

Broadband mm-Wave Transceivers for Sensing and Communication

Andrew Townley



Electrical Engineering and Computer Sciences
University of California at Berkeley

Technical Report No. UCB/EECS-2020-25

<http://www2.eecs.berkeley.edu/Pubs/TechRpts/2020/EECS-2020-25.html>

May 1, 2020

Copyright © 2020, by the author(s).
All rights reserved.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission.

**Broadband mm-Wave Transceivers for
Sensing and Communication**

by

Andrew Townley

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy

in

Engineering — Electrical Engineering and Computer Sciences

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Ali M. Niknejad, Chair
Professor Elad Alon
Professor Martin White

Spring 2018

The dissertation of Andrew Townley, titled Broadband mm-Wave Transceivers for Sensing and Communication, is approved:

Chair	_____	Date	_____
	_____	Date	_____
	_____	Date	_____

University of California, Berkeley

**Broadband mm-Wave Transceivers for
Sensing and Communication**

Copyright 2018
by
Andrew Townley

Abstract

Broadband mm-Wave Transceivers for
Sensing and Communication

by

Andrew Townley

Doctor of Philosophy in Engineering — Electrical Engineering and Computer Sciences

University of California, Berkeley

Professor Ali M. Niknejad, Chair

Scaling in silicon semiconductor process technology, although driven by digital applications, has also enabled the operation of analog integrated circuits (ICs) at higher and higher frequencies. Over the past 15-20 years, ICs operating at millimeter-wave (mm-Wave), the frequency range between 30 GHz and 300 GHz, have been demonstrated with increasing performance and complexity.

One important application for mm-Wave ICs has been in automotive radar, where they have been used to make accurate measurements of distance and velocity with minimal processing required. There also has been significant interest in adapting this technology for non-vehicular applications, such as gesture recognition, room occupancy detection, or heart-rate monitoring, where performance and energy efficiency are both important. The first part of this thesis describes a custom IC for gesture recognition radar demonstrating state-of-the-art energy efficiency. The IC consists of four transmitters and four receivers with shared frequency generation circuitry, and is packaged onto a 1.2x1.2cm antenna module containing eight antennas.

The other key application for mm-Wave technology has been for wireless communication. Products are finally coming to the market now that offer nearly 5 Gigabits per second of wireless data throughput for indoor wireless LAN applications, and mm-Wave technology will likely play a role in the next generation of wireless cellular standards as well. To demonstrate the possibility for yet-higher data rates to be achieved, a broad-bandwidth custom integrated circuit transceiver has been designed targeting a factor of 10 improvement in wireless data throughput beyond commercially available technology. The second half of this thesis will discuss the details of a transceiver and antenna design for broad-bandwidth and high data rate operation.

To my parents,
without whose love and encouragement I never could have made it this far.

Contents

Contents	ii
List of Figures	iv
List of Tables	ix
1 Introduction	1
1.1 Motivation	2
2 Millimeter-Wave Background	4
2.1 Technology: Bipolar vs MOSFET	4
2.2 LO Generation and Distribution	5
2.3 Modulation Techniques	5
2.3.1 Modulation Techniques for Radar	5
2.3.2 Modulation Techniques for Digital Communication	7
2.4 Phased Array Techniques	10
3 FMCW Radar Phased-Array Transceiver Design	11
3.1 Proposed System Architecture	11
3.2 Transmitter	13
3.3 Phase Shifter	16
3.4 LO Distribution	18
3.5 Receiver	22
3.6 Packaging	25
4 Radar IC Measurements	28
4.1 Probe Station Measurements	29
4.1.1 LO	29
4.1.2 Transmitter	29
4.1.3 Receiver	30
4.2 Packaged Measurements	31
4.2.1 Array Characterization	31
4.2.2 Radar Measurements and Characterization	35

4.3	Conclusion	37
5	Wideband mm-Wave Transceiver Design	40
5.1	Motivation and System Architecture	40
5.1.1	System Architecture	41
5.1.2	Packaging Approach	42
5.2	Building Blocks	44
5.2.1	Modular Common-Source Neutralized Amplifier Layout	44
5.2.2	Coupled Resonators using Low-K Transformers	45
5.2.3	Low-K Transformers with Lossy Inductors	49
5.3	Transmitter	51
5.3.1	Modulator Design	51
5.3.2	Power Amplifier Design	68
5.3.3	TX Chain Simulated Results	71
5.4	Receiver	74
5.4.1	Baseband Amplification	74
5.4.2	Active Mixer with TIA load	75
5.4.3	LNA	81
5.4.4	RX Chain Simulated Performance	83
5.5	PCB and Antenna Design	84
5.5.1	Folded Dipole Antenna	85
5.5.2	Surface Waves	88
5.5.3	Methods for Dealing with Surface Waves	89
5.5.4	Planar Balun on PCB	94
5.6	Fabricated Transceiver and Antenna PCB	98
5.7	Performance Comparison	101
6	Conclusion	102
6.1	Summary of Thesis	102
6.2	Future Work	103
6.2.1	Carrier Recovery	103
6.2.2	Equalization	104
	Bibliography	105

List of Figures

1.1	United States Frequency Allocations Chart, 30–300GHz band.	2
3.1	Complete phased-array transceiver block diagram. mm-wave IOs use a single-ended, ground-signal-ground (GSG) pad configuration. Ground pads are shared between adjacent phased-array elements to reduce die area.	12
3.2	Three-stage power amplifier schematic. R_5 is chosen to result in an emitter voltage of about 100mV under small-signal bias conditions. The annotated DC currents correspond to the operating points in small-signal (left of arrow) and saturated large-signal (right of arrow) conditions.	13
3.3	Full 3D EM model of PA interstage and output transformers.	14
3.4	Simulated PA power gain and power-added efficiency at 94GHz, plotted versus output power.	15
3.5	Simulated PA saturated output power, power gain, and peak power-added efficiency vs. frequency.	16
3.6	Phase shifter schematic.	17
3.7	Quadrature hybrid HFSS model. The dimensions of the simulated region are $197\mu\text{m}$ by $124\mu\text{m}$	18
3.8	Quadrature hybrid simulated IQ mismatch.	18
3.9	Phase shifter constellation points, showing combinations of phase and amplitude settings, and the phase and amplitude error for a desired phase angle.	19
3.10	Schematic of LO generation and distribution circuitry.	20
3.11	LO distribution network and power divider, based on lumped-element artificial transmission line. Each artificial quarter-wavelength line consists of a 1.5 turn inductor in series with a 1.25 turn inductor, which interfaces with a 50ohm transmission line.	20
3.12	Power splitter topologies considered for LO distribution.	22
3.13	Receiver schematic.	22
3.14	Input impedance locus of the LNA, at various stages of the matching network.	23
3.15	Receiver HFSS model.	24
3.16	Simulated receiver conversion gain and noise figure vs mixer LO frequency, for all 5 gain control settings.	24

3.17	(a) Percentage contributions of different blocks to total output noise. The contributions are separated by sideband. (b) Noise circles	25
3.18	HFSS simulations of antenna module.	25
3.19	Levels of integration of radar IC. Bare die (a) is flip-chip chip packaged onto BGA antenna module (b), which is integrated onto a test PCB (c). The test PCB is 10.2cm x 10.2cm, sized to accomodate connectors to a separate FPGA board. If a microcontroller were used instead of an FPGA, the board area could be reduced to a much smaller size.	26
3.20	Photograph of 1.2cm×1.2cm BGA antenna module with die attached, chip (left) and antenna (right) sides.	27
4.1	Die photograph of fabricated chip.	28
4.2	VCO tuning curve, and measured phase noise of PLL at 94GHz.	29
4.3	Measured transmitter output power vs frequency.	30
4.4	The measured LNA input impedance (a) and PA output impedance (b) showed good agreement with simulation.	30
4.5	Measured receiver single-sideband noise figure (probe station) vs frequency. . . .	31
4.6	EIRP at broadside vs frequency, for 1, 2, and 4 PAs enabled. All possible combinations of the 1 and 2 PA cases are plotted.	32
4.7	Beam steering of transmitter at 94GHz, characterized manually at BWRC. . . .	33
4.8	RX array conversion gain vs frequency, for 1, 2, and 4 LNAs enabled. All possible combinations of the 1 and 2 LNA cases are plotted.	33
4.9	Measured phase shifter constellation points, with circle showing amplitude level with least average amplitude error.	34
4.10	The measurement setup at University of Nice-Sophia Antipolis. The test PCB is placed with the radiating side downward, and the arm with the receive antenna is swept across phi and theta angles. The measured radiation pattern data is imported into HFSS and plotted. The axes shown in (a) correspond to the axes in the HFSS plots (subfigures b and c)	34
4.11	3D radiation pattern measurements at various beam steering angles (performed at UNS).	35
4.12	EIRP measurement along H-plane (performed at UNS).	36
4.13	EIRP measurement along E-plane (performed at UNS).	36
4.14	Radar measurements with a single target, at various distances. (a) with wide RF sweep bandwidth (b) with narrower RF sweep bandwidth.	37
4.15	Measured DC power consumption by supply domain.	39
5.1	Block diagram of 120GHz dual-channel communication transceiver.	42
5.2	Link budget, assuming a 10cm transmit distance.	43
5.3	Bit error rate versus SNR for various modulation techniques.	43
5.4	Angle view of neutralized unit cell layout	44
5.5	Angle view of amplifier with 7 unit cells	45

5.6	Magnetically coupled LC resonators	46
5.7	$ Z_{21} $ transfer functions for various values of the uncoupled resonator Q	47
5.8	Frequencies of maxima and minimum of coupled RLC circuit transimpedance for various values of Q.	48
5.9	Ripple of the magnitude of the transimpedance vs coupling coefficient.	49
5.10	Magnetically coupled LC resonators with finite Q inductors	50
5.11	$ Z_{21} $ transfer functions with finite-Q inductors	51
5.12	Conventional upconversion mixer topology typically used at low frequencies. . .	52
5.13	Unit element of common-gate RF DAC.	52
5.14	DAC operation for 8-level amplitude modulation. The DAC is depicted at code 101, so the total current at the output is $+\frac{4}{7} - \frac{2}{7} + \frac{1}{7} = \frac{3}{7}$ of the full scale value.	53
5.15	Small signal model of common-gate amplifier.	54
5.16	DAC model with continuous current steering. w represents the proportion of DAC unit elements steered in the “+1” state, so then $1 - w$ represents the proportion of DACs in the “-1” state.	56
5.17	Phase of $\tilde{y}_{21}\tilde{y}_{12}$	59
5.18	Magnitude of DAC loop gain under various loading conditions	60
5.19	AM-AM and AM-PM nonlinearity of DAC	60
5.20	I and Q DACs with current-combining at output	61
5.21	DAC schematic highlighting “half-DAC” cells.	62
5.22	DAC schematic highlighting segmentation of bits within DAC.	63
5.23	Layout of DAC cell including top level metal routing.	64
5.24	65
5.25	66
5.26	Design iterations for quadrature hybrid	66
5.27	Quadrature hybrid layout in between flip-chip pads	68
5.28	Simulated S-parameters of final quadrature hybrid design.	69
5.29	DAC output transformer / PA input transformer	69
5.30	Power Amplifier Schematic	70
5.31	Simulated PA input and output return losses.	70
5.32	Simulated PA gain and group delay.	70
5.33	Simulated PA efficiency and gain compression curves.	71
5.34	Static TX constellations from 110GHz to 120GHz	72
5.35	Simulated QPSK eye diagram and constellations for 40ps symbol period, corresponding to 50Gb/s data rate	73
5.36	Simulated 16QAM eye diagram and constellations for 80ps symbol period, corresponding to 50Gb/s data rate	73
5.37	Simulated 64QAM eye diagram and constellations for 120ps symbol period, corresponding to 50Gb/s data rate	73
5.38	Baseband amplifiers and output driver	74
5.39	Voltage gain of baseband amplifier and output driver (including transmission line routing)	75

5.40	Double balanced active mixer core schematic	76
5.41	Different strategies for the DC load of the mixer: Resistor load (a), PMOS load (b), PMOS load with TIA (c).	77
5.42	Range of gain and bandwidth possibilities for different mixer baseband loads. . .	78
5.43	Finalized mixer schematic.	79
5.44	Mixer layout floorplan. Due to IO pad constraints, the I and Q baseband amps both need to be on the same side of the mixers, necessitating routing the mixer baseband outputs a long distance (shown in thin cyan lines).	80
5.45	LNA schematic	81
5.46	Simulated LNA input impedance seen from PCB (including chip-to-PCB transition model)	82
5.47	Simulated LNA Gain and Group Delay (including 3dB loss from I-Q split) . . .	82
5.48	Simulated LNA Noise Figure	83
5.49	Simulated conversion gain of the full receiver chain, including LNA, mixer, and baseband amplification. The x-axis represents the frequency offset from the 115GHz LO signal and the baseband tone frequency that it downconverts to. . .	84
5.50	Simulated double sideband noise figure for the full RX chain with 115GHz LO frequency.	84
5.51	Selected PCB stackup.	85
5.53	Simulated antenna radiation pattern (θ representing the polar angle from the z-axis) versus frequency for thick and thin substrates	87
5.54	Dielectric slab which supports TM_0 surface waves.	88
5.55	Calculated H field of grounded slab TM_0 surface wave mode according to theory. The horizontal axis is the direction of propagation, and the vertical distance from the bottom of the image. The white horizontal line represents the air-dielectric interface.	88
5.56	Sievenpiper “mushroom” EBG structure, side view.	89
5.57	Patch antenna gain (at 115GHz) with and without guard ring, as the PCB substrate size is varied.	90
5.58	Antenna gain (at 115GHz) with two guard rings, as PCB substrate size is varied. . .	91
5.59	Antenna gain and group delay at broadside with one and two guard rings	91
5.60	Patch with grounded via ring and partial substrate removal	92
5.61	Different surface wave suppression techniques. Trench/substrate removal areas are shown in yellow.	93
5.62	Worst-case gain variation (minimum to maximum) of various surface wave mitigation techniques, as the (square) substrate size is varied from 12 to 15mm. . . .	94
5.63	Fabricated antenna with partial substrate removal.	95
5.64	Input impedance of final antenna design, shown for various substrate dimensions . . .	95
5.65	Design process for optimized rat-race hybrid without sum port termination. . .	96
5.66	Final balun design and simulations	97
5.67	Fabricated rat-race hybrid on PCB.	97
5.68	Die photo.	98

5.69	Flip-chip footprint, differential routing, and mm-Wave baluns as fabricated on PCB.	99
5.70	Antenna PCB with chip attached.	100

List of Tables

4.1	Performance Comparison of Published W-Band Phased Array Transceivers . . .	38
5.1	Mapping from DAC bits to amplitude values	53
5.2	Characteristics of Various PCB Materials	86
5.3	Performance and Energy Efficiency of mm-Wave Wideband Transmitters	101
5.4	Performance and Energy Efficiency of mm-Wave Wideband Receivers	101

Acknowledgments

First and foremost, I cannot thank my advisor Ali Niknejad enough for his technical advice and project guidance. I’ve had the luxury and flexibility of choosing my projects throughout most of my PhD, and I think few advisors would be willing to do that. It’s been fantastic to work for someone who has such a depth of technical expertise. Ali has also been a great advocate for my ambitious system-scale projects, and hopefully they will continue on in the center after I leave. I also want to thank Elad Alon who basically acted as an informal coadvisor during the design of the “Hummingbird” 120GHz chip, and provided many valuable insights.

I had the benefit of having great lecturers for all of the courses I’ve taken at Berkeley, and thank Ali, Elad, Bernhard Boser, Simone Gambini, Bora Nikolić, Chris Hull, Joseph Orenstein, and Martin White. Extra thanks to Elad, Bora, and Martin for serving on my quals committee and to Elad and Martin as readers for this thesis. Also an extra thanks to Simone, for helping me figure out what I wanted out of grad school over probably the most important beer of my career.

I would like to thank Nokia Research / Nokia Labs for research funding and support on the radar project, specifically Vason Srimi and Klaus Doppler. I also appreciated the various technical discussions I had during my internship with Michael Reiha and my fellow co-intern Paul Swirhun. I owe many thanks to Paul for his design efforts on the phased-array radar IC. I also would like to thank the NSF-EARS program for its financial support for me and my research. Thank you to the BWRC member companies and sponsors as well.

I also enjoyed my summer internships at Analog Devices and Google, and learned a lot from them. Thanks to John Cowles, Todd Weigandt, Prabir Saha, Daryl Carbonari, Joel Dobler, Monica Cordrey, and Barrie Gilbert for a really enjoyable summer up in the northwest. And thank you to Ben Mossawir and Will Wesson for mentoring me during my summer at Google.

In my (admittedly biased) opinion, BWRC is probably the best possible place in the universe to do a PhD in integrated circuits. A big part of that is the people. Firstly, I would like to thank all of the BWRC “senior students” who I have bugged over the years with various questions: Amin Arbabian, Jungdong Park, Jiashu Chen, Lingkai Kong, Matt Spencer, Sahar Tabesh, Siva Thyagarajan, Shinwon Kang, Steven Callender, and Jun-Chau Chien. I also really enjoyed all of the technical feedback, discussions, and collaborations with my peers in the Niknejad and Alon groups. In particular, I would like to thank Nathan Narevsky, Greg LaCaille, Luke Calderin, Constantine Sideris, Pramod Murali, Nai-Chung Kuo, and Sashank Krishnamurthy. Costis and Sashank also contributed quite a bit to making the 120GHz transceiver a reality and it would not have been possible without them.

Over the years at BWRC, we have also had great support from the staff, and I would be remiss if I did not thank them profusely for all that they do for the center and its students. At the EE department level, I also really would like to thank Shirley Salanio for her tireless efforts on keeping the entire department of grad students on track, myself included.

I owe a big thanks to my research collaborators at the University of Nice Sophia Antipolis and STMicroelectronics: Cyril Luxey, Aimeric Bisognin, Diane Titz, Fred Giancesello, and Romain Pilard. I also really appreciate Cyril and Diane's hospitality during my visit to Nice for the radar active measurements, during a very difficult week for the city of Nice.

On a personal level, I want to thank my fellow Shattuck house roommates, for a really enjoyable five and a half years. Special thanks to Daniel Gerber, for reminding me how much I love the outdoors and many camping and climbing trips. And thanks to Joe Corea for showing me the importance of having a good hobby or two (or three).

To Allison, thank you so much for your love, support, and understanding throughout the final years of my PhD, even when they took twice as long as I thought they would.

Finally, thank you to my parents for their support and encouragement, and for everything you have done for me that got me to this point. Thank you for cheering me on to the finish line from the other side of the country.

Chapter 1

Introduction

Since the introduction of digital wireless communication into the industrial and consumer space, the evolution of wireless technology has been a self-reinforcing cycle, with each new generation of technology generating the demand for the next. Wireless LAN on laptops first paved the way for internet access without being tied to a desktop PC. This was followed by the popularization of smartphones, which provided users wireless connectivity anywhere a cellular connection could be found. In turn, the maturation of smartphone technology has taken place alongside dramatic increases in cellular data capacity. Basic internet connectivity and picture sharing has been joined by video messaging and live streaming, and perhaps in the near future will be joined by augmented or virtual reality.

To enable these developments to take place, wireless technology has had to advance continuously. But satisfying the demand for higher and higher wireless data rates, at some point, will inevitably be constrained by physical limitations. Once that happens, there are really only two ways to improve the situation, according to Shannon's capacity theorem [1]:

$$C = BW \cdot \log_2 \left[1 + \frac{S}{N} \right] \quad (1.1)$$

When up against fundamental limits, either the signal-to-noise ratio or the signal bandwidth should be increased. Increasing signal-to-noise ratio allows for the use of more complex modulation schemes, but requires a higher transmit power and/or a lower receiver noise level to achieve. Because of the logarithm in the capacity equation, there are also diminishing returns in increasing the SNR. For high SNR, a further doubling of SNR only leads to a linear increase in channel capacity. At some point, increasing the SNR becomes prohibitive from a DC power consumption point of view, or even simply impossible due to regulatory constraints on transmitted power level or interference from other users.

If that is the case, the only available route forward is to scale bandwidth. This, too, has costs when it comes to the physical implementation of the communication link. Transmitters, receivers, and antennas cannot be made arbitrarily broadband without taking some performance hits (or causing significant headaches for the circuit, antenna, and system design engineer). The performance hit (or the size of the engineer's headache) is not related



Figure 1.1: United States Frequency Allocations Chart, 30–300GHz band.

to the absolute bandwidth required, but in fact the *fractional* bandwidth — the ratio of the bandwidth to the center frequency. It is quite difficult to design a 5 GHz carrier frequency wireless link with 5 GHz of bandwidth, while a link bandwidth of 5 GHz is readily achievable at 60GHz carrier frequency.

Shannon’s law does not say anything about the center frequency of the available channel, so any 5 GHz chunk of spectrum is equally usable. But in order to more practically and efficiently take advantage of broad bandwidth channels, we need to operate at higher frequency. This is the motivation behind operating circuits and systems in the so-called “millimeter-Wave” band.

Millimeter-Wave (mm-Wave) refers to radio frequencies with a wavelength of between 1 and 10mm, or equivalently frequencies between 30 and 300GHz (Figure 1.1). Until fairly recently, this spectrum was limited to military and radioastronomy applications. However, in the last 15-20 years, there has been significant development of circuits and systems operating in the mm-Wave band. Initially, 60GHz-band wireless communications were the driving force, culminating in the development of fully integrated transceivers supporting large antenna arrays, supporting up to 7Gb/s of peak data rate [2][3].

More recently, the major application driving mm-Wave development has been automotive radar [4][5][6]. Mm-Wave radar was initially proposed for driver-assist technologies such as adaptive cruise control and lane change warning systems, which both require the high bandwidth available at mm-Wave frequencies for accurate distance resolution. Getting automotive radar systems into the average car required driving the cost down, necessitating a change from the very first modules using gallium-arsenide (GaAs) integrated circuits, to lower-cost silicon solutions. Fully integrated CMOS systems will soon be reaching the market [6], and as interest in fully autonomous vehicles continues to grow, development of this technology will accelerate further.

1.1 Motivation

The goal of this thesis is to explore the design tradeoffs and challenges in the design of broad-bandwidth millimeter-wave integrated transceivers, with an emphasis on the demonstration of practical, complete systems. In this thesis, two transceiver designs are described, one intended for mm-Wave short-range radar, and another for short-range communication. Due to the complexity of routing signals on and off chip at millimeter-wave, it is also necessary to consider the packaging and antenna design as part of the system design itself. Both

transceivers are integrated using flip-chip die attach technology directly onto the printed circuit boards (PCBs) with included antennas.

Chapter 2

Millimeter-Wave Background

2.1 Technology: Bipolar vs MOSFET

A substantial amount of design work at millimeter-wave has been done using silicon germanium (SiGe) bipolar transistor technologies. These tend to be fairly inexpensive technologies to use, since the critical dimension for high performance is the defined not by lithography but by layer thickness (width of the base region of the bipolar transistor, in a conventional vertical bipolar device). In addition to reduced fabrication mask cost, bipolar transistors also have a better $\frac{g_m}{I_C}$ than CMOS devices, and the g_m efficiency is constant versus bias, rather than trading off with f_T as in CMOS.

Modern bipolar process offerings do almost always include some CMOS transistors as well — this type of process is referred to as SiGe BiCMOS — though the feature size is usually several process nodes behind the cutting edge [7][8]. Because of this, BiCMOS is best suited for analog-heavy applications such as power amplifiers [9], where low-speed digital circuitry may sometimes be included but it is not critical to the performance of the system.

For digitally intensive circuits, CMOS clearly has the advantage. High data-rate digital transmitters have been demonstrated with intensive digital filtering while still maintaining high efficiency [10][11]. This certainly would not be possible using bipolar transistor logic, although it could perhaps be matched in performance with some efficiency cost using a sufficiently finely scaled BiCMOS technology. Since SiGe BiCMOS scaling has not caught up with pure CMOS scaling, it is not able to efficiently provide the high performance digital functionality necessary to the operation of a complex mm-Wave system. CMOS is the de facto technology of choice for any millimeter-wave SoC intended for mass-market production [6].

As CMOS technology scales further into the FinFET regime, more challenges are introduced: thinner and more resistive routing layers, higher parasitic capacitances due to non-planar device geometry, and stringent metal fill requirements. Initial investigations published in the literature suggest that these challenges can be overcome, and millimeter wave CMOS can still achieve high performance even scaled below 28nm [12]. Once FinFET

processes become more widely available and the design challenges better understood, it is possible that performance can be improved even further — the f_T of FinFET devices in subthreshold operation may be sufficient to achieve high gain at millimeter wave, which translates to an efficiency boost due to improved $\frac{g_m}{I_D}$.

2.2 LO Generation and Distribution

One of the most challenging aspects of generating an LO signal at millimeter-wave is designing a high-performance oscillator with wide tuning range. The poor analog varactor Q at high frequency degrades oscillator phase noise [13][14]. Therefore, it is advantageous to generate the LO frequency at a lower frequency, and use a nonlinear frequency multiplier to scale it up by the desired amount.

In a multi-channel transceiver, where the LO signal is shared across multiple channels, a transmission-line-based routing network is usually required to keep impedances well-matched over long distances. This can also be a source of high power consumption, since large LO buffers are required to drive the fairly low transmission line impedances that can be realized on chip (which typically do not exceed 100Ω) without incurring high losses.

2.3 Modulation Techniques

2.3.1 Modulation Techniques for Radar

Of the various modulation schemes used for radar, all have the same main idea in common. A signal is transmitted with a known time-varying modulation, which reflects off of various scatterers in the environment, and some small amount of energy is reflected back to the receiver. Because the received signal is simply a time-delayed and attenuated version of the transmitted signal, it can be compared with the known transmitted waveform to determine the time delay. From the time delay, the round-trip distance is trivial to compute: $d = \frac{c}{t_d}$.

Also, because the transmitted signal and the received signal are both derived from the same clock, their phase noise will largely cancel out. When the offset frequency of the phase noise is equal to $\frac{1}{T_{\text{roundtrip}}}$, the phase noise will add coherently, and increase the noise floor of the receiver. This is not as much of an issue for targets that are far away — the signal will be heavily attenuated. The greatest concern is with high amplitude signals from nearby targets. As these signals (often referred to as “clutter”) also tend to degrade the linearity of the receiver, it is obviously desirable to avoid them.

In practice, the transmitted signal tends to be periodic, which potentially introduces some ambiguity. If the period of the transmitted signal is T , then a given receive waveform could correspond to a round-trip time delay of $\tau, \tau + T, \tau + 2T, \dots$. For millimeter-Wave radar, this tends to not be an issue because the high signal attenuation tends to push the reflected signals from longer distances below the noise floor of the receiver.

A critical parameter for the performance of a radar system is the bandwidth of the modulated signal. In most types of radar system, the distance resolution is inversely proportional to the bandwidth of the radar signal: a signal with larger bandwidth can better resolve two close-together targets [15]. This tradeoff is broadly applicable, since time (and therefore also distance, because of the constancy of the speed of light) and frequency are related by the Fourier transform — a signal that is confined to a small region in time is necessarily broad in frequency space.

2.3.1.1 Pulse-Based Radar

Pulse-based radar is straightforward to understand: a short pulse is transmitted, the pulse reflects off of various objects in the environment, and a series of pulses is received corresponding to reflections from objects at different distances. The narrower the pulse, the more accurately closely spaced objects can be resolved. One nice feature is that the pulse need not be detected coherently, since the information is contained only in the amplitude of the signal. This approach is commonly used at optical frequencies in LiDAR (typically referred to as “time of flight” imaging) [16]. It has also been demonstrated in mm-Wave integrated circuit form, with significant design effort required to ensure sufficiently narrow pulses [17].

2.3.1.2 Continuous Wave Radar

In contrast with pulse-based radar, continuous wave radar is always transmitting a signal. This presents significant challenges for the receiver, as it must either be carefully isolated from the transmitter to avoid saturation at its output, or it must somehow tolerate the large transmitter signal without compromising linearity and noise figure.

Continuous wave radar has no amplitude modulation, so in order to be able to measure range, phase or frequency modulation must be added. Phase modulated continuous wave radar (PMCW radar) is attractive in that it can be implemented in an extremely simple and efficient manner, and orthogonal pseudo-random sequences can be used at different times to reduce the probability of inter-user interference from multiple radar systems operating in close proximity [18]. The bandwidth in a PMCW system is set by the sample rate at which the phase is modulated. Unfortunately, this also requires the same high bandwidth in the receiver baseband amplification chain [18], which can become very challenging in high bandwidth / high resolution applications.

Frequency modulation, on the other hand, requires low baseband bandwidth [19][6], and low instantaneous transceiver bandwidth. In a frequency modulated continuous wave radar (FMCW radar) system, high resolution still requires high bandwidth, but the bandwidth does not come from fast frequency steps. Instead, high bandwidth is achieved using a slow, smooth frequency ramp over multiple GHz. By removing the need for an instantaneously wideband signal, FMCW radar relaxes the specifications on the circuit design. This leads to a more efficient overall implementation.

2.3.1.3 Advantages of Linear FMCW Radar for Millimeter-Wave

One specific type of FMCW waveform has some very desirable properties: linear FMCW. If the transmitted signal increases linearly in frequency, it can simply be mixed with the received reflected signal to yield an IF tone whose frequency is proportional to round-trip time delay [19]. Linear frequency modulated continuous wave (FMCW) radar is an attractive radar modulation scheme for energy-efficient mm-wave applications for a few main reasons: constant-envelope and simplicity.

Firstly, because FMCW is a constant-envelope modulation scheme, transmitter linearity is not a concern. This allows for use of linear power amplifiers close to saturation, or even nonlinear switching power amplifiers, either of which will improve the overall transmitter efficiency. Also, because the modulated LO signal is constant-envelope, it can be generated at a low frequency and scaled up to a higher frequency using a nonlinear frequency multiplier, without negative impacts from the nonlinearity of the multiplier.

In a phased-array system, the modulated LO signal can be generated centrally, scaled in frequency, and routed out to all elements. As the LO routing network likely needed to be there anyway, there is no penalty for doing this.

Secondly, because the modulation is simple and shared across all elements, the frequency modulation can be incorporated into the PLL that is likely present in the system anyway. For a pulsed radar system, a high-bandwidth modulation requires fast on/off times to achieve a short pulse width [17], which means the circuit creating the modulation needs to be carefully designed to support that bandwidth. In a FMCW radar system, although large overall bandwidth is needed in the RF transmit and receive chains (as in the pulsed radar case), a large *instantaneous* bandwidth is not necessarily needed, since it is the overall bandwidth of the sweep itself that determines the resolution. So, a slowly modulated signal can be used, as long as the frequency of the signal varies across the full bandwidth over time. Multiple targets at different distances correspond to multiple tones, and can be easily distinguished by applying a Fourier transform to the received IF data.

2.3.2 Modulation Techniques for Digital Communication

The most basic modulation technique, and perhaps the first modulation technique to ever be used [20], is on-off keying. This simply consists of turning on and off the RF signal in time:

$$v(t) = A(t) \cos(\omega t) \quad (2.1)$$

Where:

$$A(t) = \sum_{n=-\infty}^{\infty} [u(t - nT) - u(t - (n + 1)T)] a[n]$$

$$a[n] \in \{0, 1\}$$

and $u(t)$ is the unit step function, which is used to mathematically model the zero-order hold interpolation of the discrete-time digital data sequence $a[n]$.

Note that this inherently assumes that the digital data is sampled at a regular interval T . Since this waveform transmits one new bit of information every T seconds, we can say that the data rate is $1/T$ bits/second.

2.3.2.1 Amplitude Modulation

To improve the data rate, instead of transmitting only a one or a zero, more signal levels can be used to convey more information per bit period. For example, if there are four possible signal levels instead of two, $\log_2(4) = 2$ bits can be transmitted every T seconds. In this case, $a[n]$ can take a broader range of values. If k bits per period are desired:

$$a[n] \in \left\{ 0, \frac{1}{2^k - 1}, \frac{2}{2^k - 1}, \dots, \frac{2^k - 2}{2^k - 1}, 1 \right\} \quad (2.2)$$

Without even doing any math, it's obvious that there is a tradeoff for doing this: Shannon's equation (Equation 1.1) says that if the channel bandwidth is the same (which it approximately is, if T is held constant), then the signal to noise ratio needs to increase in order to provide increased capacity. So, in order to actually achieve the increase in data rate, we can deduce that the SNR has to increase proportionally.

2.3.2.2 Phase Modulation

Other than by changing the amplitude of a sinusoid, the only other way to send data is by modulating the phase or frequency of the wave. In an analog communication system, these are essentially the same thing: since frequency is the derivative of phase, in a continuous-time system it makes little difference for modulation and demodulation if the signal of interest is the phase or its derivative. For a sampled, digital communication system, there is in fact a meaningful distinction, because the phase and frequency now have to change in discretized steps rather than smoothly.

Although it has been demonstrated at mm-Wave [21], digital frequency modulation suffers from poor bandwidth efficiency [22] and would be extremely challenging to implement in a broadband, coherent link. Because of this, mm-Wave transceivers generally have used almost exclusively phase modulation rather than frequency modulation.

The simplest version of phase modulation is binary phase shift keying (BPSK). Mathematically, this takes the form of a sinusoid that alternates between inverted and non-inverted phase.

$$v(t) = \cos(\omega t + \phi(t)) \quad (2.3)$$

Where:

$$\phi(t) = \sum_{n=-\infty}^{\infty} [u(t - nT) - u(t - (n + 1)T)] \phi[n]$$

$$\phi[n] \in \{0, \pi\}$$

The concept can be extended to k bits per symbol by increasing the number of possible values that $\phi[n]$ can take:

$$\phi[n] \in \left\{ 0, \frac{2\pi}{2^k}, 2 \cdot \frac{2\pi}{2^k}, \dots, (2^k - 1) \cdot \frac{2\pi}{2^k} \right\}$$

This technique has the advantage of being constant envelope, which (as discussed in Section 2.3.1.3) eases some implementation challenges arising from nonlinearity of the transmitter. Similarly to the amplitude modulation case, there is a tradeoff between SNR requirement and bandwidth efficiency of the modulation scheme.

2.3.2.3 Quadrature Amplitude Modulation

By applying both phase and amplitude modulation, arbitrary linear combinations of sine and cosine can be transmitted. One of the most useful implementations of that idea is known as quadrature amplitude modulation (QAM). In this modulation scheme, both sine and cosine are scaled independently by signed amplitude values.

Typically the case of interest is when sine and cosine share the same set of possible amplitude values, that the amplitude levels are evenly spaced, and there are a power of two levels in total (Equation 2.4, where k is the bits per symbol):

$$v(t) = I(t) \cos(\omega t) + Q(t) \sin(\omega t) \quad (2.4)$$

Where:

$$I(t) = \sum_{n=-\infty}^{\infty} [u(t - nT) - u(t - (n + 1)T)] a_I[n]$$

$$Q(t) = \sum_{n=-\infty}^{\infty} [u(t - nT) - u(t - (n + 1)T)] a_Q[n]$$

$$a_I[n], a_Q[n] \in \left\{ -1, -\frac{2^{k/2}-3}{2^{k/2}-1}, -\frac{2^{k/2}-5}{2^{k/2}-1}, \dots, -\frac{1}{2^{k/2}-1}, \frac{1}{2^{k/2}-1}, \dots, \frac{2^{k/2}-5}{2^{k/2}-1}, \frac{2^{k/2}-3}{2^{k/2}-1}, 1 \right\}$$

In a completely linear system, it can be shown that for the same data rate, power level, and bandwidth efficiency, QAM can provide the same bit error rate at lower SNR than the equivalent PSK modulation [22]. On the other hand, if the transmitter has strong nonlinearity, the variable amplitude levels of the QAM signal will be distorted. This lowers the effective SNR at the receiver.

Due to the complexity in implementing PSK vs QAM at mm-Wave frequencies, QAM is used far more frequently. One exception is 4-QAM, which is also equivalent to 4-PSK (otherwise known as quadrature phase shift keying - QPSK).

2.4 Phased Array Techniques

The main challenge in any mm-Wave system design is meeting link budget requirements, in light of the large free space path loss at high frequencies. One way of efficiently addressing the link budget problem is by leveraging phased-array techniques to reduce the total transceiver DC power [23]. For an N -element phased array, transmitter EIRP is increased by a factor of N^2 , since electric and magnetic fields, not power, are summed, and power density is proportional to $\vec{E} \times \vec{H}$. Due to reciprocity, for the receiver array there will also be a benefit of N^2 in conversion gain.

Receiver SNR will increase as well, but only proportional to N : since the noise in each receiver element is uncorrelated¹, the total noise at the output will increase proportional to N , resulting in an SNR increase of $N^2/N = N$. Because these system-level performance metrics are improved in a phased-array compared to the single-element case, it is possible to reduce performance (and correspondingly, DC power) while still meeting system requirements derived from the link budget.

Consider an RF power amplifier output stage, designed to operate close to saturation, and optimized to drive a load impedance of Z_0 . If the device sizes in the power amplifier are reduced by half, the power amplifier should be able to achieve the same efficiency at saturation while driving a load impedance of $2Z_0$, and delivering half of the power to that load. This scaled-down power amplifier can be used along with a matching network with an impedance transformation ratio of 2 (or if a matching network is already present, modifying its impedance transformation ratio) to drive the original load impedance of Z_0 , while delivering half of the power at the same efficiency. With a real matching network, there will be some additional losses, so the efficiency and output power will in practice be degraded somewhat. In a phased-array system, this strategy can be used to reduce DC power and per-element performance without sacrificing efficiency.

A similar scaling approach can be used on the receiver side. Consider an LNA designed for power and noise matching to an impedance Z_0 . Because the current density for minimum noise figure is largely invariant of emitter length [25] (or similarly, transistor width in CMOS technologies), the LNA device sizes can be reduced by half, resulting in an LNA with the same NF_{min} matched to an impedance of $2Z_0$. As in the transmitter case, a matching network can be used to match the LNA back to the original Z_0 input impedance. The new LNA has half of the DC power consumption, and slightly higher noise figure due to the added matching network losses. Of course, due to matching network complexity, added losses, and the bandwidth narrowing effect of high-Q matching networks, it is not possible to continue this scaling arbitrarily. Architecture or circuit topology changes must then be used to reduce power consumption further.

¹In general, correlated noise across receiver elements will mitigate some of this SNR improvement (refer to [24] for a more detailed discussion). However, if the receiver elements are well isolated, and the noise of the combiner is small relative to the noise from the individual receivers, the factor of N scaling will hold.

Chapter 3

FMCW Radar Phased-Array Transceiver Design

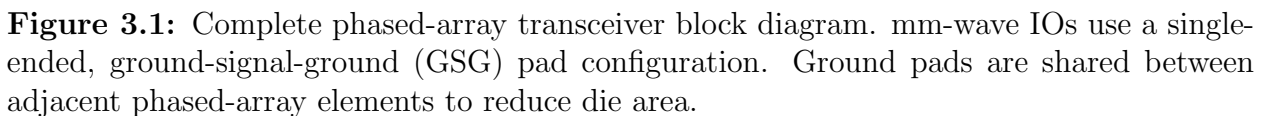
Highly integrated millimeter-wave transceivers, enabled by advances in CMOS and SiGe BiCMOS process technology over the last decade, have found what is seemingly a perfect niche in automotive radar. With many GHz of absolute bandwidth available, and a compact antenna size due to the small wavelength at mm-wave, the W band matches up well with the requirements for adaptive cruise control and similar technologies [26]. The development of mature, low-cost SiGe and CMOS technologies with f_t and f_{max} of 150GHz and beyond has brought down the cost of such driver-assist technology and with it, widespread adoption.

More recently, mm-wave radar has also received increasing attention for short-range applications such as gesture recognition, occupancy detection, and remote heart-rate monitoring [27][28][29][30][31]. However, existing mm-wave radar solutions intended for automotive use are power-hungry and often bulky. These drawbacks pose a problem for mobile, power-constrained applications. Towards this goal, in this work, a compact antenna-in-package FMCW radar phased array solution at 94GHz with record-low per-element power consumption is proposed and demonstrated.

Although some promising progress has been made on gesture recognition radar at 60GHz[30], the higher frequency at 94GHz allows the possibility for larger sweep bandwidths (which improves depth resolution of the radar) and smaller antenna sizes. However, the higher frequency also presents a challenge from the circuit design point of view, which has a negative impact on efficiency and achievable SNR.

3.1 Proposed System Architecture

Several works have demonstrated state-of-the-art synthesizers with integrated frequency modulation using a digital-PLL-based architecture [19][32][33][34]. The focus of this work is on energy-efficient array implementation and FMCW radar demonstration, so an external synthesizer is used to generate the frequency-modulated LO waveform. The chip includes



Because it is critical to minimize TX-RX leakage for an FMCW radar, an architecture with separate TX and RX antennas was selected. Although it is possible to use an integrated isolating coupler to achieve some degree of isolation [35][36], even an ideal coupler will have 3dB insertion loss due to the power-splitting nature of the coupler.

For phase shifting, LO path phase shifters are used [37]. LO path phase shifting is attractive here because it removes phase shifter degradations such as nonlinearity and noise from the signal path. This increases efficiency because amplifiers on the LO path can be designed to operate close to compression, as the LO signal is constant envelope. Baseband phase shifting is also attractive from a power consumption perspective, but requires two mixers for complex downconversion. This is not necessary for a linear FMCW system; since the TX and RX frequencies are always slightly offset, the mixer strictly speaking is not truly

operating as a direct conversion mixer, and therefore power can be saved by only using a single mixer.

3.2 Transmitter

To meet link budget requirements, a power amplifier was designed to provide approximately +9dBm of output power to a single-ended 50Ω antenna port. A single-ended antenna interface was selected to minimize mm-Wave IO count, which keeps the die area small and relaxes routing constraints within the antenna module.

The main PA gain stage is based on a cascode amplifier. Because of the high output impedance of the cascode, it is hard to achieve a good power-added efficiency (PAE) using a cascode output stage – the load-line impedance is significantly different from the small-signal impedance. However, the per-stage gain is still quite high relative to a simple common emitter amplifier, which has high PAE, but low gain when driven close to saturation. The amplifier core uses a differential topology to reduce sensitivity to modeling errors associated with the impedance seen at the cascode node. The bases of the cascode devices in a differential pair can be shorted directly together using local routing only, and therefore present a virtual short circuit in differential mode. In common mode, gain is not a concern, so low-Q bypass capacitors are used to prevent any common-mode stability problems associated with the impedance at the base of the cascode device.

To get both high gain and moderate efficiency, two cascode driver stages are used to drive a common-source output stage (see Fig. 5.30). A minimum supply voltage of 1.8V is needed to get good cascode performance, but is slightly above the open-base V_{CE} breakdown voltage

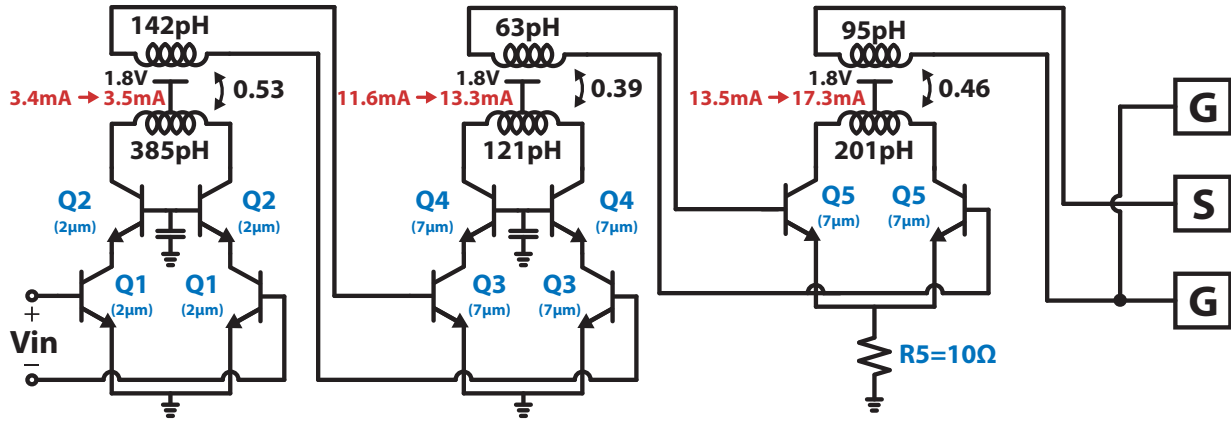


Figure 3.2: Three-stage power amplifier schematic. R_5 is chosen to result in an emitter voltage of about 100mV under small-signal bias conditions. The annotated DC currents correspond to the operating points in small-signal (left of arrow) and saturated large-signal (right of arrow) conditions.

of a single device. For the non-cascoded output stage, a moderate impedance is provided to the base via the bias network to extend the V_{CE} breakdown range beyond the open-base limit of BV_{CEO} and allow operation from a single 1.8V PA supply [38]. For additional robustness to V_{CE} breakdown with the 1.8V supply voltage, a small series emitter degeneration resistor is added at the tail. This helps improve reliability issues and has no impact on gain since it appears only in common-mode.

It is difficult to power match at the output of the cascode due to the high real part of the output impedance, on the order of $1k\Omega$. Additionally, the real part of the input impedance of the cascode amplifiers is fairly small (tens of ohms), leading to a large required transformation ratio. The available area for matching networks is constrained due to the phased-array element pitch and the internal power-supply flip-chip bumps in between the phased-array lanes, making it impossible to fit a transmission-line based matching network into the small area available for the PA. So, for a moderate impedance transformation ratio given the area constraints, coupling between PA gain stages is best achieved using 2:1 transformers with moderate to low coupling factors.

