

Flexible Integrated Architectures for Frequency Division Duplex Communication

Lucas Calderin



Electrical Engineering and Computer Sciences
University of California at Berkeley

Technical Report No. UCB/EECS-2017-55

<http://www2.eecs.berkeley.edu/Pubs/TechRpts/2017/EECS-2017-55.html>

May 11, 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.

**Flexible Integrated Architectures for Frequency Division Duplex
Communication**

by

Lucas Albert Calderin

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

in

Engineering - Electrical Engineering and Computer Sciences

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Ali M. Niknejad, Chair
Professor Borivoje Nikolić
Professor Martin White

Spring 2017

**Flexible Integrated Architectures for Frequency Division Duplex
Communication**

Copyright 2017
by
Lucas Albert Calderin

Abstract

Flexible Integrated Architectures for Frequency Division Duplex Communication

by

Lucas Albert Calderin

Doctor of Philosophy in Engineering - Electrical Engineering and Computer Sciences

University of California, Berkeley

Professor Ali M. Niknejad, Chair

The explosion in demand for wireless capacity has been a strong driver for cellular network modification. Previous solutions to this problem have involved an increase in available channel spectrum to each user and an increase in bands used for cellular communication. This practice has led to an extremely crowded low-frequency spectrum, where the significant problems now are not processing high bandwidths in a mobile device, but instead managing the interference caused by high-traffic scenarios.

One of the most challenging instantiations of this interference mitigation problem is Frequency Division Duplex (FDD) communication in the Long Term Evolution (LTE) standard, where a cellular device both transmits and receives information simultaneously, but in separate frequency bands. Globally, there are 40 LTE FDD bands, and in the worst cases, the center-to-center spacing of transmit (TX) and receive (RX) bands is only twice the bandwidth. With 120dB of dynamic range between the TX and RX signals, without significant isolation, the RX is heavily desensitized or even damaged by the high power TX signal. This problem is currently solved with fixed external duplexers, which provide high TX/RX isolation, but at the cost of frequency tunability. Due to the large number of TX/RX frequency band pairings present in the LTE standard globally, using fixed filters for isolation is infeasible if a phone is to operate well internationally.

In this thesis, a fully-integrated, highly frequency-flexible method for TX self-interference cancellation is proposed, where a mixed-signal canceller is employed at the input of the RX. This technique is shown to allow the receiver to tolerate high TX power levels over a large variety of channel and transceiver nonidealities. The deterministic TX-band interference signal is shown to be mitigated within this system, along with the non-deterministic sources of interference from TX phase noise and canceller thermal noise. Finally, further reduction in TX interference in the digital backend using digital modelling of the PA, canceller, and duplex network is shown.

To Button

Contents

Contents	ii
List of Figures	iv
List of Tables	viii
1 Introduction	1
1.0.1 Prior Art	7
1.0.2 General Interference Cancellation	7
1.0.3 Self-Interference Cancellation	9
2 Theoretical Framework	12
2.1 Conceptual Overview	12
2.2 System Advantages to a Current DAC	16
2.3 DAC Power Consumption	17
2.4 Noise Sources	18
2.4.1 Noise Overview	18
2.4.2 TX Thermal Noise	19
2.4.3 TX Phase Noise	21
2.4.4 DAC Thermal Noise	22
2.4.5 Noise Summary	28
2.5 TX Insertion Loss From RX	29
2.6 TX Efficiency Degradation Mechanisms	32
2.7 DAC Linearity	33
2.7.1 IQ Nonlinearity - Linear Adaptation	37
2.8 I/Q Nonlinearity Due to Buffer Chain	38
2.9 Residual Sampling Methods	40
2.9.1 Synchronous Sampling	41
2.9.2 Asynchronous Sampling	42
2.9.3 Asynchronous Sampling Rejection	46
2.10 Digital Backend Cancellation	49

3	A CMOS Transceiver with Integrated FDD Support Up to +12.6dBm TX Leakage	50
3.1	Chip Implementation	50
3.1.1	Transmitter	51
3.1.2	TX Efficiency Degradation	54
3.1.3	Receiver	55
3.2	DAC Design	56
3.3	DAC Thermal Noise Cancellation	59
3.4	TX Efficiency Degradation Mechanisms	63
3.5	Measurement Results	65
3.5.1	Isolated Measurements	65
3.5.2	System Measurements	68
3.5.3	DAC Thermal Noise Cancellation	74
3.6	Digital Cancellation Measurements	76
4	A Transceiver with >64dB TX Signal Cancellation and Thermal/Phase Noise Rejection	87
4.1	Overview	87
4.2	Passive Mixer First RX	87
4.3	TX/RX Passive Network	89
4.4	RX Capacitor DAC Design	93
4.5	Column Cascodes	98
4.5.1	Column Cascode Nonlinearity	99
4.5.2	Noise	101
4.6	Measurement Results	103
4.6.1	Phase Noise	105
4.6.2	Noise With Cancellation	106
4.6.3	RX Compression With Cancellation	107
4.6.4	Comparison With Prior Art	109
5	Conclusion	110
5.1	Thesis Summary	110
5.2	Future Directions	111
	Bibliography	112

List of Figures

1.1	Projected monthly global traffic for mobile data [1].	1
1.2	Comparison of technology use 8 years apart.	2
1.3	LTE frame with resource blocks highlighted.	3
1.4	LTE global frequency allocations.	4
1.5	FDD vs. TDD.	4
1.6	Dynamic range between TX power and RX sensitivity.	5
1.7	SAW and BAW mechanisms [10].	5
1.8	Carrier aggregation: combining multiple bands for higher datarates.	6
1.9	Amazon Fire phone multi-standard transceivers.	7
1.10	Active series LR N-Path for programmable notch.	8
1.11	Feedback technique applied to TX [28].	9
1.12	Hybrid with balancing impedance [34].	10
1.13	Active cancellation methods.	10
2.1	Naive TX and RX single antenna interface.	13
2.2	Impedance modification through current source.	13
2.3	Top level conceptual diagram of cancellation architecture.	14
2.4	Breakdown of conceptual diagram.	15
2.5	Advantages of mixed signal cancellation.	15
2.6	Cancellation signal domains.	16
2.7	Generation of replica in voltage and current domains.	17
2.8	Power consumption of canceller versus TX.	18
2.9	Significant transceiver noise contributions.	20
2.10	Noise figure due to TX only.	21
2.11	Feedforward cancellation of TX/DAC shared phase noise.	22
2.12	Phase interpolator phase noise.	22
2.13	DAC noise network.	23
2.14	Unit cell output states.	25
2.15	DAC foise figure versus TX power.	26
2.16	Binary and thermometer output current and associated noise.	27
2.17	Thermal noise contours vs. TX power contours.	27
2.18	Noise figure with nonidealities added.	28

2.19	Circuit with mutually coupled inductors.	30
2.20	RX transformer with input port shorted.	30
2.21	Simplified transformer model [49].	31
2.22	TX transformer performance degradation with RX inclusion.	32
2.23	Definitions of DAC static mismatch.	34
2.24	2 and 5 bit binary constellations with 10% I/Q summation magnitude nonlinearity.	35
2.25	2 and 5 bit thermometer constellations with 10% I/Q summation magnitude nonlinearity.	35
2.26	Rejection vs. DAC bits for I/Q summation nonlinearity.	36
2.27	Rejection vs. DAC segmentation for I/Q summation nonlinearity.	37
2.28	Buffer I/Q summation nonlinearity due to nonzero settling time.	38
2.29	Rejection vs. buffer BW over buffer sizes (no predistortion).	39
2.30	IQ nonlinearity vs. buffer BW over buffer sizes (digital predistortion).	40
2.31	Sampling modes and their effect on the DAC filter.	40
2.32	Maximum zero magnitude for baseband Butterworth and Chebyshev Type I filters.	42
2.33	Comparison of adaptation and calculation of DAC taps.	44
2.34	Convergence to A_n as N_{Taps} increases.	44
2.35	TX and DAC spectrum after sampling.	45
2.36	Narrowband channel residual.	48
2.37	Narrowband channel residual.	48
2.38	Digital backend cancellation.	49
3.1	Die photo of first chip.	51
3.2	Chip top level.	51
3.3	Operation of switched-capacitor power amplifier [45].	52
3.4	Available constellation regions for PA architectures.	53
3.5	Theoretical drain efficiency of polar versus Cartesian SCPA [48].	54
3.6	TX top level.	54
3.7	Simulated TX insertion loss due to RX transformer.	55
3.8	RX top level.	56
3.9	Top level of DAC.	56
3.10	Mismatch propagation for switches and tail device.	57
3.11	DAC output current waveforms with and without reset.	58
3.12	DAC high level structure.	59
3.13	DAC with noise feedback highlighted.	60
3.14	Current DAC noise propagation models.	62
3.15	Feedback rejection and normalized output noise.	63
3.16	Simulated TX insertion loss due to RX transformer.	64
3.17	Total insertion loss and desensitization for architecture.	64
3.18	System test setup.	66
3.19	TX measurements.	66
3.20	RX measurements.	67

3.21	DAC quadrant I constellation up to 4 bits.	68
3.22	TX rejection vs. DAC oversampling ratio.	69
3.23	Initial TX cancellation measurements.	69
3.24	Residual vs. TX frequency.	70
3.25	VSWR test setup and results.	71
3.26	NF degradation vs. TX power.	71
3.27	Test setup for phase noise cancellation measurement.	72
3.28	Phase noise cancellation measurement with single-tone, narrowband, and wideband injection.	73
3.29	Phase noise cancellation with 1m cable.	73
3.30	Test setup for phase noise folding measurement.	74
3.31	Measurement results for LO spur injection relative phase and amplitude.	74
3.32	Source phase noise cancelled through sharing LO between TX/RX.	75
3.33	DAC thermal noise cancellation measurements.	75
3.34	DAC model.	77
3.35	Baseband channel and effective baseband pulse response.	77
3.36	Constellation refinement procedure.	78
3.37	Comparison of measured constellations before after channel de-convolution.	78
3.38	Comparison of full measured constellation with a reconstruction using 2% of the total measured points.	79
3.39	Matching of 15MHz data.	80
3.40	TX passes supply ripple to output.	80
3.41	Simulated constellation with memoryless PA supply nonlinearity.	81
3.42	TX unit step compared with supply ripple.	82
3.43	Effect of lowering PA supply network Q	82
3.44	TX supply, arbitrary sequence with deadtime.	83
3.45	PA model including supply modulation.	83
3.46	Constellation comparison using supply modulation model.	84
3.47	Comparison of multiple constellation packets.	85
3.48	Comparison of arbitrary data measurement with reconstruction.	85
4.1	Top level of passive mixer first RX architecture.	88
4.2	Baseband cross-coupling for imaginary input impedance synthesis [64].	88
4.3	NF vs. susceptance for $R_{Match} = 200\Omega$	89
4.4	Possible locations for capacitor DAC.	90
4.5	Representative 1:2 transformer available gain ($Q_P = Q_S = 10, k = 0.9$).	90
4.6	Total NF vs. capacitor DAC location.	91
4.7	Total NF vs. C_{On}/C_{Off}	91
4.8	Total NF vs. capacitor DAC bits.	92
4.9	Total NF vs. capacitor DAC Q	93
4.10	Total NF vs. harmonic trap series inductance.	93
4.11	Capacitor DAC unit cell voltage swing for +20dBm.	94

4.12 Stacked capacitor DAC, disabled unit cell.	95
4.13 Simple biasing of unit cell.	96
4.14 Effect of bias resistor size on system noise figure.	96
4.15 PMOS switch voltage swings.	97
4.16 System noise figure with series harmonic trap included.	97
4.17 Matching network efficiency vs. DAC cap.	98
4.18 DAC columns with cascodes.	98
4.19 Cascode code-dependent bandwidth.	100
4.20 Bleeder currents for column cascodes.	100
4.21 Simulations of column cascode-induced nonlinearities.	101
4.22 DNL due to column cascode (500fF source cap, 300uA bleeder).	101
4.23 Illustration of cascode and bleeder noise current injection.	102
4.24 NF with and without feedback for column cascode.	102
4.25 Die photo of second version.	103
4.26 S_{11} measurement vs. simulation over capacitor DAC code sweep.	104
4.27 S_{11} resonance frequency vs. capacitor DAC code.	105
4.28 Measured S_{11} versus frequency with capacitor DAC tuning.	106
4.29 RX linearity metrics.	106
4.30 Improved phase noise cancellation bandwidth.	107
4.31 Breakdown of noise vs. cancellation power.	107
4.32 Gain compression vs. TX power.	108
4.33 RX input harmonic power normalized to fundamental.	108
4.34 Harmonic compression testing setup.	109

List of Tables

1.1	LTE duplexer specifications [9].	4
3.1	Comparison with prior art.	76
4.1	Comparison with prior art.	109

Acknowledgments

This is my third time trying to write this section, and I don't think I'll ever be satisfied with it because there will always be more warm memories, of wonderful people and experiences, which resurface. This five-year journey has brought me love, excitement, anxiety, stress, confidence, pride, happiness, and a sense of overwhelming gratitude for all those who have been a part of my life up to this point.

To my parents, Joe Calderin and Dolores Harter, I cannot possibly express the amount of love and support and nurturing you have both given me over the years. In these last five years, I feel that our relationship has deepened considerably and it gives me such pride and satisfaction that I can finally begin to reciprocate even a little bit of your boundless encouragement as you both strive to realize your dreams. Dad, the little science experiments we used to do in the garage, among all the clutter endlessly fascinating to my young self, getting to help do house repairs, and watching as you flitted from one huge project to the next in your overalls, these experiences instilled within me a great energy towards creation for which I am forever grateful. Watching as you take a 40-year dream of building a hot shop to fruition is so inspiring. I've cherished our talks recently. Mom, no one else has given me such limitless enthusiasm and support for what I do. You are always the first to share in my struggles, my triumphs, my life, and I am so grateful. You always listen. I don't think I will ever fully repay that incredible gift you've given me, but I am so glad you live so close so that I can at least try.

My sister Athena, you and Brandon (and now Coda, too) are so welcoming and the love I have for you three is always increasing. Your house in Concord is a place where I feel completely accepted and free. There is no sense of entitlement to that space, just a sense of deep comfort which helps me to forget about anything outside of those walls, outside of our family. Our friendship, our understanding of one another, our closeness, these have all evolved and matured and I can't wait to see what the next five years brings for us.

Auntie Kathryn, I wish I had known you for longer, but I'm so thankful that I knew you growing up. Your incredible enthusiasm was contagious, and I am afflicted with this boundless energy even today. Every time I got to see you was such a treat; remembering your voice and your warm smile makes me very happy. Memories of you giving me a periodic table t-shirt when I was 5, the Disco Dan strobe light project, teaching me about telepathy, telekinesis, astral projection, and the Aakashic Records, your wonderful humor, and your stories about working at NASA, these make me feel full of light and love. I cherish that illustrated booklet for Samuel Taylor Coleridge's *Kubla Khan* that you made, and to this day I'm still amazed that at the coincidence.

Auntie Becky, Uncle Chuck, Grandma, Granny, and all other members of my family, I'm grateful for all the support you've given me over the years and have wonderful memories of all.

Ozzy, my roommate for nearly all this time, we have seen one another through so much, and I can't imagine that stopping anytime soon. We have experienced incredibly stressful, sad, hilarious, joyous, and mundane times together. There is no one else so willing to make

a fool of themselves with me. You are just as likely to recruit me to do something crazy as I you, and that match is truly rare and wonderfully effortless. I'll never forget what Fiona at Anchor and Hope said to us on our first Snowman Day, her astonishment that after spending 6 weeks of 18-hour days right next to one another, both infinitely on-edge from caffeine and adrenaline, we chose to spend the entirety of our first day off together, sharing new experiences. I hadn't even considered sharing those moments with anyone else, and that's truly amazing.

Sameet, were I to spend ten pages thanking you, I don't think it would be enough. The level of collaboration between us over our 4 years of research on this project is unparalleled. Every idea is our idea. You are someone I respect and look up to so highly, and yet in our discussions I feel like an equal. In technical discussions, the clarity of your thought truly amazes me; it is able to cut through the haze of half-baked ideas of my own and get to strong pieces of insight. Your dedication to work and your organization are prime examples of things I aspire to. Every single day of our research, I have wanted to work together, and I can't imagine many other people are lucky enough to feel like that. It is impossible to give advice to younger students based on my experiences here; the case of our partnership is just so rare, and so effective, that I honestly cannot imagine these years any other way. All the pains and triumphs of our work, we shared so much so equally. My friendship with you is one of the best I have, and I'm so thankful for this.

Ali, I cannot thank you enough for allowing me to transfer into the doctorate program. EE142 was the best class I have taken at Berkeley. It cemented my desire to pursue a Master's degree. Then, a single semester in your group made me realize how exciting original research is. Your complete lack of hesitation when I asked you to be a part of your group for a doctorate program gave me an incredible sense of acceptance for which I will be forever thankful. You are an incredible wealth of knowledge and our discussions have been integral to Sameet and myself reaching this point in our work.

Elad, the day that Sameet and I came to your office when we first came up with our active cancellation architecture is something I'll never forget. While you were not either of our advisors, you never hesitated to meet with us, give us suggestions and encouragement. You set a high bar for professors everywhere.

Bora, your advice, encouragement, and enthusiasm towards our work was incredibly motivating, and it is clear that you genuinely care about the well-being of your students.

Antonio, I'm very glad we've become good friends. The many hours talking with you on the train ride from Portland to Oakland stands out in my mind. Your incredible intelligence and eloquence, complemented with your great personality and sense of humor, makes working with you a great joy. That night before the ISSCC deadline with you and Sameet, staying up all hours testing, writing, breaking, more testing, with fluorescent light making us forget what time it was, I'm glad we were all part of that emotional rollercoaster.

Bonjern, our friendship has convinced me that music is a true force for good in the world. I love our discussions, and the fact that these discussions have brought us closer together as friends is wonderful. You've pushed me beyond my boundaries so many times and have been one of the strongest forces for shaping how I view and interpret music.

Lorenzo, while we haven't known one another for very long, I'm very glad to count you as a good friend and appreciate both our technical and non-technical conversations. Your enthusiasm is easy to match and I'm incredibly glad we share an interest in a weirder side of sound art.

Steven Callender, I'm forever grateful for you taking Ozzy and me under your wing and showing us what someone with true drive and motivation looks like. It was inspiring (over)working with you.

Nathan, the post-apocalyptic dystopian wasteland spring break we had is something I always look back on fondly.

James Dunn, you are one of the most helpful people I've ever met. You truly go above and beyond every part of your job description to assist anyone at BWRC with nearly any issue.

Krishna, Pavan, Katerina, Rachel, and so many others at BWRC, I'm really glad I've gotten to know you all.

James Friedberg, I cannot thank you enough for being my introduction to Public Glass. I respect so much your incredible command of glass and your ability to distill your knowledge into portions I can understand. Glassblowing is such a huge departure from what I do on a day-to-day basis, but you've always made me feel included and I am so appreciative of that.

Sarah Jordan, our friendship is something I cherish, and your realness is an incredibly potent cure for my overthinking and neuroses. I'm so proud of you for moving across the country to pursue nursing. It's hard to believe it's been two years since you left BWRC, but it's not hard at all to believe how well you've taken to New York; you could fit in anywhere and I'm so glad you're someone who I know.

Ellison, having known you since about my first day of high school, it's been great becoming "adults" with you. While the details of our paths and the nuances of our personal growths may be different, I've always felt like on the whole we have very similar emotional reasonings, and our discussions have been such a great thing for me. Each time we hang out, no matter how long it has been, things feel comfortable and fun, and after every conversation I'm still wishing we had more time to catch up.

Mitchell, I've known you for longer than any other friends on here. Living half a block away from you until we were 18 was a great experience, and I'm so thankful that your parents somehow put up with me being over all the time. Our Mod-Sports, our unicycling and rollerblading adventures, pressing coins on the train tracks, all wonderful times which I look back fondly.

Kapuscik, were it not for our electronics projects in high school and undergrad, I would probably have stuck with computer science. I would be making far more money and the total amount of stress in my life would be far lower. Thank you for putting me on the right path.

For everyone who I've interacted with in my life, I am so thankful for each and every event, as it has brought me to this point in my life.

Chapter 1

Introduction

Over the last five years, global data usage on mobile devices has grown at a truly incredible rate, an annual 47% increase [1]. The principal driver of this growth is mobile-based video streaming and sharing, which accounts for 60% of total mobile traffic, projected in Fig. 1.1 to make up 80% by 2021. Social media giants Facebook, Snapchat, and YouTube boast incredible statistics for video use: Facebook users collectively view 8 billion videos each day, while Snapchat users are projected to view 18 billion videos per day by the end of 2017 [2, 3]. Each minute, nearly 40 days of Snapchat video is uploaded per minute, while YouTube sees 12 days uploaded per minute [3, 4].

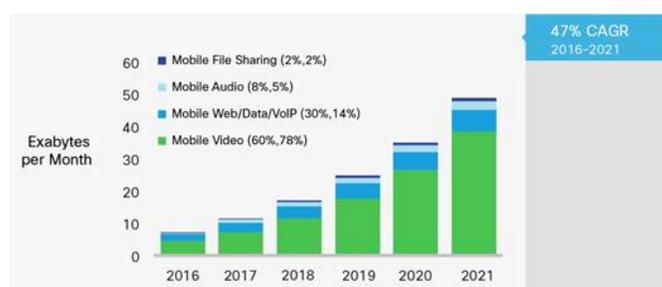


Figure 1.1: Projected monthly global traffic for mobile data [1].

