High Linearity Broadband RF Vector Multiplier for Analog/RF Pre-distortion

by

Yiling Xu

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Applied Science
in
Electrical and Computer Engineering

Waterloo, Ontario, Canada, 2016

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Wireless communication systems are moving towards a heterogeneous solution, where small-cell base stations such as pico-cells and femto-cells are used concurrently with macro-cell base stations in high data traffic areas. Small-cell networks are expected to provide much larger wireless data rates and capacity in small areas while only consuming a fraction of the power. However, power amplifier nonlinearity does not scale down with the size of the base station; a similar degree of nonlinearity correction is required in both small-cell and macro-cell base stations, meaning that the power consumed by the signal linearization circuits is the same. An analog-radio frequency pre-distortion (ARF-PD) solution, operating at a fraction of a conventional digital pre-distortion's power consumption, has been proposed to support the unrestrained growth of wireless communication.

This thesis forms part of an on going research project aimed at developing a fully integrated ARF-PD solution – a promising, low-power alternative to digital pre-distortion for future wireless communications. Specifically, it focuses on delivering an integrated design of a low-power high-linearity broadband radio frequency (RF) vector multiplier, which can be used as part of the ARF-PD solution. An RF vector multiplier is considered one of the major function blocks in analog pre-distortion solutions, as it allows the analog pre-distorter to interface with the undistorted signal in the RF domain.

In the thesis, two RF vector multiplier designs are proposed and implemented in integrated circuits. In the first implementation, the RF vector multiplier is designed to directly apply pre-distortion to the RF signal. This architecture imposes a need for high gain in the RF vector multiplier, which results in large transistor size and high power consumption in the output stage. The design is able to achieve promising simulation results, however, performance limitations and disadvantages are also clearly exposed compared to commercial products. To resolve the issues discovered, an alternative ARF-PD architecture is adopted to relax the output power level needed from the RF vector multiplier. In addition, a self-linearized variable gain amplifier topology is proposed to improve system linearity. Overall, the second design shows significant improvement in bandwidth, linearity and output noise level, while only consuming half of the power consumed by the first design.

Ultimately, simulation results have shown satisfying performance for both RF vector multipliers as part of an ARF-PD system. However, both of the proposed integrated circuit designs should be validated by measurement.

Acknowledgements

I would like to thank Dr. Boumaiza for providing me with opportunities to be part of this group and this project. To Peter, Hai, and Anik - I would not be here without your guidance and constructive criticism.

Last, but not least, to Mom and Dad - thanks for taking care of me; to you two, I dedicate my thesis.

Table of Contents

List of Tables

List of Figures

ix

List of Abbreviations

ADC   Analog to digital converter

ADS   Advanced Design System is a electronic design software developed by Keysight for RF, microwave, and high speed digital applications

AM   Amplitude modulation

ARF-PD   Analog/radio frequency pre-distortion

CMOS   Complementary metal-oxide semiconductor

DAC   Digital to analog converter

dgnfet   Dual-gate negative-channel field effect transistor which has a minimum of 240nm gate length provided by the IBM 130nm design kit

dgpfet   Dual-gate positive-channel field effect transistor which has a minimum of 240nm gate length provided by the IBM 130nm design kit

DPD   Digital pre-distortion

DSP   Digital signal processing

IMD3   Third order intermodulation

IMD   Intermodulation distortion

IMRR   Image rejection ratio

MIM   Metal-insulator-metal capacitor provided by the IBM CMOS 130nm design kit

nfettw	A triple well nfet device provedes nfet within a p-well that is isolated from the substrate
nfet	Single-gate negative-channel field effect transistor which has a minimum of 120nm gate length provided by the IBM 130nm design kit
OP1dB	Output 1 dB compression point
OPG	Quadrature phase generation
oprppres	unsilicide precision polysilicon resistor provided by the IBM CMOS 130nm design kit
PAPR	Peak to average power ratio
PM	Phase modulation
PPF	Poly-phase filter
RFPA	Radio frequency power amplifiers
RFVM	Vector multiplier for radio frequency
RMS	root mean square
SBGM	Single balanced Gilbert multiplier
SCN	Small-cell base station network
VGA	Variable gain amplifier

Chapter 1

Introduction

1.1 Motivation

In modern wireless communication systems, radio frequency (RF) transmitters face many design challenges. RF power amplifiers (RFPAs) are one of the most power hungry and nonlinear components in the transmitter chain. Their performance has an overpowering effect over the system. The inherent trade-off between linearity and power efficiency is one of the most notable challenges for RFPA designers. Driven by the exponential growth in demand for higher wireless data rates and capacity, 4G and (future) 5G communication system designers are moving towards deploying highly dense small-cell base station networks (SCNs), improving spectral efficiency, and widening the modulation bandwidths of communication signals to handle high data traffic areas. These trends have necessitated an explicit need for linear and highly efficient RFPAs.

The need to support higher data rates and capacity has led to the adoption of such measures as widening modulation bandwidth, carrier aggregation, and complex modulation schemes such as 64 quadrature amplitude modulation (QAM). However, the adoption of widening signal modulation bandwidths and carrier aggregation has led to aggravated non-linearity behaviour and memory effects in RFPAs, especially in advanced RFPA topologies. Furthermore, complex modulation schemes enable higher spectral efficiency, by imposing stringent requirements on a transmitted signal's quality and this also impacts RFPA design. Concurrently, the deployment of viable cost-effective high SCNs in high data traffic areas, while capable of significantly increasing the communication network capacity and at the same time reducing power consumption[3], requires the deployment of many self-organizing, low cost, low power small-cell base stations that pose many new design challenges on RF-

PAs, especially in terms of amplifier efficiency and power consumption. In addition, the estimated carbon emissions from fixed and mobile telecommunications contribute 0.6% to the total global carbon emissions in 2010, and that percentage is increasing at an alarming rate each year[4]. The increasing pressure to reduce carbon emissions has created even more incentives for the development of highly efficient communication systems. RFPA topologies, such as Doherty[5] and Envelope Tracking[6] have been attracting attention. These topologies are used to improve both peak and back-off power efficiency for signals with high peak to average power ratios (PAPR). However, the improved efficiency is at the cost of linearity, due to the more complicated circuit architectures of advanced topologies.

In summary, recent trends in the development of communication systems are leading to more stringent requirements on both the efficiency and linearity of RFPAs. Efficiency improving RFPA topologies, combined with linearization techniques (used to counteract an RFPA's nonlinearity), can simultaneously achieve gains in both linearity and efficiency. Many linearization techniques have been extensively researched and they can be categorized into three groups based on the principles they follow: feedback linearization, feed-forward linearization, and pre-distortion [7]. Among them, baseband digital pre-distortion (DPD) has widely used for RFPA linearization in recent years [8, 9, 10]. However, the adoption of wider modulation bandwidths and small-cell base stations have made the power consumption of DPD linearization non-negligible, and it is difficult to reduce this without compromising its practicality. Hence, a low power analog-RF pre-distortion (ARF-PD) solution that preserves the capabilities of existing DPD solutions has been proposed [1]. As part of the on-going development of ARF-PD solutions, this thesis will focus on the development of an RF vector multiplier (RFVM), one of the key building blocks in an ARF-PD system. Eventually, an ARF-PD system will deliver a fully-integrated energy-efficient analog pre-distortion solution for future SCNs.

1.2 Thesis Organization

This thesis is devoted to the development of an RFVM in the scope of an ARF-PD solution, and the content is organized as follows.

Chapter 2 outlines the relevant background knowledge on the linearity-efficiency trade-off, and the effect of nonlinearity on RFPAs. Pre-distortion techniques are introduced and the way in which they counteract nonlinearities is explained, followed by a detailed description of the ARF-PD system and its design challenges. Lastly, a summary of literature on existing vector multipliers and those on the market is provided as a benchmark for this work.

Chapter 3 defines the design scope of possible RFVM systems in the proposed ARF-PD system configuration. The RFVM systems are composed of 3-stages: 1. a broadband multi-stage resister-capacitor (RC) poly-phase filter (PPF) as input quadrature phase generation (QPG); 2. a complex variable gain amplifier (VGA) core, comprised of two parallel VGAs; and 3. the output stage employed as a power booster and a load driver. Lastly, system level simulation results and some preliminary measurement results are presented, and the measurement results match closely to the simulated results. The design has achieved satisfactory performance, except the power consumption requirement has not been met.

Chapter 4 describes an alternative ARF-PD system configuration, which reduces the output power level needed from the RFVM system. A second RFVM design is implemented, targeting a wider RF operating bandwidth. An off-chip $90°$ hybrid coupler is used for QPG, which enables a flexible operating bandwidth at the cost of system integration. The improved RFVM system is also 3-stages: 1. a wideband active balun topology, simultaneously achieving single to differential conversion and input matching; 2. a modified complex VGA core design proposed to improve linearity; and 3. an output stage responsible for delivering the desired output power level. Lastly, system level simulation results are presented to show the improvement over the first design.

Chapter 5 concludes the thesis by summarizing and comparing both designs. Limitations of and potential improvements to both designs are provided. Future work to improve the feasibility of extending the design to more challenging and higher frequency scenarios is discussed.

Chapter 2

Literature Review

Despite decades of research on RFPAs, they remain one of the most challenging components in a RF transmitter from a design point of view, transmitters often contain more than one power amplifier. The growing demand for efficient and linear RFPAs has been constantly pushing the boundaries of RFPA research. New RFPA topologies and linearization techniques are being actively researched in an attempt to overcome the inherent limitation of an efficiency/linearity trade-off. Section 2.1 provides a review of distortions in RFPAs, followed by Section 2.2 which outlines techniques to linearize them. Section 2.3 provides an overview of ARF-PD and discusses its potential for next generation wireless communication systems (i.e., within SCNs). Section 2.4 explains the role of RFVMs in ARF-PD systems and highlights existing work and research related to their implementation.

2.1 Distortions in RFPAs

Modern RFPAs commonly contain a biasing network, input and output matching networks, and one or more active devices to perform signal amplification. Common RFPAs can be categorized into different classes [11]. Classes of operation differ in their biasing conditions, power efficiency, output linearity, etc. For instance, a Class-A power amplifier is biased so that it is operating in the most linear region, while suffering from poor drain efficiency at peak power compared to other power amplifier classes. On the other hand, Class-D power amplifier is biased and driven into the saturation region of the device, such that it is working in the switch mode. Class-D operation can achieve an ideal drain efficiency of 100%, but is extremely nonlinear due to significant harmonics generated at the output. Power amplifiers, the most power-hungry components in the transmitter chain, have inevitably

driven RFPA designers to pursue high efficiency topologies. However these topologies, while achieving greater efficiency, are accompanied by many forms of distortion. Distortions caused by RPFAs can be categorized into static nonlinearity or dynamic nonlinearity. The source of static nonlinearity for any RF active device is harmonic distortion, which can be summarized with the following equation:

$$y(t) \approx \alpha_1 x(t) + \alpha_2 x^2(t) + \alpha_3 x^3(t)... \tag{2.1}$$

In equation (2.1), if input $x(t)$ is a sinusoidal signal $Acos(\omega t)$, then

$$
\begin{aligned}
y(t) &= \alpha_1 Acos(\omega t) + \frac{\alpha_2 A^2}{2}\big(1 + cos(2\omega t)\big) + \frac{\alpha_3 A^3}{4}\big(3cos(\omega t) + cos(3\omega t)\big)... \\
&= \frac{\alpha_2 A^2}{2} + \left(\alpha_1 A + \frac{3\alpha_3 A^3}{4}\right)cos(\omega t) + \frac{\alpha_2 A^2}{2}cos(2\omega t) + \frac{\alpha_3 A^3}{4}cos(3\omega t)...
\end{aligned}
\tag{2.2}
$$

where the second term is the fundamental frequency which is distorted by the third harmonic distortion and, hence, no longer in a linear relationship with the input signal. Harmonic distortion can be expanded into two nonlinearity phenomena: gain compression and intermodulation distortion (IMD). Each of these phenomena have been extensively discussed in many modern microwave textbooks[11]. Gain compression is well illustrated in Fig. 2.1. Gain decreases as the amplitude of the input signal increases due to harmonic distortion. This effect can be quantified by the 1 dB compression point, where the actual gain of the power amplifier falls 1 dB below the ideal linear gain.
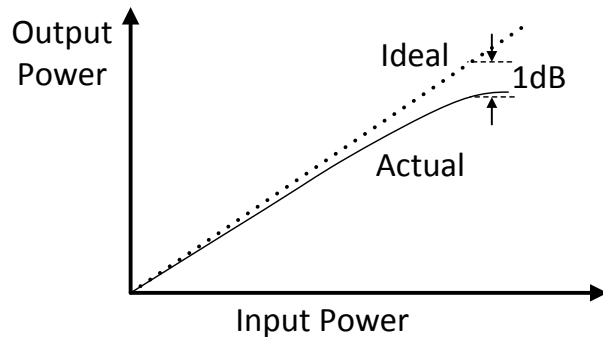
Figure 2.1: 1 dB gain compression of a RFPA.

Intermodulation in an RFPA causes spectral regrowth, where the output signal has high power content in the adjacent channels and causes interference. This effect can be captured by measuring adjacent channel leakage ratio, or by taking a simpler measure of third order intermodulation distortion (IMD3), as illustrated in Fig. 2.2(a) and 2.2(b), respectively.

(a)

(b)
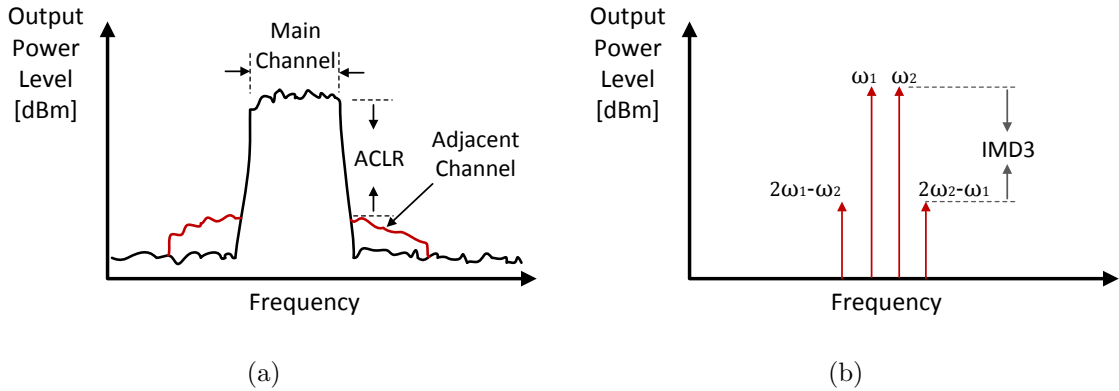
Figure 2.2: (a) ACLR for a modulated signal and (b) IMD3 for a two tone signal with frequencies ω_1 and ω_2.

(a)

(b)
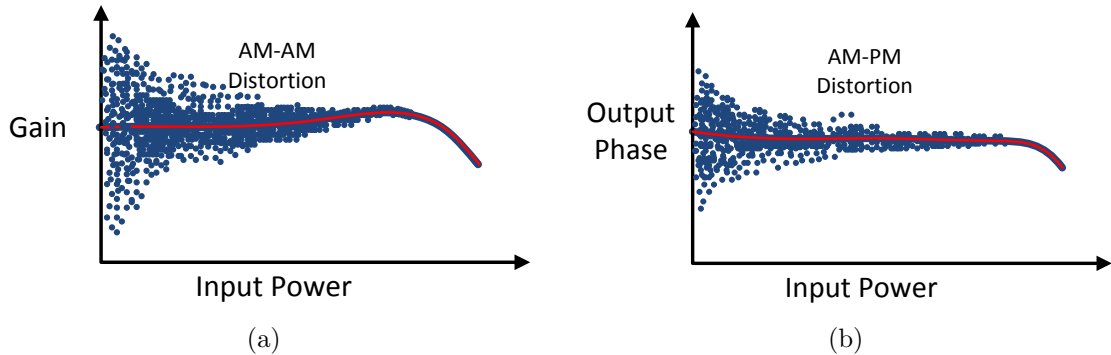
Figure 2.3: (a) AM-AM distortion and (b) AM-PM distortion.

Dynamic distortion or memory effects, only appear in time-variant systems where the output signal is not only a function of the present input signal but is also affected by the past input signal. Short term memory effects are primarily caused by extrinsic and intrinsic parasitic elements such as capacitance and inductance. Long term memory effects can be triggered by thermal effects and charge trapping. Regardless of the physical origins

6

of these memory effects, it is important to recognize that memory effects can introduce dynamic behavior in both the linear and nonlinear regions of operation. This effect is best captured by examining the amplitude and phase modulation (AM; PM). The AM/AM distortion captures the deviation of gain corresponding to the memory effect as illustrated in Fig. 2.3(a). Here, the redline represents gain affected by the static distortion present in the system and the blue dots represent scattered gain due to dynamic distortion. The AM/PM distortion captures the conversion of the amplitude modulation of the input signal to the phase modulation at the output signal, therefore capturing the amplitude modulation corruption of the phase of the output signal. Fig. 2.3(b) illustrates an AM-PM distortion; the red line showing the ideal output phase versus input power, and the blue dots representing the distorted phase of the output signal.

2.2   Linearization Techniques for RFPAs

<div style="text-align:center">(a)   (b)</div>
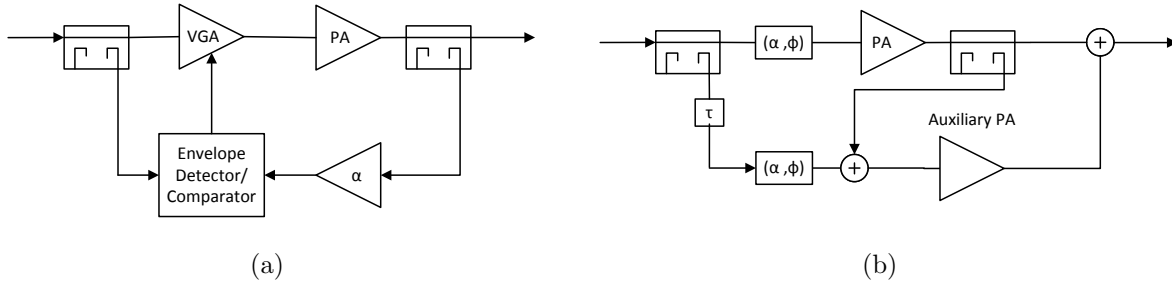
Figure 2.4: Block diagrams of (a) an envelope feedback system and (b) a feedforward system.

As stated earlier, linearization techniques can be broadly categorized into three different groups based on the techniques and principles followed[7, 12, 13]. The direct feedback technique uses a feedback loop to feed a portion of the output signal to the input of the power amplifier. It is rarely used in practical RFPA design due to the instability and loss of gain it causes in the system. Alternatively, feedback techniques such as Cartesian feedback and envelope feedback are more widely used in modern RFPA designs. Fig. 2.4(a) illustrates a block diagram of an envelope feedback system. In this type of system, the envelopes of the input and output signals are compared and, then the difference between the two is used to control the VGA to correct the distortion. However, feedback techniques

in general have limited bandwidth and are therefore impractical for wideband applications. In contrast, feedforward techniques are able to achieve good linearization across wideband signals. Fig. 2.4(b) illustrates a block diagram of a feedforward system. Here, a sample of the output signal is combined with a sample of the input signal to obtain a cancellation signal through attenuation and phase shifting. The cancellation signal, with proper scaling through the auxiliary power amplifier, is able to subtract the distortion component from the output signal. However, the need for an auxiliary amplifier incurs significant power overhead for the system, making this approach unsuitable for SCNs.

