advantages of the simulation-based approach are that the designer has more flexibility to try variations of the spiral layouts or even optimize the spiral layouts so that a desired behavior is obtained. The design cycle is also much shorter as it is independent on wafer runs [41]. An optimization procedure for the spiral inductor is proposed in [46], shown in Figure 4.3, which can be an excellent guidance when doing the design tradeoffs among the inductor parameters (e.g. inductance, peak Q value, peak Q frequency, self-resonant frequency and area).

Besides, it is also desirable to have a derived model that can be used efficiently and accurately in the design process of the circuit blocks. Hence, in this work, an equivalent circuit model, using the lumped RLC elements, is extracted according to the S-parameter model from EM simulation. The circuit can be initially designed and optimized using these lumped inductor models which in the last stage replaced by the exact EM simulation models. By doing this, iterative EM simulations are avoided in largest extent, which results in significant reduction of design time.

![Figure 4.3: Optimization procedure for inductors [46]](image-url)
4.2 Active Device Modeling

4.2.1 Review of PDK Offering

For most of previous attempts to model active device in 60 GHz design, the compact models from PDK in those cases are just not desirable and accurate for the mm-wave operation regime. Thus, many efforts have been taken to model and predict those problematic parasitic components that contribute a lot at 60 GHz operation.

The situation here is much better. The NFET model provided by this IBM 65nm PDK does offer some advanced feather dedicated for the mm-wave design. In other words, some modeling efforts conducted in the published works before, have already been included in the offered transistor RF model. Therefore, it is indispensable to have good knowledge of the RF FET model scope before the device modeling.

The MOSFET device from the IBM design kits is modeled by the BSIM4.5 core with sub-circuit wrapper, shown in Figure 4.4 [19].

![Figure 4.4: Sub-circuit schematic of RF FET model [19]](image)

This model features these high frequency parasitic parameters [19]. 1) The gate resistance includes the resistance of PC over RX, PC gate extension resistance, and gate contact resistances through CA and M1. The NQS effect is modeled by fitting the BSIM4 NQS model parameters to hardware data. 2) The source and drain diffusion resistances and wiring contact resistance
through CA-M1 are factored. 3) The wiring capacitances include M1-M1, M1-PC, CA-PC, CA-CA, and CA-M1 capacitances within the FET RX region, plus the PC strap to RX drain/source, PC strap to the substrate ring. 4) The substrate resistances inside the substrate ring plus wiring resistance, including CA and M1, on the substrate ring. Substrate resistance is dependent upon the device dimensions and the substrate contact placement.

The PDK FET model scope is also clearly informed. It offers nice match to the measured data in large signal sweep of $I_{ds}/G_{ds}/G_{m}$ as function of the bias voltage. It operates over a fair range of geometry, in other words it is scalable to a range of device sizes. It has been validated against hardware measurement over the frequency range from 0.2 to 110 GHz. It involves the physical effects such as non-quasi-static effect (NQS), stress effects from shallow-trench isolation (STI), N-well proximity effect and other process-dependent effects etc. Besides, it provides flicker and thermal noise model their correlation to hardware.

### 4.2.2 Transistor Modeling Strategies

Since the existing compact physical model accounts for transistor parasitics within the active regions (RX), up to metal1 (M1), the things missing are mainly the capacitances caused by the metal wiring stacks for the three terminals - drain, gate and source (Here is an assumption that the source and body terminals are tied together).

![Figure 4.5: A NFET layout with wiring stacks](image)

Figure 4.5 shows a specific layout for an 8 fingers NFET with total width of 16µm. In order to ensure the model accuracy, the layout under metal 1 level is strictly compatible to the RF FET PCELL (parameterized cell) style. The gate is fully strapped on two sides through PC (poly) with
double row CA (contact) on gate stripes-aim to keep gate resistance at minimum. The substrate RX ring is fully strapped with CA and M1 - for the best fitting quality on substrate resistance. Besides, a metal4 (M4) ring is made for gate contact to reduce capacitive coupling to substrate and to keep away from the metal stacks of drain and source. The width of the metal5 (M5) bar used to connect the drain and source fingers should be wide enough to cope with 2~3mA current. The design rule check of the PDK requires wide top metal interconnections which results in three big metal-via structures that are on the top levels.

There are, two part of the terminal-connection metallization, which primarily contribute to the wiring-stack capacitive effect. One portion is the level from V1 (via in between metal 1 and 2) to M5, illustrated in Figure 4.6, used to stack the drain and source bars to the fifth metal. The source to drain capacitance mainly dominates in this case due to the vertical and horizontal nature finger capacitances form the source and drain metal stack. The other part is the inter-capacitive coupling among the three huge terminals formed by the top thick metals and vias, depicted in Figure 4.7. In this case, the capacitances between each two ones should be relatively similar.

Figure 4.6: 3D view of the V1-M5 stack in transistor layout

Figure 4.7: 3D view of the WT-LD stack in transistor layout
The general idea is that the transistor layout, including the stack levels from via1 (V1) to last metal (LD), excluding the section below metal1 (M1), is imported into EMDS and simulated as three-port-network in the frequency range from 57 GHz to 64 GHz.

Two problems are supposed to be solved in advance, before we move into next stage. One is EMDS’s deficiency in handling complicated geometries. This EM solver tends to break down when trying to generate mesh for transistor drain and source finger stacks more than three metal layers with default via arrays, shown in Figure 4.8 (a). The corresponding solution is to group each row of via-array as via-bar, shown in Figure 4.8 (b). There is good reason for doing so: this “vias simplification” shows good match in smith chart compared with the original via configuration, demonstrated in Figure 4.9.

![Figure 4.8: 3D view demonstration of the “via simplification”](image)

Figure 4.8: 3D view demonstration of the “via simplification”

The other issue is not much problematic but to find a method to avoid the duplicating the metal1 M1 layer. The fact that both the RF model and the wiring-stack model include the effect of M1 would result in inaccurate modeling. The solution to this problem is simply that when doing the EM simulation for the wiring-stack, mask the M1 layout layer in substrate set-up.

![Figure 4.9: Comparison result as validation for the “via simplification”](image)

Figure 4.9: Comparison result as validation for the “via simplification”
4.2.3 Operation Procedures and Results

The capacitive-effect characterization procedure for this specific transistor is carried out by following three steps:

1) Simulate the supplied RF FET model with interest of parasitic capacitances, which ideally should include the wiring capacitances within the RF FET PCELL, through metal1 level, together with intrinsic components inside the BSIM4 model. The schematic for the Y-parameter-based extraction is given in Figure 4.10, and followed by the simulation results depicted in Figure 4.11. The NMOS transistor has the same geometry as the above shown layout, biased at 0.9\( V_{gs} \), 1\( V_{ds} \) and treated as 3-port component in which the body is connected to source. The important model parameters are defined as follow: “rgatemod” set as 3 to include the gate resistance with NQS effect; “rbodymod” set as 1 to turn on the substrate resistance network; “cwire” set as 1 to account for the wiring capacitance.

![Figure 4.10: Schematic of capacitance extraction for RF NFET model](image)

2) Manually construct the metal-vias stack to enable terminals up to last metal. The stack depicted in Figure 4.12, composed of metal-via stack from via1 to topmost metal. Its S-parameter model is then obtained to check with the capacitive coupling effects (shown in Figure 4.13).

![Figure 4.11: The RF NFET model capacitive effects](image)
3) Finally, wrap the RF FET in first step with the stack model. This new three-port NMOS model would valid throughout Back-End-Of-Line (BEOL) and Front-End-Of-Line (FEOL), covering the frequency range from 57 to 64 GHz. The total capacitive elements are presented in Figure 4.14.
4.3 Passive Device Modeling

4.3.1 Requirement Analysis and Implementation Strategies

Passive components such as inductances, capacitances, and resistances can find ubiquitous usage in the RF front-end circuits. The implementation and modeling strategies of the passives are strongly requirement-dependent.

Inductive elements with high quality factor are required for both the LNA and mixer in the matching networks. Specifically, the estimated inductance values are range from ca. 100pH to 300pH for RF and LO resonating and around 0.5 to 1nH for IF matching.

Generally, the best-performing spirals are selected on criteria, such as a maximum quality factor at the operating frequency in combination with the desired inductance value and available floor-planning space. In most cases, optimization of one of these parameters often entails compromise in one or more of the others. Thus, in order to arrive at the optimal inductor design for a particular application, knowledge of the inter-dependence of the various design parameters is required.
varieties of plots that show the relationships between the various design parameters are presented. They can help deciding which inductor parameters to alter.

The big inductance at 5-GHz IF can be implemented by either the high quality value parallel-stacked inductor or a high inductance density series-stacked inductor. While the small matching inductances (with inductance below 150 pH) could adopt the accurate coplanar waveguide which is demonstrated to be valid up to 110 GHz with acceptable layout length. As duly stated in section 4.2.2, neither the parallel nor the series connected inductors provided by design kits are suitable for resonating in the interested frequency range around 60 GHz. Therefore, constructing, optimizing and modeling of a high resonance frequency and high Q spiral inductor would be a proper solution to the most of impedance matching inductances in the circuits.