Because of the low coupling factor, the effective turns ratio is slightly less than 2:1 in practice, reducing the impedance transformation ratio. However, the leakage inductance of the transformer primary can be used to increase the impedance transformation ratio of the transformer, by treating it as an additional inductance in series with the transformer, which acts to increase the impedance seen at the ports of the primary transformer. The output stage does not use neutralization because of the extra capacitive load it would present to the output balun, which is also a 2:1 transformer. The output balun also provides ESD protection to the signal pad, as at low frequencies it provides a low impedance path to ground for the signal pads through the center tap of the secondary.

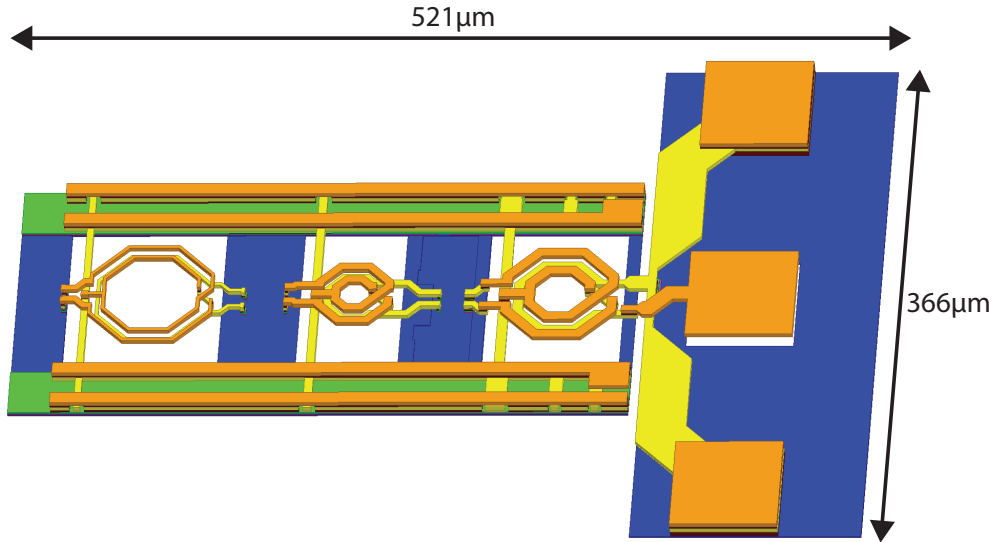


Figure 3.3: Full 3D EM model of PA interstage and output transformers.

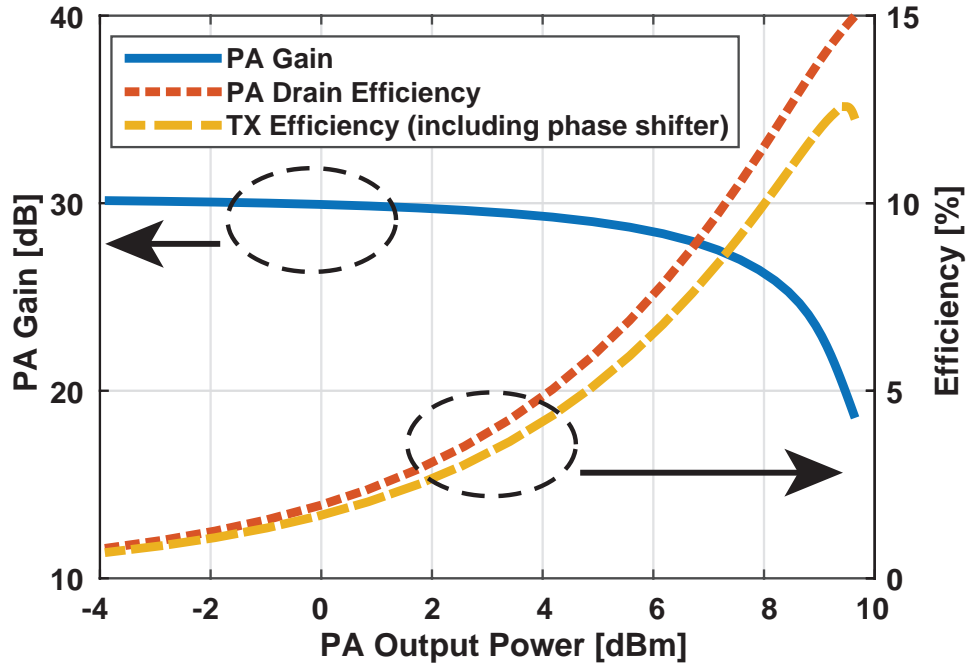


Figure 3.4: Simulated PA power gain and power-added efficiency at 94GHz, plotted versus output power.

A 3D HFSS model of the full PA, including the output GSG pads and the three stages of transformers, is shown in Figure 3.3. The individual transformers were first designed separately using HFSS. As a final verification, all three transformers were simulated together along with the output pads. When incorporated back into circuit-level simulations, this PA-scale EM model predicted nearly identical performance when compared to the simulations using separately modeled transformers.

At the intended carrier frequency of 94GHz, simulations show a small-signal power gain of 31dB, and a peak PAE of 15% at an output power of 9.6dBm (Fig. 3.4). If the DC power of the phase shifter driving the PA is included in the efficiency calculation, the PAE of the full chain drops to 12%. Because of the high small-signal gain of the cascode amplifiers, the gain starts to compress well before the output stage is fully saturated. As a result, the peak PAE is reached well beyond the P_{-1dB} of the amplifier.

Large-signal simulations show that the saturated output power and efficiency are relatively flat across frequency (Fig.3.5) with P_{sat} above 10dBm from 85GHz to 98GHz, and the peak PAE of the PA is nearly a constant 15% from 85GHz to 95GHz. The peak small-signal gain is 30.8dB at 93GHz, and the small-signal 3dB bandwidth is 12GHz (from 86GHz to 98GHz).

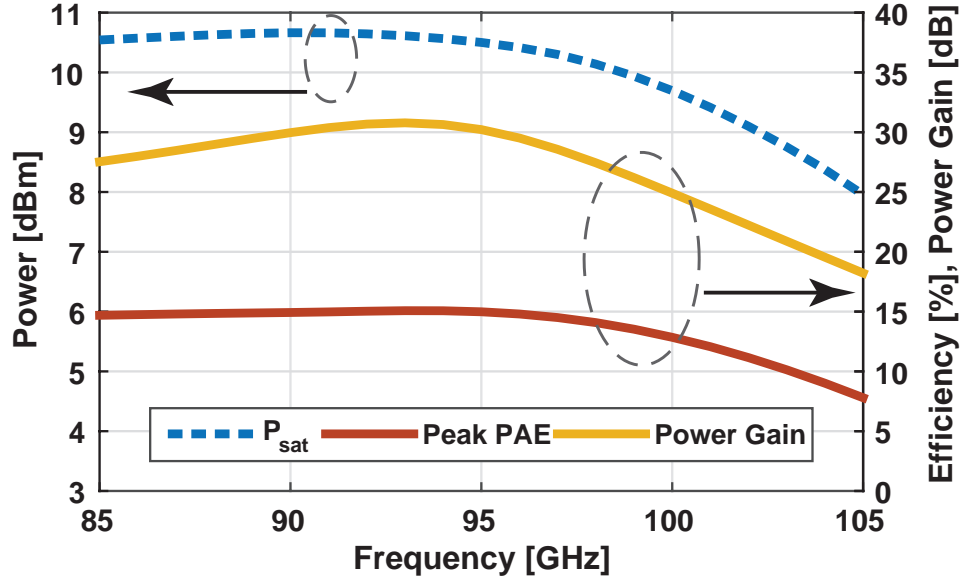


Figure 3.5: Simulated PA saturated output power, power gain, and peak power-added efficiency vs. frequency.

3.3 Phase Shifter

Passive reflection type phase shifters are an attractive option for low-power design, as they do not consume any DC power and multiple stages can be cascaded to achieve the desired phase-steering range [39][40]. However, this also implies the insertion loss trades off with phase shift range. To overcome this loss, an additional amplifier stage is needed, which consumes DC power.

Instead, in this design, a Cartesian architecture is used to ensure a full 360 degrees of phase shift. Weighted combinations of the I and Q LO waveforms are current-combined at the phase shifter output, meaning that the phase should be the same regardless of process variations (provided the input quadrature matching is sufficiently accurate). The VGA functionality is achieved by current-steering using the cascode devices (Q5 through Q8) rather than the g_m devices of the Gilbert cell. This ensures a more constant input impedance vs code. As the hybrid is single-ended, not differential, a balun stage is needed to drive the phase shifter LO inputs. A single-ended cascode amplifier with transformer load provides single-ended to differential conversion, and further isolates the hybrid from any variations in phase-shifter input impedance.

3.3.0.1 Quadrature Coupler

The quadrature coupler is designed based on the approach in [41] to provide good phase accuracy over a broad bandwidth. It consists of three high-impedance transmission lines



3.3.0.2 Phase Interpolator Simulated Results

The phase shifter and transmitter power amplifier were simulated together to characterize the effects of gain compression on the phase shifter resolution. The phase shifter control voltages are driven by differential voltages on the I and Q input terminals. Each I and Q voltage is controlled by a 4-bit DAC, so there are 256 possible codes that can be used in the phase shifter. In practice, only the largest amplitude codes will be used. The constellation of available gain and phase combinations at the PA output is shown in Figure 3.9a. Gain and phase errors are then computed for each possible desired phase angle. The worst-case phase error is about 4.5° , and the worst-case amplitude error 0.55dB, which both occur when both the differential I and Q voltages are at their largest (45° phase shift). This is because at the highest/lowest DAC codes, the the differential pairs (Q5-Q8 in Figure 3.6) are no longer in the linear range.

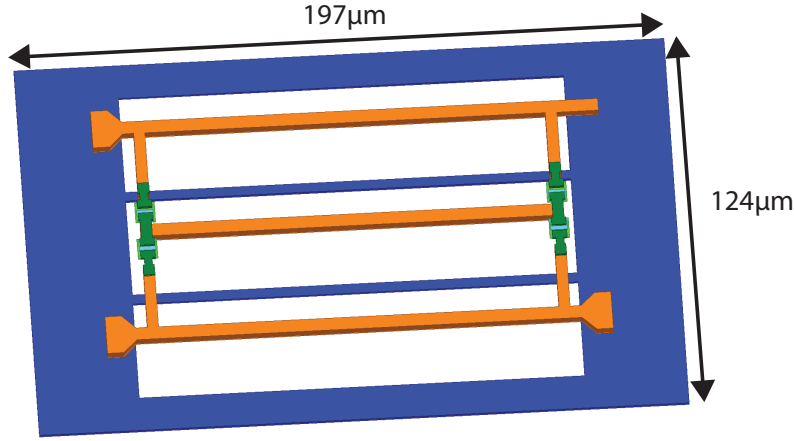


Figure 3.7: Quadrature hybrid HFSS model. The dimensions of the simulated region are $197\mu\text{m}$ by $124\mu\text{m}$.

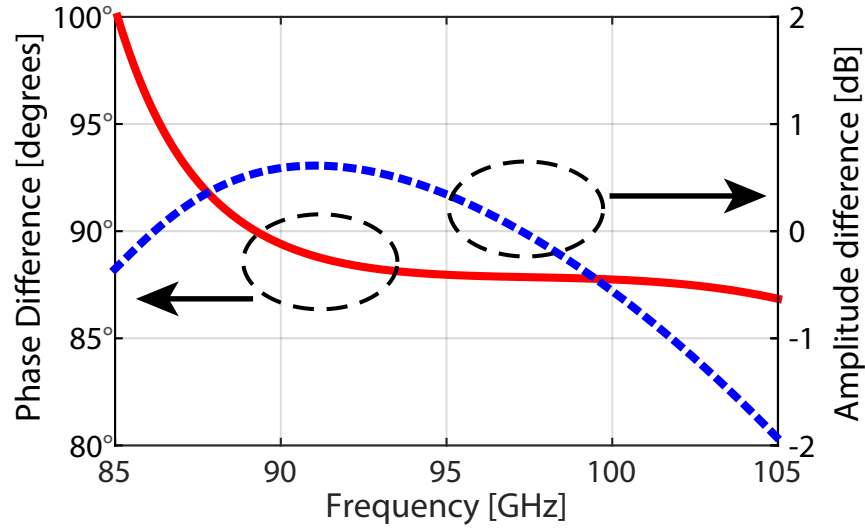


Figure 3.8: Quadrature hybrid simulated IQ mismatch.

3.4 LO Distribution

The LO generation circuitry includes an integrated VCO, frequency multiplier, frequency dividers, and LO buffers. The integrated VCO is designed to operate at half of the RF frequency, and is buffered and sent to a frequency doubler to generate the RF carrier waveform. The frequency-doubled LO waveform is buffered and distributed to the phased-array elements.

There is a routing penalty in distributing the LO signal after the frequency multiplier,

rather than before, since the absolute losses per millimeter are worse at higher frequency. However, it can be more efficient overall to accept those losses, since the alternative is to have equally many power-hungry frequency multipliers as there are phased-array elements. Since frequency multipliers typically have poor (if any) conversion gain, low output power, and low efficiency, it makes sense to have as few as possible, as long as a moderately efficient LO buffer can be placed afterwards to overcome the LO routing and distribution losses. This approach does not work as the multiplier output frequency approaches the f_{max} of the technology, since it becomes impossible to get any more power gain, but this is not the case at 94GHz.

The integrated frequency divider chain has five cascaded divide-by-two stages for a 32x total division, resulting a nominal output frequency of 1.46875GHz if the VCO is operated at 47GHz. A discrete fractional-N PLL chip is used along with an active loop filter on the test PCB to complete the LO chain externally.

3.4.0.1 47GHz VCO

The VCO core is a capacitively cross-coupled NPN pair. Rather than using a tail current source to bias the VCO, a series tail resistor is used instead. This greatly improves the simulated $1/f^3$ phase noise corner.

3.4.0.2 Frequency Dividers

The first few stages are bipolar-based static CML dividers for robust high-speed performance. No inductive peaking was used for the high-speed bipolar dividers. The last two stages are CML dividers that use 130nm CMOS devices and consume much less power. A final NPN buffer amplifier is used to drive the LO signal off chip.

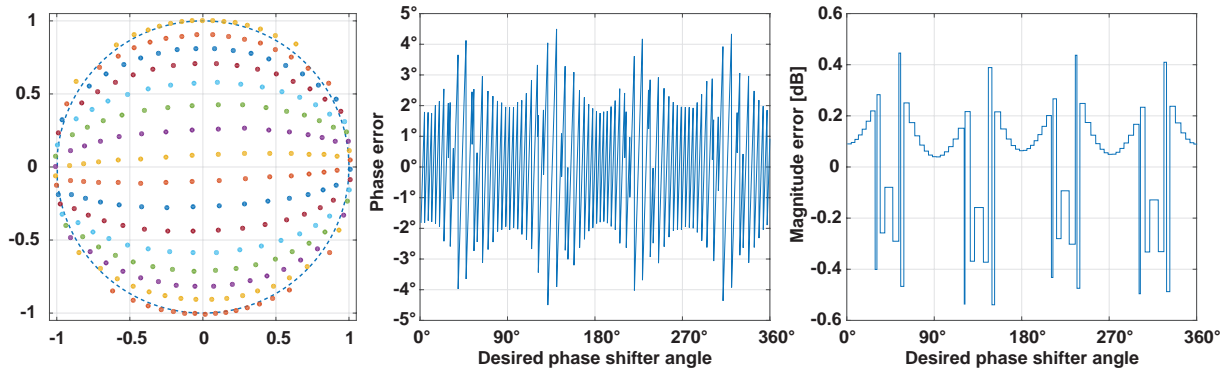


Figure 3.9: Phase shifter constellation points, showing combinations of phase and amplitude settings, and the phase and amplitude error for a desired phase angle.

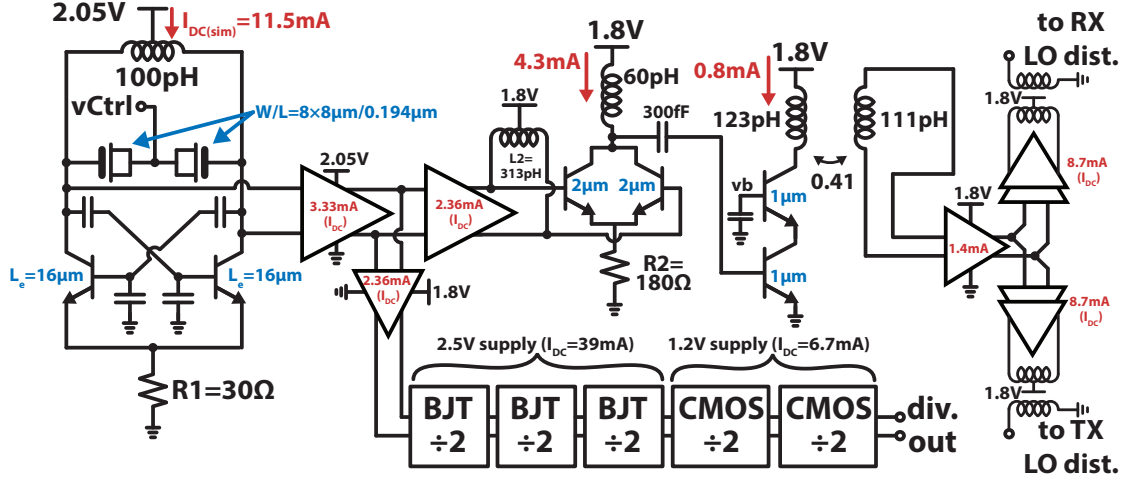


Figure 3.10: Schematic of LO generation and distribution circuitry.

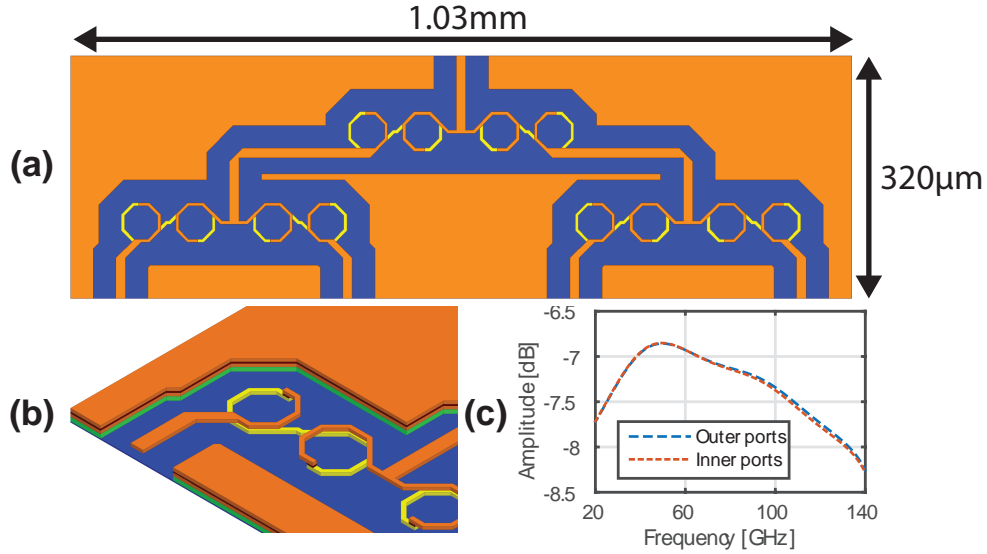


Figure 3.11: LO distribution network and power divider, based on lumped-element artificial transmission line. Each artificial quarter-wavelength line consists of a 1.5 turn inductor in series with a 1.25 turn inductor, which interfaces with a 50ohm transmission line.

3.4.0.3 Frequency Doubler

The frequency doubler uses a push-push topology with an inductive load [42]. A common-mode tail resistor is used instead of a current source, as simulations showed it provided slight enhancement of the 2nd harmonic output current. Simulations also indicated common-mode stability problems when using a tail current source in the frequency doubler, related to the high-Q capacitance that it presents at the tail node. The common-mode stability issues are ameliorated by using the tail resistor.

3.4.0.4 LO Distribution Amplifiers

After the frequency doubler, the 94GHz LO waveform must be distributed to the TX and RX phased array elements. Separate distribution networks are used for the TX and RX elements, so that the power levels can be separately controlled. Even a lossless LO power-splitting network will inherently represent a reduction in power level, since the input power is divided equally to all output paths. If LO buffers are used after the LO splitting network, the efficiency of any LO buffers will be quite poor, since the signal level will be very small due to the splitting loss. To avoid suffering that efficiency penalty, moderate-power LO buffers are used drive the input of the power splitting network. This results in the same power levels at the output, but a higher efficiency and reduced overall DC power.

The LO distribution amplifiers were designed by re-using the first PA gain stage (for both amplifier stages of the LO buffers) and redesigning the interstage matching networks. The output matching network is a transformer balun, to drive the single-ended LO distribution network.

3.4.0.5 Lumped-Element Power Divider Network

To simplify the design of the divider network, two nested 1:2 power splitters are used. The 1:2 splitter uses quarter wavelength $Z_0 = 70.7\Omega$ lines to enable use in a cascade; when terminated with 50Ω loads, the input impedance of the splitter is also 50Ω . Typically, a Wilkinson power splitter is used as a 1:2 power splitter at millimeter wave [43]. The Wilkinson splitter is an isolating power divider, which will prevent potential crosstalk between elements. A differential-mode termination resistor is needed to provide this isolation, but requires that the outputs of the Wilkinson are physically close. Because the inputs of the phased-array channels are spaced apart by the array element pitch ($300\mu\text{m}$), additional routing is required to distribute the Wilkinson outputs to the phased-array elements (Figure 3.12a), which requires additional area and increases losses.

If a non-isolating network is used, it can also provide the required routing to the input of each array element (Figure 3.12b). Since the LO distribution network is terminated in the passive quadrature hybrid load whose input impedance is constant vs phase code, there is no opportunity for crosstalk between elements even though a non-isolating splitter is used.

However, the splitter must fit within the array element pitch of $300\mu\text{m}$, but the length of a single quarter-wavelength line on-chip is approximately $400\mu\text{m}$. Clearly, a straight

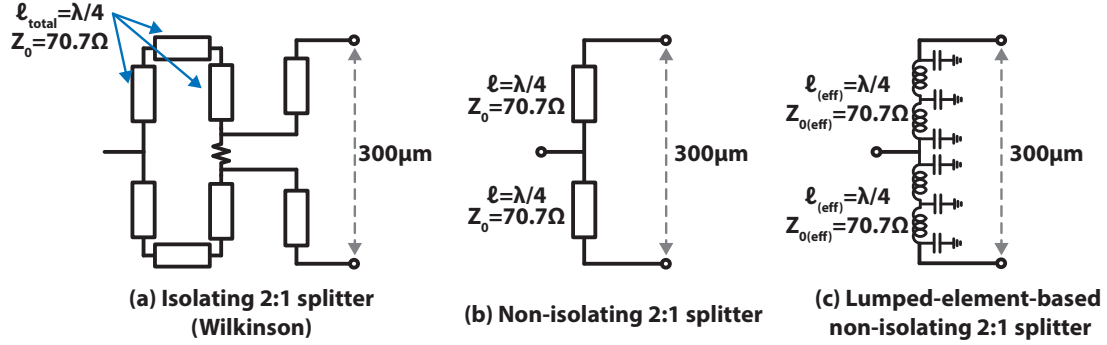


Figure 3.12: Power splitter topologies considered for LO distribution.

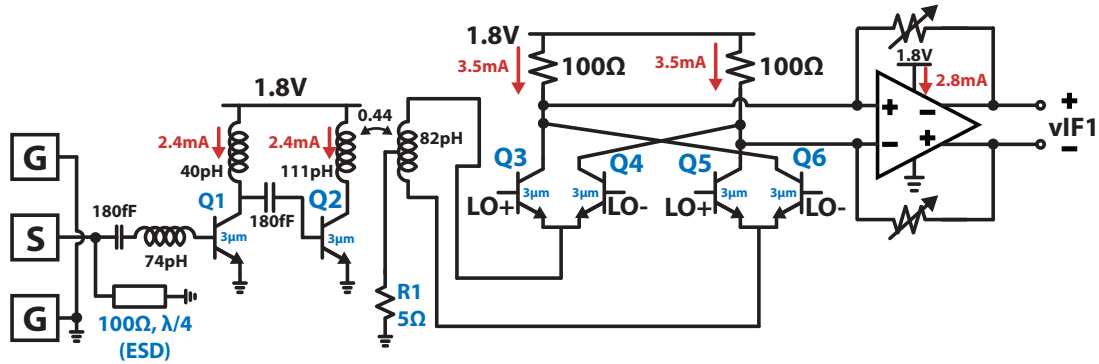


Figure 3.13: Receiver schematic.

transmission line will not be able to fit at this array element pitch - the transmission line would need to be meandered to fit. Instead of using a meandered transmission line, the splitter instead uses a lumped-element artificial transmission line to reduce the area of the network (Figure 3.12c).

The insertion loss for an ideal 1 to 4 splitter would be 6dB - for this design, EM simulations show excess insertion losses of 7.1dB to 7.4dB (1.1 to 1.4dB higher than an ideal power splitter) from 80-100GHz. The simulated amplitude mismatch is less than 0.04dB, with the outer ports receiving slightly more power than the inner ports, and less than 0.4° phase mismatch up to 100GHz.

3.5 Receiver

The number of LNA stages in the receiver was limited to keep DC power consumption low. Because of this, no inductive degeneration was used in the receiver input device (Q1 in Figure 3.13), as this would reduce the gain of the LNA and increase the noise contribution of the mixer to the overall receiver noise figure. Effectively, this means the input of the LNA is

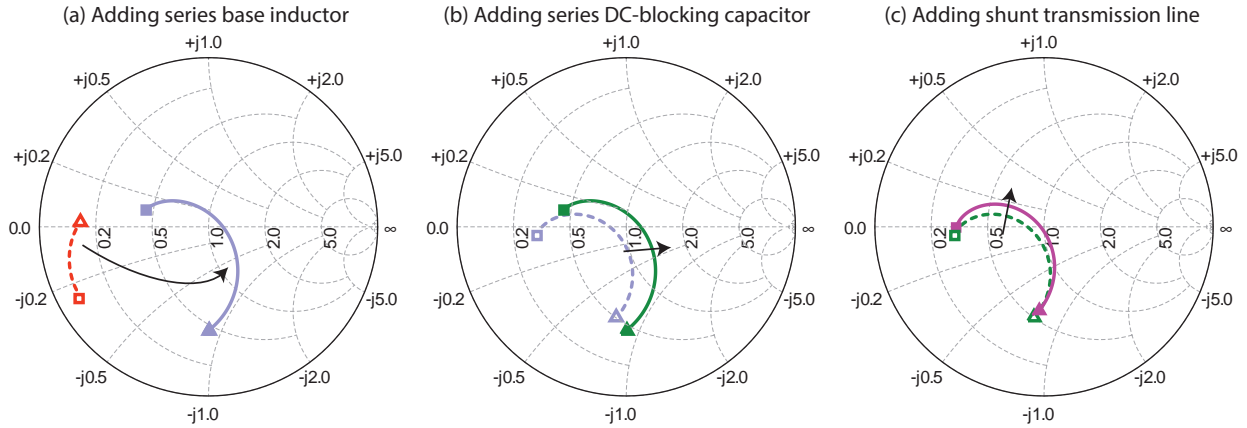


Figure 3.14: Input impedance locus of the LNA, at various stages of the matching network.

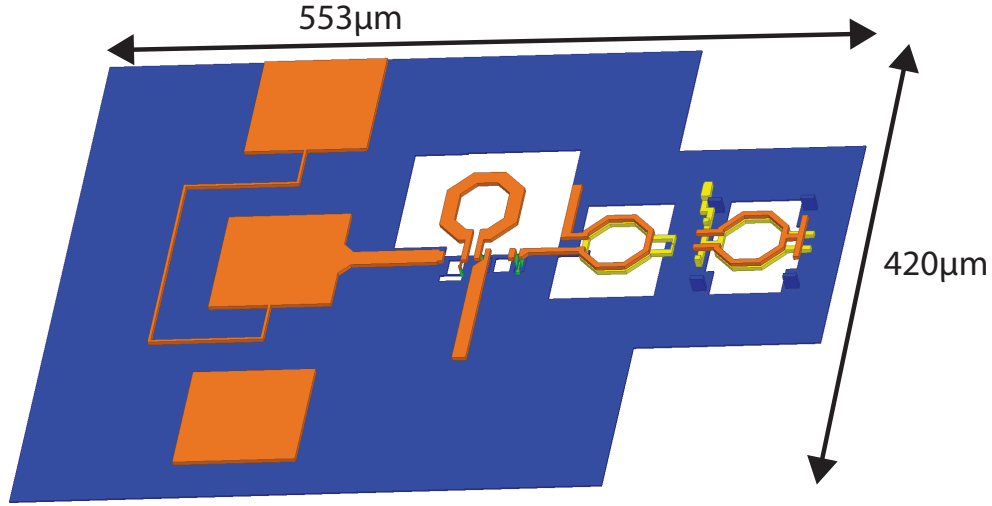
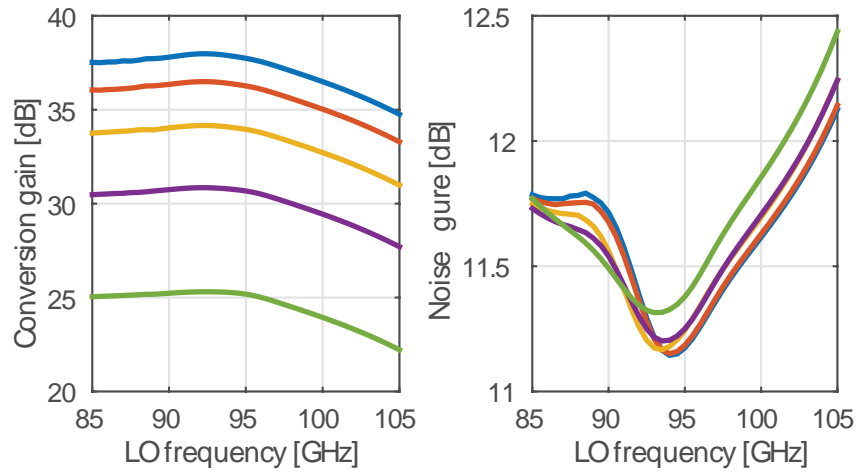
designed for a power match, rather than a noise match. This adds a small amount (0.3dB) to the overall noise figure but limits the mixer noise figure contribution (see Fig. 3.17b).

Since the current density for minimum noise figure is typically 6-10 times smaller than the current density for peak f_T [25], biasing for NF_{min} does have a gain penalty. As a compromise between gain and noise figure, the LNA input stage is biased at about half of the current density for peak f_T . This leads to an increase of 0.5dB above the estimated minimum noise figure of 3.2dB, but also allows the device to operate at nearly peak f_T .

The input matching network uses a series inductor, a DC blocking capacitor, and a quarter-wavelength transmission line shunted to ground. Nearly all of the impedance matching is provided via the series base inductor. Because of the series inductor's capacitance to ground, it acts more like a transmission line than a simple series inductor. So, on a Smith chart, this looks like a rotation, rather than moving on a line of constant resistance (Figure 3.14a). The DC-block capacitor is a small series impedance at RF and is only needed to separate the bias points at the signal pad and LNA input (Figure 3.14b). The shunt transmission line provides a low-impedance path from the pad to ground at low frequencies for ESD robustness, and contributes a small amount of shunt inductance at RF (Figure 3.14c).

The first LNA stage has an inductor load designed to resonate with its output capacitance, and is AC-coupled to the input of the second stage (Q2). The AC coupling capacitor between the first and second stages is implemented using the standard MIM capacitor offered in the process. Its value is large so that it contributes a negligible series reactance, reducing design sensitivity to the modeling accuracy of the capacitor. The second LNA stage connects to the differential mixer input using a transformer, which provides single-ended to differential conversion.

The mixer itself is a double-balanced switching core (Q3-Q6), with RF signals coupled in at the emitters, and LO signals at the bases. Because headroom is limited, the mixer is implemented as a pseudo-differential rather than differential pair, and there is no RF transconductor device in the mixer stack. The limited headroom also makes it difficult

**Figure 3.15:** Receiver HFSS model.**Figure 3.16:** Simulated receiver conversion gain and noise figure vs mixer LO frequency, for all 5 gain control settings.

to use a high-impedance active load that is sufficiently saturated. A resistive load is used instead, largely to set the bias point, and a transimpedance amplifier (TIA) is used to provide a low impedance at the mixer baseband output. Using this topology we are able to improve the voltage gain of the mixer, while the headroom is constrained. Additionally, the resistive load has improved noise performance over the active load, and suffers from less capacitive parasitics: the PMOS f_t is low relative to that of the NPN devices, and no high-speed PNP is available in this technology. Gain control is achieved by varying the feedback resistance in the baseband TIA via switched resistor segments. The TIA itself consists of a high-speed op-amp using SiGe NPN devices.

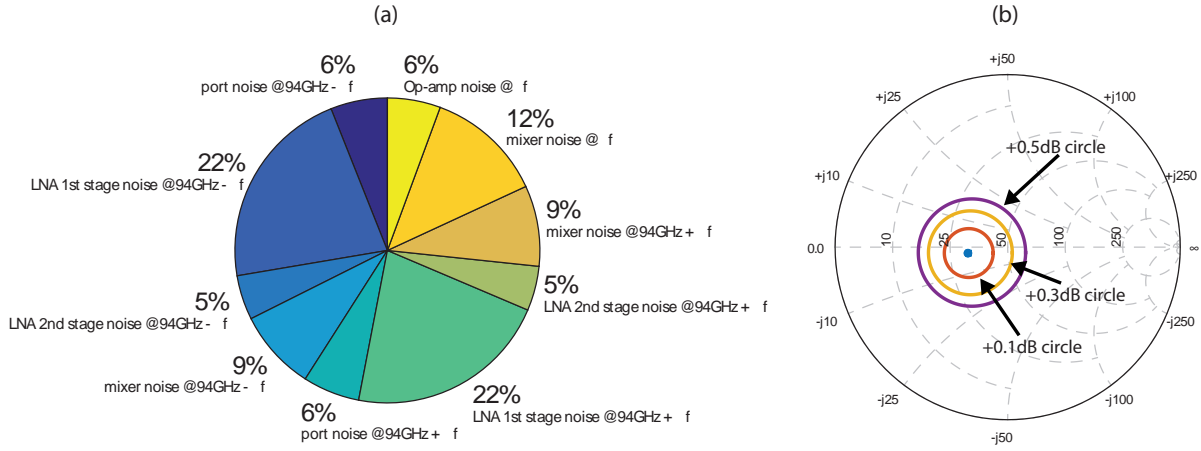


Figure 3.17: (a) Percentage contributions of different blocks to total output noise. The contributions are separated by sideband. (b) Noise circles

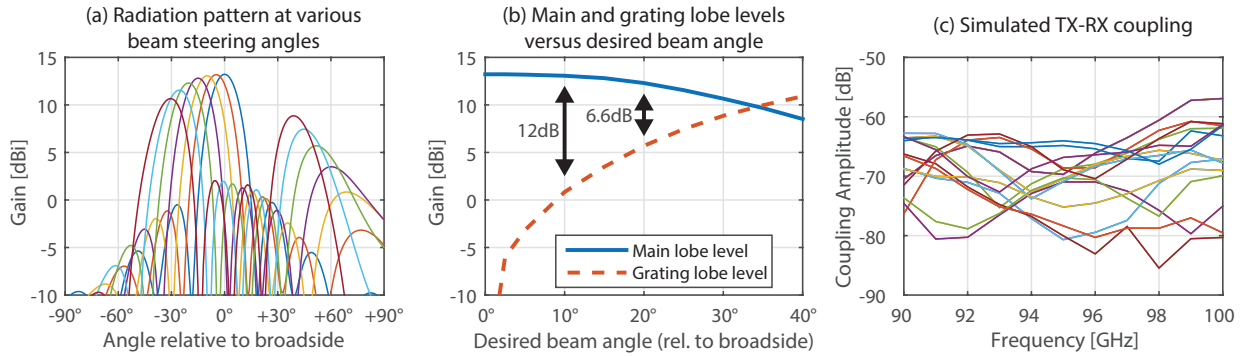


Figure 3.18: HFSS simulations of antenna module.

Simulations show a receiver conversion gain of 25-38dB and a noise figure ranging from 11.1 to 11.3dB (single-sideband) at 94GHz, depending on gain control settings. Since the gain control is implemented at the IF amplification stage, it has little impact on noise figure because of the front-end gain in the preceding stages. As can be seen in Figure 3.17a, approximately 54% of the total noise at the receiver output is due to the LNA, 30% from the mixer, 6% from the baseband amplification, and 12% from the reference noise of the input port.

3.6 Packaging

The die was flip-chip packaged using stud bumps onto an antenna module fabricated using an organic HDI substrate [44]. The antenna module is based on a standard BGA footprint

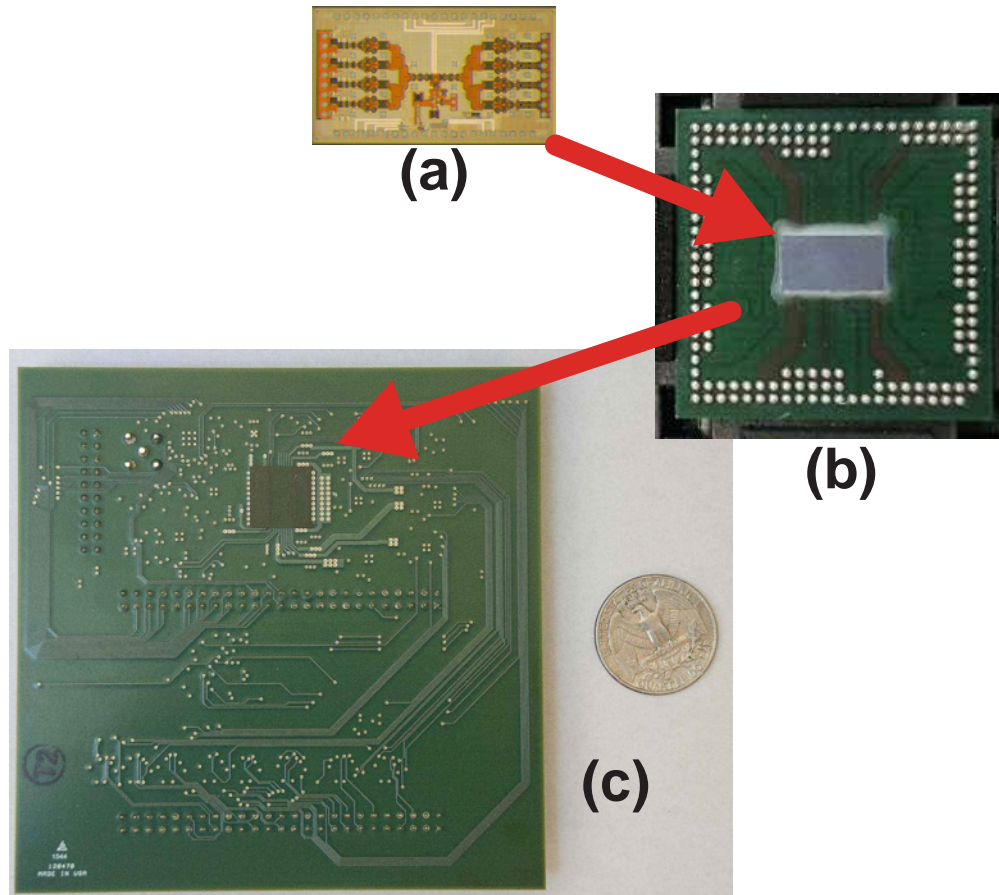


Figure 3.19: Levels of integration of radar IC. Bare die (a) is flip-chip chip packaged onto BGA antenna module (b), which is integrated onto a test PCB (c). The test PCB is 10.2cm x 10.2cm, sized to accommodate connectors to a separate FPGA board. If a microcontroller were used instead of an FPGA, the board area could be reduced to a much smaller size.

and low-frequency connections fan out to BGA balls for integration onto a larger test PCB (Figure 3.19). The substrate, 1.2cm by 1.2cm, contains two linear arrays of 4 patches each - one array for the transmitter elements and one array for the receiver elements. For high TX to RX isolation, the TX and RX antenna arrays are located on opposite sides of the module. The antennas themselves are aperture-coupled patch antennas with linear polarization. To reduce element-to-element coupling, the antenna spacing within each array is set to 0.8λ at 94GHz.

Because the antenna spacing is greater than half-wavelength, grating lobes will appear for large beam-steering angles. HFSS simulations of the antenna array predict that a beam-steering range of $\pm 27^\circ$ is possible for a grating lobe level of 3dB below the main lobe (Figure 3.18). At a simulated beam-steering angle of $\pm 34^\circ$, the grating lobe level is the same as that of the main lobe. The simulated 3dB beam width of the main beam is about 16° in the E

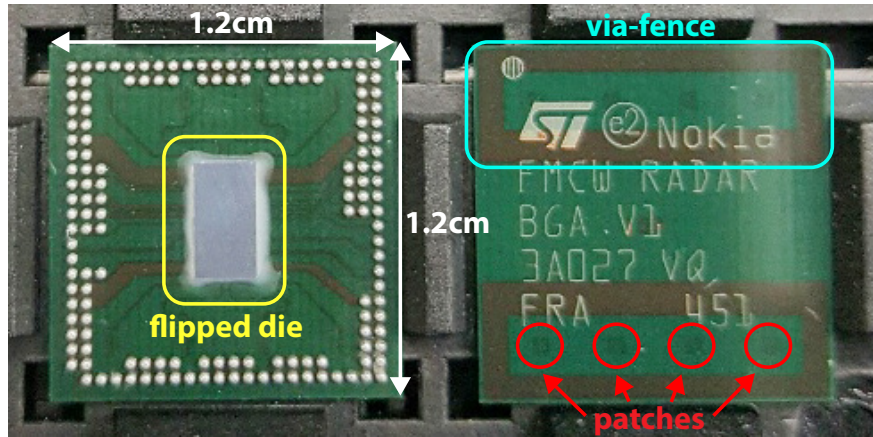


Figure 3.20: Photograph of 1.2cm \times 1.2cm BGA antenna module with die attached, chip (left) and antenna (right) sides.

plane (XY plane in Figure 4.10) and 90° in the H-plane (XZ plane in Figure 4.10). HFSS simulations of the package show a TX-RX isolation of 60dB from 90-98GHz (Figure 3.18c), which is sufficient given the input linearity simulations of the receiver.

Chapter 4

Radar IC Measurements

The 3.7x2.2mm chip was fabricated in a 130nm SiGe BiCMOS process (Fig. 4.1). The TX and RX arrays are on opposite sides of the chip. At the TX and RX arrays, the ground-signal-ground mm-Wave IO pads are placed vertically running up the sides of the chip. Adjacent ground pads are shared between elements to reduce die size. Shared LO generation circuitry is at the center of the chip and feeds in to the power divider networks, which in turn feed directly in to the phase shifters for the transmitters and receivers. The chip was tested both using mm-wave probes in a chip-on-board configuration, and using the packaged antenna module for wireless measurements.

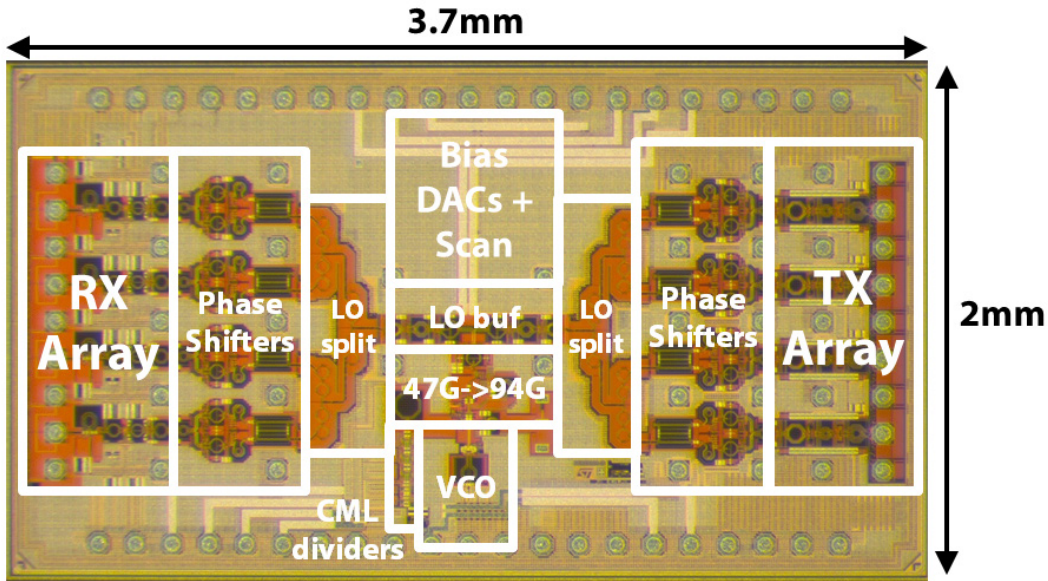


Figure 4.1: Die photograph of fabricated chip.