The basic concept behind pre-distortion technique is as follows: if the transfer characteristic of an RFPA is known, an artificial distortion can be applied to the input signal, such that the distorted input cancels the distortion produced by the RFPA. Hence the name, pre-distortion. The concept is illustrated in Fig. 2.5, where the red curve represents the distortion created by the RPFA, and the blue curve is the pre-distorted input signal. The pre-distortion techniques can be applied in both digital and analog domains and are capable of achieving wide signal bandwidths[7].
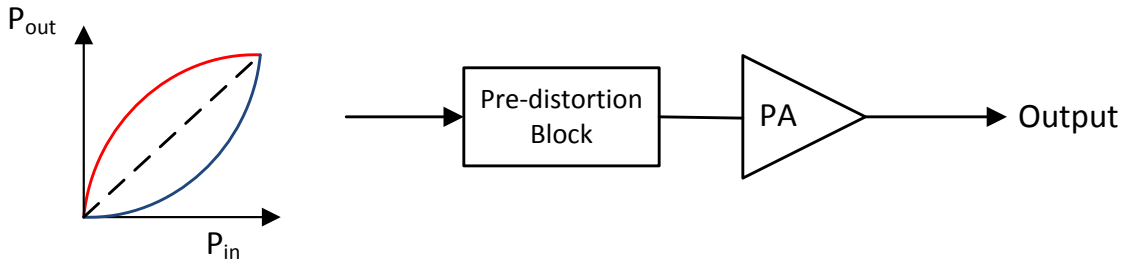
Figure 2.5: Pre-distortion concept.

However, nonlinearity in RFPAs is not a fixed constant; the fixed correction provided by the basic pre-distortion technique lacks the ability to account for nonlinearity deviations due to process, temperature, and load impedance variations. This drawback can be eliminated by employing advanced polynominal based pre-distortion techniques, where the pre-distortion system requires adaptive feedback to track the RFPA's nonlinear transfer characteristics.

DPD uses a considerable amount of digital signal processing (DSP) power to generate the pre-distorted signal. With the rapid advancement in semiconductor devices and the increasing accessibility of digital circuitry, DPD can be easily integrated with the baseband front end of wireless transmitters. Among the many advanced pre-distortion techniques

available, DPD has become the predominant linearization solution for wireless network base stations in recent years [14, 9]. Fig. 2.6 illustrates a typical DPD system. A sample of the distorted output signal is down converted to the digital domain through the feedback path and passed on to the DSP unit. The DSP unit generates the nonlinear function coefficients for the pre-distortion engine through either direct[15] or indirect learning[16]. Coefficients are generated whenever there is a change in the operation condition of the power amplifier. The pre-distortion engine generates the pre-distorted signals based on the coefficients to be up-converted to the RF domain through digital to analog converters (DACs) and mixers. Despite its popularity within the current generation of wireless communications, DPD faces many challenges when considering its use in future SCNs. The biggest of these being caused by the memory polynomial function, which is used to model PA distortions and generate the pre-distortion coefficients used by the pre-distortion engine. The engine expands the signal bandwidth by five-fold and effectively increases the DAC speed by five times, resulting in significant power consumption. In the current generation of wireless networks, most deployed base stations are macro-cells with tens and hundreds watts of power consumption in which the power consumption of a DPD system has minimal impact on the overall system efficiency. Unfortunately, the power consumed by DPD does not scale down with RFPA output power level. In small-cell base stations with less than 2 watts of power consumptions, such DPD power consumption would significantly negatively impact the system. Alternative low power solutions with linearization capabilities on par with DPD are urgently needed.
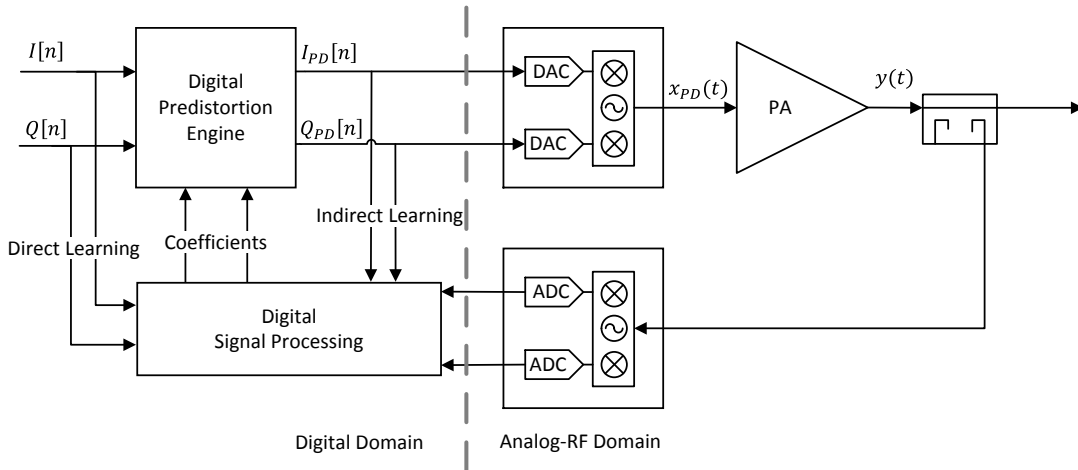
Figure 2.6: Typical DPD system block diagram with direction and/or indirect learning.

2.3   Analog-RF Pre-distortion

In a DPD system, complicated digital circuitry is adopted to synthesize the pre-distortion functions and fast speed DACs are needed to convert the signal from the digital to the analog domain. As per the Shannon-Nyquist theorem, the sampling rate of a DAC has to be at least twice the baseband bandwidth to avoid signal integrity degradation during conversion and an even higher sampling rate is needed as a certain safety margin. As discussed previously, the pre-distorted signal bandwidth is expanded by five times the original baseband signal due to the nonlinear functions[17], which further increases the speed requirement of the DACs along with rest of the digital circuitry. For the feedback path, high speed analog to digital converters (ADCs) are needed to digitize the transmitted signal, which its bandwidth has been expanded due to the spectral regrowth in the RFPA. Both the high speed DACs and ADCs in a DPD system consume a significant amount of power in a small-cell base station that cannot be overlooked. In response to the emerging trends of widening bandwidth and low power consumption solutions, an ARF-PD system with pre-distortion formulations has recently been proposed[1, 18]. The aim of this system is to offer the wireless communication field a solution that provides similar linearization capability to DPD, but more manageable power consumption.
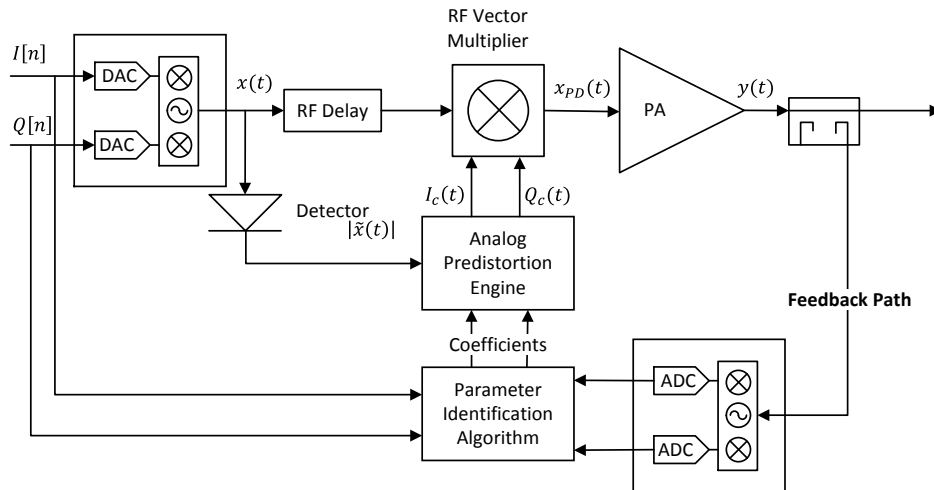
Figure 2.7: Proposed ARF-PD system block diagram[1] .

The proposed ARF-PD architecture is illustrated in Fig. 2.7, showing the pre-distortion engine is being realized in the analog domain as opposed to the digital domain. This allows the signal to be pre-distorted after the DACs, which effectively reduces the sampling speed

10

of the DACs and their power consumption. However, the architecture of the feedback path remains the same as in DPD, and improvements in the feedback path needs be addressed by future work in this area. The pre-distortion engine adopts the proposed formulations in [1] to generate control signals $I_c(t)$ and $Q_c(t)$. This requires static coefficients and the envelope of the undistorted RF signal $x(t)$, which can be obtained from the envelope detector. The power consumed by the digital circuitry when generating the static coefficient does not degrade the overall power consumption since it is only needed when operation conditions change in the PA. The control signals $I_c(t)$ and $Q_c(t)$ are functions of the envelope of the undistorted signal $x(t)$. Since the envelope of a modulated RF signal only varies at the frequency of the baseband signal, the analog pre-distortion engine is only required to operate at the baseband frequency. One unique and important block in an ARF-PD solution is the RFVM. It is used to apply the control signals generated by the distorter to the undistorted signal $x(t)$ by adjusting its amplitude and phase. However, unlike the ease of digital circuitry implementation, the design of analog and RF circuitry faces many challenges such as noise performance, circuit non-idealities, device mismatches, and so on. This thesis focuses solely on the design and development of an RFVM for an ARF-PD system.

2.4 RF Vector Multiplier

As discussed previously, many formulations used by DPD systems are incompatible with an analog pre-distorter. New pre-distortion formulations have been proposed to accommodate the restrictions imposed by the hardware design challenges of the proposed ARF-PD architecture. The theory behind RFVMs is fairly straight forward and is illustrated in Fig. 2.8.
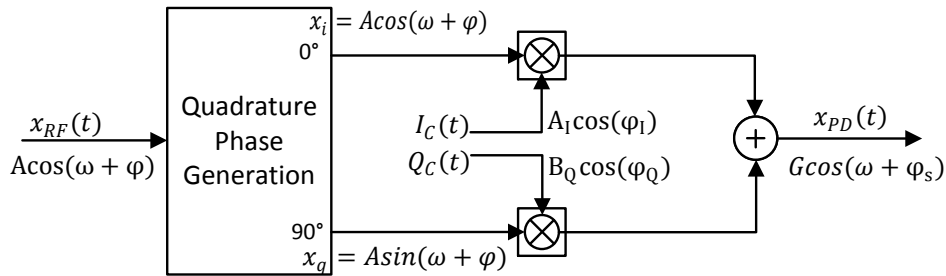
Figure 2.8: Basic operation of an RF vector multiplier.

It can be seen that the undistorted signal $x_{RF}(t)$ is split through a QPG block into two signals which are 90° out of phase: the in phase signal $x_i(t)$ and quadrature signal $x_q(t)$. The two signals are then multiplied by their respective control signals $I_c(t)$ and $Q_c(t)$. This may result in signals' independent attenuation or amplification before they are recombined using an in-phase combiner. By manipulating the magnitude of $x_i(t)$ and $x_q(t)$, the effect of amplitude and phase control can be described by equations (2.4) and (2.5).

$$x_{PD}(t) = x(t)(Ae^{0j} + Be^{-\frac{j\pi}{2}}) = x(t)Ge^{j(\varphi)} \tag{2.3}$$

$$\text{Gain Control: } G = \sqrt{A^2 + B^2} \tag{2.4}$$

$$\text{Phase Control: } \varphi = \arctan(\frac{B}{A}) \tag{2.5}$$

where A and B are magnitude modifiers of $x_i(t)$ and $x_q(t)$ respectively, dictated by the generated control signals.

Vector-sum topology has been used extensively to realize low-cost integrated phase shifters, one of the most important components for phased array antennas in wireless communications[19, 20, 21, 22], but it is rarely mentioned in the literature on pre-distortion[23]. Despite the difference in applications, vector-sum topologies developed for phase shifters can be easily extended to RFVMs.

In the concept of vector-sum theory, the key step is the generation of in-phase and quadrature phase signals. Any mismatches in amplitude or phase between the signals would result in degradation of the signal integrity. The QPG network also imposes strict restriction on the achievable operation bandwidth of the vector-sum topology. In [21], a passive 90° hybrid was implemented using branch-line couplers. It achieved 15% bandwidth at 45 GHz center frequency with $0.5dB$ gain error and 5° of phase error. However, passive couplers are not as attractive in lower RF applications due to their narrow bandwidth and large area consumption. In [20], inductor-capacitor (LC) all pass filters have been realized for IQ generation. It was possible to achieve an operation bandwidth from 2.3 GHz to 4.8 GHz with a root-mean-square (RMS) gain and phase errors of less than 1.1 dB and 1.4°. Although, the achieved bandwidth is much better when using LC all pass filters rather than passive couplers, the area required for implementing inductors on an integrated circuit (IC) is still significant. An active solution is proposed in [19] using operational transconductance amplifier integrator to achieve IQ generation. However, the achieved bandwidth is much

smaller than in [20] and the RMS gain and phase errors are also worse. In [23], a phase-shifter was used for phase pre-distortion and the QPG was implemented using RC PPF. This approach was able to achieve an operation bandwidth from 1 GHz to 2.1 GHz while only using a two-stage RC PPF. The bandwidth could be further extended using a higher number of stages, and the RMS amplitude and phase errors could be kept relatively low at the cost of an additional insertion loss.

In both [22] and [20], Gilbert-cell based VGAs were implemented as gain-variable block following the QPG block. In both cases, the researchers were able to achieve a $360°$ continuous phase shift over a very broad bandwidth. In [22], by using an off-chip $90°$ hybrid coupler being for IQ generation, measurements showed an achieved bandwidth of 0.5 GHz to 6 GHz.

The Aforementioned designs are summarized and compared with two off-the-shelf packaged RFVMs offered by Analog Devices, Inc. and Maxim Integrated in Table 2.1.

Table 2.1: Reviewed Phase Shifter and Vector Multiplier Performance Summary and Comparison

	[20]	[22]	[23]	ADL5390[24]	MAX2045[25]
Gain (dB)	-8 to -3	8 to 10	4.8	5 to 30*	6.5 to -8.5*
Phase Range	$360°$	$360°$	$90°$	$360°$	$360°$
RF bandwidth (GHz)	2.3-4.8	0.5-6**	1-2.1	0.04-2.4**	2.04-2.24
Control Bandwidth (MHz)	-	-	-	200	230
Power Consumption (mW)	19	28	4	675	800
Gain Imbalance (dB)	1.10	0.95	1.50	0.80	0.20
Phase Imbalance	$1.4°$	$7.0°$	$2.5°$	$2.5°$	$1°$
OP1dB (dBm)	-6.0	-14.0	-2.0	9.6	13.2
OIP3 (dBm)	-	-	10.0	18.7	21.5

*Gain control range for vector multiplier.

**RF operation bandwidth excluding quadrature phase generation.

Chapter 3

Proposed Design 1

The RFVM acts as an interface between the analog pre-distortion engine and the RF signal path for an APD-RF system as illustrated in Fig. 3.1. The system will be implemented in IBM CMOS 130nm technology and it has to support: 1. broadband RF operation for sub-3GHz application; 2. up to 400 MHz baseband bandwidth; 3. optimized linearity with -3 dBm maximum input power and; 4. a power consumption below 200 mW which is at 10% power consumption of typical micro-cell base stations.
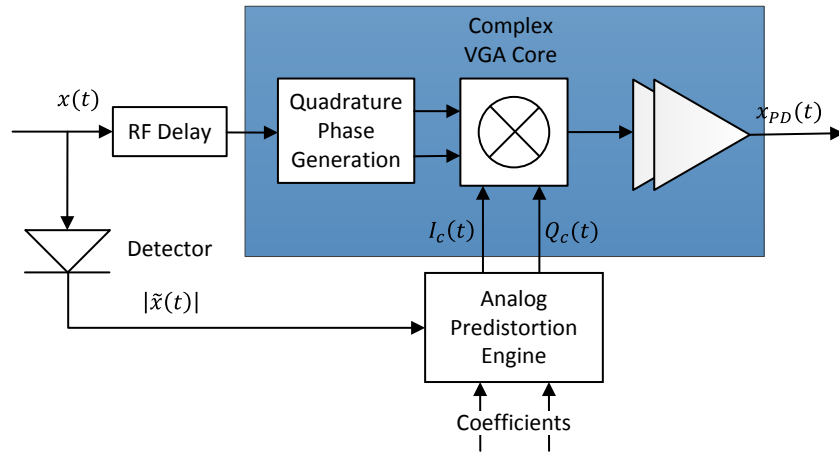
Figure 3.1: Block diagram of the APD-RF system configuration.

3.1 Multi-stage RC Poly-phase Filter Network

Quadrature signal generation is a vital part of the modern RF front-end, most commonly used for generating local oscillator signals in quadrature for direct-conversion and low intermediate frequency transceivers. In Chapter 2, multiple QPG solutions have been discussed and compared. For a broad-band integrated solution, a CMOS multi-stage RC PPF network was chosen to implement the quadrature generation function in the RFVM system.

3.1.1 RF Poly-phase Filter Fundamental

An RC PPF network is constructed of a sequenced asymmetric network of resistors and capacitors. PPFs can be differentiated into two basic types, Type I and Type II, which are slightly different from each other, as illustrated in Fig. 3.2.
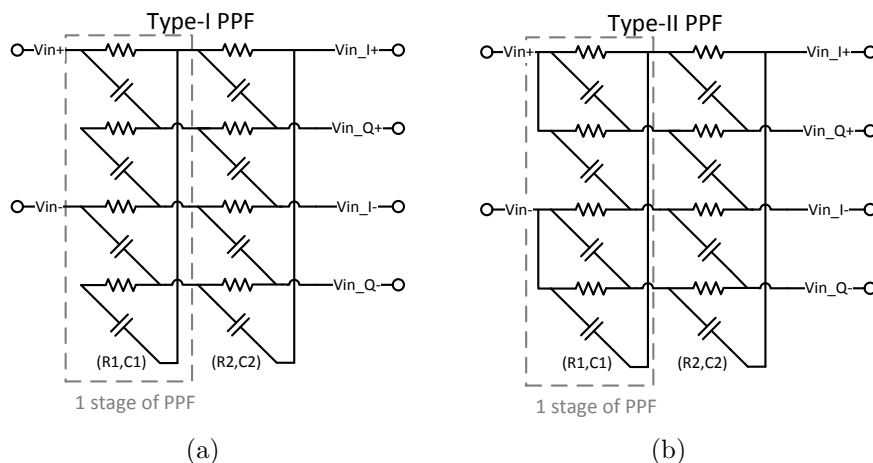
Figure 3.2: 2-stage RC poly-phase filter: (a) Type I and (b) Type II.