Fig. 1.2 illustrates this meteoric rise of video recording due to the prevalence of smartphones. Taken eight years apart, these two pictures are from the inaugural speeches of Popes Benedict and Francis, with the road leading to Francis' speech sharply illuminated with a sea of screens, continuously recording video. While this 2013 audience was content with saving video to their phones for later uploading, 2016 marked the rise of Facebook's Live feature, now accounting for 20% of all Facebook videos, empowering users to instantly stream significant events [5]. These events, which would put incredible strain on mobile network infrastructure, are rare enough, but smaller-scale events happen everyday, when there is a sharp increase in video demand at night relative to average [1].



(a) Pope Benedict's inaugural speech, 2005 [6].



(b) Pope Francis' inaugural speech, 2013 [7].

Figure 1.2: Comparison of technology use 8 years apart.

Along with current smartphone users consuming more data than ever before, the number of mobile devices is far outstripping population growth. In 2016 alone, nearly half a billion smartphones were added to mobile networks [1], and by 2021, it is estimated that there will be 50% more mobile devices than humans on the planet [1].

Forecasting this tremendous increase in data usage, the 3rd Generation Partnership Project (3GPP) created the specifications for Long-Term Evolution (LTE) in mobile communication,

where a long series of “Releases”, each approximately 1 year apart, augmented the existing mobile communication standard with more features enabling higher datarates and denser deployments. While the first 4G LTE release was published in 2006, the final 4G release, Release 14 will not have its features frozen until June 2017 [8]. This upcoming release represents the final instantiation of the 4G standard, but it will be many years until the 5G standard will be ubiquitous; it is estimated that by 2021, only 0.2% of mobile devices will be 5G capable, with 56% of devices communicating through 4G [1]. It is therefore accurate to state that 4G technology will take up the majority of high-definition live video streaming and data consumption in the next five years.

In LTE, in order to serve many users within the same sector, multiplexing systems are implemented where users are assigned a “resource block” in time and frequency, shown in Fig. 1.3, where they can download or upload data. Each mobile carrier has specific frequency bands which they allocate to their users at different times in the form of these resource blocks. Additionally, different regions of the world have different frequencies where they allow mobile communication, summarized in Fig. 1.4. Needless to say, the LTE spectrum is highly splintered, meaning that a phone meant to be operated internationally must support a wide range of frequencies in order to achieve good performance in each region.

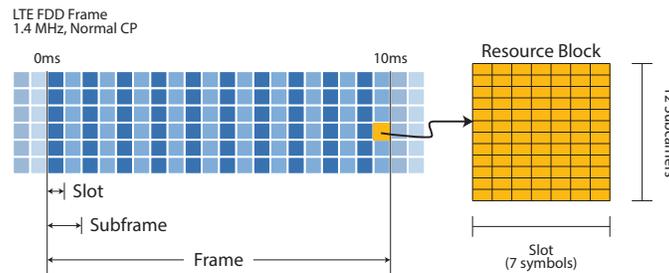


Figure 1.3: LTE frame with resource blocks highlighted.

One of the most challenging barriers to this goal of global LTE operation is the use of Frequency Division Duplexing (FDD), where transmission and reception happen simultaneously, but in separate frequencies. This is contrasted with Time Division Duplexing (TDD), where the same frequency is used, but at different times. An illustration of these two duplexing scenarios is shown in Fig. 1.5. For FDD, strong filtering is required due to the power of the transmitter (TX), the sensitivity of the receiver (RX), and the relative spacing between these two bands, which is only twice the bandwidth in the worst case. Nearly 120dB of dynamic range exists between a high powered TX signal and the sensitivity floor of the RX, as shown in Fig. 1.6. At least 30dB of isolation is required between the TX and RX bands to prevent compression, with additional isolation of the TX in the RX band to prevent desensitization due to PA noise and out of band interference. This isolation is normally performed using Surface Acoustic Wave (SAW) or Bulk Acoustic Wave (BAW) duplexers, offering very high isolation between TX and RX ports, as shown in Table 1.1.



Figure 1.4: LTE global frequency allocations.

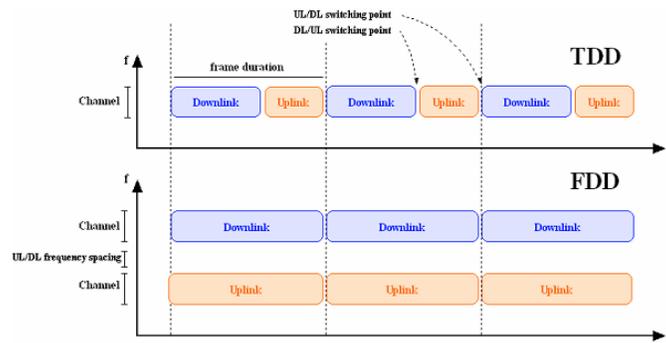


Figure 1.5: FDD vs. TDD.

TX Band Isolation	50dB
RX Band Isolation	45dB
TX Insertion Loss	2.5dB
RX Insertion Loss	2.5dB
Area	1.7mm ²

Table 1.1: LTE duplexer specifications [9].

SAW and BAW devices work in very similar ways, where they effectively use piezoelectric materials to excite vibrational substrate waves (either within or on the surface of the substrate) shown in Fig. 1.7, which have a much smaller wavelength than in air due to the reduced propagation speed. This allows high-order transmission line networks to be built in a modest space which achieve very sharp filter cutoffs. Their main drawbacks are their size, manufacturing cost, and frequency inflexibility. Because the filters rely on precise separation between piezoelectric elements to operate, modifying their operating frequency while maintaining good quality factor has not yet been achieved. These filters are created for specific

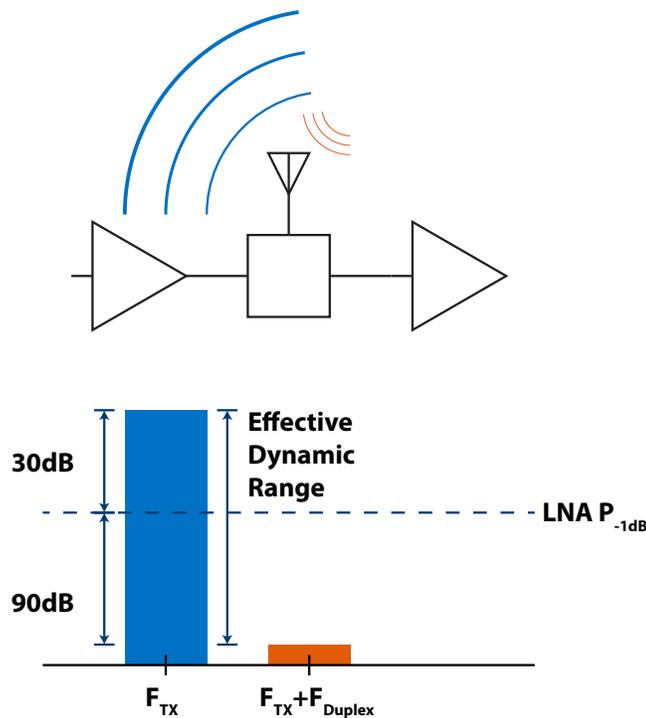


Figure 1.6: Dynamic range between TX power and RX sensitivity.

FDD bands, of which there are over 30 globally, making it extremely challenging to create a phone operating well in all regions.

SAW & BAW Technology Comparison

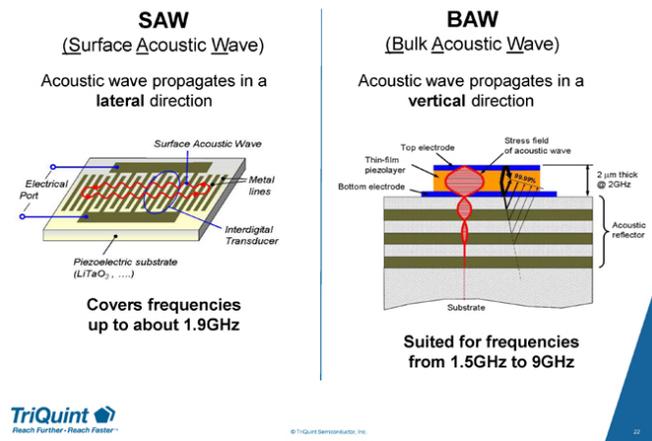


Figure 1.7: SAW and BAW mechanisms [10].

To make matters worse for fixed filters, carrier aggregation (CA), combining of multiple

communication bands to provide higher instantaneous bandwidth, is an integral component of the 4G standard, illustrated in Fig. 1.8. In these scenarios, filters are required for not only the individual bands, but any CA combinations supported. Currently, 3x CA (a combination of 3 frequency bands) is becoming widespread, significantly increasing the interconnect requirements for smartphone boards. CA is traditionally enabled through a large number of switched filter stages, with or without multiple antennas, and several shared transceivers [11, 12]. The number of interactions between the bands grows rapidly, and the RX must be shielded from both the large TX output power in its band, as well as the TX interference which falls in the RX band. For a 3x CA scenario, 6 filters are needed, and these filters must be verified over 18 different band pairings. For an N x carrier aggregation scenario, $2N^2$ isolation conditions must be met. Korean telecommunications company SK Telecom have recently shown in trials 5x CA [13] and announced that they are planning on supporting 6x CA by 2018 [14], which requires 12 filters and 72 different band pairings, severely constraining filter design. Research has been done [15] which shows that it is possible to create integrated, highly tunable filtering approaches for downlink CA in the presence of moderate blockers, but strong TX/RX isolation would still be required to include uplink. Building a system with a large number of fixed filters, where a small fraction will be enabled at any given time, is wasteful and creates a complex interconnect network, adding significant insertion loss and sensitivity reduction to the overall system.

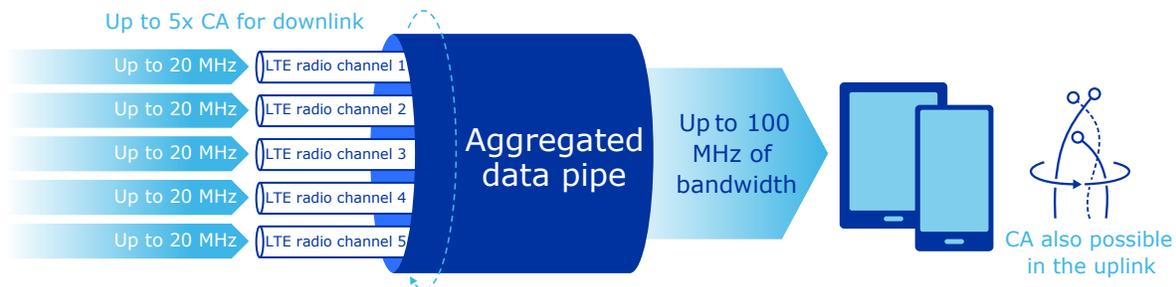


Figure 1.8: Carrier aggregation: combining multiple bands for higher datarates.

Beyond the numerous issues handling LTE FDD in a global manner, there is the common scenario of multiple wireless communication standards (Wi-Fi, GPS, Bluetooth, GSM, CDMA, LTE, etc.) operating simultaneously within the ISM band. Shown in Fig. 1.9 is the Amazon Fire smartphone, with its main RF circuit board featuring many transceiver chips which operate multiple wireless standards heavily overlapping with one another in frequency. Voice Over LTE (VoLTE) is becoming more common, but cellular users otherwise use CDMA or GSM for voice. Imagine a scenario where someone with wireless earbuds is calling a friend they are meeting at a restaurant they have never heard of. They could be using Bluetooth for their earbuds, GSM for voice, GPS for location services, and LTE for maps data, all simultaneously. Because of the tightly packed chips within their phone's frame, this is an extremely challenging environment from an interference management perspective. According

to the UMTS specification for wireless voice [16], the maximum transmit power for a handset is +33dBm, which must not interfere or damage the LTE receiver. This is essentially an FDD scenario, but with a leakage channel between chips or between sections of the same chip.

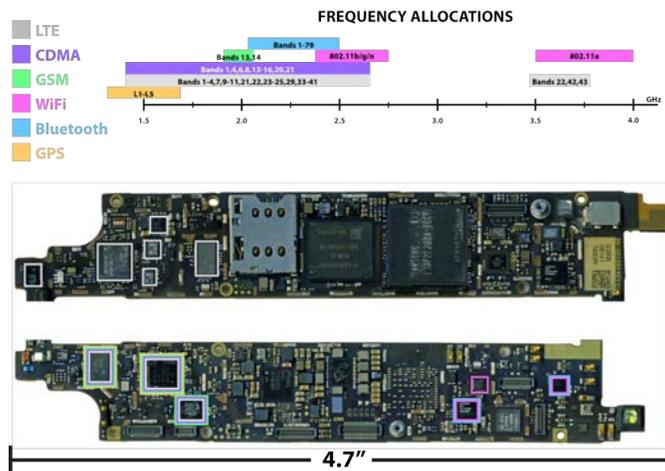


Figure 1.9: Amazon Fire phone multi-standard transceivers.

It is because of these numerous challenges that alternatives to fixed filter interference management are very attractive. Presented here are different methods for integrated, wide-band or frequency-flexible interference mitigation.

1.0.1 Prior Art

At a high level, previous works can be grouped into two separate sections: general interference cancellation and self-interference cancellation.

1.0.2 General Interference Cancellation

Here, no knowledge of the interfering signal is assumed, beyond the fact that it is not located in the RX band. A classic example of this type of interference is a blocker, an aggressor with worst-case amplitude and frequency spacing defined by the standard. Because the on-chip TX signal acts like a blocker, these techniques could also be used for self-interference cancellation. Multiple methods exist for general interference cancellation such as N-path filtering and feedback techniques.

N-Path Filtering

First written about in the early 1950s [17], but not made practical (due to low quality switches) until very recently [18], passive mixers translate baseband impedances to the switch

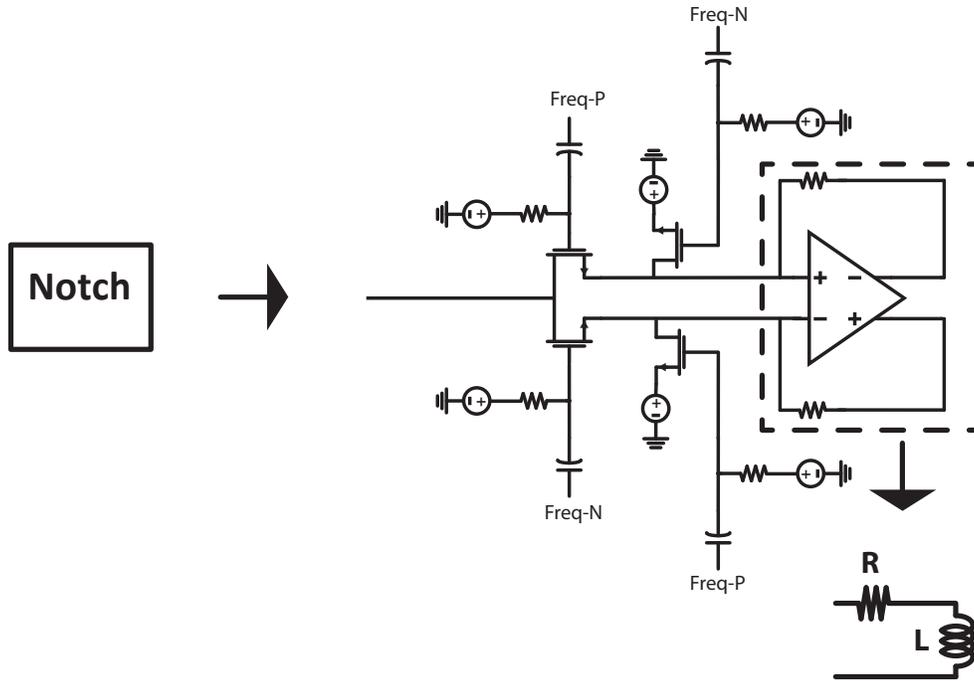


Figure 1.10: Active series LR N-Path for programmable notch.

LO frequency. This allows the construction of frequency-reconfigurable high-Q filters. For example, in Fig. 1.10, an active series LR impedance is upconverted to create a notch around the frequency of an incoming blocker. The main advantage to this filter configuration is the ability to create resonant filter effects without the need for bulky inductors or ultra-thick metal processes [19, 20, 21, 22, 23, 24]. Reconfigurability also allows a single transceiver to operate in many different bands without the need for extra hardware.

While these advantages seemingly make it an attractive option for self-interference cancellation, a number of disadvantages prevent it from being a viable candidate on its own. First and foremost, if the baseband is created using passive elements, there exists an inherent tradeoff between RX insertion loss and TX cancellation for small duplex spacings, which are prevalent in the LTE standard [25]. This is because for RL and RC networks, poles and zeros must have alternating real parts, preventing more than 20dB/decade rise or fall in impedance. LRC networks can have small regions where the impedance changes faster, but these regions are too small to be practically used [26]. Recently, [15] showcased a system using active elements which exhibits an impedance proportional to $|s|^2$ using cross-coupled G_m cells and a capacitor. This allows higher order impedances to be created, but the use of active devices limits system linearity.

Feedback Techniques

In a feedback network self-interference cancellation scheme, a network with high loop gain in the interference band is placed early on in the receiver chain to minimize the interference signal. The action of a feedback network is to reduce the error signal to a level determined by the total gain in the loop. By creating networks such that the error signal is the TX interference, a large loop gain will reduce this interference such that it does not compress later stages. An example of such a network is shown in Fig. 1.11. The main drawback of this technique is its sensitivity to gain and phase variations in the loop transfer function, which can destabilize the loop. This problem is exacerbated by the use of filters to reject the RX signal from the feedback path; the phase shift added by multiple poles or a high- Q resonance can easily destabilize the feedback path. Many works use a frequency-translational filtering technique [27, 28, 29, 30, 31, 32], where the baseband filter can be constructed such that it does not affect stability significantly, and the two mixers in the loop effectively upconvert the baseband filter to be centered around the LO frequency.

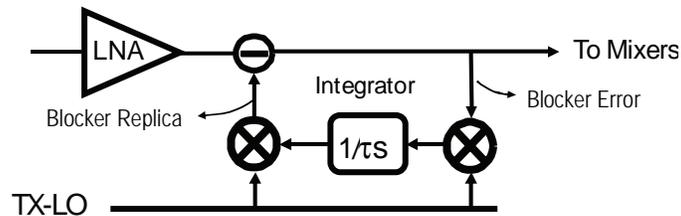


Figure 1.11: Feedback technique applied to TX [28].

1.0.3 Self-Interference Cancellation

In cases where the interference source is co-located with the receiver, Frequency Division Duplex or Full Duplex communication are prime examples, other techniques can be used which offer better performance or flexibility than general interference cancellation.

Hybrids

Hybrids are passive networks capable of isolating TX and RX through causing the TX signal to appear as a common mode for the RX port. Constructed with coupled inductor networks, the common mode signal on one side of the transformer will not leak to the other side, neglecting inter-winding capacitance. These structures have wide instantaneous rejection bandwidth and can isolate high power blocker signals. A balancing impedance, shown in Fig. 1.12, is required for hybrids to operate properly, where the voltage swing present on these tunable impedance nodes normally sets the linearity of the full system, allowing for high power TX rejection [33, 34, 35, 36, 37, 38]. One major drawback exists for hybrids, where there is a direct tradeoff between insertion losses on the TX and RX paths, where

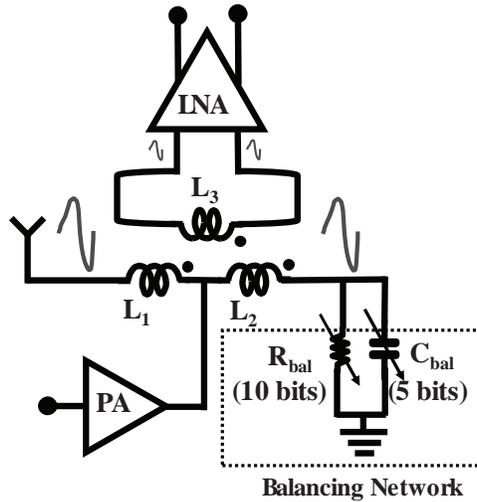


Figure 1.12: Hybrid with balancing impedance [34].

$IL_{TX} = 6\text{dB} - IL_{RX}$ [39]. Practical instantiations of this technique normally measure 4dB for both the TX and RX paths a significant penalty for the transceiver.

Active Cancellation

Offsetting the insertion loss penalty of hybrids, a replica of the TX signal can be generated with high accuracy, either by coupling a portion of the interferer’s signal directly or by using the interferer’s baseband data.

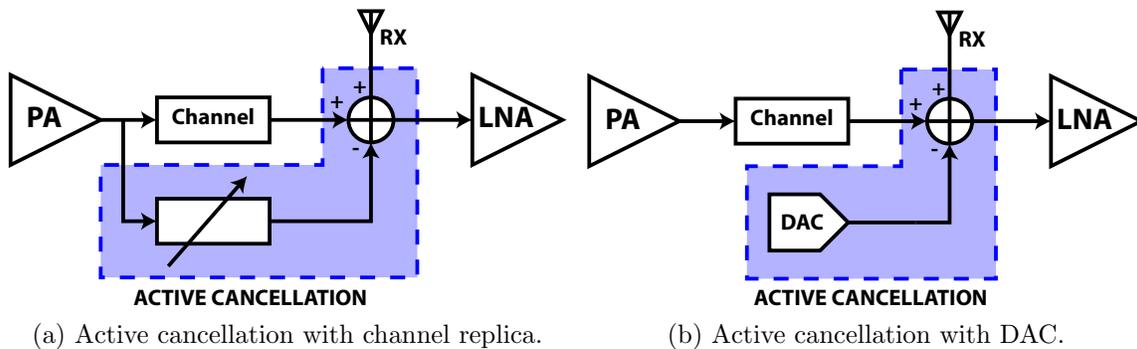


Figure 1.13: Active cancellation methods.

