

Design of Integrated Full-Duplex Wireless Transceivers

*Sameet Ramakrishnan
Borivoje Nikolic*

Electrical Engineering and Computer Sciences
University of California at Berkeley

Technical Report No. UCB/EECS-2017-24

<http://www2.eecs.berkeley.edu/Pubs/TechRpts/2017/EECS-2017-24.html>

May 1, 2017



Copyright © 2017, by the author(s).
All rights reserved.

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

Design of Integrated Full-Duplex Wireless Transceivers

by

Sameet Ramakrishnan

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

in

Engineering - Electrical Engineering and Computer Science

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Borivoje Nikolic, Chair
Professor Ali Niknejad
Professor Paul Wright

Fall 2016

Design of Integrated Full-Duplex Wireless Transceivers

Copyright 2016
by
Sameet Ramakrishnan

Abstract

Design of Integrated Full-Duplex Wireless Transceivers

by

Sameet Ramakrishnan

Doctor of Philosophy in Engineering - Electrical Engineering and Computer Science

University of California, Berkeley

Professor Borivoje Nikolic, Chair

Demand for mobile data traffic is projected to exceed 30 exabytes per month in 2020, representing an over 100x increase since 2010. Prior generations of cellular deployments have serviced increased demand largely through use of more bandwidth - from 200KHz in GSM, to now 100MHz in CA-LTE. This method of scaling is closed, as low frequency spectrum has crowded and saturated. A proposed technique to enhance spectrum access in 5G deployments is agile full-duplex (FD) transceivers, which can transmit and receive at overlapped frequencies, or tune to arbitrarily spaced transmit/receive(TX/RX) frequency division duplexed (FDD) channels, to make use of available spectrum. The key problem in such a system is mitigating the interference the system's own transmitter creates for its receiver during simultaneous operation. Current implementations mitigate TX to RX interference at the antenna interface using off-chip, fixed-frequency duplexers, limiting a device's spectrum access to a handful of pre-defined, widely separated TX/RX band combinations. Accordingly, a universal mobile device tunable across global carrier band combinations does not exist.

This work develops a transceiver architecture enabling simultaneous transmission and reception on a single shared antenna, over a wide frequency tuning range, for FD/FDD systems. The architecture is enabled by an active TX replica which cancels interference at the RX input, a highly linear passive-mixer first receiver design based on class-AB transconductors which operates linearly in the presence of residual TX interference, and digital adaptation techniques which match the interference over time-varying operating conditions. Analysis is presented for the system's fundamental performance bounds in power and sensitivity, leading to noise mitigation techniques which minimize receiver degradation in the presence of the cancellation circuits. The analysis is validated by two chip prototypes, which demonstrate over >50dB cancellation of a +16dBm peak 20MHz TX signal, from 1GHz to 2GHz, up to an antenna VSWR of 5:1. This work demonstrates the potential for a fully integrated, frequency-tunable FD/FDD transceiver system, which could ultimately double existing mobile network capacity, and enable a universal duplexer-less radio.

Dedication
To Mom and Dad.

Contents

Contents	ii
List of Figures	iv
List of Tables	viii
1 Introduction	1
1.1 Motivation	1
1.1.1 Boosted Spectral Efficiency	3
1.1.2 Universal Radio	3
1.1.3 Simplified Spectral Planning	4
1.1.4 Backhaul and Relaying	6
1.1.5 Control Planes	6
1.2 Duplexer Specifications	6
1.3 Integrated Self Interference Cancellation: Prior Work	10
1.3.1 Transformer Hybrids	10
1.3.2 Active Cancellation	12
1.4 Wideband Transceivers	15
1.5 Research Goals, Scope, and Organization	16
2 Analysis of the Active Cancellation System	18
2.1 Description of Proposed System	18
2.2 Replica Power Consumption	23
2.3 Thermal Noise	25
2.3.1 TX Thermal Noise	25
2.3.2 Replica DAC Thermal Noise	26
2.4 Phase Noise	34
2.4.1 Background	34
2.4.2 A Quick Note on Simulation	38
2.4.3 Impact of Correlated Phase Noise	38
2.4.4 Impact of Uncorrelated Phase Noise	43
2.4.5 Phase Noise of an Inverter Chain	44

2.4.6	Sizing Under Constrained Phase Noise	47
2.5	Noise Summary	49
2.6	Quantization Noise	49
2.7	Sampling Rates	51
3	Fully Integrated FDD Transceiver Implementation	53
3.1	Switched Capacitor Power Amplifier	53
3.1.1	Topology Motivation	53
3.1.2	SC PA Operation Principle	53
3.1.3	Cartesian Implementation	55
3.1.4	Sizing	62
3.2	RX Design	65
3.3	Cancellation DAC	69
3.4	Test Setup	69
3.4.1	PCB Design	71
3.4.2	Test Equipment	73
3.5	Measurements	74
3.5.1	RX	74
3.5.2	TX	76
3.5.3	System	77
3.5.4	Antenna Mismatch	83
3.5.5	Phase Noise Measurements	85
4	Receiver Design for FD/FDD Systems	93
4.1	Design Motivation	93
4.2	Passive Mixer First Receiver	96
4.2.1	Input Matching	96
4.2.2	Noise Analysis	100
4.3	First Stage RX Amplifier Design	106
4.4	RX Second Stage Biquad	116
4.5	Chip Implementation	117
4.6	Measurements	120
5	Conclusion	129
5.1	Thesis Contributions	129
5.2	Future Work	130
	Bibliography	132

List of Figures

1.1	Frequency allocation in the United States.	2
1.2	Spectral efficiency of various standards from 2007 to 2013 - full duplex represents an opportunity to double.	3
1.3	Self-interference cancellation for universal radio.	4
1.4	Comparison of TDD and FDD implementations.	5
1.5	Duplexer functionality.	7
1.6	Wireless hybrid as an integrated duplexer.	10
1.7	Distributed Tline duplexer.	14
1.8	Limitations of prior work.	14
1.9	Frequency spectrum (top) and board components (bottom) for various standards.	16
1.10	Requirements for two stage (analog/digital) cancellation.	17
2.1	Simultaneous TX/RX interface.	19
2.2	Models for TX operation (left) and RX operation (right).	20
2.3	Current mode (left) and voltage mode (right) cancellation.	22
2.4	Digital feedforward adaptation possibilities.	22
2.5	TX vs. replica power consumption.	23
2.6	DAC's current and voltage.	24
2.7	TX vs. replica power consumption.	26
2.8	Noise model for the TX.	27
2.9	Model for DAC noise analysis.	28
2.10	Reduction of the high frequency noise via degeneration.	28
2.11	Contribution of DAC $2F_{LO}$ noise with inductive degeneration.	30
2.12	DAC thermal noise contour vs. TX Power.	31
2.13	Baseband noise feedback loop.	32
2.14	Model for analysis of DAC noise feedback.	33
2.15	Reduction of the low frequency tail noise via feedback.	34
2.16	PM and AM noise components.	35
2.17	Phase noise as voltage pulses.	36
2.18	Phase noise through digital and analog paths.	39
2.19	Phase noise through the LO divider.	40
2.20	Clock divider uncorrelated phase noise.	41

2.21	Inverter noise model.	44
2.22	Noise at the inverter edge crossing.	45
2.23	Minimum power for a length of 5 digital inverters.	48
2.24	Per stage fanouts to achieve minimum power for target phase noise.	48
2.25	Receiver noise scaling vs. TX power. Right - DAC thermal noise appears at moderate powers. Left - Uncorrelated phase noise appears first.	49
2.26	Top: DAC code vs voltage characteristic, Left: data sequence Right: quantization error is data correlated.	51
3.1	SC PA as a class D PA.	54
3.2	SC PA operating principle.	55
3.3	Cartesian PA model.	56
3.4	Polar (right) vs Cartesian (left) PA.	56
3.5	Simulated phase interpolator phase noise.	57
3.6	25% Cartesian LO combination.	58
3.7	Charging model for the Q phase.	60
3.8	Cartesian(4 bits binary, 4 bits thermo) vs. Polar efficiency contours. The polar is the constant efficiency at constant amplitude circles. The black lines correspond to achievable powers for cartesian.	61
3.9	Top level TX schematic.	62
3.10	TX transformer.	63
3.11	SC PA drain kickback.	64
3.12	Worst case drain voltage vs. time.	65
3.13	Top level RX schematic.	66
3.14	LNTA schematic.	67
3.15	LNTA CMFB schematic.	67
3.16	TIA schematic.	68
3.17	Mixer/LO path schematic.	68
3.18	Top level DAC schematic.	69
3.19	Chip top level schematic.	70
3.20	Die Photo.	70
3.21	Measurement PCB.	71
3.22	PCB configuration for isolated (left) and system testing (right).	72
3.23	Inductor parameters with top 4 layers of PCB cut.	73
3.24	2FLO Noise Reduction.	73
3.25	Test Setup.	74
3.26	S21 of the SMA and output transmission line.	75
3.27	RX S21 (left) and S11(right).	75
3.28	Various RX gain and bandwidth settings.	76
3.29	RX P1dB (left) and IIP3 (right).	77
3.30	TX power (left) and current (right) vs. code.	77
3.31	TX output vs frequency.	78

3.32	TX DNL.	78
3.33	TX residual referred at RX input vs TX output power.	80
3.34	RX band spectrum before and after cancellation.	80
3.35	TX/DAC codes after adaptation.	81
3.36	TX residual vs. center frequency.	81
3.37	Noise figure degradation measurement vs curve fit of constant + 6dB/decade.	82
3.38	Noise Figure degradation vs TX LO power.	83
3.39	Comparison Table.	84
3.40	Measurement setup for VSWR.	85
3.41	Antenna impedance points measured.	86
3.42	Phase relationship of I and Q spurs for injected input spurs.	86
3.43	Spur cancellation vs offset frequency.	87
3.44	Test setup for phase noise cancellation.	88
3.45	Filtered input noise spectrum at output before and after cancellation.	89
3.46	Wideband input noise spectrum at output before and after cancellation.	89
3.47	Phase shift of TX to RX coupling path.	90
3.48	TX phase noise cancellation in RX band as the TX-RX leakage path delay increases.	91
3.49	Schematic of the input clock receiver which limits phase noise performance.	91
3.50	Simulated noise contributions of the LO divider.	92
4.1	Third harmonic power vs TX power at antenna.	94
4.2	Potential options for TX third harmonic cancellation.	94
4.3	Available gain of the transformer network vs. DAC cap.	95
4.4	Passive-mixer-first input match model.	97
4.5	Voltage on RF side of the mixer.	98
4.6	Equivalent linear impedance model of the RX.	98
4.7	S11 offset of the receiver.	99
4.8	Cross-coupled baseband for complex input impedance creation.	99
4.9	S11 with and without complex feedback.	100
4.10	Model for noise analysis of the baseband.	101
4.11	Noise model of the cross-coupled feedback path.	102
4.12	RX NF vs number of mixer phases.	104
4.13	Noise figure vs. mixer switch size.	104
4.14	Noise figure (left) and power consumption (right) vs. inverter size.	105
4.15	Noise figure vs frequency.	105
4.16	Noise figure for different RF cap values.	106
4.17	Noise figure with a noiseless baseband amp.	107
4.18	Noise figure vs amplifier input referred noise.	108
4.19	Complementary first stage schematic.	108
4.20	Baseband amplifier with shunt feedback second stage.	109
4.21	Baseband amplifier schematic.	110
4.22	Latchup path for baseband amplifier.	111

4.23	Common mode diode IV curves.	111
4.24	Common mode pulldown diodes.	112
4.25	Baseband amplifier transient startup curves.	113
4.26	Local common mode feedback loop gain.	113
4.27	Global common mode loop gain.	114
4.28	Differential mode loop gain.	115
4.29	Baseband amplifier forward gain.	115
4.30	MFB/harmonic recombination amplifier.	118
4.31	MFB filter response.	118
4.32	Receiver top level diagram.	119
4.33	RX power breakdown.	119
4.34	Receiver top level layout.	120
4.35	Die Photo.	121
4.36	S11 vs. LO frequency for fixed matching network cap DAC setting.	122
4.37	S11 vs. feedback and feedforward resistor settings.	123
4.38	S21 vs. baseband frequency at 1GHz center frequency	123
4.39	IIP3 as a function of tone spacing.	124
4.40	Receiver 1dB compression vs frequency offset.	125
4.41	Receiver third and fifth harmonic rejection.	126
4.42	RX comparison table.	127
4.43	RX compression vs. TX power.	128

List of Tables

1.1	Example design specifications from LTE standard.	9
1.2	Operating frequency range of recently published wideband receivers.	15
1.3	Operating frequency range of recently published wideband transmitters.	15

Acknowledgments

I owe a great number of people for their support to this point.

First and foremost, I would like to thank my mom and brother Girish. Mom, I can never repay your unconditional love, and sacrifice in putting me above all else.

Professor Nikolic, you're truly an advisor in the purest sense of the word. Your direction on all things, academic and life, since I was 17 years old, have immensely shaped me. As a mentor you always have my best interests in mind. I always hear from people "Bora seems to have a high opinion of you" - I hope I can one day live up to that expectation. Professor Alon, I can't thank you enough for your wisdom, guidance, and support. Its very rare to find a professor who would accommodate weekly meetings over the summer, with a student who wasn't his own - especially when the only time that worked was 12am (not a typo) on Wednesdays. Your ability to listen to a problem I've been thinking about for weeks, dissect it into its simplest form, and get three steps past me in a matter of minutes has astounded me for years. I also thank you for your repeated attempts to teach me circuit design in EE141, EE240, and EE290C. Professor Niknejad, thank you not only for your great course and the constant technical lessons, but also for your friendship outside the classroom. In particular, I admire your natural enthusiasm, whether for new research problems, or on the soccer field. Professor Wright, thank you for serving on my qualifying exam and dissertation committees, and providing feedback. Professor Muller, thank you for conversation and support both while I was deciding on graduate school, and while I was ta-ing EE140.

The majority of this work was done in extremely close collaboration with Lucas Calderin. Luke, I can't imagine working with anyone else for this long, and I've learned an incredible amount from your creativity and work ethic. I hope you've enjoyed our collaboration as much as I have, and I look forward to working together on our next adventures.

Antonio, thank you for the research collaboration and various technical discussions (or arguments), late night shenanigans in BWRC, and the subsequent late night walks home. Bonjern, thank you for for over six years of loyal friendship. Greg (Ozzy), thank you for giving me an entirely new (and sometimes unwanted) perspective on life. Krishna, thank you for the all the Warriors games we watched together, for unconditionally having my back, and for making me smile when I need it most. Nathan, thank you for the friendship and conversations over many lunches.

I'd like to thank others in BWRC, Amy, Andrew, Angie, Naichung, Nandish, and Pavan, for friendship, technical discussion and support, or both. Several older students at BWRC have provided mentorship, guidance, and support, in particular Milos Jorgovanovic, Ling kai Kong, and Charles Wu. Also thanks to members of the ComIC group for making BWRC more enjoyable, including Amanda, Ben, Brian, Dajana, Han-Chih, Jaehwa, John, Kate-rina, Keertana, Luis, Marko, Matt, Miki, Mira, Nick, Paul, Pifeng, Rachel, Sharon, Stevo, Vinayak, and Vladimir.

The staff at BWRC has made working, taping out, and testing much easier, and I'd like to thank Ajith, Amber, Bira, Brian, Candy, Prof. Dave, Erin, Fred, James, Leslie, Olivia, Sarah, and Yessica.

Thanks to Michael Reiha, Tommi Ylamurto, Hans Daneels, and Vason Srini from Nokia, and Farhana Sheikh and Dan Schwartz from Intel for technical discussion, feedback, and funding.

I'd like to thank my friends outside BWRC, who have kept me sane during my time here. Dustin and Rohan, thank you for being my brothers. Thanks to Arjun for top notch Warriors analysis. Phil, thank you for getting angry on my behalf first, asking questions second, and cracking up in laughter third. Alex, Arvind, Jay, thank you for 20 (or more?) years of friendship. Kathy, thank you for Gob.

Finally, the biggest thank you to my dad. I hope this work would make you proud.

This work is funded in part by DARPA RF-FPGA program under contract HR0011-12-9-0013 in collaboration with Nokia and Boeing, and in part by Intel, and Qualcomm.

Chapter 1

Introduction

1.1 Motivation

Since the advent of mobile devices, the demand for wireless data has grown tremendously. Current projections estimate that global mobile data traffic in 2020 will reach 30 exabytes per month, up 120x from 2010. This corresponds to 28 daily images and 2.5 daily video clips per person on earth. In fact, by 2020, more people are projected to have mobile phones than running water or electricity at home [1].

Prior generations of cellular deployments have serviced this demand largely through increase in bandwidth, from 200KHz in 2G GSM networks focusing on only voice, to 1.25MHz and 5MHz in 3G CDMA/WCDMA systems, to 20 and even 100MHz in 4G LTE. This method of scaling is largely closed - as seen in the US spectrum allocations in figure 1.1, low frequency spectrum is crowded and saturated [2]. Symptomatic of this spectral crowding, radio spectrum has become one of the most valuable commodities on earth. In 2014, the US government sold 65MHz of the so called AWS-3 spectrum around 1700MHz for over 30 billion dollars [3]. Spectrum ownership now represents the dominant cost for mobile operators, over 40% of the 10 year cost of ownership.

A significant hardware issue resulting from overcrowded spectrum is interference between closely spaced radios operating concurrently. In particular, the largest source of interference for a radio receiver (RX) is often a system's own transmitter (TX). For example, in LTE deployments, the transmitter and receiver operate simultaneously, sharing the same antenna. The high power transmit signal must be isolated from the sensitive receiver, in order to maintain RX sensitivity, and prevent RX damage. This problem is exacerbated in a modern handset, which contains transmitters for multiple standards such as WiFi, GPS, and LTE, all operating at the same time. These radios create self-interference, defined as the interference generated by a single system's own transmitter operating simultaneously with the collocated receiver.

In today's systems, the self interference issue is addressed by separating the TX and RX frequency channels, and adding frequency selective filters between the TX and RX

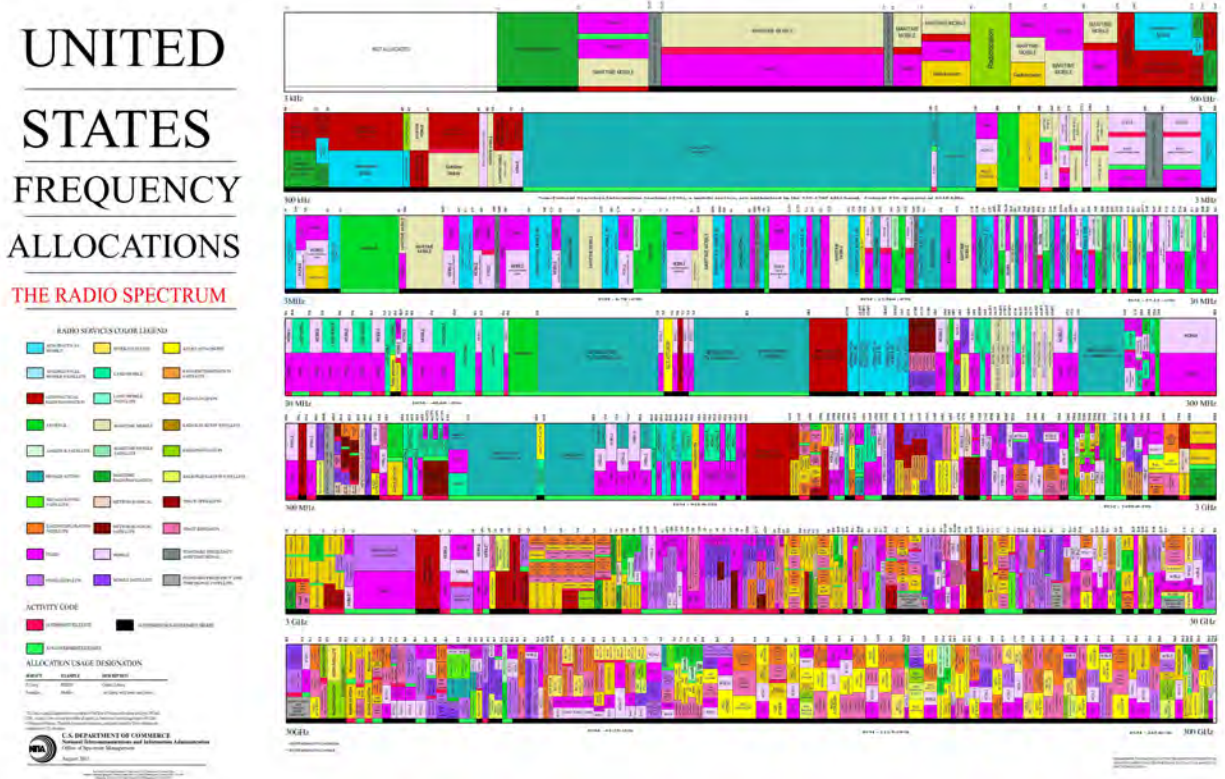


Figure 1.1: Frequency allocation in the United States.

circuits. This scheme of separated TX/RX frequency allocations is known as frequency division duplexing (FDD). TX/RX isolation filters require a large quality factor (Q), in order to achieve sufficiently steep roll-off for closely spaced frequency bands. Due to the high Q required, the filters are placed off the main transceiver chip, and are intrinsically narrowband and difficult to tune. Accordingly, a separate set of filters is required for each band, and a limited number of bands are supported in a single device. This is limited not only by the area and cost constraints of fitting these discrete filters onto the printed circuit board (PCB), but also by the loss incurred from the multiplexing paths through the filter bank.

Additionally, these filters rely on a frequency separation between the transmit and receive bands to provide isolation. It would be beneficial for a transceiver to be able to remove, or “actively cancel” its own self interference, in a frequency agnostic manner, i.e. with arbitrary TX and RX spacings, bandwidths, or even overlapped (full duplex, FD) channels. Several benefits of such a transceiver are highlighted below.

1.1.1 Boosted Spectral Efficiency

Fully overlapped transmit and receive frequency channels can increase the system spectral efficiency, defined in data rate per unit of spectrum allocated (Bits/sec/Hz). Given a finite availability of spectrum, increased data demand can only be met through an increase in spectral efficiency. However, existing standards employing the latest coding and modulation techniques are saturating in their ability to continue to extract further efficiency from the spectrum. This saturation is seen in Fig. 1.2, showing the efficiency of various standards over the past decade.

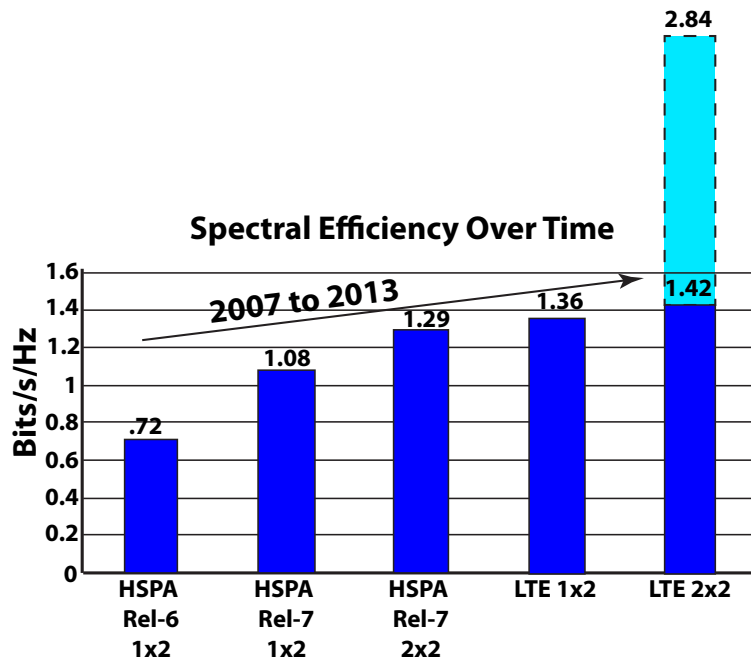


Figure 1.2: Spectral efficiency of various standards from 2007 to 2013 - full duplex represents an opportunity to double.

Fundamentally, operating the transmitter and receiver simultaneously to uplink and downlink data on overlapped spectrum can as much as double the efficiency per unit of frequency. This represents a method to break the saturation of Fig. 1.2, and would be enabled by self-interference cancellation.

1.1.2 Universal Radio

The FDD resource allocation mode has forced a paradigm of fixed TX/RX band pairings. For example, the LTE standard supports over 40 bands worldwide[4]. A separate discrete, off chip filter is needed to provide the self-interference isolation for each unique band pairing.

Accordingly, different subsets are supported in different versions of handset devices depending on the intended geographic region of operation - a universal LTE phone does not exist [5].

A tunable self-interference cancellation circuit, as shown in Fig. 1.3, would enable removal of these off-chip filters in current systems. This would save original equipment manufacturer (OEM) cost and area, and would enable consumers to use any phone across any carrier in any global market. The GSM Association (GSMA) estimates the economic impact of this globally harmonized spectrum access as hundreds of billions of dollars [6].

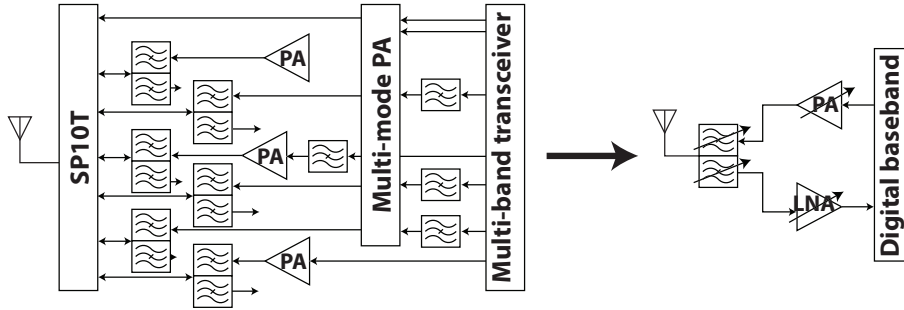


Figure 1.3: Self-interference cancellation for universal radio.

Additionally, the latest releases of LTE [4] support a new resource allocation mode, carrier aggregation (CA), wherein the radio combines disjoint pieces of spectrum when more bandwidth is needed. When paired with FDD, carrier aggregation results in an exponentially increasing number of these fixed frequency self-interference filters - not only must specific TX/RX band pairings be supported, but filters are needed for combinations as well. A circuit which could adapt the frequency of its interference rejection would unlock flexible use of disjoint spectrum in carrier aggregation scenarios.

Finally, high TX output powers at cellular basestations necessitate bulky, expensive, fixed-frequency cavity filters to be installed even between separate chips, to suppress the transmitter's self-interference into the receiver. Some tunable self interference-cancellation circuit, even one which only provides partial TX/RX isolation, could reduce size and cost of the cellular infrastructure deployment.

1.1.3 Simplified Spectral Planning

As mentioned before, the TX and RX in FDD systems occupy separate frequency bands, enabling frequency selective filters to handle the self interference during simultaneous operation. The overhead of acquiring paired TX/RX spectrum has driven many operators to time division duplex (TDD) standards, namely buying a single unpaired chunk of spectrum to reuse for transmission and reception. These TDD standards are less efficient, as shown

in data presented by Nokia-Siemens Networks in Fig. 1.4. There can be a lack of coordination between adjacent cell base-stations, which consequently interfere with each other due to transmission and reception at the same frequency. Additionally, the switching time from transmission to reception adds network overhead, eating into the capacity. However, network operators accept this inefficiency in TDD systems, due to the difficulty in building hardware to support various FDD channel pairs, and the cost of good FDD bonded spectrum.

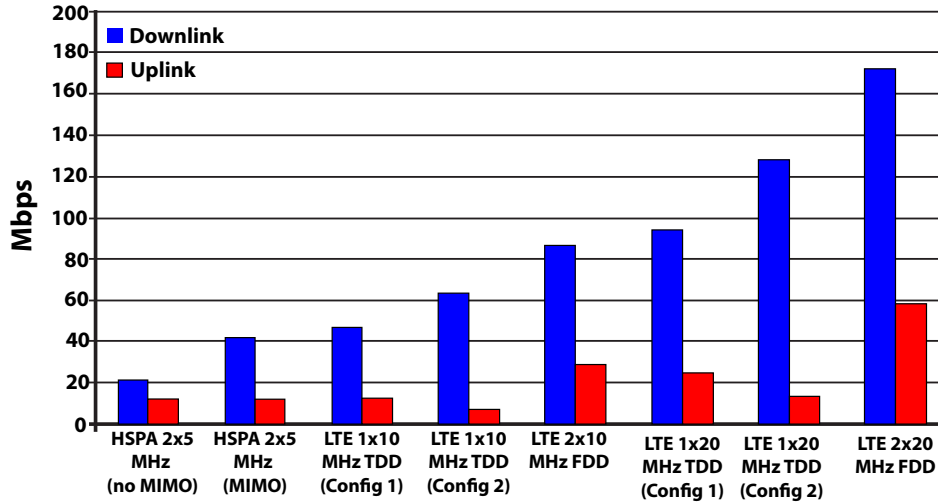


Figure 1.4: Comparison of TDD and FDD implementations.

A self-interference cancelling transceiver in essence removes the distinction between TDD and FDD systems, creating what some people have termed any-division duplexing (ADD) [7]. Old TDD systems with un-bonded spectrum could be replaced by systems which operate on the same spectrum simultaneously in transmit and receive mode. FDD systems could bond arbitrary channels together, significantly simplifying spectral planning constraints.

Additionally, in order to ease the self-interference problem, guard bands are placed between channels in the same frequency band, due to out of band TX nonlinearity and noise which leak into the systems receiver. This manifests itself in WiFi networks, where a single access point (AP) does not transmit and receive simultaneously, even on separated channels. This problem is expected to be exacerbated in upcoming proposals of LTE-U, which plans to deploy LTE in the unlicensed WiFi bands. A single AP would not be able to service both LTE-U and WiFi in a spectrally dense deployment due to this transmit leakage. A self interference cancelling transceiver would shrink or eliminate guard band requirements, and enable flexible single-device cross standard operation, again simplifying frequency planning.

1.1.4 Backhaul and Relaying

In an effort to increase network capacity, network operators have trended towards smaller, denser cell site deployment. However, backhauling the small cell sites to the main network remains a challenge. Fiber backhaul offers the best performance, but is impractical to widely deploy due to cost. High-frequency millimeter wave backhaul is still limited to line-of-sight (LOS) propagation. New low frequency spectrum is unavailable, and re-using the existing spectrum for backhaul is difficult, as the small cell station must serve its users and backhaul itself at the same time. This is therefore another instantiation of the self interference problem.

A self-interference cancellation transceiver enables a self-backhauled network, where a base station simultaneously receives from its users and backhauls itself on shared, or even overlapped spectrum. In fact, it has been shown in [8] that self-backhauled networks can approach fiber levels of performance. This is because existing LTE macro networks are heavily under-leveraged, at only 25% utilization. Self-backhauled small cells re-use the existing macro network capacity to relay back its information, improving return on the huge investment the mobile operators have already sunk into their deployment.

1.1.5 Control Planes

The control plane is a separate low-rate link between elements in a network used for network coordination purposes. In particular, this coordination information is critical for new cooperative interference mitigation schemes, such as coordinated multi point (CoMP). Interference coordination imposes tight latency requirements on the control plane data. Further, allocating new spectrum for the control plane links is a costly proposition. Simultaneous transmission of data to an end user and control plane information to another base station over shared spectrum represents a low cost, low latency method of building such control planes. This is yet another instantiation of the self interference problem, which could be enabled by frequency flexible active cancellation circuits.

As seen above, self-interference manifests across many wireless applications. Existing hardware has no ability to configure the interference rejection over uplink and downlink frequency bands, which results in bulky, expensive, and functionally limiting implementations. The possibility of overlapping the TX and RX bands opens still further advantages.

This work accordingly focuses on fully integrated transceiver design which can, in frequency agnostic manner, isolate a transmitter and receiver operating on the same antenna.

1.2 Duplexer Specifications

A duplexer is defined as a three-port device, pictured in Fig. 1.5, which interfaces the transmitter and receiver to the antenna, and mitigates the self interference. This work, in essence, attempts to integrate the duplexer's functionality directly onto the transceiver chip, using techniques independent of the transmit and receive frequencies, spacings, or bandwidths. To provide fair comparison between existing frequency inflexible off-chip duplexers, research

work, and this work, the duplexer's performance is evaluated by the specifications described in this section.

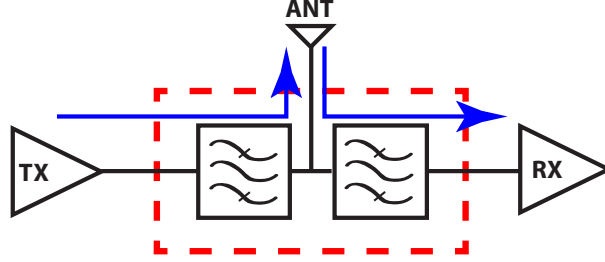


Figure 1.5: Duplexer functionality.

A receiver front-end compresses nonlinearly under a large signal, leading to distortion of the desired receive signal. As the radiated transmit signal at the antenna is significantly larger than this compression point, the transmit signal at the antenna must be isolated from the receiver input to prevent such distortion. This is reflected in the TX-RX isolation metric, which is defined in dB as

$$TX - RX \text{ Isolation} = 10 \log_{10} \left(\frac{P_{tx,rx}}{P_{tx,ant}} \right) \quad (1.1)$$

where $P_{tx,rx}$ is the TX power present at the RX input, and $P_{tx,ant}$ is the TX power present at the antenna port. This isolation in a duplexer is provided by a frequency selective filter between the antenna and receiver, as shown in 1.5. This filter passes signals in the RX band, while attenuating TX band signals. Additionally, this filter provides some attenuation of external out-of-band RX blockers.

TX-RX isolation must be provided while minimizing the Antenna to RX band attenuation, referred to as RX band insertion loss. As the duplexer is matched to 50 ohms on both sides, any loss directly degrades the signal level while maintaining a constant noise level, thus adding dB for dB to the receiver's noise figure. This insertion loss is defined as

$$RX \text{ Insertion Loss (RX IL)} = 10 \log_{10} \left(\frac{P_{rx,rx}}{P_{rx,ant}} \right) \quad (1.2)$$

where $P_{rx,ant}$ is the receive signal power present at the antenna, and $P_{rx,rx}$ is defined as the receive signal power present at the receiver's input.

Due to the dynamic nonlinearity and data quantization noise in a digital transmitter, the TX produces spurious RX band emissions, which can desensitize the receiver. This RX band noise can not be isolated by the Antenna to RX port filter - accordingly, a separate filter is needed on the TX to Antenna interface, rejecting the RX band while passing the TX band, to clean up the TX out of band emissions. This filter additionally ensures the TX

meets any spectral mask requirements imposed by the standard. The RX band filtering is defined as

$$TX's\ RXBand\ Filtering = 10\log_{10}\left(\frac{P_{tx(rxband),tx}}{P_{tx(rxband),ant}}\right) \quad (1.3)$$

where $P_{tx(rxband),tx}$ is the power of the TX spurious emission in the receive band at the transmitters output, and $P_{tx(rxband),ant}$ is defined as the power of the TX spurious emission in the receive band at the antenna.

This filter must also minimize the amount of loss from the power amplifier output to the antenna input, referred to as TX band insertion loss. This is quite critical, as the TX output power level can be on the order of a Watt, making even a few dB of loss a huge wasted power. For example, 2dB of loss on a 1 Watt TX signal corresponds to 350mW of lost power, enough to power approximately seven receive chains. Additionally, there has been recent research and commercial interest in integrated CMOS power amplifiers (PA) to lower system cost. These CMOS PAs are limited in their ability to deliver high power as compared to non-CMOS counterparts, due to the limited supply voltage and breakdown of the CMOS process. Any loss in the duplexer must be compensated by producing higher output power at the transceiver chip. This can have a super-linear power penalty, due to the need for cascoding or other circuit techniques which reduce the core PA efficiency. The TX band insertion loss is defined as

$$TX\ Insertion\ Loss\ (TX\ IL) = 10\log_{10}\left(\frac{P_{tx,ant}}{P_{tx,tx}}\right) \quad (1.4)$$

where $P_{tx,ant}$ is the transmit signal power present at the antenna, and $P_{tx,tx}$ is defined as the transmit signal power present at the transmitter's output.

Lastly, the duplexer provides a matched interface at all three ports, preventing the transmitter and receiver from excessively loading one another.

The LTE standard is taken as a good representative example to guide the specifications targeted for this work. The relevant standard level specifications are summarized in Table 1.1. Note that if a filter-based duplexer is used, the combination of these specs set the require filter order, with the ratio of TX peak power to RX noise figure setting the required rolloff, and the ratio of channel bandwidth to duplex spacing setting the transition band over which this rolloff must occur. The challenge is due to the large dynamic range difference (23dBm TX power vs -100dBm RX sensitivity) between the transmitter and receiver, which must be filtered within a very sharp stop band. In the worst case for the LTE standard, the filter must reach the stop band within 2x the signal bandwidth. This is fundamentally why narrow-band, discrete, high-Q components must be used, and integrated frequency agnostic duplexers do not exist.

The MEMS community has actively attempted to miniaturize these off-chip duplexers into a CMOS process for integration with the radio front-end. For example, bulk acoustic wave (BAW) resonators on top of silicon have been demonstrated and integrated with

Table 1.1: Example design specifications from LTE standard.

Channel BW	<20MHz
Duplex Spacing	30MHz-700MHz
Min (Spacing/BW)	2
TX Peak Power	23dBm
RX Noise Figure	<15dB
RX OOB Blocker P1dB (designed, not specced)	~0dBm

CMOS [9]. In fact, recent demonstrations have even shown that arraying many such MEMS filters (particularly those based on capacitive transduction) can address the characteristic impedance and power handling issues that plagued filters based on individual devices [10]. However, the resonant frequency of electromechanical filters is determined by physical dimensions and properties of the materials. Tuning the filter center frequency thus adds intolerable loss due to reduced network Q, or is difficult using electronically controlled means.

The duplexer performance specifications for a commercial Avago LTE duplexer correspond cleanly with the above LTE standard specifications. The TX-RX isolation is about 45-55dB. With a 1Watt (30dBm) transmit signal, this leaves around -20dBm at the RX input. As a state of the art receiver can be designed to compress by 1dB for an out of band blocker of around 0dBm, this leaves sufficient margin to process the RX signal and nearby blockers. The RX band attenuation of the TX signal is around 50dB with for example a 30dBm transmitter quantized to 11 bits, spread out over a 1GHz sampling frequency, the RX out of band quantization level is

$$30dBm - 6 \times 11bits - 10 \times \log_{10}(10^9) = -126dBm. \quad (1.5)$$

An additional 55dB of filtering puts this below the thermal noise floor level, so as to avoid RX desensitization.

The RX insertion loss is around 2dB, and the TX insertion loss is around 1dB.

Is worth noting that the same metrics apply to proposed full-duplex systems with fully overlapped transmit and receive channels. However, these metrics can not be provided through relying on frequency discrimination, and must be handled through another approach. For fully overlapped transmit and receive channels, a circulator can be used to provide an antenna interface. However, these circulators generally only provide around 20dB of isolation, so self interference cancellation is still necessary. Further, circulators require magnetic components which can not be integrated on to the integrated circuit, and are very expensive - accordingly, a technique that reduces the required isolation from the circulator and enables integration is the key enabler for a FD system.

1.3 Integrated Self Interference Cancellation: Prior Work

There have been several prior approaches to integrating the duplexer's functionality using techniques agnostic to center frequency and duplex spacing. This section details the operating principles of these approaches, and the main limitations which this work attempts to overcome.

1.3.1 Transformer Hybrids

The passive hybrid technique originated in telephone networks, and has recently re-emerged in the wireless context as an integrated duplexer. Hybrids networks [11][12][13][14][15] are reciprocal passive 4-port networks, which use a portion of the TX signal to cancel TX interference at the RX port. Often times, these networks are constructed such that the TX signal appears as a common mode perturbation across the RX, while the desired receive signal from the antenna appears in the differential mode.

A key benefit of this structure is that if the balancing network and antenna impedance are sufficiently wide-band, the structure can suppress not only the main TX band signal but also provides some isolation for the TX nonlinearity, quantization noise, and phase noise that falls in the RX band.

Any network composed of entirely reciprocal elements must also be reciprocal. Namely, for any ports i and j , $s_{ij} = s_{ji}$ where s represents the scattering parameter, or effectively the power flow, between ports. Accordingly, it is impossible for $s_{tx,ant} = 1$ and $s_{ant,rx} = 1$ simultaneously, and there is a fundamental insertion loss associated with the structure. This is normally modeled as a fourth port with a loss resistor.

To more closely understand the operation of the hybrid network, and the associated tradeoffs, prior art from [16], shown in Fig. 1.6, is examined.

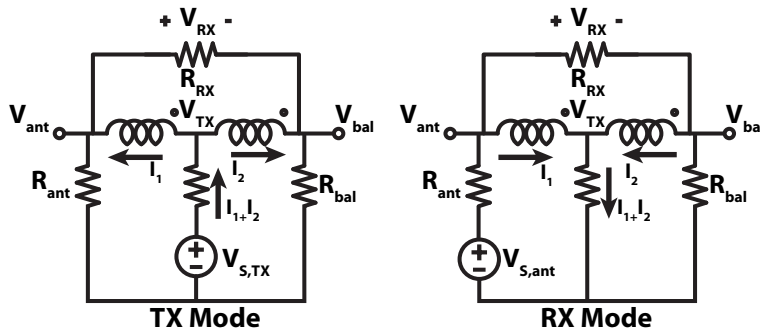


Figure 1.6: Wireless hybrid as an integrated duplexer.

Hybrid Operation Principle

For the TX signal, as shown in Fig. 1.6, with perfectly symmetric transformer, the structure attempts to enforce zero differential voltage across the receiver. This condition is true if half the TX current flows into each of the balancing and antenna branches. However, this results in a 3dB insertion loss on the transmitter. To lower the wasted TX power while maintaining zero differential voltage across the RX port, the R_{bal} resistor must be increased to flow less current, and the autotransformer turns ratio must be skewed to compensate for the imbalance. As will be shown below, increasing R_{bal} results in a noise penalty for the receive signal.

For the RX signal, in the balanced condition, the voltage at the balancing port is 0 for a stimulus applied in series with the antenna port. Accordingly, any current in the receive port must flow through I_2 in the autotransformer, and must be induced in I_1 in the other branch of the autotransformer. The current flowing through the TX must be $2I_{rx}$ by KCL, and the voltage must be half the voltage across the RX by the vertical symmetry across the transmitter - this forces a relationship that $R_{rx} = 4R_{tx}$.

To evaluate the noise penalty of the hybrid in the receive side, it is sufficient to find the Thevenin equivalent circuit looking out from the RX port for a stimulus in series with the antenna. The Thevenin resistance can be found by noting that in the balanced condition, a differential RX current should be isolated from the TX port. By KCL, the autotransformer currents must then be 0. This is somewhat confusing in that an ideal transformer can support a voltage across it with 0 current - but it can simply be imagined as a large impedance in the frequency range of interest. The Thevenin resistance is simply the series combination of the antenna and balancing resistors, which sets the available gain, and accordingly the SNR.

Note that the NF penalty of the passive network is only a function of the available gain (the Thevenin resistance), while the insertion loss is a function of both the Thevenin resistance as well as the antenna load. Accordingly, if the RX port is not matched, then RX IL is not necessarily equal to the RX NF degradation. However, there is still a tradeoff between the RX NF and TX IL. R_{bal} must be large to reduce the TX insertion loss (minimize I_2 in Fig. 1.6), but a large R_{bal} contributes a large noise voltage relative to the antenna impedance at the receiver input.

Performance

The duplexer in [16] achieves 2.5dB TX insertion loss, with >50dB isolation in the TX band, >45dB isolation in the RX band, a cascaded RX NF of 5dB, and an antenna S11 of -8dB. The isolation is maintained up to a maximum VSWR of 2:1 on a 50ohm antenna over 1.5GHz to 2.2GHz. In [16], the antenna is intentionally not matched to the RX (S11 of -8dB) in order to lower the RX noise figure.

Structure Limitations

A big limitation inherent to the structure is the previously described tradeoff between TX insertion loss vs. RX noise figure. In particular, the wasted power consumption of such a structure for a transmitter of several hundred milliwatts to a Watt is of the same order of magnitude. This is quite intuitive, as no active elements are used in the isolation structure, so the tapped off wasted portion of the TX signal must be as large as the interference to be cancelled.

Secondly, it is difficult to balance the network with a real antenna across a wide range of operating conditions. Using a 10-bit capacitor/resistor DAC as the balancing impedance, a relatively limited VSWR of 2:1 can be covered. Furthermore, it is difficult to handle high TX powers with such an impedance network, as the CMOS switches used in the capacitive DAC must handle a large voltage swing. While this can be partially overcome through the use of stacked devices, and floating wells, these structures add a signal dependent nonlinearity to the balancing network, and reduce the center frequency tuning range due to reduced $\frac{C_{on}}{C_{off}}$ of the capacitive DAC.

1.3.2 Active Cancellation

Active cancellation originates in wireline links as a method for transmitting and receiving simultaneously on a single fiber. In this technique, a portion of the transmit signal is tapped off or synthesized, and injected into the receiver to perform a feed-forward cancellation of the TX leakage.

The technique has recently gained interest in the wireless community [17] [18][19], due to its high potential for on-chip integration, and its flexibility across TX/RX frequency channels and spacing. This frequency independence makes active cancellation a good candidate for tunable duplexer implementation, and for fully overlapped transmit/receive wireless channels for in-band full duplex systems.

The difficulty in adapting this technique from wireline systems to wireless contexts stems from the higher transmit power levels, high-order modulation schemes, lower received signal power which requires high RX sensitivity, and more reflective or longer dispersion TX-RX leakage channels.