Analyses on RC PPFs have been presented many times in the literature [26] [27]; For a single-stage Type I PPF, the transfer functions of the in-phase signal output and quadrature signal output can be derived as equations (3.1) and (3.2) respectively, where Z_L is the load impedance or input impedance of the next stage. To analyze the relationship between I and Q output signals, one transfer function is divided by the other which results in (3.3). The new transfer function (3.3) is purely imaginary, meaning there is constant $90°$ phase difference between I and Q output signals across all frequency. Its only pole at

$\frac{1}{RC}$ gives a unity magnitude of the I and Q signals. The magnitude difference between the two signals starts to deviate as frequency moves away from the frequency pole at $\frac{1}{RC}$.

$$\Delta V_{out,I} = \frac{Z_L}{R_1 + Z_L + R_1 Z_1 j\omega C_1} \tag{3.1}$$

$$\Delta V_{out,Q} = \frac{j\omega R_1 C_1 Z_L}{R_1 + Z_L + R_1 Z_1 j\omega C_1} \tag{3.2}$$

$$\frac{\Delta V_{out,I}}{\Delta V_{out,Q}} = \frac{1}{j\omega R_1 C_1} \quad \text{and} \quad \left| \frac{\Delta V_{out,I}}{\Delta V_{out,Q}} \right|_{\omega = \frac{1}{RC}} = 1 \tag{3.3}$$

For a single-stage Type II PPF, the transfer functions of the in-phase signal output and the quadrature signal output can be derived as equations (3.4) and (3.5) respectively. As the opposite of Type I topology, transfer function (3.6) reveals that Type II PPFs can achieve a unity magnitude balance across all frequencies, however the phase difference between the I and Q output signals is only 90° at $\frac{1}{RC}$ frequency and deviates as the frequency changes.

$$\Delta V_{out,I} = \frac{Z_L}{R_1 + Z_L + R_1 Z_1 j\omega C_1}(1 - j\omega R_1 C_1) \tag{3.4}$$

$$\Delta V_{out,Q} = \frac{Z_L}{R_1 + Z_L + R_1 Z_1 j\omega C_1}(1 + j\omega R_1 C_1) \tag{3.5}$$

$$\frac{\Delta V_{out,I}}{\Delta V_{out,Q}} = \frac{(1 - j\omega R_1 C_1)}{(1 + j\omega R_1 C_1)} \quad \text{and} \quad \left| \frac{\Delta V_{out,I}}{\Delta V_{out,Q}} \right| = 1 \tag{3.6}$$

At $\frac{1}{RC}$ pole frequency, where a unity magnitude and 90° phase difference are achieved for both PPF topologies, equations (3.1) and (3.4) can be simplified to (3.7) and (3.8), respectively. The output magnitude for a Type I PPF is $\sqrt{2}$ smaller, from which it can be concluded that a Type II PPF inherently has 3dB lower insertion loss than a Type I PPF.

$$\Delta V_{out,I} = \frac{Z_L}{R_1 + Z_L + jZ_1} = \left| \frac{1}{1 + j} \right|_{Z_L \to \infty} = \frac{1}{\sqrt{2}} \tag{3.7}$$

16

$$\Delta V_{out,I} = \frac{Z_L}{R_1 + Z_L + jZ_L}(1-j) = |\frac{1-j}{1+j}|\bigg|_{Z_L \to \infty} = 1 \qquad (3.8)$$

The image rejection ratio (IMRR) is the most commonly used metric in the context of receivers to evaluate the suppression level of the undesired image sideband, and is defined by the magnitude ratio of the desired sideband to the image sideband. The amplitude and phase imbalance of the quadrature signals in receivers directly contribute to image sideband. The IMRR was first calculated by Norgaard[28] and is expressed as

$$IMRR = \frac{1 + 2\alpha\cos(\Delta\theta) + \alpha^2}{1 - 2\alpha\cos(\Delta\theta) + \alpha^2} \qquad (3.9)$$

where α is the amplitude imbalance and $\Delta\theta$ is the phase imbalance. For A PPF, The IMRR is also an important specification to evaluate its performance. A typical narrow-band 3dB hybrid coupler(i.e. Anaren Xinger) has an IMRR between -25dB and -30dB.

3.1.2 Multi-stage RC Poly-phase Filter Design Guidelines

In 2008, Kaukovuori [27] proposed detailed guidelines for designing multi-stage RC PPFs. The proposed guidelines can be summarized into multiple design steps which optimize intrinsic and extrinsic insertion losses of the multi-stage PPF within the targeted operating bandwidth and help to achieve a high IMRR over this bandwidth:

Step 1: Determine the number of cascaded RC stages required to achieve the targeted operating bandwidth and IMRR over the bandwidth.

Step 2: Relative bandwidth should be calculated: $BW_{rel} = \frac{\omega_{max}}{\omega_{min}}$.

Step 3: Determine the source impedance Z_S and load impedance Z_L. The source impedance is the impedance looking into the stage driving the PPF network. The load impedance is the input impedance of the following stage. Generally both source and load impedances are frequency dependent, average values should be taken based on corner frequencies of the operating bandwidth.

Step 4: Pole frequency is the $1/RC$ frequency of each PPF stage. The pole frequency location of each stage can be optimized to achieve lower loss and higher IMRR over the entire bandwidth. For example, a three-stage PPF would have three

frequency poles. Optimized pole locations would satisfy the following relationship: $\omega_1 = k\omega_2 = k^2\omega_3$. Parameter k can be determined by relative bandwidth: $BW_{rel,3-stg} = 2k^2 - 1.9k + 0.9$. To optimize the IMRR over entire operating bandwidth, Kaukovuori showed that it was best to place the ω_2 pole at the geometric center frequency, which can be determined by: $\omega_c = \sqrt{\omega_{min}\omega_{max}} = \sqrt{BW_{rel}}\omega_{min} = \frac{\omega_{max}}{\sqrt{BW_{rel}}}$.

Step 5: The resistor value of each stage can be determined base on Z_S and Z_L and used to minimize the total insertion loss of the PPF. For example, a three-stage PFF as demonstrated in [27], has capacitors of equal value at each stage which results in minimum intrinsic loss. The resistor value of the first stage can be calculated as $R_1 = \frac{|Z_L|}{k_L}$, where parameter k_L can be calculated differentially for Type I and Type II PPFs:

$$k_{L,3-stg}^{TypeI} = k\sqrt{2k_z} \tag{3.10}$$

$$k_{L,3-stg}^{TypeII} = k\sqrt{k_z} \tag{3.11}$$

where k_z is simply the ratio of Z_L over Z_S. Then the capacitor value can be calculated using $\frac{1}{RC}$ pole frequencies as determined in step 4 above.

3.1.3 3-stage Poly-phase Filter Schematic Design

The overall system specifications included an operating bandwidth from 600 MHz to 2.8 GHz, results in a relative bandwidth BW_{rel} of 4.6. A three-stage Type II CMOS PPF is sufficient to achieve more than -35dB of IMRR for the given operating bandwidth. The minimum IMRR for a three-stage PPF can be estimated using equation (3.12)[27]. A Type II PPF was chosen over Type I to reduce the overall insertion loss, which means there would be less impact on power consumption and output noise of the overall system.

$$IMRR_{min,3-stg} = \left(\frac{\sqrt{k}+1}{\sqrt{k}-1}\right)^3 \left(\frac{k-\sqrt{k}+1}{k+\sqrt{k}+1}\right) \tag{3.12}$$

The resistor and capacitor values of the three-stage Type II PPF are summarized in Table 3.1. The source impedance is assumed to be 50 Ω. The load impedance is mostly capacitive, and is estimated to be 80 fF. The Z_L at 2.8 GHz for a 80 fF capacitive load

18

is used for calculating resistor values: $|Z_L| = |\frac{1}{j2\pi(2.8GHz)(80fF)}| = 710$ Ω. As discussed previously, the CMOS 130nm technology provided by IBM was used. The design kit provides various types of resistors and capacitors. An unsilicided polysilicon resistor (oprppres) was used as it provides the highest sheet resistivity and allows the most compact layout design with the lowest absolute resistance tolerance ($\pm 8\%$). The absolute resistance tolerance can be further improved by using larger resistors connected in parallel. Metal-Insulator-Metal (MIM) capacitor was used as it provides the best linearity and quality factor with $\frac{L}{W}$ ratio between $\frac{1}{3}$ and $\frac{1}{2}$.

Table 3.1: Design Parameters for three-stage Type II Poly-phase Filter

Components	Width	Length	Multiplicity	Unit Value	Effective Value
R1	$0.96\mu m$	$0.83\mu m$	2	300.58Ω	150.29Ω
R2	$0.97\mu m$	$0.80\mu m$	1	290.1Ω	290.1Ω
R3	$0.76\mu m$	$1.39\mu m$	1	558.19Ω	558.19Ω
C	$20.21\mu m$	$10\mu m$	1	$423.79fF$	$423.79fF$

Schematic level AC simulations, sweeping through 0.6 GHz to 2.8 GHz, were performed to evaluate the designed three-stage Type II PPF using keysight's Advanced Design System's (ADS) RFIC Dynamic Link with Cadence. The source and load impedance were set as 50 Ω and 80 fF respectively. Amplitude and phase Imbalances are illustrated in Fig. 3.3(a) with the respective IMRR in Fig. 3.3(b).
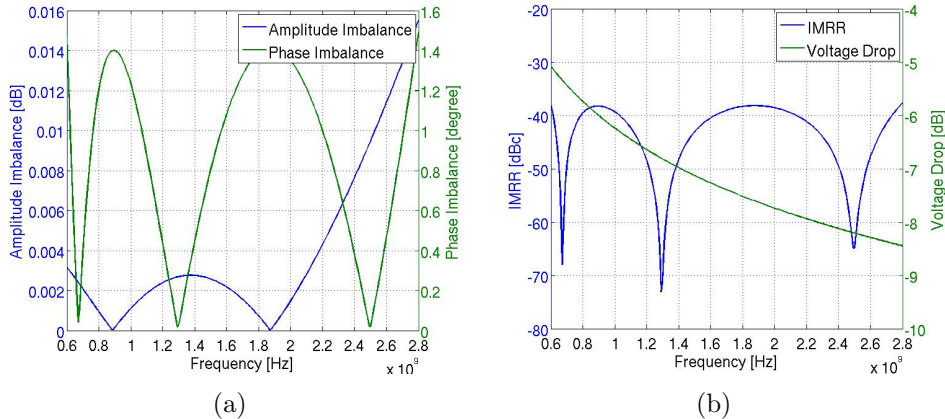
Figure 3.3: 3-stage poly-phase filter (a) amplitude (blue) and phase (green) imbalance. (b) IMRR (blue) and overall voltage gain (green) when filter driving 80 fF capacitive load.

The overall insertion loss of the 3-stage PPF is shown in Fig. 3.3(b). There is a maximum of -8.5 dB insertion loss at 2.8 GHz; post amplification is needed at the output stage to meet the required output power level. Further tuning for the 3-stage PPF was necessary when combining with the layout and following stages.

3.1.4   3-stage Poly-phase Filter Physical Layout Design

Due to the asymmetrical nature of the RC PPF, its physical layout required extra attention to match the line properties of each signal branch. Dummy resistors were used to minimize the boundary effects in resistor banks as shown in Fig. 3.4. The same method was not used for capacitor banks due to area restriction. All of the components were placed in close proximity and at the same orientation to reduce the effects of process variation and component mismatch. A common-centroid layout structure was used, shown in Fig. 3.4 to ensure good matching between signal paths.
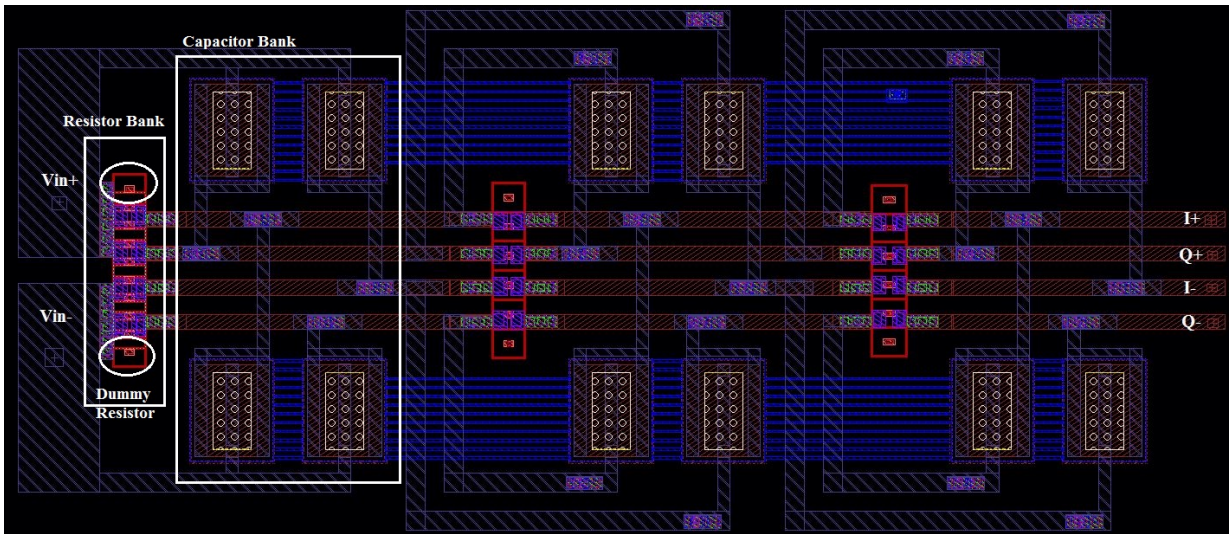
Figure 3.4: Layout of the 3-stage Type II PPF

3.2   Complex Variable Gain Amplifier Core

The core of the vector multiplier is two parallel VGAs, taking in the complex RF signals generated by the QPG network. The two VGAs are controlled by separated baseband

signals. In the context of ARF-PD system, the complex VGA core must achieve the following:

- RF operating bandwidth beyond 2.8 GHz,

- baseband bandwidth up to 400 MHz,

- high baseband and RF-band linearity.

As discussed in Chapter 2, Gilbert-cell multipliers are the most commonly used and most suitable architecture for VGAs; they were also selected for use in this thesis design work.

3.2.1 Single Balanced Gilbert Multiplier

A standard single balanced Gilbert multiplier (SBGM) has two inputs and one output as illustrated in Fig: 3.5(a), allowing for two possible input configurations.

(a) (b)

Figure 3.5: (a) A single balance Gilbert multiplier and (b) a source coupled differential pair.

The circuit topology is first analyzed in large signal. The differential output current of a source-coupled differential amplifier with a topology similar to that of a SBGM, as illustrated in Fig: 3.5(b), can be described by equation (3.13)[2], where W and L are the width and length of the transistor and I_{ss} is the total tail current. By replacing the

current source in Fig: 3.5(b) with a third transistor and input V_2 as in Fig: 3.5(a), the current flowing through the bottom transistor will be a function of V_2, thus leading to equation (3.15), where v_{th} is the threshold voltage and k' is the technology constant of the transistor. If the magnitude of V_1 is significantly smaller than V_2, an approximation can be applied to simplify the equation (3.15) to (3.16), where the output current ΔI_{out} is a linear multiplication of input signals V_1 and V_2.

$$\Delta I_{out} = \frac{k'}{2}\frac{W}{L}V_1\sqrt{\frac{4I_{ss}}{k'\frac{W}{L}} - V_1^2} \tag{3.13}$$

$$I_{ss} = \frac{k'}{2}\frac{W_3}{L_3}(V_2 - v_{th})^2 \tag{3.14}$$

$$\Delta I_{out} = \frac{k'}{2}\frac{W_1}{L_1}V_1\sqrt{2\frac{W_3}{W_1}(V_2 - v_{th})^2 - V_1^2} \tag{3.15}$$

$$\Delta I_{out} = \left(\frac{k'}{2}\frac{W_1}{L_1}\sqrt{2\frac{W_3}{W_1}}\right)V_2V_1 \tag{3.16}$$

As discussed in the previous section, the maximum input power level of the system is -3 dBm according to the design specification, which is equivalent to 0.45 V_{pp}. Taking into account the insertion loss at the three-stage PPF, the input voltage level at the RF port of the complex VGA core is estimated to be 0.2 V_{pp}. The system also specifies a baseband input voltage swing of 1 V_{pp}. Hence, to ensure a linear operation of the Gilbert multiplier, input V_1 should be the RF port and the input V_2 should be the baseband port. Linearity issues will still exist when large attenuations are needed, or where a large input swing of V_1 and a small input swing of V_2 are present.

In addition to the linearity issue, feed-through is another well-known problem present in the SBGM topology. By applying AC signal analysis, equation (3.17) can describe the current generated by the bottom transistor (M3) with input V_2; it includes both DC and AC components. Assuming the multiplier is operating in a linear region, the output current can be described by equation (3.18). The first term is the undesired feed-through that contaminates the signal quality. The second term is the desired linear product of the two input signals.

$$I_{SS} = I_{D3} + i_{d3} = I_{D3} + V_2 gm_{M3}, \quad where \quad \frac{1}{gm_{M3}} = \frac{V_{eff}}{2I_D} \tag{3.17}$$

$$\Delta I_{out} = \frac{v_1}{2V_{eff}} I_{D3} + \frac{v_1 v_2 gm_{M3}}{2V_{eff}} \tag{3.18}$$

The constant phase shift is also an important specification for VGAs. Unfortunately, the SBGM topology suffers from gain dependent phase shift due to the change of I_{ss} for various gain settings[29]. The RF signal fed through the VGA experiences different phase shifts with different gain settings, leading to undesired and uncontrolled phase shift and causing phase distortion in the signal.

3.2.2 Double Balanced Gilbert Multiplier

Although the SBGM topology suffers from many limitations when used for the VGA block, most of those limitations can be eliminated by extending to a double balanced Gilbert multiplier (DBMG) as illustrated in Fig. 3.6 [2].
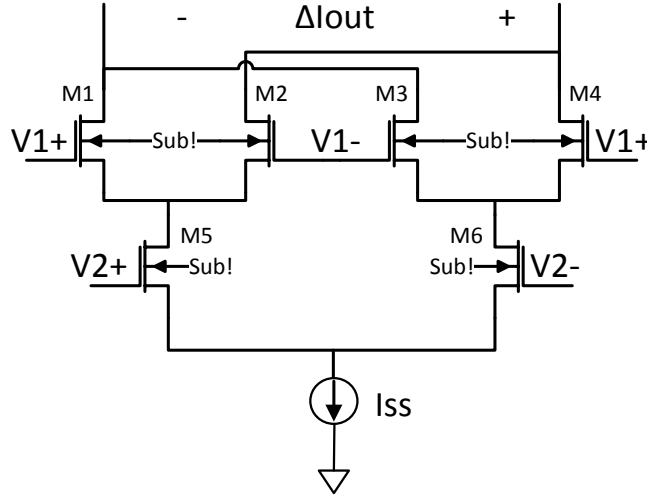

Figure 3.6: Double balance Gilbert multiplier.

Essentially, two SBGMs with an anti-parallel connection eliminate the direct feed-through of the RF input to output as demonstrated in equation (3.19), where $v_1 = v_1^+ - v_1^-$ and $v_2 = v_2^+ - v_2^-$.

$$\Delta I_{out} = (\frac{v_1}{2V_{eff}}I_{D3} + \frac{v_1 v_2^+ gm_{M3}}{2V_{eff}}) + (\frac{-v_1}{2V_{eff}}I_{D3} + \frac{v_1 v_2^- gm_{M3}}{2V_{eff}})$$

$$= \frac{(v_1 v_2^+ - v_1 v_2^-)gm_{M3}}{2V_{eff}} \tag{3.19}$$

$$= \frac{v_1 v_2 gm_{M3}}{2V_{eff}}$$

The linearity of a DBGM is also improved. For large signals, although the output current retains the same mathematical relationship as in equation (3.15), V_2 is a differential signal formed by V_c^+ and V_c^- with a common mode voltage. With a large gain setting, either V_c^+ or V_c^- is large, making at least one of the multiplier pairs work in the linear region and the other pair either turn off or operate with a small bias current. With a small gain setting, both V_c^+ and V_c^- would be close to the common mode voltage, thus neither would be small in the case that if a relatively larger common mode voltage is chosen for the design.

In addition, a constant phase shift with different gain settings is achieved with a DBGM topology as found by Klumperink in [29]. The gain dependent phase shift of an SBGM is mostly due to the transistor parameters such as tran-conductance, parasitic junction capacitance, and drain to source resistance, and it varies with bias current I_{ss}. Due to the anti-parallel connection of the DBGM, the sum of the bias current at the output nodes is constant with respect to different gain settings. Since the output's common mode voltage is determined by the bias current flowing through the resistive loading, the junction capacitance (largely dependent on the output common mode voltage) would become constant, as the sum of the output current stays constant. The total trans-conductance at each output node is also proportional to the bias current, hence, would also become constant.

