Silicon-based Microwave/Millimeter-wave Monolithic Power Amplifiers

Talha I. Haque

Thesis submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Masters of Science
in
Electrical Engineering

Sanjay Raman, Chair
Willem Odendaal
Timothy Pratt

January 18th, 2007

Blacksburg, Virginia

Keywords: Power Amplifier, linearity, point-to-point, mm-wave, monolithic, integrated circuit, Finite ground coplanar waveguide, 30 GHz

Copyright 2007, Talha Haque

Silicon-based Microwave/Millimeter-wave Monolithic Power Amplifiers

Talha Haque

(ABSTRACT)

There has been increased interest in exploring high frequency (mm-wave) spectrum (particularly the 30 and 60 GHz ranges), and utilizing silicon-based technology for reduced-cost monolithic millimeter integrated circuits (MMIC), for applications such as WLAN, inter-vehicle communication (IVC) automotive radar and local multipoint distribution system (LMDS). Although there has been a significant increase in silicon-based implementations recently, this area still has significant need for research and development. For example, one microwave/mm-wave front-end component that has seen little development in silicon is the power amplifier (PA).

Two potential technologies exist for providing a solution for low-cost microwave/mm-wave power amplifiers: 1) Silicon-Germanium (SiGe) HBT and 2) Complementary metal-oxide semiconductor (CMOS). SiGe HBT has become a viable candidate for PA development since it exhibits higher gain and higher breakdown voltage limits compared to CMOS, while remaining compatible with BiCMOS technology. Also, SiGe is potentially lower in cost compared to other compound semiconductor technologies that are currently used in power amplifier design. Hence, this research focuses on design of millimeter-wave power amplifiers in SiGe HBT technology.

The work presented in this thesis will focus on design of different power amplifiers for millimeter-wave operating frequencies. Amplifiers present the fundamental trade-off between linearity and efficiency. Applications at frequencies highlighted above tend to be point-to-point, and hence high linearity is required at the cost of lowered efficiency for these power amplifiers. The designed power amplifiers are fully differential topologies based on finite ground coplanar waveguide (FGC) transmission line technology, and have on-chip matching networks and bias circuits. The selection and design of FGC lines is supported through full-wave EM simulations. Tuned single stub matching networks are realized using FGC technology and utilized for input and output matching networks.

Two 30-GHz range SiGe HBT PA designs were carried out in Atmel SiGe2RF and IBM BiCMOS 8HP IC technologies. The designs were characterized first by simulations. The performance of the Atmel PA design was characterized using microwave/mm-wave on wafer test measurement setup. The IBM 8HP design is awaiting fabrication. The measured results indicated high linearity, targeted output power range, and expected efficiency performance were achieved. This validates the selection of SiGe HBT as the technology of choice of high frequency point-to-point applications. The results show that it is possible to design power amplifiers that can effectively work at millimeter-wave frequencies at lower cost for applications such as mm-wave WLAN and IVC where linearity is important and required transmitted power is much lower than in cellular handset power amplifiers. Moreover, recommendations are made for future research steps to improve upon the presented designs.

Acknowledgment

First and foremost, I thank Allah (God) who blessed me with guidance, good health and excellent opportunities.

I am grateful to Dr. Sanjay Raman for providing me with the wonderful opportunity to work in the Wireless Microsystems Lab (WML) and for being a constant source of guidance and motivation. Also, I appreciate him suggesting me to M/A-COM for an internship, and then working with M/A-COM to set up funding for this research. I thank Dr. Raman for his patience in reading my thesis and for making many useful suggestions, and also for suffering through it! I shall always cherish my relationship with him. I am also thankful to Dr. Timothy Pratt and Dr. Willem Odendaal for serving on my thesis committee and reviewing my thesis.

I would like to thank George Studtmann of M/A-COM (Roanoke, VA) for providing funds for this research and for the opportunity to work as an intern at M/A-COM. I am also grateful to George for supporting this research in many other ways—by helping in coming up with this research topic, trusting me by lending several expensive equipment for measurements and for his constant encouragement. I also appreciate all the other people at M/A-COM who made my internship experiences a memorable one and were always ready to answer any of my questions. Especially, I would like to thank Dr. Thomas Winslow for being a great tutor during my internship days.

I don't believe my time at Virginia Tech could have been quite as memorable if not for my colleagues at WML. In alphabetical order I list their names: Ibrahim Chamas, Mark Lehne, Krishna Vummidi and Jun Zhao. I appreciate their help and suggestions and for taking time to answer my many questions. I would also like to thank all my other friends that made my stay in Blacksburg a memorable one.

Last but not the least, I would like to thank my respected parents: my father, Irfanul Haque, and my mother, Shahla Irfan; my two sisters, my younger brother and my two brothers-in-law. I want to extend special thanks to my uncle, Syed Basit Asghar, for providing me with an opportunity to come to the US and study. Lastly, I extend thanks to the rest of my family, my friends and to all the people who supported me at various stages of life.

Contents

Contents **vi**

List of Figures **xi**

List of Tables **xvi**

1 Introduction **1**

1.1 Power Amplifiers at Millimeter-wave Frequencies 3

1.2 Technology Comparison 4

1.2.1 III-V vs. Silicon Technologies 5

1.2.2 SiGe HBT 6

1.2.3 RFCMOS 7

1.2.4 SiGe HBT versus other Silicon Device Technologies 8

1.3 Objective and Overview of Thesis 9

2 Design Issues for Microwave/Millimeter-Wave Power Amplifiers **11**

2.1 Large-signal vs. Small-signal Design 12

2.2 Gain vs. Output Power 13

2.3 Linearity 16

2.3.1 1-dB Compression Point (P_{1-dB}) 17

2.3.2 AM-PM Distortion 18
2.3.3 Intermodulation Distortion (IMD) 18
2.3.4 Intercept Point 21
2.3.5 Error Vector Magnitude (EVM) 23
2.4 Efficiency 24
2.5 Power Amplifier Classification 26
2.6 Biasing 27
2.6.1 Constant Current 28
2.6.2 Constant Voltage 28
2.6.3 Temperature-Compensation Biasing 29
2.6.4 Current Source Biasing 30
2.7 Design Goals 31
2.7.1 Linearity 32
2.7.2 Output Power 32
2.7.3 Efficiency 33
2.7.4 Die Size 33
2.8 Approach 33
2.8.1 Monolithic Design 34
2.8.2 Single-stage Design 34
2.8.3 PA Class Choice 35
2.8.4 Differential Topology 35

3 Finite Ground Coplanar Waveguide Transmission Lines 38

3.1 Conventional MMIC Transmission Lines 39
3.2 Transmission Lines in Si Technology 40

3.2.1 Truncated Ground Planes 41

3.2.2 Slotline Mode Suppression 41

3.3 FGC Design Considerations 42

3.4 Simulation of FGC Lines 45

3.5 Thin-film Microstrip vs. Finite-ground Coplanar Waveguide in Atmel Technology 46

4 Atmel SiGe2RF PA Design and Simulation 51

4.1 Iterative Design Process 51

4.2 Optimum Load Resistance 52

4.3 Selection of Device Size 53

4.4 Loadpull 54

4.4.1 Loadpull Theory 56

4.5 Matching Networks 58

4.5.1 FGC Transmission Line Matching Networks 60

4.5.2 Input and Output Matching 62

4.5.3 Bias Network 63

4.6 Power Amplifier Simulation Results 64

4.6.1 Small-signal Simulations 64

4.6.2 Large-signal Simulations 65

5 Atmel SiGe2RF Power Amplifier Fabrication and Measurement 71

5.1 Layout 71

5.1.1 Resistors 72

5.1.2 Capacitor 72

5.1.3 Inductors 73

5.1.4 Symmetry 73

5.1.5 Electromigration 75

5.1.6 Final Layout 75

5.2 Small-signal Measurements 76

5.3 Large-signal Measurements 79

6 IBM 8HP PA Design and Simulation 83

6.1 Circuit Design 83

6.1.1 Matching Networks 84

6.1.2 Bias Circuit 86

6.2 Simulation Results 87

6.2.1 Small-signal Simulations 87

6.2.2 Large-signal Simulations 88

6.3 Layout 90

6.3.1 Resistors 90

6.3.2 Capacitors 91

6.3.3 Inductors 91

6.3.4 Final Layout 92

7 Conclusion and Future work 94

7.1 Summary 94

7.2 Future Work 95

7.2.1 Power Gain and Output Power Improvement 96

7.2.2 Linearity Improvement by Forced Input Mismatch 96

7.2.3 60 GHz Two-stage Power Amplifier 97

7.2.4 Shielded-substrate Coplanar waveguide 97

Bibliography **99**

Vita **105**

List of Figures

1.1 A 64-QAM constellation diagram. 2

1.2 Cross section of typical (a) SiGe HBT and (b) Si BJT. 7

2.1 Block diagram for calculating s-parameters. 12

2.2 Small- and large-signal amplification in bipolar device. 14

2.3 P_{out} vs. P_{in} characteristic curve showing regions of weak and strong nonlinearities 14

2.4 Single-stage power amplifier circuit block. 15

2.5 Generalized P_{out} vs. P_{in} curves for an amplifier comparing output conjugate match to output power match. 16

2.6 Typical AM-PM plots for high- and low-linearity PAs 19

2.7 Output spectrum showing intermodulation products for two-tone excitation. 21

2.8 Second- and third-order intercept points depiction using interpolation of fundamental and IM2, and fundamental and IM3 curves. 22

2.9 Illustration of Error Vector Magnitude for a quadrature signal. 24

2.10 Constellation diagrams for a 16-QAM modulated signal for: (a) an ideal signal, (b) signal with EVM of 0.5%, (c) signal with EVM of 5.1%, (d) with EVM of 8.5%. 25

2.11 Loadlines for different classes on IV curves with collector (output) voltage and current waveforms. 27

2.12 Constant current bias circuit. Q_1 has area that is n times larger than that of Q_2. 29

2.13 Typical temperature-compensation bias circuit. 30

2.14 Biasing of differential pair using a current source connected to the emitters. 31

2.15 (a) Single-ended toplogy, (b) differential topology showing waveform combination. 36

3.1 (a) Standard microstrip, (b) standard coplanar waveguide CPW with lower ground plane. 39

3.2 (a) Thin-film microstrip, and (b) finite ground CPW on a silicon substrate. 41

3.3 Methods of slotline mode suppression: (a) bond wire, (b) airbrdige, (c) dielectric crossover. Permission granted by Michael Chapman to reproduce the figure [42]. 43

3.4 Connection of coplanar ground plane using inter-metal substrate vias. 43

3.5 FGC substrate definition and dimensions used for simulation in IE3D. 44

3.6 Simulated attenuation loss vs. frequency for various center conductor widths. 47

3.7 Simulated effective permittivity vs. frequency for various center conductor widths. 47

3.8 Simulated characteristic impedance and attenuation loss vs. ratio of the ground plane width (W_g) to center conductor width (w) for $w = 30$ μm, $g = 15$ μm. 48

3.9 Simulated characteristic impedance of TFMS vs. frequency for various conductor widths. 49

3.10 Simulated line attenuation of TFMS vs. frequency for various conductor width 49

3.11 Line Attenuation for FGC vs. line attenuation for TFMS. 50

4.1 Differential architecture with tail current source and differential collector load shown. 52

4.2 f_T versus collector current of different transistor device sizes. Device 1 = $5 \times .5 \times 10$ μm^2. Device 2 = $2 \times .5 \times 20$ μm^2. Device 3 = $5 \times .5 \times 15$ μm^2. 55

4.3 MAG and Rollet Stability factor, K, vs. frequency for the selected device size. 55

4.4 Generation of constant output power contour: (a) current-limited power, (b) voltage-limited power, (c) combination of the two. 57

4.5 Power delivered contours for the $2 \times 0.5 \times 20$ μm^2 device size. Contours are in .5 dB step, ending in 10.5 dBm. 58

4.6 Power delivered contours for the selected device; size is $5 \times 0.5 \times 10$ μm^2. Contours are in .5 dB step, ending in 12 dBm. 59

4.7 Different matching networks: (a) L- (b) π- (c) T-transformer. X_1, X_2, X_3 represent the reactances of the passive elements. 60

4.8 Equivalent lumped and distributed elements 62

4.9 Simplified circuit diagram 65

4.10 Simulated S_{11} and S_{22}. The -10 dB RL bandwidth is 2.3 GHz. 66

4.11 Simulated small-signal K factor and gain vs. frequency 66

4.12 Simulted P_{out} vs. P_{in} at 30 GHz. Gain vs. P_{in} is also shown. The 1-dB compression point is approximately P_{in}=10 dBm. 68

4.13 Simulated "large-signal" K factor vs. P_{in}. 68

4.14 Simulated power added efficiency (PAE) vs. P_{in}. 69

4.15 Simulated AM-PM vs. P_{in}. 69

4.16 Simulated Output spectrum vs. frequency. $P_{in} = 10$ dBm. 70

4.17 Simulated EVM vs. P_{in}. 70

5.1 Simulated inductance vs. frequency using of the chosen inductor with parameters summarized in Table 5.2. Note that the SRF for this inductor design is around 36 GHz. 74

5.2 Complete layout of the Atmel SiGe2RF power amplifier design. Die area = 1.75×1.15 μm^2. 76

5.3 Actual die photograph of the fabricated PA with Area=1.75×1.15 mm^2. 77

5.4 Small-signal measurement setup. The DUT is the power amplifier under test. 78

5.5 Measured S_{11} and S_{22} verses frequency of the fabricated power amplifier. 79

5.6 Measured differential S_{21}verses frequency of the fabricated power amplifier 80

5.7 Large-signal measurement setup. The DUT is the power amplifier chip under test. 81

5.8 Measured P_{out} vs. P_{in} of the fabricated power amplifier at 30 GHz. . 82

6.1 f_T versus collector current of different transistor device sizes.Device 1 $= 4 \times .12 \times 6$ μm^2.Device 2 $= 3 \times .12 \times 12$ μm^2.Device 3 $= 3 \times .12 \times 18$ μm^2. 84

6.2 Power delivered contours for the device size of $3 \times 0.12 \times 18$ μm^2. Contours are in 0.5 dB step; maximum P_{out} is 13 dBm. 85

6.3 Simplified PA circuit diagram. Resistances are in Ohms. 86

6.4 Simulated S_{11} and S_{22} verses frequency. 87

6.5 Simulated S_{21} and K verses frequency. 88

6.6 Simulated P_{out} and gain vs. P_{in}. 89

6.7 Simulated PAE vs. $P_{in.}$ 89

6.8 Simulated output spectrum versus harmonics. P_{in} = 7.5 dBm 90

6.9 Complete layout of the IBM 8HP power amplifier. Die area is $1.74 \times 0.89\ mm^2$. Note that components of the layout are symmetric about the plane of symmetry. 93

7.1 Gain expansion by mismatching. 97

7.2 Shielded-substrate coplanar waveguide can be used to reduce attenuation loss caused by lossy silicon substrate. Permission granted by Abbas Komijani to reproduce the figure [12]. 98

List of Tables

1.1 List of different Si-based power amplifiers. 4

1.2 Comparison of fundamental material properties of Si and III-V technologies. 5

1.3 Comparison of fundamental material properties of SiGe HBT and RF CMOS 8

2.1 Atmel SiGe2RF power amplifier design goals 33

4.1 Electrical and corresponding physical lengths of stubs for input and output matching network at 30 GHz. 63

5.1 Resistors used in fabrication of bias circuit. 72

5.2 Chosen inductor parameters. 74

5.3 Comparison between simulated and measured results at 30 GHz . . . 79

6.1 Input and output matching network parameters. 85

6.2 Resistors used in fabrication of the bias circuit. 91

6.3 Chosen inductor parameters. 92

Chapter 1

Introduction

The field of wireless communications has seen rapid growth in the past few decades. In particular, the demand for wireless communication systems with high data transmission capabilities has greatly increased, resulting in the development of new wideband communications formats and standards. In order to accommodate both the high speed and wider bandwidth requirements, as well as to overcome spectrum congestion at lower frequencies, portions of the millimeter-wave (mm-wave) spectrum (normally 30-300 GHz), particularly around 30 and 60 GHz, have been proposed as a solution. This spectrum is relatively underutilized, and hence can be used for broadband, high speed wireless data applications, including broadband video services at 30-40 GHz and indoor wireless local area networks (WLAN) at 60 GHz [1].

Millimeter-wave frequencies present challenges for both system-level and hardware-level design. For instance, the mm-wave spectrum is characterized by high propagation loss, which substantially increases attenuation of signals. This propagation loss is due to higher rain attenuation levels, and the presence of oxygen molecules that absorb electromagnetic energy. This typically limits the range of applications that utilize the mm-wave spectrum to short range, up to 1000 meters, but usually between 100-500 meters [1] [2]. On the other hand, this higher attenuation could enable more secure networks and greater frequency reuse for multi-cell communication systems. Another challenge is the requirement for high linearity front-end hardware due to the complex digital modulation schemes typically used in mm-wave applications. For example, 64-point quadrature amplitude modulation (64-QAM) achieves higher

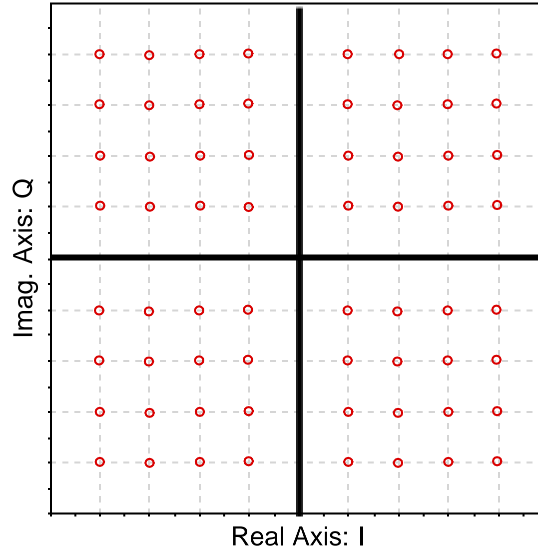

Figure 1.1: A 64-QAM constellation diagram.

spectral efficiency by encoding information in multiple levels of amplitude and phase. Figure 1.1 shows a standard constellation diagram for a 64-QAM modulation. Since the constellation points are packed close together, even some degree of nonlinearity, causing for example, unwanted phase shifts, could result in bit detection errors.

Although many applications have been under development in the 30-60 GHz range, three in particular are considered the most commercially viable by the industry:

- *Inter-Vehicle Communications (IVC).* IVC is considered a part of a more general application—Intelligent Transportation Systems (ITS). The two main purposes of inter-vehicle communication are to increase traffic safety and traffic efficiency [3], maximizing comfort and reducing traffic congestion and traffic-related accidents. It should be noted that 5.8 GHz frequency band has also been assigned for IVC; however, the 59-62 GHz or 60-64 GHz bands have the advantages of security and frequency reuse that make them more suitable for IVC applications [1].

- *Local Multipoint Distribution Service (LMDS).* This is a high data speed indoor/outdoor wireless application that is part of the IEEE802.16 standard, addressing the first-mile/last-mile problem. Applications include distribution of broad-band services such as data, video, voice and internet applications to

fixed users, primarily Interactive Video Services (IVS's) and Video on Demand (VoD) [1] [4]. The operating frequency range for LMDS is the lower mm-wave spectrum from 30-40 GHz. Currently, proposed LMDS systems use a 64-QAM signal modulation scheme.

- *Wireless Local Area Network (WLAN).* WLAN plays an important role within the multimedia communication services and is primarily a short-range, indoor service. The benefits of WLAN to end users are obviously mobility and installation flexibility. The demand for high speed wireless communication has stimulated particular interest in the design of high-quality, high-speed standards with data transfer rate in excess of 54 Mbit/s, up to 155 Mbit/s [1].

1.1 Power Amplifiers at Millimeter-wave Frequencies

The design of power amplifiers at microwave/millimeter-wave frequencies presents a significant challenge. As will be explained below in Section 1.2, the dominant semiconductor technologies for PA design at these frequencies have been the compound III-V materials such as GaAs, and more recently InGaP. These technologies have semi-insulating substrate, which enables fabrication of low-loss on-chip passive components. On the other hand, silicon-based technologies are attractive from the point of view of cost and large available design and fabrication infrastructure. However, standard silicon technologies have much lower resistivity substrate which degrades passive component performance.

One of the issues that has hampered advances in circuit design at mm-wave range, especially in silicon technologies, is the lack of accurate device modeling at such frequencies. Compared to lower frequencies where many parasitic elements and transmission line effects do not play a significant role, at millimeter-wave frequencies such effects become more dominant, and then accurate modeling becomes crucial. To address this issue, work has been ongoing to improve mm-wave frequency device models [5] [6].

Power amplifiers in silicon technologies not only suffer from the lack of accurate

Table 1.1: List of different Si-based power amplifiers.

Power Amplifiers	**Frequency**	**Gain**	**Output Power**	**Output Type**
N. Kinayman [11]	24 GHz	7 dB	14.5 dBm	Single-ended
A. Komijani [12]	24 GHz	18 dB	12 dBm	Differential
S. Chartier [13]	40 GHz	18	6 dBm	Single-ended
B. Floyd [14]	60 GHz	10.8 dB	11.2 dBm	Differential

small-signal models at mm-wave frequencies, but also from a critical lack of large-signal/nonlinear models. Lack of good large-signal models has led to longer, more iterative, and consequently more expensive, design cycles. IC processes, including those available from Atmel and IBM that were used for PA designs in this work, have yet to provide large-signal measurement data (and hence better models) for their design kits. Nevertheless, some work has been done to improve these models [7] [8], and even without data for large-signal models, such design kits can be used, albeit with caution [9]. Another issue that needs to be taken into account is that of breakdown voltage [10]. Transistor models need to be able to accurately predict regions of breakdown due to larger expected voltage swings in PAs.

Silicon-based mm-wave PA design is still in its infancy and hence not many such PAs have been demonstrated. Table 1.1 provides a list of state-of-the-art power amplifiers at, or close to, mm-wave frequencies.

1.2 Technology Comparison

There are several technologies that can be utilized to fabricate monolithic mm-wave integrated circuits (MMICs) and lower frequency radio frequency integrated circuits (RFICs). Typically for power amplifiers, devices are classified in terms of f_T, breakdown voltage, thermal conductivity, integration and cost. The f_T determines the gain capability, breakdown voltage determines the power handling limits and affects linearity, and thermal conductivity impacts reliability. Integration and cost go hand in hand because high level of integration of a PA with other transmitter or transceiver blocks potentially results in a commensurate reduction in cost. Compound III-V technologies have historically dominated this area of RF circuit design. However, recent advances in silicon technology, and demand for greater integration and

Table 1.2: Comparison of fundamental material properties of Si and III-V technologies.

Properties	**Silicon**	**GaAs**	**InP**
Breakdown Field (V/cm)	$\approx 3\text{x}10^5$	$\approx 4\text{x}10^5$	$\approx 5\text{x}10^5$
Low-field Electron Mobility ($cm^2/V\text{-}sec$)	≈ 1500	≈ 8500	≈ 5400
Thermal Conductivity ($watt/cm\text{-}^{\circ}C$)	≈ 1.45	≈ 0.45	≈ 0.71
1/f Noise Corner Frequency (Hz)			
BJT/HBT	$10 - 10^3$	$10^4 - 10^6$	$10^3 - 10^4$
MOSFET/MESFET	$10^3 - 10^5$	$10^6 - 10^8$	$10^4 - 10^5$
Intrinsic Substrate Resistivity ($\Omega\text{-}cm$)	10^3	10^8	10^8

reductions in cost, have made silicon-based technologies a more viable candidate for MMIC/RFIC.

1.2.1 III-V vs. Silicon Technologies

To date, the most widely used III-V technologies for power amplifier design have been Gallium-Arsenide (GaAs) and more recently Indium Phosphide (InP/InGaP) [15] due to their higher low-field electron mobility, transit frequency (f_T) and breakdown voltage. Consequently, GaAs and InGaP power transistors have higher gain and output drive capability compared to silicon-based technologies. Also, GaAs power transistors have higher power added efficiency (PAE—discussed in Section 2.4) because of their higher gain compared to silicon technologies, which would tend to result in higher PAE for an overall amplifier design. Due to the low-loss semi-insulating substrate, passive components with high-Q values (low-loss) can be fabricated directly on the substrate. However, one of the disadvantages of GaAs and InP substrates is their lower thermal conductivity compared to Si. Hence, thermal management issues require GaAs/InGaP wafers to be ultra-thin (less than 100 μm typically), resulting in problems with handling and durability of the wafers, which in turn increases the cost of such technologies [16]. Table 1.2 [17] summarizes the key relative differences between Si, GaAs and InP for RF applications.

1.2.2 SiGe HBT

Recent advancements in Si-based RF technologies, particularly the advent of SiGe Heterojunction Bipolar Transistors (HBTs), offers significant challenge to GaAs HBT technology in the PA area. The unity current gain frequencies (f_T) of SiGe HBTs are now comparable to those available in III-V technologies. In fact, SiGe HBTs have been demonstrated with f_Ts beyond 350 GHz. GaAs technology still has a significant performance advantage due to the higher low-field mobility, which improves the minimum noise figure of receiver font-end amplifiers, but this is not a major concern in PA design. Meanwhile, the availability of PNP transistors and MOSFETs in SiGe BiCMOS technologies offer increased design flexibility in PA circuits. SiGe technology also offers a potential cost advantage over III-V technologies due to compatibility with existing Silicon fabrication infrastructure and potential for higher fabrication yield. Finally, a high level of integration can be achieved in SiGe ICs with other transceiver circuits that are usually on silicon substrate. This in turn reduces the packaging complexity and thus may reduce the overall cost even further. However, there are still challenges in integrating PA circuits into silicon systems-on-a-chip (SOCs) such as signal integrity, power supply transients, etc. Also, although cost-effectiveness is highest for CMOS technology, SiGe is a more expensive technology compared to CMOS.

The SiGe HBT is essentially a compromise between GaAs HBT and RFCMOS devices, and was readily available for this research. Therefore, prototype power amplifiers were designed in this technology. It is important to discuss briefly how SiGe HBT differs from the classical Si BJT. The term heterojunction means that semiconductor materials with different bandgaps form a junction within a transistor. In this case, Ge is added to the silicon in the base of the transistor. Figure 1.2 shows a cross section of a standard SiGe HBT and a BJT [18] [19]. The primary result of adding Ge in the base is that the overall bandgap of the base is reduced because Ge has bandgap of .66 eV compared to Si bandgap of 1.12 eV, which improves the carrier mobilities with respect to Si-only devices [20]. Small base widths are required to reduce transit time (increase f_T), and since SiGe HBT is a vertical structure, epitaxy is used to fabricate very thin device layers. Doping the base with Ge provides other advantages over Si BJT, which will be detailed later in this chapter.

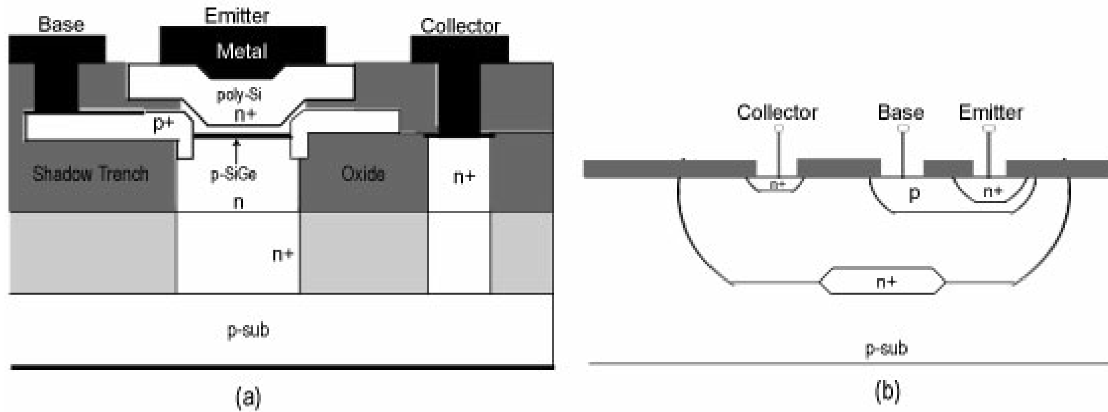

Figure 1.2: Cross section of typical (a) SiGe HBT and (b) Si BJT.

1.2.3 RFCMOS

"RFCMOS" has recently become a popular technology option for RFICs. RFCMOS is different from standard CMOS in that optimized high frequency (RF) models for active and passive components are included in a RFCMOS design kit. Also, more exotic passive components such as spiral inductors, metal-insulator-metal (MIM) capacitors and different resistors for high frequency operation are included in RFCMOS [21]. Other than cost, however, RFCMOS does not compare well with SiGe HBT technology for power amplifier design. The breakdown voltage of a RFCMOS device with comparable f_T is lower than that of SiGe HBTs, which implies lower power handling capability, and thus limits the use of RFCMOS for mm-wave PA design [22]. Moreover, for the power amplifier to sustain high output power, large device sizes are typically needed. While in current advanced SiGe HBT technologies, where vertical transistors are used, this is less of a problem, the footprint in RFCMOS technology due to large MOSFETs can add up quite rapidly, reducing the cost advantages of using RFCMOS technology.