4.1 Probe Station Measurements

4.1.1 LO

The measured VCO tuning range is 11%, from 44 to 49GHz . The VCO center frequency is about 5% lower than simulated, but due to the large designed tuning range it still covers the desired center frequency of 47GHz. The PLL locking range is from 89-95GHz, slightly reduced from the VCO tuning range. It is limited by k_{VCO} at low tuning voltages, and by the output swing of the active loop filter at high tuning voltages. At 94GHz, the closed-loop phase noise is -76dBc/Hz at 1MHz offset, measured at the PA output .

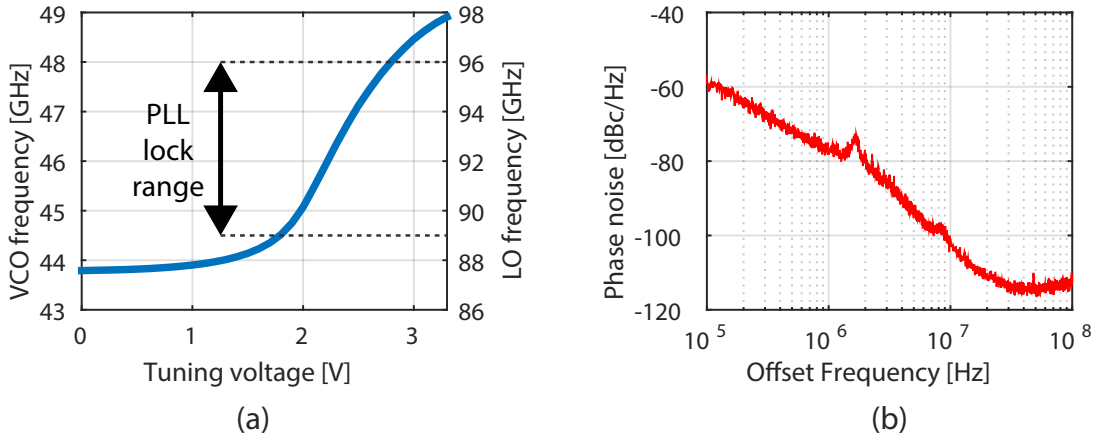


Figure 4.2: VCO tuning curve, and measured phase noise of PLL at 94GHz.

4.1.2 Transmitter

Probe station measurements were used to characterize the output power of individual PAs, along with a W-band power meter. At 94GHz, the output power varies over a range of about 0.5dB across elements, from 6.3dBm to 6.8dBm . The outer TX elements (1 and 4) have the highest output power, and the inner elements have lower output power. The output power was lower than the simulated $P_{sat} = +10.5$ dBm of the PA, as well as the simulated +9dBm output power with expected LO signal levels. The decrease in output power for inner TX elements suggests IR drop in the supply network may be a partial cause. Additionally, the output power decreases beyond 90GHz, whereas in simulations, it decreased after about 94GHz. Unfortunately, due to the integrated LO chain and limited VCO tuning range, it is not possible to measure outside of the 89-95GHz frequency range, so it cannot be verified if the output power increases further at lower frequencies. The decreased output power, and trend of power vs PA element, was consistently observed across multiple chips. The measured PA output impedance matches closely with simulation (Fig. 4.4b). The average output power of 6.5dBm results in a transmitter efficiency of 6.3% ($P_{TX,RF}/P_{TX,DC}$)

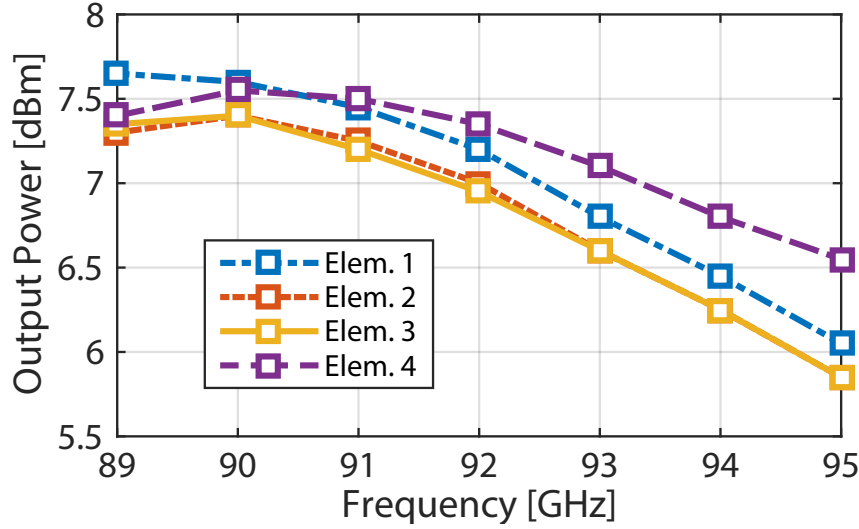


Figure 4.3: Measured transmitter output power vs frequency.

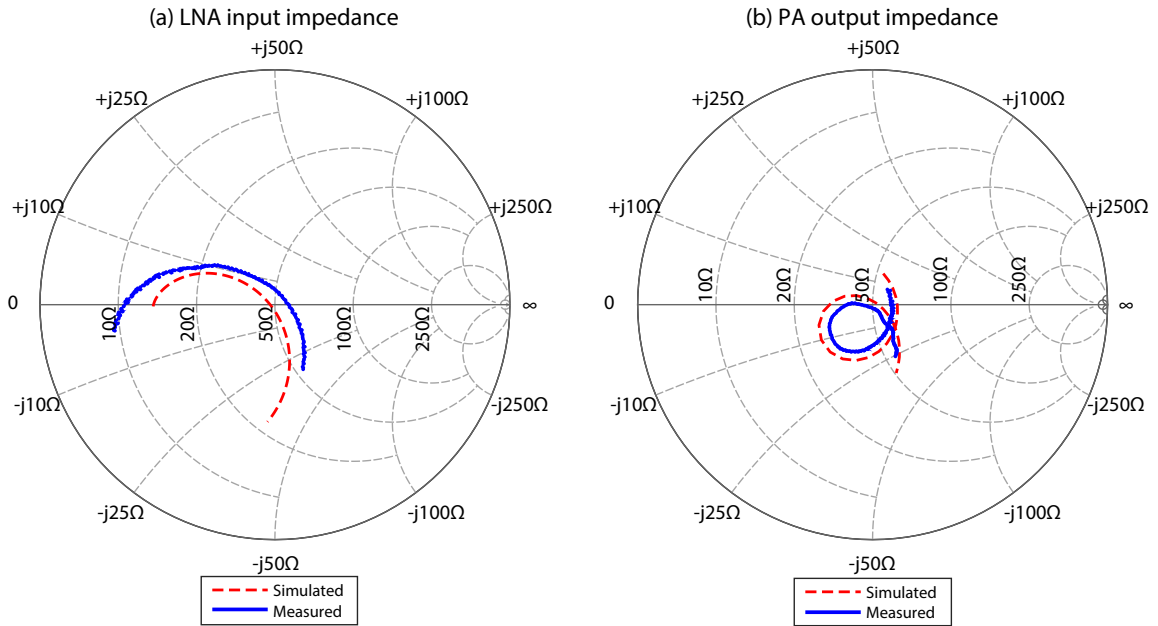


Figure 4.4: The measured LNA input impedance (a) and PA output impedance (b) showed good agreement with simulation.

4.1.3 Receiver

The receiver noise figure was measured with the hot/cold method, using an Agilent N8974A noise figure meter and W-band WR-10 noise source, connected to the receiver input via a W-band GSG probe (Fig. 4.5). In the RX case, element-to-element mismatch in noise figure

is quite small. We do not see a consistent trend of increased noise figure in the interior elements, which makes sense because noise figure should not be as sensitive to IR drop. We also see that the RX noise figure is improving vs frequency, unlike in the TX case where the best performance was at lower frequency. The measured 12.5dB (SSB) noise figure at 94GHz is higher than the simulated 11.1dB (SSB) at 94GHz. Below 94GHz, the noise figure increases quite drastically. This is potentially due to insufficient LO power at the RX mixer, as the LO distribution network is tuned slightly high due to EM modeling error and its output power drops off below 93GHz. The measured LNA impedance matches somewhat closely with simulation, although detuned slightly (Fig. 4.4a). The measured P_{1dB} of the receiver is -19dBm at 94GHz.

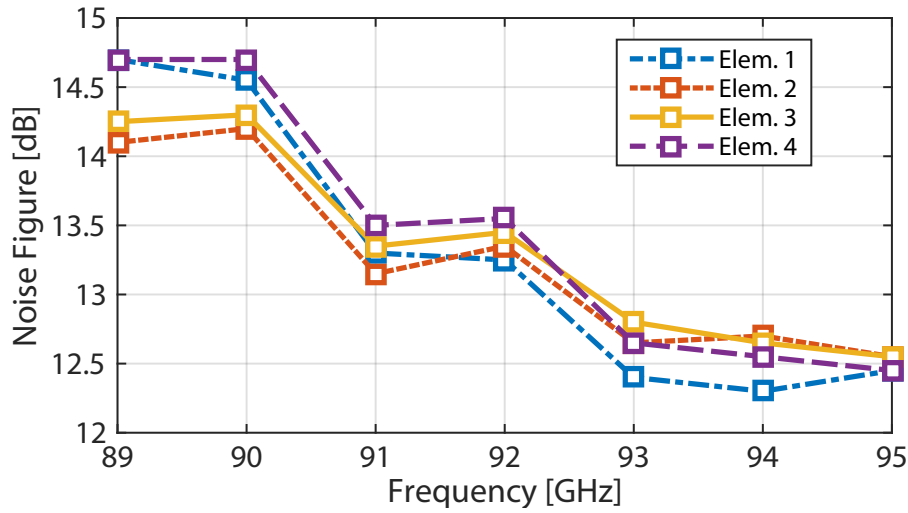


Figure 4.5: Measured receiver single-sideband noise figure (probe station) vs frequency.

4.2 Packaged Measurements

4.2.1 Array Characterization

Initial measurements of the phased array were performed at broadside, with zero phase shift between elements, to verify that transmitter spatial combining works as expected. EIRP measurements of the transmit array at broadside (Fig 4.6) indicate an EIRP of approximately +22dBm at 94GHz. The measured EIRP roughly increases by 6dB when the number of enabled PAs is doubled, which is consistent with the theoretical N^2 factor for in-phase spatial combining.

To further demonstrate phased array operation, the TX radiation pattern was measured at various beam steering angles. First, the TX radiation pattern was characterized in the lab at BWRC, by using a W-band power meter head (connected to a horn antenna) manually

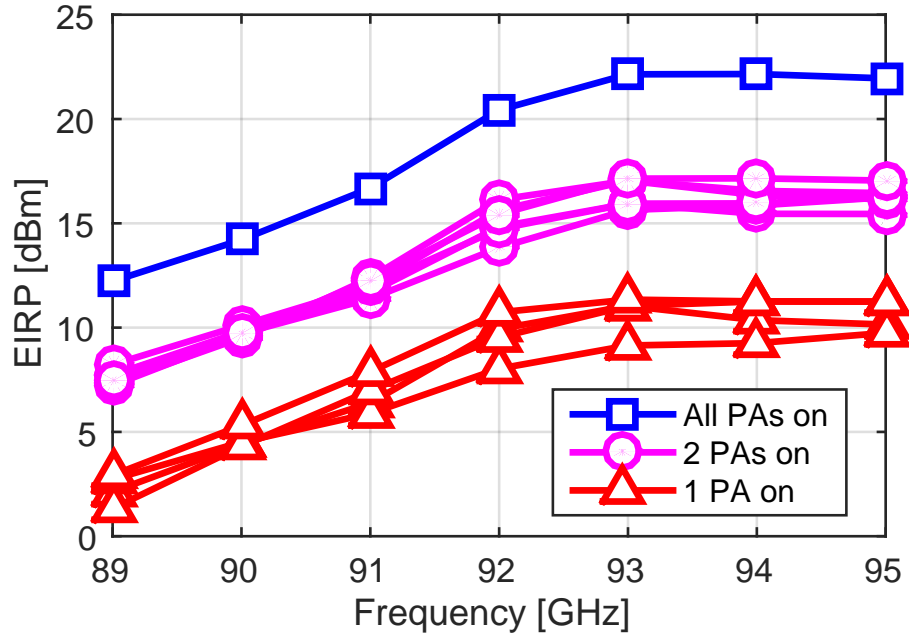


Figure 4.6: EIRP at broadside vs frequency, for 1, 2, and 4 PAs enabled. All possible combinations of the 1 and 2 PA cases are plotted.

placed at different angles relative to the broadside of the array. Because the measurement was done manually, the measurements points are in 5° increments, over a range of -60° to 60° from the broadside of the array (Figure 4.7).

The RX conversion gain pattern was also characterized using a similar approach, but instead using a 94GHz source and horn antenna placed at various angles, instead of the power meter. Similarly to the case of the TX EIRP, the conversion gain roughly increases by 6dB when the number of enabled LNAs is doubled.

However, for both the RX and TX cases, this improvised setup led to very slow measurements, required manual intervention to move the power meter or source from angle to angle, and was only capable of doing measurements along a single cut plane.

An indirect phase shifter characterization was completed based on beamforming and power-combining measurements at different phase shifter codes (Fig. 4.9). Because the PA is not fully driven into saturation, amplitude error is not suppressed. The measured peak phase and amplitude error are 9 degrees and 1.4dB, respectively.

The TX radiation pattern was also characterized at University of Nice Sophia-Antipolis (UNS), using the measurement setup first described in [45] and extended to 90-140GHz in [46] (see Fig. 4.10a). The digitally controllable arm can move along both ϕ and θ angles, so it is possible to capture nearly a full hemisphere of the radiation pattern with single-degree-level steps. An F-band subharmonic mixer is placed on the end of the moveable arm and the IF output connected to a spectrum analyzer. The power level is measured by observing the

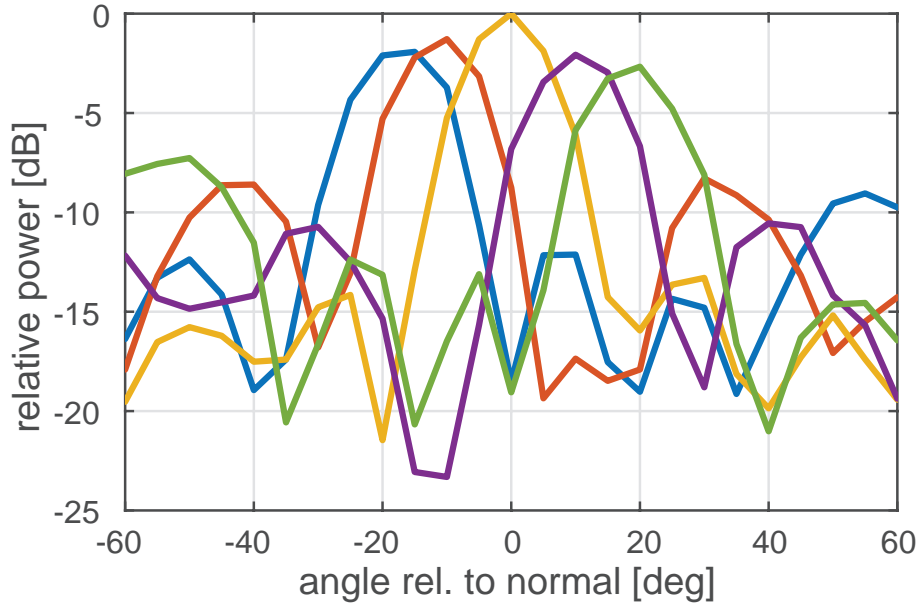


Figure 4.7: Beam steering of transmitter at 94GHz, characterized manually at BWRC.

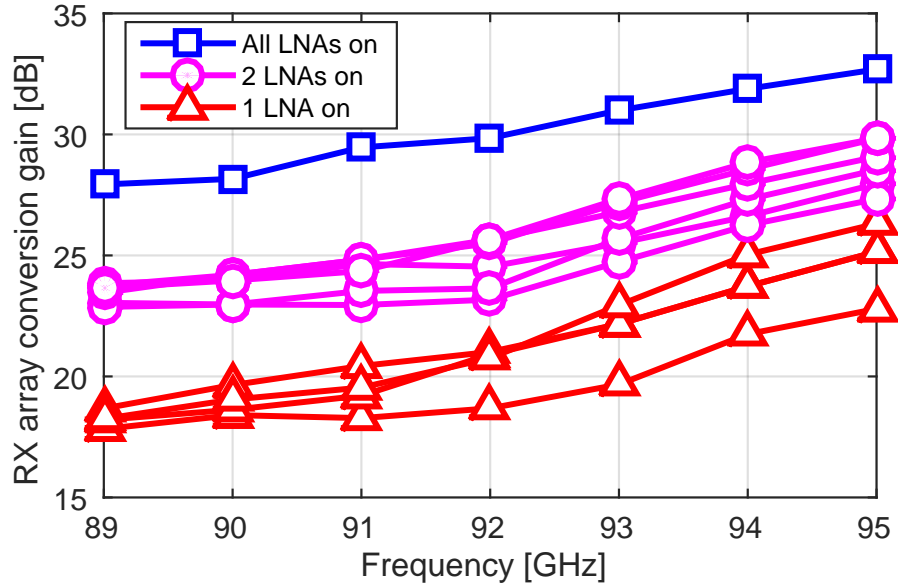


Figure 4.8: RX array conversion gain vs frequency, for 1, 2, and 4 LNAs enabled. All possible combinations of the 1 and 2 LNA cases are plotted.

peak level of the downconverted signal level on the spectrum analyzer. Due to the large size of the F-band signal source, it was not possible to mount the signal source on the moveable arm to measure the RX beam steering pattern. However, the results should be similar as

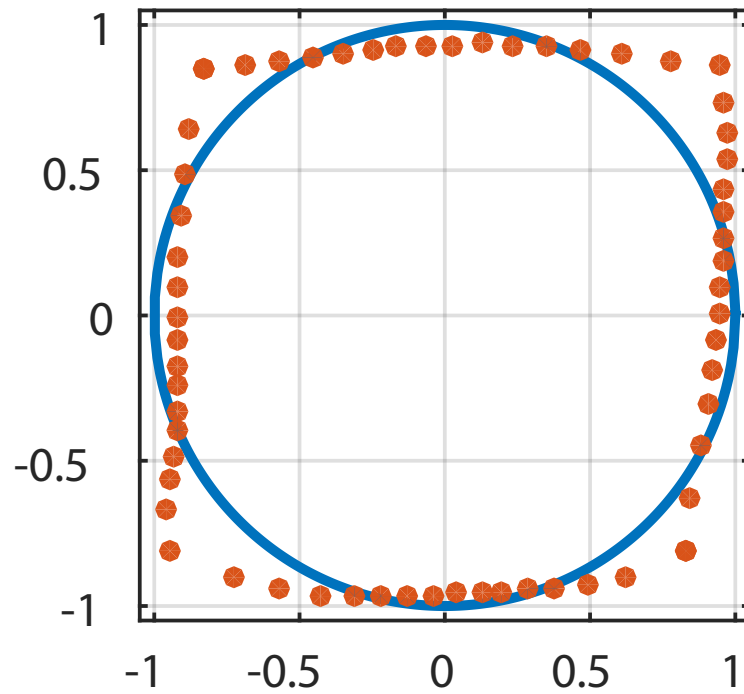


Figure 4.9: Measured phase shifter constellation points, with circle showing amplitude level with least average amplitude error.

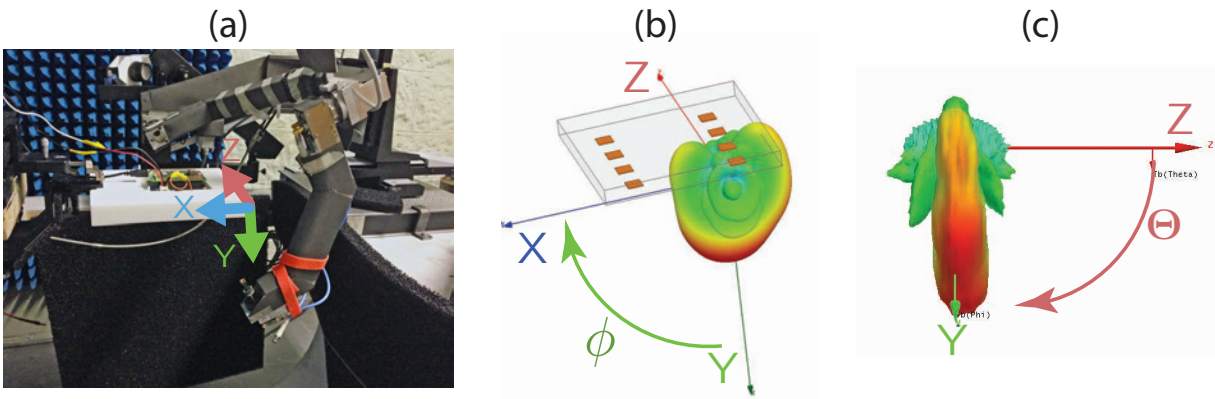


Figure 4.10: The measurement setup at University of Nice-Sophia Antipolis. The test PCB is placed with the radiating side downward, and the arm with the receive antenna is swept across ϕ and θ angles. The measured radiation pattern data is imported into HFSS and plotted. The axes shown in (a) correspond to the axes in the HFSS plots (subfigures b and c)

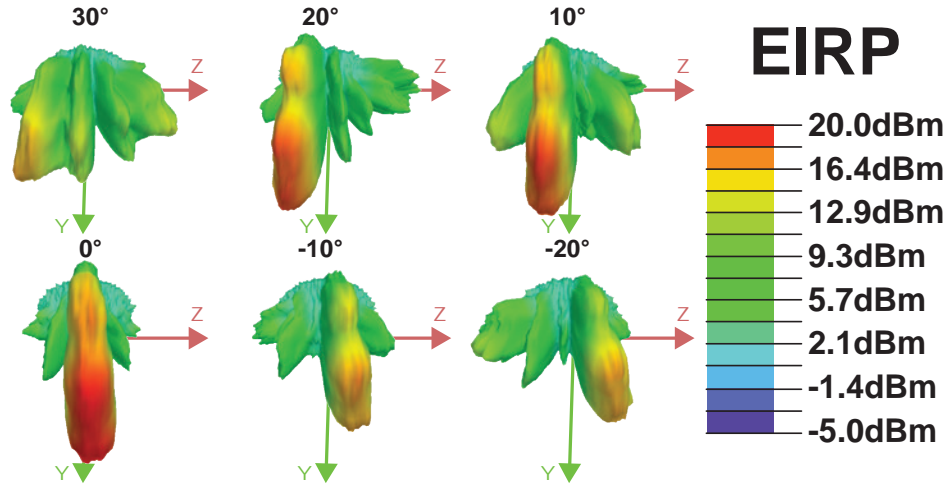


Figure 4.11: 3D radiation pattern measurements at various beam steering angles (performed at UNS).

the antenna arrays and phase shifters are identical for both the TX and RX.

The measured 3D radiation pattern for various beam steering angles is plotted in Figure 4.11. Similar to the HFSS simulations of the array, there are significant side lobes at the 30° beam steering angle. The beamforming measurements show a steering range of about $\pm 20^\circ$ degrees while maintaining 3dB main lobe to grating lobe levels. The beam drop-off is about 2-3dB at $+20^\circ$ and 4 dB at -20° (Figure 4.12). This is slightly worse than predicted by HFSS simulations of the antenna array, which showed only 1 to 1.5dB of beam drop off. Measurements along the E-plane of the array show little variation in radiation pattern versus beam steering angle (Figure 4.13).

4.2.2 Radar Measurements and Characterization

As a basic demonstration of radar capability, a simple experiment was performed using a metal reflector at different distances. A triangular frequency modulation was applied to the reference signal of the PLL, with both 4% and 2% sweep bandwidths. At 4% sweep bandwidth (3.68GHz RF bandwidth), the FFT of the IF waveform shows distinct peaks in different range bins when the object is moved by 5cm (Figure 4.14a). At 2% sweep bandwidth (1.86GHz RF bandwidth; for the 2% sweep, the center frequency was retuned slightly higher), the peaks are still present, but in some cases the reflected signal occupies multiple range bins (Figure 4.14b). According to the range resolution equation, there should be a range resolution of $\frac{c}{2B} = 4.1\text{cm}$ in the 4% sweep case, and 8.1cm in the 2% sweep case. This is consistent with the measured results of the radar experiment.

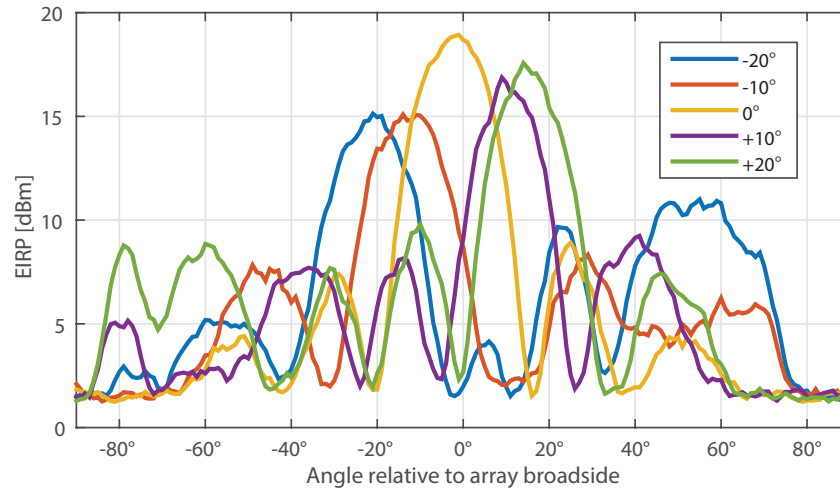


Figure 4.12: EIRP measurement along H-plane (performed at UNS).

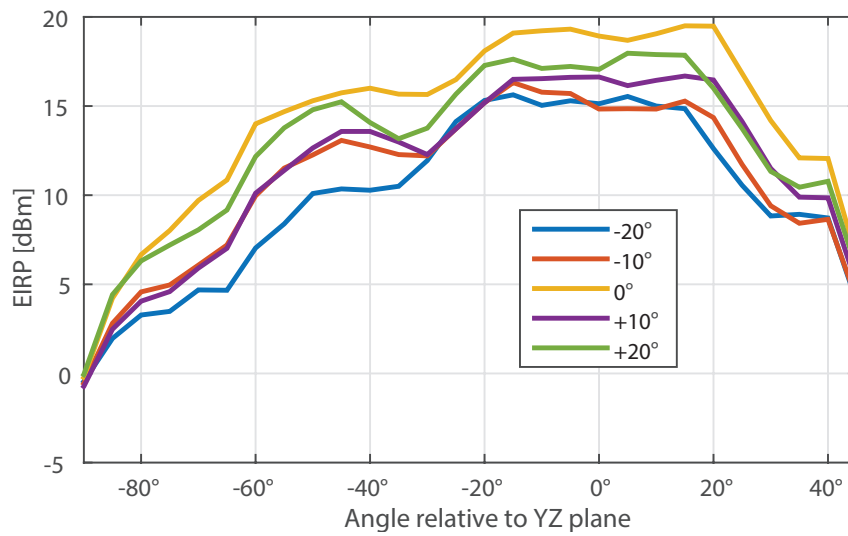


Figure 4.13: EIRP measurement along E-plane (performed at UNS).

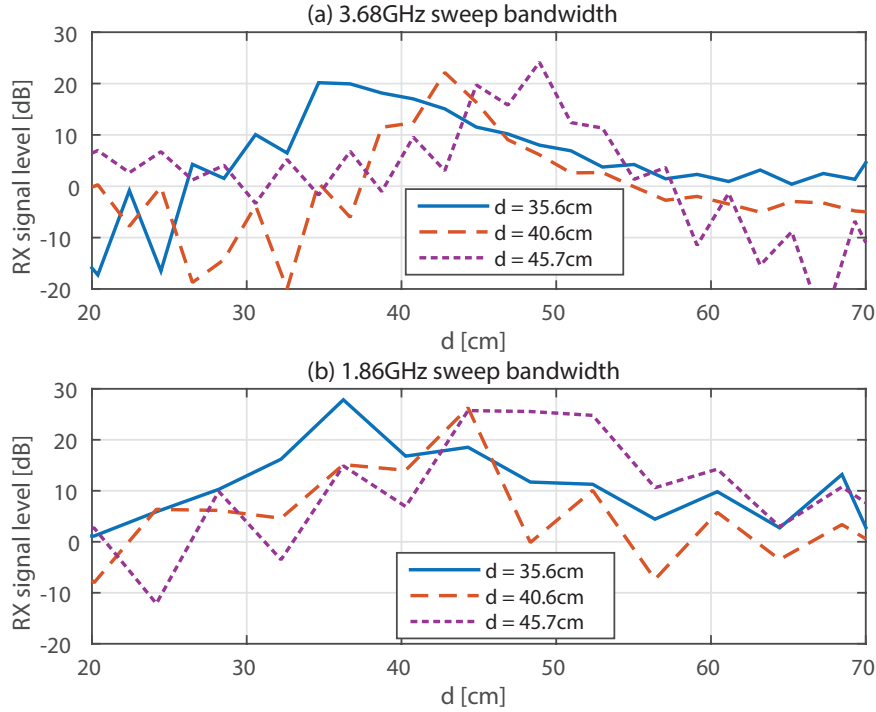


Figure 4.14: Radar measurements with a single target, at various distances. (a) with wide RF sweep bandwidth (b) with narrower RF sweep bandwidth.

4.3 Conclusion

A comparison of this work with state-of-the-art shows that this transceiver has achieved the lowest per-element DC power, including LO overhead power, while maintaining comparable per-element performance and demonstrating a high level of integration (Table 4.1). Further reduction in DC power could easily be achieved by using injection-locked frequency dividers instead of static BJT dividers (Figure 4.15).

In summary, a compact, highly integrated FMCW radar phased array transceiver module has been demonstrated. The integrated antenna-in-package allows for a small form factor of $1.2\text{cm} \times 1.2\text{cm}$ for the complete module, including all TX and RX antennas. Beam steering has been demonstrated across a $\pm 20^\circ$ range, and FMCW experiments show ranging functionality in line with theoretical resolution limits.

Table 4.1: Performance Comparison of Published W-Band Phased Array Transceivers

	This work	[39][47]	[48]	[35]	[49]	[50][51]
Technology	130nm SiGe	120nm SiGe	180nm SiGe	130nm SiGe	120nm SiGe	65nm CMOS
# of elements	4TX,4RX	16TX, 32RX	4TX,4RX	4TX, 4RX	16TX, 32RX	4TX, 4RX
Frequency [GHz]	94	94	70–100	94	90–102	94
Psat [dBm] @ 94GHz	6.4	3.2†‡	7.5†	6.5	-5‡	-5
RX NF [dB] @ 94GHz	12.5	8.2†‡	6†	16.5	9–9.5‡	10
TX DC power per ch. [mW] (no LO included)	71	—	—	86	137.5	—
TX DC power per ch. [mW] (LO included)	106* 135**	181	250	147	—	150
RX DC power per ch. [mW] (no LO included)	56	—	—	40	137.5	—
RX DC power per ch. [mW] (LO included)	91* 130**	138 (16 elem) 106 (32 elem)	250	101	—	90
Phase shifter resolution	9°	11.25°	22.5°	—	4°	—
IC size	3.7×2mm ²	6.6×6.7mm ²	1.7×2mm ²	4 chips × 1 × 1.1 mm ²	6.6 × 5.9 mm ²	2.1 × 3.6 mm ²
Packaging approach	Antenna-in-package	Antenna-in-package	Antenna on board	Antenna on board	probed testing only	Antenna-in-package
Level of integration	TX + RX + downconverter + VCO + dividers	TX + upconverter + RX + downconverter + PLL	TX + upconverter + RX + downconverter + freq. doubler	TX + RX + downconverter + LO buffers	TX + RX	TX + RX + LO

†Estimated from plot

‡Includes loss from TX/RX switch

*Simultaneous TX+RX: $P_{elem} = \frac{1}{4}(P_{RX \text{ or } TX} + \frac{1}{2}P_{LO})$ **TX or RX only mode: $P_{elem} = \frac{1}{4}(P_{RX \text{ or } TX} + P_{LO})$

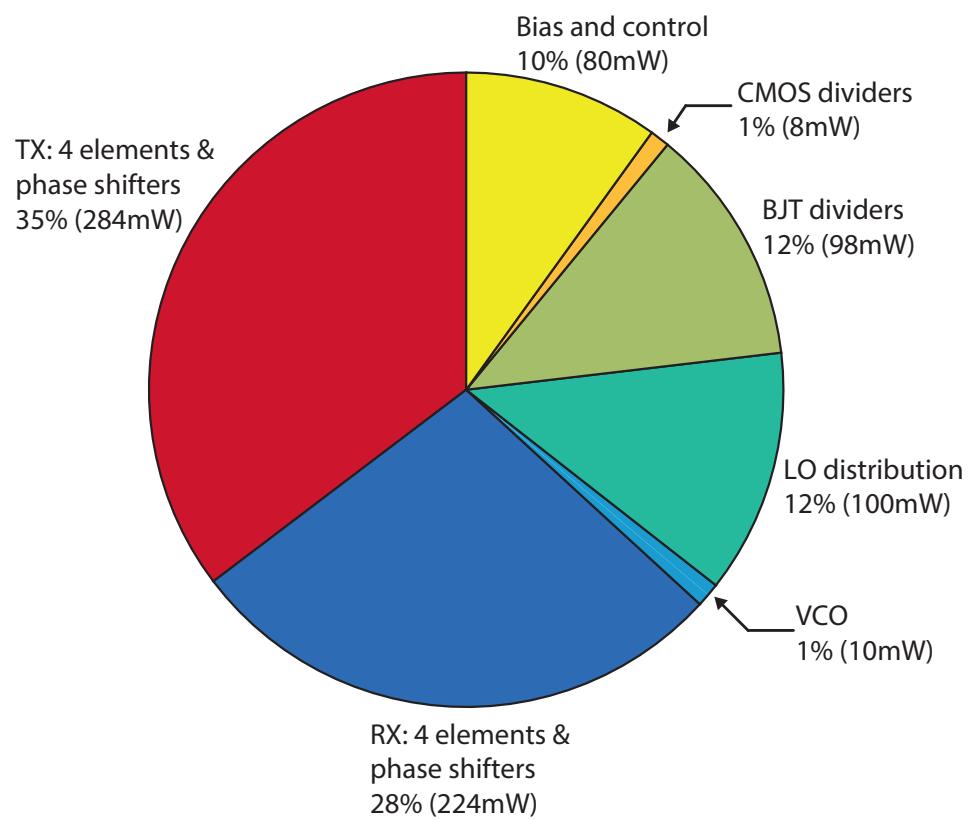


Figure 4.15: Measured DC power consumption by supply domain.

Chapter 5

Wideband mm-Wave Transceiver Design

5.1 Motivation and System Architecture

There are many examples of broadband data-rate mm-Wave transmitters and receivers that make use of simple modulation schemes to achieve high data rates. Yu *et al* have proposed a wideband on-off keying (OOK) transmitter [52] and receiver [53] pair that independently support up to 16Gb/s and 18.7Gb/s, respectively. Dolatsha *et al* demonstrate a full wireless link at 130GHz supporting 12.5Gb/s, similarly using OOK modulation [54]. Prior work at Berkeley has also used on-off keying to transmit and receive at 260GHz, demonstrating a 10Gb/s wireless link [55].

The appeal of OOK modulation is ease of implementation. It does not require anything more complex than a switch to modulate data onto an RF carrier. The receiver design is also simplified — only the amplitude of the received signal needs to be detected, and can be recovered using a simple peak detector circuit. On either side of the link, the circuits operate without any dependence on the phase (or frequency) of the RF carrier signal. Because of this, OOK-based links can be implemented with high energy efficiency (characterized by the energy to transmit one bit, or equivalently, the data rate divided by the DC power).

To increase the data rate transmitted by an OOK link, the bandwidth also needs to increase proportionally. If the system uses a fixed carrier frequency, this makes the circuit design increasingly challenging, resulting in an efficiency penalty. Wide bandwidth circuits can be designed more easily by using a higher carrier frequency, but the efficiency suffers if the frequency becomes too high due to poor transistor performance [55]. So, the path forward to maximize data rate while still maintaining reasonable overall efficiency requires switching to a different modulation technique.

In the 60GHz domain, the IEEE standard (802.11ad) is built around this idea. QAM-based modulation schemes — 16QAM and QPSK — are supported as part of this standard over wide bandwidths. The broad available 60GHz bandwidth is nominally split into several

channels spanning the full band. So-called “channel bonding” transmitters can transmit data over the full bandwidth for maximum data rate.

Recent research efforts in 60GHz transceiver design have focused on increasing data rate by supporting higher order modulations such as 64QAM and 128QAM with the broad bandwidth available from channel bonding [56][57], yielding impressive data rates. The drawback is that significant amounts of equalization are required to correct for non-flat gain and group delay in the amplifiers, circuitry, and antennas that comprise the system. Because of the high SNR requirement of the higher order modulation schemes, the equalization must be able to reduce the intersymbol interference of the channel to a very fine degree. While allowing for higher data rates at 60GHz, this degrades the efficiency of the link.

Meanwhile, scaling of silicon IC technology has enabled significant improvement in the performance of transceivers in the 100GHz to 200GHz range. Initial circuit performance seems to be following a similar trend as in early 60GHz designs. Complete transmitters and receivers have been demonstrated that show the promise for extremely wideband operation [58][59].

Therefore, there is little to suggest that high performance should not be achievable at frequencies above 60GHz. As a compromise between the higher absolute bandwidth available at higher frequencies, and the dropoff of transistor performance when getting close to the f_T of the technology, a center frequency of 120GHz was selected for this transceiver design as a compromise between the two contrasting requirements.

One final means to improve bandwidth efficiency as well as overall link efficiency, is by taking advantage of polarization diversity [60][61]. There are two possible orthogonal polarization states for electromagnetic waves, so a system designed to separately transmit or receive on both polarizations simultaneously could double its data rate without requiring any increase in circuit bandwidth (as would be the case for frequency-interleaved transceivers [62]). The efficiency (pJ/bit) can be potentially increased as well, since many components can be shared across both transceiver channels. The key challenge in this type of system is to have antennas with extremely high polarization purity, or an energy-efficient way of equalizing out cross-polarization.

5.1.1 System Architecture

To achieve the design goal of a high-data rate, high-efficiency mm-Wave wireless link, a dual-channel 120GHz 50Gb/s/channel transceiver IC is proposed (Figure 5.1). The IC is designed in 28nm CMOS to leverage the improved f_T and f_{max} relative to longer-channel CMOS technologies. The transceiver has fully differential mm-Wave inputs and outputs, and differential I and Q baseband outputs for each channel. To avoid the difficulties of feeding digital data to the transmitter at high enough speeds to fill the link capacity, an on-chip pseudo-random bit sequence (PRBS) generator is used with an external clock signal. The mm-Wave LO is generated from an externally fed 30GHz reference signal, which is quadrupled to reach the desired 120GHz carrier frequency.

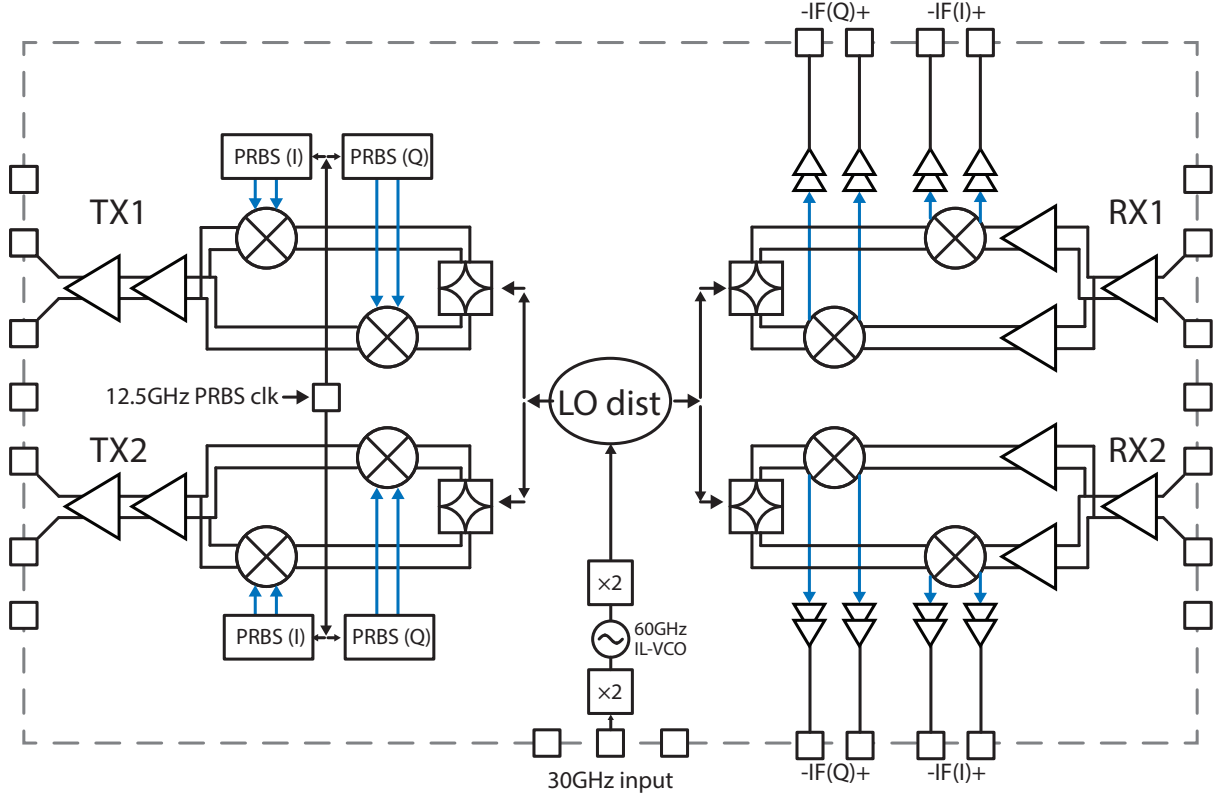


Figure 5.1: Block diagram of 120GHz dual-channel communication transceiver.

Based on rough estimates of achievable performance, a link budget was computed to determine the feasibility of this system (Figure 5.2). The link budget assumes a 10cm communication distance and (conservatively) a 25GHz noise bandwidth. The 22dB SNR requirement builds in 5dB margin beyond the 17dB required for QPSK with a bit error rate (BER) of 10^{-12} . If the same system were used for QAM16, 24dB of SNR would be required for the same bit error rate (Figure 5.3). Given that a 7dB increase in SNR is required to double the data rate without increasing BER, this demonstrates that 16QAM is not an energy efficient way to increase data rate — moving to 16QAM would likely require doubling the power consumption in both the TX and RX to meet the same link budget.

5.1.2 Packaging Approach

High-performance mm-Wave flip-chip packaging typically uses the smallest possible manufacturable flip-chip bump pitch to minimize the mutual inductance between signal and reference conductors. Stud bumps or copper pillars can offer higher performance than solder balls at a pitch of $200\mu\text{m}$ or below, down to perhaps around $100\mu\text{m}$, but with these fine pitches, fabricating the printed circuit board or substrate that the chip will be attached to

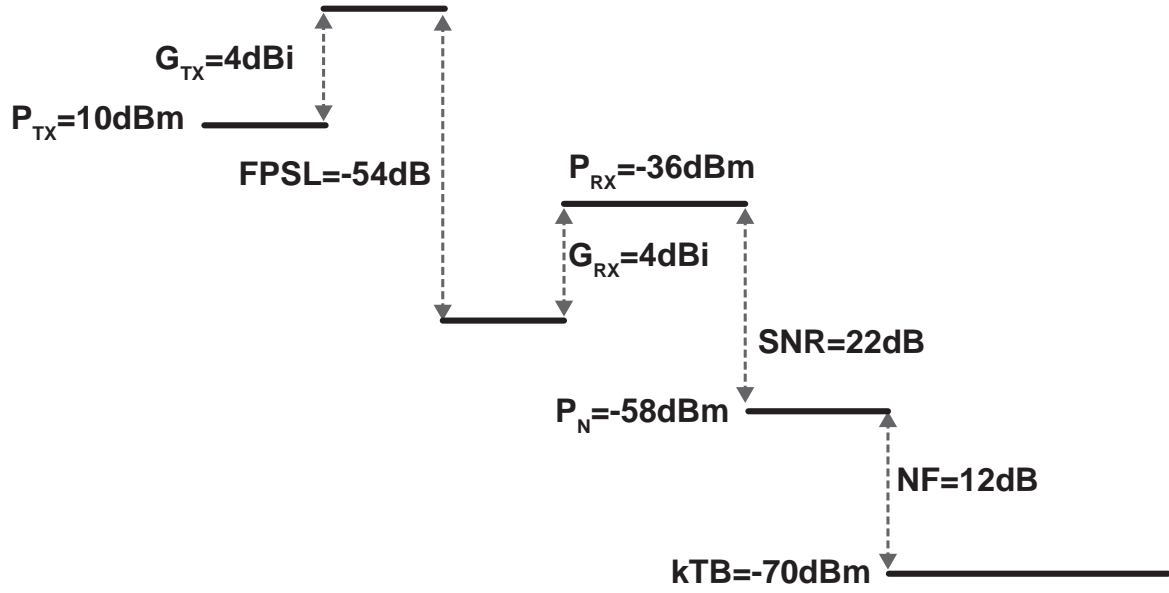


Figure 5.2: Link budget, assuming a 10cm transmit distance.