With all of the advantages summarized above, it is important to note that there are a few disadvantages to the DBGM topology such as increased supply voltage, power consumption, and total noise output. The increased power consumption and total noise output are due to the circuit's greater relative complexity and the larger number of transistors being used. These are the trade-offs that must be made in order to eliminate the previously mentioned problems of the SBGM topology. A folded topology can be adopted to reduce the required supply voltage. The topology was first introduced in [30]. In this structure, the bottom differential transistors are replaced with a PMOS differential pair that acts to steer the current into the main branches. The schematic is as demonstrated in Fig. 3.7.

24
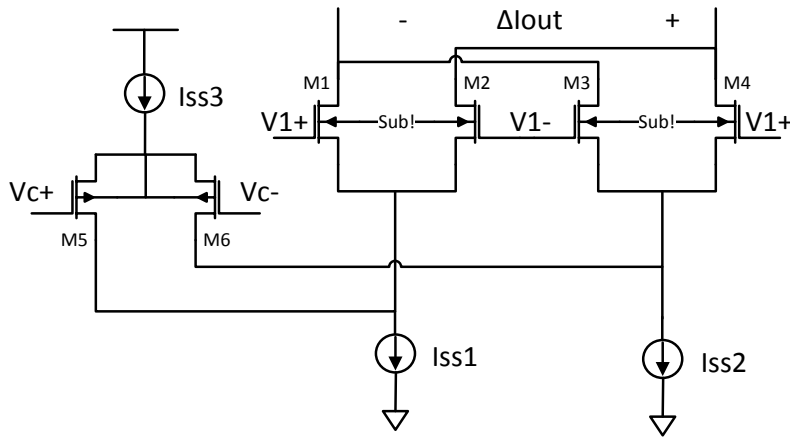
Figure 3.7: Schematic of a folded DBGM.

3.2.3 Complex VGA Core Schematic Design and Physical Layout Design

The first specification given to the VGA core is the RF bandwidth. The VGA core has to drive a capacitive load (the input capacitance of the following stage), giving the simplified equivalent circuit as illustrated in Fig: 3.8.

Figure 3.8: Equivalent circuit of a resistive source driving a capacitive load.

Its fundamental matching bandwidth limitations were derived by Bode in [31]. The bandwidth limitation is determined by the dominant pole frequency at ω_1 which can be calculated using equation (3.20), where C_M is the Miller capacitance. The effective gain

25

of the equivalent circuit starts to decrease at a rate of 20 dB per decade of frequency after the dominant pole frequency. To extend the RF operating bandwidth, the domina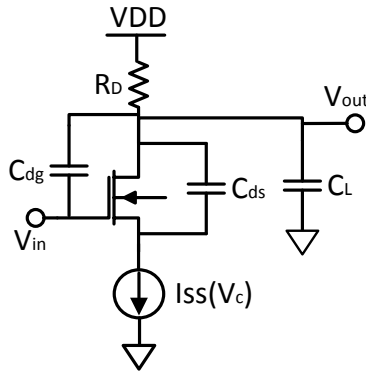nt pole frequency ω_1 has to be increased by reducing the size of the following: R_D, C_c, or C_L. The parasitic capacitance C_c can be calculated using equation (3.21)[2], where C_{gd} and C_{ds} are the junction capacitances from the drain to gate and drain to source, respectively. Both the trans-conductance g_m and the junction capacitance C_c are proportional to the transistor size and effective gain of the circuit. The relationships described above all lead to a clear design trade-off between bandwidth and gain.

$$\omega_1 = \frac{1}{R_D(C_c + C_L)} \tag{3.20}$$

$$C_c = C_{dg} + C_{ds} \tag{3.21}$$

As illustrated in Fig. 3.7, there are total of three source-coupled differential pairs, each with a dedicated tail current source. To ensure proper operation of the DBGM, two of the bottom tail current sources have to be exactly matched (i.e., $I_{ss1} = I_{ss2}$). The top current source has to provide a DC biasing current greater than or equal to the bottom tail current sources (i.e., $I_{ss3} \geqslant I_{ss1}$). When transistor M_5 is fully on and transistor M_6 is fully off, all of the DC current I_{ss3} would be pushed into I_{ss1}, effectively forcing the differential pair (M_1,M_2) to turn off. This leaves the second differential pair, M_3 and M_4, operating at maximum gain. When transistors M_5 and M_6 are both turned on at the same voltage level, i.e. $v_c+ = v_c-$, half of the DC biasing current I_{ss3} would be pushed into I_{ss1} and the other half would be pushed into I_{ss2}. This would effectively force both differential pairs (M_1,M_2 and M_3,M_4) operate in the same condition. The output AC current, ΔI_{out}, would become zero due to the anti-parallel connection of the two differential pairs.

In Fig. 3.9, DC transfer characteristic of a MOS source-coupled pair is illustrated. While the tail current is kept constant, the overdrive voltage V_{OV} determines the input range of the differential pair. To ensure that all three differential pairs operate linearly with large input voltage swings, the transistors have to be designed with a large V_{OV}. The V_{OV} of a transistor can be estimated using equation (3.22), which shows that the overdrive voltage can be increased by increasing the DC bias current or by reducing the transistor size. These relationships lead to a clear design trade-off between linearity, power consumption, and gain.

$$V_{OV} = \sqrt{\frac{2I_{d1}}{k'(W/L)}} \tag{3.22}$$
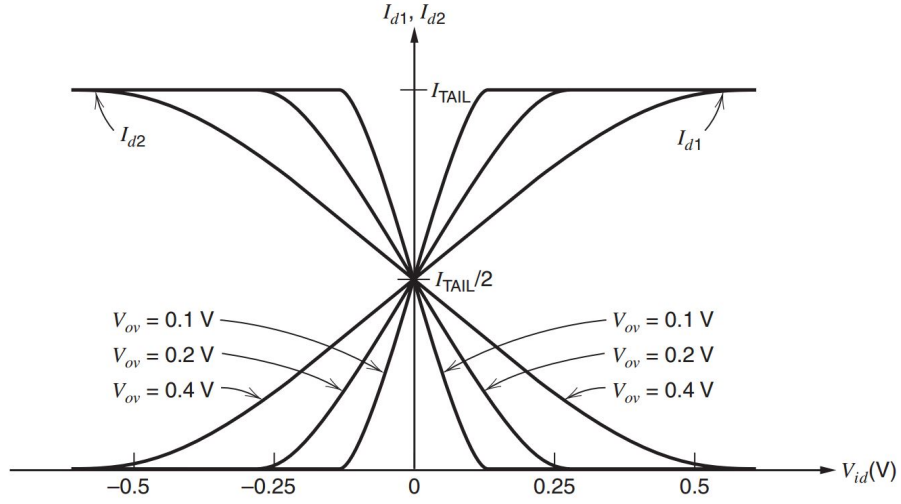
Figure 3.9: DC transfer characteristic of a MOS source-coupled pair[2].

The final schematic design of the complex VGA core is illustrated in Fig. 3.10, and all of the transistor sizes are summarized in Table 3.2 where V_i and V_q are the input quadrature signals and V_{ci} and V_{cq} are the baseband control signals for the complex VGA core. On-chip shunt peaking inductors, L_D, have been used for bandwidth extension[32].

Figure 3.10: Final schematic of the complex VGA core.

Table 3.2: Design Parameters for Complex VGA Core

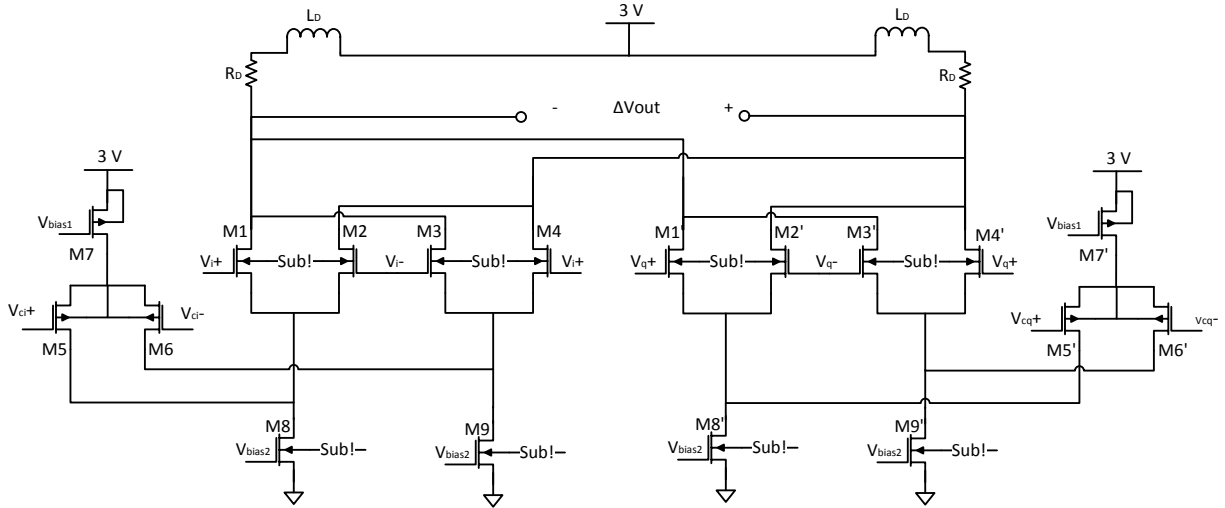| Transistors	Unit Length (nm)	Unit Width (μm)	Number of Fingers	Effective Width (μm)	Transistor Model
M1($M1'$)	240	4	8	32	dgnfet
M2($M2'$)	240	4	8	32	dgnfet
M3($M3'$)	240	4	8	32	dgnfet
M4($M4'$)	240	4	8	32	dgnfet
M5($M5'$)	240	2.5	10	25	dgpfet
M6($M6'$)	240	2.5	10	25	dgpfet
M7($M7'$)	240	4	20	80	dgpfet
M8($M8'$)	240	3	10	30	dgnfet
M9($M9'$)	240	3	10	30	dgnfet
Resistor	**Width**	**Length**	**Multiplier**	**Resistance**	**Model**
R_D	4.5 μm	16 μm	5	167.29 Ω	oprppres
Symmetrical Inductor	**Outer Dimension**	**Turns**	**Coil Width**	**Space**	**Effective Inductance**
L_D	260 μm	5	8.5 μm	5 μm	5.7 nH

The final layout of the complex VGA core is shown in Fig. 3.11, which follows a common-centroid layout structure.
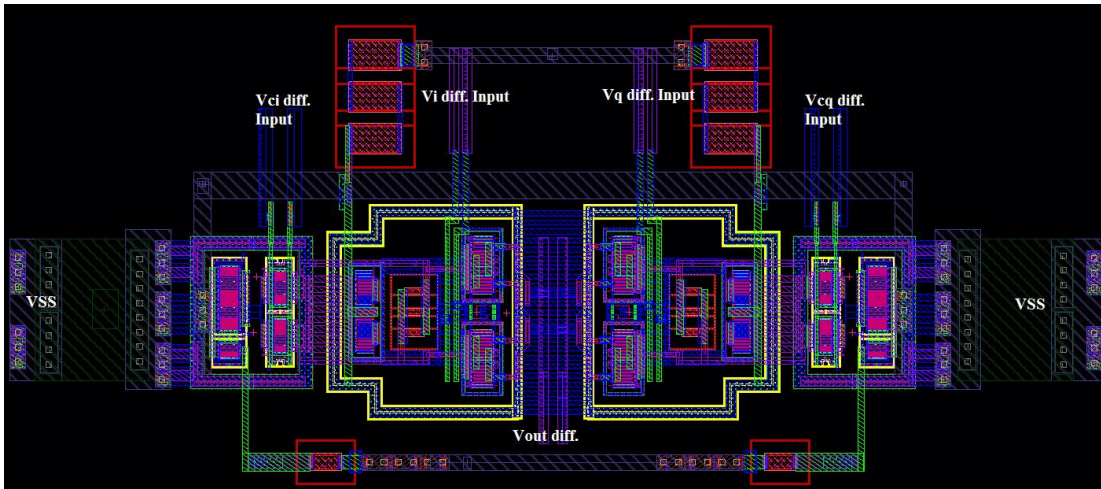

Figure 3.11: Final layout of the complex VGA core.

28

3.3 Output Stage

The ARF-PD system requires a maximum gain of 5 dB and has a maximum input power level of -3 dBm to the RFVM. For a differential system, the output swing voltage has to be 1.15 V_{pp} over a 100 ohm load for 2 dBm output power. The output voltage level at the complex VGA core is estimated to be 0.3 V_{pp}. At least 11.5 dB of gain is required from the output stage.

An amplifier was included in the output stage to compensate the insertion loss due to the PPF network, as well as to provide enough output power to drive a 50 Ω load. To provide enough gain, a 2-stage cascaded common source amplifier was implemented, and the schematic as illustrated in Fig. 3.12(a). Since the amplifier is driving a 100 Ω load, a large DC current has to flow through the transistors. Hence, large transistors are used to provide enough power to drive the load. As discussed in previous sections, large transistor size effectively reduces the circuit bandwidth; on-chip shunt inductive peaking was used to extend bandwidth of the 2-stage amplifier.
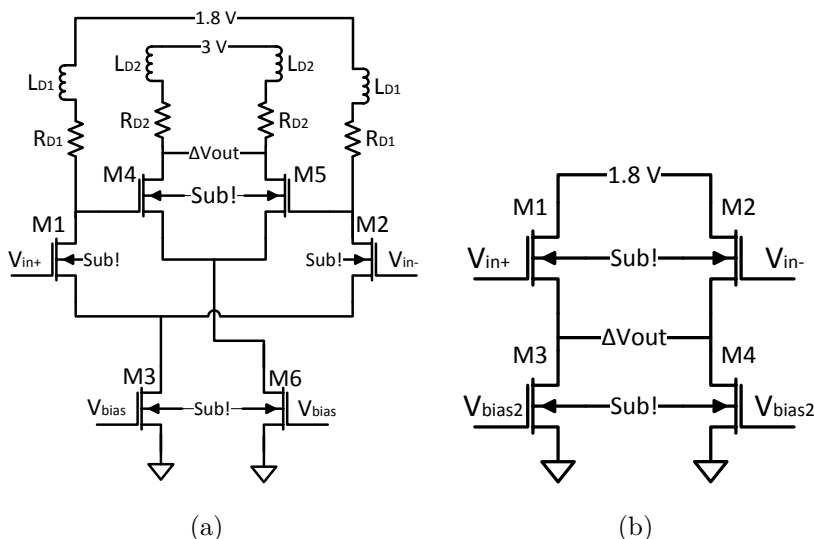
Figure 3.12: (a) 2-stage cascaded common-source amplifier and (b) common drain buffer.

To avoid having the complex VGA core directly driving the 2-stage amplifier(i.e., a large source impedance driving a huge capacitive load), a buffer stage as illustrated in Fig. 3.12(b) was implemented. The buffer stage provides a small capacitive load to the complex VGA core, and a small source impedance to drive the 2-stage amplifier, effectively

29

improving circuit's bandwidth. All output stage transistor sizes are summarized in Table 3.3.

Table 3.3: Design Parameters for the Output Stage

Device	Unit Length (nm)	Unit Width (μm)	Number of Fingers	Effective Width (μm)	Transistor Model
Amplifier					
M1	120	4	32	128	nfet
M2	120	4	32	128	nfet
M3	120	4	80	320	dgnfet
M4	120	4	66	264	dgnfet
M5	240	4	66	264	dgnfet
M6	240	4	200	800	dgnfet
Buffer					
M1	120	4	24	96	nfet
M2	120	4	24	96	nfet
M3	240	4	22	88	dgnfet
M4	240	4	22	88	dgnfet
Resistor	**Width (μm)**	**Length (μm)**	**Multiplier**	**Resistance (Ω)**	**Model**
R_{D1}	8.5	10	5	55.97	oprppres
R_{D2}	17	21	5	57.49	oprppres
Peaking Inductor	**Outer Dimension (μm)**	**Turns**	**Coil Width (μm)**	**Space (μm)**	**Effective Inductance (nH)**
L_{D1}	260	5	8.5	5	5.7
L_{D2}	270	5	8.5	5	6.2

3.4 Final Layout

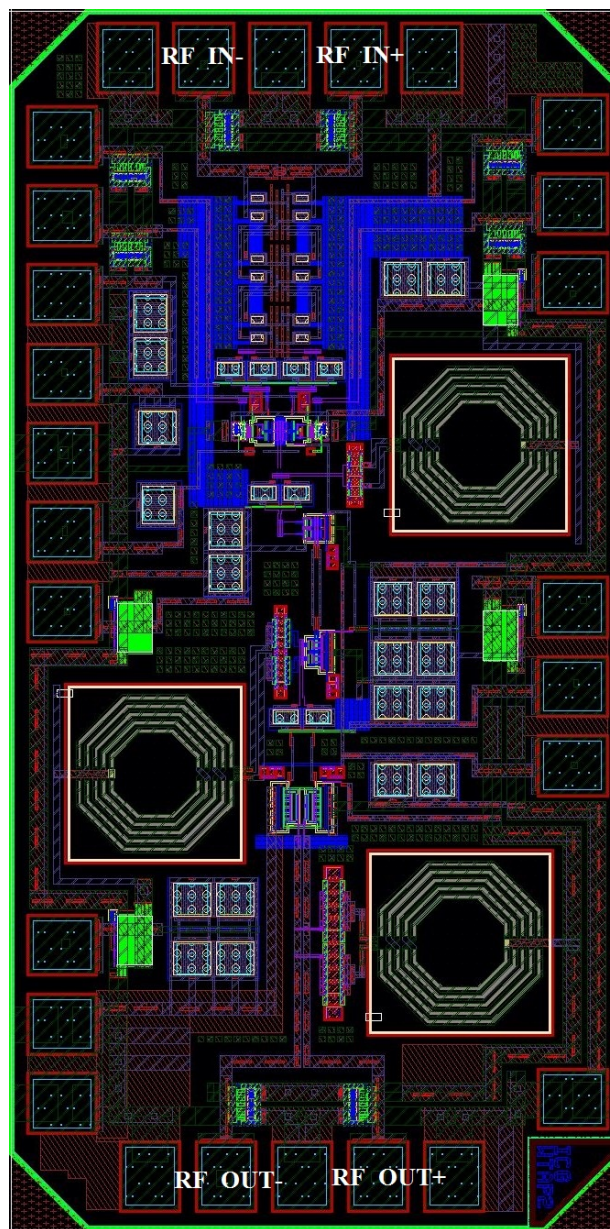The final layout occupying an area of 2x1 mm^2 is presented in the following:

Figure 3.13: Finaly layout of the first VM design.

3.5   Overall System Simulation Results

In this section, the overall performance of the RFVM was evaluated against similar components which are readily available on the market. Unfortunately, it was not possible to complete measurements of the fabricated chip at the time of writing of this thesis. Instead, post-layout simulation results have been used for the performance evaluation. The model of the simulated chip is a combination of the electromagnetic (EM) model and the RC extracted model. All of the active devices were modeled using the RC model extracted by Calibre, and the remainder of the chip, including interconnects, transmission lines, bond pads, and inductors, were EM simulated using Momentum. The printed circuit board (PCB) for testing was also EM simulated to further improve the simulation's accuracy. Each wire-bond connection from the IC to the PCB was assumed to be 0.5 nH. The ADS harmonic balance simulation testbench, as illustrated in Fig. 3.14 was setup to obtain the following parameters:
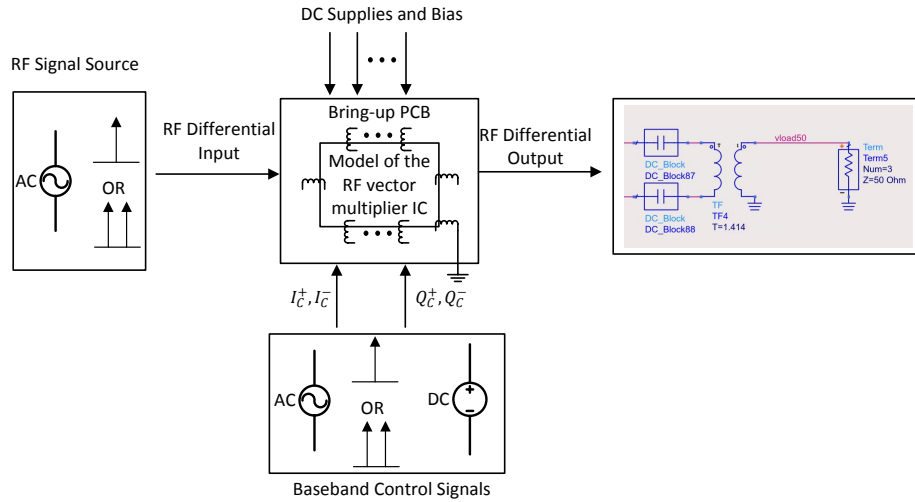
Figure 3.14: Block diagram for simulation testbench.