Still, if RFCMOS can provide comparable gain performance, the breakdown voltage limit can be offset using circuit techniques such as sharing voltage swing among several transistors, power combination, differential topologies, etc [23].

Table 1.3: Comparison of fundamental material properties of SiGe HBT and RF CMOS

Properties	**SiGe HBT**	**RFCMOS**
f_T (GHz)	≈ 200	≈ 110
Breakdown voltage (V)	≈ 1.8 (BV_{CE_O})	≈ 4.3[1]
Thermal Conductivity $(watt/cm\text{-}°C)$	≈ 1.5	≈ 1.5
Linearity	*good*	*good*
Power Gain (dB) @ 30 GHz	≈ 10 (*very good*)	≈ 3 (*poor*)
Integration Level	*good*	*very good*

1.2.4 SiGe HBT versus other Silicon Device Technologies

SiGe HBTs have some distinct advantages compared to RFCMOS. Table 1.3 summarizes some of the parameters for state-of-the-art SiGe HBT and RFCMOS technologies [24].

Furthermore, some important differences between SiGe HBT and Si BJT and RFCMOS are summarized below:

- Size and power consumption are critical to mobile applications. For a given f_T, a SiGe HBT would require only about a third of the collector current as compared to a Si BJT. Also since MOSFETs historically have lower g_m per mA of bias current as compared to BJTs/HBTs, the current required to obtain the same f_T would be somewhat higher using MOSFETs.
- Breakdown voltages of SiGe HBTs are about twice those of Si BJTs for same f_T. As far as RFCMOS technology is concerned, the breakdown voltage, as well as the maximum operating voltage, are much more limited due to thin gate oxide breakdown and hot-carrier effects. Moreover, it has been shown [11] that for SiGe HBTs, collector-emitter voltage swing can exceed the breakdown limit by ˜$1.5\times BV_{CE_O}$ to ˜$2\times BV_{CE_O}$ if impedance on the order of less than 500 Ω is seen from the base. Such low impedance allows reverse avalanche current to exit through the bias circuit to ground instead of entering the emitter.

[1] Breakdown voltage for RFCMOS is shown as 4.3V in Table 1.3. For PAs this is quite low and presents a much lower limit than 1.8 V in HBTs because at the highest voltage excursion of $2V_{DD}$ (V_{DD} is 1.8 $\Rightarrow$ $2\times V_{DD}$ = 3.6V), the gate voltage is at the lowest point. So, only a $\pm$.7 V gate voltage excursion could cause breakdown.

- HBTs have better noise performance since they have higher current gain (β) and f_T as compared to BJTs. Since SiGe HBTs typically have higher base doping, the base resistance of a SiGe HBT can also be lower than the Si BJT counterpart, which improves their f_{max}. CMOS devices typically have higher minimum noise figure as compared to bipolar devices when biased at the same current density.
- By optimizing the Ge profile in the base of a HBT, much higher early voltage V_A can be obtained than in BJTs. This results in higher r_o which maximizes gain and also helps in improving stability. Short-channel RF-MOSFETs have much lower r_o than SiGe HBTs due to increased channel-length modulation effects.
- HBTs have higher current gain than conventional BJTs as a result of which improved linearization by feedback is possible. Cancellation of the base-emitter heterojunction capacitance improves the intermodulation performance of HBTs as compared to MESFET and HEMT.

Hence, due to the advantages of SiGe HBTs over RFCMOS, such as higher current gain (β), early voltage, breakdown voltage, and since f_T and f_{max} that are competitive with GaAs/InP technologies can be achieved at potentially lower cost, SiGe technology was selected for the circuit designs in this work.

1.3 Objective and Overview of Thesis

This research focuses on the design of monolithic power amplifier designs in silicon technology, specifically SiGe HBT technology. The designs are to be developed for millimeter-wave frequency range point-to-point applications around 30 GHz. Two different power amplifiers will be developed—one in Atmel SiGe2RF technology and the other in IBM SiGe 8HP BiCMOS. The proof-of-concept power amplifier designs are single stage, differential amplifiers and having either fully distributed (Atmel) or lumped/distributed mixed (IBM) input/output matching networks.

This chapter has provided a brief introduction and background on this research topic. Chapter 2 provides a brief overview of power amplifier fundamentals such as efficiency, linearity specifications, bias classes, etc. This chapter also discusses the major design

steps taken into consideration for the two power amplifier designs and details the approach used for the designs. Chapter 3 discusses the design and implementation of finite ground coplanar waveguides, which were used in the design of distributed matching networks. Chapter 4 describes the implementation of the Atmel SiGe2RF power amplifier design based on design steps and approach outlined in Chapter 2, and provides the simulation results. Chapter 5 details the layout and fabrication of the PA IC and small- and large-signal measured results. Chapter 6 describes the implementation of the IBM 8HP PA design. Simulation results of the prototype IBM PA are also found in this chapter. Finally, the thesis is concluded by a brief summary of accomplishments and a discussion of future research directions.

Chapter 2

Design Issues for Microwave/Millimeter-Wave Power Amplifiers

Power amplifier design is characterized by the need to meet simultaneous requirements of output power, reasonable gain, stability of operation, avoiding permanent transistor damage due to breakdown effects, and a compromise between linearity and efficiency that satisfies the given specifications. In order to achieve these goals, accurate linear and nonlinear models, robust input and output impedance matching and effective current handling capability are required. In the case of the power amplifiers designs in this research, a high level of integration is desired in the minimum possible die area, suggesting a monolithic implementation including on-chip bias circuits and matching networks.

This chapter will briefly discuss non-linear effects in RF amplifiers and will summarize commonly used parameters that characterize a power amplifier design such as: output power; gain; power added efficiency (PAE); and linearity measurements such as 1-dB compression point, intermodulation distortion, input referred intercept points (IIP) and error vector magnitude (EVM). Then, this chapter will present design goals and approaches for the power amplifiers in this thesis.

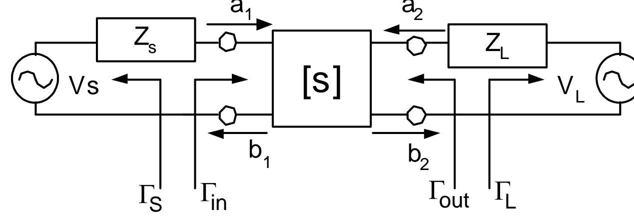

Figure 2.1: Block diagram for calculating s-parameters.

2.1 Large-signal vs. Small-signal Design

Since power amplifiers need to amplify signals to a level that can be reliably transmitted by the air channel over some distance, they must provide large voltage and/or current swings to the load antenna. This is explicitly different from small-signal RF amplifiers, such as a low-noise amplifier (LNA), that exhibit much smaller voltage and/or current swings around a DC operating point. Since strong nonlinearities resulting, for example, from gain compression are not a factor in small-signal amplifiers, they are usually treated/designed as linear amplifiers.

The standard s-parameter-based design of RF amplifiers is based upon assumption of linear or small-signal operation. Consider the two port network representation in Figure 2.1. The input reflection $s_{11}^{'} = \Gamma_{in}$ when the network is presented with load reflection Γ_L can be written as:

$$s_{11}^{'} = \frac{b_1}{a_1} = s_{11} + \frac{s_{12}s_{21}\Gamma_L}{1 - s_{22}\Gamma_L} \tag{2.1}$$

Similarly, the output reflection coefficient $s_{22}^{'} = \Gamma_{out}$ when the network is presented with input reflection Γ_S can be written as:

$$s_{22}^{'} = s_{22} + \frac{s_{12}s_{21}\Gamma_S}{1 - s_{11}\Gamma_S} \tag{2.2}$$

It is well known that for linear small-signal amplifier design, conjugate matches at the

input and output such that $s'_{11} = \Gamma_S^*$ and $s'_{22} = \Gamma_L^*$ result in maximizing the overall gain. On the other hand, large-signal amplifiers exhibit nonlinear effects because of larger voltage/current swings cause cutoff, compression, etc. (Figure 2.2). Therefore, standard s-parameters design does not work and is not applicable; rather the design is typically focused at finding optimal load impedance for maximum output power. However, "large-signal" s-parameters can also be calculated for nonlinear circuits such as power amplifiers. Unlike standard s-parameters, "large-signal" s-parameters solutions include nonlinear effects and are power dependent—the s-parameters change with power levels.

Nonlinearities can usually be divided into weak and strong nonlinearities. Weak nonlinearities are due to the inherent nonlinearity in the I-V characteristics of a device, and give rise to intermodulation, the phenomenon of signals of different frequencies mixing together. This situation is dominant when the PA is usually at an operation point backed off from the 1-dB compression point, as shown in Figure 2.3. On the other hand, strong nonlinearities occur when physical transistor limitations such as clipping at saturation or power supply limits are encountered. The nonlinearity can be modeled by a power series representation of output of an amplifier:

$$v_o = a_1 v_i + a_2 v_i^2 + a_3 v_i^3 + a_4 v_i^4 + ..., \tag{2.3}$$

where v_i and v_o are the input and output signals. However, such a mathematical model is restricted to lower order terms and thus does not approximate strong non-linearities well. A different mathematical scheme such as "envelope analysis", could estimate strong nonlinearities more accurately [25].

2.2 Gain vs. Output Power

Two types of power gain are commonly used to characterize an amplifier. Considering the two port amplifier block diagram shown in Figure 2.4, they are defined as:

- Transducer power gain, which is the ratio of power delivered to the load to the

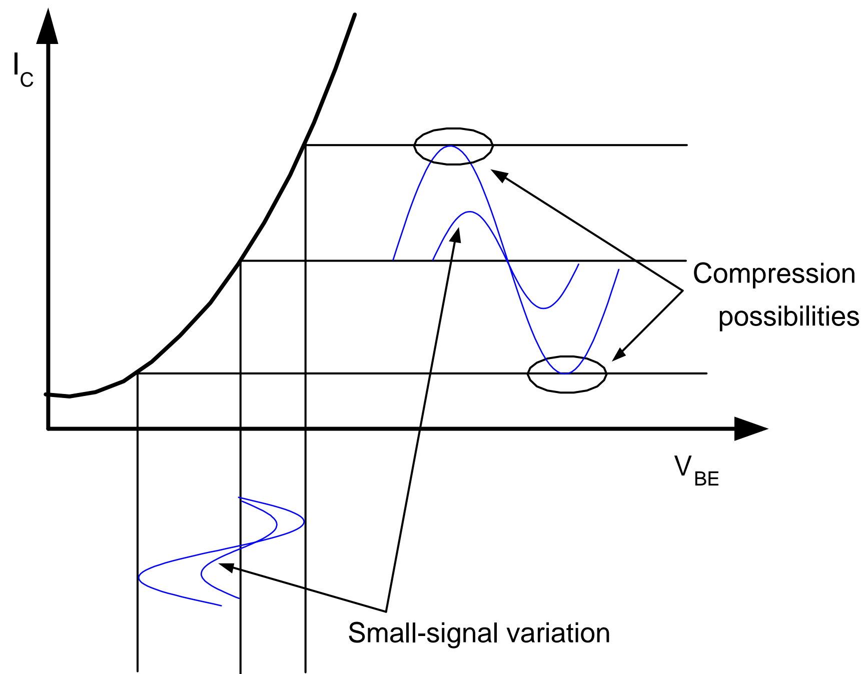

Figure 2.2: Small- and large-signal amplification in bipolar device.

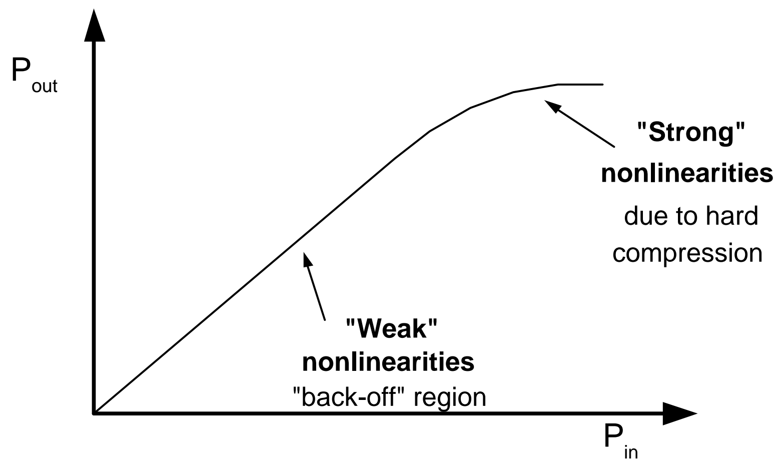

Figure 2.3: P_{out} vs. P_{in} characteristic curve showing regions of weak and strong nonlinearities

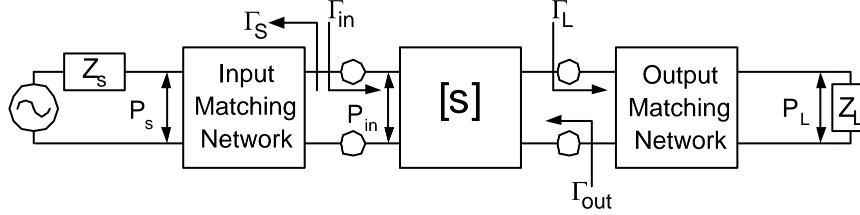

Figure 2.4: Single-stage power amplifier circuit block.

power available from the source or $G_T = P_L/P_S$

- Operating power gain, which is the ratio of power delivered to the load to the power delivered at the input of the active device or $G_P = P_L/P_{in}$

Although G_P is often used because it is important to know what power is being delivered at the input of a device, power amplifiers are usually defined by G_T because it is more useful to consider how much power is actually delivered to the load given a particular power from the source. The transducer power gain can be written as:

$$G_T = \frac{P_L}{P_S} = \frac{(1-|\Gamma_S|^2)(|S_{21}|^2)(1-|\Gamma_L|^2)}{|1-\Gamma_{in}\Gamma_S|^2|1-S_{22}\Gamma_L|^2} \tag{2.4}$$

Equation 2.4 shows that the transducer gain is dependent upon several factors, and to maximize this gain, the input and output need to be conjugately matched. However, doing so will typically result in current/voltage waveforms that do not have enough swing for maximum output power. Power at the load can be defined as:

$$P_L = \frac{1}{2}\frac{V_L^2}{Z_L} \tag{2.5}$$

Hence, if the voltage at the device output has maximum possible swing (from 2×voltage supply[1] to zero) but with a load, Γ_L or Z_L, that does not allow the highest current swing, then the output power will not be maximized. This is usually the case when G_T is maximized—in other words, the load presented to the transistor resulting from

[1]The output (collector/drain) actually swings above the positive supply. An ideal inductor (inductor as RF choke) cannot have any DC voltage across it. Hence, if the output swings from supply voltage to zero voltage, then it must also swing in positive direction from the supply voltage, which means reaching possible peak of $2\times V_{cc}$ or V_{dd} [26].

conjugate output match does not support maximum output power. Therefore, it can be said that there is a trade-off between maximum transducer gain and maximum output power. Thus, a load, often called the optimal load, is needed that can provide maximum voltage and current excursion. By simple circuit theory this optimal load can be calculated as

$$R_{opt} = \frac{V_{\max}}{I_{\max}}, \tag{2.6}$$

where $V_{\max}$ is the peak voltage at device output and $I_{\max}$ is the peak collector/drain current.

Figure 2.5 gives a visual representation of the cases where the load resistance is achieved using conjugate match versus optimal load calculations. Note that conjugate match results in higher gain, while the optimal load results in higher output power and 1-dB compression at a certain high power level.

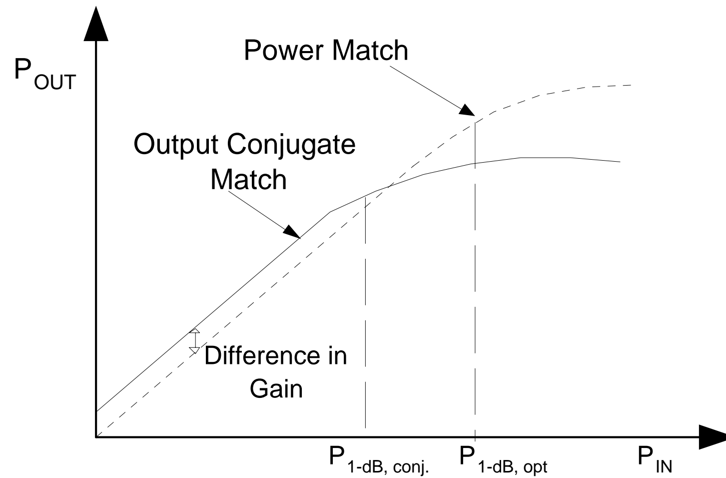

Figure 2.5: Generalized P$_{out}$ vs. P$_{in}$ curves for an amplifier comparing output conjugate match to output power match.

2.3 Linearity

As discussed above, the power amplifier can be viewed as a nonlinear circuit block, which exhibits strong nonlinearity effects when driven at high input levels. Nonlinear

effects can be mitigated by careful design, such as: optimal load selection that achieves maximum output power before entering in compression region; providing load tuning at frequencies of concern; or using circuit techniques that would cancel particular harmonics.

Power amplifiers exhibit nonlinearities in both amplitude and phase. Although amplitude nonlinearity can be well understood as waveform clipping or compression, phase distortion is not as obvious. Compression or clipping limits the gain of a transistor and consequently the output power level. Phase distortion occurs due to nonlinear phase shifting of a signal as it passes through the device. It is entirely possible to have strong nonlinearity in phase but have little nonlinearity in amplitude at a given range of input drive, or vice versa. In some modulation schemes only the amplitude carries the information and in others only the phase is modulated. For example, in a system where most information is carried by the phase of the signal, such as phase-shift keying modulation, a circuit design that minimizes phase distortion at the cost of amplitude nonlinearities might be regarded as preferable. In complex modulation schemes, information is carried by both the amplitude and the phase. Therefore, both amplitude and phase linearities are important; however, one may be preferred over the other depending upon whether amplitude or phase carries most of the information.

Obviously, in a transceiver chain, the relevant nonlinearities need to be understood well and their effects minimized. It is also important to relate the nonlinearities to overall system performance. This leads to different figures of merit for defining power amplifier nonlinearities. Some specifically relate only devices and circuit elements, while others also depend on a communication system characteristics such as the modulation scheme being used.

2.3.1 1-dB Compression Point (P_{1-dB})

The 1-dB compression point is primarily used to quantify the gain compression characteristic of a power amplifier. It has been seen that transistors have physical limits and as such maximum voltage or current waveforms corresponding to a certain drive level. Above, this drive level waveform clipping occurs and gain decreases. The point at which the gain drops to 1 dB below the small-signal/linear gain is referred to as the

1-dB compression point, P_{1-dB},which refers to the input level. The corresponding output power at P_{1-dB} is referred as $P_{out,1-dB}$.

P_{1-dB} does not completely defines PA linearity. The 1-dB compression point is a single-tone figure of merit and only indicates *nonlinearity in amplitude* (AM-AM). It is most useful and practical in modulation schemes where information is send by modulating the amplitude. In other words, a RF system where only phase is modulated, the AM-AM figure-of-merit is not very useful. Other system-related linearity figure-of-merit, explained below, fully define a PA linearity. Nevertheless, P_{1-dB} can be well correlated with other figure-of-merits, as explained later in this chapter.

2.3.2 AM-PM Distortion

In power amplifiers, it is also important to observe phase shifts versus input power drive levels for modulation schemes where information is transmitted by modulating the phase. This is usually referred to as AM-PM distortion, where *PM* stands for *phase modulation*. A typical AM-PM plot is shown in Figure 2.6 [25]. The unit used for this figure of merit is degree of phase/dB. Ideally, phase shift should be constant as input drive is increased, indicating little phase distortion with increasing input power levels. However, in reality, severe phase distortion can occur.

It has been suggested that in bipolars such as HBTs, the nonlinear base-collector capacitance is a crucial source of input level-dependent phase shift [25]. This dependence on transistor parasitic element can lead to phase distortion in what is usually assumed to be the low-level, linear region of operation for power amplifiers. For example, Figure 2.6 shows a typical AM-PM plot of a low-linearity PA where phase shift is increasing steadily even at low drive levels.

2.3.3 Intermodulation Distortion (IMD)

The AM-AM and AM-PM figures of merit described above are derived from single-tone signals. However, amplifiers are often analyzed using two-tone input because this is representative of common frequency spectra, where signals are spaced together

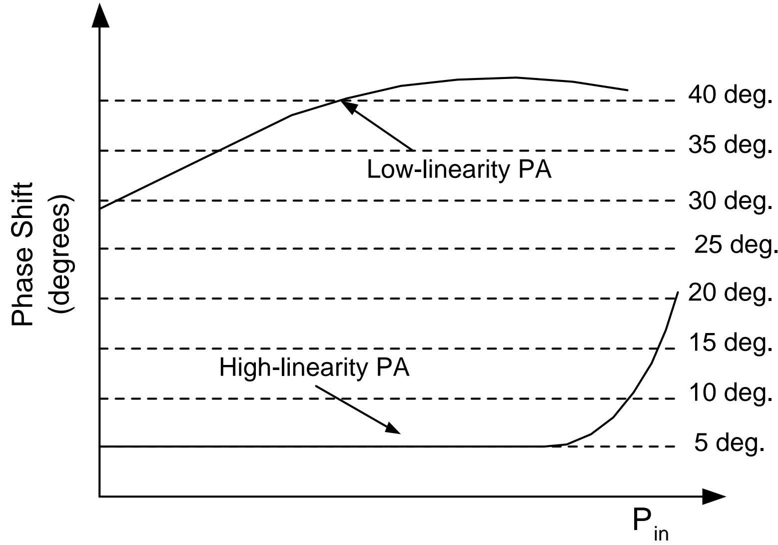

Figure 2.6: Typical AM-PM plots for high- and low-linearity PAs

within a band. Power amplifier nonlinearities produce harmonics at multiples of a fundamental frequency, ω_o. Intermodulation distortion (IMD) occurs due to mixing of any two higher harmonic frequencies. This can give rise to unwanted IM products that fall in the frequency band of interest and corrupt the desired information.

Earlier, a simple power series was described (Equation 2.3) that showed the relationship between the output and the input signal. Assume that the input voltage is a two tone input (w_1 and w_2) where

$$v_i(t) = V\cos(\omega_1 t) + V\cos(\omega_2 t) \tag{2.7}$$

Using Equation 2.7 and Equation 2.3 and appropriate trigonometric identities the output voltage can be written as:

$$v_o(t) = a_1 v[\cos(\omega_1 t) + \cos(\omega_2 t)] + a_2 v^2[\cos(\omega_1 t) + \cos(\omega_2 t)]^2 + ..., \tag{2.8}$$

where the terms are infinite and can be expanded to show IM products explicitly.

Equation 2.8 can be expanded to show different intermodulation products, such as

second-, third- and fifth-order intermodulation products ($2w_1$, $2w_2$, $w_1 \pm w_2$; $2\omega_1 \pm \omega_2$, $2\omega_2 \pm \omega_1$; and $3\omega_1 \pm 2\omega_2$, $3\omega_2 \pm 2\omega_1$, respectively) that fall in band as shown below, where part of expanded $v_o(t)$ is shown:

$$v_o(t) \ : \ a_2 v^2 \cos(\omega_1 + \omega_2)t \ + a_2 v^2 \cos(\omega_1 - \omega_2)t \tag{2.9}$$

$$v_o(t) \ : \ \frac{3a_3 v^3}{4} \cos(2\omega_1 + \omega_2)t \ + \frac{3a_3 v^3}{4} \cos(2\omega_1 - \omega_2)t \tag{2.10}$$

$$v_o(t) \ : \ \frac{3a_3 v^3}{4} \cos(2\omega_2 + \omega_1)t \ + \frac{3a_3 v^3}{4} \cos(2\omega_2 - \omega_1)t \tag{2.11}$$

$$v_o(t) \ : \ \frac{5a_5 v^5}{8} \cos(3\omega_1 + 2\omega_2)t \ + \frac{5a_5 v^5}{8} \cos(3\omega_1 - 2\omega_2)t \tag{2.12}$$

$$v_o(t) \ : \ \frac{5a_5 v^5}{8} \cos(3\omega_2 + 2\omega_1)t \ + \frac{5a_5 v^5}{8} \cos(3\omega_2 - 2\omega_1)t \tag{2.13}$$

Figure 2.7 is a graphical representation of some of these components. Note the decrease in amplitude for higher intermodulation products. The third-order IM products $2w_1 - w_2$ and $2w_2 - w_1$ are of most interest because are close to the fundamental frequencies. The ratio of IM3 amplitude, $\frac{3a_3v^3}{4}$, to $a_1 v$ defines the level of IM distortion. Detailed analysis in [25] has shown that the AM-PM distortion generates additional terms at the same frequencies as the normal third-order intermodulation (IM3) products. This means that they will be added to IM3 products generated by the amplitude distortion, making IM3 products even more prominent in the frequency band. Even-order IM products of closely spaced and equal amplitude fundamental frequencies ω_1 and ω_2 such as $2\omega_1$, $2\omega_2$, $\omega_1 \pm \omega_2$ are generally far away from center/operating frequency. Thus, they are usually less of a concern compared to odd-order IM products. However, depending on system bandwidth, it is possible that they may fall within the spectrum of interest. Nevertheless, it is advantageous to suppress even-order harmonics and their intermodulation products as much as possible especially for large bandwidth amplifiers that are proposed for mm-wave applications.

Intermodulation distortion is critical and very common in RF system, and identifying its presence can lead to better component, i.e. PA, design to mitigate its effects through unique circuit implementations. Also, IM is the source of an important performance metric for RF components, nth-order intercept point, which is explained

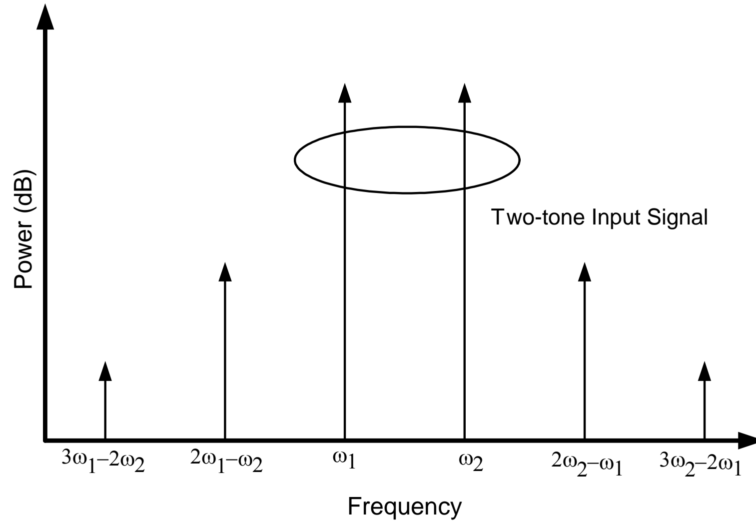

Figure 2.7: Output spectrum showing intermodulation products for two-tone excitation.

below.

2.3.4 Intercept Point

The 1-dB compression point characterizes the amplitude nonlinearity at the amplitude fundamental frequency. A similar figure of merit can be obtained for higher harmonics and intermodulation products. From Equation 2.9 it can be seen that the third-order IM product has an amplitude term $\frac{3}{4}a_3v^3$ while the fundamental term is a_1v (corresponding to the case of no nonlinearity). If the power magnitudes of these two terms are plotted versus input power on a log scale (Figure 2.8, [27]) the fictitious point at which the linear asymptotes of the two curves would intersect is called the third-order intercept point (IP3). Similarly, second-order IM product, with amplitude term a_2v^2, and the fundamental term intersect at some fictitious point. This point is known as the second-order intercept point (IP2), which is also shown on Figure 2.8.

These points cannot actually be measured since in reality the fundamental, the second-order, and the third-order IM responses compress significantly before this point, and thus the two curves do not actually intersect. A quick way to estimate the input referred second- and third-order intercept points (IIP2 and IIP3, respectively)

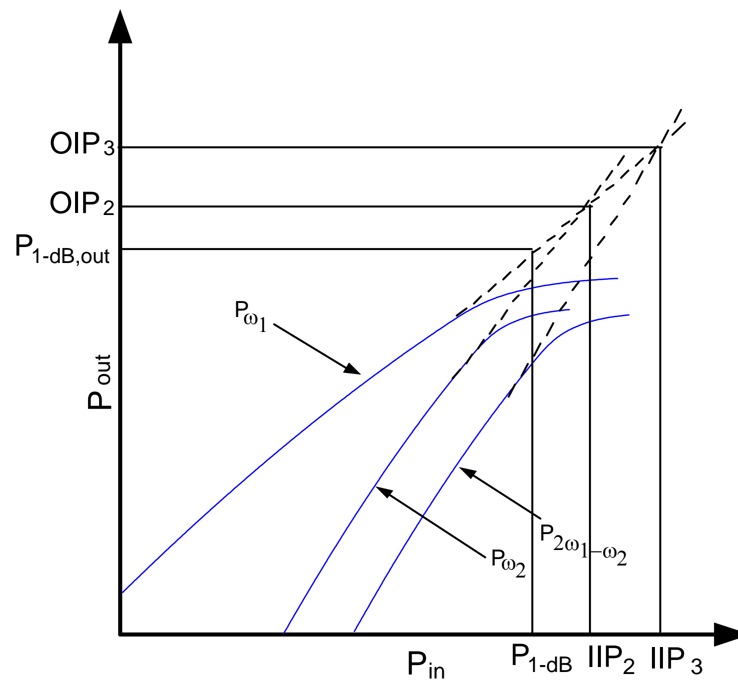

Figure 2.8: Second- and third-order intercept points depiction using interpolation of fundamental and IM2, and fundamental and IM3 curves.

from P_{1-dB} is [25]:

$$IIP2(dB) = P_{1-dB} + 19.6 \quad (2.14)$$
$$IIP3(dB) = P_{1-dB} + 9.6 \quad (2.15)$$

However, it should be noted that the above relationship is a rule of thumb and not a substitute for actual simulations or measurements.