Those which couple the TX signal directly [40, 41, 42] use a bank of analog filters to generate a replica of the leakage channel, and then subtract it directly at the input of the RX, as shown in Fig. 1.13a. These have the advantage that any nonlinearity present in the transmitter is inherently preserved through the leakage channel. One disadvantage is that the power coupled into the replica network directly adds to insertion loss on the PA. A

major disadvantage is the linearity requirements for the replica leakage channel components, which directly sets the maximum cancelable TX power. Furthermore, because the channel is replicated using analog components, there is a limited bandwidth over which the filtration can be adjusted to match the leakage channel, making it challenging to achieve high cancellation over a wide bandwidth.

Finally, mixed signal techniques which reproduce the interference signal by using base-band data and a DAC, the category under which this research falls, have been used in the ethernet domain [43] to cancel interference from multiple simultaneous network streams. This technique is illustrated in Fig. 1.13b and has the advantage that no power is removed from the transmit path. Additionally, arbitrary filtering can be applied in the digital domain, widening the bandwidth over which active cancellation is effective.

Chapter 2

Theoretical Framework

The goal of this research is to either obviate the need for off-chip duplexers, or significantly reduce their requirements. Therefore, the canceller architecture should have the potential to meet or exceed the specifications laid out in Table 1.1. Additionally, a single antenna interface is highly desired to minimize the effective footprint of this architecture. Finally, frequency-flexible operation is needed. These requirements drive the novel choice of architecture presented in this research.

2.1 Conceptual Overview

Starting with the requirements for a single antenna interface, as well as low TX insertion loss, a key element of the architecture may be motivated. From the perspective of the TX, the duplexer would ideally look like a direct connection to the antenna, with no other components in series or parallel. An observation can be made that if $Z_{In,RX} = 0$, the receiver could be placed in series with the TX. Similarly, if $Z_{In,RX} = \infty$, then the receiver could be placed in parallel with the TX. Both configurations are shown in Fig. 2.1.

Next, consider only the series connection between this transmitter and receiver pair. If a current source is placed in parallel with the receiver, it is possible to modify the effective RX input impedance, due to the fact that the source shunts current away from the RX input. If this current source shunts the same current waveform which flows through the antenna, then $Z_{In,RX} = 0$, regardless of the actual input impedance of the receiver, shown in Fig. 2.2a. Taking this one step further, because the current source causes the RX to look like a short, rather than controlling the current source based off the current flowing through the antenna, the source can be controlled by another PA connected directly to an antenna, like in Fig. 2.2b. In this way, the current source causes the PA to see a short, but any signal incoming on the antenna sees the full impedance of the RX, a key observation of this work. It is important to note that in the real architecture, only one PA is used; the additional PA used here is simply to aid explanation.

Putting these together, conceptually the duplexing network consists of a series stack of

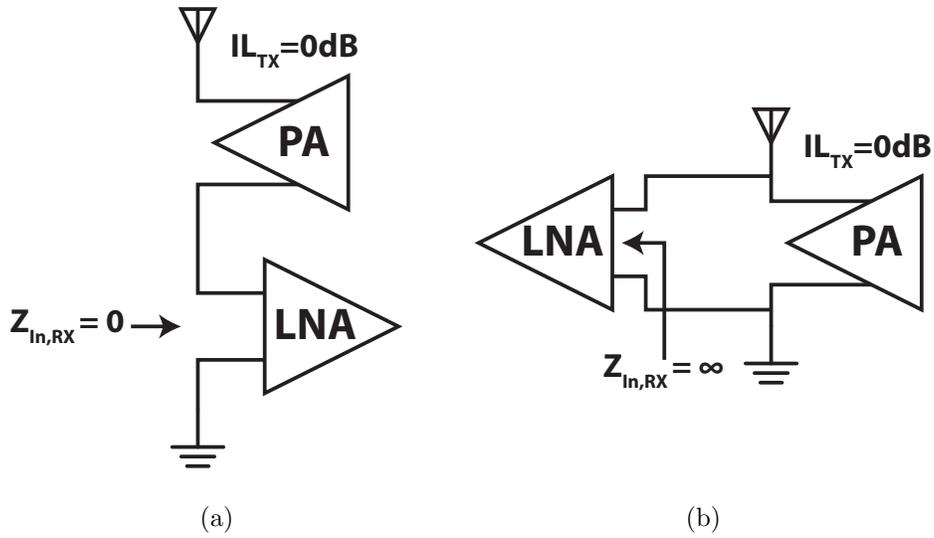
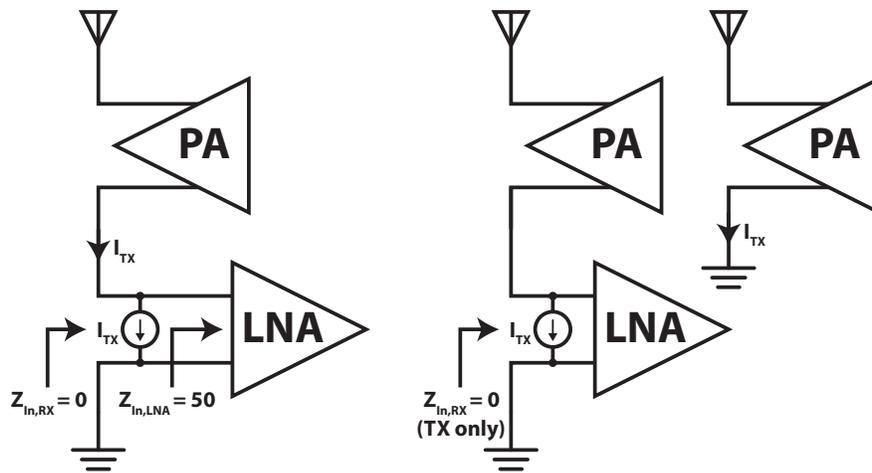


Figure 2.1: Naive TX and RX single antenna interface.



(a) Controlled current source to modify $Z_{In,RX}$. (b) Controlling source from an isolated PA.

Figure 2.2: Impedance modification through current source.

the antenna, the TX, and the RX. There is a replica current source in parallel with the RX which circulates only TX current, shown in Fig. 2.3. Focusing on the TX signals specifically in Fig. 2.4a, the current source's modification of the input impedance of the RX zeros the differential voltage across the RX, creating a virtual ground at the bottom of the TX balun.

Because the replica DAC sinks only TX current, RX loss is unaffected beyond any parasitics at the DAC output nodes and the output impedance of the TX, as shown in Fig. 2.4b. With todays switching PAs, the output impedance of a transmitter can be made on the order of a few ohms, allowing for sub-1dB loss for the receive path.

Provided that the DAC sampling rate and fullscale are high enough that the DAC can track the current excursions of the TX waveform, the residual error signal will be bounded to $\pm\text{LSB}_{DAC}$, regardless of the TX power. Therefore, the mean power of the TX residual will be equal to $P_{TX} - 6N_{Bits,DAC}$, shown in Fig. 2.4c.

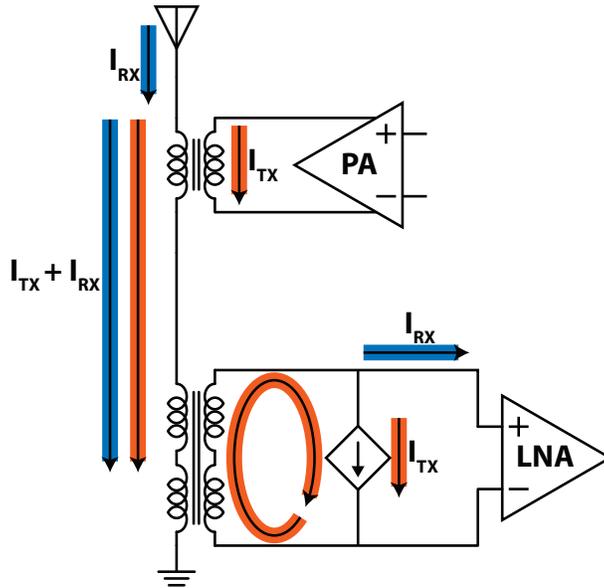


Figure 2.3: Top level conceptual diagram of cancellation architecture.

Compared with an analog cancellation system, this mixed-signal cancellation architecture is far more tolerant of channel and TX nonidealities. Because the DAC is fed by a digital baseband, an arbitrary number of filter taps may be applied, with no limit to their full scale, enabling the channel to be approximated over a wide bandwidth. Similarly, because the taps can be arbitrarily spaced, long echoes may be easily accounted for. Nonlinear predistortion of the baseband data can account for nonidealities present within the TX or DAC. Finally, assuming the sampling rate and instantaneous bandwidth of the DAC are high enough, multiple self-interference sources at different frequencies can be handled by making an equivalent baseband signal from multiple data sequences multiplied by $e^{j2\pi\Delta f T_s}$. These advantages are summarized in Fig. 2.5. For the rest of this chapter, further detail on the advantages and considerations of this architecture are given.

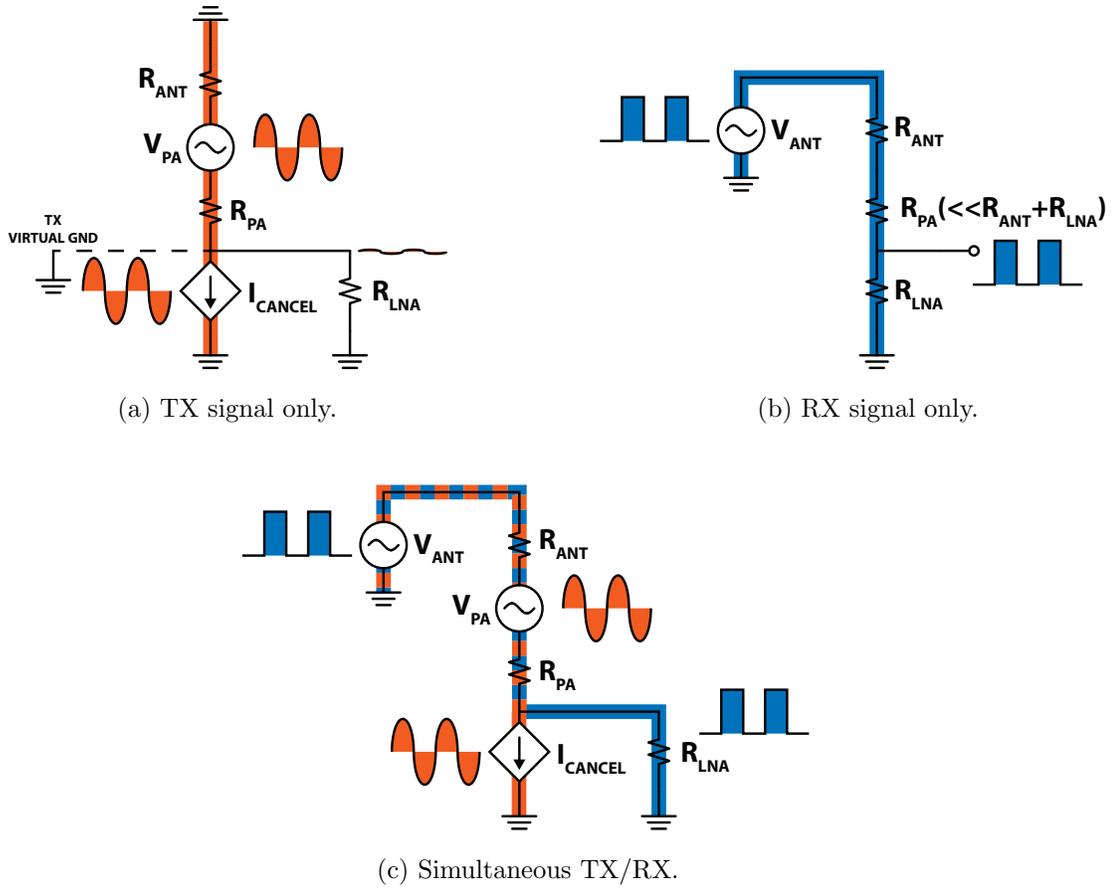


Figure 2.4: Breakdown of conceptual diagram.

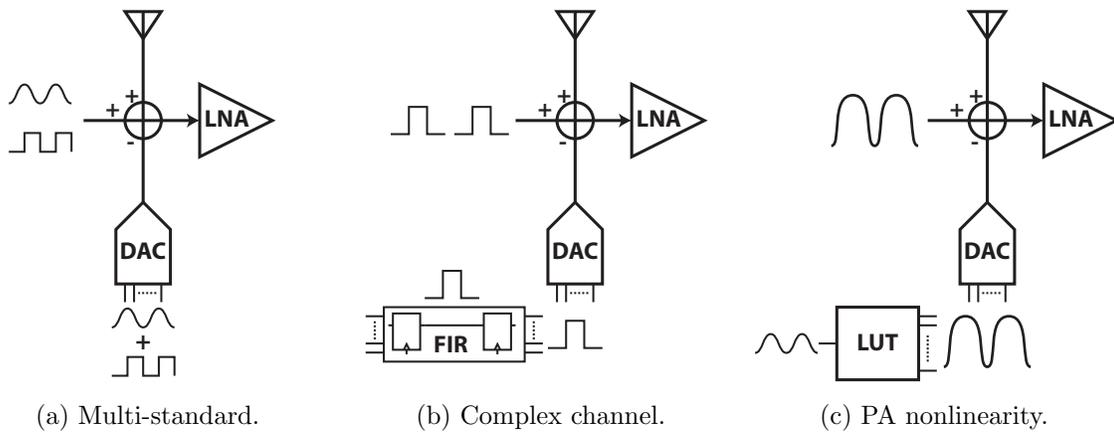


Figure 2.5: Advantages of mixed signal cancellation.

2.2 System Advantages to a Current DAC

At a high level, the objective of the cancellation source is to minimize the signal which is incident on the RX input. There are two signal domains available for cancellation, voltage and current, and the choice of domain necessitates a specific architecture for the system. These two cancellation domains are shown in Fig. 2.6. It is important to note that in both cases, the current through and voltage across the RX input is identically 0, but this does not mean that the voltage swing seen by the TX is equal to 0. In the current domain, Fig. 2.6a, the cancellation source is directly in parallel with the RX, so zero swing across the RX translates to zero swing seen by the TX, meaning the RX is seen as a short. In the voltage domain, Fig. 2.6b, the cancellation source is in series with the RX, meaning that the TX sees no current flow into the RX, and therefore the RX is seen as an open. In current domain cancellation, the low RX impedance (seen by the TX) necessitates a series connection between the TX, RX/DAC, and antenna, in order to incur little TX insertion loss. Voltage domain cancellation requires a parallel connection between the three elements.

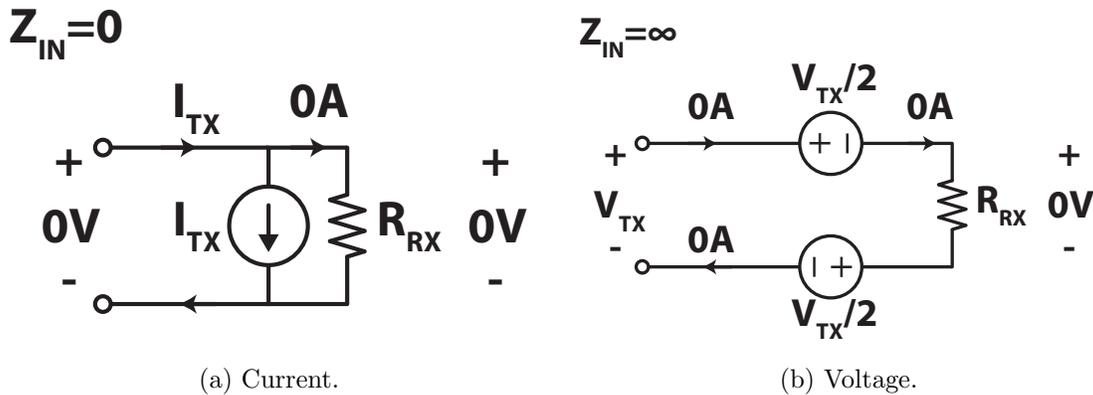


Figure 2.6: Cancellation signal domains.

There are a number of reasons for why current domain cancellation is more practical than voltage domain. First, while it is simple to create an effective floating differential current source, it is much less practical to create a floating voltage source which is balanced across the RX input port. While the RX is assumed to have common mode rejection, the high power of the TX signal may still desensitize the receiver in common mode due to imbalanced cancellation. Secondly, it is far more difficult to create large voltages on chip than large currents. Voltage mode CMOS power amplifiers commonly require stacks of transistors to reach powers beyond +20dBm without damaging the devices[44, 45] due to the high voltages incident on the transistors. Even if a 2:1 transformer is used for the RX to halve the voltage requirement of the replica, >3V peak-to-peak swing is needed for a +20dBm signal. In the case of current domain cancellation, the fact that there is ideally no voltage swing across the RX can be exploited such that the current DAC array can be simply scaled to cancel higher TX powers. To first order, the replica sees a short at its output, regardless of current. The

electrical dual of this scenario, zero current but high voltage, cannot be exploited as simply due to transistor breakdown.

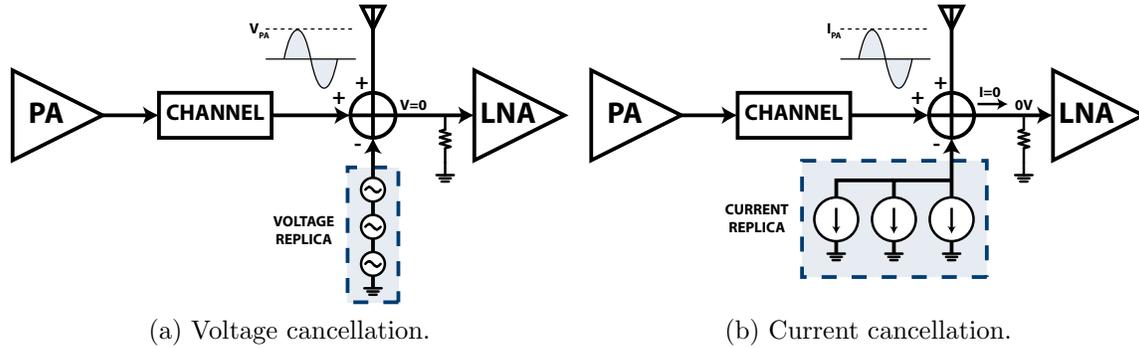


Figure 2.7: Generation of replica in voltage and current domains.

2.3 DAC Power Consumption

In the conceptual image of the system from Fig. 2.3, the current source experiences zero differential voltage swing. If it were possible to create an ideal floating current source and use this as the canceller, the cancellation system would consume no additional power. Because there is not a way of creating such an ideal floating current source, the DAC is instead designed as a hard-switched differential pair powered from the RX center tap. The center tap voltage is set by the static headroom requirement for high output impedance on the DAC. It should be noted that the replica DAC cancels the *current* of the TX, rather than the power. This allows the DAC power consumption to be much smaller than the TX as power levels rise.

At a high level, the DAC modulates its DC tail current, drawn from the RX center tap, around F_{TX} to cancel the TX current incident on the RX input. For a TX power of P_{TX} , an antenna impedance of R_{Ant} , and an RX balun turns ratio of N_{Turns} , the sinusoidal current amplitude flowing through a short at the RX input port is equal to:

$$I_{TX,RX} = \sqrt{\frac{2P_{TX}}{R_{Ant}}} \frac{1}{N_{Turns}} \quad (2.1)$$

The DAC replicates this current using a waveform with a fundamental amplitude of $2A_{Conv}I_{DAC}$, where I_{DAC} is the tail current and A_{Conv} is DAC conversion gain. The factor of 2 is due to the fact that conversion gain is referenced to complex exponential magnitude. Therefore,

$$P_{DAC} = V_{DD,DAC} \sqrt{\frac{P_{TX}}{2R_{Ant}}} \frac{1}{N_{Turns}A_{Conv}} \quad (2.2)$$

where $V_{DD,DAC}$ is the center tap supply. An important observation is that the power consumption of the DAC is proportional to $\sqrt{P_{TX}}$ due to the fact that the DAC is pulling its replica current from a constant supply.