- Maximum gain: -3 dBm single-tone RF signal (maximum input power level) was fed to the input of the RFVM, while the baseband control signals $I_c = I_c^+ - I_c^-$ and $Q_c = Q_c^+ - Q_c^-$ were set to the maximum DC voltage of 1 V. The frequency was swept from 0.6 GHz to 2.8 GHz which was the targeted frequency range for this design. The simulated results are displayed below:
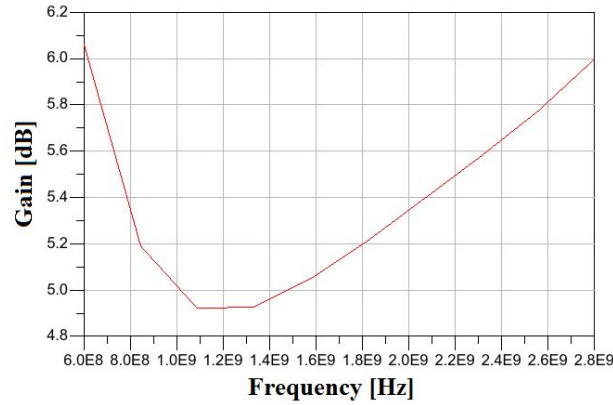
Figure 3.15: Maximum gain versus frequency.

It can be seen that the gain variation within the RF operation range is 1 dB, and the parabolic shape of the gain curve is due to gain roll-off and inductive peaking.

- Gain control range: -3 dBm single-tone RF signal (maximum input power level) was fed to the input of the RFVM, while the baseband control voltages V_{ci} and V_{cq} were swept from -1 V to 1 V. The simulated results for 1.8 GHz are illustrated below:
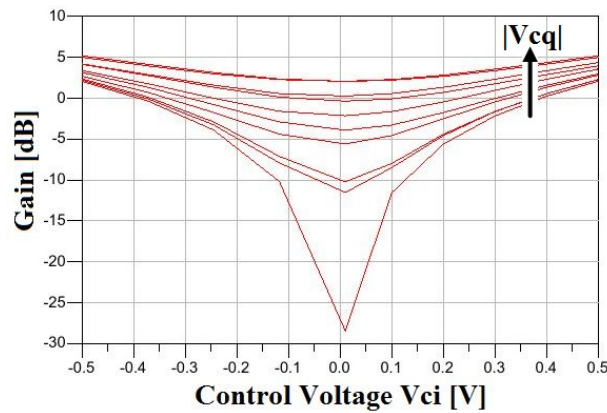

Figure 3.16: Plot of gain control range with various gain settings.

In the figure, each line represents one different voltage setting for Vcq and the x-axis represents the gain settings for Vci. The design is able to achieve a gain control range from 5 dB to -28 dB.

33

- Baseband control bandwidth: obtained by simply fixing the single-tone RF input power and frequency, while feeding a baseband single-tone signal through the desired frequency range. The highest output power of the mixing products changes as the baseband frequency increases. The baseband bandwidth can be determined when the change is more than 3 dB. The design is able to achieve a minimum baseband control bandwidth with less than 1 dB variation at 200 MHz with 2 GHz RF input frequency as illustrated in Fig. 3.17.
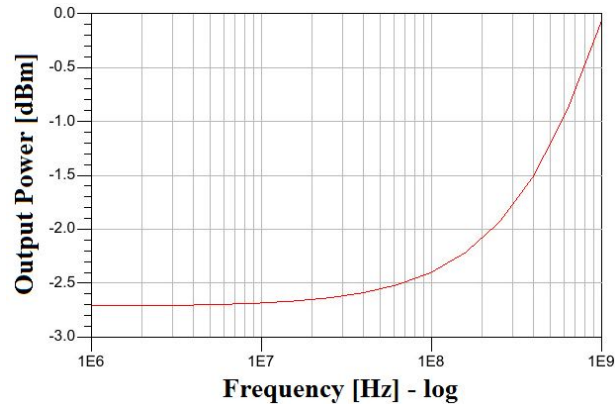

Figure 3.17: Plot of baseband bandwidth with -5 dBm RF input at 2 GHz.

- Output 1 dB compression (OP1dB) point at maximum gain: obtained by sweeping input power for a single-tone RF signal, while maintaining the maximum gain setting and setting both V_{ci} and V_{cq} at 1 V. the OP1dB can be obtained at the point where the maximum output power has dropped by 1 dB. Simulation results show that the RFVM has the worst OP1dB (6.4 dBm) at 0.6 GHz and the best OP1dB (9.6 dBm) at 2.8 GHz.

- Third-order intermodulation distortion (IMD3) versus different gain settings: measured by stimulating the RF input with a two-tone signal, where the peak voltage of the two-tone signal was equivalent to the peak voltage of a -3 dBm single-tone signal. The IMD3 with respect to the gain setting profile can be obtained by sweeping the baseband control voltages V_{ci} and V_{cq}. Fig. 3.18(a) shows the simulated IMD3 with all gain settings at 1.8 GHz, where good linearity is achieved, and with all the gain settings tat the maximum, where the worst linearity is seen (48 dBc). However, the linearity is not particularly good at lower frequencies either, as illustrated in Fig. 3.18(b); the worst IMD3 is 43 dBc at 0.6 GHz.
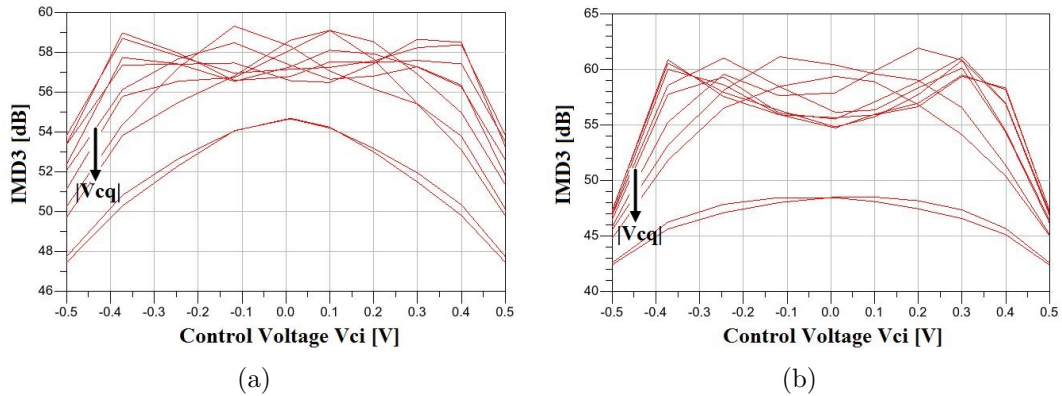
Figure 3.18: IMD3 plot versus gain setting at (a) 1.8 GHz and (b) 0.6 GHz.

- Output third-order intercept point (OIP3) at maximum gain: obtained using a similar setup as to IMD3 simulation, except OIP3 is measured at a low input power such that the system is operating in an extremely linear region where only third-order distortion is present in the system. Fig. 3.19 illustrates how the simulated OIP3 remains constant at low input power.
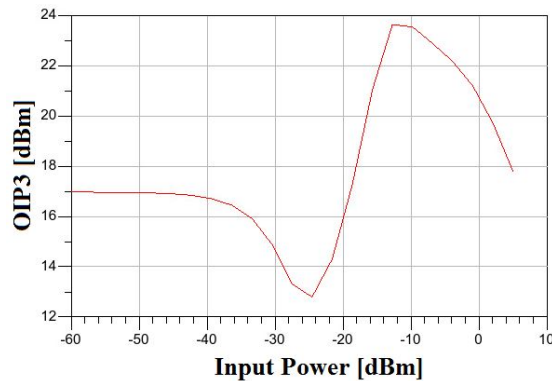
Figure 3.19: Plot of OIP3 versus input power at 2.8 GHz.

- Output noise power: obtained by stimulating the RF input with a 20 MHz noise signal centered around the RF frequency using the harmonic noise controller offered by ADS. The total noise output power can be obtained by integrating the output noise voltage as illustrated in Fig. 3.20, where Fig. 3.20(a) is the output noise voltage without RF stimulus, and Fig. 3.20(b) is the output spectrum with a -3 dBm single-

tone input at 2.8 GHz. The total noise output power over 20 MHz with and without applied RF stimulus are -141 dBm/Hz and - 138 dBm/Hz respectively.
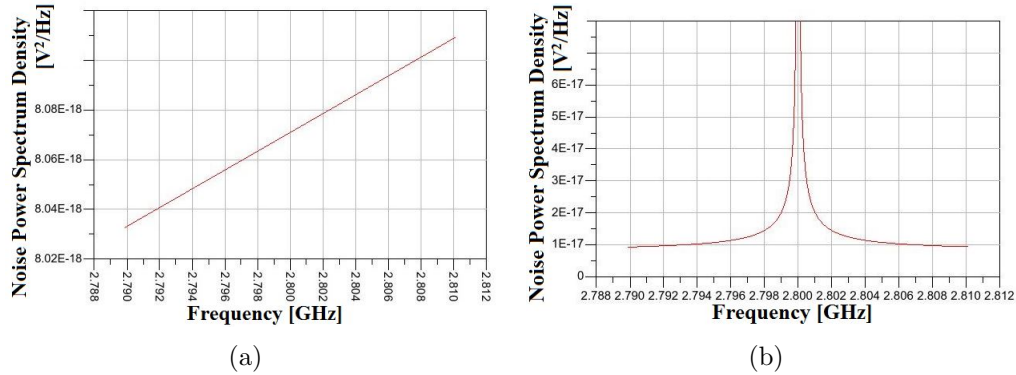
Figure 3.20: Output noise voltage (a) without RF stimulus and (b) with single tone RF stimulus.

- Gain and phase imbalances: obtained by measuring the amplitude and phase imbalances of the differential output signals. The simulated results are displayed in Fig. 3.21; the worst amplitude and phase imbalances are 0.8 dB and 3.7° respectively.

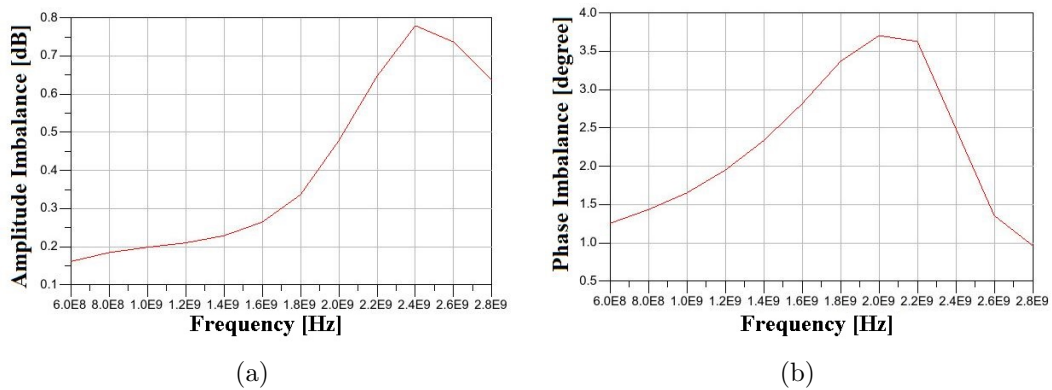Figure 3.21: Plots of (a) amplitude imbalance and (b) phase imbalance of the differential output signals.

36

The overall performance of the RFVM is summarized in Table 3.4 and compared with two off-the-shelf products. The MAX2045 has the best linearity, however it has a very narrow bandwidth and consumes the most power (more than twice that consumed by the proposed design). It is difficult to compare the ADL5390's performance to the other as it lacks of quadrature phase generation. However, it is a good design reference for this work. Overall, the proposed design shows the best RF and baseband bandwidth when compared to both ADL5390 and MAX2045, and also has the lowest power consumption at 300 mW. However, it has not met the power consumption requirement of 200mW and it has the worst output noise and linearity performances due to large insertion loss of the multi-stage PPF.

Table 3.4: Proposed and Commercial Vector Multiplier Performance Summary and Comparison

	Current Work	ADL5390[24]	MAX2045[25]
Maximum Gain (dB)	5 to -28	5 to 30	6.5 to -8.5
Phase Range	360°	360°	360°
RF Bandwidth (GHz)	0.6 - 2.8	0.04-2.4*	2.04-2.24
$3dB$ Control Bandwidth (MHz)	>400	200	230
Power Consumption (mW)	300	675	800
OP1dB (dBm)	6.4 (0.6 GHz)	11.5 (0.9 GHz)	-
	9.6 (2.8 GHz)	9.6 (2.4 GHz)	13.2 (2.1 GHz)
OIP3 (dBm)	17.4 (0.6 GHz)	23 (0.9 GHz)	-
	16.8 (2.8 GHz)	18.7 (2.4 GHz)	21.5 (2.1 GHz)
Output Noise (dBm/Hz)	-141	-148	-146.8

*RF operation bandwidth excluding quadrature phase generation.

3.6 Preliminary Measurement Results

In this section, preliminary measurement results are presented and compared to the simulated results. The measurement process involves two sets of measurement and each involves a different testbench setup. The first testbench setup is illustrated in Fig. 3.22, where the vector network analyzer (PNA-X) allows amplitude and phase measurement using differential single tone stimulus.
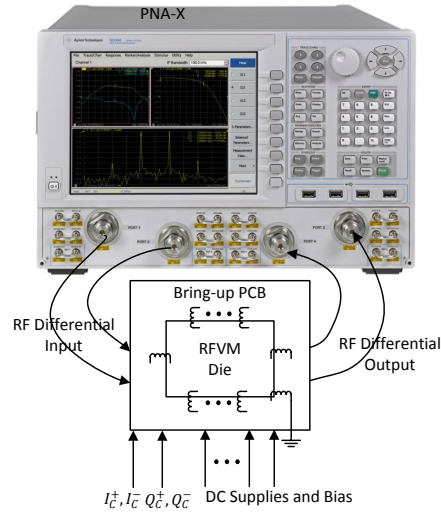
Figure 3.22: Block diagram for measurement testbench.

In Fig. 3.23, constellation plots are illustrated to show the performance of amplitude and phase control of the RFVM. It is obtained by sweeping V_{ci} and V_{cq}, and each step of constellation point is 0.2 V.

(a)   (b)
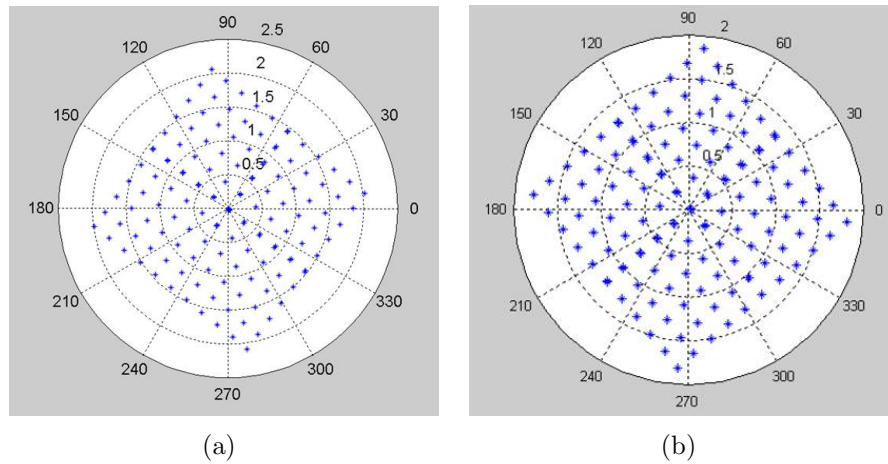
Figure 3.23: Measured constellation plots: (a) at 0.6 GHz and (b) 2.8 GHz at -3 dBm input power.

In Fig. 3.24, maximum gain of the RFVM is measured across the frequency range from

0.6 GHz to 2.8 GHz. Both baseband control voltages V_{ci} and V_{cq} are set to 1 V. The measured gain matches closely to the simulation results at lower frequency range. As frequency increases, the maximum gain starts to degrade, but still follows the trend in simulation. This degradation is expected, as extra parasitic capacitance of the ESD diode at the input is not accounted in the simulation and its effect on the overall gain of the RFVM becomes noticeable at the higher frequency band.


Figure 3.24: Measured maximum gain versus frequency.
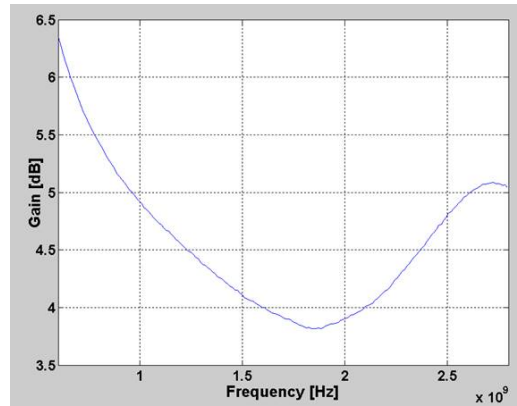
The second testbench setup is illustrated in Fig. 3.25, where the arbitrary waveform generator (M8190A) produces differential two-tone stimulus and spectrum analyzer (N9030A PXA) which captures the output spectrum of the RFVM. This setup is used to measurement the IMD3 performance of the RFVM using two-tone stimulus.
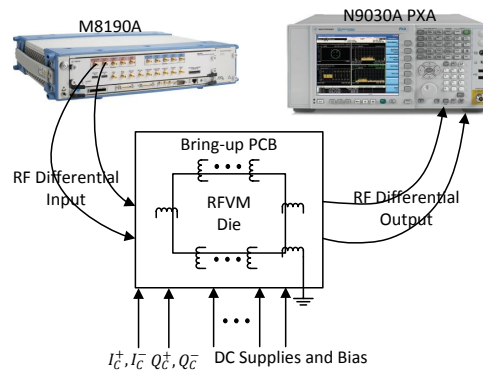
Figure 3.25: Block diagram for measurement testbench.

39

The measurement results and simulation results of the first RFVM design are summarized in Table 3.5. The measured performance of the RFVM matches closesly to the simulated performance in terms of RF bandwidth, overall gain, and linearity performance. Further measurement is still required to obtain baseband bandwidth and linearity with continuous wave signals.

Table 3.5: Measurement and Simulation Performance Summary and Comparison

	Measurements	Simulation
Maximum Gain (dB)	6.4 - 3.8	6.1 - 4.9
Phase Range	360°	360°
RF Bandwidth (GHz)	0.6 - 2.8	0.6 - 2.8
Power Consumption (mW)	300	300
IMD3 (dBc)(@ -3 dBm P_{in})	41.4 (@ 0.6 GHz)	43 (@ 0.6 GHz)
	49.5 (@ 1.8 GHz)	48 (@ 1.8 GHz)
	49.6 (@ 2.5 GHz)	49 (@ 2.8 GHz)

Chapter 4

Proposed Design 2

Many design challenges and limitations were revealed in Chapter 3. A few outstanding issues are listed as following:

- ARF-PD systems require high output power levels, thus consuming more power.

