

©Copyright 2013
Venumadhav Bhagavatula

Power and area optimization techniques for ultra-wideband millimeter-wave CMOS transceivers

Venumadhav Bhagavatula

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2013

Reading Committee:

Jacques C. Rudell, Chair

David J. Allstot

Sudip Shekhar

Program Authorized to Offer Degree
Department of Electrical Engineering

Abstract

Power and area optimization techniques for ultra-wideband millimeter-wave CMOS transceivers

Venumadhav Bhagavatula

Chair of the Supervisory Committee:
Professor Jacques C. Rudell
Department of Electrical Engineering

Over the past decade, opportunities for utilizing the broadband spectrum available at millimeter-wave (mm-wave) frequencies has motivated research on both short and long-range, highly-integrated complementary metal oxide semiconductor (CMOS) transceivers. Prototype mm-wave CMOS transceivers have been demonstrated for application in high-speed data transfer (57-64 GHz), wireless back-haul (71-76 GHz), automotive radar (77GHz) and medical imaging (90 GHz) systems. However, in spite of promising results, large scale deployment of mm-wave CMOS transceivers in portable and hand-held electronics is currently hindered by front-end power-consumptions on the order of several watts. Moreover, as a first order approximation, power consumption is directly proportional to system bandwidth. Therefore, as the bandwidth requirements of systems increase, the challenge with on-chip power consumption will become increasingly difficult to solve.

In this dissertation, techniques for optimizing the power and area of ultra-wideband millimeter-wave transceivers are described. This work resulted in the fabrication of three mm-wave integrated circuits (IC), all of which were realized in a 6-metal layer 40-nm CMOS process. The first IC is a multi-stage transformer-feedback based 11-to-13 GHz direct-conversion receiver. The device achieves a 16% fractional-bandwidth, a peak power-gain of 27.6dB, and noise-figure of 5.3dB while consuming 28.8mW from a 0.9V supply. Second, a compact 24-54GHz 2-stage bandpass distributed amplifier which utilizes dual mirror-symmetric Norton transformations to reduce inductor component values allowing efficient layout to occupy an active area of 0.15mm². The device has a 77% fractional-bandwidth, an overall gain of 6.3dB, a minimum in-band IIP3 of 11dBm, while consuming 34mA from a 1V supply. The third, and the IC which includes the most integration among the three, is an ultra-broadband single-element heterodyne receiver intended for use in low-power phased-array systems. The receiver maintains 17GHz of bandwidth from the mm-wave front end, through a high-IF stage, and to the baseband output. The device occupies 1.2mm² and exploits properties of gain-equalized transformers throughout the signal path to achieve an overall 17GHz bandwidth 20dB gain with a flat in-band response, 7.8dB DSB NF, and a P_{-1dB} of -24dBm, while consuming 104mW off a 1.1V supply.

Table of Contents

1	Introduction	1
1.1	Research objectives	4
1.2	Overview and organization of the thesis	4
2	Millimeter-wave Systems	8
2.1	Phased-array receivers.....	8
2.2	Operating Principle	10
2.3	Phased-array receiver architecture	12
2.4	Element scalability and impact on total power	18
2.5	Ultra-wideband receiver architecture	23
3	Transformer Feedback.....	28
3.1	Current Feedback	30
3.2	Source-to-Gate Transformer Feedback Based Matching Networks	34
3.3	Matching Network Design	37
3.4	Transconductance and Noise.....	41
3.5	Circuit Design	47
3.5	Measurement Results and Comparison.....	54
3.6	Conclusion.....	56
	Appendix-3.1	57

Appendix-3.2.....	58
Appendix-3.3.....	58
4 Band-Pass Distributed-Amplifier.....	65
4.1 Canonical form of a BPDA.....	66
4.3 Norton Transformation.....	70
4.4 Gain-cell design.....	73
4.5 Implementation.....	73
4.5 Measurement Results	76
4.6 Conclusions	81
5 Ultra-Wideband Millimeter-wave Receiver.....	83
5.1 Second-order resonant tanks	84
5.2 Coupled Resonant tanks	89
5.3 Receiver architecture, modeling and design	104
5.4 Measured results.....	110
5.5 Conclusions	116
Appendix 5.1	118
6 Conclusion	122
6.1 Thesis Summary.....	122
6.2 Future Directions.....	123

List of Figures

Fig. 2.1 Plane-wave incident on a two-element phased-array receiver.....	10
Fig. 2.2 (a) Signal path and (b) LO-path of a direct-conversion phased-array receiver with signal-path phase-shifting	13
Fig. 2.3 (a) Signal-path and (b) LO-path of a heterodyne phased-array receiver with signal path phase-shifting.....	14
Fig. 2.4 (a) Signal-path and (b) LO-path of a direct-conversion phased-array receiver with LO-path phase-shifting	15
Fig. 2.5 (a) Signal-path and (b) LO-path of a heterodyne phased-array receiver with LO-path phase-shifting	17
Fig. 2.6. LO distribution network for a phased-array with one and two elements.....	19
Fig. 2.7 (a) Total power consumption of a direct-conversion (blue) and heterodyne (red) phased-array receiver (b) Power saving in a heterodyne receiver as a function of a number of array elements.....	22
Fig. 2.8. Millimeter-wave spectrum targeted in this receiver	23
Fig. 2.9. Frequency plan for wideband heterodyne architecture.....	24
Fig. 2.10. Block diagram of the ultra-wideband mm-wave receiver implemented in a 40nm CMOS.....	24
Fig. 3.1. Heterodyne phased-array receiver: low-noise amplifier (LNA), power-combiner (PC) followed by 11-to-13 GHz IF stage comprised of an IF amplifier, quadrature down-conversion mixer, and lumped-element Lange coupler for I/Q generation.....	29
Fig. 3.2. (a) Generic feedback circuits (b) transformer-feedback circuit.....	29
Fig. 3.3. (a) Common-source (CS) amplifier with inductor degeneration (b) Inductively degenerated amplifier with current feedback (c) Input impedance Z_{in} with $g_{m1} = g_{m2}$ (d) Z_{in} with $g_{m1} = g_{m2}(1 + \varphi)$	31
Fig. 3.4. Schematic and small-signal model of an SGTxvFB amplifier	33
Fig. 3.5 Real and imaginary admittance: model versus circuit simulation	36
Fig. 3.6. Design space for matching TC1 to a 50- Ω source resistance with $g_m = (25mS, 35mS, 50mS, 75mS, 125mS)$ and $0.3 < k < 0.7$ (a) transformer turns-ratio (b) matching network Q factor	40
Fig. 3.7 Small-signal model for noise calculation	41
Fig. 3.8 (a) Effective trans-conductance (b) Thermal noise contribution in TC1.....	42
Fig. 3.9. Two-stage stagger-tuned IF-amplifier with SGTxvFB driving the mixer transconductance.....	47
Fig. 3.10. Simulated frequency response of IFA1 and IFA2	48

Fig. 3.11 Compact floor-plan for multiple transformer design (a) transformer-coupled circuit (b) layout of multiple transformer-coupled stages (c) transformer-feedback circuit (d) layout of multiple transformer-feedback stages	49
Fig. 3.12 Quadrature down-conversion IF-mixer	50
Fig. 3.13 Layout of the three-winding transformer which couples the mixer transconductance to the switching stages	51
Fig. 3.14. Transformer-based lumped-element Lange coupler	52
Fig. 3.15. Chip micrograph	53
Fig. 3.16. (a) Input matching S_{11} (dB) (b) IF-section down-conversion gain	54
Fig. 3.17 (a) noise figure measurement set-up (b) NF versus frequency	55
Fig. 3.18 Comparison with state-of-the-art	56
Fig. 4.1. (a) Canonical form of the LPDA and BPDA (b) Low-pass to band-pass filter transformation	66
Fig. 4.2. Canonical form of the BPDA	67
Fig. 4.3. Norton transformation of a series floating inductor	70
Fig. 4.4. Derivation of the compact-area bandpass filter from the canonical bandpass filter through the application of mirror-symmetric dual Norton-transforms	71
Fig. 4.5. Input and output impedances of the gain-cells	72
Fig. 4.6. Circuit diagram of the BPDA	74
Fig. 4.7. Chip Micrograph for the BPDA	75
Fig. 4.8. Setup for BPDA S-Parameter measurement	76
Fig. 4.9. Measured S-parameters	77
Fig. 4.10. Setup for BPDA noise characterization using the N8975A noise-figure analyzer	78
Fig. 4.11. Compression-point, group-delay, and IIP3 characterization versus frequency	79
Fig. 5.1. Ideal amplifier with second-order R-L-C tank	84
Fig. 5.2. Cascaded amplifier with R-L-C load	86
Fig. 5.3. (a) Gain (A_{v1}) of transistor amplifier with transconductance g_m , and R-L-C load (b) Gain (A_{v1}) of transistor amplifier with transconductance αg_m . Value of inductor scaled down to maintain a fixed resonance frequency. (c) Plot of $A_{v1}A_{v2}$ as a function of scaling-factor α	88

Fig. 5.4 <i>Dr</i> – Driver Amplifier, <i>Ld</i> - Load Amplifier. (a) Single inductor to resonant capacitor $C_D = C$ or, and load capacitor $C_L = C$ (b) Load resonant tanks separated by a large DC-Block capacitor C_{BL} (c) coupled by series resonant tank L_2 and C_2 (d) coupled by series-capacitor C_c (e) coupled by mutual-inductor M (f) coupled by M and C_c (gain-equalized transformer).....	89
Fig. 5.5 Magnetically coupled resonant tanks	91
Fig. 5.6 Electrically-coupled resonant tanks.....	95
Fig. 5.7 Trans-resistance of a magnetically-coupled (MC), and electrically-coupled (CC) resonant tanks.....	98
Fig. 5.8 Electric and magnetically coupled resonant tank	99
Fig. 5.9. Trans-resistance of a gain-equalized transformer.....	101
Fig. 5.10. (a) Schematic of magnetically-coupled (MC) transformer (b) Schematic of gain-equalized (MC-CC) transformer (c) Comparison of the trans-resistance of circuit-a and circuit-b.....	102
Fig. 5.11 HFSS model of the gain-equalized transformer	103
Fig. 5.12 Block diagram of the wideband, heterodyne millimeter-wave receiver	105
Fig. 5.13. HFSS model of the G-S-G pad and the input balun	106
Fig. 5.14. Schematic of the three-stage low-noise amplifier	106
Fig. 5.15 Schematic of RF-mixer, two-stage IF-amplifier and IF-mixer.....	108
Fig. 5.16. Schematic of the LO distribution network.....	108
Fig. 5.17 Power break-up for the different circuit blocks	109
Fig. 5.18. Die micrograph of the millimeter-wave CMOS receiver	109
Fig. 5.19 Test setup for conversion-gain and linearity measurements.....	111
Fig. 5.20 Receiver frequency response measured at the baseband output. Referred to the receiver front-end, a gain of $20 \pm 1.5\text{dB}$ is maintained across a 51-to-68 GHz bandwidth.	112
Fig. 5.21 Measured input matching, noise-figure, input compression point.....	113
Fig. 5.22 Test setup for noise-figure measurement.....	114

Acknowledgements

This work is a culmination of collaborations with many wonderful and talented people and would not have been possible without their contributions. First and foremost, I would like to thank my advisor Dr. Rudell; working with him has been a pleasure and a privilege. I wish him the very best, as he leads the lab towards greater heights. Over the years he has given me the freedom to grow as a researcher, and the flexibility to explore areas that I found interesting. He has been, and will remain, a constant source of motivation for my career.

I would like to thank Dr. Allstot for serving on my PhD supervisory committee. While I have not had the chance to work in close quarters with Dr. Allstot, I have benefited immensely by studying the work done by him and his students. The depth and breadth of his contributions has set a high standard for all of us at the university. Thanks are also due to Dr. Shekhar and Dr. Wang for reviewing my thesis and serving on my committee.

Life in graduate school, especially in an IC design group, is bound to be stressful at times. However, even when completely submerged in work, spending time in the lab was enjoyable thanks to my friends and lab-mates. I would like to thank Will, Tong, Apsara, Eric, Samrat, Soonkyun, Jason for all the good times. They all have played a significant role in my academic progression through graduate school. My first project was done in collaboration with Will, while Tong and I spent countless hours in the lab designing and measuring the millimeter-wave test-chips. I wish the very best to the younger lab members Daniel, Chenxhi and Yongdong.

As I took initial steps in the millimeter-wave IC design-space I have had the luxury of being mentored by several exceptional circuit designers. I am very grateful to Michael Boers for

hosting me during my six-month internship at Broadcom. Mike took a personal interest to ensure I was provided opportunities to learn. I would like to thank M. Nariman, S. Sarkar, B. Perumana, P. Sen, B. Afhsar and E. Adabi for patiently answering all my questions. I am very grateful to my friends K. Agarwal (Stanford) and A. Chakrabarti (Columbia University) for their guidance and support when I was designing my millimeter-wave test-chips at school. I have also gained a lot from my discussions with M. Taghivand, M. Wiklund and R. Brockenborough during my visits to Qualcomm.

Success in graduate school is not merely an academic endeavor, but a personal journey as well. Here too I have been extremely fortunate to have many of ‘my’ people by my side. My friends Ullas and Sulbha, have been family to me in this foreign land. I have had the luxury of having great housemates over the years – Manohar, Nihar, Virag, Jaymin, Kiran, Ullas, Roshan and Deepak. I would like to thank Prasad Bhosale and Srikar Bhagavatula for being around for a chat at all times.

I am grateful to my in-laws, Ravish and Suvarna, for the immense amount of faith they placed in me. With the PhD out of the way, my next big project is to learn to speak fluent Kannada.

Without the support and encouragement of my elder brothers, Rama and Vamsi, and their better halves, I would never have come back to graduate school. I am just a couple of months away from joining them on the wrong-side of 30, but I remain the ‘kid’ brother whom they watch out for. I look forward to spending more time with my wonderful nephews, Pranav and Dhruva, and niece, Tanvi

Special thanks go out to my wife, Apsara, for all her patience, love and encouragement. None of this work would have been possible without her support. She deserves a giant share of the credit for any success that I might achieve. I couldn't have asked for a better woman to share my life with.

Although I mention them last, my *Amma* and *Nannagaru* have made the greatest contribution to this work. Over the past forty years, my parents have provided limitless help and opportunities to whomsoever they have met, not just their own children. Ma, Nanna - There are so many people in this world who owe everything to your guidance; I humbly join at the end of the line. Although I will never be able to repay you for all that you have given to me over the years, this dissertation is but a small token of appreciation. For that, I dedicate this work to you with all my love.

To,

Nannagaru and Amma

and the ideals they stand for

To,

My family - *The Bhagavatula(s)*

1 INTRODUCTION

The applications for single-chip CMOS electronics in the millimeter-wave and terahertz spectrum promises to provide antenna, circuit, device and system engineers with some of the most exciting opportunities for innovation over the next decade. Not long ago, the *radio frequency* (RF) band between 500MHz and 10GHz was considered **high** frequency. However, the astonishingly rapid developments in the science and technology of device fabrication, and the resultant scaling of minimum device dimensions, has extended maximum device operating frequency in excess of several hundreds of gigahertz. While silicon-germanium and indium-phosphide devices have superior high-frequency electrical characteristics compared to CMOS, for low-cost integrated systems where analog and digital circuitry must coexist, CMOS has emerged as the technology of choice. The millimeter-wave (30GHz to 300GHz) and terahertz (300GHz to 3000GHz) spectrum, hitherto unexplored in CMOS, provides researchers with the luxury of virtually limitless bandwidth.

Systems designed for the RF-band enabled the semiconductor industry to power a global boom in low-cost wireless communication and revolutionized the way individuals and machines communicate with one another. As recently as 1990, cellular-phones were a luxury-item and a very basic form of wireless connectivity was observed in consumer-electronics. Now, twenty years on, it is common for an individual to carry a laptop, cellular-phone, tablet all of which are wirelessly connected to the world-wide web. In short, improvements in semiconductor technology manufacturing and innovations in circuit design have played a role in making life

simpler by providing increased mobile connectivity and ushering in a new generation of mobile applications.

The next major thrust for the semiconductor industry could possibly come from the growing bio-instrumentation industry. Currently, instruments used in medical clinics and hospitals across the world, tend to be expensive. This in turn makes the cost of health-care prohibitively high; a problem more prevalent in under-developed, and developing countries. Moreover, the lack of mobility in bulky equipment limits the ability of the doctor operating in remote and inaccessible locations. CMOS based integrated circuits could potentially bring the benefits of small form-factor and low-cost to the bio-instrumentation domain.

The ultra-wide bandwidth available at millimeter-wave frequencies brings significant advantages for both consumer-electronics and biomedical applications. In high-speed wireless communication market, data-rate is the golden performance metric. Data-rate, in turn, can be increased by (a) higher channel-bandwidth and (b) high-order complex modulation techniques. At RF, the spectrum has to be shared between multiple standards such as Bluetooth, Wi-Fi, GSM and Zigbee. As a result, complex time- and frequency-multiplexed schemes are required to maximize the bandwidth utilization efficiency. However, high-order modulation techniques increase the complexity of the RF front-end and baseband digital circuitry. At millimeter-wave, even with a simple modulation technique such as FSK and QPSK, high data-rates can be achieved due to the ultra-wide bandwidth available.

Ultra-wideband circuits will also play a pivotal role in integrated imaging-systems for medical and security applications. The operation of active imaging systems is very similar to the operation of pulse-based radar systems. A narrow-pulse is shot at the object, and the information

in the pulse reflected by the object is used to create the image. A narrow-pulse results in improved timing resolution, which is reflected in improved spatial-resolution in the reconstructed image. A narrow-pulse in time-domain translates to a higher bandwidth in the frequency-domain. Thus, high-resolution imaging systems will also benefit from wideband millimeter-wave circuits.

The large number of applications for millimeter-wave integrated-circuits has made this topic a very active area of research. The seminal paper by Emami et.al [1] is among the earliest reports of CMOS device modeling and circuit design at 60-GHz. Initial research focused on the implementation of single-blocks such as mixers [2] and low-noise amplifiers [3]. Implementation of direct-conversion [4][5] and heterodyne receivers [6] followed. The growing confidence in mm-wave CMOS is reflected in the increasing levels of integration, for example, receivers with on-chip mm-wave frequency-synthesis [7] and phase-locked loops operating at frequencies even beyond the 60GHz standard [8]. Another confirmation on the feasibility of 60-GHz CMOS has been interest not only from academia, but also from industry [9][10]. While the 60-GHz standard provided the impetus, research efforts occurring in parallel have pushed the operating frequencies of mm-wave integrated transceivers beyond 100-GHz [11], [12]. Integrated millimeter-wave imaging has been shown to be useful for detecting minute tumor tissues in the body during cancer therapy and detecting small concealed weapons at security screening check-point. The recent demonstration of a 90-GHz imaging-radar [13], designed for breast cancer diagnosis, is one such example of a millimeter-wave IC solving important and socially-relevant problems.

In summary, moving forward, the biggest challenge in CMOS circuits - RF, mm-wave or terahertz - is how to design circuits capable of handling signals with ultra-wide bandwidth. As described earlier, the demands for wireless data transfer and the need for low-cost medical

electronics is only going to increase with time. For example, to achieve 100 Gbps wireless communication using simple QPSK modulation, transceiver bandwidths exceeding 50 GHz are necessary. Similarly, it can be proven that to reliably resolve moving objects (such as a malignant tumors) with a resolution of 4 mm requires a input pulse generation circuitry with a bandwidth of 40 GHz [13] or higher. Therefore, circuit and system-level techniques which address the issue of achieving extremely high bandwidth, and more importantly, address the issue in a manner that is area and power efficient, will be very important.

1.1 Research objectives

This thesis explores issues with respect to realizing ultra-wideband circuits operating at millimeter-wave frequencies. Traditionally, wideband circuits utilize either multiple inductors (resulting in large area), or low quality-factor load structures (resulting in large power). Therefore, circuit and system-level optimization techniques to minimize area and power-consumption of wideband integrated-circuit receivers are described. While the prototype chips reported in this thesis have been fabricated in CMOS technology, the techniques reported are equally valid in bipolar technologies such as SiGe and GaAs.

1.2 Overview and organization of the thesis

The thesis has the following organization:

Chapter 2: A brief overview of phased-array based millimeter-wave receiver architectures is provided. Different techniques for introducing phase-shift on the signal/LO path and the related trade-offs involved are described. A receiver architecture and frequency plan with the goal of reducing LO power is detailed, along with the challenges introduced by this architecture.

Chapter 3: Feedback is a popular technique to design wideband circuits at analog-frequencies. Extending feedback-techniques for radio frequencies using transformer-based reactive-feedback was first proposed in [14]. In Chapter 3, the design and analysis of a complete direct-conversion receiver which uses multi-stage source-to-gate transformer-feedback to receive signals in the 11-to-13 GHz band is described. The measurement results from a prototype IC designed in a 40nm CMOS process are described.

Chapter 4: The design paradigm of distributed amplifiers (DA) is extended to obtain high fractional-bandwidth band-pass signal amplification. A well-known drawback of the DA topology has been the required silicon-area. Chapter 4 proposes a DA which employs multiple Norton-transformations to reduce the area of the IC. A 24-54GHz wideband band-pass distributed amplifier (BPDA) was designed in a CMOS 40-nm process using this principle. The design and measurements results from the prototype chip are described in detail.

Chapter 5: This chapter explores the use of coupled resonant circuits - magnetic coupling only, electrical coupling only, combination of magnetic and electric coupling (gain-equalized load) – for wideband signal amplification. Using the gain-equalized transformer as the core building block, the design of a 50-to-70 GHz heterodyne millimeter-wave receiver is described. The measurement results from the prototype IC are described in this section.

- [1] C. H. Doan, S. Emami, A. M. Niknejad and R. W. Broderson, "Millimeter-wave CMOS design," in *IEEE Journal of Solid-State Circuits*, vol 40, no.1, pp.144-155, 2005.
- [2] S. Emami, C. H. Doan, A. M. Niknejad, and R. W. Broderson, "A 60-GHz down-converting CMOS single-gate mixer," in *IEEE RFIC Symposium, Digest of Technical Papers*, pp.163-166, 2005.
- [3] A. Natarajan, S. Nicolson, M.-D. Tsai, and B. Flyod, "A 60-GHz variable-gain LNA in 65nm CMOS," in *IEEE ASSCC, Proc. of Technical Papers*, pp.117-120, 2008.
- [4] B. Razavi, "A 60 GHz CMOS receiver front-end," in *IEEE Journal of Solid-State Circuits*, vol.41, no.1, pp.17-22, 2006.
- [5] B. Afshar, and A.M. Niknejad, "A robust 24mW 60GHz receiver in 90nm standard CMOS," in *IEEE ISSCC, Digest of Technical Papers*, pp.182-183, 2008.
- [6] B. Razavi, "A millimeter-wave CMOS heterodyne receiver with on-chip LO and divider," in *IEEE Journal of Solid-State Circuits*, vol. 43, no.2, pp.477-485, 2008.
- [7] T. Mitomo et.al., "A 60-GHz CMOS receiver front-end with frequency synthesizer," in *IEEE Journal of Solid-State Circuits*, vol.43, no.4, pp.1030-1037, 2008.
- [8] J. Lee, M. Liu, and H. Wang, "A 75-GHz phase-locked loop in 90-nm CMOS technology," in *IEEE Journal of Solid-State Circuits*, vol.43, no.6, pp.1414-1426, 2008.
- [9] K. Okada et.al., "A 60-GHz 16QAM/8PSK/QPSK/BPSK direct-conversion transceiver for IEEE802.15.3c," in *IEEE Journal of Solid-State Circuits*, vol. 46, no.12, pp.2988-3004, 2011.
- [10] A. Siligaris et.al. "A 65-nm CMOS fully integrated transceiver module for 60-GHz wireless HD applications," in *IEEE Journal of Solid-State Circuits*, vol.46, no.12, pp.3005-3017, 2011.

- [11] B. Heydari, M. Bohsali, E. Adabi, A. M. Niknejad, "Millimetre-wave devices and circuit blocks up to 104 GHz in 90nm CMOS," in *IEEE Journal of Solid-State Circuits*, vol.42, no.12, pp.2893-2903, 2007.
- [12] E. Laskin, M. Khanpour, S. T. Nicholson, A. Tomkins, P. Garcia, A. Cathelin, D. Belot, and S. P. Voinigescu, "Nanoscale CMOS transceiver design in the 90-170-GHz range," in *IEEE Transactions on Microwave Theory and Techniques*, vol.57, no.12, part:2, pp.3477-3490, 2009.
- [13] Arbabian, S. Callender, S. Kang, B. Afshar, J-C Chien, and A. M. Niknejad, "A 90 GHz hybrid switching pulsed-transmitter for medical imaging," in *IEEE Journal of Solid-State Circuits*, vol.45, no.12, pp.2667-2681, 2010.
- [14] M. T. Reih, and J. R. Long, "A 1.2V reactive-feedback 3.1-10.6 GHz low-noise amplifier in 0.13um CMOS," in *IEEE Journal of Solid-State Circuits*, vol.42, no.5, pp.1023-1033, 2012.

2 MILLIMETER-WAVE SYSTEMS

At millimeter-wave (mm-wave) frequencies spectrum is in abundance while the utilization is still relatively scarce. The limited usage of mm-wave, as opposed to RF, can be attributed to two main reasons. First, before CMOS fabrication technologies reached the 130nm process node, very limited power-gain could be obtained from the CMOS device at mm-wave frequencies. Second, stringent rules from regulatory authorities such as the FCC (in the USA) govern the spectrum usage, and restricted millimeter-wave operation. However, with CMOS minimum device length scaling down to 40nm and beyond, and governments worldwide releasing an unlicensed 60-GHz band for commercial applications, millimeter-wave CMOS is now firmly placed on the technology roadmap.

2.1 Phased-array receivers

One of the biggest challenges in mm-wave system design is the high path-loss associated with these frequencies. The path-loss increases in proportion to the square of the carrier-frequency; therefore, loss observed in mm-wave signal transmission can be significantly higher than at RF. On the receiver side, a high path loss translates to reduced sensitivity and on the transmitter-side a reduction in the transmission-range relative to RF bands. As a result, to improve the signal-transmission efficiency at mm-wave frequencies, narrow directed beams are preferable over isotropic radiation. For mobile applications, in which the relative position between the transmitter and receiver could possibly vary with time, it is important for this beam to be electronically steerable. Phased-arrays, a class of multiple-antenna, multi-element systems, were introduced by the Bell Labs in the 1930s for receivers “*capable of being steered to meet the varying angle at which short radio waves arrive at the receiving location*” [1]. The first phased-

arrays with fully electronic beam steering were developed for the military during the second world-war. Over the course of the next eight decades, phased-arrays have been employed for airborne, space, surface, and ground-based applications [2]. However, the high cost associated with discrete microwave components precluded their widespread use in consumer applications. To drive down the cost, integration of the entire phased-array transceiver, antenna and signal-path, onto a single silicon chip was the next important step.

The large area required by passive components such as inductors, transmission-lines, and antennas is the biggest drawback of phased-array systems. In advance process nodes, fabrication cost are evolving as a major barrier to future consumer electronics, thus realizing minimum size solutions is imperative. However, the size of the passives is inversely proportional to the operating frequency. Thus, from an area perspective, transceivers operating at mm-wave or terahertz carrier frequencies are more amenable to an integrated implementation. The shrinking size of passives and antennas as the carrier frequency increases opens the door to **single-chip** multiple-element, multiple-antenna systems. Gordon Moore's prophecy [3] that "*successful realization of such items as phased-array antennas, for example, using a multiplicity of integrated microwave power sources, could completely revolutionize radar*" finally came true in 2004, when Guan [4] reported the first fully integrated eight-element phased-array receiver for automotive-radar applications in a SiGe-BiCMOS process. More recently, SiGe and CMOS realizations of 24-GHz automotive-radar [4][5], 77-GHz automotive radar [5], 60-GHz short-range data-communication [7], and E-band long-range data communication [9] phased-array systems have been reported.

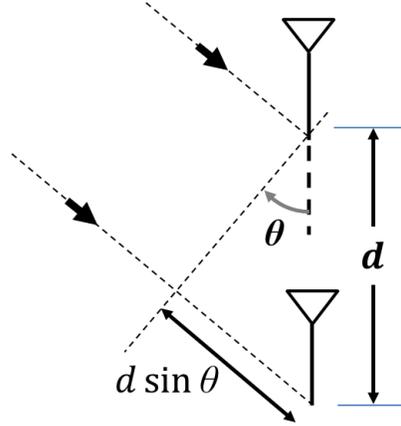


Fig. 2.1 Plane-wave incident on a two-element phased-array receiver

2.2 Operating Principle

The block diagram of a two-element phased-array with an antenna-spacing (d) is shown in Fig. 2.1. For a plane-wave that is incident onto the phased-array receiver, the time delay (τ) between the arrivals at the adjacent antennas is a function of d , the speed of light (c) and the angle of incidence (θ),

$$\tau = d \sin \theta / c \quad (2-1)$$

Assuming a carrier frequency w_c , and that information in the received signal is encoded onto both the amplitude and phase, the signal received at the first antenna $S_o(t)$ and the $(k+1)^{\text{th}}$ antenna $S_k(t)$ can be expressed as,

$$S_o(t) = A(t) \cos(w_c t + \varphi(t)) \quad (2-2)$$

$$S_k(t) = A(t - k\tau) \cos(w_c t - w_c k\tau + \varphi(t - k\tau)) \quad (2-3)$$

In order to ensure constructive-interference between the signals $S_o(t)$ and $S_k(t)$, a time delay of $k\tau$ has to be introduced after $S_o(t)$. From (2-1), it can be observed that the time-delay is

a function of θ . Thus, ideally, electronic beam-steering requires an on-chip tunable time-delay element. However, non-ideal effects such as noise, loss, and non-linearity make tunable time-delay elements difficult to implement at RF frequencies [4]. While tapped LC-ladder based true-time delay circuits have been proposed in [10] and [11], the large silicon area (16mm^2) required makes them less favorable for large array implementations.

For narrow-band systems it is possible to approximate the time-delay element by a phase-shift. To illustrate the validity of this approximation, consider an N-element phased-array with maximum expected antenna-to-antenna time delay, $N\tau$, receiving modulated data with a symbol rate F_{symbol} . If, $N\tau \ll 1/F_{\text{symbol}}$, then

$$A(t - k\tau) \approx A(t) \quad (2-4)$$

$$\varphi(t - k\tau) = \varphi(t) \quad (2-5)$$

Applying (2-4) and (2-5), (2-3) can be simplified to

$$S_k(t) = A(t) \cos(w_c t + \varphi(t) - w_c k\tau) \quad (2-6)$$

Thus, to ensure constructive interference between $S_o(t)$ and $S_k(t)$, a phase-shift $\varphi_k(t) = w_c k\tau$ radians has to be introduced after $S_o(t)$. It is important to note that the phase-shift $\varphi_k(t)$ introduces a correct time delay only at the carrier frequency w_c . Moreover, the error in the time delay is frequency dependent.

The technique to apply a narrow-band phase-shift approximation instead of a true time-delay has been universally applied in transceivers dealing with modulated signals with small fractional bandwidth (fBW). A study of the impact of this approximation on the EVM degradation of wideband signals has been presented in [5].

2.3 Phased-array receiver architecture

One major classification of phased-array architectures is based on whether the phase-delay element is introduced in the signal-path or the local-oscillator (LO) path. The two phase-shift techniques, coupled with choice of heterodyne and direct-conversion architectures yields a total of four possible combinations for phased-array systems. In this section, we will explore the merits and demerits of the following architectures.

- Direct conversion with signal-path phase shifting- [7], [9]
- Heterodyne with signal-path phase shifting- [7], [12]
- Direct conversion with LO-path phase shifting
- Heterodyne with LO-path phase shifting - [5], [5]

The block diagram for a four-element direct-conversion phased-array receiver with signal-path phase-shifting is shown in Fig. 2.2. The front-end comprises of four low-noise amplifiers which drive four phase-shifters. Ideally, each phase-shifter provides a programmable phase-shift over the range of 0 to 2π radians. The signal from each phase-shifter is combined using a passive-combiner such as a Wilkinson hybrid. The output of the combiner drives a quadrature RF-mixer to generate the baseband I/Q signals, BB_I and BB_Q .

One of the key advantages of RF-power combining approach is that the signal path comprises of only a single-set I/Q RF-mixers; independent of the number of elements in the array. This architecture is *element-scalable* because LO distribution power does not increase as the number of the elements in the array increases. However, the architecture is not *frequency-scalable* because the LO frequency ($RF-LO_I$ and $RF-LO_Q$) is equal to the carrier frequency. Thus, for systems that desire to exploit the abundant bandwidth available at hundreds of the GHz, a 100+

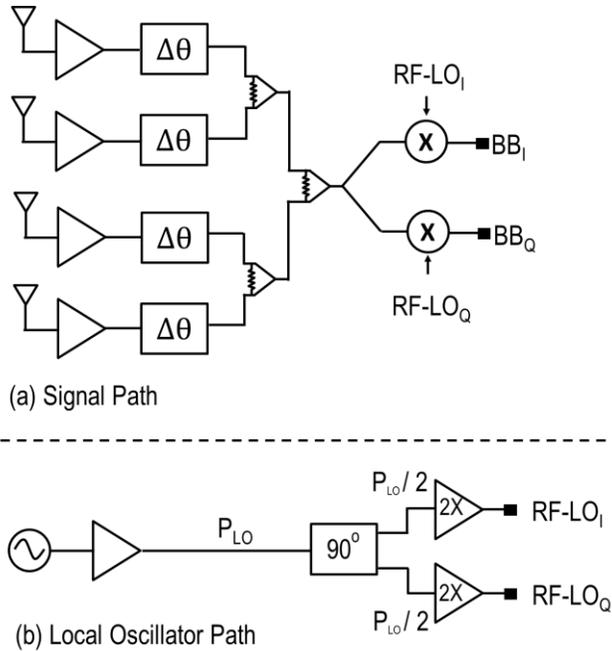


Fig. 2.2 (a) Signal path and (b) LO-path of a direct-conversion phased-array receiver with signal-path phase-shifting

GHz LO has to be generated. Moreover, for quadrature demodulation, this receiver architecture requires quadrature oscillator signals, a huge challenge at mm-wave frequencies.

The heterodyne phased-array receiver, shown in Fig. 2.3, addresses some of the concerns with LO generation and distribution. Similar to the direct-conversion phased-array, the front-end comprises of four low-noise amplifiers and phase-shifters. However, in the heterodyne receiver, the RF-mixer is not in quadrature phase, and therefore, I/Q generation of the high-frequency LO is not required. The mixer down-converts the received signal to an intermediate frequency (IF). After amplification in the IF-stage, the signal is finally down-converted to baseband through a quadrature IF-mixer to generate BB_I and BB_Q. The heterodyne architecture is more *frequency scalable* than direct-conversion, because RF-LO, IF-LO_I and IF-LO_Q can be selected to minimize receiver power.