In the aspect of on-chip capacitance, the RF-shunt capacitance should be large enough to play insignificant effect on the input and output matching. Same as the inductor, resonating capacitor at IF also requires a large capacitance with decent Q. As small as 10~20fF capacitances would be used to implement the DC-block/RF&LO resonating capacitors.

Instead of transmission lines, lumped capacitors are employed in the matching network, AC coupling and DC bypass. As the rule of thumb, an ideal structure would have infinite impedance at DC and zero impedance at the frequency of interest. Thus, large high quality factor capacitors with self-resonance frequency situated above 60 GHz are desirable.

The big capacitance (ca. 1pF) can be implemented by the high-area-efficiency and high-precision MIM capacitors (in order to get reasonable quality factor especially at RF frequency, it is wise to keep the length-to-width ratio possibly large). Because of the minimum geometry constrain of the top thick metals, it turns out impossible to deal the capacitance smaller than about 80fF with the MIM capacitor. Hence, there are two possible alternatives to find the way out for the small matching capacitances. One is to utilize the parasitic capacitance to ground plane in the bond pad model; while the other is to build finger capacitors using higher level metal stacks (to minimize capacitive coupling to substrate).

**4.3.2 Spiral Inductor Modeling Procedure**

The following procedures are carried out gradually to extract an electrical model for the spiral.

1) Substrate-Stack Validation

As stated in section 2.2.1 that, the Momentum modeling is initiated by entering the substrate details for the BEOL. However, in order to ensure a reliable modeling, this is far from enough. A
validation for the substrate stack is of primary importance. Thus, a simple and efficient test structure, a PDK inductor model layout, is imported to the Momentum. Its EM simulation result is then compared to the one given by the model from PDK. A series-stacked inductor with specific geometry is simulated in both Cadence and ADS, and finally in the Momentum through the frequency up to 60 GHz. The S-parameters shown in Figure 4.15 indicate that good matching is achieved.

![Figure 4.15: Comparison of S-parameters for a PDK inductor](image)

2) Building the Spiral Inductor

Once the substrate set-up has been proved valid, a single level spiral structure is built manually according to the process physical design rule and then optimized to achieve better high frequency performance. As shown in Figure 4.16, the spiral is built using the topmost metal with thickness of 4.125 µm. 3-µm-thick copper line is used as the underpass to connect the center of the spiral. The dielectric thickness between spiral and substrate is around 8.9 µm and the substrate has a resistance of 8-10 Ω/µm.

![Figure 4.16: 3D view of the single-level spiral](image)
There are good reasons to choose the topmost aluminum as the drawing layer. For one thing, in order to reduce the loss and improve the quality factor, the metallization layer with the lowest loss would be preferable. For another, it is advantageous to put this layer as far away from the Silicon substrate to reduce eddy current losses in the substrate and to reduce the capacitive coupling to the substrate.

Furthermore, a patterned ground shield using an array of M1 wires is adopted here, with the purpose of isolating the inductor from substrate noise and preventing induced currents in the substrate, which will lead to degraded quality value.

3) FoM and Performance Evaluation

Just before characterizing this newly built inductor, the evaluation fashion should be made clear to avoid confusion in quantifying the performance. There are two figures-of-merit (FoM) that can be used to characterize the performance of inductors: inductance (L) and quality factor (Q).

For an integrated inductor, the most commonly used model is a π-configuration, from which the individual elements can be extracted from the Y- and/or Z-parameters [44]. This π-configuration can be connected in a single-ended, two-port, or differential configuration. As stated in [44], IBM uses the two-port configuration to avoid the confusion of load condition definition and due to the fact that the differential impedance is not commonly used to describe effective inductance. Shown in Figure 4.17 [44], the effective inductance is extracted from $Z_3$: $L_{\text{eff}} = \frac{\text{Im}(Z_3)}{\omega}$, which is specified as the imaginary component of an impedance divided by the radian frequency $\omega$.

![Figure 4.17: The two-port configuration to extract inductance [44]](image)

As for the quality factor, things are much simpler: a commonly used conventional definition for the Q of integrated inductor can be reasonably adopted here, which is derived from the short circuit input admittance $Q_{\text{conv}} = \frac{\text{Im}(Y_{11})}{\text{Re}(Y_{11})}$. It relates to the difference between the magnetic and electrical energies stored [44]. In addition, it can be easily derived from the measured or
simulated network parameters, which makes it convenient for comparing measurements and simulations of inductors.

4) Characterization Using Momentum

The two-port S-parameters of the inductor are simulated using Momentum over the frequency range from 1 GHz to 60 GHz. The S11 parameter and the estimated inductance together with quality value are illustrated in Figure 4.18 and Figure 4.19, respectively.

![Figure 4.18: S-parameters of the spiral](image1)

![Figure 4.19: Inductance and quality value for the spiral](image2)

Seen from the above figures, 60GHz is far away from inductor’s resonating frequency, which makes the estimation of L and Q more reliable. Besides, over the frequency range from 57 to 64GHz, more than 40 quality values value can be gained with the inductance around 180pH.

5) The Lumped RLC Model

In practice, the S-parameter model as it is obtained directly from the EM simulation can be used during the design phase later. However, it would be more efficient to obtain a derived physical model. The current equivalent circuit commonly adopted the single-π structure, in which series metal resistance and inductance, feed through capacitance, dielectric isolation, and substrate effects are modeled. The cross-section of a spiral inductor together with its equivalent π model is
illustrated in Figure 4.28 (a) and (b) [46]. \( L_s \) consists of the self-inductance, positive mutual inductance, and negative mutual inductance. \( C_s \) is the capacitance between metal lines. \( R_s \) is the series resistance of the metal line. \( C_{ox} \) is the capacitance of oxide layer underneath the spiral. \( R_{sub} \) and \( C_{sub} \) are the coupling resistance and capacitance associated with Si substrate. In the optimization point of view, except for the series inductance, all components in the model are parasitics of the inductor and need to be minimized.

![Diagram of inductor and its equivalent circuit](image)

(a)

(b)

Figure 4.20: Cross-section of an inductor and its equivalent circuit [46]

These parameters can be roughly calculated using the formulas [46] [47] [48], listed in Table I, which would serve as starting point of simulation and optimization. After optimization, its S-parameter is simulated as illustrated in Figure 4.21. Compare with the original ones from EM-simulation, it shows acceptable matching below approximately 30 GHz.

The reason to explain its mismatch at high frequencies is mainly the skin-depth effect. In the low frequencies condition, the DC current is uniformly distributed inside the conductor. Hence, it is reasonable to use only a series resistor and inductor. However, as the frequency goes up to high gigahertz, the depth of current penetrating into the metal (skin depth) becomes comparable to or even smaller than the cross-sectional dimensions of the line [49]. The skin effect pushes the AC current toward the surface of the conductor. To capture this effect, additional RL branch, which models the surface layer resistance and inductance, is introduced as in Figure 4.22. Another high frequency effect, the proximity effect, which refers to the inductive coupling between the neighboring lines, is not going to be accounted for in this case. The reason is simply that here we only use one turn for the spiral to ensure a reasonable Q factor.
<table>
<thead>
<tr>
<th>Parameters</th>
<th>Analytical Equation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$L_S$</td>
<td>$\frac{9\mu_n^2d_{\text{mean}}^2}{11d_{\text{out}} - 7d_{\text{mean}}}$</td>
</tr>
<tr>
<td>$R_S$</td>
<td>$\frac{1}{w\sigma\delta(1 - e^{-\frac{t}{\delta}})}$, $\delta = \sqrt{\frac{2}{\omega\mu_0\sigma}}$</td>
</tr>
<tr>
<td>$C_S$</td>
<td>$\frac{nw^2\epsilon_{\text{ox}}}{t_{\text{ox}}}$</td>
</tr>
<tr>
<td>$C_{OX}$</td>
<td>$\frac{w \cdot l \cdot \epsilon_{\text{ox}}}{2 \cdot d_{\text{ms}}}$</td>
</tr>
<tr>
<td>$R_{\text{sub}}$</td>
<td>$\frac{\rho_{\text{sub}} \cdot l}{d \cdot w}$</td>
</tr>
<tr>
<td>$C_{\text{sub}}$</td>
<td>$\frac{w \cdot l \cdot C_{\text{sub}}}{2}$</td>
</tr>
</tbody>
</table>

Table 4.1: Formulas for RLC parameters

Figure 4.21: Comparison of S-parameters for the original spiral with the pi model
6) Model Comparison and Contrast

The model in Figure 4.22 is under fitting to the original S-parameter model presented in step 4. It is facilitated by the internal optimization capabilities in ADS. The S-parameters of this model with the EM simulation data show good match in wide band (Figure 4.23). Therefore, this model successfully captures most of the behavior of the spiral.

![Figure 4.23: Comparison of S-parameters for the original spiral with the derived lumped model over the frequency band from 1~60 GHz](image)
Chapter 5 Receiver Sub-block Design

In the last chapter, by building a library of active and passive components for the 60 GHz, we have formed the foundation for the predictable and robust circuit design. Then here comes the specific circuit design details of the Rx building blocks – a low noise amplifier and a down-conversion mixer, which can, inversely serve as the verification of the device modeling validity.