This power consumption can be normalized by the power consumption of the TX using η_{TX} , the PA efficiency, to produce an effective cancellation efficiency metric.

$$\eta_{Cancellation} = \frac{A_{Conv} N_{Turns} \sqrt{P_{TX} R_{Ant}}}{\sqrt{2} \eta_{TX} V_{DD,DAC}} \quad (2.3)$$

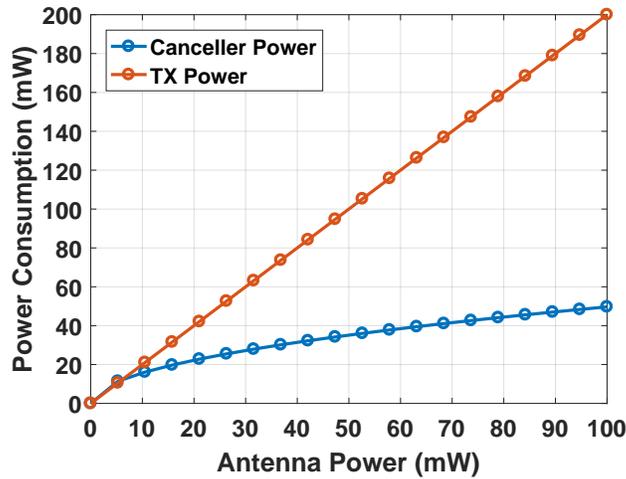


Figure 2.8: Power consumption of canceller versus TX.

Plotted in Fig. 2.8 is the power consumption for the DAC versus the TX for a range of TX output power levels. This concrete example is representative of the design where $V_{DD,DAC} = 1V$, $N_{Turns} = 2$, $\eta_{TX} = 50\%$. The DAC current waveform is a 50% duty-cycle square wave of amplitude $\frac{I_{DAC}}{2}$, meaning $A_{Conv} = \frac{1}{\pi}$. Around the 100mW power level, or +20dBm, the replica DAC uses 25% of the TX power.

2.4 Noise Sources

2.4.1 Noise Overview

Any elements exhibiting power gain or loss exhibit non-deterministic fluctuations in their voltage/current which is referred to as noise. One important metric for a receiver system is the noise figure (NF), which takes into account both the gain of the system and the noise output from the system. Noise figure is defined as

$$F = \frac{\text{SNR}_{\text{In}}}{\text{SNR}_{\text{Out}}} \quad (2.4)$$

$$= 1 + \frac{\overline{v^2}_{\text{In},n}}{\overline{v^2}_{\text{Source},n}} \quad (2.5)$$

where $\text{SNR}_{\text{In,Out}}$ are the signal-to-noise ratios at the input of the receiver and output of the receiver, respectively. For a general RX, it is assumed that the transmitted signal has extremely high SNR, such that the only noise at the input is due to the antenna noise. Interestingly, the noise due to the antenna does not come from the antenna itself (ideally, the antenna is a purely reactive element, aside from radiative properties), but in fact comes from the antenna picking up thermal radiation from the room in which it resides, with an average power density equal to kTW/Hz , this power density can be rewritten in terms of a voltage variance density $\overline{v^2}_{\text{Source},n}$, whereas the noise of the circuitry within the RX is referred to the input of the RX as $\overline{v^2}_{\text{In},n}$.

It is very important that the duplexing system minimally increase the RX noise figure, as noise figure directly impacts receive distance for a given TX power. The noise in this system can be lumped into 4 categories: RX thermal noise, PA thermal noise, DAC thermal noise, and TX/DAC uncorrelated phase noise. The RX and PA thermal noise are independent of the TX power and form the base sensitivity of the network, while the DAC thermal noise and TX/DAC phase noise inject more noise as the TX power grows. This increase in desensitization, along with compression due to the TX third harmonic, is what sets the practical limit on TX output power. It will be shown that PA thermal noise is negligible, so the majority of the design work should be focused on minimizing the DAC thermal noise and the TX/DAC uncorrelated phase noise. All significant noise sources are shown in Fig. 2.9.

A final source of noise which is worth commenting on is RX LO phase noise. Through reciprocal mixing, interference outside of the RX band is mixed by the phase noise of the RX LO, spreading its energy to the RX baseband. Because this cancellation technique subtracts the TX interference before it is downconverted by the RX, reciprocal mixing of this small residual produces negligible desensitization.

2.4.2 TX Thermal Noise

The architecture for the TX must have low output impedance in order to not add considerable insertion loss due to the series combining network. In this particular system, the PA is a switching power amplifier, and because the transistors are hard-switched, there are only two sources of noise output from the PA: phase noise of the input signal and switch thermal noise. Given the use of a switching power amplifier as the TX, the PA can be thought of as a passive linear time-varying network, where power modulation is simply due to a code-dependent voltage divider off of the PA power supply. Using thermodynamic arguments [46], it can be shown that a passive network with output impedance Z outputs the same level

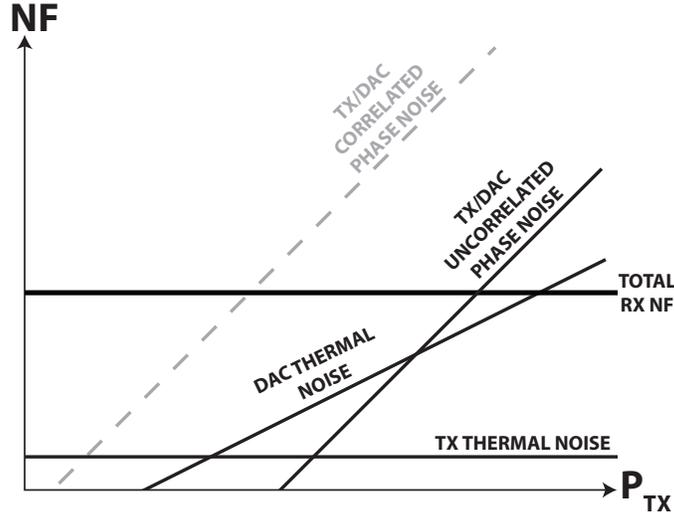


Figure 2.9: Significant transceiver noise contributions.

of thermal noise as a resistor of resistance $\text{Re}(Z)$. The noise figure is shown in Eq. (2.6), where $\overline{v^2}_{RX,n}$ is the input-referred voltage noise of the RX, and $\overline{v^2}_{Ant,n}$ is the antenna noise referred to the input of the RX. Requiring the RX insertion loss due to the TX to be small necessitates that the real part of the TX output impedance, $R_{TX} \ll R_{Ant}$, meaning the noise figure penalty is similarly small. Shown in Fig. 2.10, a simulation using transformer and PA parameters from the first version of the chip, the noise figure due to the TX alone is small, impacting the total noise figure by $<1\text{dB}$.

$$F = 1 + \frac{R_{TX}}{R_{Ant}} + \frac{\overline{v^2}_{RX,n}}{\overline{v^2}_{Ant,n}} \quad (2.6)$$

It is worth noting that the effects of TX noise added to the system and the loss due to the TX output impedance are one and the same, and should not be considered as independent degradations. Using the Friis cascade noise figure expression Eq. (2.10) to determine overall noise figure:

$$\overline{v^2}_{Ant,TX,n} = 4kT(R_{Ant} + R_{TX}) \quad (2.7)$$

$$G_{A,TX} = \frac{R_{Ant}}{R_{Ant} + R_{TX}} \quad (2.8)$$

$$F_{TX} = 1 + \frac{R_{TX}}{R_{Ant}} \quad (2.9)$$

$$F_{Total} = 1 + (F_{TX} - 1) + \frac{\left(1 + \frac{\overline{v^2}_{RX,n}}{\overline{v^2}_{Ant,TX,n}}\right) - 1}{G_{A,TX}} \quad (2.10)$$

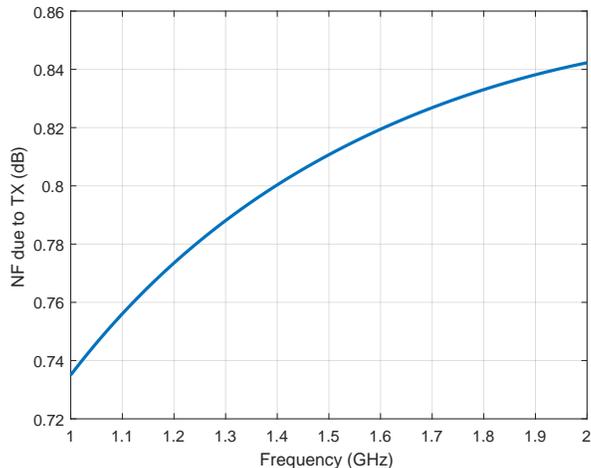


Figure 2.10: Noise figure due to TX only.

the total noise figure is exactly the same as taking into account only the effect of the noise voltage of the TX, as shown in Eq. (2.6).

2.4.3 TX Phase Noise

Phase noise is defined as fluctuations in the zero-crossing points of an LO. Just as any noise profile can be deconstructed into real and imaginary parts about some center point, an arbitrary noise profile can also be separated into phase and amplitude noise. For purposes here, the most important aspect of phase noise is the fact that unlike additive noise, the SNR due to phase noise is constant with signal power, since a mixer can be thought of as multiplying a noisy LO with a desired signal, so as the signal power increases, so does the interference from the multiplication. Phase noise from the TX is a very strong RX desensitization mechanism for the canceller system because the effective RX band noise due to TX phase noise increases dB for dB of TX output power. To give a concrete example, a phase noise level of -150dBc/Hz for the TX LO when transmitting at $+20\text{dBm}$ power gives -140dBm/Hz at the RX input, corresponding to a 34dB noise figure. In this cancellation system, however, phase noise on the TX can be cancelled in the same way as the main TX signal by the cancellation DAC if the DAC and TX LOs have identical phase noise profiles [47]. An in-depth analysis of the conditions and limitations for this phase noise cancellation effect is given in [48], but at a high level, the design decisions are as follows.

First, the bandwidth of the phase noise cancellation is set by the bandwidth over which the leakage network has relatively similar amplitude and phase shift as the main tone experiences, illustrated in Fig. 2.11. In this cancellation architecture, the leakage channel between TX and RX is tightly controlled because both are closely spaced to one another. Furthermore, reflections from the environment negligibly affect the phase noise cancellation, since any reflection off the environment will have a significant attenuation associated with it. A

useful property of this cancellation regime is that any impedances in parallel with the RX input, for example an N-path filter used to attenuate out of band blockers, does not affect the phase noise or signal transfer function because the DAC creates a virtual short.

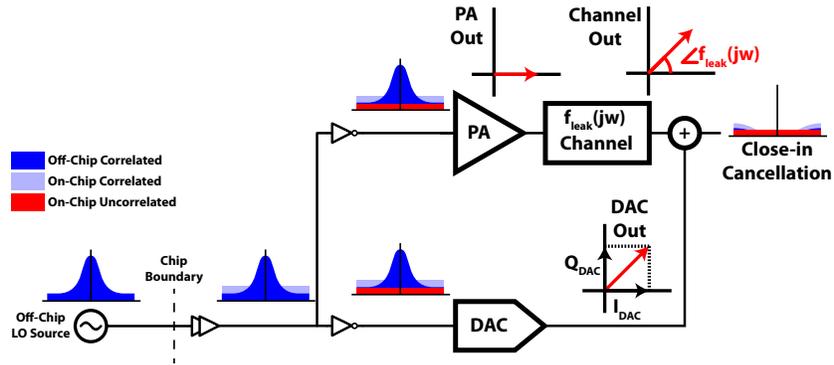


Figure 2.11: Feedforward cancellation of TX/DAC shared phase noise.

Secondly, the number of unshared buffers between TX and DAC must be minimized to produce a DAC LO highly correlated with the TX LO. This necessitates the use of a Cartesian PA and cancellation DAC because a polar architecture would require separate phase interpolators, leading to significant uncorrelated phase noise. In Fig. 2.12, the phase noise of a representative phase interpolator is shown, which would lead to a 30dB noise figure for +20dBm TX output power.

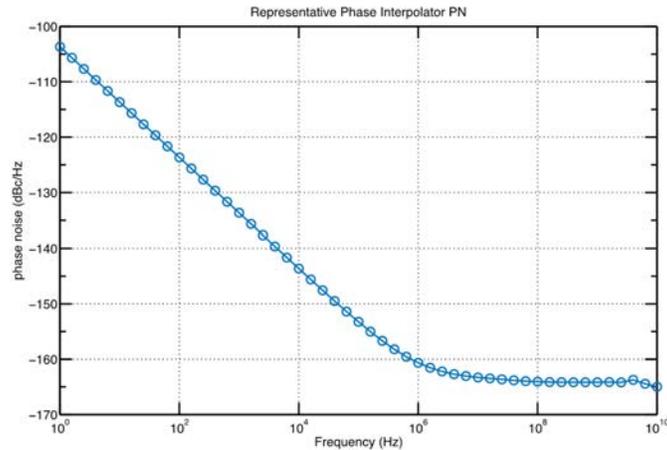


Figure 2.12: Phase interpolator phase noise.

2.4.4 DAC Thermal Noise

The DAC output is directly connected to the differential input of the RX, and its effect on the RX noise figure can be derived by analyzing Fig. 2.13. Here, the TX is assumed to be

zero output impedance, and therefore contributes zero thermal noise power. Due to the high output impedance of the DAC, its noise is represented as a current source. Additional noise sources included are the input-referred current noise of the RX and the current noise due to the antenna. Transferring all noise sources to the RX side, the total noise figure in terms of DAC noise and F_{RX} , the receiver noise figure, may be computed as:

$$F = 1 + N^2 \frac{\overline{i_{n,DAC}^2} + \overline{i_{n,RX}^2}}{\overline{i_{n,Ant}^2}} \quad (2.11)$$

$$= F_{RX} + N^2 \frac{\overline{i_{n,DAC}^2}}{\overline{i_{n,Ant}^2}} \quad (2.12)$$

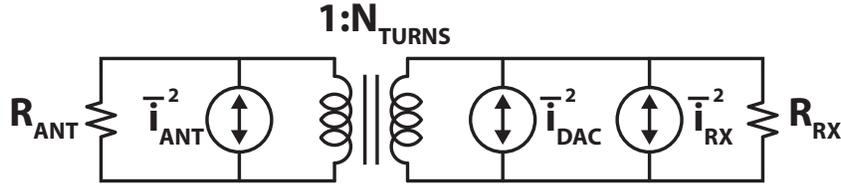


Figure 2.13: DAC noise network.

It is clear from this equation that the DAC current noise adds directly to the RX noise figure. Therefore, it is important to accurately model this noise current. A general model for the DAC output noise current as a function of the TX leakage signal can be created by considering the DAC as a noisy tail device connected to a noiseless mixer driven by a square wave. The vast majority of RF current DACs conform to this model due to their construction as a tail transistor with hard driven switches.

A DAC unit cell's noise is the thermal noise of the tail device, mixed by the noiseless mixer. This mixer has a conversion gain of A_{Conv} , and the tail thermal noise, in terms of the tail transconductance g_m , noise factor γ , and overdrive voltage v_{Ov} , can be written as:

$$\frac{\overline{i_{n,Tail}^2}}{\Delta f} = 4kT\gamma g_m \quad (2.13)$$

$$= 4kT\gamma \frac{2I_{Tail,Unit}}{v_{Ov}} \quad (2.14)$$

$$\frac{\overline{i_{n,Unit}^2}}{\Delta f} = 4kT\gamma \frac{2I_{Tail,Unit}}{v_{Ov}} A_{Conv}^2 \quad (2.15)$$

Each unit cell possesses a separate tail device, therefore all DAC noise sources are uncorrelated with one another, leading them to add together in variance. If the assumption that the DAC is created with all thermometer cells, then overall analysis can be simplified

considerably. The cases of polar and Cartesian DACs (with and without I/Q cell sharing) will be shown separately. Starting with the polar DAC, the phase of the output is controlled by the phase of the LO. The output amplitude of a unit cell is independent of the desired phase, and therefore $A_{Conv} \neq f(\phi_{Leak})$, where ϕ_{Leak} is the phase of the TX leakage at the input of the RX. As stated earlier, $I_{TX} = \sqrt{2 \frac{P_{TX}}{R_{Ant}} \frac{1}{N_{Turns}}}$, and $I_{Tail,Total} = \frac{I_{TX}}{2A_{Conv}}$, where the factor of 2 is due to the fact that I_{TX} is the sinusoidal amplitude of the TX tone. The unit cells are all the same phase, so they add in current: $I_{Tail,Total} = n_{enabled} I_{Tail,Unit}$, where n is the number of unit cells enabled. The total output noise can then be found as a function of P_{TX} :

$$\frac{\bar{i}_{n,Total}^2}{\Delta f} = 8kT\gamma \frac{n_{enabled} I_{Tail,Unit}}{v_{Ov}} A_{Conv}^2 \quad (2.16)$$

$$= 8kT\gamma \frac{\sqrt{2 \frac{P_{TX}}{R_{Ant}}}}{N_{Turns} 2A_{Conv} v_{Ov}} A_{Conv}^2 \quad (2.17)$$

$$= 4kT\gamma \frac{\sqrt{2 \frac{P_{TX}}{R_{Ant}}}}{N_{Turns} v_{Ov}} A_{Conv} \quad (2.18)$$

Using Eq. (2.12), the noise figure due to the thermal noise of the DAC only is equal to

$$F = 1 + \gamma \sqrt{2P_{TX} R_{Ant}} \frac{N_{Turns} A_{Conv}}{v_{Ov}} \quad (2.19)$$

Before continuing to the more involved case of the Cartesian DAC, it is worth noting an important trade off between noise and power consumption in the replica DAC which is identical for all architectures shown here. For a fixed process, antenna impedance, and TX power level, there are 3 design variables which affect noise figure: N_{Turns} , A_{Conv} , and v_{Ov} . DAC current consumption is inversely proportional to N_{Turns} and A_{Conv} , while the DAC supply voltage is roughly proportional to v_{Ov} for a fixed output impedance requirement.

In the case of a Cartesian DAC with I/Q cell-sharing, where each unit cell can output either I , Q , or both phases through modulation of the unit cell LO, shown in Fig. 2.14, the situation is complicated by the fact that in a unit cell, $A_{Conv} = f(\phi_{TX})$. This is because a code of $(1, 1)$ will create a 50% duty-cycle square wave, whereas $(1, 0)$ will create a 25% square wave. These two square waves modulate the same tail current, meaning that unlike the Cartesian case without cell sharing, the noise from I and noise from Q cannot be considered separately, and their segmentation matters. Define $n_{enabled,I}$ and $n_{enabled,Q}$ as the number of unit cells with I and Q phases enabled, respectively. Without loss of generality, take $|n_{enabled,I}| \leq |n_{enabled,Q}|$. In an all-thermometer DAC, $n_{enabled,I}$ cells will have 50% duty-cycle waveforms, while $n_{enabled,Q} - n_{enabled,I}$ cells will have 25%. A_{Conv} for the 50% case is $\sqrt{2}$ higher than the 25% case. Adding the noise of all these cells together:

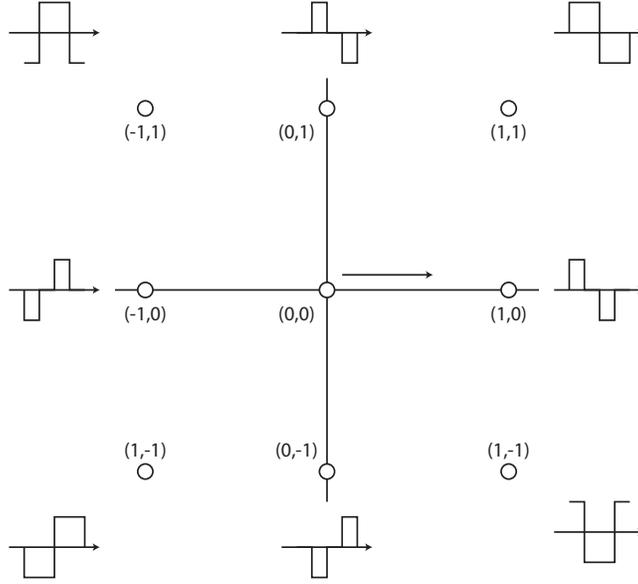


Figure 2.14: Unit cell output states.

$$\frac{\overline{i_{n,Total}^2}}{\Delta f} = 8kT\gamma \frac{n_{enabled,I} I_{Tail,Unit}}{v_{Ov}} \left(\sqrt{2} A_{Conv,25} \right)^2 \quad (2.20)$$

$$+ 8kT\gamma \frac{(n_{enabled,Q} - n_{enabled,I}) I_{Tail,Unit}}{v_{Ov}} A_{Conv,25}^2 \quad (2.21)$$

$$= 8kT\gamma \frac{A_{Conv,25}^2 I_{Tail,Unit}}{v_{Ov}} (n_{enabled,I} + (n_{enabled,Q} - n_{enabled,I})) \quad (2.22)$$

$$= 8kT\gamma \frac{A_{Conv,25}^2 I_{Tail,Unit}}{v_{Ov}} (n_{enabled,I} + n_{enabled,Q}) \quad (2.23)$$

$$= 8kT\gamma \frac{A_{Conv,25}}{v_{Ov}} A_{Conv,25} I_{Tail,Unit} (|I| + |Q|) \quad (2.24)$$

Using similar logic to the polar case,

$$I_{Tail,Unit} |I + jQ| 2A_{Conv,25} = \frac{1}{N_{Turns}} \sqrt{2 \frac{P_{TX}}{R_{Ant}}} \quad (2.25)$$

Combining Eq. (2.25) with Eq. (2.24) the total DAC noise can be written in terms of the TX signal power and phase:

$$\frac{\overline{i_{n,Total}^2}}{\Delta f} = 4kT\gamma \frac{A_{Conv,25}}{v_{Ov}} \frac{1}{N_{Turns}} \sqrt{2 \frac{P_{TX}}{R_{Ant}}} (|\cos \phi_{TX}| + |\sin \phi_{TX}|) \quad (2.26)$$

Noise from higher order harmonics will add further noise, but because the conversion gain of a square wave drops of as $\frac{1}{N_{\text{Harmonic}}}$, higher frequency noise contributes $<1\text{dB}$ to the total noise figure relative to the fundamental. However, noise which is downconverted from $2F_{TX}$ will still have the same conversion gain as upconverted baseband noise. An impedance resonating at $2F_{TX}$ is used to degenerate the source of the tail, rejecting the $2F_{TX}$ noise. A simulation of NF with a noiseless TX/RX, taking into account all harmonics but $2F_{TX}$, is shown in Fig. 2.15.