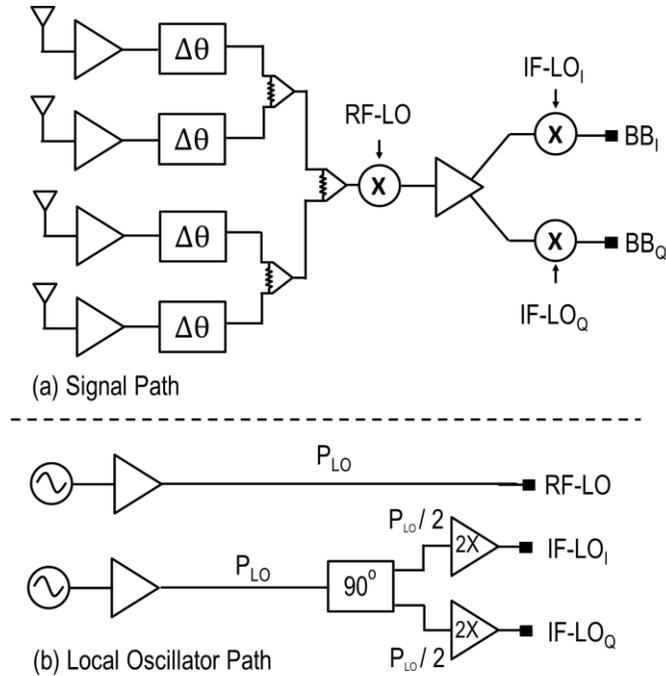


Fig. 2.3 (a) Signal-path and (b) LO-path of a heterodyne phased-array receiver with signal path phase-shifting

Both the direct and heterodyne receivers, described in Fig. 2.2 and Fig. 2.3, suffer from one big disadvantage; the phase-shifter is in the signal-path. For wideband systems, signal-path phase-shifters must exhibit low amplitude and phase imbalance across a wide range of frequencies. As an example, a phased-array receiver operating over a bandwidth of 50 to 70-GHz would require a signal-path phase-shifter with a fixed phase-shift over a bandwidth of 20 GHz, or a fBW of 30%. To the best of the author’s knowledge, a fully integrated programmable phase-shifter operating over such a wide frequency band has not been reported in prior-art.

Moreover, intuitively, sharp phase transitions are associated with high-Q circuits, while a flat phase response is indicative of a low-Q circuit. In addition, low-Q circuits result in higher insertion loss as compared to high-Q circuits. Thus, it is foreseeable that an integrated phase-shifter with a flat phase-shift over a wide frequency range would introduce large insertion loss in the signal-path.

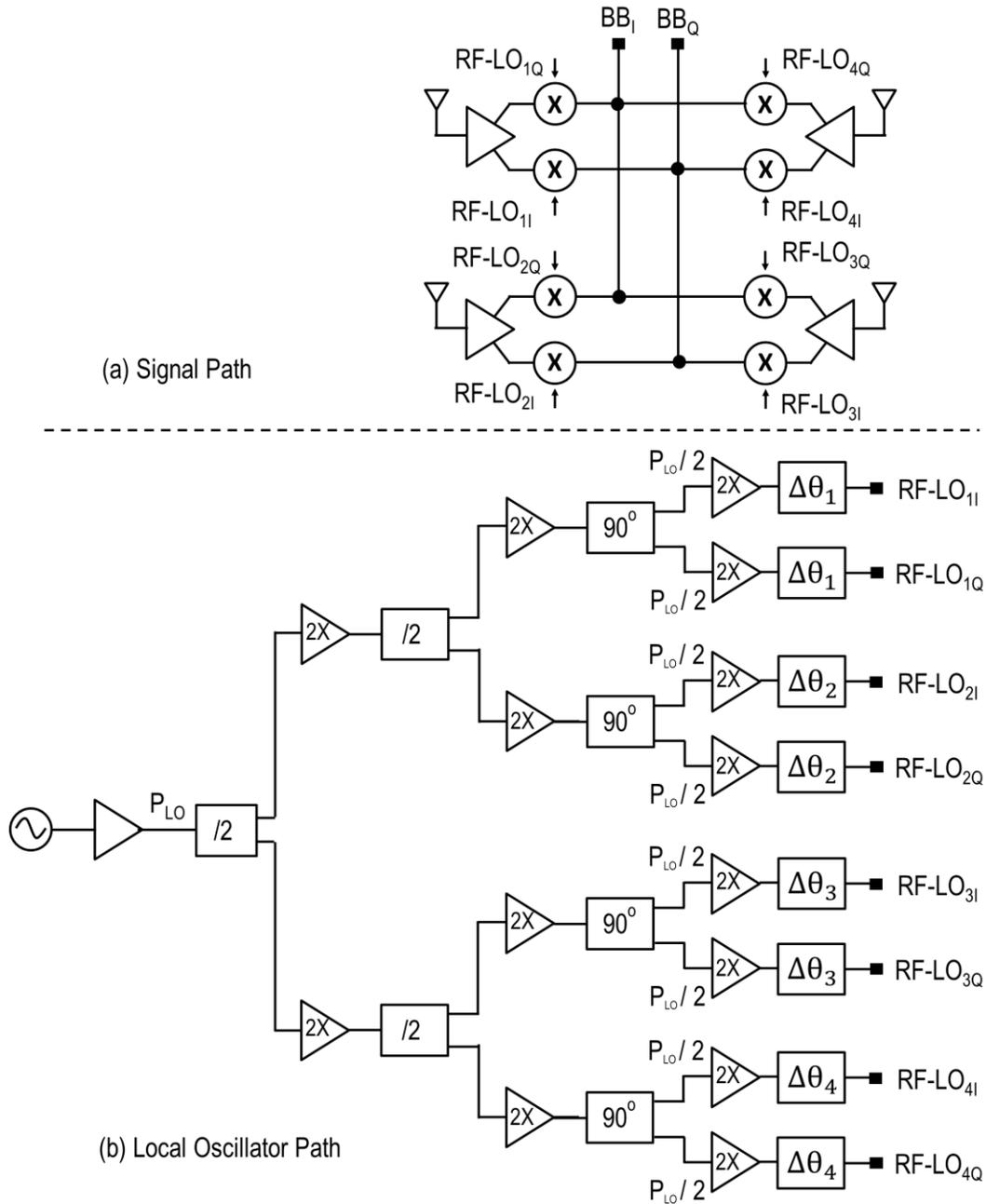


Fig. 2.4 (a) Signal-path and (b) LO-path of a direct-conversion phased-array receiver with LO-path phase-shifting

An alternate solution is to introduce a phase-shift onto the local-oscillator (LO) signal. The output phase of a mixer is a linear combination of the phase of the RF-input and LO, thus, any phase-shift introduced onto the LO-path is transferred on to the signal path during down-

conversion. A phase-shift on the local-oscillator signal instead of the signal path circumvents the bandwidth issue described above. An ideal LO is a single-tone and therefore, independent of the bandwidth of the signal being processed, the phase-accuracy has to be maintained only at a single-frequency. The block-diagram of a four-element direct-conversion receiver with LO-phase shifting is shown in Fig. 2.4. Each LNA drives a quadrature RF mixer. The relative phase-shift is introduced between the oscillator signals driving the RF-mixer. At the output of the RF mixer, the baseband signal can be combined in current-domain to obtain BB_I and BB_Q .

In the LO-phase shift approach, the insertion-loss due to the phase-shifter results in a reduction in the LO-voltage swing appearing at the mixer switches. However, if LO-buffers are used to drive the mixer switches then the sensitivity of mixer-gain to variation in LO amplitude will be low. Thus, the LO-phase shift approach is less sensitive to insertion loss than signal-path approaches. However, an N-element phased-array requires 2N mixers, as compared to a single mixer of the signal-path phase-shifting approach. Thus, the increase in immunity to signal-path phase imbalance comes at the expense of increased LO-power consumption.

Similar to the signal-path phase shifting receivers, LO-phase shifting can be implemented in both direct-conversion and heterodyne architectures. The block diagram of a four-element heterodyne receiver with LO-path phase shifting is shown in Fig. 2.5. The earlier discussion on *frequency-scalability* and *element-scalability* is equally applicable to direct and heterodyne LO-phase shifting based phased-array receivers.

In summary, the optimal architecture for integrated phased-array receivers is highly application dependent. For low-power RF systems direct conversion has evolved as the architecture of choice. However, for millimeter-wave N-element phased-array receivers, which

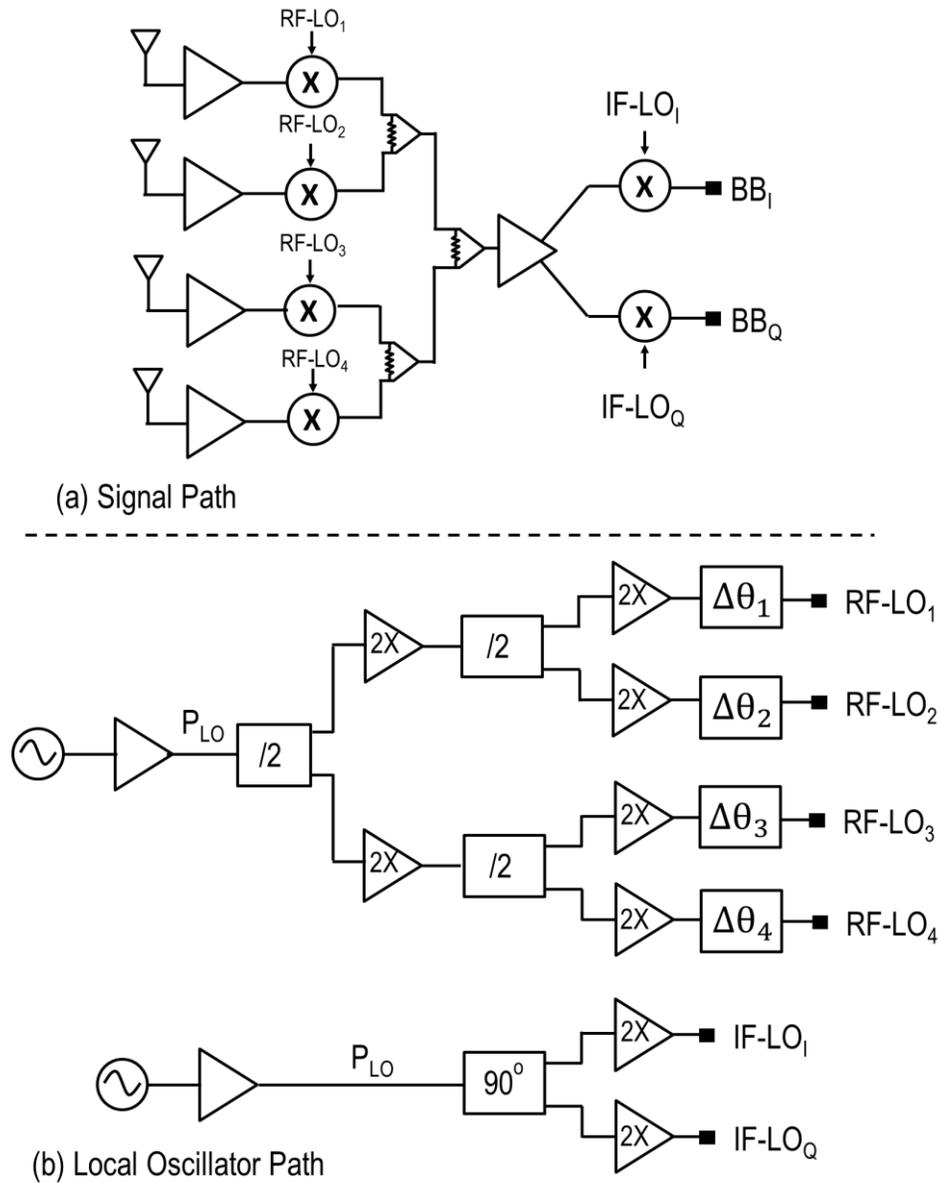


Fig. 2.5 (a) Signal-path and (b) LO-path of a heterodyne phased-array receiver with LO-path phase-shifting

comprise of N parallel signal paths (in addition to signal-combiners and phase-shifters), the choice is not so obvious.

In narrow-band systems the direct-conversion signal-path phase-shifting approach is most useful. Whereas in ultra-wideband systems, in which the in-band phase and amplitude

imbalance introduced by the signal-path phase shifter is unacceptable, the LO-phase shift architecture is preferable.

Even amongst LO-phase shift based phased-arrays, the choice of direct-conversion and heterodyne depends on the application. In systems requiring a high beam-steering granularity (for example, centimeter or meter-range chip-to-chip communication) the number of elements in the phased-array is low. In such cases, the direct-conversion approach might be more power optimal. However, systems targeting high beam resolution require a large number of array elements; therefore, the heterodyne approach could be more optimal. Similarly, for systems targeting a high carrier frequency, the heterodyne approach could be more power optimal.

2.4 Element scalability and impact on total power

In order to select the ‘optimal’ phased-array architecture for a given application, it is important to develop a strong analytical-model and understand the trade-offs between the different architectural techniques. Section 2.3 provides a qualitative comparison between the direct/heterodyne receivers with signal-path/LO-path phase-shifter. For ultra-wideband receivers it was concluded that LO-path phase shifting is preferable because it precludes the need for an ultra-wideband phase-shifter. Next, to study the impact of LO distribution power on *element-scalability*, in more quantitative terms, an analytic-model for the receiver power-budget as a function of the number of phased-array elements (N) is described next.

In the transmit/receive signal path, the nodes LO and LOB would need to drive the switching quad of a mixer in the signal path. The conversion-gain of the mixer is function of the LO voltage swing. Therefore, to make a fair comparison, it is assumed that the LO voltage-swing required in circuits (a) and (b) is identical. In addition, the buffers are assumed to be linear

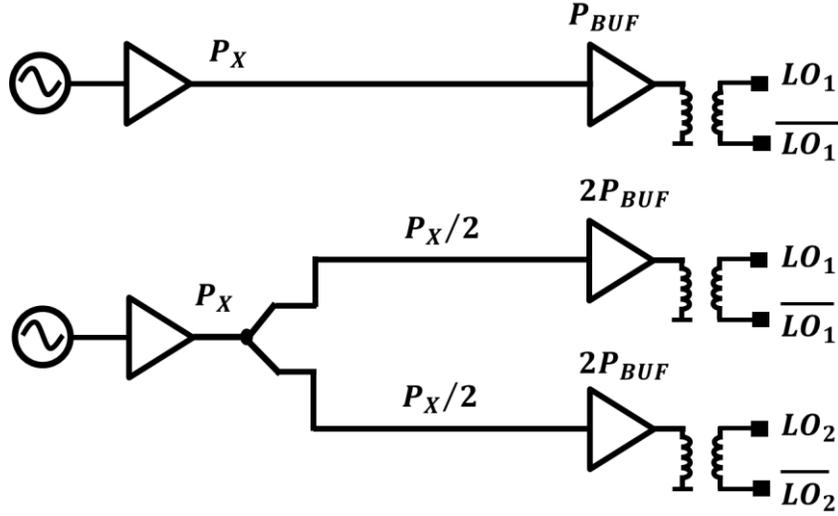


Fig. 2.6. LO distribution network for a phased-array with one and two elements.

amplifiers operating in the non-saturated state (the buffer is operating at power levels below the P_{sat})

As a first step, consider the two LO distribution networks shown in fig 2.6. The circuits in fig 2.6 (a) and (b) correspond to one possible implementation of a LO distribution network in which $N_1=1$ and $N_2=2$. As the number of elements scales by (N_2/N_1) , the total LO distribution power scales up by a factor of $(N_2/N_1)^2$. Two factors contribute to the ‘square’ in the scaling-factor. First, the number of buffers increases by a factor of (N_2/N_1) . Second, since the LO is distributed equally among the multiple paths, the input power to each buffer scales down by a factor of (N_2/N_1) ; therefore, to achieve the desired voltage swing at the buffer output, the voltage-gain has to be increased by (N_2/N_1) . As a first order approximation, to increase the voltage-gain, the DC-current of the buffer has to be increased by (N_2/N_1) .

In fig 2.6, the two-element receiver requires two buffers in place of one (in the $N_1 = 1$ case), resulting in an increase in a DC power by a factor of 2. Assuming an ideal power-splitter (with an insertion-loss of 0dB), the input power to the buffer in (b) is 3dB lower than the input

power in (a). Therefore, to double the power-gain the LO-buffers in (b) need to burn twice the DC-power in comparison to (a).

It is the N^2 scaling-factor that causes LO distribution power to dominate the total power consumption in phased-arrays with high element count (N). To illustrate this, consider the block-diagrams of direct-conversion and heterodyne phased-array receivers in Fig. 2.4(a) and Fig. 2.5(a), respectively.

In a direct conversion receiver, the signal path comprises of N low-noise amplifiers, N I/Q mixers (a total of $2N$ mixers). To drive the

$$P_{DC-SIGPATH} = NP_{LNA} + 2NP_{RFMX,60-BB} \quad (2-7)$$

$$P_{DC-LOPATH} = (2N)^2 P_{RFMLOBUF} \quad (2-8)$$

In an N -element heterodyne phased-array receiver, the mm-wave front-end comprises of N low-noise amplifiers and N single-phase RF mixers. The outputs of the N parallel mm-wave front-end paths are combined using an active/passive signal-combiner. The single power-combined output drives a single intermediate-frequency (IF) amplifier. The output of the IF-amplifier drives a set of I/Q IF mixers which down-converts the signal to baseband. In comparison with the signal-path power of a direct-conversion receiver (2-7), the heterodyne architecture incurs the additional power consumption of the IF-amplifier in the signal path.

$$P_{H-SIGPATH} = NP_{LNA} + NP_{RFMX} + P_{IFA} + 2P_{IFMX} \quad (2-9)$$

Table 2.1 Estimated power consumption of individual block blocks

	Power	Comments
P_{LNA}	8mA	Single-stage of a common-source amplifier
$P_{RF-LOBUF, 30}$	0.6mA	Cascode amplifier
$P_{RF-LOBUF, 60}$	1.8mA	Cascode amplifier
$P_{RFMX,60-BB}$	1mA	Single-balanced active mixer with a baseband load
$P_{RFMX,60-30}$	3mA	Single-balanced active mixer with tuned load
$P_{RFMX,30-BB}$	1mA	Single-balanced active mixer with a baseband load
P_{IFA}	10mA	Cascode amplifier

However, it is important to note that the power-penalty of the IF-stage is a fixed quantity and independent of the number of elements in the array. In other words, as the number of elements in the array increases, signal-path power per unit element ($P_{H-SIGPATH}/N$)

Next, consider the LO distribution path comprising of a total of $(N+2) - N$ buffers to drive the N single-phase mm-wave mixers and 2 buffers to drive the I/Q IF mixer and 2. In contrast to the direct-conversion receiver, the only N (rather than $2N$) buffers are required at the mm-wave front-end. Moreover, the N buffers of the heterodyne will operate at lower-frequency than the $2N$ buffers of the direct-conversion receiver.

$$P_{H-LOPATH} = N^2 P_{RFMLOBUF} + 2P_{IF-LOBUF} \quad (2-10)$$

Finally, from (2-7)-(2-10), it can be shown that the total power consumption of the direct-conversion (P_{DC}) and heterodyne (P_H) is given by,

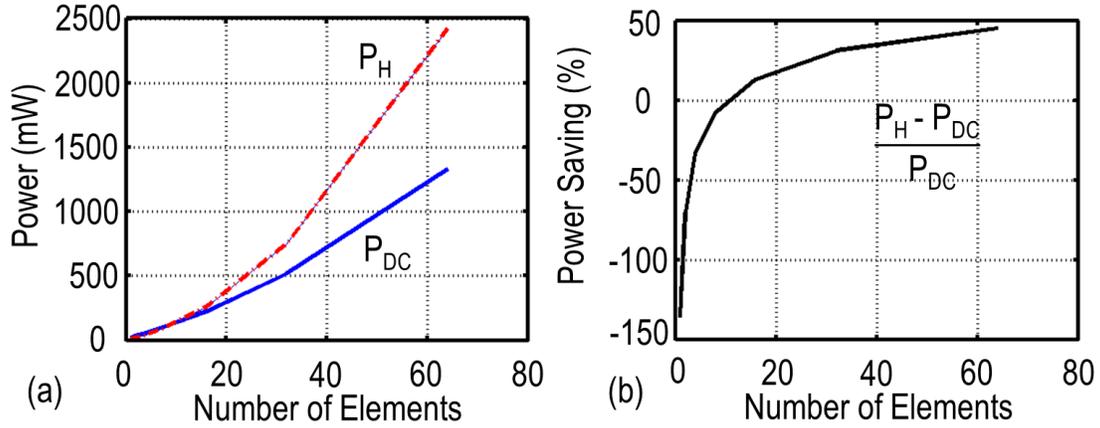


Fig. 2.7 (a) Total power consumption of a direct-conversion (blue) and heterodyne (red) phased-array receiver (b) Power saving in a heterodyne receiver as a function of a number of array elements

$$P_H = NP_{LNA} + N^2 P_{RF-LOBUF, 30} + NP_{RFMX} + P_{IFA} + 2P_{IF-LOBUF} + 2P_{IFMX} \quad (2-11)$$

$$P_{DC} = NP_{LNA} + (2N)^2 P_{RF-LOBUF, 60} + 2NP_{RFMX,60-BB} \quad (2-12)$$

In summary, (2-7) and (2-9) suggests the signal-path of a heterodyne receiver consumes more power than the equivalent direct-conversion receiver. In contrast, from (2-8) and (2-9), the LO-path of the direct-conversion system is higher than the heterodyne. For low-element phased-array systems the signal-path power dominates, however, as the number of elements increase the LO-path begins to dominate. The *point of inflection* for the number of elements in an array where a heterodyne receiver becomes more power efficient than a direct-conversion receiver, is the point where the power consumption of the LO-buffers in a direct-conversion receiver becomes significantly larger than the power consumes by the IF-amplifier in the equivalent heterodyne. As an example, consider some typical power numbers for the circuit blocks in the receive-chain shown in Table 2.1. Using the values in Table 2.1, P_H and P_{DC} are plotted as plot as a function of N in Fig. 2.7. For $N < 15$, the direct-conversion is more optimal. However, for $N > 15$, the

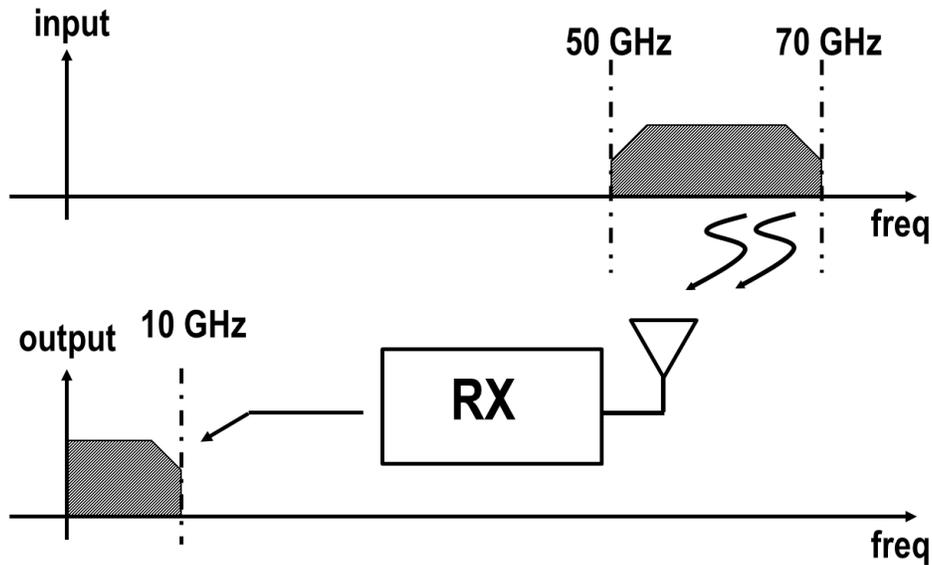


Fig. 2.8. Millimeter-wave spectrum targeted in this receiver

power-saving achieved by selecting a heterodyne architecture compensates for the additional power of the IF-amplifier.

2.5 Ultra-wideband receiver architecture

This dissertation is part of a research effort to design a fully-integrated wideband phased-array receiver with very fine beam-steering resolution. Fine beam-steering resolution, in turn, necessitates a high element count. From the preceding analysis it can be concluded that,

- For ultra-wideband receivers, the phase-shifter is easier to implement in the LO-phase shifting approach as compared to signal-path phase-shifters.
- The heterodyne architecture is more element-scalable than the direct-conversion approach for systems in which the LO distribution power is non-negligible.

Therefore, the wideband mm-wave receiver architecture proposed in this work will be a part of a heterodyne receiver with LO-phase shifting. The receiver has been designed for a channel

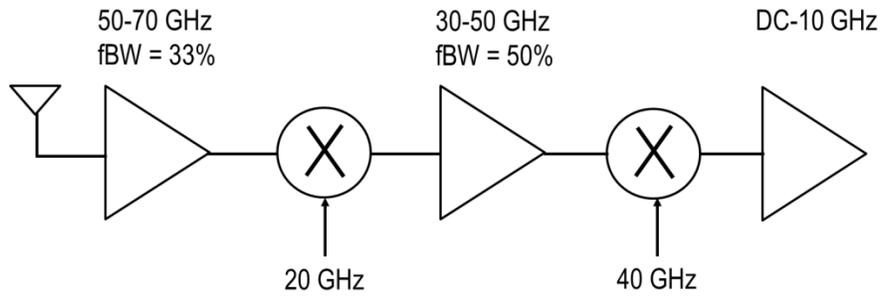


Fig. 2.9. Frequency plan for wideband heterodyne architecture

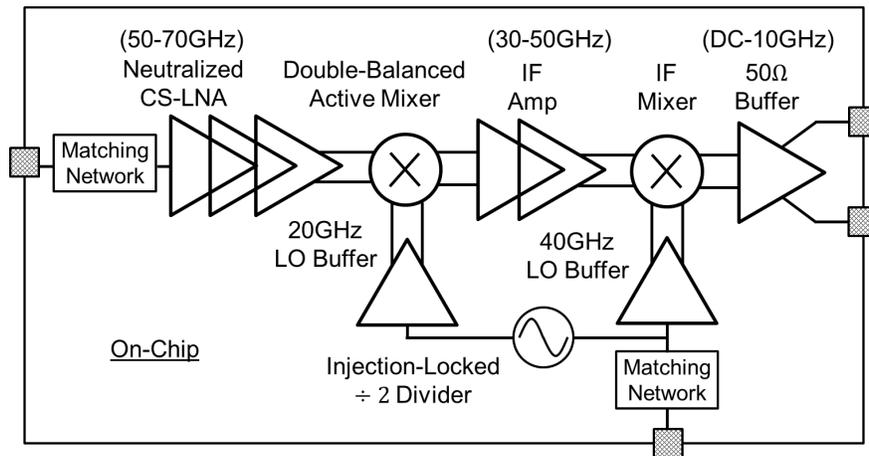


Fig. 2.10. Block diagram of the ultra-wideband mm-wave receiver implemented in a 40nm CMOS

bandwidth of 20 GHz (pass-band) or 10 GHz (baseband). This bandwidth is almost 10 times larger than the channel-bandwidth of transceivers designed for the 60 GHz standards. The target millimeter-wave bandwidth was selected to be 50-to-70 GHz shown in Fig. 2.8. An upper frequency limit of 70GHz was selected to ensure the receiver IC could be fully characterized using the measurement facilities available at the University of Washington.

While the heterodyne architecture allows the designer to scale down the local-oscillator frequency for the RF-mixer, for an ultra-wideband system it introduces new challenges in the signal path design. To illustrate this, consider the frequency-plan shown in Fig. 2.9. The input

signal has a bandwidth of 20 GHz and is modulated onto a carrier frequency of 60 GHz. At the front-end, the required fractional-bandwidth of the circuit is 33%. Assume that the first down-converter mixes the input with a 20 GHz oscillator signal. The input is brought down to intermediate-frequency (IF) of 40 GHz and signal amplification at IF requires circuits with a fractional bandwidth of 50%! As a signal progresses from the mm-wave front-end towards the baseband, fractional bandwidth becomes larger even though the absolute bandwidth remains constant.

The focus of this dissertation is to propose area and power optimal solutions for high fractional-bandwidth receiver designs. The techniques proposed in this work have been used to design the ultra-wideband mm-wave receiver shown in Fig. 2.10 (realized in a 40nm CMOS process), based on the frequency plan described in Fig. 2.9. To achieve high fractional-bandwidths, three wideband circuit techniques: transformer-feedback, bandpass distributed amplifiers, and gain-equalized transformers, have been explored, and are discussed in subsequent chapters.

- [1] H. T. Friss, and C. B. Feldman, "A multiple unit steerable antenna for short-wave reception," in *Proceedings of the IRE*, vol.25, no.7, pp.841-917, 1937.
- [2] D. Parker, and D. C. Zimmermann, "Phased arrays-part 1: theory and architectures," in *IEEE Transactions on Microwave Theory and Techniques*, vol.50, no.3, pp.678-687, 2002.
- [3] G. E. Moore, "Cramming more components onto integrated circuits," in *Electronics*, vol.38, no.8, pp.114-117, 1965.
- [4] X. Guan, H. Hashemi, and A. Hajimiri, "A fully integrated 24-GHz eight element phased-array receiver in silicon" in *IEEE Journal of Solid-State Circuits*, vol.39, no.12, pp.2311-2320, 2004.
- [5] H. Hashemi, X. Guan, A. Komijani, and A. Hajimiri, "A 24-GHz SiGe phased-array receiver- LO phase-shifting approach," in *IEEE Transactions on Microwave Theory and Techniques*, vol.53, no.2, pp.614-626, 2005.
- [6] A. Babakhani, X. Guan, A. Komijani, A. Natarajan, A. Hajimiri, "A 77-GHz phased-array transceiver with on-chip antennas in silicon: receiver and antennas," in *IEEE Journal of Solid-State Circuits*, vol.41, no.12, pp.2795-2806, 2006.
- [7] M. Tabesh, J. Chen, C. Marcu, L. Kong, S. Kang, A. Niknejad, and E. Alon, "A 65nm CMOS 4-element sub-34mW/element 60GHz phased-array transceiver," in *IEEE Journal of Solid-State Circuits*, vol.46, no.12, pp.3018-3032, 2011.
- [8] A. Natarajan, et.al., "A fully-integrated 16-element phased-array receiver in SiGe BiCMOS for 60-GHz communications," in *IEEE Journal of Solid-State Circuits*, vol.46, no.5, pp. 1059-1075, 2011.

- [9] S. Shahramian, Y. Baeyens, N. Kaneda, Y.-K. Chen, “A 70-100GHz direct-conversion transmitter and receiver phased array chipset demonstrating 10 Gb/s wireless link,” in *IEEE Journal of Solid-State Circuits*, vol.48, no.5, pp.1113-1125, 2013.
- [10] T.Chu, J. Roderick, and H. Hashemi, “An integrated ultra-wideband timed array receiver in 0.13um CMOS using a path-sharing true time delay architecture”, in *IEEE Journal of Solid-State Circuits*, vol.42, no.12, pp.2834-2850, 2007.
- [11] T. Chu, H. Hashemi, “True-time-delay-based multi-beam arrays,” in *IEEE Transactions on Microwave Theory and Techniques*, vol.61, no.8, pp.3072-3082, 2013.
- [12] E. Cohen, C. Jakobsen, S. Ravid, and D. Ritter, “A thirty two element phased-array transceiver at 60GHz with RF-IF conversion block in 90nm flip chip CMOS process,” in *IEEE Radio Frequency Integrated Circuit Symp. (RFIC) Dig.*, pp.457-460, 2010.

3 TRANSFORMER FEEDBACK

Feedback is one of the most commonly employed techniques for designing wideband circuits. In low-frequency or baseband analog circuits it is relatively easy to design circuits with a high open-loop gain. Feedback provides a mechanism to trade-off a high open-loop gain which is process-dependent, narrow-band with a low closed-loop gain which is process-invariant and wideband. In older technology nodes, the low power gain associated with CMOS devices at radio frequencies, precluded the use of feedback. However, scaling of the minimum device dimensions in advanced technology nodes has resulted in the device unity power-gain frequency (F_{\max}) to exceed several hundred gigahertz. This allows the possibility of exploiting resistive or reactive feedback to trade off the extra open-loop gain for a wider bandwidth. While resistive feedback can provide a small form-factor solution, the noise introduced by the feedback resistor can be prohibitively large for applications which require high noise sensitivity. In this chapter we analyze reactive feedback techniques, using integrated transformers, for high fractional-bandwidth circuit design.

Feedback using fully-integrated transformers, in which the magnetically coupled windings provide a path for current-sense current-feedback, has received considerable interest [8]-[9]. For the three-terminal MOS device, there are three fundamental transformer-feedback topologies [11]: drain-to-source, drain-to-gate, and source-to-gate. The first two, drain-to-source and drain-to-gate, have been applied in single-ended amplifiers to neutralize the C_{GD} device capacitance and improve reverse-isolation over a wide bandwidth. The source-to-gate transformer feedback (SGTxFB) topology is more suited for wideband matching network design

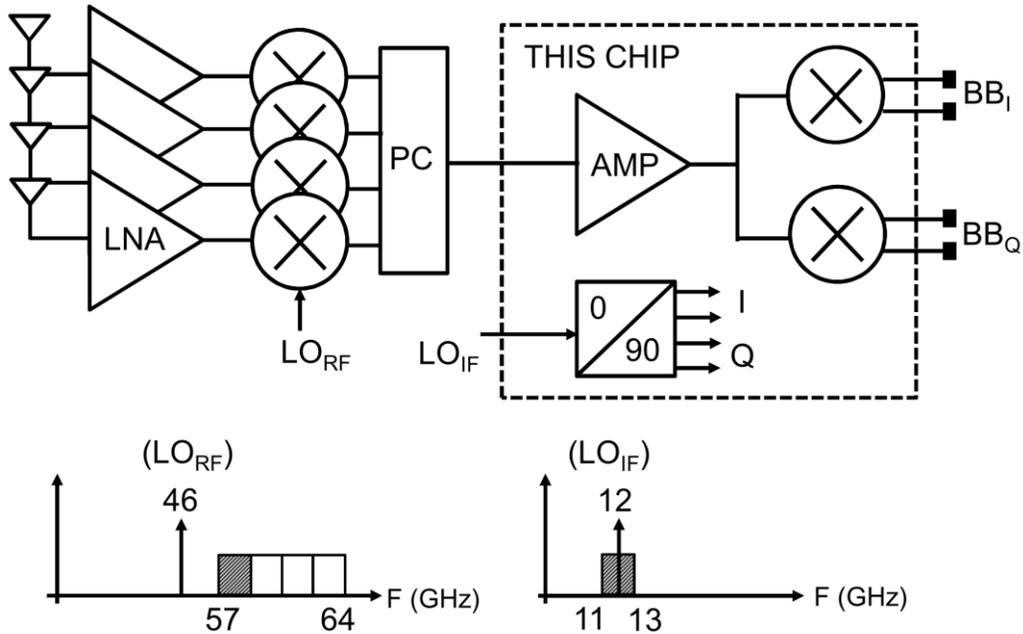


Fig. 3.1. Heterodyne phased-array receiver: low-noise amplifier (LNA), power-combiner (PC) followed by 11-to-13 GHz IF stage comprised of an IF amplifier, quadrature down-conversion mixer, and lumped-element Lange coupler for I/Q generation.

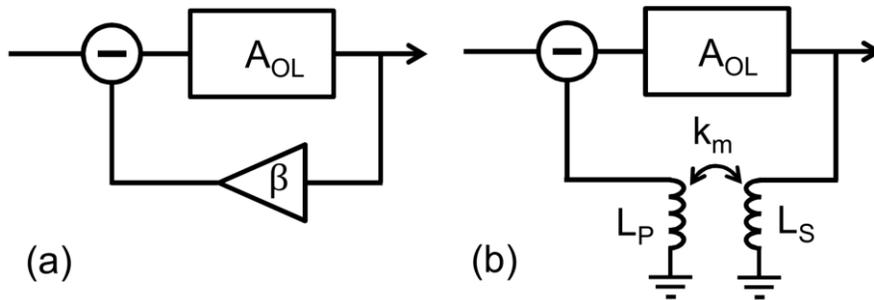


Fig. 3.2. (a) Generic feedback circuits (b) transformer-feedback circuit.

and has been applied over a wide range of operating frequencies, ranging from UWB [8] to W-Band [9].

However, currently available analytic models for SGTxFB-based matching-network design provide little intuition for optimization. In this chapter, a generic and systematic approach to the design of SGTxFB amplifiers is presented. The input admittance is modeled as a function

of transformer and transistor parameters. The models are then used to assess the impact of the circuit parameters on the bandwidth, noise figure, and gain. The design and measured results from a prototype test-chip employing multiple stages of SGTxFB is described. The test-chip was designed to operate as the intermediated-frequency stage of the heterodyne 60-GHz receiver shown in Fig. 3.1. The challenges involved in designing a multi-stage SGTxFB down-converter are described.