- The bandwidth is limited by the PPF network.

- The PPF leads to significant insertion loss and added noise.

- Linearity is limited at lower frequency bands due to excessive inductive peaking.

Hence, an alternative ARF-PD system configuration is adopted as illustrated in Fig. 4.1 to mitigate these issues. The system is targeted to have an overall operating bandwidth from 0.6 GHz to 6 GHz for sub-6 GHz application(Phase-1 of future 5G wireless communication). The input signal $x(t)$ has a maximum power level of 0 dBm, which is split into two paths by a 3 dB coupler. Each path has a -3.5 dB power level, assuming the coupler has 0.5 dB loss. To achieve a better noise performance and wider bandwidth, a design decision was made to replace the integrated PPF network with off-chip banded 90° 3 dB hybrid couplers to generate quadrature signals. In Fig. 4.1, the blue box includes all of the RFVM system's function blocks would be implemented in an IC using IBM CMOS 130nm technology. The RFVM system generates a correction signal (inverse of the error component in a distorted signal) which then combines with the modulated RF signal in the upper path through a 10 dB coupler to create a pre-distorted signal. The tunable delay line in the upper path matches the signal delay on the lower path. The correction signal only needs to match the power level of the error component, meaning that it has

a much lower power level than the modulated RF signal. System level simulation with analog pre-distortion algorithm indicated that the maximum required output power of the correction signal is at least 10 dB lower than the RF signal. Hence, a decision was made to use a -10 dB coupler and -3 dbm output power for the RFVM design.
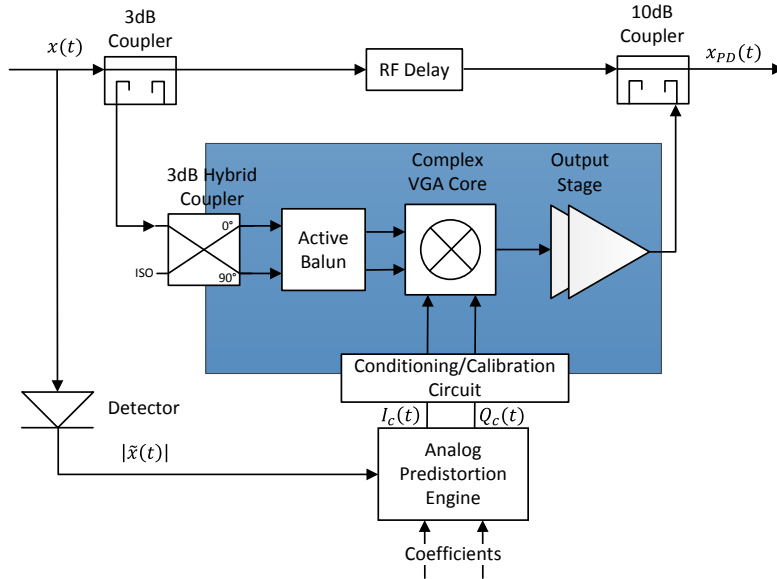
Figure 4.1: Block diagram of the second APD-RF system configuration.

4.1 Wideband Active Balun and Input Matching

4.1.1 Active Balun Topology

A 90° hybrid coupler has a single-ended input and a single-ended output configuration, but the complex VGA core requires differential input signals. A single to differential converter has been included in the system. In addition, since the off-chip hybrid coupler requires 50 ohm loading for proper operation, input matching becomes necessary part of the RFVM system. Single to differential conversion can be achieved by using either a passive balun transformer or an active balun topology. If using an off-chip passive balun, similar to the off-chip hybrid coupler, multiple banded passive baluns would be required to cover the entire RF bandwidth, and additional input buffer circuitry would have to be added before

the complex VGA core for input matching. A wideband active balun circuit can achieve both single to differential conversion and input matching with similar power consumption to input buffer circuitry. Given the options, the wideband active balun topology was selected for this design.
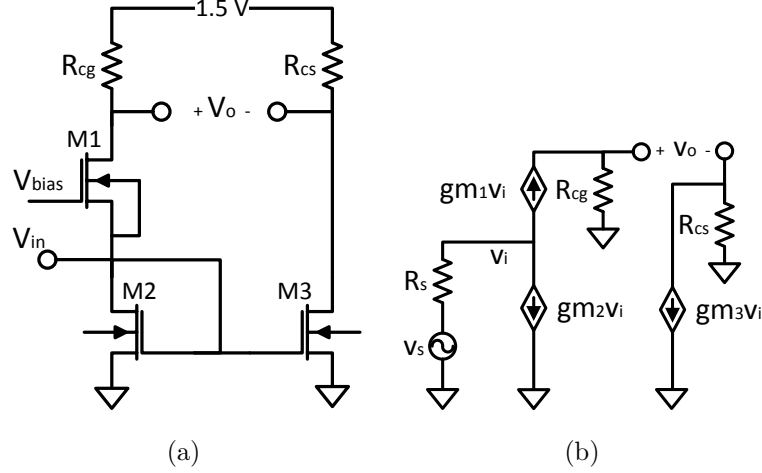

Figure 4.2: (a) Common-gate and common-source active balun topology and (b) equivalent small signal model.

Common-gate (CG) common-source (CS) topology has been intensively used in balun based low noise amplifier (LNA) circuit receiver designs[33]. For this RFVM system,CG-CS topology is only implemented with very small gain due to limited input dynamic range of the complex VGA core. Though the small gain of the active balun circuit would increase the noise figure of the overall system, its impact on the noise floor is insignificant. A quick harmonic balance noise simulation shows an 0.5 dBm/Hz increase in the noise floor. This topology uses a CG in parallel with a CS as illustrated in Fig. 4.2(a). Its small signal equivalent circuit is illustrated in Fig. 4.2(b). The transfer functions of CG and CS are described in equations (4.1) and (4.2) respectively. The input impedance can be estimated by equation (4.3).

$$\frac{V_{o+}}{V_i} = gm_1 R_{cg} \tag{4.1}$$

$$\frac{V_{o-}}{V_i} = -gm_3 R_{cs} \tag{4.2}$$

43

$$R_{in} = \frac{1}{gm_1} \bigg\| \frac{1}{gm_2} \qquad (4.3)$$

4.1.2 Active Balun Desgin

To achieve a balanced differential output signal, both transfer functions (4.1) and (4.2) have to be of equal magnitude. Since both the CG and CS are driving similar capacitive loads, to ensure matching loading conditions, R_{cg} and R_{cs} are designed to be of equal value. Hence, for a balanced differential output signal gm_1 and gm_3 must be equal. This can be achieved by designing both transistors $M1$ and $M2$ with a similar size and bias condition. Transistors $M2$ and $M3$ form a current mirror to ensure the same DC biasing condition between $M1$ and $M3$, where $gm_1 = gm_2 = gm3$. For a 50 ohm input impedance, gm_1 and gm_2 can be calculated using equation (4.3). Loading resistances R_{cg} and R_{cs} can be then calculated based on the desired gain.
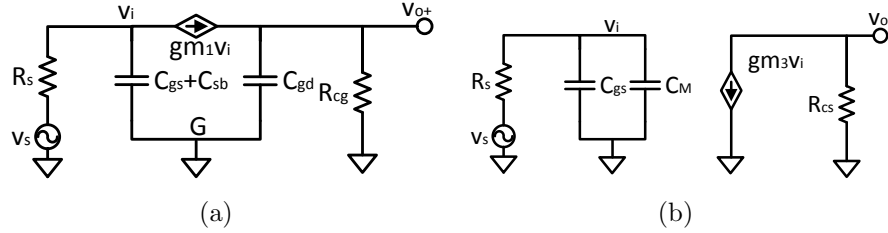

Figure 4.3: Small signal model of active balun topology: (a) common-gate and (b) common-source.

Realistically, the parasitic capacitance should be included in the small signal model as illustrated in Fig. 4.3, where C_M is the Miller capacitance. The dominant pole frequency ω_{1-cg} of the CG circuit can be estimated by $\frac{gm}{c_{gs}}$; the dominant pole frequency ω_{1-cs} of the CS circuit can be estimated by $\frac{1}{R_s C_M}$. Based on the equations above, the CG has a higher pole frequency than CS and, hence, better gain bandwidth. As the frequency increases, the magnitude and phase imbalance between the output nodes v_{o+} and v_{o-} increase. In addition, the ground inductance increases source degeneration for CS circuit operating at high frequency, which further increases the magnitude of the output imbalance.

The final schematic of the active balun is illustrated in Fig. 4.2(a) and all component sizes are summarized in Table 4.1. A triple-well NMOS is used for CG transistor. This allows a bulk to source connection to reduce body-effects and improving linearity.

Table 4.1: Design Parameters for the Active Balun

Transistors	Unit Length (nm)	Unit Width (μm)	Number of Fingers	Effective Width (μm)	Transistor Model
M1	120	2	11	22	nfettw
M2	120	2	12	24	nfet
M3	120	2	11	22	nfet
Resistor	Width (μm)	Length (μm)	Multiplier	Resistance (Ω)	Component Model
R_{cg}	2	2.42	5	65.19	oprppres
R_{cs}	2	2.42	5	65.19	oprppres

4.2 Self-linearized VGA Core

Although a DBGM has better linearity and wider input dynamic range than a SBGM, the linearity of the VGA core still needs to be improved. This is particularly important since the PPF network has been replaced with a hybrid coupler. The maximum input voltage swing has been effectively increased and will drive the VGA core deeper into the non-linear region.

4.2.1 Linearization of Source-coupled Differential Pair

As discussed in the previous chapter, the linear input range of a source-coupled differential pair is limited by the tail current due to the square law nature of the CMOS transistor. In 1985, Babanezhad introduced a technique for linearizing source-coupled differential pairs[34]. A squaring circuit is proposed to generate a squared product of the input signal as illustrated in Fig 4.4(a). The output current of the squaring circuit is then injected into the tail current source of the source-coupled differential pair through series of current mirrors. By injecting the V_1^2 term into tail current as illustrated in Fig. 4.4(b), the output current can be linearized as demonstrated in equation (4.4).

$$
\begin{aligned}
\Delta I_{out} &= \frac{k'}{2}\frac{W_1}{L_1}V_1\sqrt{\frac{2(I_{ss}+\frac{1}{2}kV_1^2)}{k}-V_1^2} \\
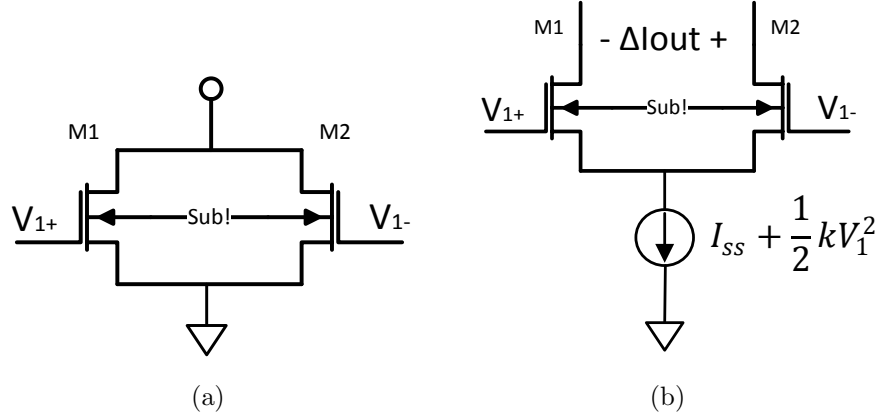&= V_1\sqrt{2k\frac{W_1}{L_1}I_{ss}}
\end{aligned}
\tag{4.4}
$$

45

Figure 4.4: (a) Squaring circuit and (b) linearized source-coupled differential pair.

4.2.2 Proposed Topology - DC and AC Analysis

The above technique can be adapted to improve the linearity of the complex VGA core. The proposed circuit presented in [34] is too complex for wideband RF applications. A simplified circuit is proposed and illustrated in Fig. 4.5, where the tail current source is directly replaced by the squaring circuit to avoid using PMOS current mirrors that are slow-speed and band-limiting. This approach also reduces the group delay difference between the differential pair and injection circuit inputs.
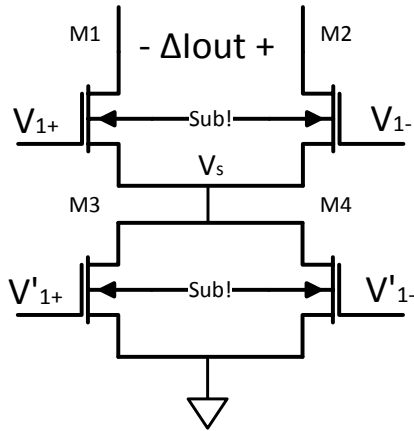

Figure 4.5: The proposed modified linearized source-coupled differential pair.

The differential pair $M1$ and $M2$, and the squaring circuit $M3$ and $M4$, have the same

46

RF input signal, but a different DC bias level. The input differential signals of the squaring circuit are defined as follows: $V'_{1+} = V_{cm} + \frac{V_1}{2}$ and $V'_{1-} = V_{cm} - \frac{V_1}{2}$. The common mode voltage V_{cm} determines the biasing current, I_{ss}, for the differential pair as illustrated in equation (4.5).

$$
\begin{aligned}
I_{ss} &= I_{d3} + I_{d4} \\
&= \frac{k'}{2}(\frac{W}{L})_3(V_{cm} + \frac{V_1}{2} - v_t)^2 + \frac{k'}{2}(\frac{W}{L})_4)(V_{cm} - \frac{V_1}{2} - v_t)^2 \\
&= \frac{k'}{2}(\frac{W}{L})_3(V_{cm} - v_t)^2 + \frac{k'}{4}(V_1)^2
\end{aligned}
\tag{4.5}
$$

where the first component is the biasing current I_{ss} and the second component is the cancellation current to be injected into the differential pair.

DC simulation results show a significant improvement in input dynamic range of the linearized source-coupled differential pair as illustrated in Fig. 4.6. In Fig. 4.6(a), dashed lines represent the DC-IV curves of a source coupled differential pair, and solid lines presents the DC-IV curves of a linearized source-coupled differential pair. Fig. 4.6(b) shows the DC trans-conductance of the two topologies. Both figures show that the proposed topology has improved linearity and input dynamic range.
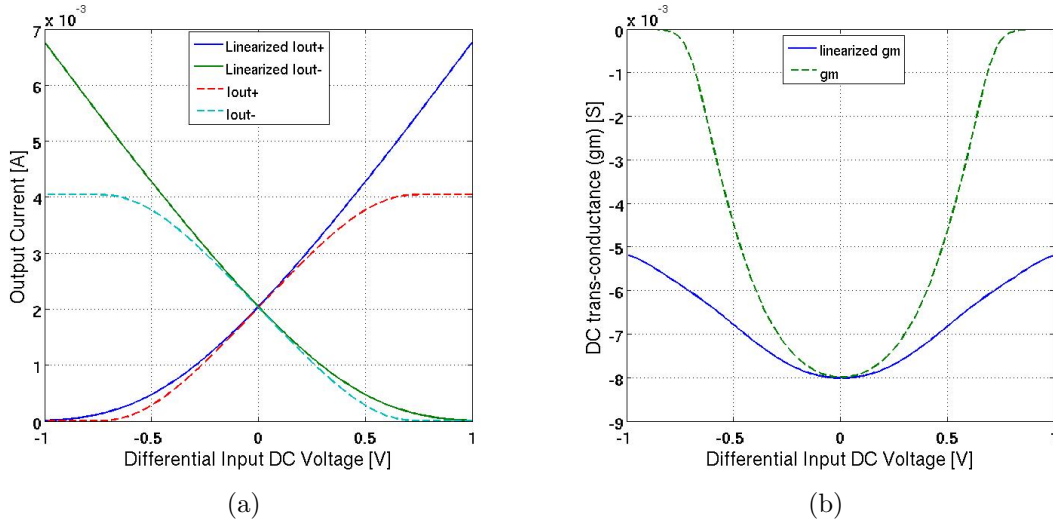

Figure 4.6: Source-coupled differential pairs: (a) DC-IV plots and (b) DC trans-conductance gm plots.

To further study the squaring injection circuit and its effect on linearity, the circuit

behavior was also analyzed with an AC signal input. If the input signal is a single-tone continuous wave, $v_i = cos(\omega_1 t)$, and the transistors' non-linearity creates amplified copies of the input signal at the fundamental frequency and its harmonics as illustrated in equation (4.6), then gm_1 is the gain at the fundamental frequency, gm_2 is the gain at second harmonic frequency, gm_3 is the gain at third harmonic frequency, and so on.

$$i_{ds} = gm_1 cos(\omega_1 t) + gm_2 cos(2\omega_1 t) + gm_3 cos(3\omega_1 t) + ... \tag{4.6}$$

The current flowing through transistors $M3$ and $M4$ is summed at the source node v_s of the differential pair, which can be defined by equation (4.7).

$$
\begin{aligned}
i_{ds}(v_i)_{M3} &= gm_{1,M3}(v_i) + gm_{2,M3}(v_i)^2 + gm_{3,M3}(v_i)^3 + .. \\
i_{ds}(-v_i)_{M4} &= gm_{1,M4}(-v_i) + gm_{2,M4}(-v_i)^2 + gm_{3,M4}(-v_i)^3 + .. \\
i_{v_s} &= i_{ds}(v_i)_{M3} + i_{ds}(-v_i)_{M4} \\
&= 2gm_{2,M3}(v_i)^2
\end{aligned}
\tag{4.7}
$$

The voltage swing on node v_s can be estimated by multiplying the current i_{v_s} by the impedance looking into the source of the differential pair $M1$ and $M2$. The impedance looking into node v_s can be estimated as $\frac{1}{gm_{1,M1}}||\frac{1}{gm_{1,M2}}$. Hence, the voltage v_s can be solved to be:

$$
\begin{aligned}
v_s &= (2gm_{2,M3}v_i^2)(\frac{1}{2gm_{1,M1}}) \\
&= \frac{gm_{2,M3}v_i^2}{gm_{1,M1}}
\end{aligned}
\tag{4.8}
$$

The current flowing through transistor $M1$ is expressed in equation (4.9).

$$
\begin{aligned}
i_{ds}(v_i)_{M1} &= gm_{1,M1}(v_i - v_s) + gm_{2,M1}(v_i - v_s)^2 + gm_{3,M1}(v_i - v_s)^3 + .. \\
&= gm_{1,M1}(v_i - v_s) + gm_{2,M1}(v_i^2 - 2v_iv_s + v_s^2) + gm_{3,M1}(v_i^3 - 3v_i^2v_s + 3v_iv_s^2 - v_s^3)
\end{aligned}
\tag{4.9}
$$

by substituting v_s into equation (4.10):

$$
\begin{aligned}
i_{ds}(v_i)_{M1} &= gm_{1,M1}\left(v_i - \frac{gm_{2,M3}v_i^2}{gm_{1,M1}}\right) + gm_{2,M1}\left(v_i^2 - \frac{2gm_{2,M3}v_i^3}{gm_{1,M1}} + (\frac{gm_{2,M3}v_i^2}{gm_{1,M1}})^2\right) \\
&+ gm_{3,M1}\left(v_i^3 - \frac{3gm_{2,M3}v_i^4}{gm_{1,M1}} + 3v_i(\frac{gm_{2,M3}v_i^2}{gm_{1,M1}})^2 - (\frac{gm_{2,M3}v_i^2}{gm_{1,M1}})^3\right)
\end{aligned}
\tag{4.10}
$$

A further simplification can be made if equation (4.10) only keeps terms up to the third harmonic:

$$i_{ds}(v_i)_{M1} = (gm_{1,M1})v_i + \big(gm_{2,M1} - gm_{2,M3}\big)v_i^2 + \Big(gm_{3,M1} - \frac{2gm_{2,M1}gm_{2,M3}}{gm_{1,M1}}\Big)v_i^3 \quad (4.11)$$

where the fundamental component remains as $gm_{1,M1}$, and both the second harmonic and third harmonic components are changed by the current injected at node v_s. Since the second harmonic component can be eliminated by the differential output, the squaring circuit can be designed to eliminate the third harmonic component, when the condition is set to be $gm_{2,M3} = \frac{gm_{3,M1}gm_{1,M1}}{2gm_{2,M1}}$.