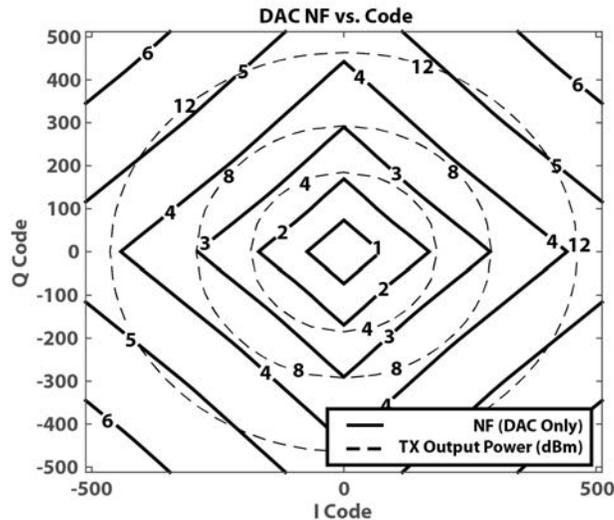


Figure 2.15: DAC foise figure versus TX power.

Segmentation of the current DAC has a strong influence on the thermal noise contour. In fact, if one considers a binary-only segmentation of the DAC, the noise figure can be decreased, though at the cost of a large amount of power. Both the increased power and decreased noise figure are intrinsic to the design of the current DAC binary unit cells. Binary unit cells have the same tail current as the thermometer cells, but differing fractions of this tail current are brought to the differential RX input versus shunted to the center tap. This partitions the current as well as the current noise, such that in an all-binary implementation, the current variance for code (2, 2) is 4x the current variance for code (1, 1). This is contrasted with thermometer cells, which are all identical and add in variance, where the current variance for code (2, 2) is 2x the current variance for code (1, 1). This effect is illustrated in Fig. 2.16.

The theoretical noise figures due to fundamental only for all-binary and all-thermometer segmentation are shown in Fig. 2.17. At high codes, the noise figure can be reduced by nearly 2dB for an all-binary implementation. The step increases in noise figure for the binary contour are due to the fact that for $2^n - 1$, many cells add in variance, but for 2^n , there is a single cell outputting a noisy current.

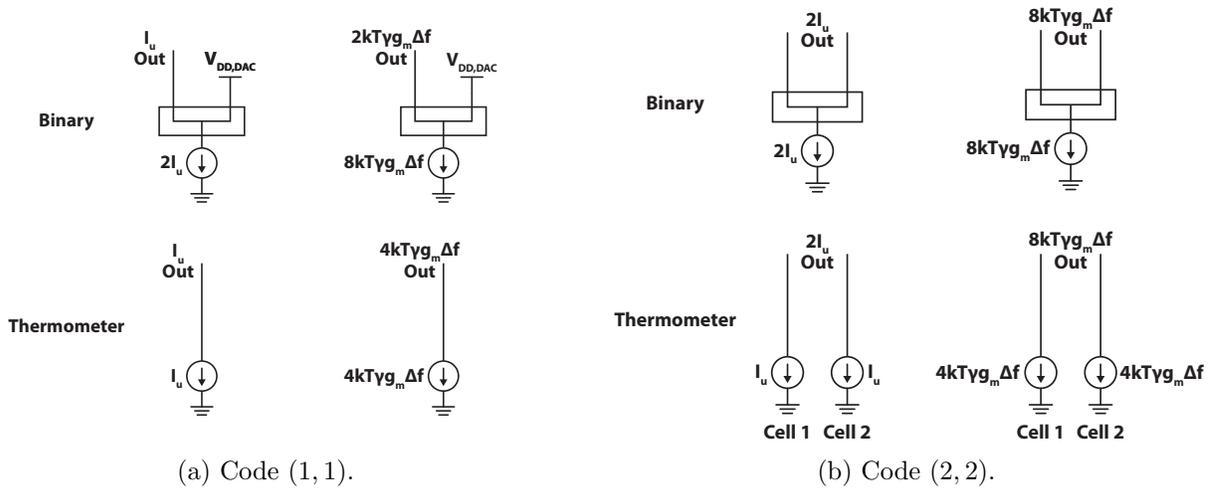


Figure 2.16: Binary and thermometer output current and associated noise.

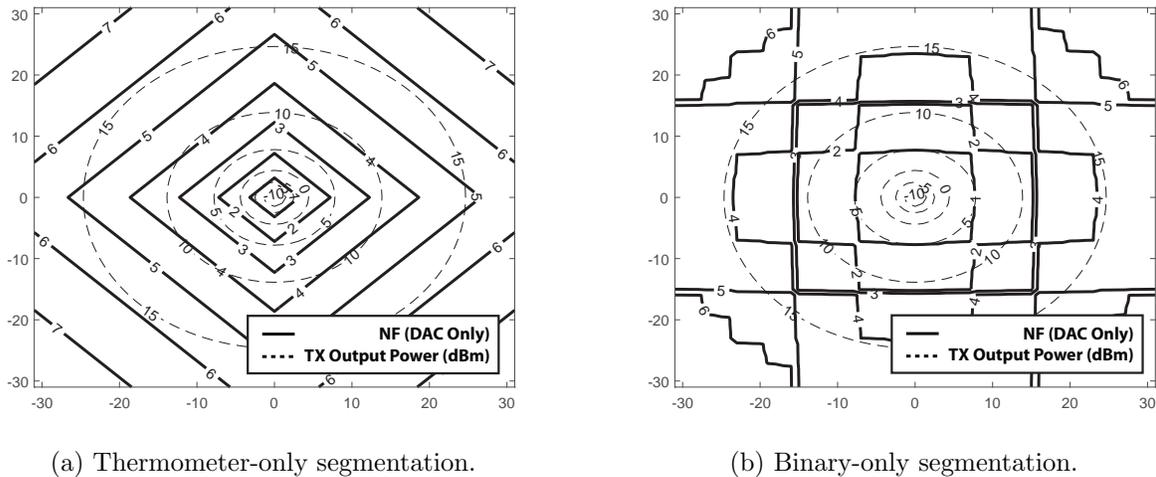


Figure 2.17: Thermal noise contours vs. TX power contours.

Segmenting the entire array as binary does have a significant power penalty, in that each of the B cells draw approximately half the required tail current from the supply, while a thermometer-only array draws exactly the required current from the supply, leading to a $B/2$ increase in power draw for the binary-only segmentation.

2.4.5 Noise Summary

Putting each of these sources of noise figure degradation together, a more complete picture of the effect of this duplexing network on the RX can be shown. Again, there are 4 sources of noise figure degradation in this system: RX transformer loss, DAC thermal noise, TX phase noise, and loss due to the series TX/RX connection. Rather than being fundamentally limited, these degradation mechanisms are highly implementation dependent. For example, the process option of ultra-thick metals can significantly reduce RX transformer loss, and more advanced process nodes can lower the output impedance of a switching PA due to the lowered R_{on} of the switching devices. The replica DAC power budget constrains both the headroom available for the DAC (thereby placing an upper-bound on the tail device v_{Ov}) and the minimum allowable RX turns ratio. Finally, improved co-location of the TX and DAC, reducing the length of the independent TX and DAC buffer chains will reduce the uncorrelated phase noise between the TX and DAC, significantly reducing noise figure degradation in high TX power regimes.

Illustrating the heavy dependence on implementation choice, Fig. 2.18 presents three scenarios, where the first is an implementation with moderate RX noise figure, TX phase noise, and DAC overdrive, the second has improvements to nominal RX noise figure, and the third is an aggressive design for minimizing practical noise figure, where uncorrelated phase noise is improved, as well as TX output impedance, DAC overdrive, and RX winding loss.

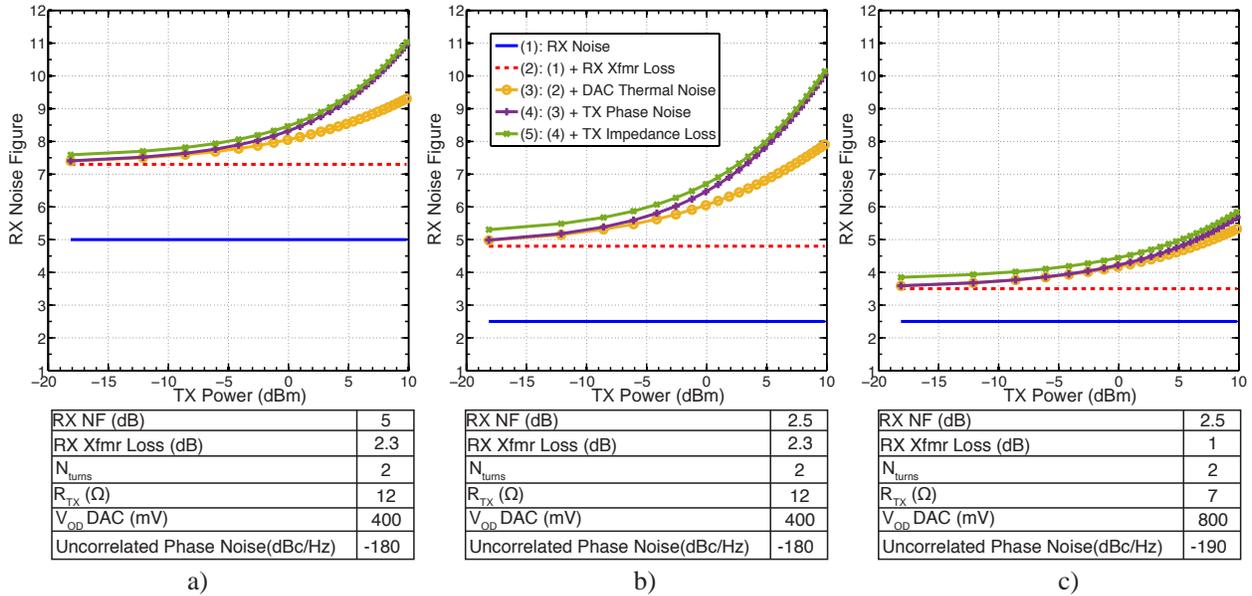


Figure 2.18: Noise figure with nonidealities added.

2.5 TX Insertion Loss From RX

The series connection between TX and RX reduces transmitter efficiency because resistive elements in the series balun dissipate power normally incident on the antenna. The cancellation current source creates a virtual short at the RX input port, but due to finite transformer Q , this short is not directly transferred to the antenna side. The winding resistances of both the RX side and the antenna side loops set the real part of the antenna-side RX impedance R_{RX} . The antenna-side winding resistance appears directly in series with the TX, while the RX-side resistance is transformed to the antenna side through inductive coupling.

Inductive coupling is a physical property of any two current-carrying loops. A magnetic field is generated by current flowing through Loop 1, and the total field captured by Loop 2 is called the magnetic flux. When this magnetic flux changes (e.g. the current in Loop 1 changes), this change in magnetic flux induces an electromotive force (EMF) on the charges in Loop 2, which serve to counteract the flux change. The amount of EMF generated, normalized by the amplitude and frequency of the current through Loop 1 is called the mutual inductance M_{12} . Interestingly $M_{12} = M_{21} = M$, regardless of the shape, size, or separation of the two loops. This definition of mutual inductance can be applied to the same loop, where it is called self inductance L . A network consisting of two loops with self-inductances L_1 and L_2 and a mutual inductance M can be described using the Z-Parameter matrix Eq. (2.27).

$$\begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = \begin{bmatrix} j\omega L_1 & j\omega M \\ j\omega M & j\omega L_2 \end{bmatrix} \begin{bmatrix} I_1 \\ I_2 \end{bmatrix} \quad (2.27)$$

If mutual inductance is compared with the self-inductances, it is seen that $|M| \leq \sqrt{L_1 L_2}$ always. Intuitively, this is clear because mutual inductance serves to cancel the change in magnetic flux, so the best that the mutual inductance can do is cancel out the magnetic flux exactly, where $M = \sqrt{L_1 L_2}$. M can therefore be rewritten in terms of self-inductance as $k\sqrt{L_1 L_2} = kL_1\sqrt{n}$, where n is the turns ratio of the inductors and $|k| \leq 1$ is known as the coupling coefficient. The sign of k is determined by the direction of the windings of L_1 and L_2 . $|k| = 1$ is the case of an ideal transformer, where the same amount of magnetic flux captured by the first loop is captured by the second loop. With this formulation, the Z parameter network can be reinterpreted as the circuit shown in Fig. 2.19.

Valid to first order, the transformer losses are represented as two series resistors, one at each transformer port, which defines Q for each winding as $Q = \frac{\omega L}{R}$. The network for a finite Q transformer network with one side shorted (representative of the RX balun network, from the perspective of the TX signal) is shown in Fig. 2.20. The impedance seen on the non-shortened port, Z_{RX} is equal to:

$$Z_{RX} = \frac{R_p R_s - \omega^2 L_p L_s (1 - k^2) + j\omega (R_s L_p + R_p L_s)}{R_s + j\omega L_s} \quad (2.28)$$

A well-designed on-chip transformer may have a coupling factor $|k| \approx 0.9$, simplifying the real part of Z_{RX} :

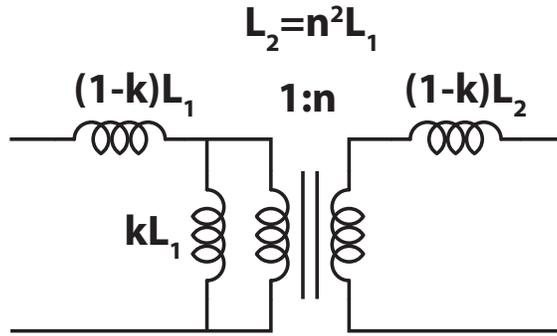


Figure 2.19: Circuit with mutually coupled inductors.

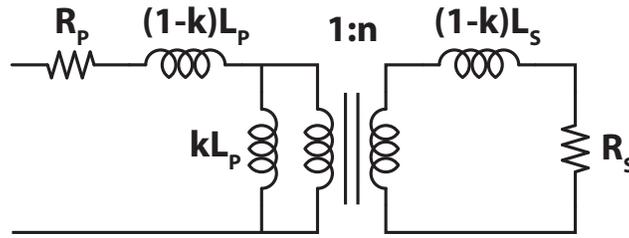


Figure 2.20: RX transformer with input port shorted.

$$\text{Re}(Z_{RX}) \approx R_p + \frac{1}{n^2} \frac{R_s}{1 + Q_s^2} \quad (2.29)$$

To quantify the TX efficiency loss due to the series Z_{RX} , the TX transformer efficiency G_P is derived for an optimized network, then this is compared to G_P for the same network with the addition of Z_{RX} . The model Fig. 2.21 and the optimization formulae from [49] are used to optimize G_P for a given assumed coupling coefficient and winding Q .

The design equations for this model are as follows:

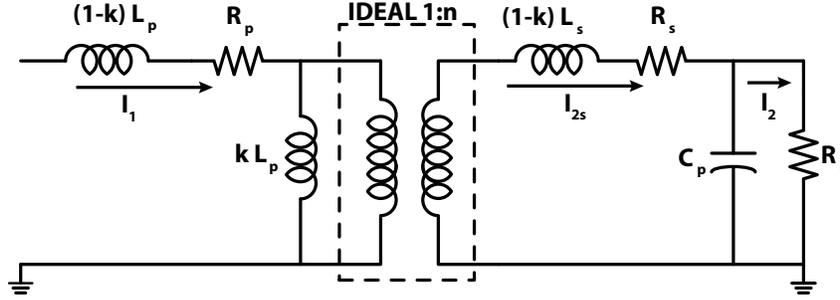


Figure 2.21: Simplified transformer model [49].

$$C_p = \frac{1}{2}C_{eq} \left(1 \pm \sqrt{1 - \frac{4}{(\omega R_L C_{eq})^2}} \right) \quad (2.30)$$

$$C_{eq} = \frac{1}{\omega^2 L_s} \quad (2.31)$$

$$L_s = n^2 L_p \quad (2.32)$$

$$L_p = \frac{1}{\omega} \frac{\alpha}{1 + \alpha^2} \frac{R_L}{n^2} \quad (2.33)$$

$$\alpha = \frac{1}{\sqrt{\frac{1}{Q_s^2} + \frac{Q_p}{Q_s} k^2}} \quad (2.34)$$

Along the same lines as [49], the expression for G_p can be simplified as follows:

$$G_p = \frac{P_{Ant}}{P_{Ant} + P_{Dissipated}} \quad (2.35)$$

$$= \frac{R_{eq}}{(R_{eq} + R_s + R_{RX}) + n^2 R_p \left| \frac{Z_s + ksL_p}{ksL_p} \right|^2} \quad (2.36)$$

with,

$$Z_s = s(1-k)L_p + \frac{1}{n^2}R_s + Z_{RX} + \frac{R_{Ant}}{1 + sR_{Ant}C_p} \quad (2.37)$$

Fig. 2.22 compares G_p with and without Z_{RX} for 1:2 TX and RX baluns with primary inductances of 550pH and 2nH, respectively. Both have quality factors of approximately 8 at 1.5GHz. As can be seen, there is a large frequency range over which the losses are less than 1dB. It should be noted that this Fig. 2.22 represents a somewhat pessimistic view of the efficiency losses due to the series RX balun because this assumes that there are no off-chip routing losses, which offset the relative difference in efficiency.

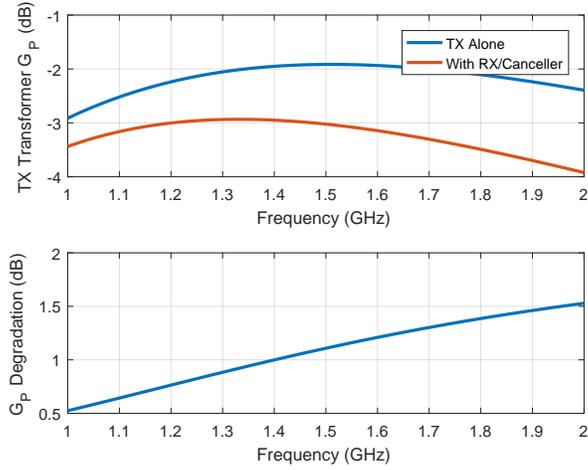


Figure 2.22: TX transformer performance degradation with RX inclusion.

2.6 TX Efficiency Degradation Mechanisms

Because this work creates a novel interface between the TX and RX, any effective reduction in TX efficiency must also be considered. Here, two main sources of degradation to the effective TX efficiency exist: RX winding loss and the power consumption of the canceller power and digital predistortion (DPD).

In the previous section, TX insertion loss due to the RX balun nonidealities was detailed. The second form of efficiency degradation is due to the increased system power consumption due to the power draw of the canceller as well as that of the DPD and filtering schemes which are used to improve cancellation of the TX signal at the RX input. A conservative estimate of requirements for the digital filter is 8 taps for 200MS/s with 10 bit coefficients. According to [50], the power consumption for this filter is 10mW in a 65nm process. Digital predistortion is achieved through a lookup table, estimated to cost 15mW in power. There is additional power consumed to run the adaptation algorithms to change these filters and lookup tables as the network or other nonidealities change, but because the dynamics of the channel are far slower than the data rate, the power consumption of these digital algorithms can be amortized over a very large operation time, making their average power consumption negligible. Additionally, the power consumption of the DAC is a dominant source of power consumption. As stated before, the current required from the DAC supply is proportional to $\sqrt{P_{TX}}$, though power at backoff for modulated data depends also on the type of backoff available for the DAC. In Fig. 3.17a, degradation of the TX efficiency due to these mechanisms, for a modulated data signal with 6dB PAPR, is plotted with class-A and class-B DAC backoff.

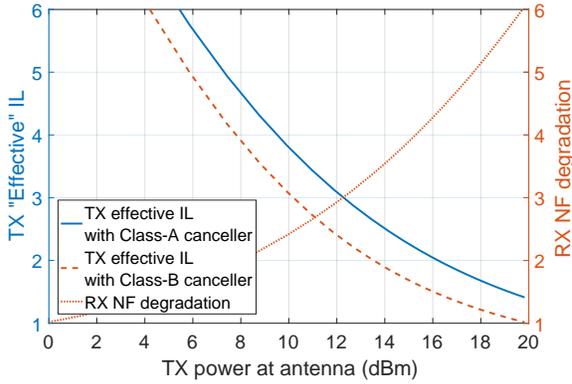
Combining the effects of TX efficiency degradation and RX noise figure degradation, a fair comparison can be made with the hybrid technique. RX noise figure degradation can be cleanly converted to RX insertion loss, and TX efficiency degradation to effective TX insertion loss. In practical hybrids, TX and RX insertion losses are typically both 4dB.

TX loss parameters.