A series of active cancellation works were presented by Columbia University [20] [21] [22] [23] [24] [25]. In the first iteration, the TX signal is injected at the gate of a common gate LNTA, providing a cancellation across the V_{gs} of the low noise transconductance amplifier (LNTA) input device. A noise-cancelling common source branch is included, as proposed in [26] and rediscovered in [27][28][29]. The work proposes an additional injection point after the LNTA for cancellation of the TX noise in the RX band. The work is limited to $<+2\text{dBm}$ of TX power, due to the nonlinearity induced in the common gate for a larger injection signal. In this topology, the drain voltage of the LNTA input transistor is still pinned, so only a limited source or gate swing can be suppressed despite the cancellation. Additionally, the work does not integrate the transmitter, and it remains unclear how the TX signal and

the TX noise at the second injection point are coupled from the transmitter. Finally, this work only tests suppression for single tone TX signals without any modulation.

In the second iteration, the TX signal is passed through an n-path filter [30] to control the amplitude and phase of the injected TX signal at the RX input, extending the cancellation to a modulated TX bandwidth. This approach is limited to $<-4\text{dBm}$ of TX power, potentially due to the nonlinearity induced in the n-path based filter. For 20MHz TX modulation, only 20dB of isolation is provided. Lastly, this work also does not integrate the transmitter, and where the TX signal is coupled from the transmit path in practice is unclear.

In [25], a circulator based on [31] is implemented in CMOS, again using n-path filters. While this work provides 42dB of TX/RX isolation, is it limited to -6dBm of TX power, and incurs a relatively high 10.9dB noise figure. This work again does not integrate the transmitter, and accordingly TX to RX leakage channel may be inaccurately modeled. Many secondary coupling paths, for example via the substrate, can arise when both transmitter and receiver are integrated on the same die, adding a difficult frequency selective portion to the TX/RX coupling path.

Another work targeting a full-duplex implementation is described in [32][33]. This work is relatively similar to [21], where a vector modulator couples the TX signal into the baseband virtual ground of a passive mixer first receiver. This work provides around 27dB of isolation for TX signals, at a limited TX power of $<+1.5\text{dBm}$. Similar to the Columbia work, this does not integrate a transmitter, and accordingly does not address the issue of where the TX signal is coupled.

A different active cancellation approach is presented in [34]. The work segments and embeds the PA within an artificial transmission line, with the antenna placed on one side of the transmission line, and the receiver on the other as in Fig. 1.7. The individual PA signals are manipulated such that the TX signals sum out of phase on the antenna side, and in phase on the RX side, providing TX/RX isolation. Additionally, as the TX is in shunt with the receiver, an n-path based degeneration is proposed on the transmitter output stage to avoid excessive RX loss or loading.

This work can operate the RX and TX simultaneously up to around $+14\text{dBm}$ of TX output power, while providing around 25dB of TX-RX isolation. One main downside is this limited isolation, potentially due to the required matching or resolution between sub-PAs to achieve out-of-phase summing at the receive port. The work also demonstrates a rather high $>12\text{dB}$ NF. One guess is the transmitter's phase noise falling in the receive band does not sum out of phase at the receive port. Additionally, the work is measured at a rather large TX/RX duplex spacing of 115MHz. This is potentially because the work relies on a first order n-path based degeneration to raise the TX impedance present in the RX band, which reduces shunt loss for the receive signal. At a closer TX/RX spacing, the lowered TX impedance may further desensitize the receiver.

A chart summarizing the above, and other works [35][36][37] [38][39] in active cancellation is shown in Fig. 1.8, where the y-axis plots the TX-RX isolation in dB, and the x-axis shows the maximum TX power the cancellation network is able to process. This work attempts to develop a system pushing to the top right of the chart, with high cancellation at high power,

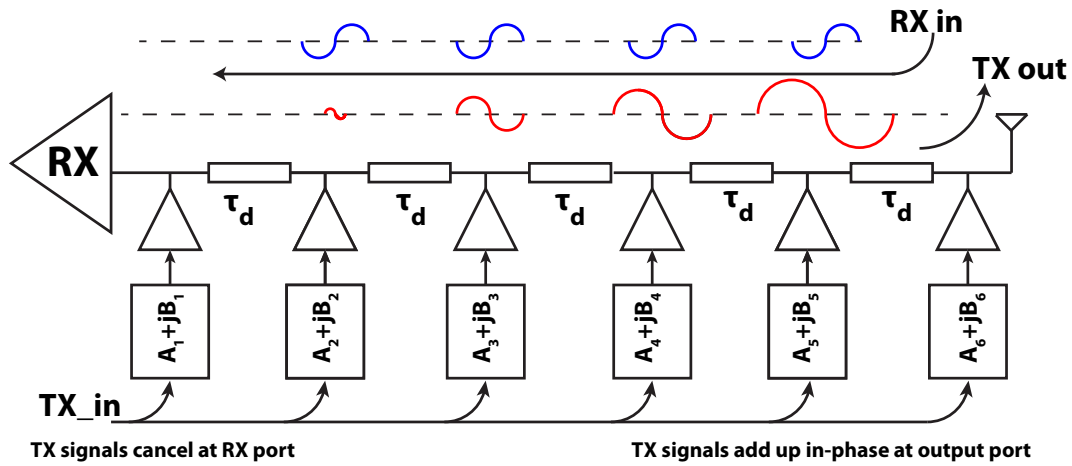


Figure 1.7: Distributed Tline duplexer.

towards the goal of a practical active cancellation transceiver meeting existing standard specifications. This requires around 50dB of isolation at greater than 20dBm of TX output power.

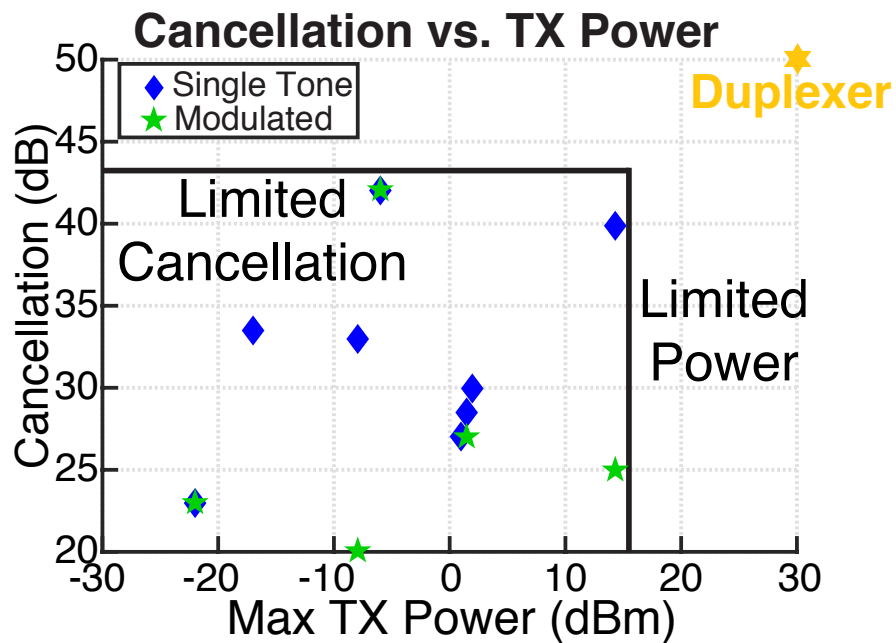


Figure 1.8: Limitations of prior work.

Table 1.2: Operating frequency range of recently published wideband receivers.

Receivers	Murphy, ISSCC12	Fabiano, ISSCC13	Lin, ISSCC15	Borremans, JSSC11	Andrews, JSSC10	Wu, RFIC15
Operating Frequency (GHz)	.08-2.7	1.8-2.4	.1-1.5	.4-6	.1-2.4	.4-3.4

Table 1.3: Operating frequency range of recently published wideband transmitters.

Transmitter	Imanishi, ASSCC09	Wang, RFIC14	Kuo, ASSCC14
Operating Frequency (GHz)	.9-3	.5-6	.6-2.4

1.4 Wideband Transceivers

It is worth questioning if self-interference is truly the limiting factor in enabling an agile universal radio, capable of handling multiple bands and standards across the spectrum. Namely, is it possible that the transmitter and receiver chains themselves simply cannot be built to be tuned across a wide frequency range?

This problem of designing a universal transceiver has generated research interest over the past 5-10 years. On the receiver side, the primary push has been for higher linearity receivers, which can successfully receive a desired signal in the presence of blockers without front-end off-chip filters. This is a difficult problem in a wideband frequency agile system, due to the large tuning range over which blocker rejection must be provided. Research in this space has focused on current-mode receivers, where the linearity issue stemming from limited voltage headroom in a CMOS process is avoided by first converting the input signal into a current. This current is down-converted by highly linear switching passive mixers [40][41], which drive a low-input impedance filter to maintain linearity prior to re-converting to a voltage signal. Variants on this technique include work by [42][43] to eliminate the front-end current converter (LNTA) and drive the passive mixers directly from the antenna to push the linearity even further, as well as work by [29] [28] which attempts to reduce the noise penalty in the passive-mixer-first approach with a differential noise sensing and cancellation path. Such work has demonstrated widely frequency tunable radio receiver designs, at least in a research context, summarized in Table 1.2.

On the transmit side, the difficulty is in maintaining high efficiency, high output power, and low harmonics and spectral emissions without bandwidth limiting high Q passive components for impedance transformation. A few of these are summarized in Table 1.3 [44] [45][46][47][48].

As seen through the above tables, relatively mature research methods can push the transceiver chain to cover the majority of the low frequency wireless bands. It is not the receiver or transmitter which provides limits for a reconfigurable radio [49] [50]. A look at the board components in Fig. 1.9 suggests that the main limitation is due to the narrowband,

discrete, off-chip components and filters used to provide filtering and interference rejection for the receiver at the antenna interface [51]. Integration of the interference rejection filters would unlock the potential of these widely frequency-tunable transceivers.

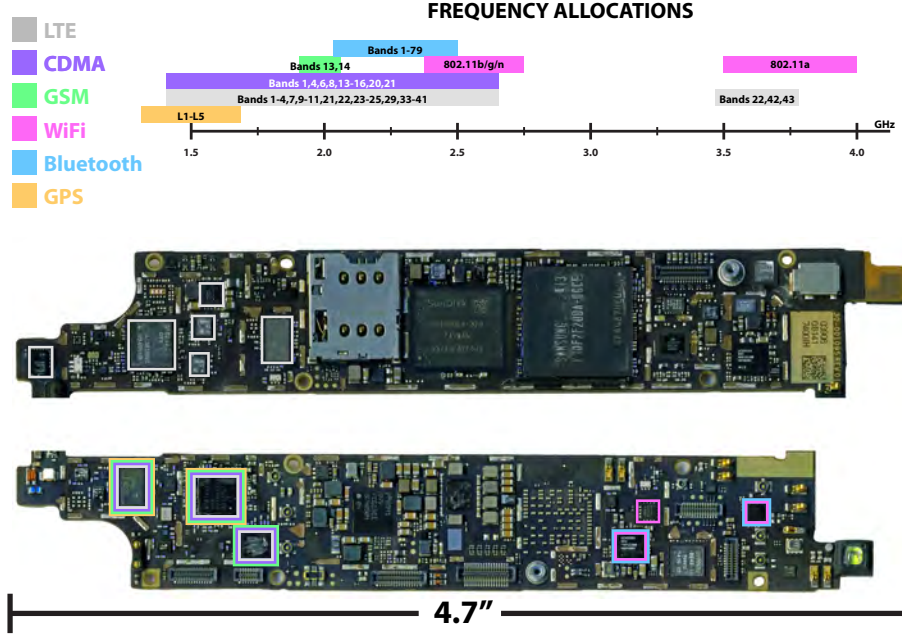


Figure 1.9: Frequency spectrum (top) and board components (bottom) for various standards.

1.5 Research Goals, Scope, and Organization

The key goals of this thesis are 1) the description of an active cancellation system which pushes towards the goal of fully integrating the front-end duplexer on chip 2) analysis of the key fundamental performance bounds of this system, many of which are generalizable considerations in all active cancellation systems 3) implementation of the system in a chip prototype 4) design of a new high linearity receiver in a second test chip, boosting the system performance. It is worth noting that the system described in this thesis could be used as a transceiver with overlapped TX/RX frequency channels, though not directly tested as such. The performance of the test prototype pushes the state of the art in electronic subtraction systems in maximum handled TX power, TX/RX isolation, as well TX/RX bandwidth and spacings.

This work attempts to suppress the transmit signal to below the compression point of the receiver, without introducing significant analog desensitization of the receiver. In particular, thermal noise and phase noise mitigation are key considerations. While still additional TX residual suppression, particularly TX quantization noise in the RX band, is needed to

maintain receiver sensitivity, this residual is sufficiently small such that the receiver still operates linearly in its presence. Accordingly, further digital back-end cancellation at the receiver output can be designed with relatively simple models of the circuit components to subtract the remaining residual. This two stage analog/digital cancellation approach of Fig. 1.10 has been described in [17].

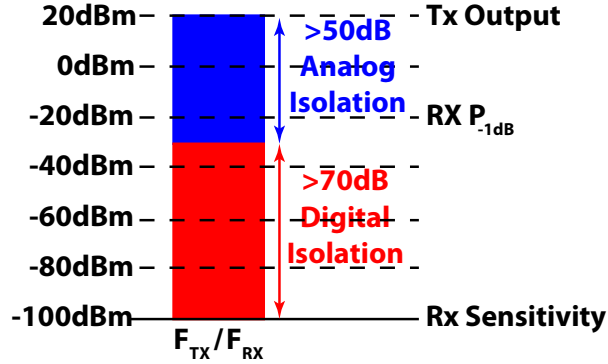


Figure 1.10: Requirements for two stage (analog/digital) cancellation.

The remainder of this work focuses on the analog cancellation, while preliminary exploration of digital cancellation is described in [52].

Towards this end, description of the analog front-end of a transceiver system and its fundamental performance bounds are described in Chapter 2. In particular, schemes to mitigate the impact of thermal noise, and phase noise, and the power tradeoffs in the system are described.

A prototype chip implementing the system is detailed in Chapter 3, demonstrating over 50dB cancellation for 20MHz modulated TX signals at +12.6dBm.

The performance was found to be limited by receiver linearity - accordingly, a second chip prototype is designed, as described in Chapter 4. A passive mixer first receiver with complementary class-AB amplifiers results in an IIP3 of +25Bm, enabling cancellation of a TX power up to +16dBm with <1dB of RX gain compression.

Chapter 5 provides some concluding remarks and future directions for this work.

Many aspects of this system were designed in collaboration - accordingly [52] is cited repeatedly throughout this thesis for relevant details.

Chapter 2

Analysis of the Active Cancellation System

In this chapter, a transceiver architecture is proposed, which enables simultaneous operation of the TX and RX on a single antenna at closely spaced or even overlapped frequencies. Section 1 describes the architecture, while the remainder of the chapter describes the challenges in mitigating the impact of the transmitter and cancellation circuits on the receiver sensitivity. In particular, this work describes techniques to manage the thermal and phase noise.

As will be apparent from the architecture, care must be taken in sampling the signal, and a digital backend must apply further digital cancellation of the residual TX signal to restore full receiver performance. This digital correction is described in further detail in [52]

2.1 Description of Proposed System

The proposed frequency division duplex transceiver is shown in Fig. 2.1.

The transmitter and receiver are connected in series with the antenna through series stacked transformers for impedance tuning. An RF current DAC, as in [53] and [54], is placed in shunt with the receiver to provide the isolation.

The cancellation current DAC acts as a controlled current source, reproducing the TX current flowing through the antenna that would have to be induced on the RX side of the transformer. As the current induced in the RX secondary is shunted by the cancellation DAC, only the residual difference in current, a small amount set by the resolution of the cancellation DAC, flows into the RX load impedance. This is the only portion of the current which can generate a swing across the RX port, and accordingly, the TX swing induced in the receiver is simply the magnitude of this residual current times the receive load impedance.

As a simple intuition, for a cancellation DAC with $N_{BITS,DAC}$ bits of resolution:

$$P_{res} = (P_{TX,MAX} - 6N_{BITS,DAC})dBm. \quad (2.1)$$

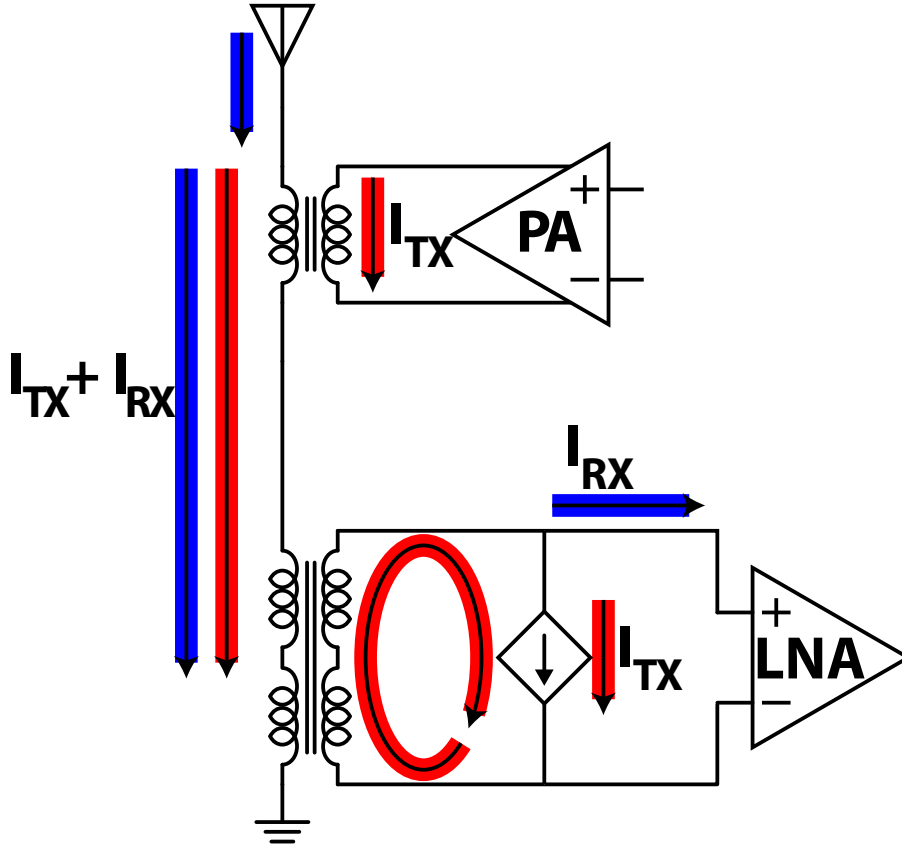


Figure 2.1: Simultaneous TX/RX interface.

Note that this scheme produces a constant residual interference at the RX input, set by the resolution of the delivered cancellation current, with no dependence on the instantaneous TX power level. The $P_{TX,MAX}$ term only appears as a constraint that the DAC full-scale current must be sized to handle this power. Intuitively, this differs from a standard duplexer which provides a fixed amount of rejection independent of TX power, rather than a fixed residual power.

As little swing is generated by the transmitter across the receive port, the port appears as a TX signal-dependent virtual ground, effectively shielding the receiver from the large transmit swing. Note that for any other current, the cancellation DAC does not create the virtual ground condition. The virtual ground condition also helps to maintain cancellation DAC linearity. For the DAC current to scale linearly across amplitude codeword, the DAC must maintain a high output impedance relative to the load impedance, to avoid a code-dependent current division. When cancellation is enabled, the DAC drives a low impedance

virtual ground node, mitigating the current division and enabling linear performance for even moderate DAC output impedance.

In addition to shielding the receiver from the large transmit swing, the virtual ground also shields the transmitter from the receiver's load impedance. Note that due to the shunt cancellation, the transmitter can not induce a voltage change across the receive port, independent of its current or the RX load impedance. Accordingly, the transmitter appears to be connected directly between the antenna and ground, with no receiver present, allowing the transmitter's efficiency to be unaffected by the series connection.

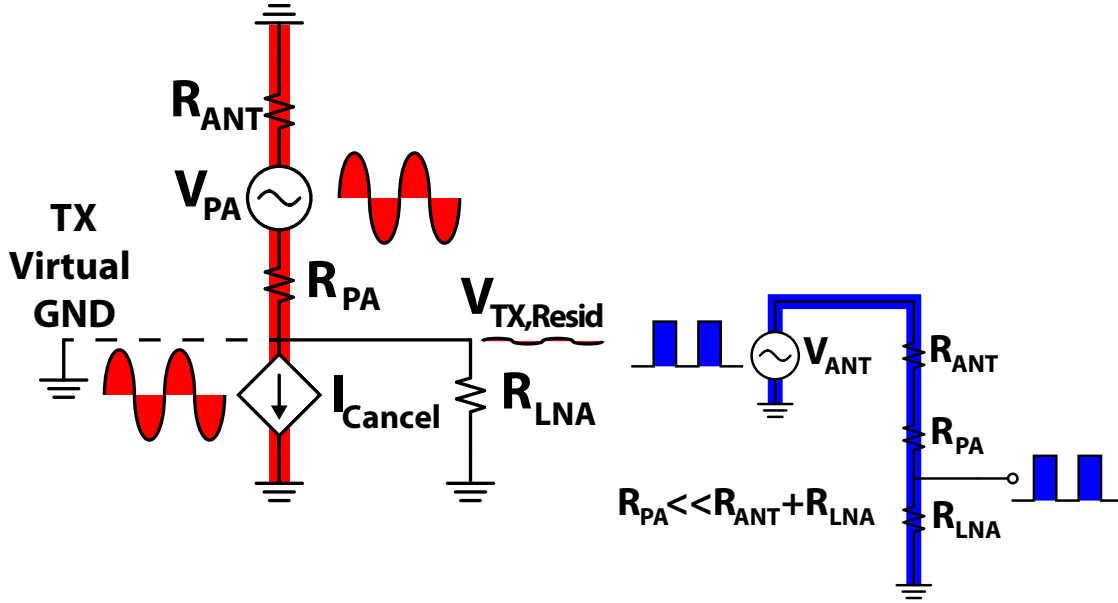


Figure 2.2: Models for TX operation (left) and RX operation (right).

On the receive side, the transmitter's impedance appears in series with the antenna connection, and the DAC's impedance appears in shunt. Accordingly, a transmitter with low series output impedance in the RX band is desired. A large impedance forms a voltage divider, resulting in RX side insertion loss. Additionally, it is desired that this TX impedance is TX amplitude independent, to avoid a time varying voltage division. For example, the voltage at the receiver input port can be simply expressed as

$$V_{input} = \frac{V_{rx} R_{LNA}}{R_{LNA} + R_{ant} + R_{pa}(A_{pa})} \quad (2.2)$$

where the $R_{PA}(A_{PA})$ term results in a TX/RX distortion product. For this work, the switched capacitor RF-DAC, described in Chapter 3, is chosen as a transmitter in order to

maintain a low, amplitude-independent impedance.

The current DAC impedance must ideally be significantly higher than the receiver input impedance in the RX band to minimize shunt loss. In this work, this is handled by cascoding the DAC, and resonating it's capacitance with the transformer inductance around the desired operation frequency.

At first glance it may be conceptually confusing that the cancellation DAC and transmitter both act to create the same current at the RX port, and are placed symmetrically with respect to the antenna on a series connected transformer coil. The question arises of how a cancellation can occur across the receive port, while not at the antenna port. Intuitively, the DAC acts as a controlled current source, which breaks the superposition based intuition leading to the above conclusion. Because the same current would be present through the RX transformer secondary with or without the presence of the DAC, the DAC does not excite any changing flux across the transformer, and the signal at the antenna is not disrupted. Equivalently the symmetry is broken by the impedance difference between the main TX structure and the cancellation structure.

In order to realize the floating current source of Fig. 2.1, this work builds a differential DAC, with a DC common mode current drawn from the transformer center tap. The current source headroom can be reduced by lowering this center-tap supply, and the cancellation could occur with little additional power consumption. The key point is that the voltage swing across the DAC is not fundamentally coupled to the TX output voltage, but rather the residual voltage swing after cancellation. This allows the DAC's power consumption to be substantially reduced with respect to the transmitter. In practice, some voltage headroom must be maintained across the transistor such that it behaves as current source. This results in a power/noise tradeoff, in that reducing the supply headroom and squashing the current source increases its noise and lowers its output impedance, desensitizing the receiver. This tradeoff is analyzed in subsequent sections.

It is worth noting that an equivalent dual structure could be used, where the PA and replica are connected in shunt with the antenna connection, and a series voltage replica is used to provide the isolation, conceptually shown in Fig. 2.3. This allows the use of a high (as opposed to low) impedance transmitter, if desired. This series cancellation scheme is harder to realize than its shunt counterpart, due to the difficulty in designing the floating voltage DAC in series with the receive input, as well as generating a large voltage swing potentially above the transistor breakdown voltage.

The notion of a shunt cancellation DAC placed across the receiver to cancel the TX could also be applied even without the TX and RX directly stacked in series, for example, in a system with multiple antennas, or with a relaxed isolation external duplexer interface.

It can be conceptually argued that this architecture is capable of overcoming two main limitations of prior electronic subtraction works, namely, the cancellation of high TX powers, and cancellation for wide modulation bandwidth TX signals.

As mentioned above but re-iterated here, in the proposed topology, the replica current is fed directly into the low impedance TX virtual-ground node. Accordingly, no single node in the system, aside from the antenna side of the transformer, which contains no active

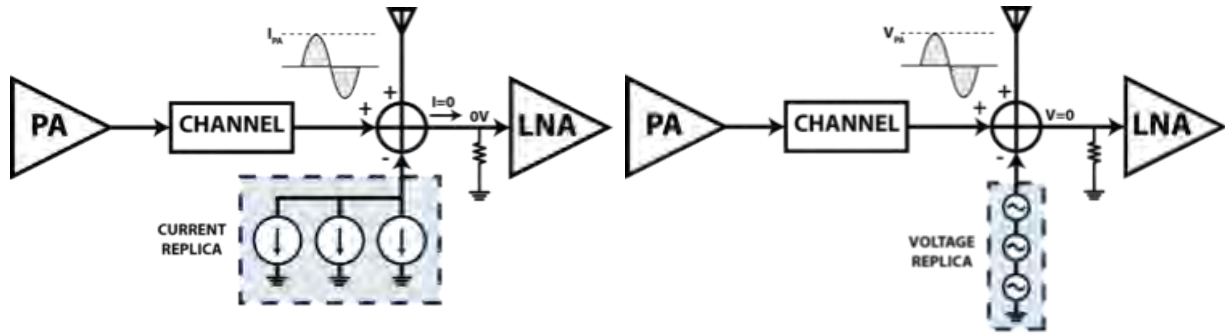


Figure 2.3: Current mode (left) and voltage mode (right) cancellation.

devices, experiences a large voltage swing. In order to handle a +20dBm TX signal, the DAC must slew a current on the order of 50mA, which is possible to implement linearly in CMOS technology.

The difficulty in providing cancellation across a wide modulation bandwidth stems from the need for the cancellation network to mimic the frequency dependence of the TX/RX coupling path [21]. As the TX residual at the RX input must fall well below the receiver compression point, the analog front-end canceller, rather than a digital back-end canceller, must handle the frequency dependence.

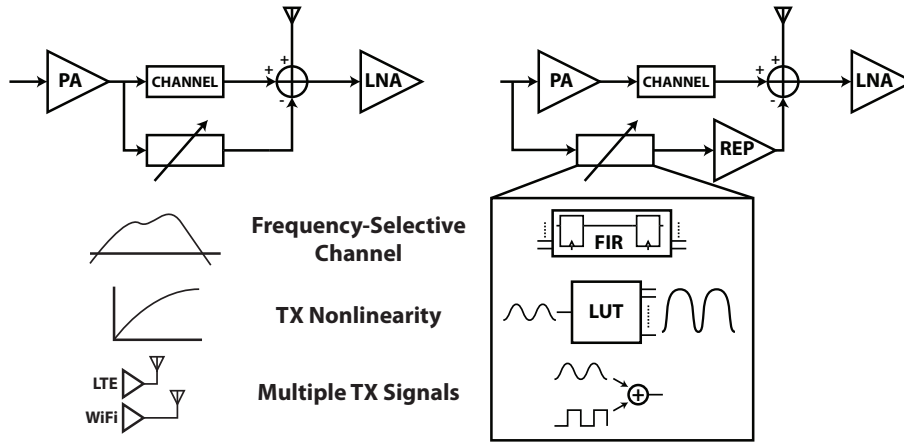


Figure 2.4: Digital feedforward adaptation possibilities.

As a DAC is used to provide the cancellation, the input data can be digitally adapted or pre-distorted to match the frequency dependence. For example, a digital FIR filter can shape the DAC baseband data to match the frequency dependency of the leakage channel. Nonlinearity present in the PA or main TX to RX coupling path can be matched with a lookup-table, or other nonlinear DAC pre-distortion mechanism. Furthermore, if a system

contains multiple co-located transmitters (for example, a mobile terminal with LTE and WiFi), the transmit signals could be digitally summed at the DAC input to provide isolation for both aggressors simultaneously. Further elaboration of these concepts, and in particular the implication on the number of filter taps and the DAC sampling rate in this system are described in [52].

2.2 Replica Power Consumption

The replica cancellation must drive a current with no explicit voltage swing requirement, whereas the transmitter's 50Ω drive requirement couples its current and voltage swing, as shown in Fig. 2.5. Accordingly, there is no theoretical lower bound on the amount of power the replica needs to consume. In practice, as described in subsequent sections, lowering the cancellation DAC power trades off with the system's noise figure.

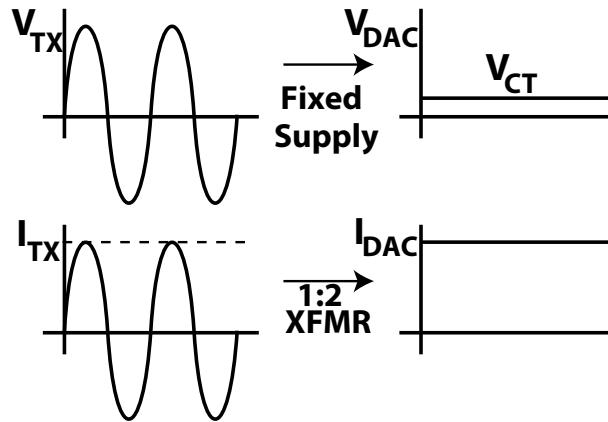


Figure 2.5: TX vs. replica power consumption.

Nonetheless, despite the noise implication, the cancellation DAC can still consume significantly less power than the transmitter. In particular, as will be described, the DAC's power consumption scales with square root of the TX power consumption, making it a smaller percentage of the total system power at higher TX powers.

The circuit model of Fig. 2.6 is used to describe the replica canceller's power as a function of the TX power from the supply. The model consists of a baseband current up-converted by an RF mixer, with the DAC current drawn from the RX transformer's center tap at a voltage V_{CT} .

Note that this structure implements a floating current between the transformer terminals with 2 currents to ground in a differential structure, resulting in a power penalty due to the common mode component of the current. If the single-ended DAC produces a total current of I_{DAC} , which is switched from one differential side to the other, the output current consists

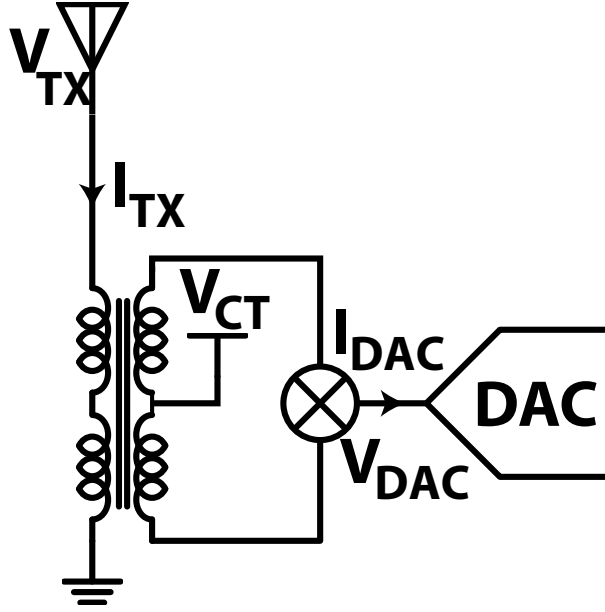


Figure 2.6: DAC's current and voltage.

of a common mode component of $\frac{I_{DAC}}{2}$, and a differential mode component of $\pm \frac{I_{DAC}}{2}$. This is the differential amplitude of a square-wave current, due to the assumption of a hard switched mixer. The amplitude of a sine wave with equivalent content in the fundamental frequency as this square wave is $4/\pi$ times larger. Accordingly, if the TX current I_{TX} induces a differential current $I_{TX \rightarrow RX}$ on the RX side of the transformer, the required current for the DAC to cancel this is:

$$I_{DAC} = 2I_{TX \rightarrow RX} \times \frac{\pi}{4}. \quad (2.3)$$

The TX current on the antenna and RX sides of the transformer are related very simply by the turns ratio

$$I_{TX \rightarrow RX} = \frac{I_{TX}}{N}. \quad (2.4)$$

The TX power is the TX current scaled by the TX drain, efficiency, as

$$P_{TX} = \frac{1}{2} \times I_{TX}^2 R_{ANT} \times \frac{1}{\eta_{drain}}. \quad (2.5)$$

Accordingly, the TX power normalized by the cancellation power is

$$\eta_{cancellation} = \frac{N \times I_{TX} \times R_{ANT}}{\pi \times \eta_{drain} \times V_{CT}}. \quad (2.6)$$

$$\text{Rewriting } I_{Tx} = \sqrt{\frac{2P_{TX}}{R_{ANT}}} ,$$

$$\eta_{cancellation} = \frac{N \times \sqrt{2P_{TX}} \times \sqrt{R_{ANT}}}{\pi \times \eta_{drain} \times V_{CT}} . \quad (2.7)$$

This efficiency increases with increased turns ratio and increased TX power at the antenna impedance, and decreases with TX drain efficiency and transformer center tap voltage. The efficiency increase with decreased TX drain efficiency is a bit misleading, as this simply states that the cancellation contributes less to the total system power as compared with the transmitter as the transmitter performance is reduced. The transformer turns ratio and center-tap supply voltage are the only free design parameters, as TX power and antenna impedance are fixed by the system.

Both of transformer turns ratio and transformer center-tap voltage trade-off against other system considerations. As will be shown later, decreased center tap supply voltage can result in decreased RX sensitivity. Increased transformer turns ratio incurs several difficulties. Realizing a high turns ratio transformer with low loss is difficult. The requirement on the DAC output impedance to maintain low RX insertion loss from the shunt DAC increases quadratically, difficult to support with limited V_{CT} . Additionally, increasing the turns ratio linearly increases the RX voltage swing, stressing the RX linearity requirements given spurious TX emissions or RX blockers.

The TX power from the supply, and replica power consumption, are plotted against the TX power at the antenna in Fig. 2.7, under the specific assumptions of a 1:2 turns ratio, 1V center tap supply, and 50% peak drain efficiency transmitter.

2.3 Thermal Noise

The thermal noise falling into the RX band due to the transmitter, as well as the cancellation replica, is analyzed here.

2.3.1 TX Thermal Noise

As stated, earlier, the transmitter must provides a low, code-independent output impedance in order to minimize the insertion loss on the receive side. The noise variance due to thermal noise produced by the transmitter is linearly proportional to the real part of the series output impedance. Accordingly, the appropriate noise model pictured in Fig. 2.8, with the series combination the transmitter, antenna, and receiver.

The noise figure could be computed as the transmitters noise at the receiver normalized by the antenna's noise, as:

$$10 \times \log_{10} \left(1 + \frac{N_{RX}^2}{N_{ant}^2} + \frac{4kT\gamma R_{TX} \left(\frac{R_{RX}}{R_{ant} + R_{RX} + R_{TX}} \right)^2}{4kT R_{ant} \left(\frac{R_{RX}}{R_{ant} + R_{RX} + R_{TX}} \right)^2} \right) \quad (2.8)$$

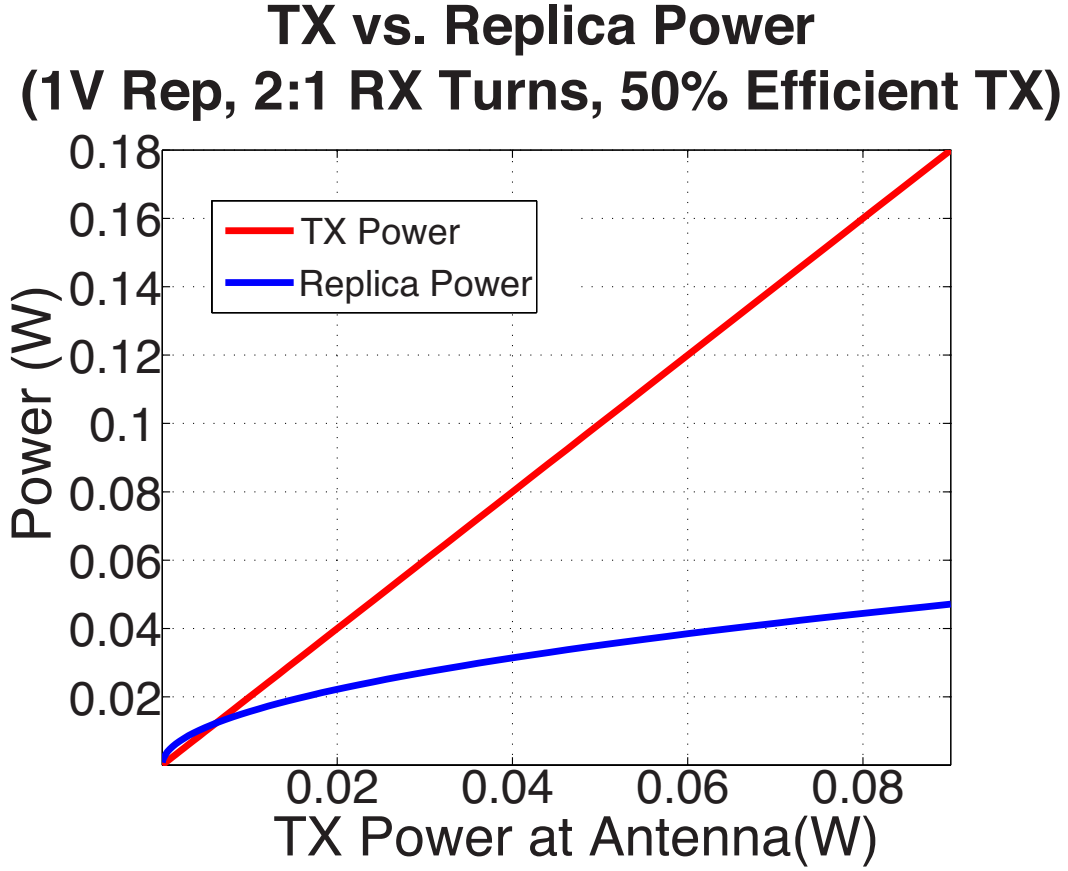


Figure 2.7: TX vs. replica power consumption.

$$= 10 \times \log_{10} \left(1 + \frac{R_{TX}}{R_{ant}} + \frac{N_{RX}^2}{N_{ant}^2} \right) \quad (2.9)$$

where $\frac{N_{RX}^2}{N_{ant}^2}$ is the ratio of the receiver's input referred noise variance normalized to the antenna noise variance. Because $R_{TX} \ll R_{ant}$, this term is relatively small. Concretely, if the receiver nominally had a 3dB noise figure $\frac{N_{RX}^2}{N_{ant}^2} = 1$, and $\frac{R_{TX}}{R_{ant}} = \frac{10\Omega}{50\Omega}$, the added noise figure would be $< .5dB$.

2.3.2 Replica DAC Thermal Noise

Due to the large current required from the cancellation DAC, and its position at the RF input before any active signal gain, the DAC's thermal noise can be significant in dictating the RX sensitivity.

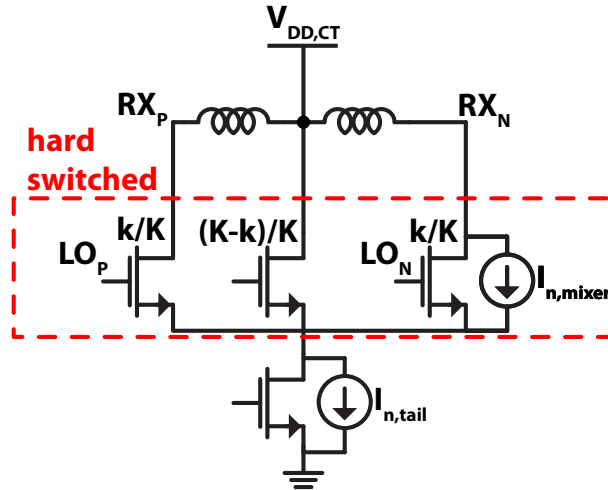


Figure 2.9: Model for DAC noise analysis.

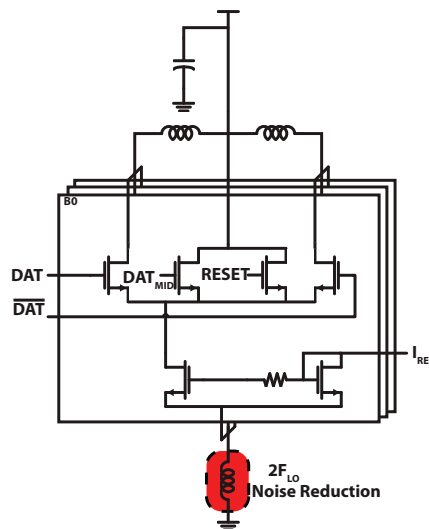


Figure 2.10: Reduction of the high frequency noise via degeneration.

cancel a given TX power P_{TX} and transformer turns ratio N_{turns} is

$$I_{DAC} = \sqrt{\frac{2P_{TX}}{R_{ant}}} \frac{\pi}{2N_{turns}}. \quad (2.10)$$

Due to the mixer switches, the differential noise at F_{RX} at the output comes from tail current source noise at F_{duplex} up-converted by the TX LO, and noise at $2F_{TX} + F_{duplex}$ down-converted by the TX LO.

The down-converted $2F_{TX} + F_{duplex}$ noise can be handled by a resonant degeneration of the tail, which circulates the high frequency tail noise without affecting the low frequency DC tail current. Note that in the DAC model, the tail transistor is constantly on, and the amplitude codeword simply controls the amount of current routed to the output. Accordingly, the tail device transconductance is simply dependent on the maximum current. Given an LC degeneration, the tail noise experiences a current division between $\frac{1}{gm}$ and the degeneration impedance Z_s , resulting in a noise division factor of

$$T_{noise,2flo} = (1 + 2 \frac{\sqrt{\frac{2P_{TX,max}}{R_{ant}}} \frac{2}{N_{turns}\pi}}{V^*} \times Z_s)^2 \quad (2.11)$$

In order to operate over a fundamental frequency of 1GHz to 2GHz, the tail must provide a resonance from 2GHz to 4GHz to mitigate the $2F_{LO}$ noise, implying a Q of roughly 1.5. This high frequency tail noise is downconverted by a $\frac{1}{\pi}$ conversion gain. Substituting for Z_s at resonance yields a total $2F_{LO}$ noise as a function of P_{TX} , assuming the tail is sized to cancel $P_{TX,max}$, of

$$i_{noise,2flo}^2 = \frac{4kT\gamma}{\pi^2} \frac{2 \frac{\sqrt{\frac{2P_{TX,max}}{R_{ant}}} \frac{2}{N_{turns}\pi}}{V^*}}{(1 + 2 \frac{\sqrt{\frac{2P_{TX,max}}{R_{ant}}} \frac{2}{N_{turns}\pi}}{V^*} \times \frac{w_0 L}{Q} (1 + Q^2))^2}. \quad (2.12)$$

The added noise figure degradation due to this term, assuming a 3dB nominal RX NF is shown in Fig. 2.11. The plot demonstrates that the degeneration provides a good reduction of noise at high codes. Even at the peak value, this terms degradation can be made <.5dB using this technique.

Folding of higher harmonics contribute <1.5dB due to their lower conversion gain.

Focusing now on the low frequency component of the DAC noise upconverted to the RX frequency, the gm noise at the RX input is expressed as:

$$\frac{i_{n,unit}^2}{\Delta f} = 8kT\gamma \frac{I_{DAC}}{V^*} A_{conv}^2 \quad (2.13)$$

The baseband noise is unconverted by the upper sideband of the switching square wave to the RX frequency. Due to the hard switching, half of the noise appears half in differential mode across the RX, and half in common mode. As the fundamental component of the

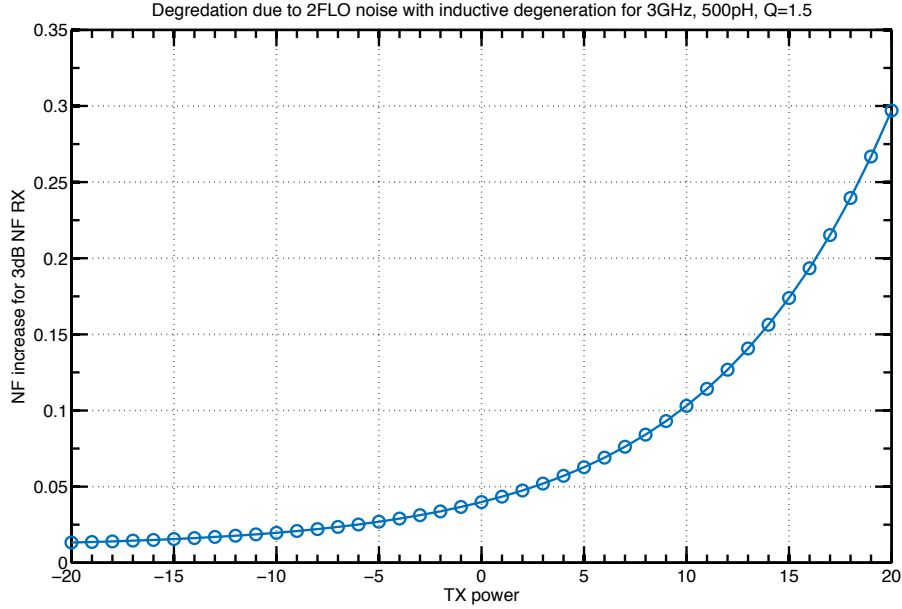


Figure 2.11: Contribution of DAC $2F_{LO}$ noise with inductive degeneration.

square wave from -1 to 1 is $\frac{4}{\pi} \cos(w_o t)$, the upper sideband component is $\frac{2}{\pi} e^{-jw_o t}$, and the differential component is $\frac{1}{\pi} e^{-jw_o t}$, resulting in a conversion gain of $A_{conv} = \frac{1}{\pi}$.