As discussed in the previous chapter, IMD3 is used to evaluate the linearity of the circuit. Since third order harmonic product is a major contributor to IMD3, the proposed circuit that has been proved to be able to reduce third order harmonic product, can also reduce the IMD3 component. A similar mathematical derivation can be approached by replacing v_i with a two-tone signal $cos(\omega_1 t) + cos(\omega_2 t)$.

An AC simulation was performed to evaluate the proposed circuit against standard source-coupled differential pairs. The results showed an improvement of more than $20dB$ in IMD3 as illustrated in Fig. 4.7.

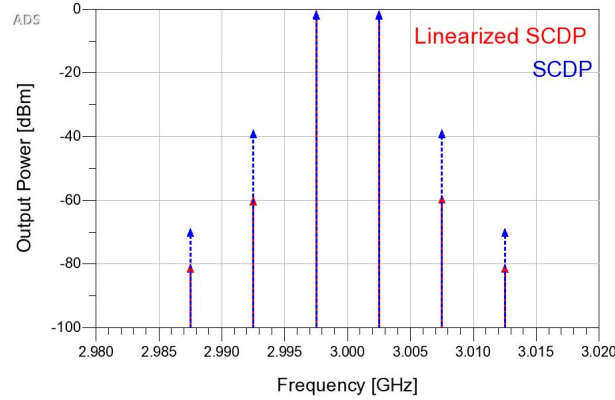Figure 4.7: AC simulation using two-tone signal with $5MHz$ spacing, center frequency at 3 GHz.

Although both the DC and AC simulations have shown promising results, they have also revealed the approach's limitations. In Fig. 4.6(a), as the input voltage increases, the total DC current start to increase due to the unbalanced operation of $M3$ and $M4$. A higher

49

common mode voltage V_{cm} is needed to reduce this effect, which leads to increased supply voltage. In addition, with the current injection at the source node of the differential pair $M1$ and $M2$, a voltage swing is presented at the node calculated by equation (4.8). With a large input signal introducing a large swing at the node v_s, the supply voltage needs to be further increased to ensure transistors are operating in the saturation region. However, this problem can be ignored since the input signal swing is limited in this application.

4.2.3 Proposed Complex VGA Core Design

The proposed linearization technique can be extended to DBGM by replacing the bottom current sources with squaring circuits as illustrated in Fig. 4.8. Although the modified folded DBGM shows improved linearity as shown in Fig. 4.9, the linearization ability is fairly limited compared to the IMD3 improvement demonstrated for the source-coupled differential pair. The limitation is due to current steering in the RF differential pair. As the baseband input controls V_{c+} and V_{c-} change, the total bias current for the RF differential pairs also varies, effectively changing the linearization condition.
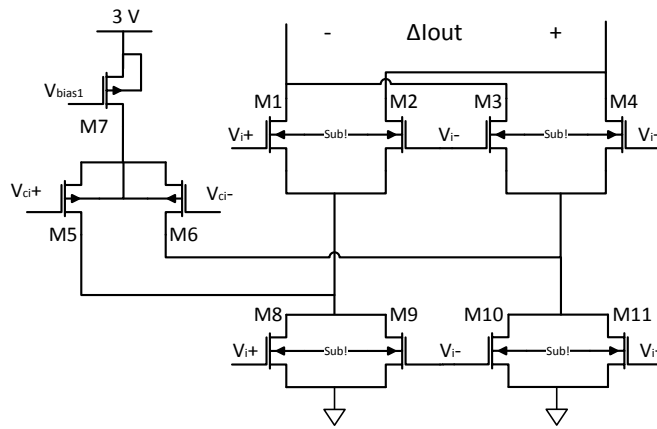
Figure 4.8: Modified folded double balance Gilbert multiplier.

Furthermore, to investigate the impact of devices mismatch on the linearization performance, 10% device mismatch is used between the first squaring circuit($M8$ and $M9$) and the second squaring circuit($M10$ and $M11$). The two devices in each squaring circuit are presumed to be identical due to their layout configuration. In Fig. 4.9, the blue dotted curve shows the linearity degradation at the low gain regions. However, this degradation on low gain region will have minimal impact on the overall performance of the RFVM.

The sizes for bottom transistors $M8$ to $M11$ have to be swept in simulation to determine a compromise condition that gives improved IMD3 over the desired range of gain settings.
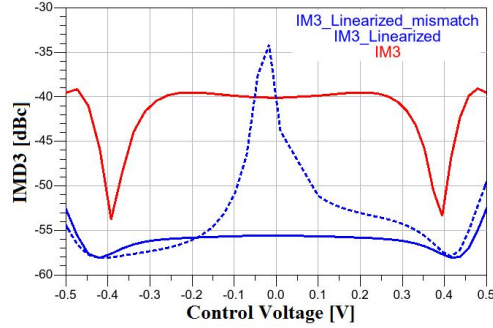
Figure 4.9: IM3 simulation for standard FDBGM and modified FDBGM using harmonic balance simulation with $0.4V_{pp}$ input voltage swing.

Since the squaring circuits also require input signals, this will inevitably increase the input parasitic capacitance and decrease the RF bandwidth. For $M8$ to $M11$, $120nm$ transistors were used to reduce parasitic capacitance while still providing the required current injection. The final schematic design of the proposed complex VGA core is illustrated in Fig. 4.10 with all transistor sizes summarized in Table 4.2.

Figure 4.10: Final schematic of the proposed complex VGA core.

Table 4.2: Design Parameters of The Proposed Complex VGA Core

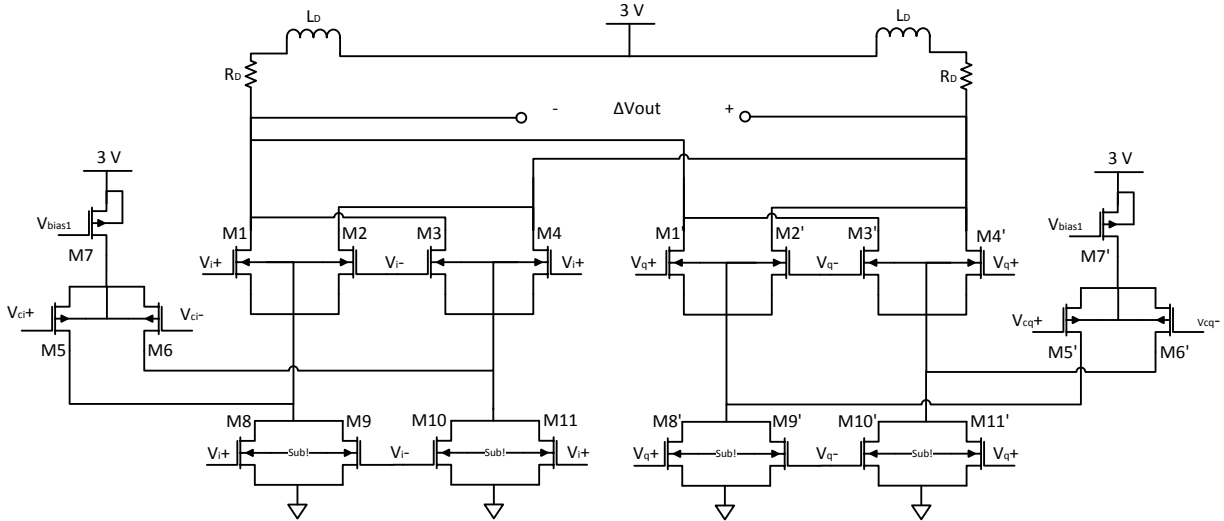| Transistors	Unit Length (nm)	Unit Width (μm)	Number of Fingers	Effective Width (μm)	Transistor Model
M1($M1'$)	240	3.25	8	26	dgnfettw
M2($M2'$)	240	3.25	8	26	dgnfettw
M3($M3'$)	240	3.25	8	26	dgnfettw
M4($M4'$)	240	3.25	8	26	dgnfettw
M5($M5'$)	240	2.5	10	25	dgpfet
M6($M6'$)	240	2.5	10	25	dgpfet
M7($M7'$)	240	4	20	80	dgpfet
M8($M8'$)	120	1.75	8	14	nfet
M9($M9'$)	120	1.75	8	14	nfet
M10($M10'$)	120	1.75	8	14	nfet
M11($M11'$)	120	1.75	8	14	nfet
Resistor	**Width**	**Length**	**Multiplier**	**Resistance**	**Model**
R_D	2 μm	7 μm	5	171.25 Ω	oprppres
Peaking Inductor	**Outer Dimension**	**Turns**	**Coil Width**	**Space**	**Effective Inductance**
L_D	240 μm	5	8.5 μm	5 μm	4.9 nH

The final layout of the proposed complex VGA core are shown in Fig. 4.11, which follows a common-centroid layout structure.
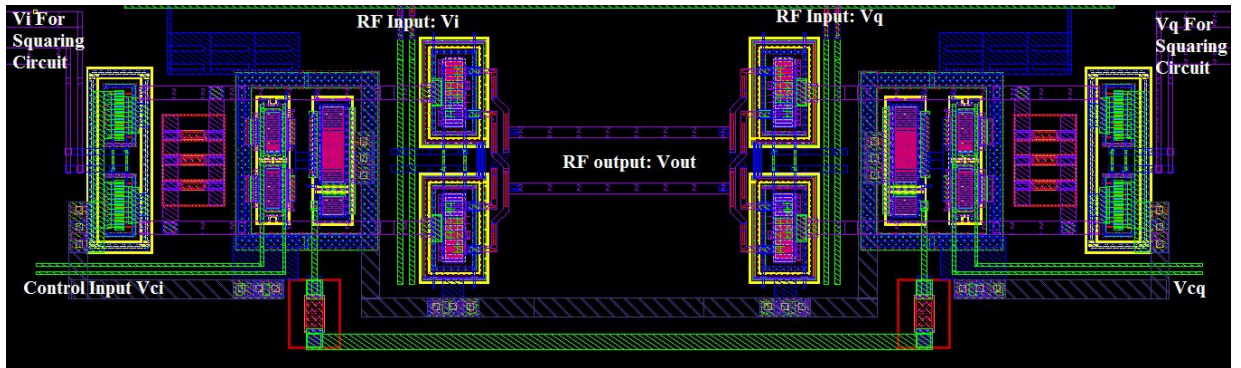

Figure 4.11: Final layout of the proposed complex VGA core.

4.3 Output Stage

The new ARF-PD system configuration requires a maximum of -3 dBm output power level from the RFVM system. For a differential system, the output swing voltage has to be 0.633 V_{pp} over a 100 ohm load for -3 dBm output power. The output voltage level at the complex VGA core is estimated to be 0.3 V_{pp}. hence, the output stage requires at least $7dB$ of gain. Since low output power is required, the sizing and power consumption of the 2-stage amplifier can reduced significantly compared to the design in Chapter 3. In addition, the buffer stage proceeding the amplifier stage is no longer required, the lower input capacitance of the 2-stage amplifier providing a high enough loading impedance. Instead, a buffer is added after the amplifier stage to drive a 100ohm load, which is estimated to have a 5dB loss. The total output stage achieved $8dB$ of gain and 80 mw of power consumption.

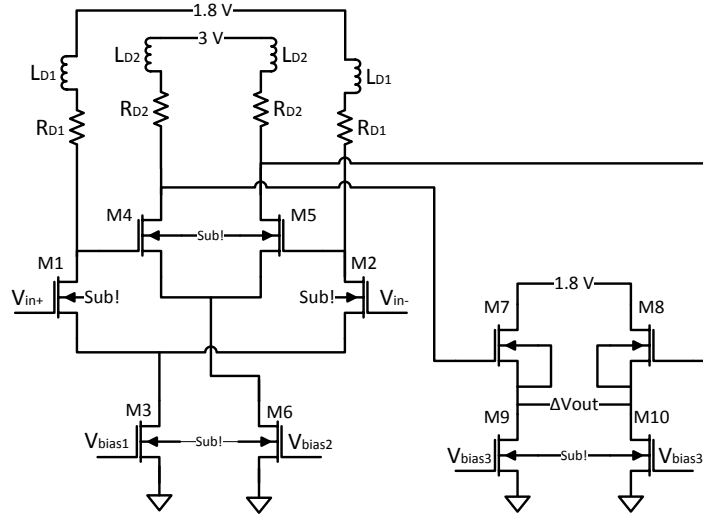The circuit implementation is illustrated in Fig: 4.12, with all transistor sizes summarized in Table 4.3.

Figure 4.12: Final schematic for the output stage: two-stage amplifier and buffer.

Table 4.3: Design Parameters for the Output Stage

Transistors	Unit Length (nm)	Unit Width (μm)	Number of Fingers	Effective Width (μm)	Transistor Model
Amplifier					
M1	120	2	10	20	nfet
M2	120	2	10	20	nfet
M3	120	3	24	72	nfet
M4	240	3	14	42	dgnfet
M5	240	3	14	42	dgnfet
M6	120	3	32	96	nfet
Buffer					
M7	120	4	18	72	nfettw
M8	120	4	18	72	nfettw
M9	120	2	25	50	nfet
M10	120	2	25	50	nfet
Resistor	**Width (μm)**	**Length (μm)**	**Multiplier**	**Resistance (Ω)**	**Model**
R_{D1}	2	7	5	171.25	oprppres
R_{D2}	3	12	5	190.37	oprppres
Peaking Inductor	**Outer Dimension (μm)**	**Turns**	**Coil Width (μm)**	**Space (μm)**	**Effective Inductance (nH)**
L_{D1}	225	5	8.5	5	4.3
L_{D2}	235	5	8.5	5	4.7

4.4 Final Layout

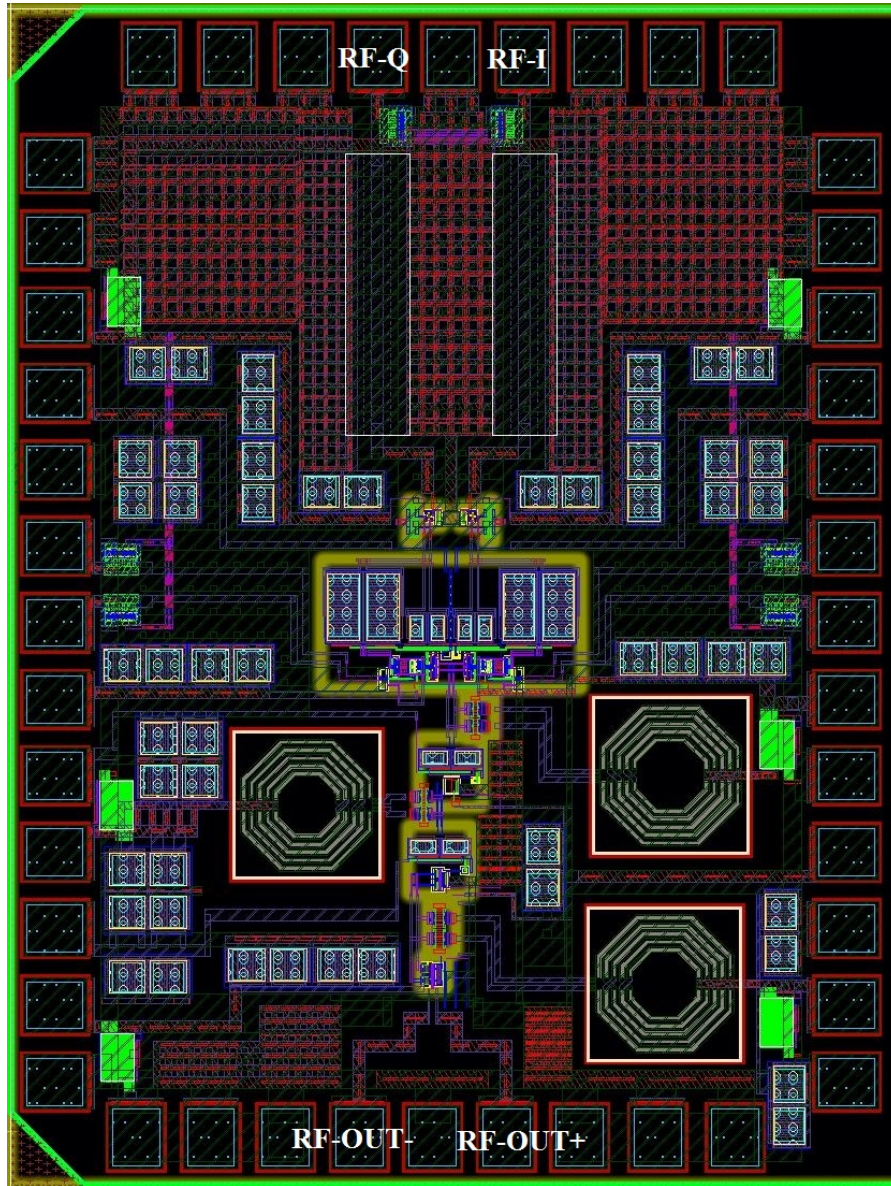The final layout occupying an area of 2x1.5 mm^2 is presented in the following:

Figure 4.13: Finaly layout of the first VM design.

4.5 Overall System Simulation Results

In this section, the overall performance of the second design is evaluated against the first design. At the time of writing this thesis, the chip has not yet been returned from fabrication, hence, only post-layout simulation results are available. Again the model of the simulated chip was a combination of the EM and the RC extracted models. All of the active devices were modeled using the RC model extracted by Calibre, and the remainder of the chip including the interconnects, transmission lines, bond pads, and inductors were EM simulated using Momentum. Each wire-bond connection on the chip bond pad was assumed to be $0.5nH$. ADS harmonic balance simulation illustrated in Fig. 4.14 was performed as explained in Chapter 3.4. In this test bench, an off-chip 90° hybrid coupler was used for quadrature phase generation. The coupler parameters were set to match the typical performance of hybrid couplers offered by the Anaren Xinger brand: 0.5 dB loss, 0.5 dB gain imbalance, and 3° phase imbalance.
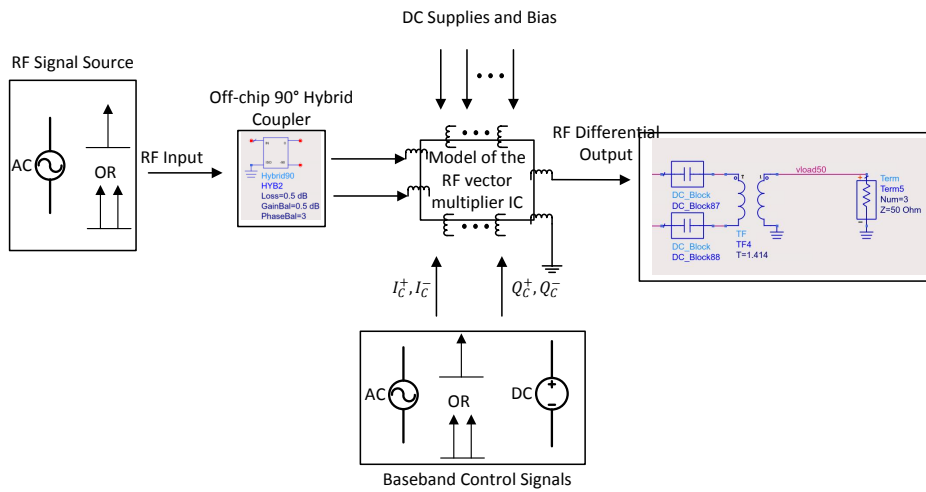
Figure 4.14: Block diagram for simulation testbench.