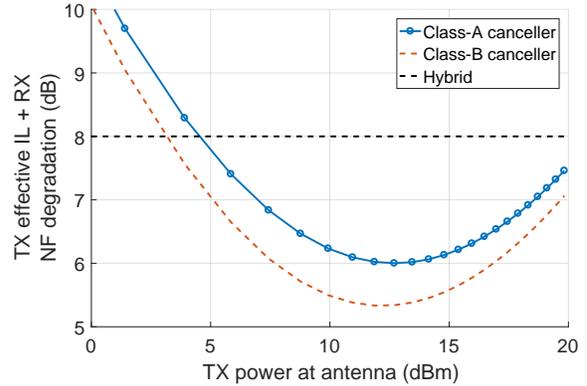
TX Average Backoff (dB)	6
TX Average PAE	25%
Digital Filter Power (mW)	10
Canceller DPD Power (mW)	15
IL From RX Winding (dB)	-0.35
DAC Supply Voltage (V)	1

RX NF Parameters.

RX NF (dB)	2.5
RX XFMR IL (dB)	1
RX XFMR N_{Turns}	2
$R_{TX}(\omega)$	7
DAC V_{ov} (mV)	800
Uncorrelated Phase Noise (dBc/Hz)	-190



(a) TX/RX degradation versus output power.



(b) Total degradation comparison with hybrid.

In Fig. 3.17b, the combined effective TX and RX insertion losses for the aggressive design point specified in Fig. 2.18 are combined and compared with an 8dB combined loss from the hybrid. As shown, from approximately +5dBm to above +20dBm, the replica DAC gives better performance than the hybrid. Additionally, the tunability of the replica canceller over a wide range of antenna VSWR makes this a more attractive solution than the hybrid even for power levels where the effective combined insertion loss is close to or exceeds that of the hybrid.

2.7 DAC Linearity

While it may first appear that the cancellation DAC must be highly linear in order to provide large TX cancellation, it is actually found that as long as the DAC has a sufficient number of bits, its required effective number of bits (ENOB) is far lower.

Three metrics for nonlinearity within an RF DAC are quadrature angle mismatch, and I/Q summation angle and magnitude mismatch. Quadrature angle mismatch is simply the angle between codes (1, 0) and (0, 1) compared with 90° , shown in Fig. 2.23a. I/Q magnitude and angle mismatch, shown in Fig. 2.23b, affect how well I and Q add together. Because of the I/Q cell-sharing technique implemented in the cancellation DAC, where a single active unit cell can output a 9QAM signal shown in Fig. 2.14, duty cycle distortion between the 25% and 50% signals in a unit cell can lead to (1, 1) not equaling (1, 0) + (0, 1). It is worth

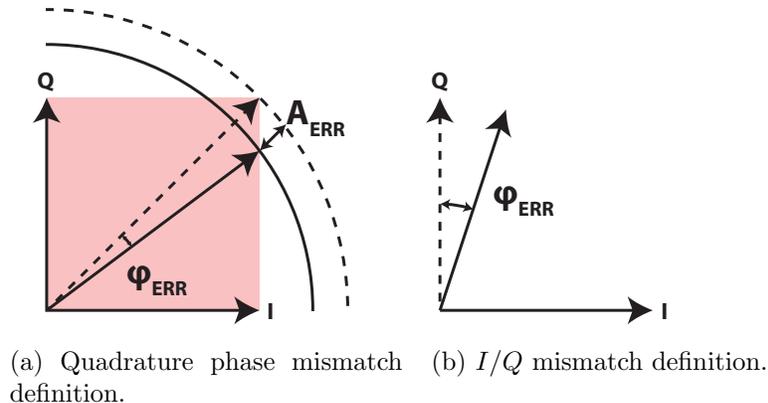


Figure 2.23: Definitions of DAC static mismatch.

noting that both of these nonlinearities are present only at the unit-cell level in the first version of the system because of the minimal swing across the output node of the DAC during cancellation.

To understand the effect that these nonlinear mechanisms have on the overall constellation, first consider quadrature angle mismatch. This is due to nonidealities within generation of I and Q phases of the LO. In this cancellation system, there is no amplitude mismatch between I and Q due to the symmetric way in which the 25% clock phases are generated. I/Q summation mismatch has both amplitude and phase components, and measures the error between $(C, 0) + (0, C)$ and (C, C) . I/Q summation mismatch is an important nonlinearity to characterize because it creates “holes” in the DAC constellation, shown in Fig. 2.24b. In other words, it unevenly distributes the spacing between DAC constellation points in a discontinuous way, compared with a “softer” nonlinearity like compression.

Spacing discontinuities arise from the fact that individual cells add very linearly, but unit cells possess I/Q summation nonlinearity. Consider the simple case of a 2-bit binary DAC with I/Q summation magnitude nonlinearity, shown in Fig. 2.24a. Note that the red dots are points which are unaffected by this unit cell nonlinearity. The points along the edge are intuitive because they are from unit cells with either only I enabled or only Q enabled, so I/Q summation nonlinearity is not a factor. The two in the middle of the constellation come from the fact that the I and Q components of the code have non-overlapping binary representations ($\text{AND}(C_I, C_Q) = 0$). Therefore, the linear summation of unit cells prevents these points from being affected by summation nonlinearity as well. Because some cells are affected by this nonlinearity, whereas others are not, and because ones which are affected are directly next to ones which are not affected, holes in the constellation develop. Additionally, as the number of binary bits increases, the size of the gaps remains constant, while the ideal LSB spacing decreases, leading to a limit in rejection in those regions.

Thermometer arrays do not have this same issue and distort the constellation in a some-

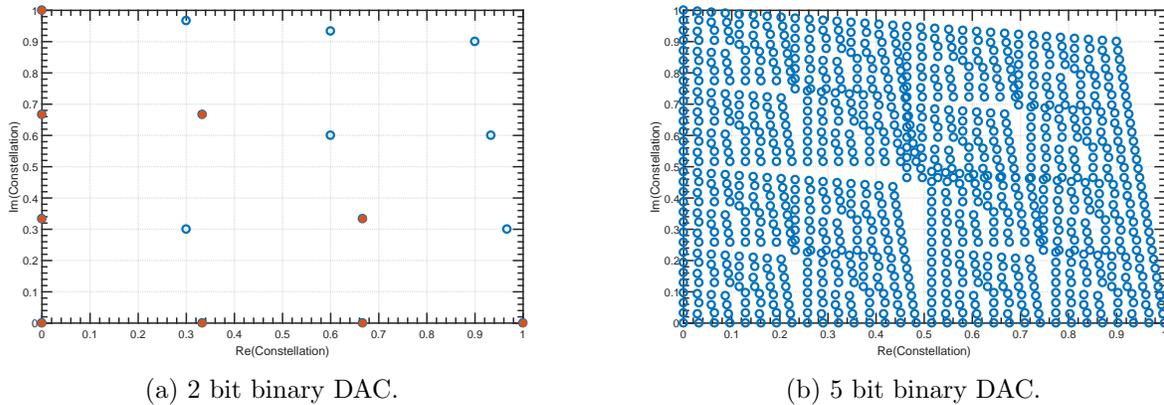


Figure 2.24: 2 and 5 bit binary constellations with 10% I/Q summation magnitude nonlinearity.

what continuous way. Consider the same 2 bit DAC, but with thermometer-only segmentation in Fig. 2.25a. Here, the only points which are unaffected by I/Q summation nonlinearity are those at the edges, due to the fact that unit cells with both I and Q enabled are up to $\min(C_I, C_Q)$, irrespective of binary representation. Therefore, it can be thought of as a continuum of nonlinearity expression, leading to no holes. The constellations with 5 bits in Figs. 2.24b and 2.25b showcase this difference.

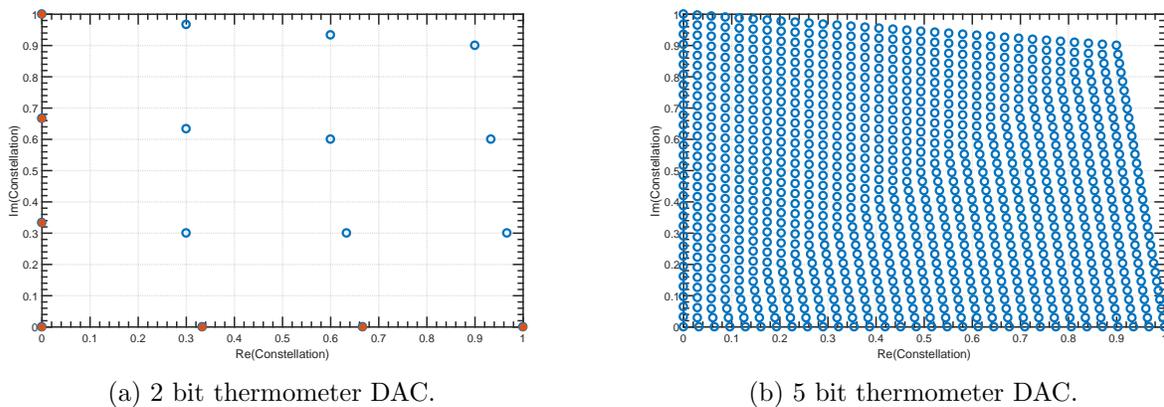


Figure 2.25: 2 and 5 bit thermometer constellations with 10% I/Q summation magnitude nonlinearity.

To quantify the effect that these constellation impairments have on the cancellation system when nonlinear predistortion is used, a series of simulations of rejection versus DAC bits and DAC nonlinearities were carried out, shown in Fig. 2.26. Here, DAC constellations were generated using the impairments from Fig. 2.23. Then, for each baseband equivalent TX

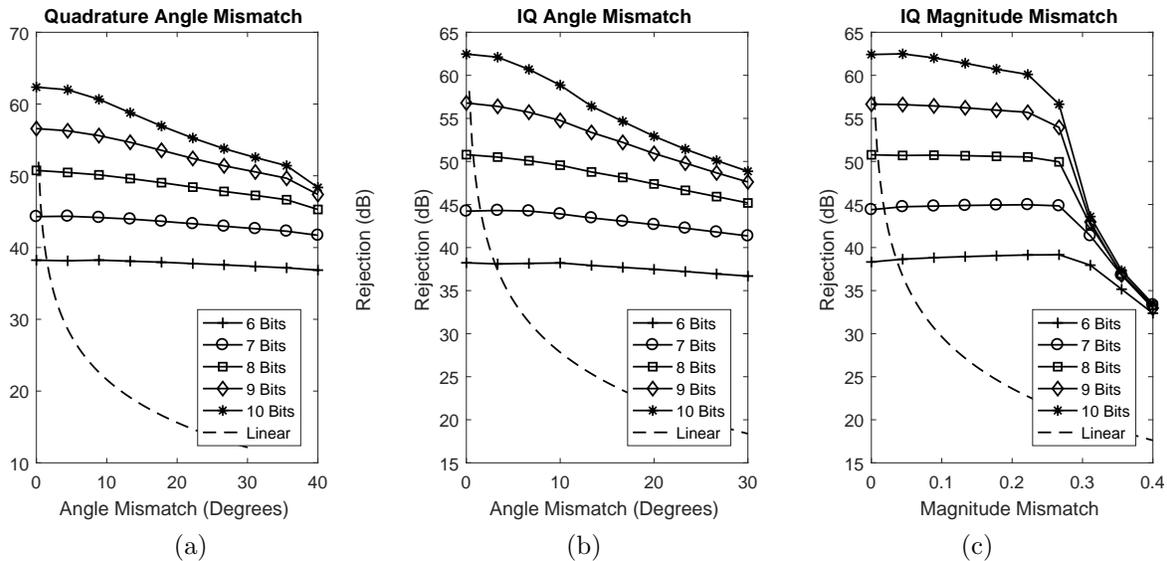


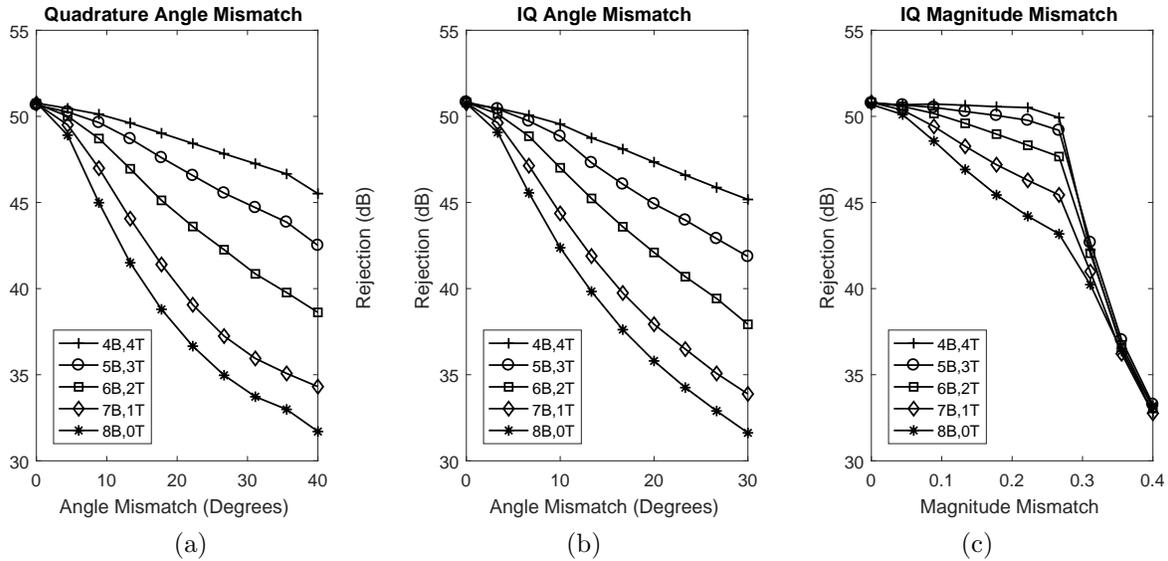
Figure 2.26: Rejection vs. DAC bits for I/Q summation nonlinearity.

leakage symbol in a sequence, the closest DAC constellation point was found and subtracted from the TX sequence. The residual was then compared with the full power of the TX leakage signal to determine rejection. Given this nonlinear predistortion, large distortions in the DAC constellation still produced high TX signal rejection.

An additional tradeoff exists between binary and thermometer segmentation, where DNL is worse and constellation holes are larger. As an example, Fig. 2.27 shows the results of a simulation where the number of bits was held constant, but the segmentation of binary vs. thermometer was modified while sweeping the magnitude of I/Q summation error. As the number of binary bits increases, the rejection falls for a fixed summation nonlinearity.

Comparing this error sequence to the error sequence using linear FIR adaptation, where the DAC signal is convolved an adapted set of filter taps, it is seen that the power of the error with predistortion is significantly below that of using a linear FIR. It can be seen that a 30% IQ magnitude mismatch and 30° I/Q and quadrature angle mismatch can be tolerated while still achieving >53 dB TX cancellation (taking into account the fact that these errors could add together). The higher constellation density afforded by increased physical bits lowers the sensitivity of system cancellation to DAC impairments.

Accordingly, there is a tradeoff between cancellation and replica DAC power consumption. There are two main ways of achieving higher cancellation, assuming the unit cells are not perfectly linear: reduce the cancellation sensitivity to impairment for a fixed unit cell linearity, or increase the unit cell linearity. Increasing the number of physical bits also incurs penalties once the DAC driver power consumption is routing dominated, a reasonable assumption for >8 bit arrays. Due to other analog impairments such as supply variation and unit cell mismatch, it is reasonable to target a higher number of bits with lowered unit


 Figure 2.27: Rejection vs. DAC segmentation for I/Q summation nonlinearity.

cell requirements for this system.

2.7.1 IQ Nonlinearity - Linear Adaptation

Confirming the relaxation of requirements that nonlinear predistortion offers, this section is dedicated to estimating the rejection vs. I/Q nonlinearity if only linear filtering is used. A general form of the equivalent complex baseband output for the DAC is $O_{Nonlinear} = I + jQ + AP(I, Q)$, where A is a complex-valued factor representing the I/Q summation nonlinearity and $P(I, Q)$ is a function used to count the portion of cells where $I = Q$. Because of how the TX and DAC are constructed, it is accurate to represent the summation nonlinearity with this constant. This function may depend on the total number of bits, segmentation, or other factors. For example, in the case of thermometer-only segmentation:

$$P(I, Q) = \min(I, Q) \quad (2.38)$$

and for binary-only segmentation:

$$P(I, Q) = \text{AND}(I, Q) \quad (2.39)$$

where AND is defined in a bitwise manner.

Assuming no channel impairments and a fully linear TX, DAC I/Q summation nonlinearity leads to a residual signal of $R = O_{Nonlinear} - (I + jQ) = AP(I, Q)$ without adaptation. Therefore, the mean error power is:

$$E(|R|^2) = |A|^2 E(|P(I, Q)|^2) \tag{2.40}$$

If instead, a single tap filter is used, the residual is now $R_{adapt} = CO_{Nonlinear} - (I + jQ)$. While this assumes no quantization of $CO_{Nonlinear}$, this accurately models a case where the number of bits are sufficient to make I/Q summation nonlinearity the dominant rejection-limiting mechanism. It can be shown that the minimum residual using this linear adaptation is equal to:

$$E(R_{min}^2) = E(R_{nom}^2) \left(1 - \frac{|E(P(I, Q)D)|^2}{E(|D|^2) E(P(I, Q)^2)} \right) \tag{2.41}$$

Here, $D = I + jQ$. Additionally it can be shown that, regardless of DAC segmentation, the difference between Eq. (2.40) and Eq. (2.41) is approximately 6dB. Displayed in Fig. 2.26 as the dotted lines, a very strict bound on I/Q summation nonlinearity for good cancellation is present in the case of linear adaptation, requiring very careful design of the DAC if nonlinear predistortion is not allowed.

2.8 I/Q Nonlinearity Due to Buffer Chain

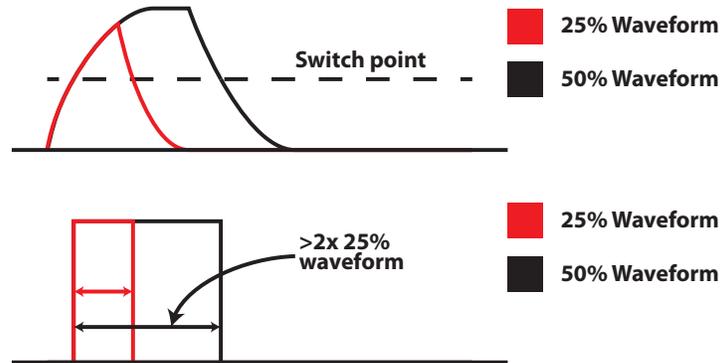


Figure 2.28: Buffer I/Q summation nonlinearity due to nonzero settling time.

One source of I/Q summation nonlinearity for the DAC and TX is the limited bandwidth of the LO buffer chain. Consider the use of a Cartesian, I/Q cell-sharing topology. This topology is sensitive to LO duty-cycle mismatch, where the energy and phase of a 25% LO, representing codes where $I = 0$ or $Q = 0$, differs from a 50% LO, representing codes where $I = Q$. Assuming the DAC added no duty-cycle distortion of its own, the limited bandwidth of the LO buffers leading to the DAC fundamentally causes mismatch. Consider an ideal buffer with switch point $V_{DD}/2$ and time constant τ . A 50% waveform has nearly half the LO period to settle to V_{DD} before the falling edge comes, while a 25% LO only has a

quarter period between rising and falling edges, leading to different settling behavior, shown in Fig. 2.28.

This issue is exacerbated by the number of buffers throughout the chain. Simulating a variable-length buffer chain where all buffers have the same τ , representing a constant fanout across stages, it is shown in Fig. 2.29 that achieving 50dB of cancellation using a linear FIR approach requires 6GHz of bandwidth for a chain of 4 buffers with an LO frequency of 2GHz.

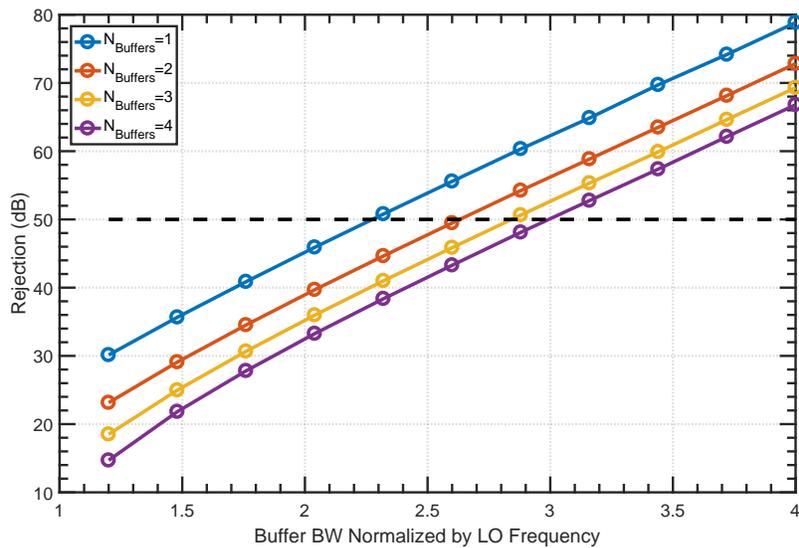


Figure 2.29: Rejection vs. buffer BW over buffer sizes (no predistortion).

As shown in the previous section, digital predistortion's significantly relaxed I/Q linearity requirements are 25% magnitude and 30° angle static mismatch. Shown in Fig. 2.30, 3GHz of bandwidth is required for 4 buffers at an LO frequency of 2GHz. In the case of linear adaptation without predistortion, a 3dB corner of 6GHz is required for 50dB cancellation. In 65nm technology, a fanout of 4 buffer chain has 15ps of 10-90% risetime, equating to a bandwidth of approximately 30GHz. Therefore, I/Q summation nonlinearity due to the buffer chain is not an issue.