Substituting the DAC current in terms of the TX power, and the conversion gain, yields a noise of:

$$\frac{i_n^2}{\Delta f} = 4kT\gamma \sqrt{\frac{2P_{TX}}{R_{ant}}} \frac{1}{N_{turns}\pi} \frac{1}{V^*} \quad (2.14)$$

The DAC in this work uses a 25% I/Q cell-sharing technique, where each unit cell outputs a 50% (I=Q) or 25% (I=0 or Q=0) square wave, causing the output noise to be dependent on the phase of the output current. This is described in more detail in Chapter 3. In the above scheme at 0° phase angle, a 25% pulse is sent, and the expression for output noise for a phase of 0° is $\sqrt{2}$ lower than for 45° . If the DAC is segmented as all thermometer bits, a reasonable simplification because thermometer cells are the largest and therefore dominate the noise, the total DAC noise as a function of phase angle can be rewritten as:

$$\frac{i_n^2}{\Delta f} = 4kT\gamma \sqrt{\frac{2P_{TX}}{R_{ant}}} \frac{1}{N_{turns}\pi} \frac{1}{V^*} \times (|\cos(\phi)| + |\sin(\phi)|), \quad (2.15)$$

and the final DAC noise figure over phase angle can be written as

$$F_{tot} = F_{RX} + \gamma \frac{2N}{\pi} \frac{1}{V^*} \sqrt{R_A P_{TX}} (|\cos(\phi)| + |\sin(\phi)|). \quad (2.16)$$

A plot of this DAC noise contribution vs the TX output power at the antenna is shown in Fig. 2.12.

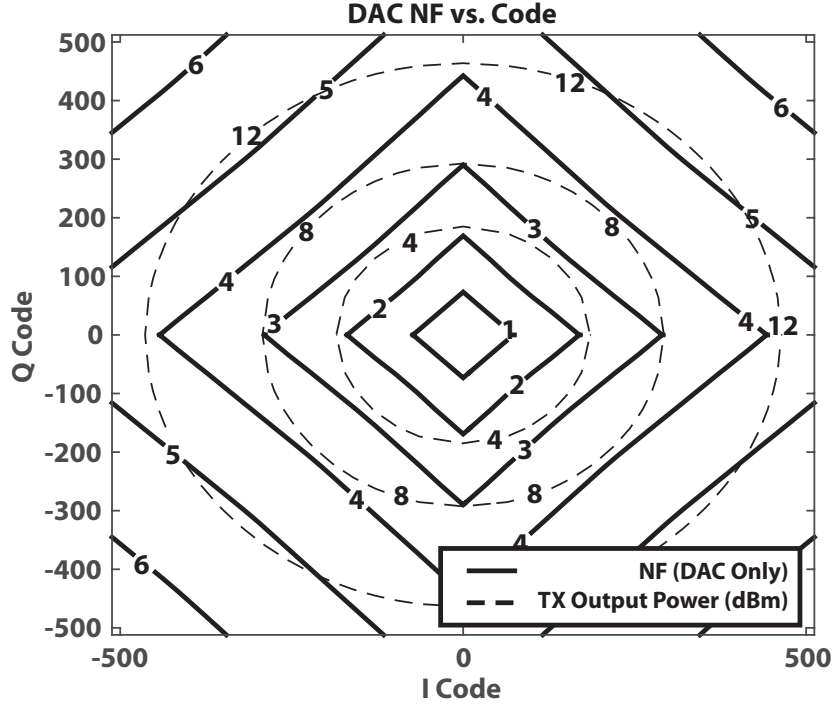


Figure 2.12: DAC thermal noise contour vs. TX Power.

Note that the overdrive of the transistor, V^* , appears in the denominator of the noise figure expression. In order to maintain the high impedance of the current source device, its supply voltage must be at least this voltage V^* . This results in a noise/power tradeoff in the DAC design - reduction of the supply necessitates a corresponding decrease in the transistor overdrive, which increases the current source noise. Additionally, the noise figure scales with the square-root TX power, as opposed to linearly with power. As will be shown in the following sections, this contrasts with phase noise, making phase noise the dominant effect at high TX output powers.

Reduction of this DAC's fundamental thermal noise falling in the RX band is possible through exploiting the frequency and spatial separation of the noise and the signal within the DAC. As the tail current is constantly on, the effective tail signal is at DC. The tail noise sources contributions to the RX band are at higher frequency, providing some opportunity for filtering of the noise.

Note that the tail current flows through the transformer center tap without any frequency translation. If the DAC is class-A and maintains a constant drain current, then any non-DC component in this current corresponds to thermal noise of the tail source. A low frequency

supply resonance at the duplex spacing translates the RX band noise current to a proportional voltage at the transformer center-tap. This noise voltage could be used to suppress the tail noise in either in a feed-forward cancellation, or by wrapping the tail source in a feedback loop. In this work, this center tail voltage is tied to the tail current source gate in a feedback loop, to provide the noise reduction, as shown in Fig. 2.13.

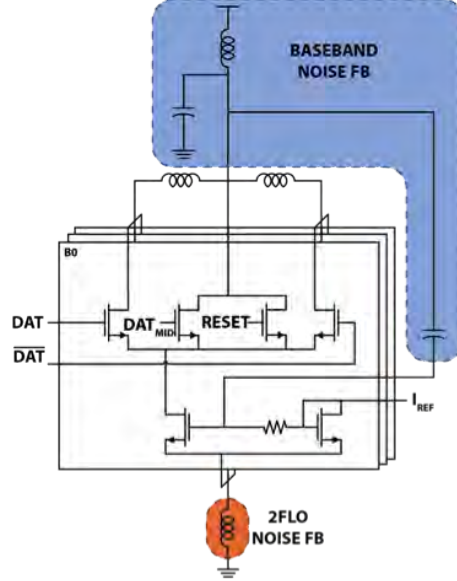


Figure 2.13: Baseband noise feedback loop.

There are two intuitive explanations for how this connection provides noise reduction. Wrapping the noise in a feedback loop reduces its contribution by $1 + T$, where T is the loop gain, set by the DAC tail source gm and the resonant impedance. Equivalently, in the limit, the resonant impedance acts as a current source, providing a high impedance at the duplex frequency. The DAC tail source then looks like a diode-connected device. While the noise of the tail current source remains the same, the noise reduction at the output comes from the lowered DAC impedance presented at the RX frequency ($1/gm$ as compared with the original r_o) relative to the receiver's input impedance. The resulting current division acts to circulate the thermal noise current, providing a reduction in noise figure.

To analyze this noise reduction technique, and in particular its performance vs. DAC code, the model in Fig. 2.14 is used.

The voltage at the center tap V_X can be written as:

$$V_X = (i_{n,diff} - gm_{diff}V_X)Z_{CT} + (i_{n,mid} - gm_{mid}V_X)Z_{CT}. \quad (2.17)$$

Solving for V_X yields:

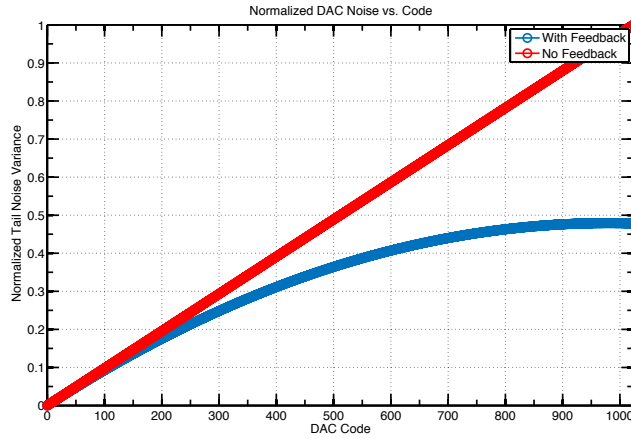


Figure 2.15: Reduction of the low frequency tail noise via feedback.

all codes, the true noise improvement can be quantified by overlaying the plot with the expected signal statistics.

Finally, note that there is a tradeoff between the bandwidth and magnitude of the noise reduction through the Q factor of the center-tap resonance. Lower Q increases the bandwidth of the resonance, but adds a noise component from the lossy center tap impedance. The noise reduction technique, its implications on the DAC circuit design, selection of passive values for the center tap impedance, and measurement results with the noise feedback enabled, are described in further detail in [52].

2.4 Phase Noise

2.4.1 Background

The transmitter's phase noise, which can fall in the receive band and enter the receiver, can be a dominant source of desensitization in this architecture. To give context for the analysis of phase noise in this work, some relevant definitions of phase noise are described here. A more fundamental analysis can be found in [55] and [56].

In particular, the relevant intuitions are the difference between phase noise and voltage noise, the latter of which is measured directly from a spectrum analyzer, and the effect when a square wave rather than a sinusoidal waveform is used for mixing.

Phase noise is a model of the difference between the imperfect LO with the ideal LO. A sinusoidal LO can be written as an ideal cosine plus a voltage noise, as

$$V(t) = \cos(\omega t) + n(t). \quad (2.22)$$

This can be decomposed into a phase noise and amplitude noise term, in terms of the noise voltage $n(t)$, as

$$\cos(\omega t) + n(t) = (1 + A(t)) \cos(\omega t + \phi(t)) \quad (2.23)$$

where the term $\phi(t)$ is the phase noise term, and the term $A(t)$ describes the amplitude noise. In an assumption where the $\phi(t)$ term is small, we can approximate this expression as

$$\cos(\omega t) + n(t) \quad (2.24)$$

$$= (1 + A(t)) \cos(\omega t + \phi(t)) \quad (2.25)$$

$$\approx (1 + A(t))(\cos(\omega t) + \phi(t) \sin(\omega t)) \quad (2.26)$$

$$\approx \cos(\omega t) + A(t) \cos(\omega t) + \phi(t) \sin(\omega t) \quad (2.27)$$

and therefore

$$n(t) = A(t) \cos(\omega t) + \phi(t) \sin(\omega t). \quad (2.28)$$

Accordingly, because the amplitude and phase noise terms are up-converted with sine and cosine, they can be decomposed as the even and odd portions of the voltage noise spectrum up-converted around the carrier, as

$$A(t) = \frac{N(f - fc) + N(-f + fc)}{2} \quad (2.29)$$

$$\phi(t) = \frac{N(f - fc) - N(-f + fc)}{2} \quad (2.30)$$

$$. \quad (2.31)$$

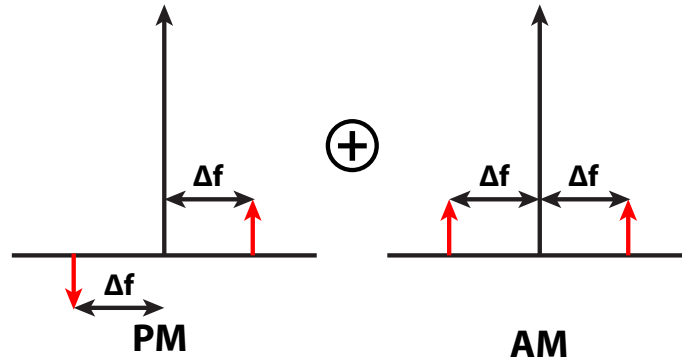


Figure 2.16: PM and AM noise components.

The takeaway is the conversion between voltage noise of the LO signal (as measured by a spectrum analyzer) and the LO's phase noise. The phase error is the down-converted odd portion of the voltage noise spectrum.

This decomposition is relevant, because when the LO waveform passes through a limiting buffer, such as an inverter, or drives a hard switched transistor as in most mixers, the amplitude component is rejected by the voltage limiting, and only the phase component is passed. Accordingly, the phase noise determines the remaining noise at the output.

In a parallel intuition for square waves, once the amplitude component of the error has been rejected, the voltage error signal can be described as a series of pulses spaced T_{period} apart, with the pulse width of the k^{th} period as ΔT_k . This error signal could be further approximated as a series of deltas spaced T_{period} apart, with power proportional to the product of ΔT and full-scale voltage.

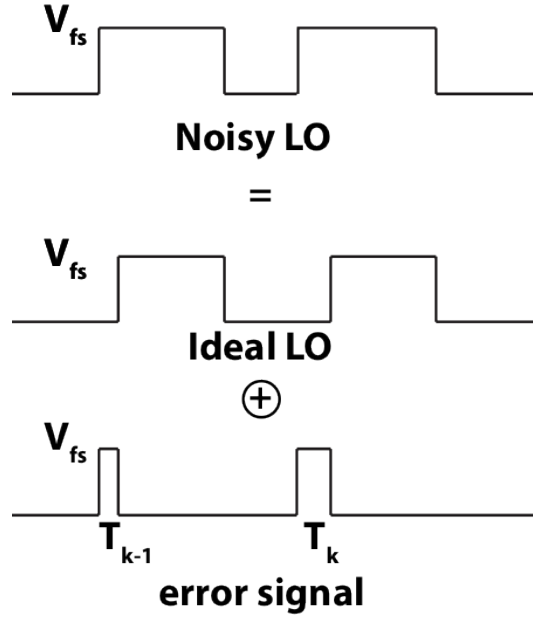


Figure 2.17: Phase noise as voltage pulses.

The discrete-time phase error signal at edge crossing k is defined as proportional to the time ΔT as

$$\phi(k) = \Delta T_k \times \omega \quad (2.32)$$

and can thus be written in terms of the sampled voltage noise power with scaling factors of ω and the full scale voltage.

Accordingly, in this model, where the amplitude component of the noise is rejected (for example if the difference between the high and low voltage levels is much greater than the

variance of the noise process), the voltage noise and phase noise can be thought of as roughly the same to within constant scaling factors, and this quantity can, for example, be measured directly with a spectrum analyzer.

In a hard-switched mixer, the output voltage can be written as

$$V_{RF} \times (LO_{ideal} + LO_{error}) \quad (2.33)$$

If the model above where amplitude noise is rejected holds true, and the voltage noise power is a scalar times the sampled phase noise process, the relevant phase error spectrum is the power spectral density (PSD) of the sampled phase noise process, ie. the Fourier transform of the autocorrelation of the discrete-time sequence of errors. Its worth nothing that this phase noise process is periodic in frequency over f_s , because it is a sampled process. This contrasts with the initial sinusoidal cases, where the phase noise is defined over all time points, not just the edge crossings. Accordingly, in the sinusoidal case, the continuous autocorrelation function is the relevant metric, which results in an aperiodic PSD.

To get the sampled phase noise spectrum at the edge crossings for a square wave LO from a sinusoidal LO, we can just sample the continuous time phase at the edge crossings. If the noise spectrum falls off sharply by f_s , then sampling aliases are small, and the continuous and discrete-time phase noise spectra are effectively the same.

Equivalently, the voltage noise could be sampled at the edge crossings with the above scaling applied. This links the square wave pulse model noise as a sampled version of the sinusoidal noise with the amplitude component rejected. If amplitude noise in the sinusoidal model is removed, resulting only in phase noise, and if this noise is bandlimited to $f_s/2$, the sampling folds this noise to around DC, giving a discrete-time phase noise spectrum which is equivalent to sampling the original continuous time $n(t)$ spectrum at the edge crossings.

It is often convenient to summarize the noise of the square wave in a single number, the jitter. There are many jitter metrics - period jitter is the deviation of the LO from the ideal LO, consistent with the variance ΔT_k in the above model. Accordingly, given the link between autocorrelation function and PSD, the jitter is the integral of the sampled phase noise process times a constant scaling term. Equivalently, the jitter can then be written in terms of the voltage noise variance at the edge crossing (solved through integrating the voltage noise PSD), normalized by the waveform slope (in V/s), which describes how much time uncertainty a voltage perturbation adds.

Finally, the LO phase noise is normally quoted in terms of the dBc/Hz, as the voltage signal's noise power in a one Hz bin normalized by the carrier power.

$$L(\Delta\omega) = 10 \times \log\left(\frac{P_{sideband}(\omega_0 + \Delta\omega, 1Hz)}{P_{carrier}}\right) \quad (2.34)$$

If the measured waveform is not amplitude limited, then this definition, somewhat confusingly, actually contains both amplitude and phase noise components. If the waveform is amplitude limited it is a direct measure of the phase noise, as the amplitude noise component has been rejected. This formulation is convenient firstly because it can be directly measured

with a spectrum analyzer, and secondly because the normalization simplifies calculations when this LO is mixed with blocking signals. As the time domain output of the mixing operation is a multiplication of the LO with the blocking waveform, the frequency domain output is a convolution. This allows us to simply add the phase noise power spectral density to the blocker power and integrate over the relevant blocker bandwidth. For example, for a phase noise of -173dBc/Hz, a 0dBm LO, and a 10dBm blocker, the resulting noise would have a total power -163dBm spread over the blocker's bandwidth.

2.4.2 A Quick Note on Simulation

There are two types of simulations relevant for phase noise characterization. The first is to measure average PSD, which is the average of the fourier transform of the instantaneous autocorrelation functions $R(t, \tau)$, τ going from $-\infty$ to ∞ . This is essentially what is measured by spectrum analyzer. As the resolution bandwidth of the spectrum analyzer is generally much smaller than the periodicity of the noise sequence, the spectrum analyzer averages over a long time window relative to the noise periodicity. Accordingly the time varying nature of the noise statistics is hidden.

The second type of simulation is the instantaneous PSD, which is a family of PSD's of discrete-time signals sampled T_s apart.

The first measurement is useful if the circuit cares about the error over all instances, as opposed to just the edge crossing. For example if the waveform is used in an ideal multiplying mixer or is sampled asynchronously with the period, average PSD determines the output noise characteristics. The second measurement is useful if only the noise at the edge crossings is relevant, for example in amplitude limited waveforms.

Both time and frequency domain representations for average and sampled PSD's are needed for this work, and can be simulated via Spectre's pnoise simulation. Selection of pnoise:sources provides the average PSD. Selecting pnoise:timedomain:outputnoise:spectrum provides a family of sampled PSDs for each sampling phase, and outputnoise:integ noise power:0 to fs gives the time domain noise variance at each of the sampling phases. Selecting pnoise:jitter:tdnoise: integ output noise is the integrated noise voltage from $\frac{f_s}{2}$ to $\frac{f_s}{2}$ which is essentially the same as the integrated phase noise power above, with a slope factor scaling from time to voltage error.

2.4.3 Impact of Correlated Phase Noise

Given that the TX and DAC can share LO's, the phase noise at the TX and DAC output are correlated. Accordingly, the opportunity exists for feedforward cancellation of the TX LO source phase noise, just as the main signal is cancelled. If not accounted for, the LO source phase noise in the RX band would heavily desensitize the receiver.

For a single-tone input, the analog channel that the TX signal experiences is matched by I/Q weighting of the DAC. For example, for a 90° phase shift in the channel, an I sent on the PA is cancelled by a Q output on the DAC. Accordingly, in order for the phase noise to

cancel, the phase noise present on I and Q must have the same relationship as the I and Q tones - ie. the I phase noise with 90° phase shift must equal the Q DAC noise.

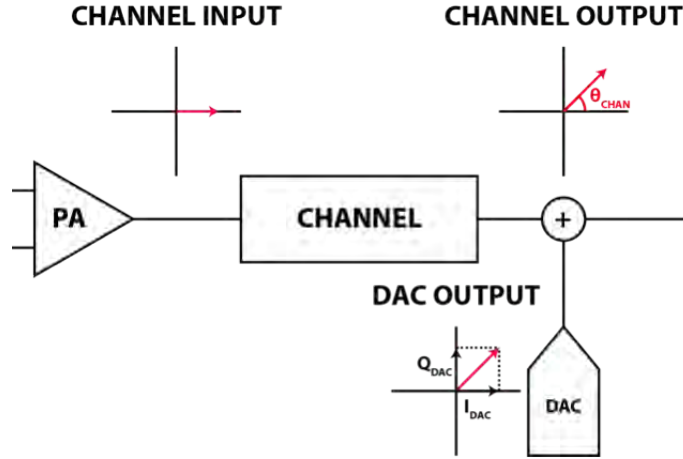


Figure 2.18: Phase noise through digital and analog paths.

Given that a standard on-chip LO chain consists of a 2x clock divider, and a 25% duty cycle generator, a spur input into the LO chain can be tracked to check its phase relationship at the I and Q outputs.

The divider and 25% generation circuit sample the input edges of the source waveform to generate the edges of the differential 25% I and Q waveforms. Some input edges map to the rising edges at the output, and some to falling edges at the output.

For an input rising edge which maps to rising edge on either I or Q , or a falling edge which maps to a falling edge on either I or Q , adding a positive phase delay at the input results in an equivalent positive phase difference at the output. However, for an edge which maps to the opposite transition, adding a positive phase delay at the input results in an equivalent negative phase delta at the output.

This can be seen slightly more intuitively in the voltage domain picture of Fig. ??, by examining the output effect due to a positive voltage noise at the input, and remembering that the voltage noise and phase noise at an edge are simply related by the edge rate. Any positive voltage noise on a rising edge results in the same voltage noise at the output edge if the edge is not inverted, and negative noise if the edge is inverted.

Consequently, given the picture above, the relationship between the input phase noise and output phase noise on the I LO can be computed as taking the sampled phase noise sequence at the input edges and multiplying by the $(\dots + 1 + 1 - 1 - 1 \dots)$ periodic sequence shown on the far right. Similarly, the the output phase noise on the Q LO is the sampled phase noise sequence at the input edges times the $(\dots - 1 + 1 + 1 - 1 \dots)$ periodic sequence on the far right.

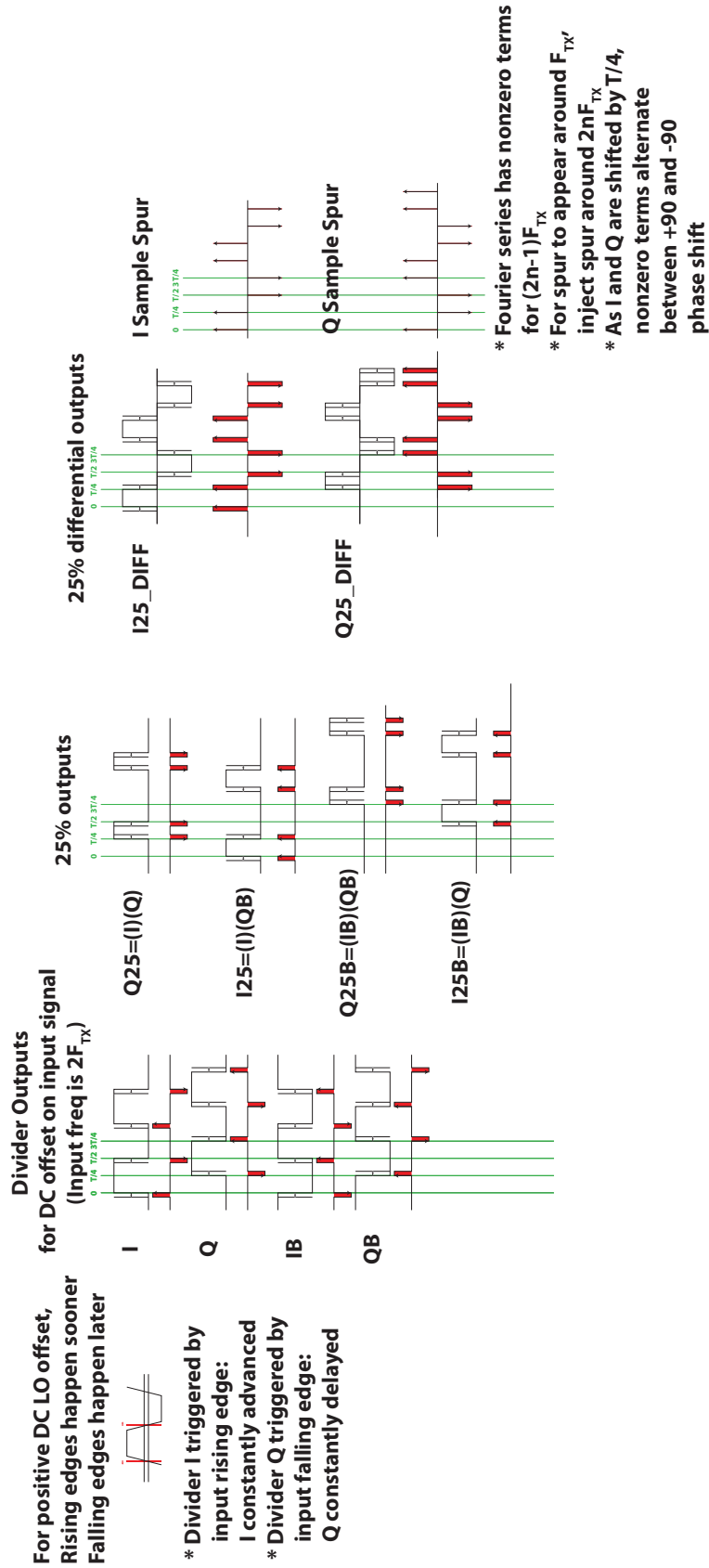


Figure 2.19: Phase noise through the LO divider.

In the frequency domain, this multiplication is convolution of the input phase noise spectrum with an impulse train with strengths given by the Fourier series coefficients of the corresponding sequence.

The Fourier series coefficients for the I sampling sequence and Q sampling sequence can then computed as

$$I_{samplespur} = 2(1 - j^n) \text{ for } n \text{ odd, } 0 \text{ for } n \text{ even} \quad (2.35)$$

and

$$Q_{samplespur} = 2(1 + j^n) \text{ for } n \text{ odd, } 0 \text{ for } n \text{ even.} \quad (2.36)$$

The Fourier series coefficients give insight into which source noise terms fold to around the fundamental, and the phase relationship of the noise between the I LO and the Q LO. The sequences have no DC term, and contain only odd frequency components (at $\pm 1F_{TX}$, $\pm 3F_{TX}$, $\pm 5F_{TX}$, etc). Accordingly, source spurs, or equivalently, phase noise, around even sidebands ($2F_{TX}$, $4F_{TX}$, $6F_{TX}$) are folded to the TX frequency. For $2F_{TX} - F_{offset}$ term to fold to $F_{TX} - F_{offset}$, it is mixed with the -1 sideband, which has a 90° phase shift between I and Q . For $-2F_{TX} - F_{offset}$ term to fold to $F_{TX} - F_{offset}$, it is mixed with the 3rd sideband, which also has a 90° phase shift between I and Q . Accordingly source noise terms around $2F_{TX}$ fold around the carrier with the correct phase shift to be cancelled. Noise around alternating harmonics (ie. $2F_{TX} \pm 4k \times F_{TX}$) will also all fold with the 90° phase difference into I LO and Q LO, due to the alternating 90° / -90° relationship between the fourier series coefficients of $I_{samplespur}$ and $Q_{samplespur}$. However, noise around the other harmonics ($\pm 4k \times F_{TX}$) will fold with a -90° relationship, and accordingly can not be cancelled.

Additionally, it's worth noting that noise injected by a clock divider itself (for example, designed in Fig. 2.20), will not maintain this phase relationship. Because the differential pair input has settled as the clock transitions, there is no noise propagation from the I latch to the Q latch. Rather, the latch tail sources independently add noise to the I and Q LO outputs, and there will be no correlation.

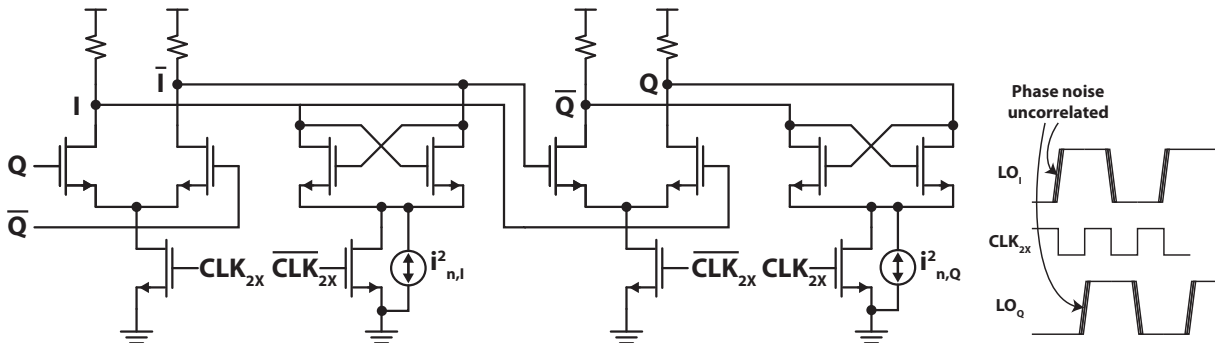


Figure 2.20: Clock divider uncorrelated phase noise.

The summary of the above analysis is that

1) Phase noise components at the source close to the source clock fundamental frequency, and alternating even harmonics (ie. $2F_{TX} \pm 4k \times F_{TX}$) can be feed-forward cancelled

2) Phase noise components at odd harmonics (eg. at $\pm 1F_{TX}$, $\pm 3F_{TX}$, $\pm 5F_{TX}$, etc) do not fold down to around F_{TX} , and do not affect the output noise.

3) Phase noise components at $4F_{TX} \pm 4k \times F_{TX}$ do not fold around the fundamental with the appropriate 90° I/Q phase relationship, and thus this source noise can not be cancelled.

4) Phase noise injected by the clock divider, and any subsequent LO chain components is not appropriately correlated between the TX and DAC, and can not be feed-forward cancelled.

There is additionally a bandwidth limitation on cancellation of the phase, as set by the bandwidth of the analog leakage channel. The DAC I and Q amplitude weights are set to cancel the phase shift at the TX frequency. Accordingly, noise at some offset frequency with a different phase shift will not be reproduced by the cancellation DAC with the correct weights for cancellation. It is beneficial for the leakage network to have as wide a bandwidth as possible.

To give some quantitative intuition to this effect, a single tap delay on a discrete-time sample point has the resulting error:

$$h_1 x(t - \tau) e^{j(\omega_{TX}(t-\tau) + \phi_{jitter}(t-\tau))} - x^*(t) e^{j(\omega_{TX}t + \phi_{jitter}(t))} \quad (2.37)$$

where τ is the delay, h_1 is the tap weight for delay τ , $x(t)$ is the TX signal, and $x^*(t)$ is the chosen DAC signal.

To separate the phase noise decorrelation error from quantization error in $x^*(t)$, the assumption is made that the best possible $x^*(t) = h_1 x(t - \tau) x^{-j\omega_{TX}\tau}$ is chosen. With this choice of $x^*(t)$, the residual error signal from the phase noise de-correlation can be found as:

$$h_1 x(t - \tau) e^{j\omega_{TX}(t-\tau)} (e^{j\phi_{jitter}(t-\tau)} - e^{j\phi_{jitter}(t)}) \quad (2.38)$$

$$\approx h_1 x(t - \tau) e^{j\omega_{TX}(t-\tau)} (1 + j\phi_{jitter}(t - \tau) - 1 - \phi_{jitter}(t)) \quad (2.39)$$

and its power is given by

$$|h_1 x(t - \tau)|^2 (E[(j\phi_{jitter}(t - \tau) - \phi_{jitter}(t))^2]) - E[j\phi_{jitter}(t - \tau) - \phi_{jitter}(t)]^2 \quad (2.40)$$

$$= 2 \times |h_1 x(t - \tau)|^2 (R_{\phi_{jitter}}(\tau) - R_{\phi_{jitter}}(0)) \quad (2.41)$$

Intuitively, the error term is proportional to how much the phase noise changes in a time τ , and the strength of the channel for the delay τ . In the proposed system, the dominant tap is the direct on-chip transformer connection, which has a negligible electrical delay. The bandwidth of phase noise cancellation, further described in the chip measurements section, is then set by the secondary reflections with longer delay [57].

The relevant design heuristics from this section are therefor to minimize the amount of uncorrelated noise between the PA and DAC, (through pushing the clock tree split point as

close to the end of the chain as possible), to minimize the amount of noise added in the clock generation which does not maintain the 90° phase relationship between I noise and Q noise (high frequency phase noise, as well as noise in the clock divider and subsequent buffers), and to design a high bandwidth leakage network if possible.

2.4.4 Impact of Uncorrelated Phase Noise

While the TX/DAC correlated phase noise can achieve a feedforward cancellation, as described in the previous section, the the uncorrelated portion of the noise simply desensitizes the receiver. The constraints on this uncorrelated noise is described in this section.

For the purpose of this cancellation, quantization effects and the channel filtering are decoupled. The channel filtering has a larger effect on the correlated phase noise, as analyzed above. The difference $R(t)$ between the TX and cancellation DAC outputs given an uncorrelated phase noise component $\phi_{jitter}(t)$ and an input signal $x(t)$ is

$$R(t) = x(t)e^{j\omega_c(t)+\phi_{TX}(t)}(1 - e^{j\phi_{jitter}(t)}). \quad (2.42)$$

Given a small jitter term, by a Taylor expansion:

$$e^{j\phi_{jitter}(t)} \approx 1 + j\phi_{jitter}(t) \quad (2.43)$$

and

$$R(t) \approx -j\phi_{jitter}(t)(x(t)e^{j\omega_c t + \phi_{jitter}(t)}). \quad (2.44)$$

Accordingly, the residual error PSD is the convolution of the phase noise PSD with the modulated TX signal PSD.

Because the PSD is a convolution with a band-limited signal, only the phase noise at the TX-RX duplex offset results in output voltage noise in the RX band. In a scenario where there are a few MHz offset between the two, it is reasonable to assume that the phase noise spectrum is white at this offset (ie. flicker noise is not dominant here). The error power in a 1Hz bin at the offset frequency is then given by

$$\int_{\omega_{offset} - \frac{\omega_{bw}}{2}}^{\omega_{offset} + \frac{\omega_{bw}}{2}} 10^{\frac{P_{TX}}{10}} \frac{1}{\omega_{bw}} |\phi_{jitter}(\omega)| d\omega. \quad (2.45)$$

More intuitively, given the white phase noise, the signal power at any frequency is spread to all other frequencies, with the PSD of the phase noise as a weighting factor. Thus, in order to meet some sensitivity requirement P_d , the uncorrelated phase noise must, in dBc/Hz, meet

$$\phi_{jitter}(t) = P_d - P_{TX} \quad (2.46)$$

Concretely, in order to leave the residual error from phase noise at -174dBm/Hz, at the thermal noise floor, for a 13dBm average signal power, the uncorrelated phase noise between the transmitter and replica in the receiver band must be at -187dBc/Hz. Any increase from this adds dB for dB to the RX noise figure.

2.4.5 Phase Noise of an Inverter Chain

In this project, the dominant uncorrelated phase noise is set by phase noise injected after the TX/DAC LO distribution chain split point. This section details an approximation for computing the phase noise of a chain of digital gates, leading to a framework for sizing such gates.

As the inverter chain is a driven system, the noise added to each edge of the inverter chain is independent. Accordingly, the phase noise should not have any correlation from edge to edge, resulting in a white phase noise profile.

This white phase noise profile allows a computation of the phase noise level from a calculation of the jitter. As the jitter is an integration of the phase noise, it is not in general possible to obtain the phase noise spectrum from the jitter. However given a white phase noise profile, the jitter power can simply be spread out over the $\frac{f_s}{2}$ bandwidth to obtain the white phase noise density.

As will be apparent from the equations that follow, jitter of an arbitrary digital gate is approximately the same as an inverter with equivalent rise time and load capacitance. Accordingly, the jitter calculation of a single inverter is sufficiently general to give design intuition.

Finally, note that each inverter in a chain adds an independent jitter to each clock edge. Accordingly, the total jitter at the output is the sum in variance of the jitter added by each gate, and the chain jitter follows from the per-stage jitter.

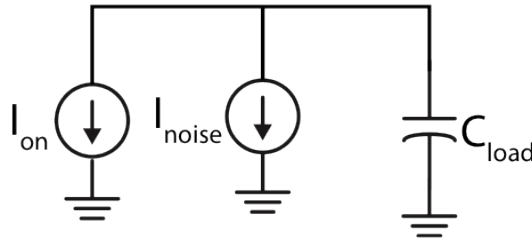


Figure 2.21: Inverter noise model.

During the half supply output voltage transition, an inverter can be modeled as a constant current source charging or discharging a capacitance. The effective load capacitance, which is really nonlinear across the voltage swing, is approximated here as a fixed capacitor which results in the same average edge rate as the nonlinear capacitance. It follows that the edge is then approximated as a linear ramp charging and discharging between 0 and V_{dd} at a rate of

$$T_{edge} = \frac{C_{load}V_{dd}}{I_{on}}. \quad (2.47)$$

Recall that the jitter is the variance of voltage noise at the edge crossing, normalized by the slope factor, which maps a voltage uncertainty to a timing uncertainty. Finding the jitter

is then equivalent to finding the voltage noise variance at the edge crossing. From Fig. 2.22, this voltage noise is composed of 2 components - the sampled voltage noise at the instant the inverter switches, and the noisy switch current integrated onto the load capacitor over the transition period of $T_{edge}/2$. Note that the first effect is due to the PMOS pulling the capacitor up to the power supply, and the second is due to the NMOS discharge current. To obtain the total voltage noise variance, these two effects can simply be added in variance, as the sampled noise and the NMOS noise current are uncorrelated error sources.

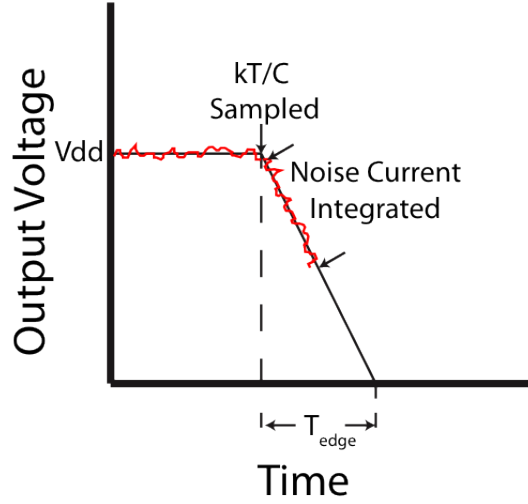


Figure 2.22: Noise at the inverter edge crossing.

The first component of this voltage noise has variance $\sigma_{v,n}^2 = \frac{kT}{C_{load}}$, so the resulting jitter is

$$\sigma_{jitter,1}^2 = \frac{\sigma_{v,n}^2}{S^2}, S = \frac{V_{dd}}{T_{edge}} \quad (2.48)$$

$$= \frac{kT}{C_{load}} \times \frac{T_{edge}^2}{V_{dd}^2} = kT \times \frac{V_{dd}}{I_{on} T_{edge}} \times \frac{T_{edge}^2}{V_{dd}^2} \quad (2.49)$$

$$\sigma_{jitter,1}^2 = \frac{kT}{I_{on} V_{dd}} \times T_{edge}. \quad (2.50)$$

There are two things to note from the form of this first jitter term. Firstly, the jitter variance increases linearly with the edge rate. A slower edge corresponds to a larger timing uncertainty for the same voltage noise. However this is a linear penalty in terms of variance. Secondly, the denominator term refers roughly to a power consumption, implying that reducing jitter requires an increase in power consumption, for example by adding more capacitance and maintaining a fixed edge rate.

The second jitter term can be computed similarly. From the model of Fig. 2.22, the voltage noise at the edge crossing is a windowed integration of the noisy current source onto the load capacitance over a time window of $\frac{T_{edge}}{2}$. The noise current is well known in its spectral density, as

$$\sigma_{i,n}^2 = 4kT\gamma gm\Delta f \quad (2.51)$$

As the time domain voltage noise is an integration of the current noise over a time window corresponding to half the edge, in the frequency domain, the voltage noise spectral density is $\frac{1}{C_{load}}$ times a sinc filter. The voltage noise variance is then the integration of this voltage noise spectral density from 0 to ∞ . Performing this integration to get the noise voltage:

$$\sigma_{v,n}^2 = \frac{1}{C_{load}^2} \times \int_0^\infty 4kT\gamma gm \left| \int_0^{\frac{T_{edge}}{2}} e^{-j\omega t} dt \right|^2 df \quad (2.52)$$

$$\sigma_{v,n}^2 = \frac{1}{C_{load}^2} 4kT\gamma gm \times \frac{T_{edge}}{2} \quad (2.53)$$

Using the above definition,

$$\sigma_{jitter,2}^2 = \frac{\sigma_{v,n}^2}{V_{dd}^2} T_{edge}^2 \quad (2.54)$$

and the relations

$$gm = \frac{2I_{on}}{V_{dd} - V_{th}} \quad (2.55)$$

$$C_{load} = \frac{I_{on} T_{edge}}{V_{dd}} \quad (2.56)$$

yields

$$\sigma_{jitter,2}^2 = \frac{4kT\gamma}{I_{on}(V_{dd} - V_{th})} \times T_{edge}. \quad (2.57)$$

This jitter variance is again decreased by increasing power, and increases linearly with the edge rate. The total jitter of a chain of stage is then the sum of jitter across each stage.

As previously mentioned, because the phase noise in this case is simply the jitter power spread over a bandwidth of $\frac{f_{lo}}{2}$, this equation allows for a guide to sizing the LO distribution chain to achieve a desired phase noise target.

2.4.6 Sizing Under Constrained Phase Noise

Given the above jitter expression for a single stage inverter, a digital chain can be sized to minimize its power consumption for a desired target phase noise. Again the jitter of a digital gate is approximated to be the same as that of an inverter with equivalent delay and load capacitance. Accordingly, the gate complexity is only factored into its additional power consumption. The single stage delay and power consumptions are simply expressed as below, where t_{inv} represents the technology RC delay, p represents the stage intrinsic delay, LE is the logical effort, f is the stage fanout, γ_{cap} represents the ratio of drain capacitance to gate capacitance. From the above model, α is 2, but is curve fit against circuit simulation for an inverter to account for the time-varying I_{on} across the output transition.

$$t_d = t_{inv}(p + LE \times f) \quad (2.58)$$

$$P = (C_{self} + C_{load}) \times V_{dd}^2 \times f \quad (2.59)$$

$$\phi_n^2 \propto k \times \frac{t_d^\alpha}{(C_{self} + C_{load})} \quad (2.60)$$

The problem is accordingly to choose the fanouts f_i for each of the i stages which minimize the $\Sigma(P_i)$ under a constrained $\Sigma\phi_n^2 < C$. Finding a closed form optimum for an arbitrary number of stages is difficult. To provide intuition, the problem is solved numerically for a representative example of length 4 inverter chain, driving a 1pF load capacitance. This is done through enumerating the total jitter and power consumption while sweeping the fanout of each stage from 1 to 3.

First, the minimum power over the possible fanout combinations for the 4 stage chain is plotted against the desired phase noise in Fig. 2.23. At a high phase noise target, the stage fanouts are simply pushed to the largest fanouts taken in the optimization. Note that in the limit of a very high phase noise constraint, simply minimum-sized gates would be selected. Accordingly, the power consumption flattens out towards the power required to switch the explicit load capacitance and the minimum per stage intrinsic capacitance. At low phase noise targets, the gate edge rate can not be readily improved, as it is set by the self loaded limit with low per stage fanouts. In this regime, the phase noise is only decreased by increasing the gate size to boost the per stage capacitance, resulting in a super-linear power increase for a linear phase noise improvement.

Additionally, the power consumption for fanouts which provide an equal per stage phase noise is superimposed on the curve of Fig. 2.23 in red. A clear heuristic arises that to minimize power for a given phase noise constraint, the per stage phase noise contributions should be made roughly equivalent.

The stage fanouts required to obtain the minimum power for a given phase noise are shown in Fig. 2.24, where stage 1 is the first stage, closest to the input, and stage 4 is closest to the load capacitance. The curves show that to minimize power consumption for a fixed phase noise, the stages should be tapered from end of the chain to beginning of chain. This is an intuitive result - as the stage phase noise is inversely proportional to load

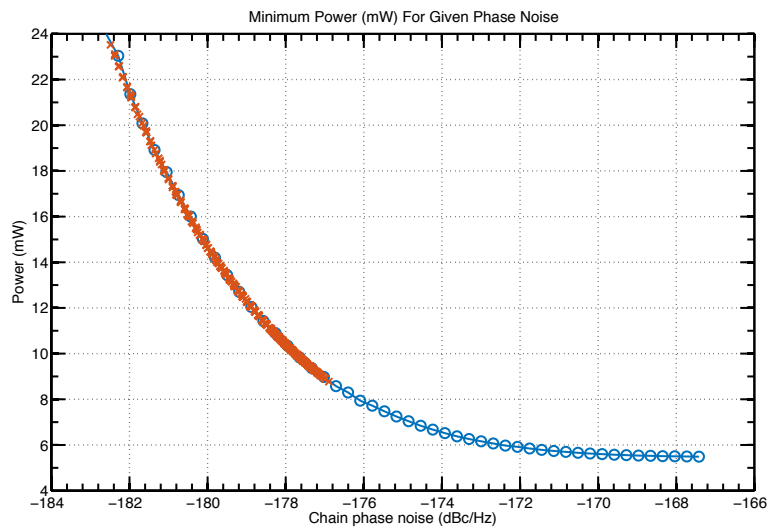


Figure 2.23: Minimum power for a length of 5 digital inverters.

capacitance, stages with higher absolute load capacitance can be designed with higher fanout while incurring a relatively lower phase noise penalty.