This chapter is organized as follows. First, an ideal current- feedback amplifier model is introduced, and related to a simplified first-order SGTxFB stage. Next, guidelines for matching network design and accurate input admittance models for SGTxFB amplifiers are discussed in Section 3.2. This is followed with a derivation of analytic expressions for the noise and gain of SGTxFB amplifiers in Section 3.3. The design of an IF-stage, operating over a frequency range of 11-to-13 GHz, is described in Section 3.4. Measured results from a prototype chip implemented in a 40nm CMOS process are presented in Section 3.5.

3.1 Current Feedback

The generic model of a feedback system is shown in Fig. 3.2(a). The forward path consists of a high-gain amplifier, A_{OL} . A fraction of the output voltage (or current) of A_{OL} is sampled by the feedback circuit β and fed back to the input. In the circuit shown in fig. 2(b) the feedback circuit β is a transformer, hence the name *transformer-feedback*. The secondary windings of the transformer samples the output current of A_{OL} and the current induced in the primary is fed back to the input. In the specific case of source-to-gate transformer-feedback, the output current is sampled at the source and induced current is fed back to the gate of the MOSFET.

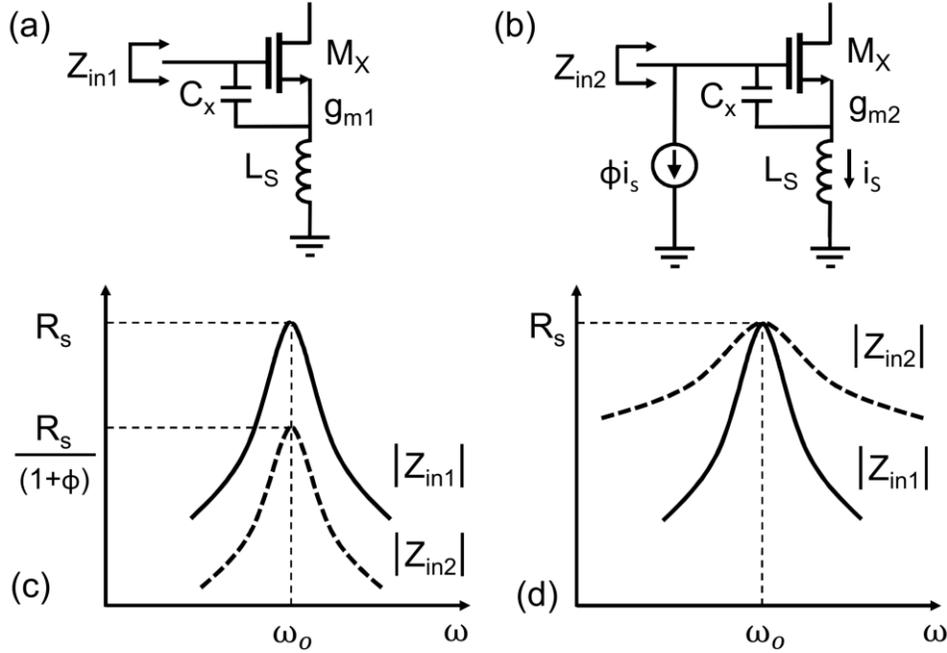


Fig. 3.3. (a) Common-source (CS) amplifier with inductor degeneration (b) Inductively degenerated amplifier with current feedback (c) Input impedance (Z_{in}) with $g_{m1} = g_{m2}$ (d) Z_{in} with $g_{m1} = g_{m2}(1 + \phi)$

As a first order approximation, a SGTxFB amplifier can be modeled as the current-feedback amplifier (CFA) shown in Fig. 3.2 (b). In this ideal current-feedback model, the feedback current induced in the primary (L_p) is included; however, the feed-forward current induced in secondary (L_s) is ignored. To study the impact of current-feedback on bandwidth, consider the inductor-degenerated common-source amplifier without and with current-feedback shown in Fig. 3.3 (a) and Fig. 3.3 (b), respectively. Starting with the circuit in Fig. 3.3(a), it is straightforward to prove that the input impedance Z_{in1} is described by,

$$Z_{in1} = \frac{L_s g_{m1}}{C_x} + j\omega_o L_s + \frac{1}{j\omega_o C_x} \quad (3-1)$$

Accordingly, the circuit appears to be a series-RLC network with $Re\{Z_{in1}\} = (g_{m1} L_s) / C_x$, resonant frequency $\omega_o = 1 / \sqrt{L_s C_x}$ and a quality factor Q_1 , where

$$Q_1 = \frac{\omega_o L_s}{Re(Z_{in1})} = \frac{1}{g_{m1}} \sqrt{\frac{C_X}{L_s}} \quad (3-2)$$

Next, consider the CFA in Fig. 3.3 (b), where ϕ describes the ratio of the source current i_s to the current fed back to the gate. The CFA and inductor-degenerated amplifier have identical L_s , C_X and M_X , however, the devices are biased differently and hence have different transconductance.

The input impedance Z_{in2} is given by,

$$Z_{in2} = \frac{1}{sC_X} + \frac{\frac{g_{m2}}{sC_X} + 1}{1 + \phi \left(\frac{g_{m2}}{sC_X} + 1 \right)} sL_s - \frac{\frac{g_{m2}}{sC_X} + 1}{1 + \phi \left(\frac{g_{m2}}{sC_X} + 1 \right)} \frac{\phi}{sC_X} \quad (3-3)$$

At $s = j\omega_o$, $g_{m2}/(sC_X) = 1/Q$. Accordingly, for circuits with quality factor $Q > 4$, the complex term in the denominator $|1 - jg_{m2}/\omega_o C_X|$ is approximately equal to 1; reducing (3-3) to,

$$Z_{in2} = \frac{1}{j\omega_o C_X} \frac{1}{(1 + \phi)} + \frac{g_{m2} L_s}{C_X (1 + \phi)} + \frac{j\omega_o L_s}{1 + \phi} \quad (3-4)$$

Thus, similar to (3-1), the input impedance of the CFA appears as a series RLC resonant circuit with a resonance frequency of ω_o . In addition, it can be observed that the CFA has a input impedance $Re\{Z_{in2}\} = g_{m2} L_s / (C_X (1 + \phi))$ and quality factor Q_2 , where

$$Q_2 = \frac{1}{g_{m2}} \sqrt{\frac{C_X}{L_s}} \quad (3-5)$$

Several important observations can be made based on the above result. Assume the circuit in Fig. 3.3 (a) is designed to match with an antenna with resistance R_S i.e. $R_S = Re(Z_{in1})$.

The first case to consider is wherein the transconductance of M_X is identical in both

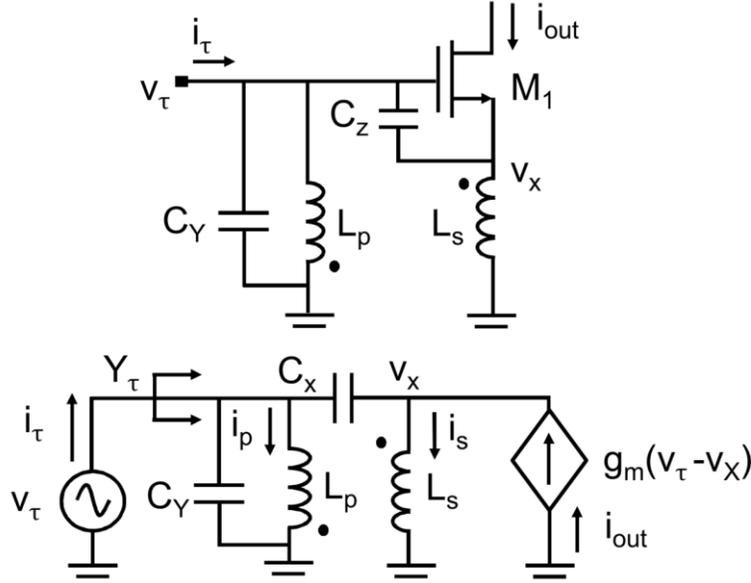


Fig. 3.4. Schematic and small-signal model of an SGTx FB amplifier

circuits ($g_{m1} = g_{m2}$). The resulting input impedance as a function of frequency is plotted in Fig. 3.3 (c). From (3-2) and (3-5), one can observe that with equal transconductance both the amplifiers display an identical quality-factor ($Q_1 = Q_2$). However, $Re(Z_{in2}) = R_S/(1 + \phi)$, as a result the CFA is not power-matched to the antenna

To correct the antenna mismatch, without altering the value of passive components L_S and C_X , transistor M_X in the current-feedback amplifier is biased such that $g_{m2} = g_{m1}(1 + \phi)$. From (3-4) one observes that the new bias-condition ensures $Re(Z_{in2}) = Re(Z_{in1})$. Furthermore, since the quality factor of the matching network is inversely proportional to the transconductance, the g_m -boost results in a $(1 + \phi)$ reduction in the quality factor. Therefore, as shown in Fig. 3.3 (d), compared to the inductor-degenerated common-source amplifier, the current-feedback amplifier effectively achieves a $(1 + \phi)$ higher matching bandwidth at the expense of $(1 + \phi)^2$ higher current (assuming square-law devices).

The desire to exploit current-feedback to achieve a wide input matching bandwidth motivates application of source-gate transformer feedback (SGTxFB). The SGTxFB amplifier in Fig. 3.4 contains two feedback loops. In the first feedback-loop, at frequency ω_o , L_s senses the (output) current flowing through M_1 and converts it to a voltage which controls the (input) V_{GS} of the device. In the second loop, current through L_s is fed back to the input via the anti-phase mutual magnetic coupling between L_p and L_s . In effect, the transformer formed by L_s and L_p provides current-sense current-feedback.

To design a matching network using SGTxFB, an accurate model of the input impedance of the amplifier as a function of transformer and transistor parameters is derived in the next section.

3.2 Source-to-Gate Transformer Feedback Based Matching Networks

Although receiver front-end circuitry realized with SGTxFB amplifiers has been reported in recent literature [8]-[9], insightful and compact analytic expressions that assist matching network design and model the noise performance are yet to be presented. This is primarily due to relatively open design space, comprising multiple variables such as device transconductance (g_m), self-inductance (L_p, L_s) and mutual inductance (M). To simplify the calculations, it is common to assume perfect magnetic coupling; a coupling-coefficient ($k = M/\sqrt{L_p L_s}$) of one. However, for a large turns-ratio ($n = \sqrt{L_p/L_s}$) a coupling-coefficient close to unity is difficult to achieve. To investigate the trade-offs involved in the design of a SGTxFB, an input admittance model $Y_\tau = f\{\omega_o, n, k, g_m, L_s\}$ is computed next.

Input Admittance

The small-signal model of the SGTxFB amplifier is presented in Fig. 3.4. Transistor M_1 is assumed to have zero output-conductance. In the small-signal model, C_X is the parallel combination of the gate-to-source capacitance (C_{GS}) of M_1 and an extra capacitance C_Z . For the input admittance analysis, capacitance C_Y which appears in parallel with the ideal voltage source, v_τ , is ignored. The transformer model of [12] is adopted and the body effect is neglected. Initially, to simplify the mathematical analysis, inductors L_p and L_s are assumed to be ideal. However, after deriving the model, a technique to include inductor non-idealities will also be described. Applying KCL to the circuit in Fig. 3.4 yields,

$$v_\tau = i_p s L_p - i_s s M \quad (3-6)$$

$$v_x = -i_p s M + i_s s L_s \quad (3-7)$$

$$i_\tau = i_p + (v_\tau - v_x) s C_X \quad (3-8)$$

$$i_s = (v_\tau - v_x)(g_m + s C_X) \quad (3-9)$$

$$Y_\tau(s) = \frac{1}{s L_p (1 - k^2)} + s C_X - \left\{ \frac{nk}{s L_p (1 - k^2)} - s C_X \right\} \frac{\left\{ \frac{\frac{k}{n}}{s L_s (1 - k^2)} - (g_m + s C_X) \right\}}{\left\{ \frac{1}{s L_s (1 - k^2)} + (g_m + s C_X) \right\}} \quad (3-10)$$

Solving (3-6)-(3-9), the input admittance $Y_\tau = i_\tau / v_\tau$ can be shown to be given by (3-10).

To verify (3-10) a test circuit (TC1) with $\{\omega_o, n, k, g_m, L_s\} = \{12\text{GHz}, 1.16, 0.5, 50\text{mS}, 800\text{pH}\}$

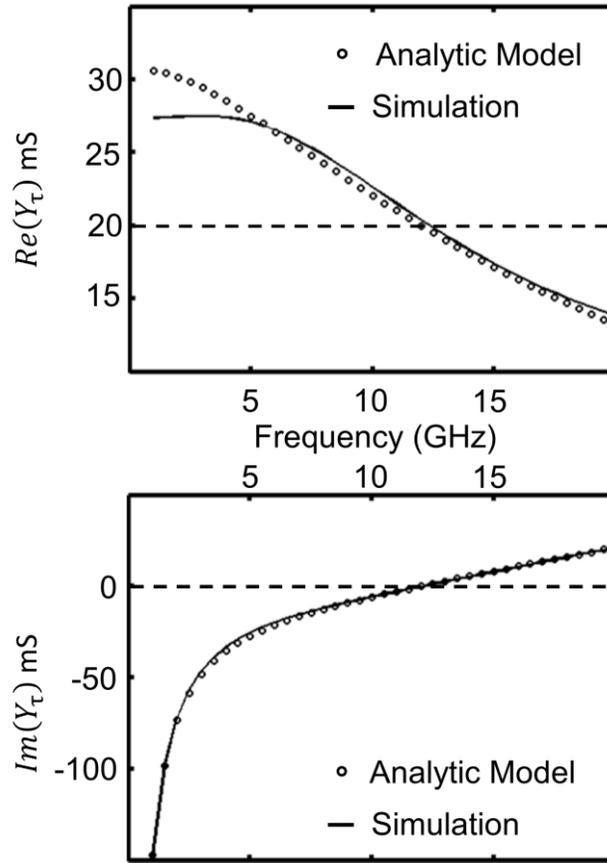


Fig. 3.5 Real and imaginary admittance: model versus circuit simulation

was designed. A comparison of $Re(Y_t)$ and $Im(Y_t)$ obtained from circuit simulation and from (3-10) are plotted as a function of frequency in Fig. 3.5.

However, while (3-10) is an accurate and exact solution to the KCL equations, it fails to provide insight on how to select the component values to achieve the target admittance.

To simplify (3-10), the design space must be constrained as to reflect normal operating conditions. Towards this goal, a *resonance-condition* is enforced: At frequency ω_o , C_X resonates with leakage inductance, $L_s(1 - k^2)$. As mentioned earlier, C_X includes a shunt capacitance C_Z which can be appropriately selected to ensure C_X satisfies the resonance condition mathematically described in (3-11). Using (3-11), the simplifications given in (3-12)-(3-14) can

be made. Finally, using (3-12)-(3-14), a simpler expression for Y_τ is derived in (3-15). Equation (3-15) has a clear physical interpretation: the input impedance of a SGTxFB amplifier appears as a parallel R-L circuit, where R and L are functions of $\{\omega_o, n, k, g_m, L_s\}$.

$$s^2 L_s (1 - k^2) C_X = - \omega_o^2 L_s (1 - k^2) C_X = -1 \quad (3-11)$$

$$\left\{ \frac{nk}{sL_p(1-k^2)} - sC_X \right\} = \frac{n(n+k)}{j\omega_o L_p(1-k^2)} \quad (3-12)$$

$$\left\{ \frac{\frac{k}{n}}{sL_s(1-k^2)} - (g_m + sC_X) \right\} = \frac{\left(1 + \frac{k}{n}\right)}{j\omega_o L_s(1-k^2)} - g_m \quad (3-13)$$

$$\left\{ \frac{1}{sL_s(1-k^2)} + (g_m + sC_X) \right\} = g_m \quad (3-14)$$

$$\begin{aligned} Y_\tau(s = j\omega_o) &= \frac{1 + n(n+k)}{j\omega_o L_p(1-k^2)} + j\omega_o C_X + \frac{(n+k)}{\omega_o L_p(1-k^2)} \frac{(n+k)}{\omega_o L_s(1-k^2)} \frac{1}{g_m} \\ &= \frac{1 + nk}{j\omega_o L_p(1-k^2)} + \left(\frac{C_X}{g_m L_s} \right) \frac{\left(1 + \frac{k}{n}\right)^2}{(1-k^2)} \end{aligned} \quad (3-15)$$

3.3 Matching Network Design

Since input impedance of the SGTxFB amplifier is a function of the transformer parameters, applying transformer-feedback to match the amplifier with its preceding driver-stage is possible. The driver-stage could be an on-chip pre-amplifier, mixer, off-chip transmission line driver, or antenna. To maximize the power gain of the driver-stage with an output conductance G_s , a matching network is designed to ensure $Re(Y_\tau) = G_s$. In order to match the SGTxFB amplifier to a driver with purely-real admittance, all imaginary terms in (9b) must be eliminated.

Manipulating the design variables L_s and C_x to achieve cancellation of imaginary terms is not possible due to the already established resonance condition (3-11). Therefore, an additional capacitance C_Y must be added in parallel to L_p to achieve $Im(Y_\tau) = 0$. From (9b) it can be shown that,

$$C_Y = \frac{1 + nk}{\omega_o^2 L_p (1 - k^2)} \quad (3-16)$$

After the addition of C_Y , the input admittance of the circuit is purely real and given by,

$$Re(Y_\tau) = \frac{(n + k)}{\omega_o L_p (1 - k^2)} \frac{(n + k)}{\omega_o L_s (1 - k^2)} \frac{1}{g_m} \quad (3-17)$$

Further intuition regarding the impact of feedback for bandwidth extension can be obtained by using (3-11) to reformulate (3-17), resulting in (3-18) (derivation in Appendix. II),

$$R_\tau(k, n) = \frac{1}{Y_\tau} = \left(\frac{g_m L_s}{C_x} \right) \frac{(1 - k^2)}{\left(1 + \frac{k}{n}\right)^2} \quad (3-18)$$

In (3-18), one should notice the expression derived is similar to the input impedance of a common-source amplifier with inductor degeneration. For the circuit in Fig. 3.3(a), from (3-1), the real component of the input impedance is given by $(g_m L_s / C_x) = R_\tau (k = 0)$. Thus, verifying that the narrow-band inductor degeneration is a special-case of transformer feedback. In the SGTxFB amplifier, the magnetic-coupling between the windings of the transformer reduces the shunt input impedance R_τ and provides a wideband match.

While the aforementioned results have been derived using an ideal model inductor L_p and L_s , the results can be extended to include the effects of a finite Q-factor and self-resonance frequency. The loss in L_p can be modeled by a shunt resistance $R_{pL} = Q_p \omega L_p$, which appears in parallel with input resistance R_τ derived in (3-18). The loss in L_s can be modeled as a series resistance $R_{sL} = \omega L_s / Q_s$. As a first order approximation, R_{sL} can be absorbed into the amplifier transconductance by defining an effective $g_m' = g_m / (1 + g_m R_{sL})$. Finally, the parasitic capacitance associated with L_p and the windings between L_p and L_s can also be absorbed in capacitances C_Y and C_X , respectively.

Quality Factor

A key metric for wideband matching networks is the Q-factor. For a parallel RLC circuit, the quality factor is $(R/\omega_o L)$ or $(\omega_o C R)$. From (3-15), the Q of the matching network can be proven to be,

$$Q(s = j\omega_o) = g_m \sqrt{L_s(1 - k^2)/C_X} \frac{1 + nk}{(n + k)^2} \quad (3-19)$$

As described earlier, the Q-factor and $Re\{Y_\tau\}$ are functions of $\{\omega_o, n, k, g_m, L_s\}$. However, in order to design for $G_s = Y_\tau(j\omega_o)$, only three among $\{n, k, g_m, L_s\}$ can be uniquely specified. The g_m is primarily limited by the current budget of the target application and an upper limit on L_s is placed by the inductor's self-resonance frequency. As a result, to obtain an optimal power-match, the transformer parameters $\{n, k\}$ cannot be selected independently. Using (3-17) as a starting point, the following interdependence between n and k can be shown,

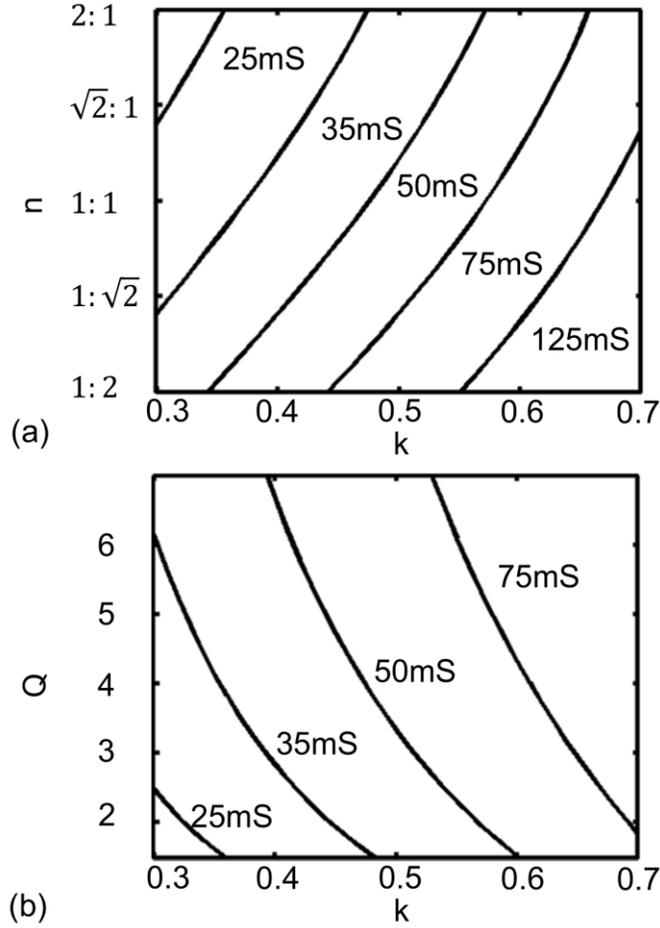


Fig. 3.6. Design space for matching TC1 to a 50-Ω source resistance with $g_m = (25mS, 35mS, 50mS, 75mS, 125mS)$ and $0.3 < k < 0.7$ (a) transformer turns-ratio (b) matching network Q factor

$$n = \frac{k}{\omega_o L_s (1 - k^2) \sqrt{Re(Y_\tau) g_m} - 1} \quad (3-20)$$

Defining the ‘design-space’ of the circuit as all sets of $\{n, k, g_m, L_s\}$ which satisfy (3-20), the design space for the test circuit TC1, with $L_s = 800pH$ at $\omega_o = 12GHz$, is plotted in Fig. 3.6. For values of g_m ranging from 25mS to 125mS, Fig. 3.6(a) plots n as a function of k for a SGTx FB circuit designed to match a 50Ω driving source. Two important observations can be made from this graph. First, for a fixed turns-ratio (n), an amplifier with higher current

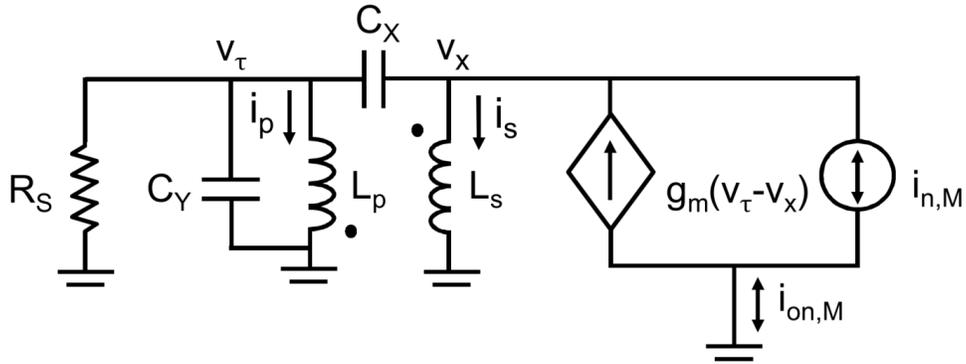


Fig. 3.7 Small-signal model for noise calculation

(higher g_m) requires larger k to achieve a power-match. This is important because obtaining a high k using spiral-inductor based transformers for a non-unity turns ratio is quite challenging. Second, for a fixed coupling coefficient, the power-match achieved by increasing values of n is accompanied by a reduction in g_m , thereby reducing the gain of the amplifier.

Using the values of $\{n, k\}$ obtained from (3-20) to solve for the Q-factor of the matching network (3-19), a plot of the Q-factor as a function of k is plotted in the Fig. 3.6(b). As expected from the result given in (3-18), increasing k of the transformer, while maintaining a power-match, results in a lower Q. In addition, it is interesting to note that as g_m increases, so too does the matching network Q, thereby requiring a high- k transformer for a wideband match.

3.4 Transconductance and Noise

Effective Transconductance

The bandwidth extension provided by transformer-feedback, described in the previous section, is accompanied by a suppression of the ‘closed-loop’ effective-transconductance ($G_m = i_d/v_\tau$) of the SGTxFB amplifier. G_m can be derived using the small-signal model described in Fig. 3.4. Assuming a transistor with infinite output impedance, the short-circuit

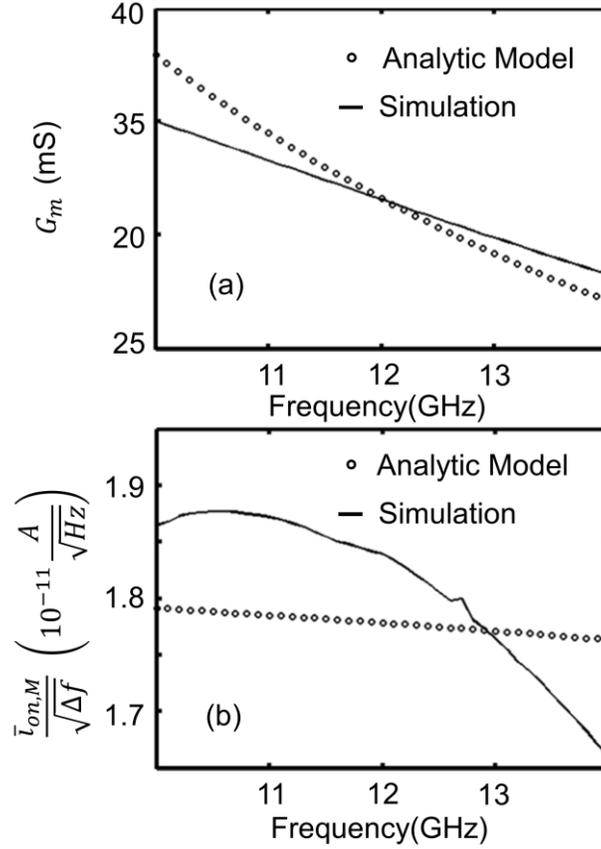


Fig. 3.8 (a) Effective trans-conductance (b) Thermal noise contribution in TC1.

output current (i_d) of the amplifier is a product of g_m and gate-to-source voltage ($v_\tau - v_x$). The relationship between the voltage at the primary (v_τ) and secondary (v_x) of the transformer is derived in Appendix.1. Using (A.3) and (3-11), it can be proven that ($v_\tau - v_x$) is inversely proportional to g_m . As a result, at $s = j\omega_0$,

$$|G_m| = \frac{\left(1 + \frac{k}{n}\right)}{\omega_0 L_s (1 - k^2)} \quad (3-21)$$

With $k = 0$, the effective-transconductance reduces to $1/\omega_0 L_s$. This is similar to the inductively degenerated matching network, fig.3(a), where,

$$|G_m| = g_m Q = g_m \frac{1/\omega_0 C_X}{g_m L_S / C_X} = \frac{1}{\omega_0 L_S} \quad (3-22)$$

It is important to note, though G_m is not an explicit function of g_m in (3-21), specifying $\{\omega_0, n, k, L_S\}$ implicitly constrains g_m . For the test circuit TC1, the G_m obtained from circuit simulation is compared with the analytic model (3-21) in Fig. 3.8(a).

Further insight regarding the relationship between G_m and g_m can be obtained by considering the expression for $|G_m|^2$, as given in (3-23). If the SGTx FB amplifier is perfectly matched to G_S , using (3-15) and (3-23), it is straightforward to prove that G_m is the geometric mean of g_m and G_S .

$$|G_m|^2 = \frac{(n+k)^2}{\omega_0 L_S (1-k^2) \omega_0 L_p (1-k^2)} \quad (3-23)$$

$$|G_m| = \sqrt{g_m G_S} \quad (3-24)$$

Equation (3-24) offers intuition on the impact of SGTx FB loading on the gain of a two-stage amplifier. Consider a two-stage amplifier, the first being $(k-1)^{th}$ stage followed by a k^{th} stage. The input admittance of k^{th} stage ($G_{S,k}$) is designed to be wideband using transformer feedback in order to provide a wideband load to $(k-1)^{th}$ stage. The gain of the $(k-1)^{th}$ stage, {transconductance * load impedance}, is inversely proportional to $G_{S,k}$. However, from (3-24), G_m or the gain of k^{th} stage is directly proportional to $\sqrt{G_{S,k}}$. As a result, the cascaded-gain of the k^{th} and $(k-1)^{th}$ stage is only proportional to $(1/\sqrt{G_{S,k}})$.

Noise Contributions

To employ SGTxFB at the front-end of the receiver, the impact of transformer-feedback on the noise-performance must be considered. Two important noise contributors, thermal noise from the source resistance ($\bar{i}_{on,RS}$) and transistor M_1 ($\bar{i}_{on,M}$), are considered in this section.

Assuming a driving-source with an input resistance R_s , noise power ($v_{n,RS}^2$), and a SGTxFB perfectly matched to the source resistance, it is straightforward to prove that the available noise power density is ($v_{n,RS}^2/4$) [13]. Using the effective transconductance defined previously, the source-induced current-noise power density ($\bar{i}_{on,RS}^2$) at the output of the SGTxFB amplifier is given by,

$$\bar{i}_{on,RS}^2 = G_m^2 \frac{v_{n,RS}^2}{4} \quad (3-25)$$

To analyze the current-noise contributed by M_1 ($\bar{i}_{on,M}$), consider the small-signal model in fig. 7. The source impedance of the driver-stage is modeled by the resistor R_s . Due to transformer-feedback, the thermal noise in the drain-current ($i_{n,M}$) is coupled to the gate of the transistor. The resulting gate voltage-noise is amplified, inverted, and fed back to the drain. The drain-current noise and the current-noise fed back via SGTxFB are fully correlated. As a result the inverted phase of the two noise components suppresses the output noise contribution of M_1 . Applying KCL and KVL,

$$-\frac{v_\tau}{R_s} = i_p + sC_X(v_\tau - v_x) \quad (3-26)$$

$$\bar{i}_{on,M} + sC_X(v_\tau - v_x) = i_s \quad (3-27)$$

$$\bar{i}_{on,M} = g_m(v_r - v_x) + i_{n,M} \quad (3-28)$$

With the help of (A.1), (A.2) and (3-26), it can be proven that (derived in the Appendix.III),

$$\bar{i}_{on,M}^2 = \frac{i_{n,M}^2}{4 + \left(\frac{g_m}{G_m}\right)^2 \left(\frac{nk + 1}{n(n+k)}\right)^2} \quad (3-29)$$

To validate (3-29), the output thermal-noise power density for test circuit TC1 is calculated and compared with the noise-simulation result. The results are shown in Fig. 3.8(b). At 12GHz, the output noise is modeled with an accuracy of $\pm 5\%$ of the simulation result.

Noise Figure

The noise figure (NF) of the SGTxFB amplifier is given by,

$$NF = 1 + \frac{\bar{i}_{on,M}^2}{\bar{i}_{on,RS}^2} \quad (3-30)$$

Substituting, (3-25) and (3-29) in (3-30),

$$NF = 1 + \left(\frac{i_{n,M}^2}{v_{n,RS}^2}\right) \frac{1}{G_m^2 + \frac{g_m^2}{4} \left(\frac{nk + 1}{n(n+k)}\right)^2} \quad (3-31)$$

For a circuit that is perfectly matched to the source resistance R_s (conductance G_s), G_m is governed by (3-24). Thus, with a source noise power, $v_{n,RS}^2 = 4kTR_s \Delta f$, and drain thermal noise current, $i_{n,M}^2 = 4kT\gamma g_m \Delta f$,

$$NF = 1 + \frac{\gamma}{1 + \frac{g_m R_s}{4} \left(\frac{nk + 1}{n(n+k)}\right)^2} \quad (3-32)$$

The noise figure of the SGTxFB is a function of g_m and the transformer parameters (n, k) . As a point of reference, consider a common-gate (CGA) and a resistor-terminated common-source amplifier (CSA) with $NF_{CGA} = (1 + \gamma)$ and $NF_{CSA} = (1 + \gamma/g_m R_s)$, respectively. To highlight the noise contribution of the MOSFET, a noiseless resistor termination has been assumed in the CSA.

In a CGA, the transistor g_m is uniquely specified by the admittance of the driving-source, i.e. $g_m = 1/R_s$. For a power-matched circuit the minimum NF_{CGA} is independent of g_m and only a function of γ . From (3-32) one observes that $NF < NF_{CGA}$ for all values of (n, k) .

For a resistively-terminated CSA, NF_{CSA} is inversely proportional to g_m and therefore suffers from a noise-versus-power trade-off. In order to minimize the noise figure, the current must be maximized while maintaining power and linearity requirements. The SGTxFB relaxes this trade-off by introducing additional design variables (n, k) via the feedback transformer. However, a qualitative comparison between the relative noise-performance of the CSA and the SGTxFB amplifier is difficult because g_m is also dependent on the choice of (n, k) .

In summary, analytic closed-form expressions for the input admittance, quality factor, effective transconductance and noise figure have been derived in this section. As a test-vehicle to validate the results, the design of a three-stage SGTxFB based IF down-converter operating over a frequency range of 11 to 13 GHz is described next. The challenges involved in the design and physical-implementation of a multi-stage SGTxFB is presented in the following section.

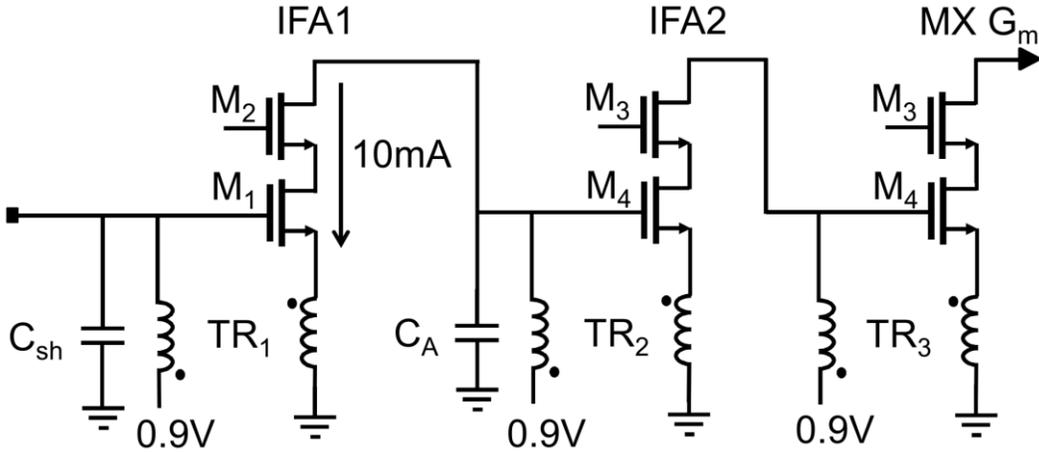


Fig. 3.9. Two-stage stagger-tuned IF-amplifier with SGTxFB driving the mixer transconductance.