5.1 State-of-the-Art and Performance Metrics

5.1.1 State-of-the-Art of 60GHz Sub-circuits

LNA Solutions

Generally, the 60 GHz circuit designs follow one of the following methodologies. One traces back to the traditional “Microwave Means”, which prefer impedance matching with transmission lines; while the other developed from the “RFIC Ways” point of view, which tends to employ lumped components. In state-of-the-art mm-wave designs, the majority adopt the microwave means for its modeling accuracy and well established methods at such a high GHz frequency, for instance the LNAs presented in [8][11][14][32][39][50][51][52][53][54]. In pursuit of low power and limited area performance, some high-gigahertz designs began to choose the lumped inductive elements for matching and resonating, as reported in [10][29][36][55][56][57][58][59]. Additionally, there seems existing a third approach, “Hybrid Method”, that somewhat converges the previous two-design in the circuit way yet implement the inductance with the more reliable distributed elements, as shown in [9][39][60][61].

The circuit topologies of low noise amplifier generally use a cascade of simple amplification stages. They commonly contain two to four gain stages, and each stage employs either CS (common source) or cascode designs. The most convenient solutions are the cascode amplifiers which are extensively used in the low gigahertz LNA designs [8][10][11][29][32][34][50][51][53][58][56][59][61][62]. This architecture is easy to match simultaneously for noise and impedance, capable of reducing Miller capacitance, good for reverse isolation and unconditionally stable. Nevertheless, at higher frequencies for instance 60 GHz, the pole at the cascode node shunts a considerable portion of the RF current to ground, thereby lowering the gain and raising the noise contributed by the cascode device [60]. In contrary, the common source stage, as reported in [7][36][39][52][54][55][57][58], has high linearity and output power resulting from large current
and voltage swings. However, it suffers from poor reverse isolation and insufficient gain. Moreover, it is potentially unstable thus special efforts have to be carried out to insure circuit stability in all the frequency bands [10][11]. Apart from the former two popular topologies, a common gate stage can also be found in some particular works, such as in [9][39]. Although it provides a wide input match, the CG topology suffers from relatively larger noise and lower gain.

Single-ended topology dominates in 60 GHz LNA designs for its advantages in terms of noise performance as compared to their differential counterparts. However, there are some differential topologies utilized in published designs [8][14][56], seeking for the characteristics of less susceptible to parasitic feedback loops. Additionally, differential outputs provide the opportunity to connect directly the LNA to double-balanced Gilbert-type mixer cells, which avoid the difficulties in designing on chip balun.

**Mixer Solutions**

In general, two kinds of circuit topologies of CMOS mixer have been published [68]. The first category is the active mixer, which obtained with Gilbert cells or simple Gm-cells, implemented either with single ended or balanced RF and LO inputs. The second one is passive resistive mixers, consisting of a single NMOS transistor LO-pumped either on the gate or on the source, while the RF is applied either on the drain, gate or combined together with LO on gate terminal.

The advantages of active mixer, reported in [2][8][9][10][12][14][51][61][63][64][65], are the relative high value of the conversion gain, the reduced LO power and the good noise figure. However, its power is huge and its electrical performance suffers from a limited linearity. Besides, wide-band LO&RF ports matching is difficult to obtain. The passive mixer, reported in [10][37][39][66], presents a virtual zero DC power consumption, making it attractive for the integration in deep-submicron CMOS, low-supply-voltage technologies. It also presents a linear behavior and superior inter-modulation properties. The main drawbacks of the passive implementation consist of conversion losses and sometimes the need of a higher LO power with respect to the active mixers.

The mixer designs reported in [9][10][12][37][39][50][51][60][61][63][66] having a differential LO signal but a single-ended RF signal, are categorized as “single-balanced mixer”. While the mixer designs reported in [2][8][64][65] having both differential LO and RF input, are classified as “double-balanced mixer”. The single-balanced topology exhibits less input-referred noise for a given power dissipation yet is more susceptible to noise in the LO signal, which may lead to the LO-IF feed-through. The double-balanced mixer generates less even-order distortion, which
relaxes the half-IF issue in the heterodyne receivers [67]. Nevertheless, since the RF signal from the LNA is usually single ended, it makes the single-balanced favorable. The design goal of low-noise, together with mm-wave CMOS modeling difficulties, limits the use of complex mixer topologies [13]. Although the Gilbert cell mixer is ubiquitous in low Gigahertz CMOS designs, due to its complexity, a differential LO would be required. Therefore, simpler architectures are preferred. Low-loss diodes are dominantly utilized in traditional microwave design but are not available in a standard digital CMOS process. Thus, other alternatives such as dual-gate mixers [14], or single-transistor [12] [50], have been investigated.

5.1.2 Performance Evaluation Standards

In most cases of IC designs, an efficient way to evaluate the performance of a circuit block is to use FoM (Figure-of-Merit). The FOM of a radio front-end circuit is commonly in terms of the following factors: operation frequency, gain to noise ratio and large signal property with consumed power. It can be expressed by Eq. 5.1.

\[
FOM = f_c[GHz] \cdot \frac{G[dB]}{NF[dB]} \cdot \frac{IIP3[dBm]}{P_{ac}[dBm]}
\] (5.1)

However, the LNA and mixer would definitely not give same focus on those aspects, due to their different roles in the receiver chain. Since the LNA is the first stage in receiver front-end chain, its noise figure takes dominate role in the overall system noise performance, according to the Friis formula [69]. Moreover, another implication from the Friis formula is that if sufficient gain can be provide by the LNA, then noise contribution of the following stages will be reduced significantly. In addition, to provide enough dynamic range, large signal measures represented by the compression or inter-modulation should be adequate [46]. Followed by LNA, a down-conversion mixer is made there to convert the RF carrier back down to IF. Since the signal amplified by the LNA is applied to the RF port of mixer, this port must exhibit sufficiently low noise. Assume a high gain from LNA, mixer NF is not that critical as in LNA, since its contribution on the system noise is relatively weak. However, it dose necessitate high linearity. According to the total IP3 expression of a cascaded system given in [67], the linearity of latter stages will dominate the total receiver linearity. Same as the reason to exempt a mixer from low noise, the conversion gain performance is also not that significant and even a conversion loss is normal and acceptable.
5.2 Front-end Circuits Design Flow

As described in section 3.2, the design and implementation of the circuit blocks would follow a certain proposed design stream, (shown in Figure 3.3). Given that the active and passive models are already valid in vicinity of 60 GHz, the design flows for both LNA and mixer are presented respectively in this section, which are followed by detailed steps and results in section 5.3 and 5.4.

5.2.1 LNA Design Flow

As depicted in Figure 5.1 (a), the LNA design begins with the choice of circuit topology. Optimal bias point is selected mainly in terms of minimum $F_{\text{min}}$ and then transistor is sizing to achieve maximum $F_t$ and $F_{\text{max}}$. The noise performance and power transferring are then optimized to find a good balance of the two factors. Once the circuit schematic has been settled and its parameters estimated, the initial simulation is conducted. This characterization is divided into two sub-steps: simulation with ideal models and then with modeled passives. When incorporating with real passive models, some adjustments in passives are required. Another parameter modification in need is that when accounting for the passive achievability, which is limited majorly by the process design rules. After the physical layout, some revisions are still demanded to optimize the floor planning and wire routing. Finally, given the design rules checked, a co-simulation using the combination of EM-models and PDK models is carried out.

![Diagram](image)

Figure 5.1: Design flow of frontend circuit blocks. (a) LNA (b) mixer
5.2.2 Mixer Design Flow

The design procedure for mixer, shown in Figure 5.1 (b), is of no significant difference with the LNA flow. Some distinctions appear in initial steps. When bias the transistor, the purpose is not approaching the maximum available gain but to find good nonlinearity. For the transistor sizing, conversion gain has to tradeoff with large parasitic capacitances. In addition, noise performance is not the first-priority criteria anymore.

5.3 Low Noise Amplifier

5.3.1 Discussion of Circuit Topology

As stated in the introduction section, there are three fundamental differences of mm-wave design compared with lower frequency designs. The circuit topology for 60 GHz LNA, in this sense, prefers a simple structure. Based on the analysis in section 5.1.1, a single-stage cascode topology is utilized for the LNA.

The benefits for using cascode in this 60 GHz design are the followings. Compared to a CS or CG stage, the cascode circuits have higher power gain, which is attributed to the reduced Miller effect and the increased output resistance [70]. The reduced S12 lowers the frequency for which the stability factor k becomes larger than one (unconditional stable) [6]. In other words, it provides better reverse isolation and ensures unconditional stability at 60 GHz.

Unfortunately, however, there is one severe drawback of the cascode configuration that prevents its widespread use in mm-wave spectrum. Traditional cascode transistors have a large parasitic capacitance on the inter-stage node. This capacitance will short-circuit the small signal current at 60 GHz, thus reducing the power gain and consequently intensify the noise contribution of the cascode transistor. One commonly adopted solution is placing inductance between the CS and CG transistors to tune out the parasitic capacitances. Generally, there are two techniques to do this inter-stage matching: one is through a short section of transmission line and an alternative is to put a series inductor between the devices.

In low gigahertz LNA design, inductive source degeneration is frequently used to make the input impedance of the transistor more resistive, thereby simplifying the input matching. However, this technique can hardly be implied in this design, since it degrades the equivalent transconductance and hence reduces the gain. The effective transconductance after adding the source inductance can be approximated by Eq. 5.2[70].