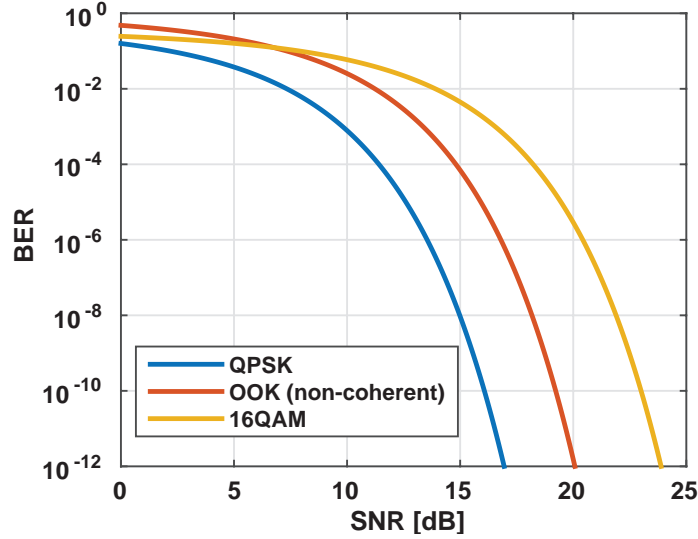


Figure 5.3: Bit error rate versus SNR for various modulation techniques.

can become very costly. For a low-cost solution, an approach using conventional solder ball flip chip technology at a relaxed pitch of 10mil ($254\mu\text{m}$) is proposed, that achieves moderate packaging losses even at 120GHz.

5.2 Building Blocks

5.2.1 Modular Common-Source Neutralized Amplifier Layout

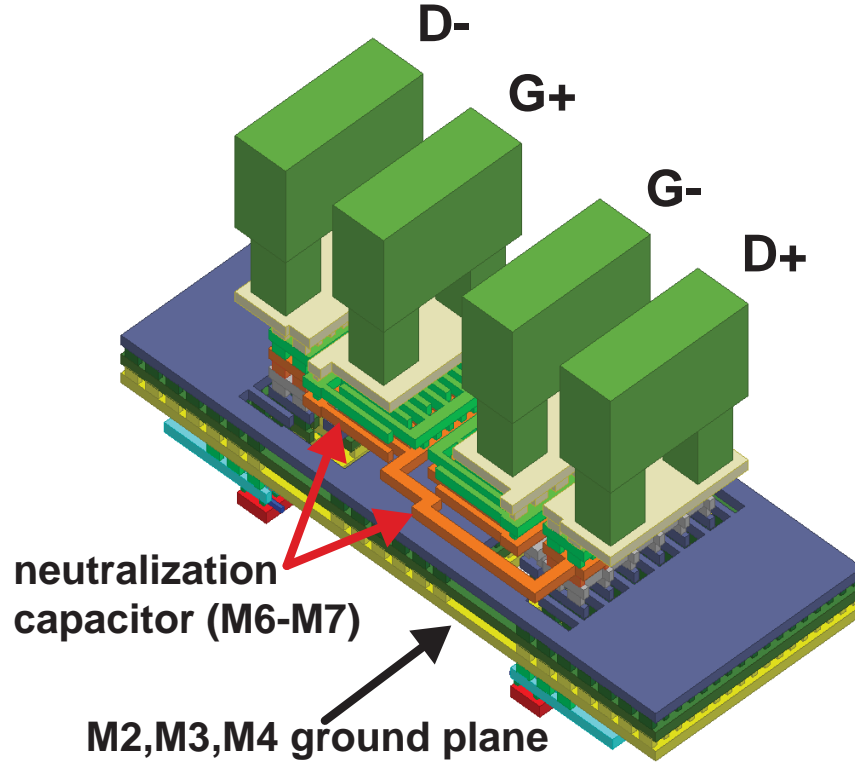


Figure 5.4: Angle view of neutralized unit cell layout

At millimeter-wave frequencies, achieving both robust stability and high power gain is made more difficult by the gate-to-drain capacitance of the transistor [63]. To mitigate this effect, neutralization techniques can be used to make the transistor unilateral - in other words, modify the circuit so that the effective y_{12} goes to zero. Transformer-based neutralization methods can be used [64][65], but rely on matching ratios of inductors to ratios of capacitances, which can be challenging given the care with which millimeter-wave passives must be modeled. Additionally, the transformer coils will have finite Q , lowering the potential gain that can be achieved.

On the other hand, capacitive neutralization via cross-coupling does not require lossy passives (MOM capacitors can be fairly high Q even at 100GHz), although it does come at the cost of doubling the input and output capacitances of the circuit [66]. This halves the f_T of the transistor, effectively trading off f_T for power gain. One downside is that

this technique is only applicable to balanced or differential circuits, which may make it less appealing for low-power designs.

To simplify the layout and reduce design iteration time given the complexity of layout rules in 28nm CMOS, a modular layout approach was adopted. Amplifiers are constructed out of a number of unit cells connected in parallel, and implemented in layout as a linear array of unit cells. A 3D representation of a typical unit cell layout is depicted below, in Figure 5.4.

This approach also helps with the tradeoff between high frequency modeling accuracy and complexity. The slices themselves are extracted using local parasitic extraction only, but the interconnection between the slices can be modeled using an electromagnetic field solver (in this case, Integrand EMX). If the interconnect pattern is also based on unit cells, a scalable lumped-element-based model (self and mutual inductances, and capacitances between nets) can be fitted to the simulation results.

A representative example of a full layout cell with amplifiers and routing is shown in Figure 5.5.

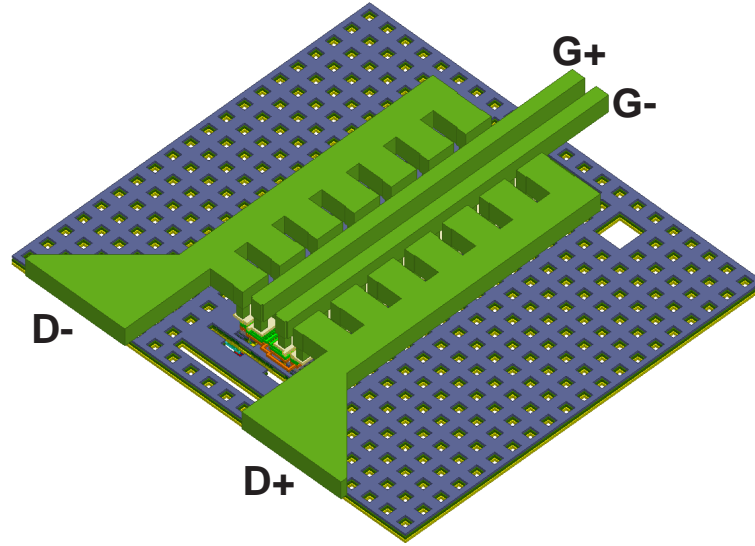


Figure 5.5: Angle view of amplifier with 7 unit cells

5.2.2 Coupled Resonators using Low-K Transformers

To understand the behavior of low-K transformer-based matching networks, we can study a simplified/idealized version first (Figure 5.6). This has two identical coupled inductors with equal resistances and capacitances in parallel.

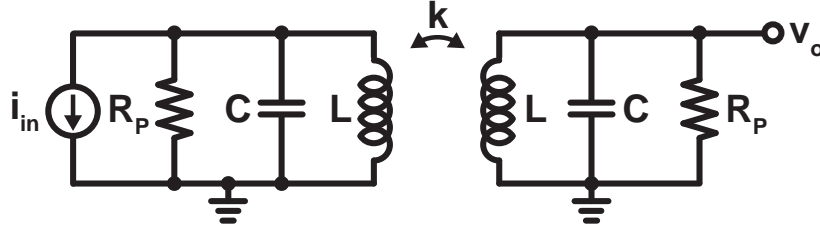


Figure 5.6: Magnetically coupled LC resonators

For a cascade of common-source amplifiers, we typically will be most interested in the **transimpedance** ($Z_T = \frac{v_o}{i_{in}}$) of this circuit.

$$Z_T(s) = \frac{kLs}{\left[1 + \frac{L}{R_P}(1-k)s + LC(1-k)s^2\right] \left[1 + \frac{L}{R_P}(1+k)s + LC(1+k)s^2\right]} \quad (5.1)$$

In terms of the parallel tank $Q = \frac{R_P}{\omega_0 L}$ and $\omega_0 = \frac{1}{\sqrt{LC}}$ of the individual, uncoupled RLC circuits, the transfer function is:

$$Z(s) = R_P \frac{k \frac{s}{Q\omega_0}}{\left[1 + \frac{1-k}{Q\omega_0}s + \frac{1-k}{\omega_0^2}s^2\right] \left[1 + \frac{1+k}{Q\omega_0}s + \frac{1+k}{\omega_0^2}s^2\right]} \quad (5.2)$$

To give an idea of what this transfer function looks like, some representative examples of this transfer function have been plotted below in Figure 5.7 for some select values of k and Q .

We can verify that as $|k| \rightarrow 1$, the transfer function takes the limiting case of

$$Z(s) = \pm \frac{R_P}{2} \frac{\frac{2}{\omega_0 Q} s}{1 + \frac{2}{\omega_0 Q} s + \frac{2}{\omega_0^2} s^2} \quad (5.3)$$

as expected, with the final sign coming from the direction of mutual coupling (sign of k). The new complex pole pairs of the coupled inductors are:

$$p_{1,2} = \frac{\omega_0}{2Q} \left[-1 \pm j \sqrt{\frac{4Q^2}{1+k} - 1} \right] \quad (5.4a)$$

$$\omega_{0(1,2)} = \frac{\omega_0}{\sqrt{1+k}} \quad Q_{1,2} = \frac{Q}{\sqrt{1+k}} \quad (5.4b)$$

$$p_{3,4} = \frac{\omega_0}{2Q} \left[-1 \pm j \sqrt{\frac{4Q^2}{1-k} - 1} \right] \quad (5.4c)$$

$$\omega_{0(3,4)} = \frac{\omega_0}{\sqrt{1-k}} \quad Q_{3,4} = \frac{Q}{\sqrt{1-k}} \quad (5.4d)$$

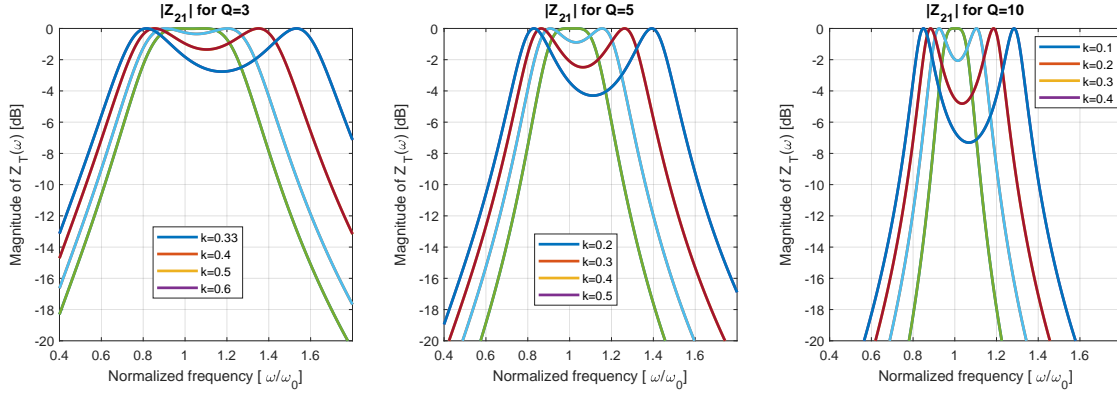


Figure 5.7: $|Z_{21}|$ transfer functions for various values of the uncoupled resonator Q

The frequencies of the two resonant peaks and transimpedance minimum can be exactly computed:

$$\omega_{max1} = \omega_0 \sqrt{\frac{-1 + k^2 + 2Q^2 - \sqrt{1 + k^4 - 4Q^2 + k^2(-2 + 4Q^2 + 4Q^4)}}{2(1 - k^2)Q^2}} \quad (5.5a)$$

$$\omega_{max2} = \omega_0 \sqrt{\frac{-1 + k^2 + 2Q^2 + \sqrt{1 + k^4 - 4Q^2 + k^2(-2 + 4Q^2 + 4Q^4)}}{2(1 - k^2)Q^2}} \quad (5.5b)$$

$$\omega_{min} = \omega_0 \sqrt{\frac{-1 + k^2 + 2Q^2 + \sqrt{1 + k^4 - 4Q^2 + 16Q^4 - 2k^2(1 - 2Q^2 + 6Q^4)}}{6(1 - k^2)Q^2}} \quad (5.5c)$$

For large Q these can be approximated as:

$$\omega_{max1} = \frac{\omega_0}{\sqrt{1 + k}} \left(1 - \frac{1 - k^2}{4kQ^2} - \frac{(2 - k)(1 - k^2)^2}{32k^3Q^4} + \dots \right) \approx \frac{\omega_0}{\sqrt{1 + k}} \quad (5.6a)$$

$$\omega_{max2} = \frac{\omega_0}{\sqrt{1 - k}} \left(1 - \frac{1 - k^2}{4kQ^2} - \frac{(2 + k)(1 - k^2)^2}{32k^3Q^4} + \dots \right) \approx \frac{\omega_0}{\sqrt{1 - k}} \quad (5.6b)$$

$$\omega_{min} \approx \frac{\omega_0}{\sqrt{3}} \sqrt{\frac{1 + \sqrt{4 - 3k^2}}{1 - k^2}} \quad (5.6c)$$

The two resonant frequencies are plotted below in Figure 5.8 as the coupling frequency is varied, for various tank Q factors.

The magnitude of the transfer function at each of the two peaks is (exactly) $\frac{R_P}{2}$ and is independent of k .

The magnitude of the transfer function at the minimum can be calculated, but the exact

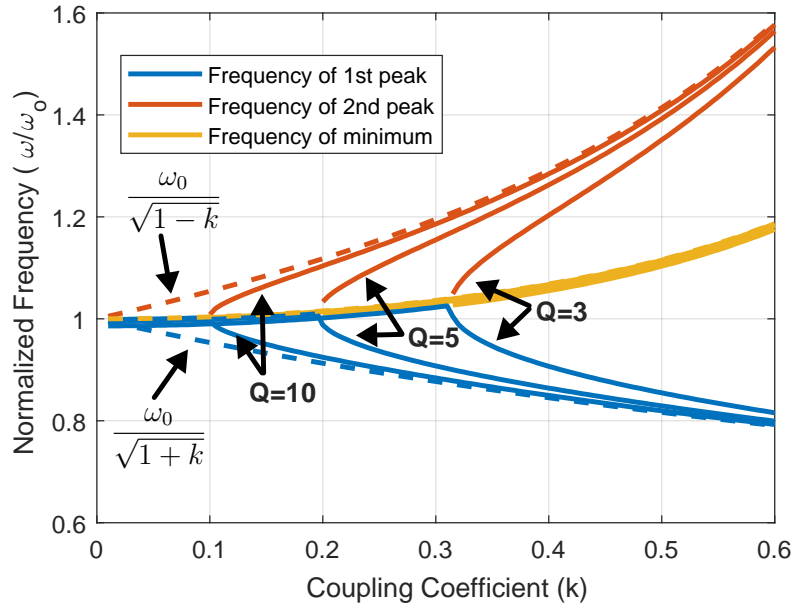


Figure 5.8: Frequencies of maxima and minimum of coupled RLC circuit transimpedance for various values of Q .

result is extremely complicated:

$$|Z(j\omega_{min})| = \frac{1}{G} \sqrt{\frac{3}{2}} k Q^2 \sqrt{\frac{1-k^2}{\beta + \alpha^{3/2}}} \quad (5.7)$$

Where α and β are given by:

$$\alpha = 4(4 - 3k^2)Q^4 - 4(1 - k^2)Q^2 + (1 - k^2)^2 \quad (5.8a)$$

$$\beta = -8(8 - 9k^2)Q^6 + 6(1 - k)(4 + k)Q^4 + 6(1 - k^2)^2Q^2 + k^2(1 - k^2)^2 \quad (5.8b)$$

Because of the complexity of this expression, it is not straightforward to derive a simple equation for the ripple as a function of the coupling coefficient. Instead, the ripple is plotted below for various values of Q (Figure 5.9).

For a maximally flat response, $\omega_{max1} = \omega_{max2}$. This occurs at:

$$k = \pm \sqrt{1 - 2Q^2 - 2Q^4 + 2\sqrt{2Q^6 + Q^8}} \quad (5.9)$$

This can be approximated (for $Q \gg 1$) as follows using the Taylor series for $\sqrt{1+x}$ around $x = 0$:

$$k = \pm \sqrt{1 - 2Q^2 - 2Q^4 + 2Q^4 \sqrt{1 + \frac{2}{Q^2}}} \quad (5.10)$$

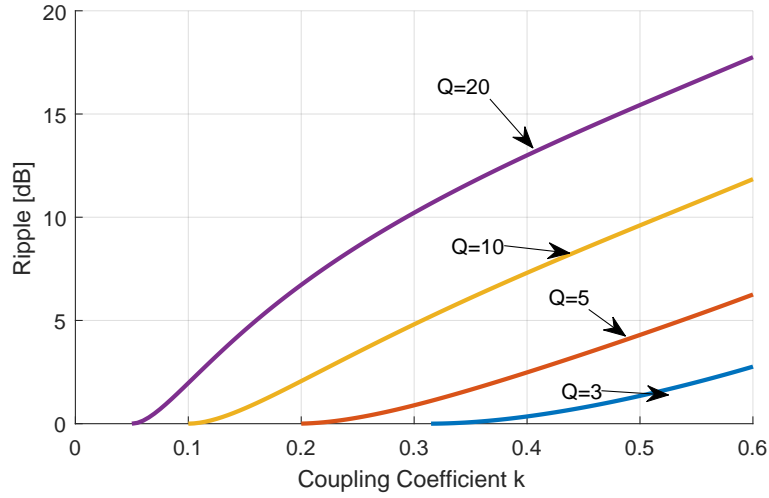


Figure 5.9: Ripple of the magnitude of the transimpedance vs coupling coefficient.

Then, the term under the inner square root can be expanded via Taylor series:

$$k = \pm \sqrt{1 - 2Q^2 - 2Q^4 + 2Q^4 \left(1 + \frac{1}{2} \cdot \frac{2}{Q^2} - \frac{1}{8} \cdot \frac{4}{Q^4} + \frac{1}{16} \cdot \frac{8}{Q^6} - \frac{5}{128} \cdot \frac{16}{Q^8} + \frac{7}{256} \cdot \frac{32}{Q^{10}} + \dots \right)} \quad (5.11)$$

Canceling terms:

$$k = \pm \sqrt{\frac{1}{Q^2} - \frac{5}{8Q^4} + \frac{7}{8Q^6} + \dots} = \pm \frac{1}{Q} \sqrt{1 - \frac{5}{8Q^2} + \frac{7}{8Q^4} \dots} \quad (5.12)$$

So for “large” Q , maximum flatness is achieved with:

$$k \approx \pm \frac{1}{Q} \quad (5.13)$$

This approximation is actually quite good, even for Q as low as 3. Visually, it can also be confirmed from Figure 5.8 — the frequencies of the two maxima become equal at about $k \approx \frac{1}{Q}$.

5.2.3 Low-K Transformers with Lossy Inductors

Not surprisingly, the analysis is not as straightforward when the inductors themselves have finite Q . In that case, the circuit¹. looks more like that of Figure 5.10, where $R_S = \frac{\omega_0 L}{Q_L}$

¹More generally, R_S increases versus frequency due to the skin effect, but this is not a significant source of modeling error for moderate bandwidth circuits.

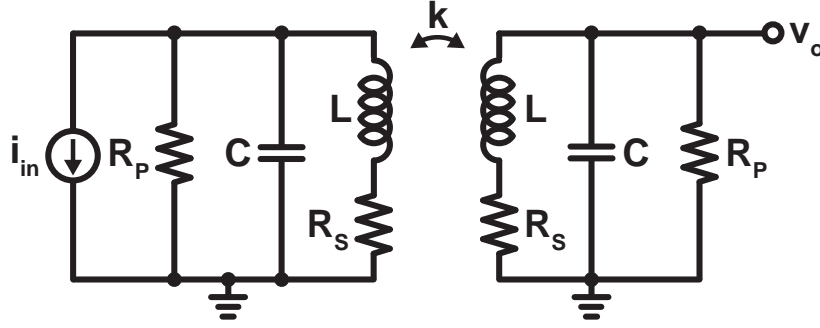


Figure 5.10: Magnetically coupled LC resonators with finite Q inductors

For narrowband circuits, it's possible to do a series-parallel transformation and lump the effect of R_S into the tank Q . However, to correctly describe broadband behavior, this is not strictly correct - the transfer function has to be re-computed.

With series inductor resistance, the complex pole pairs of the circuit shift to:

$$\omega_{0(1,2)} = \frac{\omega_0}{\sqrt{1+k}} \sqrt{1 + \frac{1}{QQ_L}} \quad (5.14a)$$

$$Q_{1,2} = \frac{Q}{\sqrt{1+k}} \sqrt{1 + \frac{1}{QQ_L}} \frac{1}{1 + \frac{Q/Q_L}{1+k}} \quad (5.14b)$$

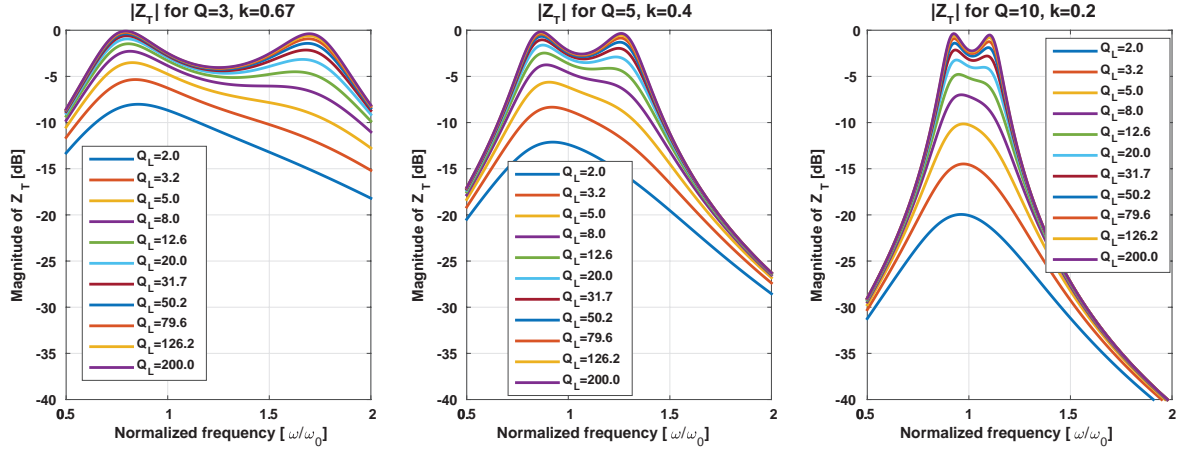
$$\omega_{0(3,4)} = \frac{\omega_0}{\sqrt{1-k}} \sqrt{1 + \frac{1}{QQ_L}} \quad (5.14c)$$

$$Q_{3,4} = \frac{Q}{\sqrt{1-k}} \sqrt{1 + \frac{1}{QQ_L}} \frac{1}{1 + \frac{Q/Q_L}{1-k}} \quad (5.14d)$$

The last part of the Q term is what is problematic. Both pole pairs see a reduction in Q , but the higher frequency pole pair associated with $(1 - k)$ sees a greater reduction in Q . This leads to a non-flat gain response (Figure 5.11).

There are several potential solutions to this problem that have been proposed in the literature. One solution proposes the use of capacitive coupling, in addition to the magnetic coupling between the tanks [67]. Although attractive from a schematic point of view, adding capacitors bridging between the two LC tanks significantly complicates the layout. Also, this method adds loss as well, due to the long series routing lines for the bridging capacitors.

A more practical approach is to scale the sizes of the inductances, while maintaining the same coupling factor [68]. This effectively “pre-distorts” the frequency response, so that when the Q of the second resonance is lowered, it results in an overall flat gain response.

Figure 5.11: $|Z_{21}|$ transfer functions with finite- Q inductors

5.3 Transmitter

5.3.1 Modulator Design

5.3.1.1 Common-Gate RF DAC

In a conventional RF upconversion mixer, analog baseband data is fed to the circuit at the tail of a hard-switching current commuting pair driven at the LO frequency (Figure 5.12). To achieve high bandwidth and linearity without sacrificing gain, the modulator architecture presented here departs significantly from the conventional design.

Instead of using a baseband DAC to drive the transconductor stage (devices M_1 and M_2 of Figure 5.12), the modulator in this work uses an RF DAC architecture which directly modulates the mm-Wave carrier with digital data [43]. This architecture minimizes the number of analog blocks between the digital I and Q data streams and the modulated RF signal, avoiding any added bandwidth and linearity degradations from the baseband DAC.

Another unusual design choice in this modulator is the swapping of source and gate for the LO and baseband ports. A more conventional upconversion mixer has the LO connected to the gate, and relies on hard-switching of the mixer devices to ensure linear upconversion of the baseband signal at the sources. However, at mm-Wave it is difficult to achieve hard-switching in the mixer since the LO drive is largely sinusoidal. Additionally, it can be difficult to impedance match to the gate of the mixer which is a high- Q capacitance. Both of these factors make it difficult to have high conversion gain with the conventional topology.

Since it is simpler to achieve high swing with a digital baseband signal, in this design, the baseband data is used to hard-switch the mixer devices in each unit cell to control the polarity of the RF signal. Because the LO input of the mixer is at the source rather than the gate, impedance matching at the input is simplified. This is similar to the “power mixer” proposed by Dasgupta [69] and the RF DAC architecture of Shopov [70].

The unit element of the DAC is a doubly-balanced Gilbert cell (Figure 5.13). The mm-

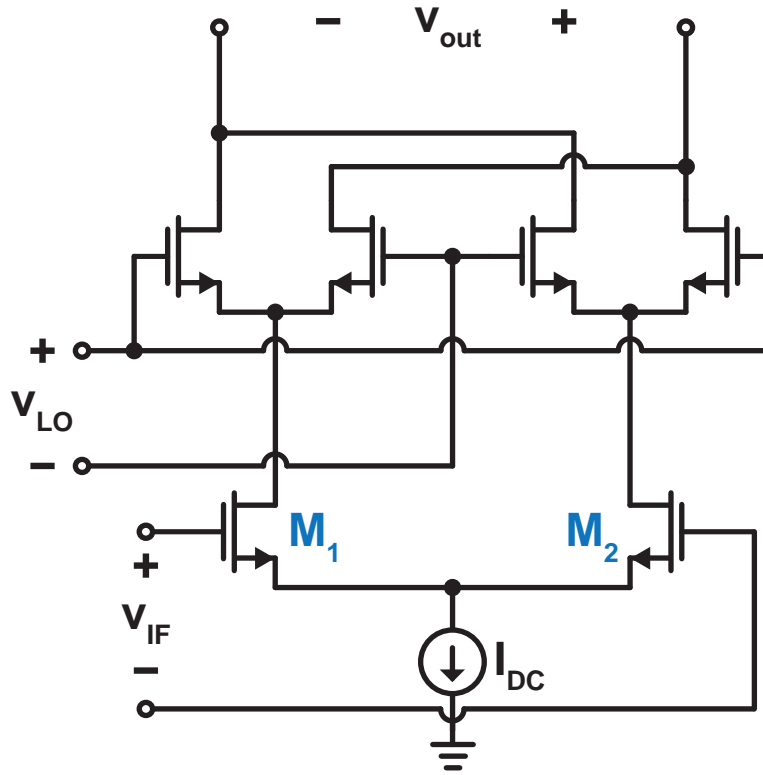


Figure 5.12: Conventional upconversion mixer topology typically used at low frequencies.

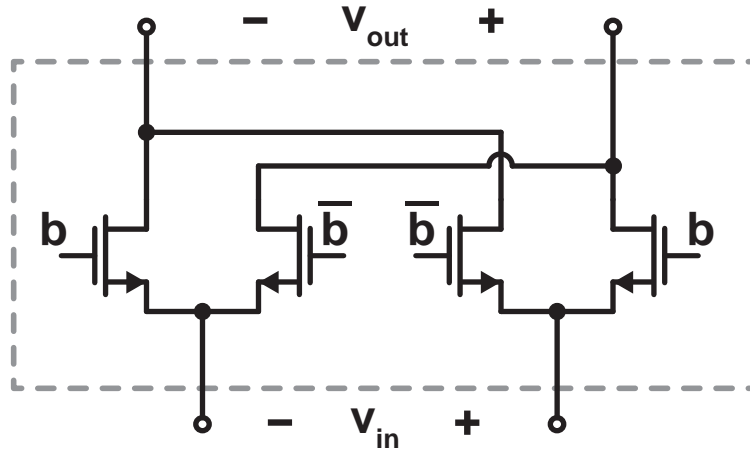


Figure 5.13: Unit element of common-gate RF DAC.

Wave sinusoidal LO signal is current steered by the digital baseband signal **D**. For multi-bit DACs, unit elements are grouped together in binary-weighted fashion [71]. Full-scale DAC codes are reached by steering all of the LO signal to one drain or the other. Intermediate codes rely on steering some LO signal to the “positive” drain and some to the “negative”

drain so that the LO signal partially cancels at the DAC output (Figure 5.14).

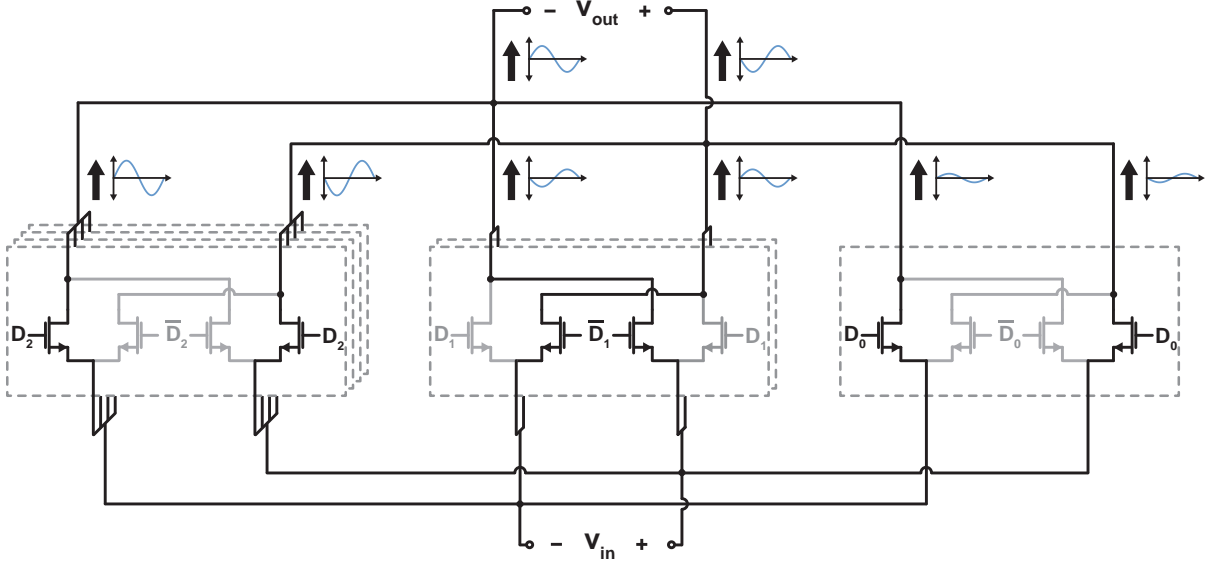


Figure 5.14: DAC operation for 8-level amplitude modulation. The DAC is depicted at code 101, so the total current at the output is $+\frac{4}{7} - \frac{2}{7} + \frac{1}{7} = \frac{3}{7}$ of the full scale value.

For a 3-bit DAC, the current gain is as follows:

$$\frac{i_{out}}{i_{in}} = \frac{1}{7} (-1)^{D_0} + \frac{2}{7} (-1)^{D_1} + \frac{4}{7} (-1)^{D_2} \quad (5.15)$$

The full mapping of DAC bits to amplitude values is shown below in Table 5.1.

D_2	D_1	D_0	DAC Amplitude
0	0	0	$\frac{4}{7} + \frac{2}{7} + \frac{1}{7} = 1$
0	0	1	$\frac{4}{7} + \frac{2}{7} - \frac{1}{7} = \frac{5}{7}$
0	1	0	$\frac{4}{7} - \frac{2}{7} + \frac{1}{7} = \frac{3}{7}$
0	1	1	$\frac{4}{7} - \frac{2}{7} - \frac{1}{7} = \frac{1}{7}$
1	0	0	$-\frac{4}{7} + \frac{2}{7} + \frac{1}{7} = -\frac{1}{7}$
1	0	1	$-\frac{4}{7} + \frac{2}{7} - \frac{1}{7} = -\frac{3}{7}$
1	1	0	$-\frac{4}{7} - \frac{2}{7} + \frac{1}{7} = -\frac{5}{7}$
1	1	1	$-\frac{4}{7} - \frac{2}{7} - \frac{1}{7} = -1$

Table 5.1: Mapping from DAC bits to amplitude values

The DAC architecture of Figure 5.14 is a Class-A-like mode of operation, in that the DC current of the DAC is constant versus code. An alternative Class-B-like design could turn

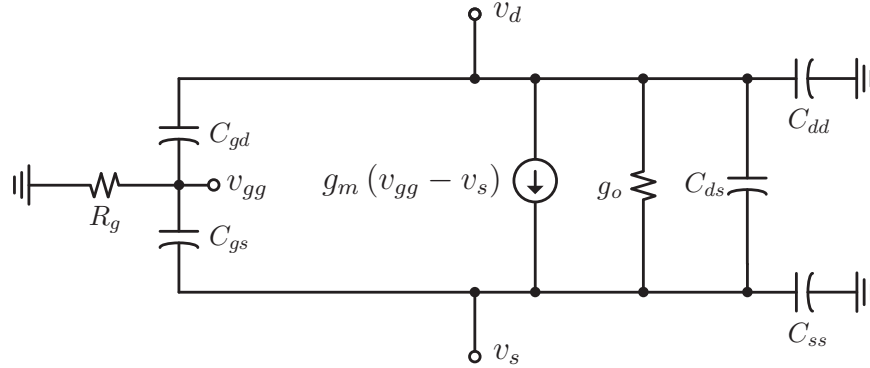


Figure 5.15: Small signal model of common-gate amplifier.

off DAC elements for intermediate codes instead. This is potentially problematic in that the DAC input impedance will vary vs code (since a different number of “on” devices are seen in each state). Also, each code transition would inject a common-mode current step into the output matching network which is undesirable. Although inefficient, the Class-A DAC does ensure that the input impedance is roughly constant across all codes, and minimizes problems from common-mode transients.

5.3.1.2 CG DAC Model

To model potential non-idealities of the DAC, we first derive a common gate equivalent circuit model, based on the small signal model of Figure 5.15.

The Y-parameters of a common-gate circuit are shown in Equations 5.16a-5.16d.

$$y_{11} = g_m + g_o + sC_{ss} + sC_{ds} + sC_{gs} \frac{1 - g_m R_g + sR_g C_{gd}}{1 + sR_g (C_{gs} + C_{gd})} \quad (5.16a)$$

$$y_{21} = -g_m - g_o - sC_{ds} + sC_{gs} \frac{g_m R_g + sR_g C_{gd}}{1 + sR_g (C_{gs} + C_{gd})} \quad (5.16b)$$

$$y_{12} = -g_o - sC_{ds} - sC_{gd} \frac{g_m R_g + sR_g C_{gs}}{1 + sR_g (C_{gs} + C_{gd})} \quad (5.16c)$$

$$y_{22} = g_o + sC_{dd} + sC_{ds} + sC_{gd} \frac{1 + g_m R_g + sR_g C_{gs}}{1 + sR_g (C_{gs} + C_{gd})} \quad (5.16d)$$

Because the gate resistance is nonzero, the internal gate voltage v_{gg} is also nonzero. The internal gate node v_{gg} , through the action of the g_m dependent source, ends up modulating the Y-parameters of the device.

So, with non-zero gate resistance, the transfer functions from source to v_{gg} ($\frac{v_{gg}}{v_s}$) and drain to v_{gg} ($\frac{v_{gg}}{v_d}$) form a C-R filter and therefore have a high-pass characteristic.

The end result is that the gate resistance modifies the capacitances seen in the Y-parameters of the circuit (Equations 5.18a-5.18d). Note that the pole/zero frequencies of the

sC_{gd} and sC_{gs} terms of the Y-parameters are close to or beyond the f_t of the technology, so we can simplify each the Y-parameters to an equivalent capacitive and real part (Equation 5.17). Because of the direction of the g_m source, the result is an added negative capacitance if the drain is grounded (Equations 5.18a,b) or an added positive capacitance if the source is grounded (Equations 5.18c,d).

$$y_{11} = g_m + g_o + sC_{11} \quad (5.17a)$$

$$y_{21} = -g_m - g_o - sC_{21} \quad (5.17b)$$

$$y_{12} = -g_o - sC_{12} \quad (5.17c)$$

$$y_{22} = g_o + sC_{22} \quad (5.17d)$$

where

$$C_{11} = C_{ss} + C_{ds} + C_{gs} (1 - g_m R_g) \quad (5.18a)$$

$$C_{21} = C_{ds} - C_{gs} g_m R_g \quad (5.18b)$$

$$C_{12} = C_{ds} + C_{gd} g_m R_g \quad (5.18c)$$

$$C_{22} = C_{dd} + C_{ds} + C_{gd} (1 + g_m R_g) \quad (5.18d)$$

Simulations show that the simplified Y-parameters are valid (although not perfect) approximations at 120GHz.

One way of visualizing the systematic nonlinearity in the operation of this DAC is by modeling an infinite resolution DAC. This allows for simple equations that can be used to analytically predict AM-AM and AM-PM distortion. The Y-parameters of the CG amplifier are simulated in the “on” case ($V_g = 1V$) and the “off” case ($V_g = 0V$). Then, they are used to compute Y-parameters (Equation 5.19) of the “infinite resolution” DAC model of Figure 5.16. In this model, the input current is continuously steered so that the DAC achieves a current gain of between -1 (for $w = 0$) to 1 (for $w = 1$).

The 4-port Y-parameters of the DAC model of Figure 5.16 are:

$$\begin{bmatrix} i_1 \\ i_2 \\ i_3 \\ i_4 \end{bmatrix} = \begin{bmatrix} y_{11}^{ON} + y_{11}^{OFF} & w y_{12}^{ON} + (1-w) y_{12}^{OFF} & 0 & (1-w) y_{12}^{ON} + w y_{12}^{OFF} \\ w y_{21}^{ON} + (1-w) y_{21}^{OFF} & y_{22}^{ON} + y_{22}^{OFF} & (1-w) y_{21}^{ON} + w y_{21}^{OFF} & 0 \\ 0 & (1-w) y_{12}^{ON} + w y_{12}^{OFF} & y_{11}^{ON} + y_{11}^{OFF} & w y_{12}^{ON} + (1-w) y_{12}^{OFF} \\ (1-w) y_{21}^{ON} + w y_{21}^{OFF} & 0 & w y_{21}^{ON} + (1-w) y_{21}^{OFF} & y_{22}^{ON} + y_{22}^{OFF} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{bmatrix} \quad (5.19)$$

These Y parameters are then transformed into mixed-mode Y-parameters:

$$\begin{bmatrix} i_{dm1} \\ i_{dm2} \\ i_{cm1} \\ i_{cm2} \end{bmatrix} = \begin{bmatrix} \frac{1}{2} (y_{11}^{ON} + y_{11}^{OFF}) & \frac{2w-1}{2} (y_{12}^{ON} - y_{12}^{OFF}) & 0 & 0 \\ \frac{2w-1}{2} (y_{21}^{ON} - y_{21}^{OFF}) & \frac{1}{2} (y_{22}^{ON} + y_{22}^{OFF}) & 0 & 0 \\ 0 & 0 & 2 (y_{11}^{ON} + y_{11}^{OFF}) & 2 (y_{12}^{ON} + y_{12}^{OFF}) \\ 0 & 0 & 2 (y_{21}^{ON} + y_{21}^{OFF}) & 2 (y_{22}^{ON} + y_{22}^{OFF}) \end{bmatrix} \begin{bmatrix} v_{dm1} \\ v_{dm2} \\ v_{cm1} \\ v_{cm2} \end{bmatrix} \quad (5.20)$$

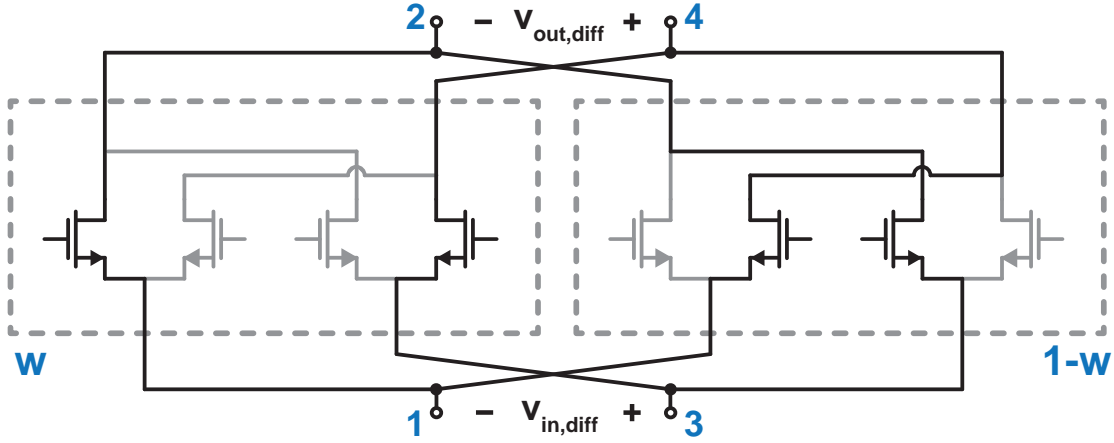


Figure 5.16: DAC model with continuous current steering. w represents the proportion of DAC unit elements steered in the “+1” state, so then $1 - w$ represents the proportion of DACs in the “-1” state.

Because the DAC is a fully balanced mixer, there is no mode conversion and the common mode can be ignored. So, only the upper-left block of Eq. 5.20 needs to be considered. Then we are left with equation 5.21.

$$\begin{bmatrix} i_{dm1} \\ i_{dm2} \end{bmatrix} = \begin{bmatrix} \tilde{y}_{11} & \tilde{y}_{12} \\ \tilde{y}_{21} & \tilde{y}_{22} \end{bmatrix} \begin{bmatrix} v_{dm1} \\ v_{dm2} \end{bmatrix} = \begin{bmatrix} \frac{1}{2} (y_{11}^{ON} + y_{11}^{OFF}) & \frac{2w-1}{2} (y_{12}^{ON} - y_{12}^{OFF}) \\ \frac{2w-1}{2} (y_{21}^{ON} - y_{21}^{OFF}) & \frac{1}{2} (y_{22}^{ON} + y_{22}^{OFF}) \end{bmatrix} \begin{bmatrix} v_{dm1} \\ v_{dm2} \end{bmatrix} \quad (5.21)$$

To avoid confusion with the Y-parameters of the CG transistor, the Y-parameters of the DAC are denoted with a tilde above the y .

The DAC Y-parameters reveal a couple of things. Firstly, the reason the DAC works is because the Y-parameters of the common-gate transistor change between its off and on states. If this were not the case, the differential mode y_{21} would vanish.

Secondly, the DAC is unilateral if the y_{12} of the common-gate transistor is the same in both the off and on states. This is not inherently true (there is a finite transistor g_o that varies strongly with operating point), and there are some implications to this that are discussed in the following sections.

Leaving the unilaterality of the DAC aside for now, the source and load impedances for a bi-conjugate match to the DAC can be computed from its Y-parameters [72].

$$K = \frac{2 \operatorname{Re}(Y_{11}) \operatorname{Re}(Y_{22}) - \operatorname{Re}(Y_{12}Y_{21})}{|Y_{12}Y_{21}|} \quad (5.22)$$

$$Y_S^* = y_{11} - \frac{y_{12}y_{21}}{Y_L + y_{22}} \quad (5.23a)$$

$$Y_L^* = y_{22} - \frac{y_{12}y_{21}}{Y_S + y_{11}} \quad (5.23b)$$

$$Y_{L_{opt}} = -j\tilde{b}_{22} \left[1 - \frac{\text{Im}[\tilde{y}_{12}\tilde{y}_{21}]}{2\tilde{g}_{11}\tilde{b}_{22}} \right] + \tilde{g}_{22} \sqrt{1 - \frac{\text{Re}[\tilde{y}_{12}\tilde{y}_{21}]}{\tilde{g}_{11}\tilde{g}_{22}} - \left(\frac{\text{Im}[\tilde{y}_{12}\tilde{y}_{21}]}{2\tilde{g}_{11}\tilde{g}_{22}} \right)^2} \quad (5.24)$$

where $\tilde{b}_{ij} = \text{Im}[\tilde{y}_{ij}]$ and $\tilde{g}_{ij} = \text{Re}[\tilde{y}_{ij}]$.

So, if the DAC is unilateral (i.e. $\tilde{y}_{12} = 0$) the optimal load impedance will just be the complex conjugate of the \tilde{y}_{22} of the DAC. Unfortunately, due to the impacts of C_{ds} , r_o , and the nonzero gate resistance of the common-gate amplifiers that make up the DAC, \tilde{y}_{12} is non-zero. In fact, the problem is somewhat worse than that - because \tilde{y}_{12} and \tilde{y}_{21} are proportional to the DAC code, the variation in $Y_{L_{opt}}$ will be related to the DAC code.