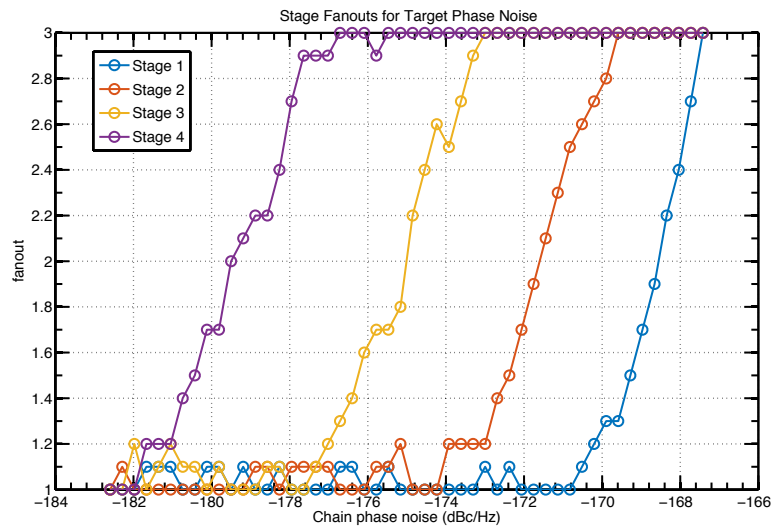


Figure 2.24: Per stage fanouts to achieve minimum power for target phase noise.

2.5 Noise Summary

A graphical summary of the above noise analysis is shown in in Fig. 2.25. The noise figure degradation of the receiver is shown in red as the TX power is swept, decomposed as a superposition of thermal and phase noise components.

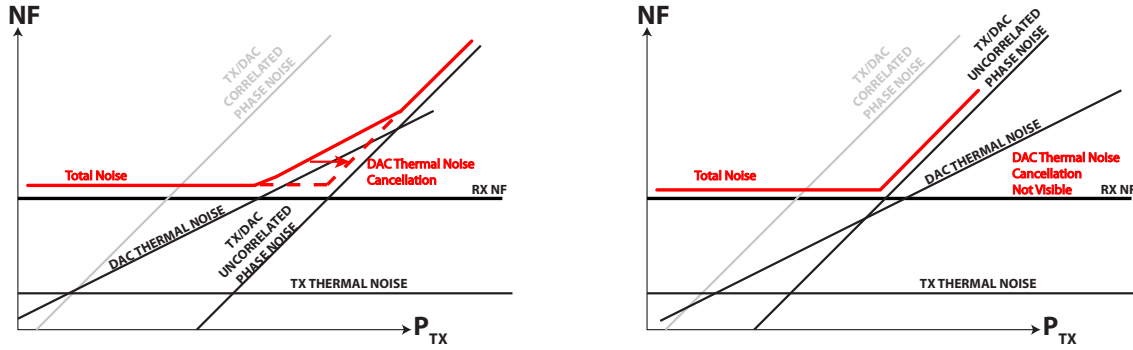


Figure 2.25: Receiver noise scaling vs. TX power. Right - DAC thermal noise appears at moderate powers. Left - Uncorrelated phase noise appears first.

The receiver contributes a constant noise floor, dominating the sensitivity at low TX output powers. The TX thermal noise is a small, code-independent contribution. As TX power increases, the TX and DAC phase noise in the RX band increases. However, this phase noise is initially dominated by the source LO phase noise, which is correlated between the TX and DAC and can be feedforward cancelled. At moderate TX powers, either the uncorrelated TX/DAC phase noise, or the DAC thermal noise appears, depending on the on design parameters chosen. In particular, both uncorrelated phase noise and DAC thermal noise can be pushed lower with increased power consumption. If DAC thermal noise dominates at moderate output powers, shown in Fig. 2.25 on the left, then the described thermal noise feedback loop can reduce the noise figure in this region, pushing the curve to the right, towards the phase noise dominated regime. Instead, if uncorrelated phase noise appears first, as on the right in Fig. 2.25, the thermal noise feedback loop has limited benefit. Finally, recall the the phase noise contribution scales dB for dB with TX power, while the DAC contribution scales .5dB for dB, proportional instead to TX current. Accordingly, in the high power regime, the uncorrelated phase noise dominates the noise figure increase.

2.6 Quantization Noise

The wide-band quantization noise of the DAC is another source of desensitization of the receiver. In particular, if the DAC could be designed with high enough resolution, the

residual error between the DAC output and the TX signal could be made less than the noise floor. In practice, this is quite difficult.

The power spectral density of the quantization noise floor set by the DAC as referred to the antenna port is expressed as

$$P_{quant} = 10\log_{10}\left(\frac{1}{2} \times I_{DAC}^2 R_{ant}\right) - 10\log_{10}\left(\frac{F_{sample}}{2}\right) - 10\log_{10}\left(\frac{1}{2}\right)^2 \times N_{bits,DAC} \text{ dBm/Hz} \quad (2.61)$$

Taking a representative case of 10bits, 1GSPS, and a max TX power of 20dBm, the quantization noise floor is still 30-40dB above the RX noise floor.

In theory, given the high sampling rate of the DAC relative to the 20MHz modulation bandwidth, this difference could potentially be bridged with noise shaping via a digital delta-sigma modulator, though this is certainly nontrivial. In particular, a digital delta-sigma with multi-bit DAC is challenging (due to the linearity requirements on the DAC or alternately, necessary modification of the digital quantizer in the loop), the noise must be shaped at an offset frequency (TX band signal vs. RX band noise), and DAC out of band emissions due to the noise shaping may be radiated through the antenna, violating the TX spectral mask requirement. Resolution of these challenges is an option that could be pursued in future works.

Another option is to exploit the fact that quantization error as shown in Fig. 2.26 is theoretically a data correlated time sequence, and to do further cancellation of the residual signal in the digital domain. For example, for a fixed code, the static error of a DAC is a (perhaps slowly time varying) fixed error. Accordingly, if the DAC code to signal characteristic is measured with sufficiently high resolution, this code dependent error can be subtracted digitally from the residual signal in the digital domain.

This is consistent with other active cancellation works, in that the limited TX isolation leaves a residual signal, which must be cleaned up digitally.

Accordingly, in this work first focuses on reducing the TX signal in the TX band in the RF/analog domain far enough to avoid receiver desensitization, while not injecting significant unknown noise sources (thermal, phase, etc.) which can not be handled in the digital back-end. These noise sources which can not be handled later are quoted as the noise figure degradation in the remained of this work. This can be separated from quantization error by sending periodic (such as sinusoidal single-tone) sequences, such that the quantization error is also periodic tones, and can distinguished from the residual noise floor.

If this is accomplished, the TX signal in the RX band is sufficiently small so as to not de-linearize the receiver, and can accordingly be digitized with reasonable dynamic range. It is assumed that a digital block can then be built, which will predict this TX signal in the RX band from TX data sequence, and perform a cancellation in the digital domain.

Note that although the TX quantization noise is ignored in the above description, the problem is conceptually identical - the true quantization error is, rather than simply the DAC quantization error, the difference between the TX and DAC quantization errors. Accordingly, the problem remains the same - a table mapping input TX and DAC code to total TX minus

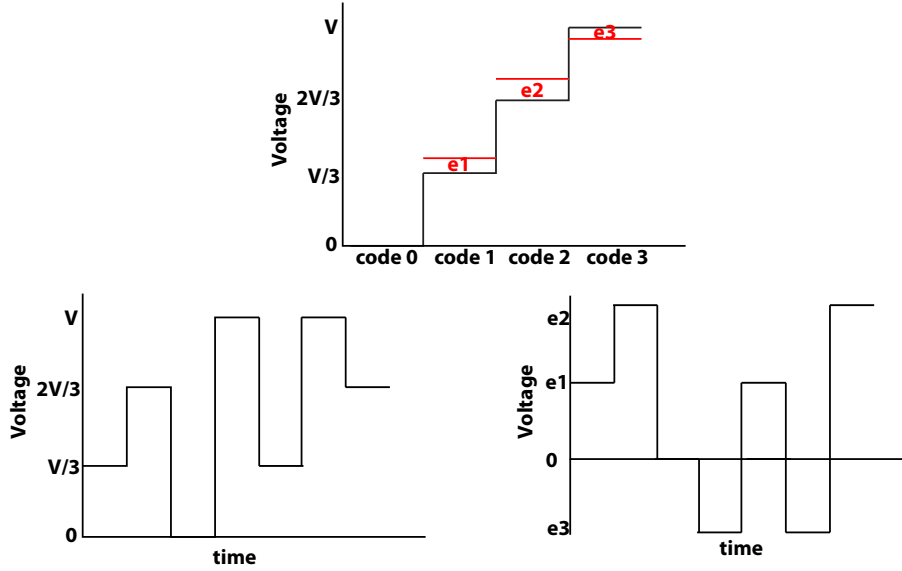


Figure 2.26: Top: DAC code vs voltage characteristic, Left: data sequence Right: quantization error is data correlated.

DAC quantization error can be used to predict the residual error after analog cancellation. We have used this approach in [52] to demonstrate an additional 25 to 30dB cancellation of the residual signal at the RX output in the digital domain, corresponding to lower the DAC's quantization noise floor by 4-5 bits.

Further work on digital correction techniques as applied to this work are covered in [52].

A final solution to the quantization noise may be to provide some partial external isolation at the antenna interface. If 15dB of external isolation was provided, the DAC full scale current could be reduced, and the quantization noise could be pushed below the RX sensitivity level for a 20MHz modulated bandwidth through design of a 12 to 13 bit DAC with 50x oversampling.

2.7 Sampling Rates

In this system, the analog leakage channel of the PA waveform to the input of the receiver is matched with a digital filter applied to the input of the cancellation DAC. Because the DAC outputs a zero-order value over an entire sample period, while the analog leakage signal can vary over this time window, the difference is non zero in between sample spaced points. Intuitively, this error is a function of the data and channel bandwidth, or how fast the analog waveform varies in a sample period, and how fast the DAC changes it's output sample. This raises three questions in the design of the digital processing, and the DAC:

- 1) Over what frequency range does this error appear? 2) What is the optimal digital

processing to apply at the DAC input, as a function of the leakage channel and the DAC sample rate? 3) At what minimum sampling speed does the DAC needs to operate?

Calderin [52] describes this in further detail, describing that if the cancellation is restricted to a limited bandwidth, then the DAC can cancel the continuous time waveform to the quantization noise floor, despite this discretized zero-order held output. The choice of sampling speed question is also discussed there. To summarize the result, cancelling at the sample spaced instances enforces no guarantees on the error at non-sample spaced time intervals. If the residual error between sample spaced times is too large, though it is not sampled by the ADC, it may desensitize the receiver. This residual error is a function of the leakage network and the sampling rate - increased sampling speed reduces the time difference over which the error is controlled, limiting the maximum possible error. Accordingly, the sample rate must be chosen to keep the power during off-sample time instances sufficiently low.

Chapter 3

Fully Integrated FDD Transceiver Implementation

This chapter presents a circuit realization of the previously described architecture, and the resulting measurements.

3.1 Switched Capacitor Power Amplifier

3.1.1 Topology Motivation

The switched capacitor power amplifier was first proposed in [58] in 2011. The topology is a digital PA (DPA) architecture, in which a number of unit element PAs are tiled, as controlled by a digital code word, to perform a direct digital to RF conversion[59].

This topology was motivated as a result of two characteristics of prior digital PAs. Firstly, due to the parallel combination of RF-DAC unit cells, most such DPAs rely on current summing of high impedance unit elements. These elements are often cascoded to boost DAC accuracy, which can increase required voltage headroom and decrease the core PA efficiency. Additionally, most other DPAs have a saturated output characteristic at high power due to the nonlinear summing of unit elements, which requires extensive digital predistortion and additional DAC resolution to linearize [59][60].

The motivation for using the switched-capacitor topology in this work is the nonlinearity of the output impedance characteristic of prior such DPAs. In contrast, as described here, the switched-capacitor topology achieves a low impedance which is, to first order, amplitude-codeword independent. This impedance characteristic enables the series TX/RX combination network with low RX insertion loss.

3.1.2 SC PA Operation Principle

The SCPA topology is a voltage mode class D PA, consisting of a switched inverter driving a band-pass matching network. While standard class D PAs employ pulse density modulation

to control output amplitude, in the SCPA, the driving inverter and series capacitor in the band-pass network are segmented into unit cells. The number of capacitors switched versus connected to ground set a capacitive divide ratio, which controls the amplitude. In the original implementation, the input LO waveforms are modulated with a phase interpolator to control the output phase.

Like many switching power amplifiers, the ideal power-added efficiency is 100% at the highest code. An analysis for the power added efficiency vs output is provided in [58]. A simple intuition for the ideal 100% peak efficiency is that in the absence of drain parasitic cap from the inverter, no capacitance to ground is switched, so any charge drawn from the supply is delivered to the load.

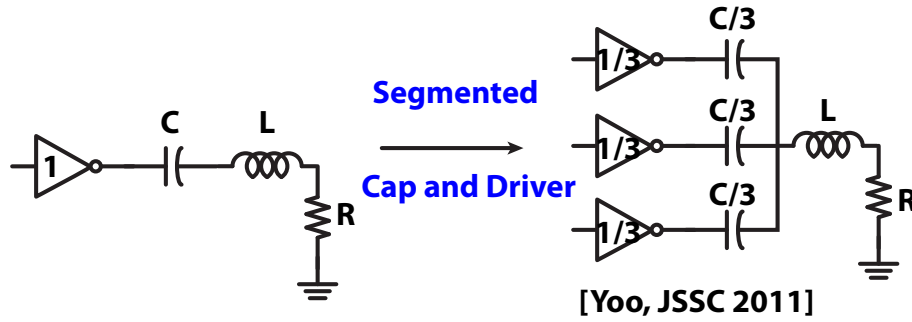


Figure 3.1: SC PA as a class D PA.

The output impedance of the transmitter looking back from the antenna side is roughly input amplitude codeword independent. Shown in Fig. 3.2, when a unit cell is disabled, it's input is statically pulled to the supply. As all capacitor bottom plates are connected to an AC ground independent of the amplitude code-word, the full PA array capacitance always appears in shunt with the transformer inductance. Around the center frequency, the LC resonance thus provides a low impedance across all amplitude code-words. The series switch resistance can be made code-independent by sizing the PMOS and NMOS transistors to have similar on-resistance. The residual code dependence comes from the mismatch of PMOS and NMOS resistances in the inverter, making the switch resistance to AC ground slightly codeword dependent, as well as from the finite edge rate of the switching inverters, which can briefly place the inverters in a high impedance state during the transition time. However, because these devices are sized for low on-resistance to achieve high efficiency, any residual mismatch and code-dependence are small when compared to the resistance due to the Q of the transformer.

The constant output impedance contributes to the linearity of this transmitter topology. The low amplitude independent impedance of this voltage-mode PA also enables several other interesting PA techniques. An impedance-inverter-less Doherty PA, which exhibits Doherty efficiency over a wide LO center frequency range, can be implemented due to the voltage mode nature of the switched cap PA [61]. Additionally, the low PA series impedance enables

direct series transformer combination for efficiency enhancement without the shorting switch traditionally used as in [62]. As the large shorting switch is slow to transition, it limits the TX modulation bandwidth over which efficiency enhancement can be achieved. Lastly, due to the linear summing of unit elements, mixed signal DAC filtering and harmonic rejection techniques, as in [63], are heavily improved.

Finally, it is worth noting that the PA efficiency and linearity are largely set by the precision of capacitor ratios and by switch parasitics, both of which track well to scaled technology nodes.

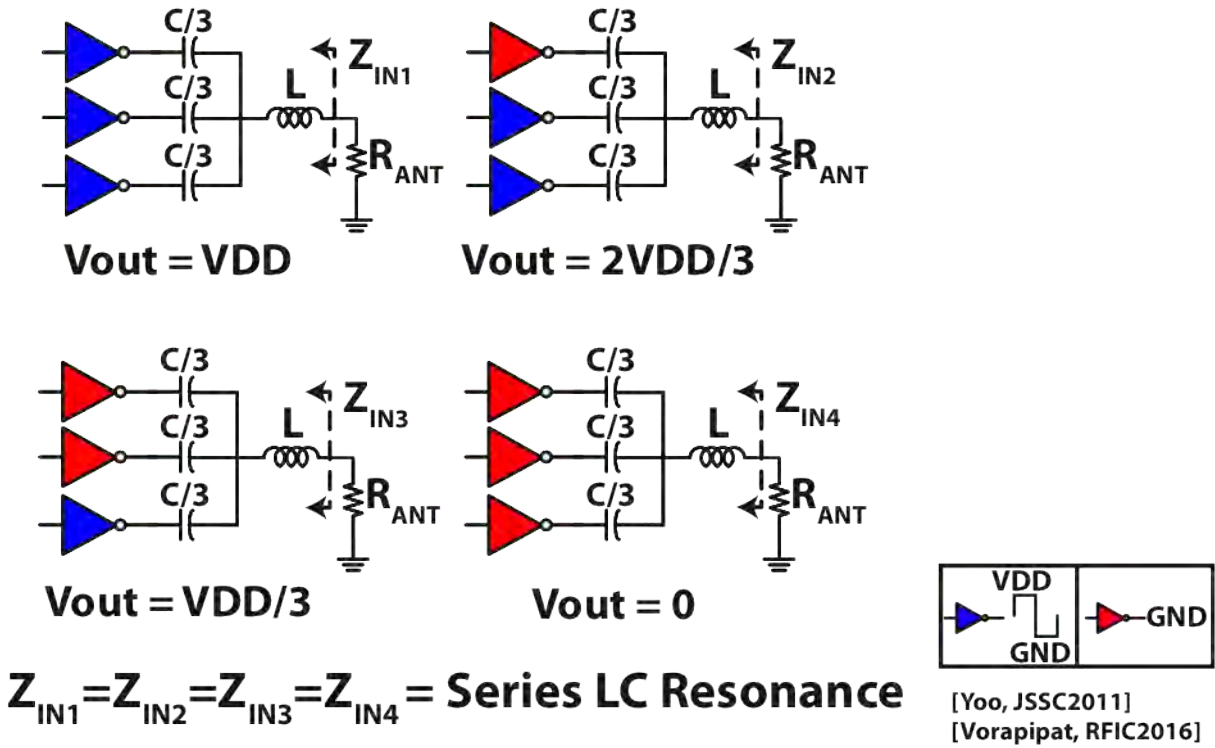


Figure 3.2: SC PA operating principle.

3.1.3 Cartesian Implementation

A Cartesian PA topology as shown in Fig. 3.3 is desired to maximize the amount of available LO sharing between replica cancellation DAC and transmit chains. A polar implementation of the TX and cancellation DAC would require independent phase interpolators for the two paths, in order to adjust the DAC output to handle the frequency selective TX-RX network. These phase interpolators can add a significant uncorrelated phase noise component, adding to the RX noise floor.

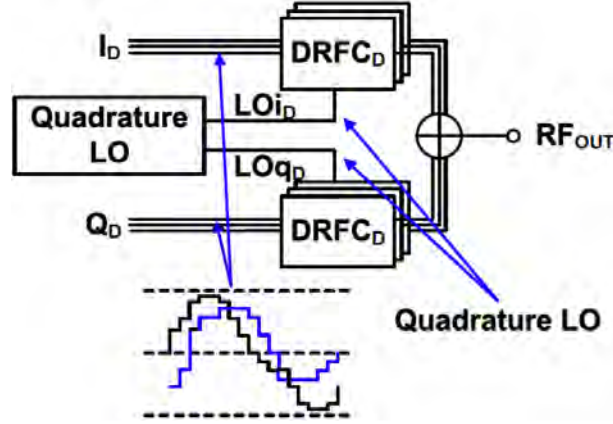


Figure 3.3: Cartesian PA model.

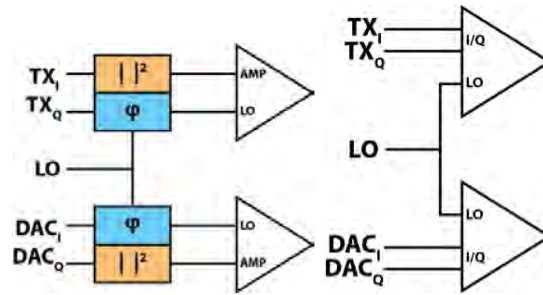


Figure 3.4: Polar (right) vs Cartesian (left) PA.

The phase interpolator operates by slowing down the input clock edge rate in order to phase combine quadrature LO's into a relatively continuous sinusoid. Accordingly, as the phase noise power is essentially the thermal noise power at the edge crossing normalized by the rate of voltage change over time, the slow edge translates thermal noise around the threshold crossing to a large time uncertainty. More precisely, recall that the jitter variance of a digital gate increases linearly with the clock edge rate. The phase noise of a representative phase interpolator consuming a large 20mA is simulated as shown in Fig. 3.5. Note that for a 0dBm TX output, this uncorrelated phase noise of -164dBm/Hz immediately translates to a 10dB noise figure in the RX band at 40MHz offset.

Additionally, the Cartesian topology alleviates some complications in the polar implementation such as bandwidth expansion on the phase path, which requires a high bandwidth phase modulator, and delay mismatch on the amplitude and phase paths which result in out-of-band distortion.

If the two RF-DAC segments in Fig. 3.3 each draw half the DC power as an equivalent polar transmitter, the cartesian PA efficiency is significantly degraded, as the amplitude cells

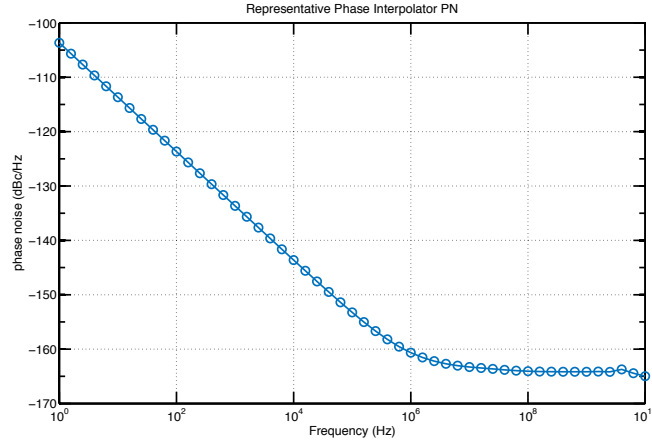


Figure 3.5: Simulated phase interpolator phase noise.

are summed out of phase. As the combined power is $\sqrt{2}$ less than the summed powers of I and Q , the topology would have a $\sqrt{2}$ efficiency hit as compared with a polar architecture consuming the same power.

In order to overcome this efficiency penalty, use of a 25% duty cycle LO is proposed here, as similarly demonstrated in [64][65]. In this technique, the I and Q phases are time multiplexed on the same physical PA unit cell, by use of 25% waveforms. Note that the effective 3 level LO's from [64] can be implemented directly in a differential PA structure using 2 level 25% duty cycle LOs, which removes the need for their proposed 2 level implementation with the disable opposite cell enhancement technique. This is accomplished by switching each unit cell in either the 25% I phase, 25% Q phase, or 45° phase, depending on the decoded codeword. The decoding is simply implemented by noting that if the 25% duty cycle I and Q waveforms are provided as input, the unit cell should be switched if $DATA_I$ AND LO_I are active, OR $DATA_Q$ AND LO_Q are active. This maps cleanly to the equivalent combinational logic of 3 NANDs, as shown in Fig. 3.6.

As the I phase is sent via a 25% pulse, while 45° by a combination of I and Q , resulting in a 50% pulse, the maximum total power in the I or Q phase is less than at 45° phase angle, resulting in a square shaped achievable constellation. This is in contrast to a polar topology, which results in a circular constellation, corresponding to constant maximum amplitude output with tunable phase.

In the peak power case, when I and Q are both enabled for transmission at 45° phase angle, all the PA unit cells drive all cap cells with a 50% duty cycle waveform. This is same configuration as a polar topology at max code, and the peak efficiency is the same as in the polar case. When driven in just the I or Q phase, the ideal peak efficiency is still 100%, as it would be in the polar case, once again given the intuition that no capacitance to ground is switched. Any charge drawn from the supply is thus delivered to the load. Note that in

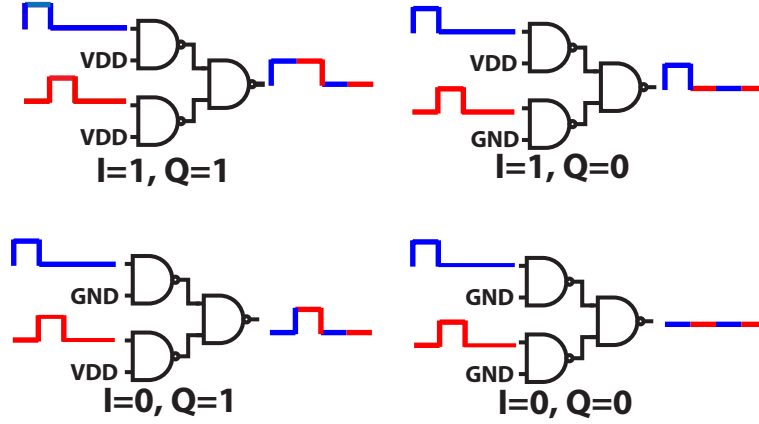


Figure 3.6: 25% Cartesian LO combination.

practice, the efficiency at the 0° and 90° phase angles will be lower than the efficiency at the 45° phase angle. The same amount of parasitic drain capacitance at the inverter output is switched in both cases, but is normalized against a higher output power in the 45° phase.

At other phase angles and power levels, the ideal efficiency of the combination can be derived as follows.

The bandpass matching network transforms the 50 ohm load resistor to some real value R_{load} , while passing only the first harmonic of the square wave. Accordingly, the output power can be written as:

$$P_{out} = \frac{1}{2} \frac{V_{amp}^2}{R_{load}} \quad (3.1)$$

$$V_{amp} = \sqrt{V_{amp,i}^2 + V_{amp,q}^2} \quad (3.2)$$

$V_{amp,i}$ and $V_{amp,q}$ are set by the capacitive divide ratio in terms of the number of units on, n , and the total number of units N , as

$$V_{amp,i} = \frac{4}{\sqrt{2\pi}} \frac{n_i}{N} V_{dd}^2 \quad (3.3)$$

$$V_{amp,q} = \frac{4}{\sqrt{2\pi}} \frac{n_q}{N} V_{dd}^2 \quad (3.4)$$

$$(3.5)$$

where the $\frac{4}{\sqrt{2\pi}}$ is the fundamental Fourier series coefficient of a differential 25% LO. Accordingly,

$$P_{out} = \frac{\frac{4}{\pi^2} \frac{V_{dd}^2}{N^2} (n_i^2 + n_q^2)}{R_{load}} \quad (3.6)$$

During the transition times, the series inductance of bandpass matching network presents a high impedance, and accordingly, the PA simply drives the captive divider. Because the I and Q phases are time interleaved onto the same capacitor bank, and I phase is switched first, followed by the Q phase, the energy consumed can be computed as follows.

The I phase first draws a certain amount of charge from the supply set by the capacitive divide ratio it sees, as

$$C_i = \frac{n_i(N - n_i)}{N^2} C_{total} \quad (3.7)$$

$$(3.8)$$

where

$$C_{total} = \frac{1}{(2\pi f R_{load} Q_{loaded})}. \quad (3.9)$$

The power required to switch this capacitance is

$$P_{sc,i} = C_i \times V_{dd}^2 f. \quad (3.10)$$

The Q phase then draws a charge from the supply, related to the initial and final states of the capacitors. To express this generally, note that in transitioning from I to Q , a capacitor can experience one of four states, shown in Fig. 3.7. The capacitance can be held at V_{dd} (as C_3), held at ground (as C_4), turned on with a transition from ground to V_{dd} (as C_1), or turned off with a transition from V_{dd} to ground (as C_2).

Only the charge drawn from the supply is relevant, as charge returned to ground does not consume energy. The additional charge drawn from the supply during the Q phase can be computed as the the charge on the top plates of C_3 and the bottom plate of C_1 after the Q phase, minus the charge already there from the I phase. Calling the middle node V_{out} , these charges can be computed as follows:

$$V_{out,initial} = \frac{C_3 + C_2}{C_{total}} V_{dd} \quad (3.11)$$

$$V_{out,final} = \frac{C_1 + C_3}{C_{total}} V_{dd} \quad (3.12)$$

$$Q_{C3,final} = C_3 \times (V_{out,final} - V_{dd}) \quad (3.13)$$

$$Q_{C3,initial} = C_3 \times (V_{out,initial} - V_{dd}) \quad (3.14)$$

$$Q_{C1,final} = C_1 \times (V_{out,final} - V_{dd}) \quad (3.15)$$

$$Q_{C1,initial} = C_1 \times V_{out,initial} \quad (3.16)$$

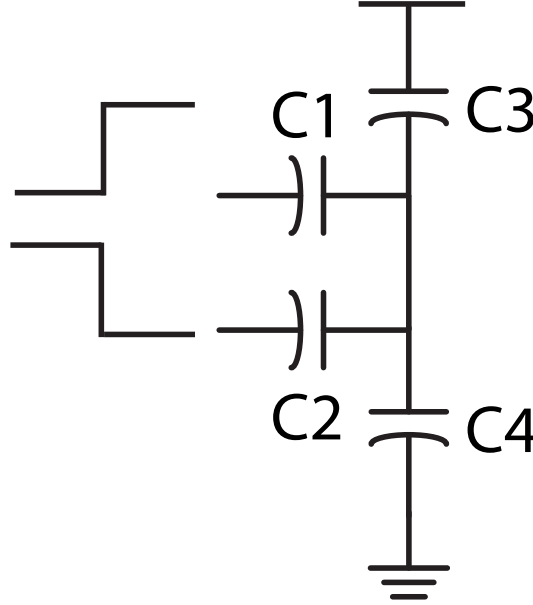


Figure 3.7: Charging model for the Q phase.

and accordingly, after simplification

$$\Delta Q = Q_{C3,final} + Q_{C1,final} - Q_{C3,initial} - Q_{C1,initial} \quad (3.17)$$

$$= \frac{-2C_1C_2 - C_2C_3 - C_1C_4}{C_{total}} V_{dd} \quad (3.18)$$

and the total power switching is

$$P_{sc} = V_{dd}^2 \times f \times (C_i + \frac{2C_1C_2 + C_2C_3 + C_1C_4}{C_{total}}) \quad (3.19)$$

Accordingly, a fixed power penalty is used to switch the I capacitors. The power penalty in the Q phase is proportional to the number of capacitors switched in transitioning from the I to Q phase. This formulation reveals that the contour efficiency is a function of the encoding scheme of the capacitor bank. In particular, a thermometer encoding is more efficient than binary, as it minimizes the amount of capacitors switched for a given code transition; for a thermometer encoding, either C_1 or C_2 is always zero.

To provide some intuition, note that in the polar case, only the first term in P_{sc} exists, as there is simply one switching transition per cycle. In the $Q=0$ phase, $C_1 = C_3 = 0$, and no additional power is consumed relative to the polar case. Similarly, in the $I=0$ phase, $C_2 = C_3 = 0$, $C_i = 0$, and the total capacitance seen is simply the series combination of C_1 and C_4 , identical to the polar case. Finally, at the 45 degree phase angle, $C_1 = C_2 = 0$, and no

additional power is consumed relative the polar case, consistent with the previous intuition. At all constellation points in between these phase angles, the second term represents an efficiency penalty.

The ideal efficiency is then simply calculated as:

$$\eta_{scpa} = \frac{P_{out}}{P_{out} + P_{sc}}. \quad (3.20)$$

The efficiency contours for a 4 bit binary/4 bit thermometer 25% duty cycle Cartesian, and a polar PA normalized to achieve the same peak output power, are shown in Fig. 3.8. Note that, as expected, the two contours match at 45% phase angle, and the Cartesian peak is 100% at (I,Q) values of (max, max), (0, max), and (max,0). Interestingly, the Cartesian implementation performs better at high power, while the polar PA performs better at back-off. To map this directly to overall efficiency, the data statistics of the desired modulation scheme could be passed through the contour map.

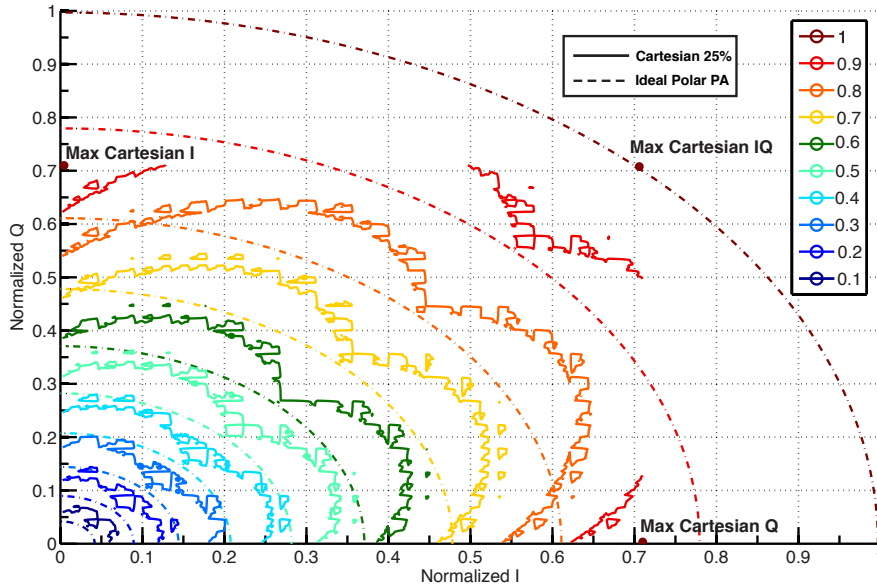


Figure 3.8: Cartesian(4 bits binary, 4 bits thermo) vs. Polar efficiency contours. The polar is the constant efficiency at constant amplitude circles. The black lines correspond to achievable powers for cartesian.

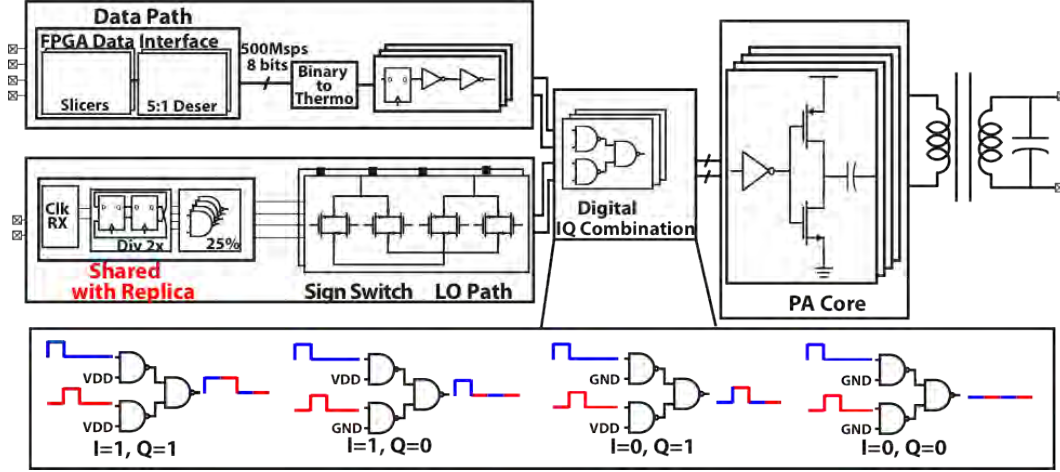


Figure 3.9: Top level TX schematic.

3.1.4 Sizing

In this implementation, single stacked output devices are used, as the target output power of 20dBm can be achieved with a differential implementation, a 1:2 transformer turns ratio, and a 1.2V supply. This aids the system design, as stacked cascode devices contribute to AM/PM and AM/AM distortion through the inclusion of additional nonlinear parasitic components in the stack, contribute higher on resistance leading to lower efficiency, and slow down the edge rates leading to longer time in the high impedance driver state un-accounted for by the analysis.

The main loss mechanisms in the PA are the power to switch the PA parasitic capacitances, resistive loss in the series PA switches, and the loss in the transformer. Given the process technology, the transformer network is first optimized to maximize its G_p , as derived in [66]. This transformer topology and parameters are shown below

Applying the heuristic presented in [66], a shunt capacitor is placed on the antenna side to resonate the transformer secondary inductance. Then, the series PA capacitance is chosen to resonate the remaining imaginary part on the PA side. Given this passive network design, the total PA NMOS switch width W_{tot} can be sized to maximize the PA efficiency, trading the series resistive loss against the capacitive switching power, as shown below.

Defining ρ_p as the PMOS relative size as compared with the NMOS (in this case 2), C_d as the drain capacitance per unit NMOS width, C_g as the gate capacitance per unit NMOS width, V_{sw} as the voltage swing across these capacitors (in this case V_{dd}), and V_{dd} as the power rail, the effective capacitance being switched is

$$C_{eff} = C_d \times (1 + \rho_p) \times \frac{V_{sw}}{V_{dd}} + C_g \times (1 + \rho_p) \times \frac{V_{sw}}{V_{dd}} \quad (3.21)$$

where the first term represents the drain parasitic capacitance, and the second term the

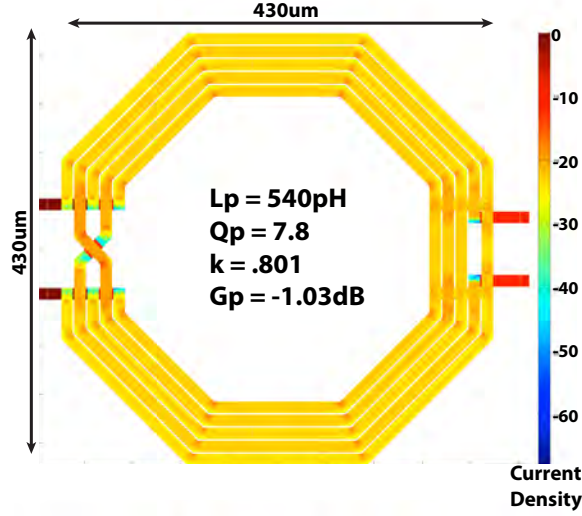


Figure 3.10: TX transformer.

input gate capacitance that must be driven. Define the frequency as f_{sw} , the number of unit elements n_{elems} , the fixed routing cap to ground at the transistor output as C_{shunt} , and routing parasitic cap to ground between the series cap and transformer network C_{par} . The capacitive switching power is then

$$P_{loss,C} = (W_{tot}C_{eff} + n_{elems}C_{shunt}) \times V_{dd}^2 \times f_{sw} \quad (3.22)$$

Further defining $R_{pa,u}$ as the resistance of the switch times a unit width, the resistive loss can be found by first solving for the output voltage at the $n:1$ transformer output, and consequently, the current.

$$R_{pa} = \frac{R_{pa,u}}{W_{tot}} \quad (3.23)$$

$$V_{out} = \frac{4}{\pi} \times n \times \frac{R_L}{(R_L + 2n \times R_{pa})} \times \frac{C_{ser}}{C_{ser} + C_{par}} V_{dd} \quad (3.24)$$

$$I_{load} = \frac{V_{out}}{R_L} \quad (3.25)$$

$$P_{loss,R} = \frac{1}{2} \times (I_{load} \times n)^2 \times R_{pa} \quad (3.26)$$

Accordingly, as $P_{loss,R}$ decreases with increasing W_{tot} , and $P_{loss,C}$ increases, an optimal value of W_{tot} for minimizing loss can be solved, substituting n with 2 as

$$W_{tot,min} = \frac{4}{R_L} \times \sqrt{\frac{R_{pa,u}}{C_{eff}f_{sw}} \times \frac{(\frac{4}{\pi})^2 + 8C_{eff}f_{sw}R_{pa,u}}{2}}. \quad (3.27)$$

The output power and drain efficiency can then be computed as below, where A_{xfrm} represents the loss in the transformer

$$P_{out} = I_{load}^2 \times \frac{R_L}{2} \quad (3.28)$$

$$P_{load} = P_{out} \times A_{xfrm} \quad (3.29)$$

$$\eta_{drain} = \frac{P_{load}}{P_{out} + P_{loss,R} + P_{loss,C}} \quad (3.30)$$

Using this methodology, the transmitter is sized with a total of 1.6 mm NMOS switches, 3.2mm PMOS switches, and 20pF series (40pF on each differential side) capacitance.

One last effect to note in the SCPA is a drain-node kickback that can result on disabled unit-cell inverter outputs. The worst case example of this for an 8 bit PA is shown in Fig. 3.11.

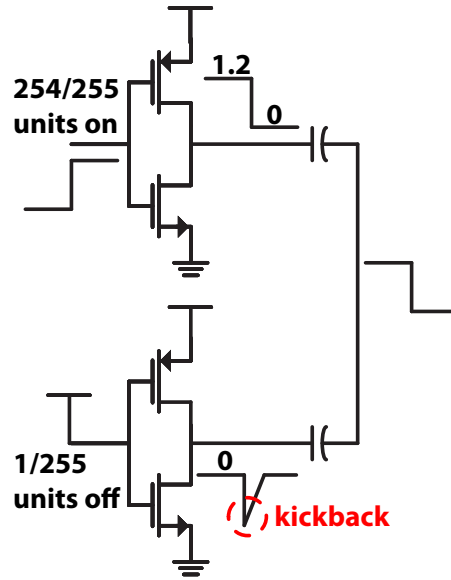


Figure 3.11: SC PA drain kickback.

Note that for a 8 bit PA, if code 254 is sent, the LSB unit cell is disabled, while the rest of the cells are enabled. The series capacitance for each cell is connected to the shared output node. At the end of the LO cycle, the on cells transition from V_{dd} to 0. The series capacitor from the off cell can not change rapidly change its voltage, and accordingly, there is a sharp downward spike in the off cell's drain voltage. As this voltage can swing below ground, it must be ensured to fall within a safe operating range. The kickback can be mitigated at an efficiency penalty by increasing the driver size. This serves to lower the switch resistance and increase the parasitic drain capacitance to ground, which both attenuate the kickback. For

the sizing used in this work, the worst case kickback, shown in Fig. 3.12, is seen to produce a voltage spike of -380mV, deemed safe to avoid a diode forward bias condition.

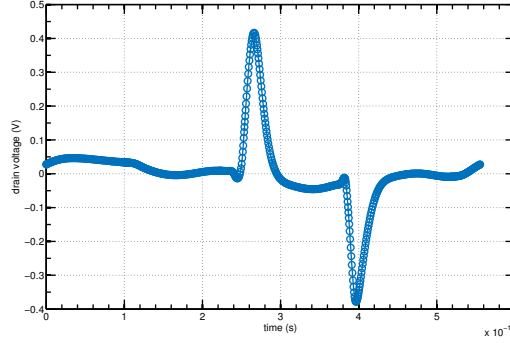


Figure 3.12: Worst case drain voltage vs. time.

3.2 RX Design

In order to measure the efficacy of the proposed cancellation scheme in a realistic environment, a receiver is implemented on the same die. A top level schematic of the receive chain is pictured in Fig. 3.13. It consists of a complementary common gate/common source low noise transconductance amplifier (LNTA), followed by current-mode passive mixers, driving first order shunt feedback transimpedance amplifiers (TIAs).

The receiver for this chip prioritized linearity and measurement flexibility to test a variety of cancellation scenarios.

The receiver gain is targeted for 20dB, to knock down the -160dBm/Hz noise floor of the spectrum analyzer used for measurement to below the thermal noise floor.

The out-of-band linearity is specified in order to handle the largest residual TX signal, under the assumption of a 20dBm TX fundamental, and a pessimistic 30dB of cancellation, resulting in -10dBm at the RX input at a 40MHz offset spacing from the RX center frequency.

The option to disable the mixers and directly observe the TX residual at RF is desired, for testing. Accordingly, to drive this RF signal directly off chip to a spectrum analyzer, 50 Ω on-chip pad drivers are included on-chip. The full receiver schematic is shown in Fig. 3.13.

The LNTA consists of a complementary common gate/common source amplifier, shown in Fig. 3.14. The common source portion of this structure decouples the amplifier transconductance from the input matching condition. This allows a higher transconductance to be designed, reducing the input referred noise of the subsequent stages in the chain. If the gate of the common gate amplifier pair is also driven with the cross-coupled input, the LNTA noise figure can be dropped below 3dB. However, this results in a 6dB drop in the linearity

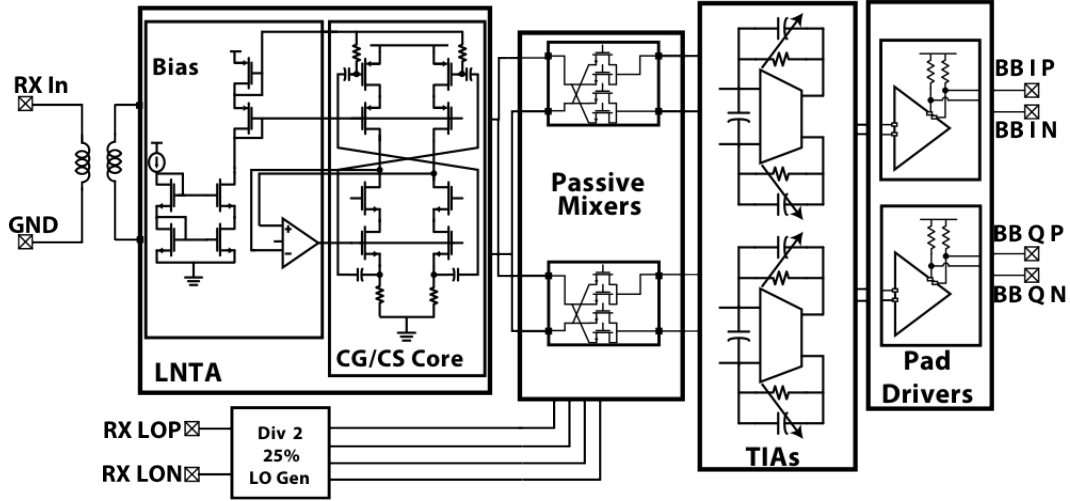


Figure 3.13: Top level RX schematic.

by doubling the voltage swing across the common gate stage, and accordingly is not suitable for this work. The LNTA is biased with a replica bias network, and a common mode feedback loop with a differential-difference CMFB amp, shown in Fig. 3.15, regulating the cascode gate voltage.