- 39 -
\[ g'_m = \frac{1}{1 + j\omega L_s (g_m + j\omega C_{gs})} g_m \] (5.2)

The reduction of gain is intolerable in this case, since there is naturally not sufficient gain for the one stage amplifier especially at such high frequencies. Therefore, the degenerate source topology has been ruled out.

In summary, the circuit topology adopted in this design would be a cascode with inter-stage tuning but without the source degeneration. In addition, input and output impedance matching networks are applied at both ports to match them to 50Ω (detailed in section 5.3.3).

### 5.3.2 Transistor Sizing and Biasing

With the circuit topology settled down, the device size and DC bias should be optimized before the circuit design.

![Transconductance versus drain and source bias voltages](image1)

*Figure 5.2: Transconductance versus drain and source bias voltages*

![Current voltage characteristics of the NFET](image2)

*Figure 5.3: Current voltage characteristics of the NFET*
To reach optimum gain and NF performance, the smallest gate length of this technology, 65 nm (effective channel length $L_{\text{eff}}$ equals 60nm), is employed. When the gate length is determined, the number of gate fingers ($N_f$) will influence the noise figure due to the variation of each finger length that leads to variation of gate resistance. The swept of total Gate width vs. $NF_{\text{min}}$ at 60 GHz is given in Figure 5.4 (a). It is clear from the figure that the $NF_{\text{min}}$ increases linearly versus $W_g$. In the noise point of view, however, the gate width is supposed to be short, since the shorter the total finger width, less gate resistance exists. This can also be explained by the Eq. 5.3, which modeled the gate resistance [29].

$$R_G = \frac{R_{\text{gdr}}}{3} \cdot \frac{W_f}{N_f L_g} + \frac{R_{\text{cont}}}{N_f L_f} + \frac{R_{\text{gdr}}}{N_f} \cdot \frac{L_{\text{access}}}{L_g} \quad (5.3)$$

Intuitively, one conclusion can be drawing from the equation that the gate resistance $R_G$ is proportional to the finger width $W_f$ and inversely proportional to number of fingers, $N_f$. However, the small size transistor leads to inadequate $g_m$ (transconductance) hence not enough gain. Thus, here is the tradeoff between gain and noise. Figure 5.4 (b) provides the relation between $N_f$ and $NF_{\text{min}}$. As expected, when total finger width is fixed, the NF can be reduced significantly with increased finger numbers. However, when the number is getting bigger, NF reduction tends to be saturated. In addition, an extraordinarily large finger numbers will increase the parasitic capacitances and complicates the layout and modeling for the transistor. Therefore, a good compromise should be achieved here between reduced gate resistance and increased capacitance.

![Figure 5.4: Simulated $NF_{\text{min}}$ as a function of (a) $W_g$ and (b) $N_f$ at 60 GHz](image)

There are two small-signal-speed-parameters majorly used to characterize the transistors: transit frequency ($f_t$) and maximum frequency of oscillation ($f_{\text{max}}$). Those two measures are of great important for the transistors gain performances.
The transit frequency is defined as the frequency where the current gain equals one. Hence, $f_t$ serves as the measure for current switches and it can be expressed using Eq. 5.4 [46].

$$f_t = \frac{g_m}{2\pi \cdot C_{gs}} \tag{5.4}$$

Three useful implications can be concluded from this equation [46]. One is that since $g_m$ and $C_{gs}$ are both proportional to the gate width, $f_t$ is independent of $W_g$. The other is that since $g_m$ is inversely proportional to $L_g$, whereas $C_{gs}$ is proportional to $l_g$, $f_t$ increases with $(1/L_g)^2$. This can serve as the proof for the fact that gate width scaling improves the transit frequency. The third implication suggests the bias condition can take effect on $f_t$, since $g_m = dI_d/V_{gs}$, $f_t$ can be improved by increasing the drain current. After bias optimizing, a $f_t$ of ca. 180 GHz is extracted from simulation for the PDK RF NMOS transistor, as shown in Figure 5.5. In addition, a current gain of 8.2 dB can be achieved at 60 GHz under this bias condition and transistor size.

![Figure 5.5: Current gain of the 65nm CMOS RF NFET](image)

The maximum frequency of oscillation is defined as the frequency where the maximum available gain (MAG) equals one. Hence, $f_{\text{max}}$ serves as the measure for power gain achieved at conjugate matching and it can be expressed using Eq. 5.5 [46].

$$f_{\text{max}} = \frac{g_t}{8\pi \cdot R_{gs} \cdot C_{gd}} \tag{5.5}$$

This equation reveals that $f_{\text{max}}$ can be improved by minimizing $R_{gs}$ and $C_{gd}$ [46]. As illustrated previously, $R_{gs}$ can be reduced by employing multiple gate fingers. However, too much fingers will result in the increasing of $C_{gd}$. Thus, here is the similar tradeoff situation as the one for optimized NF. we need to work out a good balance. It is for sure that $f_{\text{max}}$ should be at least larger than the operation frequency to provide certain gain. As depicted in Figure 5.6, a $f_{\text{max}}$ of ca.
97 GHz is estimated from small signal simulation and a 5.3 dB power gain at 60 GHz can be achieved for a single RF NFET from PDK.

![Power Gain of the 65nm CMOS RF NFET](image)

Figure 5.6: Power gain of the 65nm CMOS RF NFET

### 5.3.3 Noise and Impedance Matching

The essence of any LNA design is the input impedance match, which is optimized for minimum noise factor together with maximum power transfer [6]. Fortunately, the cascode topology in use enables the coincident matching of power and noise. In addition, the good reverse isolation of cascode amplifiers makes the output impedance matching does not change the input impedance.

Eq. 5.6 gives the noise figure of a two-port amplifier [71].

\[
F = F_{\text{min}} + \frac{G_s}{R_s} \left| Z_s - Z_{\text{opt}} \right|^2
\]

(5.6)

The noise factor reaches a minimum value, \(F_{\text{min}}\), at the optimum input impedance match, expressed in optimal input impedance \(Z_{\text{opt}}\). To get a low noise factor, a matching network that transforms the original source impedance \(Z_{\text{source}}\) to \(Z_{\text{opt}}\), where \(Z_{\text{opt}}\) should be designed to coincide with the matching condition for maximum power transfer. To make the power and noise match coincident, \(Z_{\text{in}}\) and \(Z_{\text{opt}}\) should be complex conjugates of one another. Failing to make the noise match and power match coincident results in sub-optimal noise figure [6].

### 5.3.4 Circuit Schematic and Parameters

Based on the above illustration and analysis, a LNA design is presented for the 60 GHz front-end. In Figure 5.7, the circuit schematic of the LNA is depicted, including the matching and bias networks.
The elements enclosed in the red-dashed square compose of the core of this amplifier circuit. It is a cascode with inter-stage matching and without source degeneration. The reasons to adopt such a topology have been duly stated in section 5.3.1. Briefly, cascode structure reduces Miller effect and increased output resistance; no inductive source degeneration avoids the $g_m$ reduction; and CS and CG inter-stage inductance resonant with capacitance seen at the cascode point.

Apart from this, matching and biasing network are required for the amplifier design. For the bias part, inductive bias is selected instead of the traditional resistive bias, in order to achieve low noise performance and low voltage drop from DC supply. For the match part, two LC tank at both RF input and output transfer the impedance to 50 Ω. The capacitors in matching networks ($C_{in}$ and $C_{out}$) are reused as DC block capacitances. The inductors in matching networks ($L_{in}$ and $L_{out}$) are reused as the shunted biasing inductances mentioned above. Besides, big capacitances ($C_1$ and $C_2$) are placed at supply points to shunts undesired RF signals from DC supply.

The dimension of the transistors is determined with respect to several aspects, maximum achievable gain, noise factor, power consumption, linearity etc. According to section 5.3.2, after doing tradeoff among those factors, a transistor size of 16 µm total width with 8 fingers is designated for both cascode devices M1 and M2. The gate bias voltage $V_g$ is selected at 0.9 volts.