Since the equation is quite complex, it's reasonable to start understanding this type of behavior in the low-frequency limit and then see how it extends to higher frequencies. So, first consider the case where $\tilde{b}_{11} = \tilde{b}_{12} = \tilde{b}_{21} = \tilde{b}_{22} = 0$. Then:

$$Y_{L_{opt}} = \tilde{g}_{22} \sqrt{1 - \frac{\tilde{g}_{12}\tilde{g}_{21}}{\tilde{g}_{11}\tilde{g}_{22}}} \quad (5.25)$$

Substituting in values from Equation 5.21 and Eq 5.24 gives:

$$Y_{L_{opt}} = \frac{1}{2} (g_o^{ON} + g_o^{OFF}) \sqrt{1 - (2w - 1)^2 \frac{(g_o^{ON} - g_o^{OFF})(g_o^{ON} + g_m^{ON} - g_o^{OFF} - g_m^{OFF})}{(g_o^{ON} + g_o^{OFF})(g_o^{ON} + g_m^{ON} + g_o^{OFF} + g_m^{OFF})}} \quad (5.26)$$

If the “off” transistor is truly off (with $g_m = g_o = 0$), then:

$$Y_{L_{opt}} = g_o^{ON} \sqrt{w - w^2} \quad (5.27)$$

which implies that the optimal load impedance varies versus code from $2r_o$ (at $w = 1/2$) to infinite² (at $w = 0$ or $w = 1$).fP This means the DAC will have gain expansion, as the output impedance increases with code.

To get some intuition about whether this actually matters or not in practice, the transimpedance of the DAC can be considered instead.

The transimpedance of the DAC cell is:

$$Z_T = -\frac{\tilde{y}_{21}}{(Y_L + \tilde{y}_{22})(Y_S + \tilde{y}_{11}) - \tilde{y}_{12}\tilde{y}_{21}} \quad (5.28)$$

²Of course, this should not be too surprising — the optimal output impedance to maximize power gain for a common-gate amplifier (at low frequency) is also infinite.

Again assuming the low-frequency limit for the DAC Y-parameters, the DAC transimpedance for an arbitrary load conductance and source admittance is then:

$$Z_T = -\frac{\frac{1}{2}(2w-1)(g_m^{ON} + g_o^{ON})}{(G_L + \frac{1}{2}g_o^{ON})(G_S + \frac{1}{2}g_o^{ON} + \frac{1}{2}g_m^{ON}) - (\frac{2w-1}{2})^2(g_o^{ON} + g_m^{ON})g_o^{ON}} \quad (5.29)$$

If the DAC were linear, it would have a gain proportional to $\frac{1}{2}(2w-1)$. By normalizing to this gain, the effect of gain expansion is clear:

$$\frac{Z_T(w)}{2w-1} = \frac{2(g_m^{ON} + g_o^{ON})}{(2G_L + g_o^{ON})(2G_S + g_o^{ON} + g_m^{ON})} \frac{1}{1+T} \quad (5.30)$$

Where T , given below in Equation 5.31, represents the loop gain of the two-port circuit[73].

$$T = -(2w-1)^2 \frac{g_o^{ON}}{g_o^{ON} + 2G_L} \frac{g_o^{ON} + g_m^{ON}}{g_o^{ON} + g_m^{ON} + 2G_S} \quad (5.31)$$

The derivation can be repeated for a general set of DAC Y-parameters to arrive at the following:

$$Z_T(w) = -\frac{\tilde{y}_{21}}{(\tilde{y}_{11} + Y_S)(\tilde{y}_{22} + Y_L)} \frac{1}{1+T} \quad (5.32)$$

$$T = -\frac{\tilde{y}_{21}\tilde{y}_{12}}{(\tilde{y}_{11} + Y_S)(\tilde{y}_{22} + Y_L)} \quad (5.33)$$

From this equation, it is clear that if T is real-valued, then there will only be AM-AM distortion in the DAC.

There are two possibilities that can contribute to a non-zero imaginary part for T , and therefore contribute to AM-PM distortion:

1. Nonzero susceptance at input/output ($\text{Im}[Y_L + \tilde{y}_{22}] \neq 0$ or $\text{Im}[Y_S + \tilde{y}_{11}] \neq 0$)
2. The product $\tilde{y}_{21}\tilde{y}_{12}$ has a nonzero imaginary part.

The first possibility is easy to guarantee with proper input/output matching at the desired output frequency, although it is interesting to note that this implies that the AM-PM characteristic of the DAC will degrade away from the center frequency of the match.

A sufficient (but not necessary) ³ condition to rule out the second possibility is to ensure that \tilde{y}_{21} and \tilde{y}_{12} have imaginary parts equal to zero. Referring back to equations 5.17, 5.18, and 5.21, we want the following statements to be true:

$$\text{Im}[\tilde{y}_{21}] = \omega \frac{2w-1}{2} (C_{21}^{ON} - C_{21}^{OFF}) = 0 \quad (5.34a)$$

$$\text{Im}[\tilde{y}_{12}] = \omega \frac{2w-1}{2} (C_{12}^{ON} - C_{12}^{OFF}) = 0 \quad (5.34b)$$

³In particular, the necessary condition is only that $\text{Im}[\tilde{y}_{21}\tilde{y}_{12}] = 0$, which is true when $\text{Im}[\tilde{y}_{21}] = \text{Im}[\tilde{y}_{12}] = 0$, but could also be true (though improbable) if $\sin(\text{Arg}[\tilde{y}_{21}] + \text{Arg}[\tilde{y}_{12}]) = 0$.

Substituting in values from 5.18:

$$C_{21}^{ON} - C_{21}^{OFF} = C_{ds}^{ON} - C_{gs}^{ON} g_m^{ON} R_g^{ON} - C_{ds}^{OFF} + C_{gs}^{OFF} g_m^{OFF} R_g^{OFF} = 0 \quad (5.35a)$$

$$C_{12}^{ON} - C_{12}^{OFF} = C_{ds}^{ON} + C_{gd}^{ON} g_m^{ON} R_g^{ON} - C_{ds}^{OFF} - C_{gd}^{OFF} g_m^{OFF} R_g^{OFF} = 0 \quad (5.35b)$$

C_{ds} is a purely linear parasitic capacitance, and will be the same in both the on and off states of the CG unit element. So, the conditions become:

$$C_{21}^{ON} - C_{21}^{OFF} = -C_{gs}^{ON} g_m^{ON} R_g^{ON} + C_{gs}^{OFF} g_m^{OFF} R_g^{OFF} = 0 \quad (5.36a)$$

$$C_{12}^{ON} - C_{12}^{OFF} = C_{gd}^{ON} g_m^{ON} R_g^{ON} - C_{gd}^{OFF} g_m^{OFF} R_g^{OFF} = 0 \quad (5.36b)$$

Since C_{gs} , C_{gd} , g_m , and R_g will change somewhat depending on the bias point, not necessarily in an easily analyzed way, it seems unlikely that the conditions of Equation 5.37 will be met exactly. This can be verified from simulation by computing the phase angle of $\tilde{y}_{12}\tilde{y}_{21}$. Although the phase is not zero, it is fairly small (6.8° at 120GHz as shown in Figure 5.17) indicating that the imaginary part of $\tilde{y}_{12}\tilde{y}_{21}$ is about 8.4 times as small as the real part.

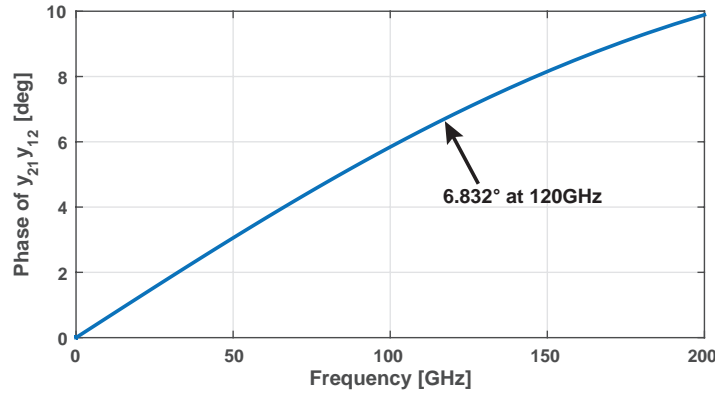


Figure 5.17: Phase of $\tilde{y}_{21}\tilde{y}_{12}$

If the susceptances are matched at the DAC input and outputs, then the phase of the loop gain will be equal to the phase of $\tilde{y}_{21}\tilde{y}_{12}$. Of course, the amount of AM-PM conversion this results in depends on the magnitude of T . To try to estimate this, assume the real part of the load admittance is provided by the parallel equivalent resistance of an inductor at the load ($G_L = \frac{\omega C_{22}}{Q_L}$) and the source impedance is matched according to Equation 5.23a. The magnitude of the resulting loop gain versus frequency under various conditions is shown in Figure 5.18.

The loop gain is then used to calculate AM-AM and AM-PM nonlinearity of the DAC itself (Figure 5.19). It can be seen that with realistic Q tuned loads at the gate and source of the DAC, the AM-PM linearity becomes quite small, and the AM-AM linearity improves, although it is still problematic. This is perhaps not catastrophic, though, since the transmitter power amplifier will likely have a compressive gain characteristic, suggesting there is potential to cancel out the nonlinearity.

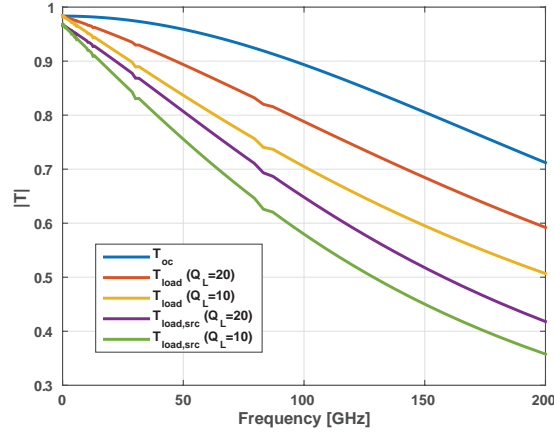


Figure 5.18: Magnitude of DAC loop gain under various loading conditions

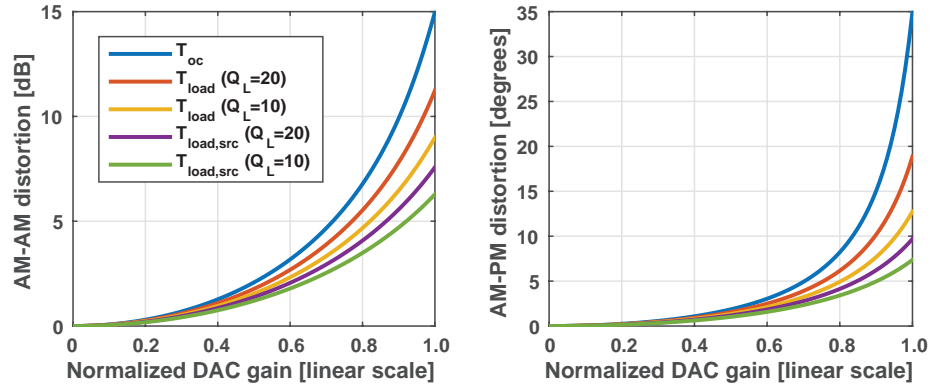


Figure 5.19: AM-AM and AM-PM nonlinearity of DAC

5.3.1.3 IQ DAC Model

The above analysis has focused on amplitude modulation only, but can easily be extended to QAM. To achieve quadrature modulation, separate I and Q DACs are used, and current-combined at their outputs (Figure 5.20). The KCL equations governing this circuit are

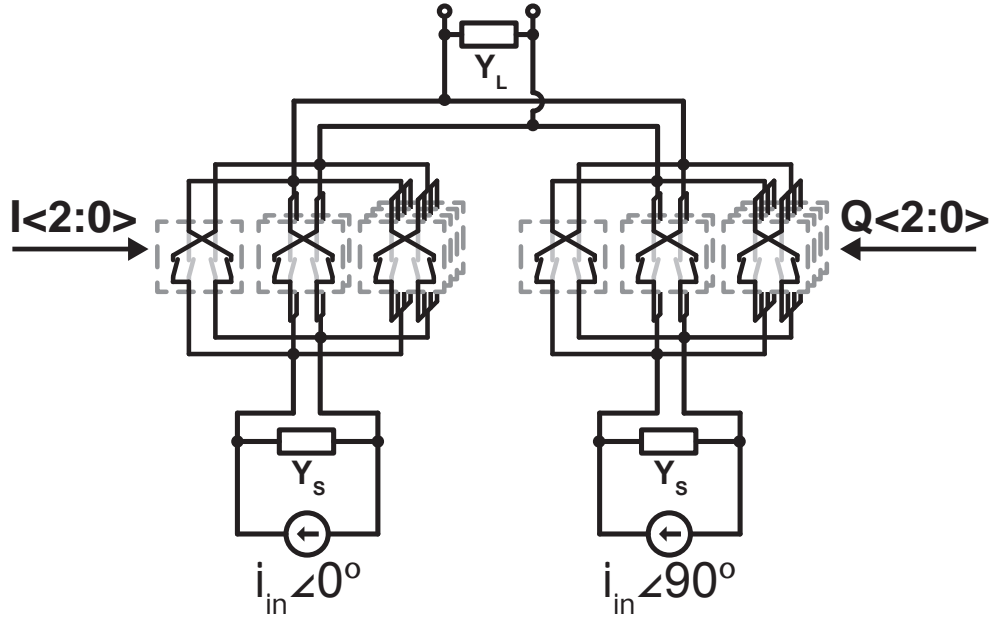


Figure 5.20: I and Q DACs with current-combining at output

derived and used to compute the transimpedance ($Z_T = v_o/i_{in}$).

$$\begin{bmatrix} i_1^I \\ i_2^I \end{bmatrix} = \tilde{\mathbf{Y}}^I \begin{bmatrix} v_1^I \\ v_o \end{bmatrix} \quad (5.37a)$$

$$i_1^I = i_{in} e^{j0^\circ} - v_1^I Y_S \quad (5.37b)$$

$$\begin{bmatrix} i_1^Q \\ i_2^Q \end{bmatrix} = \tilde{\mathbf{Y}}^Q \begin{bmatrix} v_1^Q \\ v_o \end{bmatrix} \quad (5.37c)$$

$$i_1^Q = i_{in} e^{j90^\circ} - v_1^Q Y_S \quad (5.37d)$$

$$v_o Y_L = -i_2^Q - i_2^I \quad (5.37e)$$

Simplifying the above, the following expression results:

$$Z_{T,complex} = -\frac{\tilde{y}_{21}^I + j\tilde{y}_{21}^Q}{(Y_L + 2\tilde{y}_{22})(Y_S + \tilde{y}_{11}) - \tilde{y}_{21}^I \tilde{y}_{12}^I - \tilde{y}_{21}^Q \tilde{y}_{12}^Q} \quad (5.38)$$

Or, written in terms of the loop gain:

$$Z_{T,complex} = -\frac{\tilde{y}_{21}^I + j\tilde{y}_{21}^Q}{(Y_L + 2\tilde{y}_{22})(Y_S + \tilde{y}_{11})} \frac{1}{1 + T_{IQ}} \quad (5.39)$$

Where

$$T_{IQ} = -\frac{\tilde{y}_{21}^I \tilde{y}_{12}^I + \tilde{y}_{21}^Q \tilde{y}_{12}^Q}{(Y_L + 2\tilde{y}_{22})(Y_S + \tilde{y}_{11})} \quad (5.40)$$

5.3.1.4 DAC Layout

The final RF DAC design uses three bits for each of I and Q, so that QPSK, 16QAM, and 64QAM modulations can be supported.

The 3-bit RF DAC consists two “half-DAC” cells, each of which consists of seven unit elements (Figure 5.21). Each “half-DAC” has two source connections but only one drain connection. This means that drain connections of adjacent devices can be shared, which minimizes parasitic capacitances on the drain and results in a higher bandwidth at the output. Elements are grouped together in a binary weighted fashion (Figure 5.22). The LSB consists of only one unit element and is located in the center of the DAC. The upper two DAC bits are distributed around the center of the DAC in a symmetric way. CMOS inverters are included in each unit element so that the impedance seen by the gate of each transistor are all identical.

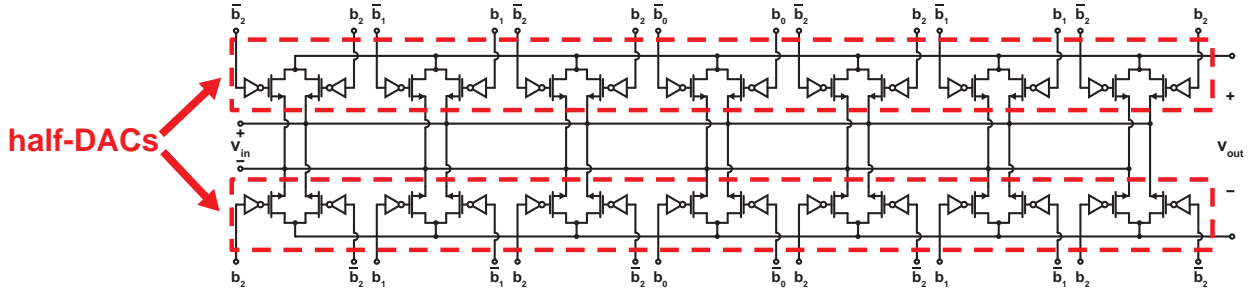


Figure 5.21: DAC schematic highlighting “half-DAC” cells.

Because the DAC layout is fairly long, the series inductance of the drain and source routing lines between each unit element may have an effect on overall DAC performance. To model this effect without adding significant simulation time and complexity, local drain/source routing at the unit element level is extracted using the PDK’s standard RC parasitic extraction flow. Intermediate routing, including the long source/drain buses on thick metal layers (Figure 5.23), is modeled using Integrand EMX which accounts for self and mutual inductances between the routing buses.

5.3.1.5 Current Re-Use

This DAC structure can be used on its own, with the sources of the devices DC-connected to ground and separate I and Q LO signals magnetically coupled into the sources (Figure 5.24a). This is not ideal from an energy efficiency point-of-view if the DAC gates are driven by full-rail CMOS digital signals. A V_{GS} of 1V is enough to drive a 28nm transistor well into velocity saturation, which means the g_m/I_D of the transistor will be reduced. This might be an acceptable penalty if the f_T of the transistor showed a corresponding improvement, but in fact the f_T only shows a small increase beyond the onset of velocity saturation, because the slope of the transistor $I_D - V_{GS}$ curve (i.e., g_m) is nearly constant in that region. One

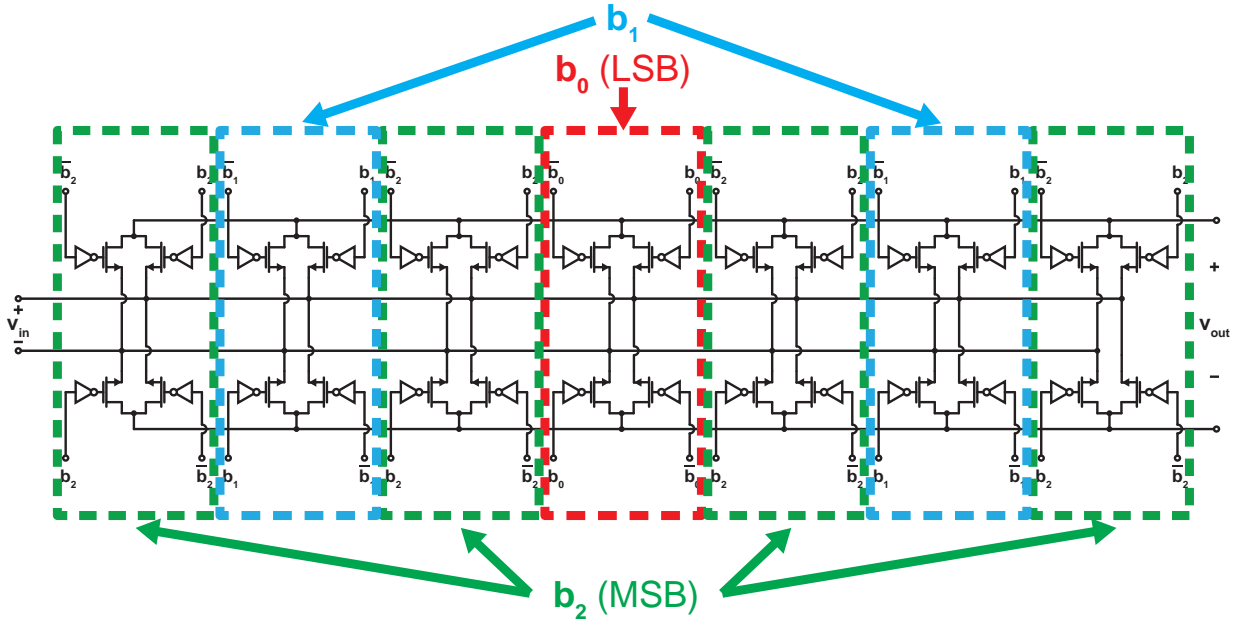


Figure 5.22: DAC schematic highlighting segmentation of bits within DAC.

way to improve the efficiency of the circuit could be to switch to a higher V_T transistor flavor for the DAC, reducing the effective overdrive and bias current, but reducing the headroom at the output without getting much benefit in return.

Another way of improving energy efficiency (and the approach used here) is to use device stacking to reduce the overdrive in the DAC devices. A neutralized common-source amplifier can be stacked underneath each I and Q DAC to provide additional power gain while reusing the same current from the DAC devices (Figure 5.24b) [70]. This actually consumes less DC current than the non-stacked version (assuming 1V gate bias on the DAC devices) as the overdrive of the common-gate devices is reduced, but f_T is only slightly degraded. Since the amplifier is not a true cascode, but is being switched periodically in time, neutralization of the common-source amplifier is helpful in ensuring stability of the modulator, and also reduces sensitivity to the impedance driving the amplifier.

The power gain of the stacked amplifier-modulator can be further increased by adding a matching network between the amplifier and modulator stages. This matching network should perform two goals: to resonate out the capacitances of the amplifier and modulator, and to transform the real part of the input and output admittances of the two blocks. A simple solution is to add a differential series inductance or transmission line between the two blocks, which will (along with the parasitic capacitances of the devices) form an artificial quarter-wavelength transmission line (Figure 5.24c).

Because the common-gate DAC is terminated with a non-zero load impedance, its input

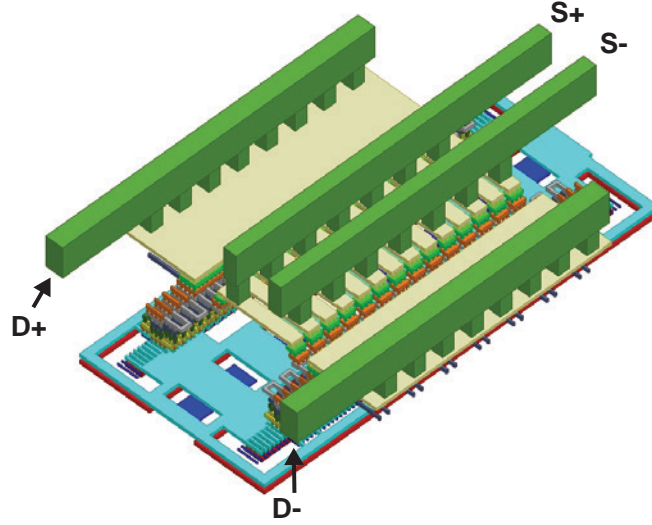


Figure 5.23: Layout of DAC cell including top level metal routing.

impedance will be higher than the $1/g_m$ that is usually assumed [74]:

$$Z_{in} = \frac{R_L + r_o}{1 + (g_m + g_{mb})r_o} \quad (5.41)$$

For the DAC design selected above, this works out to about 100Ω with a realistic Q inductive load. A differential, neutralized common-source amplifier biased at the same current density has a real part of its output impedance of about 200Ω . So, the required line impedance for a quarter wave match is 141Ω , or even higher if it is necessary to absorb the device capacitances into the effective transmission line. It is not practical to achieve impedances that high on-chip without incurring significant losses.

A way of getting around this problem is by combining the two common-source amplifiers into one, and using the matching network itself to generate the quadrature LO signals (Figure 5.24d).

5.3.1.6 Lumped-Element Differential Quadrature Hybrid

Since the quadrature hybrid is stacked in between the common-gate modulator and common-source amplifier, it should provide a suitable DC connection from one to the other. A branchline hybrid [75] provides the right DC connections and is easily implemented on-chip using transmission lines (Figure 5.25a). Branchline hybrids can also be constructed out of artificial transmission lines [76][77] to reduce the area required, or to absorb parasitic capacitances.

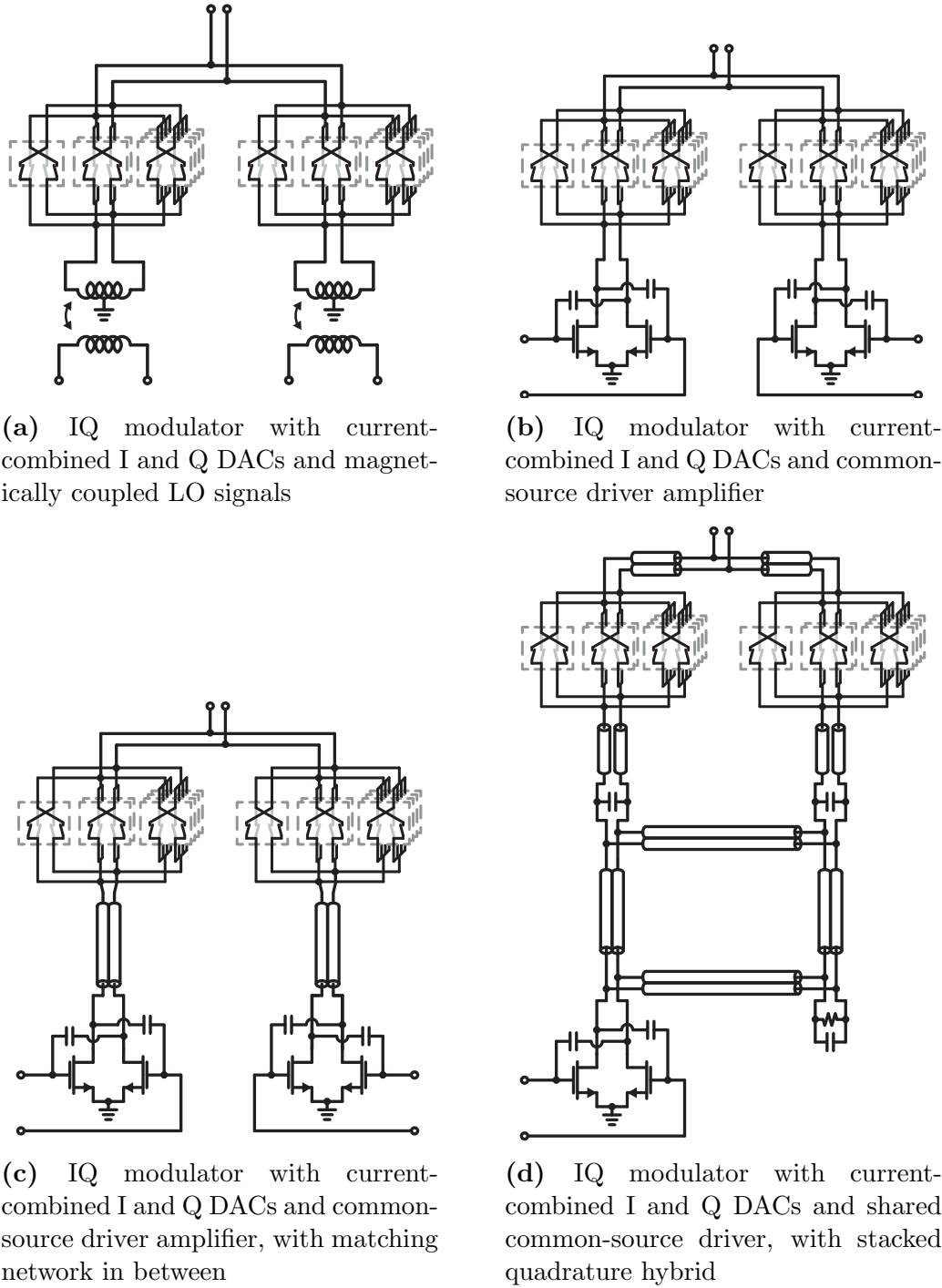


Figure 5.24

The initial quadrature hybrid design was a simple differential lumped-element branchline hybrid (Figure 5.26a), similar in concept to the one designed by Kang and Thyagarajan [77].

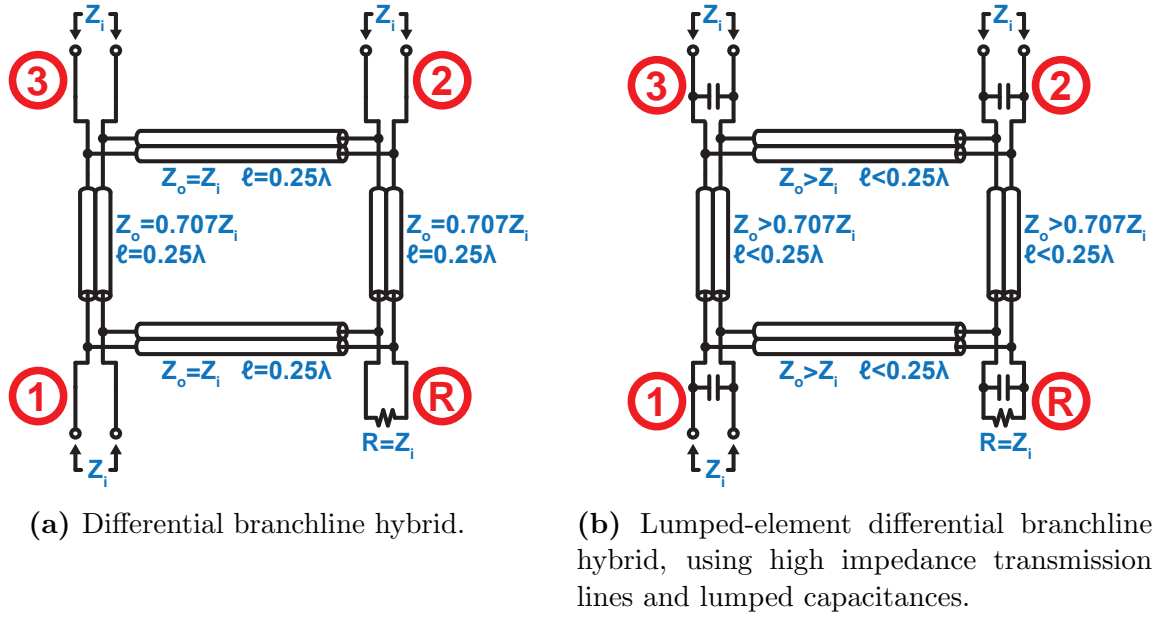


Figure 5.25

However, to satisfy the stringent metal density rules in 28nm, the differential transmission lines are surrounded by grounded metal regions on both sides (becoming differential coplanar waveguide (CPW), rather than coplanar stripline). Thin metal strips periodically connect the two ground conductors of the differential CPW to suppress the slotline mode, and also help to satisfy metal density requirements on the lower metal layers.

The gap between the two conductors in the transmission lines is fairly large to achieve

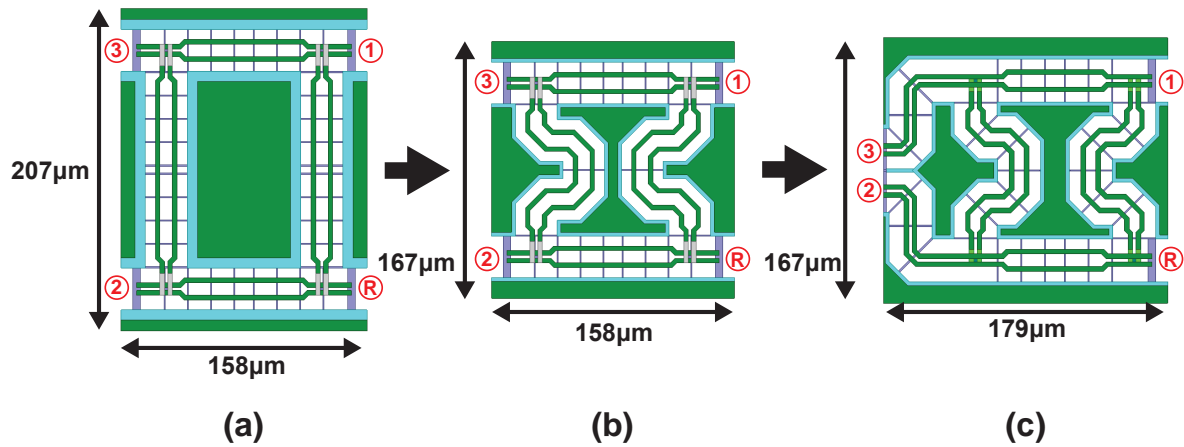


Figure 5.26: Design iterations for quadrature hybrid

high impedance, but this means the “tee” junctions between the horizontal/vertical lines will have a large amount of parasitic inductance which is not symmetric across the two conductors. To avoid this asymmetry and the DM-CM mode conversion issues it would present, the transmission line gaps taper close to the junctions to minimize the added parasitic inductances and their inherent asymmetry.

Unfortunately, the initial design (Figure 5.26a) has the wrong aspect ratio for the chip floorplan. In particular, it is too large to be able to fit between two flip-chip bumps. This problem can easily be resolved by meandering the longer transmission lines to reduce the size of the hybrid in the vertical dimension (Figure 5.26b). The revised design just barely fits between the two flip-chip bumps (Figure 5.27).

There is one more issue with the hybrid design: the I and Q DACs will be fairly far apart (nearly $150\mu m$). This means it will be difficult to current-combine the outputs of the I and Q DACs without introducing significant routing parasitics. The solution shown here is to add another short transmission line routing segment to each of the I and Q hybrid outputs, that moves the two outputs closer together (Figure 5.26c). This also can perform a little impedance transformation, stepping up the 80Ω impedance of the hybrid to a slightly higher 100Ω . By stepping up the impedance the DAC sizing can be reduced by the same factor, lowering both the DC power of the DAC itself and the digital switching power used to drive the DAC.

The final simulated S-parameters are plotted in Figure 5.28, along with the gain and phase mismatch. An insertion loss of between 5 and 5.5dB to each of the I and Q ports is achieved over the 110GHz to 120GHz band. I and Q output amplitudes are within 0.4dB and phase error is within 5° (-3° to $+5^\circ$) between 110GHz to 120GHz. The input match is quite broadband, as is the isolation — both are below -10dB over the extended range of 100GHz to 130GHz.

5.3.1.7 Output Matching

To couple to the input stage of the power amplifier, a low-k transformer is used. The primary goal of the matching network design is to provide a broad-bandwidth resistive load around 120GHz to the modulator. As described in section 5.3.1.2, it is important to minimize $\text{Im}[Y_{in}^{balun} + \tilde{y}_{22}^{DAC}]$ to avoid any AM-PM distortion and associated IQ pulling effects. So, the input impedance of the balun (loaded with the PA input capacitance) was designed to have dual resonances surrounding the center frequency to provide a broadband match (Figure 5.31a).

A long series transmission line connects the I and Q DAC output terminals to the primary of the balun. This transmission line is necessary because a large space is required to route the DAC digital input signals. A shielding ground layer is added underneath the transmission lines and above the DAC digital signals to prevent any coupling and improve the overall modeling accuracy.

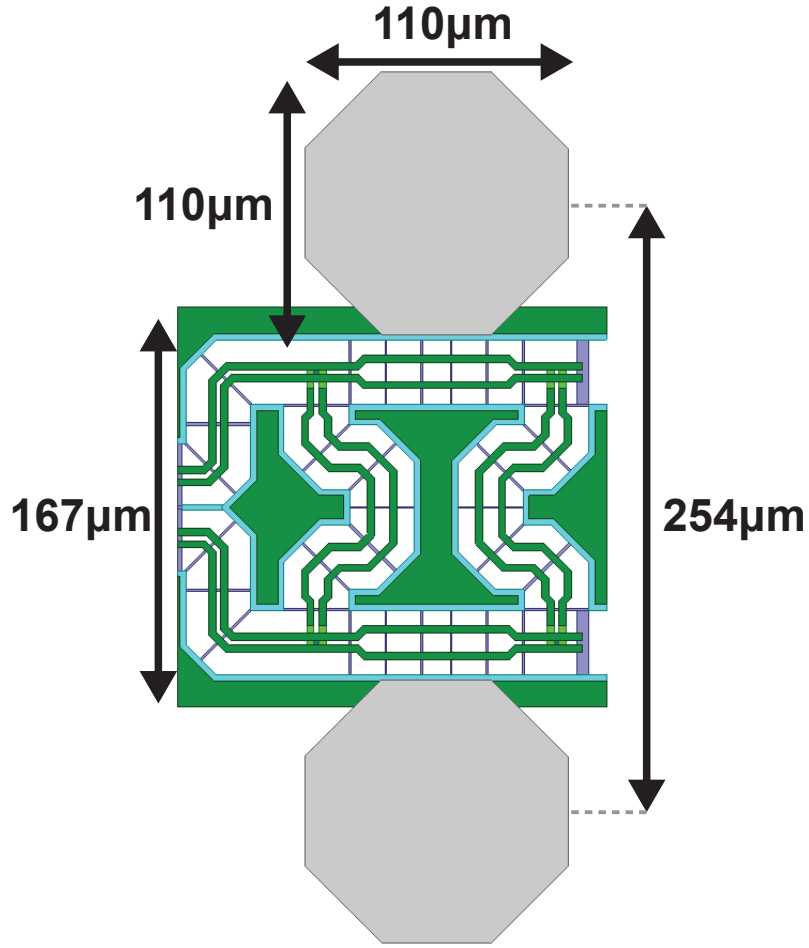


Figure 5.27: Quadrature hybrid layout in between flip-chip pads

5.3.2 Power Amplifier Design

The power amplifier uses a two stage design (Figure 5.30), with each stage based around the neutralized common source unit cell described in Section 5.2.1. The first amplifier stage sizing was determined by the output capacitance of the modulator — equalizing capacitances on both sides of the balun maximized the bandwidth. For the second amplifier stage, the sizing was increased relative to the first stage. This increased capacitance lowered the bandwidth of the interstage matching network, with the benefit of increasing the saturated output power of the amplifier.

The simulated input (S_{11}) and output return loss (S_{22}) of the power amplifier are shown below in Figure 5.31. The output return loss also includes an electromagnetic simulation model of the chip-to-PCB packaging transition, so the S_{22} plotted below is actually the return loss seen on the PCB after the transition. Because of the low output impedance of the PA, and the use of low-K transformer matching, an extremely broad output match has

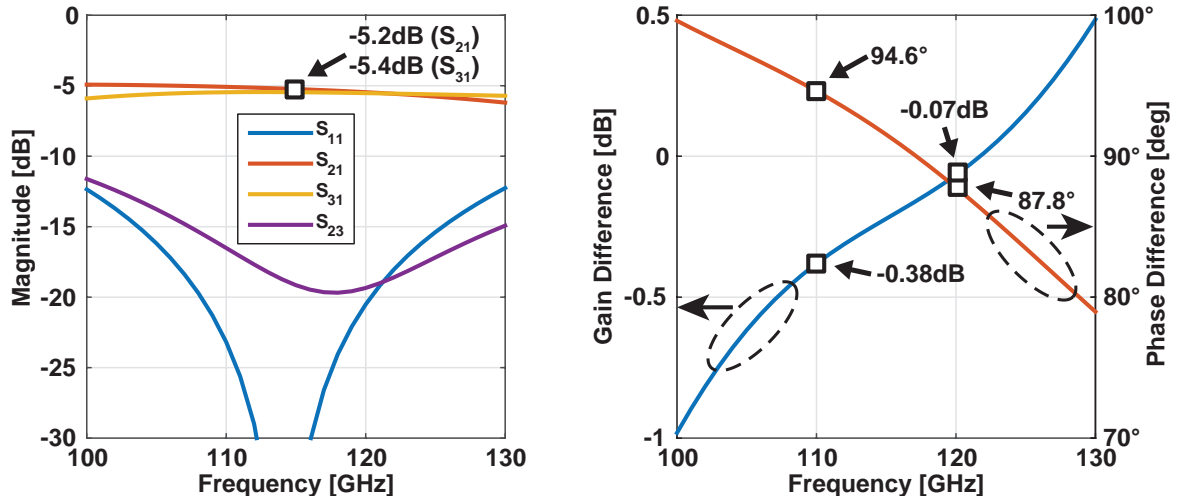


Figure 5.28: Simulated S-parameters of final quadrature hybrid design.

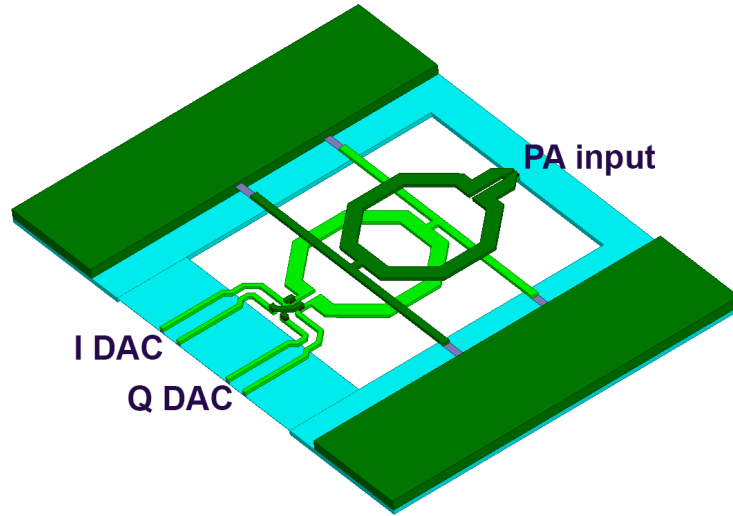


Figure 5.29: DAC output transformer / PA input transformer

been achieved, with over 30GHz of -10dB S_{22} bandwidth.

The small-signal gain and group delay of the PA are plotted in Figure 5.32, again including the effects of the chip-to-PCB transition. The power amplifier has a simulated peak gain of 9.5dB, with a 3dB bandwidth of 17GHz (from 108 to 125 GHz). Over this band, the PA exhibits a group delay variation of only 5ps.

The large signal performance of the PA peaks at 115GHz, where it has a simulated peak PAE of 9.5% (Figure 5.33a). The efficiency curve vs frequency tracks the small-signal gain characteristic of the PA (Figure 5.33b).

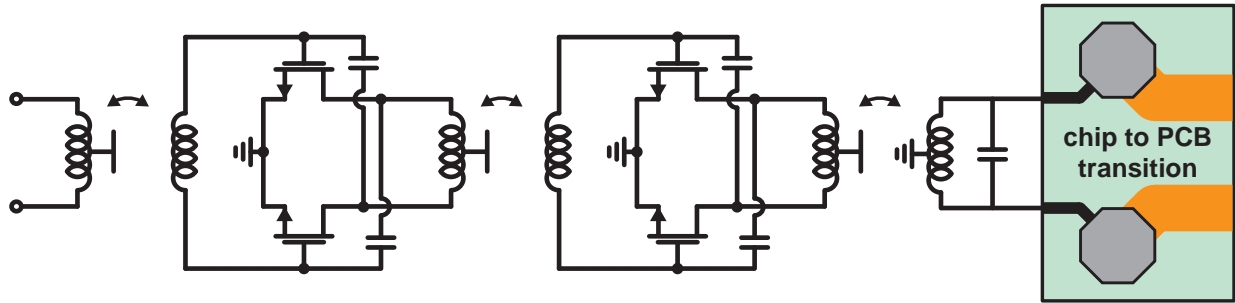


Figure 5.30: Power Amplifier Schematic

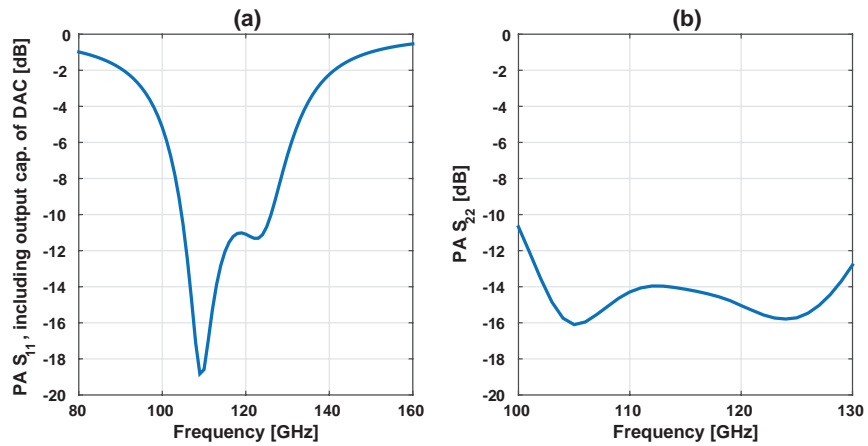


Figure 5.31: Simulated PA input and output return losses.