It may be intuitively expected that in the matched condition, the common gate portion would contribute a 3dB noise figure, and the common source would add additional noise, raising the overall noise figure to above 3dB. One interesting artifact of this complementary amplifier is that the common source branch provides partial cancellation of the common gate device's noise, due to the inverted gain polarity of the common gate noise relative to the signal path. A full noise analysis shows that in the matched condition, this cancellation results in a total LNTA noise figure due to the common gate and common source devices of 3dB, independent of the common source transconductance. This allows the LNTA to provide a large transconductance gain to reduce subsequent noise in the chain, independent of the matching condition.

The TIAs consist of a simple cascoded amplifier, shown in Fig. 3.16. A similar differential difference amplifier as in the LNTA regulates the TIA common mode. Two bit feedback resistor and capacitive DACs in the shunt feedback path set the closed loop bandwidth, while a resistive DAC to ground is used to set the input common mode.

The RX LO path, is shown in Fig. 3.17. The LO is brought in at 2x the fundamental frequency, and divided by two to get quadrature phases using a current mode logic (CML) clock divider.

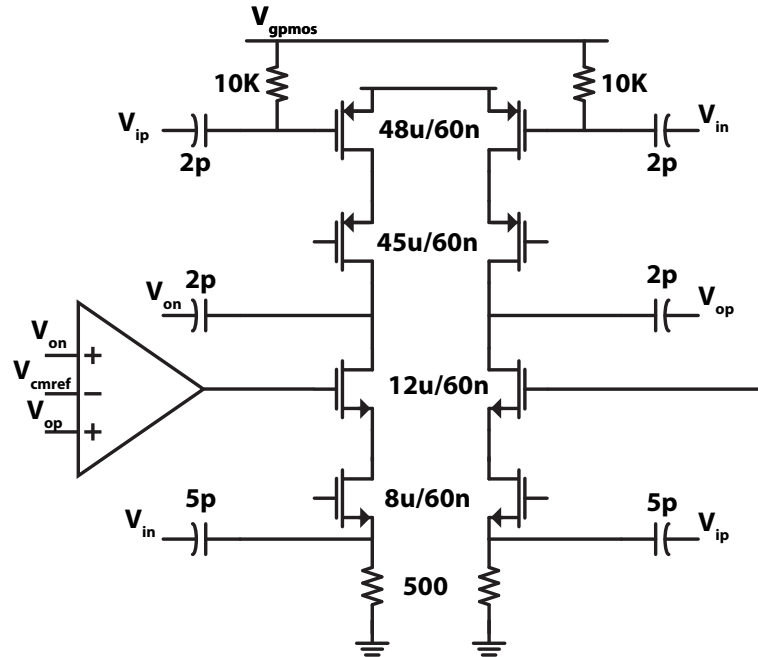


Figure 3.14: LNTA schematic.

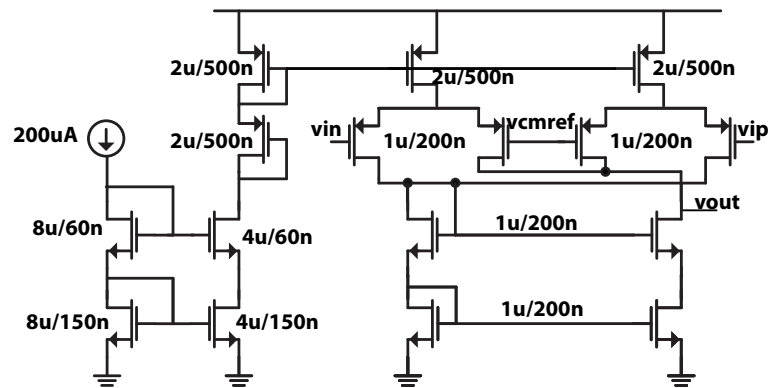
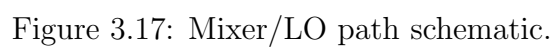


Figure 3.15: LNTA CMFB schematic.



3.3 Cancellation DAC

The replica cancellation DAC is implemented as a 10-bit (5 binary, 5 thermometer) class-A RF current steering DAC, operated up to 500MSPS. Inactive currents are shunted through a switch to the center tap, to maintain class A operation for dynamic linearity. As in the transmitter, the 25% duty cycle LO scheme with IQ cell sharing is once again used, allowing re-use of the TX LO path for the cancellation replica. As mentioned previously, LO path re-use enabled maximum correlation of TX/DAC phase noise, enabling feedforward phase noise cancellation. The DAC full scale current is sized at 60mA, to handle a TX full scale of +20dBm. The full DAC schematic is shown in Fig. 3.18 - more implementation details can be found in [52]

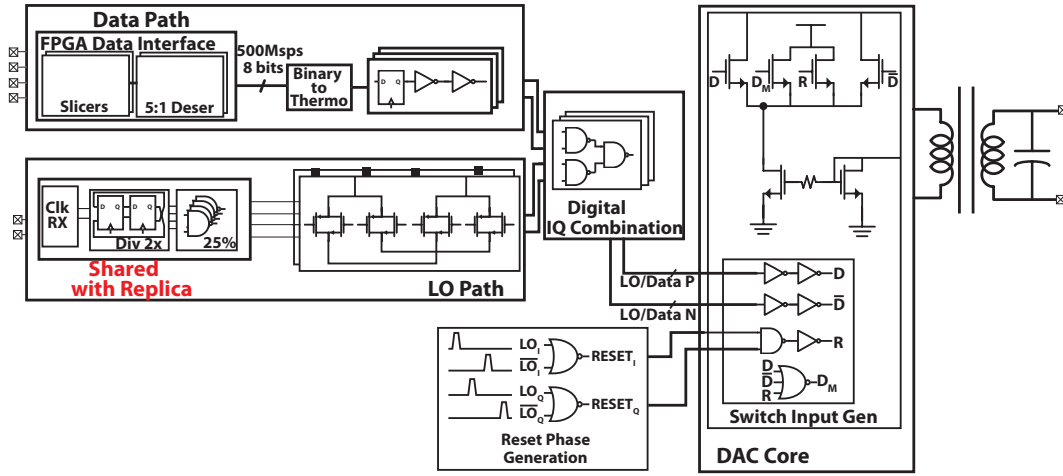


Figure 3.18: Top level DAC schematic.

The TX and DAC data are clocked into parallel length five shift registers on both positive and negative edges of an external LO, and clocked out in parallel by a 5x slower clock to achieve a 10:1 deserialization.

The full chip schematic, shown in Fig. 3.19 was taped out in TSMC65 GP process. The 2.5mm x 2.5mm die is shown in Fig. 3.20 .

3.4 Test Setup

Measurement of such a system is difficult for two reasons. Firstly, test setup must support characterization of the system blocks in-situ. Secondly, separation of measurement equipment non-idealities (noise, distortion) from device-under-test non-idealities is difficult in such a system, due the large dynamic range of the measurement (RX noise floor to TX power). The test setup is discussed in the context of these two considerations.

3.4.1 PCB Design

The die was bonded to a 6-layer PCB using flip-chip assembly. The PCB contains LDO's to generate five chip supplies from a common voltage, current sense amplifiers to enable power measurements on each supply, a tunable reference current for chip biasing, and an FMC connection to digital data links on a Xilinx VC707 board for both 5Gbps and low speed scan chain interfaces. Board level component reconfiguration is used to provide various testing modes, described next.

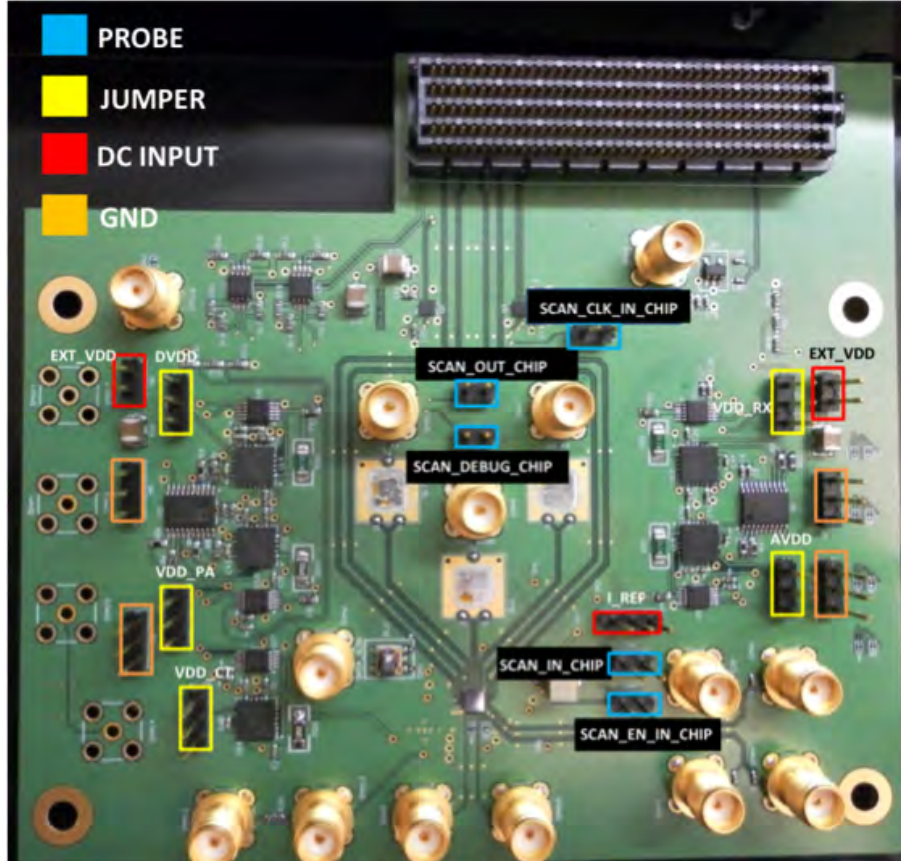


Figure 3.21: Measurement PCB.

To enable isolated testing of the TX and RX, a shunt zero-ohm resistor to ground was optionally included on the board at the chip's output center pin, with positive and negative chip outputs connected to SMA's via 50ohm single ended transmission lines on the board. To test the full system, the shunt zero-ohm resistor was de-soldered, and a zero-ohm resistor to ground was inserted on the chip minus output, as in Fig. 3.22. To minimize parasitics, 0201 surface mount resistors were used. This scheme was chosen to provide wideband isolation. In practice, however the isolation is limited by the series inductance of the 0 ohm resistor.

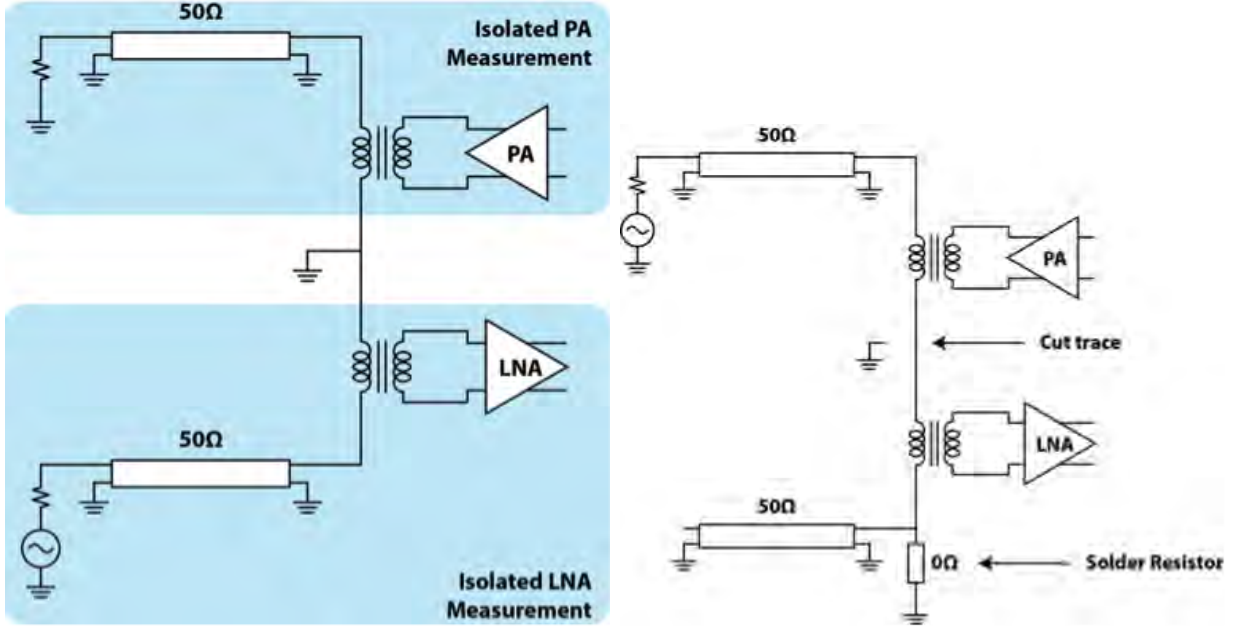


Figure 3.22: PCB configuration for isolated (left) and system testing (right).

The thermal noise filter on DAC supply on the RX center tap was also designed for board configurability via component replacement. SMAs with a series resistor on the AC side of the bias tee provide the ability to inject signals into the center tap, as well as measuring the DAC noise current. The board is nominally configured with a zero-ohm resistor soldered in shunt with the tank filter, and decoupling capacitance included on the center tap. To enable the noise cancellation, the zero-ohm resistor and decap are removed, and the SMA AC connections are grounded.

A few EM simulations relevant to the board chip interface are shown below.

As the chip is flipped, the ground planes on the PCB under the transformer can affect the Q and SRF. The top 4 layers on the PCB were removed so as to maintain performance to within 10% below the SRF, shown in Fig. 3.23.

The DAC tail degeneration inductance for $2F_{LO}$ noise reduction is implemented as an on board PCB via inductance. The DAC source via to the bottom ground plane, the high frequency return path through the center tap through the decoupling capacitance, and the center tap inductor were simulated in ADS Momentum, and the closest standard board thickness was selected to provide the required inductance. The simulation impedance matches well to a simple lumped model of just the source and drain pad cap and via inductance, and when the on-chip cap is included, the network resonates at the desired 4GHz.

The RF transmission lines at the TX output and RX input are replicated on the board with SMA connections on each side for de-embedding. SMAs connected to load, short, and open configurations were also included. This allowed direct de-embedding of measurements

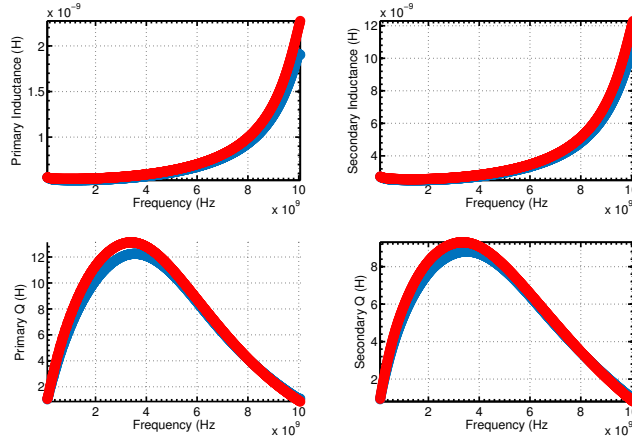


Figure 3.23: Inductor parameters with top 4 layers of PCB cut.

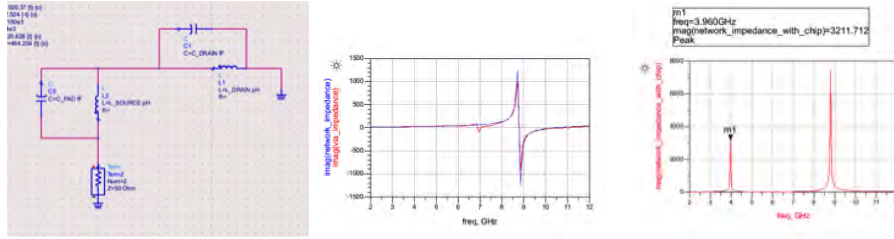


Figure 3.24: 2FLO Noise Reduction.

through the VNA fixture simulation/de-embedding options, described later.

3.4.2 Test Equipment

The Xilinx VC707 FPGA board with 10Gbps SERDES interfaces was used to generate the the digital TX/DAC data for the chip. An Agilent E5071C VNA was used for receiver chain measurements, and transmission line de-embedding. Rhode & Schwarz SME03 and HP 83711B signal generators were used as TX and RX LO's. HP 8563E spectrum analyzer was used to measure the TX signal at the antenna output. The Agilent N9030A spectrum analyzer was used for RX measurements, due to its ability to reach a -173dBm/Hz noise floor to enable noise measurements. This spectrum analyzer was read into a laptop via GPIB, and the data fed back into the FPGA to close an adaptive loop controlling the cancellation DAC data. A Keysight E4428C vector signal generator was driven into the antenna output to simulate a receive signal while the transmitter was operating. A Lorch tunable bandpass filter and Micronetics white noise generator were used for noise transfer function measurements.

The test setup is pictured in Fig. 3.25.

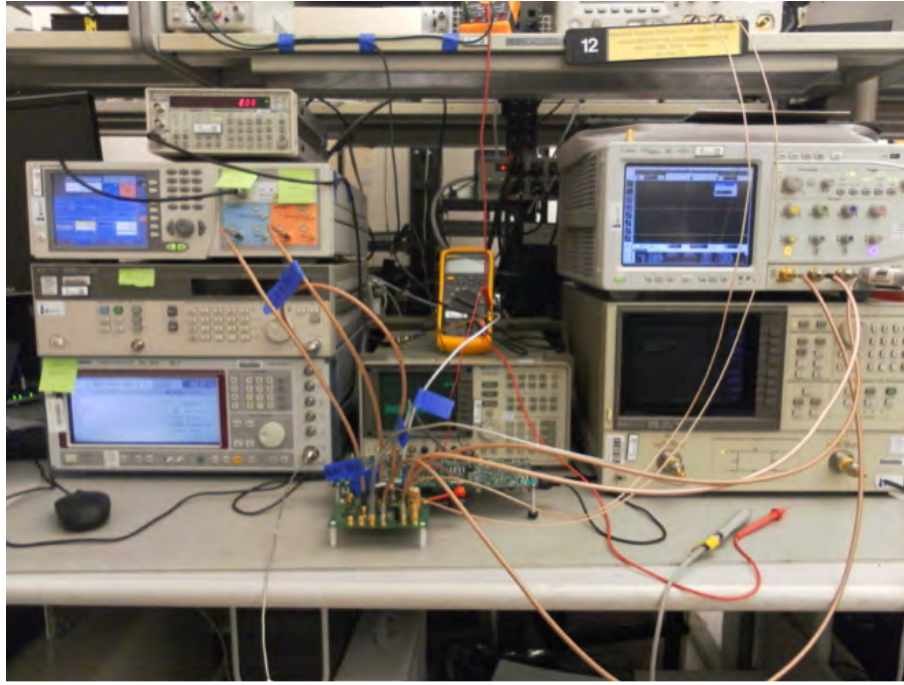


Figure 3.25: Test Setup.

3.5 Measurements

Details of the measurement setup and results for block level measurements are shown below. Block level measurements include de-embedding of the low bandwidth of the cascaded SMA and antenna transmission line, measured below in Fig. 3.26.

3.5.1 RX

Receiver measurements were performed after de-embedding of the input RF transmission line and SMA connection through the following structures on the board: 1) SMA connected to an open, 2) SMA connected to a short, 3) SMA connected to a load 4) SMA connected to replica RF transmission line connected to another SMA. To de-embed a single SMA's s-parameters were measured through the first three structures. Then, using the VNA's fixture removal option, the s-parameters of the SMA were de-embedded from structure 4, to obtain the s-parameters of a single SMA connected to the replica transmission line, shown below. Finally, again using the VNA's fixture removal option, the s-parameters of the SMA/replica

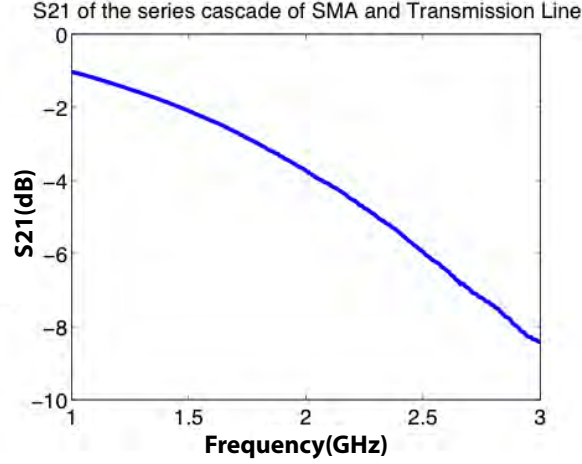


Figure 3.26: S21 of the SMA and output transmission line.

transmission line were removed from the receiver measurements. The final receiver measurements were obtained using the frequency translational s-parameter option of the VNA with the TX on, accurately representing the TX's series effects on the RX input matching.

The receiver S21 and input matching for highest gain setting are shown in Fig. 3.27. The input match falls below -10dB over a relatively narrow bandwidth. This is a design flaw, and the result is predicted in simulation. The DAC final capacitance was not correctly predicted until close to the tape-out deadline, and the RX matching network was designed for a lower assumed capacitance. The matching network was not redesigned, and accordingly performs quite poorly, providing nearly 3dB of loss. This is a simple and fixable design error that can be corrected in a redesign.

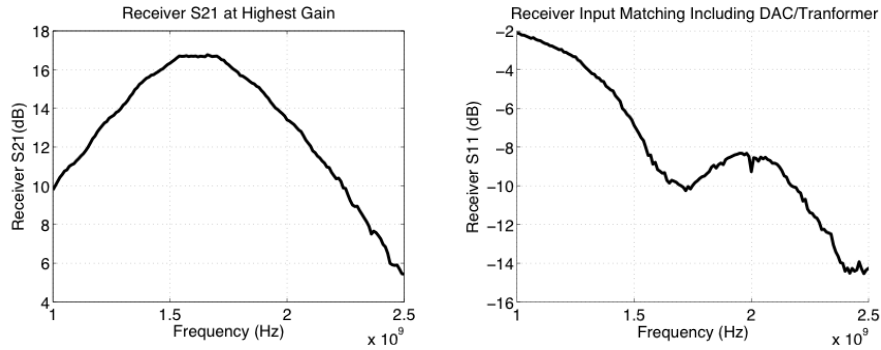


Figure 3.27: RX S21 (left) and S11(right).

The baseband gain and bandwidth settings of the TIA were measured, again through the frequency translational option on the VNA by fixing an RF input tone and sweeping the

down-converting mixer LO frequency to isolate this from an RF bandwidth measurement. The gain steps in 8 steps from 6dB to 18dB. At the peak gain setting, the bandwidth ranges from 15MHz to 140MHz - the lowest bandwidth setting matches very closely with simulation. The highest bandwidth case is 22% lower than expected, likely due to imperfect capacitance extraction at the output node. However, this does not limit system testing. Additionally, the highest bandwidth settings enable direct observation of the RF residual through bypassing the mixer.

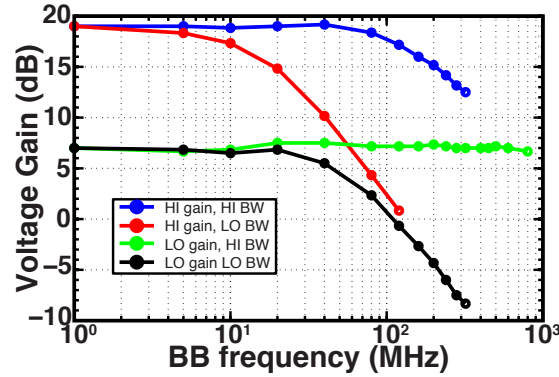


Figure 3.28: Various RX gain and bandwidth settings.

The RX noise figure is computed through averaging the output spot noise over 10MHz receive bandwidth, and input referring by the above S21. The NF of the receiver is 4.7dB, with a total NF of 7.6dB due to the matching network loss.

Receiver linearity was measured as well, in the form of two-tone in-band IIP3 of -7.5dBm, and in band P1dB of -18.7dBm, corresponding to 1.5V peak to peak swing at the baseband output, as shown in figure 3.29. The IIP3 is measured after the glitch, which corresponds to a calibration error in the attenuator step of the input source. The receiver draws 40mA from a 2.5V supply, including bias currents.

3.5.2 TX

The transmitter current and power vs. code characteristics are measured by sweeping TX code and measuring the output at the antenna port. A zero-ohm resistor is soldered at the series transformer's center tap to isolate the transmitter from the receiver's loading. The maximum measured TX output power is +19.5dBm at 1.2GHz, shown in Fig. 3.31. The DNL is shown to be <1LSB over codes in Fig. 3.32, with the downward slope indicative of a slightly compressive characteristic.

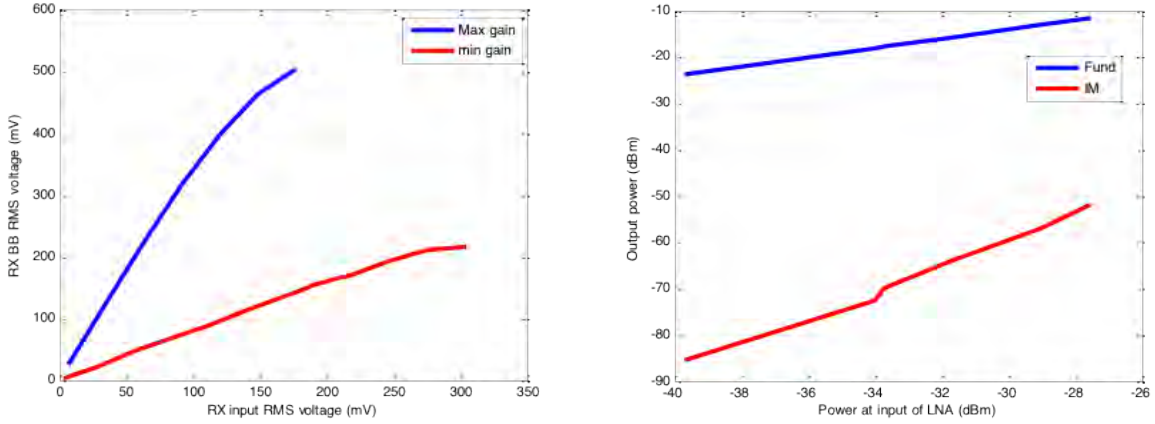


Figure 3.29: RX P1dB (left) and IIP3 (right).

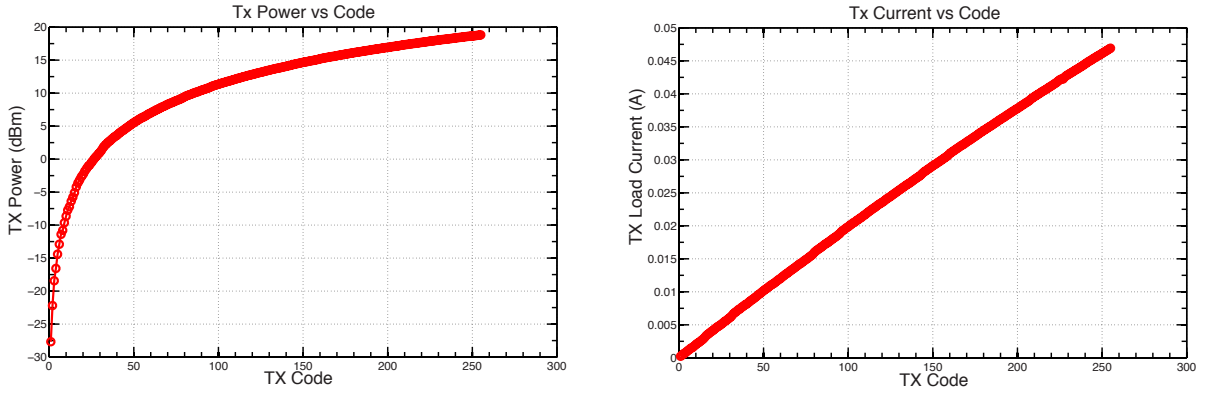


Figure 3.30: TX power (left) and current (right) vs. code.

3.5.3 System

The test setup and system level measurements for cancellation and noise are described below.

First, single tone cancellation is measured by sweeping the TX power in the 45° phase angle ($I = Q$), shown in Fig. 3.33. The power plotted on the x-axis is the power on the antenna port present after the optimal DAC cancellation code is found through the procedure described below. Measuring the antenna power with cancellation enabled is a valid reported operating condition, as the TX virtual ground across the receiver from cancellation is necessary to get an accurate TX power measurement without the RX loading effects.

To characterize the cancellation, an initial assumption is made on the leakage channel amplitude and phase, to form an initial guess for the cancellation DAC code. This is impor-

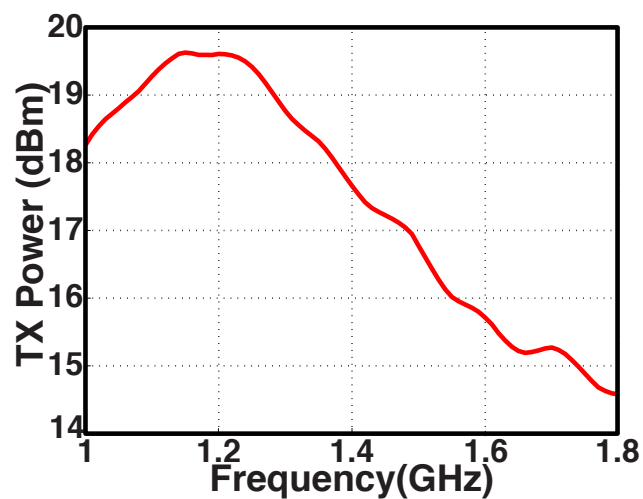


Figure 3.31: TX output vs frequency.

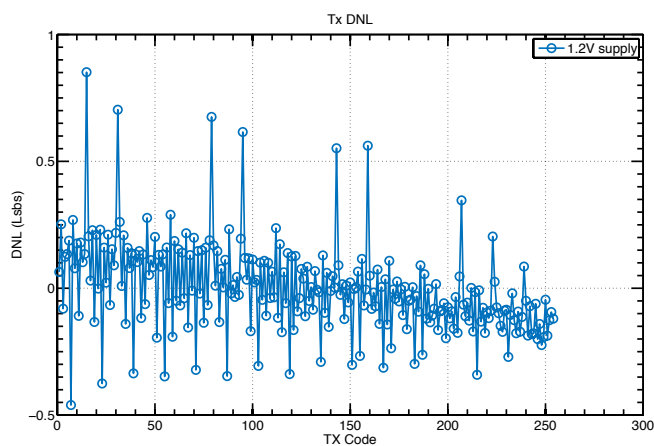


Figure 3.32: TX DNL.

tant in order to obtain some partial cancellation at the receiver input, to avoid damaging the RX input transistors. The TX output power is measured at RF through a spectrum analyzer on the antenna port, and the RX down-converted residual is simultaneously measured on a second spectrum analyzer. An RX band spur is coupled into the antenna port at an offset, and the gain is monitored in order to ensure that the RX is not compressing under the measured TX power/cancellation. The RX spectrum analyzer output is measured through GPIB to feedback to a computer running Matlab. A Matlab script performs a gradient-descent based feedback, modifying the DAC code to find the code maximizing cancellation. The residual with this final code is input referred by the receiver gain to produce the blue curve in Fig. 3.33.

The red curve in the plot shows the TX signal present at the input of the receiver without any cancellation. This is measured at lower codes which do not damage the receiver, and is then linearly extrapolated. This is 6dB lower than the TX power with cancellation (the x-axis), as without cancellation the PA is loaded by the RX matching network impedance, the voltage is divided between the antenna and the receiver, and there is some loss through the RX matching network.

The blue curve representing the residual at the input of the receiver is independent of TX output power. Intuitively, the residual is set simply by the LSB of the DAC, assuming the appropriate cancellation code has been found. At the highest measured single tone power of 12.6dBm, greater than 50dB of cancellation is observed.

Note that this procedure for finding the optimal DAC code is not real-time. Nonlinearity in the DAC can be handled while still maintaining cancellation, as long as the DAC's constellation space is sufficiently dense to cover all TX codes, as proven in [52]. However, DAC nonlinearity significantly complicates the predistortion/search to find the optimal DAC code. In practice, its is likely necessary to perform a linearity calibration on the DAC, and employ a look-up table based approach to finding a sufficiently good, though possibly not optimal, DAC code.

Cancellation with 20MHz modulated data is also measured. The FPGA code driving the 10 bit TX and DAC data is programmed to send periodic sequences of length 128 at a line rate of 2.5Gbps. Accordingly, the test setup limits the TX and DAC to produce sequences with Fourier coefficients at $2.5\text{Gbps}/10\text{lanes}/128 = 2\text{MHz}$ spaced tones. This is suitable for testing purposes, as it corresponds to 10 random power tones spaced within a 20MHz bandwidth. Additionally, the sequences were designed to produce a peak to average power ratio of around 6dB, to represent a relatively realistic complex modulated signal.

The current procedure to find the best code for cancellation is not real time. An initial guess of the optimal DAC sequence is made as the best DAC code for each TX code in the sequence, assuming a single tone is being sent. This initial guess neglects any memory affects from symbol to symbol, introduced by the frequency selective leakage channel. This effect is then included by iterating through each code in the length 128 code sequence and performing a gradient descent based search, minimizing the average energy in the TX band at the RX output measured through the spectrum analyzer. The DAC codes for a given TX sequence following the adaptation procedure is pictures in Fig. 3.35.

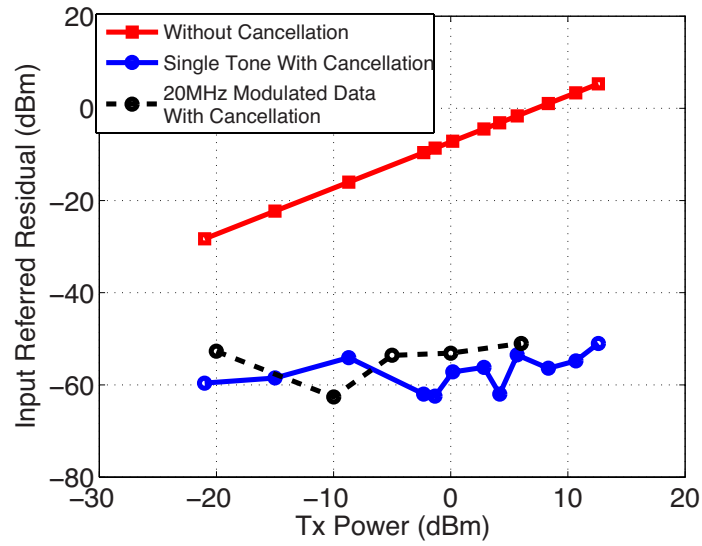


Figure 3.33: TX residual referred at RX input vs TX output power.

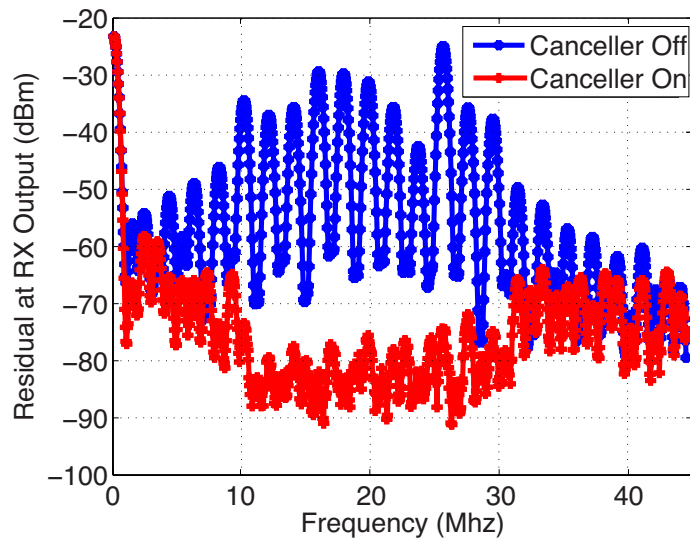


Figure 3.34: RX band spectrum before and after cancellation.

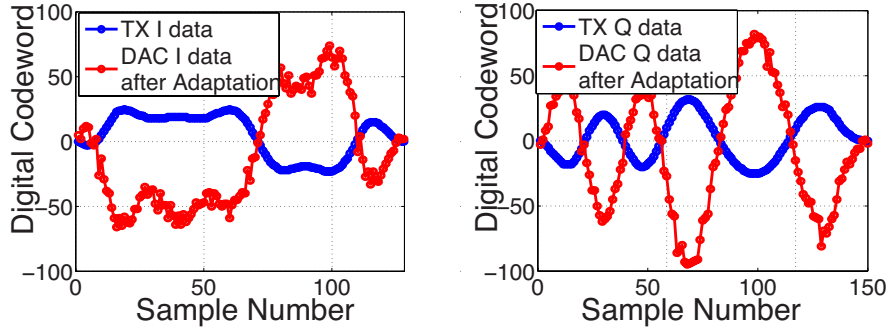


Figure 3.35: TX/DAC codes after adaptation.

Once again, in practice, this approach is offline, and the cancellation sequence should ideally be found through an adaptive filter and a lookup table based model of the TX/DAC static nonlinearity. However, this approach proves the existence of optimal DAC codes, and shows the resulting cancellation. Real-time modeling and adaptation of similar systems has been pursued in other works, such as [67].

From the original figure, we can see that for a +6dBm average signal with 12dBm peak, the cancellation is still limited by the DAC LSB, and >50dB cancellation is still achieved in the modulated case.

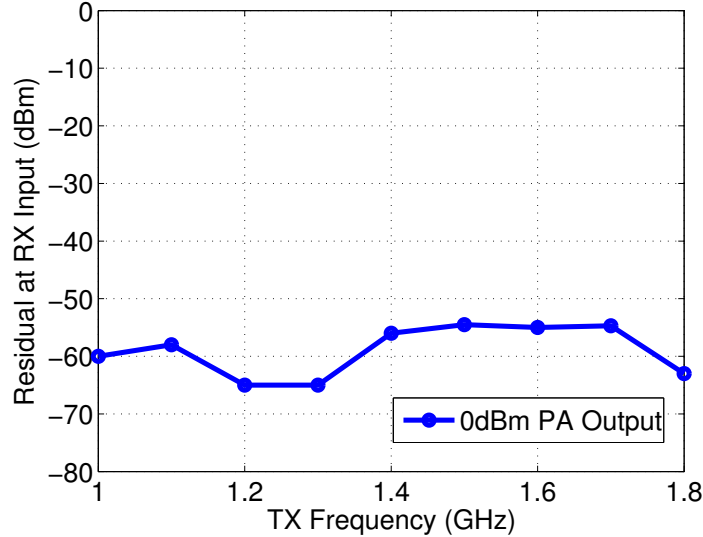


Figure 3.36: TX residual vs. center frequency.

The residual is also measured across frequency from 1GHz to 1.8GHz by setting the TX

code to produce a 0dBm output as the TX LO is swept. In all cases, the residual is still limited by the DAC LSB, demonstrating that the cancellation system is wideband, due to the wideband nature of the DAC, and the low impedance of the summing junction with respect to the DAC signal.

Noise degradation measurements are performed as TX power is swept. A Keysight 9030A spectrum analyzer, which can reach a noise floor of -165dBm/Hz is used to measure the noise spectrum. First, a TX single tone is input, and the optimum DAC cancellation code is found, observing the TX band. Then the spectrum analyzer bandwidth is restricted to the RX band, to lower the noise floor and to ensure the ADC dynamic range is not exercised.

Intuitively, 3 noise contributions are expected as a function of TX power. For low TX power, the receiver noise floor should dominate, and the noise figure degradation should be constant against TX power. For moderate TX power, the DAC's thermal noise should dominate. As the DAC noise variance is linearly proportional to the cancellation current, so square root proportional to the TX power, this would provide an increase of 3dB per doubling of TX power. For larger TX power, the uncorrelated component of TX phase noise falling in the RX band should dominate the RX noise figure. As this noise is proportional to the TX power, and not current, it should provide an increase of 6dB per doubling of TX power.

From the fit shown in Fig. 3.37, the measured curve is well modeled with a constant plus 6dB/octave model indicated that the degraded noise figure is due mostly to un-cancelled phase noise that falls in the RX band, and the DAC thermal noise limit has not been reached.

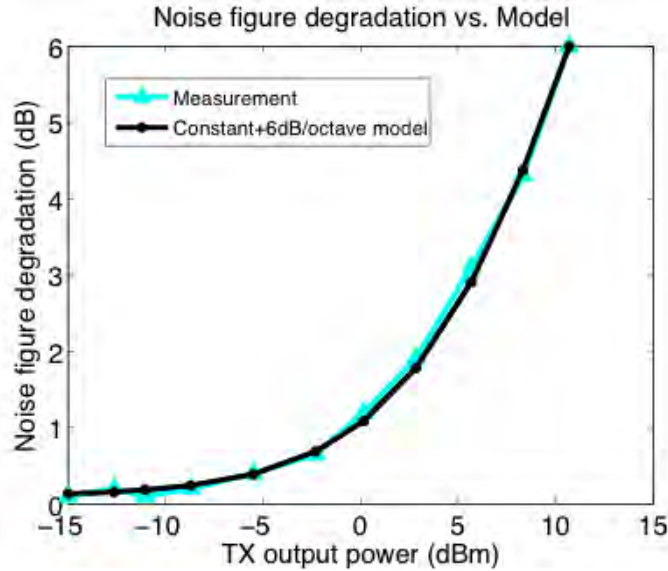


Figure 3.37: Noise figure degradation measurement vs curve fit of constant + 6dB/decade.

This is corroborated through noticing that in the Fig. 3.38, even up to a maximum power

input into the chip clock port, the noise figure degradation continues to drop. This indicates that further improvement could be reached with lower on-chip LO path phase noise. This is likely due to three issues - the on board loss of the clock routing, suboptimal performance of the clock receiver due to limited output swing from the diode connected PMOS load (as verified in simulation), and high frequency noise folding in the clock divider. Further investigation of these mechanisms are presented at the end of Section 3.5.5.

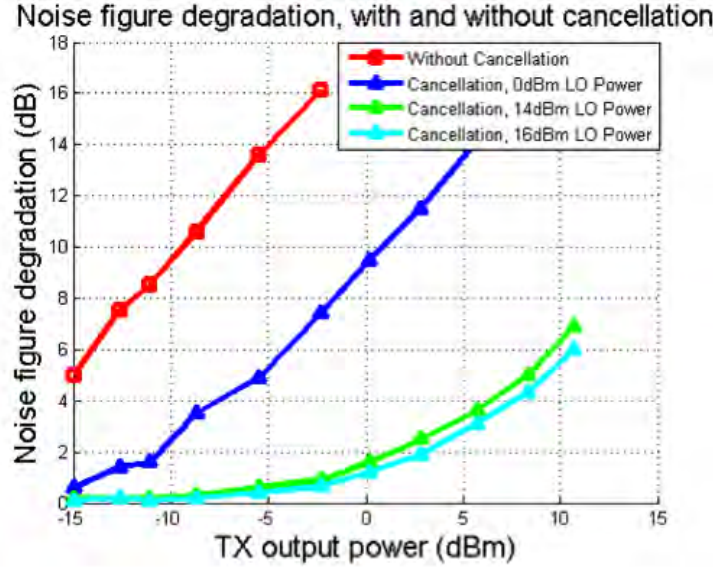


Figure 3.38: Noise Figure degradation vs TX LO power.

At a +2dBm TX output power the system incurs a moderate 1.7dB noise figure penalty, with a 4.3dB penalty at +10.6dBm. Compared to other state of the art active cancellation networks, [38][20][21][32][34][22], this system provides isolation for higher TX powers, with over 15dB isolation for single tone, and over 20dB higher isolation for 20MHz modulated bandwidth, limited by the DAC resolution. Additionally, this is the only work to fully integrate both a transmitter and receiver, with a single antenna interface with no external isolation. This is important for two reasons. Firstly the system realistically captures the TX to RX coupling mechanisms, as opposed to test setups which inject an off-chip generated interferer. Secondly, many of the prior demonstrations relied on narrowband external isolation in the form of board level changes to function over their reported operating frequency ranges. The performance is summarized in Fig. 3.39.

3.5.4 Antenna Mismatch

Mismatch of the antenna impedance can change the TX-RX coupling path in a frequency selective manner, making cancellation difficult over a wide TX modulation bandwidth. In this

	[20]	[21]	[33]	[22]	[38]	[34]	[12]	[13]	This Work
Architecture	CG/CS LNA canceller	RF-FD Equalizer	BB Vector Modulator	N-path Circulator	Feed-forward Canceller	Distributed Amp	Xfmr Hybrid	Xfmr Hybrid	Cancellation DAC
Frequency (GHz)	.3-1.7	.8-1.4	.15-3.5	.6-8	2.3-2.5	.3-1.6	1.5-2.1	1.9-2.2	1.0-1.8
TX/RX offset(MHz)	-	110	0	0	-	115			40
Max TX Power leakage (dBm)	+2	-8	+1.5	-6	0	+14 ¹	<+12	>+27	+12.6
Cancellation at Max TX Power (dB)	>30	33	>27	42	90	>40	50	50	>50
Cancellation over 20MHz Modulated Bandwidth (dB)	-	20	27	46 ²	50	>25	50	50	>50
RX NF (dB)	4.2	7.5 ³	6.3	5.0	4.7	8-12	5	3.9 (no RX)	7.6
NF Degradation at +2dBm TX power	.8	.9 ⁴	4	5.9	.25	4-5 ⁵	-	-	1.1
Fully-Integrated TX+RX	No	No	Yes	No	No	Yes	No	No	Yes
Canceller Power (mW)	13-72	44-182	23-56	89	-	0 ⁶	0	0	60
TX Insertion Loss (dB)	-	-	-	1.7	5	0	2.5	3.7	0
Noise Cancellation in RX band (dB)	13		-			15	50	50	19
Active Area (mm²)	1.2	4.8	2	1.44	1	7.2	.1	1.75	3.9
Technology	65nm	65nm	65nm	65nm	130nm	65nm	65nm	180nm SOI	65nm

1) With RX-band TX degeneration 2) 12MHz modulation bandwidth 3) Includes 2.7dB LC duplexer loss 4) TX power for this measurement not reported 5) NF degradation reported here for moderate TX power 6) Does not include power for RX tracking noise degeneration

Figure 3.39: Comparison Table.

work, the cancellation DAC input sequence is adjusted to match the frequency selectivity of the antenna mismatch. This is done by varying the antenna impedance, and doing a gradient descent on the DAC input data to re-minimize the power in the RX band. At a fixed 1.4GHz center frequency, the antenna impedance is varied over impedance points shown in Fig. 3.41 by adding a tunable length transmission line terminated with a short in shunt with the antenna, as Fig. 3.40, to create a VSWR up to 5:1. Note that the tested impedance points represent a limitation in the test setup, rather than the chip itself. After digital adaptation of the DAC input sequence, the residual TX signal is still limited by the DAC LSB, demonstrating the flexibility of the cancellation system due to the wide-band nature of the DAC, the low impedance of the summing junction, and digital adaptation of the input data.