As for the passive elements, the initial values of those circuit parameters can be estimated by some equations. As stated in section 4.3, the RF-shunt capacitors C1 and C2 should be large enough and here can be on order of 1 pf. To matching both the input to 50Ω, it requires $Re\{Z_{in}\}=50$ Ω, and $Im\{Z_{in}\}= 0$. The impedance of output termination is not necessarily 50 Ω. In this chapter, in order to characterize the LNA individually, we make its output match to 50 Ω as well. When doing the Rx integration later, which is the case the LNA needs to be loaded with the
mixer in next stage, we have to rebuild the matching network then (details in section 6.1.1). The input impedance of the LNA can be expressed as in Eq. 5.7,

$$Z_{in} = \frac{1}{j\omega C_{in}} + j\omega L_{in} \left( R_{gs1} + \frac{1}{j\omega C_{gs1}} \right) \left(\frac{1}{j\omega C_{gd1}}\right)$$

(5.7)

where $C_{gs1}$, $R_{gs1}$ and $C_{gd1}$ can obtained from the capacitive parasitic extraction of the NFET in section 4.2.3. Likewise, the output impedance of the LNA can be expressed with Eq. 5.8,

$$Z_{out} = \frac{1}{j\omega C_{out}} + j\omega L_{out} \left( g_{m2} \cdot Z_{out1} \right)$$

(5.8)

where $g_{m2}$ is the transconductance of the common gate stage M2 and $Z_{out1}$ can be gotten by simulating the output impedance of the common source stage M1. The series inductor $L_m$ forms an artificial transmission line with $C_{gd1}$ and $C_{gs2}$, whose characteristic impedance $Z_c$ can be estimated using Eq. 5.9 and Eq. 5.10 [29].

$$Z_c = Z_{out1} = \sqrt{\frac{L_m}{C_{gs2}}}$$

(5.9)

$$Z_c = Z_{in2} = \frac{1}{g_{m2}} + \frac{\omega L_{out} Q}{1 + g_{m2} \cdot r_n}$$

(5.10)

### 5.3.5 Simulation Results for Ideal and Modeled Passives

With the calculated values, the LNA is firstly simulated and optimized using the modified RF NFET transistors together with the ideal passive components. After tuning the input, output and inter-stage matching networks, the performances are summarized in Table 5.1.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency</td>
<td>60.0 GHz</td>
</tr>
<tr>
<td>DC Power</td>
<td>4.34 mW</td>
</tr>
<tr>
<td>$S_{21}$</td>
<td>10.01 dB</td>
</tr>
<tr>
<td>$S_{11}$</td>
<td>-27.48 dB</td>
</tr>
<tr>
<td>$S_{22}$</td>
<td>-28.29 dB</td>
</tr>
<tr>
<td>NF</td>
<td>3.91 dB</td>
</tr>
<tr>
<td>$I_{IP3}$</td>
<td>-2.9 dBm</td>
</tr>
</tbody>
</table>

Table 5.1: Performance summary of LNA with ideal passives
Once the merits are acceptable, the simulation in the following step is to incorporate the passive elements modeled by LRC equivalent network (one part of the models come from PDK, and the other part come from the derived models in section 4.3). The simulation results are given in Figure 5.7~Figure 5.12 (The S-parameter simulations are in the frequency range 57~64 GHz).

Figure 5.8 gives the return losses of the two RF ports. S11 refers to the input reflection coefficient and S22 refers to the output reflection coefficient. As shown on the Smith chart, the input and output impedances are approaching the perfect match to $50\, \Omega$. At 60 GHz, the S11 and S22 are -34.4 and 27.5, respectively.

![Smith chart showing input and output impedances approaching 50\,\Omega](image)

Figure 5.8: Return losses at input and output ports

Figure 5.9 presents the S21 and noise figure of the LNA. In range of 57~64 GHz, the forward transmission coefficient S21 keeps above 5.4 dB and the noise figure keeps below 4 dB. The stability is checked in Figure 5.10. The k-factor is above one, which implicates that the circuit is unconditional stable around 60 GHz. Since in the interested frequency spectrum, $K \geq 1$, the MAG (maximum available gain) is defined for the transducer power gain ($G_T$) at optimum source and load terminations. The calculation for K factor and MAG are given in Eq. 5.11 and Eq. 5.12 [71].

$$K = \frac{1 - |S_{11}|^2 - |S_{22}|^2 + |S_{11}S_{22} - S_{12}S_{21}|^2}{2|S_{12}S_{21}|}$$

\[ (5.11) \]
\[ MAG = \left. G_{T \max} \right|_{K \geq 1} = \frac{S_{21}}{S_{12}} \left( K - \sqrt{K^2 - 1} \right) \]  

(5.12)

Figure 5.9: S21 and noise figure

Figure 5.10: Stability factor and maximum available gain

The input and output voltage amplitudes in a transient analysis are depicted in Figure 5.11. Based on the data from the peak-markers, we can estimate the voltage gain of this LNA as:

\[ G_V = \frac{V_{RF_{out}}}{V_{RF_{in}}} = 20\log\left(\frac{0.4 - 0.287}{0.032 \times 2}\right) \approx 4.96 dB \]  

(5.13)

Figure 5.11: Transient responses of RF input and output
Figure 5.12 gives the 1-dB compression point, which is ca. -11.2 dBm of input power level and Figure 5.13 shows an extrapolated IIP3 (input third order intercept point) value of -2.9 dBm.

5.3.6 Physical Layout of the LNA

The simulation results in the last section comes from the circuit using equivalent circuit models, which are composed of the provided PDK models for transistor core and capacitances, and the derived lumped models for inductances and interconnections. However, when it comes to the real floor-planning and physical layout, some passive components should be replaced or modified to account for the passive achievability, to comply with the design rules, and to cope with the floor-planning and wiring optimizations.

The first step is to check if the passive components employed in circuit is possibly to achieve in the real layout practice. For instance, some of the inductances might be optimal for matching, yet cannot be realized by the practical physical spiral or can be achieved but with poor quality factor. Hence, taking into account the layout factor, the inductances in impedance matching design cannot be arbitrarily chosen, but are selected from a range of well-characterized and low loss EM
models. The choice of inductance implementation is also limited by the layout floor plan. For instance, in order to achieve good process match, it is advisable to make the cascode devices as close as possible. In this case, the inter-stage matching inductance $L_{m}$ is preferable to be a spiral inductor rather than a transmission line. The input and output ports of the spiral should come out at same side.

In addition, the three huge inductors should be separated in a sufficiently save distance to ensure no significant inductive coupling among them. For 60-GHz application, the rule of thumb is to keep them ca. 50 $\mu$m away from each other. In the same way, in order to minimize the proximity effect, it is advisable to keep wires especially on the same metal level away from inductor. The rule of thumb is to place wires at least five line width from the inductors [73].

Shown in Figure 5.14, decoupling capacitors implemented by MIM-Cap are connected between the voltage supply and ground. By doing this, high frequency noises are shorted to ground plane. Substrate contacts are placed all around. They offer the passives and transistors good connection to ground plane, which would avoid latch up effects and help to minimize substrate noise [74]. The minimum width of each level of metal wires is determined by the sheet resistance, length, thickness together with current handing capability of the corresponding metal layers.

![Figure 5.14: Layout of the LNA including pads and power rails](image-url)
5.3.7 EM Co-Simulation Results

The final step simulation and validation for the LNA is carried out in the pattern that has been demonstrated in section 3.1.2. Figure 5.15 gives the EM co-simulation results, and performances are summarized in Table 5.2 and being compared with state-of-the-art 60 GHz LNA designs.

![Figure 5.15: LNA EM Co-simulation Results](image)

<table>
<thead>
<tr>
<th><strong>Reference</strong></th>
<th>[34]</th>
<th>[29]</th>
<th>[31]</th>
<th>[52]</th>
<th>[39]</th>
<th>This work*</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CMOS Process</strong></td>
<td>130-nm</td>
<td>90-nm</td>
<td>90-nm</td>
<td>90-nm</td>
<td>65-nm</td>
<td>65-nm</td>
</tr>
<tr>
<td><strong>Frequency [GHz]</strong></td>
<td>60</td>
<td>58</td>
<td>64</td>
<td>60</td>
<td>60</td>
<td>60</td>
</tr>
<tr>
<td><strong>LNA Topology</strong></td>
<td>3-stage cascode</td>
<td>2-stage cascode</td>
<td>2-stage cascode</td>
<td>2-stage CS</td>
<td>3-stage CS</td>
<td>Single-stage cascode</td>
</tr>
<tr>
<td><strong>Gain (S21) [dB]</strong></td>
<td>12</td>
<td>14.6</td>
<td>15.5</td>
<td>12.2</td>
<td>11.5</td>
<td>7.9</td>
</tr>
<tr>
<td><strong>S11/S22 [dB]</strong></td>
<td>-15/-15</td>
<td>-22/-20</td>
<td>-26/-23</td>
<td>-12/-30</td>
<td>-9/-7</td>
<td>-27.2/-30.8</td>
</tr>
<tr>
<td><strong>NF [dB]</strong></td>
<td>8.8</td>
<td>5.5</td>
<td>6.5</td>
<td>6</td>
<td>5.6</td>
<td>4.2</td>
</tr>
<tr>
<td><strong>IIP3 [dBm]</strong></td>
<td>-</td>
<td>-6.8</td>
<td>-</td>
<td>14</td>
<td>-</td>
<td>-7</td>
</tr>
<tr>
<td><strong>P1dB [dBm]</strong></td>
<td>2</td>
<td>-</td>
<td>3.8</td>
<td>-</td>
<td>1.5</td>
<td>-</td>
</tr>
<tr>
<td><strong>DC Power [mW]</strong></td>
<td>54</td>
<td>24</td>
<td>86</td>
<td>10.5</td>
<td>72</td>
<td>5.5</td>
</tr>
</tbody>
</table>

*Results come from EM co-simulation not measurement data.

Table 5.2: Performance comparison of CMOS LNAs around 60 GHz
5.4 Down-conversion Mixer

5.4.1 Choice of Circuit Topology

As illustrated in section 5.1.1, down-conversion mm-wave mixers can be categorized into two families, active and passive mixers. Essentially, mixer modulates either the $g_m$ of an amplifier or the resistance of a switch to produce the mixing action through time-varying mechanism [6]. Same as the case in 60 GHz LNA design, the design goal of low-noise and good-linearity, together with mm-wave modeling difficulties, limits the use of complex mixer topologies. In this application, a single-gate transconductance mixer is adopted, which employs the nonlinear resistive characteristics of the transistor at zero drain-bias voltage.