3.5 Circuit Design

IF-Amplifier

The IF-amplifier (IFA) consists of two stagger-tuned SGTxFB amplifiers, IFA1 and IFA2. The circuit diagram of the IFA and mixer transconductance-stage is shown in Fig. 3.9. In this circuit, three transformers are included for bandwidth extension. The transformer-feedback network in IFA1 is designed to match the amplifier to the 50Ω impedance of the off-chip measurement circuitry. The overlay-transformer in the input-matching network (TR_1) has a coupling coefficient of 0.7 and is formed using spiral inductors of 780pH ($Q=14$) and 2nH ($Q=10$). At the interface between IFA1 and IFA2, TR_2 is designed to provide a wide bandwidth, high gain load for IFA1. From (3-20), it can be observed that to increase the input impedance of IFA2, the transformer turns ratio should be increased. Thus, TR_2 and TR_3 are 1-to-4 transformers to maximize the gain of IFA1 and IFA2. The simulated frequency-response of IFA1 and IFA2 are shown in Fig. 3.10. The center-frequencies of IFA1 and IFA2 are tuned to 11 and 12GHz, respectively, to reduce the in-band gain variation. Furthermore, to mitigate the impact of the cascode-pole on the frequency response, transistor M_1 and M_2 are sized equally to allow a

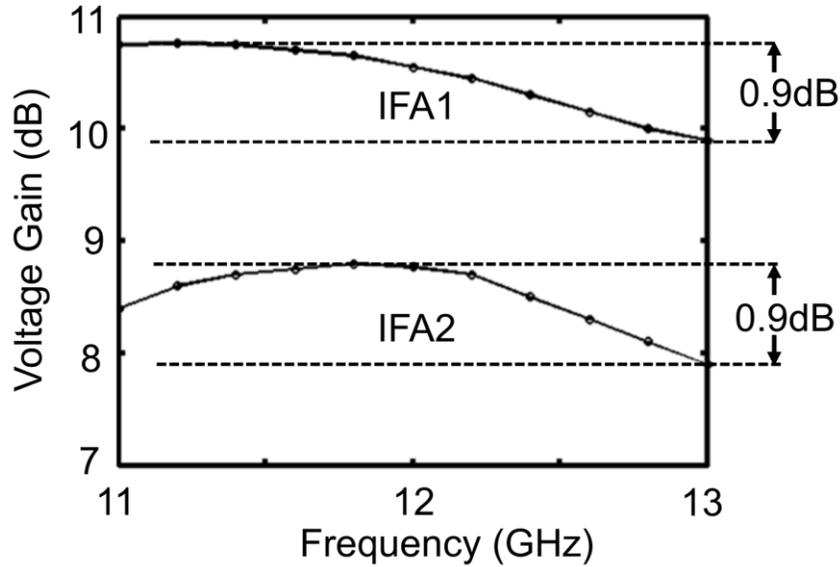


Fig. 3.10. Simulated frequency response of IFA1 and IFA2

shared-junction layout. The gain of each stage of the amplifier varies by less than 1dB over the 2GHz of signal bandwidth. The two-stage IFA achieves a peak gain of 19.2dB while consuming 20mA of current from a 0.9V supply.

Theoretically, the gain of a generic N-stage amplifier can be increased by increasing the number of stages cascaded. However, in the case of the SGT_xFB based IFA, cascading stages becomes challenging from the perspective of unwanted parasitic elements due to the rather complicated routing between transistors and transformers from stage to stage. This problem is better illustrated by drawing a parallel between transformer-coupled and transformer-feedback amplifiers. In the three-stage transformer-coupled amplifier shown in fig. 11(a), the output of stage 1 (drain of the amplifier) and the input of stage 2 (gate of the amplifier) are completely isolated by transformer TR1. A popular-technique to efficiently layout the cascaded configuration relies on using the transformer to route the output of one stage to the input of the next stage, hence the name *transformer-coupled* amplifiers [14][15]; this is shown in fig. 11(b),

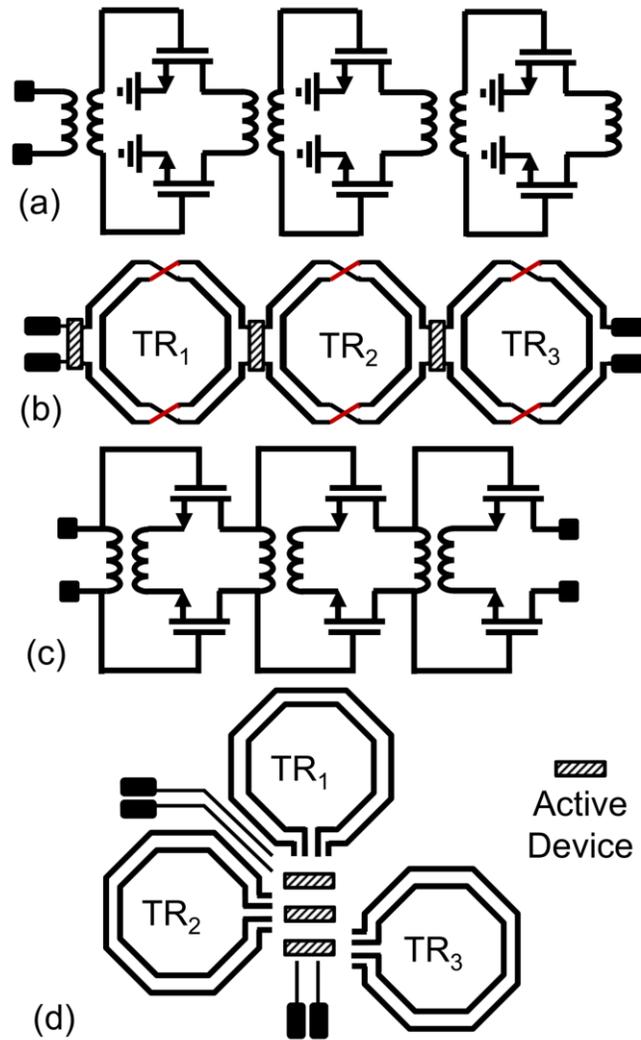


Fig. 3.11 Compact floor-plan for multiple transformer design (a) transformer-coupled circuit (b) layout of multiple transformer-coupled stages (c) transformer-feedback circuit (d) layout of multiple transformer-feedback stages

with the transformers between the active devices. The parasitic routing from the amplifier and transformer can be tightly controlled and minimized. Conversely, in the SGTxFB amplifier, fig. 11(c), both the primary and secondary of a given transformer must be routed to both the gate and source of one set of devices associated with a single stage. As a result, the standard layout from fig. 11(b) applied to SGTxFB results in a significant amount of extra routing and a high corresponding parasitic capacitance. An alternate approach, presented in this work, minimizes

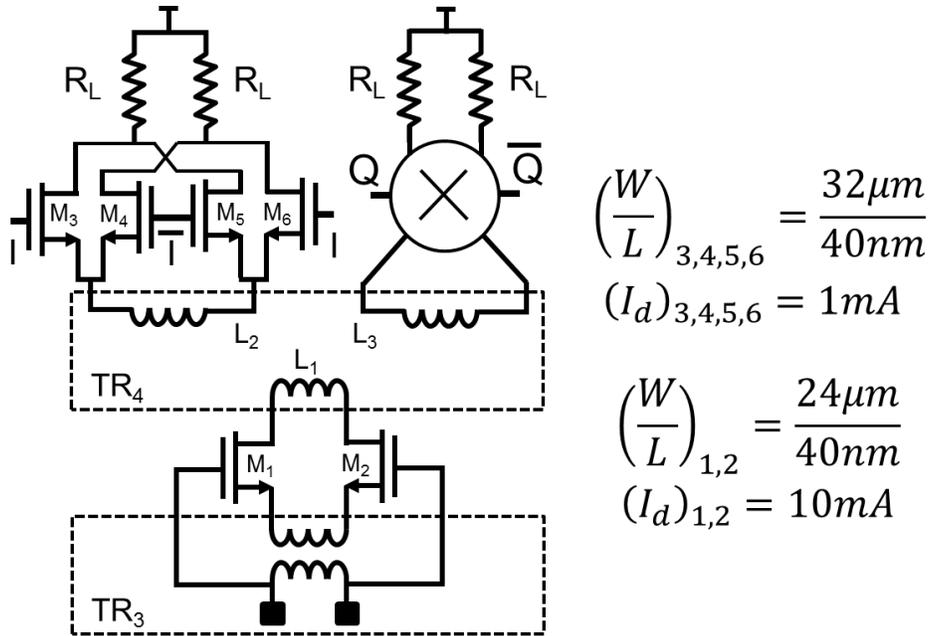


Fig. 3.12 Quadrature down-conversion IF-mixer

the routing between the active devices and the transformers by placing the MOSFET devices of each cascaded stage in a centralized cluster, fig. 11(d). All of the transformers are then wrapped around the clump of active devices, with the primary and secondary of each transformer oriented to facilitate transistor access. This compact layout for a cascaded SGTxFB optimizes the layout for minimal routing between stages; for example between IFA1, IFA2, and the mixer transconductance stage.

IF-Mixer

The output of the IFA is mixed with a 12 GHz I/Q LO signal and down-converted to baseband. The schematic of the IF-mixer is shown in fig. 12. Two transformers, TR₃ and TR₄, are included within the mixer for wideband down-conversion. The SGTxFB network using transformer TR₃ at the interface, between the mixer transconductance-stage and IFA2 has been

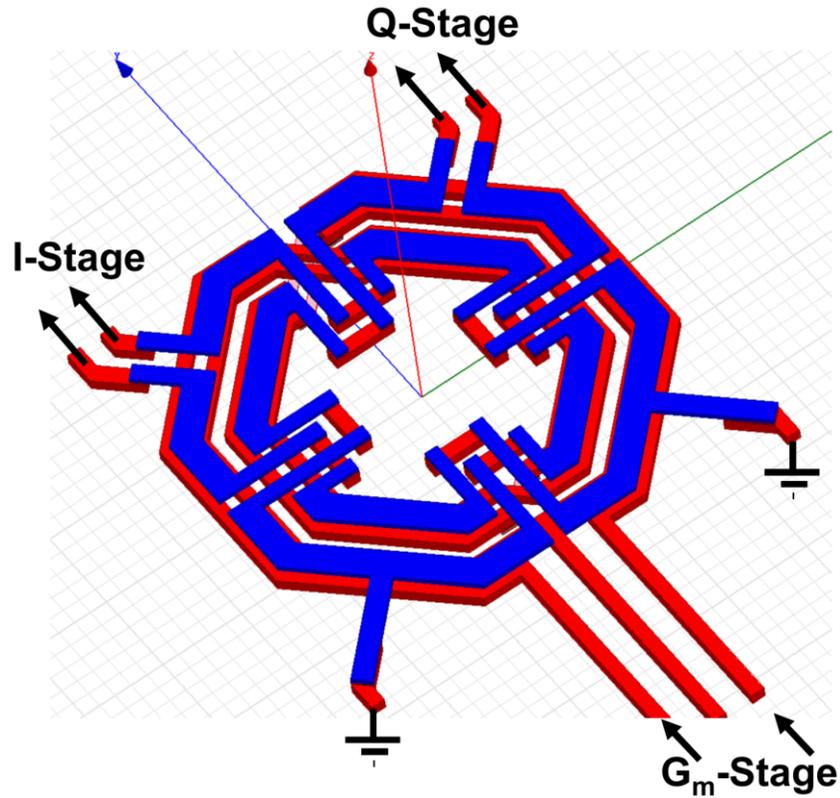


Fig. 3.13 Layout of the three-winding transformer which couples the mixer transconductance to the switching stages.

described earlier. The three-winding transformer (TR_4) couples the output of the transconductance-stage into the switching stage [16]. TR_4 performs two important functions.

First, inductance L_1 resonates with the drain-to-bulk capacitance of $M_{1,2}$ and L_2 resonates with the device capacitance of the switching transistors $M_{3,4,5,6}$. Secondly, the isolation provided by TR_4 prevents the flow of DC current from $M_{1,2}$ into the switching transistors, thereby reducing the flicker-noise contribution of the switch at the baseband output. In addition, isolating the DC-current from the switching stage allows for a higher load-resistance R_L and gain.

The layout of the TR_4 is shown in Fig. 3.13. An overlay transformer structure was used to increase the coupling between the inductor pairs L_1, L_2 and L_1, L_3 . Compared to the switching

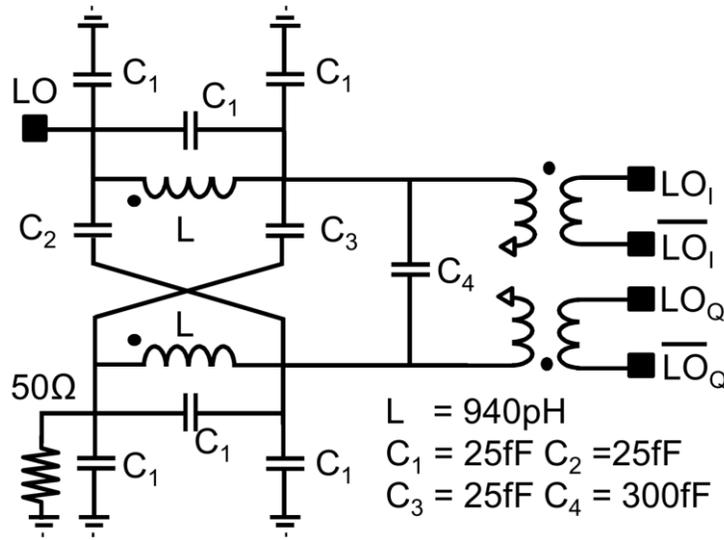


Fig. 3.14. Transformer-based lumped-element Lange coupler

stage, higher current flows through the transconductance stage, thus L_1 has been implemented in an ultra-thick metal layer. L_2 and L_3 carry significantly lower switching current and have been implemented in the Aluminum passivation (AP) layer.

Quadrature LO-Generation

To simplify the measurement set-up, quadrature LO-signals for the mixer are generated on-chip using a single external 12GHz sinusoidal signal source. Several active [16] and passive [18]-[20] techniques for generating quadrature signals have been proposed in literature. Passive I/Q generation circuits including RC poly-phase filters and transmission-line (T-line) based branch-line and Lange couplers are more favorable. However, both T-line and RC-based filters are associated with significant design challenges. In cases where the C values are more than an order of magnitude higher than the layout parasitics, RC poly-phase filters are suitable for I/Q generation. In contrast, T-line based structures are better suited for frequencies above 60GHz; frequencies at which the physical dimensions of the T-line is sufficiently small to allow

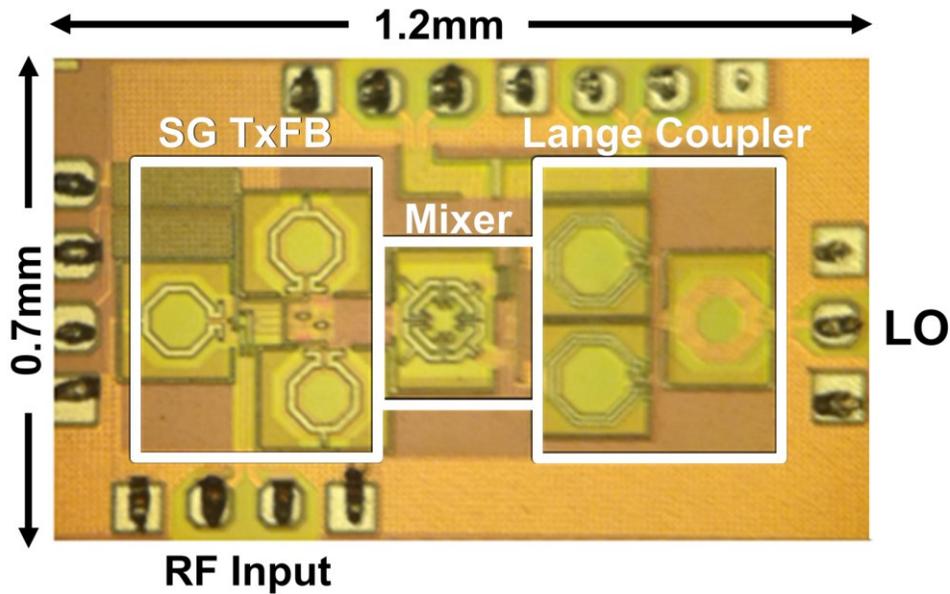


Fig. 3.15. Chip micrograph

implementation on-chip. At 12GHz, the RC poly-phase implementation is highly sensitive to parasitic routing capacitances, and the T-line based quadrature generation technique is area-intensive. At the intersection of these two approaches, the lumped-element implementation of a T-line based I/Q generator is found to be the most optimal.

Two important lumped-element coupler topologies are the branch-line and Lange coupler [20]. Branch-line couplers use only capacitive coupling, have narrower bandwidth, and require large area to ensure zero magnetic coupling between the inductors in the circuit. With the goal of optimizing area, a Lange-coupler has been implemented in this design. Lange-couplers use both capacitive and inductive coupling for the I/Q generation. However, to ensure the amplitude of the I and Q outputs are perfectly matched, a well-controlled magnetic coupling-coefficient (k_m) is required between the inductors. Values of k_m deviating from 0.707 increase the amplitude mismatch. The schematic for the Lange-coupler and balun is shown in Fig. 3.14. From EM

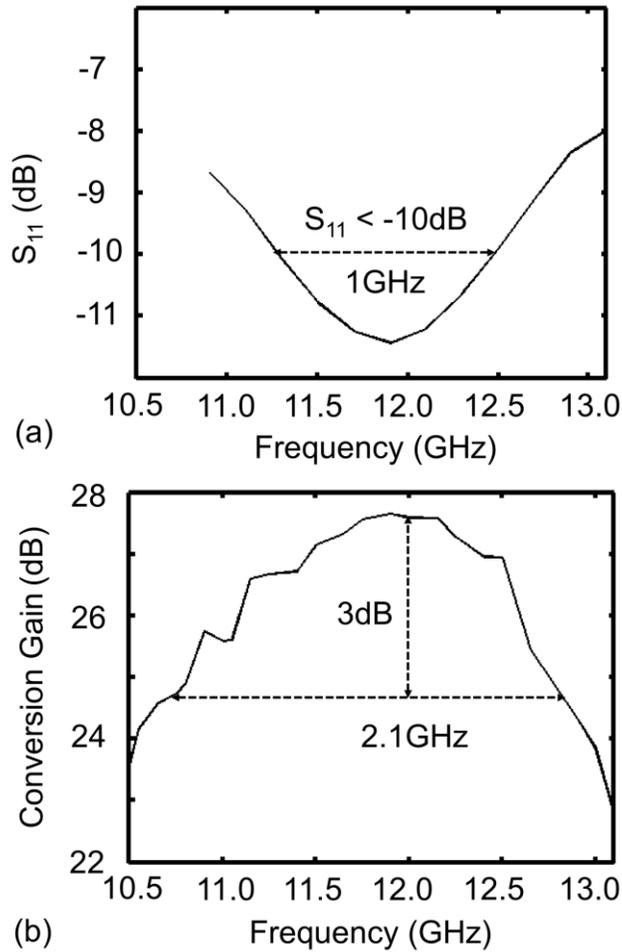


Fig. 3.16. (a) Input matching S_{11} (dB) (b) IF-section down-conversion gain

simulations, LO signals have quadrature accuracy within ± 0.5 degree and an amplitude mismatch of less than 0.8dB.

3.5 Measurement Results and Comparison

The chip [21] was fabricated in a six-metal layer 40nm CMOS process with a top level ultra-thick metal (UTM) layer. The die photo is shown in Fig. 3.15 and occupies an area of 1mm x 0.6mm (pads included). The three transformers used for SGTxFB, the three-winding transformer in the mixer, and the Lange-coupler based quadrature generator circuit are highlighted on the chip micrograph. On-chip wafer probing was done to measure the

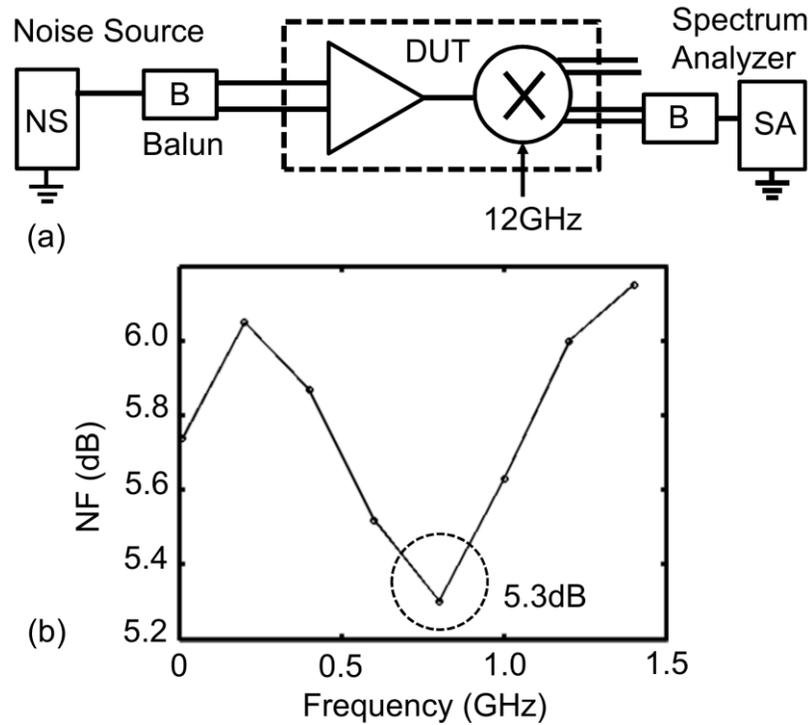


Fig. 3.17 (a) noise figure measurement set-up (b) NF versus frequency

performance. A balun-probe provides a differential input signal; and a single-ended off-chip 12GHz LO signal drives the I/Q generation circuitry.

The chip consumes 30mA of current from a 0.9V supply. The input matching (S_{11}) of the SGTxFB amplifier is shown in Fig. 3.16(a). The matching bandwidth ($S_{11} < -10\text{dB}$) is 1GHz. The conversion-gain of the IF-section is plotted as a function of frequency in Fig. 3.16 (b). The measured peak conversion gain is 27.6dB and the 3dB bandwidth is 2.1GHz. The center frequency of the IF-section is at 11.6GHz; a 400MHz offset from the desired frequency. The noise figure (NF) is measured using the test set-up shown in Fig. 3.17(a). Over the baseband signal bandwidth of 1.08GHz, the total NF variation is less than 0.8dB with a peak NF of 6.1dB and a minimum of 5.3dB. The linearity of the receiver is characterized by a two-tone test. The measured IIP3, with two tones at 10MHz offset from 12GHz, is -22dBm.

Reference	CMOS Tech	V _{DD} (V)	P _{DC} (mW)	RF Freq (GHz)	fBW (%)	S ₁₁ (dB)	Gain (dB)	IIP3 (dBm)	NF (dB)
[22]	180nm	1.8	51 ^a	21.3-29	30.6	< -14.5	38.1	-9	5.5-7.4
[23]	65nm	N/A	40.8	23-25	8.3	N/A	31.5	-13	6.7
[24]	65nm	1-1.2	40	54-67	21	< -14	35.5 (high)	-39 ^b	5.6-6.5
							14 (low)	-21 ^b	
[25]	180nm	1.2	5.2	20-30	40	< -17.6	18.7	-7.6	7.1-14.2
This Work	40nm	0.9	28.8	10.6-12.7	18	< -9	27.6	-22	5.3-6.2

Fig. 3.18 Comparison with state-of-the-art

The performance of the transformer-feedback based wideband receiver has been compared with high fBW K-, K_a- band and 60-GHz direct-conversion receivers in Table I. The 30% fBW UWB pulsed-radar proposed in [22] employs high-order LC bandpass filters for input and inter-stage matching. As a result, the two-stage LNA occupies an on-chip area of 0.93mm² (estimated from die micrograph), approximately 8x larger than the area consumed by the 3 transformers TR₁, TR₂ and TR₃ in Fig. 3.15. The two-stage, single-ended LNA presented in [22] uses shunt-peaking to achieve a fBW of 8%. In addition, the inductor-degenerated input matching-network results in narrow-band input power-match. [24] and [25] propose wideband amplifiers using capacitively-coupled and magnetically coupled resonant tanks, respectively. While [24] achieves a gain, fBW and noise figure similar to this work, [25] achieves double the fBW. However, it is important to note that [25] only includes a single-stage LNA and has 9dB lower gain. Cascading multiple stages to enhance the gain would result in a reduction of the bandwidth.

3.6 Conclusion

Analytic expressions for the input-admittance, quality-factor and noise figure of SGTx_{FB} amplifiers are derived as a function of design variables $\{\omega_o, n, k, g_m, L_s\}$. The impact of high and

low k on the quality factor of the matching network is described. Using transformer-feedback based bandwidth extension techniques, a 16% fBW IF-section consisting of a two-stage stagger-tuned IF-amplifier, a transformer-coupled quadrature mixer and a Lange-coupler is presented. The challenges associated with the layout in multi-stage SGTxFB are highlighted and a strategy for compact layout has been proposed.

Appendix-3.1

The primary and secondary transformer currents can be expressed as a function of v_τ and v_x using (3-6)-(3-9) to obtain (3-33) and (3-34),

$$i_s = \frac{\left(v_\tau \frac{k}{n} + v_x\right) sL_p}{s^2 L_p L_s (1 - k^2)} \quad (3-33)$$

$$i_p = \frac{(v_\tau + v_x nk) sL_s}{s^2 L_p L_s (1 - k^2)} \quad (3-34)$$

From (3-33) and (3-6), the relation between v_τ and v_x can be expressed as,

$$v_x = -v_\tau \frac{\left\{ \frac{\frac{k}{n} sL_p}{s^2 L_p L_s (1 - k^2)} - (g_m + sC_X) \right\}}{\left\{ \frac{sL_p}{s^2 L_p L_s (1 - k^2)} + (g_m + sC_X) \right\}} \quad (3-35)$$

Substituting (3-34) in (3-8),

$$i_\tau = v_\tau \left\{ \frac{sL_s}{s^2 L_p L_s (1 - k^2)} + sC_X \right\} + v_x \left\{ \frac{sL_s nk}{s^2 L_p L_s (1 - k^2)} - sC_X \right\} \quad (3-36)$$

Finally, substituting (3-35) in (3-36) to obtain (3-10)

Appendix-3.2

From (3-15), input-admittance of the SGTxFB amplifier at the resonant frequency ω_0 is,

$$Re(Y_\tau) = \frac{(n+k)}{\omega_0 L_p (1-k^2)} \frac{(n+k)}{\omega_0 L_s (1-k^2)} \frac{1}{g_m} \quad (3-37)$$

Rearranging the terms in (3-37) and with $n = \sqrt{L_p/L_s}$,

$$Re(Y_\tau) = \frac{n \left(1 + \frac{k}{n}\right)}{\omega_0 n^2 L_s (1-k^2)} \frac{n \left(1 + \frac{k}{n}\right)}{\omega_0 L_s (1-k^2)} \frac{1}{g_m} \quad (3-38)$$

From (3-38), and $\{1/\omega_0^2 C_X L_p (1-k^2)\} = 1$, it can be shown that,

$$Re(Y_\tau(s = j\omega_0)) = \frac{\left(1 + \frac{k}{n}\right)^2}{(1-k^2)} \frac{C_X}{L_s} \frac{1}{g_m} \quad (3-39)$$

Appendix-3.3

The output noise analysis is performed using the small-signal model shown in Fig. 3.7. The transformer current equations (3-33) and (3-34) are valid even for this model. Substituting (3-33) in (3-26),

$$\bar{i}_{on,M} + sC_X(v_\tau - v_x) = \frac{\left(v_\tau \frac{k}{n} + v_x\right)}{sL_s(1-k^2)} \quad (3-40)$$

At $s = j\omega_0$, with the help of (3-11) in (3-40) can be simplified to,

$$v_\tau = \frac{\bar{i}_{o,n} j\omega_0 L_s (1 - k^2)}{1 + \frac{k}{n}} \quad (3-41)$$

Next, from (3-26) and (3-34),

$$-\frac{v_\tau}{R_s} - sC_X(v_\tau - v_x) = \frac{(v_\tau + v_x nk)}{sL_p(1 - k^2)} \quad (3-42)$$

Using $n^2 = L_p/L_s$, the terms in (3-42) can be re-arranged to express v_x as a function of v_τ .

Again, using (3-11), it can be shown that

$$v_x = -v_\tau \frac{\left\{ (1 - n^2) + \frac{1}{R_s} j\omega_0 L_p (1 - k^2) \right\}}{(n + k)n} \quad (3-43)$$

$$\bar{i}_{on,M} = g_m v_\tau \left(1 + \frac{(1 - n^2) + \frac{1}{R_s} j\omega_0 L_p (1 - k^2)}{(n + k)n} \right) + \bar{i}_n \quad (3-44)$$

$$\bar{i}_{on,M} = g_m \frac{\bar{i}_{o,n} j\omega_0 L_s (1 - k^2)}{1 + \frac{k}{n}} \left(\frac{nk + 1 + \frac{1}{R_s} j\omega_0 L_p (1 - k^2)}{(n + k)n} \right) + \bar{i}_n \quad (3-45)$$

$$\begin{aligned} \bar{i}_{on,M} - \bar{i}_{on,M} \left(\frac{jg_m \omega_0 L_s (1 - k^2)(nk + 1)}{(n + k)^2} \right. \\ \left. - \frac{1}{R_s} \frac{g_m \omega_0 L_p (1 - k^2) \omega_0 L_s (1 - k^2)}{(n + k)^2} \right) = \bar{i}_{n,M} \end{aligned} \quad (3-46)$$

Now, with the help of (3-29), (3-41), and (3-43), $\bar{i}_{on,M}$ is expressed as a function of $\bar{i}_{n,M}$ only in (3-46). Since the SGTxvB amplifier is perfectly matched to the source resistor R_s , using $Re(Y_\tau)$ from (3-15), it can be shown that,

$$\bar{i}_{on,M} = \frac{\bar{i}_{n,M}}{\left(2 - \frac{jg_m\omega_0 L_s(1-k^2)(nk+1)}{(n+k)^2}\right)} \quad (3-47)$$

Finally, to compute the noise-power

$$\bar{i}_{on,M}^2 = \frac{i_{n,M}^2}{4 + g_m^2 \left\{ \frac{\omega_0 L_s(1-k^2)}{\left(1 + \frac{k}{n}\right)} \right\}^2 \left(\frac{nk+1}{n(n+k)} \right)^2} \quad (3-48)$$

- [1] E. Laskin, M. Khanpour, S. T. Nicholson, A. Tomkins, P. Garcia, A. Cathelin, D. Belot, and S. P. Voinigescu, "Nanoscale CMOS transceiver design in the 90-170-GHz range," in *IEEE Transactions on Microwave Theory and Techniques*, vol.57, no.12, part:2, pp.3477-3490, 2009.
- [2] K. Okada et.al., "A 60-GHz 16QAM/8PSK/QPSK/BPSK direct-conversion transceiver for IEEE802.15.3c," in *IEEE Journal of Solid-State Circuits*, vol. 46, no.12, pp.2988-3004, 2011.
- [3] A. Siligaris et.al. "A 65-nm CMOS fully integrated transceiver module for 60-GHz wireless HD applications," in *IEEE Journal of Solid-State Circuits*, vol.46, no.12, pp.3005-3017, 2011.
- [4] Arbabian, S. Callender, S. Kang, B. Afshar, J-C Chien, and A. M. Niknejad, "A 90 GHz hybrid switching pulsed-transmitter for medical imaging," in *IEEE Journal of Solid-State Circuits*, vol.45, no.12, pp.2667-2681, 2010.
- [5] J. W. May, and G.M. Rebeiz, "Design and characterization of W-Band SiGe RFICs for passive millimeter-wave imaging," in *IEEE Transactions on Microwave Theory and Techniques*, vol.58, no.5, part:2, pp.1420-1430, 2010.
- [6] A. Ismail, and A. A. Abidi, "A 3-10GHz low-noise amplifier with wideband LC-ladder matching networks," in *IEEE Journal of Solid-State Circuits*, vol.39, no.12, pp.2269-2277, 2004.
- [7] Bevilacqua, and A. M. Niknejad, "An ultrawideband CMOS low-noise amplifier for 3.1-10.6-GHz wireless receivers," in *IEEE Journal of Solid-State Circuits*, vol.39, no.12, pp.2259-2268, 2004.

- [8] M. T. Reiha, and J. R. Long, "A 1.2V reactive-feedback 3.1-10.6 GHz low-noise amplifier in 0.13 μ m CMOS," in *IEEE Journal of Solid-State Circuits*, vol.42, no.5, pp.1023-1033, 2007.
- [9] M. Khanpour, K. W. Tang, P. Garcia and S. P. Voinigescu, "A wideband W-band receiver front-end in 65-nm CMOS," in *IEEE Journal of Solid-State Circuits*, vol.43, no.8, pp.1717-1730, 2008.
- [10] A. C. Heiberg, T. W. Brown, T. S. Fiez and K. Mayaram, "A 250mV, 352 μ W GPS receiver RF front-end in 130nm CMOS," in *IEEE Journal of Solid-State Circuits*, vol.46, no.4, pp.938-949, 2011.
- [11] V. Bhagavatula and J. C. Rudell, "Transformer-feedback based CMOS amplifiers," in *IEEE ISCAS*, 2012, pp.237-240.
- [12] J. R. Long, "Monolithic transformers for silicon RF IC design," in *IEEE Journal of Solid-State Circuits*, vol. 35, no.9, pp.1368-1382, 2000.
- [13] J. C. Rudell, J. A. Weldon, J. J. Ou, L. Lin, and P. Gray, "An integrated GSM/DECT receiver: design specifications," Univ. of Cal., ERL Memo, No. UCB/ERL M97/82.
- [14] M. Boers, "A 60 GHz transformer coupled amplifier in 65nm digital CMOS," in *IEEE RFIC Symp. Tech. Dig.*, 2010, pp.343-346.
- [15] D. Chowdhury, P. Reynaert and A. M. Niknejad, "Design considerations for 60 GHz transformer coupled CMOS power amplifiers," in *IEEE Journal of Solid-State Circuits*, vol.44, no.10, pp.2733-2744, 2009.
- [16] J. Paramesh, R. Bishop, K. Soumyanathan and D. J. Allstot, "A four-antenna receiver in 90-nm CMOS for beamforming and spatial diversity," in *IEEE Journal of Solid-State Circuits*, vol.40, no.12, pp.2515-2524, 2005.

- [17] A. Rofougaran, J. Rael, M. Rofougaran and A. Abidi, "A 900 MHz CMOS LC-oscillator with quadrature outputs," in *IEEE ISSCC Dig. Of Tech. Papers*, pp.392-393, 1996.
- [18] F. Behbahani, Y. Kishigami, J. Leete and A. A. Abidi, "CMOS mixers and polyphase filters for large image rejection," in *IEEE Journal of Solid-State Circuits*, vol.36, no.6, pp.873-887, 2001.
- [19] R. C. Frye, S. Kapur and R. C. Melville, "A 2-GHz quadrature hybrid implemented in CMOS technology," in *IEEE Journal of Solid-State Circuits*, vol.38, no.3, pp.550-555, 2003.
- [20] D. Ozis, J. Paramesh and D. J. Allstot, "Integrated quadrature couplers and their applications in image-reject receivers," in *IEEE Journal of Solid-State Circuits*, vol.44, no.5, pp.1464-1476, 2009.
- [21] V. Bhagavatula, M. Boers, and J. C. Rudell, "A transformer-feedback based wideband IF-amplifier and mixer for a heterodyne 60 GHz receiver in 40 nm CMOS," in *IEEE RFIC Symp. Tech. Dig.*, 2012, pp.167-170.
- [22] V. Jain, S. Sundararaman, and P. Heydari, "A 22-29-GHz UWB pulse-radar receiver front-end in 0.18-um CMOS," in *IEEE Transactions on Microwave Theory and Techniques*, vol.57, no.8, pp.1903-1914, 2009.
- [23] A. Mazzanti, M. Sosio, M. Repposi, and F. Svelto, "A 24 GHz subharmonic direct conversion receiver in 65nm CMOS," in *IEEE Transactions on Circuits and Systems-I:Regular Papers*, vol.58, no.1, pp.88-97, 2011.
- [24] F. Vecchi et al., "A wideband receiver for multi-Gbit/s communications in 65nm CMOS," in *IEEE Journal of Solid-State Circuits*, vol.46, no.3, pp.551-561, 2011.