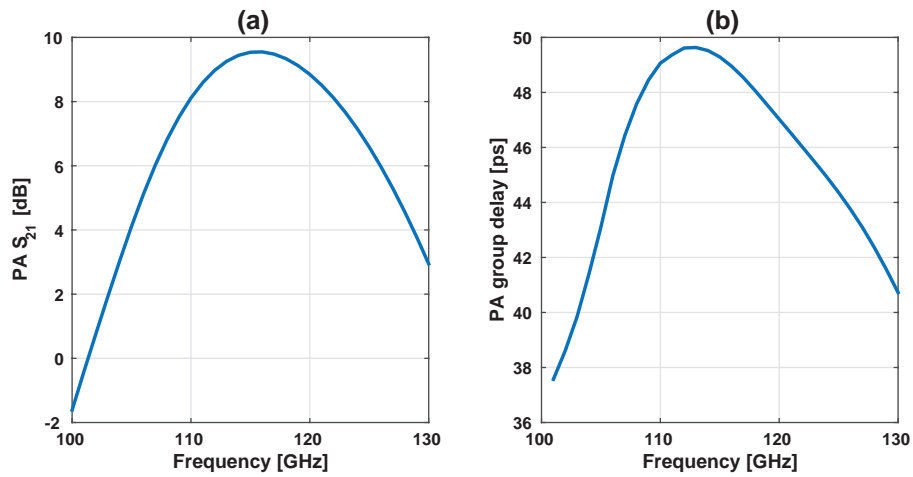


Figure 5.32: Simulated PA gain and group delay.

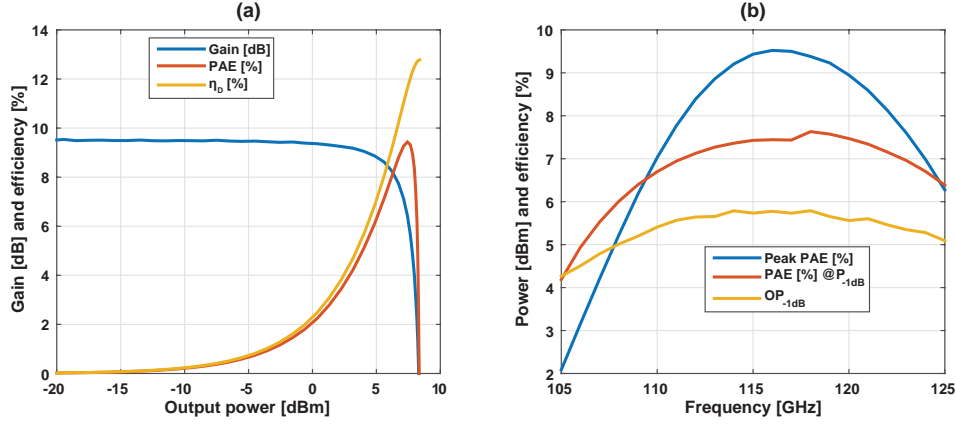


Figure 5.33: Simulated PA efficiency and gain compression curves.

5.3.3 TX Chain Simulated Results

To make sure that the DAC-PA combination can adequately support broadband modulation, the full transmitter chain was simulated in each of the three supported modulation settings (QPSK, 16QAM, and 64QAM).

5.3.3.1 TX Chain Static Characterization

First, to verify the DAC provides sufficient linearity across frequency, a quasi-static small-signal characterization was performed. This computes the gain and phase in steady state if the TX were to continuously transmit the same code indefinitely. Each DAC code was simulated at frequencies from 110GHz to 120GHz (in 1GHz steps) to verify that the I-Q pulling effects in the DAC do not significantly degrade its EVM. The results are shown below in Figure 5.34. The ideal constellation points for each modulation scheme are represented by the black triangles, and the simulated complex gain in each DAC state is represented by the multicolored dots (each color represents the gain at a different frequency).

By visual inspection, it is clear that there is significant gain expansion in the 64QAM constellation at higher codes. This was predicted by the proposed DAC nonlinearity mechanism described earlier. In QAM16 mode, the gain expansion is suppressed somewhat because the LSBs of the I and Q DACs are totally disabled, reducing the loop gain of the DAC and somewhat mitigating the gain expansion problem. In QPSK mode, gain expansion affects all codes equally, so it has little impact on the overall linearity. Some other DAC nonidealities can be noticed in the QPSK constellation, such as DC offset due to some inherent asymmetries in the DAC layout, and I-Q amplitude imbalance from the quadrature hybrid.

5.3.3.2 TX Chain Dynamic Characterization

Finally, to verify that the target data rate can be supported by the transmitter, the TX chain was simulated in time domain with a pseudo-random symbol sequence in each modulation

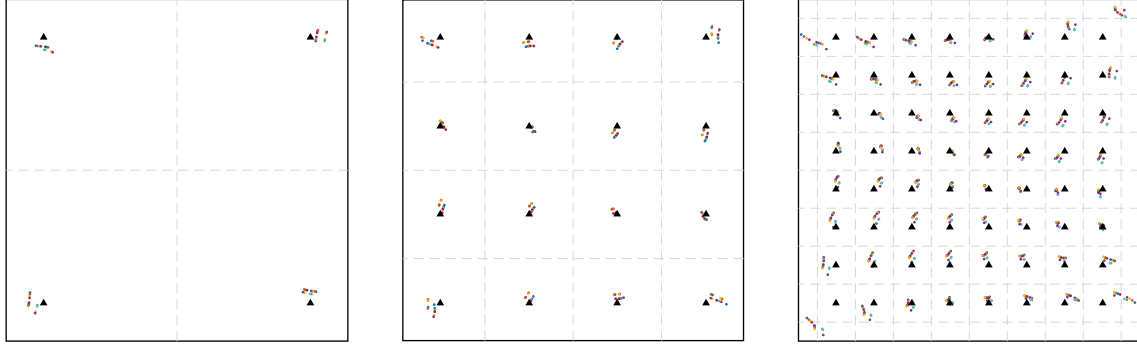


Figure 5.34: Static TX constellations from 110GHz to 120GHz

setting.

Simulations predict that 50Gb/s data rate should be achievable using either QPSK or 16QAM modulation. In QPSK mode, there is a fair amount of inter-symbol interference (ISI) coming from the transmitter itself (Figure 5.35) which suggests that this cannot be pushed much faster. However, in 16QAM mode, due to the lower symbol rate, the ISI is reduced (Figure 5.36). This represents the potential to improve the data rate of the link further, if the SNR at the receiver is sufficiently high.

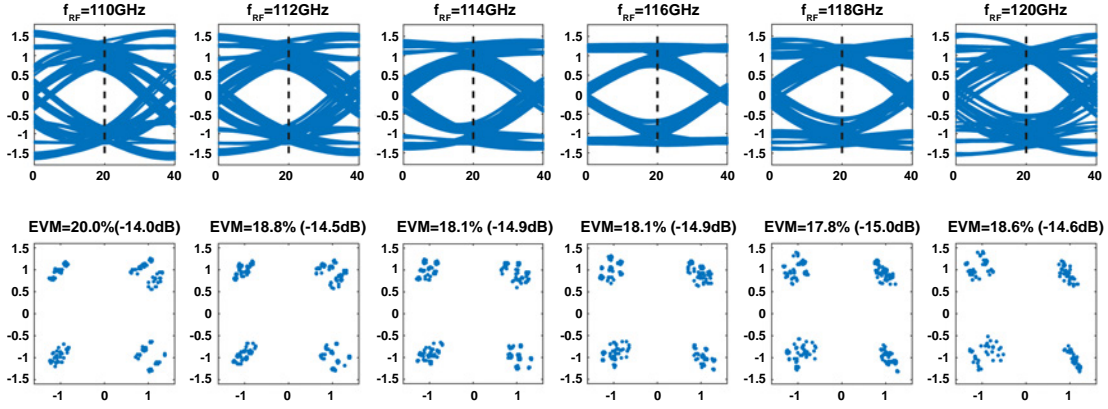


Figure 5.35: Simulated QPSK eye diagram and constellations for 40ps symbol period, corresponding to 50Gb/s data rate

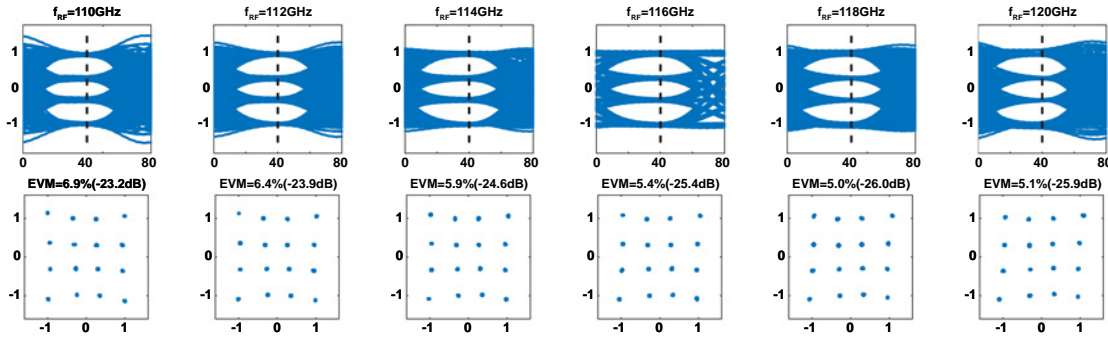


Figure 5.36: Simulated 16QAM eye diagram and constellations for 80ps symbol period, corresponding to 50Gb/s data rate

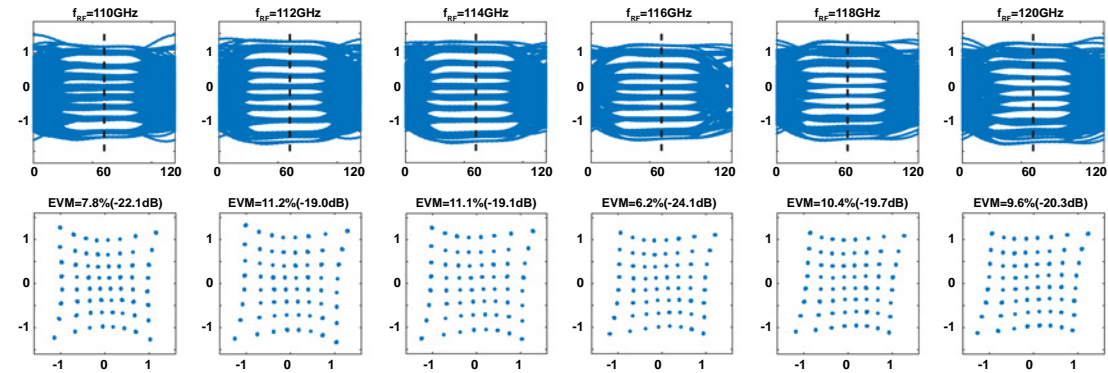


Figure 5.37: Simulated 64QAM eye diagram and constellations for 120ps symbol period, corresponding to 50Gb/s data rate

5.4 Receiver

5.4.1 Baseband Amplification

The baseband amplification chain and output driver were designed by Sashank Krishnamurthy, and are briefly summarized here. The main amplifier chain is designed around an inverter-based Cherry-Hooper amplifier core [78]. This allows maximizing both gain and bandwidth, rather than the usual gain-bandwidth tradeoff of a single-stage amplifier. Because the inverter-based amplifier does not have any inherent common-mode rejection, a unity-gain differential buffer amplifier has been inserted before each Cherry-Hooper stage to provide additional common-mode rejection. Series DC-blocking capacitors are used to decouple common-mode voltages between stages and reduce the impact of DC offset voltages in each amplifier stage.

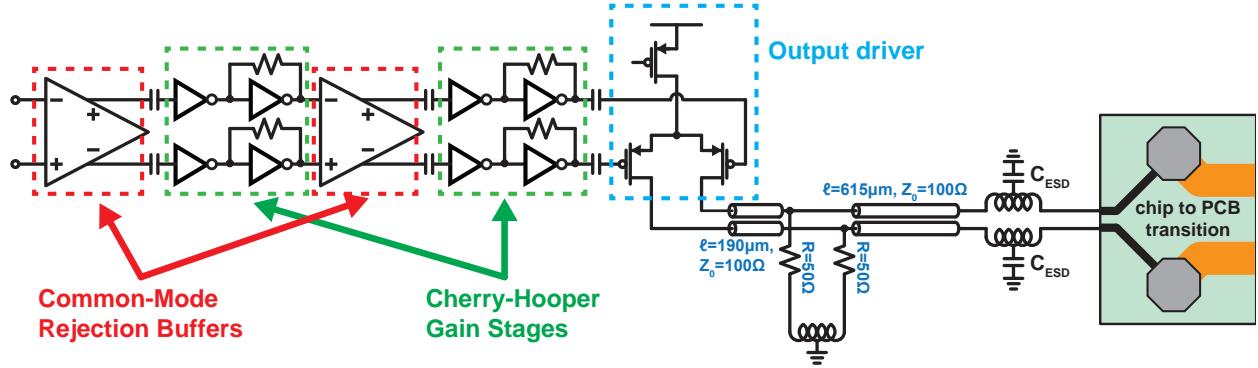


Figure 5.38: Baseband amplifiers and output driver

Following the Cherry-Hooper stages, a PMOS output driver is used to drive the differential signal off chip. A PMOS output driver was selected so that the output signal from the chip will be truly ground referenced, rather than referenced to ground through an on-chip bypass capacitor (as in the NMOS case). A shunt peaking inductor is placed along the transmission line to extend the bandwidth of the chain. ESD protection at the output is provided using a series inductor with a capacitor at the center tap. This behaves as an artificial transmission line and absorbs the impact of the ESD capacitance without impacting the bandwidth too much.

The simulated performance of the baseband amplification chain predicts an in-band gain of 24.7dB, with a 3dB bandwidth of 13.5GHz. Because of additional losses induced by the skin effect in the peaking inductor and transmission line routing, the gain drops slightly from 24.7dB to 24.5dB from 100MHz and 5GHz. The DC blocking capacitors in between stages result in a low-frequency 3dB cutoff of 7.6MHz.

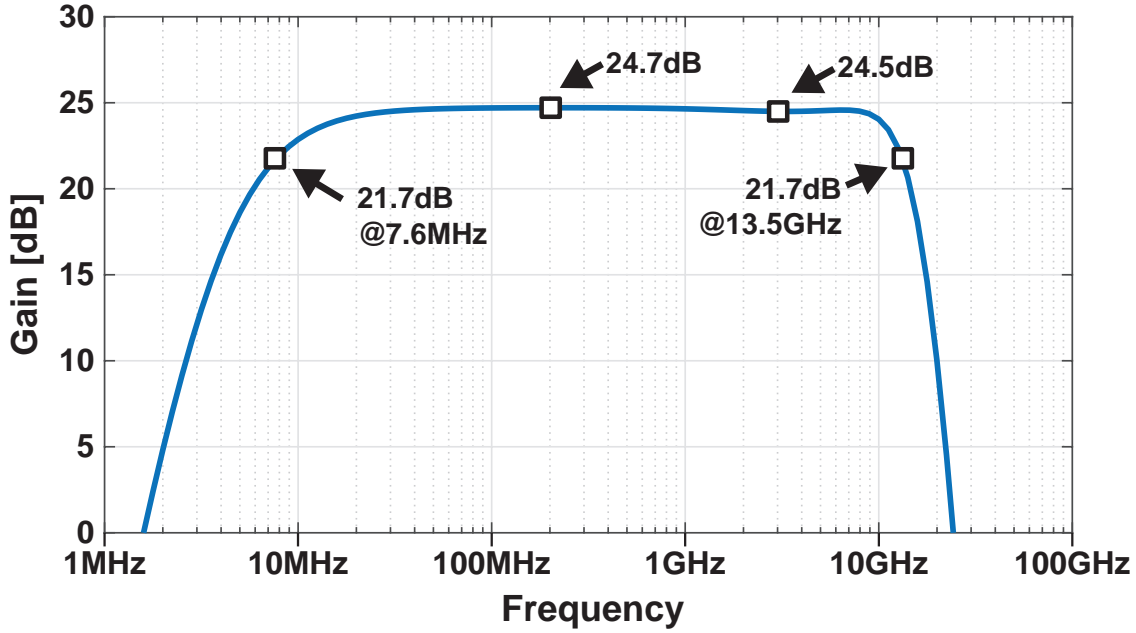


Figure 5.39: Voltage gain of baseband amplifier and output driver (including transmission line routing)

5.4.2 Active Mixer with TIA load

It has been demonstrated that as the carrier frequency approaches the f_{max} of the technology, a passive mixer can provide better noise performance than an active mixer [79]. Additionally, because the passive mixer input impedance is an upconverted version of its baseband load impedance, it is also straightforward to provide a wideband input impedance for the mixer. However, in 28nm at 120GHz, the operation frequency is still only about half of f_{max} , so there may still be some benefit in using an active mixer instead of a passive mixer.

For a fair comparison between active and passive mixers, the transconductor device typically present in an active Gilbert-cell mixer is omitted here. Instead, the RF signal can be transformer coupled into the sources of the Gilbert-cell. This can easily be done for both passive and active mixers.

If the mixer operation is viewed as purely current-commutating, there is no significant difference in conversion gain between a passive and active mixer [80]. All of the RF current is commutated to baseband and flows through the baseband load. Therefore, the transimpedance gain of the mixer is set only by the baseband load resistance, regardless of whether it is an active or passive mixer. However, this assumes the input signal is a true current source, with infinite source impedance.

If the source impedance is finite, then there is a current divider between the source impedance and mixer input impedance. There is also effectively a “current commutation gain” which represents the fraction of current that gets downconverted to baseband, depen-

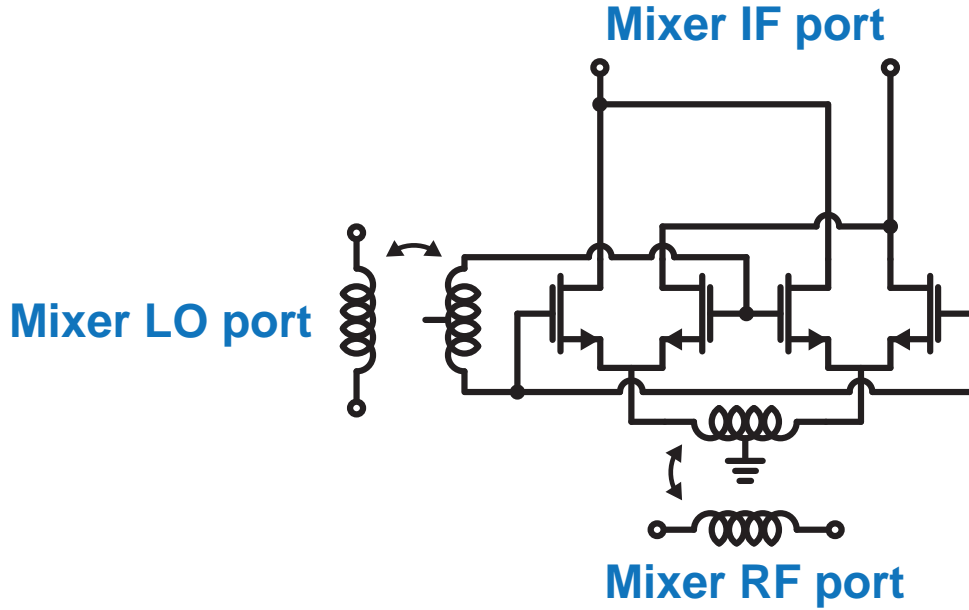


Figure 5.40: Double balanced active mixer core schematic

dent on the swing on the LO port. So, the transimpedance gain is determined by the product of these two factors, multiplied by the baseband load impedance (Equation 5.42).

$$Z_T = \frac{Z_{out(LNA)}}{Z_{in(mixer)} + Z_{out(LNA)}} \cdot G_I \cdot Z_{L,BB} \quad (5.42)$$

Unlike in a passive mixer, an active mixer's input impedance has little dependence on its load impedance. So, it is possible to reduce the input impedance without changing the load impedance, leading to a higher overall effective transimpedance gain.

One final drawback of a passive mixer is that its conversion gain is more sensitive to the LO swing. For an efficiency-optimized high frequency design, it is advantageous to use an active mixer instead to save power on the mixer LO driver.

Typically, a current-commutating mixer assumes a hard-switching LO waveform - usually a large-amplitude square wave with fast rise and fall times. However, the mixer LO waveform at millimeter-wave will be unlikely to be a square wave. A noise and gain analysis by Terrovitis assumes square wave LO drive, with some approximations for large sinusoidal LO signals [81]. For moderate or low LO signal levels, these approximations do not yield accurate predictions.

Another paper by Melly *et al* provides more detailed modeling of sinusoidally driven mixers [82]. Conversion gain can be analytically computed in the subthreshold (weak inversion) and strong inversion modes of operation (Equations 5.4.2). The analysis becomes too difficult for operation in moderate inversion or operation that crosses the boundaries of

these regions of operation. For large LO swing in any region of operation, the conversion gain is simply $\frac{2}{\pi}$.

$$G_I = \frac{V_{LO}}{2n \frac{kT}{q}} \quad (\text{Weak inversion, low LO swing}) \quad (5.43a)$$

$$G_I = \frac{V_{LO}}{4n \frac{kT}{q}} \sqrt{\frac{2I_S}{I_{tail}}} \quad (\text{Strong inversion, low LO swing}) \quad (5.43b)$$

$$G_I = \frac{2}{\pi} \quad (\text{High LO swing}) \quad (5.43c)$$

There are a few caveats to the equations above [82]:

1. Quasistatic devices are assumed, and therefore the equations may not necessarily agree with simulations at higher frequencies.
2. Truly differential pairs are assumed, rather than the pseudodifferential pairs of Figure 5.40.

The “typical” CMOS Gilbert-cell mixer design has been analyzed in detail for both conversion gain and noise figure. However, this design is not quite typical in that this Gilbert-cell mixer does not use a tail device to set the bias point, nor does it use transconductor devices to provide current gain. To save supply headroom, instead the sources of the mixer devices are all DC-connected to ground. The RF signal can be magnetically coupled into the sources using a transformer.

Because of this choice of topology, with large-signal LO drive, the DC current of the mixer will shift away from the quiescent bias point. Under these large signal conditions, the conventional mixer analysis of equation is not strictly valid, but it serves as a good design approximation.

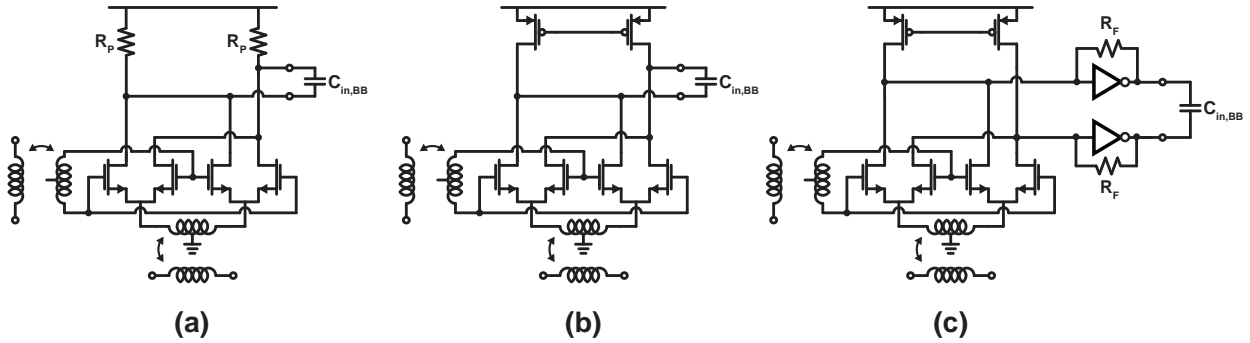


Figure 5.41: Different strategies for the DC load of the mixer: Resistor load (a), PMOS load (b), PMOS load with TIA (c).

The output of the mixer needs to be DC-connected to either a resistive (Figure 5.41a) or active (Figure 5.41b) load. According to Equation 5.42, the mixer load impedance should be maximized to maximize its transimpedance gain. With a fixed load capacitance (set by the input capacitance of the baseband amplification), this then determines the gain-bandwidth tradeoff of the mixer. However, if large bandwidth is desired, there is not a significant difference in the range of possibilities between active and resistive loads (Figure 5.42a). So, given that the output of the mixer goes directly to the pseudo-differential baseband amplification stage, the active load is preferred since it can easily be adjusted to set the output common mode to the desired level.

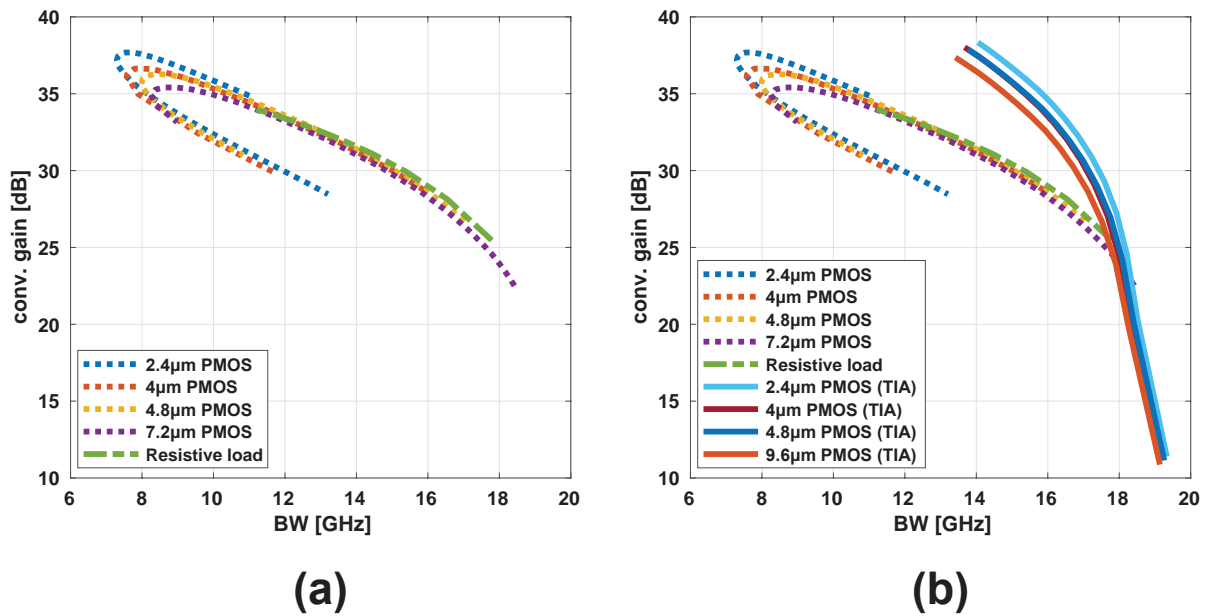


Figure 5.42: Range of gain and bandwidth possibilities for different mixer baseband loads.

However, there is another option that can break the gain-bandwidth tradeoff. An intermediate gain stage, in the form of a transimpedance amplifier, can provide a low input impedance to the mixer output, while also driving the input capacitance of the baseband amplifier with high bandwidth. By varying the feedback resistance R_F in the inverter-based TIA, a curve of possible values of gain and bandwidth can be traced out (Figure 5.42b). For a desired mixer bandwidth of 18GHz or more, the feedback resistance is low and the added loading of the TIA becomes harmful rather than helpful. However, for a mixer bandwidth of 14-17GHz, the TIA is helpful in providing higher gain for the same bandwidth. Similarly, for a given gain requirement, the bandwidth in the TIA-based mixer is strictly higher. The only penalty is DC power, which (for a small size inverter) is not a significant increase overall.

Unfortunately, due to layout-related complications, some more effort is required to achieve this performance. To support broad bandwidth at the baseband output, the output pads themselves need to be on the perimeter of the die, so they can be easily and cleanly routed

away from the chip on the PCB. With a limited number of IO bumps available, this meant that the I and Q output pads for a single channel needed to be on the same side of the chip. To make this happen, the I signal (at either LO, RF, or baseband) needed to cross underneath the Q signal. The least painful and most symmetric way to do this is at baseband, though it still requires some care in layout.

The layout floorplan for the mixer (Figure 5.44) illustrates the dilemma. The baseband output signal from the “Q” mixer is routed underneath the “I” LO input. For matching of performance between I and Q, the I mixer is routed in a similar way, with a matched path length. Effectively, this adds an RC routing line at the output of the mixer (Figure 5.43), which has a gain and bandwidth impact. This can be overcome somewhat by adding a second TIA at the end of the RC routing, to absorb some of the routing capacitance with a lower impedance termination and drive the baseband amplifier input with a high bandwidth.

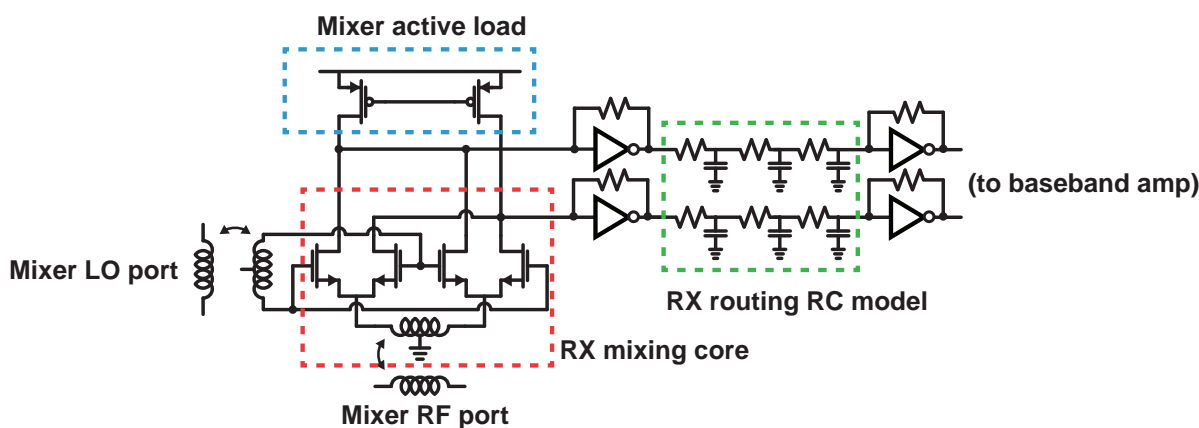


Figure 5.43: Finalized mixer schematic.

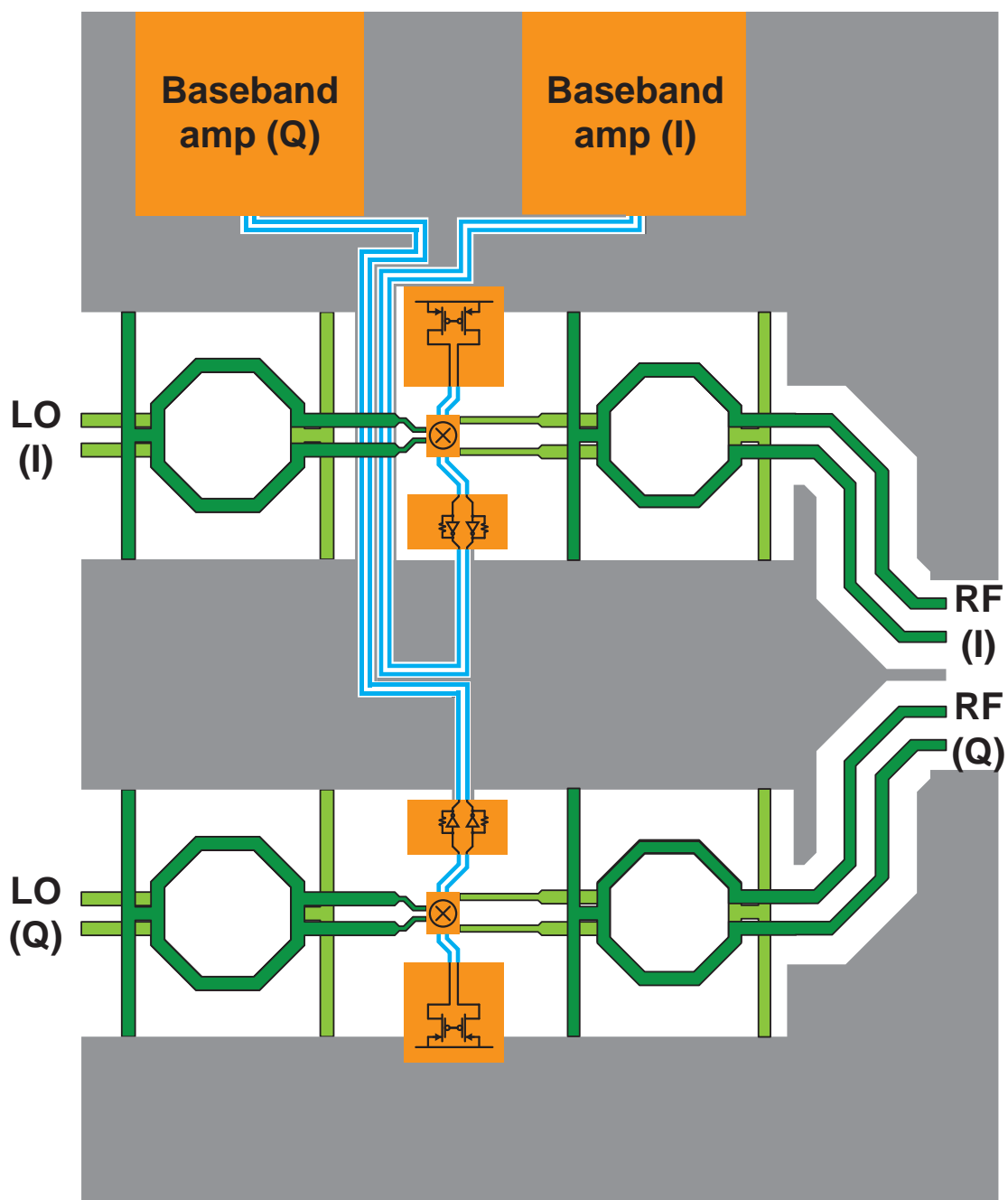


Figure 5.44: Mixer layout floorplan. Due to IO pad constraints, the I and Q baseband amps both need to be on the same side of the mixers, necessitating routing the mixer baseband outputs a long distance (shown in thin cyan lines).

5.4.3 LNA

The LNA design is based around a common-source neutralized differential amplifier, using the same layout neutralization style as in the power amplifier. The extremely broad-band input match (Figure 5.46 is provided mostly by the loss of the input transformer and the chip-to-PCB transition (the chip-to-PCB model is included in all receiver simulations).

After the first LNA stage, the signal is split into two separate amplification paths. This is to isolate the I and Q mixers from each other and prevent the possibility of any LO or noise coupling between them. The total LNA gain to a single output, including the 3dB loss from the split, is simulated to be 8dB at the center frequency of 115GHz (Figure 5.47). The LNA has a 3dB bandwidth of 31GHz, but this comes at the cost of increased group delay ripple of 30ps.

The simulated LNA noise figure is less than 9dB over the full 100-130GHz band (Figure 5.48). The noise figure peaks significantly towards the higher end of this band. This is a consequence of increasing packaging losses above 130GHz.

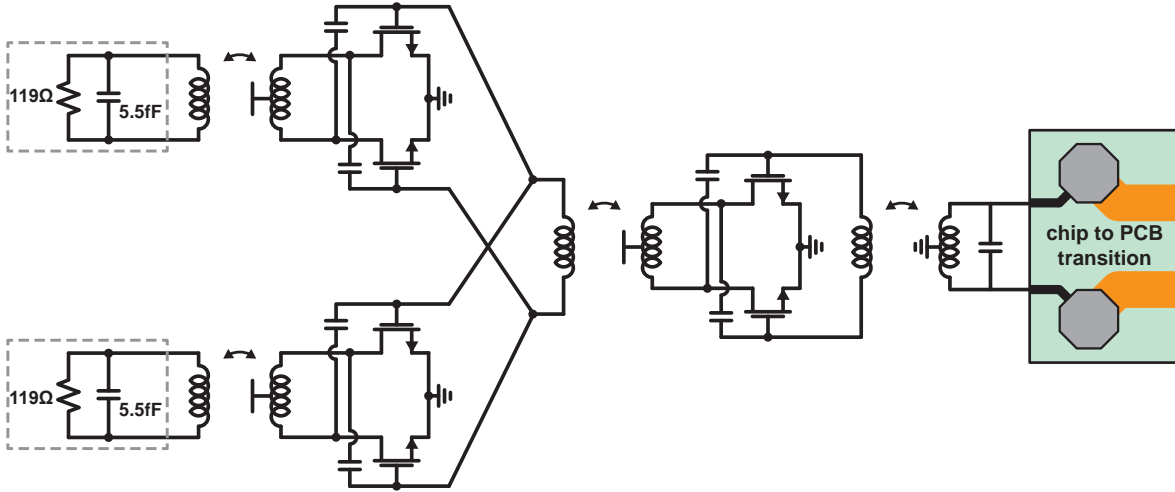


Figure 5.45: LNA schematic

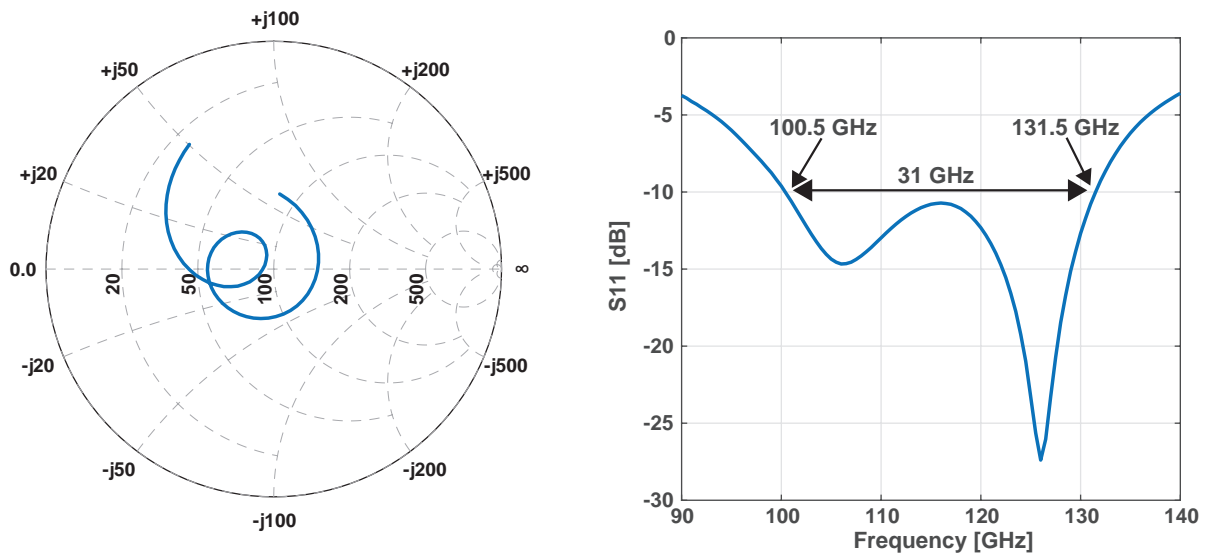


Figure 5.46: Simulated LNA input impedance seen from PCB (including chip-to-PCB transition model)

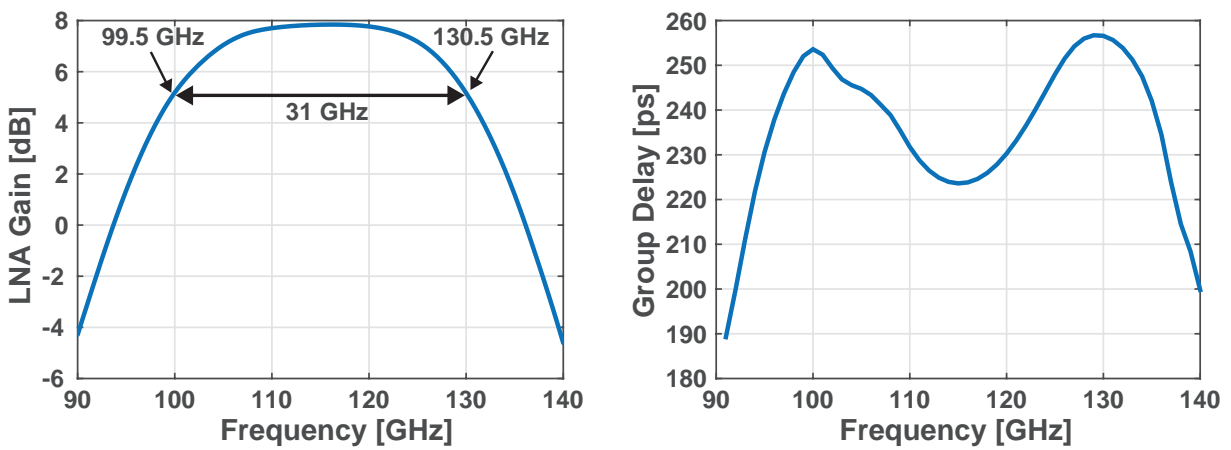


Figure 5.47: Simulated LNA Gain and Group Delay (including 3dB loss from I-Q split)

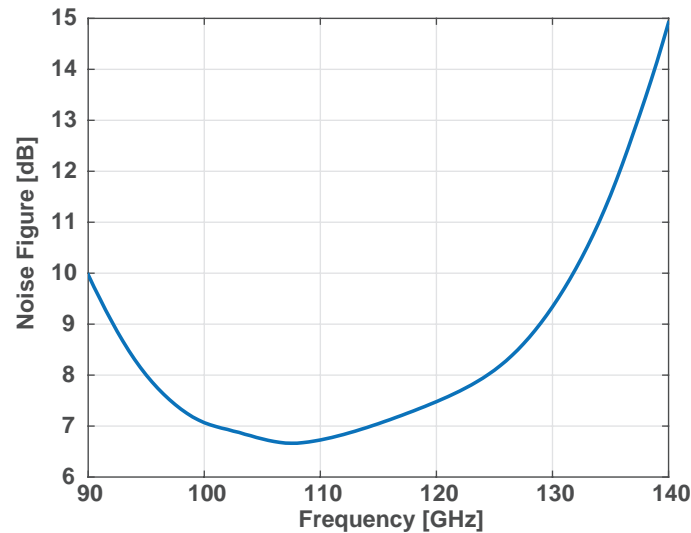


Figure 5.48: Simulated LNA Noise Figure

5.4.4 RX Chain Simulated Performance

For verification of the complete system, the full receiver chain was simulated together using an LO frequency of 115GHz. The transceiver has nearly 44dB of conversion gain (8dB LNA + 11dB mixer + 25dB baseband), with a 3dB bandwidth of 10GHz (Figure 5.49).

The (double sideband) noise figure in-band for the complete chain is simulated to be 11.2dB (Figure 5.50). Because of the AC coupling capacitance in the baseband, and the presence of flicker noise, the noise figure increases dramatically below 25MHz.

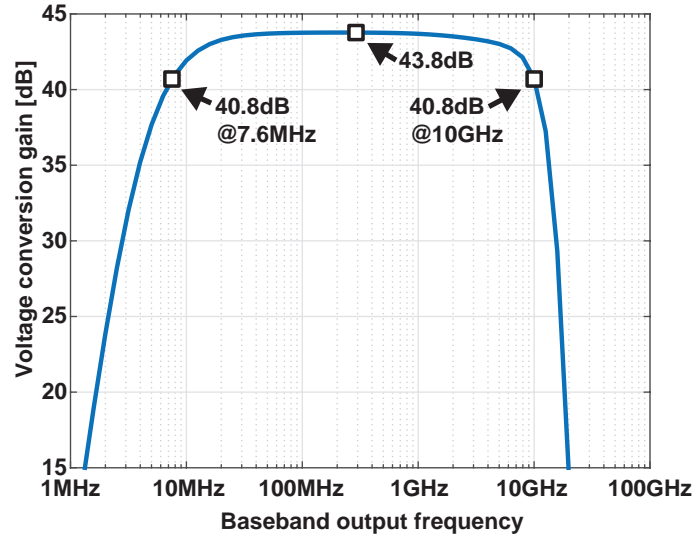


Figure 5.49: Simulated conversion gain of the full receiver chain, including LNA, mixer, and baseband amplification. The x-axis represents the frequency offset from the 115GHz LO signal and the baseband tone frequency that it downconverts to.

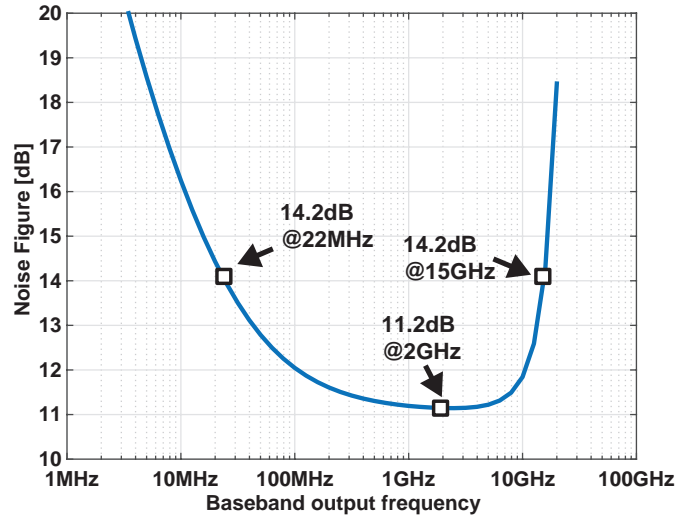


Figure 5.50: Simulated double sideband noise figure for the full RX chain with 115GHz LO frequency.