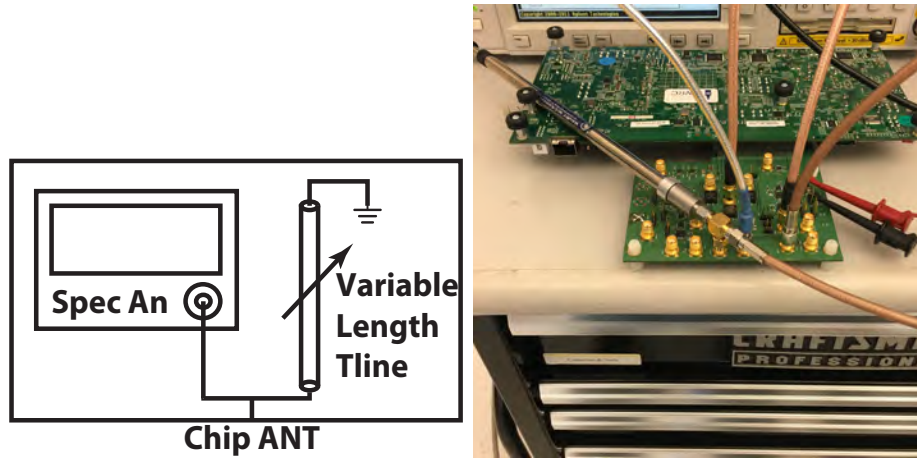


Figure 3.40: Measurement setup for VSWR.

3.5.5 Phase Noise Measurements

Several tests were run to measure the efficacy of the proposed feedforward phase noise cancellation.

In the first test, spurs are injected at the input of the shared TX/DAC LO port, to verify the folding relationships into the I and Q LO's given by the Fourier series coefficients derived in Chapter 2. The amplitude component of the injected noise is rejected by the clock receiver, which provides rail to rail swing at the input of the divider. The resulting spurs on the I LO, and the Q LO are measured separately through setting the PA to transmit in the I phase or the Q phase. The phase relationship between these spurs is measured and shown in Fig. 3.42, and are consistent with the predictions given by the Fourier series coefficients.

Next, a single sideband spur at $2F_{TX} + F_{offset}$ is injected into the shared LO port input. Given the analysis of Chapter 2, the -1 sideband down-converts the spur to $F_{TX} + F_{offset}$.

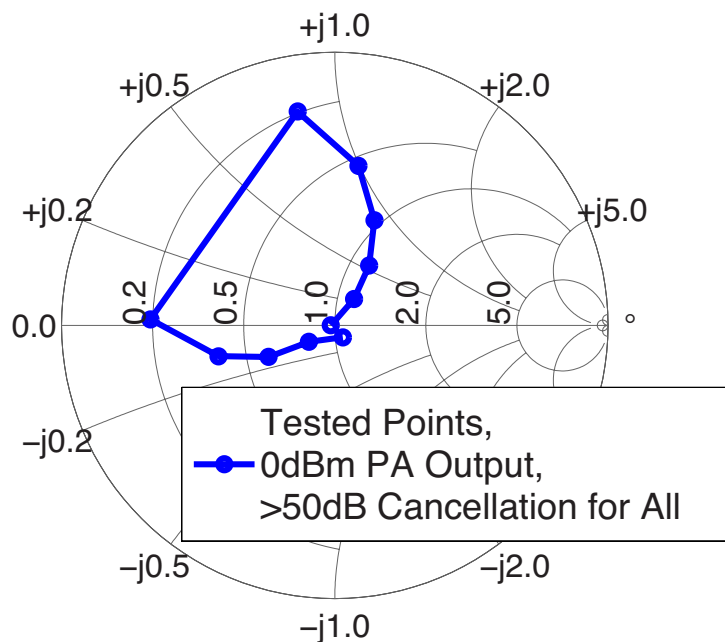


Figure 3.41: Antenna impedance points measured.

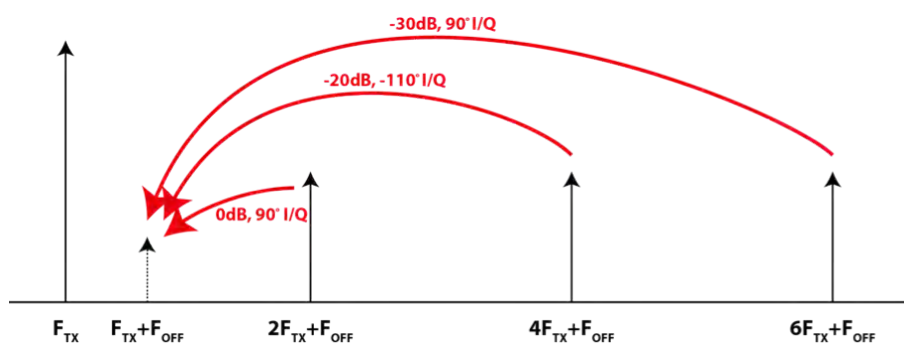


Figure 3.42: Phase relationship of I and Q spurs for injected input spurs.

The spur is measured with and without the cancellation DAC enabled, and the result is shown in Fig. 3.43. The close-in spur cancellation is $>35\text{dB}$, while the cancellation of a spur in the RX band at 40MHz spacing is $>20\text{dB}$.

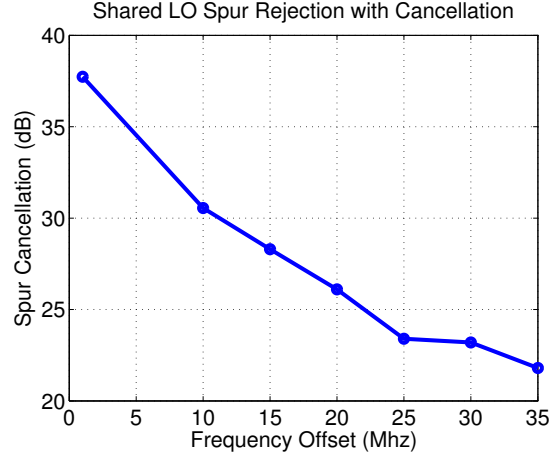


Figure 3.43: Spur cancellation vs offset frequency.

Intuitively, the cancellation at close-in spur offset is high, approaching the cancellation for a single tone itself. The spur cancellation drops off vs. frequency - the I and Q values at the DAC output are fixed to cancel the main TX tone, but a spur at an offset frequency experiences a different analog amplitude and phase shift, and accordingly is not perfectly cancelled by the fixed shift. This cancellation is effectively a function of the channel bandwidth, and could be extended for a wider bandwidth network.

White noise from a source generator, filtered by a 100MHz tunable bandpass filter is then injected into the LO port. As the noise is filtered to a narrow bandwidth around $2F_{TX}$, no additional noise folding terms should contribute. The noise can simply be thought of as a sum of spurs, so intuitively cancellation result should closely match the result for a single LO spur. The measurement shows this to be the case - noise close-in is cancelled by around 35 to 40dB , while noise at 40MHz offset is cancelled by around 20dB .

Wideband phase noise with 3GHz 3dB corner is injected into the LO port, to emulate the case of wideband source noise. Noise around $4F_{TX} + F_{offset}$ is folded by the -3 sideband with -90 degree phase relationship between I and Q, and accordingly is un-cancelled. From the spur folding measurement above, the strength of this folding term is -20dB down relative to the strength of the noise term folded from $2F_{TX} + F_{offset}$. As this term remains after cancellation, the cancellation is limited to 20dB even close-in, verified in the wide-band measurement in figure 3.46.

Finally, the effect of a delay in the leakage network from TX to RX is verified in measurement. This is measured by comparing two cases. First the TX and RX are directly coupled on chip through the series stacked transformer network, as in normal operation. In

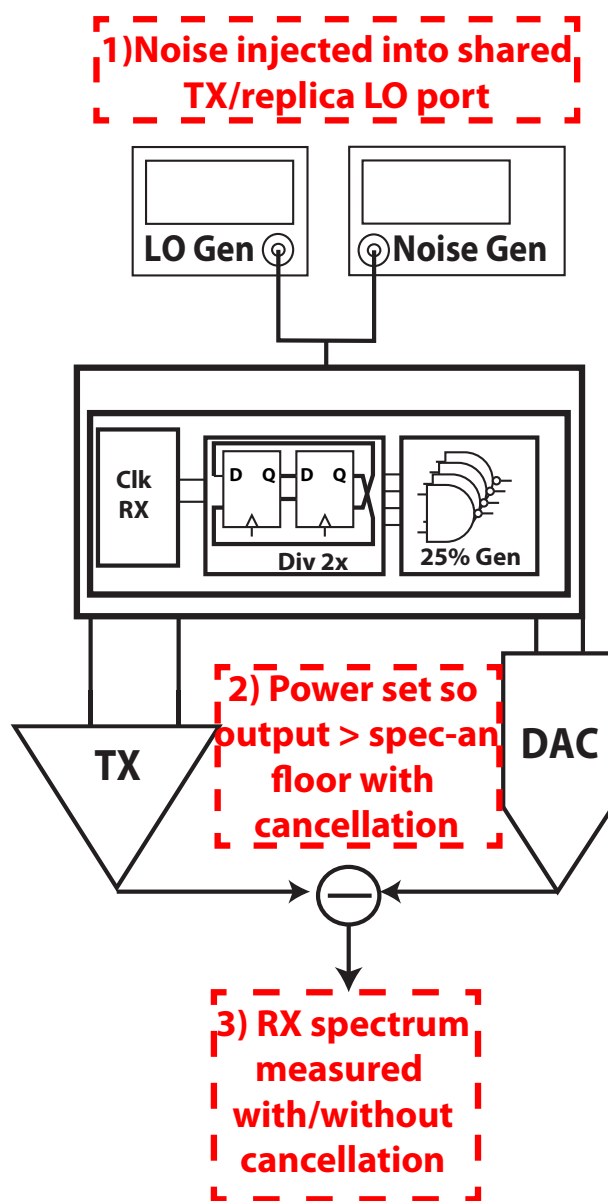


Figure 3.44: Test setup for phase noise cancellation.

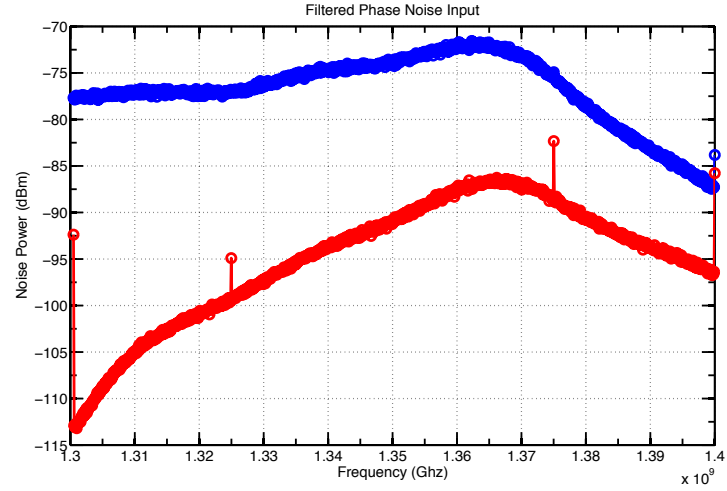


Figure 3.45: Filtered input noise spectrum at output before and after cancellation.

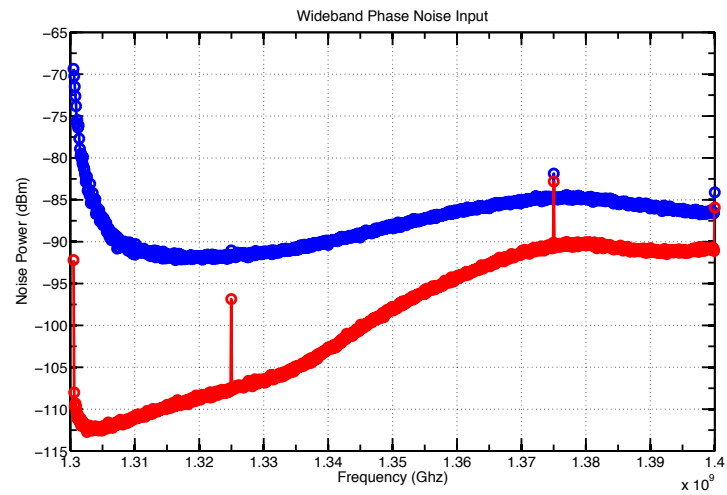


Figure 3.46: Wideband input noise spectrum at output before and after cancellation.

this case, the transformer path, with low electrical delay, is the dominant path. Next, the mid-point of the transformer between TX and RX is grounded on the PCB, isolating the TX and RX. The TX is then coupled to the RX on board through a series 50Ω resistor and an SMA cable, providing a longer electrical delay, as shown in the phase response of Fig. 3.47.

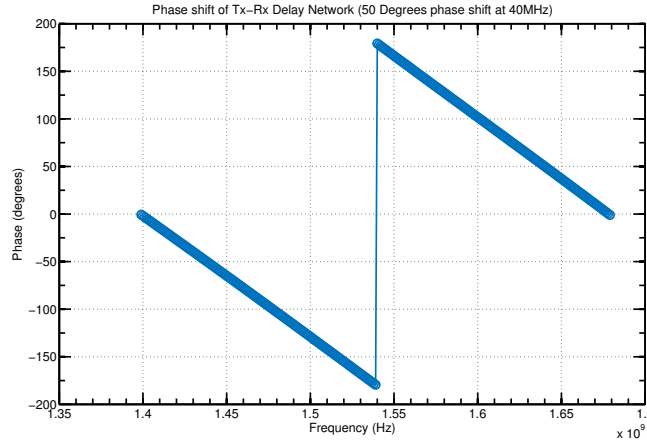


Figure 3.47: Phase shift of TX to RX coupling path.

By adding a delay which produces a significant frequency dependent phase shift (50° at the 40MHz duplex offset), the TX phase noise cancellation becomes narrowband, shown in Fig. 3.48, as predicted by analysis. Once again, the digital cartesian DAC weights its I and Q LO's to match the main TX tone. The phase noise is only cancelled over the limited frequency range where the phase difference between the main tone and the phase noise at the offset is the same. Under this condition the DAC's phase noise is produced with the appropriate phase shift to match the TX phase noise. At large frequency offset, the cancellation DAC and TX phase noise are phase shifted, and do not cancel. Instead, as seen on the right of Fig. 3.48 they can add constructively, and the noise can be worse than just the TX phase noise on its own. Accordingly, the TX-RX leakage network must be made wide-band to maximize the bandwidth of phase noise cancellation. Any longer delay TX-RX paths should be heavily attenuated to prevent this desensitization.

NF Degradation vs. LO Power

It is observed that RX noise figure degradation decreased heavily with increased LO swing into the TX LO port. This is due to design of on chip input clock receiver, shown in figure 3.49.

The phase noise at the output of the clock divider is simulated at two input swings on the board (450mV and 5V amplitude), with results shown below. For the low swing, the high overdrive voltage of the input devices keep the input clock receiver differential pair

transistors (M0 and M1) in saturation, and their noise is dominant. The diode connected PMOS loads (M4 and M5) on the input differential pair limit the input swing on the top end, and their noise is also significant. When the input swing is increased, one side of the NMOS input differential pair is cut off, while the other side acts in triode operation, with its noise circulated by the high impedance tail current source - thus their noise is reduced. The PMOS load noise contribution is also decreased, as it is referred against a larger output swing. The dominant sources become the tail reference mirror (M3) noise - which could be reduced with appropriate decoupling and filtering on the reference, and the CML to CMOS converter following the clock divider. This noise could be reduced by increasing inverter size.

Contributor	Noise $\frac{V}{\sqrt{Hz}}$	% Contribution
M3	7.87e-16	6.31%
M1	7.01e-16	5.62%
M0	2.57e-16	2.86%

Contributor	Noise $\frac{V}{\sqrt{Hz}}$	% Contribution
M1	1.17e-15	4.50%
M0	1.12e-15	4.29%
M5	9.64e-16	3.69%
M9	8.52e-15	3.28%
M4	7.11e-15	2.73%

Figure 3.50: Simulated noise contributions of the LO divider.

Chapter 4

Receiver Design for FD/FDD Systems

4.1 Design Motivation

The receiver's linearity is a key limitation in operating a TX and RX simultaneously at high TX output powers. This chapter describes a highly linear RF receiver design suitable for FD/FDD systems.

Fig. 4.1 summarizes why the transceiver chip described in the previous chapter can support only +10dBm TX power, even though the scheme provides over 50dB of TX signal rejection. The cancellation DAC only operates at the TX fundamental frequency. As a digital switching PA is used with limited TX filtering, a strong TX third harmonic content is produced, which reaches the RX input. This un-cancelled 3rd harmonic is large enough to desensitize the receiver at >10dBm TX power, as seen from Fig. 4.1. In particular, the third harmonic at the transmitter output is only $20 \times \log_{10}(3) = 9.5\text{dB}$ lower than the fundamental, as the Fourier series coefficient for the N^{th} harmonic in a square wave scales as $\frac{1}{N}$. At the RX input, while this is filtered by the front-end transformer network, the power is reduced by only an additional 10dB. In particular, the effective leakage channel response at the third harmonic need not match the response at the fundamental, so tuning the cancellation DAC to provide rejection at the fundamental will not cancel the third harmonic.

There are several options for rejecting the TX third within the system, as show in Fig. 4.2. Techniques which reject the third harmonic within the PA itself have recently been actively researched [68][63][69], due to the desire for widely frequency tunable transmitters which meet spectral mask requirements. One approach of harmonic rejection is through PA segmentation and output recombining with a 60 degree fundamental phase shift, resulting in a 180 degree phase shift at the third harmonic. The cancellation DAC's third harmonic output could potentially be shaped either in a similar manner, or through resonant impedances which tune the third harmonic phase shift and achieve feed forward cancellation of the third harmonic with the PA. A resonant trap could be placed in shunt with the receiver to provide a low impedance filtering path for the far-out third harmonic. This could be implemented either as a passive tunable trap, or with an n-path filter as in [30], though such an n-path filter

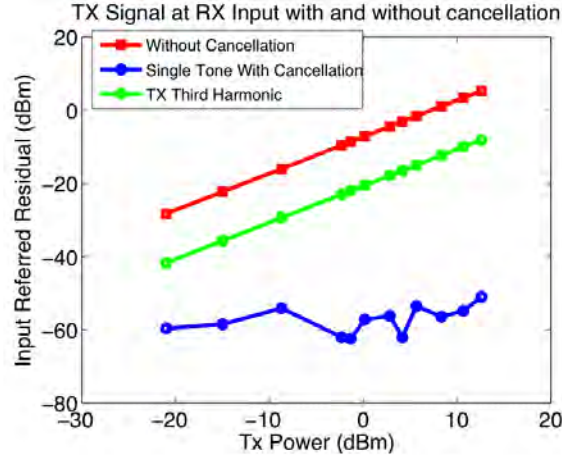


Figure 4.1: Third harmonic power vs TX power at antenna.

is likely to produce spurious re-radiated emissions at the antenna. A receiver with higher OOB linearity also enables RX operation in the presence of higher power third harmonic content.

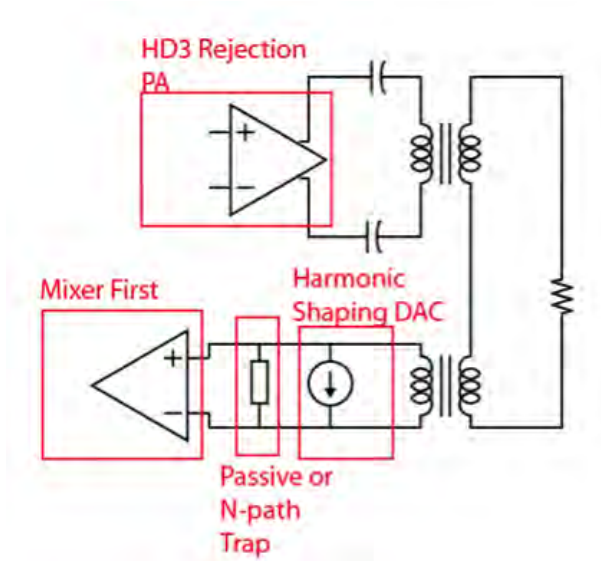


Figure 4.2: Potential options for TX third harmonic cancellation.

In this chapter, a passive-mixer-first receiver design targeting high out-of-band linearity is described. In particular, design of high linearity receivers which can down-convert both the TX residual and desired signal while maintaining acceptable noise figure and power

consumption is a key challenge of FD/FDD transceiver systems. This is apparent from noting the nominally high RX noise figure and power consumption of the previously described transceiver chip. As will be shown, the passive-mixer-first receiver provides a unique knob for tuning linearity and noise performance against power consumption, making it a good candidate for a FD or FDD transceiver.

In the first version of the transceiver, the transformer matching network is designed with an inductance well above the optimum for maximum available gain, due to the requirement to resonate the DAC's output capacitance at the desired center frequency, as shown in Fig. 4.3.

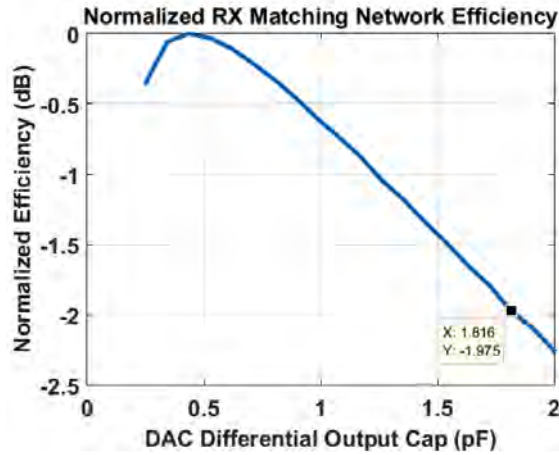


Figure 4.3: Available gain of the transformer network vs. DAC cap.

Accordingly, the system stands to gain almost 2dB in overall noise figure from improvement in passive design. This can be accomplished either through reduction of the DAC differential output capacitance, or with a receiver providing an inductive component in its input impedance to partially resonate the DAC capacitance. The former is described in [52]. Here, it will be shown that through baseband cross-coupling of I and Q paths, the passive-mixer-first receiver can present a tunable complex impedance which can resonate the DAC capacitance to optimize the passive network efficiency. In general, the tunable input impedance allows the receiver to absorb some of the parasitics of front-end cancellation circuits, further justifying its use in a FD/FDD transceiver system.

A receiver design is described in this chapter, employing complementary class-AB amplifiers to achieve a favorable noise/linearity tradeoff. Other transceiver improvements, primarily design of the tunable third harmonic TX trap to push maximum handled TX power, are described in [52]. Section 4.2 describes the fundamentals of the passive-mixer-first receiver. Section 4.3 discusses the proposed amplifier design, and Section 4.4 describes merging the harmonic recombination amplifiers with baseband biquads. The measurement results for the receiver, as well as its impact on the full duplex transceiver system, are summarized in

Section 4.6. In particular, the receiver boosts system performance, enabling cancellation of higher TX powers at lower overall system noise figure.

4.2 Passive Mixer First Receiver

The passive-mixer-first receiver is a promising option for an FD or FDD system, given the high linearity provided by up-converted baseband filter to the antenna input. Many of the derivations are thoroughly documented in [70], so just the key derivations and insights, and a few extensions, are documented here.

4.2.1 Input Matching

The RX input impedance is tunable through the baseband impedance, due to the transparency of the passive mixer. As the voltage on antenna side is always connected to one of the baseband outputs, the baseband impedance sets the RF impedance. The design equations for the baseband impedance are derived in [70]. The derivation is summarized below, and follows 3 steps, first assuming a single tone input in steady state.

1) Set up a charge balance equation current to solve for $V_{c,b}$ in the schematic of Fig. 4.4. In steady state, the current from the RF port must equal current discharged through the baseband resistor R_b , and accordingly

$$R'_a = R_a + R_{sw} \quad (4.1)$$

$$\frac{V_{c,b}}{R_B} T_{LO} = \int_{\frac{bT_{LO}}{4} - \frac{T_{LO}}{8}}^{\frac{(b+1)T_{LO}}{4} - \frac{T_{LO}}{8}} \frac{V_{rf}(T) - V_{c,b}}{R'_a} dt. \quad (4.2)$$

For an input cosine,

$$V_{rf}(t) = A \cos(\omega t + \phi) \quad (4.3)$$

$$V_{c,b} = \frac{AR_B}{R_B + NR'_a} \text{sinc}\left(\frac{\pi}{N}\right) \cos\left(\phi + \frac{2\pi b}{N}\right). \quad (4.4)$$

2) V_x , the node on the RF side, can be written as a piecewise composition of the $V_{c,b}$ terms in periods of $\frac{T_{LO}}{N}$, as pictured in Fig. 4.5. The first harmonic term of the Fourier series can be extracted and used to find the first harmonic input impedance, as

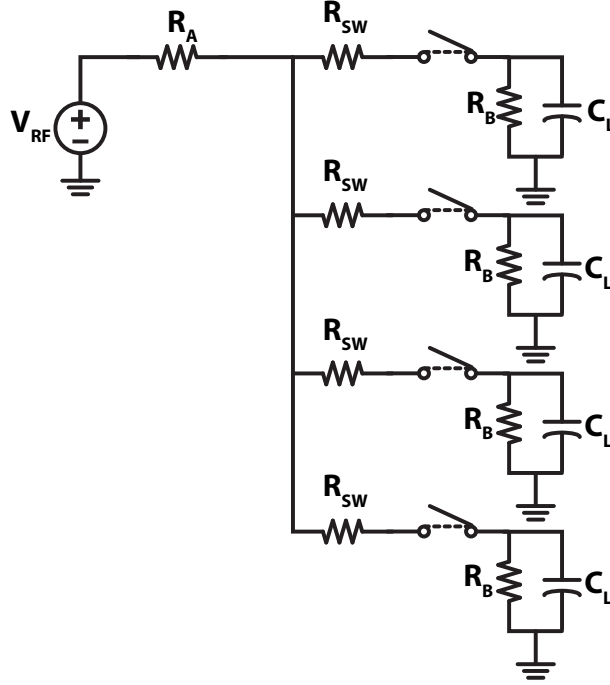


Figure 4.4: Passive-mixer-first input match model.

$$v_{x,fund}(t) = v_{rf}(t) \times \frac{R_b \operatorname{sinc}(\frac{\pi}{N})}{R_B + NR'_a} \quad (4.5)$$

$$Z_{in} = \frac{v_{rf}(t) - v_{x,fund}(t)}{R'_a} \quad (4.6)$$

$$Z_{in} = \frac{(NR'_a + (1 - \operatorname{sinc}^2(\frac{\pi}{N}))R_B)}{R'_a R_B + NR_a'^2} \quad (4.7)$$

3) This input impedance for the linear time-varying (LTV) system can be matched with the linear time-invariant (LTI model) shown below, for appropriate values of R_{sh} and γ . The γ term essentially represents the first harmonic conversion gain, and can be written as a function of the number of LO phases, and R_{sh} represents the loss mechanism due to re-up conversion of the baseband voltage through LO harmonics dissipated in the antenna resistance.

The model can be generalized using the same procedure above to determine the frequency dependence of R_{sh} as a function of Z_{ant} at harmonics of the LO and R_b , resulting in the following relationships:

$$\gamma = \frac{1}{N} \operatorname{sinc}^2(\frac{\pi}{N}) \quad (4.8)$$

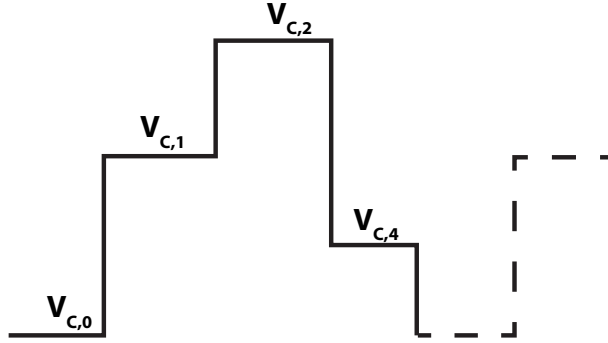


Figure 4.5: Voltage on RF side of the mixer.

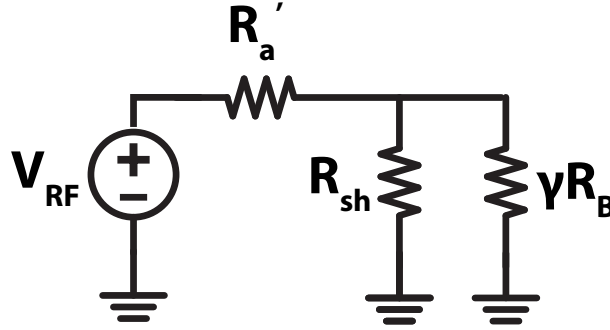


Figure 4.6: Equivalent linear impedance model of the RX.

$$Z_{sh} = \sum_{n=(N-1), (2N-1), \dots}^{\infty} \left(\frac{1}{n^2 Z_{ant,n}} \right)^* + \sum_{n=(N-1), (2N-1), \dots}^{\infty} \left(\frac{1}{n^2 Z_{ant,n}} \right) \quad (4.9)$$

It is worth noting that the baseband capacitive reactance of Z_{sh} and C_{bb} look inductive at negative IF frequency. This inductance resonates with capacitance on the antenna, shifting the impedance match to slightly below the LO frequency. For example, the S11 shown for a representative design with 8-phase $30\mu m$ mixer switches with 40pF per phase baseband cap, corresponding to a baseband bandwidth of 20MHz resonates roughly 20MHz below the 1GHz LO, as shown in Fig. 4.7.

The resonance can be tuned to match at the LO frequency by introducing a tunable reactive baseband element whose phase is unaffected by the IF frequency. A phase dependent element would create an effective capacitance on one sideband of the LO and an inductance on the other sideband. Such a reactive impedance can be created by cross-coupling the 90 degree phase-shifted baseband outputs to the input of the opposite branch, as shown in Fig. 4.8 [70].

Intuitively, cross-coupling introduces a shunt current at the input of the baseband ampli-

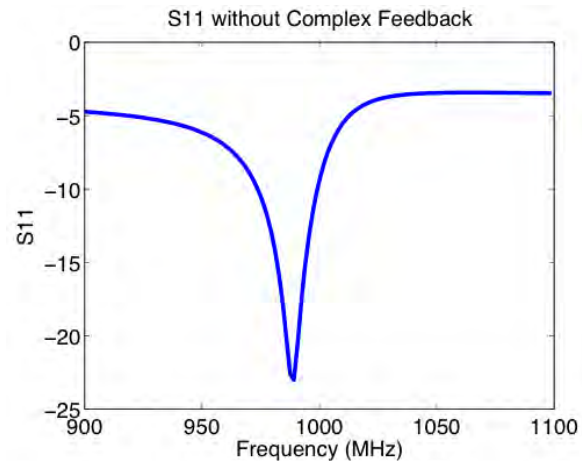


Figure 4.7: S11 offset of the receiver.

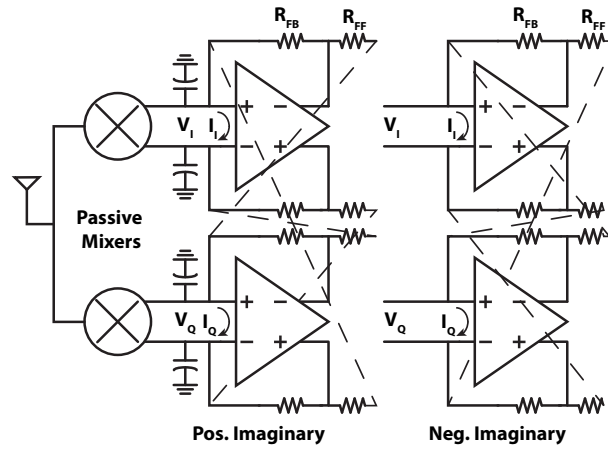


Figure 4.8: Cross-coupled baseband for complex input impedance creation.

fier that is 90 degrees out of phase with the signal. The strength of this quadrature current component is set by the frequency independent R_{ff} resistor. When upconverted by the LO, looking in from the antenna side, this creates a complex impedance, as a sine component of input current is drawn for an applied cosine voltage. The impedance can be written as

$$Z_{bb} = \left(\frac{1+A}{R_{fb}} + \frac{1}{R_{ff}} \pm j \times \frac{A}{R_{ff}} \right)^{-1} \quad (4.10)$$

where A represents the gain of the first stage baseband amplifier.

i

$$R_{ff} = \frac{A}{\omega C_{rf} \gamma} \quad (4.11)$$

$$R_{fb} = \frac{(1+A)R_B R_{ff}}{R_{ff} - R_B} \quad (4.12)$$

$$R_B = \frac{1}{\gamma} \times \frac{(R_{ant} - R_{sw})R_{sh}}{R_{sh} - (R_{ant} - R_{sw})} \quad (4.13)$$

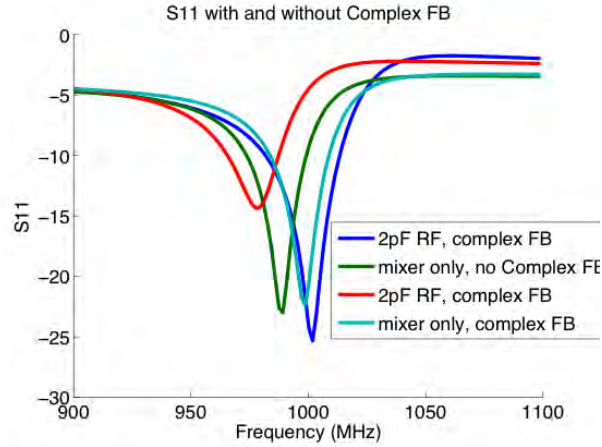


Figure 4.9: S11 with and without complex feedback.

For high switch resistance, the R_{sw} dependent transformation of the complex baseband impedance makes an analytical solution for R_{fb} and R_{ff} difficult, but they can be solved numerically. This technique allows absorption of the RF capacitance, and cancellation network parasitics.

4.2.2 Noise Analysis

One of the key insights from [70] is that the linear time-varying (LTV) mixer-first system can be equivalently modeled for noise purposes as the same figure 4.6 LTI system used for

input impedance, where the antenna impedance sees R_{sw} in series with shunt Z_{sh} and shunt γZ_b .

There are a few questions that arise from this claim 1) is it correct to consider this LTV process with an LTI model, only considering a single baseband? 2) Where does the γ term on the baseband noise power come from? 3) What is the significance of Z_{sh} ?

From the antenna port, the cyclostationary nature of the noise process is hidden by the fact that in steady state, each baseband path is equivalent for noise analysis, even if some memory is included in the baseband. For example, one could do a thought experiment of an observer viewing the noise voltage of N resistors of the same value R_b being switched into the measurement node periodically. The observer shouldn't be able to distinguish this from just constantly observing the noise of a single resistor of value R_b . Accordingly, only a single baseband's noise needs to be considered.

The γ factor applied to the noise baseband power is more intuitively applied to the antenna side as the first harmonic's conversion gain. Accordingly, the baseband noise should be referenced to $1/\gamma$ multiplied by the antenna noise power.

The shunt Z_{sh} is a fictitious noise source which produces the same noise power at the baseband as the antenna noise at LO harmonics folded by the passive mixer. This can be proven by verifying the following relationship:

$$Z_{sh} = \sum \left(\frac{4kT}{n^2 R'_a(w_{LO})} \right) = \frac{4kT}{Re(Z_{sh})} \quad (4.14)$$

Notably missing from [70] is the effect of the feedforward branch of the complex impedance matching path on the noise figure. As the noise analysis is done by considering the the input referred voltage and current noise of a single baseband output branch while the other mixer switches are off, the baseband feedforward branch is driven by an open circuit for noise analysis. The input referred voltage and current noise, or equivalently open circuit and short circuit input referred voltage noises in the model of Fig. 4.10, can be derived.

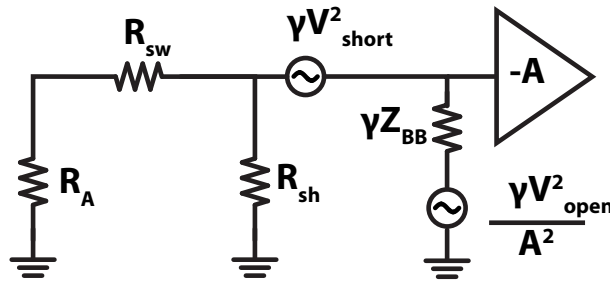


Figure 4.10: Model for noise analysis of the baseband.

The circuit of Fig. 4.8 can be simply modeled by a differential half circuit of Fig. 4.11 for noise analysis, if the magic circuit element (-1) shown below has the property of inverting

both the sign of the voltage with reference to ground and the current direction across it. Such a model maintains the same KCL equations as the original circuit.

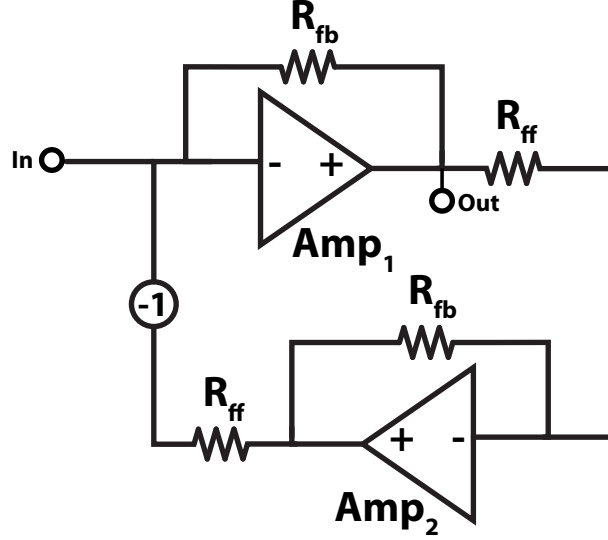


Figure 4.11: Noise model of the cross-coupled feedback path.

The short circuit noise voltage is simply the input referred voltage noise of the main branch amp.

$$V_{short}^2 = V_{n1}^2 \quad (4.15)$$

The open circuit voltage noise has contributions from both amplifiers and all 4 resistors, with the transfer functions are shown below:

$$V_{out_{Rff1}}^2 = V_{Rff1}^2 \left(\frac{A^2 R_{fb}^2}{A^2 R_{fb}^2 + (R_{fb} + (1 + A)R_{ff})^2} \right)^2 \quad (4.16)$$

$$V_{out_{Rff2}}^2 = V_{Rff2}^2 \left(\frac{AR_{fb}(R_{fb} + (1 + A)R_{ff})}{A^2 R_{fb}^2 + (R_{fb} + (1 + A)R_{ff})^2} \right)^2 \quad (4.17)$$

$$V_{out_{Rfb1}}^2 = V_{Rfb1}^2 \left(\frac{AR_{ff}(R_{fb} + (1 + A)R_{ff})}{A^2 R_{fb}^2 + (R_{fb} + (1 + A)R_{ff})^2} \right)^2 \quad (4.18)$$

$$V_{out_{Rfb2}}^2 = V_{Rfb2}^2 \left(\frac{A^2 R_{ff} R_{fb}}{A^2 R_{fb}^2 + (R_{fb} + (1 + A)R_{ff})^2} \right)^2 \quad (4.19)$$

$$V_{out_{n1}}^2 = V_{n1}^2 \left(\frac{A(R_{fb} + R_{ff})(R_{fb} + (1 + A)R_{ff})}{A^2 R_{fb}^2 + (R_{fb} + (1 + A)R_{ff})^2} \right)^2 \quad (4.20)$$

$$V_{out_{n2}}^2 = V_{n2}^2 \left(\frac{A^2 R_{fb}(AR_{ff} - (R_{fb} + (1 + A)R_{ff}))}{A^2 R_{fb}^2 + (R_{fb} + (1 + A)R_{ff})^2} \right)^2 \quad (4.21)$$

$$V_{open}^2 = V_{out_{Rff1}}^2 + V_{out_{Rff2}}^2 + V_{out_{Rfb1}}^2 + V_{out_{vn1}}^2 + V_{out_{vn2}}^2 \quad (4.22)$$

The noise figure can then be written by referring all noise sources to the baseband node.

$$\begin{aligned} NF = 10\log_{10} & \left(1 + \frac{R_{sw}}{\text{real}(Z_{ant})} + \right. \\ & \frac{\text{real}(Z_{sh})}{\text{real}(Z_{ant})} \left| \frac{Z_{ant} + R_{sw}}{Z_{sh}} \right|^2 + \\ & \frac{\gamma V_{short}^2}{4kT \times \text{real}(Z_{ant})} \left| \frac{Z_{ant} + R_{sw} + Z_{sh}}{Z_{sh}} \right|^2 + \\ & \left. \frac{\gamma V_{open}^2}{4kT \times \text{real}(Z_{ant}) A^2} \left| \frac{Z_{ant} + R_{sw}}{\gamma Z_{bb}} \right|^2 \right) \end{aligned}$$

From the noise figure equation, there are a few insights. The expression indicates that there is an optimal number of clock phases to minimize noise figure, for a given technology's switch fT . From the expression, the NF is monotonically decreased with increased Z_{sh} , so it is desirable to maximize the antenna impedance at odd harmonics beginning from N-1. Increasing N, the number of phases, can serve to increase Z_{sh} , by eliminating noise folding from the first N-2 harmonics. However, for a fixed R_{sw} , more phases corresponds to more mixer switches, adding additional cap on the RF side which decreases Z_{sh} . Fig. 4.12 shows NF assuming noiseless baseband amplifier as number of mixer phases is swept. In this technology, 16 phases is the optimal for noise. As the LO generation and distribution for 16 phases is power intensive, 8 phases is a more practical choice, as <.1dB penalty is incurred.

Secondly, the mixer size can be optimized for noise performance, due to the tradeoff between R_{sw} and Z_{sh} which is fixed given switch fT (or more precisely, technology parameters R_{on} and C_{drain}). Assume a minimum length device for a given technology with has an on resistance of $330\Omega \times \mu m$, and a C_{drain} of $1.6fF/\mu m$ with 1V V_{gs} with 0V V_{ds} . Fixing the requirement for complex impedance matching, and ignoring the noise from the baseband amplifier, the noise vs. mixer switch width with 8 phase mixer is plotted in Fig. 4.13 at 1GHz for varying baseband gains. The curve is shown when just the mixer capacitance is included on the RF side, and with an additional 2pF of RF cap added for the DAC and parasitics on the antenna interface. The optimum transistor width is roughly $70\mu m$ in all cases but this is a very shallow optimum. The optimum mixer size is slightly increased when the explicit cap is added, which is intuitive as its effect on Z_{sh} is amortized by the explicit RF cap.

Thirdly, noise figure can be traded for baseband amplifier power consumption. This tradeoff can be demonstrated by considering an inverter as the baseband amplifier. Fig. 4.14 shows the tradeoff of noise figure vs. power through the free parameter of inverter width, at a fixed inverter length of 640nm, at 1GHz. As gm is increased and more power is used by increasing width, the noise figure approaches the bound of noise figure with a noiseless

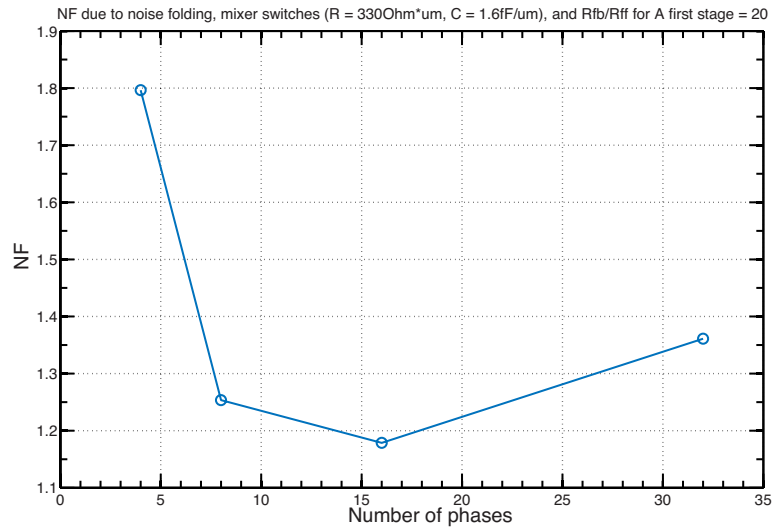


Figure 4.12: RX NF vs number of mixer phases.