- [25] H. Li, C. N. Kuo, and M. C. Kuo, "A 1.2-V, 5.2-mW, 20-30GHz wideband receiver front-end in 0.18um CMOS," in *IEEE Transactions on Microwave Theory and Techniques*, vol.60, no.11, pp.3502-3512, 2012.

4 BAND-PASS DISTRIBUTED-AMPLIFIER

Practical amplifying devices are fundamentally bound by a finite gain-bandwidth product. Using the MOS transistor as an example, doubling the width of the device to increase the transconductance is invariably accompanied with a proportional increase in the gate-to-source and the drain-to-bulk capacitance. The increase in capacitance at the input and output nodes results in a bandwidth reduction. The distributed amplifier (DA) topology, originally described in a british patent disclosure by Percival [1] in 1938, is a technique to relax this gain-bandwidth constraint. In a DA topology the gain stages are connected such that the capacitances are isolated while the output currents add in-phase. Over the decades, this technique has been applied for wideband amplifier design in vacuum-tube [2], bipolar and CMOS [3] technologies. Not surprisingly, one of the earliest millimeter-wave CMOS circuits, reported at the ISSCC 2004 [4], was a DA.

The distributed-amplifier (DA) topology has been extensively applied in systems that require signal amplification starting from DC to several tens of giga-hertz. For wireline and optical applications, for example, to compensate for signal loss during long-distance transmission using optical-fibers transmission, high bandwidth DAs have been demonstrated in [4][6]. However, in wireless transceivers, where signal amplification beyond the desired channel is detrimental rather than beneficial, DA based circuits have not been applied. Front-end circuits for narrow-band cellular and wireless standards circuits are narrow-band and require off-chip filters with a high-quality factor to reject out-of-band blockers and image signals. In contrast, the receiver architecture proposed in Chapter-1 is unique in that it requires an intermediate frequency (IF) amplifier with an fBW in excess of 50%. In this chapter we explore techniques to extend the

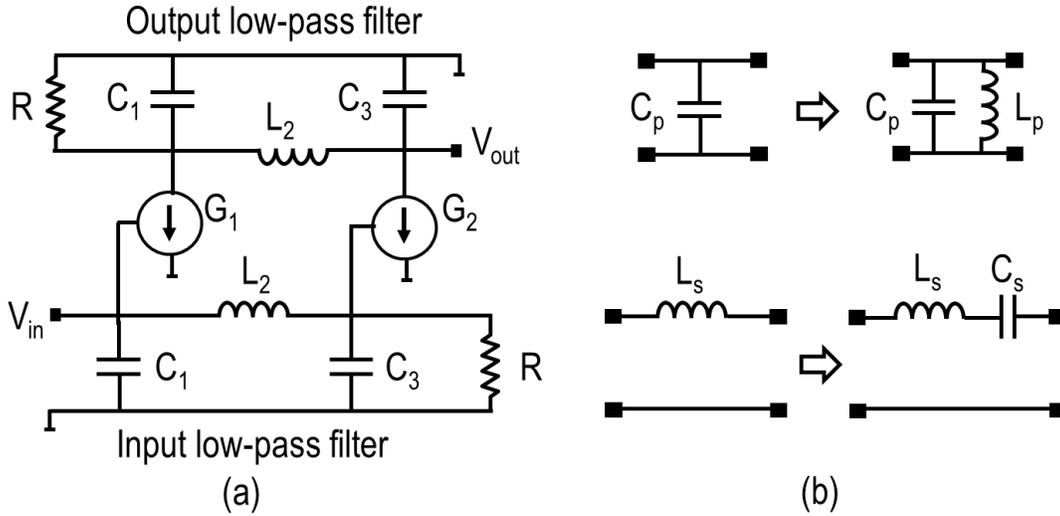


Fig. 4.1. (a) Canonical form of the LPDA and BPDA (b) Low-pass to band-pass filter transformation.

principles of LPDAs for band-pass signal amplification over a wide frequency-range in integrated radio transceivers.

The biggest barrier to the practical implementation of an integrated DA is the large area associated with multiple on-chip spiral-inductors or transmission-lines. The limitations of a canonical band-pass distributed amplifier (BPDA) structure derived using standard filter transformations will be described in section 4.2. In section 4.3, a methodology to apply dual mirror-symmetric Norton transformations on the canonical BPDA to derive a more compact BPDA while preserving the frequency response characteristics is presented. A detailed description of a prototype BPDA test-chip implemented in a 40nm CMOS process is presented in Section 4.4. Finally, the measured results and a comparison with prior-art are presented in Section 4.5.

4.1 Canonical form of a BPDA

The LPDA is comprised of two low-pass filters (LPFs); the first at the input node and the second at the output node. A simplified schematic of an LPDA with two third-order LPFs is

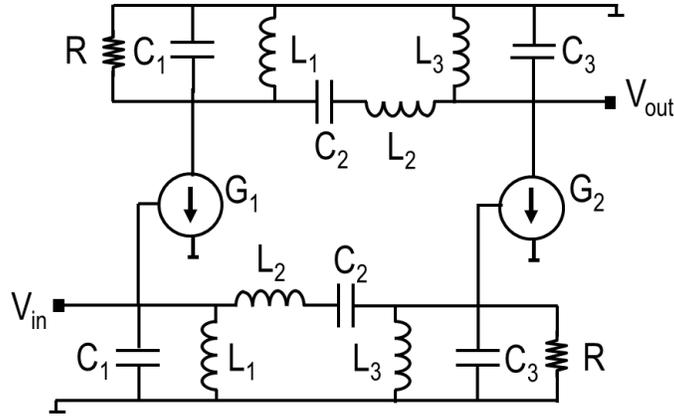


Fig. 4.2. Canonical form of the BPDA

shown in Fig. 4.1(a). The gain-cells have been approximated by ideal voltage-controlled current-source or trans-conductance G_1 and G_2 . In a BPDA, each low-pass filter of the LPDA is replaced by a band-pass filter.

A BPF can be derived from an LPF using the filter transformation shown in Fig. 4.1(b). Each series inductor L is replaced by a series L - C and each shunt capacitor C by a shunt L - C . The resulting canonical implementation of the BPDA is shown in Fig. 4.2. The biggest drawback in this particular implementation of a BPDA is the increase in the number of the inductors. If we start with an LPDA with N inductors, the number of inductors in the canonical-BPDA is $2N+1$; thereby exacerbating the area-challenges which plague all LPDA designs. However, a more subtle, but equally important, drawback in canonical BPDA is related to the scaling of inductive components (L_1 , L_2 and L_3) as a function of bandwidth (BW) and center frequency (ω_c). To illustrate these drawbacks, the procedure to design an LPDA and BPDA is described next.

For the sake of simplicity, we again return to the third-order LPDA of Fig. 4.1(a). The circuit comprises of two ideal trans-conductance cells G_1 and G_2 . Gain-cell G_1 has (equal) input and output capacitance of C_1 . Similarly, the gain-cell G_2 has an input and output capacitance of

Table 4.1. Element values for maximally flat low-pass filter prototype ($g_0 = 1, \omega_c = 1, K = 1$ to 8)

K	g_1	g_2	g_3	g_4	g_5	g_6	g_7	g_8	g_9
1	2.0000	1.0000							
2	1.4142	1.4142	1.0000						
3	1.0000	2.0000	1.0000	1.0000					
4	0.7654	1.8478	1.8478	0.7654	1.0000				
5	0.6180	1.6180	2.0000	1.6180	0.6180	1.0000			
6	0.5176	1.4142	1.9318	1.9318	1.4142	0.5176	1.0000		
7	0.4450	1.2470	1.8019	2.0000	1.8019	1.2470	0.4450	1.0000	
8	0.3902	1.1111	1.6629	1.9615	1.9615	1.6629	1.1111	0.3902	1.0000

C_2 . A series inductor L_2 separates the gain-cells (and by extension, the input and output capacitances) to form two low-pass C-L-C filters, one at that input and other at the output of the LPDA. The bandwidth of the DA is a determined by the bandwidth of the C-L-C filter. Fortunately, there exists a well-documented [7] algorithmic approach to select component values for the desired bandwidth and transfer function. As an example, a filter-coefficients table for obtaining a Butterworth filter response (up to the eight-order) is shown in Table 4.1. In this table, K corresponds to the order of the filter and g_k the normalized filter component values. Using Fig. 4.1(a) as a reference, for $N=3$ the filter coefficients $g_1=1, g_2=2$ and $g_3=1$ map to C_1, L_2 and C_3 , respectively. The coefficient values are normalized to a unit termination impedance and a unit bandwidth. To design a filter with a termination impedance of 50-ohms and a bandwidth BW , the component values are

$$R_1 = g_1 * Z_0 \quad (4-1)$$

$$C_1 = \frac{g_1}{BW * Z_0} \quad (4-2)$$

$$L_2 = \frac{g_2 * Z_0}{BW} \quad (4-3)$$

$$C_3 = \frac{g_3}{BW * Z_0} \quad (4-4)$$

$$R_4 = g_4 * Z_0 \quad (4-5)$$

From (4-3) it can be observed that as the BW specification of the filter increases, the value of inductor L_2 decreases. Therefore, from an area perspective, as the BW of the LPDA increases, the area of the circuit reduces.

Also, from (4-2) and (4-4), it is interesting to note that the BW increase is accompanied by a reduction in the value of capacitors C_1 and C_2 . As described previously, the basic idea of a ‘distributed-amplification’ is to absorb the input and output capacitances of the gain cell into the low-pass C-L-C filter. Assuming a gain-cell implemented using MOS or BJT, the input/output capacitance is directly proportional to the trans-conductance of the device. Thus, an increase in BW translates into a smaller transistor and a lower-gain from the DA.

The BPDA in Fig. 4.2 consists of capacitors (C_1, C_2, C_3) and inductors (L_1, L_2, L_3). For, three among these six components (C_1, L_2, C_3) the equations (4-1) to (4-5) are applicable. Therefore, the aforementioned BW scalability of the L and C values is equally applicable in the case of a BPDA design. In addition, it is important to note that L_2 is independent of ω_c . Therefore, as the center-frequency of the BPDA increases the inductor L_2 operates closer to its

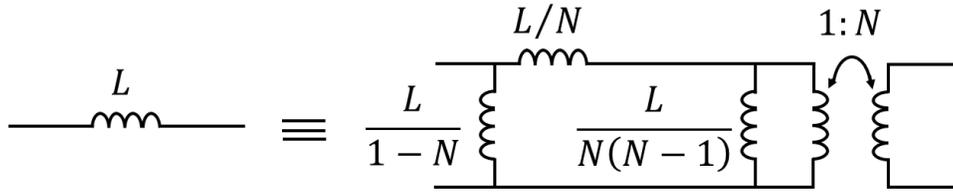


Fig. 4.3. Norton transformation of a series floating inductor

self-resonance frequency (SRF). The three additional components of the BPDA (L_1 , C_2 , L_3) can be shown to be,

$$L_1 = \frac{g_1 * BW * Z_0}{\omega_c^2} \quad (4-6)$$

$$C_2 = \frac{BW}{g_2 * \omega_c^2 * Z_0} \quad (4-7)$$

$$L_3 = \frac{g_3 * BW * Z_0}{\omega_c^2} \quad (4-8)$$

From (4-7) the shunt inductance L_1 is inversely proportional to $1/\omega_c^2$, but directly proportional to the target BW. Thus, for a given center-frequency, the shunt-inductor L_1 operates closer to its SRF as the BW increases.

In summary, at millimeter-wave frequencies, the SRF of the inductors $L_{1,2}$ ultimately limits ω_c and the BW achievable for a BPDA implementation. In the next section, we will describe how Norton transformations can be applied to substantially reduce the value of the series inductances required.

4.3 Norton Transformation

The basic idea of a Norton transform (NT) is illustrated in Fig. 4.3. The series floating inductance L is replaced with an electrical equivalent circuit using the inductors

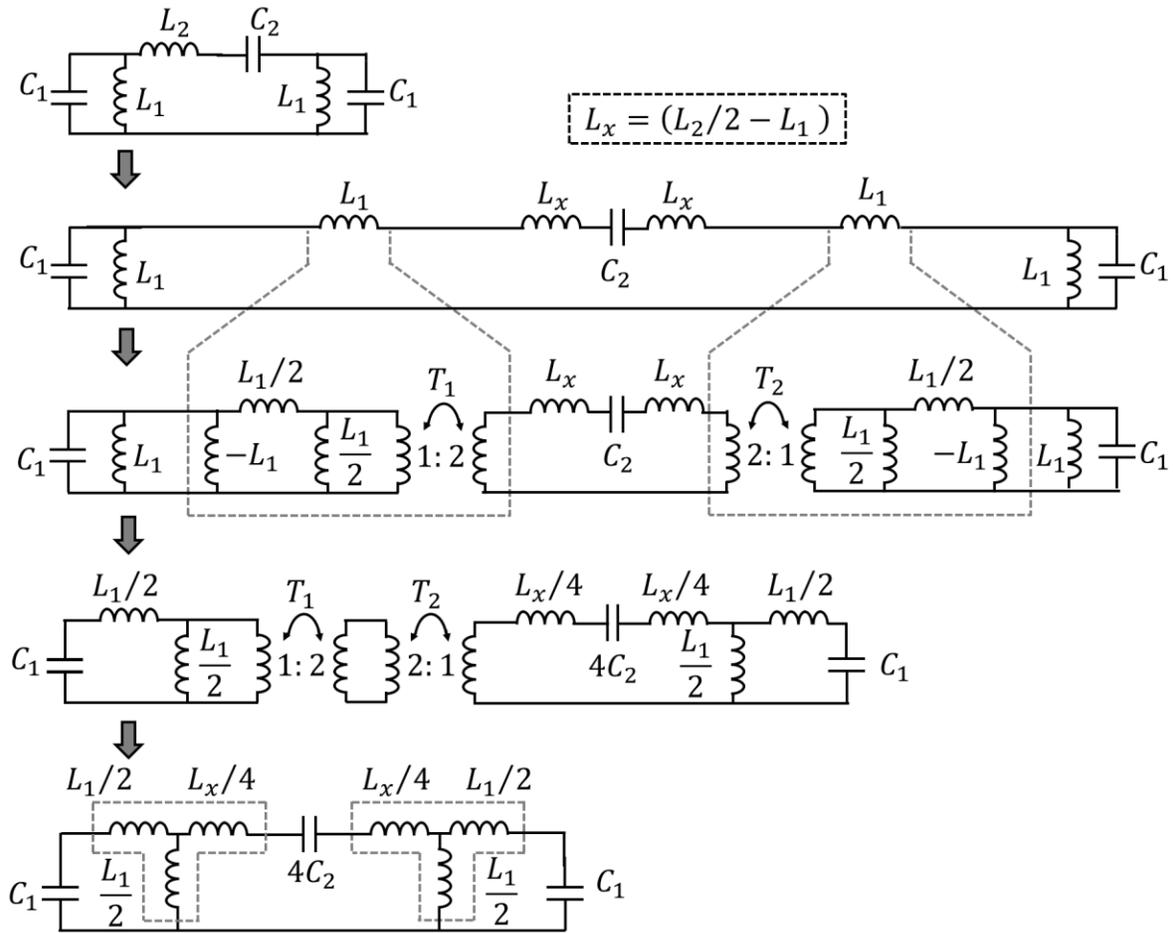


Fig. 4.4. Derivation of the compact-area bandpass filter from the canonical bandpass filter through the application of mirror-symmetric dual Norton-transforms.

labeled $L/(1 - N)$, L/N , $L/\{N(N - 1)\}$, and a single $1:N$ transformer. Although at first glance the NT appears to increase the number of inductors in the circuit, a component reduction may be found by setting N equal to 2. The presence of an effective negative inductance $L/(1 - N)$ can be exploited to further reduce the number of inductors in the BPDA. Starting with the BPF canonical form, shown in Fig. 4.4, component reduction may be found by setting N equal to 2. Inductor L_2 is first split into two equal valued series inductors and placed symmetrically about capacitance C_2 . Next, the left most $L_2/2$ inductor is further split into two inductances, L_1 and $L_x = L_2/2 - L_1$. Applying an NT on the series inductor L_1 ($N = 2$), results in an effective

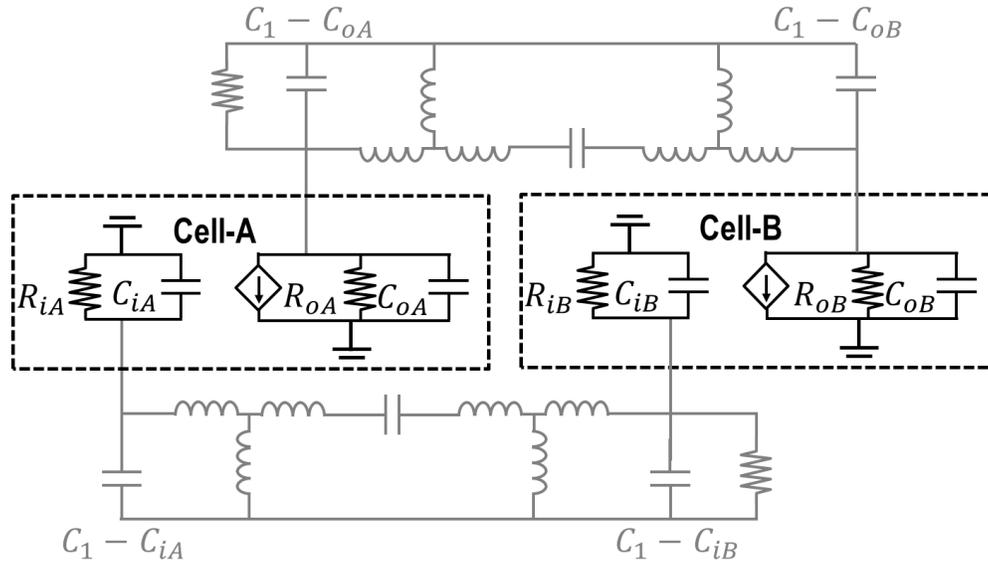


Fig. 4.5. Input and output impedances of the gain-cells

negative inductance, $-L_1$, in parallel with a shunt inductor, L_1 ; note $L_1 // -L_1$ is an open circuit, effectively eliminating both inductors. The NT also introduces a 1:2 transformer T_1 .

The same process is repeated by replacing the second half-inductance $L_2/2$ with L_x and L_1 . The synthesis continues by reflecting the two residual series inductances, L_x , and capacitance, C_2 , across the windings of transformer T_2 . This scales the required value of inductor L_x by $(1/N)^2$. Finally, the residual transformers T_1 and T_2 , produced by the mirror-symmetric NTs, now appear in cascade, and have equal turns ratios of 1: N and N :1 respectively, which effectively neutralize (eliminates) one another. In summary, inductor L_2 in the canonical form of the bandpass filter is reduced to two series inductances of $\{(L_2/2) - L_1\}/N^2$ after recursive transformations are applied; a minimum reduction of 75% ($N = 2$) in the value of L_2 . The shunt-inductor inductor L_1 is split into two equal halves along the series and shunt signal paths.

4.4 Gain-cell design

The preceding discussion assumes that the gain-cells only present a capacitive load on the LC band-pass filter. However, the input impedance of gain-cells implemented using MOS or BJT trans-conductors has a significant resistive component as well. Fig. 4.5 shows a simplified model of a BPDA in which the input and output resistances of the gain-cells are highlighted. Gain-cell A has an input impedance $R_{iA} \parallel C_{iA}$, and an output impedance $R_{oA} \parallel C_{oA}$. Similarly, gain-cell B has an input impedance $R_{iB} \parallel C_{iB}$, and an output impedance $R_{oB} \parallel C_{oB}$. The following constraints can be placed on the gain-cells.

- The input capacitors (C_{iA} , C_{iB}) and output capacitors (C_{oA} , C_{oB}) load the input and output band-pass filters, respectively. Therefore, these capacitance values have to be lower than the upper-limit specified by (4-2) and (4-4).
- The input resistance of cell-B (R_{iB}) is in shunt with the termination resistance on the input BPF. Therefore, the designer has the flexibility to choose large size devices to achieve a high-gain at the expense of a reduced R_{iB} . The value of the termination resistance can be chosen in conjunction with R_{iB} to ensure a good 50-ohm termination.
- Similar to R_{iB} , the output resistance of the cell-A (R_{oA}) can be low if necessary because it appears in parallel with the output termination resistance
- Cell-A (R_{iA}) is directly driven by the external 50-ohm source. Therefore, the input resistance of cell-A (R_{iB}) must be high.
- Similar to R_{iA} , the output impedance of Cell-B must be high.

4.5 Implementation

The circuit diagram of the BPDA is shown in Fig. 4.6. For the sake of clarity only the input (gate) BPF is shown explicitly in Fig. 4.6; the output (drain) BPF is identical in structure.

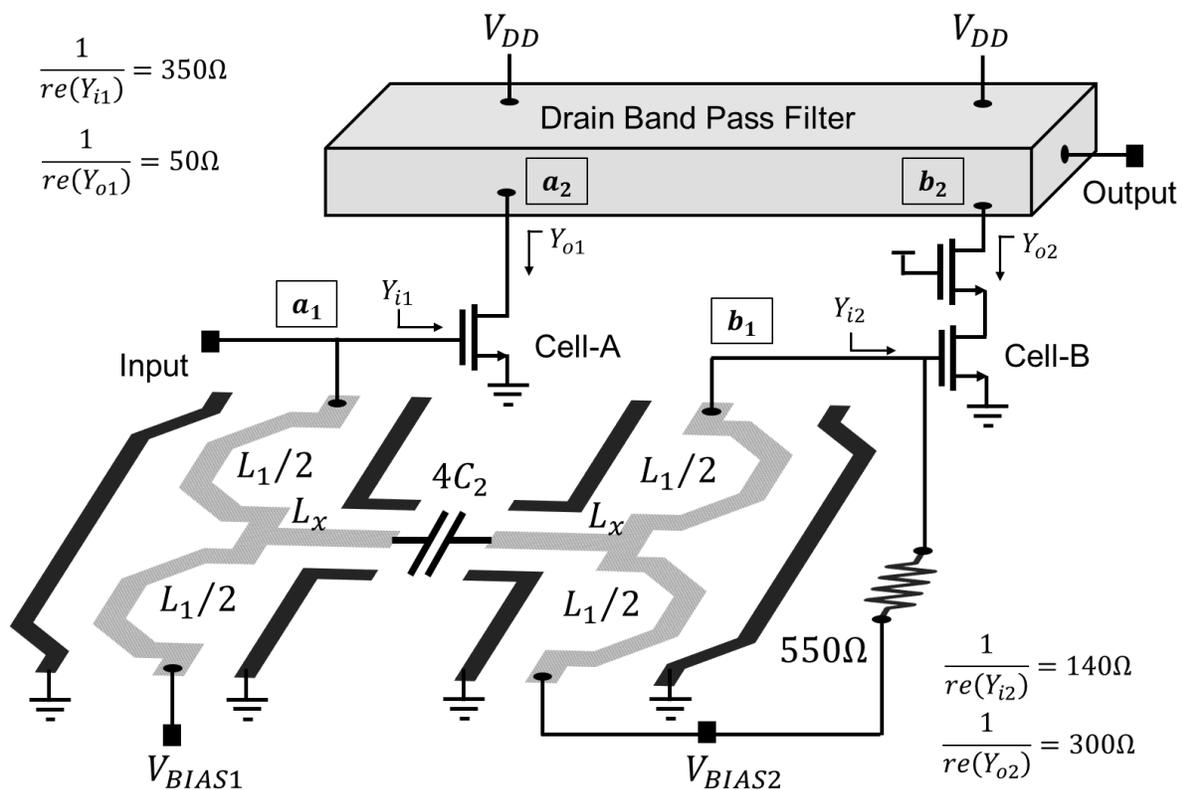


Fig. 4.6. Circuit diagram of the BPDA

The BPDA reported in this work consists of two doubly-terminated band-pass filters; each filter has a Butterworth-filter response. Through the application of Norton transforms, the inductance values of $L_1=180\text{pH}$ and $L_2=800\text{pH}$ in the canonical structure are reduced to 90pH , 90pH , and 55pH , respectively. The 800pH series inductor has been reduced to two inductances of 55pH each. The band-pass section is now realized as two symmetric T-sections; highlighted in Fig. 4.4. Thus, although the Norton transform (NT) appears to increase the number of inductors in the circuit, the value of these inductors are at least an order of magnitude lower. In addition to the size advantage, it is noteworthy to mention the 55pH inductor will have a significantly higher SRF as compared to an 800pH inductor, further ensuring proper operation at mm-wave frequencies.

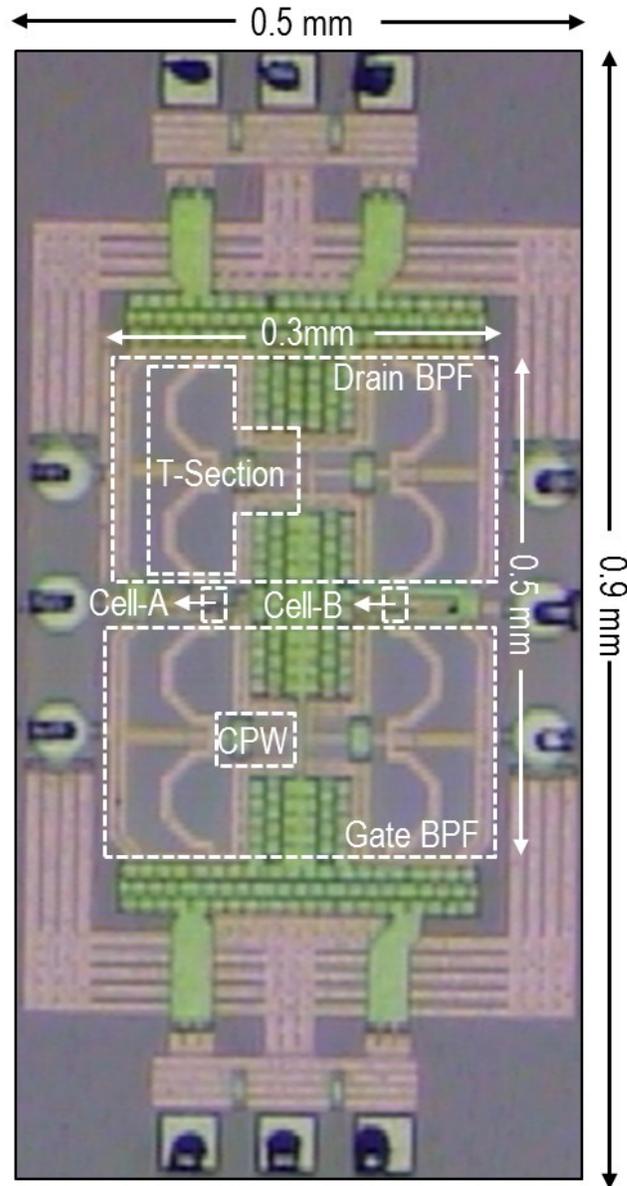


Fig. 4.7. Chip Micrograph for the BPDA

Although reducing the BPDA inductance values to as low as 55pH provides a significant area advantage, it makes the transfer function more sensitive to routing-dependent parasitic inductance and capacitance. To alleviate this concern, the scaled inductor $L_x/4$ is realized as a coplanar waveguide (CPW) and used as routing between two symmetric T-sections in the simplified BPDA structure, Fig. 4.6. The series and shunt inductors $L_1/2$ are implemented as

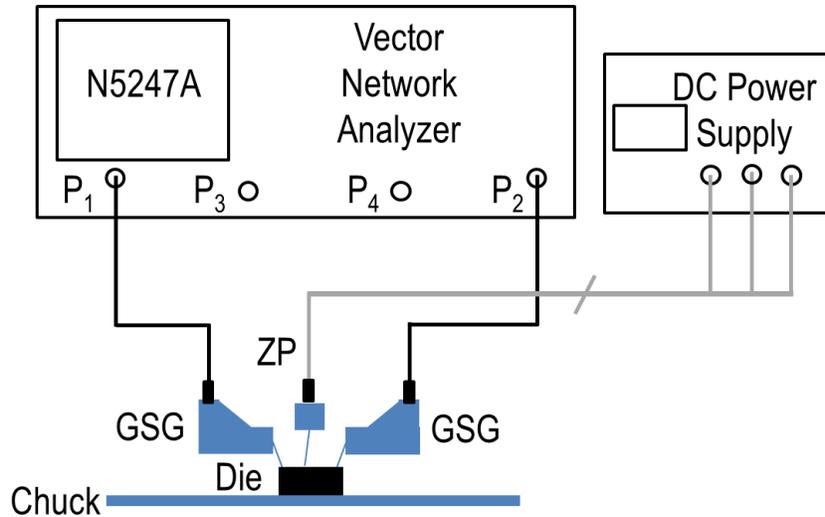


Fig. 4.8. Setup for BPDA S-Parameter measurement

half-turn spiral inductors to eliminate additional routing to the supply and bias pads, Fig. 4.6. To account for stray parasitic capacitance and mutual magnetic-coupling, the three inductors in each T-section were modeled and simulated as a single, three-port passive structure.

The design of the gain-cell is based on the criteria specified in section 4.4. For gain cell-A, the output impedance can be low but the input impedance has to be high; therefore, a common-source trans-conductance stage with length of 40 μ m and width of 64 μ m was selected. cell-A has an $R_{in}=350\Omega$ and low $R_{out}=50\Omega$. For gain-cell B, the input impedance can be low but the output impedance has to be high; therefore a cascode device was selected. The cascode gain-cell B has an $R_{in}=140\Omega$ and a high $R_{out}=300\Omega$. Both cells consume a current of 17mA from a 1V supply.

4.5 Measurement Results

The die photograph of the prototype mm-wave receiver chip fabricated in a TSMC 40nm CMOS process is shown in Fig. 4.7. The BEOL consists of a 6 metal stack with 1 ultra-thick

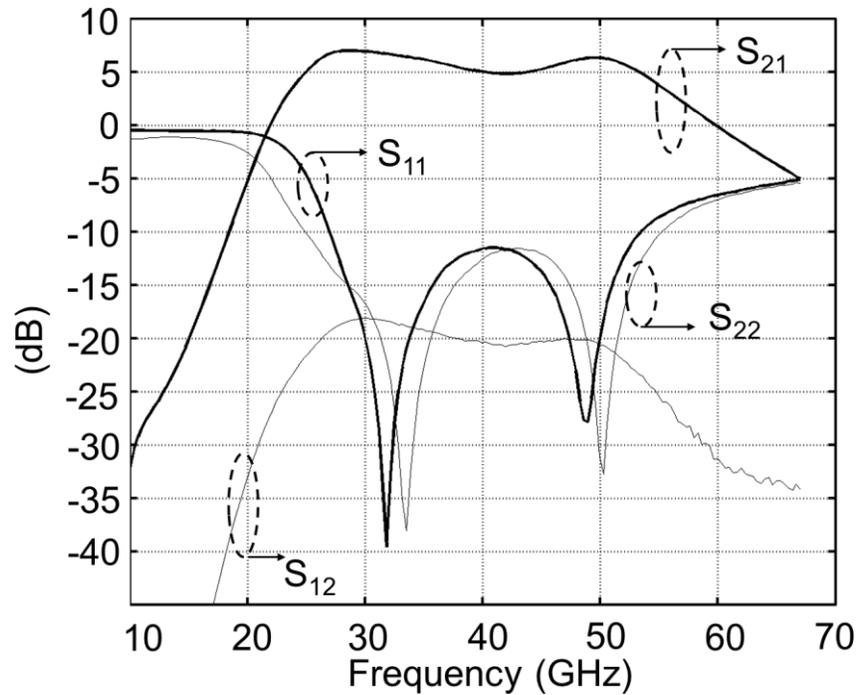


Fig. 4.9. Measured S-parameters

metal (UTM) and 1 aluminum passivation layer (AP). The 3.5 μm thick UTM is the top-most copper layer 2.3775 μm above the surface of the substrate. The compact core DA occupies an area of less than 0.5mm x 0.3mm.

The chip was characterized using Cascade's 12000AP Summit on-wafer probe station. The DC-supply is brought on chip using a 3-pin Z-probe on the north-end. The gate bias voltages are brought on-chip through a Z-probe from the south-side. The millimeter-wave input and output are brought on-chip using GSG probes on the west and east-side, respectively.

Measurements were performed using Agilent's N5247A PNA-X. For accurate S-parameter measurements, a two-port SOLT calibration was done on the Impedance Standard Substrate (ISS). For linearity metrics, such as IIP3 and P1dB, which require accurate power measurement, power-calibration was performed up to the end of the cables. The millimeter-wave

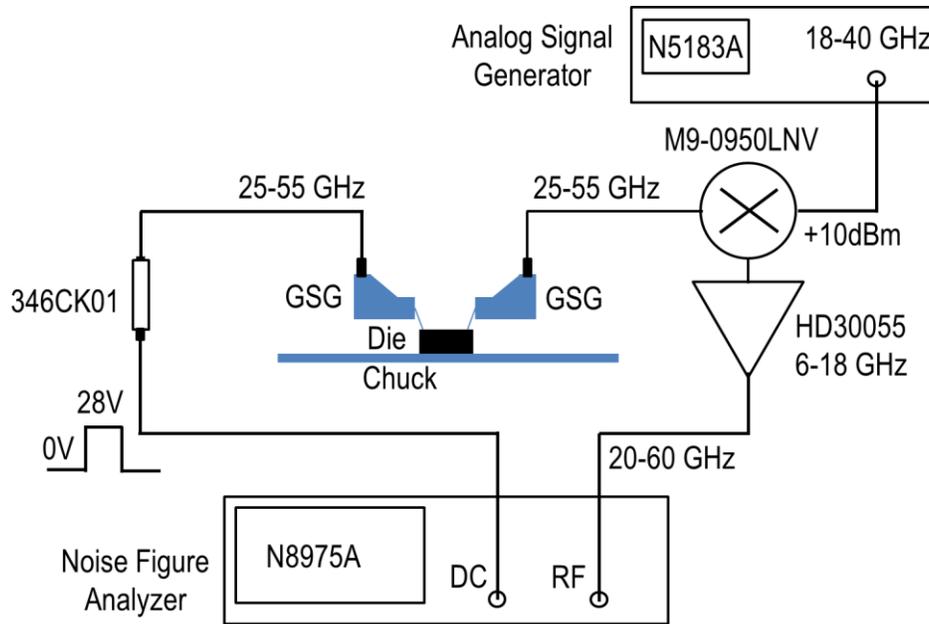


Fig. 4.10. Setup for BPDA noise characterization using the N8975A noise-figure analyzer input and output were provided through 67-GHz GSG Cascade infinity probes, and all the instruments were interconnected using 1.85mm high-frequency co-axial cables.

The test setup for S-parameter and linearity measurements is shown in Fig. 4.8. Only a single-probe has been shown in the diagram for the sake of clarity. The measured S-parameters of the BPDA are plotted in Fig. 4.9. The S_{11} and S_{22} are less than 10dB across a frequency range of 26.8GHz-to-54GHz, and 24.8GHz-to-55GHz, respectively. The peak-gain of the amplifier is 7dB with a 3-dB pass-band from 24GHz to 54GHz; a BW of 30GHz, or an fBW of 77%. The total in-band gain-variation is 2dB.