2.3.5 Error Vector Magnitude (EVM)

The classical linearity figures of merit discussed above do not take into consideration the actual communication system modulation scheme. Alternatively, the error vector magnitude (EVM) *is* dependent upon the type of modulated information that is passed through the power amplifier. Millimeter-wave applications such as LMDS at 30 GHz use complex modulation schemes such as Quadrature Amplitude Modulation (QAM) to further increase spectral efficiency. QAM modulates the amplitude and the phase of two carrier signals that are 90° out of phase with each other (I and Q signals).

The EVM measurement is a primary measure of transmitted modulation accuracy [28]. Error Vector Magnitude is defined as the difference between an reference (ideal) waveform or signal and the measured (detected) signal as shown in Equation 2.16 [17] and depicted in Figure 2.9 [29] for a general case.

$$EVM = \sqrt{(I - I_{actual})^2 + (Q - Q_{actual})^2} \;, \quad (2.16)$$

where I and Q are the quadrature components of the measured signal, and I_{actual} and Q_{actual} are the quadrature components of the ideal signal. Also, note that EVM takes into account both the amplitude and phase error. Thus, since power amplifier designed in this work would operate in a RF system using modulation scheme such as 64-QAM, EVM simulations are required during design to verify high linearity.

Figure 2.10 shows a qualitative example of constellation diagrams for a 16-QAM signal [30]. A constellation diagram shows the position of signal points in I/Q space, and the deviation of these points from their ideal locations represents errors/distortion

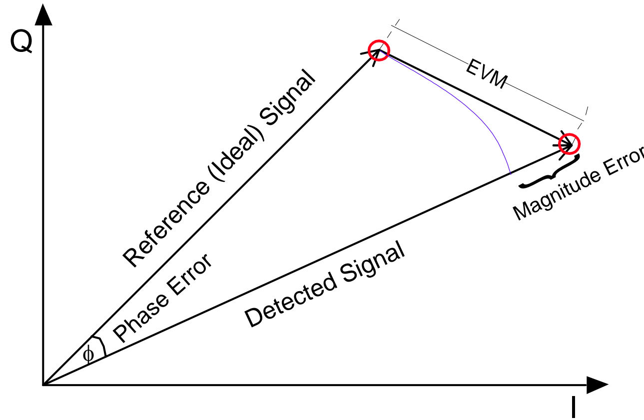

Figure 2.9: Illustration of Error Vector Magnitude for a quadrature signal.

in the signal. For example, Figure 2.10(a) shows how the ideal 16-QAM signal would look like with no error (or EVM of zero). This case would occur if there are no nonlinearities, and thus no errors in amplitude or phase of the signals. Figure 2.10(b) shows distortion that my occur from gain compression of a power amplifier, not considering any other nonlinearity. This results in worst-case EVM of 0.5%. Note that the constellation points have become skewed from the ideal case. Figure 2.10(c) shows what a more realistic signal would look like, with higher nonlinearities; this resulted in worst-case EVM of 5.1%. Finally, Figure 2.10(d) shows highly nonlinear signal transmission, where the worst-case EVM is 8.5%. Such high power amplifier EVM results are unacceptable for practical system implementations.

2.4 Efficiency

The goal of an RF power amplifier is to drive a load (e.g. an antenna) with enough power to reliably transmit a signal. The ratio of transmitted power to the DC power consumed by the amplifying device is called efficiency. Thus, the efficiency, also known as DC to RF efficiency or collector drain efficiency, can be defined as:

$$Efficiency = \eta = \frac{P_{RF}}{P_{DC}} \tag{2.17}$$

An ideal power amplifier has an efficiency of $\eta = 1$ or 100%. Of course, that is not the

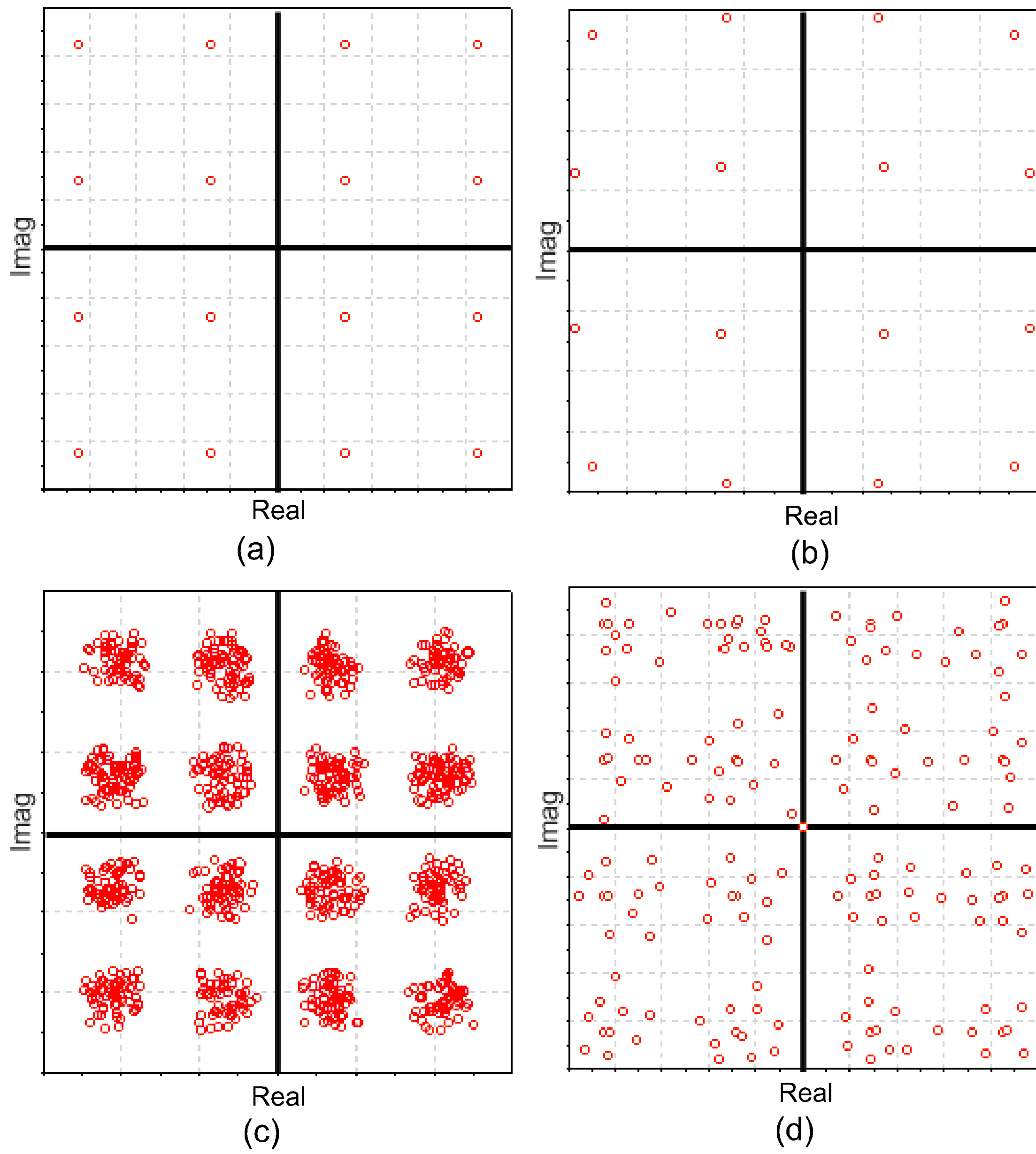

Figure 2.10: Constellation diagrams for a 16-QAM modulated signal for: (a) an ideal signal, (b) signal with EVM of 0.5%, (c) signal with EVM of 5.1%, (d) with EVM of 8.5%.

case in reality—efficiency is always below 100% because some power is dissipated in the active device and other PA components such as matching networks. At mm-wave frequencies and in silicon-based technologies, efficiency drops even more significantly. This limitation can be further understood by considering the Power-Added Efficiency (PAE).

Note that the above efficiency definition does not account for the input drive level. On the other hand, PAE is defined as

$$PAE = \frac{P_{out} - P_{in}}{P_{DC}} = \frac{P_{out} - \frac{P_{out}}{G_T}}{P_{DC}} = \frac{P_{out}}{P_{DC}}(1 - \frac{1}{G_T}) \tag{2.18}$$

which accounts for the gain (G_T) through the PA. This definition is more insightful than the collector efficiency definition because it indicates what input drive level is needed to achieve a particular efficiency value. Hence, to increase efficiency while keeping the same output power the gain needs to be increased. However, for higher amplification (higher gain), more DC power is required.

Efficiency is one of the more important parameters that define a power amplifier, especially in mobile communications where battery life is very important. However, currently envisioned commercial mm-wave applications are not designed for portable devices or for systems where battery life is crucial. In such cases, efficiency is not as critical a specification, although it is always desirable to optimize efficiency while meeting the other specifications.

2.5 Power Amplifier Classification

Power amplifier classes are well known and are based on where the quiescent current or voltage is biased in relation to the peak current excursion. The PA classes range start from most linear (Class A) to highest efficiency (Class F, G, etc.). Unlike Class A, AB, B and C, Class D, E, F, etc. are known as switching amplifiers, where transistors operate like a switch. In other words, the collector/drain voltage as well as the current are non-sinusoidal. Figure 2.11 shows IV-loadline curves for Class A, AB, B and C and the corresponding output voltage and current swings. It can be seen that to achieve high efficiency, the transistor is biased such that the

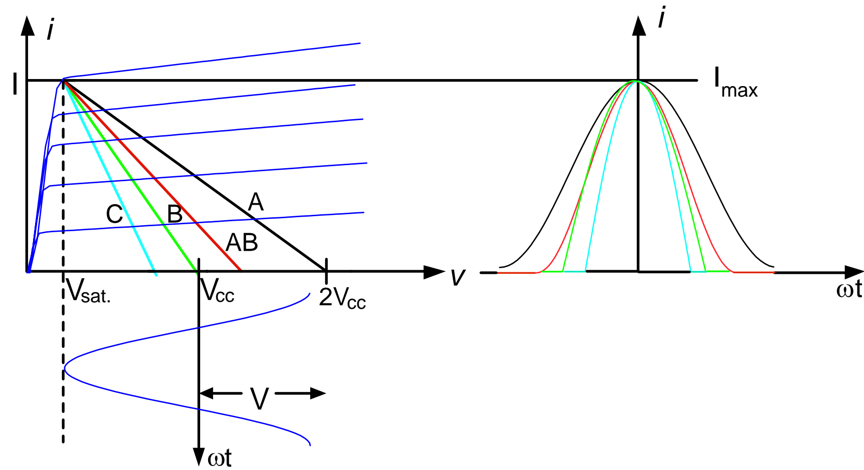

Figure 2.11: Loadlines for different classes on IV curves with collector (output) voltage and current waveforms.

waveform excursion turns the transistor off ($i_c = 0$) for a certain amount of time in each cycle/period; power dissipation is zero during that time, thus increasing the efficiency. However, the non-sinusiodal waveforms introduce significant non-linearity. It should be noted that, theoretically, higher classes like B and C produce the same output power as Class A because the peak voltage and current should be the same (Figure 2.11). Also, observe that for higher classes (non-black lines) the slope of the loadline is steeper, which indicates a smaller load resistance. Therefore, as the transistor is biased closer to the cutoff threshold, such as in deep Class AB, B and C, the load resistance *must be reduced* to produce the maximum output power possible.

2.6 Biasing

Providing reliable *dc* bias to a RF power transistor is a crucial step in PA design process. As shown in the section above, the PA operating classes are strongly related to the bias point selected, and different classes can be achieved by simply changing the bias point. Consequently, any change in the bias point can result in moving into an undesirable region of operation for a power amplifier from the point of view of efficiency or linearity. Low-frequency oscillations (instability at very low frequencies) are also an important issue, and bias circuits need to be designed to prevent such

oscillations. Based on the technology/transistor type being used, bias circuits may also need to perform additional tasks. For example, as will be discussed further later, in SiGe HBT technology the bias circuits are usually designed to provide impedance looking from the base of the HBT to be no more than 500 Ω [31] [32]. Other issues that are important in bias design are insensitivity to temperature variations, and good electrical isolation between *dc* and RF signals. Two basic types of biasing techniques are used: Constant Current Biasing and Constant Voltage Biasing

2.6.1 Constant Current

When a bias circuit provides constant current base bias to a BJT/HBT, it is categorized as Constant Current Biasing (CCB). A typical bias circuit of this type is shown in Figure 2.12. Bias transistor Q_2 effectively forms a current mirror with the power transistor, Q_1, and is scaled in area by the ratio of the reference current to the desired power transistor collector current. $R_1$and R_2 are known as ballasting resistors, and the ratio between their values should be equal to that of Q_1/Q_2 [27]. Ideally, this bias circuit provides constant current to the transistor base, and this is usually the case with linear small-signal amplifiers. However, due to the high temperature and high current exhibited in power amplifiers, this bias circuit deviates from its ideal performance and at some bias levels may not be able to provide the required current [33]. Constant current biasing may also degrade linearity since as the drive level increases, the base current does not change because it is set by the constant current biasing, while the base-emitter bias voltage, V_{BE}, drops, which drops the collector current bias point, and thus causes the lower swing of collector current to clip at high input levels.

2.6.2 Constant Voltage

The other basic biasing approach is to provide constant voltage across the base-emitter junction of a BJT or HBT; thus, this is referred to as Constant Voltage Biasing (CVB). CVB does not suffer as much from the clipping problem seen with CCB because as input power increases, the base *dc* current of CVB increases, and so does the collector *dc* current. However, in monolithic designs, the standard CVB

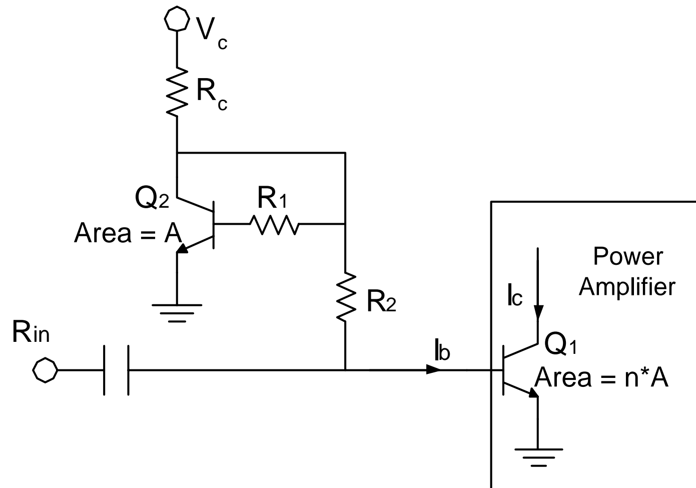

Figure 2.12: Constant current bias circuit. Q_1 has area that is n times larger than that of Q_2.

technique is not feasible because of the need for a large inductor as an RF choke at the base. This has lead to use of an op-amp to provide a constant voltage at the base, doing away with the need for an RF choke [18]. However, such a scheme is difficult to implement at mm-wave frequencies, where the output impedance of op-amp would drop sharply. Although bias through a resistive divider network could be used in place of an inductive RF choke, it is undesirable because resistors are susceptible to temperature variations, and thus constant bias source might not be achievable at high power levels. Also, isolation between RF and *dc* might be degraded. In many cases where the impedance looking into the input of an active device is large, a resistive divider with prohibitively large equivalent resistance might be required to prevent RF signal from leaking through the bias network.

2.6.3 Temperature-Compensation Biasing

As mentioned above, undesirable shifts in bias can occur as the drive level is increased; in BJT/HBT designs this problem is usually associated with junction temperature increase with drive level, which causes the V_{BE} to drop. The drop in V_{BE} can significantly change the PA operation region. For example, a PA biased in Class B, close to cut-off threshold, can easily enter the more nonlinear Class C operation with a

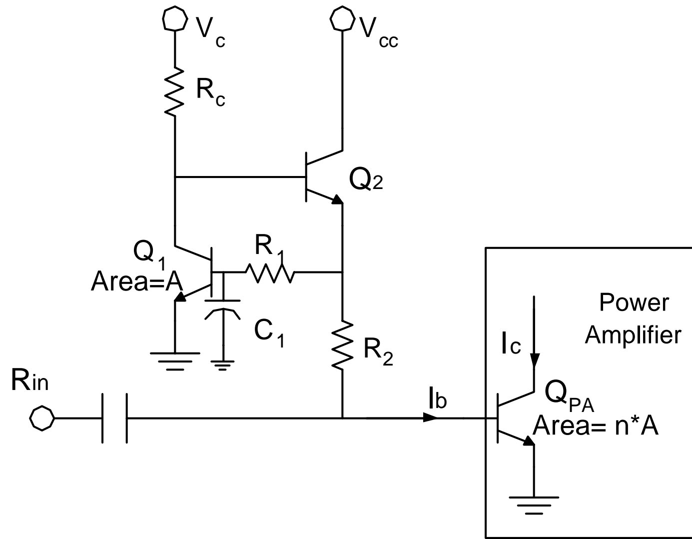

Figure 2.13: Typical temperature-compensation bias circuit.

slight drop of V_{BE}. The bias circuit in Figure 2.12 is highly sensitive to this problem. A small modification and careful selection of ballasting resistors $R_1$and R_2 can lead to a temperature compensating V_{BE} bias (Figure 2.13), where V_{BE} drop is lower and more gradual. This gradual drop is achieved because the ballasting resistor *share* the base voltage with the base-emitter junction. Hence, variation in base voltage is not fully reflected on V_{BE}. Note that the capacitor C_1 is advantageous because it forms a low-pass RC filter, isolating the RF signal from the bias circuit.

The required values of base current is relatively high in power amplifiers, and usually the supply voltage V_c in Figure 2.12 is often unable to provide such current levels. Thus, device Q_2 is used to increase Q_{PA} (Figure 2.13) base current levels [33]. More complex and robust temperature and process insensitive bias circuits can be designed and are used frequently in appropriate applications [33] [31].

2.6.4 Current Source Biasing

In RFIC design, differential pairs are typically biased by a current source connected to the emitters of the differential pair, as shown in Figure 2.14. However, this bias implementation is often not used in power amplifier design, where highest output voltage swing is desired, due to the fact that the available output voltage headroom

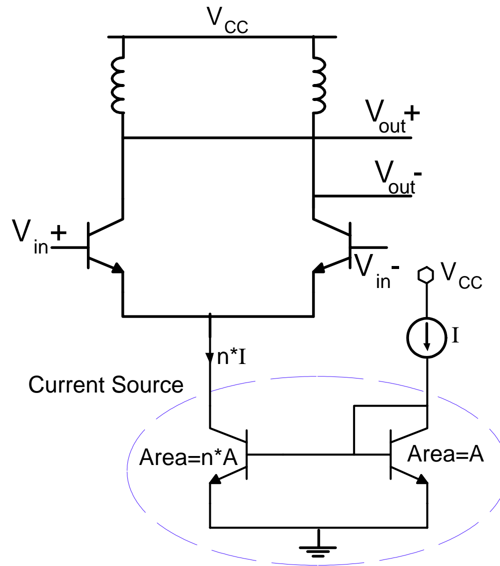

Figure 2.14: Biasing of differential pair using a current source connected to the emitters.

is reduced. In addition, the current source increases the standby power dissipation which in turn degrades the efficiency. Nevertheless, such an implementation has some benefits such as improving common-mode rejection ratio (CMRR), stability, and improving RF to *dc* isolation.

2.7 Design Goals

The purpose of this research is to further the development of silicon-based power amplifier design at microwave/mm-wave frequencies. To this end, silicon-based technologies such as RFCMOS or Silicon Germanium (SiGe) are considered. The power amplifiers designed in this research use Silicon Germanium Heterojunction Bipolar Transistor (HBT) technology. Even though silicon-based technologies lag behind GaAs in terms of PA design maturity and have some drawbacks compared to III-V technologies as discussed in Chapter 1, the market is ultimately driven by cost-effectiveness, even at the expense of lesser performance. Therefore, the possibility of higher level of integration, which would significantly reduce cost, has been driving research in this area.

Before starting any design it is necessary to understand what the goals and parameters

that needs to be met in order to achieve a successful product. Since the use of silicon-based technology for microwave power amplifiers is still in its infancy, the main objective of this work is to demonstrate a working circuit design at 30 GHz with reasonable output power while exhibiting high linearity. Efficiency is not as important due to the fact that applications such as LMDS, WLAN, etc. do not operate on a portable power supply, and thus, power consumption is not a crucial aspect. Moreover, since the specific application at mm-wave frequencies was open at this time, detailed system other specifications were not established. However, some general goals were defined for the power amplifiers.

2.7.1 Linearity

Figures of merit concerning linearity have been defined earlier. The two figures of merit used here are 1-dB compression point, P_{1-dB} and EVM. One future application at 30 GHz is that for local multipoint distribution service (LMDS), which is a high data rate application based on IEEE802.16 standard. For such a system, an EVM value of 3-5% or less is required for a PA.

2.7.2 Output Power

Point-to-point applications usually operate at distance of 1000 meters or less. Therefore, lower output powers are required, compared to a mobile phone which requires high output power to transmit signal to base stations that may be on the order of a km or more away. If the PA is operated at saturation, where output power does change much with increase in input power, nonlinearities would be too high and EVM requirement would be difficult to meet (Figure 2.10). The output power for millimeter-wave applications, such as LMDS, range between 12 to 15 dBm [1] [11]. Hence, the goal is to operate the PA at its 1-dB compression point with output power in the vicinity of 13 dBm or approximately 20 mW.

Table 2.1: Atmel SiGe2RF power amplifier design goals

Design Parameters	**Goal**
Output Power	$\approx 12 - 13$ dBm @ P_{1-dB}
Linearity	EVM of $3 - 5\%$ @ P_{1-dB}
Efficiency	$undefined$
Die Size	$< 2.0mm^2$

2.7.3 Efficiency

Battery life is not as critical an issue in fixed point-to-point applications as it is with mobile wireless devices such as a cellular phones. Applications such as inter-vehicle communication, LMDS or WLAN typically do not depend on a limited battery supply. Thus, there is no specific efficiency target for this project. However, since in these applications linearity is a much higher priority, in this work efforts are made to optimize linearity, which will obviously degrade the efficiency. Note that, however, it is always beneficial to improve efficiency as long as all other targets are met.

2.7.4 Die Size

Even today some commercial GaAs PA ICs, mostly at lower frequencies, still have some parts off chip, such as matching networks. The motivation to design microwave power amplifiers in silicon-based technologies is to demonstrate integration possibilities of PA with other transceiver components. Therefore, in this research fully monolithic power amplifiers are being designed where the bias circuits, RF chokes, DC blocks and matching networks are on-chip. However, available wafer space is limited. Therefore, the target die area is less than 2.0 mm^2.

Table 2.1 summarizes the design goals explained above.

2.8 Approach

For millimeter-wave range power amplifier design the most fundamental design, step is the selection of a technology that can produce reasonable gain at 30 GHz and

above. Amplification is directly related to the maximum output power that can be achieved, and together with the requirement of high linearity, the overall approach is established. SiGe HBT technology has been discussed compared with GaAs and RFCMOS in Chapter 1, and its merits over GaAs and CMOS have been explained, leading to the selection of SiGe HBT as the technology of choice. The remaining discussion in this section is devoted to explaining the overall approach such as number of amplification stages, PA classification, monolithic matching network implementation and circuit topology for the power amplifier. As has been mentioned earlier, two different power amplifiers were designed, one in Atmel SiGe2RF technology and the other in IBM 8HP technology; the design approaches detailed below generally apply to both cases.

2.8.1 Monolithic Design

Monolithic designs in silicon are a challenge because of the higher loss incurred in passive structures, such as on-chip input and output matching networks. These losses will degrade the gain, the output power, and also whatever PAE that can be achieved. As shown in Chapter 3, matching networks will be designed using CPW lines and the loss (dB/mm) has been estimated for various structures. Provided that reasonable gain can still be maintained and all the parameters are met, it is safer to over-estimate the loss during computer simulations to avoid any unexpected results during measurements.

2.8.2 Single-stage Design

When the required gain cannot be achieved from a single amplification stage, the logical extension is to use multi-stage amplification, where a number of stages are cascaded together to produce required overall gain. Also, power amplifier output stages usually have large peripheries for carrying high currents and thus have larger parasitic capacitance, which presents problems in integrating the PA with other transmitter or transceiver blocks. Thus, it is seen that generally a multi-stage PA designed is preferred.

For short-range point-to-point mm-wave applications, the required output power is

only in the range of 12-13 dBm, and advanced SiGe HBT technologies like IBM's 8HP or 9HP provide enough gain with single transistor that acceptable low drive levels can achieve the target output power. Hence, considering wafer space limitations for this research, it was decided to use single-stage amplification for the power amplifiers.

2.8.3 PA Class Choice

The two fundamental parameters that need to be met in this PA design work were the output power and high linearity. Power amplifier classifications were briefly mentioned in Section 2.5, and it was seen that Class A is the most linear region of operation. As has been detailed earlier, the efficiency is not as important a parameter in this work and hence it was decided to bias the power amplifiers in Class A for maximum linearity possible. Also, one underlying issue that has been discussed earlier is the lack of accurate large-signal/nonlinear modeling of power transistors. A side benefit of operation in Class A is that the simulation results from the given model in the design kit can be trusted to a greater degree than if the PA were to be biased in more nonlinear classes such as AB or B. However, caution is still needed for this mm-wave frequency design even with Class A biasing.

2.8.4 Differential Topology

The last step in finalizing the design approach was to select a single ended or differential (balanced) topology. In particular, the design will be fully differential one with both input and output ports being differential. Since power amplifiers typically do drive a single-ended antenna, they are often designed as single-ended. However, most of the integrated transmitter blocks are differential topologies, and especially the system that was the basis of this design was fully differential as well. One of the most important advantages of using a differential topology, especially in this design where linearity is critical, is the suppression of even-order harmonics that would be present in a single-ended topology.

Using balanced circuit design is also a way of power combining. Large voltage swings are needed to produce desired output power. However, the breakdown limits in SiGe are much lower than that of GaAs and a single SiGe power transistor might not be

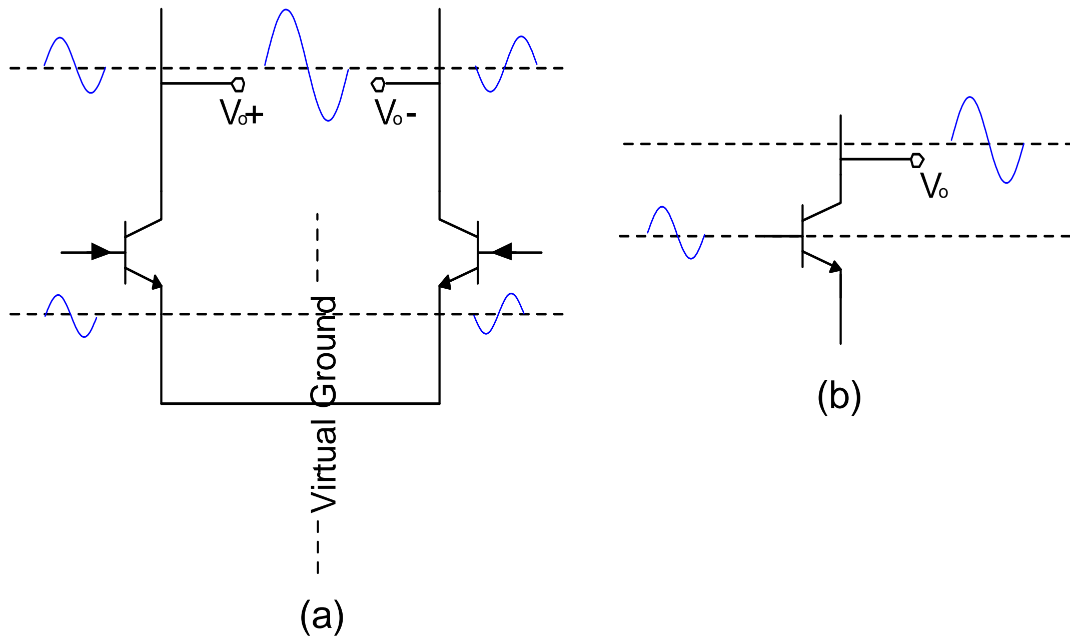

Figure 2.15: (a) Single-ended toplogy, (b) differential topology showing waveform combination.

able to handle the entire swing. In this case, a differential topology helps in easing the voltage requirement at all the nodes including the output (collectors). In an ideally balanced circuit, each side of the differential pair carries identical 180° out-of-phase signals with equal amplitude. The differential output is the difference of these two signals, as can be seen from Figure 2.15(a), but since they are out-of-phase each side will add to give differential output signal of twice that of each transistor. Hence, required output power can be obtained by adding power generated from each device. On the other hand, generating the required output power through just one device (Figure 2.15(b)) may lead to device breakdown due to the larger required output voltage waveform.