5.5 PCB and Antenna Design

The choice of PCB materials and stackup is influenced by several factors. The dielectric for the top routing layer of the PCB should be thin to minimize surface wave excitation from the chip-PCB transition. For a symmetric board stackup, the dielectric for the bottom

routing layer should therefore also be thin — asymmetric stackups are possible but prone to warpage, creating issues for the assembly process.

To keep costs low, the board has only a four layer construction, with two routing layers and two plane layers. So, the only dimension left undetermined is the thickness of the internal dielectric between the two plane layers. A thicker dielectric means a thicker PCB which is more rigid, and is simpler to assemble and more robust. But coupling into surface waves from the antenna is worsened with a thicker dielectric.

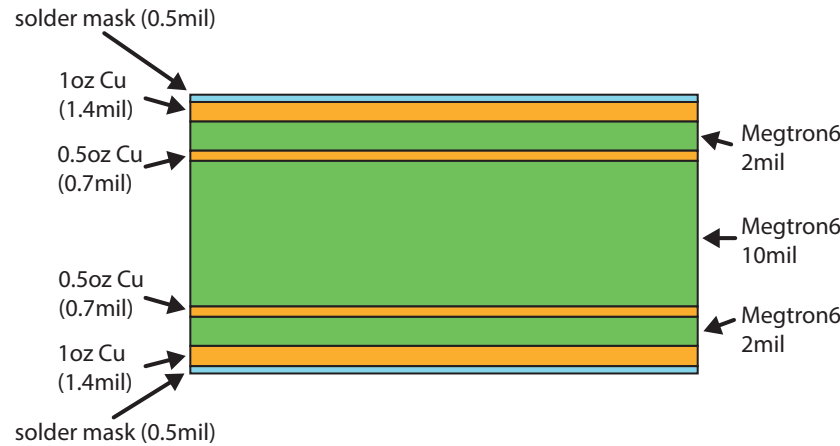


Figure 5.51: Selected PCB stackup.

Several classes of high-performance PCB materials are available, few of which have been characterized at millimeter-wave. Table 5.2 lists a variety of commercially available materials, with loss tangents ranging over two orders of magnitude. The lowest loss material, CuFlon from PolyFlon, is a pure PTFE (Teflon) material, while nearly all other PTFE based materials are reinforced with glass fibers or ceramic filled and have higher losses. The trade-off is that pure PTFE is very difficult to work with, with reinforced PTFE being somewhat simpler. A standard off-the-shelf FR-4 material, Isola 370HR, is the highest loss, but simplest to assemble into a multilayer board. The Megtron line of materials offers a compromise between complexity of assembly (similar to FR-4) and high-frequency performance, so it was selected for the final antenna design.

To interface smoothly with the differential I/O from the transceiver IC, it is desirable to have a differential antenna as well. The antenna should have broadband operation (wide S_{11} bandwidth, and flat gain vs frequency) while also having high polarization purity to enable dual-polarization data transmission with two antennas.

5.5.1 Folded Dipole Antenna

Initial design investigations showed that a folded dipole antenna on a thin substrate can provide moderate S_{11} -10dB bandwidth of 12GHz (Figure 5.52a). Although this is not

Table 5.2: Characteristics of Various PCB Materials

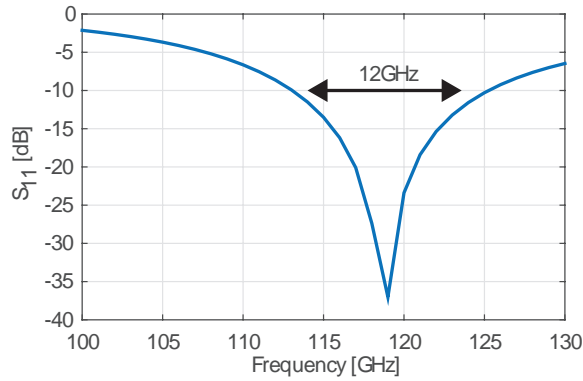
PCB Core Material	Dielectric Constant	Loss Tangent @ 10GHz	Loss Tangent @120GHz (estimated)	Minimum Thickness	Material Type
Rogers 5880	2.20	0.0009	0.0015 - 0.0022[83]	5mil	Glass microfiber reinforced PTFE
Rogers XT/duroid 8000	3.23	0.0035	—	2mil	Thermoplastic
Rogers ULTRALAM 3850	3.14	0.0025	0.007 [84]	1mil	Liquid crystalline polymer (LCP)
Nelco NY9208	2.08	0.0006	—	30mil	Reinforced PTFE laminate
Nelco NY9217	2.17	0.0008	—	5mil	Reinforced PTFE laminate
Taconic FastRise	2.74	0.0014	—	2mil	Ceramic, thermoset and PTFE
Taconic TLP-3	2.33	0.0009	—	5mil	Fiberglass fabric coated with PTFE
Taconic TLX-9	2.5	0.0015	—	2mil	PTFE fiberglass laminate
Megtron7 R-5785(N)	3.20	0.002	0.008	2mil	PPE blend resin system
Megtron6 R-5775(N)	3.22	0.004	0.011	2mil	PPE blend resin system
Polyflon CuFlon	2.05	0.00045*	—	0.25mil	Pure PTFE
Isola 370HR	3.63	0.03	—	2mil	FR-4 (Epoxy resin)

*Datasheet reports loss tangent at 18GHz

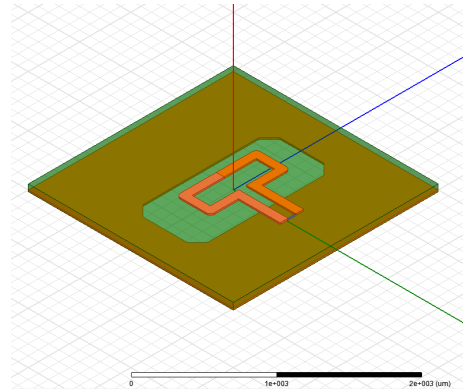
sufficient to operate at the maximum designed data rate, this antenna structure was a good enough starting point to justify further investigation.

The antenna, depicted in Figure 5.52b, is designed using a single thin sheet (2mil) of dielectric, with a differential microstrip transmission line feed referenced to a ground plane underneath the dielectric. There is a ground plane cutout around the antenna region to allow the antenna to radiate properly. Without the cutout, image currents will form in the ground plane that cancel out the currents in the antenna, which would significantly decrease the radiated power.

However, it's not possible to package the entire chip on such a thin substrate, since multiple routing layers are required for all of the other non-RF I/O from the chip. A thicker, multilayer substrate presents additional challenges. The added dielectric material underneath the antenna effectively increases the capacitance of the distributed structure, lowering its input impedance. It also impacts the front to back ratio of the antenna, as it becomes difficult to prevent radiation out of the back side of the antenna. Most problematically, the



(a) Input impedance of folded dipole antenna



(b) 3D model of folded dipole antenna

antenna couples its energy into surface waves much more easily on the thicker substrate. This can be seen by looking at the radiation pattern of the antenna at different frequencies (Figure 5.53).

The surface waves impact the radiation pattern by traveling through the substrate until they reach a discontinuity (a via, piece of metal, or the edge of the board). At the point of discontinuity, the surface wave diffracts and couples some energy into free space [85]. Because the surface waves undergo a different phase shift at different frequencies, this effect appears as a “ripple” superimposed on a plot of gain versus frequency. In Figure 5.53, this effect appears as smaller side lobes of the radiation pattern that change location with frequency.

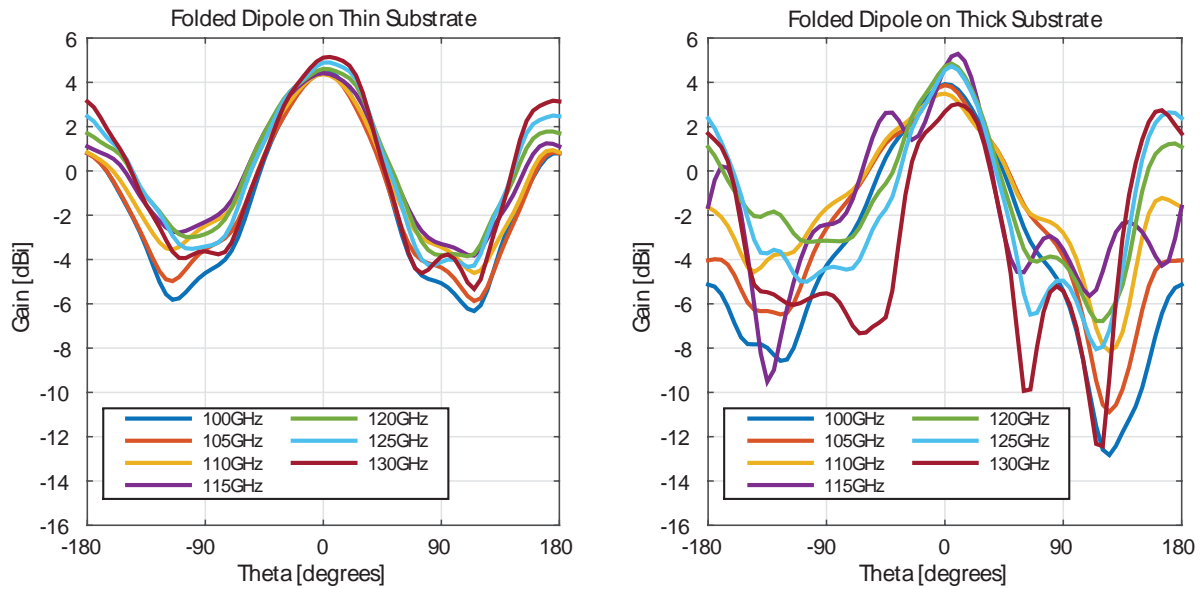


Figure 5.53: Simulated antenna radiation pattern (θ representing the polar angle from the z-axis) versus frequency for thick and thin substrates

5.5.2 Surface Waves

In general, a surface wave is defined as a wave that propagates at the interface of two dielectrics [75]. The fields in a surface wave drop off exponentially when moving away from the interface surface [86]. Surface waves can propagate with various configurations of dielectrics and conductors, but the most relevant surface wave configuration for this PCB structure is the grounded dielectric sheet (Figure 5.55). This structure supports both TM and TE modes, and interestingly (but problematically) the lowest order TM mode has a cutoff frequency of zero (Eq. 5.44) [75].

$$f_{c,TM} = \frac{nc}{2d\sqrt{\epsilon - 1}}, \quad n = 0, 1, 2, \dots \quad (5.44a)$$

$$f_{c,TE} = \frac{(2n - 1)c}{4d\sqrt{\epsilon - 1}}, \quad n = 1, 2, 3, \dots \quad (5.44b)$$

For a dielectric with a relative permittivity of 3.2, and a thickness of 12mil, the cutoff frequency for the first TE mode is 166GHz. So, the only surface wave mode of concern is the TM_0 mode.

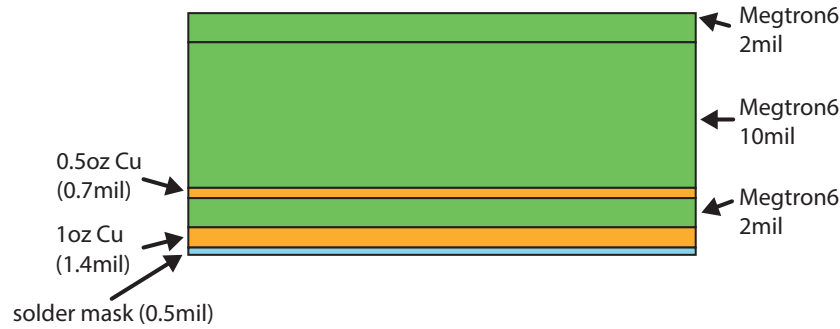


Figure 5.54: Dielectric slab which supports TM_0 surface waves.

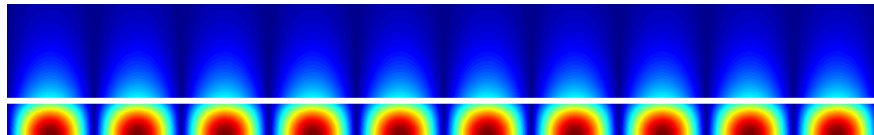


Figure 5.55: Calculated H field of grounded slab TM_0 surface wave mode according to theory. The horizontal axis is the direction of propagation, and the vertical distance from the bottom of the image. The white horizontal line represents the air-dielectric interface.

5.5.3 Methods for Dealing with Surface Waves

Because of the difficult-to-predict and problematic impact of surface waves on the radiation pattern of the antenna, several methods have been proposed to mitigate those issues.

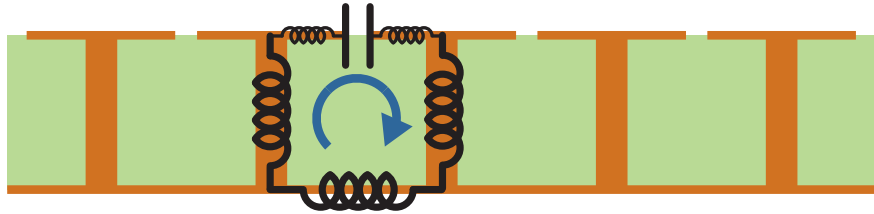


Figure 5.56: Sievenpiper “mushroom” EBG structure, side view.

Sievenpiper *et al.* proposed an artificial surface made of sub-wavelength structures that prevents surface wave propagation within a forbidden band [85]. This structure, also known as an “artificial magnetic conductor”, behaves similarly to a theoretical magnetic conductor over the band of operation. The exact structure proposed uses conductive patches on the top metal layer of a PCB, connected to the ground plane through a central via (Figure 5.56), and is referred to as a “mushroom” or “Sievenpiper” EBG (electromagnetic band gap) in the literature. Resonance between the inductance of the vias/ground plane and the capacitance between the patches results in the desired EBG behavior around the resonant frequency [87].

There are few examples in the literature of this type of EBG structure at mm-Wave; only a handful of 60GHz designs have made use of the technique [88][89]. For operation above 100GHz, the large electrical size of the available vias in this PCB technology (both diameter and height) makes it prohibitive to use this approach. Artificial magnetic conductor techniques have been demonstrated that do not use vias [90] for the purpose of replacing a back reflector ground plane in an antenna. However, it is not clear if these types of constructions can also be used to suppress surface wave propagation.

Another approach is to try to capture and re-radiate whatever surface waves are launched by the antenna structure [91]. This has the added benefit of increasing the gain of the antenna because of the larger effective aperture that is contributing to free-space radiation. A potential drawback is that, due to the nonzero propagation delay of the surface wave from the main antenna to the re-radiation structure, the two radiated signals will only be in phase at one frequency. Away from that frequency, the gain will drop off as the two waves start to interfere destructively.

Based on the design from University of Nice[91], an aperture coupled patch antenna with a grounded-via guard ring was designed using the PCB stackup in Figure 5.51. By looking at the radiation pattern as the substrate size is swept, the effectiveness of the ring in capturing surface waves is clear (Figure 5.57). In contrast, the radiation pattern of the antenna without the guard ring is very unstable as the substrate size is varied, indicating a large amount of energy coupled into surface waves.

To capture even more surface waves, a second via ring can be added, although this does not provide a significant increase in antenna gain. It does create a radiation pattern that is very stable versus PCB size, so there is little energy that propagates outside the structure as surface waves (Figure 5.57). The downside is a more narrow-band radiation pattern. Because the distance to the second ring is further, this added time delay means the phase of its radiated field, relative to that of the main patch, will change more quickly versus frequency. This also leads to greater and steeper group delay variation versus frequency (Figure 5.59).

Another potential solution to the surface wave problem is to surround the antenna with a trench. This will prevent surface waves from propagating, since there is no media for the wave to travel through any more (in practice, some nonzero amount of energy will still couple across the gap). A thorough investigation based on a design at 2.4GHz has shown that broadside gain can be enhanced by up to 2.7dB by removing the substrate around the patch [92]. The substrate is removed right up to the edge of the patch, so the fringing fields extend out of the substrate material into the air. Partial and complete removal is investigated in

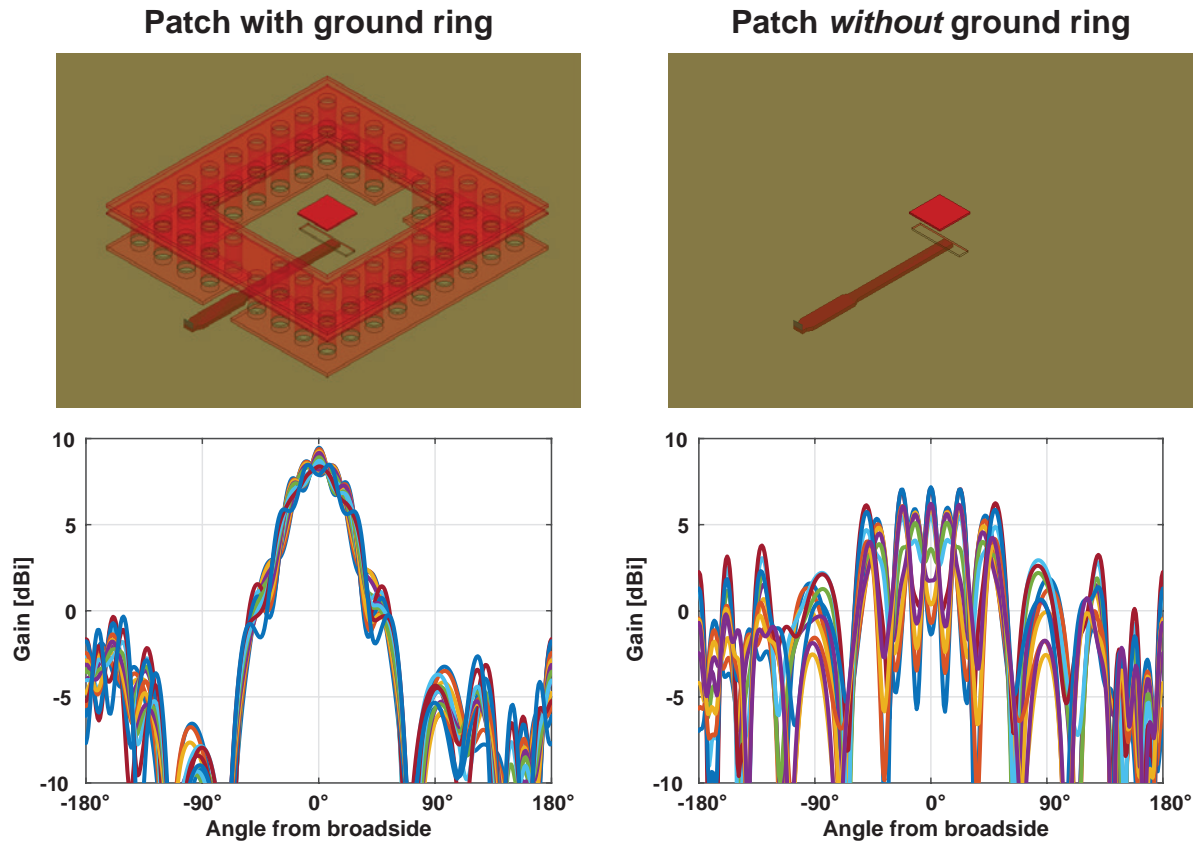


Figure 5.57: Patch antenna gain (at 115GHz) with and without guard ring, as the PCB substrate size is varied.

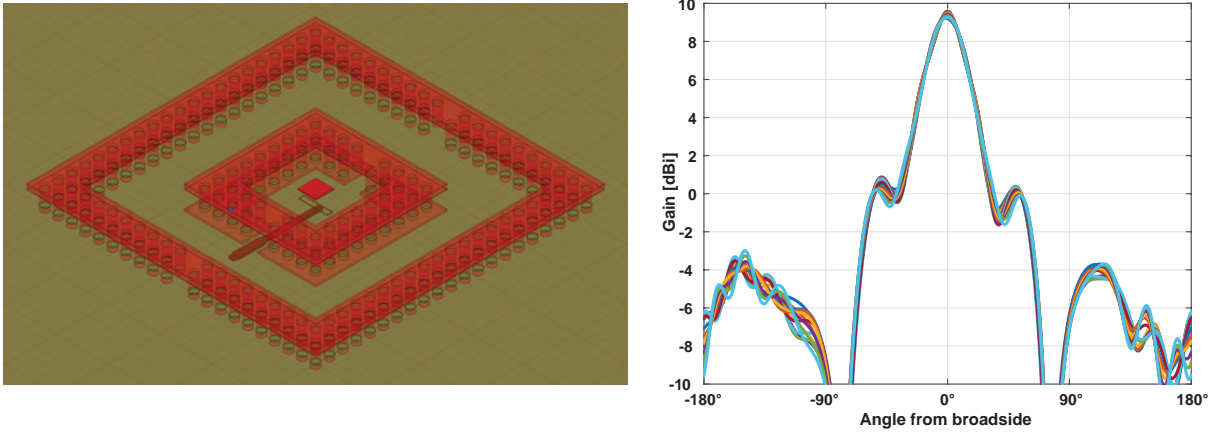


Figure 5.58: Antenna gain (at 115GHz) with two guard rings, as PCB substrate size is varied.

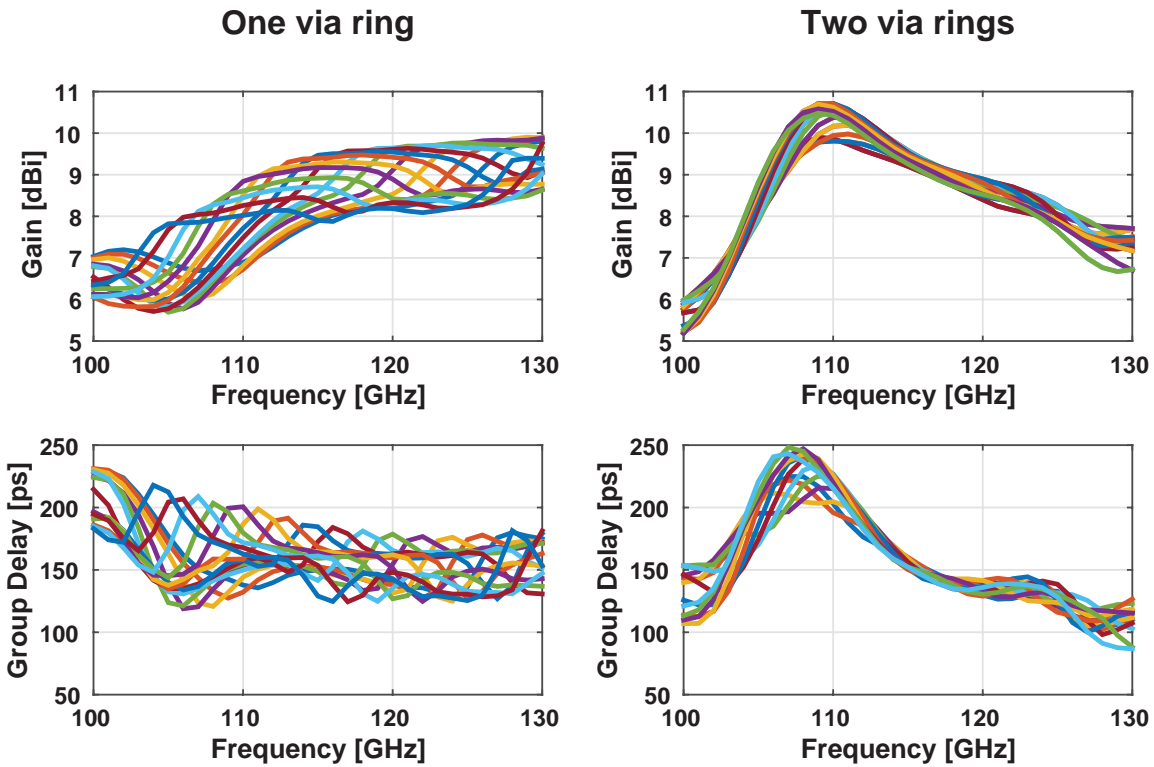


Figure 5.59: Antenna gain and group delay at broadside with one and two guard rings

the paper, including the effect of removing substrate on only the radiating or non-radiating sides of the patch, with the conclusion that substrate removal on the radiating ends of the patch is more effective in suppressing surface waves.

Similar techniques have been tried at 60GHz [93], and a 2.3-2.8dB increase in gain is reported for the removed-substrate antennas. A downside is that this effect also appears to correspond with a reduction in impedance bandwidth, hypothesized by the authors of that publication to be because of higher field concentration near the edge of the removed region that increases the Q of the antenna resonance.

Because of this bandwidth reduction effect, and concerns about the sensitivity of antenna input impedance to the tolerances of the substrate removal process, the substrate removal strategy used here was somewhat different than that has been shown in the literature. Instead, the grounded via ring described earlier is included in the design, and the substrate is only removed in the area outside of the ring (Figure 5.60), forming a “trench” around the ring. This ensures that any surface waves that are not captured by the ring are re-radiated in a controlled way, and do not travel further away from the antenna. Simulations show that there is little impact on the antenna input impedance from PCB features outside the ring, so surface waves can be suppressed effectively by the cavity without any unpredictable impacts on the input impedance.

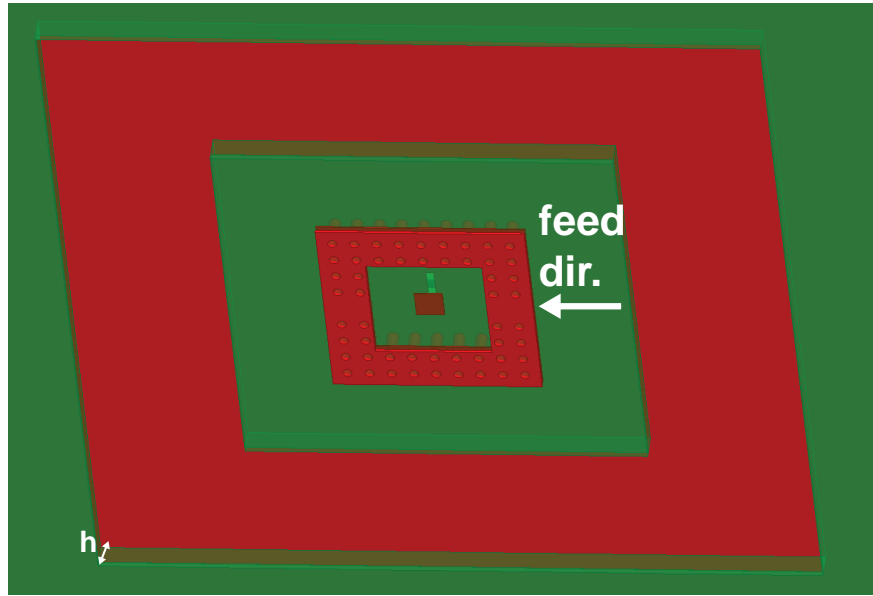


Figure 5.60: Patch with grounded via ring and partial substrate removal

Simulated results predict that a trench surrounding a patch with a single via ring (design B) is fairly effective at suppressing surface waves (Figure 5.62). The simulations capture the effect of surface waves by proxy: the substrate dimensions of the PCB material are varied and the variation in antenna gain at broadside is recorded. Notably, the “trench all around” (B) and “trench only on radiating sides” (E) have nearly identical surface wave suppression, while having the trench only on the *non-radiating* sides (D) is the same or worse than no trench at all (A). These results are consistent with what has been reported in the literature [92]. None of these solutions are as effective at suppressing surface waves as having two

via rings, but they do not have the negative side effect of significantly narrowed 3dB gain bandwidth.

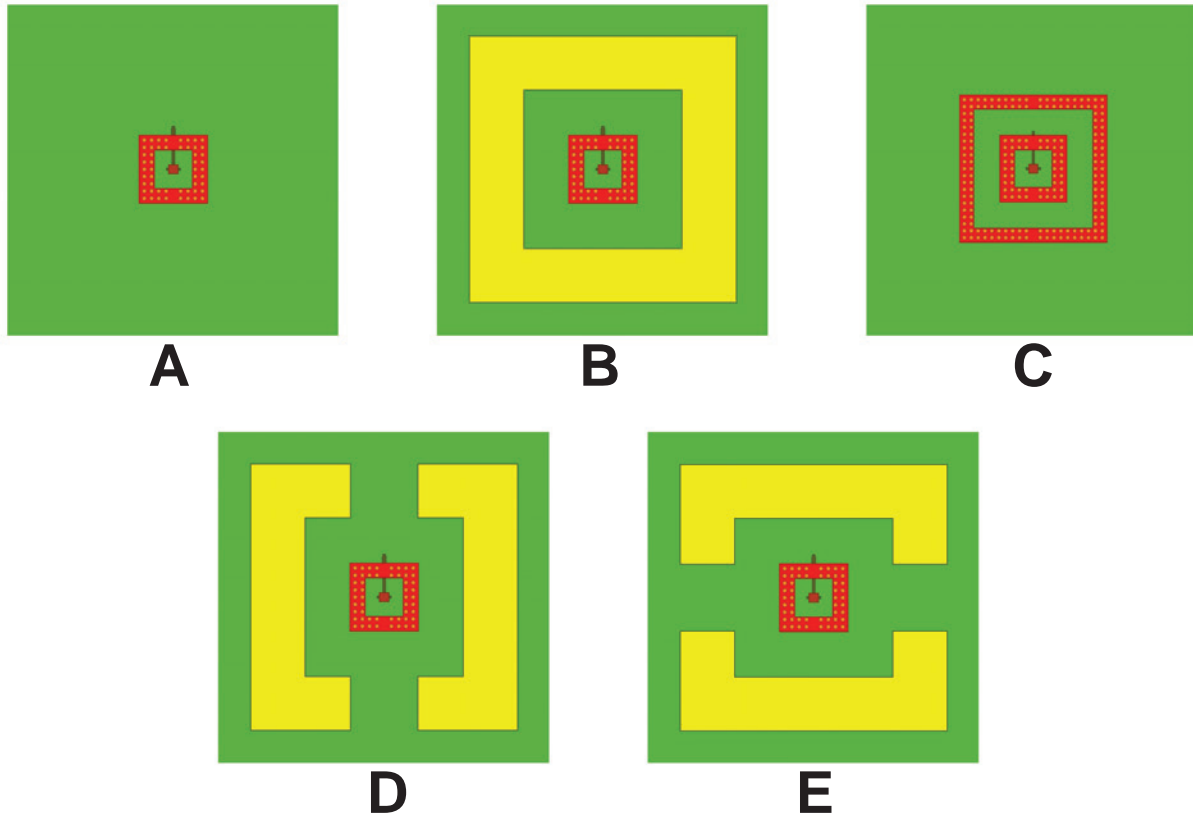


Figure 5.61: Different surface wave suppression techniques. Trench/substrate removal areas are shown in yellow.

Unfortunately, due to the fabrication constraints of the PCB vendor, it was not possible to have the trench fully surround the antenna. The solution to this issue was to keep the antenna region supported on the left and right (along the direction of the feedline), so the feedline is mechanically well-supported during the PCB multilayer lamination process. Unfortunately, this also implies that the substrate cannot be removed along the radiating edges of the patch, leading to reduced overall surface wave suppression. To form a complete trench, the last bit of substrate material would need to be removed by an external vendor after the multilayer PCB has been pressed together and vias added (the vias would be necessary to mechanically support the patch region after complete trenching). A photo of the final antenna design, with partial substrate removal, is shown in Figure 5.63.

The final simulated input impedance is plotted below in Figure 5.64, and has minimal dependence on the size of the substrate (i.e., is not impacted by surface waves).

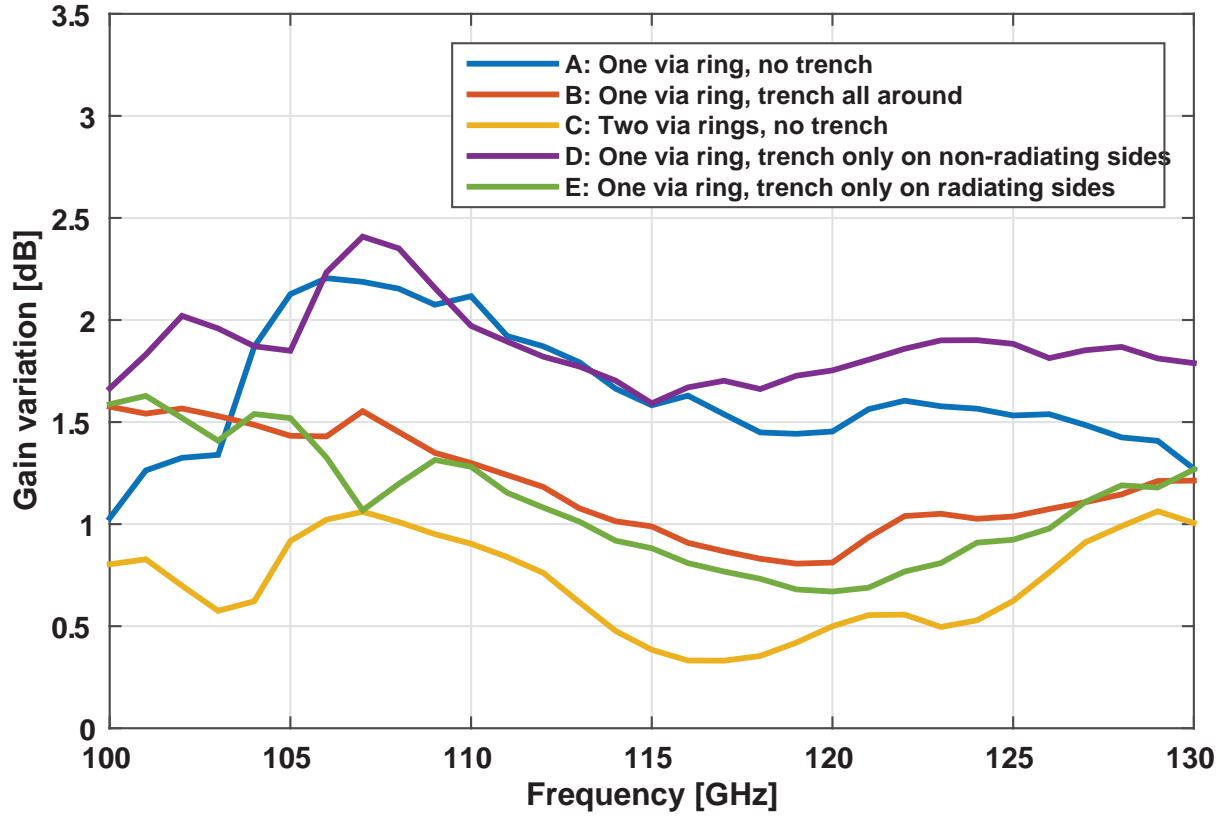


Figure 5.62: Worst-case gain variation (minimum to maximum) of various surface wave mitigation techniques, as the (square) substrate size is varied from 12 to 15mm.

5.5.4 Planar Balun on PCB

Because the patch antenna is single-ended, but the chip has a fully differential interface, a balun is required to interface between the two. A planar microwave solution to this problem is the so-called “rat-race” hybrid [75] (Figure 5.65a).

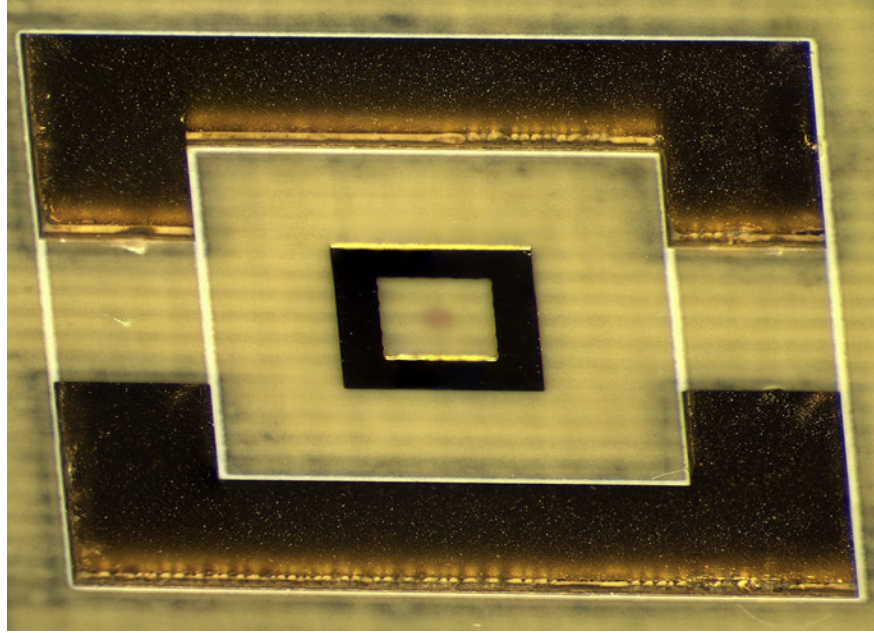


Figure 5.63: Fabricated antenna with partial substrate removal.

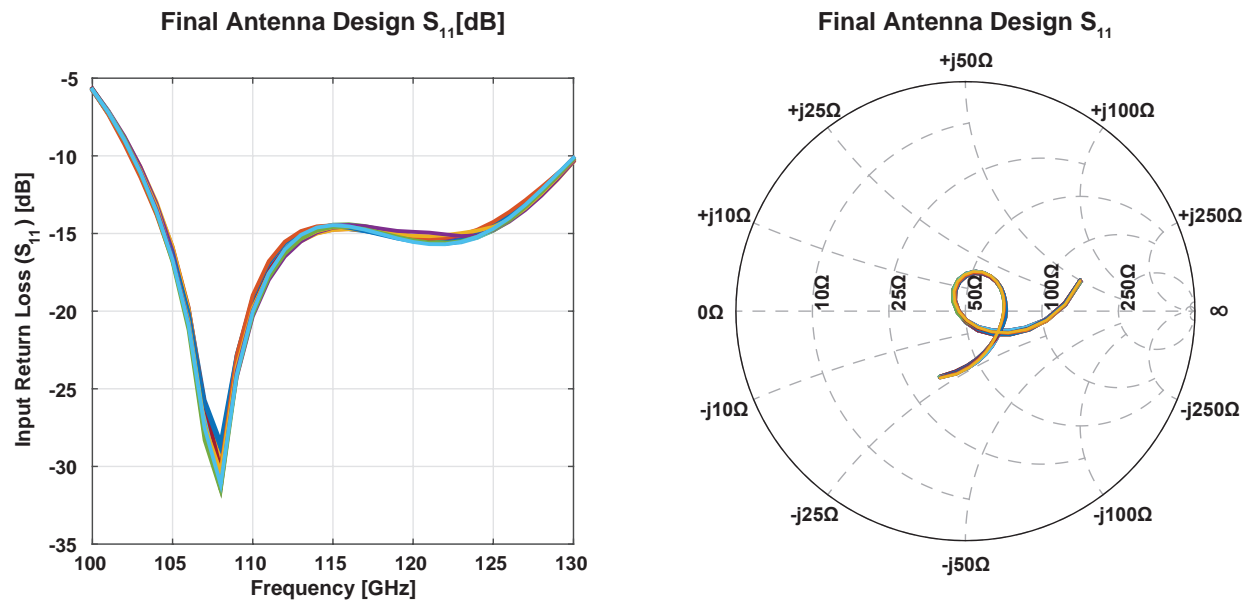


Figure 5.64: Input impedance of final antenna design, shown for various substrate dimensions

For a 50 Ω hybrid, the impedance of the transmission line of the ring should be $50\sqrt{2}\Omega = 70.7\Omega$. The line width required to implement a microstrip line of this impedance would be smaller than the minimum trace width of 3mil for the moderate cost PCB fabrication run.

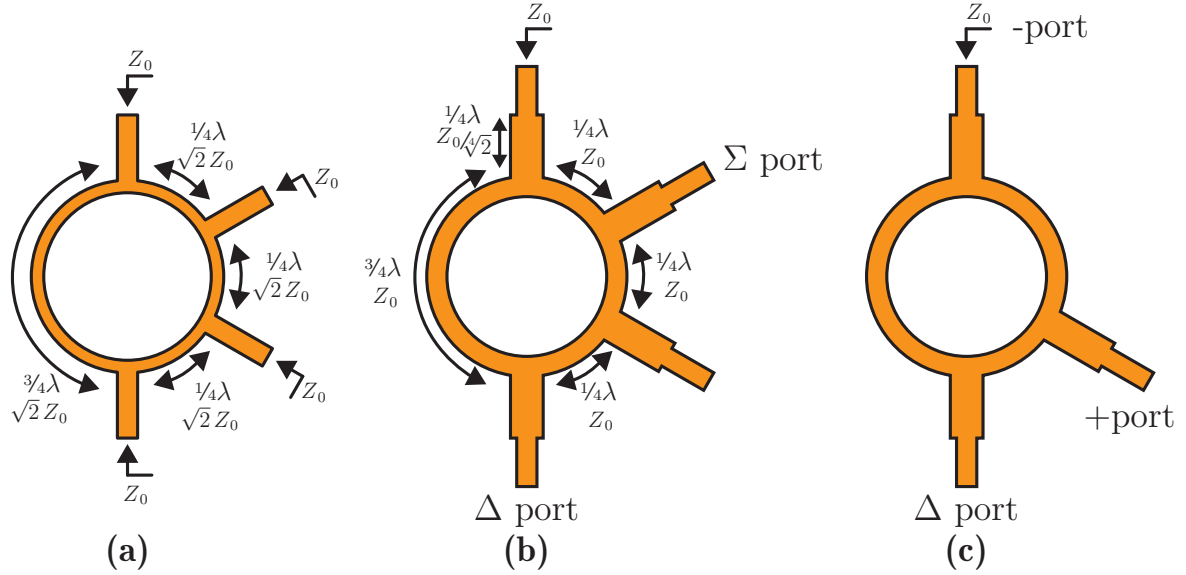
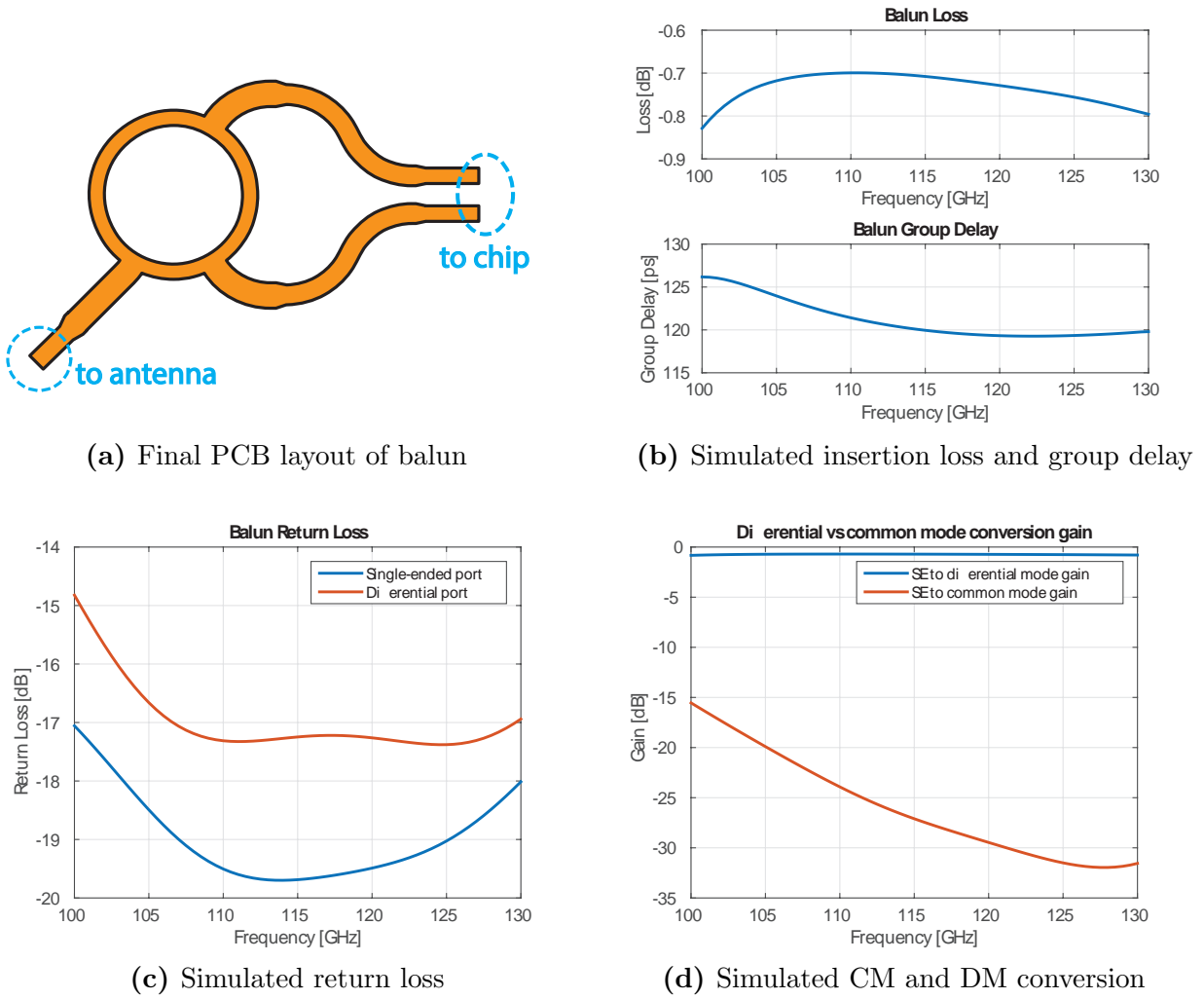
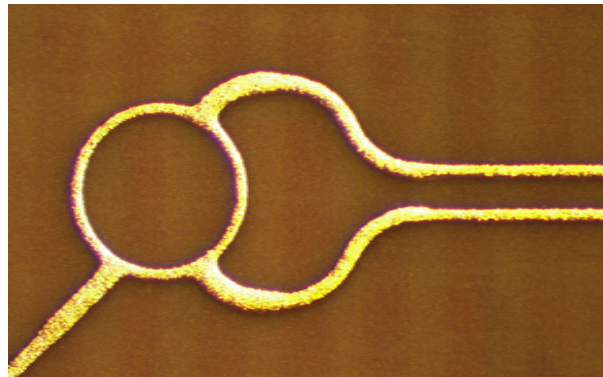


Figure 5.65: Design process for optimized rat-race hybrid without sum port termination.