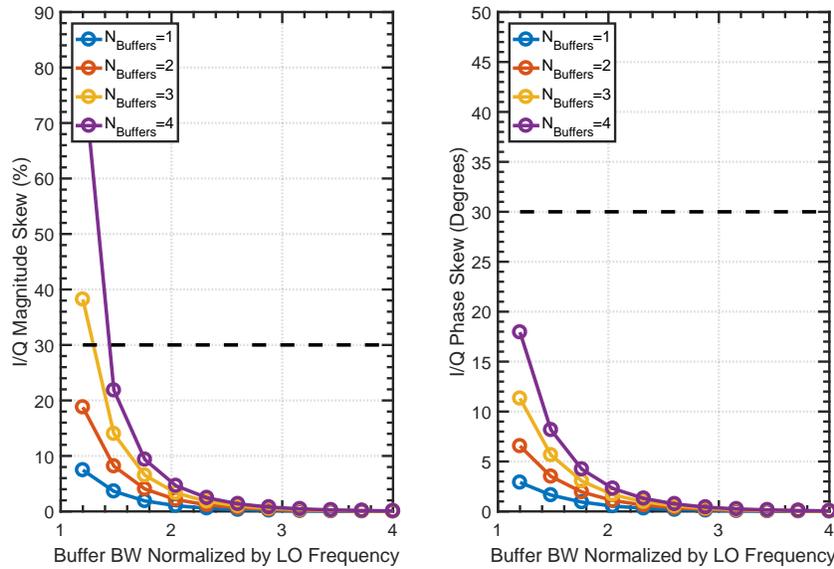


Figure 2.30: IQ nonlinearity vs. buffer BW over buffer sizes (digital predistortion).

2.9 Residual Sampling Methods

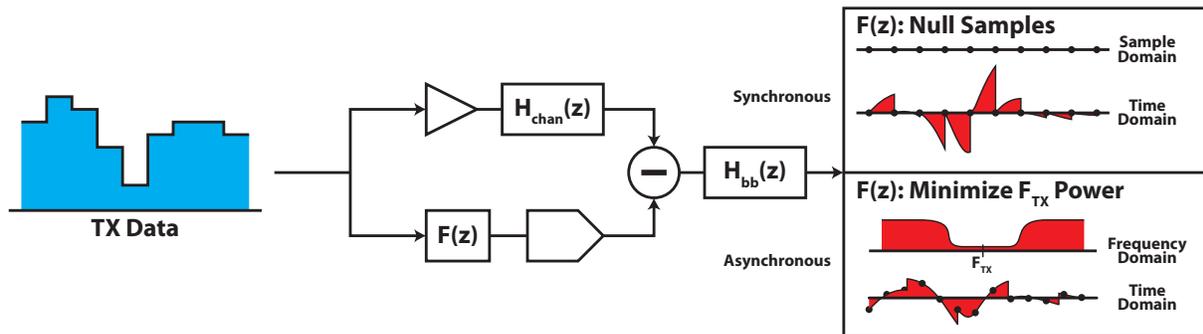


Figure 2.31: Sampling modes and their effect on the DAC filter.

The TX and DAC are both sampled systems, meaning that even if there were infinite bits of precision on the DAC and zero I/Q nonlinearity, the residual would not be zero over all time. Additionally, the RX is a sampled system as well, so considerations about the way in which the receiver samples the residual should be made to determine if there exists a best policy. In general, there are two policies for TX/DAC vs. RX sampling: RX clock synchronous with TX/DAC, and RX clock asynchronous with TX/DAC. Filtering methods are associated with these sampling types in order to minimize the error under some metric. At a high level, synchronous sampling allows the error to ideally be set to zero at every

sampling instance of the RX, but the imposes strict requirements on the anti-aliasing filter used in the RX baseband. Filtering with asynchronous sampling is best thought of in the frequency domain, where the interference in the TX band may be nulled completely, rather than ensuring zero sampled residual. Illustrating the difference between these sampling modes is Fig. 2.31

2.9.1 Synchronous Sampling

If the RX ADC is synchronized with the clock of the TX and DAC, then a very intuitive answer can be found for how to minimize the sampled error. If the DAC outputs the same current as the TX at the RX sampling instance, then the residual will be zero, shown in Fig. 2.31. The case of an anti-aliasing filter on the baseband chain is considered, along with a derivation of requirements for DAC digital filter stability.

In discrete time, with $H_{tot} = H_{chan}H_{bb}$, this relation holds:

$$H_{tot,p}(z) - F(z)H_{bb,p}(z) = 0 \quad (2.42)$$

$$F(z) = \frac{H_{tot,p}(z)}{H_{bb,p}(z)} \quad (2.43)$$

$$= \frac{H_{tot,u}(z)}{H_{bb,u}(z)} \quad (2.44)$$

Here, $H_{.,u}$ and $H_{.,p}$ are unit step and pulse responses, respectively. The last line follows from the line above because a pulse response is identical to a delayed unit step response subtracted from a normal unit step response, and this factor is applied to both numerator and denominator.

A filter $F(z)$ can always be found to satisfy these conditions, but it may be the case that this filter is not stable. $H_{tot,u}(z)$ is stable on its own because it is a sampled version of $H_{tot,u}(s)$, which is stable, so has poles with $\text{Re}(s_0) < 0$, which corresponds to $|z_0| < 1$. Therefore, it is the zeros of $H_{bb,u}(z)$ which will determine the stability of $F(z)$. This is convenient because this shows that the synchronous sampling filter stability is independent from the channel transfer function. Here, different anti-aliasing filter transfer functions, Butterworth and Chebyshev Type 1 of varying filter orders, are shown and evaluated based on their stability.

Because $F(z)$ is a discrete-time filter which must mimic the time-domain response of the channel, the unit step responses of H_{tot} and H_{bb} must be transformed to the Z-domain through impulse invariance. This is a technique for deriving discrete-time filters from continuous-time by sampling the response. In the frequency domain, partial fraction expansion is performed and the individual responses are transformed accordingly. A single pole at $s = s_0$ becomes a pole at $z = e^{s_0}$. Because the pole locations are moved in this partial fraction expansion, the zeros of the Z-domain transfer function are not related to the S-domain in a straightforward manner and must be attacked on a case-by-case basis.

For proper anti-aliasing and blocker rejection for the received signal, it is assumed that the baseband filter should have $>40\text{dB}$ of rejection at a channel spacing away. For the cases of Butterworth and Chebyshev Type I filters, this presents a problem, as the discrete-time versions of these filters possess zeros which are outside of the unit circle for any filter order higher than second, as shown in Fig. 2.32. This means that $F(z)$ will be unstable due to having poles outside of the unit circle. As the sampling frequency is increased for a given bandwidth, the zeros stay inside the unit circle or outside, without ever crossing the boundary, meaning stability of $F(z)$ is not a function of the filter oversampling ratio. For all oversampling ratios, any filter order higher than second will produce an unstable $F(z)$. A second-order filter is inadequate for 40dB of rejection, so it is not practical to attempt to zero the residual synchronously with the sampling clock. Therefore, asynchronous filtering options should be explored.

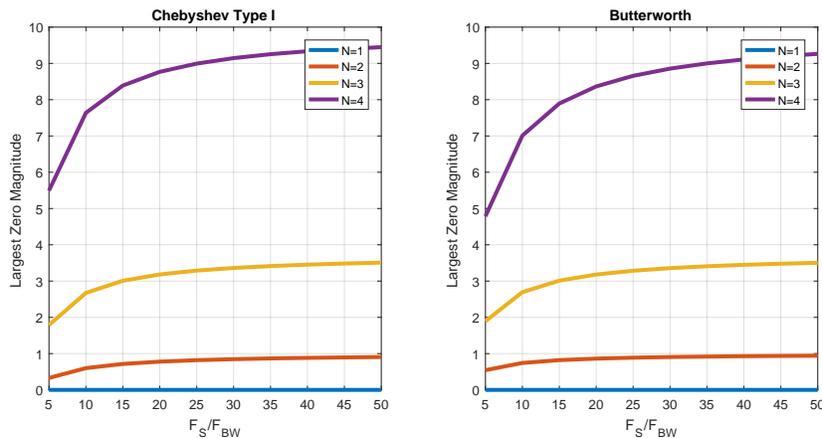


Figure 2.32: Maximum zero magnitude for baseband Butterworth and Chebyshev Type I filters.

2.9.2 Asynchronous Sampling

While synchronous sampling may seem to provide the best bound on residual, where the each sample is no greater in magnitude than the DAC LSB, asynchronous sampling can also provide similar levels of rejection in the TX spectrum. Two methods are shown here, the first is a causal system, while the second is a non-causal filter taking advantage of knowing the TX data beforehand.

Asynchronous Sampling - Limited Taps, Causal

First, the case of limited filter taps with the requirement of causality. The error will also be minimized across the full spectrum. A few definitions need to be put in place first to start

analysis. Define the TX current signal at the RX input as O_{TX} , the DAC current signal as O_{DAC} and the mean squared error as P_{err} .

$$O_{TX}(t) = \sum_{m=-\infty}^{\infty} d[m]h_p(t - mT) \quad (2.45)$$

$$O_{DAC}(t) = \sum_{q=-\infty}^{\infty} \sum_{k=0}^{L-1} c_k d[q - k]p(t - qT) \quad (2.46)$$

$$P_{err} = E \left(\frac{1}{T} \int_0^T (O_{TX}(t) - O_{DAC}(t))^2 dt \right) \quad (2.47)$$

Here, $h_p(t)$ is the pulse response of the TX leakage channel, $p(t)$ is the pulse response of the DAC signal (a pulse of width T , the symbol period, d is the TX data, and c_k are the DAC FIR taps. The expectation of $P_{err,TX} = O_{TX}^2$ is fixed, but the expectations of $P_{err,DAC} = O_{DAC}^2$ and $P_{err,TX,DAC} = O_{DAC}O_{TX}$ can be minimized by the correct choice of c_k . These expectations, $P_{err,DAC}$ and $P_{err,TX,DAC}$, can be simplified into discrete time versions:

$$P_{err,DAC} = \sum_{k=0}^{L-1} \sum_{l=0}^{L-1} c_k c_l R[k - l] \quad (2.48)$$

$$P_{err,TX,DAC} = \sum_{k=0}^{L-1} \sum_{m=0}^{\infty} c_k R[m - k] A_m \quad (2.49)$$

where $R[n]$ is the autocorrelation function of the TX data and where:

$$A_m = \frac{1}{T} \int_{mT}^{(m+1)T} h_p(t) dt \quad (2.50)$$

which is the TX pulse response integrated over a symbol period.

Setting the partial derivatives of these quantities to zero creates this system of equations:

$$\sum_{k=0}^{L-1} c_{k,opt} R[l - k] = \sum_{m=0}^{\infty} A_m R[l - m] \quad (2.51)$$

Since there are L taps, Eq. (2.51) represents a system of L equations which have $c_{k,opt}$ as their solutions. Compared with an LMS adaptation loop simulation, where an arbitrary leakage channel is used, the converged values match those calculated to a high degree of precision, as shown in Fig. 2.33.

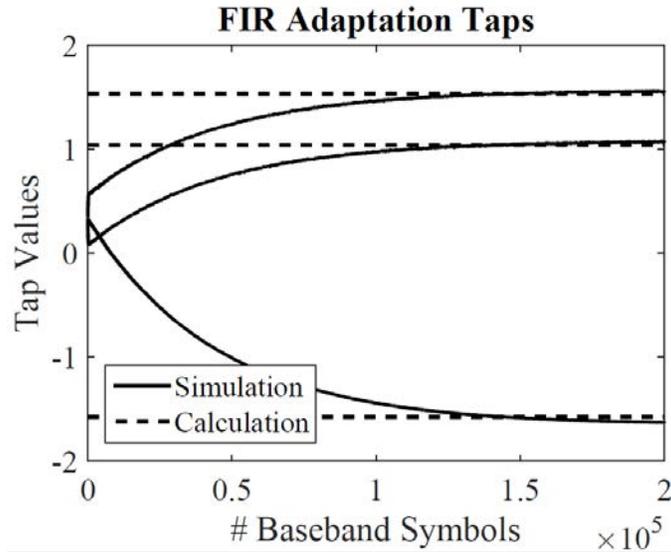


Figure 2.33: Comparison of adaptation and calculation of DAC taps.

Asynchronous Sampling - Unlimited Taps, Causal

By observing Eq. (2.51), it is apparent that if $L \rightarrow \infty$, then the two sides of the equation would be the same, meaning:

$$c_k = A_k \tag{2.52}$$

which is shown to converge in Fig. 2.34.

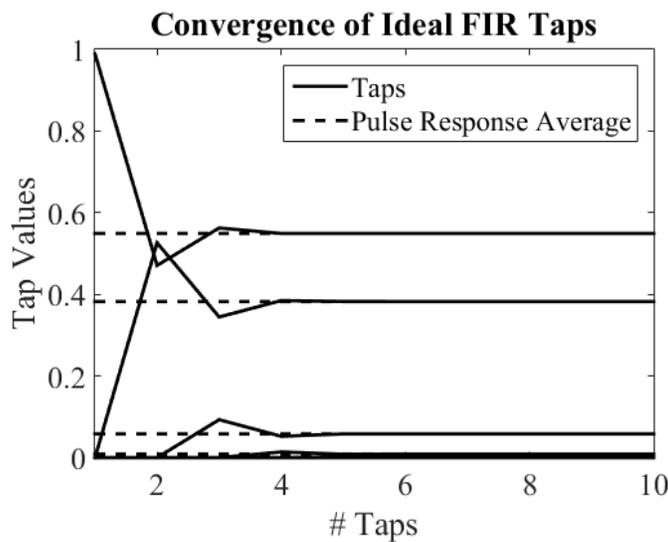


Figure 2.34: Convergence to A_n as N_{Taps} increases.

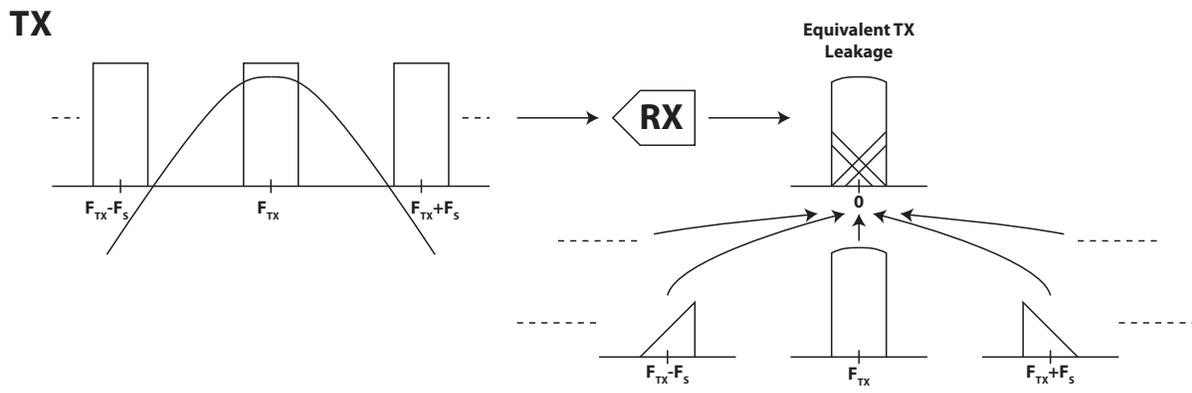
Unlike in the limited tap case, the unlimited tap case does not require knowledge of the autocorrelation function of the TX data, meaning that the TX data could be any bandwidth from 0 to $\frac{1}{T}$ and the residual would be minimized with this set of taps.

Furthermore, if all constraints are removed from the DAC data, where it can be non-causal and completely independent from the TX data, it is found that the exact same relation holds, that for minimization of the TX residual:

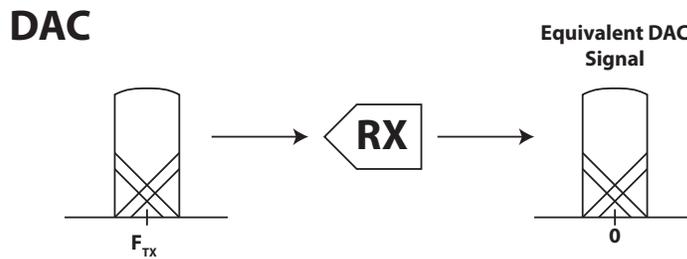
$$d_{DAC}(n) = \sum_{m=0}^{\infty} d_{TX}[n-m]A_m \quad (2.53)$$

It is important to note that while there was no enforcement of causality applied to this derivation, the answer is still causal. This is due to the fact that the leakage channel itself is causal.

Asynchronous Sampling - Unlimited Taps, Non-Causal



(a) TX leakage spectrum.



(b) DAC spectrum to match TX in-band.

Figure 2.35: TX and DAC spectrum after sampling.

Unlike the previous derivations outlined, in this case it is easiest to see the result in the frequency domain. Consider the process of going from TX data to TX leakage at the RX

input. Oversampled TX data is sent to the PA, which is effectively a zero-order hold and a modulation by F_{TX} . Then, the channel response, $H_{Leak}(j\omega)$, is applied to this signal. In the frequency domain, this translates to copies of the TX spectrum at multiples of F_S , multiplied by a sinc, then multiplied by the leakage transfer function. These copies spaced F_S apart are then aliased down to DC during sampling by the RX. This baseband signal, $Y_{TX}(j\omega)$ is proportional to the TX data spectrum $D_{TX}(j\omega)$, and can be written as:

$$Y_{TX}(j\omega) \propto D_{TX}(j\omega) \sum_{k=-\infty}^{k=\infty} \text{sinc}\left(\frac{\omega}{F_S} + 2\pi k\right) H_{Leak}(j\omega + j\omega_{TX} + 2\pi k F_S) \quad (2.54)$$

$$+ \text{sinc}\left(\frac{\omega}{F_S} + 2\pi k\right) H_{Leak}(j\omega - j\omega_{TX} + 2\pi k F_S) \quad (2.55)$$

The DAC is the same, except that there is no leakage transfer function due to the high bandwidth of the cancellation node. Therefore, if the DAC data is simply created from a bandlimited version of the TX and leakage transfer function spectrum, then within that bandwidth, the DAC will perfectly cancel the TX, shown in Fig. 2.35. In this figure, sinc filtering on the TX/DAC, as well as aliasing of the high frequency components of the DAC signal, are ignored because these do not affect the validity of this intuition. Assuming a TX signal of length L , the DAC samples are an inverse DFT of the TX spectrum.

2.9.3 Asynchronous Sampling Rejection

In the case of asynchronous sampling, there are two different ways the error can be quantified. The first is across the full bandwidth, the second is across only the TX bandwidth. The first gives a sense of how low the overall power is, useful for determining whether the RX may be compressed before sampling, while the second gives information about the interference the TX creates.

Both are very dependent on the type of channel which is present, as well as the oversampling ratio on the data. Quantization errors will not be introduced, and instead a study will be done of the effects of the channel and oversampling ratio.

The simplest case for in-band residual is the non-causal filter technique, where the zero-order held output of the DAC exactly matches the in-band TX leakage, assuming infinite DAC resolution. Here, quantization noise is the only limiting factor for the residual, so for the purposes of this section, its rejection is infinite.

In the casual case, as derived in Eq. (2.52), the ideal filter to minimize the total mean-squared error is composed of windowed integrations of the leakage channel's pulse response. In this section, a thorough analysis of this filter will be performed with regard to its error signal characteristics.

The transfer function from TX data to TX leakage can be described in three steps. First, the TX data is oversampled via some process. Next, the data is zero-order held to the oversampled symbol period. Finally, the leakage channel is applied to the zero-order

held signal. Defining $D(j\omega)$ as a delta-train version of the TX oversampled data spectrum, $D(j\Omega)$, the TX output can be expressed as:

$$Y_{TX} = H(j\omega) D(j\omega) \operatorname{sinc}\left(\frac{\omega T}{2}\right) e^{-j\frac{\omega T}{2}} \quad (2.56)$$

To generate the DAC data, the oversampled TX data is filtered with Eq. (2.50), then zero-order held. Creating A_m can be split into two steps: sampling an integrated pulse response of the leakage channel and subtracting neighboring samples to create a windowed integration. In the frequency domain, the first step can be expressed as:

$$H_1(j\Omega) = H\left(j\frac{\Omega}{T}\right) e^{-j\frac{\Omega}{2}} \operatorname{sinc}\left(\frac{\Omega}{2}\right) \frac{1}{j\frac{\Omega}{T}} \quad (2.57)$$

where $|H(j\omega)|$ is assumed to be negligible past $\frac{1}{T}$, a reasonable assumption when the oversampling ratio on the TX data is high. The second step creates a filter expressed as:

$$H_2(j\Omega) = H_1(j\Omega) e^{j\frac{\Omega}{2}} 2j \sin\left(\frac{\Omega}{2}\right) \quad (2.58)$$

$$= H\left(j\frac{\Omega}{T}\right) \left(\operatorname{sinc}\left(\frac{\Omega}{2}\right)\right)^2 \quad (2.59)$$

When zero-order held and referred to the output of the DAC:

$$Y_{DAC}(j\omega) = H(j\omega) D(j\omega) \left(\operatorname{sinc}\left(\frac{\omega T}{2}\right)\right)^2 \operatorname{sinc}\left(\frac{\omega T}{2}\right) \quad (2.60)$$

Finally, subtracting the two and normalizing by Y_{TX} , approximating $\operatorname{sinc}(x) \approx 1 - \frac{x^2}{6}$:

$$\frac{Y_{TX} - Y_{DAC}}{Y_{TX}} \propto H(j\omega) D(j\omega) \left(\frac{\omega T}{2}\right)^2 \quad (2.61)$$

which is valid for $\omega T \ll 1$, true over the TX bandwidth for large oversampling ratios. The normalized residual is proportional to T^2 , meaning that the average residual power within the TX bandwidth decreases at 12dB/octave of oversampling. Fig. 2.36 shows that across oversampling ratios, the TX band residual lowers at 12dB/octave.