The noise-figure is measured using Agilent's 346CK01 noise-source and Agilent's N8975A 26.5 GHz noise-figure analyzer (NFA). The NC5115 is a 1.85mm co-axial calibrated noise-source operating over a bandwidth of 10MHz-50GHz. The output frequency of the BPDA, 26-to-56 GHz, lies outside the measurement range of the NFA. Thus, an external down-

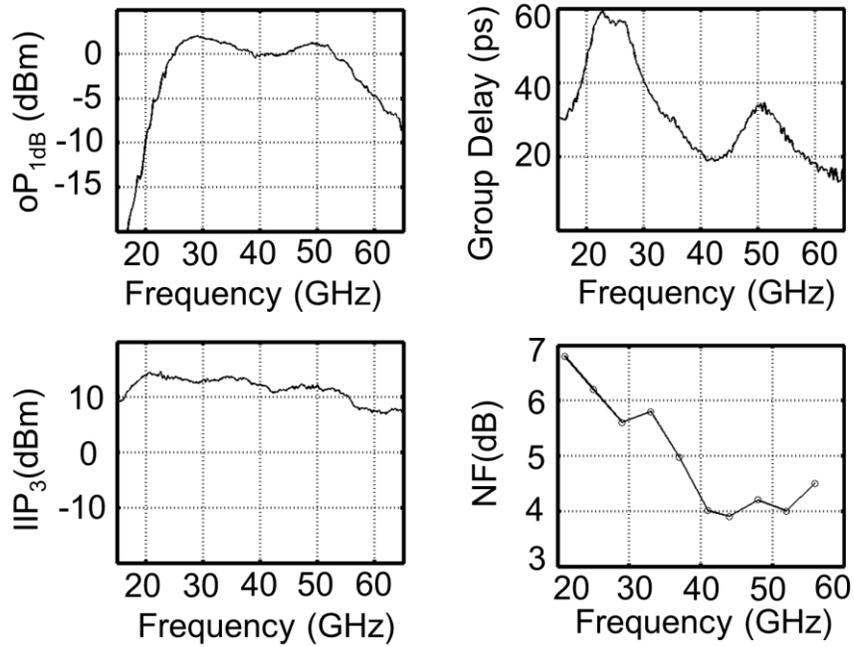


Fig. 4.11. Compression-point, group-delay, and IIP3 characterization versus frequency

conversion is required to bring the output noise of the NFA within the measurement band of the NFA.

The NF of the receiver is measured using the Y-factor method. The block diagram of the test setup is shown in Fig. 4.10. A 0/28V pulse-source inside the NFA drives the millimeter noise-source into cold/hot state. The output of the BPDA drives Marki Microwave's M9-0950 wideband passive mixer. To compensate for the 10dB in-band conversion loss of the mixer, the signal is provided 20dB of amplification using HDCom's HD30055 6-18GHz low noise amplifier. The output of the amplifier is measured by the NFA. Using two measured data-points, power in the hot state and power in the cold-state, and the effective noise-ratio (ENR) of the noise-source, the NF is calculated. This process is repeated at each frequency step. The variation in NF as a function of frequency is plotted in Fig. 4.11. The minimum NF is 3.9dB. The NF remains less than 6.2dB over the frequency range of 24 to 54 GHz.

Table 4.2 Comparison and performance summary for the BPDA

Parameters	[8] IMS'10	[9] JSSC '10	[10] JSSC '13	This Work
Technique	High-pass DA	Coupled resonator	T-type matching	BPDA
Bandwidth (GHz)	21-42.5	23-32	47-77	24-54
ω_c (GHz)	31.75	9	30	39
fBW	67.7%	32%	48%	77%
Gain (dB)	8.3 Power gain	12 Voltage gain	22.5 Power gain	7dB Power gain
Power (mW)	28	13	52	34
IIP3 (dBm)	-x-	-6.3 to -4.5	-x-	10 to 13
OP ₋₁ (dBm)	0	-x-	4.5 (simulated)	-0.5 to 2
Area(mm ²) Pads not included	0.28	0.25	0.55	0.15
Technology	120nm SiGe-BiCMOS	180nm BiCMOS	250nm BiCMOS	40nm CMOS

To characterize the wideband linearity of the BPDA, the compression-point and intercept point were measured across frequency. The N5247A contains two internal signal-sources, thereby simplifying two-tone testing. The linearity measurement setup is identical to the one used for S-parameter measurement, Fig. 4.8. The minimum and maximum in-band output 1dB compression points are -0.5dBm and 2dBm, respectively. The minimum in-band IIP3 for a 100MHz tone-spacing is 11dBm. As shown in Fig. 4.11, the in-band group-delay varies between 20ps and 55ps over the frequency-range of interest.

The performance of the BPDA is summarized in Table 4.2. The CMOS BPDA reported in this work achieves a higher fBW, while consuming the lowest silicon area, in comparison to prior implementation in both bipolar and BiCMOS technologies.

4.6 Conclusions

Design paradigms for LPDA were extended for bandpass signal amplification. The limitations of the canonical BPDA were discussed. A technique to design small form-factor bandpass distributed amplifiers using dual mirror-symmetric Norton transformations was described in this chapter. In a 50-ohm input/output environment, the prototype test chip demonstrated a peak-gain of 7dB, an in-band ripple of 2dB, and finally a 3dB bandwidth greater than 77%, while consuming a core area of 0.15mm² only.

Ultra-wideband communication systems provide a steady stream of opportunities for single-chip mm-wave electronics. Potential applications for the BPDA described in this chapter include automotive-radar systems (22-29 GHz), phase-array systems for satellite communication in the K_a (26.5 GHz) and Q-band (30-50 GHz). The 77% fBW CMOS BPDA covers both the K_a-band and the Q-band.

- [1] W. S. Percival, "Thermionic valve circuits," British Patent 460 562, Jan. 25, 1937.
- [2] E. L. Ginzton, W. R. Hewlett, J. H. Jasberg, and J. D. Noe, "Distributed amplification," Proc. IRE, vol. 36, pp. 956-969, Aug. 1948.
- [3] B. M. Ballweber, R. Gupta, and D. J. Allstot, "A fully integrated 0.5-5.5-GHz CMOS distributed amplifier," in *IEEE Journal of Solid-State Circuits*, vol. 35, no.2, pp.231-239, 2000.
- [4] H. Shigematsu, M. Sato, T. Hirose, F. Brewer, and M. Rodwell, "40Gb/s CMOS distributed amplifier for fiber-optic communication systems," in *IEEE ISSCC*, 2004.
- [5] Arbabian, A. M. Niknejad, "A three-stage cascaded distributed amplifier with GBW exceeding 1.5THz," in *IEEE RFIC*, pp.211-214, 2012.
- [6] K. Moez, M. Elmasry, "A 10dB 44GHz loss-compensated CMOS distributed amplifier," in *IEEE ISSCC*, pp.548-549, 2007.
- [7] G. L. Matthaei, L. Young, E. M. T. Jones, "Microwave filters, impedance-matching networks, and coupling structures," (Dedham, Mass.: Artech House, 1980.
- [8] T. Gathman et.al., "A Ka-band High Pass Distributed Amplifiers in 130nm SiGe BiCMOS" in *Microwave Symposium Digest*, pp.952-955, 2010.
- [9] H. Wang et.al., "A 5.2-to-13GHz Class-AB power amplifier with a 25.2dBm peak output power at 21.6% PAE" in *ISSCC Dig. Tech. Papers*, pp.44-45, 2010.
- [10] M. El-Nozahi et.al. "A millimeter-wave (23-32GHz) wideband BiCMOS low-noise amplifier", in *IEEE JSSC*, vol.45, no.2, pp.289-299, 2010.
- [11] G. Liu, et.al., "Broadband millimeter-wave LNA (47-77 GHz and 70-140 GHz) using a T-type matching topology", in *IEEE JSSC*, vol.48, no.9, 2022-2029, 2013.

5 ULTRA-WIDEBAND MILLIMETER-WAVE RECEIVER

The high-frequency performance of integrated circuits is primarily limited by the intrinsic and extrinsic parasitic capacitances of the amplifying device. As the frequency of interest increases, the impedance presented by the parasitic capacitance reduces and eventually short-circuits the entire signal current to ground. However, completely eliminating these capacitances is not feasible because they are inherent to the device operation (in CMOS and bipolar transistors). Therefore, to extend the frequency-range of integrated-circuits it is common to use a single shunt inductor to ‘tune’ the amplifier to operate at the desired center frequency. An amplifier, with a parasitic capacitor C , can be loaded with an inductor L such that $L||C$ appears as a perfect open circuit at a tuned frequency. The resulting second-order L-C based load has been extensively used for **narrow-band** tuned amplifier design at both RF and mm-wave frequency bands. In this chapter, the properties of second-order networks which restrict efficient wideband circuit design are explored. It will be shown that higher-order load networks, formed using multiple reactive elements, can overcome the limitations of second-order load networks. To avoid the area-penalty of additional inductors, techniques to synthesize higher-order loads by coupling two resonant tanks are analyzed. This analysis leads to a ‘gain-equalized’ transformer load, in which electrical **and** magnetic coupling is introduced between two resonant tanks to achieve wideband operation. Chapter-4 describes the design of a 77% fBW single-ended BPDA for wideband signal amplification using a third-order LC load. In contrast, in this chapter we will use coupled resonant-loads to design a complete ultra-wideband differential heterodyne receiver.

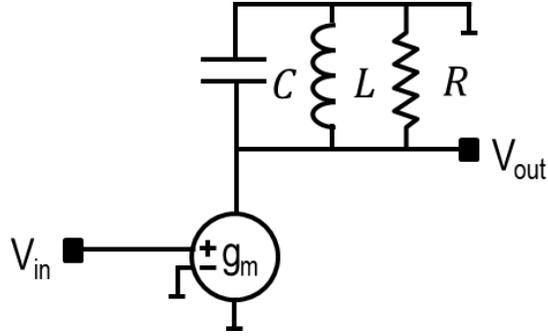


Fig. 5.1. Ideal amplifier with second-order R-L-C tank

5.1 Second-order resonant tanks

First, a brief review of the important design constraints in an ideal amplifier loaded with a parallel R-L-C load are presented. An ideal amplifier is defined as a transconductor with infinite output-impedance and zero output capacitance. Next, the impact of finite-Q inductor, driver-amplifier output capacitance, and load-amplifier amplifier input capacitance are included into the model.

Case-1: Amplifier with no output capacitance, and inductor with infinite quality-factor

With an ideal trans-conductor driving an R-L-C tank, shown in Fig. 5.1, the input-output transfer-function $H(s)$ has the canonical-form of a second-order bandpass filter is,

$$H(s) = \frac{V_{out}(s)}{V_{in}(s)} = g_m Z(s) = g_m R \frac{s \frac{1}{RC}}{s^2 + s \frac{1}{RC} + \frac{1}{LC}} = A_v \frac{\frac{\omega_c}{Q} s}{s^2 + \frac{\omega_c}{Q} s + \omega_n^2} \quad (5-1)$$

In this dimensionless transfer function ω_c and Q are called the natural ‘resonant’ frequency and quality-factor, respectively.

$$\omega_c = 1/\sqrt{LC} \quad (5-2)$$

$$Q = R\sqrt{L/C} \quad (5-3)$$

From (5-1) it can be shown that load impedance $Z(s)$ is a function of the frequency, and

$$Z(s) = \frac{R}{\sqrt{2}} @ s = \omega_c \pm \frac{1}{2RC} \quad (5-4)$$

Thus, for an R-L-C load driven by an ideal voltage-controlled-current-source (*vccs*) with a transconductance g_m , the two key amplifier specifications - the peak voltage-gain at ω_c and 3-dB bandwidth - are,

$$A_v = g_m R \quad (5-5)$$

$$f_{3dB} = \left(\omega_c + \frac{1}{2RC} \right) - \left(\omega_c - \frac{1}{2RC} \right) = \frac{1}{RC} = \frac{\omega_c}{Q} \quad (5-6)$$

Two important results emerge from (5-5) and (5-6). First, from (5-5), for a fixed g_m the only way to increase the peak gain A_v is by increasing R. However, the increase in R results in an increase in the Q, and therefore, from (5-6), a reduction in the f_{3dB} .

Second, for the ideal circuit considered in Case-1, the gain-bandwidth product is equal to (g_m / C) . Thus, for a fixed ω_n and L, the gain-bandwidth product can be increased *indefinitely* by increasing g_m of the amplifying device. Unfortunately, this result does not hold true in real tuned amplifiers. The second-order effects that must be considered to understand the limitation of the L-C load are described next.

Case-2: Amplifier with output capacitance C_D , load capacitance C_L , and inductor with quality-factor Q_L

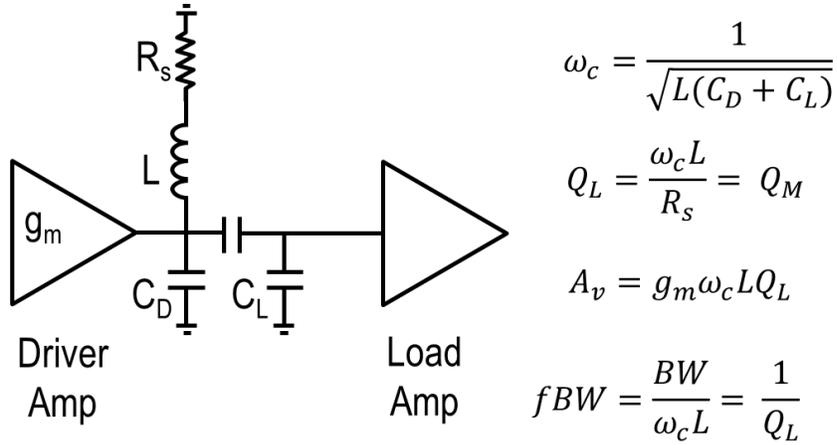


Fig. 5.2. Cascaded amplifier with R-L-C load

In the above analysis the capacitance C and the resistance R are considered independent parameters. However, in integrated tuned amplifiers (a) R is a function of the finite quality-factor of the inductor (b) C depends on the output capacitance C_D and input capacitance C_L of the driver and load amplifiers.

$$C = C_D + C_L \quad (5-7)$$

$$R \approx Q_L^2 R_s = \omega_c L Q_L \quad (5-8)$$

For the circuit in Fig. 5.2,

$$\omega_c = \frac{1}{\sqrt{L(C_D + C_L)}} \quad (5-9)$$

$$Q = Q_L \quad (5-10)$$

$$A_{v1} = g_m \omega_c L Q_L \quad (5-11)$$

Now, to study the impact of increasing the amplifier trans-conductance on the gain of the amplifying device, consider the two circuits shown in Fig. 5.3. In Fig. 5.3(a), the amplifier has a

trans-conductance g_m , gain A_{v1} and a fractional bandwidth of $1/Q_L$. In Fig. 5.3(b), the trans-conductance of the driver amplifier is scaled-up by a factor α . In a common-source amplifier this could be implemented by increasing the (W/L) aspect-ratio of the transistor while keeping the bias conditions fixed. The key point here is that the increase in the amplifier trans-conductance is accompanied by a proportional increase in the output capacitance to αC_D . In lieu of the increased output capacitance, to maintain the resonant frequency, the inductor L has to scale down to L' ,

$$L' = \frac{L}{\alpha(1 - \beta) + \beta} \quad (5-12)$$

where $\beta = C_L/(C_L + C_D)$ is the ratio of the load capacitance to the total capacitance. The gain of the amplifier in Fig. 5.3(b) can be shown to be,

$$A_{v2} = \frac{\alpha}{\alpha(1 - \beta) + \beta} g_m \omega_c L Q_L = \frac{\alpha}{\alpha(1 - \beta) + \beta} A_{v1} \quad (5-13)$$

Thus, for a circuit designed to operate at a fixed ω_c and Q_L , the benefit of increasing the trans-conductance of the amplifier is offset by a reduction in L . In other terms, the gain-bandwidth product **cannot** be indefinitely increased by burning more current (higher g_m).

From (5-11) and (5-13), the relative increase in the voltage-gain,

$$\frac{A_{v2}}{A_{v1}} = \frac{\alpha}{\alpha(1 - \beta) + \beta} \quad (5-14)$$

The variation in A_{v2}/A_{v1} as a function of α is plotted in Fig. 5.3. It can be seen that as β reduces, or the driving capacitance becomes a bigger factor of the total load capacitance, there are diminishing returns in terms of increasing the total gain by burning more power in the amplifier.

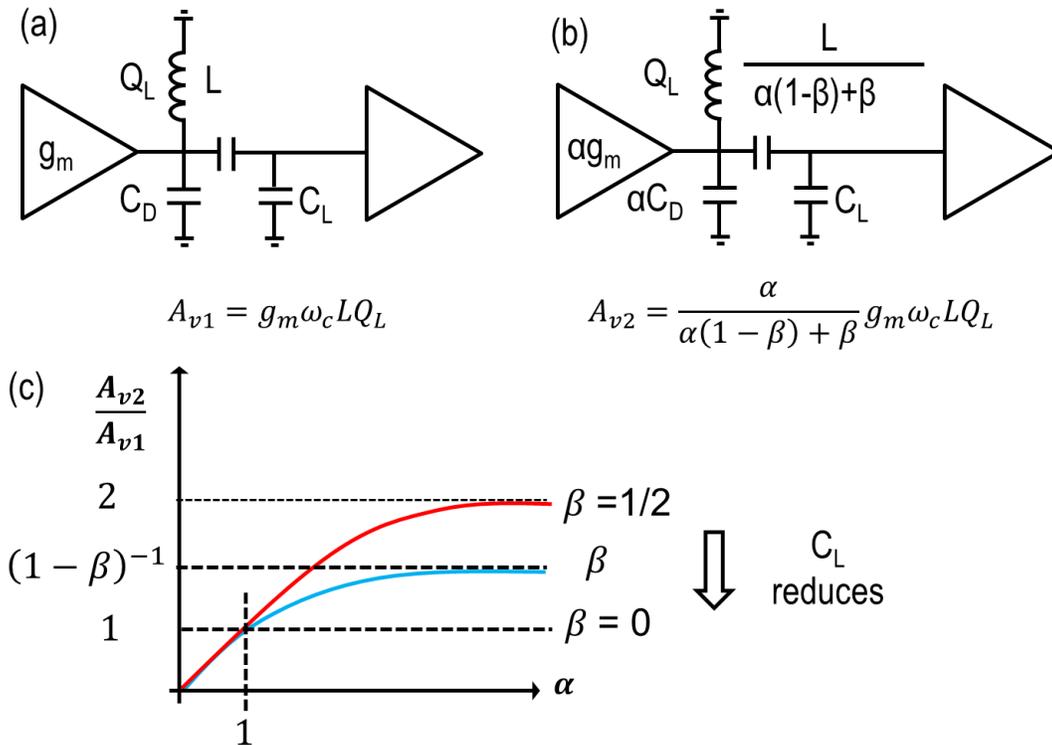


Fig. 5.3. (a) Gain (A_{v1}) of transistor amplifier with transconductance g_m , and R-L-C load (b) Gain (A_{v2}) of transistor amplifier with transconductance αg_m . Value of inductor scaled down to maintain a fixed resonance frequency. (c) Plot of A_{v2}/A_{v1} as a function of scaling-factor α

In summary, a second-order LC load constrains the amplifier to have a single-peak frequency response, resulting in the bandwidth which is tightly coupled to the Q of the load network. Higher-order resonant tanks break the strong dependence of bandwidth on Q by introducing multiple resonant peaks in the frequency response. In contrast to a second-order tank, a higher-order tank can control the location of the resonant peaks, thus enabling bandwidth extension. However, higher-order loads require additional inductors, which not only will incur additional losses, but also increase the chip area. Techniques to generate higher-order resonant loads, without incurring the cost of additional inductors, will be described in the next section.

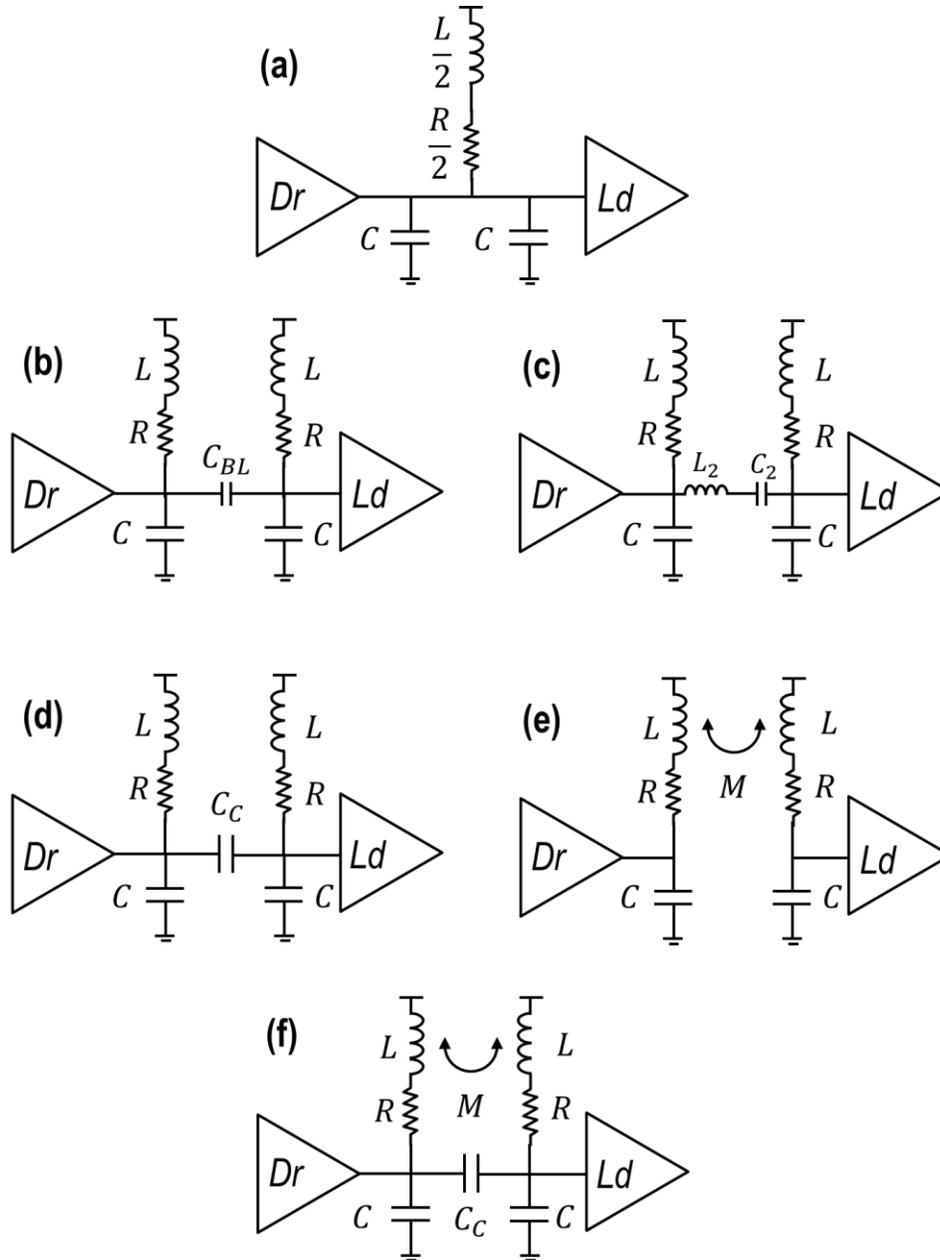


Fig. 5.4 Dr – Driver Amplifier, Ld - Load Amplifier. (a) Single inductor to resonant capacitor $C_D = C$ or, and load capacitor $C_L = C$ (b) Load resonant tanks separated by a large DC-Block capacitor C_{BL} (c) coupled by series resonant tank L_2 and C_2 (d) coupled by series-capacitor C_c (e) coupled by mutual-inductor M (f) coupled by M and C_c (gain-equalized transformer)

5.2 Coupled Resonant tanks

Consider the driver amplifier (Dr) with output capacitance $C_D = C$ and load amplifier (Ld) with input capacitance $C_L = C$, shown in Fig. 5.4. In Fig. 5.4(a), the total capacitance on the

output-node $2C$ is resonated at the operating frequency (ω_c) by a single inductor $L/2$. The inductor ($L/2$) and capacitor ($2C$) form a second-order network; therefore the only way to increase the bandwidth is by reducing the Q of the inductor. Another drawback of this topology is the absence of a DC-block in series with the signal-path. As a result, the DC voltage at the output of Dr (nominally the supply-voltage) has to be equal to the bias-voltage of the Ld stage.

To block the DC-signal a large DC-block capacitor C_{BL} , shown in Fig. 5.4(b), could be used. However, C_{BL} also results in unwanted signal attenuation in the signal-path. To minimize the signal attenuation, C_{BL} must be at least ten times larger than the load-capacitance. This, in turn, compromises the mm-wave circuit performance because the Q and self-resonance frequency of large capacitors.

Fig. 5.4(c) describes an amplifier configuration in which the two resonant-tanks are coupled with a series-resonant circuit comprised of L_2 and C_2 . The resulting load network bears a striking resemblance to the canonical band-pass filter discussed in Chapter-4. Thus, the algorithmic design approach (using the filter-coefficient table such as the one described in Table 4.1) can be applied to synthesize the desired frequency response and select appropriate L and C values. The major drawback of this circuit is the need for the additional inductor L_2 and the resulting area penalty.

It is possible to generate a higher-order frequency response without adding an explicit inductor. Three techniques are described in Fig. 5.4 (d), (e), and (f). The goal in these structures is to approximate the higher-order system of Fig. 5.4 (c) by introducing a coupling mechanism between the resonant tanks. In magnetically-coupled tanks (Fig. 5.4 (d)) the interaction between the magnetic-flux of two closely placed inductors results in a mutual inductance M . In

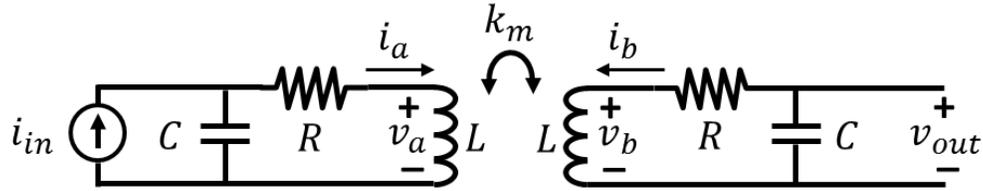


Fig. 5.5 Magnetically coupled resonant tanks

capacitive-coupling, two magnetically-isolated tanks are electrically-coupled with an explicit capacitance C_c . Finally, in Fig. 5.4(f), the load network utilizes both magnetic, and electrical coupling for synthesizing a higher-order response.

In this section, we will describe higher-order loads generated by coupling two second-order systems. We discuss three coupled structures,

- Magnetically-coupled resonant tanks
- Capacitively-coupled resonant tanks
- Magnetically **and** capacitively coupled resonant tanks (gain-equalized)

For each structure, the parameter of interest is the trans-resistance Z , where Z is the ratio of the output-voltage to the input-current.

Magnetically-coupled resonant tanks

The circuit-diagram of two magnetically-coupled resonant tanks is shown in Fig. 5.5. Each resonant tank comprises of an inductor L and capacitor C . The finite quality-factor Q of the inductor is modeled by a series resistance R . The resonant-frequency of the LC-tank is $\omega_n = 1/\sqrt{LC}$. From a node-analysis perspective, this circuit can be modeled using 7 variables $v_a, v_b, v_{out}, v_1, i_a, i_b$ and i_{in} and therefore, requires 6 equations to solve for v_{out}/i_{in} .

By applying KVL and KCL to the circuit in Fig. 5.5, equations (5-15) to (5-20) can be derived,

$$v_a = i_a sL + i_b sk_mL \quad (5-15)$$

$$v_b = i_a sk_mL + i_b sL \quad (5-16)$$

$$v_{out} = -i_b \frac{1}{sC} \quad (5-17)$$

$$v_{out} = v_b + i_b R \quad (5-18)$$

$$i_{in} = i_a + v_1 sC \quad (5-19)$$

$$v_1 = v_a + i_a R \quad (5-20)$$

The detailed description of the steps to derive the trans-resistance is provided in Appendix-5.1. While the sequence of steps to solve the series of linear-equations is cumbersome, the final solution is remarkably simple.

$$Z(s) = \frac{v_{out}}{i_{in}} = \frac{sk_mL}{\{1 + sCR + s^2LC(1 - k_m)\}\{1 + sCR + s^2LC(1 + k_m)\}} \quad (5-21)$$

Equation (5-21) can also be expressed in the canonical form,

$$Z(s) = sLk_m \left\{ \frac{Q\omega_n \frac{\omega_{n1}}{Q_1}}{\omega_{n1}^2 + s\frac{\omega_{n1}}{Q_1} + s^2} \right\} \left\{ \frac{Q\omega_n \frac{\omega_{n2}}{Q_2}}{\omega_{n1}^2 + s\frac{\omega_{n1}}{Q_1} + s^2} \right\} \quad (5-22)$$

Where,

$$\omega_n = 1/\sqrt{LC} \quad (5-23)$$

$$\omega_{n1} = 1/\sqrt{L(1 - k_m)C} \quad (5-24)$$

$$\omega_{n2} = 1/\sqrt{L(1 + k_m)C} \quad (5-25)$$

$$Q = \frac{1}{R} \sqrt{\frac{L}{C}} \quad (5-26)$$

$$Q_1 = \frac{1}{R} \sqrt{\frac{L(1 - k_m)}{C}} \quad (5-27)$$

$$Q_2 = \frac{1}{R} \sqrt{\frac{L(1 + k_m)}{C}} \quad (5-28)$$

Thus, magnetically-coupling two second-order tank with a self-resonant frequency ω_n , and quality-factor Q , results in a higher-order load which exhibits a trans-resistance with two natural resonant frequencies, ω_{n1} and ω_{n2} . From (5-22) and (5-24), the trans-resistance of the circuit at $\omega = \omega_{n1}$ can be expressed as,

$$\begin{aligned} |Z(j\omega)_{\omega=\omega_{n1}}| &= \frac{j\omega_{n1}k_mL}{\{1 + j\omega_{n1}CR - \omega_{n1}^2LC(1 - k_m)\}\{1 + j\omega_{n1}CR - \omega_{n1}^2LC(1 + k_m)\}} \\ &= \frac{k_mL}{CR\{1 + j\omega_{n1}CR - \omega_{n1}^2LC(1 + k_m)\}} \\ &= \frac{k_mL}{CR \left\{ 1 + j \frac{\sqrt{C}R}{\sqrt{L(1 - k_m)}} - \frac{1 + k_m}{1 - k_m} \right\}} \\ &= \frac{k_mL}{CR \left\{ j \frac{1}{Q} - \frac{2k_m}{1 - k_m} \right\}} \end{aligned}$$

Under the assumption that $Q > 1$, $|Z(j\omega)_{\omega=\omega_{n1}}|$ can be further simplified to,

$$|Z(j\omega)_{\omega=\omega_{n1}}| \approx \frac{k_m L}{CR \left\{ -\frac{2k_m}{1-k_m} \right\}}$$

$$\Rightarrow |Z(j\omega)_{\omega=\omega_{n1}}| = \frac{1}{2} \frac{L(1-k_m)}{CR} \quad (5-29)$$

Similarly, it can be shown that,

$$|Z(j\omega)_{\omega=\omega_{n2}}| = \frac{1}{2} \frac{L(1+k_m)}{CR} \quad (5-30)$$

While it has been mathematically proven that the trans-resistance of magnetically-coupled tanks at the natural resonance frequencies are given by (5-29) and (5-30), a more intuitive interpretation follows by considering the terms Q_1 and Q_2 defined in (5-26) and (5-27). Due to the mutual-inductance $M (=k_m L)$ induced by magnetic-coupling, the effective-inductance of the load increases from L to $L + M = L(1 + k_m)$. Therefore, assuming a fixed series loss-resistance R , the quality-factor of the load has increased. At ω_{n1} , it is straightforward to prove the effective quality-factor is given by,

$$|Q_{LC}|_{\omega=\omega_{n1}} = \frac{\omega_{n1} L(1+k_m)}{R} = \frac{1}{R} \sqrt{\frac{L(1-k_m)}{C}} = Q_1 \quad (5-31)$$

The effective parallel load resistance due to a single LC tank at the frequency ω_{n1} can be computed using the standard series-to-parallel impedance transformation,

$$R_{LC} = (1 + Q_1^2)R \approx Q_1^2 R \quad (5-32)$$

Finally, since there are two LC-resonant tanks, the effective trans-resistance is the parallel-combination of two resistance of value R_{LC} to obtain,

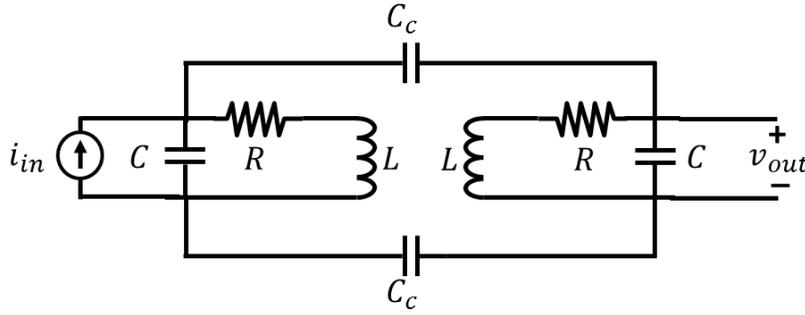


Fig. 5.6 Electrically-coupled resonant tanks

$$|Z(j\omega)_{\omega=\omega_{n1}}| = \frac{1}{2} Q_1^2 R \quad (5-33)$$

$$|Z(j\omega)_{\omega=\omega_{n2}}| = \frac{1}{2} Q_2^2 R \quad (5-34)$$

The results obtained by this more intuitive approach, (5-33) and (5-34), agree well with the more rigorous mathematical results derived in (5-29) and (5-30).

Peak splitting in magnetically-coupled resonant-tanks is achieved by increasing the mutual-inductance or magnetic-coupling coefficient (k_m) between the resonant-tanks. However, from (5-29) and (5-30), $|Z(j\omega)_{\omega=\omega_{n1}}|$ and $|Z(j\omega)_{\omega=\omega_{n2}}|$ are equal if and only if $k_m = 0$. Thus, *peak-splitting* based solely on magnetic-coupling exhibits an inherent amplitude mismatch.