- Maximum gain: $-3.5dBm$ single-tone RF signal (maximum input power level) was fed to the input of the RFVM, while the baseband control signals $I_c = I_c^+ - I_c^-$ and $Q_c = Q_c^+ - Q_c^-$ were set to the maximum DC voltage of 1.2 V. The frequency were swept from 0.6 GHz to 6 GHz, the targeted frequency range for this design. The simulated results are displayed below:
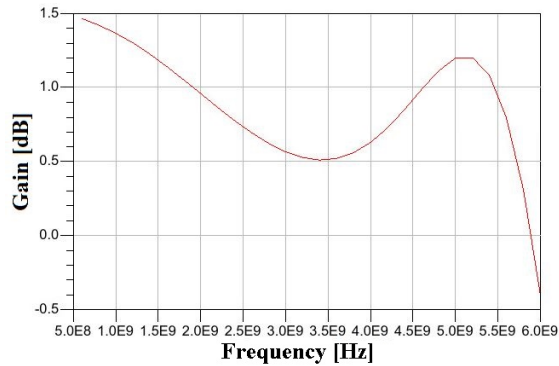
56

Figure 4.15: Maximum gain versus frequency.

It can be seen that the gain variation within the RF operation range is less than 2 dB. The gain peaking at 5 GHz is due to inductive peaking and is expected to be reduced in measurements.

- Gain control range: $-3.5dBm$ single-tone RF signal (maximum input power level) was fed to the input of the RFVM, while the baseband control voltages V_{ci} and V_{cq} were swept from -1.2 V to 1.2 V at various RF frequency points. The simulated results for 3 GHz are illustrated below:
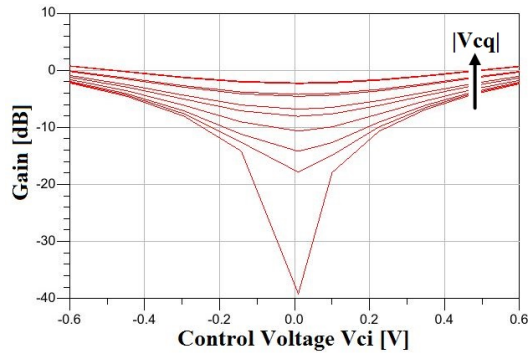

Figure 4.16: Plot of gain control range with various gain settings.

In the plot, each line indicates one different voltage setting for Vcq and the x-axis represents the gain settings for Vci. The design is able to achieve a gain control range of 0 dB to -38 dB.

- Baseband control bandwidth: The design is able to achieve a minimum baseband

57

control bandwidth with less than 1 dB variation at 200 MHz with 6 GHz RF input frequency as illustrated in Fig. 4.17. The gain variation is the worst around the RF frequency of 6 GHz as shown in Fig. 4.15, resulting in the worst baseband bandwidth.
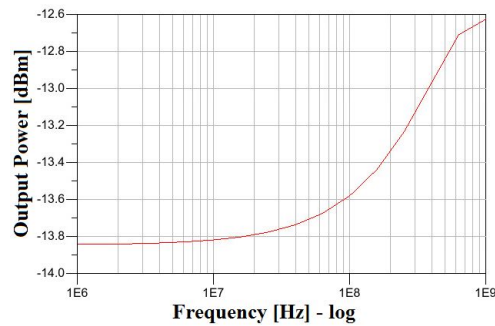
Figure 4.17: Plot of baseband bandwidth with $-5dBm$ RF input at 6 GHz.

- IMD3: displayed in Fig. 4.18, almost 50 dBc of IMD3 is achieved from 0.6 GHz to 6 GHz for most of the gain setting combinations, except at very low gain settings. However, the overall system linearity is not affected by the linearity degradation of the RFVM at low control voltages (high input power and low output power). When the RFVM is operating with high attenuation, the correction signal's power level is a few decades lower than the undistorted RF signal and results in almost no impact on overall signal quality.

(a)   (b)

Figure 4.18: IMD3 plot versus gain setting with -3.5 dBm input power at (a) 0.6 GHz and (b) 3 GHz.

- OP1dB at maximum gain: obtained by sweeping input power for a single-tone RF signal, while maintaining the maximum gain setting of 1.2 V at both V_{ci} and V_{cq}. OP1dB can be obtained at a point where the maxim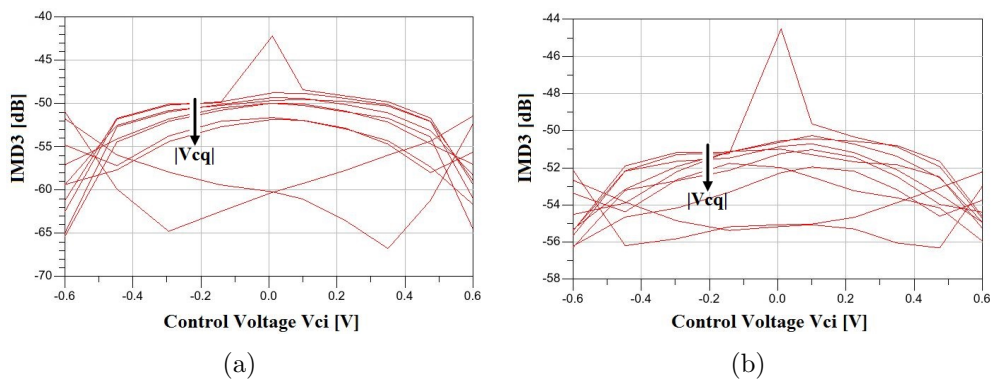um output power has dropped by 1 dB. Simulation results show that the RFVM has an identical OP1dB across the entire RF operation bandwidth: 2.2 dBm at 0.6 GHz, 2.2 dBm at 2.8 GHz, and 2.3 dBm at 6 GHz.

- OIP3 at maximum gain: the simulated results show an OIP3 of 19.5 dBm at 0.6 GHz, 17.3 dBm at 2.8 GHz, and 17.5 dBm at 6 GHz. Fig. 4.19 illustrates simulated OIP3 versus different input power at 2.8 GHz,
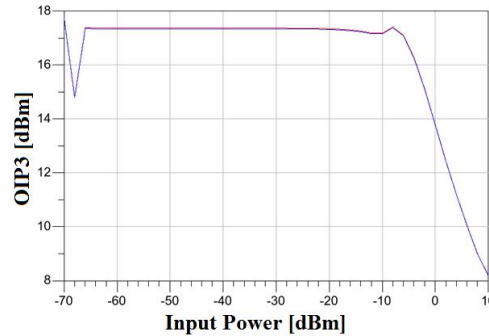
Figure 4.19: Plot of OIP3 versus input power at 2.8 GHz.

- Output noise power: The simulated output noise power was found to be -149 dbm/Hz at 6 GHz, which is the worst case across the RF frequency bandwidth.

- Gain and phase imbalances: the simulated results are displayed in Fig. 4.20; the worst amplitude and phase imbalances are 0.8 dB and 1.3° respectively.

The overall performances of the two RFVM designs are summarized in Table 4.4. The second design showed a significant improvement in RF operation bandwidth, linearity, and output noise power, and the power consumption is only half of the first design. The second design has achieved the design goals for the new ARF-PD system configuration, proven by post-layout simulation results. However, the improvement is at the cost of integration, since the second design uses discrete 90° hybrid couplers for quadrature phase generation. In addition, the RF delay block required to match the signal delays between the two paths over a wide RF bandwidth will be a challenge for future work.
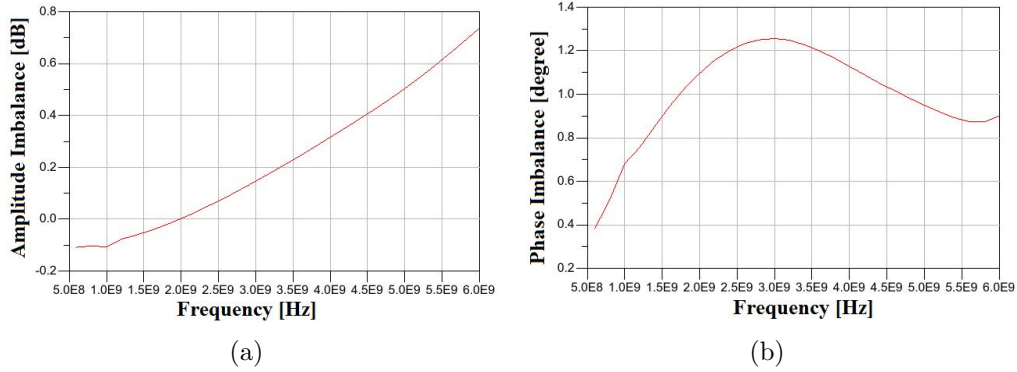
Figure 4.20: Plots of (a) amplitude imbalance and (b) phase imbalance of the differential output signals.

Table 4.4: Design 1 and 2 Performance Summary and Comparison

	Design 2	Design 1
Maximum Gain (dB)	0 to -38	5 to -28
Phase Range	360°	360°
RF Bandwidth (GHz)	0.6 - 6	0.6 - 2.8
$3dB$ Control Bandwidth (MHz)	>400	>400
Power Consumption (mW)	150	300
OP1dB (dBm)	2.2(@ 0.6 GHz)	6.4 (@ 0.6 GHz)
	2.2 (@ 2.8 GHz)	9.6 (@ 2.8 GHz)
	2.3 (@ 6 GHz)	-
IMD3 (dBc)(-3 dBm P_{in})	50 (@ 0.6 GHz)	43 (@ 0.6 GHz)
	49.5 (@ 2.8 GHz)	49 (@ 2.8 GHz)
	50 (@ 6 GHz)	-
OIP3 (dBm)	19.5 (@ 0.6 GHz)	17.4 (@ 0.6 GHz)
	17.3 (@ 2.8 GHz)	16.8 (@ 2.8 GHz)
	17.5 (@ 6 GHz)	-
Output Noise (dBm/Hz)	-149.5	-141

Chapter 5

Conclusions and Future Work

Driven by the exponential growth in demand for higher wireless data rates and capacity, communication systems are moving towards deploying densely packed small-cell base stations and widening the modulation bandwidths of communication signal to handle high data traffic areas. As PAs are the most nonlinear unit within the base station, PA linearization techniques have become a necessity. Baseband DPD has been a major candidate for power amplifier linearization. However, the adoption of wider modulation bandwidths and small-cell base stations have made the power consumption of conventional DPD solutions unsupportable. To overcome this problem, a low power ARF-PD solution has been proposed which preserves the capabilities of existing DPD solutions. The objective of this thesis was to develop a low power high linearity broadband RFVM for an ARF-PD system. The RFVM is a key building block in an ARF-PD system and acts as the interface between the pre-distortion engine and the undistorted signal in the RF domain. Based on an analysis of two designs proposed, and simulated results of the RFVM function, the following conclusions can be drawn:

Design I – In the first ARF-PD configuration, the pre-distortion function was applied directly to the undistorted signal; this imposed the need for high gain in the RFVM to ensure proper pre-distortion. A maximum 5 dB gain was targeted for the design, but this resulted in large transistor sizes and high power consumption in the output stage. This design also targeted integrated input QPG using a multi-stage RC PPF for broadband IQ generation from 0.6 GHz to 2.8 GHz. Unfortunately, this design suffered from a large noise contribution and high insertion loss due to the PPF network which further increased the power consumption in the output stage. However, the design was able to achieve comparable performance to commercial products ADL5390 and MAX2045.

Design II – To resolve issues discovered in the first design, an alternative ARF-PD configuration was adopted to relax the output power levels needed from the RFVM. In addition, discrete 90° hybrid couplers were used to mitigate limitations imposed by the multi-stage RC PPF. An active balun topology was implemented to provide wideband S2D and input matching. The conversion was needed due to the fact that the discrete hybrid couplers were single-ended but differential signals were needed for a continuous 360° phase control. In addition, a self-linearized VGA topology was proposed to improve the linearity of the complex VGA core. Overall, the second design showed significantly improved linearity and output noise levels, covering a much broader bandwidth from 0.6 GHz to 6 GHz, while only consuming half as much power as the first design.

Finally, while simulation results have shown satisfactory performance for both the RFVM designs when used as part of the ARF-PD system, their performance may still change drastically under measurement. Ultimately, measurements for both IC designs should be completed in the near future, in order to for validate the simulation results.

References

[1] H. Huang, J. Xia, A. Islam, E. Ng, P. M. Levine, and S. Boumaiza, "Digitally assisted analog/RF predistorter with a small-signal-assisted parameter identification algorithm," vol. 63, pp. 4297–4305, Dec. 2015.

[2] P. R. Gray, P. Hurst, R. G. Meyer, and S. Lewis, *Analysis and design of analog integrated circuits*. Wiley, 2001.

[3] J. Hoydis, M. Kobayashi, and M. Debbah, "Green small-cell networks," vol. 6, pp. 37–43, Mar. 2011.

[4] H. Dunn, "The carbon footprint of icts," *Global Information Society Watch*, 2010.

[5] H. Golestaneh, F. A. Malekzadeh, and S. Boumaiza, "An extended-bandwidth three-way doherty power amplifier," vol. 61, pp. 3318–3328, Sept. 2013.

[6] H. Sarbishaei, Y. Hu, B. Fehri, and S. Boumaiza, "Concurrent dual-band envelope tracking power amplifier for carrier aggregated systems," in *Proc. IEEE MTT-S Int. Microwave Symp. (IMS2014)*, pp. 1–4, June 2014.

[7] F. H. Raab, P. Asbeck, S. Cripps, P. B. Kenington, Z. B. Popovic, N. Pothecary, J. F. Sevic, and N. O. Sokal, "Power amplifiers and transmitters for RF and microwave," vol. 50, pp. 814–826, Mar. 2002.

[8] D. Y. T. Wu, F. Mkadem, and S. Boumaiza, "Design of a broadband and highly efficient 45W GaN power amplifier via simplified real frequency technique," in *Proc. IEEE MTT-S Int. Microwave Symp. Digest (MTT)*, pp. 1090–1093, May 2010.

[9] H. Sarbishaei, B. Fehri, Y. Hu, and S. Boumaiza, "Dual-band Volterra series digital pre-distortion for envelope tracking power amplifiers," vol. 24, pp. 430–432, June 2014.

[10] M. N. A. Abadi, H. Golestaneh, H. Sarbishaei, and S. Boumaiza, "Doherty power amplifier with extended bandwidth and improved linearizability under carrier-aggregated signal stimuli," vol. 26, pp. 358–360, May 2016.

[11] B. Razavi and R. Behzad, *RF microelectronics*, vol. 1. Prentice Hall New Jersey, 1998.

[12] N. Pothecary, *Feedforward linear power amplifiers.* Artech House Publishers, 1999.

[13] P. B. Kenington, *High linearity RF amplifier design.* Artech House, Inc., 2000.

[14] L. Ding, G. T. Zhou, D. R. Morgan, Z. Ma, J. S. Kenney, J. Kim, and C. R. Giardina, "A robust digital baseband predistorter constructed using memory polynomials," vol. 52, pp. 159–165, Jan. 2004.

[15] G. Yang, H. Wang, L. Li, and F. Liu, "One-step model extraction method for direct learning digital predistortion," *Electron Lett*, vol. 50, pp. 1148–1150, July 2014.

[16] R. N. Braithwaite and S. Carichner, "An improved doherty amplifier using cascaded digital predistortion and digital gate voltage enhancement," vol. 57, pp. 3118–3126, Dec. 2009.

[17] Maxim Integrated Product Inc., "RF predistortion (RFPD) vs. digital predistortion (DPD)," Sept 2015.

[18] J. Xia, A. Islam, H. Huang, and S. Boumaiza, "Envelope memory polynomial reformulation for hardware optimization of analog-rf predistortion," *IEEE Microwave and Wireless Components Letters*, vol. 25, no. 6, pp. 415–417, 2015.

[19] Y. Zheng and C. E. Saavedra, "Full 360 Vector-sum phase-shifter for microwave system applications," vol. 57, pp. 752–758, Apr. 2010.

[20] A. Asoodeh and M. Atarodi, "A full 360 Vector-sum phase shifter with very low RMS phase error over a wide bandwidth," vol. 60, pp. 1626–1634, June 2012.

[21] E. Juntunen, D. Dawn, J. Laskar, and J. Papapolymerou, "CMOS 45 GHz vector modulator with gain/phase correction through calibration," *Electron Lett*, vol. 49, pp. 267–269, Feb. 2013.

[22] H. Zijie and K. Mouthaan, "A 0.5 –6 GHz 360 deg vector-sum phase shifter in 0.13-μm CMOS," in *Proc. IEEE MTT-S Int. Microwave Symp. (IMS2014)*, pp. 1–3, June 2014.

[23] Y. Y. Huang, H. Jeon, Y. Yoon, W. Woo, C. H. Lee, and J. S. Kenney, "An ultra-compact, linearly-controlled variable phase shifter designed with a novel RC poly-phase filter," vol. 60, pp. 301–310, Feb. 2012.

[24] A. Devices, "Adl5390 rf/if vector multiplier," *ADL5390 RF/IF Vector Multiplier Rev. 0 (datasheet)*, 2009.

[25] M. I. P. Inc., "High gain vector multiplier." Device datasheet, 2003.

[26] F. Behbahani, Y. Kishigami, J. Leete, and A. A. Abidi, "CMOS mixers and polyphase filters for large image rejection," vol. 36, pp. 873–887, June 2001.

[27] J. Kaukovuori, K. Stadius, J. Ryynanen, and K. A. I. Halonen, "Analysis and design of passive polyphase filters," vol. 55, pp. 3023–3037, Nov. 2008.

[28] D. E. Norgaard, "The phase-shift method of single-sideband signal reception," *Proc IRE*, vol. 44, pp. 1735–1743, Dec. 1956.

[29] E. A. M. Klumperink, C. T. Klein, B. Ruggeberg, and E. J. M. van Tuijl, "AM suppression with low AM-pm conversion with the aid of a variable-gain amplifier," vol. 31, pp. 625–633, May 1996.

[30] J. Ramirez-Angulo and S. Ming-Shen, "The folded gilbert cell: a low voltage high performance CMOS multiplier," in *Proc. 35th Midwest Symp. Circuits and Systems*, pp. 20–23 vol.1, Aug. 1992.

[31] H. W. Bode *et al.*, "Network analysis and feedback amplifier design," 1945.

[32] S. S. Mohan, M. D. M. Hershenson, S. P. Boyd, and T. H. Lee, "Bandwidth extension in CMOS with optimized on-chip inductors," vol. 35, pp. 346–355, Mar. 2000.

[33] S. C. Blaakmeer, E. A. M. Klumperink, D. M. W. Leenaerts, and B. Nauta, "Wide-band balun-LNA with simultaneous output balancing, noise-canceling and distortion-canceling," vol. 43, pp. 1341–1350, June 2008.

[34] J. N. Babanezhad and G. C. Temes, "A 20-V four-quadrant CMOS analog multiplier," vol. 20, pp. 1158–1168, Dec. 1985.

[35] M. Goossens, F. Mittelbach, and A. Samarin, *The LaTeX Companion*. Reading, Massachusetts: Addison-Wesley, 1994.

[36] D. Knuth, *The TeXbook*. Reading, Massachusetts: Addison-Wesley, 1986.

[37] L. Lamport, *LaTeX — A Document Preparation System*. Reading, Massachusetts: Addison-Wesley, second ed., 1994.

[38] E. Juntunen, D. Dawn, J. Laskar, and J. Papapolymerou, "CMOS 45 GHz vector modulator with gain/phase correction through calibration," *Electron Lett*, vol. 49, pp. 267–269, Feb. 2013.

[39] A. Kerr, "Theoretical and practical limits on the rf bandwidth of sis mixers," *RN*, vol. 1, pp. 0–71, 1994.