The functional mechanism of $g_m$-type mixer (whose schematic simply depicted in Figure 5.13), is basically that either the gate or drain voltage can be pumped by the LO signal leading to a time-variant function of $g_m$, which generates the nonlinearities required for mixing [46].

![Figure 5.16: Simplified schematic of $g_m$ mixers (a) gate-pumped and (b) drain-pumped](image)

Shown in Figure 5.16 (a), the gate-pumped mixer operates in the saturation region with $V_{gs}$ biased close to the threshold voltage $V_{th}$, where non-linear variation of $g_m$ are achieved [46]. Here, the LO is applied at the gate, thus $g_m = f(V_{gs})$. Whereas the drain-pumped mixer, depicted in Figure 5.16 (b), is biased at the transition between the linear and saturation region, where the good nonlinearities can be provided by the transistor [46]. Here, the LO is fed at the drain, so that $g_m = f(V_{ds})$. As will illustrated in the next section, this topology requires nearly no bias-voltage at drain which leads to a zero power consumption. In addition, using only one common-source transistor promises a good noise figure at mm-wave frequencies. Moreover, because the mixer is effectively a CS device, it is easy to model with simple transistor structure [6].

In this design, we choose the drain-pumped type, since the RF and LO are placed at different transistor terminals. It makes the high frequency port filtering easier and improves the LO to RF isolation, which is an important factor for overall system performances. In addition, low loss is achieved here by reusing the LO power to drive the transistor into active region instead of put some bias voltage at drain [46].
5.4.2 Device Sizing and Bias Optimizing

Due to the natural simplicity of this single-gate topology, only one single transistor needs to be investigated in terms of biasing and sizing.

To get high nonlinearities hence the ability to generate frequency-mixing product, the bias point should be set to enable a small LO power to pump the transistor into the switching between linear and saturation region. Attribute to extremely small drain-source saturation voltage $V_{ds}$ of the 65 nm CMOS technology, (around 0.2V in this case, refers to Figure 5.17), nearly no bias voltage is required since the applied LO power is sufficient to drive the transistor from zero $V_{ds}$ to saturation voltage. From the figure, it is worth noticing that the lower the gate bias ($V_{gs}$), the smaller the drain to source voltage it requires for saturation. The small $V_{ds,sat}$ will result in a lower LO power level to drive the transistor. Hence, it is possible to generate the nonlinear

![Figure 5.17: Current-voltage characterization of NFET](image1)

![Figure 5.18: NFET characterization of drain current vs. gate-source voltage](image2)
transition with a relatively low LO power, as long as the $V_{gs}$ is set just above the threshold voltage. The threshold voltage in this case is around 0.4V, as shown in Figure 5.18.

However, a low $V_{gs}$ will decrease the transconductance, as indicated from Figure 5.19, which consequently leads to a decreased conversion gain for the mixer. In order to find the optimal gate bias point, the conversion gain is swept as the function of $V_{gs}$, as given in Figure 5.20 (a). Therefore, biasing is a tradeoff among gain, LO power and DC power consumption.

![Figure 5.19: NFET characterization of $g_m$ versus drain and source bias voltages](image1)

![Figure 5.20: Conversion-gain sweeping as function of (a) $V_{gs}$ and (b) $W_g$](image2)

Given the proper bias point, the transistor is then sized aiming to achieve the optimal gain, since it is challenging to avoid conversion loss for this mixer type. The conversion gain is swept as the function of $W_g$, as depicted in Figure 5.20 (b). One thing needs notice is that larger devices would suffer from large capacitance (for instance a big $C_{gd}$) that will degrade the LO to RF isolation. So that if the LO to RF isolation matters, here is tradeoff to be considered. Furthermore, the chosen gate width should allow relatively simple impedance matching and filtering at the specific RF, LO and IF center frequencies of 60 GHz, 55 GHz and 5 GHz, respectively.
5.4.3 Filter Networks for Ports

The impedance matching for mixer is not as complicated as in LNA design since noise is not a critical factor for mixer performance. In this case, filtering networks are required at all three ports.

At RF port, same as the LNA input-matching network, the input impedance of the mixer $Z_{in}$ is converted towards the source impedance $Z_s$ through a LC network, to achieve maximum power transfer. In this chapter, in order to characterize the mixer individually, the impedance needs to be matched to 50 Ω. However, this impedance matching is not necessary when integrating the LNA and mixer together (will give details in chapter 6). Additionally, this network should also features a high pass filtering, which will suppress the IF frequency. Furthermore, it should also function as gate bias and input DC block. At the IF port, the network should have low-pass feature in order to filter out both the LO and RF frequencies. Impedance matching at 5 GHz IF is also required here to match the output to 50 Ω here. (Of cause, the impedance matching is not necessary when the mixer is followed by a next stage for instance an IF buffer. Then the output impedance in that case should be as high as possible to ensure enough gain.) At LO port, a band pass filter is utilized to select the pass band around 55 GHz, meanwhile do the impedance matching and DC blocking for LO feeding.

5.4.4 Circuit Schematic and Parameters

According to above explanations, a transconductance mixer, depicted in Figure 5.21, is designed to down-convert the 60 GHz RF to a low-gigahertz IF of 5 GHz.

![Circuit schematic of a transconductance mixer](image)

Figure 5.21: Circuit schematic of a transconductance mixer

The gate-source voltage $V_{gs}$ is chosen as 0.9V, about twice as the threshold voltage, in the aim of achieving optimal gain. The device is sized as $16 \times 2 \mu m/65nm$. To decrease passive losses, the filter network should be reused for the other functions, which are required by different ports stated above. This results in the minimum numbers of passive elements for three ports, as shown.
in the red-dashed squares. Due to the reason that, in this case, all the three ports are going to be matched to 50 $\Omega$, then the parameters in those LC networks can be obtained efficiently using the Smith-chart tuning tools offered by ADS. The impedance transformation procedures for all three ports are depicted in Figure 5.22. The ideal passive components are firstly used for tuning and then replaced by the real models.

![Impedance transformation for RF, IF and LO ports represented by S11, S22 and S33](image)

**5.4.5 Simulation with Real Models**

To get an intuitive view of how the mixer modulates the RF carrier into IF signal, a transient analysis is conducted for 0.4 $\mu$s. The voltage amplitude of both RF input and IF output are given in Figure 5.23. It can be indicated from the figure that the mixer suffers from several dBs conversion loss, and the signal frequency has been modulated around 12 times slower.

![Transient amplitude at RF input and IF output](image)

The return losses are shown in Figure 5.24. Values of -53.1, -46.5 and -7.5 are simulated at the RF, IF, and LO ports, and the corresponding frequencies of 60, 5, and 55 GHz, respectively.
Figure 5.25 shows the conversion gain of the mixer for different LO power levels. In practical, the LO power is usually much higher than the RF power, because the RF signal is attenuated by propagation in the air, whereas the LO is locally fed. Conversion loss of 2.2 dB is achieved for the RF input power of -40 dBm and LO power of 6dBm.

![Conversion Gain vs. LO Power](image)

Figure 5.24: Return loss for RF, LO and IF ports

The noise performance for mixer is more complicated to evaluate due to the fact that the IF noise comes from both $f_{RF}$ and $f_{image}$. However, there is generally no desired signal at the image frequency. In the usual case where the desired signal exists at only one frequency, the noise figure measured is called the single-side band (SSB) NF. Whereas the rare case, where both the “main” RF and image signals contain useful information, leads to double-side band (DSB) NF [7]. Clearly, the SSB NF is greater than the DSB case, since both have the same IF noise but former has signal power in only a single sideband.

The SSB and DSB NF of the mixer are given in Figure 5.26. At 6dBm LO power, a SSB NF of 6.4dB is the reasonable noise performance to be reported instead of the DSB NF.

The IF output power versus swept RF input shows in Figure 5.27. The input 1-dB compression point of the mixer is approximately -7.6 dBm at 60 GHz.
To investigate the intermodulation property of the mixer, a two-tone test is used with the frequency offset of ±200MHz. Figure 5.28 shows an approximated 19dBm IIP3 for RF two-tone of 60.2 GHz and 59.8 GHz.

Figure 5.26: SSB and DSB NF vs. the LO power

Figure 5.27: RF input power versus IF output power to capture ICP

Figure 5.28: RF input power versus IF output power to capture IIP3
5.4.6 Physical Layout of the Mixer

Most of the floor-plan considerations and layout guidelines explained in section 5.3.6 can also be applied to this part. However, the placement of the inductances is always a headache, potentially limiting the achievability of the passives. Illustrated in Figure 5.29, the existence of the huge IF spiral inductor reduces the possibility of using spirals for RF and LO matching, since the interconnects of the transistor terminals and inductances should be as short as possible. Alternatively, two sections of CPWs are deployed as the inductance in matching networks.

It can be observed that one big issue lies in this mixer layout. The signal pads, e.g. RF and IF are failed to be separated apart; same story happens in the LNA layout for the RF input and output pads. The required matching passive elements at each port are the main reason to explain for the difficulty to wire the signal flow in an optimal way. This problem can be solved when doing the integration of the LNA-mixer cascade, where some of the matching inductances are not there (reference to section 6.2).