Electrically-coupled resonant tanks

In the previous section, the trans-resistance of magnetically coupled resonant tanks was computed by solving KVL and KCL equations. While the nodal-analysis approach can also be applied on the electrically-coupled tanks, (shown in Fig. 5.6) for this circuit the analysis can be simplified by using two-port parameters instead. For a two-port network characterized by $Y = [y_{11} \ y_{12} ; y_{21} \ y_{22}]$, the trans-resistance is given by,

$$Z(s) = \frac{y_{21}}{y_{11}y_{22} - y_{12}y_{21}} \quad (5-35)$$

The Y-parameters of a network shown in Fig. 5.6, two parallel LC resonant tanks coupled with a capacitance C_c , can be shown to be,

$$y_{11} = \left. \frac{I_1}{V_1} \right|_{V_2=0} = s \left(\frac{C_c}{2} + C \right) + \frac{1}{sL + R} \quad (5-36)$$

$$y_{22} = \left. \frac{I_2}{V_2} \right|_{V_1=0} = s \left(\frac{C_c}{2} + C \right) + \frac{1}{sL + R} \quad (5-37)$$

$$y_{12} = \left. \frac{I_1}{V_2} \right|_{V_1=0} = -s \frac{C_c}{2} \quad (5-38)$$

$$y_{21} = \left. \frac{I_2}{V_1} \right|_{V_2=0} = -s \frac{C_c}{2} \quad (5-39)$$

Using, equations (5-35) to (5-39), the trans-resistance of the electrically-coupled resonant tanks can be shown to be,

$$\begin{aligned} Z(s) &= \frac{-s \frac{C_c}{2}}{\left(s \left(\frac{C_c}{2} + C \right) + \frac{1}{sL + R} \right)^2 - \left(s \frac{C_c}{2} \right)^2} \\ &= \frac{-s \frac{C_c}{2}}{\left\{ s(C_c + C) + \frac{1}{sL + R} \right\} \left\{ sC_c + \frac{1}{sL + R} \right\}} \\ &= \frac{1}{2} \frac{-sC_c(sL + R)^2}{\{s^2L(C_c + C) + sR(C_c + C) + 1\} \{s^2LC + sRC + 1\}} \end{aligned} \quad (5-40)$$

Equation (5-40) can also be expressed in the canonical form,

$$Z(s) = \frac{1}{2} \omega_{n1}^2 \omega_{n2}^2 \frac{-sC_c(sL + R)^2}{\left\{s^2 + s\frac{\omega_{n1}}{Q_1} + \omega_{n1}^2\right\} \left\{s^2 + s\frac{\omega_{n2}}{Q_2} + \omega_{n2}^2\right\}} \quad (5-41)$$

Thus, the electrically-coupled tank exhibits a trans-resistance with two natural resonant frequencies, ω_{n1} and ω_{n2} , and associated quality-factors Q_1 and Q_2 , where

$$\omega_{n1} = 1/\sqrt{LC} \quad (5-42)$$

$$\omega_{n2} = 1/\sqrt{L(C + C_c)} \quad (5-43)$$

$$Q_1 = \frac{1}{R} \sqrt{\frac{L}{C}} \quad (5-44)$$

$$Q_2 = \frac{1}{R} \sqrt{\frac{L}{C + C_c}} \quad (5-45)$$

The spacing between the two resonant-peaks can be increased by increasing the value of the coupling capacitance C_c . In contrast to the magnetically coupled tank which changes the location of both the poles ω_{n1} and ω_{n2} , electric-coupling only moves the location of ω_{n2} , while ω_{n1} remains fixed.

Starting from (5-40), and using (5-42) and (5-45), the trans-resistance at ω_{n2} can be shown to be,

$$\begin{aligned} |Z(j\omega)_{\omega=\omega_{n2}}| &= \frac{1}{2} \frac{-j\omega_{n2}C_c(j\omega_{n2}L + R)^2}{\{-\omega_{n2}^2L(C_c + C) + j\omega_{n2}R(C_c + C) + 1\} \{-\omega_{n2}^2LC + j\omega_{n2}RC + 1\}} \\ &= \frac{1}{2} \frac{-j\frac{C_c}{\sqrt{L(C + C_c)}} R^2 \left(1 + j\sqrt{\frac{L}{(C + C_c)}} \frac{1}{R}\right)^2}{jR\sqrt{\frac{C_c + C}{L}} * \frac{C_c}{C_c + C} \left\{1 + j\frac{C}{C_c} R\sqrt{\frac{C_c + C}{L}}\right\}} \end{aligned}$$

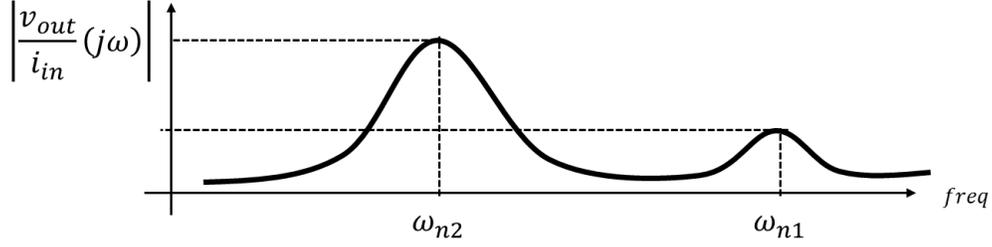


Fig. 5.7 Trans-resistance of a magnetically-coupled (MC), and electrically-coupled (CC) resonant tanks

$$= \frac{1 - R(1 + jQ_2)^2}{2 \left\{ 1 + j \frac{1}{Q_2} \right\}}$$

$$\Rightarrow |Z(j\omega)_{\omega=\omega_{n2}}| \approx \frac{1}{2} Q_2^2 R \quad (5-46)$$

It is interesting to note that in terms of the parameter Q_2 , the trans-resistance of the electrically-coupled tank in (5-46) is identical to the magnetically coupled tank. Substituting (5-45) in (5-46) the trans-resistance at the natural resonant frequency ω_{n2} can be expressed as,

$$|Z(j\omega)_{\omega=\omega_{n2}}| = \frac{1}{2} \frac{1}{R} \frac{L}{C + C_c} \quad (5-47)$$

Similarly it can be shown that,

$$|Z(j\omega)_{\omega=\omega_{n1}}| = \frac{1}{2} Q_1^2 R = \frac{1}{2} \frac{1}{R} \frac{L}{C} \quad (5-48)$$

From, (5-42) and (5-43) it can be seen that the spacing between the resonant peak frequency is a function of the coupling capacitance C_c . As C_c increases, the frequency ω_{n1} remains fixed; however ω_{n2} reduces. However, (5-47) and (5-48) indicate that equal trans-resistance at ω_{n1} and ω_{n2} can be achieved only when $C_c = 0$. Thus, similar to magnetically-coupled resonant-tanks, *peak-splitting* based solely on electric-coupling exhibits an inherent amplitude mismatch.

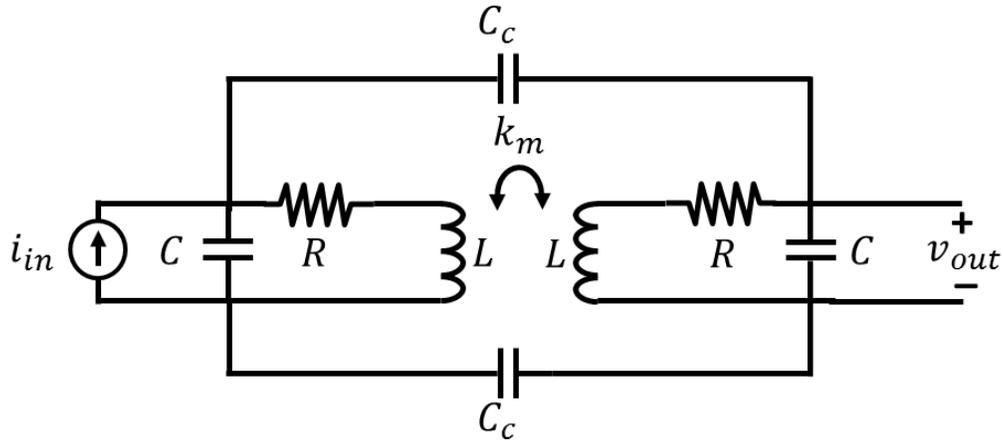


Fig. 5.8 Electric and magnetically coupled resonant tank

The fundamental limitation of magnetically-coupled and electrically-coupled is that in both cases the trans-resistance and the resonant-peak frequencies are controlled by a single design variable. To obtain equal the trans-resistance at the two resonant-peaks frequencies a *gain-equalized* transformer is described in the next section.

Electrically and magnetically coupled resonant tanks

As described previously, the root cause for the inherent amplitude mismatch in magnetically/electrically-coupled resonant tanks is the dependence on a single coupling variable; k_m (magnetically-coupling coefficient) or k_c (electrical-coupling coefficient) where

$$0 < \left\{ k_c = \frac{C_c}{C_c + C_2} \right\} < 1 \quad (5-49)$$

To overcome this limitation, the authors of [5] explored ways to to introduce both electrical **and** magnetic coupling between the resonant tanks. The resulting ‘gain-equalized transformer’ has two independent design parameters, k_m and k_c . A formal analysis of how to select k_m and k_c to equalize the trans-resistance at the resonant-frequencies will be described

next. Similar to the EC-resonant tank, Y-parameters will be used to compute the trans-resistance of the gain-equalized transformer. The Y-parameters for the circuit shown in Fig. 5.8 are,

$$y_{11} = y_{22} = s \left(\frac{C_c}{2} + C \right) + \frac{sL + R}{\{sL(1 - k_m) + R\}\{sL(1 + k_m) + R\}} \quad (5-50)$$

$$y_{12} = y_{21} = - \left(s \frac{C_c}{2} + \frac{sLk_m}{\{sL(1 - k_m) + R\}\{sL(1 + k_m) + R\}} \right) \quad (5-51)$$

From (5-35), (5-50) and (5-51) the trans-resistance of the gain-equalized transformer is,

$$\begin{aligned} Z_{MC-CC}(s) & \quad (5-52) \\ &= \frac{\left[s \frac{C_c}{2} + \frac{sLk_m}{\{sL(1 - k_m) + R\}\{sL(1 + k_m) + R\}} \right]}{\left[s \left(\frac{C_c}{2} + C \right) + \frac{sL + R}{\{sL(1 - k_m) + R\}\{sL(1 + k_m) + R\}} \right]^2 - \left[s \frac{C_c}{2} + \frac{sLk_m}{\{sL(1 - k_m) + R\}\{sL(1 + k_m) + R\}} \right]^2} \\ &= \frac{s \frac{C_c}{2} \{sL(1 - k_m) + R\}\{sL(1 + k_m) + R\} + sLk_m}{[s^2L(C_c + C)(1 - k_m) + s(C_c + C)R + 1][s^2LC(1 + k_m) + sCR + 1]} \end{aligned}$$

It is clear that due to the increase in the design variables, the expressions for trans-resistance are increasingly complicated. A good method to verify the validity of (5-52) is to analyze the expression at the two boundary conditions,

$$Z_{MC-CC}(s) \xrightarrow{C_c=0} \frac{sLk_m}{\{s^2LC(1 - k_m) + sCR + 1\}\{s^2LC(1 + k_m) + sCR + 1\}} \quad (5-53)$$

$$Z_{MC-CC}(s) \xrightarrow{k_m=0} \frac{s \frac{C_c}{2} \{sL(1 - k_m) + R\}\{sL(1 + k_m) + R\}}{\{s^2L(C_c + C) + s(C_c + C)R + 1\}\{s^2LC + sCR + 1\}} \quad (5-54)$$

Thus, in the limiting case (5-53) and (5-54) simplify to the expression for $Z_{CC}(s)$ and $Z_{MC}(s)$ given by (5-40) and (5-21), respectively. Equation (5-52) can also be expressed in the canonical form,

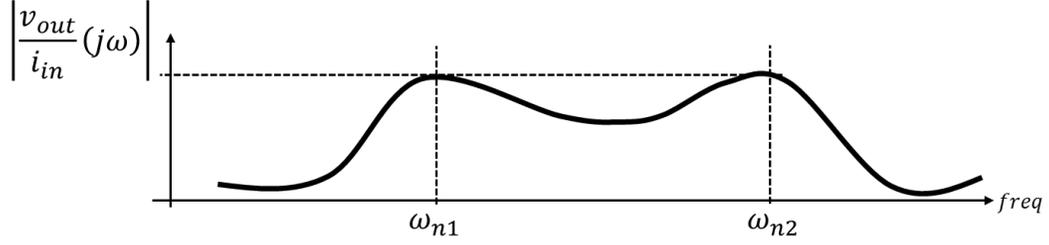


Fig. 5.9. Trans-resistance of a gain-equalized transformer

$$Z_{MC-CC}(s) = \omega_{n1}^2 \omega_{n2}^2 \frac{s \frac{C_c}{2} \{sL(1 - k_m) + R\} \{sL(1 + k_m) + R\} + sLk_m}{\left[s^2 + s \frac{\omega_{n1}}{Q_1} + \omega_{n1}^2 \right] \left[s^2 + s \frac{\omega_{n2}}{Q_2} + \omega_{n2}^2 \right]} \quad (5-55)$$

where,

$$\omega_{n1} = \frac{1}{\sqrt{LC}} \frac{1}{\sqrt{1 + k_m}} \quad (5-56)$$

$$\omega_{n2} = \frac{1}{\sqrt{L(1 - k_m)(C + C_c)}} = \frac{1}{\sqrt{LC}} \sqrt{\frac{1 - k_c}{1 - k_m}} \quad (5-57)$$

$$Q_1 = \frac{1}{R} \sqrt{\frac{L(1 + k_m)}{C}} \quad (5-58)$$

$$Q_2 = \frac{1}{R} \sqrt{\frac{L(1 - k_m)}{C + C_c}} \quad (5-59)$$

Finally, using an analysis similar to electrically, magnetically-coupled tanks, the trans-resistance at the natural resonant frequencies ω_{n1} and ω_{n2} can be shown to be,

$$|Z(j\omega)_{\omega=\omega_{n1}}| = \frac{1}{2} Q_1^2 R = \frac{1}{2} \frac{1}{R} \frac{L(1 - k_m)}{C + C_c} \quad (5-60)$$

$$|Z(j\omega)_{\omega=\omega_{n2}}| = \frac{1}{2} Q_2^2 R = \frac{1}{2} \frac{1}{R} \frac{L(1 + k_m)}{C} \quad (5-61)$$

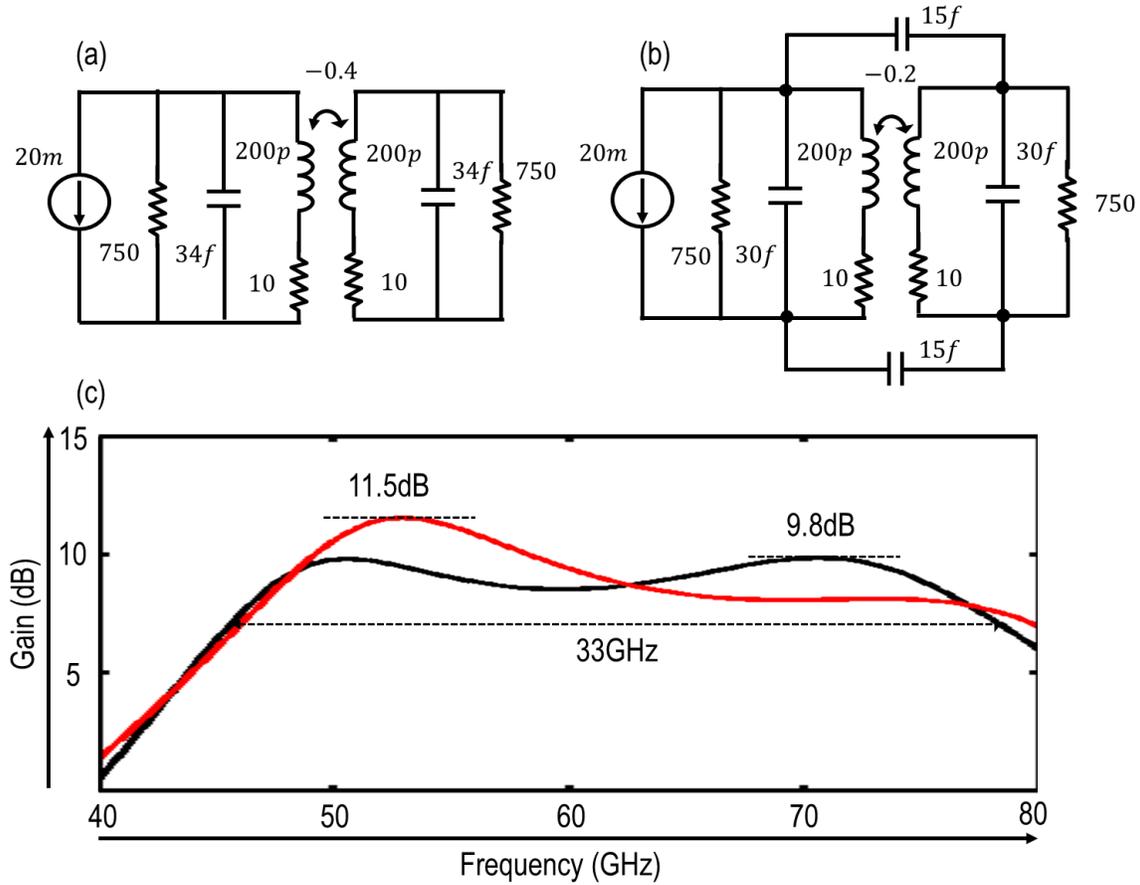


Fig. 5.10. (a) Schematic of magnetically-coupled (MC) transformer (b) Schematic of gain-equalized (MC-CC) transformer (c) Comparison of the trans-resistance of circuit-a and circuit-b

From, (5-65) and (5-72), it can be shown that for a gain-equalized transformer the upper- and lower resonant peaks will be equal if $k_c = -2k_{m2}/(1 - k_{m2})$.

$$|Z(j\omega)_{\omega=\omega_{n1}}| = |Z(j\omega)_{\omega=\omega_{n2}}|$$

$$\Rightarrow \frac{L(1 - k_{m2})}{(C_2 + C_c)R_s} = \frac{L(1 + k_{m2})}{C_2R_s} \quad (5-62)$$

$$\Rightarrow \frac{L(1 - k_{m2})}{\left[\frac{C_2}{1 - k_c}\right]R_s} = \frac{L(1 + k_{m2})}{C_2R_s} \quad (5-63)$$

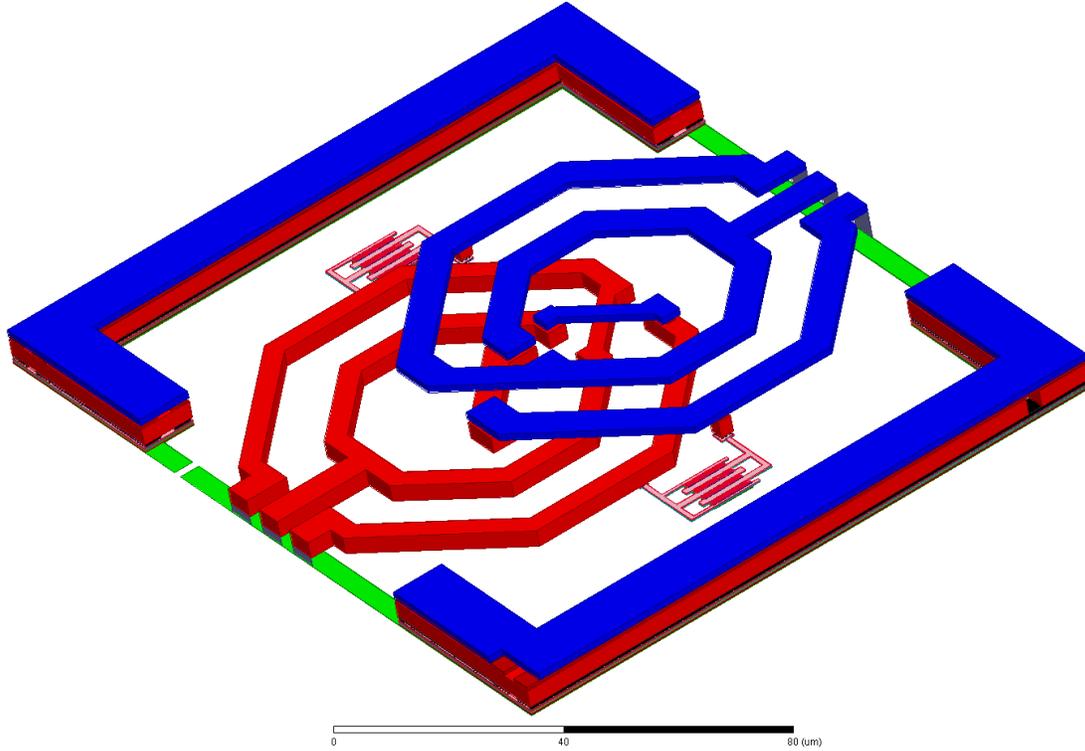


Fig. 5.11 HFSS model of the gain-equalized transformer

$$\Rightarrow 1 - k_c = \frac{1 + k_{m2}}{1 - k_{m2}} \quad (5-64)$$

$$\Rightarrow k_c = \frac{-2k_{m2}}{1 - k_{m2}} \quad (5-65)$$

An important point to note is that since k_c is bound between 0 and 1, equation (5-65) can only be satisfied when k_m is negative. In other terms, designers must be aware of the magnetic flux polarity when deciding how to connect the cross-coupling capacitor.

To illustrate the advantages of the gain-equalized transformer over a standard magnetically coupled transformer, consider the circuits shown in Fig. 5.9(a) and (b). In both the circuits, the value and quality-factor of the inductor is the same. Moreover, the values of k_c

and k_m in each circuit (for circuit-a $k_m = -0.4$ and $k_c = 0$, and for circuit-b $k_m = -0.2$ and $k_c = 1/3$) are selected to ensure the resonant-peak frequencies are equal. Circuit-b achieves a trans-resistance bandwidth of the 33GHz due to equalization of the trans-resistance, Fig. 5.9(c).

Physical Implementation

The gain-equalized transformers were characterized using a HFSS, a 3D electromagnetic (EM) simulator. The EM model used for the simulation is shown in Fig. 5.11. The primary and secondary windings of the transformer are designed in the two top metal layers. The mutual magnetic coupling between the windings of the lateral transformer is controlled by varying the degree of overlap. The cross-coupling capacitors are implemented using metal-oxide-metal (MOM) capacitors. To minimize the impact of process-variation, minimum sized capacitors have not been used. In addition each capacitor is surrounded by floating dummy capacitances to minimize abrupt metal density variation in the vicinity of the desired cross-coupling capacitors. The cross-coupling capacitors have been included within the EM model to improve the modeling accuracy. However, the dummy-capacitors have not been included to minimize the simulation time.

5.3 Receiver architecture, modeling and design

The ultra-wideband receiver proposed in this work utilizes multiple instances of the gain-equalized transformer to achieve a broadband frequency response. The block-diagram of the 50-67GHz receiver implemented in a 40nm CMOS process is shown in Fig. 5.12. The front-end is comprised of a three stage low-noise amplifier and a mm-wave mixer to down-convert the signal to a 30-to-47 GHz IF frequency. At IF, the signal is amplified by a two-stage IF-amplifier. The IF-amplifier drives an IF mixer which down-converts the signal to baseband. The LO-path

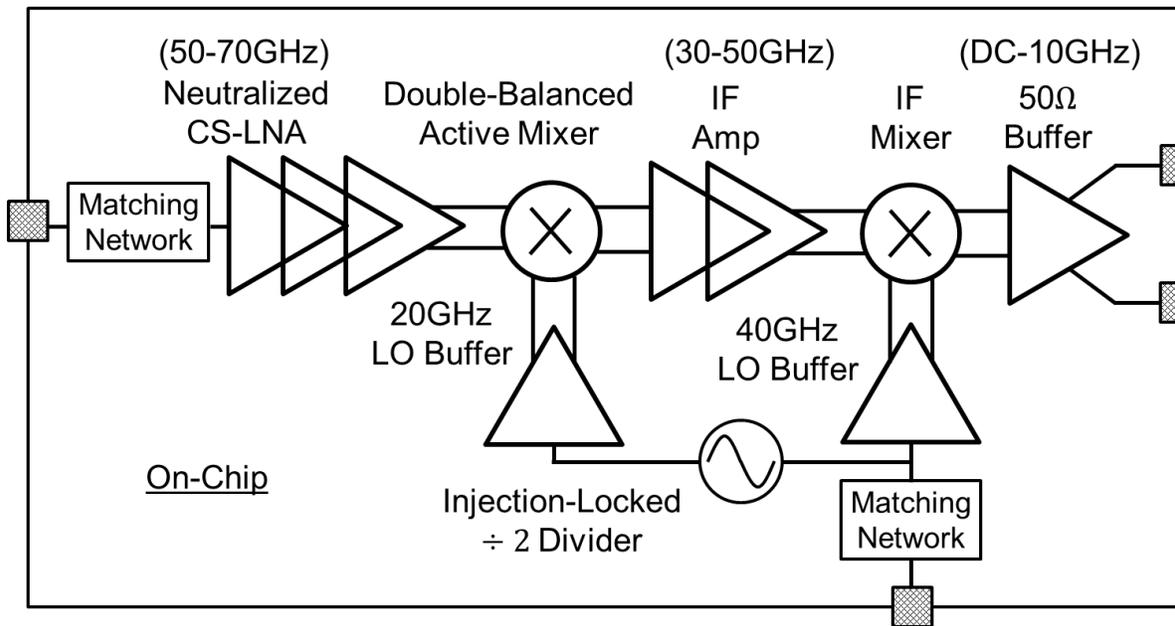


Fig. 5.12 Block diagram of the wideband, heterodyne millimeter-wave receiver

comprises of 20GHz and 40GHz LO-buffers to drive the mm-wave and IF-mixers, respectively. An important point to note is that the LO-frequency to the first mixer ($F_{LO,RF}$) is lower than the LO-frequency to the second mixer ($F_{LO,IF}$). As discussed in Chapter-2, if the receiver-architecture proposed in this work is used as the core element of a phased-array receiver with a large number of elements, the low $F_{LO,RF}$ can potentially reduce the total LO distribution power. In the prototype test-chip, a single-ended external 40GHz LO is routed on-chip via a transformer-based balun/matching network and the 20GHz LO is generated on-chip using an injection-locked divider.

The entire receiver signal and LO paths are fully-differential. A differential receiver is more robust to common-mode on the supply and ground-rails. In addition, the even-order harmonics are inherently suppressed in a well-matched differential circuit. Another big advantage of differential mm-wave circuits is related to EM modeling. In symmetric differential

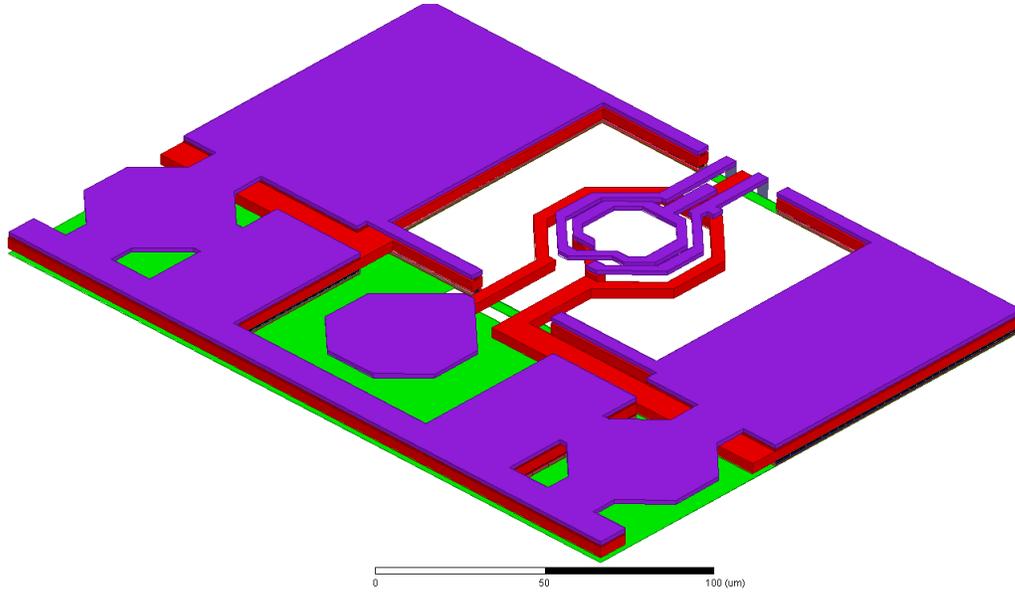


Fig. 5.13. HFSS model of the G-S-G pad and the input balun

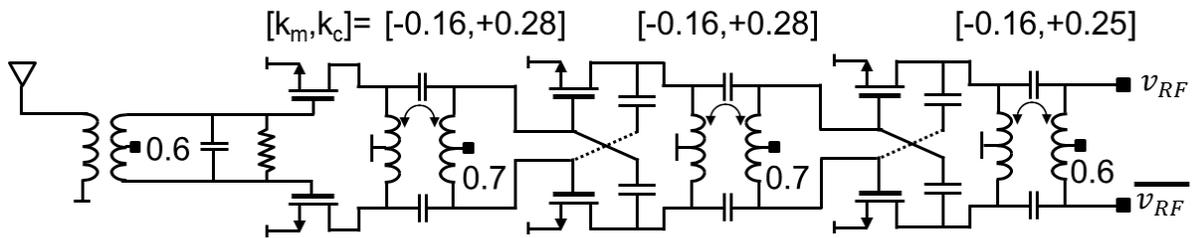


Fig. 5.14. Schematic of the three-stage low-noise amplifier

inductors and transformers it is much easier to model the current-return in comparison to single-ended inductors. Incorrect ground loop modeling can easily introduce errors of 50 to 100pH in the simulated inductance value. At radio-frequencies it is common to use inductors of value exceeding 1 to 2nH, so a 50-100 pH error does not result in severe performance degradation. However, at mm-wave frequencies, for which desired inductors could be on the order of 100-200 pH, it is critically important to minimize modeling inaccuracies.

While the entire circuit is differential, the mm-wave input is single-ended to model the real-scenario in which the LNA must interface with a single-ended antenna. An on-chip

transformer-based balun has been designed for single-ended to differential conversion. The on-chip balun has a turns-ratio of 1:2 and provides a wideband impedance match with 6dB passive voltage-gain. It is interesting to note that at mm-wave frequencies, having the balun on-chip rather than off-chip, simplifies the test setup as well. An off-chip balun would require differential mm-wave cables to bring the signal on-chip. Maintaining a phase and gain balance across two different cables is extremely difficult, if not impossible. Wideband differential *balun-probes*, in which the balun is built within the wafer-probe tip, is still a topic of active research and not commercially available.

The GSG pads and the balun are modeled as a single passive structure in the EM simulator. The HFSS model for the pad-balun combination is shown in Fig. 5.13. This method of modeling ensures that all of the parasitic effects (capacitance between the large probe-pad and the substrate, parasitic inductance within the probe-pad) are captured within the simulation framework.

The schematic for the three-stage low-noise amplifier is shown in Fig. 5.14. The first stage of the LNA is realized as a common-source topology instead of a cascode topology to mitigate the effects of noise figure degradation introduced by the cascode transistor. To improve reverse-isolation, the second and third stages of the LNA employ cross-coupled transistors for C_{GD} neutralization. The LNA is loaded with a gain equalized transformer, where $k_m = -0.16$ and $k_c = 0.28$, while achieving a gain of 18dB with 35mA current consumption.

As the signal progresses down the receiver-path from the mm-wave front-end to the IF, the fBW increases from 29% to 43%. The schematic of the RF-mixer, IF-amplifier and IF-mixer is shown in Fig. 5.15. Assuming the intrinsic parasitic effects of the device are relatively constant

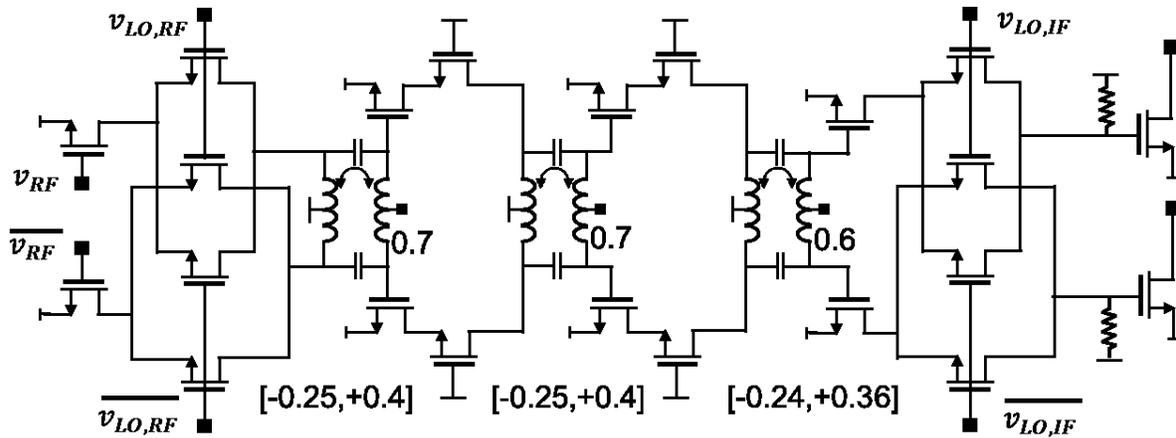


Fig. 5.15 Schematic of RF-mixer, two-stage IF-amplifier and IF-mixer

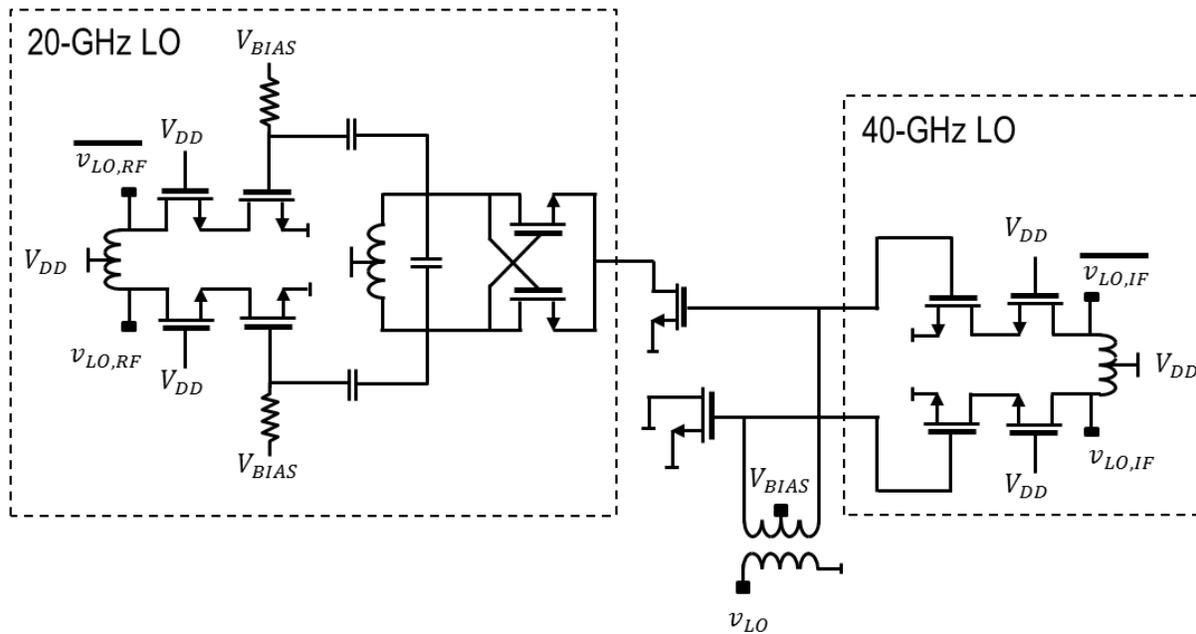


Fig. 5.16. Schematic of the LO distribution network

in the mm-wave and IF frequency bands, it becomes more challenging to generate low-power gain blocks when low-Q components are required to achieve a high fBW. Thus, the burden of the gain is placed at the LNA rather than at IF.