A differential topology has several other significant advantages. Since each device is 180° out-of-phase, a virtual ground node is established at the connected emitters of the transistors (Figure 2.15) which helps in increasing common-mode rejection ratio (CMRR). CMRR is important because undesirable signals such as noise, variations in power supply,etc. appear as common mode components.

Also, ground lead inductance is of major concern in power amplifiers. Not taking

inductance of a ground lead connected to differential pair emitters into consideration can significantly affect linearity [34]. Moreover, power amplifiers generate large *ac* current and emitter-connected lead inductance can result in large voltage spikes that can alter the potential of the common port and thus make a ground connection unreliable. With a differential implementation, such spikes appear as common-mode components and their effects would be reduced. Another advantage of having an established virtual ground, is the possibility of creating desired isolation between RF and *dc* signals by connecting bias signals/lines to this virtual ground, thus attenuating any RF signal from entering the bias circuit, as will be seen later. While integrating several devices or circuit blocks together, loading at the output due to subsequent cascade stages is a common concern. Usually, a buffer stage, such as an emitter follower, is used at the output of a device or circuit block to avoid loading issues. However, by implementing balanced topologies, which increase the input and output impedance compared to the single-ended case, the loading problem can be mitigated.

Chapter 3

Finite Ground Coplanar Waveguide Transmission Lines

Microstrip line and CPW are the two most widely used transmission structures in conventional (GaAs) millimeter-wave monolithic integrated circuits (MMIC) design. Coplanar waveguide transmission lines have advantages that make them attractive for use in MMIC compared to conventional microstrip lines. Some of these advantages are:

- CPW technology can operate in quasi-static mode up to very high frequencies because it displays less dispersion compared to microstrip on a given substrate.
- In CPW technology, ground planes are at the top plane, eliminating the need for via holes in a III-V process (which are required for microstrip structures to make connections to the backside ground plane) thus, reducing the cost.
- Coplanar ground planes improve isolation between adjacent signal lines.
- Balanced/differential circuit topologies can be realized more easily in CPW/slotline environment [35].
- Due to the coplanar ground planes, on-wafer measurements using standard ground-signal-ground (GSG) or, in case of differential measurement, ground-signal-ground-signal-ground (GSGSG) probes is highly compatible with CPW structures.

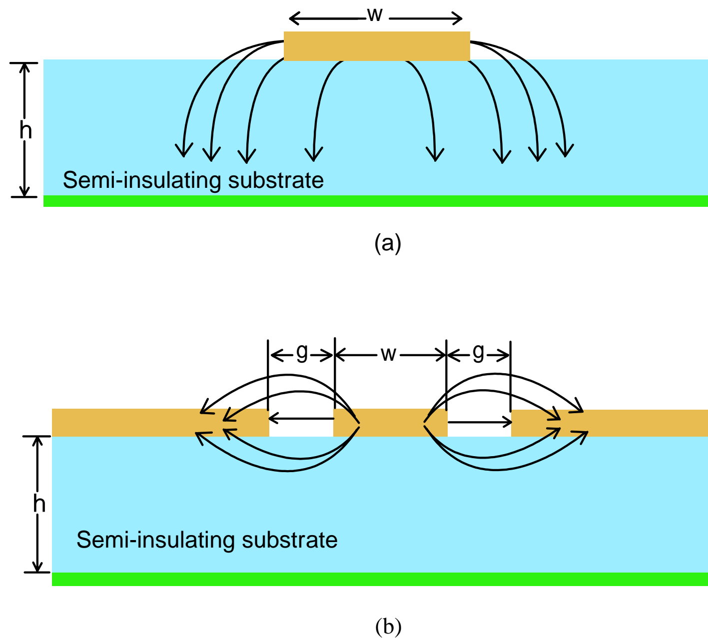

Figure 3.1: (a) Standard microstrip, (b) standard coplanar waveguide CPW with lower ground plane.

This chapter first provides an introduction of standard mm-wave transmission line structures such as those used in III-V technology, and then the design of FGC lines for use in the integrated millimeter-wave monolithic power amplifier is presented. A comparison between finite ground CPW (FGC) and TFMS used in Si IC technology is then presented.

3.1 Conventional MMIC Transmission Lines

Due to the typical high-resistivity (low-loss) substrate, transmission line structures can be directly fabricated with the bulk substrate (e.g. GaAs) acting as the dielectric. Figure 3.1 shows conventional Microstrip and CPW in GaAs technology. The distributions of electric fields for both Microstrip and CPW are also shown. Note that the electric fields penetrate the substrate more so for microstrip than for CPW. Microstrip technology has been preferred in MMIC design because it has higher effective permittivity (ϵ_{eff}) compared to CPW, which leads to smaller dimensions. CPW also has higher conductive losses than microstrip (for a given conductor width).

Coplanar waveguide transmission lines can be easily analyzed by assuming that the ground planes are infinite in width as shown in Figure 3.2(b). However, in practice this is not the case and finite ground plane width must be taken into consideration for more accurate analysis [36] [37].

3.2 Transmission Lines in Si Technology

Unlike conventional III-V technology, in standard silicon based technologies, the bulk substrate cannot be used as a dielectric for transmission line structures due to high losses at microwave/millimeter-wave frequencies, which would result in unacceptable levels of line attenuation[1]. Therefore, interconnect (metal) levels above the substrate are rather used with inter-metal dielectric as the transmission line dielectric. The two primary options are thin-film microstrip (TFMS) and finite ground CPW (FGC). Figure 3.2 shows a TFMS structure and a FGC structure. The conductor strip of a TFMS is fabricated on a higher metal layer and a ground plane is placed above the silicon substrate to contain the electric fields, preventing them from interacting with the substrate. Since the dielectric height is much reduced in TFMS, narrower conductor width is required to achieve standard 50 Ω characteristic impedance. It has been shown that at mm-wave frequencies, FGC structure can provide better performance compared to TFMS line [38] by carefully selecting the CPW dimensions for a given characteristic impedance. In addition, FGC has higher ϵ_{eff} than TFMS, which leads to smaller physical lengths for a given electrical length. Based on this, and the other advantages described above, FGC technology was chosen for the PA designs in this thesis.

The FGC structure has some unique issues that should be taken into consideration during the design. The remainder of this section discusses issues such as truncated ground planes, parallel plate resonance and slotline mode suppression.

[1] Note that some technologies provide an option for "high-resisitvity" silicon substrate. Although this would reduce losses in passive components, the cost could be high enough to offset cost effectiveness of using a silicon technology.

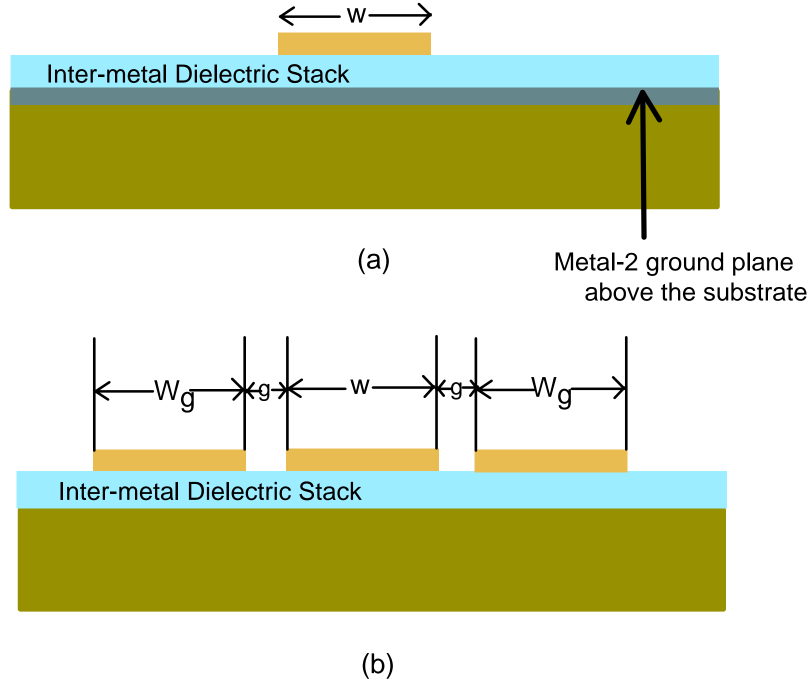

Figure 3.2: (a) Thin-film microstrip, and (b) finite ground CPW on a silicon substrate.

3.2.1 Truncated Ground Planes

In a coplanar waveguide, the majority of high-frequency current in the ground planes is conducted at their inner edges parallel to the center conductor. However, due to the presence of the lower ground plane, a parallel plate waveguide structure can be formed between the top and bottom ground planes [39]. One solution to this problem is to short the top and bottom planes together using rows of vias. However, a simpler solution is to truncate the upper ground planes as shown in Figure 3.2(b) while keeping the overall structure width $(w + 2_g + 2W_g)$ less than a quarter of a dielectric wavelength $(< \lambda_o/(4\sqrt{\epsilon}))$ [36]. This dimension ensures that no parallel plate resonance occurs below the upper limit of operating frequency.

3.2.2 Slotline Mode Suppression

Another aspect of FGC design is the need for suppression of the slotline, or even, propagation mode [40] [41] . This is a parasitic mode that can arise from voltage

differences between the two ground planes. Such potential differences can occur due to asymmetries in the slot lengths of the surrounding ground planes that result from discontinuities such as T-junctions, where transmission line stubs join the through line. It is evident that the two ground planes should be somehow connected together near any discontinuity such as T-junctions to force them to the same potential and suppress any even mode propagation. There are three common ways to provide such a connection: bond wires, airbridges and dielectric crossovers [41]. Figure 3.3 shows these types of ground plane connections [42].

At mm-wave frequencies, typical bond wire lengths result in a high impedance due to their inductive component, and thus do not provide the required potential equalization. Airbridges are prone to damage, and incorporating additional process steps for such structures in commercial IC technology can be expensive. On the other hand, dielectric crossovers do not suffer from these drawbacks and are often used. However, in order to use this method a dielectric crossover technology must be included in the IC process and be available for use. The Atmel SiGe2RF technology does not specifically provide a crossover dielectric layer. However, a straightforward alternative is to use inter-metal vias to connect the two ground planes to a lower metal layer crossunder, metal-2 in the Atmel case 3.4. The disadvantage of dielectric crossunders is the higher capacitance, which impacts matching network design. Nevertheless, the finite ground planes of all the FGC lines designed in this work used cross-under connections.

3.3 FGC Design Considerations

Before designing on-chip transmission line structures, it is important to examine the integrated circuit (IC) process in which the structures are to be fabricated. Parameters such as the substrate thickness, metal thickness, and dielectric layer primitivities are important factors that impact the selection of the type of transmission line and its structural dimensions. The Atmel SiGe2RF IC process used in the design is a three-metal process with a low-resistivity silicon substrate and silicon-dioxide dielectric between its metal layers Figure 3.5. The 180 μm-thick silicon substrate has a low resistivity of $\rho = 20\ \Omega - cm$ and permittivity of $\epsilon_r = 11.9$. The SiO_2 dielectric has a permittivity of $\epsilon_r = 3.9$. There is also a 0.4 μm thick silicon nitride passivation layer

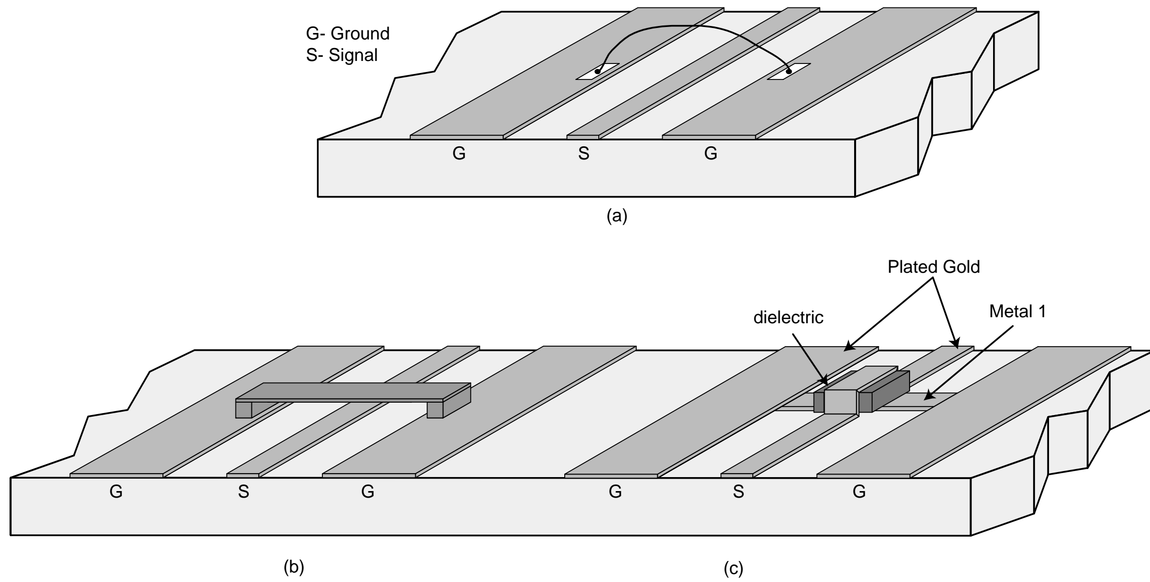

Figure 3.3: Methods of slotline mode suppression: (a) bond wire, (b) airbrdige, (c) dielectric crossover. Permission granted by Michael Chapman to reproduce the figure [42].

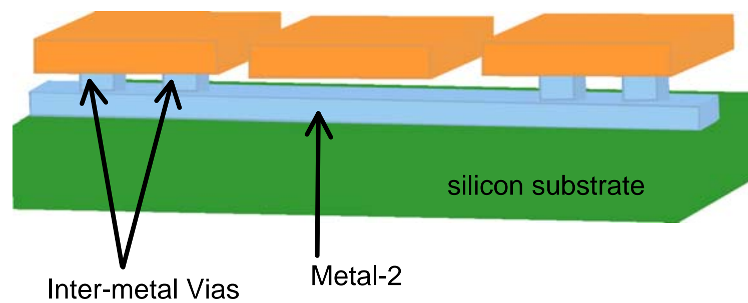

Figure 3.4: Connection of coplanar ground plane using inter-metal substrate vias.

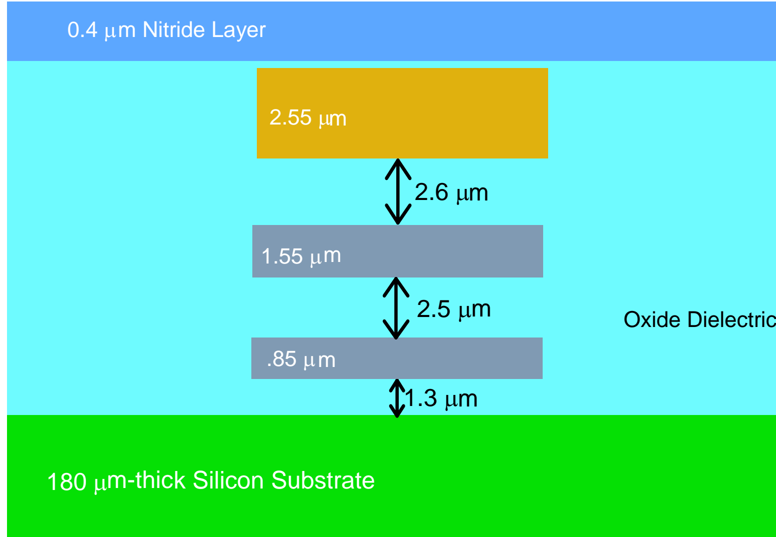

Figure 3.5: FGC substrate definition and dimensions used for simulation in IE3D.

with $\epsilon_r = 7.0$ over the chip. The top metal (Aluminum) layer is 2.55 μm thick and is 6.4 μm away from the substrate.

In designing transmission line structures, the two most important design parameters are the transmission line characteristic (Z_o) and the attenuation loss (α). The finite-ground coplanar waveguide lines were chosen to have a characteristic impedance of 50 Ω. Unlike standard CPW, Figure 3.1(b), where the characteristic impedance depends on the ratio of the signal conductor width to total cross section, FGC, Figure 3.2(b), line's characteristic impedance is determined mostly by the ratio of the signal conductor width to the overall slot width and is not as significantly affected by the width of the ground planes. Line attenuation for FGC comparable to that of CPW can be produced if the ground plane width is set to the overall slot width ($w + 2g$) or larger [42]. However, it is important to verify that indeed this results in low dispersion and CPW mode is maintained once conductor width is chosen.

It is not easy to analytically determine the characteristic impedance and the attenuation loss of a FGC line. However, EM simulations through IE3D can be used to simulate different dimensions that would result in a line with a particular Z_o (e.g. 50 Ω). The conductor width (w) and the slot width (g) are the two dimensions that are varied during simulation while the ground plane widths (W_g) are set according to the guideline $W_g = w + 2g$. Although larger combinations of conductor width and slot

width can produce required characteristic impedance, practical considerations such as high attenuation, dispersion and power loss, moding, etc. limit the range of such combinations. Line attenuation and center conductor width are important consideration but are inversely related. As the conductor width is reduced, significant line attenuation can occur because the cross sectional area of the conductor is reduced; hence lowering the conductivity of the metallic plane. Therefore, narrower conductor widths will not yield attenuation levels that would be acceptable, especially in the case of high-frequency power amplifier design where gain and output power are of premium.

3.4 Simulation of FGC Lines

The substrate in IE3D was defined as shown in Figure 3.5. Note that to reduce computation time the actual 180 μm substrate, the oxide dielectric and nitride passivation layers can all be replaced in IE3D with a single *substrate* with weighted average dielectric parameters calculated from the individual values of the layers mentioned above. However, due to the different modes supported by FGC and the need for highest level of accuracy at millimeter-wave designs, the substrate structure was defined exactly as shown in Figure 3.5. Several combinations of center conductor width and slot gap were simulated to achieve a 50 Ω FGC line. First, a conductor width was chosen and then an approximate slot gap was chosen to produce a 50 Ω characteristic impedance. The ground plane widths were set equal to $w + 2_g$. The structure was then simulated to see if the required characteristic impedance was achieved. Together with Z_o, other important transmission line parameters were calculated such as attenuation loss (α) and effective permittivity (ϵ_{eff}). Several iterations were carried out to choose an appropriate structure.

The overall dimension ($w + 2_g + 2W_g$) is restricted to $\lambda_o/(4\sqrt{\epsilon_d})$ as detailed above for the suppression of parallel plate mode. Although the expected operating frequency in this work is at 30 GHz, as a safety factor it was decided to use a frequency of 40 GHz to calculate the dielectric wavelength. For a highest frequency of interest of 40 GHz, the overall line width should be less than 490 μm. This in turn limits the conductor width to not exceed approximately 70 μm. Figure 3.6 shows the attenuation loss versus frequency. Figure 3.7 shows the effective permittivity (ϵ_{eff}) of different conductor

widths versus frequency. Dispersion can be inferred from variations in the effective permittivity. Choosing a large conductor width would enhance the possibility of even propagation mode and would require excessive die area. In order to minimize die area, which was a major design concern, and considering the simulated results for attenuation and effective permittivity, a conductor width of 30 μm was chosen, corresponding to slot gap of 15 μm and ground plane widths of 60 μm. Figure 3.7 also supports the selection of the conductor width because the 30 μm line does not exhibit high dispersion; in other words, its effective permittivity is relatively flat over the frequency range of interest.

The reason to set W_g equal to $w + 2g$ was to obtain attenuation performance comparable to that of infinite ground plane CPW. Hence it is useful to verify that the chosen FGC structure approximate ideal CPW. Figure 3.8 shows the characteristic impedance and line attenuation versus the ratio of ground conductor width to center conductor width (W_g/w), where w is fixed to 30 μm with g of 15 μm. It can be seen that both the characteristic impedance and line attenuation change significantly as (W_g/w) is varied, which is similar to standard CPW trends [36] [43]. Also, note from Figure 3.8 that beyond a W_g/w ratio of 2.1, the change in the attenuation parameter is minimal, which validates the choice of $W_g = 60$ μm.

3.5 Thin-film Microstrip vs. Finite-ground Coplanar Waveguide in Atmel Technology

A Full-wave electromagnetic (EM) simulation software, such as IE3D [44], is typically used to obtain accurate simulations for analyzing transmission line structures, such as FGC and TFMS. Acceptable line attenuation and reasonable structure dimensions for a 50 Ω characteristic impedance were the two requirements analyzed when comparing TFMS and FGC. Even though matching stubs can be designed with a characteristic impedance not equal to 50 Ω, due to the number of iterations of EM simulation required for different line geometry cases, a standard 50 Ω line geometry was used through out the design to simplify design process.

Although a TFMS with the same conductor width dimension as a FGC center conductor (w) will produce lower loss due to electric field confinement within the dielectric, a

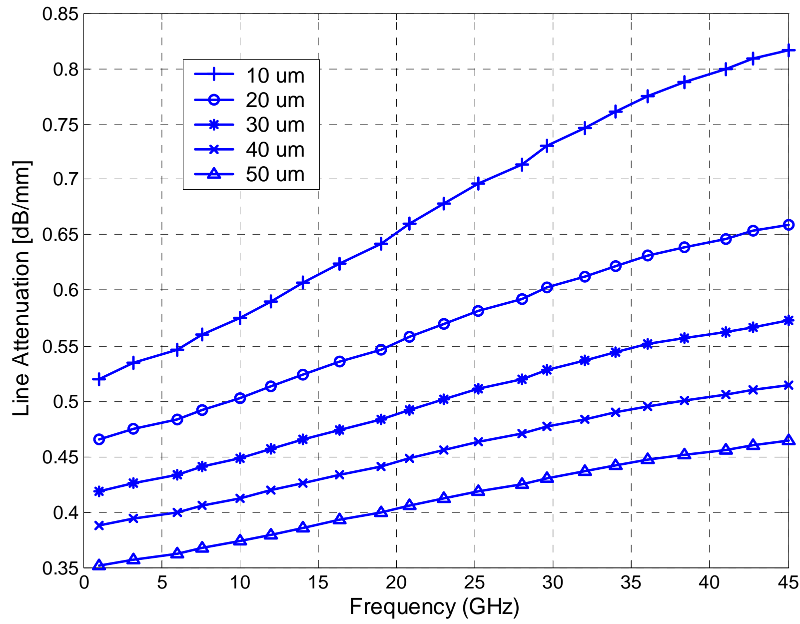

Figure 3.6: Simulated attenuation loss vs. frequency for various center conductor widths.

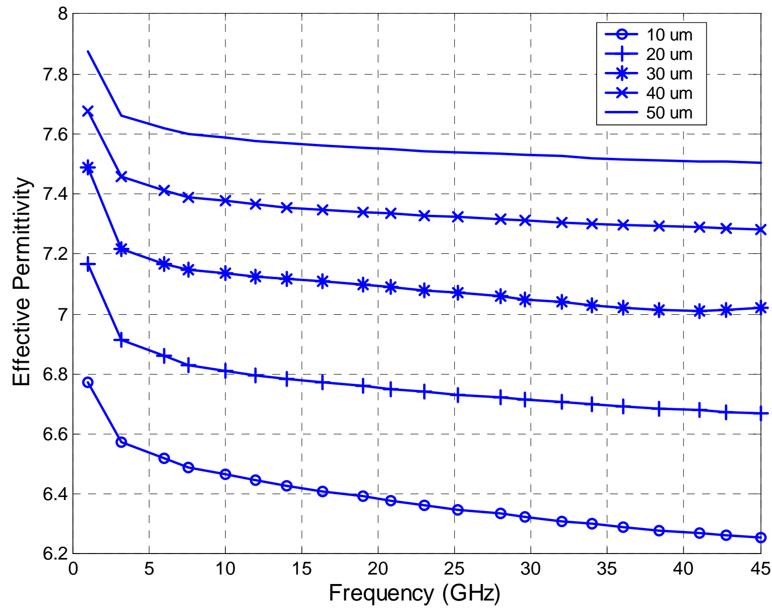

Figure 3.7: Simulated effective permittivity vs. frequency for various center conductor widths.

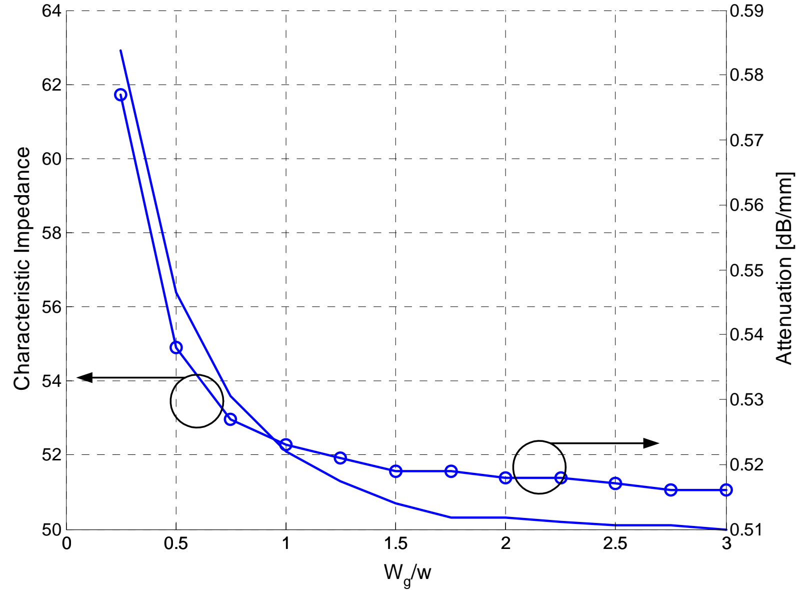

Figure 3.8: Simulated characteristic impedance and attenuation loss vs. ratio of the ground plane width (W_g) to center conductor width (w) for $w = 30$ μm, $g = 15$ μm.

much narrower conductor width is required to achieve 50 Ω characteristic impedance. (Note that in the Atmel SiGe2RF process the dielectric substrate thickness if fixed to 5.1 μm between the top metal-3 and metal-1). Figure 3.9 and 3.10 shows the characteristic impedance and attenuation loss of various TFMS conductor widths versus frequency, respectively. It can be seen that wider conductor widths result in lower characteristic impedances and to achieve a Z_o of 50 Ω, a narrower conductor width of ~7 μm is required, which increases the attenuation loss. Looking at Figure 3.11, line attenuation of 7 μm TFMS is approximately 0.7 dB/mm, which exceeds that of 30 μm FGC at approximately 30 GHz. Moreover, as explained in the introduction, using balanced transmission line structures such as CPW/FGC are more suitable for millimeter-wave, differential circuit topologies. Additionally, FGC lines have higher effective permittivity compared to TFMS, and thus the physical length would be shorter for a required electrical length. Hence, the choice of FGC over TFMS is validated.

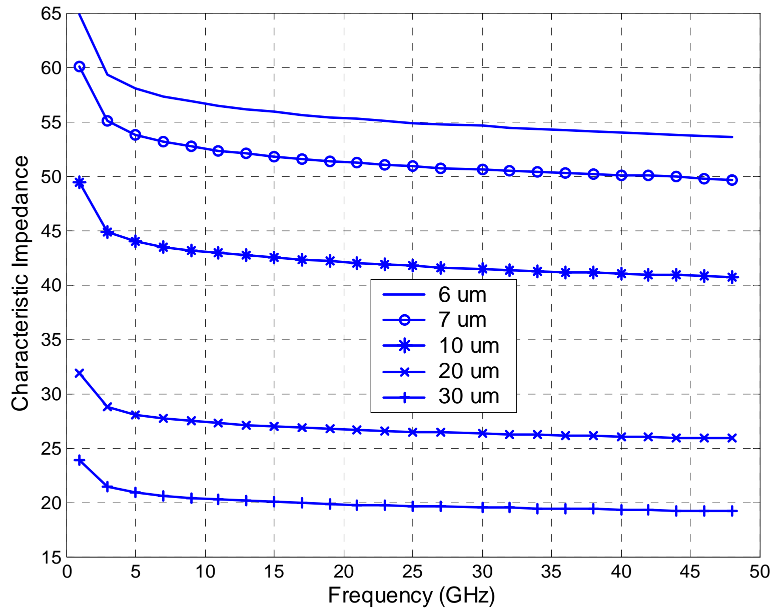

Figure 3.9: Simulated characteristic impedance of TFMS vs. frequency for various conductor widths.