Figure 5.29: Layout of the mixer including pads and power rails
5.4.7 EM Co-Simulation Results

Same as for LNA, the final step simulation and validation for mixer is carried out in the pattern in section 3.1.2. Figure 5.30 gives the EM co-simulation results, and performances are summarized in Table 5.3 and being compared with state-of-the-art 60 GHz mixer designs.

Figure 5.30: Mixer EM Co-simulation Results
<table>
<thead>
<tr>
<th>Reference</th>
<th>[12]</th>
<th>[65]</th>
<th>[64]</th>
<th>[66]</th>
<th>[39]</th>
<th>This work*</th>
</tr>
</thead>
<tbody>
<tr>
<td>CMOS Process</td>
<td>130-nm</td>
<td>130-nm</td>
<td>90-nm</td>
<td>90-nm</td>
<td>65-nm</td>
<td>65-nm</td>
</tr>
<tr>
<td>$f_{RF}/f_{IF}$ [GHz]</td>
<td>60/-</td>
<td>60/1</td>
<td>60/0.01</td>
<td>60/2</td>
<td>60/5</td>
<td>60/5</td>
</tr>
<tr>
<td><strong>Mixer Topology</strong></td>
<td>Active quadrature balanced</td>
<td>Active Gilbert cell</td>
<td>Active Gilbert cell</td>
<td>Resistive single ended</td>
<td>Resistive singly balanced</td>
<td>Drain pumped transconductance</td>
</tr>
<tr>
<td>$G_c$ [dB]</td>
<td>-2</td>
<td>3</td>
<td>2</td>
<td>-11.6</td>
<td>-12.5</td>
<td>-4.6</td>
</tr>
<tr>
<td>ICP [dBm]</td>
<td>-3.5</td>
<td>-15</td>
<td>-4.5</td>
<td>6</td>
<td>5</td>
<td>-4.7</td>
</tr>
<tr>
<td>$P_{LO}$ [dBm]</td>
<td>0</td>
<td>0</td>
<td>-</td>
<td>4</td>
<td>8.7</td>
<td>6</td>
</tr>
<tr>
<td>$P_{dc}$ [mW]</td>
<td>2.4</td>
<td>-</td>
<td>93</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>RF Retune Loss (estimated) [dB]</td>
<td>-16</td>
<td>-</td>
<td>-8</td>
<td>-</td>
<td>-5</td>
<td>-30.8</td>
</tr>
<tr>
<td>LO to RF Isolation [dB]</td>
<td>-12</td>
<td>-36</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-13.7</td>
</tr>
</tbody>
</table>

*Results come from EM co-simulation not measurement data.

Table 5.3: Performance comparison of CMOS down-conversion mixers at 60 GHz
6.1 Front End Integration

6.1.1 System Integration and Re-design Consideration

In section 5.3 and 5.4, the frontend blocks are originally designed independently because of the need to characterize them separately. In that case, impedance matching is required at all ports. This leads to signal power waste, large layout area for passives and complex circuits. If the frontend is going to be tested as a whole, system integration measures are required and some of the impedance matching is not necessary any more. One advantage for doing this is that the use of passive components, in different blocks is minimized which leads to less occupied die areas, and reduction of unwanted inter-dependence due to parasitic magnetic couplings.

![Figure 6.1: Schematic of the 60 GHz front-end](image)

The schematic diagram depicted in Figure 6.1, is the integrated Rx frontend circuit. For the LNA, its input and inter-stage matching are kept as in the individual design, while its output impedance matching network needs some adjustments. Since the LNA is directly connected to the mixer, it is not desirable to match its output to 50Ω, which will decrease its inertial voltage gain significantly.

Alternatively, between the LNA-mixer cascades, only an inductor (the inductor in red shown in the figure) is required there to tune the LNA RF output center frequency to 60 GHz. The input impedance of the mixer is mainly capacitive and can be estimated by $C_{gs} + C_{gd}(1 + g_m Z_{out})$. More accurately, we get this capacitance load from the Y-parameter simulation, as 19.62 fF. Besides, a capacitor is required in between the two blocks to make the gate bias of mixer.
unaffected by the DC point of LNA output. This DC-block capacitor should also be considered in
the LC-tank for best tuning accuracy.

Likewise, the input impedance of mixer does not need to be matched either. In this integrated
design, a resistive bias is added at the mixer RF input, instead of the parallel inductance that has
been reused as bias feeding.

Regarding the IF port, also no impedance matching is required anymore. The L-type network
just functions as a low-pass filter there. The IF output is loaded by the impedance of next stage IF
buffer input ca. 22.5fF, if a simple cascode topology is applied.

6.1.2 Receiver Front-end Characterization

The overall LNA/mixer cascade is then simulated to obtain various characteristics. Figure 6.2
plots the simulated return losses, compression point, and NF and gain as a function of frequency.
The simulation test bench is provided in Appendix B.1.

Figure 6.2: Front-end simulation results. (a) Reflection coefficients (b) Gain of LNA and
frontend (c) Total NF of front-end (d) 1-dB compression point

Figure 6.2 (a) shows the reflection coefficients at 60-GHz RF and 55-GHz LO, represented by
S11 and S33 respectively. Both the inputs are well matched at their respective frequencies. Figure
6.2 (b) illustrate the power spectrum at RF, LO and IF ports. A conversion gain of 11.3 dB is
obtained by this Rx frontend. The individual gain of the LNA and mixer are 13.9 dB and -2.6 dB, respectively. Figure 6.2 (c) gives the total noise figure of the frontend. A NF of ca. 8.2 dB is achieved for the combined LNA and mixer at 60 GHz. Figure 6.2 (d) reports the large-signal nonlinearity of the circuit. The input referred 1-dB compression point appears at ca. -22.2 dBm.

6.2 Layout and Simulation Results

Although the separated LNA and mixer layouts are available, layout for the integrated frontend is almost another matter. Due to redesign of passive networks between the LNA and mixer stages, configurations of the original layouts as well as certain passive components have to be modified to adjust the new floor plan.

The final layout of the Rx frontend is shown in Figure 6.3. It occupies a total chip area of 0.45×0.61 mm² including a large area occupied by the wide ground rings and pads. The size of the core cell is only 0.33×0.44 mm².

The three RF inductors used in LNA together with one IF inductor used in mixer dictate structure of the receiver floor plan. Generally, the inductors in LNA would follow the same pattern as there are in the individual layout (refers to section 5.4.6). They are places in close proximity to minimize the length of the 60-GHz lines. The configuration of the relatively huge inductor at the IF output needs altered accordingly. In order to connect directly the inductor output to IF pad, the two wiring connections should be separated at both sides instead of at one side as other inductors. Regarding the LO tuning inductances, it is not desirable to choose the spiral inductor, since there is no space and save room to hold a square-shaped inductances. Alternatively, a transmission line is employed, and it can just connect the source of the mixer transistor straight to the LO pad.

The LO is placed far away from the LNA input so that its leakage does not desensitize the receiver. The signal pads at RF, IF and LO ports follow a G-S-G (Ground-Signal-Ground) pattern, which is aims to comply with on-wafer probes in measurement. Due to ultra high operation frequency, when doing measurement, common practice is that only the power and ground pads are wire-bonded to the testing board, the three signal ports are tested by directly probing on the pads. Separated voltage supply pads are assigned to the LNA and mixer. Like the case in the single LNA and mixer layouts, substrate-contacts are placed around each device at minimum spaces to reduce the substrate loss. Additionally, On-chip bypass capacitors are placed between each supply and ground.
Figure 6.3: Layout of the integrated receiver frontend
As described in Figure 3.1, the co-simulation after layout of the frontend is carried out in such a way: inductance components and some critical interconnects are represented by an S-parameter box resulting from EM simulation. The core transistors are represented by their RF model and are connected to the S-parameter box. Other passives in circuits such as capacitors and bias inductor directly use the corresponding PDK models. Table 6.1 summarizes performance of the frontend as well as the contributions from both LNA and mixer stages. Here is no comparison towards state-of-the-art 60 GHz Rx frontends, since the number of circuit blocks is too small to make this design comparable to other published works.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Overall Front-end</th>
<th>LNA</th>
<th>mixer</th>
</tr>
</thead>
<tbody>
<tr>
<td>RF Port Return Loss</td>
<td>-13.7 dB</td>
<td>-13.7 dB</td>
<td>-</td>
</tr>
<tr>
<td>LO Port Return Loss</td>
<td>-13.2 dB</td>
<td>-</td>
<td>-13.2 dB</td>
</tr>
<tr>
<td>Peak Frequency</td>
<td>60 GHz RF</td>
<td>60 GHz RF</td>
<td>5 GHz IF</td>
</tr>
<tr>
<td>Gain</td>
<td>11.9 dB</td>
<td>18.9 dB</td>
<td>-7.0 dB</td>
</tr>
<tr>
<td>DSB/SSB Noise Figure</td>
<td>7.9/8.2 dB</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>1-dB Compression Point</td>
<td>-22 dBm</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>LO Power</td>
<td>8 dBm</td>
<td>-</td>
<td>8 dBm</td>
</tr>
<tr>
<td>RF Power</td>
<td>-40 dBm</td>
<td>-40 dBm</td>
<td>-</td>
</tr>
<tr>
<td>IF Capacitive Load</td>
<td>17.9 fF</td>
<td>-</td>
<td>17.9 fF</td>
</tr>
<tr>
<td>Power Consumption</td>
<td>6.11 mW</td>
<td>6.11 mW</td>
<td>0</td>
</tr>
<tr>
<td>Chip Area (excl. pads)</td>
<td>0.33×0.44 mm²</td>
<td>0.32×0.18 mm²</td>
<td>0.33×0.18 mm²</td>
</tr>
</tbody>
</table>

Table 6.1: Performance summary of the 60 GHz front-end
Chapter 7 Conclusions and Future Work

7.1 Conclusions

In this master thesis work, the following contributions have been achieved.