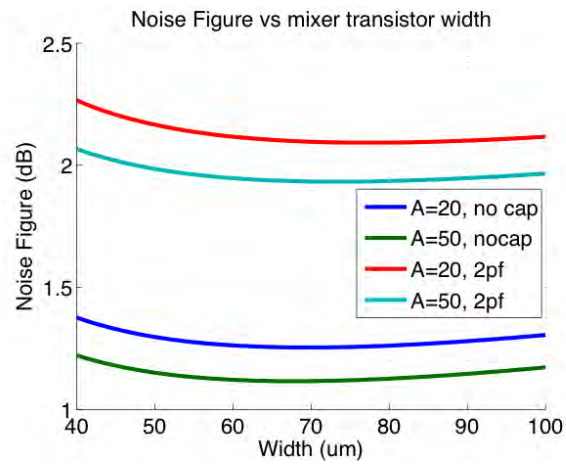


Figure 4.13: Noise figure vs. mixer switch size.

baseband amplifier. For a degradation of .5dB from the baseband amplifier, an NMOS width of $30\mu m$ (PMOS $120\mu m$) is chosen, which corresponds to a total power consumption of 20mA in all baseband amplifiers.

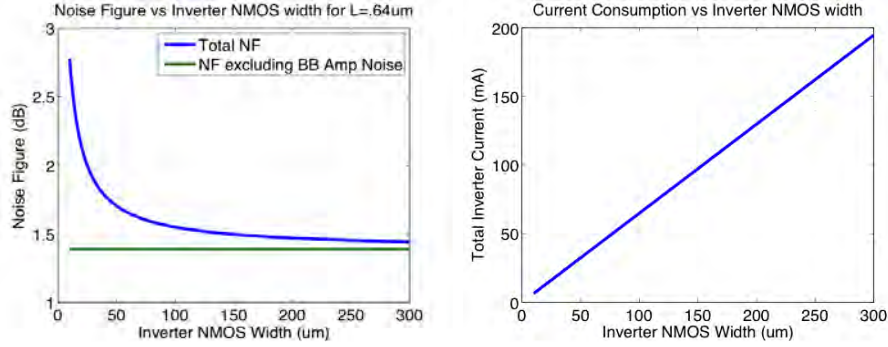


Figure 4.14: Noise figure (left) and power consumption (right) vs. inverter size.

Fourthly, the noise figure is a strong function of frequency, due to its dependence on Z_{sh} , as shown in Fig. 4.15. This is exacerbated for a large value of parasitic cap on the RF side. The noise figure vs. mixer switch width is shown in Fig. 4.16 for 2pF RF cap and no explicit RF cap (i.e. only cap from the mixer switches), as the LO frequency is moved from 1GHz to 2GHz. For the same RF cap, a smaller mixer switch optimizes the noise figure at higher frequency, corresponding to a smaller value of Z_{sh} . The optimum noise figure at 2GHz with 2pF of RF cap is around 3.1dB, again only considering the noise of the mixer switches and the matching resistors.

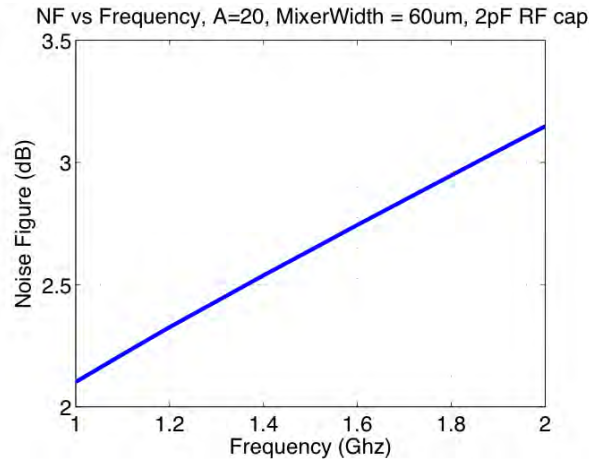


Figure 4.15: Noise figure vs frequency.

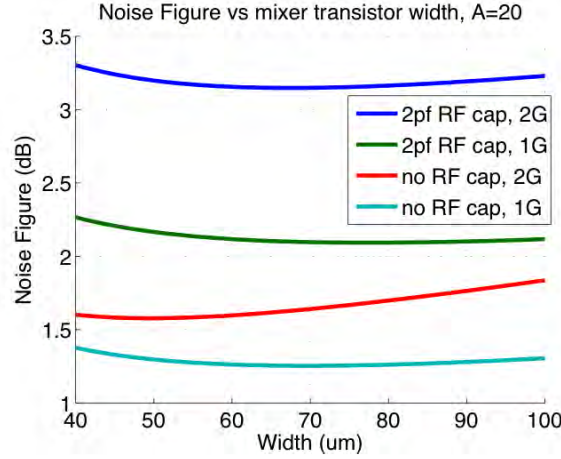


Figure 4.16: Noise figure for different RF cap values.

4.3 First Stage RX Amplifier Design

Given a fixed mixer switch size and amplifier gain, the feedback and complex feed-forward resistors in the baseband are fully constrained to achieve a matching condition. As shown earlier, the mixer switch has an optimum size to minimize noise figure, due to the tradeoff of mixer switch resistance versus capacitance. Accordingly, the main free parameter to select is the amplifier gain. It is shown here that this gain is a knob which trades off between noise and linearity.

As the Z_{in} of the amplifier is fixed in order to maintain an input match, the swing at the baseband amplifier input is fixed given a certain power at the antenna. This is rather unintuitive at first, as the amplifier's loop gain can't be increased in order to suppress its input voltage swing and increase linearity. This increase in amplifier gain must be compensated by an increase in feedback resistor, in order to maintain a fixed input match.

Further, it is preferable to use an op-amp as the baseband amplifier as opposed to an operational-transconductance-amplifier [43][71], for low noise performance. This is because the closed loop OTA input impedance is set at $\frac{1}{gm}$, independent of the feedback resistor, which couples the noise generator to the input matching. Accordingly, the topology can not achieve a NF lower than 3dB. An op-amp in feedback presents an input impedance of $\frac{R_f}{(1+A)}$. This allows the op-amp's input referred noise to be lowered for power at a fixed input match, through increasing the op-amp input pair transconductance at a fixed gain.

Assuming the op-amp noise can then be sufficiently reduced, the baseband noise is bounded by the noise contributed by the feedback resistors. As the feedback resistor noise in closed loop decreases with increased resistance, the value of the op-amp gain is essentially a parameter which selects a tradeoff point between linearity and noise, given a fixed input match. For a fixed antenna power, the matching condition constrains the input swing. Ac-

cordingly, for a fixed supply, the only way to increase linearity is to reduce the amplifier gain, reducing the feedback resistor to maintain a fixed input match, and increasing the noise.

The noise figure due to the mixer switches and feedback resistors is plotted in Fig. 4.17 vs. the amplifier gain, under the assumptions described in the figure. As a design procedure, one could either fix a target noise figure, thus setting the linearity, or vice versa.

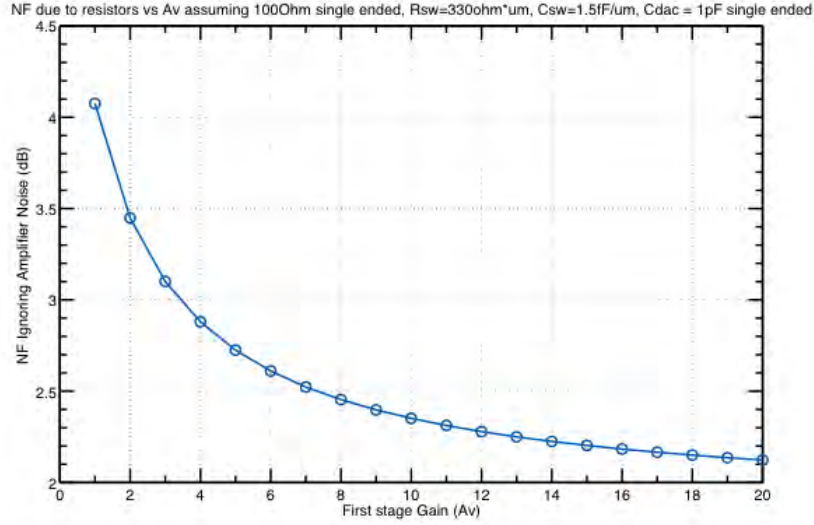


Figure 4.17: Noise figure with a noiseless baseband amp.

From the above plot, the first stage gain is fixed around 12, due to the diminishing returns in noise figure beyond that point.

Then, the required amplifier noise can be specified, to achieve a given degradation due to the amplifier on the overall total noise figure.

From Fig. 4.18, an op-amp with around $1 \frac{nV}{\sqrt{Hz}}$ of input referred noise is desired in order to degrade the overall NF by $< .5dB$ from the amplifier, and maintain $< 3dB$ NF overall.

If the first cut assumption is made that the majority of this noise contribution is set by the op-amps input pair, the input pair requires a gm of around $30mS$. For a $\frac{2}{(gm/Id)}$ of $150mV$ this requires around $2.5mA$ of current per device, close to $5mA$ differentially in just the first stage amplifier.

If a complementary structure is used for gm reuse, over the 4 baseband amplifiers (8 phases), a total of $10mA$, or $15mW$ from a $1.5V$ supply, can be saved - a significant amount as a state of the art RX front-end power consumes a total of $50mW$ [43].

The difficulty in the complementary design is the desire to bias both output and input voltages at mid-rail for maximizing linearity. The overdrive voltage of the input pair must be made low, in order to maximize the input pair gm within the current budget. In order to maintain control over the input pair overdrive voltage with a fixed input voltage bias,

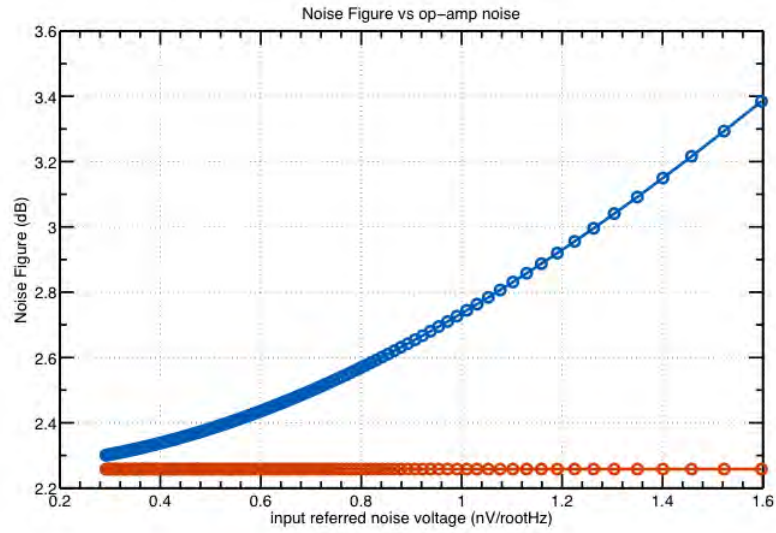


Figure 4.18: Noise figure vs amplifier input referred noise.

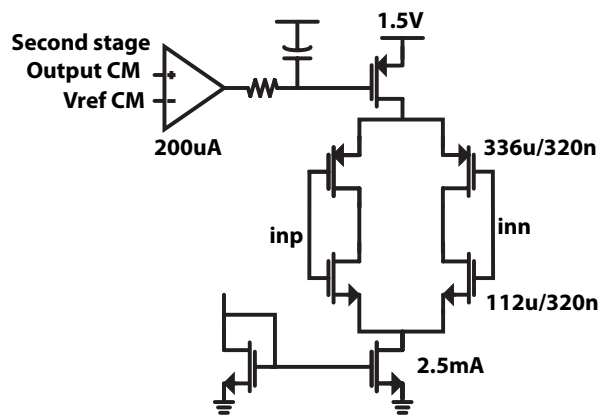


Figure 4.19: Complementary first stage schematic.

tail current sources can be used as in Fig. 4.19, as opposed to a pseudo-differential inverter-based input pair. However, this limits the output differential swing to around 500mV total. A second stage is desirable to support a higher output swing.

A local shunt-feedback second stage as in Fig. 4.20 is a promising solution, as the output swing of the first stage is suppressed by the second stage feedback. The total overall gain is set by the first stage gm times the second stage feedback resistor.

$$A_{DC} = \left(\frac{A_{amp2}}{1 + A_{amp2}} \right) gm_1 R_{fb,2} \quad (4.23)$$

Additionally, the input referred noise, shown below, is suppressed by the first stage transconductance, and the overall amplifier noise can accordingly be traded for power consumption.

$$v_{n,input}^2 = \left(\frac{A_{amp2}}{1 + A_{amp2}} \right)^2 \times \frac{1}{gm_1^2} \times \left(\frac{4kT}{R_{fb}} + \frac{v_{n,amp2}^2}{R_{fb}^2} + i_{n,amp1}^2 + 4kT\gamma gm_1 \right) \quad (4.24)$$

If a class-A amplifier is used in the second stage, then the DC current must be made twice the first stage current, in order to slew the peak current at the largest input swing. This requires an additional $5mA \times 4amplifiers \times 1.5V = 30mW$, prohibitive in a 50mW desired power budget. A class-AB amplifier is effective here to save this current in the nominal small signal condition, and provide this current under a large blocking condition. Additionally, the class AB output stage can swing nearly rail-to-rail, maximizing the receive chain's linearity.

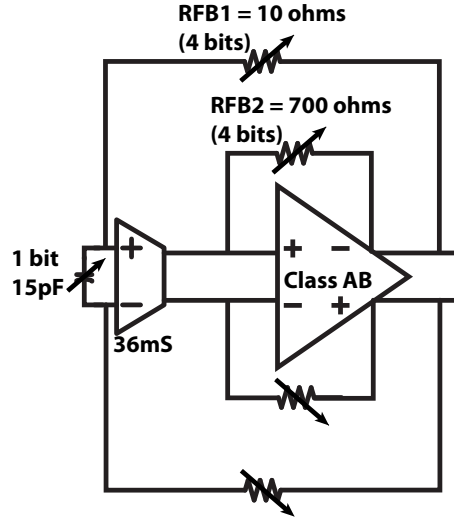


Figure 4.20: Baseband amplifier with shunt feedback second stage.

A schematic of the second stage of the first amplifier is pictured in Fig. 4.21. A second copy of the output stage is used for the other differential half. The floating battery

attempts to set the appropriate V_{gs} of the class-AB output devices to regulate the output stage quiescent current. This is accomplished as the NMOS output device and the NMOS floating battery device forms a trans-linear loop with the biasing branch to mirror the appropriate current into the output stage. An equivalent translinear loop regulates the PMOS current. Equivalently, the bias battery's V_{gs} voltages provide independent control of the output PMOS and NMOS V_{gs} , to set the output current.

This biasing battery is segmented with a portion used as a complementary common drain amplifier coupling the first amplifier to the output stage. A common drain was used to couple the two stages, as the low impedance of this amplifier presented at the class-AB gate node reduces the input referred noise contribution of the floating battery bias branch, which can otherwise be significant. The output common mode voltage is regulated at midrail through a common mode feedback loop. Note that this sacrifices some control over the quiescent current. However, the devices are sized such that both the nominal first stage and second stage output voltages, as well as the nominal floating battery bias voltages, are set around mid-rail at the quiescent operating point.

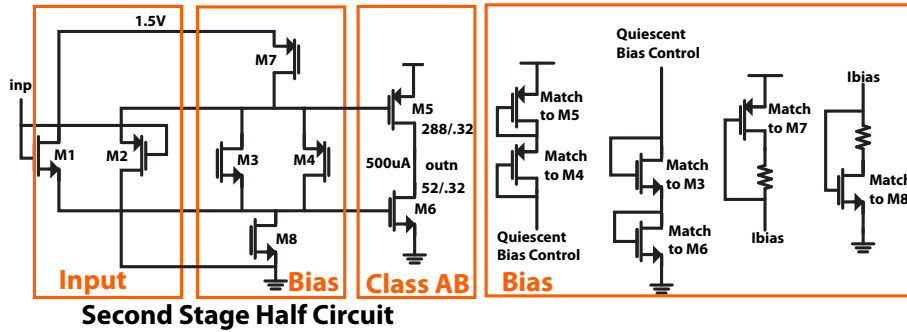


Figure 4.21: Baseband amplifier schematic.

Because the amplifier input is AC coupled on the RF side, the DC input common mode is regulated to mid-rail by the output common mode feedback loop through the feedback resistor. If the common mode loop is enabled, the DC operating point is stable, and small common mode variations within the loop bandwidth can be regulated. Additionally, due to the AC coupling cap, high frequency common mode glitches on the output can be handled, as the low antenna impedance at the amplifier input at high frequency reduces the loop gain. However, large low frequency common mode glitches can cause latching of the amplifier, due to the complementary nature. If the output common mode reaches a high voltage, say at startup, the input common mode will also settle to that voltage. This shuts off the PMOS portion of the complementary input stage, disabling the common mode feedback loop which acts through the PMOS tail current. The NMOS portion of the output stage is then disabled and the amplifier is latched. Current is slewed through the PMOS output device through the second stage shunt feedback resistor into the first stage current source, as shown in Fig. 4.22.

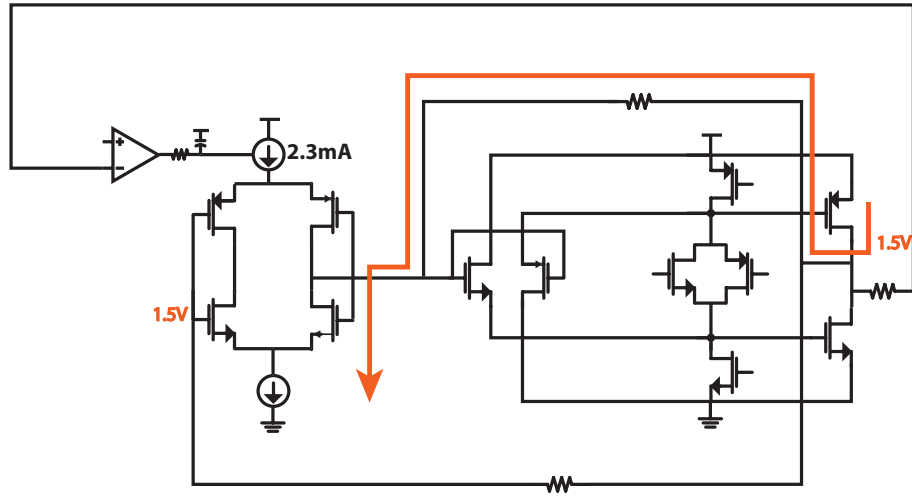


Figure 4.22: Latchup path for baseband amplifier.

In order to combat this startup issue, common mode diodes, as shown in Fig. 4.24 are placed at the output, which activate if the output common mode reaches above the diode's threshold. The diode pulldown current versus the common mode voltage is shown in Fig. 4.23 - the diodes begin to activate around a common mode voltage of 1V.

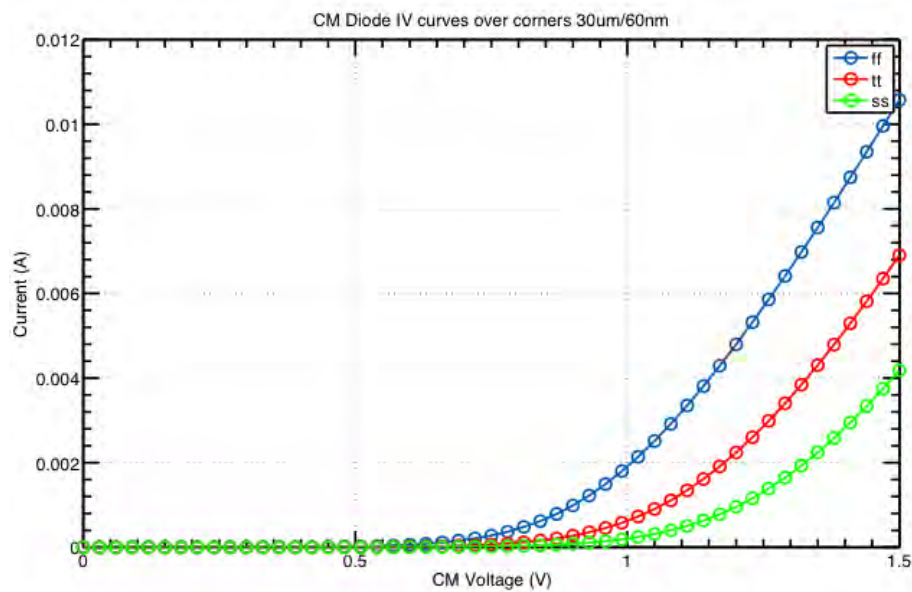


Figure 4.23: Common mode diode IV curves.

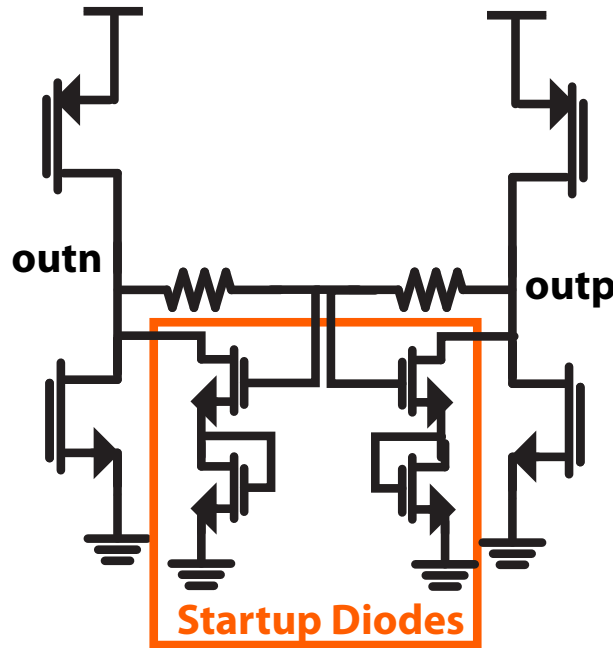


Figure 4.24: Common mode pulldown diodes.

To check the diode efficacy, a startup test can be used where the amplifier output voltage is statically pulled to $1.5V$ (V_{dd}) through low impedance switches, which are then disabled. If the diodes are effective, latchup condition is escaped as the common mode loop is able to regulate the output common mode to the desired mid-rail point. This transient startup test is shown with a strong PMOS bias mismatch in the amplifier (which can accentuate the latching condition) across corners, in Fig. 4.27.

An explicit RC narrow-banding pole at the common mode amplifier's output compensates the local common mode feedback loop. As the common feedback loop passes through 3 stages of amplification, the common mode feedback amplifier is design for just a gain of 2, to maintain stability. The loop has a total loop gain of 28.8dB and 60 degrees of phase margin at 74MHz, narrowband by an explicit 70k resistor and 200fF capacitance, shown in Fig. 4.26.

Note that as the amplifier is fully differential and has 2 inverting stages, the global common mode feedback loop due to the feedback resistors around the entire amplifier must be positive. Accordingly, in order to maintain stability, the local negative common mode feedback loop must be strong enough and have sufficient bandwidth to overcome the global positive loop. The total common mode loop gain is shown in Fig. 4.27. At low frequency, the strong local CMFB's loop surpasses the global loop, resulting in a sufficiently low loop gain. Outside of the CMFB loop bandwidth, the common mode gain begins to rise closer to 0dB. The mixer's drain capacitance and the large bottom plate capacitance of the baseband

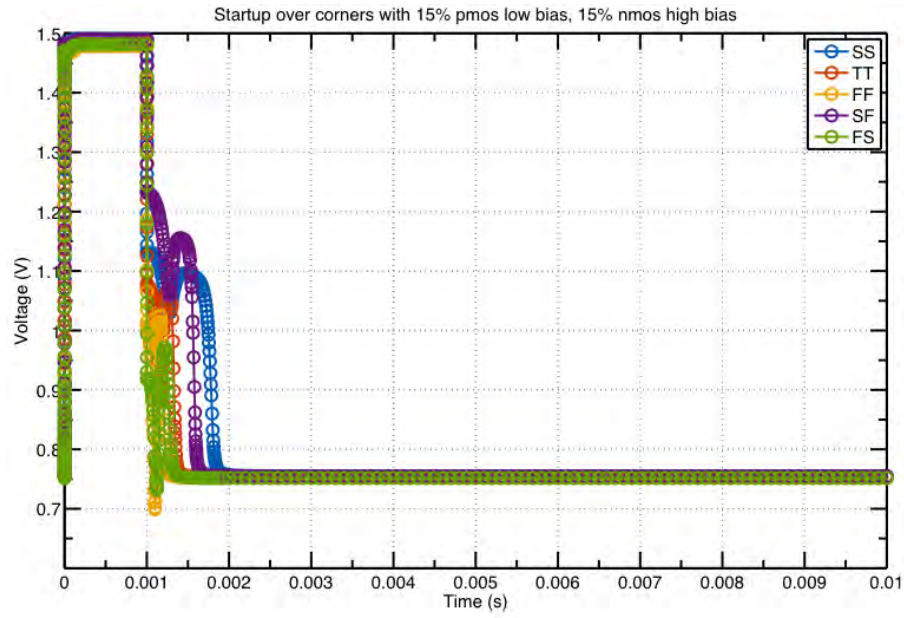


Figure 4.25: Baseband amplifier transient startup curves.

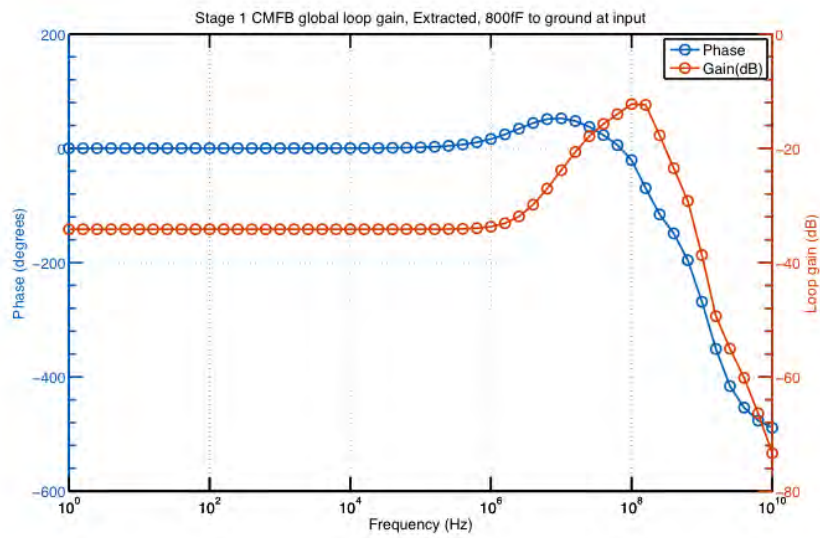


Figure 4.26: Local common mode feedback loop gain.

filtering cap at the amplifier input are used to stabilize the global feedback loop by shorting the amplifier input to ground at high frequency. To simulate the worst case stability, the lowest feedback resistor is used, setting the largest high frequency feedback factor, and a 4.4% (800fF) bottom plate capacitance as extracted from layout is used - around 12dB of gain margin is maintained. Note that if this bottom-plate capacitance is not enough, some of the baseband capacitance can be placed directly to ground, instead of differentially across the amplifier's input, setting the same differential input capacitance while increasing the common mode capacitance for common-mode stability.

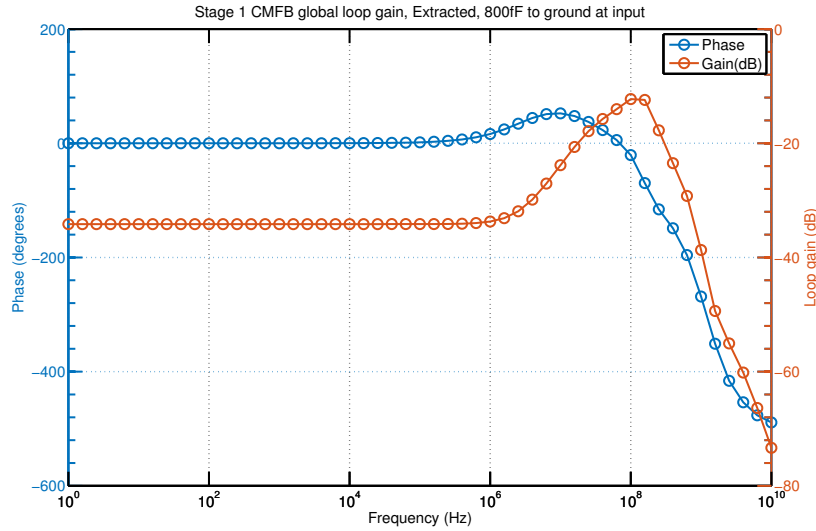


Figure 4.27: Global common mode loop gain.

For differential stability, when the mixer switch is enabled, the low antenna impedance relative to the feedback resistor effectively breaks the feedback loop by setting a very low loop gain. The critical case is when the mixer switch for a baseband amplifier is opened, putting the amplifier in a unity gain feedback configuration. However, no explicit differential compensation cap is needed. Recall that the baseband filtering capacitance sets a 20MHz corner against at the 50Ohm antenna impedance. Accordingly, when the switch is open, the node's resistance to ground increases and a very low frequency pole is set, ensuring a phase margin of close to 90°.

The resistors are implemented as 3 bit DACs, for measurement tunability. The first stage amplifier gain is shown for the nominal setting, and as stated below, is designed for a gain of 22.3dB, and has a 3dB bandwidth of around 320MHz, shown in Fig. 4.29. This bandwidth was not explicitly designed, but rather is a byproduct of the large input pair transconductance needed for noise purposes.

The baseband filtering capacitors in the passive-mixer-first receiver can either be placed at the amplifier input to ground, as in in Fig. 4.8, or as a Miller capacitance across the

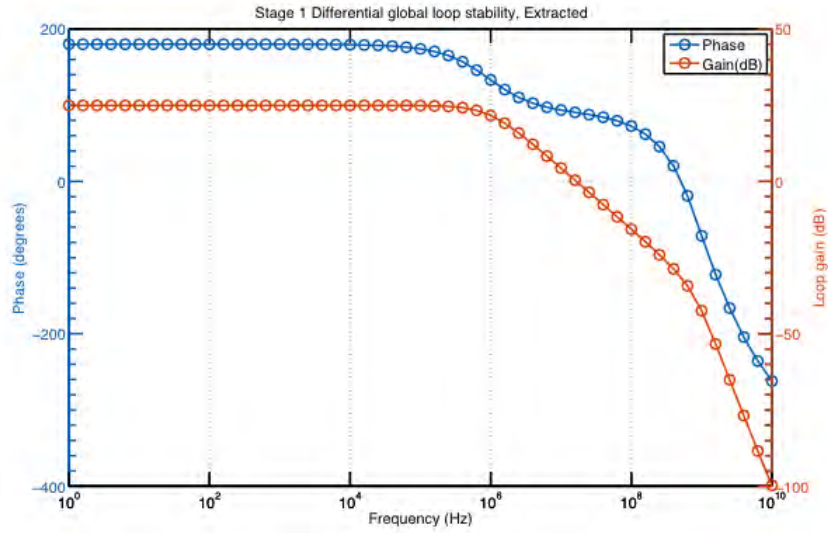


Figure 4.28: Differential mode loop gain.

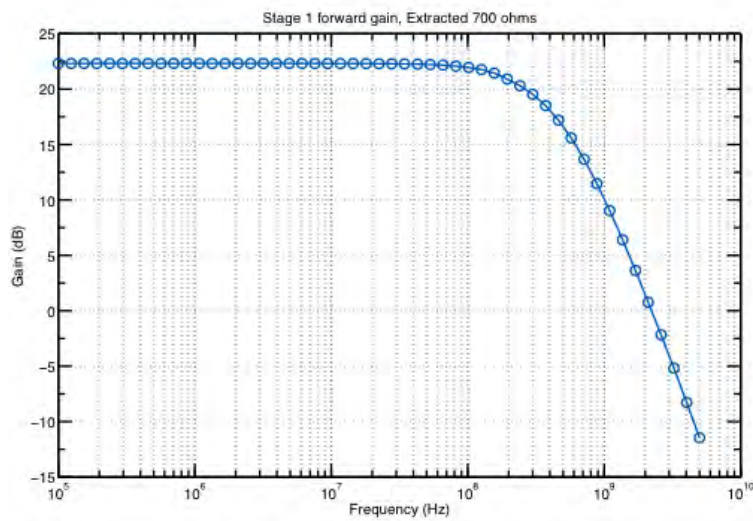


Figure 4.29: Baseband amplifier forward gain.

amplifier. The tradeoff is in the capacitor's area versus power consumption. If a Miller capacitance is used, while the voltage of out-of-band blockers is attenuated at the amplifier's output, the blocker's current path flows through the baseband amplifier's output stage. If a capacitance to ground at the amplifier input is used instead, the amplifier does not have to sink this current. Accordingly, for this work, 20pF differential capacitance is used at the amplifier's input to provide the filtering.

4.4 RX Second Stage Biquad

To improve receiver linearity, the harmonic recombination amplifiers were wrapped in a multifeedback (MFB) biquad as Fig. 4.30, to provide further out-of-band blocker filtering. This is easily accomplished, as inside the filter bandwidth, the MFB biquad acts as an inverting amplifier. The feed-forward resistor of the biquad can accordingly be split into 3 paths, scaled by the values needed for harmonic rejection [72][73], such that the separate phase signals sum in current through the filter. As the Norton equivalent resistance of each path is the same, the three paths see the same filter transfer function, simply with scaled DC gains.

The biquad design is an over specified problem, as it must meet 3 filter specifications of DC gain, cutoff frequency, and Q factor, with 5 passive components. This allows some design flexibility in the selection of the filter passives to trade area and power against noise.

The feed-forward resistor noise inside the signal bandwidth is simply added to the stage 1 amplifier noise, independent of the choice of the other passive parameter values. Accordingly a lower resistance value will have lower noise, but higher power and area consumption. Higher DC current must be slewed for the same voltage swing, and larger capacitors must be taken to fix the same cutoff frequency.

In order to contribute less than 5% additional noise power to the $1nV/\sqrt{Hz}$ first stage amplifier input referred noise, the resistor must be sized at

$$R_3 = \frac{.05 \times (1nV/\sqrt{Hz})^2 \times A_{DC,stage1}}{4kT} \quad (4.25)$$

Given the above, this resistor is chosen as 350Ω , to enable clean resistor values in the harmonic rejection ratio. Given a choice of this resistor, the feedback resistor is then fixed to achieve the desired DC gain.

$$R_1 = \frac{R_3(1 + A_{amp2})}{(\frac{A_{amp2}}{A_{DC}} - 1)} \quad (4.26)$$

The resistor R_2 essentially separates the input node, to allow the creation of 2 poles. This resistor has input referred noise which decreases monotonically with its resistance value. While this noise response does peak at the band edge, R_2 's in-band noise density can be approximated as:

$$v_{n,R_2,in}^2 = \left(\frac{1}{A_{DC}} \times \frac{A_{amp}}{1 + A_{amp}} \times \frac{R_2(R_1 + R_3)}{R_3} \right)^2 \times 4kTR_2 \quad (4.27)$$

However, a reduction in this resistance requires large (quadratically increasing) capacitors to maintain the same filter cutoff and quality factor - accordingly, the smallest resistor is taken given a maximum area budget. The two capacitor values are then fixed to set the desired frequency response.

$$C_2 = \frac{1}{2\pi f_0 \times Q\gamma\sqrt{R_1 R_2}} \quad (4.28)$$

$$C_1 = Q^2\gamma^2 C_2 \quad (4.29)$$

Using the above methodology, the harmonic rejection resistors are thus chosen as 350 ohms, and 500 ohms with a 700 ohm feedback resistor, approximating $\sqrt{2}$ with $\frac{700}{500}$.

The amplifier itself is the final portion of the filter design. The amplifier must maintain a minimum gain-bandwidth given the filter cutoff frequency and quality factor to reproduce the filter response accurately. In particular, a more aggressive Q filter, such as a Chebyshev, requires a larger gain-bandwidth product to achieve the target Q. Analysis of the sensitivity of filter cutoff and Q to the amplifier bandwidth can be found in [74].

For this work, the filter is chosen as a second order Butterworth, designed for $< 1dB$ attenuation at 10MHz, the edge of the passband. The amplifier is designed as a two stage op-amp with class AB second stage, similar to the first stage amplifier. A 300MHz gain bandwidth product amplifier design was sufficient to maintain filter performance within 1dB of the ideal response up to 40MHz. To avoid an additional output buffer stage in the receive chain, the amplifier is designed to directly drive the output pad, interfacing to an on-board, high-input impedance, 50 ohm output-impedance buffer. Accordingly, the amplifier's gain, bandwidth, and stability are verified for a 3pF load capacitance.

The simulated filter response is compared with the ideal in Fig. 4.31 provides an additional 12dB of attenuation of the TX signal 40MHz away.

4.5 Chip Implementation

The mixer-first RX in Fig. 4.32 is implemented in a redesign of the full transceiver system, again including the switched capacitor PA, cancellation current DAC, transformer network, and peripheral clock and data circuits. The cancellation DAC is redesigned with cascode devices at the top of each column, shielding the DAC output routing capacitance at the cascode's low impedance source node, and lowering the nominal RX matching network loss. Additionally, an integrated tunable LC trap is used to absorb third harmonic TX currents at the receiver input. More information on these is found in [52].

The RX consumes a total of 63.36mW from 1.5V and 1.2V supplies, as broken down in Fig. 4.33.

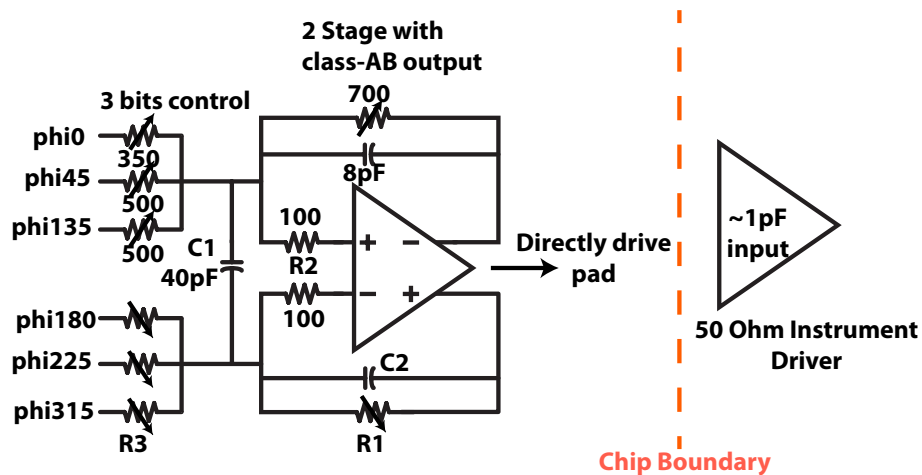


Figure 4.30: MFB/harmonic recombination amplifier.

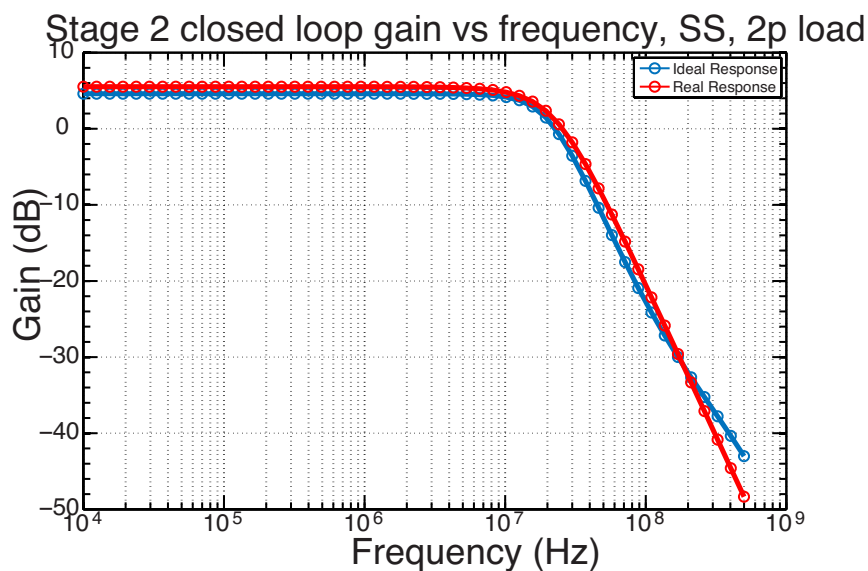


Figure 4.31: MFB filter response.

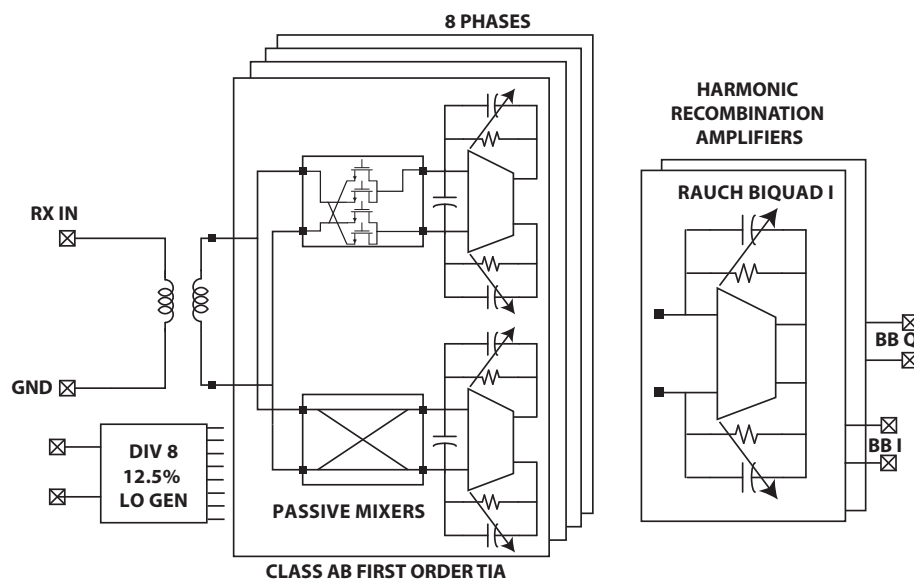


Figure 4.32: Receiver top level diagram.

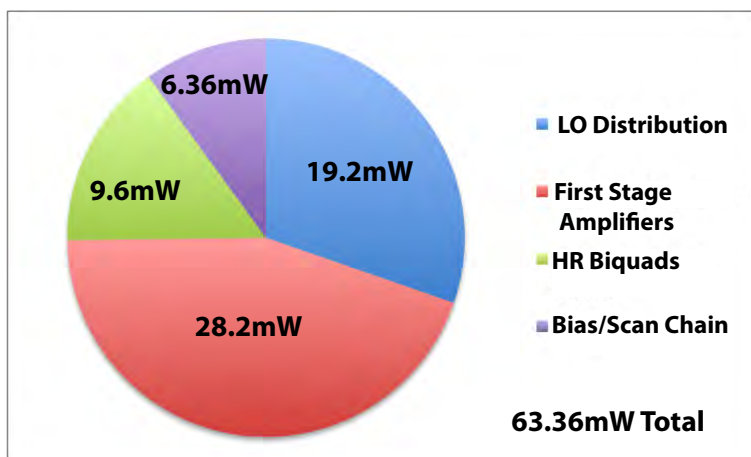


Figure 4.33: RX power breakdown.

The full layout occupies $550\mu m \times 600\mu m$ and is shown below in Fig. 4.34. The top level die photo is shown in Fig. 4.35.

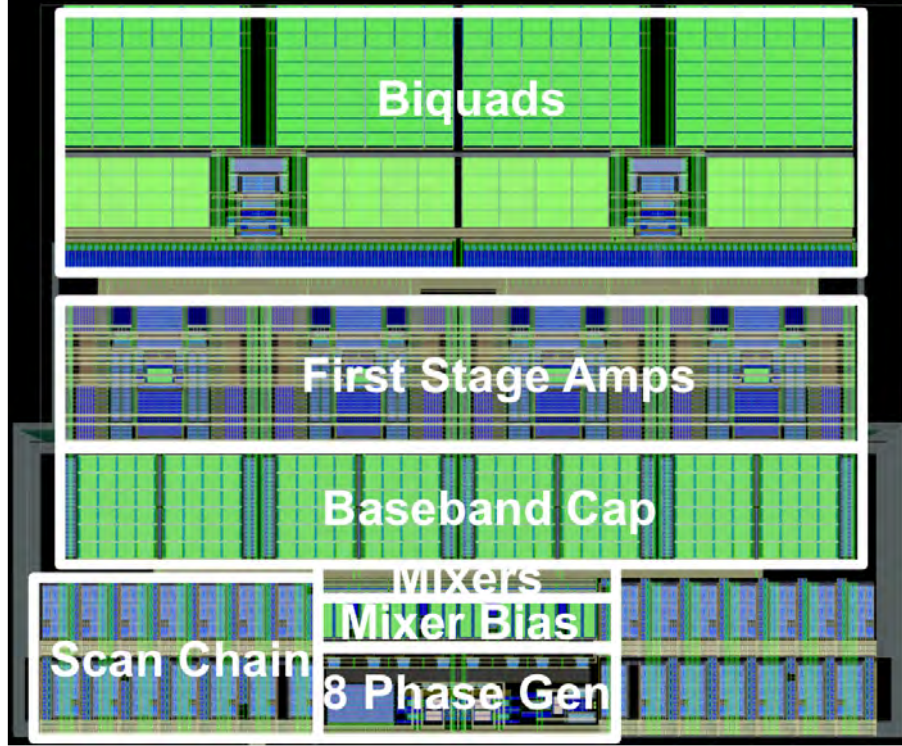


Figure 4.34: Receiver top level layout.

4.6 Measurements

The receiver S11 was first measured, to verify that the input match tracks the LO frequency. The input capacitive DAC must be adjusted at each input frequency to resonate with the RX transformer inductance. Tuning the capacitive DAC at fixed baseband resistor and gain settings, the receiver S11 exceeds -20dB from 1GHz to 2GHz, shown in Fig. 4.36. The S11 rolls off to -5dB away from the LO frequency, demonstrating the low pass to bandpass impedance up-conversion of the passive mixer receiver.

The effects of varying the feedback and feedforward resistors in the baseband to tune the real and imaginary parts of the output impedance can be seen in Fig. 4.37. Note that optimizing the feedback resistor changes the depth of the S11, while tuning the feedforward resistor shifts the center frequency of the match.