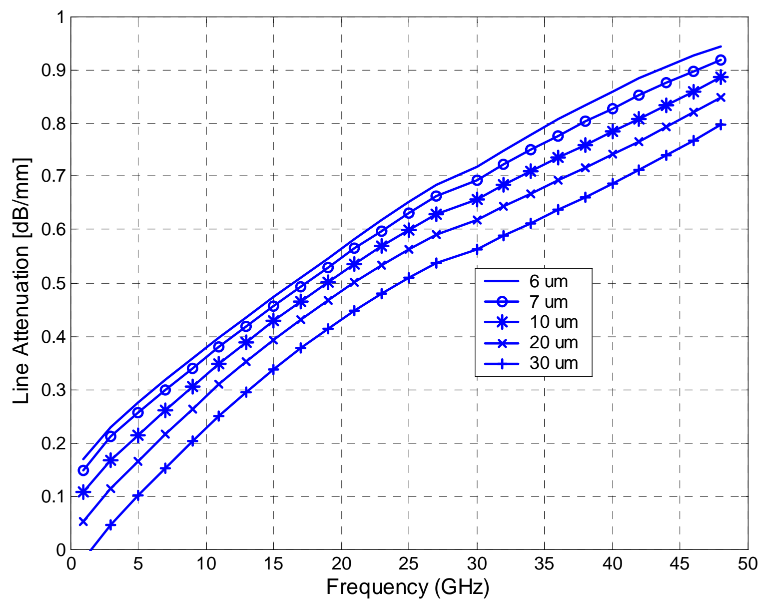

Figure 3.10: Simulated line attenuation of TFMS vs. frequency for various conductor width

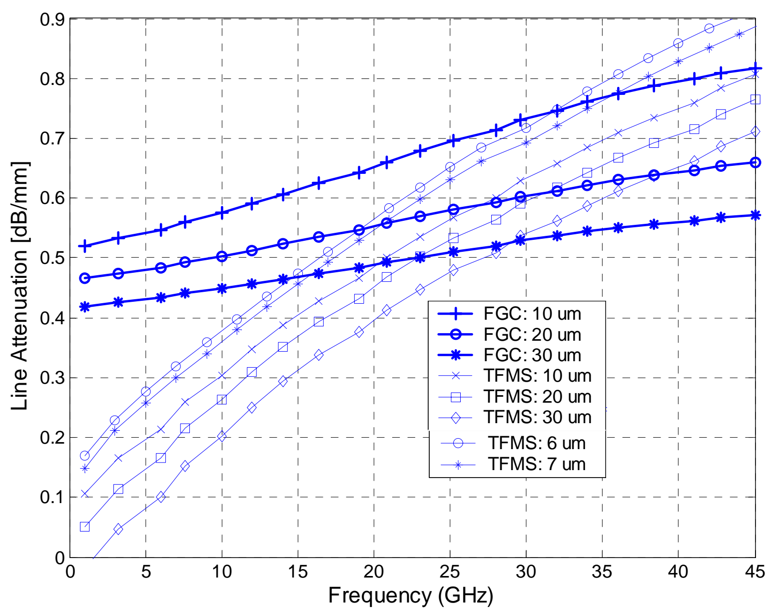

Figure 3.11: Line Attenuation for FGC vs. line attenuation for TFMS.

Chapter 4

Atmel SiGe2RF PA Design and Simulation

In Chapter 2 the design goals and approaches for the power amplifier designs in this work were detailed. If a successful design is achieved, the proposed 30 GHz SiGe HBT power amplifier will represent only a handful of silicon-based millimeter-wave monolithic, fully differential power amplifiers reported in the literature to date.

This chapter describes the detailed circuit design using the approach presented earlier in Section 2.7, as well as key simulation results for the final power amplifier circuit.

4.1 Iterative Design Process

Power amplifier designs are often quite iterative in nature. This is mainly due to the lack of robust CAD models that accurately predict the high power, large-signal behavior of such circuits, which in turn leads to longer design cycles (design-fab-measurement) and increases the overall cost of power amplifier modules. An important contributor to PA design inaccuracy, especially for high-power designs, is the prediction of optimal load for maximum output power based on device output impedance behavior under large-signal operation. This calculation is typically accomplished through the use of loadpull techniques. However, CAD based loadpull simulations are still not sufficiently accurate, primarily due to lack of precise models

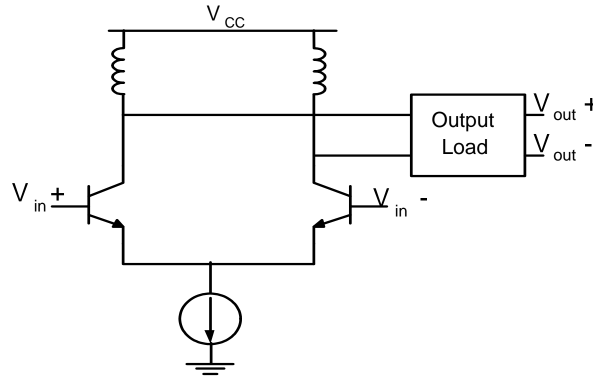

Figure 4.1: Differential architecture with tail current source and differential collector load shown.

for large-signal output impedance behavior. In any case, this SiGe2RF PA design and other PA designs are heavily dependent on loadpull results.

4.2 Optimum Load Resistance

For single stage power amplifier design, it is best to use the common emitter (CE) configuration because common emitter amplifiers have highest power gain compared to other basic circuit configurations such as common base (CB) and common collector/emitter follower [45]. However, CE suffers from lower breakdown voltages as compared to common base. In CB configuration the base is *ac* grounded and essentially a short is seen from the base. Therefore, the base-emitter voltage does not increase because the reverse current (from breakdown) ideally shunts to ground instead of entering the emitter. Since the base-emitter voltage does not increase, the collector current remains the same and the device does not push into avalanche breakdown [31]. However, since gain is at a premium, CE configuration is still the most suitable choice for this design. As discussed in Chapter 2, a differential topology was chosen; i.e. two common emitter HBTs are emitter coupled and an out-of-phase RF output is delivered. A load that provides optimal output power is connected at the collectors as shown in Figure 4.1.

As mentioned earlier, PA design is iterative in nature and hand calculations are only used as a starting point; extensive simulations are used to verify or modify calculated

design parameters. To generate an output power of around 12 dBm with V_{CE} bias of $2.4V$, an initial approximate load impedance was $R = \frac{V_{pk}^2}{2P} = \frac{1.1^2}{2(.015)} \simeq 40\ \Omega$. However, this meant that the maximum voltage swing would be $2.4V + 1.1V = 3.5V$, which is more than $1V$ higher than the CE breakdown voltage. Although, as mentioned in Chapter 1, common-emitter voltage swing can exceed BV_{CE_o} in SiGe HBTs, to be on the safe side, it was decided to limit the voltage swing to or near BV_{CE_o}. Hence, required power is generated by reducing the resistance (increasing the current). The optimal resistance is recalculated as $R_{opt} = \frac{.9^2}{2(.015)} \simeq 27\ \Omega$, where voltage swing is .9 V on a 2 V bias, which means the peak voltage will be 2.9 V, which is above but close to the 2.4 V breakdown limit.

This estimated resistance value is subject to change based on loadpull/large-signal simulations. Since power equal $\frac{1}{2} \times I \times V$ and voltage swing is only .9 V, it can be seen that most of the power will be generated from the current swing. The optimal load value calculated above is used to approximate the required peak current as:

$$I_{pk} = \sqrt{\frac{2 \times (.015)}{27}} = 33.33\ mA. \tag{4.1}$$

This is a relatively high level of current for a silicon-based RF device, and there are some important aspects to consider. First of all, the device should be able to handle this type of current density without causing any reliability issues, such as change in gain characteristics, device failure and degradation mechanisms caused by hot carrier effects, electromigration, etc [46]. In particular, electromigration—a limitation on the current handling capability of the metal layers—must be avoided by designing wider metal interconnects.

Recall from Chapter 2 that this is a Class A PA design. For the peak current excursion calculated above, the device will be biased such that $I_{pk} = I_{DC}$. Therefore, the total current will be $I_{total} = I_{pk} + I_{DC} = 2I_{pk} = 2I_{DC}$.

4.3 Selection of Device Size

The next step is the selection of appropriate active device geometry. The device size is chosen in order to maximize the gain at the expected current density (bias). In

other words, a device with its f_T close to the maximum possible at the calculated bias collector current of 33.33 mA is required. The f_T for a device in CE configuration is simulated by shorting its output and measuring the high frequency current gain; the frequency at which the current gain drops to unity or 0 dB is the f_T. This was accomplished in ADS for a range of device sizes (Figure 4.2). As explained further below, an initial device size of $2 \times 0.5 \times 20$ μm^2 was chosen; a device size of $2 \times 0.5 \times 10$ μm^2 means the device has an emitter stripe area (width x length) of 0.5×10 μm^2 and two (2) emitter stripes. The HBT device in SiGe2RF technology has two layout options: Collector-Base-Emitter (CBE) or Collector-Base-Emitter-Base (CBEB). However, for the 0.5×10 μm^2 device dimension, only the CBEB layout is available. The $2 \times 0.5 \times 20$ device ultimately proved to be insufficient. Based on Figure 4.2, it can be seen that a device size of $5 \times 0.5 \times 10$ μm^2 has maximum f_T at around 33 mA. Note that the bias value in Equation 4.1 were obtained in conjunction with the the DC/f_T and loadpull (explained below) simulations, which involved many iterations and update of the calculations or values.

Once the device size is selected, it is important to verify the maximum available gain (MAG) and stability. Figure 4.3 shows the MAG and the Rollett's stability factor, K, which should be greater than 1 [45]. The MAG of 4.8 dB at 30 GHz is close to what is given in the design manual [47], and the device is unconditionally stable around the frequency of interest. As has been discussed earlier, some gain is traded off for higher output power, but the MAG value gives an idea as to the power gain limit for the power amplifier.

4.4 Loadpull

Loadpull simulation is perhaps the most important step in power amplifier design. The R_{opt} and I_{pk} or I_{DC} values calculated above need to be optimized in order to achieve the target output power in simulation. Usually, it is seen that the simulated values are different than the calculated ones because the interaction between different factors cannot simply be modeled using $R = \frac{V_{pk}^2}{2P}$ and 4.1. The basic idea behind loadpull is straightforward. To explore the effects of different load impedances on the delivered output power, the real and imaginary parts of the load presented to the device are systematically varied in the simulator and contours of constant output

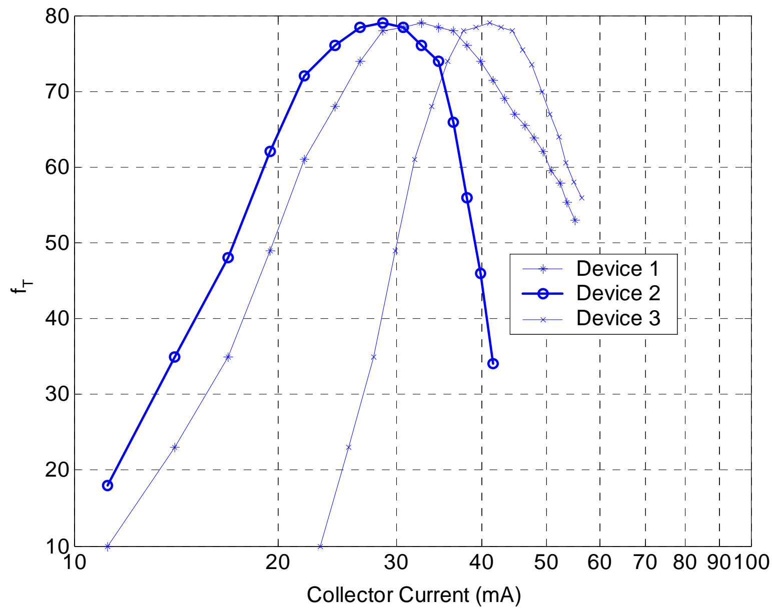

Figure 4.2: f_T versus collector current of different transistor device sizes. Device 1 = $5 \times .5 \times 10\ \mu m^2$. Device 2 = $2 \times .5 \times 20\ \mu m^2$. Device 3 = $5 \times .5 \times 15\ \mu m^2$.

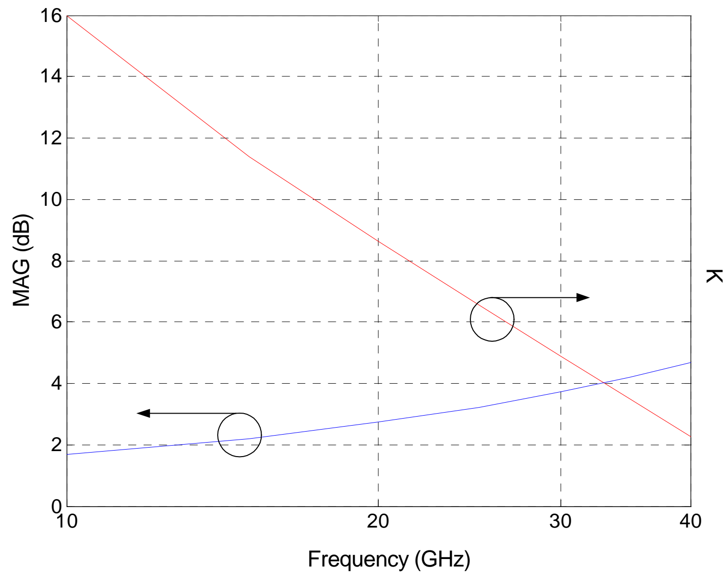

Figure 4.3: MAG and Rollet Stability factor, K, vs. frequency for the selected device size.

power are plotted on a Smith chart. It should also be noted that this can be done experimentally and dedicated load-pull systems are available. However, at mm-wave such systems are not readily available, and are extremely expensive.

4.4.1 Loadpull Theory

The detailed theory behind loadpull and the power contours that are generated is explained in [25] [48]. This process is summarized as follows:

- There are two resistive terminations that result in a maximum linear power—one smaller than R_{opt} (R_{low}) and the other larger than R_{opt} (R_{high}). This is the case because while operation in linear region (not maximum output power), both the voltage and current swing do not have to be maximized. In the case of R_{high} there is maximum voltage swing with smaller corresponding current swing; in the case of R_{low} there is maximum current swing with smaller corresponding voltage swing.
- At R_{low}, the peak output current limits the output power. The contour of constant delivered power is a line of constant load resistance with varying series reactance as shown in Figure 4.4(a). For a constant output power, since reactance is being added, the maximum current remains unchanged, but the voltage swing is increased. In other words, the voltage swing and the impedance change together to keep the output power constant ($P = V^2/|Z|$). The maximum value of series reactance, X_m, and the constant resistance R_{Low} give a complex magnitude, $\sqrt{X_m^2 + R_{Low}^2}$, equal to R_{opt}.
- At R_{high}, the peak output voltage limits the output power. Since here the power is voltage-limited (voltage is constant), it is more convenient to consider the admittance. The contour of constant delivered power is a line of constant load admittance (1/resistance) with varying shunt reactance as shown in Figure 4.4(b). For a constant output power, since susceptance (1/reactance) is being added, the maximum voltage remains unchanged, but the current swing is increased. In other words, the current swing and the impedance change together to keep the output power constant ($P = I^2|Z|$). The maximum value

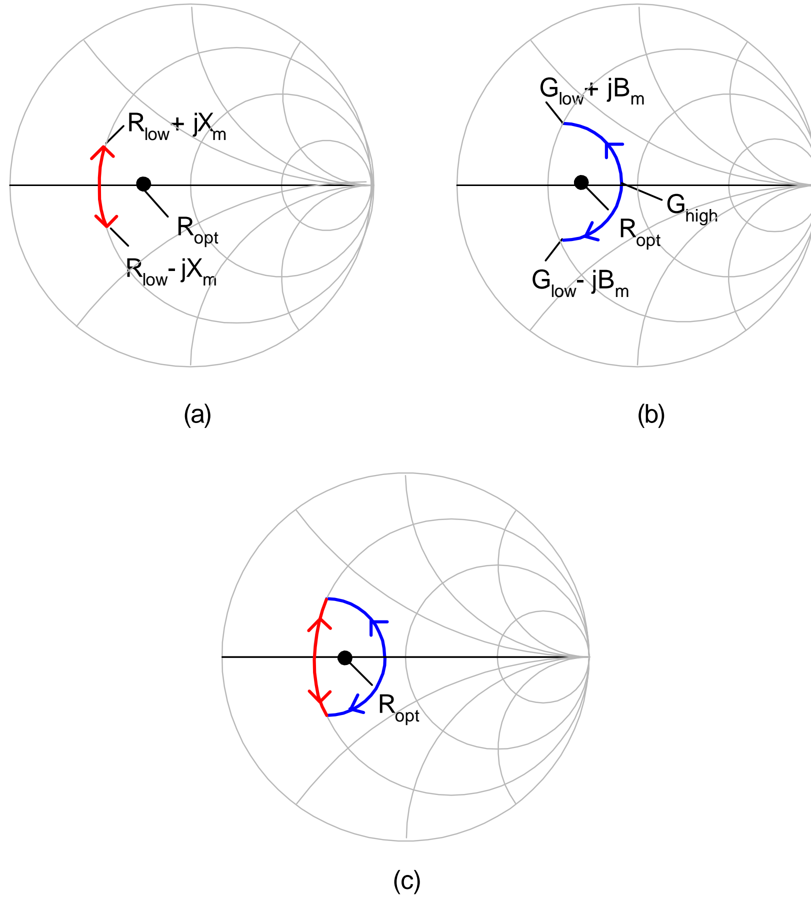

Figure 4.4: Generation of constant output power contour: (a) current-limited power, (b) voltage-limited power, (c) combination of the two.

of shunt susceptance, B_m, and the constant resistance R_{high} give a complex magnitude, $\sqrt{B_m^2 + R_{high}^2}$, equal to R_{opt}.

- The derived loadpull contour is closed by the intersection of the two arcs.

Loadpull contours are generated in simulation by varying the load resistance/conductance and load reactance/susceptance for different levels of delivered output power, which leads to different power contours. Agilent Advanced Design System (ADS) has a robust simulation engine that carries out nonlinear simulations using the harmonic balance technique [30]. Varying the load is accomplished by using a load tuner component.

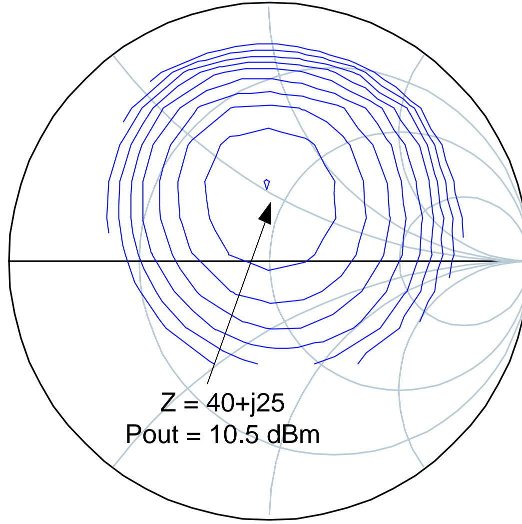

Figure 4.5: Power delivered contours for the $2 \times 0.5 \times 20$ μm^2 device size. Contours are in .5 dB step, ending in 10.5 dBm.

Initially, loadpull simulations on a device with two emitter stripes and same emitter area as mentioned earlier ($2 \times 0.5 \times 20$ μm^2) provided less than 12 dBm of output power as shown in figure 4.5. This meant more current was required to generate the output power, and after a few iterations, the device size of $5 \times 0.5 \times 20$ μm^2 was chosen in conjunction with f_T simulations. The delivered power contours for the selected device size of $5 \times 0.5 \times 20$ μm^2 generated by loadpull simulation are shown in Figure 4.6.

4.5 Matching Networks

Different matching network topologies are used in RF/microwave amplifier design and each are advantageous in particular cases. The two most common types of matching network topologies are shown in Figure 4.7. One type consists of two passive elements and the other type consists of three passive elements. As shown in Figure 4.7(a) , a matching network with two elements is called an L-transformer. On the other hand, three-element matching networks can have either π (4.7(b)) or T (4.7(c)) transformer topologies. The π-transformers are typically used for output matching networks for

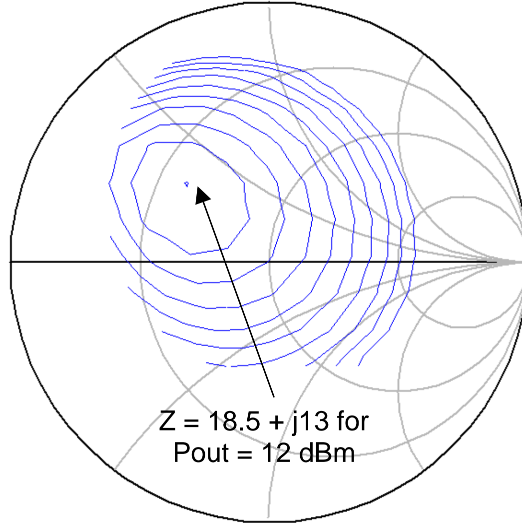

Figure 4.6: Power delivered contours for the selected device; size is $5 \times 0.5 \times 10\ \mu m^2$. Contours are in .5 dB step, ending in 12 dBm.

high power PA designs. They are also typically used as inter-stage matching networks where it might be necessary to achieve a sinusoidal voltage waveform at the input of subsequent stage by attenuating harmonics [27]. On the other hand, for high power PAs with low input impedance, T-transformers are usually used as input matching networks, especially the two-capacitor/high inductance topology, in order to generate a sinusoidal current waveform at the input.

In any matching network, the two important issues are the resonant frequency, f_o and the quality factor of the matching network, Q. Note that this is not the same as the "Q" of lossy passive components. The three-element transformers provide flexibility in choosing the resonant frequency, bandwidth and quality factor but occupy more die area. On the other hand, the L-transformer is restricted in that the Q is fixed by the load and source resistances as:

$$Q = \sqrt{\frac{R_{in}}{R_{out}} - 1}, \tag{4.2}$$

where R_{in} and R_{out} are the input and output resistances as seen by the matching network. Hence, L-transformers are usually used for narrow-band, low-power design, and in designs where a specific impedance-match Q is not required. L-transformers

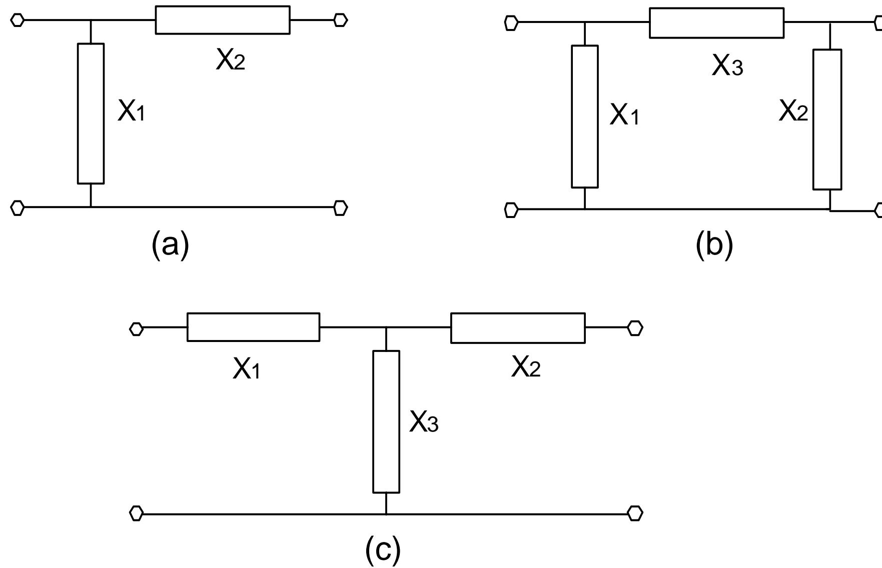

Figure 4.7: Different matching networks: (a) L- (b) π- (c) T-transformer. X_1, X_2, X_3 represent the reactances of the passive elements.

are advantageous to use here because fewer passive elements are required (and therefore less die area), and tuning can be easily achieved both at the input and output. Either a low pass or high pass topology for L-transformers can be used. A low pass topology was chosen for this design for easier implementation with transmission lines. However, this requires that *dc* blocking capacitors be placed in series with the series element (i.e. the series element X_2 in Figure 4.7).

4.5.1 FGC Transmission Line Matching Networks

At higher microwave/mm-wave frequencies, lumped elements are harder to realize because much smaller values are required, which is particularly an issue in the case of inductors. Also, the self-resonance caused by inductor parasitic capacitances is a major concern at mm-wave frequencies because at such high frequencies capacitive components become comparable to the inductance. Moreover, realizing high-Q inductors is difficult due to high parasitics and resistive loss. At 30 GHz, distributed matching networks were implemented using FGC transmission lines. The free-space wavelength at 30 GHz is:

$$\lambda = \frac{c}{f} = \frac{3 \times 10^8}{30 \times 10^9} = .01\ m = 10\ mm. \tag{4.3}$$

This is relatively large on the length scale of typical RFICs (few millimeters). In addition, transmission lines in silicon technology would suffer higher power loss at such lengths due to line attenuation. Consequently, at 30 GHz, with today's IC technologies, using lumped passive elements for designing matching network are desirable. However, the Atmel SiGe2RF technology used here has not been optimized for microwave/mm-wave design, and it was seen that the required inductor values could not be achieved with the available inductor libraries [47]. Therefore, to avoid additional uncertainties arising from custom inductor design in this technology, distributed elements were used for better control over the final design. Also, the Q of distributed elements is higher than what could be achieved with passive lumped components at these frequencies.

A low-pass L-transformer topology means that the matching network consists of a series inductance and shunt capacitance. As shown in Figure 4.8 [27], a series inductance can be modeled as a short transmission line section, while a shunt capacitance with electrical length less than 90° can be modeled as open-circuit shunt stub. Therefore, the matching networks are based on a series line and an open-circuit shunt stub. Initially, ideal ADS transmission line components were used to simulate the matching networks in the PA design, which were specified in terms of their electrical length in degrees, θ, at 30 GHz. The guided wavelength is given by:

$$\lambda_g = \frac{c}{f\sqrt{\epsilon_{eff}}}, \tag{4.4}$$

where ϵ_{eff} is the effective permittivity. Distributed elements in this design are implemented in FGC technology (Chapter 3). The simulated electrical length from ADS for each ideal stub and through line was used together with the equation below to approximate the physical length.

$$Physical\ Length = L = \frac{\theta \times \lambda_g}{360^o}. \tag{4.5}$$

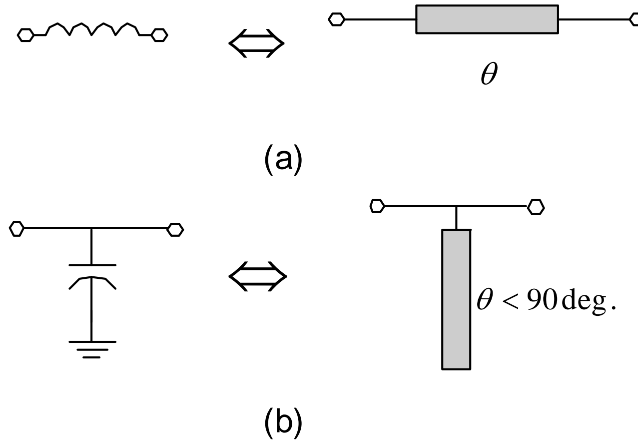

Figure 4.8: Equivalent lumped and distributed elements

Finally, extensive EM simulations using IE3D were performed on the different FGC stubs and through lines required in this design to obtain accurate physical lengths for each stub and through line. The performance (i.e., mainly phase(S_{21})) of the designed FGC stubs and through lines were similar to that of the ideal ADS transmission line components, and the ideal components were replaced by the EM simulated FGC structures using their s-parameters.

4.5.2 Input and Output Matching

A similar concept to loadpull can be applied to the source side of the amplifier. In other words, different impedances can be presented at the source or input to see what impedances or load provide the best gain and power results. However, in most PA designs, for simplicity, the source is conjugate matched at the input because optimal source impedance usually does not have any appreciable affect on the overall results. Hence, in this design the input matching is designed to conjugate match the input at 30 GHz. The L-transformer topology described above was used for the input and output matching networks. Note that, unlike the input match, the output match transform a 50 Ω load to the impedance found using the loadpull simulations.

Table 4.1 shows the simulated electrical length and the corresponding physical length for input and output match. The electrical lengths were optimized using IE3D EM simulations and included in the PA simulations.

Table 4.1: Electrical and corresponding physical lengths of stubs for input and output matching network at 30 GHz.

	Electrical Length	**Physical Length**
Input Match		
Series Line	27°	215 μm
Open-circuit Stub	72°	39.5 μm
Output Match		
Series Line	52°	525 μm
Open-circuit Stub	47°	148 μm

4.5.3 Bias Network

Several biasing networks for power amplifiers were described in Section 2.6. In the Atmel power amplifier design, a current source connected to differential pair emitters was used to provide the bias.

It was seen above that a bias current of approximately 33 mA was required to obtain desirable output power. Therefore, the current mirror needs to provide 66 mA of current for the two devices connected to the current mirror. DC simulations led to a current mirror device size of $2 \times 25 \times 0.5\ \mu m^2$ with V_{CE} bias of $2V$ for the devices. The ratio of emitter area of diode-connected device and current mirror output device defines the current gain of a current source. For example, if the output device is twice as large as the diode-connected device, then the output current (collector current for output device) is twice that of the input collector current. However, to avoid mismatch issues and breakdown due to high current densities, both the input and output devices were large and of the same size. Typically, the output current of a current source changes with variation in the output voltage. This phenomenon is typically characterized by finite output resistance of a current source, r_o. Finite r_o can lead to difference between the input and the output currents—in other words, the current source does not exactly "mirror" the current from the input to the output. Usually, the output current is greater than the input current [49]. Thus, the input current was set at a level to ensure that the current-source output current was approximately equal to the required bias.