- An extensive investigation of state-of-the-art 60 GHz radio design has been made.
- Simulation and design environmental setup has been done for the IBM 65nm technology.
- An ADS-tailored simulation approach together with its corresponding design flow has been proposed.
- Efforts on modeling and characterizing the transistors used in 60 GHz frontend circuits with EMDS have been made.
- Efforts on modeling on-chip inductances and implementing the other passive components have been made.
- A simple receiver frontend, which targets a specific 5/60 GHz WPAN application, has been demonstrated.
- Design and characterization of the frontend blocks (LNA and mixer, both schematic and layout) have been conducted.
- The LNA-mixer cascade has been integrated, implemented, layout, and characterized after the individual design of sub-circuit blocks.
- A conference paper [74], based on the proposed methodology and device modeling sections, has been accepted in the IEEE Nordic Microelectronic Conference, 2009.

7.2 Future Work

There are majorly two efforts, suggested to be on further investigation for this thesis work.

One is to characterize the coupling effects among the big spiral inductors, like the measures employed in [36] [75], and further study the paracitics in layout that have not been taken into account in the post-layout co-simulation.

The other one is to design the IF buffer depicted in Figure 2.1, which can refers to the circuit topology presented in [76].
Bibliography


68


Appendix A | BEOL Set-up for EM Simulator

A.1 Substrate Stack Setup for ADS Momentum/EMDS

In the Momentum, modeling is initiated by entering the substrate details, which can be found, from the following four sources for this 65nm technology case.

**Predefined Substrate Files from PDK**

First source is from the newly released version of the CMOS10LPe design kit. Predefined substrate files covering all of the 16 available BEOL metallization options are supplied which have been created and verified for ADS 2008 Momentum [20]. The predefined substrate files (*.slm) define the physical substrate stack-up such as each dielectric and metal thickness and properties. In our particular needs, the file “substrate5_01_00_01_LD.slm” is selected and it can be found in the following directory on the server:

/afs/it.kth.se/pkg/cadence/IBM_PDK/cmos10lpe/V1.3.0.0RF/EM/Momentum

Those files are provided by the foundry hence may give high reliability.

**MOSIS Website Information**

The second source is from the website information about the MOSIS supported options [21] and the process description [23].

**Model Reference Guide and RF Addendum**

The third source is the CMOS10LPe Model Reference Guide [16] together with its RF Addendum [18] from which physical design information such as the technology cross section and wiring resistance models can be found as well as the physical information of metal layers and vias.

**Parameters Extracted from VPCM**

The fourth source is the technology layer file extracted using the Virtuoso Passive Component Modeling (VPCM). It contains the detail information about each metal and dielectric layer.
A.2 Layer Metallization for ADS Momentum/EMDS

Combined with the previous four sources, the substrate stacks can be properly defined for the modeling of passives such as the capacitors, inductors and T-lines. The next step is then to map the layout layers into the metallization layers, which would be taken into account in simulation. When mapping the drawing-layers, several matters need to be considered.

All the metal layers are mapped as strip (conductor material); via drawing layers are mapped as via (conductor material) with finite conductivity; the area above passivation layer is always defined as open boundary with permittivity equals one; the area beneath p-sub is always treated as closed boundary with perfect conductivity.

All the metal layers are modeled as 3-D expansion thick conductors for the reason that the wiring metal layers are thick enough to be comparable with the wire width and spacing.

Three simplifications can be done for the substrate setting: 1) The p-sub is treated as uniformly doped, eliminating the effect of the transistors, 2) The dielectric loss tangent is assumed to be zero for each SiO2 layers and 3) Multiple dielectric layers are combined into one equivalent dielectric layer with a effective dielectric constant calculated by [20]:

\[ \varepsilon_{eff} = \frac{t_1 + \ldots + t_n}{\varepsilon_1 + \ldots + \varepsilon_n}. \]

The MIM capacitors are built with two extra thin metal masks QT and HT as top and bottom plate between the last copper and the top aluminum metal layers. Those two layers are not included in the metal stacks, which require newly created dielectric and metal layers. The setup is ground on the cross section view of the MIM capacitor provided by [18] and the layer information extracted from VPCM. Besides, the layers from last copper “OA” down to metal 1 can be simplified as one layer for that the capacitor doesn’t need metal grounding or metal dummies filling beneath it.

An intuitive check of the substrate can be done by the process “Momentum -> Post-Processing -> Visualization” or “EMDS -> 3D EM Preview”.

74
# A.3 IBM 65nm BEOL Stacks Set-up File

<table>
<thead>
<tr>
<th>UNIT</th>
<th>um</th>
</tr>
</thead>
</table>

**SUBNAME**

| TOP 0 0 0 0 |
| Bottom 1 1 0 0 |

**SUB0** FreeSpace 1 1 0 1 0 -1 0.000745743 0.000745743 1 0 0

**SUB1** Polyamide 1 3.40 1 1 0 3.275 0.000742468 0.000745743 1 0 3

**SUB2** Nitride 1 7 0 1 1 0 0.4 0.000742068 0.000742468 1 0 3

**SUB3** Oxide 1 4.10 1 1 0 0.45 0.000741618 0.000742068 1 0 3

**SUB4** SiO2_LD_OA 1 4.417 0 1 1 0 1.45 0.000740168 0.000741618 2 1 3

**SUB5** SiO2_OA_BA 1 3.969 0 1 1 0 1.6 0.000738568 0.000740168 2 1 3

**SUB6** SiO2_BA_M5 1 3.981 0 1 1 0 0.32 0.000738248 0.000738568 2 1 3

**SUB7** SiO2_M5_M4 1 3.292 0 1 1 0 0.167 0.000738081 0.000738248 2 1 3

**SUB8** SiO2_M4_M3 1 3.292 0 1 1 0 0.167 0.000737914 0.000738081 2 1 3

**SUB9** SiO2_M3_M2 1 3.292 0 1 1 0 0.167 0.000737747 0.000737914 2 1 3

**SUB10** SiO2_M2_M1 1 3.292 0 1 1 0 0.167 0.00073758 0.000737747 2 1 3

**SUB11** SiO2_M1_STI 1 4.268 0 1 1 0 0.58 0.000737 0.00073758 2 0 3

**SUB12** Silicon_Substrate 2 11.9 66.667 1 1 0 737 0 0.000737 1 0 3

**MET1** M1 0.00073758 0 2 3 4.5005e+07 0 Siemens/m Siemens/m 1 0.22 um

**MET2** M2 0.000737747 0 2 3 4.5005e+07 0 Siemens/m Siemens/m 1 0.22 um

**MET3** V2 0.000737914 0 4 3 9.8235e+06 0 Siemens/m Siemens/m 2 0 um

**MET4** M3 0.000737914 0 2 3 4.5005e+07 0 Siemens/m Siemens/m 1 0.22 um

**MET5** V3 0.000738081 0 4 3 9.8235e+06 0 Siemens/m Siemens/m 2 0.31 um

**MET6** V4 0.000738248 0 4 3 9.8235e+06 0 Siemens/m Siemens/m 2 0 um

**MET7** M4 0.000738081 0 2 3 4.5005e+07 0 Siemens/m Siemens/m 1 0.22 um

**MET8** M5 0.000738248 0 2 3 4.5005e+07 0 Siemens/m Siemens/m 1 0.22 um

**MET9** BA 0.000738568 0 2 3 4.6904e+07 0 Siemens/m Siemens/m 1 0.4 um

**MET10** OA 0.000740168 0 2 3 5.0505e+07 0 Siemens/m Siemens/m 1 3 um

**MET11** LD 0.000741618 0 2 3 3.2323e+07 0 Siemens/m Siemens/m 1 4.125 um

**MET12** V1 0.000737747 0 4 3 9.8235e+06 0 Siemens/m Siemens/m 2 0 um

**MET13** WT 0.000738568 0 4 3 4.31e+07 0 Siemens/m Siemens/m 2 0 um

**MET14** JT 0.000740168 0 4 3 2.2222e+07 0 Siemens/m Siemens/m 2 0 um

**MET15** VV 0.000741618 0 4 3 4.5313e+06 0 Siemens/m Siemens/m 2 0 um
Appendix B  Schematics and Test-benches

B.1 Rx Frontend Test-bench

Notes:

1. There is no dc-block capacitor placed at RF and LO inputs in the test-bench, since the capacitors in the RF and LO matching networks are reused for dc blocking.
2. The capacitor “Cblock” between the LNA and mixer stage employs a real PDK MIM capacitor model instead of using the ideal “DC Block” for the reason that this capacitance is in series with the input capacitance of mixer input to load the LNA.
3. The load of this frontend is set as 22.5 fF, which is the estimation of the input capacitance of next stage IF buffer.