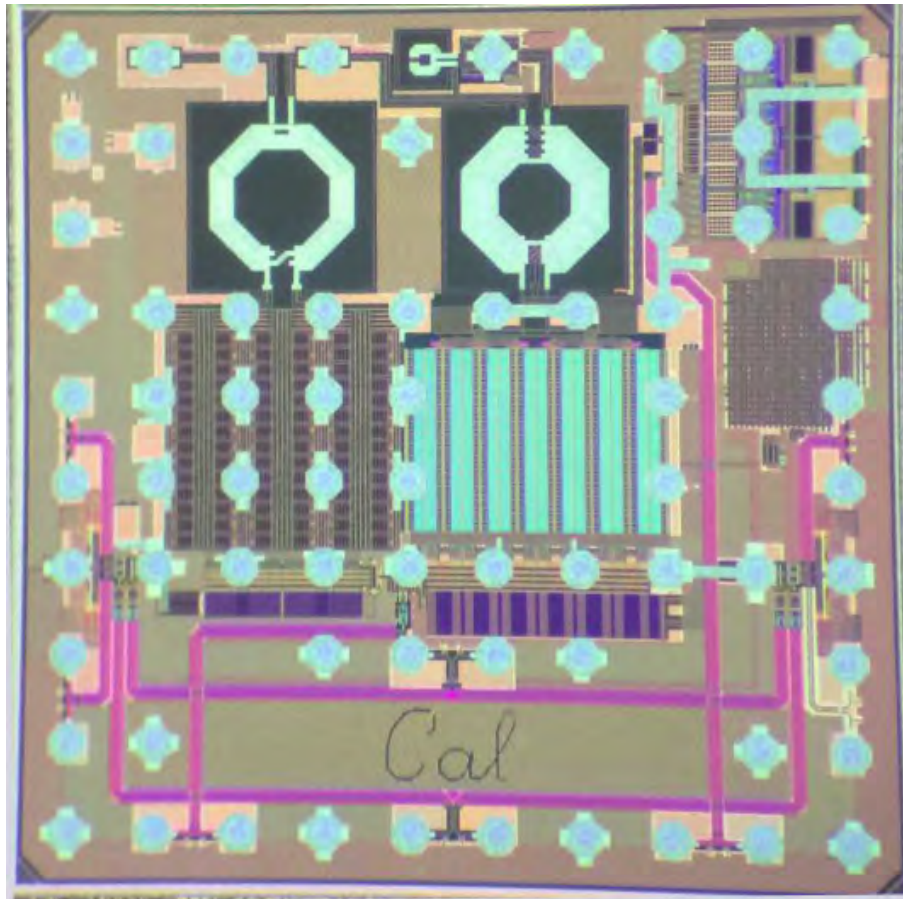


Figure 4.35: Die Photo.

The measured S_{21} is plotted in Fig. 4.38 against a simulation including the on-board PCB loss and on-chip transformer network. Including these losses, the receiver reaches an in-band gain of around 35dB.

To characterize the receiver IIP3, two test cases were used. In the first case, an in-band tone inter modulates with a nearby out-of-band tone, producing an in-band third order distortion tone. To test this, the power of a fixed in-band tone at 10MHz, and a stepped out of band tone from 11MHz to 29MHz were swept, and the in-band distortion product was measured versus power. In the second case, two out-of-band tones inter-modulate to produce an in-band tone. For this test case, the power of two out-of-band tones was swept, as the first tone moved at a rate X away from the band, and the second tone moved at a rate $2X$ away from the band, such that the intermodulation product fell in-band at a fixed frequency.

Once the power of the distortion is measured for various powers, the extrapolated intercept point can be calculated for a fixed input power as below. As the distortion scales

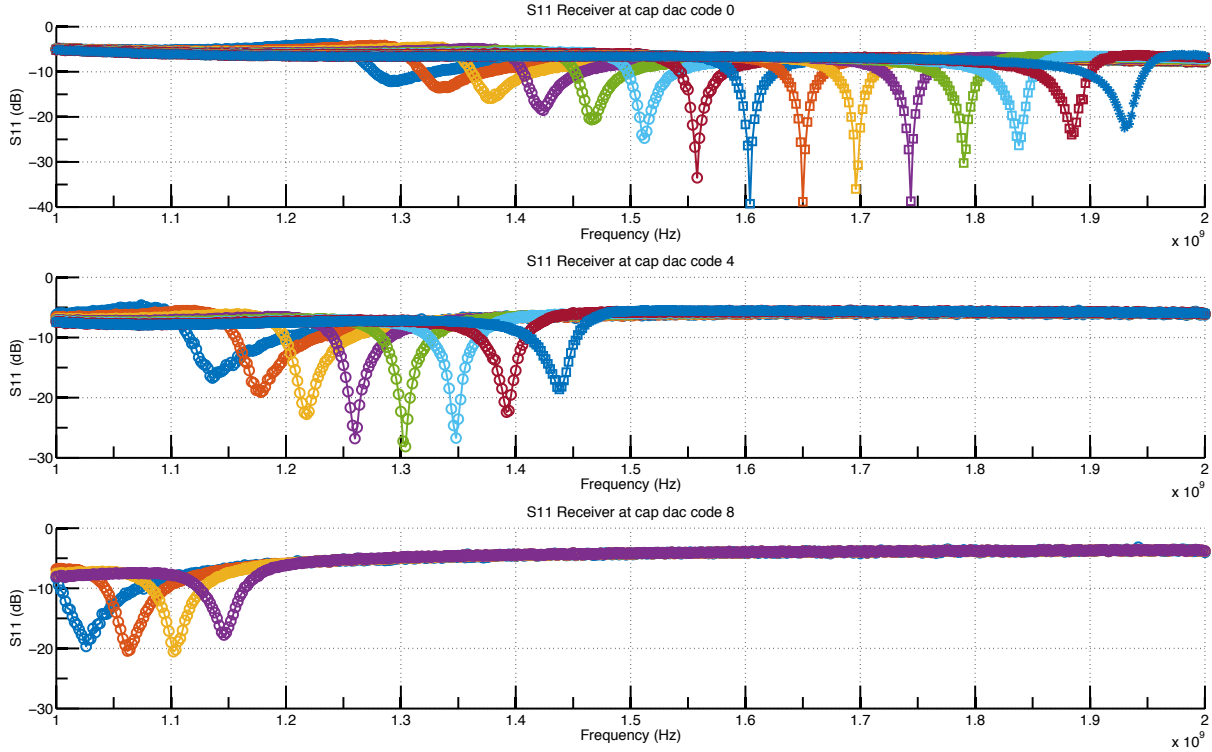


Figure 4.36: S11 vs. LO frequency for fixed matching network cap DAC setting.

3dB/dB of the input power, the gap between the distortion power and the input power shrinks by 2dB per dB of input power.

$$IIP_3 = P_{in} + \frac{P_{in} - (IM_3 + S_{21})}{2} \quad (4.30)$$

This value at each tone spacing was then averaged for several input powers, to produce the IIP3 vs tone spacing curve shown below.

The out-of-band IIP2 is similarly measured, by placing 2 out of band tones such that the second order distortion product falls in-band. In this case, the gap between the distortion power and the input power shrinks by a dB for a dB increase input power, so the IIP2 is computed as

$$IIP_2 = 2P_{in} - (IM_3 + S_{21}) \quad (4.31)$$

The receiver achieves an IIP3 of +25dBm out-of-band, and +1dBm in-band at the same fixed gain settings used for matching and noise measurements, due to the passive mixer providing out of band attenuation at the RF input, the 1.5V power supply, and the class AB baseband amplifier design. An IIP2 of 66dBm out-of-band is measured, due to independently

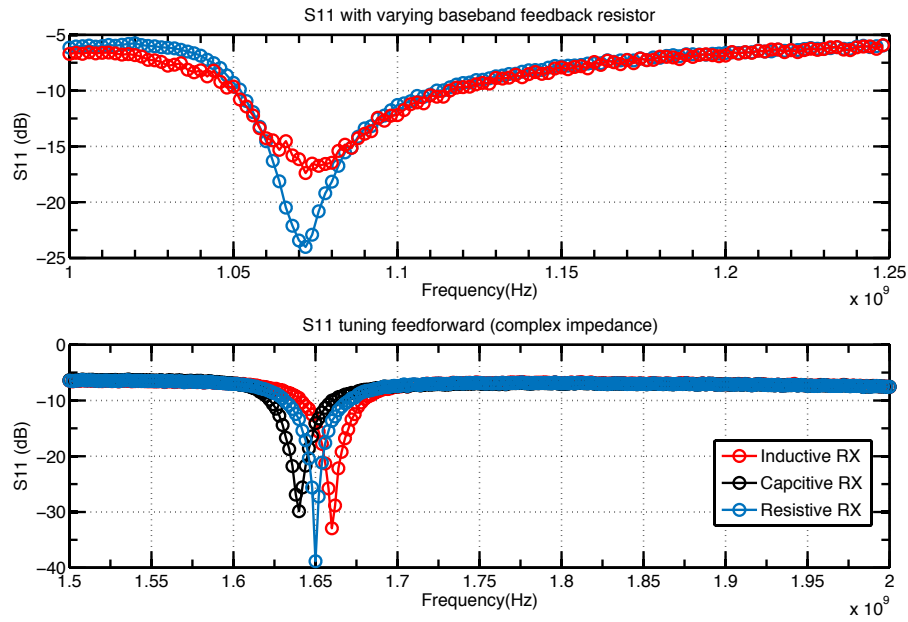


Figure 4.37: S11 vs. feedback and feedforward resistor settings.

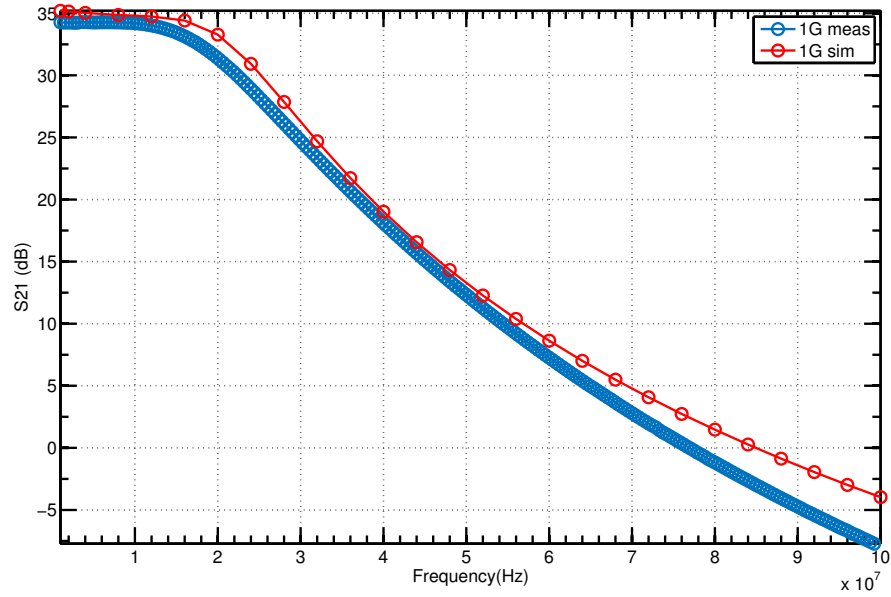


Figure 4.38: S21 vs. baseband frequency at 1GHz center frequency

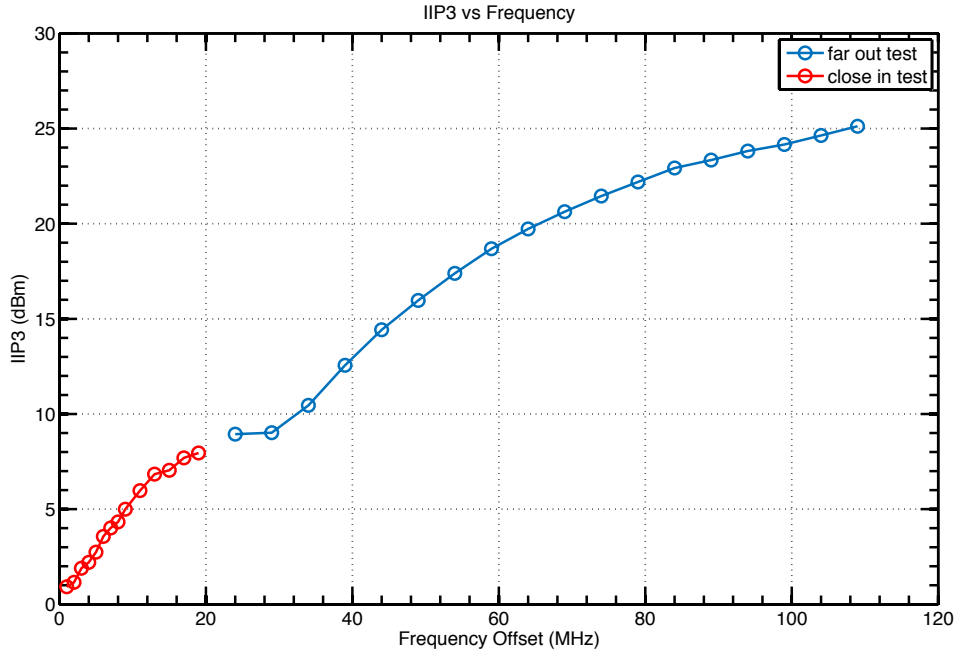


Figure 4.39: IIP3 as a function of tone spacing.

tunable bias voyages on the mixer devices for calibration, the balanced transformer at the input, and differential topology.

Care must be taken in appropriately measuring the noise figure, as the measurement is sensitive to the test setup. Loss in the input or output network can result in additional noise figure degradation. The noise figure is measured here by measuring the in-band S21 from the chip input port to its output, after de-embedding loss from the input SMA cable and PCB input routing. The output referred noise is then scaled by this in-band S21, and compared with 50 ohms to find the noise figure. Additionally, measurement of DSB noise figure can only occur after combining the I and Q paths, as the noise is correlated between the two.

The off chip 50ohm output buffers contribute an $2 \frac{nV}{\sqrt{Hz}}$ input referred noise within the RX baseband bandwidth. When input referred by a 34dB S21 of the receiver, this produces $0.04 \frac{nV}{\sqrt{Hz}}$ at the antenna, about 25 times lower than the noise of the 50 ohms source. For a 0dB noise figure receiver, this would contribute a penalty of less than .1dB, well within acceptable measurement tolerance of the setup.

The noise figure is measured as 6dB, of which approximately 3.8dB comes from the receive chain, and an additional 2.2dB comes from the front end transformer network, and capacitive DAC loss. The simulated receive chain noise figure under 3dB, so an additional 0.8dB of performance could potentially be achieved with cleaner measurement setup.

To demonstrate the efficacy of the 8 phase mixer and harmonic rejection biquads in

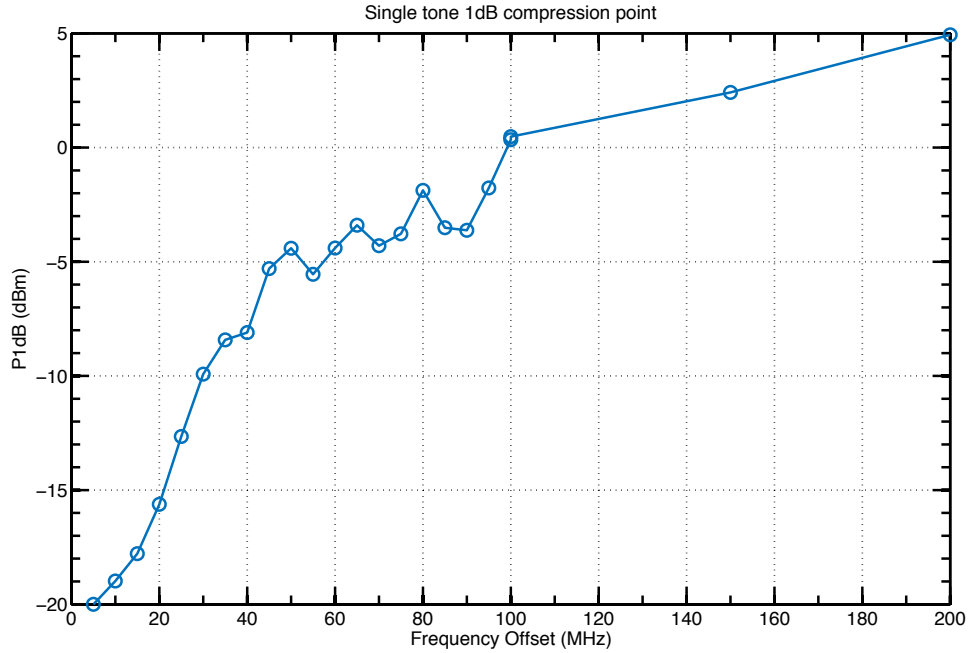


Figure 4.40: Receiver 1dB compression vs frequency offset.

rejecting harmonic blockers, tones were injected at the 3rd and 5th harmonic. These tones were down-converted inside the baseband filter bandwidth, with the harmonic recombination disabled and enabled. Fig. 4.41 demonstrates over 48dB rejection of the third harmonic, and over 53dB rejection of the fifth harmonic, limited by the 2 bit tuning resolution of the biquad resistor DACs. Note that this result does include the filtering of the transformer matching network at the third and fifth harmonic, but is de-embedded to the antenna port.

The RX performance is summarized in Fig. 4.42, with respect to other highly linear receivers [75][42][27][76][77][78][79][80][81][43][82].

When plugged into the active cancellation transceiver, the receiver's high linearity enables cancellation at a 6dB higher TX output power than a reference receiver design of Chapter 3 before compression. Shown in Fig. 4.43, the transceiver can operate up to roughly +4dBm TX output power with no TX/RX spacing, +16dBm TX output power at 40MHz duplex spacing, and +17dBm power at 80MHz, due to RX compression from un-cancelled TX harmonics. This is improved from +10.3dBm 1dB compression with the receiver design of Chapter 3.

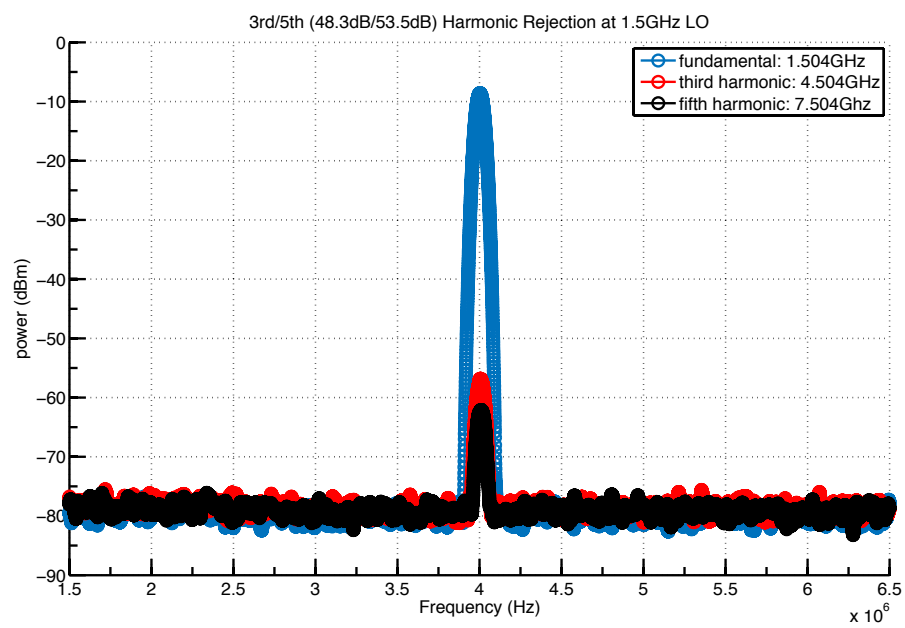


Figure 4.41: Receiver third and fifth harmonic rejection.

	Andrews 2010 ISSCC	Borremans 2011 ISSCC	Murphy 2012 ISSCC	Borremans 2013 VLSI	Murphy 2014 ISSCC	Tohidian 2014 ISSCC	Razavi 2014 ISSCC	Wu 2015 RFIC	This Work
Tech	65nm	40nm	40nm	28nm	28nm	65nm	65nm	28nm	65nm
RF Input	SE	Diff	SE	Diff	SE	Diff	Diff	SE	SE
RF Freq [GHz]	.1-2.4	.4-6	.08-2.7	.4-3	.6-3	1.8-2.5	.05-2.5	.4-3.5	1-2
BW [Hz]	2M	100K	-	-	-	-	.35- 20M	15- 50M	20M
NF [dB]	5	3	1.9	2.3-2.9	1.8 (3)	3.2-4.5	2.9	2.4- 2.6	3.8 (6*)
OOB P1dB [dBm]	+4	-8	-2	-	-6	-	-	+4.6	+5
IB P1dB [dBm]	N/A	-	-	-	-	-	-	-14.7	-20
Supply	1.2-2.5	1.1/2.5	2.8	.9	1	1.2-2	1.2	1/1.5	1.5
IB-IIP3 [dBm]	-	-	0	-	-	-7	-	+6.7	+1
OOB- IIP3 [dBm]	25	10	-	3	10	-	10	20.5	25
OOB- IIP2 [dBm]	58	70	54	85	49	85	-	64	66
HR (3rd/5th)	35/42	-	42/45	70/55	52/54	-	-	47/51	48/53
Power [mW]	37-70	30-55	35-78	40	39-70	55-65	20	38-75	64
Active Area [mm²]	2	2	1.2	.6	5	1.1	.82	.23	.33

* Including cancellation network loss

Figure 4.42: RX comparison table.

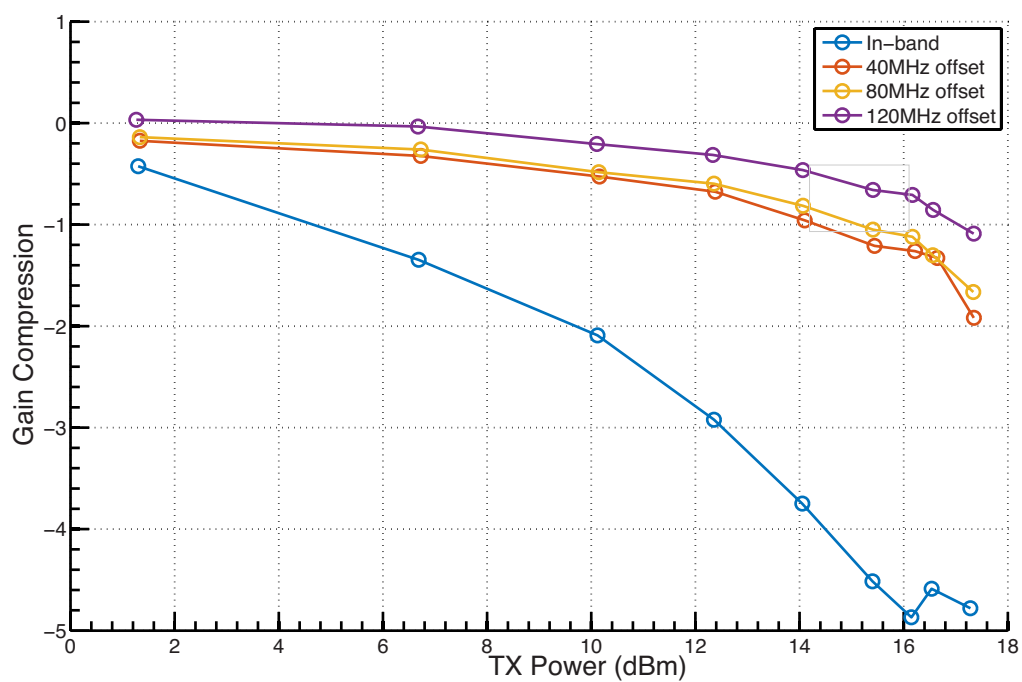


Figure 4.43: RX compression vs. TX power.

Chapter 5

Conclusion

This work addresses the problem of a frequency-agile transceiver with self-interference mitigation for FD/FDD applications. A transceiver system is demonstrated where a TX outputting up to +16dBm, shares the same antenna with an RX compressing by <1dB while simultaneously receiving. With some partial isolation, this could enable FDD standards, such as LTE, over a tunable frequency range, or serve as a transceiver for fully overlapped TX/RX frequency channels in new application scenarios. Through use of a mixed-signal digital-to-analog converter as the interference rejection element, a digital to analog adaptive loop is built to maximize interference rejection. This loop maintains the isolation over varying operating conditions, such as TX frequency, TX/RX leakage network, and antenna interface VSWR.

5.1 Thesis Contributions

This work represents a potential solution for a compact fully-integrated full-duplex transceiver. In particular, this work:

- Proposes a new TX/RX interface, where-in a transmitter and receiver can operate simultaneously on a single shared antenna.
- Identifies the Cartesian switched-capacitor power amplifier as an enabling transmitter for this interface. Proposes a simple alternative Cartesian combination scheme and shows the efficiency of the scheme vs. the polar counter-part.
- Identifies TX phase noise as a dominant desensitization mechanism in full-duplex transceivers, and quantifies the distinct impact of phase noise sources that are correlated and uncorrelated between the TX and cancellation sources.
- Proposes and analyzes LO sharing as a technique to mitigate the phase noise of the transmitter falling in the receive band. Constraints on the LO chain, and the band-

width limitation of the LO sharing technique for phase noise cancellation are shown through calculation and in silicon measurements.

- Demonstrates the system in a silicon prototype operating up to +16dBm TX output power, while maintaining over 50dB TX/TX isolation for 20MHz modulated TX signals. These numbers represent the highest achieved metrics for active cancellation systems reported to date.
- Demonstrates that digital nonlinear filtering of the cancellation input data can enable isolation over a range of TX operating conditions, including antenna VSWR, and TX center frequency. Cancellation is maintained over 5:1 VSWR over a 1GHz tuning range.
- Proposes the use of complementary class-AB TIAs within a passive mixer first receiver as a high linearity receiver for full-duplex systems. A receiver design implemented using this technique achieves state of the art +25dBm IIP3.

5.2 Future Work

Several future improvements are possible for this work:

- Further exploration and implementation of digital-back end to cancel the residual TX/DAC signal to the final receiver noise floor is the key remaining task to make self-interference cancellation possible. Preliminary work on this is documented in [52].
- In particular, building computationally tractable predictive models of TX and DAC nonlinearities would enable high resolution real time digital cancellation. Borrowing techniques from predistortion of digital power amplifiers [83][84] and DACs [85] may be useful.
- Quantization noise shaping on the cancellation DAC could be explored as another option to reduce RX-band quantization noise.
- On-chip/real-time implementation of the adaptive filter which selects the cancellation DAC data to maximize analog domain cancellation is a necessary practical step.
- Mitigation of higher DAC/TX harmonics through either passive RF filtering or more aggressive active harmonic rejection techniques could push the maximum TX power up from +16dBm.
- Noise cancellation techniques on both the DAC and TX side could enable lower RX desensitization. In particular, some preliminary noise cancellation results for the DAC are documented in [52].

- Exploration of architectures which exploit both some partial external isolation element, such as a circulator, along with this active cancellation technique could reduce the desensitization from the active cancellation network, and push to higher output powers. In particular, if a +15dBm TX signal was isolated by an off chip circulator with 25dB of isolation, this would result in -10dBm at the RX input. For a 20MHz bandwidth signal, another 90dB is needed to get to the thermal noise floor. This could be covered with the technique presented in this work, using a 13 bit DAC and oversampling the 20MHz signal to 1GHz, achieving a full high sensitivity in-band full duplex system.
- Designing a non class-A cancellation DAC could significantly save power consumption, especially in systems with high peak-to-average power transmit signals where the DAC would generally be backed off.
- Implementing the DAC as a mixed signal FIR filter, such as [86] to handle the frequency selectivity of the TX-RX coupling network could ease the requirements on the digital backend and the dynamic range requirements on the DAC itself.
- Investigation of self-interference-cancellation in MIMO systems is another interesting avenue, as there is likely to be heavy interest in massive MIMO techniques for 5G systems.

Bibliography

- [1] *Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2015-2020 White Paper*. web. 2016. URL: <http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/mobile-white-paper-c11-520862.html>.
- [2] M. McHenry et al. *Spectrum Occupancy Measurements*. web. November 2005.
- [3] *FCC AWS-3 Auction*. web. URL: http://wireless.fcc.gov/auctions/default.htm?job=auction_summary&id=97.
- [4] *LTE specifications*. URL: <http://www.3gpp.org/specifications>.
- [5] *Apple Iphone 6 technical spec*. web.
- [6] *The socio-economic benefits of greater spectrum policy harmonisation in the EU*. web. 2015. URL: <http://www.gsma.com/spectrum/wp-content/uploads/2015/12/Socioeconomic-benefits-of-harmonisation1.pdf>.
- [7] Rachid El Hattachi and Javan Erfanian. “NGMN 5G White Paper”. In: (2015). URL: https://www.ngmn.org/uploads/media/NGMN_5G_White_Paper_V1_0.pdf.
- [8] Steven Hong et al. “Applications of Self-Interference Cancellation in 5G and Beyond”. In: *IEEE Communications Magazine* 52.2 (2014), pp. 114–121.
- [9] P. Ancey. “Above 1c RF Mems and Baw Filters: Fact or Fiction?” In: *Bipolar/BiCMOS Circuits and Technology* (Oct 2006), pp. 1–5.
- [10] Sheng-Shian Li et al. “An Msi Micromechanical Differential Disk-Array Filter”. In: *Solid-State Sensors, Actuators and Microsystems Conference* (June 2007), 307311.
- [11] Mohyee Mikhemar, Hooman Darabi, and Asad Abidi. “A Tunable Integrated Duplexer with 50dB Isolation in 40nm CMOS”. In: *IEEE International Solid-State Circuits Conference* (2009).
- [12] Mohyee Mikhemar, Hooman Darabi, and Asad Abidi. “An On-Chip Wideband and Low-Loss Duplexer for 3G/4G CMOS Radios”. In: *IEEE Symposium on VLSI Circuits* (2010).
- [13] B. van Liempd et al. “A +70dBm IIP3 single-ended electrical-balance duplexer in 0.18um SOI CMOS”. In: *IEEE International Solid-State Circuits Conference* (2015).

- [14] Tong Zhang et al. “An integrated CMOS passive transmitter leakage suppression technique for FDD radios”. In: *Radio Frequency Integrated Circuits Symposium* (2014).
- [15] Sherif H. Abdelhalem, Prasad S. Gudem, and Lawrence E. Larson. “Hybrid Transformer-Based Tunable Differential Duplexer in a 90-nm CMOS Process”. In: *IEEE Transactions of Microwave Theory and Techniques* 61.3 (2013), pp. 1316–1326.
- [16] Moyhee Mikhemar. “Interference Cancellation in Software-Defined CMOS Receivers”. PhD thesis. University of California, Los Angeles, 2010.
- [17] Bharadia et al. “Full Duplex Radios”. In: *ACM SIGCOMM* (2013).
- [18] Y. S. Choi and H. Shirani-Mehr. “Simultaneous Transmission and Reception: Algorithm, Design and System Level Performance”. In: *IEEE Transactions on Wireless Communications* 12.12 (December 2013), pp. 5992–6010.
- [19] Melissa Duarte. “Full-duplex Wireless: Design, Implementation and Characterization”. PhD thesis. Rice University, 2012.
- [20] J. Zhou, P. R. Kinget, and H. Krishnaswamy. “A blocker-resilient wideband receiver with low-noise active two-point cancellation of >0dBm TX leakage and TX noise in RX band for FDD/Co-existence”. In: *IEEE International Solid-State Circuits Conference* (2014).
- [21] J. Zhou et al. “Receiver with >20MHz bandwidth self-interference cancellation suitable for FDD, co-existence and full-duplex applications”. In: *IEEE International Solid-State Circuits Conference* (2015).
- [22] J. Zhou, N. Reiskarimian, and H. Krishnaswamy. “Receiver with integrated magnetic-free N-path-filter-based non-reciprocal circulator and baseband self-interference cancellation for full-duplex wireless”. In: *IEEE International Solid-State Circuits Conference* (2016).
- [23] J Zhou et al. “Low-noise active cancellation of transmitter leakage and transmitter noise in broadband wireless receivers for FDD/co-existence”. In: *IEEE Journal of Solid-State Circuits* 49.12 (2014), pp. 3046–3062.
- [24] J Zhou et al. “Integrated wideband self-interference cancellation in the RF domain for FDD and full-duplex wireless”. In: *IEEE Journal of Solid-State Circuits* 50.12 (2015), pp. 3015–3031.
- [25] Negar Reiskarimian et al. “Analysis and design of two-port N-path band-pass filters with embedded phase shifting”. In: *IEEE Transactions on Circuits and Systems* 63.8 (2016), pp. 728 –732.
- [26] Federico Bruccoleri, Eric A. M. Klumperink, and Bram Nauta. “Wide-Band CMOS Low-Noise Amplifier Exploiting Thermal Noise Canceling”. In: *IEEE Journal of Solid-State Circuits* 39.2 (2004), pp. 275 –282.
- [27] David Murphy et al. “A blocker-tolerant wideband noise-cancelling receiver with a 2dB noise figure”. In: *IEEE International Solid-State Circuits Conference* (2012).

- [28] David Murphy. “Noise in Large-Signal, Time-Varying RF CMOS Circuits: Theory and Design”. PhD thesis. University of California, Los Angeles, 2012.
- [29] D. Murphy et al. “A Blocker-Tolerant, Noise-Cancelling Receiver Suitable for Wideband Wireless Applications”. In: *IEEE Journal of Solid-State Circuits* 45.12 (Dec. 2010), pp. 2696–2708.
- [30] Ashkan Borna. “Interference Management Techniques for Multi-Standard Wireless Receivers”. PhD thesis. University of California, Berkeley, 2012.
- [31] N. A. Estep, D. L. Sounas, and A. Alu. “Magnetless Microwave Circulators Based on Spatiotemporally Modulated Rings of Coupled Resonators”. In: *IEEE Transactions on Microwave Theory and Techniques* 64.2 (February 2016), pp. 502–518.
- [32] D. J. van den Broek, E. A. M. Klumperink, and B. Nauta. “A self-interference-cancelling receiver for in-band full-duplex wireless with low distortion under cancellation of strong TX leakage”. In: *IEEE International Solid-State Circuits Conference* (2015).
- [33] Dirk-Jan van den Broek, Eric AM Klumperink, and Bram Nauta. “An in-band full-duplex radio receiver with a passive vector modulator downmixer for self-interference cancellation”. In: *IEEE journal of solid-state circuits* 50.12 (2015), pp. 3003–3014.
- [34] D. Yang et al. “A Fully Integrated Software-Defined FDD Transceiver Tunable from .3-1.6GHz”. In: *IEEE RFIC Symposium* (2016).
- [35] Aminghasem Safarian et al. “Integrated Blocker Filtering RF Front Ends”. In: *IEEE Radio Frequency Integrated Circuits Symposium* (2007).
- [36] Hyungsoo Kim et al. “A CMOS Transmitter Leakage Canceller for WCDMA Applications”. In: *IEEE Journal of Solid-State Circuits* (2013).
- [37] Dong Yang and Alyosha Molnar. “A widely tunable active duplexing transceiver with same-channel concurrent RX/TX and 30dB RX/TX isolation”. In: *Radio Frequency Integrated Circuits Symposium* (2014).
- [38] S. A. Ayati et al. “Adaptive integrated CMOS circulator”. In: *IEEE RFIC Symposium* (2016).
- [39] Dong Yang and Alyosha Molnar. “A Wideband Highly Integrated and Widely Tunable Transceiver for In-Band Full-Duplex Communication”. In: *IEEE Journal of Solid-State Circuits* (2015).
- [40] Behzad Razavi. *RF Microelectronics*. Prentice Hall Communications Engineering, 2011.
- [41] T. Lee. *The Design of CMOS Radio-Frequency Integrated Circuits*. Cambridge Univ. Press, 2004.
- [42] C. Andrews and A. C. Molnar. “A Passive Mixer-First Receiver With Digitally Controlled and Widely Tunable RF Interface”. In: *IEEE Journal of Solid-State Circuits* 45.12 (Dec. 2010), pp. 2696–2708.

- [43] C. Wu et al. “An Interference-Resilient Wideband Mixer-First Receiver With LO Leakage Suppression and I/Q Correlated Orthogonal Calibration”. In: *IEEE Transactions on Microwave Theory and Techniques* 64.4 (April 2016), pp. 1088–1101.
- [44] Daisuke Imanishi, Kenichi Okada, and Akira Matsuzawa. “A 0.93.0 GHz fully integrated tunable CMOS power amplifier for multi-band transmitters”. In: *IEEE Asian Solid State Circuits Conference* (2009).
- [45] N. C. Kuo et al. “A frequency-reconfigurable multi-standard 65nm CMOS digital transmitter with LTCC interposers”. In: *IEEE Asian Solid-State Circuits Conference* (10-12 Nov. 2014).
- [46] B. Jin et al. “A 30.8-dBm Wideband CMOS Power Amplifier With Minimized Supply Fluctuation”. In: *IEEE Transactions on Microwave Theory and Techniques* 60.6 (June 2012), pp. 1658–1666.
- [47] Hongrui Wang and Hossein Hashemi. “A 0.5-6 GHz 25.6 dBm Fully Integrated Digital Power Amplifier in 65nm CMOS”. In: *IEEE Radio Frequency Integrated Circuits Symposium* (2014).
- [48] Z. Boos et al. “A fully digital multimode polar transmitter employing 17b RF DAC in 3G mode”. In: (2011), pp. 376–378. ISSN: 0193-6530. DOI: 10.1109/ISSCC.2011.5746361.
- [49] A. Apsel et al. “Challenges and approaches to software defined duplexing radio”. In: *IEEE International Symposium on Circuits and Systems* (2016).
- [50] B. Debaillie et al. “Analog/RF Solutions Enabling Compact Full-Duplex Radios”. In: *IEEE Journal on Selected Areas in Communications* 32.9 (Sept. 2014), pp. 1662–1673.
- [51] *Amazon Fire Phone Teardown*. web. 2014. URL: <https://www.ifixit.com/Teardown/Amazon+Fire+Phone+Teardown/27458>.
- [52] Lucas Calderin. PhD thesis. University of California, Berkeley, 2017.
- [53] E. Bechthum et al. “A 5.3GHz 16b 1.75GS/S wideband RF Mixing-DAC achieving $\text{IMD} < -82\text{dBc}$ up to 1.9GHz”. In: *IEEE International Solid-State Circuits Conference* (2015).
- [54] H. Van de Vel et al. “A 240mW 16b 3.2GS/s DAC in 65nm CMOS with $< -80\text{dBc}$ IM_3 up to 600MHz”. In: *IEEE International Solid-State Circuits Conference* (2014).
- [55] Asad Abidi. “Phase Noise and Jitter in CMOS Ring Oscillators”. In: *JSSC* 41.8 (2006), pp. 1803–1816.
- [56] Ali Hajimiri and Thomas Lee. “A General Theory of Phase Noise in Electrical Oscillators”. In: *JSSC* 33.2 (1998), pp. 179–194.
- [57] D. J. van den Broek, E. A. M. Klumperink, and B. Nauta. “A self-interference cancelling front-end for in-band full-duplex wireless and its phase noise performance”. In: *RFIC* (2015).

- [58] Sang-Min Yoo et al. "A Switched-Capacitor RF Power Amplifier". In: *JSSC* 46.12 (2011), pp. 2977–2987.
- [59] Lu Ye. "Design and Analysis of Digitally Modulated Transmitters for Efficiency Enhancement". PhD thesis. University of California, Berkeley, 2013.
- [60] Chao Lu et al. "A 24.7dBm All-Digital RF Transmitter for Multimode Broadband Applications in 40nm CMOS". In: *IEEE International Solid-State Circuits Conference* (2013).
- [61] V. Vorapipat, C. Levy, and P. Asbeck. "A wideband voltage mode Doherty power amplifier". In: *IEEE Radio Frequency Integrated Circuits Symposium* (2016).
- [62] Gang Liu et al. "Fully Integrated CMOS Power Amplifier With Efficiency Enhancement at Power Back-Off". In: *IEEE Journal of Solid State Circuits* 43.3 (2008), pp. 600–609.
- [63] Chenxi Huang et al. "A 40nm CMOS single-ended switch-capacitor harmonic-rejection power amplifier for ZigBee applications". In: *IEEE Radio Frequency Integrated Circuits Symposium* (2016).
- [64] H. Jin et al. "efficient digital quadrature transmitter based on IQ cell sharing". In: *IEEE International Solid-State Circuits Conference* (2015).
- [65] Z. Deng et al. "A dual-band digital-WiFi 802.11a/b/g/n transmitter SoC with digital I/Q combining and diamond profile mapping for compact die area and improved efficiency in 40nm CMOS". In: *IEEE International Solid-State Circuits Conference* (2016).
- [66] Debopriyo Chowdhury. "Efficient Transmitters for Wireless Communications in Nanoscale CMOS Technology". PhD thesis. University of California, Berkeley, 2010.
- [67] Vladimir Aparin et al. "An Integrated LMS Adaptive Filter of TX Leakage for CDMA Receiver Front Ends". In: *IEEE Journal of Solid-State Circuits* (2006).
- [68] Yen-Horng Chen et al. "An LTE SAW-Less Transmitter Using 33% Duty-Cycle LO Signals for Harmonic Suppression". In: *IEEE International Solid-State Circuits Conference* (2015).
- [69] Ao Ba et al. "A 2.4GHz Class-D Power Amplifier with Conduction Angle Calibration for -50dBc Harmonic Emissions". In: *IEEE Radio Frequency Integrated Circuits Symposium* (2016).
- [70] C. Andrews and A. C. Molnar. "Implications of Passive Mixer Transparency for Impedance Matching and Noise Figure in Passive Mixer-First Receivers". In: *IEEE Transactions of Circuits and Systems* 57.12 (Dec. 2010), pp. 3092–3103.
- [71] Charles Wu et al. "A passive-mixer-first receiver with LO leakage suppression, 2.6dB NF, >15dBm wide-band IIP3, 66dB IRR supporting non-contiguous carrier aggregation". In: *Radio Frequency Integrated Circuits Symposium* (2015).

- [72] Jeffrey A. Weldon et al. “A 1.75-GHz Highly Integrated Narrow-Band CMOS Transmitter With Harmonic-Rejection Mixers”. In: *IEEE Journal of Solid-State Circuits* 36.12 (2001), pp. 2003–2015.
- [73] Z. Ru et al. “A 300-800 MHz Tunable Filter and Linearized LNA Applied in a Low-Noise Harmonic-Rejection RF-Sampling Receiver”. In: *IEEE Journal of Solid-State Circuits* 45.5 (2010), pp. 967–978.
- [74] Michael Steffes. *Design Methodology for MFB Filters in ADC Interface Applications*. web. 2006. URL: <http://www.ti.com/lit/an/sboa114/sboa114.pdf>.
- [75] Caroline Andrews and Alyosha Molnar. “A Passive-Mixer-First Receiver with Baseband Controlled RF Impedance Matching, < 6dB NF, and > 27dBm Wideband IIP3”. In: *IEEE International Solid-State Circuits Conference* (2010).
- [76] Jonathan Borremans et al. “A 40nm CMOS highly linear 0.4-to-6GHz receiver resilient to 0dBm out-of-band blockers”. In: *IEEE International Solid-State Circuits Conference* (2011).
- [77] Jonathan Borremans et al. “A 0.9 V low-power 0.46 GHz linear SDR receiver in 28 nm CMOS”. In: *IEEE Symposium on VLSI Circuits* (2014).
- [78] Hao Xu, David Murphy, and Hooman Darabi. “A noise-cancelling receiver with enhanced resilience to harmonic blockers”. In: *IEEE International Solid-State Circuits Conference* (2014).
- [79] Massoud Tohidian, Iman Madadi, and Robert Bogdan Staszewski. “A Fully Integrated Highly Reconfigurable Discrete Time Super-Heterodyne Receiver”. In: *IEEE International Solid-State Circuits Conference* (2014).
- [80] Behzad Razavi Joung Won Park. “A 20mW GSM/WCDMA Receiver with RF Channel Selection”. In: *IEEE International Solid-State Circuits Conference* (2014).
- [81] Behzad Razavi Joung Won Park. “Channel Selection at RF Using Miller Bandpass Filters”. In: *IEEE Journal of Solid-State Circuits* 49.2 (2014), pp. 3063–3078.
- [82] Hugo Westerveld, Eric Klumperink, and Bram Nauta. “A Cross-Coupled Switch-RC Mixer-First Technique achieving +41dBm Out-of-Band IIP3”. In: *IEEE Radio Frequency Integrated Circuits Symposium* (2016).
- [83] Lei Guan and Anding Zhu. “Low-Cost FPGA Implementation of Volterra Series-Based Digital Predistorter for RF Power Amplifiers”. In: *IEEE Transactions of Microwave Theory and Techniques* 58.4 (2010), pp. 866–872.
- [84] Anding Zhu, John Dooley, and Thomas J. Brazil. “Simplified Volterra Series Based Behavioral Modeling of RF Power Amplifiers Using Deviation-Reduction”. In: *IEEE Transactions of Microwave Theory and Techniques* (June 2006).
- [85] Amir Amirkhany et al. “Time-Variant Characterization and Compensation of Wideband Circuits”. In: *IEEE Custom Integrated Circuits Conference* (2007).

- [86] David B. Barkin et al. “A CMOS Oversampling Bandpass Cascaded D/A Converter With Digital FIR and Current-Mode Semi-Digital Filtering”. In: *IEEE Journal of Solid State Circuits* 39.4 (2004), pp. 585–593.
- [87] P.G. Gray et al. *Analysis and Design of Analog Integrated Circuit*. 5th ed. Wiley, Jan 2009.
- [88] S. Ramakrishnan et al. “A 65nm CMOS transceiver with integrated active cancellation supporting FDD from 1GHz to 1.8GHz at +12.6dBm TX power leakage”. In: *IEEE Symposium on VLSI Circuits* (2016).