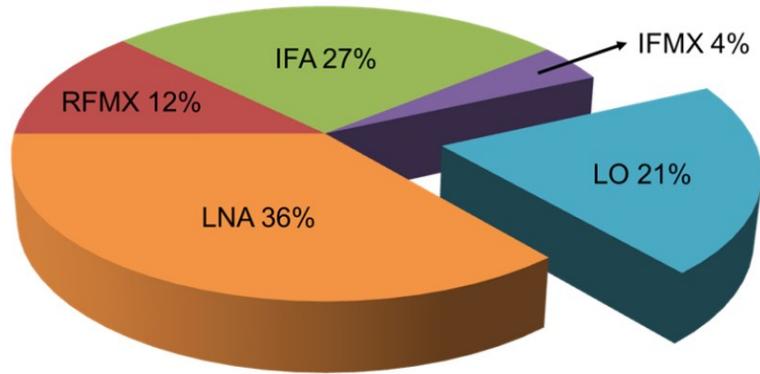


Fig. 5.17 Power break-up for the different circuit blocks

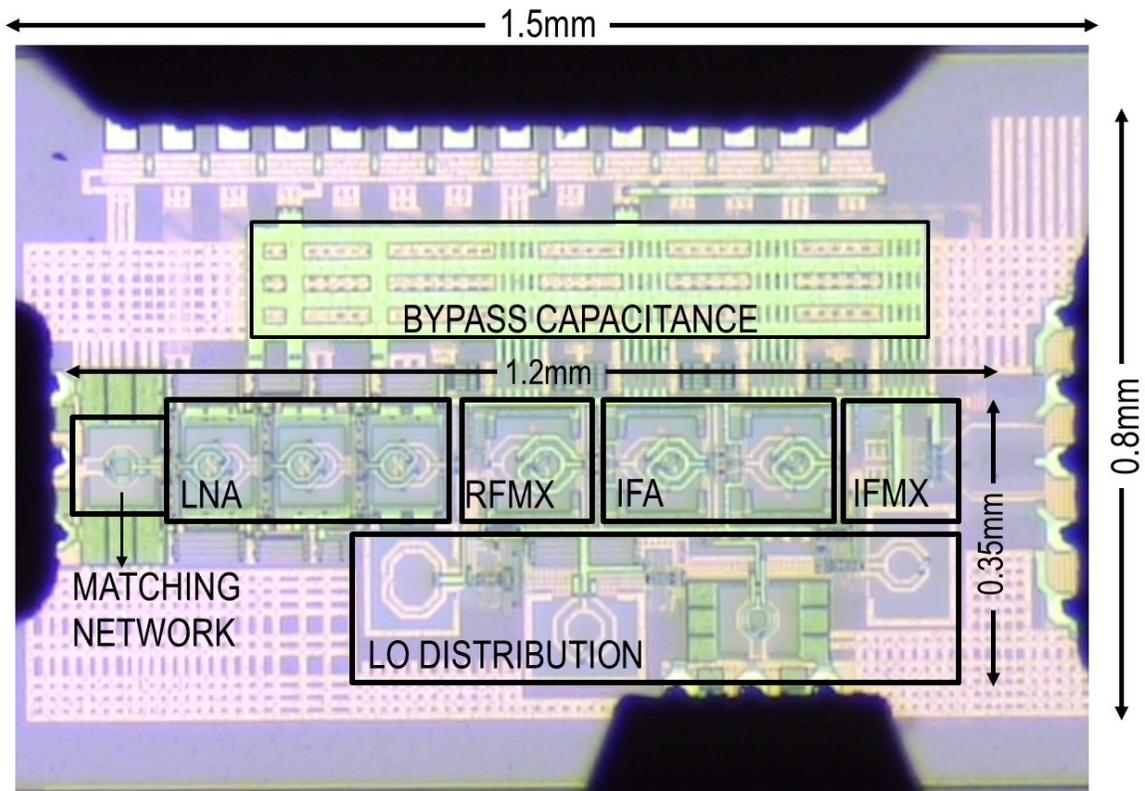


Fig. 5.18. Die micrograph of the millimeter-wave CMOS receiver

The LNA drives a double balanced active mixer, which converts the 50-68 GHz mm-wave signal to an IF of 30-to-47GHz. The local-oscillator signal to the mm-wave signal is 40 GHz. The mixer has a conversion loss of 3dB while consuming 12mA of current. A higher degree of peak-splitting is required to maintain a high fBW through the IF-stage. As a result, the

gain-equalized transformers in the IF stage have coupling coefficients of $k_m = -0.24$, and $k_c = 0.38$. The IFA consists of a two-stage cascode amplifier and provides 5dB of gain to compensate for the loss of the RF-mixer.

The schematic of the LO distribution network is shown in Fig. 5.16. A breakdown of power consumption for the various receiver blocks is shown in Fig. 5.17; note, the percentage power consumption of the LO driver for a single-element receiver is comparable to the power of the IFA. If the receiver is utilized in an N-element phased-array receiver with LO phase-shifting / IF signal-combining, the IFA power is fixed while the LO-power scales up by N^2 . The equal power consumption of the IF-amp and the LO in the single-element receiver, gives an indication that as the number of elements increases the LO power will dominate the total receiver power. Architectures which focus on minimizing LO-distribution power are crucial for future phased-array mm-wave transceivers.

5.4 Measured results

The die photograph of the prototype mm-wave receiver chip fabricated in a TSMC 40nm CMOS process is shown in Fig. 5.18. The BEOL consists of a 6 metal stack with 1 ultra-thick metal (UTM) and 1 aluminum passivation layer (AP). The 3.5um thick UTM is the top-most copper layer 2.3775um above the surface of the substrate. The entire-chip, including the mm-wave wafer-probe pads, ESD protected DC pads, bypass capacitance, and supply/bias routing occupies an area of 0.8mm x 1.5mm. The core of the receiver, comprised of the signal and LO-path, occupies an area of 1.2 mm x 0.35mm.

The chip is characterized using Cascade's 12000AP Summit on-wafer probe station. Fig. 5.18 shows the chip with all four probes landed. The DC-bias is brought on chip using a custom

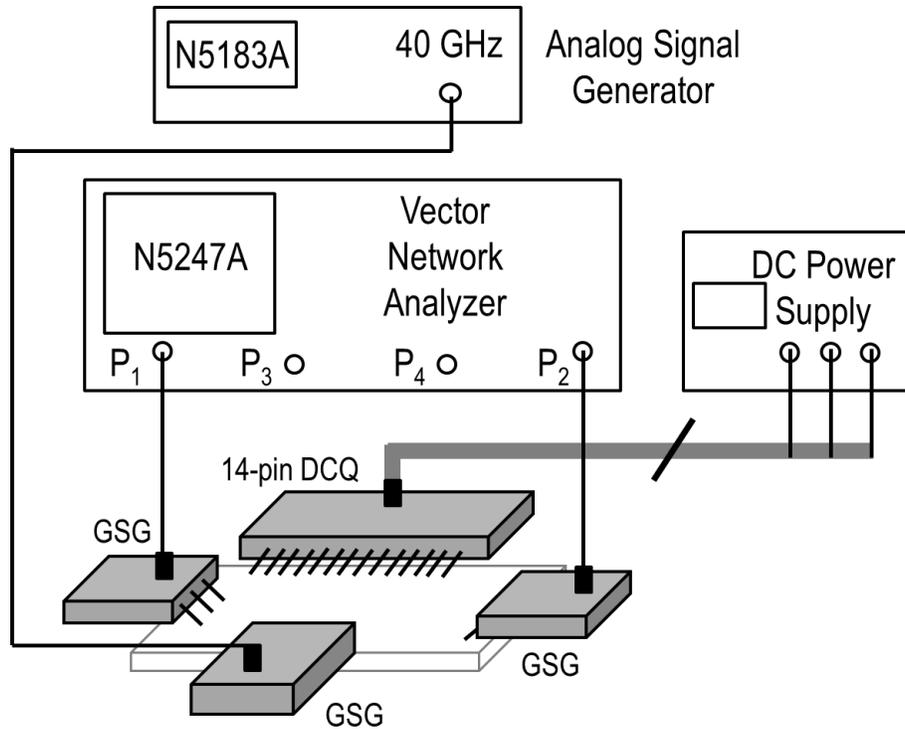


Fig. 5.19 Test setup for conversion-gain and linearity measurements

designed 14-pin DC probe on the north-end. The LO signal is brought on-chip via a 67GHz GSG probe from the south-side. The millimeter-wave input is brought on-chip using a GSG on the west-side. The baseband output drives GSGSG probes on the east side. The GSGSG pad configuration allows the flexibility of making single-ended measurements using GSG, as well as differential measurements using GSGSG. For input-matching, conversion-gain and linearity measurements, which require standard two-port calibration and scalar-mixer calibration on the VNA, the baseband output is measured single-ended. For noise-figure, the output is probed differentially.

The measurement setup is shown in Fig. 5.19. The electrical performance of the receiver was characterized using an Agilent N5247A 67GHz PNA-X network analyzer. The N5183A signal-source was used to generate the 40GHz local oscillator signal. Other equipment required

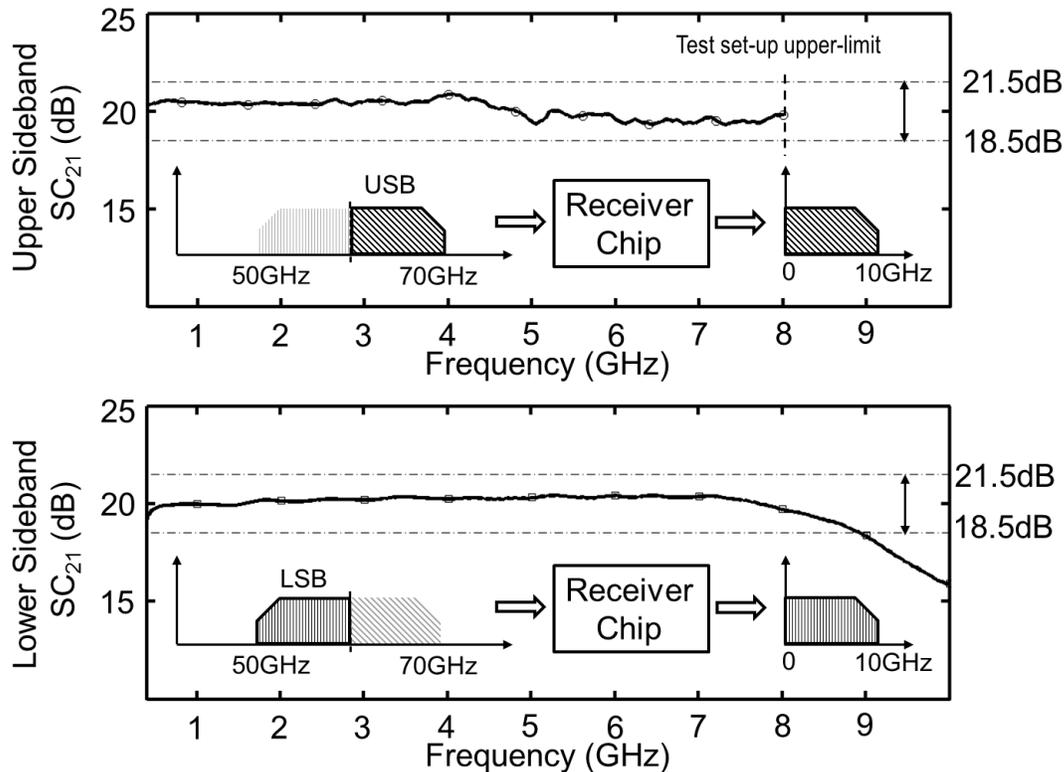


Fig. 5.20 Receiver frequency response measured at the baseband output. Referred to the receiver front-end, a gain of 20 ± 1.5 dB is maintained across a 51-to-68 GHz bandwidth.

to complete the test-setup (not included in the figure for simplicity) include a N8488A 67-GHz average thermocouple power sensor and N1913A power meter. All the equipment was interfaced using 1.85mm co-axial cables. Two-port calibration was performed using an impedance standard substrate (ISS) to shift the S-parameter measurement plane to the probe-tips.

The measured channel frequency response of the entire receiver at the baseband output is shown in Fig. 5.20. For conversion-gain a scalar mixer calibration was performed to de-embed the loss of the cables and accurately measure the mm-wave input-power and baseband output power of the receiver. The mm-wave input comprises of two components, an upper-sideband (USB) extending from 60 to 70 GHz, and the lower-sideband (LSB) from 60 to 50 GHz. At the output of the receiver, both sidebands are down-converted to the same baseband frequency (0 to

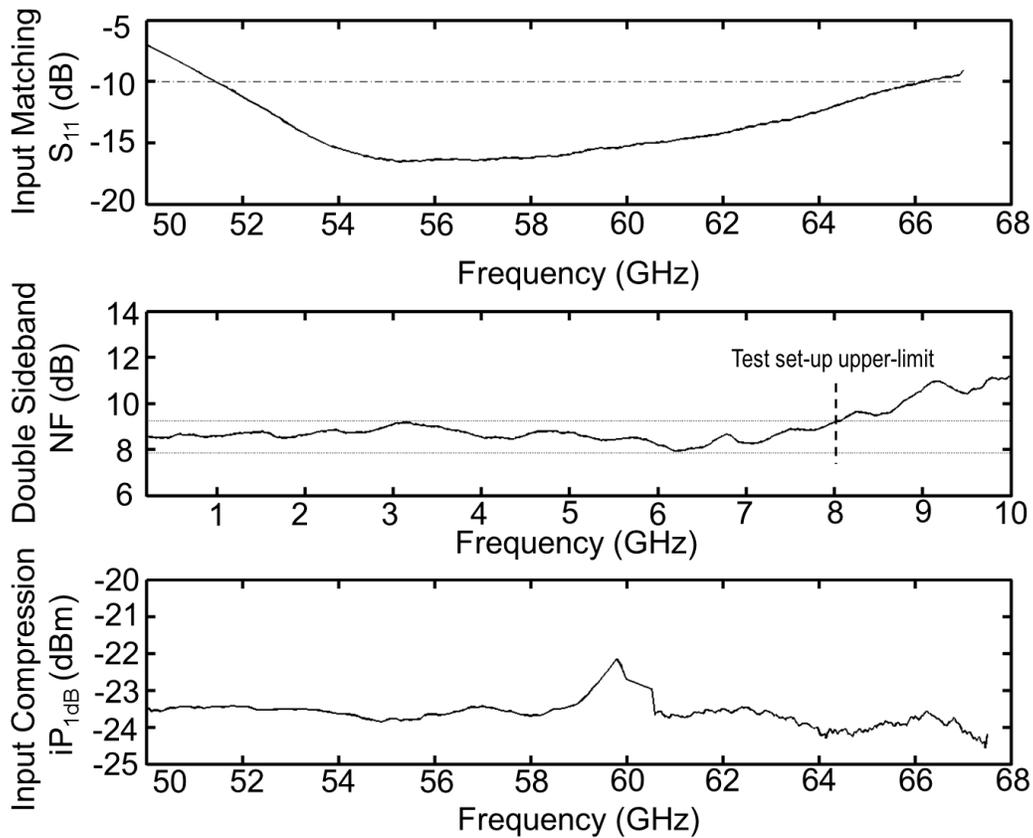


Fig. 5.21 Measured input matching, noise-figure, input compression point

10 GHz). In the USB, the receiver achieves a nominal power gain of 20dB over a frequency range of 60 to 68 GHz. In the LSB, the receiver achieves a nominal power gain of 20dB with +/- 1.5dB of gain variation over frequency-range of 51 to 60 GHz. The effective mm-wave bandwidth of the receiver is from 51 to 68 GHz. It is noteworthy that frequency-response of the multi-stage receiver has been synthesized with low pass-band gain-variation and no prominent signs of resonance peaking.

Another important performance metric of a receiver is the input power matching bandwidth. As described in section 5.3, transformer-based balun has been implemented on-chip to match the differential LNA input impedance to the single-ended antenna. At the mm-wave

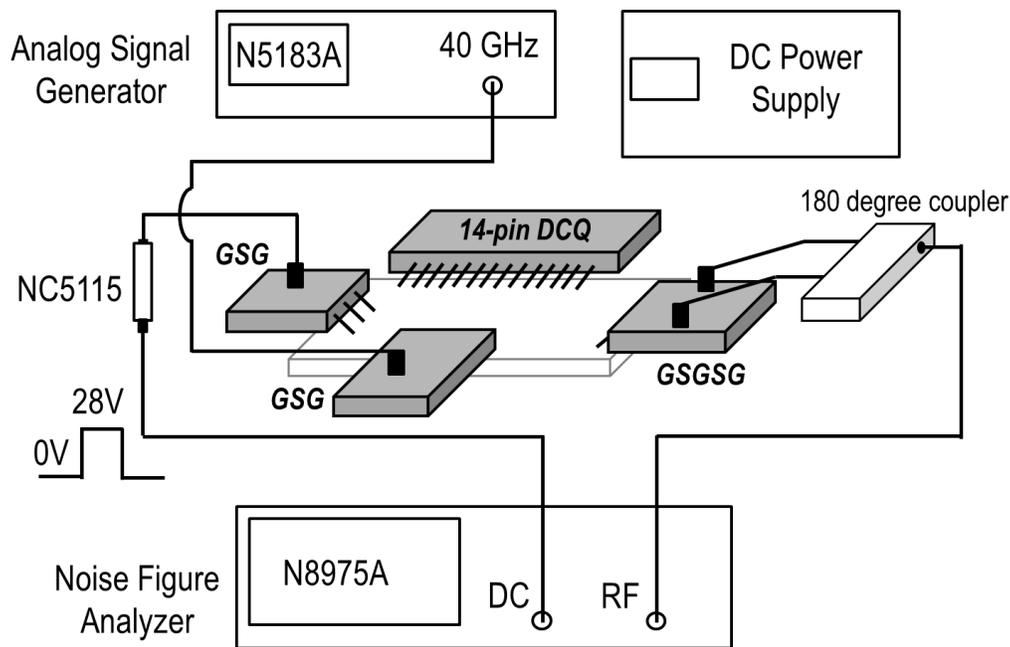


Fig. 5.22 Test setup for noise-figure measurement

front-end, the S_{11} is less than -10dB over a frequency range of 51 to 66 GHz, corresponding to a matching bandwidth of 15GHz. While the matching bandwidth does not cover the entire pass-band of interest, the bandwidth could be increased by adding de-Q resistors at the mm-wave front-end at the expense of noise-figure degradation.

The receiver noise-figure is characterized using the Y-factor method. The block diagram of the test setup is shown in Fig. 5.22. Noisecom's WR-15 waveguide based mm-wave noise-source (NC5115) functions as a calibrated noise-source operating over a bandwidth of 50-75GHz. However, the probes, waveguide-to-coaxial adapters, and 1.85mm coaxial cables limit the measurement bandwidth to 68 GHz. The noise-power at the output of the receiver (in the DC to 10GHz frequency band) is measured using an Agilent's N8975A 26.5GHz noise-figure analyzer (NFA). Agilent's N9002A, a 26.5GHz Smart Noise Source (SNS) is used to calibrate the internal low-noise receiver of the NFA.

For a two-point Y-factor noise-figure measurement, the noise-source must be toggled between two states, termed hot and cold, using a 0/28V pulsed-source inside the NFA. The single-ended noise-input is brought on-chip using the GSG probe. After down-conversion, the output of the receiver is probed differentially using a GSGSG probe, and converted to a single-ended signal using an off-chip wideband balun. The power at the output balun is measured by the NFA. Using two measured data-points, power in the hot state and power in the cold-state, and the effective noise-ratio (ENR) of the noise-source, the NF is calculated. This process is repeated at each frequency step.

The noise-power is measured at the output of the receiver, i.e. after down-conversion. Thus, the noise from the frequency-band 50-to-60GHz and 60-to-70GHz folds back into the same baseband frequency; therefore, the measured double-sideband (DSB) noise-figure is reported as a function of frequency is plotted in Fig. 5.21. The minimum DSB noise-figure of the entire receive-chain is 7.8dB. The DSB-NF remains less than 9.3dB up to a frequency of 8GHz. The noise-figure degradation at frequencies greater than 8GHz is attributed to the upper frequency limit of the measurement equipment (68GHz).

For a two-point Y-factor noise-figure measurement, the noise-source must be toggled between two states, termed hot and cold, using a 0/28V pulsed-source inside the NFA. The single-ended noise-input is brought on-chip using the GSG probe. After down-conversion, the output of the receiver is probed differentially using a GSGSG probe, and converted to a single-ended signal using an off-chip wideband balun. The power at the output balun is measured by the NFA. Using measured power in the hot/cold-state, and the effective noise-ratio (ENR) of the noise-source, the NF is calculated. This process is repeated at each frequency step.

The N5247A contains two internal signal-sources, thereby simplifying the measurement setup for a wide range of receiver two-tone characterization. The test setup for linearity measurements is identical to conversion-gain measurement setup (Fig. 5.19). The receiver achieves an input-referred 1dB compression-point of -24dBm and is plotted vs. frequency in Fig. 5.21.

The entire chip consumes 104mW from a 1.1V supply. The reported power consumption includes the power of the signal-path and the LO distribution path. The power per unit mm-wave bandwidth is (104mW/18GHz) or 6.1pW/Hz.

5.5 Conclusions

The mm-wave receiver presented in this work utilizes multiple instances of gain-equalized transformer-loads where the resonant-peaks of the mm-wave front-end and IF-stage are staggered and tuned to achieve a flat frequency response across the entire channel, from the LNA input to the baseband output. While application of capacitive cross-coupling in a transformer has previously been applied for a resonant-mode switching based oscillator design [5], this is the first use of this technique for bandwidth extension in the signal path.

The receiver IC has been implemented in a 40-nm CMOS process with a 6-metal stack and the entire chip (including the pads and bypass capacitance) occupies an area of 1.2mm². The core signal and LO-path of the receiver occupies 0.42mm² only.

Innovative circuits and architectures aimed at capturing wide absolute bandwidths in the millimeter-wave spectrum have received considerable interest over the past decade. The performance of this prototype receiver is compared with other state-of-the-art implementations in technologies such as SiGe, BiCMOS, SOI, and standard CMOS in Table 5.1. As described

Table 5.1 Comparison with state-of-the-art wideband receivers

	[1]	[2]	[3]	[4]	This Work
Tech	130um SiGe-BiCMOS	65nm CMOS	180nm BiCMOS	45nm SOI-CMOS	40nm CMOS
BEOL	6-metal, 2 UTM	7-metal, 2 UTM, 1 Aluminum	6-metal Aluminum	-NA-	6-metal, 1 UTM, 1 Aluminum
Architecture	Direct Conversion	Sliding IF Heterodyne	Direct Conversion	Direct Conversion	Heterodyne
$F_{LO,RF}, F_{LO-IF}$ (GHz)	76, -x-	37.3-42.9, 18.65-21.9	-NA-	-NA-	20, 40
RF-BW (GHz)	70-80	53 - 66	75 – 95	46 - 64	51 – 68
IF-BW (GHz)	-x-	4.5 ⁺	-x-	-x-	31 - 48
BB-BW (GHz)	6	1.2 ⁺	~ 8	1.8	8(USB)* 9(LSB)
Power (mW) (50Ω output buffer power not included)	180 ^{1a}	61 ^{2a}	250 ^{3a}	18.4 (LO power not included)	104
Gain (dB)	50	35.5 (Voltage-gain)	37	24	20
NF (dB)	< 7	5.6-6.5	< 7	< 7.1	< 9.3
iP_{1dB} (dBm)	-51	-39	-35	-22	-24
V_{DD} (V)	1.5, 2.5, 3.3	1-1.2	2.5	1.1	1.1
Area (including pads)	2.1mm ²	2.6mm ²	3.4mm ² ^{3b}	0.8mm ²	1.2mm²
$FoM = \frac{Power}{2 \times BB-BW}$	15pW/Hz	25.4pW/Hz	15.6pW/Hz	5.1pW/Hz (LO power not included)	6.1pW/Hz

⁺ Target IF-BW listed in [2], BB-BW estimate based on the 60GHz standard specifications

* Limited by 1.85mm coax based measurement setup

^{1a} Estimated power of LNA, IQ-Mixer, Baseband, ½ clock-tree based on Table-II [1]

^{2a} Estimated power of RX-front end, VCO, VCO buffers based on Table-II [2]

^{3a} Estimated per channel power based on Table-I[4] ^{3b} Area of the entire 4-element phase-array chip

previously, direct-conversion as well as heterodyne have been explored. The direct-conversion receivers reported in [1], [3], [4] have front-end circuits which achieve fractional bandwidths of 20%, 20% and 40% respectively. The mm-wave heterodyne receiver reported in [2], supports a baseband channel bandwidth of 1.2GHz only. Among all the receivers, the prototype IC

described in this chapter has the highest baseband bandwidth, 8 GHz in the USB and 9 GHz in the LSB.

The figure-of-merit used to compare the different receivers is power consumption per unit hertz. Power efficiency is a strong function of the metal stack available in the process; therefore the BEOL available in each process has been included in the Table 5.1. The receiver reported in this work has the highest baseband bandwidth in comparison with millimeter-wave receivers reported in prior-art.

This receiver consumes 104mW of power from a 1.1V supply, while providing a flat conversion-gain over an effective baseband bandwidth of $(8+9)/2$ or 8.5 GHz. The receiver FoM of 6.1pW/Hz is lower than [1], [2], [3]. The direct-conversion receiver in [4] has an FoM of 5.1pW, however, the IC lacks any on-chip LO distribution chain (the LO was driven directly from an off-chip source). In contrast to [4], the power reported in this work includes the power of the local-oscillator buffers (20 and 40 GHz), and 20 GHz injection-locked divider. For high-element phased-array receivers the LO power consumption will be a major component of the total power budget.

Appendix 5.1

There are a total of seven variables $v_a, v_b, v_{out}, v_1, i_a, i_b, i_{in}$; therefore, six equations are required to solve for v_{out}/i_{in} .

$$v_a = i_a sL + i_b sk_mL \quad (5-66)$$

$$v_b = i_a sk_mL + i_b sL \quad (5-67)$$

$$v_{out} = -i_b \frac{1}{sC} \quad (5-68)$$

$$v_{out} = v_b + i_b R \quad (5-69)$$

$$i_{in} = i_a + v_1 sC \quad (5-70)$$

$$v_1 = v_a + i_a R \quad (5-71)$$

Using (5-68) in (5-69)

$$v_b = v_{out}(1 + sCR) \quad (5-72)$$

Using (5-72) and (5-68) in (5-67),

$$v_{out}(1 + sCR) = i_a s k_m L - v_{out} sC sL \quad (5-73)$$

$$v_{out}(1 + sCR + s^2 LC) = i_a s k_m L \quad (5-74)$$

$$i_a = \frac{v_{out}(1 + sCR + s^2 LC)}{s k_m L} \quad (5-75)$$

Using (5-71) in (5-70),

$$i_{in} = i_a + (v_a + i_a R) sC \quad (5-76)$$

$$i_{in} = i_a + (i_a sL + i_b s k_m L + i_a R) sC \quad (5-77)$$

$$i_{in} = i_a(1 + s^2 LC + sRC) - v_{out} sC s^2 k_m LC \quad (5-78)$$

Finally substituting, (5-78) in (5-75)

$$i_{in} = \frac{v_{out}(1 + sCR + s^2 LC)^2}{s k_m L} - v_{out} sC s^2 k_m LC$$

$$i_{in} = \frac{v_{out}(1 + sCR + s^2LC)^2}{sk_mL} - v_{out}sC s^2k_mLCsk_mL$$

$$\frac{i_{in}}{v_{out}} = \frac{(1 + sCR + s^2LC)^2 - s^4k_m^2L^2C^2}{sk_mL}$$

$$\frac{v_{out}}{i_{in}} = \frac{sk_mL}{(1 + sCR + s^2LC)^2 - s^4k_m^2L^2C^2}$$

$$\frac{v_{out}}{i_{in}} = \frac{sk_mL}{(1 + sCR + s^2LC)^2 - (s^2k_mLC)^2}$$

Simplifying the above

$$\frac{v_{out}}{i_{in}} = \frac{sk_mL}{\{1 + sCR + s^2LC(1 - k_m)\}\{1 + sCR + s^2LC(1 + k_m)\}}$$

- [1] Sarkas, et.al.. "An 18-Gb/s, direct QPSK modulation SiGe BiCMOS transceiver for last mile links in the 70–80 GHz band", *IEEE J. Solid-State Circuits*, vol. 45, no.10, pp.1968 -1980, Oct. 2010.
- [2] F. Vecchi, et.al. "A wideband receiver for multi-Gbit/s communications in 65nm CMOS", in *IEEE J. Solid-State Circuits*, vol.46, no.3, pp.551-561, March 2011.
- [3] S. Shahramian, et.al. "A 70-100GHz direct-conversion transmitter and receiver phased array chipset demonstrating 10Gb/s wireless link", *IEEE J. Solid-State Circuits*, vol. 48, no.5, pp.1113-1125, May 2013.
- [4] S. Kundu et.al., "A supply-voltage scalable, 45nm CMOS ultra-wideband receiver for mm-wave ranging and communication", in *IEEE CICC*, pp.1-4, Sept 2012.
- [5] G. Li, et.al., "A low-phase-noise wide-tuning-range oscillator based on resonant mode switching", *IEEE J. Solid-State Circuits*, vol.47, no.6, pp.1295-1308, June 2012.

6 CONCLUSION

Over the past decade, research efforts in the millimeter-wave space have progressed in two directions – increasing the highest frequency of operation and reducing the system power consumption. Increasing demands for low-cost and high-bandwidth wireless connectivity is compelling electronic manufactures to seriously consider single-chip millimeter-wave transceivers in the next-generation of cellular-phones and tablets. With advanced CMOS process nodes (such 40 and 28nm) the maximum operating frequency of integrated transistors now exceed 200GHz. It is foreseeable that technology scaling and reduction in the transistor feature size will push the operating frequency closer to the terahertz domain in the near future.

6.1 Thesis Summary

The goal of the research described in this thesis is to explore power and area minimization techniques for multi-element phased-array receiver design. The mm-wave heterodyne receiver with a 17GHz IF bandwidth described in Chapter-2 and 5 would serve as the signal-path for a single-element in an N-element phased-array. To handle the resulting high fractional-bandwidth signals, this dissertation proposed three different wideband circuit design techniques – transformer-feedback-based design, band-pass distributed amplification, and gain-equalized transformer-loading.

The direct-conversion receiver described in chapter-3 utilizes multi-stage transformer feedback for achieving a 3dB conversion-gain bandwidth of 2GHz (11 to 13GHz). An in-depth analysis of the design of the matching-network is provided. In addition, analytic expressions for the input-resistance, quality-factor, and noise-figure as a function of the transformer turns-ratio and the magnetic-coupling coefficient are derived.

The transformed-feedback-based receiver achieved a fBW of approximately 15%, which falls short of the requirements of the 20GHz receiver originally targeted. Therefore, in chapter-4, we explore the possibility to extend the design principles of distributed amplification popular in low-pass amplifier design to high fractional-bandwidth band-pass amplifiers. A prototype chip with a 77% fBW band-pass distributed amplifier (BPDA) employing dual mirror-symmetric Norton transforms for area reduction was demonstrated in a 40-nm CMOS process.

The BPDA, while demonstrating high fBW, is essentially a single-ended amplifier optimized for systems requiring a 50-ohms input and output impedance. Therefore, application of the BPDA as the differential IF-amplifier of a heterodyne receiver is neither area nor power optimal. To capture wideband signals, chapter 5 proposes the application of multi-stage coupled resonant circuits. The limitations of using only magnetic coupling or only electrical coupling are mathematically derived. This analysis leads to a discussion on gain-equalized transformers; a technique to introduce parasitic cross-coupling between magnetically coupled resonant tanks. Finally, using the gain-equalized transformer as the core building block, the design of a 50-to-70 GHz heterodyne millimeter-wave receiver is described.

6.2 Future Directions

Research in the area of integrated circuits is entering an exciting new phase with single-chip mm-wave transceivers, slowly but steadily, entering the consumer electronics market. From the perspective of a circuits and system engineer, future research efforts will need to focus in two directions: increasing the carrier frequency and reducing the power consumption per unit bandwidth.

Increasing carrier frequency

Current research efforts have primarily focused on the 6-to-7 Gbps of wireless data-rate supported by 60-GHz standards such as the IEEE802.11ad and WiGig. However, it has been reported in [1] that the increase in speed of wireless communication has been approximately ten-fold every four years, a trend which leads us to 100Gbps by 2020. The ultra-wideband receiver presented in this dissertation applies high fractional-bandwidth design techniques to capture a channel bandwidth ten times larger than state-of-the-art *60GHz* standards. To support two orders of magnitude increase in the wireless data-rate, extending these high fBW techniques to sub-terahertz (100 to 300 GHz) and terahertz (300 GHz to 3GHz) frequencies holds significant promise. While the feasibility of sub-terahertz CMOS and bipolar transceivers has been demonstrated in [1]-[4], challenges such designing phase-locked loops with ultra-wide tuning range, generating on-chip oscillators with high output powers, high-efficiency on-chip antennas are few of the many issues yet to be fully resolved.

Reducing power consumption

For applications such as low-cost, portable medical imaging-systems, high-speed wireless data-transfer, or any battery operated mobile device, power consumption is big concern. One of the major issue with mm-wave phased-arrays reported in prior-art has been the total-power consumption. For example the millimeter-wave front-end reported in [6],[7] is a sixteen-element phased-array 60GHz transceiver and consumes approximately 5.6W of power for 3.1Gbps of RF data-rate. A more complete recent single-element transceiver (with a front-end, PHY, and MAC layer) reported in [8] consumes 1.7W of power and data rate of 1.5Gbps over one meter. In short, the high-data rate is coming with a huge power penalty. Watt-level power consumption

might be permissible long-range infrastructure applications (such as cellular base-stations, wireless backhaul networks) but would preclude usage in battery-driven mobile transceivers and mobile-devices. To minimize the total power consumption optimizations are necessary at both the circuit and the system-level. The high-IF heterodyne architecture proposed in this work was motivated by the need to reduce the local-oscillator distribution power. Another possible approach worth considering is to eliminate high-power LO distribution by using an injection-locked self-oscillating mixer topology [9]. Understanding, signal and oscillator distribution, not merely on-chip, but also off-chip (for example to the package and the antenna) will be an important research direction in the future.

While the primary focus of this dissertation has been on the mm-wave front-end, it is important to note that a complete transceiver design includes several other important building blocks. For example, sampling and digitization of 10GHz of data down-converted to the baseband is very challenging. A thorough power-budget which includes the power-consumption of the high sampling-rate analog-to-digital and digital-to-analog converters is important for justifying ultra-wideband systems. In addition, moving forward, the ability to extend techniques from discrete component microwave design for on-chip passives (phase-shifters, signal-combiners) will be important.

- [1] M. Fujishima, M. Motoyoshi, K. Katayama, K. Takano, N. Ono, and R. Fujimoto, "98mW 10Gbps wireless transceiver chipset with D-Band CMOS circuits," in *IEEE Journal of Solid-State Circuits*, vol.48, no.10, pp.2273-2284, 2013.
- [2] R. Han, and E. Afhsari, "A high-power broadband passive terahertz frequency doubler in CMOS," in *IEEE Transactions on Microwave Theory and Techniques*, vol.61, no.3, pp.1150-1160, 2013.
- [3] Z. Wang, P.-Y. Chiang, C.-C. Wang, Z. Chen, and P. Heydari, "A 210GHz fully integrated differential transceiver with fundamental frequency VCO in 32nm SOI CMOS," in *IEEE International Solid-State Circuits Conference*, Digest of Tech. Papers, pp.136-137, 2013.
- [4] 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," in *IEEE Journal of Solid-State Circuits*, vol.47, no.10, pp.2344-2354, 2012.
- [5] J.-D. Park, S. Kang, S. V. Thyagarajan, E. Alon, A. M. Niknejad, "A 260 GHz fully integrated CMOS transceiver for wireless chip-to-chip communication," in *IEEE Symposium on VLSI Circuits*, Digest of Tech. Papers, pp.48-49, 2012.
- [6] A. Natarajan et.al. "A fully-integrated 16-element phased-array receiver in SiGe BiCMOS for 60-GHz communications," in *IEEE Journal of Solid-State Circuits*, vol.46, no.5, pp.1059-1075, 2011.
- [7] A. Valdes-Garcia et.al. "A fully integrated 16-element phased-array transmitter in SiGe BiCMOS for 60-GHz communications," in *IEEE Journal of Solid-State Circuits*, vol.45, no.2, pp.2757-2773, 2010.

- [8] N. Saito et.al., "A fully-integrated 60-GHz CMOS transceiver chipset based on WiGig/IEEE 802.11ad with built-in self calibration for mobile usage" in *IEEE Journal of Solid-State Circuits*, vol.48 , no.12 , pp.3146-3159 , 2013.
- [9] M. Tedeschi, A. Liscidini, and R. Castello, "Low-power quadrature receivers for Zigbee (IEEE 802.15.4) applications," in *IEEE Journal of Solid-State Circuits*, vol.45, no.9, pp.1710-1719, 2010.