Transistors used for the current source were high power/high breakdown devices with lower f_T of around 50 GHz compared to the RF devices that have been enhanced for

high performance (high f_T of 80 GHz). A resistive divider network was connected to the base of the PA device to maintain close to 2 V at the emitter, and collector voltage was provided from the supply through an on-chip (1.3 nH) RF choke.

The final circuit design based on the approach and methods discussed above is shown in Figure 4.9. The figure shows the balanced amplifier stage. Each side is matched to 50 Ω (or 100 Ω differential) at the input and the output. The low-pass L-transformer matching networks are depicted by through lines (i.e. through line of $\theta = 27^o$) and open-circuit shunt stubs (i.e. stub of $\theta = 72^o$). Complete bias network with emitter-connected current source, resistive divider and RF choke for voltage supply feed is shown as well. Note that dc blocking capacitors were placed in series with FGC through lines at the input and output. Also, by-pass capacitors were connected to RF chokes at the voltage supply node.

4.6 Power Amplifier Simulation Results

This circuit was simulated to analyze the small-signal as well as the large signal performance at 30 GHz. It has been explained in Section 3.2 that IE3D was used for EM simulation for the FGC lines/stubs and their s-parameters were included in ADS using two-port s-parameter blocks. The rest of the circuit simulations were carried out in ADS, using both its RF engine and its system-simulation engine called Ptolemy [30]. The simulation results for the final design are discussed below.

4.6.1 Small-signal Simulations

Small-signal simulation was carried at extremely backed-off operating point—around -30 dBm input power. This ensures no effects due to nonlinearities. The input and output matching network designs were described earlier, and their small-signal behaviors are shown in Figure 4.10. S_{11} is -11.9 dB while S_{22} is -8.2 dB at 30 GHz. The S_{22} plot indicates a mismatch condition exists (referenced to 50 Ω); however, that is not of concern because the output is being matched for optimal output power. The stability of the active device was shown earlier but the overall amplifier small-signal stability needs to be checked. The Rollett stability factor, K, which should be

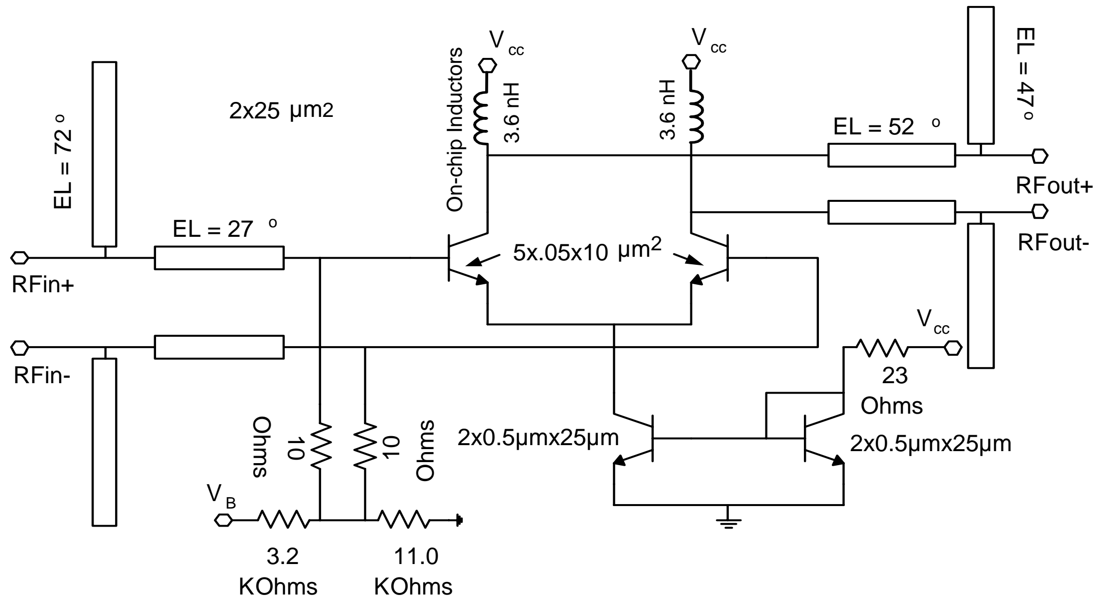

Figure 4.9: Simplified circuit diagram

greater than 1, is shown in Figure 4.11, confirms that the amplifier is stable at and around 30 GHz. The small-signal gain is around 3.3 dB as shown in Figure 4.11.

4.6.2 Large-signal Simulations

As discussed in Chapter 2 output power, linearity and efficiency are typically the main performance parameters for power amplifiers. Moreover, linearity parameters such as 1-dB compression point, harmonic suppression, AM-PM distortion and EVM were discussed. In this work, these were all simulated. Figure 4.12 shows the AM-AM (P_{out} versus P_{in}) and the gain versus the P_{in}. The output power at 1-dB compression point is 12.3 dBm . The gain versus power is 3.3 dB, as expected. Although efficiency was not a target, the Power Added Efficiency (PAE) is also simulated and is shown in Figure 4.14. As expected, it is much lower than for PA's in mobile applications.

Amplifier stability parameters are derived for small-signal operation. The Rollett factor was verified at 30 GHz but that simulation was done for the amplifier back-off into small signal operation. There is no comprehensive theory or technique to characterize stability in large-signal operation or versus power levels. Hence, only after

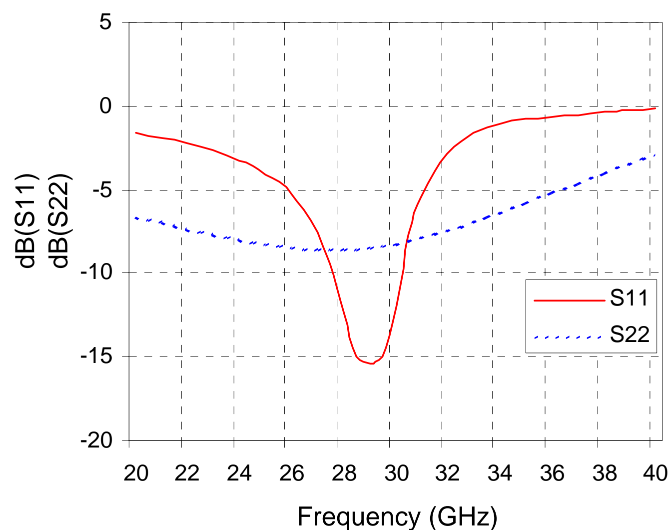

Figure 4.10: Simulated S_{11} and S_{22}. The -10 dB RL bandwidth is 2.3 GHz.

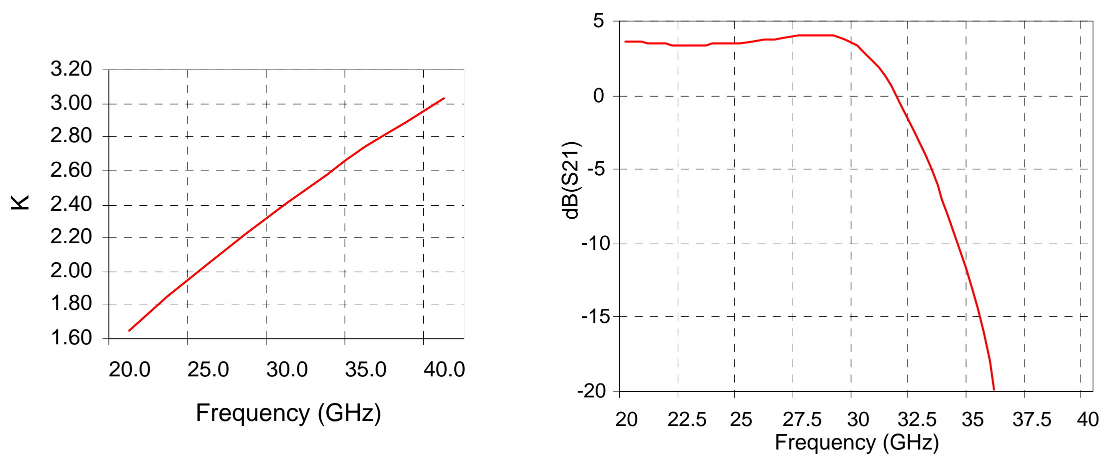

Figure 4.11: Simulated small-signal K factor and gain vs. frequency

fabrication and measurement, it is known with confidence whether a power amplifier is stable at high input power levels by observing the presence or absence of oscillations. However, a theory has been proposed by Agilent and has been incorporated in the ADS simulator. The proposed simulation for large-signal stability is supposed to accurately predict whether an amplifier oscillates or not at large signals [50]. Using this ADS simulation, the large signal K factor versus power level at 30 GHz was simulated and is shown in Figure 4.13. This K factor is to be greater than 1 as well to indicate stability.

The AM-PM plot (in other words phase distortion versus input power) is shown in Figure 4.15. This plot conforms closely with what would be expected by a Class A power amplifier. The phase shift is essentially flat until the PA starts to compress, at which point the phase distortion level increases markedly. At 1-dB compression point, the phase shift is 4 deg./dB. Figure 4.16 shows the harmonics generated by the amplifier. Due to differential operation, the even harmonics are suppressed, as is expected. The worst case harmonic distortion (due to the third harmonic) is -41 dBc.

Finally, the EVM simulation was carried out in ADS Ptolemy, which employs the ADS large-signal simulation engine (harmonic balance/envelope simulation) and ADS DSP simulator. Refer to Section 2.3.5 for more information on EVM. A 64-QAM modulation scheme was used at RF carrier frequency of 30 GHz for the EVM simulation. Figure 4.17 shows the results. At 1-dB compression point, P_{1-dB}, the EVM is 2.2%, which is well under the EVM limit set as a design goal.

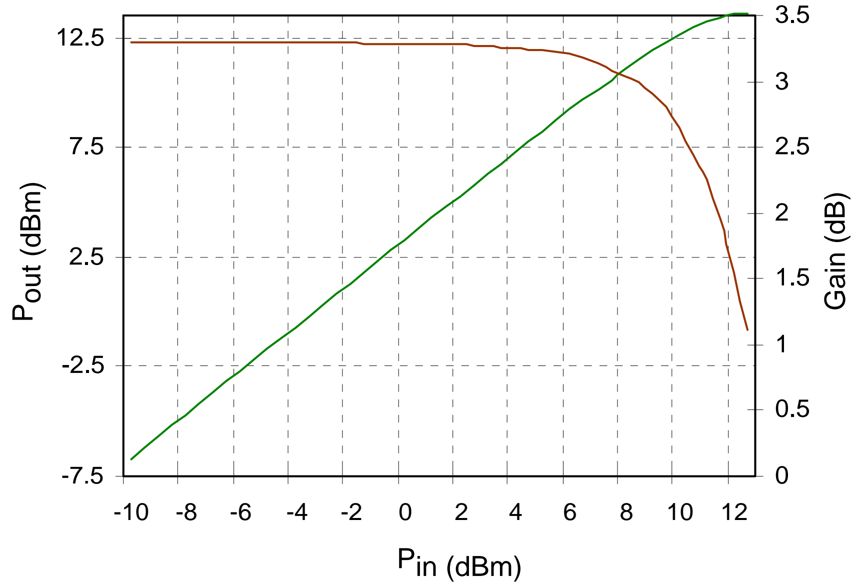

Figure 4.12: Simulted P_{out} vs. P_{in} at 30 GHz. Gain vs. P_{in} is also shown. The 1-dB compression point is approximately P_{in}=10 dBm.

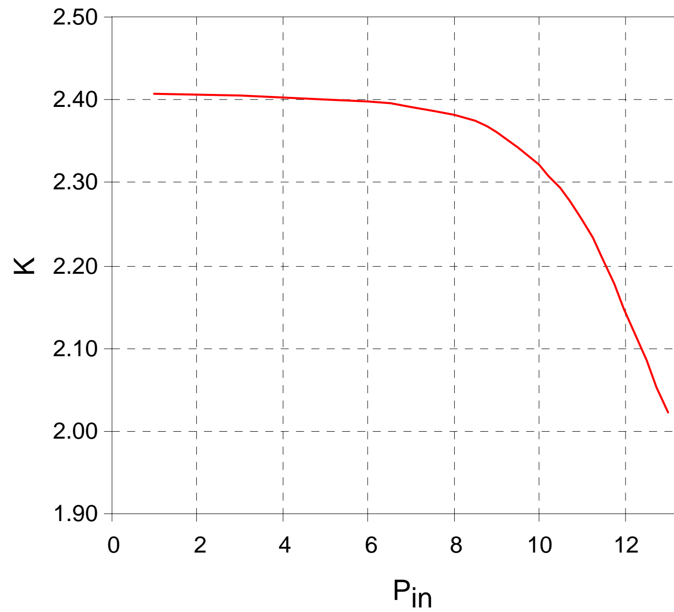

Figure 4.13: Simulated "large-signal" K factor vs. P_{in}.

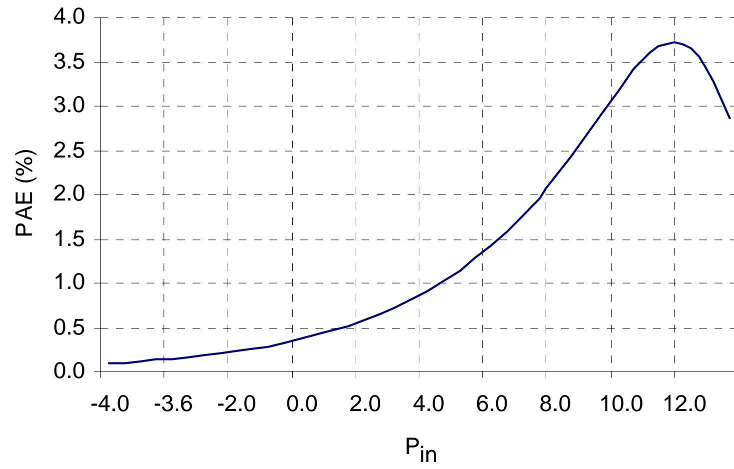

Figure 4.14: Simulated power added efficiency (PAE) vs. P_{in}.

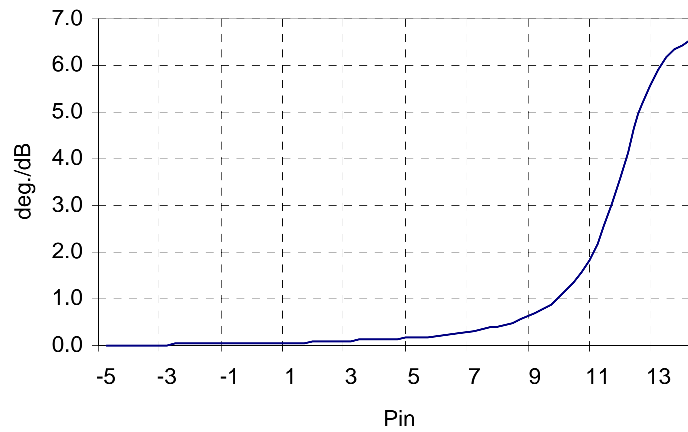

Figure 4.15: Simulated AM-PM vs. P_{in}.

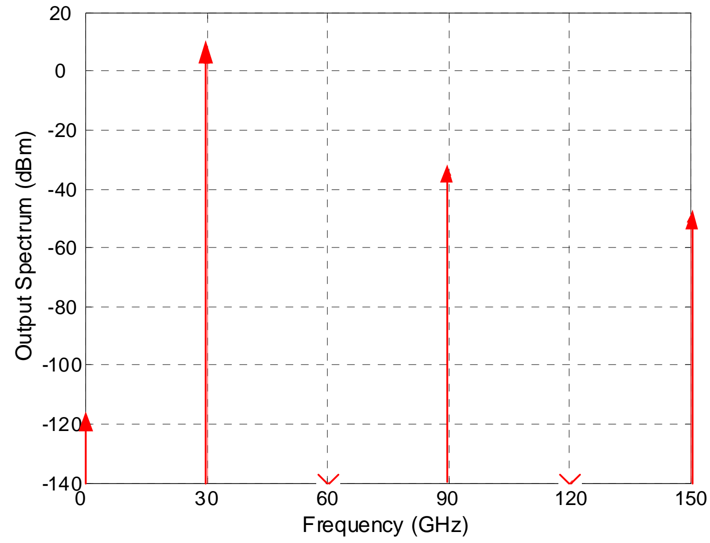

Figure 4.16: Simulated Output spectrum vs. frequency. $P_{in} = 10$ dBm.

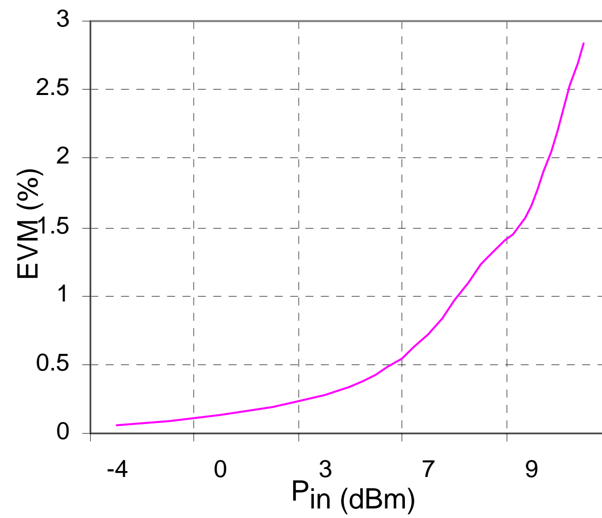

Figure 4.17: Simulated EVM vs. P_{in}.

Chapter 5

Atmel SiGe2RF Power Amplifier Fabrication and Measurement

The PA design described in the previous chapter culminate in chip layout, verification and fabrication by the Atmel foundry. After fabrication, the chip was tested at Virginia Tech using on-wafer measurement techniques. This chapter details issues involving layout of the Atmel power amplifier design, and explains test plans for both small-signal and large-signal measurements. Due to the lack of access to a fully differential test set or appropriate baluns, single-ended measurements for each side of the circuit were made independently and the results were combined to synthesize differential measurements.

5.1 Layout

In the production of MMICs or RFICs, the layout of the circuit is as important as the design itself. Poor layout can lead to degradation in performance, and sometimes even failure, if issues such as asymmetry interconnect parasitics and electromigration—are not adequately considered. For power amplifiers, high current densities and output power require more careful layout.

Table 5.1: Resistors used in fabrication of bias circuit.

Resistor	Sheet Resistance	IC Layer	Current Condition[1]
R500	425 Ω/sq	poly-1	2 mA
R1500	1450 Ω/sq	poly-2	1 mA

In Chapter 3, the SiGe2RF technology process was briefly discussed. It is a three-metal process and has a library of RF-enabled passive elements such as resistors, capacitors and inductors. For this design, transmission lines were used for distributed matching, which required custom layout and EM simulation. However, inductors and capacitors were used for RF choke to provide bias voltage from the supply and as dc-blocks respectively. Capacitors were also used for by-passing at the supply.

5.1.1 Resistors

There are four different types of resistors available in the process, ranging from low sheet resistance to very high sheet resistance. The resistors are fabricated using poly layers. In order to save space and realizing that the resistors used in bias circuits have relatively high values, high-sheet resistance resistors were used. Specifically, R500 and R1500 resistors were used. Table 5.1 shows some parameters of the resistors used.

5.1.2 Capacitor

Two distinct types of capacitors are available: Nitride MIM capacitors with 1.1 $fF/\mu m^2$ and standard MIM capacitors with 0.93 $fF/\mu m^2$. MIM capacitors in this technology are built between metal-2 and metal-1, with Nitride being the dielectric between the metal layers. Nitride capacitors are different from MIM as the upper plate is not a standard interconnect metal layer but is TiSi, located much closer to metal-1 so the dielectric is much thinner. Since the MIM capacitors are used for dc blocking in this design, relatively large capacitor values are needed to reduce the impedance seen by the RF signal. High frequency operation in the millimeter-wave

[1]Sheet resistance calculated at this current level current level.

range inherently helps in this cause as well. It was found that capacitors around 8 pF and with aspect ratio (L×W) of 1.5 produced the desired results with reasonable capacitor dimensions with regard to overall chip area.

5.1.3 Inductors

There are two types of circular spiral inductors available in this process; circular inductors are advantageous over octagonal or square inductors since they provide higher Q. At 30 GHz, the impedance of approximately 4-5 nH inductor is $|jwL| \approx 6.28 \times 30e9 \times 4e-9 \approx 755\ \Omega$. This impedance is sufficient to *choke* the RF signal and provide a low-resistance path for the *dc* bias. Hence, simulations on the circular inductor model were performed to select an appropriate inductor with 4-5 nH of inductance or more, while keeping the self resonant frequency (SRF) of the inductor well above the operating frequency of 30 GHz. Varying the width of the metal, spacing between adjacent turns, number of turns[2] and the inner radius of the inner-most turn. It was seen that an inductor with dimension values listed in Table 5.2 provided about 3.6 nH of inductance with its SRF above 30 GHz. The simulated SRF is ~36 GHz. Figure 5.1 shows the inductance versus frequency plot of the chosen inductor. The inductance was calculated by performing *ac* simulation on the circular inductor model and dividing the imaginary part of inductor's input impedance by $2 \times \pi \times f$ as shown below:

$$L = \frac{\mathrm{Im}(Z_{in})}{2 \times \pi \times f} \tag{5.1}$$

5.1.4 Symmetry

In order to maximize the advantages of a differential design discussed in Chapter 2 mismatches between the two sides of the differential amplifier must be reduced. Therefore, the most important goal in any differential circuit layout is to maximize symmetry. Here, it is critical to keep the circuit symmetric from the RF perspective.

[2]Note that in the Atmel design kit, the "number of turns" have to be specified in terms of quarter (1/4) turns. For example, 13 turns means 13/4 = 3 and 1/4 turns.

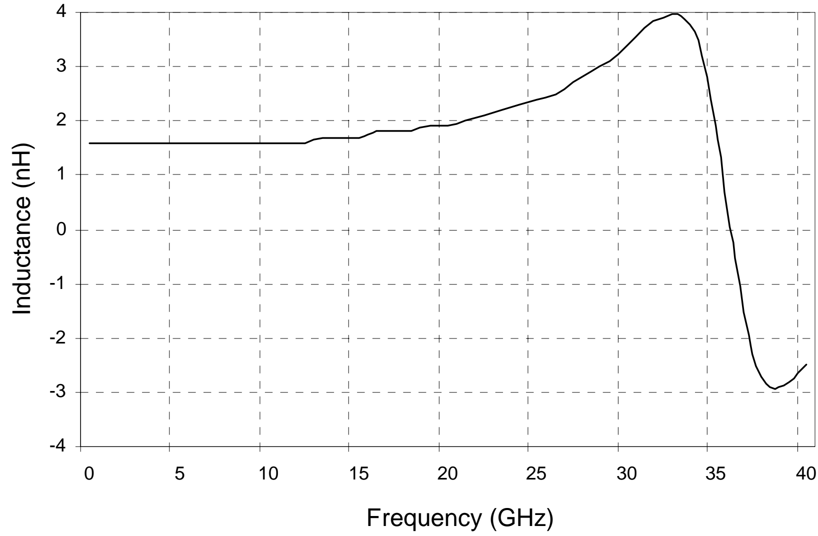

Figure 5.1: Simulated inductance vs. frequency using of the chosen inductor with parameters summarized in Table 5.2. Note that the SRF for this inductor design is around 36 GHz.

Table 5.2: Chosen inductor parameters.

Parameter	**Value**
Width of turns	6 μm
Spacing between adjacent turns	5 μm
Number of quarter turns	13
Inner radius	27.5 μm
Total inductance	4.02 nH
Self-resonant frequency	36 GHz

This means that signal lines should be of exactly the same dimensions and follow the same routing. Fortunately, in this design such symmetry is inherently achieved through the use of balanced CPW (or FGC) lines connected on both sides of the circuit. Some asymmetry arises from the *dc* bias routing because the placement of supply voltage node for resistive divider network symmetric routing with respect to current-source supply voltage node. Symmetry is also maintained for the supply voltage connections to both power devices through the respective RF chokes. However, the base voltage bias circuit (resistor divider) connected to device bases inherently results in some asymmetry because of the need to place the current mirror very close to the connected device emitters. However, this is not a significant a mismatch because the asymmetry is in the bias circuit and not in the RF path.

5.1.5 Electromigration

Electromigration is the process of metal dislocation in interconnects due to excess current densities. In power amplifiers where high current density is carried through metal interconnects, it is crucial to keep in mind the current handling limits of the metal layers, and increase the width of traces as appropriate to prevent electromigration effects. For this design, the SiGe2RF process has current density, J_{max}, limit of 8 $mA/\mu m$ of metal width for metal 3 at 125 ^oC. The chosen signal line center conductor width (Section 3.4) of 30 μm therefore has a current handling capability of $30 \times 8\ mA = 240\ mA$, which is well above the expected peak current of around 66 mA. Similarly, the *dc* lines were appropriately widened so that they could handle *dc* bias current of around 33 mA.

5.1.6 Final Layout

Taking into consideration the aspects explained above, and other issues such as Design Rule Checks (DRC) and antenna rules, the resulting final layout is shown in Figure 5.2. The circular pads are 80 μm in diameter, and the distance between each RF pad is determined by the pitch of probes used during measurements. The available ground-signal-ground-signal-ground (GSGSG) probes have a pitch of 150 μm for their pins. This means that the pads are placed such that the distance between the center

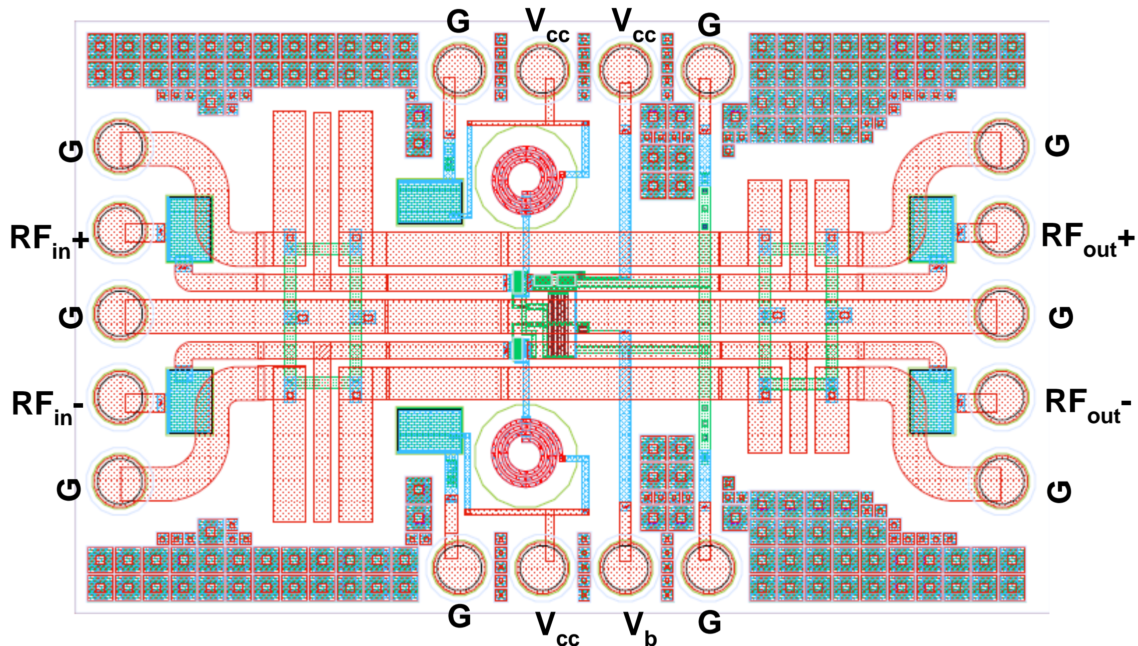

Figure 5.2: Complete layout of the Atmel SiGe2RF power amplifier design. Die area = $1.75 \times 1.15\ \mu m^2$.

of a pad to the center of the next pad is 150 μm. DC blocking capacitors are placed in the RF path near the pads. The complete layout was verified versus the schematic (LVS check) to ensure that all connections were made properly and there were no short circuits, floating nodes, etc. The layout data was subsequently submitted to Atmel for fabrication via M/A-COM. Figure 5.3 shows the actual die after fabrication. The total die area is 1.75x1.15 mm^2.

5.2 Small-signal Measurements

Small-signal measurements (S_{11}, S_{21} and S_{22}) were taken for comparison with the previously simulated results. Stability was verified since the circuit operated as an amplifier rather than having undesired oscillations.