Smaller line widths can be fabricated, but at significantly higher cost. To avoid this problem, a 35Ω hybrid is designed using a 35Ω impedance for the ring, and a series quarter wavelength line is used to match back to 50Ω (Figure 5.65b).

This requires a matched termination on all ports to provide ideal isolation and common-mode rejection. However, it is extremely difficult to create an ideal 50Ω termination at 120GHz on a PCB (even an ideal resistor would require a via to the ground plane underneath, contributing significant inductance). It has been shown that the termination on the “sum” port can be omitted at the cost of reducing the isolation of the balun [94] (Figure 5.65a).

The final balun design is shown in Figure 5.66a. The two differential ports are tapered with a smooth bend to interface with a differential microstrip transmission line. Simulations predict a loss of between 0.83 to 0.7 dB and 7ps variation in group delay over the 100 to 130GHz band (Figure 5.66b). Simulated return losses for the differential and single ended ports are below -15dB from 101 to 130GHz (Figure 5.66c). Common-mode conversion is worst at the low end of the frequency range, but still better than -15dB (Figure 5.66d). Since the common-mode impedance of the chip is not specifically designed to be matched, the common-mode performance will likely vary from this result in practice.

**Figure 5.66:** Final balun design and simulations**Figure 5.67:** Fabricated rat-race hybrid on PCB.

5.6 Fabricated Transceiver and Antenna PCB

The complete transceiver was fabricated in 28nm bulk CMOS — a photograph of the die is included below (Figure 5.68). The final die dimensions measured 2mm by 2.36mm. The chip was flip-chip packaged using solder balls onto an antenna PCB (Figure 5.69). The complete assembled board is shown in Figure 5.70 with the TX and RX antennas annotated. Chip and link measurements are in progress.

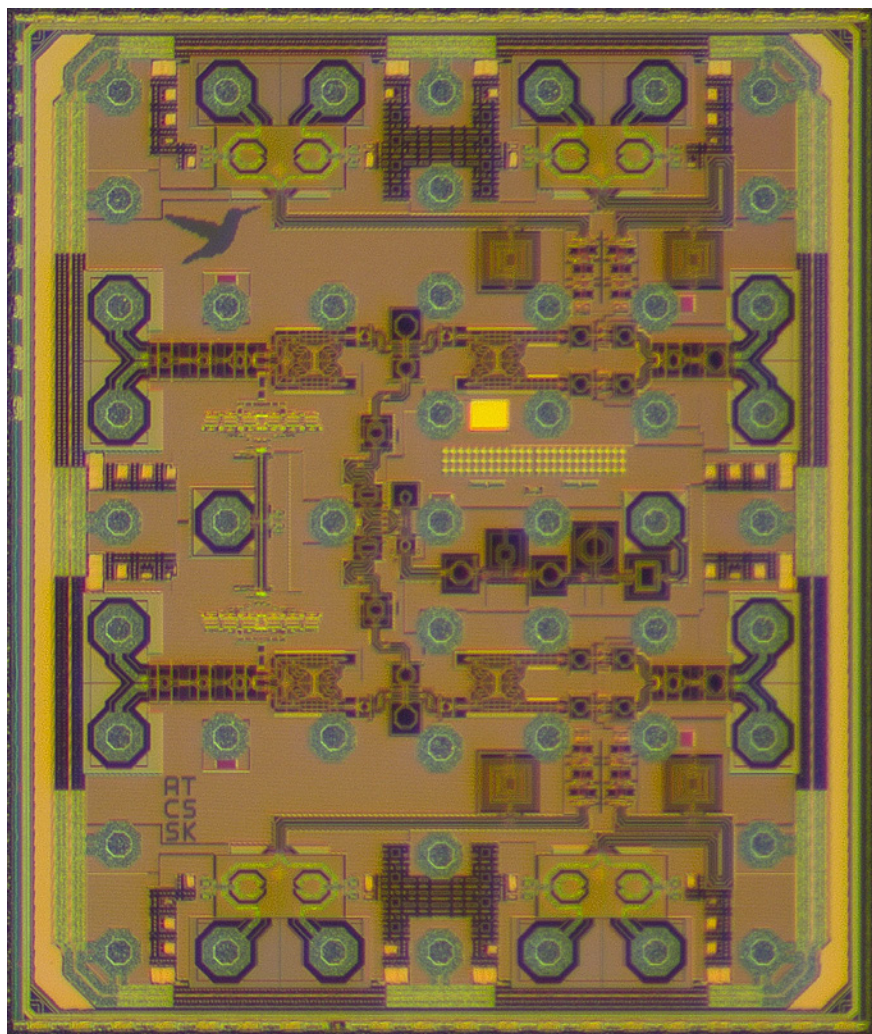


Figure 5.68: Die photo.

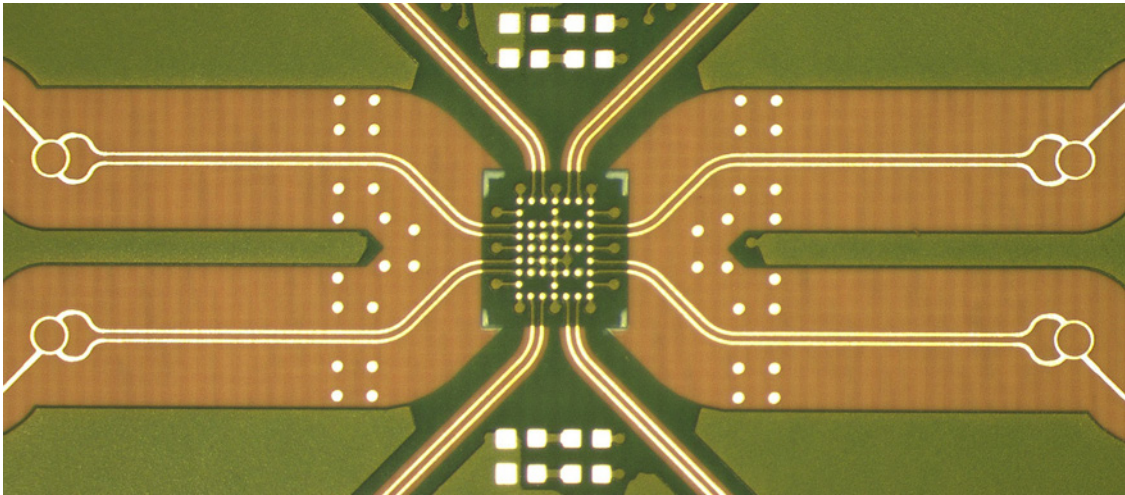


Figure 5.69: Flip-chip footprint, differential routing, and mm-Wave baluns as fabricated on PCB.



Figure 5.70: Antenna PCB with chip attached.

5.7 Performance Comparison

Table 5.3: Performance and Energy Efficiency of mm-Wave Wideband Transmitters

Reference	Carrier Frequency [GHz]	Data Rate [Gb/s]	Modulation Technique	DC power [mW]	Energy efficiency [pJ/bit]
This work	115	2x50***	QPSK, 16QAM, 64QAM	2x110 + 130 (LO)***	2.7*** (including LO)
Dolatsha, ISSCC 2017 [54]	130	12.5†	OOK	59	4.72
Yu, T-MTT 2014 [52]	60	16	OOK	19.2	1.2
Yang, RFIC 2014 [58]	155	20†	QPSK	345‡	17.25‡
Jiang, JSSC 2017 [60]	220	2x12.2, 2x40***	ASK	2x450	36.9
Tokgoz, ISSCC 2018 [62]	70,105	2x60*, 2x36**	64QAM, 16QAM, QPSK, 8-PSK	120	1*, 1.7**

†Complete link demonstrated

‡Power for TX+RX

*With equalization

**Without equalization

***Simulated

Table 5.4: Performance and Energy Efficiency of mm-Wave Wideband Receivers

Reference	Carrier Frequency [GHz]	Data Rate [Gb/s]	Modulation Technique	DC power [mW]	Energy efficiency [pJ/bit]
This work	115	2x50***	QPSK, 16QAM, 64QAM	2x83 + 130 (LO)***	2.97*** (including LO)
Dolatsha, ISSCC 2017 [54]	130	12.5†	OOK	38	3.04
Yu, TCAS-I 2015 [53]	60	18.7	OOK	11.6	0.62
Yang, RFIC 2014 [58]	155	20†	QPSK	345‡	17.25‡
Tokgoz, ISSCC 2018 [62]	70,105	2x60*, 2x36**	64QAM, 16QAM, QPSK, 8-PSK	160	1.3*, 2.22**

†Complete link demonstrated

‡Power for TX+RX

*With equalization

**Without equalization

***Simulated

Chapter 6

Conclusion

6.1 Summary of Thesis

In this thesis, the designs for two highly-integrated broadband mm-Wave transceivers have been presented. Both transceivers have been integrated with antennas to facilitate system-level demonstrations.

First, circuit design techniques for a 94GHz phased-array FMCW radar system are explored. Energy efficiency is optimized by system-level optimizations, and relaxation of individual element specifications due to phased array antenna gains. By careful design of shared LO generation and distribution circuitry, the overhead power of these components can be kept under control to achieve an efficient system implementation.

The assembled chip and antenna module were then used to demonstrate beamforming on the receiver and transmitter. Radar measurements show that basic range measurement functionality is possible with high distance resolution, corresponding to the high available circuit bandwidth.

After the presentation of the radar IC, circuit techniques focusing on extending the fractional bandwidth in millimeter-Wave amplifiers were analyzed and presented. A key contribution here is the proposal of an energy-efficient RF DAC supporting direct digital modulation of the millimeter-Wave carrier. Mechanisms for DAC nonlinearity have been analyzed and used to optimize the DAC design. A complete transmitter chain was then designed around the DAC to support 50Gb/s data transmission using QPSK, 16QAM, and 64QAM modulation schemes.

A broadband millimeter-wave downconversion receiver was also analyzed and designed using the same circuit principles. A thorough analysis of the gain-bandwidth tradeoffs in a receiver mixer reveals that a transimpedance amplifier load can be used to extend the bandwidth of the mixer without sacrificing gain. The full receiver chain was designed to support 10GHz of single-sideband RF bandwidth to support the targeted link data rate.

Finally, the antenna and packaging challenges associated with the transceiver were investigated, resulting in a new antenna design with a milled trench for surface wave suppression.

A wideband planar balun was designed to interface between the differential chip IO and the single-ended antenna port. The transceiver IC was packaged onto the antenna printed circuit board and is actively being characterized.

In summary, a nearly complete high-speed dual-polarization wireless link was proposed and implemented, including transmitter, receiver, and antennas. Simulations predict data rate and energy efficiency competitive with state-of-the-art frequency interleaved transceivers, with higher output power to support a link budget compatible with printed planar antennas.

6.2 Future Work

There are several steps that need to be taken in order to make the wideband transceiver design of this thesis into a functional standalone system. Although there is significant research activity in the area of wideband mm-Wave transceivers, there has been very little activity to date towards addressing some of these practical system issues.

6.2.1 Carrier Recovery

As it is, the receiver is fed a ~ 30 GHz LO signal from off-chip, which is frequency multiplied up to mm-Wave. In the lab, this LO signal can be shared between the transmitter and receiver for initial testing purposes. However, for a true wireless link, the receiver should be able to reconstruct the carrier of the transmitted signal and lock its own LO to the correct center frequency. It should also be able to track the close-in phase noise of the received signal, to minimize the SNR degradation from the local LO.

One possibility is a Costas-loop style carrier recovery [95]. Although strictly intended for QPSK or BPSK, it could potentially also be used for QAM modulation by discarding the interior points of the square QAM constellation (subsampling) or by using a limiting amplifier on the baseband data to turn the QAM signal into a QPSK-like data stream. It may be challenging to achieve high tracking bandwidth without compromising stability, so some innovation may be required here.

Another possibility is to use a packet-based transmission scheme with a known pilot sequence for phase synchronization. This requires more intensive signal processing and could become challenging for extremely wideband systems if the pilot sequence is used for a full channel estimation. On the other hand, if the pilot sequence is simply used as an aid for close-in carrier recovery, it can be transmitted at a significantly reduced data rate.

A key challenge in the design of a carrier recovery loop, regardless of the implementation technique, will be designing a tunable oscillator with sufficiently low phase noise. Outside the bandwidth of the carrier recovery loop, the phase noise of the RX LO will add to, rather than cancel, the phase noise of the received signal.

6.2.2 Equalization

Another obstacle towards implementing a standalone system is the equalization of finite channel bandwidth and finite polarization isolation. In its most basic form, a polarization equalizer would simply act as a “de-rotation”, by removing any polarization crossover between channels due to imperfect alignment. This is straightforward to implement given that the antenna alignment introduces no frequency dependence.

In reality, the antenna itself will have some frequency-dependent cross polarization, which will reduce the SNR of the desired polarization signal. A mixed-signal discrete time equalizer is the most straightforward way to deal with this and has already been demonstrated at 60GHz at lower data rates [96].

Bibliography

- [1] C. E. Shannon, "Communication in the presence of noise," vol. 37, no. 1, pp. 10–21, 1949.
- [2] S. Emami, R. F. Wiser, E. Ali, M. G. Forbes, M. Q. Gordon, X. Guan, S. Lo, P. T. McElwee, J. Parker, J. R. Tani, J. M. Gilbert, and C. H. Doan, "A 60GHz CMOS phased-array transceiver pair for multi-Gb/s wireless communications," in *2011 IEEE International Solid-State Circuits Conference*, 2011, pp. 164–166.
- [3] M. Boers, B. Afshar, I. Vassiliou, S. Sarkar, S. T. Nicolson, E. Adabi, B. G. Perumana, T. Chalvatzis, S. Kavvadias, P. Sen, W. L. Chan, A. H. T. Yu, A. Parsa, M. Nariman, S. Yoon, A. G. Besoli, C. A. Kyriazidou, G. Zochios, J. A. Castaneda, T. Sowlati, M. Rofougaran, and A. Rofougaran, "A 16TX/16RX 60 GHz 802.11ad Chipset With Single Coaxial Interface and Polarization Diversity," *IEEE Journal of Solid-State Circuits*, vol. 49, no. 12, pp. 3031–3045, 2014.
- [4] S. Trotta, M. Wintermantel, J. Dixon, U. Moeller, R. Jammers, T. Hauck, A. Samulak, B. Dehlink, K. Shun-Meen, H. Li, A. Ghazinour, Y. Yin, S. Pacheco, R. Reuter, S. Majied, D. Moline, T. Aaron, V. P. Trivedi, D. J. Morgan, and J. John, "An RCP Packaged Transceiver Chipset for Automotive LRR and SRR Systems in SiGe BiCMOS Technology," *IEEE Transactions on Microwave Theory and Techniques*, vol. 60, no. 3, pp. 778–794, 2012.
- [5] C. Wagner, J. Böck, M. Wojnowski, H. Jäger, J. Platz, M. Treml, F. Dober, R. Lachner, J. Minichshofer, and L. Maurer, "A 77GHz automotive radar receiver in a wafer level package," in *2012 IEEE Radio Frequency Integrated Circuits Symposium*, 2012, pp. 511–514.
- [6] B. P. Ginsburg, K. Subburaj, S. Samala, K. Ramasubramanian, J. Singh, S. Bhatara, S. Murali, D. Breen, M. Moallem, K. Dandu, S. Jalan, N. Nayak, R. Sachdev, I. Prathapan, K. Bhatia, T. Davis, E. Seok, H. Parthasarathy, R. Chatterjee, V. Srinivasan, V. Giannini, A. Kumar, R. Kulak, S. Ram, P. Gupta, Z. Parkar, S. Bhardwaj, Y. C. Rakesh, K. A. Rajagopal, A. Shrimali, and V. Rentala, "A multimode 76-to-81GHz automotive radar transceiver with autonomous monitoring," in *2018 IEEE International Solid - State Circuits Conference - (ISSCC)*, 2018, pp. 158–160.

- [7] G. Avenier, M. Diop, P. Chevalier, G. Troillard, N. Loubet, J. Bouvier, L. Depoyan, N. Derrier, M. Buczko, C. Leyris, S. Boret, S. Montusclat, A. Margain, S. Pruvost, S. T. Nicolson, K. H. K. Yau, N. Revil, D. Gloria, D. Dutartre, S. P. Voinigescu, and A. Chantre, “0.13 μm SiGe BiCMOS Technology Fully Dedicated to mm-Wave Applications,” *IEEE Journal of Solid-State Circuits*, vol. 44, no. 9, pp. 2312–2321, 2009.
- [8] P. Chevalier, G. Avenier, G. Ribes, A. Montagné, E. Canderle, D. Céli, N. Derrier, C. Deglise, C. Durand, T. Quémerais, M. Buczko, D. Gloria, O. Robin, S. Petitdidier, Y. Campidelli, F. Abbate, M. Gros-Jean, L. Berthier, J. D. Chapon, F. Leverd, C. Jenny, C. Richard, O. Gourhant, C. De-Buttet, R. Beneyton, P. Maury, S. Joblot, L. Favennec, M. Guillermet, P. Brun, K. Courouble, K. Haxaire, G. Imbert, E. Gourvest, J. Cossalter, O. Saxod, C. Tavernier, F. Foussadier, B. Ramadout, R. Bianchini, C. Julien, D. Ney, J. Rosa, S. Haendler, Y. Carminati, and B. Borot, “A 55 nm triple gate oxide 9 metal layers SiGe BiCMOS technology featuring 320 GHz f_T / 370 GHz f_{MAX} HBT and high-Q millimeter-wave passives,” in *2014 IEEE International Electron Devices Meeting*, 2014, pp. 3.9.1–3.9.3.
- [9] Y. Zhao and J. R. Long, “A Wideband, Dual-Path, Millimeter-Wave Power Amplifier With 20 dBm Output Power and PAE Above 15% in 130 nm SiGe-BiCMOS,” *IEEE Journal of Solid-State Circuits*, vol. 47, no. 9, pp. 1981–1997, 2012.
- [10] K. Khalaf, V. Vidojkovic, K. Vaesen, M. Libois, G. Mangraviti, V. Szortyka, C. Li, B. Verbruggen, M. Ingels, A. Bourdoux, C. Soens, W. V. Thillo, J. R. Long, and P. Wambacq, “Digitally Modulated CMOS Polar Transmitters for Highly-Efficient mm-Wave Wireless Communication,” *IEEE Journal of Solid-State Circuits*, vol. 51, no. 7, pp. 1579–1592, 2016.
- [11] J. Chen, L. Ye, D. Titz, F. Ganesello, R. Pilard, A. Cathelin, F. Ferrero, C. Luxey, and A. M. Niknejad, “A digitally modulated mm-Wave cartesian beamforming transmitter with quadrature spatial combining,” in *2013 International Solid-State Circuits Conf. Dig. Tech. Papers (ISSCC)*, (San Francisco, CA, USA), Feb. 2013, pp. 232–233.
- [12] S. Callender, S. Pellerano, and C. Hull, “A 73GHz PA for 5G phased arrays in 14nm FinFET CMOS,” in *2017 IEEE Radio Frequency Integrated Circuits Symposium (RFIC)*, 2017, pp. 402–405.
- [13] L. Iotti, A. Mazzanti, and F. Svelto, “Insights Into Phase-Noise Scaling in Switch-Coupled Multi-Core LC VCOs for E-Band Adaptive Modulation Links,” *IEEE Journal of Solid-State Circuits*, vol. 52, no. 7, pp. 1703–1718, 2017.
- [14] S. Kang and A. M. Niknejad, “A 100ghz active-varactor vco and a bi-directionally injection-locked loop in 65nm cmos,” in *2013 IEEE Radio Frequency Integrated Circuits Symposium (RFIC)*, 2013, pp. 231–234.
- [15] I. S. Merrill *et al.*, “Introduction to Radar Systems,” *Mc Graw-Hill*, 2001.

- [16] S. M. Han, T. Takasawa, K. Yasutomi, S. Aoyama, K. Kagawa, and S. Kawahito, "A Time-of-Flight Range Image Sensor With Background Canceling Lock-in Pixels Based on Lateral Electric Field Charge Modulation," *IEEE Journal of the Electron Devices Society*, vol. 3, no. 3, pp. 267–275, 2015.
- [17] A. Arbabian, S. Callender, S. Kang, B. Afshar, J. C. Chien, and A. M. Niknejad, "A 90 GHz Hybrid Switching Pulsed-Transmitter for Medical Imaging," *IEEE Journal of Solid-State Circuits*, vol. 45, no. 12, pp. 2667–2681, 2010.
- [18] V. Giannini, D. Guermandi, Q. Shi, A. Medra, W. V. Thillo, A. Bourdoux, and P. Wambacq, "A 79 GHz Phase-Modulated 4 GHz-BW CW Radar Transmitter in 28 nm CMOS," *IEEE Journal of Solid-State Circuits*, vol. 49, no. 12, pp. 2925–2937, 2014.
- [19] J. Lee, Y. A. Li, M. H. Hung, and S. J. Huang, "A Fully-Integrated 77-GHz FMCW Radar Transceiver in 65-nm CMOS Technology," *IEEE Journal of Solid-State Circuits*, vol. 45, no. 12, pp. 2746–2756, 2010.
- [20] T. H. Lee, *The Design of CMOS Radio-Frequency Integrated Circuits*. Cambridge University Press, 2004.
- [21] N. V. Thienen, W. Volkaerts, and P. Reynaert, "A multi-gigabit cpfsk polymer microwave fiber communication link in 40 nm cmos," *IEEE Journal of Solid-State Circuits*, vol. 51, no. 8, pp. 1952–1958, 2016.
- [22] S. Haykin, *Communication Systems*. John Wiley & Sons, 2001.
- [23] A. Puglielli, A. Townley, G. LaCaille, V. Milovanović, P. Lu, K. Trotskovsky, A. Whitcombe, N. Narevsky, G. Wright, T. Courtade, E. Alon, B. Nikolić, and A. M. Niknejad, "Design of Energy- and Cost-Efficient Massive MIMO Arrays," *Proceedings of the IEEE*, vol. 104, no. 3, pp. 586–606, 2016.
- [24] A. Valdes-Garcia, S. T. Nicolson, J. W. Lai, A. Natarajan, P. Y. Chen, S. K. Reynolds, J. H. C. Zhan, D. G. Kam, D. Liu, and B. Floyd, "A Fully Integrated 16-Element Phased-Array Transmitter in SiGe BiCMOS for 60-GHz Communications," *IEEE Journal of Solid-State Circuits*, vol. 45, no. 12, pp. 2757–2773, 2010.
- [25] S. P. Voinigescu, M. C. Maliepaard, J. L. Showell, G. E. Babcock, D. Marchesan, M. Schroter, P. Schvan, and D. L. Harame, "A Scalable High-Frequency Noise Model for Bipolar Transistors with Application to Optimal Transistor Sizing for Low-Noise Amplifier Design," *IEEE Journal of Solid-State Circuits*, vol. 32, no. 9, pp. 1430–1439, 1997.
- [26] J. Hasch, E. Topak, R. Schnabel, T. Zwick, R. Weigel, and C. Waldschmidt, "Millimeter-Wave Technology for Automotive Radar Sensors in the 77 GHz Frequency Band," *IEEE Transactions on Microwave Theory and Techniques*, vol. 60, no. 3, pp. 845–860, 2012.
- [27] P. Molchanov, S. Gupta, K. Kim, and K. Pulli, "Short-range FMCW monopulse radar for hand-gesture sensing," in *2015 IEEE Radar Conference (RadarCon)*, 2015, pp. 1491–1496.

- [28] F. K. Wang, M. C. Tang, Y. C. Chiu, and T. S. Horng, "Gesture Sensing Using Retransmitted Wireless Communication Signals Based on Doppler Radar Technology," *IEEE Transactions on Microwave Theory and Techniques*, vol. 63, no. 12, pp. 4592–4602, 2015.
- [29] T. Kiuru, M. Metso, S. Jardak, P. Pursula, J. Häkli, M. Hirvonen, and R. Sepponen, "Movement and respiration detection using statistical properties of the FMCW radar signal," in *Millimeter Waves (GSMM) & ESA Workshop on Millimetre-Wave Technology and Applications, 2016 Global Symposium on*, IEEE, 2016, pp. 1–4.
- [30] I. Nasr, R. Jungmaier, A. Baheti, D. Noppeney, J. S. Bal, M. Wojnowski, E. Karagozler, H. Raja, J. Lien, I. Poupyrev, *et al.*, "A Highly Integrated 60 GHz 6-Channel Transceiver With Antenna in Package for Smart Sensing and Short-Range Communications," *IEEE Journal of Solid-State Circuits*, vol. 51, no. 9, pp. 2066–2076, 2016.
- [31] J. Lien, N. Gillian, M. E. Karagozler, P. Amihoud, C. Schwesig, E. Olson, H. Raja, and I. Poupyrev, "Soli: Ubiquitous Gesture Sensing with Millimeter Wave Radar," *ACM Trans. Graph.*, vol. 35, no. 4, 142:1–142:19, Jul. 2016.
- [32] Y. Wang, K. Tang, Y. Zhang, L. Lou, B. Chen, S. Liu, L. Qiu, and Y. Zheng, "A Ku-band 260mW FMCW synthetic aperture radar TRX with 1.48GHz BW in 65nm CMOS for micro-UAVs," in *2016 IEEE International Solid-State Circuits Conference (ISSCC)*, 2016, pp. 240–241.
- [33] W. Wu, R. B. Staszewski, and J. R. Long, "A 56.4-to-63.4 GHz Multi-Rate All-Digital Fractional-N PLL for FMCW Radar Applications in 65 nm CMOS," *IEEE Journal of Solid-State Circuits*, vol. 49, no. 5, pp. 1081–1096, 2014.
- [34] H. Sakurai, Y. Kobayashi, T. Mitomo, O. Watanabe, and S. Otaka, "A 1.5GHz-modulation-range 10ms-modulation-period 180kHzrms-frequency-error 26MHz-reference mixed-mode FMCW synthesizer for mm-wave radar application," in *2011 IEEE International Solid-State Circuits Conference*, 2011, pp. 292–294.
- [35] M. Jahn, R. Feger, C. Wagner, Z. Tong, and A. Stelzer, "A Four-Channel 94-GHz SiGe-Based Digital Beamforming FMCW Radar," *IEEE Transactions on Microwave Theory and Techniques*, vol. 60, no. 3, pp. 861–869, 2012.
- [36] H. P. Forstner, H. Knapp, H. Jager, E. Kolmhofer, J. Platz, F. Starzer, M. Tremml, A. Schinko, G. Birschkus, J. Bock, K. Aufinger, R. Lachner, T. Meister, H. Schafer, D. Lukashevich, S. Boguth, A. Fischer, F. Reininger, L. Maurer, J. Minichshofer, and D. Steinbuch, "A 77GHz 4-Channel Automotive Radar Transceiver in SiGe," in *2008 IEEE Radio Frequency Integrated Circuits Symposium*, 2008, pp. 233–236.
- [37] H. Hashemi, X. Guan, A. Komijani, and A. Hajimiri, "A 24-GHz SiGe Phased-Array Receiver-LO Phase-Shifting Approach," *IEEE Transactions on Microwave Theory and Techniques*, vol. 53, no. 2, pp. 614–626, 2005.
- [38] H. Hashemi and S. Raman, *Mm-wave Silicon Power Amplifiers and Transmitters*. Cambridge University Press, 2016.

- [39] A. Natarajan, A. Valdes-Garcia, B. Sadhu, S. K. Reynolds, and B. D. Parker, "W - Band Dual-Polarization Phased-Array Transceiver Front-End in SiGe BiCMOS," *IEEE Transactions on Microwave Theory and Techniques*, vol. 63, no. 6, pp. 1989–2002, 2015.
- [40] M. Tabesh, A. Arbabian, and A. Niknejad, "60GHz low-loss compact phase shifters using a transformer-based hybrid in 65nm CMOS," in *2011 IEEE Custom Integrated Circuits Conference (CICC)*, 2011, pp. 1–4.
- [41] J. A. Hou and Y. H. Wang, "A Compact Quadrature Hybrid Based on High-Pass and Low-Pass Lumped Elements," *IEEE Microwave and Wireless Components Letters*, vol. 17, no. 8, pp. 595–597, 2007.
- [42] S. A. Maas, *Nonlinear Microwave and RF Circuits*. Artech House, 2003.
- [43] M. Tabesh, J. Chen, C. Marcu, L. Kong, S. Kang, A. M. Niknejad, and E. Alon, "A 65 nm CMOS 4-Element Sub-34 mW/Element 60 GHz Phased-Array Transceiver," *IEEE Journal of Solid-State Circuits*, vol. 46, no. 12, pp. 3018–3032, 2011.
- [44] R. Pilard, D. Titz, F. Giancesello, P. Calascibetta, J. M. Rivière, J. Lopez, R. Coffy, E. Saugier, A. Poulain, F. Ferrero, C. Luxey, P. Brachat, G. Jacquemod, and D. Gloria, "HDI organic technology integrating built-in antennas dedicated to 60 GHz SiP solution," in *Proceedings of the 2012 IEEE International Symposium on Antennas and Propagation*, 2012, pp. 1–2.
- [45] D. Titz, F. Ferrero, and C. Luxey, "Development of a Millimeter-Wave Measurement Setup and Dedicated Techniques to Characterize the Matching and Radiation Performance of Probe-Fed Antennas [Measurements Corner]," *IEEE Antennas and Propagation Magazine*, vol. 54, no. 4, pp. 188–203, 2012.
- [46] A. Bisognin, D. Titz, F. Ferrero, G. Jacquemod, R. Pilard, F. Giancesello, D. Gloria, C. Laporte, H. Ezzeddine, D. Lugara, and C. Luxey, "Probe-fed measurement system for F-band antennas," in *The 8th European Conference on Antennas and Propagation (EuCAP 2014)*, 2014, pp. 722–726.
- [47] A. Valdes-Garcia, A. Natarajan, D. Liu, M. Sanduleanu, X. Gu, M. Ferriss, B. Parker, C. Baks, J. O. Plouchart, H. Ainspan, B. Sadhu, M. Islam, and S. Reynolds, "A fully-integrated dual-polarization 16-element W-band phased-array transceiver in SiGe BiCMOS," in *2013 IEEE Radio Frequency Integrated Circuits Symposium (RFIC)*, 2013, pp. 375–378.
- [48] S. Shahramian, Y. Baeyens, N. Kaneda, and Y. K. Chen, "A 70–100 GHz Direct-Conversion Transmitter and Receiver Phased Array Chipset Demonstrating 10 Gb/s Wireless Link," *IEEE Journal of Solid-State Circuits*, vol. 48, no. 5, pp. 1113–1125, 2013.
- [49] F. Golcuk, T. Kanar, and G. M. Rebeiz, "A 90 - 100-GHz 4 x 4 SiGe BiCMOS Polarimetric Transmit/Receive Phased Array With Simultaneous Receive-Beams Capabilities," *IEEE Transactions on Microwave Theory and Techniques*, vol. 61, no. 8, pp. 3099–3114, 2013.

- [50] P. J. Peng, P. N. Chen, C. Kao, Y. L. Chen, and J. Lee, "A 94 GHz 3D Image Radar Engine With 4TX/4RX Beamforming Scan Technique in 65 nm CMOS Technology," *IEEE Journal of Solid-State Circuits*, vol. 50, no. 3, pp. 656–668, 2015.
- [51] P. N. Chen, P. J. Peng, C. Kao, Y. L. Chen, and J. Lee, "A 94GHz 3D-image radar engine with 4TX/4RX beamforming scan technique in 65nm CMOS," in *2013 IEEE International Solid-State Circuits Conference Digest of Technical Papers*, 2013, pp. 146–147.
- [52] X. Yu, S. P. Sah, H. Rashtian, S. Mirabbasi, P. P. Pande, and D. Heo, "A 1.2-pJ/bit 16-Gb/s 60-GHz OOK Transmitter in 65-nm CMOS for Wireless Network-On-Chip," *IEEE Transactions on Microwave Theory and Techniques*, vol. 62, no. 10, pp. 2357–2369, 2014.
- [53] X. Yu, H. Rashtian, S. Mirabbasi, P. P. Pande, and D. Heo, "An 18.7-Gb/s 60-GHz OOK Demodulator in 65-nm CMOS for Wireless Network-on-Chip," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 62, no. 3, pp. 799–806, 2015.
- [54] N. Dolatsha, B. Grave, M. Sawaby, C. Chen, A. Babveyh, S. Kananian, A. Bisognin, C. Luxey, F. Ganesello, J. Costa, C. Fernandes, and A. Arbabian, "17.8 A compact 130GHz fully packaged point-to-point wireless system with 3D-printed 26dBi lens antenna achieving 12.5Gb/s at 1.55pJ/b/m," in *2017 IEEE International Solid-State Circuits Conference (ISSCC)*, 2017, pp. 306–307.
- [55] J. D. Park, S. Kang, S. V. Thyagarajan, E. Alon, and A. M. Niknejad, "A 260 GHz fully integrated CMOS transceiver for wireless chip-to-chip communication," in *2012 Symposium on VLSI Circuits (VLSIC)*, 2012, pp. 48–49.
- [56] R. Wu, R. Minami, Y. Tsukui, S. Kawai, Y. Seo, S. Sato, K. Kimura, S. Kondo, T. Ueno, N. Fajri, S. Maki, N. Nagashima, Y. Takeuchi, T. Yamaguchi, A. Musa, K. K. Tokgoz, T. Siriburanon, B. Liu, Y. Wang, J. Pang, N. Li, M. Miyahara, K. Okada, and A. Matsuzawa, "64-QAM 60-GHz CMOS Transceivers for IEEE 802.11ad/ay," *IEEE Journal of Solid-State Circuits*, vol. 52, no. 11, pp. 2871–2891, 2017.
- [57] J. Pang, S. Maki, S. Kawai, N. Nagashima, Y. Seo, M. Dome, H. Kato, M. Katsuragi, K. Kimura, S. Kondo, Y. Terashima, H. Liu, T. Siriburanon, A. T. Narayanan, N. Fajri, T. Kaneko, T. Yoshioka, B. Liu, Y. Wang, R. Wu, N. Li, K. K. Tokgoz, M. Miyahara, K. Okada, and A. Matsuzawa, "24.9 A 128-QAM 60GHz CMOS transceiver for IEEE802.11ay with calibration of LO feedthrough and I/Q imbalance," in *2017 IEEE International Solid-State Circuits Conference (ISSCC)*, 2017, pp. 424–425.
- [58] Y. Yang, S. Zahir, H. Lin, O. Inac, W. Shin, and G. M. Rebeiz, "A 155 GHz 20 Gbit/s QPSK transceiver in 45nm CMOS," in *2014 IEEE Radio Frequency Integrated Circuits Symposium*, 2014, pp. 365–368.

- [59] D. Fritsche, G. Tretter, C. Tzschoppe, C. Carta, and F. Ellinger, "A low-power SiGe BiCMOS 190 GHz receiver with 47 dB conversion gain and 11 dB noise figure for ultra-large-bandwidth applications," in *2016 IEEE Radio Frequency Integrated Circuits Symposium (RFIC)*, 2016, pp. 290–293.
- [60] C. Jiang, A. Cathelin, and E. Afshari, "A High-Speed Efficient 220-GHz Spatial-Orthogonal ASK Transmitter in 130-nm SiGe BiCMOS," *IEEE Journal of Solid-State Circuits*, vol. 52, no. 9, pp. 2321–2334, 2017.
- [61] S. Daneshgar, K. Dasgupta, C. Thakkar, A. Chakrabarti, S. Yamada, D. Choudhury, J. Jaussi, and B. Casper, "A 27.8Gb/s 11.5pJ/b 60GHz transceiver in 28nm CMOS with polarization MIMO," in *2018 IEEE International Solid - State Circuits Conference - (ISSCC)*, 2018, pp. 166–168.
- [62] K. K. Tokgoz, S. Maki, J. Pang, N. Nagashima, I. Abdo, S. Kawai, T. Fujimura, Y. Kawano, T. Suzuki, T. Iwai, K. Okada, and A. Matsuzawa, "A 120Gb/s 16QAM CMOS millimeter-wave wireless transceiver," in *2018 IEEE International Solid - State Circuits Conference - (ISSCC)*, 2018, pp. 168–170.
- [63] D. Chowdhury, P. Reynaert, and A. M. Niknejad, "Design Considerations for 60 GHz Transformer-Coupled CMOS Power Amplifiers," *IEEE Journal of Solid-State Circuits*, vol. 44, no. 10, pp. 2733–2744, 2009.
- [64] D. J. Cassan and J. R. Long, "A 1-V transformer-feedback low-noise amplifier for 5-GHz wireless LAN in 0.18- μ m CMOS," *IEEE Journal of Solid-State Circuits*, vol. 38, no. 3, pp. 427–435, 2003.
- [65] S. Kundu and J. Paramesh, "A 17 GHz transformer-neutralized current re-use LNA and its application to a low-power RF front-end," in *2010 IEEE Radio Frequency Integrated Circuits Symposium*, 2010, pp. 307–310.
- [66] Z. Deng, "Design Techniques for High-Frequency CMOS Integrated Circuits: From 10 GHz To 100 GHz," PhD thesis, University of California, Berkeley, 2011.
- [67] V. Bhagavatula, T. Zhang, A. R. Suvarna, and J. C. Rudell, "An Ultra-Wideband IF Millimeter-Wave Receiver With a 20 GHz Channel Bandwidth Using Gain-Equalized Transformers," *IEEE Journal of Solid-State Circuits*, vol. 51, no. 2, pp. 323–331, 2016.
- [68] M. Vigilante and P. Reynaert, "On the Design of Wideband Transformer-Based Fourth Order Matching Networks for E-Band Receivers in 28-nm CMOS," *IEEE Journal of Solid-State Circuits*, vol. 52, no. 8, pp. 2071–2082, 2017.
- [69] K. Dasgupta, K. Sengupta, A. Pai, and A. Hajimiri, "A mm-Wave Segmented Power Mixer," *IEEE Transactions on Microwave Theory and Techniques*, vol. 63, no. 4, pp. 1118–1129, 2015.
- [70] S. Shopov, A. Balteanu, and S. P. Voinigescu, "A 19 dBm, 15 Gbaud, 9 bit SOI CMOS Power-DAC Cell for High-Order QAM W-Band Transmitters," *IEEE Journal of Solid-State Circuits*, vol. 49, no. 7, pp. 1653–1664, 2014.

- [71] E. Laskin, M. Khanpour, S. T. Nicolson, A. Tomkins, P. Garcia, A. Cathelin, D. Belot, and S. P. Voinigescu, "Nanoscale CMOS Transceiver Design in the 90–170-GHz Range," *IEEE Transactions on Microwave Theory and Techniques*, vol. 57, no. 12, pp. 3477–3490, 2009.
- [72] R. S. Carson, *High Frequency Amplifiers*. John Wiley and Sons, 1975.
- [73] A. M. Niknejad, *Lecture notes, EE142*, 2011.
- [74] B. Razavi, *Design of Analog CMOS Integrated Circuits*. McGraw-Hill Education, 2000.
- [75] D. M. Pozar, *Microwave Engineering*. Wiley, 2012.
- [76] J. D. Park, S. Kang, and A. M. Niknejad, "A 0.38 THz Fully Integrated Transceiver Utilizing a Quadrature Push-Push Harmonic Circuitry in SiGe BiCMOS," *IEEE Journal of Solid-State Circuits*, vol. 47, no. 10, pp. 2344–2354, 2012.
- [77] S. Kang, S. V. Thyagarajan, and A. M. Niknejad, "A 240 GHz Fully Integrated Wideband QPSK Transmitter in 65 nm CMOS," *IEEE Journal of Solid-State Circuits*, vol. 50, no. 10, pp. 2256–2267, 2015.
- [78] E. M. Cherry and D. E. Hooper, "The design of wide-band transistor feedback amplifiers," *Electrical Engineers, Proceedings of the Institution of*, vol. 110, no. 2, pp. 375–389, 1963.
- [79] S. V. Thyagarajan, S. Kang, and A. M. Niknejad, "A 240 GHz Fully Integrated Wideband QPSK Receiver in 65 nm CMOS," *IEEE Journal of Solid-State Circuits*, vol. 50, no. 10, pp. 2268–2280, 2015.
- [80] S. Chehrazai, A. Mirzaei, and A. A. Abidi, "Noise in Current-Commutating Passive FET Mixers," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 57, no. 2, pp. 332–344, 2010.
- [81] M. T. Terrovitis and R. G. Meyer, "Noise in current-commutating CMOS mixers," *IEEE Journal of Solid-State Circuits*, vol. 34, no. 6, pp. 772–783, 1999.
- [82] T. Melly, A. S. Porret, C. C. Enz, and E. A. Vittoz, "An analysis of flicker noise rejection in low-power and low-voltage CMOS mixers," *IEEE Journal of Solid-State Circuits*, vol. 36, no. 1, pp. 102–109, 2001.
- [83] J. Coonrod, Personal communication, May 23, 2013.
- [84] S. L. Smith and V. Dyadyuk, "Measurement of the dielectric properties of Rogers R/flex 3850 liquid crystalline polymer substrate in V and W band," in *2005 IEEE Antennas and Propagation Society International Symposium*, vol. 4B, 2005, 435–438 vol. 4B.
- [85] D. Sievenpiper, L. Zhang, R. F. J. Broas, N. G. Alexopolous, and E. Yablonovitch, "High-impedance electromagnetic surfaces with a forbidden frequency band," *IEEE Transactions on Microwave Theory and Techniques*, vol. 47, no. 11, pp. 2059–2074, 1999.

- [86] R. E. Collin, *Field Theory of Guided Waves*. McGraw Hill, 1960.
- [87] D. F. Sievenpiper, "High-impedance electromagnetic surfaces," PhD thesis, University of California, Los Angeles, 1999.
- [88] W. E. McKinzie, D. M. Nair, B. A. Thrasher, M. A. Smith, E. D. Hughes, and J. M. Parisi, "60-GHz 2x2 LTCC Patch Antenna Array With an Integrated EBG Structure for Gain Enhancement," *IEEE Antennas and Wireless Propagation Letters*, vol. 15, pp. 1522–1525, 2016.
- [89] M. J. Al-Hasan, T. A. Denidni, and A. R. Sebak, "Millimeter-Wave EBG-Based Aperture-Coupled Dielectric Resonator Antenna," *IEEE Transactions on Antennas and Propagation*, vol. 61, no. 8, pp. 4354–4357, 2013.
- [90] Q. Zhong, W. Choi, C. Miller, R. Henderson, and K. K. O, "25.2 A 210-to-305GHz CMOS receiver for rotational spectroscopy," in *2016 IEEE International Solid-State Circuits Conference (ISSCC)*, 2016, pp. 426–427.
- [91] A. Bisognin, N. Nachabe, C. Luxey, F. Giansello, D. Gloria, J. R. Costa, C. A. Fernandes, Y. Alvarez, A. Arboleya-Arboleya, J. Laviada, F. Las-Heras, N. Dolatsha, B. Grave, M. Sawaby, and A. Arbabian, "Ball Grid Array Module With Integrated Shaped Lens for 5G Backhaul/Fronthaul Communications in F-Band," *IEEE Transactions on Antennas and Propagation*, vol. 65, no. 12, pp. 6380–6394, 2017.
- [92] S. B. Yeap and Z. N. Chen, "Microstrip Patch Antennas With Enhanced Gain by Partial Substrate Removal," *IEEE Transactions on Antennas and Propagation*, vol. 58, no. 9, pp. 2811–2816, 2010.
- [93] S. B. Yeap, Z. N. Chen, and X. Qing, "Gain-Enhanced 60-GHz LTCC Antenna Array With Open Air Cavities," *IEEE Transactions on Antennas and Propagation*, vol. 59, no. 9, pp. 3470–3473, 2011.
- [94] K.-S. Chin, J.-A. Liu, C. C. Chang, and J.-C. Cheng, "LTCC differential-fed patch antennas with rat-race feeding structures," *Progress In Electromagnetics Research C*, vol. 32, pp. 95–108, 2012.
- [95] S. J. Huang, Y. C. Yeh, H. Wang, P. N. Chen, and J. Lee, "W -Band BPSK and QPSK Transceivers With Costas-Loop Carrier Recovery in 65-nm CMOS Technology," *IEEE Journal of Solid-State Circuits*, vol. 46, no. 12, pp. 3033–3046, 2011.
- [96] S. Kang, C. Thakkar, N. Narevsky, K. Dasgupta, S. Daneshgar, J. Jaussi, and B. Casper, "A 40Gb/s 6pJ/b RX baseband in 28nm CMOS for 60GHz polarization MIMO," in *2018 IEEE International Solid - State Circuits Conference - (ISSCC)*, 2018, pp. 164–166.