Figure 5.4 shows the measurement setup. The Agilent E8364B Vector Network Analyzer (VNA), or "PNA" Network Analyzer as it is called, was used to carry

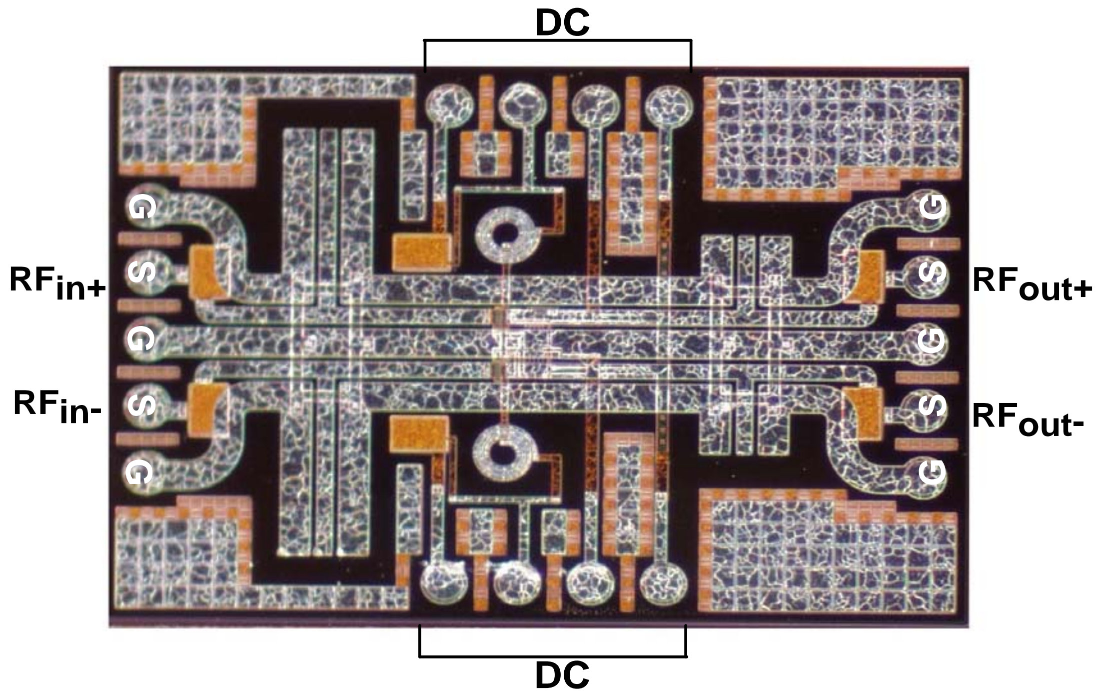

Figure 5.3: Actual die photograph of the fabricated PA with Area=1.75 × 1.15 mm^2.

out the measurements. First, the PNA was calibrated using Cascade Calibration Standard and the PNA built-in Short-Open-Load-Thru (SOLT) calibration routine for the full on-wafer 2-port system. This calibrates out the losses to the reference plane of the probe tips. The calibration was carried out over the frequency range of 20 to 40 GHz. Two 2.92 mm (K-connector) coaxial cables were used to connect each port of the PNA to each port of the device under test (DUT). A 2.92 mm cable is appropriate for this use because it has an upper limit of mode-free operation of approximately 40 GHz [51]. Infinity GSGSG differential probes from Cascade Microtech were employed. For each measurement, one side was terminated with a 50 Ω load while RF signal was passed through the other side of the circuit. The results from each measurement (each side) were combined in ADS using ideal baluns. Hence, the measurements show the synthesized results obtained from ADS.

Figure 5.5 shows the small signal measurements. S_{11} is -11.7 dB, S_{21} is 2.9 dB and S_{22} is -12.5 dB at 30 GHz. These results were obtained at input power level of -30 dBm.

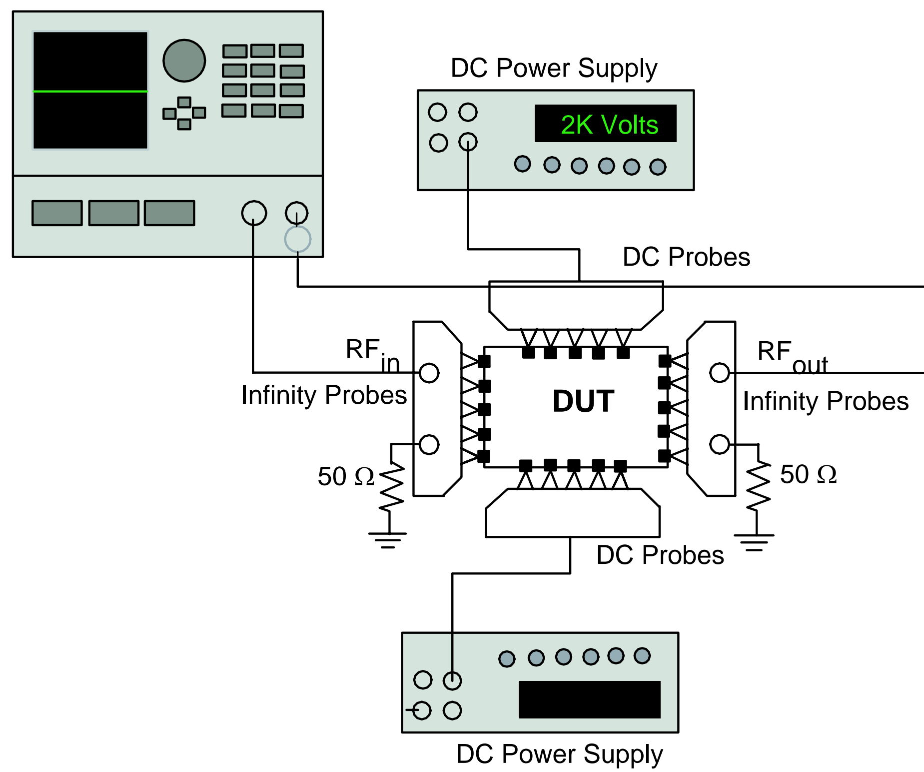

Figure 5.4: Small-signal measurement setup. The DUT is the power amplifier under test.

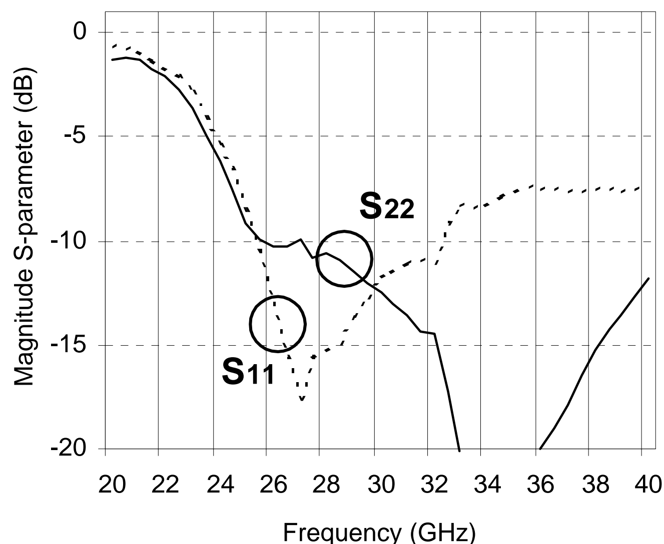

Figure 5.5: Measured S_{11} and S_{22} verses frequency of the fabricated power amplifier.

Table 5.3: Comparison between simulated and measured results at 30 GHz

Parameters	**Simulated**	**Measured**
Gain (S_{21}) $[dB]$	3.3	2.9
P_{out} @ 1-dB compression $[dBm]$	12.3	11.9
S_{11} $[dB]$	-11.9	-11.7
S_{22} $[dB]$	-8.2	-12.5

Figure 5.6 shows the measured small-signal gain. The measured gain is 0.4 dB below the simulated small-signal gain. Table 5.3 summarizes some simulated and measured parameters.

5.3 Large-signal Measurements

Large-signal output power versus input power measurements were taken to determine the measured 1-dB compression point. Other large signal measurements such as AM-PM compression, EVM and harmonics were not possible due to the lack of appropriate

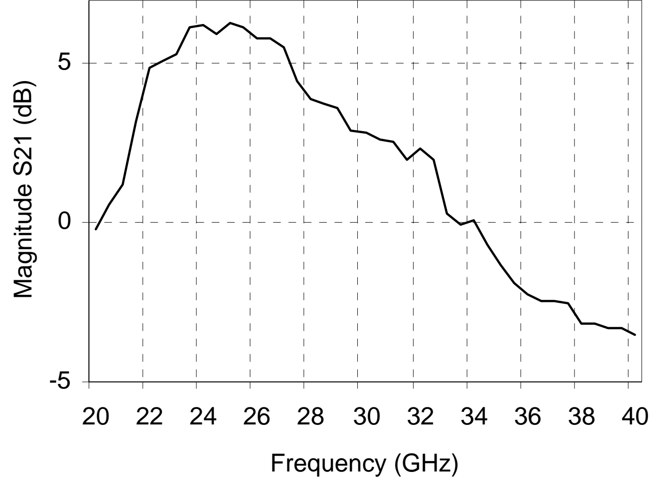

Figure 5.6: Measured differential S_{21}verses frequency of the fabricated power amplifier

equipment and time limitations. Before doing the measurement, the PNA system was SOLT calibrated again in the power domain. In other words, the power source (VNA) was calibrated to generate correct power levels at around 30 GHz so that it provided stable, accurate input power to the DUT.

Figure 5.7 shows the measurement setup. The large-signal measurement required the use of external pre-amplifier to drive the DUT sufficiently into compression, power sensor and power meter. The Agilent 8487A power sensor was calibrated using the HP E4418B power meter, which had the sensor's calibration data stored in it. The external amplifier was an HP 83050A microwave amplifier that had the upper frequency limit of 50 GHz.

Figure 5.8 shows the AM-AM plot, showing the output power versus input power. The input referred 1-dB compression point, $P_{in,1-dB}$ is approximately 10 dBm, and the output power at the compression point is approximately 12 dBm. Although harmonic content of the power amplifier is shown in simulation results, such measurement using a Harmonic Mixer and Spectrum Analyzer were not successful in the test lab. Unstable measurement were seen, and the power levels at the harmonics were nearly undetectable. Table 5.3 summarizes some simulated and measured parameters.

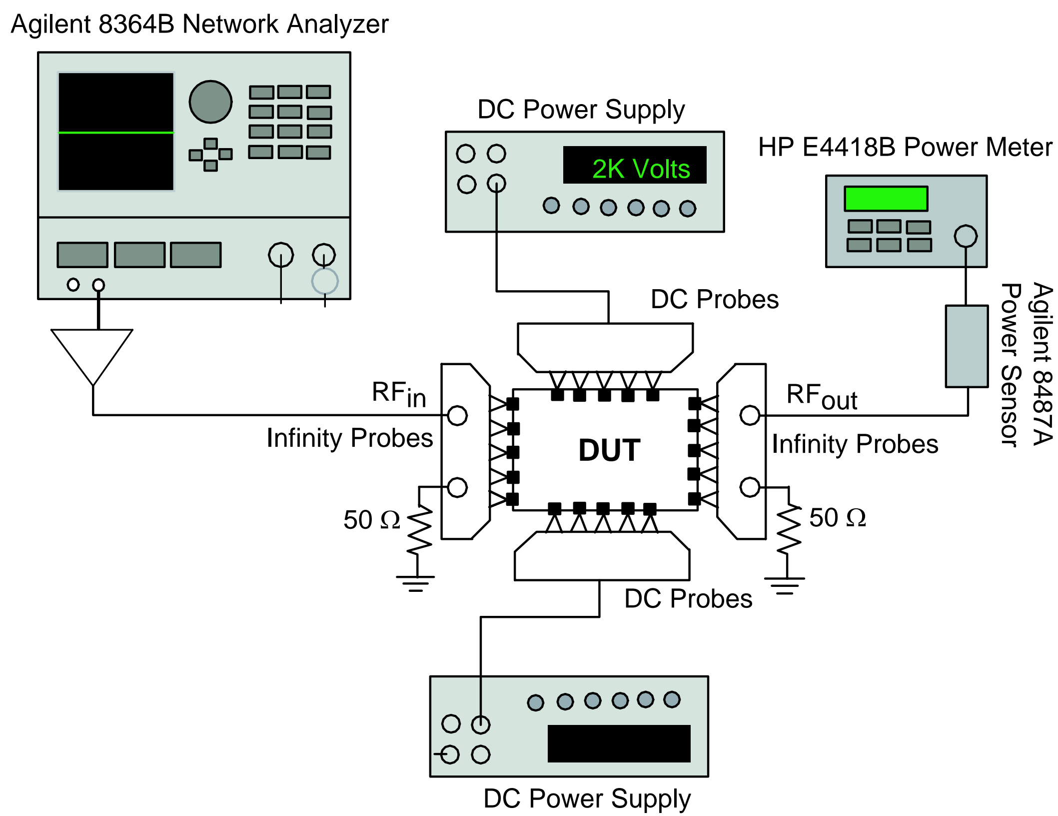

Figure 5.7: Large-signal measurement setup. The DUT is the power amplifier chip under test.

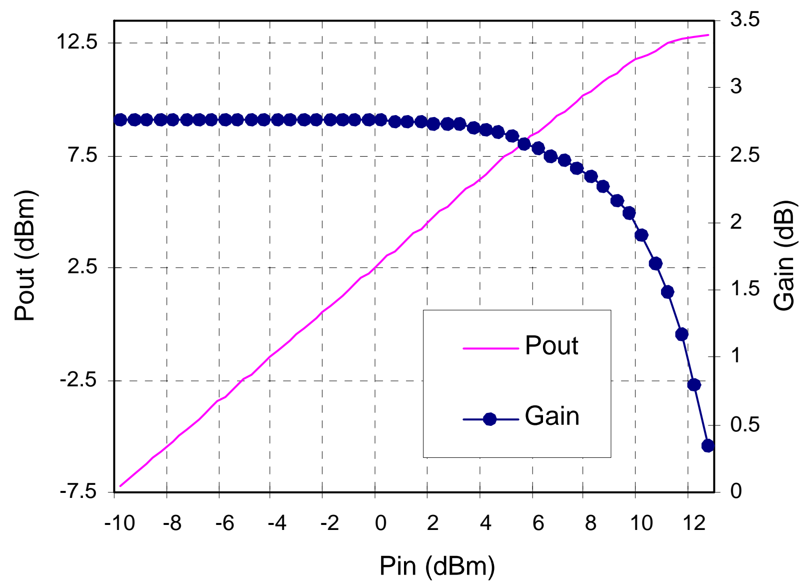

Figure 5.8: Measured P_{out} vs. P_{in} of the fabricated power amplifier at 30 GHz.

Chapter 6

IBM 8HP PA Design and Simulation

In order to achieve higher performance for potential mm-wave applications, a second PA design was completed in IBM BiCMOS 8HP technology. In this chapter, the IBM 8HP PA design will be described briefly. The basic design is very similar to that of the Atmel amplifier presented in previous chapters with same design goals. However, some changes are made in the implementation of the input and output matching networks, and the bias circuitry. Simulation results and full custom layout for this design are presented. This circuit has not yet been fabricated; therefore, no measurement data is available.

6.1 Circuit Design

In order to perform ADS simulation with the IBM 8HP design kit, schematic capture had to be done in Cadence and then exported to ADS using ADS RFIC Dynamic Link (DL) [30]. RFIC DL is an interface that allows communication between ADS and Cadence. In addition, final design was also simulated in Cadence to compare the large-signal simulations with those in ADS. Since this design is basically the same as that of the Atmel design discussed previously, only the differences will be highlighted below.

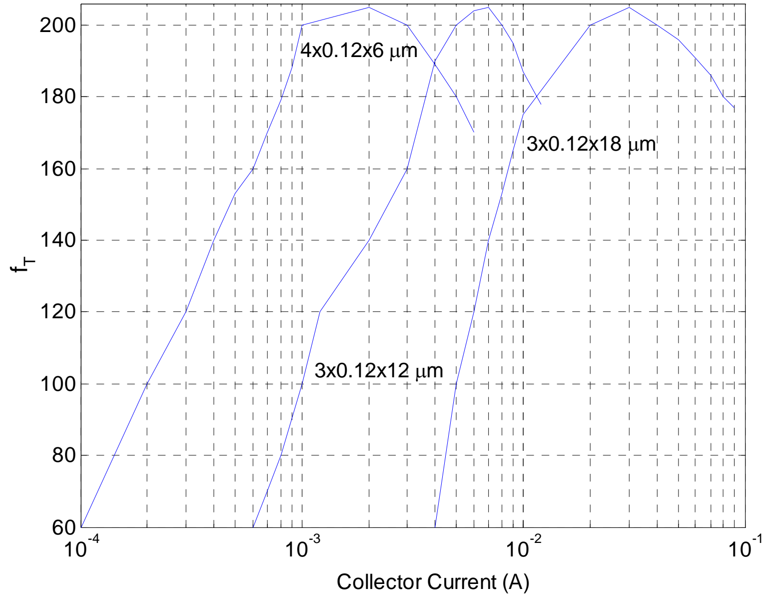

Figure 6.1: f_T versus collector current of different transistor device sizes.Device 1 $= 4\times.12\times6$ μm^2.Device 2 $= 3 \times .12 \times 12$ μm^2.Device 3 $= 3 \times .12 \times 18$ μm^2.

The HBT devices in IBM 8HP technology provide much higher gain compared to the Atmel SiGe2RF technology, with a peak f_T close to 220 GHz. Figure 6.1 shows simulated f_T for three different device sizes. The chosen device size is $3\times0.12\times18$ μm^2, which was selected on the basis of a near-peak f_T at the bias point of 30 mA. The simulated MAG at 30 GHz is 11.9 dB, and Rollett's stability factor K is 1.8; and thus the amplifier is unconditionally stable around 30 GHz. The loadpull output power contours are shown in Figure 6.2 for the selected device.

6.1.1 Matching Networks

As described in Chapter 4, a shunt capacitor is equivalent to open-circuit shunt stub, given that the electrical length of the stub is less than 90°. The shunt elements in the input and output matching networks need to provide a susceptances of 38.6 mS and 19.8 mS, respectively. The lower values of susceptances translate to small values of capacitors since capacitance $= C = \frac{B}{\omega}$, where the susceptance is denoted by

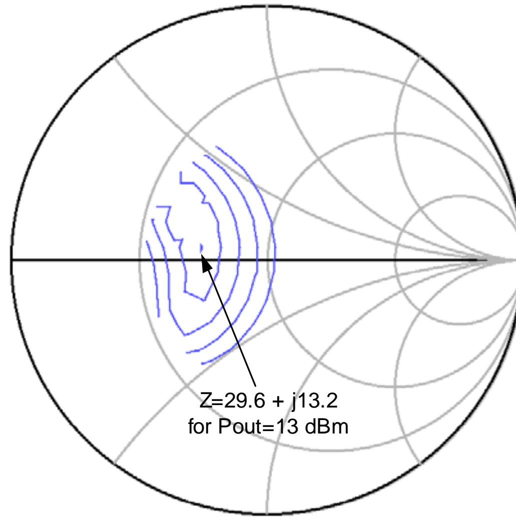

Figure 6.2: Power delivered contours for the device size of $3 \times 0.12 \times 18$ μm^2. Contours are in 0.5 dB step; maximum P_{out} is 13 dBm.

Table 6.1: Input and output matching network parameters.

	Electrical Length	**Physical Length or Value**
Input Match		
Series Line	25.2^o	290 μm
Shunt Capacitor	—	205 pF
Output Match		
Series Line	44.2^o	520.8 μm
Shunt Capacitor	—	105 pF

B. The IBM technology is enhanced for much higher frequency design, and provides excellent models and parasitic information for passive components valid well into millimeter range. Thus, small values of capacitors can be used in a design with more confidence. Therefore, in order to conserve area, mixed distributed and lumped matching networks were used at both the input and output. Again, the series FGC elements were simulated in IE3D and the results used in ADS as 2-port s-parameters blocks, which replace the ideal transmission lines used in initial simulations. Table 6.1 provides the corresponding electrical and physical lengths of the series lines and the shunt capacitor values used in the matching networks. As with the Atmel design, *dc*-blocking capacitors were placed in series with the through lines.

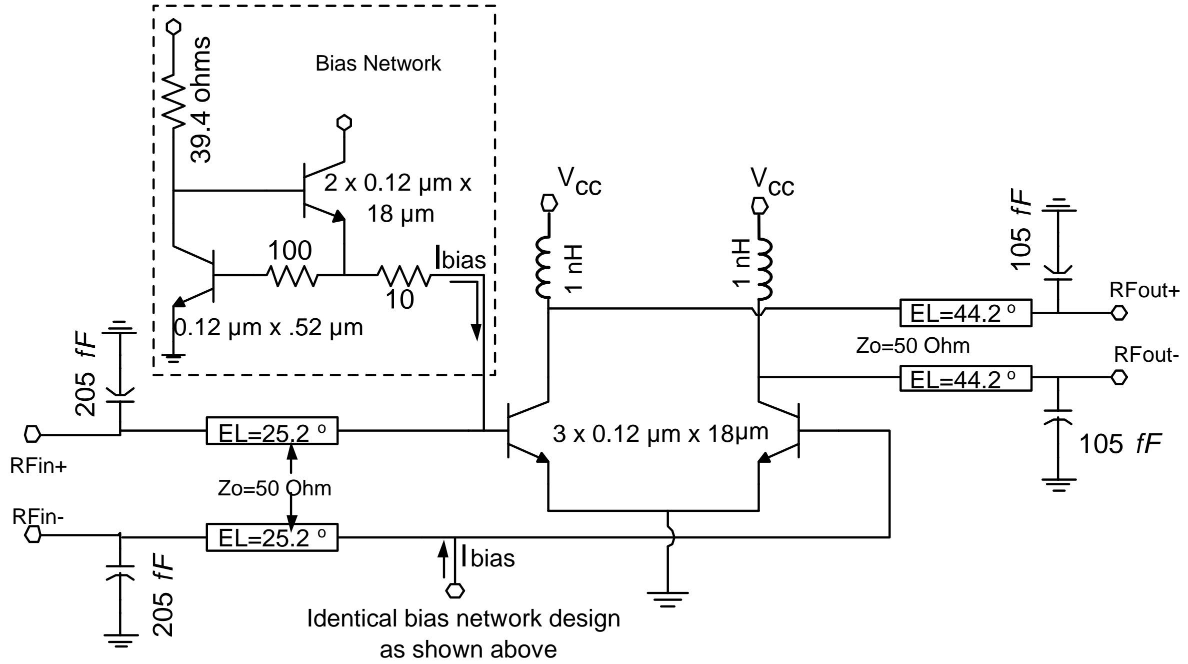

Figure 6.3: Simplified PA circuit diagram. Resistances are in Ohms.

6.1.2 Bias Circuit

Unlike the Atmel design that provided active stage bias using a current mirror, the bias circuit used in the IBM design is a more conventional one for PAs. A tail current source, such as in the Atmel design, was not used here due to the fact that the breakdown voltage of HBTs in the IBM process is lower than that of Atmel process (1.6 V to 2.4 V BV_{CE_O}), which reduces allowable supply voltage, and a tail current source reduces the output voltage swing available for generating output power. Another advantage of using this type of bias, as explained in Section 2.6, is that it keeps the base-emitter bias voltage, V_{BE}, relatively constant over input power levels; thus, preventing PA class (bias) from changing as the amplifier is driven harder.

In this case, two identical bias circuits are connected to the base of each power device in the differential pair. This is so because the designed bias circuit is connected to base of one of the device and does not support the base current/voltage for the other side of the differential pair. The bias circuit is similar to Figure 2.13 discussed in Section 2.6. Figure 6.3 shows the complete circuit for this design. For simplicity, the bias circuit is shown for only one side of the power amplifier circuit.

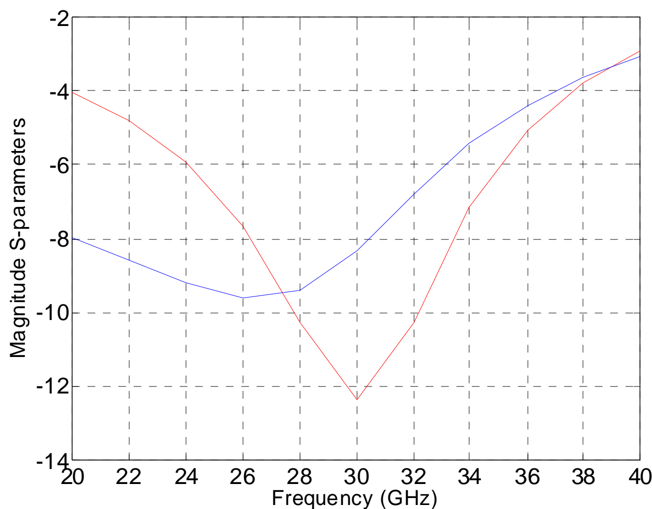

Figure 6.4: Simulated S_{11} and S_{22} verses frequency.

6.2 Simulation Results

The simulation results for the final design are shown below and were obtained using Cadence SpectreRF tools. Note that FGC lines were simulated using IE3D, and their results were included in the overall power amplifier simulation using s-parameters blocks. Both the s-paramters (small signal) and periodic steady state, PSS, (large signal) SpectreRF simulators were used to perform the simulations.

6.2.1 Small-signal Simulations

The input and output matching results are shown in Figure 6.4. S_{11} is -12.4 dB while S_{22} is -8.3 dB at 30 GHz. The Rollett's stability factor, K, is shown in Figure 6.5 along with small-signal gain. The corresponding small-signal gain is around 9.6 dB at 30 GHz and the amplifier is unconditionally stable around 30 GHz as indicated by $K > 1$. Note that Figure 6.5 is not a plot of MAG versus gain; rather, it is the plot of transducer gain, S_{21}, versus frequency of the power amplifier.

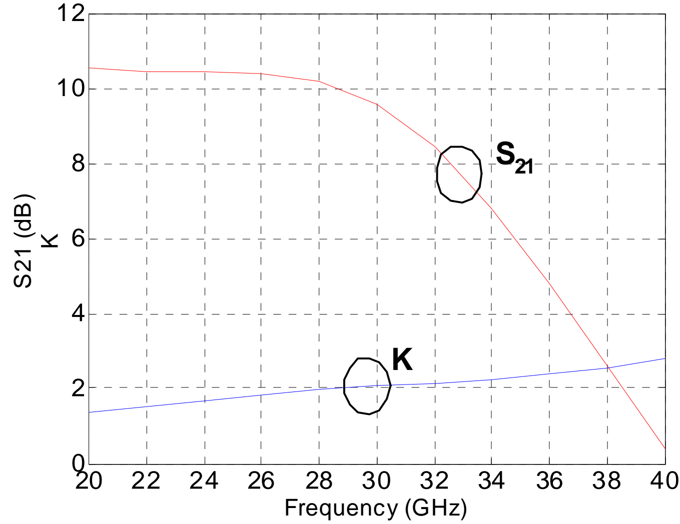

Figure 6.5: Simulated S_{21} and K verses frequency.

6.2.2 Large-signal Simulations

Large-signal simulations were done using the PSS engine in Cadence. The design procedure was carried out using RFIC DL because simulation for loadpull using Cadence's portAdapter component failed due to simulator/design kit issues. Also, Harmonic Balance simulator in ADS/RFIC DL is faster than and as accurate as PSS. Figure 6.6 shows the AM-AM plot, with the gain versus the input power super-imposed. The P_{1-dB} is 7.5 dBm and the output power at this point is 16 dBm. The peak gain is 9.6 dB at 30 GHz, which agrees with the plot shown in 6.5. Although the design was not optimized for efficiency, the Power Added Efficiency (PAE) is also simulated and is shown in Figure 6.7. As predicated, the PAE is much higher than that of the Atmel design, with a peak PAE of greater than 16%. Figure 6.8 shows the harmonics generated by the amplifier; as seen with the Atmel design, due to differential operation, the even harmonics are suppressed. The worst case harmonic distortion (due to the third harmonic) is -37 dBc.

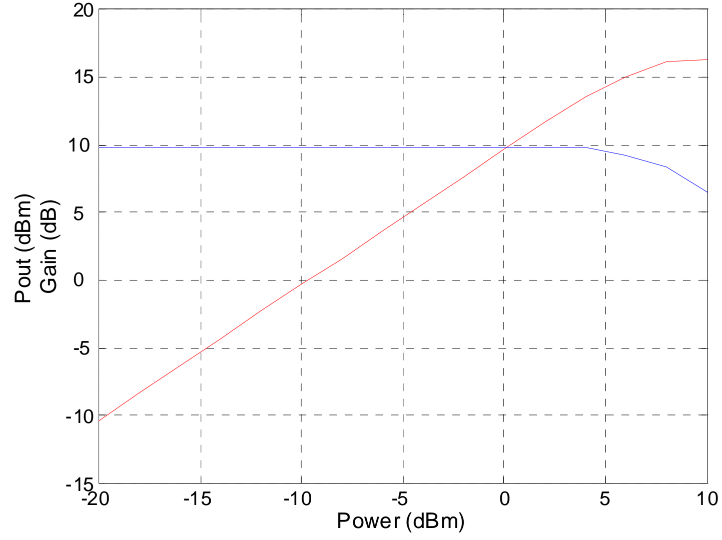

Figure 6.6: Simulated P_{out} and gain vs. P_{in}.

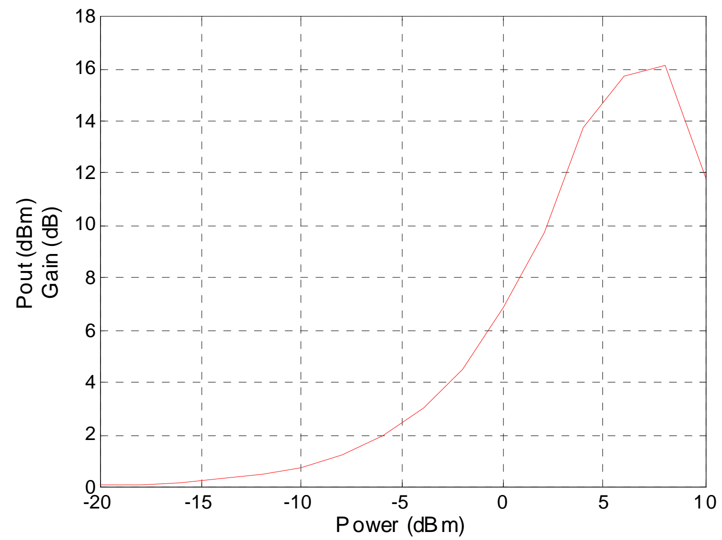

Figure 6.7: Simulated PAE vs. $P_{in.}$

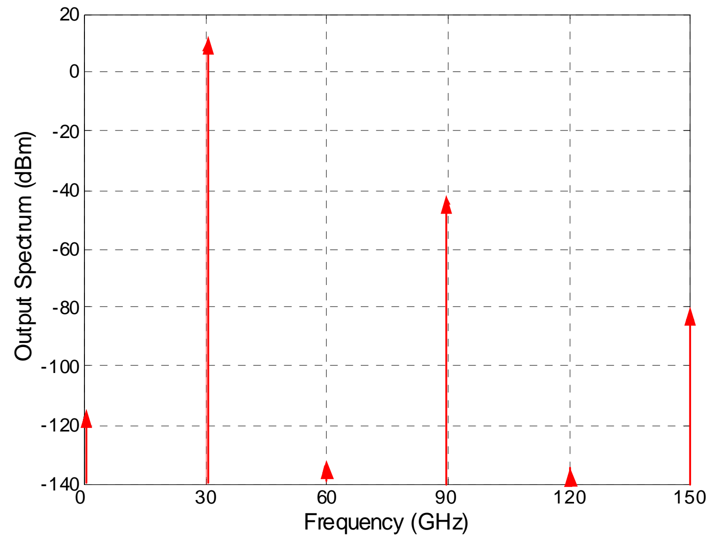

Figure 6.8: Simulated output spectrum versus harmonics. P_{in} = 7.5 dBm

6.3 Layout

As with the Atmel design, layout issues such as symmetry, parasitics and electromigration needed to be taken care of. The 8HP provides option for five, six or seven metallic layers. Seven metals layers were chosen so as to locate the passive components (transmission lines, inductors, etc.) as far as possible from the silicon substrate.

6.3.1 Resistors

There are four different types of resistors ranging from low sheet resistance to high sheet resistance. This PA design involved small resistor values in the bias circuit. Hence, a resistor type was chosen that had relatively low sheet resistance so that the resistor sizes were reasonable. The characteristics of this resistor is shown in Table 6.2.

Table 6.2: Resistors used in fabrication of the bias circuit.

Resistor	**Sheet Resistance**	**IC Layer**	**Current Handling**
KQ BEOL	60.5 Ω/sq	Above MQ (metal-5)	.5 $mA/\mu m$

6.3.2 Capacitors

Three types of capacitors are available in the 8HP process: MOS varactors, Hyper-abrupt (HA) Junction varactors and MIM capacitors. The MIM capacitors were used in matching networks as *dc* blocks because the cutoff/SRF frequencies of MOS and HA varactors were lower than or too close to the operating frequency of 30 GHz [46]. Also, varactor capacitance changes with voltage variation, introducing nonlinearity and variation in matching networks. Moreover, varactors need to biased, which might require a separate bias circuit other than the ones used for biasing power devices. The MIM capacitors in this technology are built between thin metal-5 and metal-6 layers, metal-5 being the top plate, and the capacitance per unit area for the MIM capacitors was 1.0 $fF/\mu m^2$. The shunt capacitor value needed in the input matching network resulted in capacitor dimensions with aspect ratio (W×L) of 2:1, while for output match the resultant capacitor dimensions had aspect ratio (W×L) of close to 1:1. Since the values of capacitors in the matching networks are relatively small, the parasitic bottom-plate capacitance of .0045 $fF/\mu m^2$ was taken into account for the capacitor design. Finally, 4 pF MIM capacitors were used as *dc*-blocks with aspect ratio of 1:1.

6.3.3 Inductors

There are two types of inductors available in this process. Both inductors are fabricated in 4 μm-thick analog metal (AM), which is the topmost metal, but differ in the type of ground plane used. DT inductors have deep trench as the ground plane while metal-1 is the ground plane for M1 inductors. As was the case with the Atmel design, the inductance of the different inductors was simulated using the inductor model to verify that the SRF was sufficient for use in mm-wave applications. Inductance simulations were performed by varying DT vs. M1 plane, the width of the metal, spacing between adjacent turns, number of turns and the inner radius of the

Table 6.3: Chosen inductor parameters.

Parameter	Value
Width of turns	5 μm
Spacing bet. adjacent turns	5 μm
Number of turns	2.5
Radius (outer dimension)	125 μm
Total inductance	1.1 nH
Self-resonant frequency	38 GHz
Ground plane	Deep trench

inner-most turn. It was seen that inductors with deep trench as the ground plane had higher SRFs compared to same inductors with metal-1 as the ground plane. Unlike with the Atmel technology, inductors in the 4-5 nH range were difficult to realize at mm-wave frequencies because their SRF fell below the target operating frequency of 30 GHz. Ultimately, it was found that a 1.1 nH inductance had SRF of 38 GHz while sufficiently blocking the RF signal from shunting to ac ground. Table 6.3 shows the parameters of the chosen inductor.

6.3.4 Final Layout

Taking into consideration all the aspects explained above, the final layout shown below in Figure 6.9 was realized. The layout passed all the Design Rule Checks (DRC) including antenna rules. Also, the Cadence Assura tool was used for layout versus schematic (LVS) check, which completed without errors. As with the Atmel design, the probe pitch determined the distance between the I/O pads, which were (W×L) 88×100 μm^2 in size. The total die area is 1.74×0.89 mm^2.

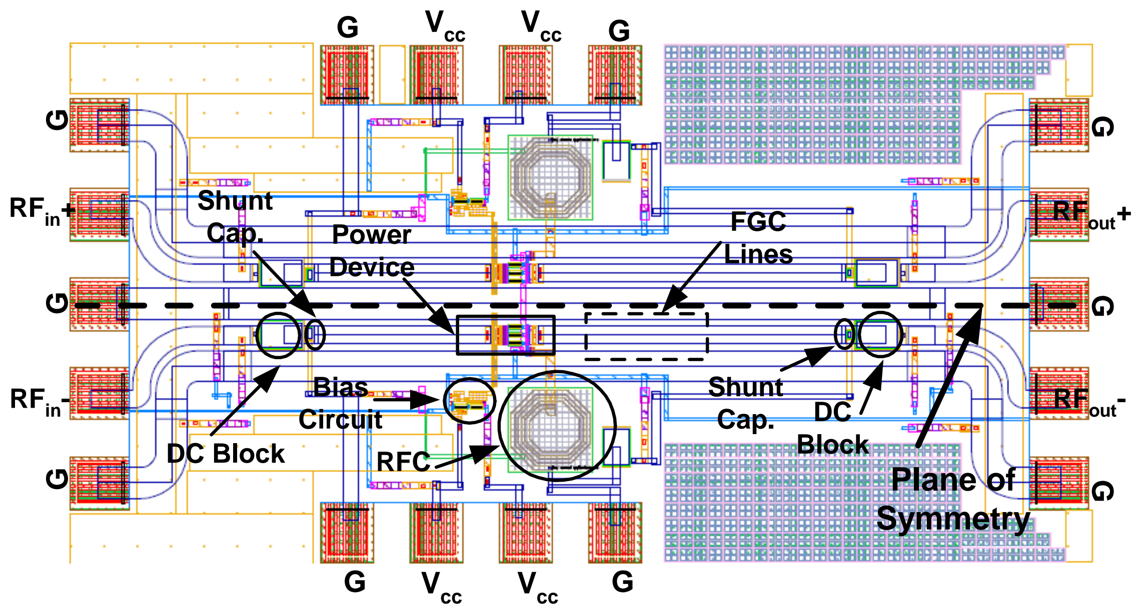

Figure 6.9: Complete layout of the IBM 8HP power amplifier. Die area is $1.74 \times 0.89\ mm^2$. Note that components of the layout are symmetric about the plane of symmetry.

Chapter 7

Conclusion and Future work

The objective of this thesis was the development of silicon-based microwave/millimeter-wave power amplifiers for use in wireless communication applications at microwave/mm-wave frequencies. Millimeter-wave applications include inter-vehicle communication (IVC), local multipoint distribution systems (LMDS) and WLAN. Applications at mm-wave frequencies tend to be point-to-point and use higher-order modulation schemes such as quadrature amplitude modulation (QAM), and hence high linearity is required at the cost of lowered efficiency for power amplifiers. The PA designs developed in this thesis demonstrated mm-wave range operation with output powers in the range of 12-16 dBm and high linearity. This chapter briefly summarizes the accomplishments of this thesis and suggests future research directions to further this work.

7.1 Summary

The SiGe HBT is historically advantageous over RFCMOS devices due to high gain availability and higher f_T, and therefore SiGe HBTs were used for the PA designs. However, the most significant advantage of using SiGe HBT is higher breakdown voltages compared to RFCMOS technology at comparable f_T. This allows for larger voltage swings and thus higher output power. The PA designs in this thesis were designed in Atmel SiGe2RF technology, which has f_T of 75-80 GHz for high frequency HBTs and IBM BiCMOS 8HP, which has f_T of around 220 GHz for high frequency

HBTs.

Several factors must be considered before selecting a device to deliver high output power, such as collector-emitter breakdown voltage, collector current density and maximum gain for a given bias. The breakdown voltage in silicon-based technologies is lower than that of GaAs or other III-V technologies. Hence, a large part of output power is developed by driving a high current into the load. Different device sizes were simulated to select those that met the specifications for collector current density and peak f_T performance, which would ensure maximum gain availability. Differential circuit topologies were used for both designs to ease the peak voltage swing on each device, and to exploit the inherent linearity advantages.

Initially, on-chip input and output tuned distributed matching networks using open-circuit shunt stubs and series lines were implemented in CPW (FGC) technology. Implementation in subsequent designs used lumped shunt capacitors in order to conserve die area. FGC structures were optimized using IE3D Full-wave EM simulator. The power amplifier designed in Atmel SiGe2RF technology was fabricated and measured, while the IBM 8HP BiCMOS design has not yet been fabricated. The most evident difference between the two design is the power gain. The Atmel PA has measured power gain of 2.9 dB at 30 GHz while the IBM design has a simulated power gain of 9.6 dB at the same frequency. Also, the circuit implementation for the IBM design resulted in better efficiency, but lower linearity as indicated by lower P_{1-dB} point of 8 dBm, whereas Atmel PA design has measured P_{1-dB} of approximately 10 dBm. High linearity of Atmel PA is seen by simulated EVM of just 2.2% at P_{1-dB}.

7.2 Future Work

This work represents initial prototypes for silicon-based power amplifiers at microwave/mm-wave frequencies (particularly 30 and 60 GHz bands). Several enhancements can be made to the designs such as higher power gain, output power, and better efficiency while keeping the high linearity. Furthermore, this work can be a basis for other silicon-based PA designs at higher frequencies.

7.2.1 Power Gain and Output Power Improvement

One of the key improvements that can be made to the designs detailed above is increasing the power gain of the PAs. SiGe HBT multi-stage PA designs have been demonstrated at very high microwave/mm-wave frequencies with gain in excess of 15 dB [13] [12]. Such power gain levels are desirable to reduce input drive level requirements. Due to the lack of wafer space, multi-stage designs were not pursued initially. In multi-stage PA designs, obtaining high linearity requires careful design of each stage, starting with the output or power stage. Careful planning is required for selecting the optimum ratio of device sizes between the different stages and designing inter-stage matching networks for the best possible linearity. However, due to the larger number of active devices in a multi-stage design, the overall efficiency would decrease.

As has been discussed in Chapter 2, the common-base configuration has higher breakdown voltage compared to common-emitter. Therefore, the CB configuration is an attractive option in silicon-based technologies, where BV_{CE_O} is typically quite small. On the other hand, due to sub-unity current gain, CB suffers from lower power gain and several stages might be required to obtain a reasonable value. Hence, CB configurations need to be carefully studied option to CE.

7.2.2 Linearity Improvement by Forced Input Mismatch

In a multi-stage design, the linearity of each of the amplification stage needs to be taken into consideration. Where output stage device is mainly constrained by the output power requirement, the driver stage can be optimized to achieve best overall linearity. As the drive level is increased beyond the small-signal regime, the input impedance will change. Hence, a match that works well for a specific power range will degrade with change in the drive level. By deliberately mismatching the input of the driver stage at a point approximately closer to the 1-dB compression point, better linearity performance can be achieved [52]. As shown in Figure 7.1 [53], this would result in some loss of gain at small-signal or linear region, while gain expansion will occur ideally at the 1-dB compression point.

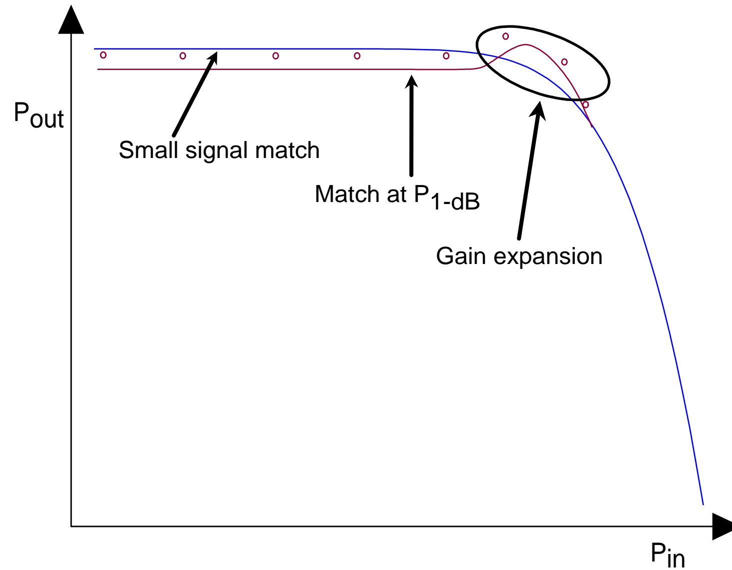

Figure 7.1: Gain expansion by mismatching.

7.2.3 60 GHz Two-stage Power Amplifier

The available unlicensed spectrum between 59 to 64 GHz has generated significant interest for high speed, wide-band communication, especially for short-range and in-door applications [54]. Hence, this thesis work can be extended to a mm-wave PA design at 60 GHz. 60 GHz and a 77 GHz power amplifiers have already been demonstrated in [14] and [55], but they suffer from very low efficiency and their linearity was not characterized. The goals of this design will be to improve efficiency while maximizing the linearity. Implementation of a prototype 60 GHz power amplifier has already been initiated. The design is biased in Class A, has two amplification stages, distributed/lumped on-chip matching networks, and a temperature and parameter insensitive bias on-chip circuit. This bias scheme will also potentially help in improving linearity.

7.2.4 Shielded-substrate Coplanar waveguide

As discussed previously, power-added efficiency is dependent upon the amplifier gain—the higher the transducer gain at a given bias, the higher the PAE. Thus, one way to improve efficiency without having to employ different efficiency-enhancing techniques

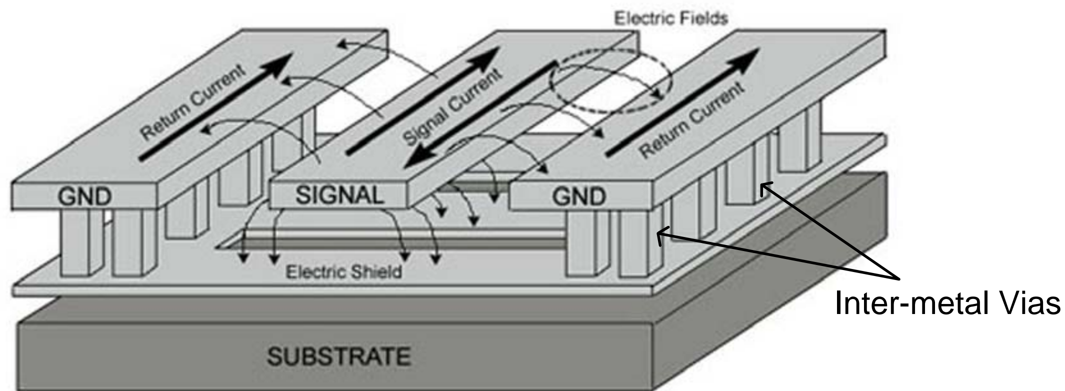

Figure 7.2: Shielded-substrate coplanar waveguide can be used to reduce attenuation loss caused by lossy silicon substrate. Permission granted by Abbas Komijani to reproduce the figure [12].

is to reduce losses in passive components. Shielded-substrate coplanar waveguide (SCPW) is one such option to reduce attenuation loss in distributed structures on silicon substrate. Figure 7.2 [12] shows the basic structure of such a SCPW line. By placing a shield below the signal line, electric fields can be prevented from penetrating into the lossy silicon substrate. Comparing this to a standard CPW or FGC where there is no shield between the top plane and substrate, stray fields can enter the substrate, thus increasing the component loss. The slotted shield is introduced to force most of the return current through the coplanar ground planes, thus SCPW current flows resemble that of conventional CPW. On the other hand, the structure is more likely to develop microstrip mode component, which requires careful modeling to achieve the desired line parameters.

Bibliography

[1] O. Andrisano, V. Tralli, R. Verdone, "Millimeter waves for short-range multimedia communication systems," *Proceedings of the IEEE*, vol. 86, pp. 1383–1401, July 1998.

[2] S.Y. Wang, "Integrating inter-vehicle communication with roadside wireless access points to provide a lower-cost message broadcasting service on highways," *IEEE 17th International Symposium on Personal, Indoor and Mobile Radio Communications*, no. 1, pp. 1–5, September 2006.

[3] W. Schafer, "Channel modelling of short-range radio links at 60 GHz for mobile intervehicle communication," *41st IEEE Vehicular Technology Conference*, no. 1, pp. 314–319, May 1991.

[4] J.R. Jones, "Baseband and passband transport systems for interactive video services," *IEEE Communication Magazine*, no. 1, pp. 90–101, May 1994.

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

[6] K. Lee, K. Choi, S. Kook, D. Cho, K. Park, and B. Kim, "Direct paramter extraction of SiGe HBTs for the VBIC bipolar compact model," *IEEE Transactions on Electron Devices*, vol. 52, no. 3, pp. 375–384, March 2005.

[7] F. X. Sinnesbichler and G. R. Olbrich, "Accurate large-signal modeling of SiGe HBTs," *IEEE MTT-S Digest*, vol. 2, pp. 749–752, June 2000.

[8] F. Arcioni, "SiGe HBT large-signal modelling and its application to the design of millimetre wave amplifiers," *Topical Meeting on Silicon Monolithic Intergrated Circuits in RF Systems Digest*, no. 1, pp. 105–112, September 1198.

[9] U. Pfeiffer, S. Reynolds, and B. Floyd, "A 77 GHz SiGe power amplifier for potential applications in automotive radar systems." Radio Frequency Integrated Circuits (RFIC) Symposium Digest of Papers, June 2004, pp. 91–94.

[10] V. Ilderem, et al., "The emergence of SiGe:c HBT technology for RF applications." International Conference on Compound Semiconductor Mfg., 2003.

[11] N. Kinayman, A. Jenkins, D. Helms and I. Gresham, "Design of 24 GHz SiGe HBT balanced power amplifier for system-on-a-chip ultra-wideband applications," *Digest of Papers IEEE Radio Frequency Integrated Circuits Symposium*, no. 1, pp. 91–94, June 2005.

[12] A. Komijani and A. Hajimiri, "A 24GHz, +14.5 dBm fully-integrated power amplifier in 0.18 um CMOS," *Proc. IEEE Custom Integrated Circuits Conference*, pp. 561–564, October 2004.

[13] S. Chartier, E. Sönmez, and H. Schumacher, "Millimeter-wave amplifiers using a 0.8 ţm Si/SiGe HBT technology," *IEEE Topical Meeting on Silicon Monolithic Integrated Circuits in RF Systems*, no. 1, January 2006.

[14] B. A. Floyd, S.K. Reynolds, U.R. Pfeiffer, T. Zwick, T. Beukema and B. GAucher, "SiGe bipolar transceiver circuits operating at 60 GHz," *IEEE Journal of Solid-State Circuits*, vol. 40, no. 1, pp. 156–167, January 2005.

[15] R. Jos, "Technology developments driving an evolution of cellular phone power amplifiers to integrated RF front-end modules," *IEEE Journal of Solid-State Circuits*, vol. 36, no. 9, pp. 1382–1389, September 2001.

[16] K. Bock, M. Bleier, O. Köthe and C. Landesberger, "New manufacturing concepts for ultra-thin silicon and gallium arsenide substrates," *International Conference on Compound Semiconductor Mfg.*, no. 1, 2003.

[17] S. Muthukrishnan, "ESD protected SiGe HBT RFIC power amplifiers," Master's thesis, Virginia Polytechnic Institute and State University, March 2005.

[18] J. Deng, "High-efficiency and high-linearity SiGe BiCMOS power amplifiers for WCDMA handset applications," Ph.D. dissertation, University of California, San Diego, 2005.

[19] J.D. Cressler, "SiGe HBT technology: A new contender for si-based RF and microwave cricuit applications," *IEEE Transactions on Microwave Theory and Techniques*, vol. 46, no. 5, pp. 572–589, May 1998.

[20] ——, "Silicon-germanium heterojunction bipolar technology: The next leap in silicon?" *IEEE International Solid-State Circuits Conference Digest of Technical Papers*, no. 5, pp. 24–27, 1994.

[21] T. Ytterdal, Y. Cheng and T. Fjeldly, *Device Modeling for Analog and RF CMOS Circuit Design*. England: John Wiley and Sons, 2003.

[22] K. Nellis and P.J. Zampardi, "A comparison of linear handset power amplifiers in different bipolar technologies," *IEEE Journal of Solid-State Circuits*, vol. 39, pp. 1746–1754, October 2004.

[23] I. Aoki, S. Kee, D. Rutledge, A. Hajimiri, "Fully integrated CMOS power amplifier design using the distributed active-transformer architecture," *IEEE Journal of Solid-State Circuits*, vol. 37, pp. 371–383, March 2002.

[24] *IBM BiCMOS 8HP Design Manual*. Essex Junction, VT: IBM, October 2005.

[25] S. C. Cripps, *RF Power Amplifiers for Wireless Communications*, 1999.

[26] T. Lee, *The Design of CMOS Radio-Frequency Integrated Circuits*, 2nd ed., 2004.

[27] A. Grebennikov, *RF and Microwave Power Amplifier Design*. United States: McGraw Hill, 2004.

[28] A. Amini, Ed., *Effects of High Peak-to-Average Ratio on 5 Ghz WLAN Power Amplifiers*. DesignCon, Agilent Technologies, 2002.

[29] Agilent PN 89400-14, Ed., *Using Error Vector Magnitude Measurements to Analyze and Troubleshoot Vector-Modulated Signals*. Agilent Technologies, 2000.

[30] *Agilent Advanced Design System Version 2003C*. Agilent Technologies, December 2003.

[31] H. Veenstra, G. Hurkx, D. Goor, H. Brekelmans, and J. Long, "Analysis and design of bias circuits tolerating output voltages above BVCEO," *IEEE Journal of Solid-State Circuits*, vol. 40, no. 10, pp. 2008–2018, October 2005.

[32] G. Freeman et. al., "40-gb/s circuits built from a 120-GHz fT SiGe technology," *IEEE Journal of Solid-State Circuits*, vol. 37, no. 9, pp. 1106–1114, September 2002.

[33] E. Jarvinen, S. Kalajo, and M. Matilainen, "Bias circuits for GaAs HBT power amplifiers," *IEEE MTT-S International Microwave Symposium Digest*, vol. 1, pp. 507–510, May 2001.

[34] J. Deng, P.S. Gudem, L.E. Larson, "Linearity analysis of SiGe HBT amplifiers using a power-dependent coefficient volterra technique." IEEE Radio and Wireless Congerence, September 2004, pp. 479–482.

[35] G.S.Dow, T.N. Ton, K. Nakano, "Q-BAND coplanar waveguide amplifier," *IEEE MTT-S Digest*, vol. 2, pp. 809–812, June 1989.

[36] G. Ponchak and L. Katehi, "Finite ground coplanar (FGC) waveguide: A better transmission line for microwave circuits," *Advancing Microelectronics*, vol. 1, pp. 15–18, May 1998.

[37] ——, "Open- and short-circuit terminated series stubs in finite-width coplanar waveguide on silicon," *IEEE Transactions on Microwave Theory and Techniques*, vol. 45, no. 6, pp. 970–976, June 1997.

[38] Robert W. Jackson, "Considerations in the use of coplanar waveguide for millimeter-wave integrated circuits," *IEEE Transactions on Microwave Theory and Techniques*, vol. 34, no. 12, pp. 1450–1456, December 1986.

[39] G. Ghione and M. Goano, "The influence of ground-plane width on the ohmic losses of coplanar waveguides with finite lateral ground planes," *IEEE Transactions on Microwave Theory and Techniques*, vol. 45, no. 9, pp. 1640–1642, September 1997.

[40] K. Belenhoff, W. Heinrich, and H. Hartnagel, "Analysis of t-junctions for coplanar MMICs," *International Symposium Digest*, no. 9, pp. 1301–1304, May 1994.

[41] C. Lee Y. Liu and T. Itoh, "The effects of coupled slot line mode and air-bridges on CPW and NLC waveguide discontinuities," *IEEE Transactions on Microwave Theory and Techniques*, vol. 43, no. 12, pp. 2759–2765, December 1995.

[42] M. Chapman, "A 60 GHz uniplanar MMIC 4x subharmonic mixer," Master's thesis, Virginia Polytechnic Institute and State University, November 2000.

[43] D. Pozar, *Microwave Engineering.* New York, NY: John Wiley and Sons, Inc., 1997.

[44] *Zeland IE3D.* Fremont, CA: Zeland Software, Inc., December, info@zeland.com.

[45] G. Gonzalez, *Microwave Transistor Amplifiers Analysis and Design*, 2nd ed. New Jersey: Prentice Hall, 1997.

[46] *BiCMOS8HP Design Manual*, IBM Microelectronics Division, June 2006.

[47] *SiGE2RF Design Manual*, Atmel, August 2005.

[48] S.L. Cripps, "A theory for the prediction of GaAs FET load-pull power contours," *IEEE MTT-S Digest*, vol. 83, pp. 221–223, May 1983.

[49] P. Gray and R. Meyer, *Analysis and Design of Analog Integrated Circuits*, 3rd ed., 1993.

[50] *Amplifier DesignGuide.* Help Document, Agilent Technologies, 2003.

[51] Microwaves101, "http://www.microwaves101.com/encyclopedia/connectors.cfm," Internet Web site.

[52] R. Mahmoudi, "A multidisciplinary design method for second and third generation mobile communication systems microwave concepts," Ph.D. dissertation, University of Twente, 2001.

[53] J.P.B Janssen, R. Mahmoudi, and A.H.M. van Roermund, "Design of a two-stage MMIC power amplifier at 30 GHz for LMDS," Eindhoven University of Technology, Tech. Rep., 2004.

[54] R. L. V. Tuyl, "Unlicensed millimeter wave communications a new opportunity for MMIC technology at 60 GHz," *Technical Digest of Gallium Arsenide Integrated Circuit (GaAs IC) Symposium*, pp. 3–5, November 1996.

[55] U.R. Pfeiffer, S. Reynolds and B. Floyd, "A 77 GHz SiGe power amplifier for potential applications in automotive radar systems," *Digest of Papers IEEE Radio Frequency Integrated Circuits Symposium*, no. 1, pp. 91–94, June 2004.

Vita

Talha Irfanul Haque was born on August 4th, 1982 in Karachi, Pakistan. He received his dual degrees in Bachelors of Science in Electrical Engineering and Bachelors of Science in Computer Engineering from the University of Missouri-Columbia with honors in May, 2004. In Fall of 2004, Talha moved to Blacksburg, Virginia to pursue Masters of Science degree in Electrical Engineering at Virginia Tech.

He worked at M/A-COM, Roanoke as an intern during the summers of 2005 and 2006. At M/A-COM, he worked on developing and improving models for nonlinear active devices and on-chip passive components. In the Fall of 2005, he joined the Wireless Microsystems Laboratory as a graduate research assistant to conduct research on wireless RF/mm-wave ICs.

Talha will have completed the requirements for the degree of Master of Science in Electrical Engineering in Spring 2007. After graduation, he will join M/A-COM in Roanoke, VA as a Staff Design Engineer.