As for the residual across all frequencies, another simplifying approximation of the residual can be made. In the time domain, the TX leakage is roughly a function with no discontinuities, while the DAC output is a zero-order held signal. The DAC waveform can be said to approximate the TX leakage with steps. The residual between these two functions can be approximated by a series of right triangles. The area of each residual triangle is inversely proportional to the square of the oversampling ratio (as the samples get closer, the DAC follows the TX waveform more and more closely). Therefore, the total residual power lowers by 6dB/octave, as shown in Fig. 2.37. Note that because both causal and non-causal

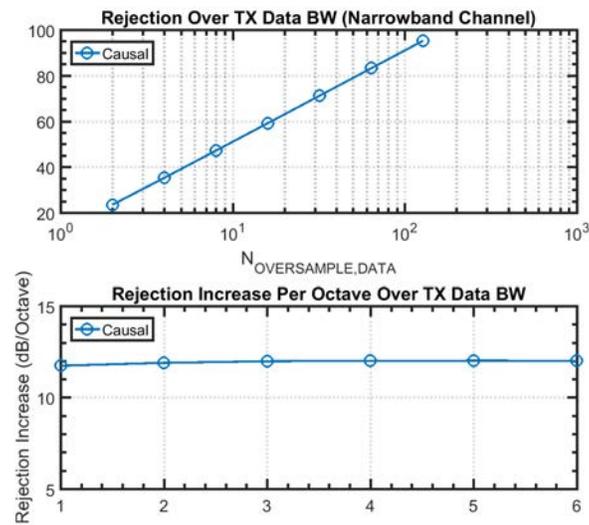


Figure 2.36: Narrowband channel residual.

methods seek to match the TX leakage to a high degree, there is no difference between the two for their total residual output power.

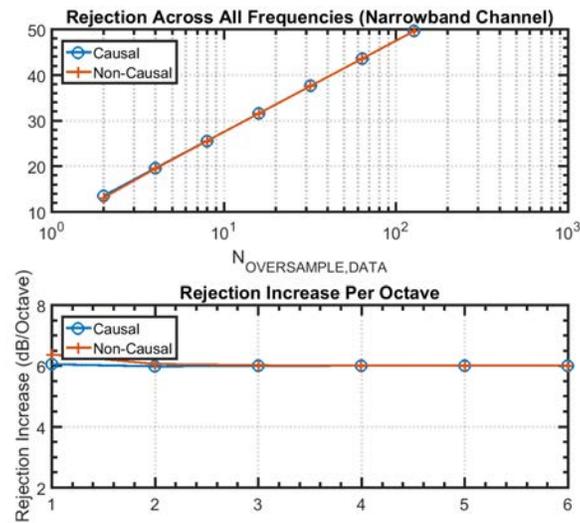


Figure 2.37: Narrowband channel residual.

2.10 Digital Backend Cancellation

While analog cancellation of 50dB is enough to prevent RX compression for signals $>+20$ dBm, it alone cannot be used for replacing filters in FDD. This is primarily due to the quantization noise added by the TX and cancellation DAC. Assuming a DAC with negligible DNL, quantization noise power density in dBm/Hz is defined as $P_{TX} - 10 \log(F_{sample}) - 6(N_{Bits} + 1)$. Assuming a noise floor of -164 dBm/Hz (10dB NF), and given 1GSps for TX and DAC, the total system precision is required to be 15 bits in order to increase the noise floor by 0.5dB. This system bit count can be segmented into partially analog cancellation and partially digital cancellation. The digital baseband sampled by the receiver ADC may be further processed using an adapted model of the TX/DAC, as well as an estimate of the leakage channel. The digital estimate of the leakage is then subtracted from the digitized baseband data to further improve the RX SINR in the presence of simultaneous transmission. This full process is illustrated in Fig. 2.38.

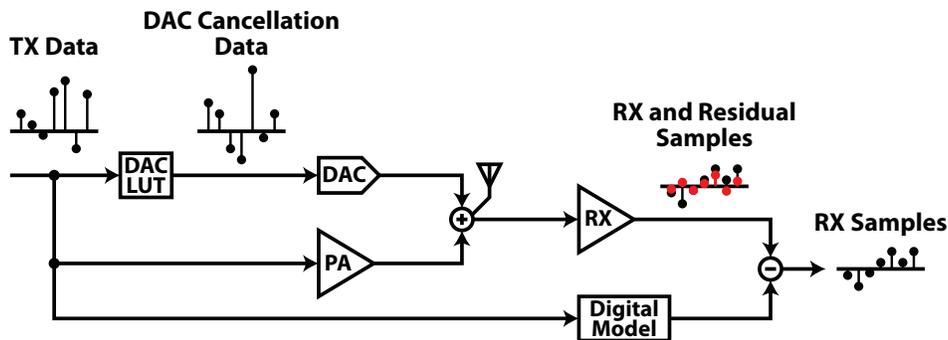


Figure 2.38: Digital backend cancellation.

The models used and their digital cancellation performance are detailed in Section 3.6. The models in that section remain unchanged through antenna VSWR, chip temperature, process variation, and other nonidealities, but their coefficients are subject to change from these factors. The change in adapted coefficients due to time variation is slow enough that it is feasible to conclude that an online calibration sequence could be developed to track these variations in real-time, but in this work, all adaptation and processing is performed offline.

Chapter 3

A CMOS Transceiver with Integrated FDD Support Up to +12.6dBm TX Leakage

3.1 Chip Implementation

While some recent works [51] include the full transceiver on chip, the vast majority of other cancellation works [41, 40, 20, 42] use some external elements in their measurements. These elements come in the form of either external antenna interfaces with isolation, external PAs, or both. For TX/RX isolation testing, it is essential to integrate as much of the system on chip as possible in order to exercise all leakage paths created by close proximity. Beyond antenna reflections, there exist substrate leakage paths, power supply noise, stray TX/RX transformer coupling, as well as other interference paths which are not captured in a design with external elements. Furthermore, signal generators or off-the-shelf PAs do not possess the nonidealities of transmitters made on chip, due to their lack of constraints in power, size, or technology. Secondly, regarding power, using a signal generator or off-chip PA as the TX makes it impossible to truly quantify the effective TX efficiency loss due to losses in the network or power consumption of the cancellation network. Finally, limited linearity of on-chip transmitters pose additional cancellation difficulties when compared with a single-tone output from a spectrum analyzer, or a high P_{SAT} PA.

With these considerations in mind, a test chip was created with the goal of integrating as much of the system as possible, while still maintaining flexibility for testing. Integrated on chip were the PA, RX (LNA, mixer, and baseband amplification), cancellation DAC, TX/RX matching networks, deserializers and retiming for 10Gb/s TX/DAC links, 25% LO generation and distribution, I/Q cell-sharing unit cell signal generation, as well as bias current DACs. There are three RF interface pins on the chip: the TX antenna port, RX antenna port, and the middle node between the two series matching network baluns. This middle pin provides the ability to isolate the TX and RX networks from one another in order

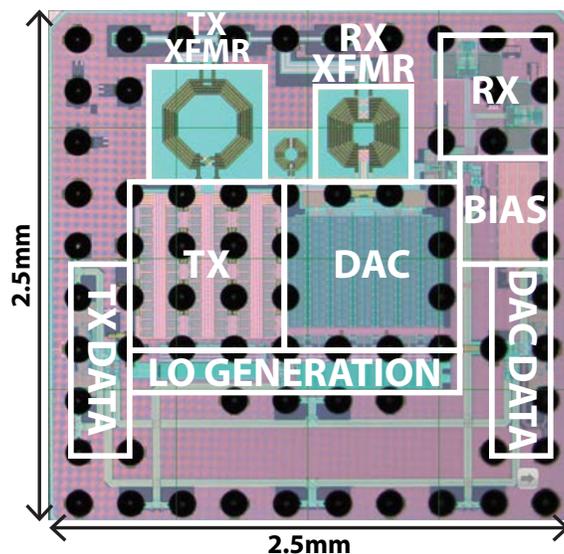


Figure 3.1: Die photo of first chip.

to test them separately. In normal cancellation operation, the middle node is opened on the PCB and the RX antenna port is grounded, while the TX antenna port is connected to the antenna load.

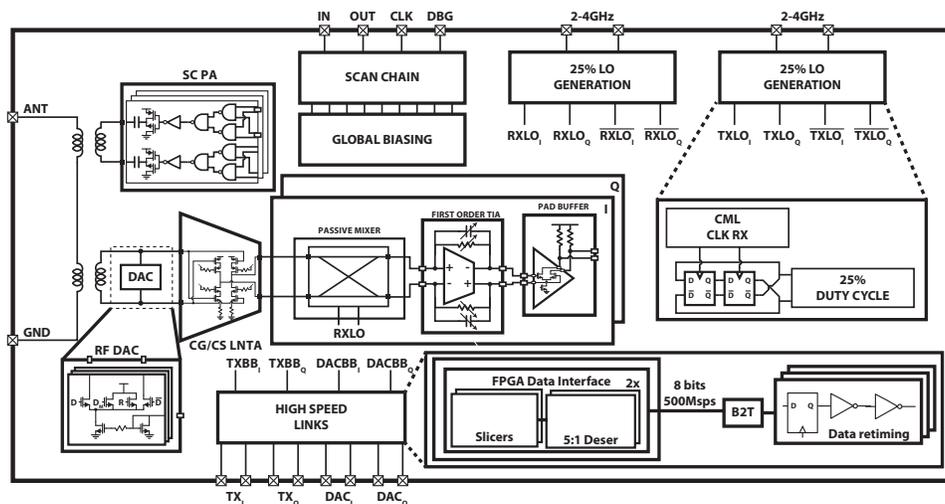


Figure 3.2: Chip top level.

3.1.1 Transmitter

As explained in Chapter 2, a low, code-independent PA output impedance lowers insertion loss/noise in the series configuration and prevents mixing of the RX signal with the TX when

outputting modulated data. A power amplifier with these characteristics, along with high signal linearity, is the switched-capacitor power amplifier (SCPA), first proposed by [52]. A variation of this architecture, a Cartesian SCPA with I/Q cell-sharing, was implemented on chip.

The basic operation of this power amplifier is depicted in Fig. 3.3. This is a voltage-mode RF DAC with unit cells consisting of an inverter and series capacitor. Enabled unit cells are driven with a square wave, while disabled unit cells are driven with a static 0 or V_{DD} signal. If the output impedance of the inverters are first assumed to be zero, then this network looks like a capacitive divider where the denominator is a constant because all cells are driven either to DC or to a square wave. The output resistance of the inverters acts like a series resistance, which scales in the same way as the capacitors. Driven by a square wave with fundamental voltage amplitude V_{DD} , the output voltage is written:

$$V_{Out} = V_{DD} \frac{Y_{On}}{Y_{Total}} \tag{3.1}$$

$$= V_{DD} \frac{n}{N} \tag{3.2}$$

where n is the number of unit cells connected to the square wave input and N is the total number of unit cells. This first-order expression is both linear and wideband, even though the unit cell output impedance varies with frequency. Furthermore, the output impedance is independent of n and is equal to Y_{Total}^{-1} . The capacitive imaginary part of the output impedance is cancelled by connecting this PA to a series inductance, provided by the balun matching network, shown in Fig. 3.3.

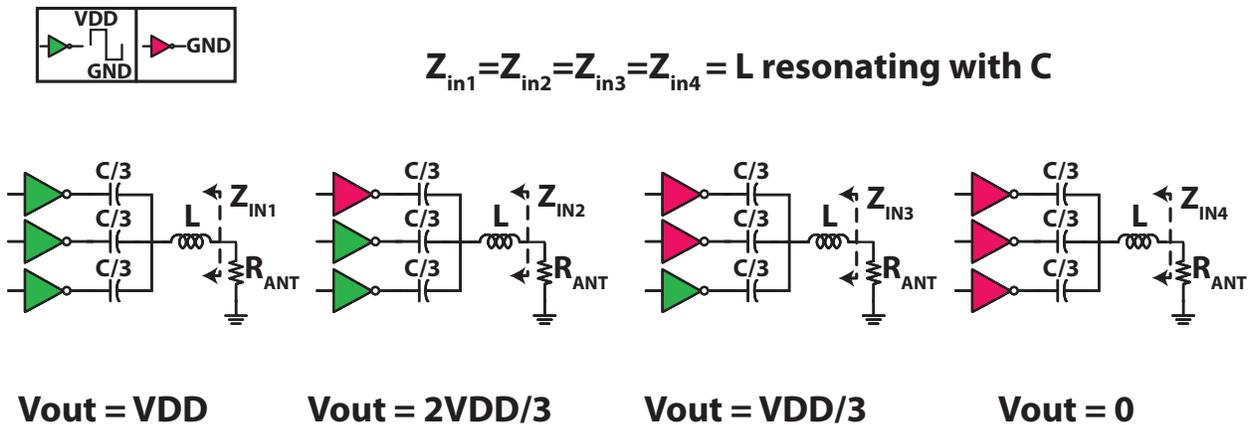


Figure 3.3: Operation of switched-capacitor power amplifier [45].

Sizing and matching network optimization for this architecture can be found in [48]. The Cartesian SCPA was implemented using an I/Q cell-sharing technique, where each enabled unit cell can output a 9QAM signal using pulse width modulation, where the $I = 0$ and $Q = 0$

waveforms are 25% duty cycle and the $|I| = |Q|$ waveforms are 50%, shown in Fig. 2.14. For the same peak output power, this method uses the same area as a polar PA and $\sqrt{2}$ less area than a conventional Cartesian PA having separate I and Q unit cells. Consider the peak power case of the I/Q cell-sharing architecture, where all unit cells are enabled and output with 50% duty cycle. This is indistinguishable from the polar PA at maximum power with the same area, since both output 50% duty cycle waveforms in this scenario. For $I = 0$ or $Q = 0$ cases, the maximum output power is half that of the polar case. The achievable constellation region is a diamond inscribing the polar architecture's circle, illustrated in black in Fig. 3.4.

If I/Q cell-sharing is not used, and instead I and Q sub-PAs are created and summed in parallel, then this configuration causes the sub-PAs to load one another, reducing their individual maximum output voltage to $V_{DD}/2$ and individual maximum power to $P_{TX}/4$. $I = 0$ and $Q = 0$ signals add in power due to the orthogonality of I and Q , so when both sub-PAs output their maximum code, the total output power is equal to half of the polar's maximum. Therefore, the output power of the parallel SCPA is strictly less than the maximum in the polar case, plotted in green in Fig. 3.4. Furthermore, the parallel configuration requires twice the area of the polar or I/Q cell-sharing case for the same number of bits. In the case where a significant portion of the PA area is due to data routing, the area of the PA is roughly proportional to the number of cells. If separate I and Q sub-PAs are required, then cells in total are needed, which doubles the area of the parallel PA.

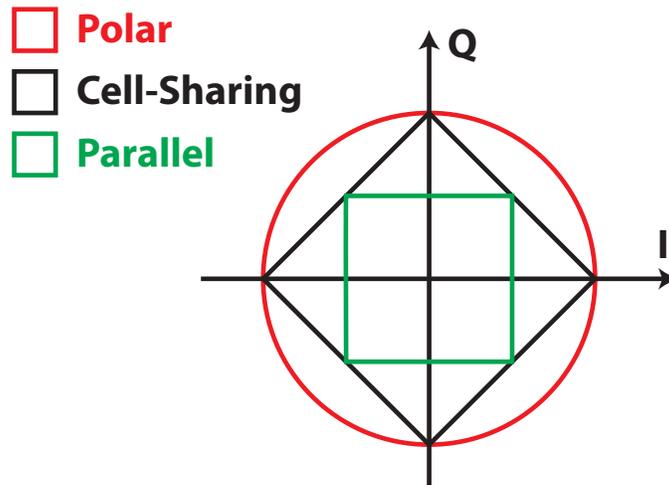


Figure 3.4: Available constellation regions for PA architectures.

Explained in detail in [48], the theoretical efficiency of the I/Q cell-sharing SCPA is compared with that of the polar implementation in Fig. 3.5. When simulated using a representative OFDM transmit signal with 6dB PAPR, the average efficiency lags the polar case by single-digit percentage points.

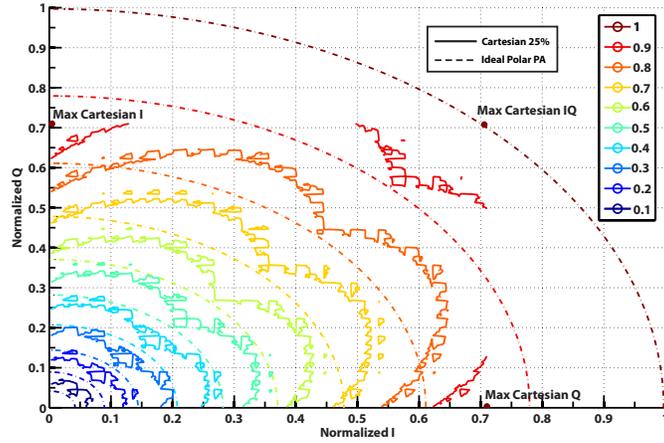


Figure 3.5: Theoretical drain efficiency of polar versus Cartesian SCPA [48].

The PA core was integrated along with the matching network balun, LO distribution, data deserializers, and retiming circuitry, shown in the top level schematic of Fig. 3.6.

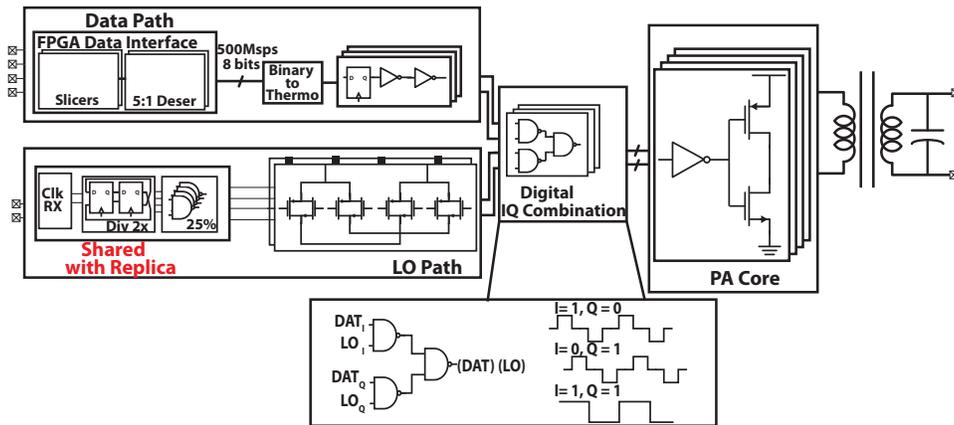


Figure 3.6: TX top level.

3.1.2 TX Efficiency Degradation

Introduced in Section 2.6, the series stack of antenna, TX, and RX creates TX insertion loss due to losses in the RX balun winding. In simulation, taking into account all on-chip passives in the TX/RX interface for the chip, the power gain of the TX has been plotted with and without the RX and cancellation DAC in Fig. 3.16. Both TX and RX baluns have a 1:2 turns ratio and a quality factor of 8 at 1.5GHz. Across the full operating frequency range, the power loss due to the RX winding is approximately 0.35dB.

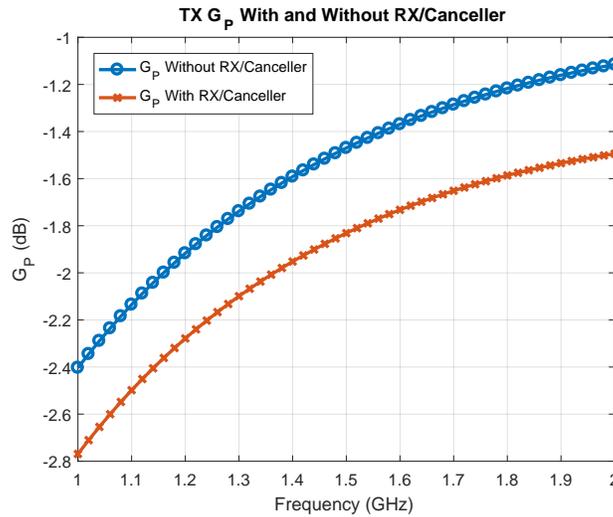


Figure 3.7: Simulated TX insertion loss due to RX transformer.

3.1.3 Receiver

The chief motivation driving the design of the on-chip receiver was flexibility of the system testing setup. The nominal gain, approximately 15dB, was chosen to elevate the thermal noise floor well above that of the measurement instruments, most notably the spectrum analyzer, such that reliable measurements of RX noise figure and canceller noise figure degradation could be made. The gain was also desired to be tunable over a wide range such that moderate leakage signals could be amplified linearly with and without cancellation. In normal operation, the baseband bandwidth should be made on the order of the channel bandwidth so as to filter blockers as well as the frequency offset TX leakage. In order to view the uncanceled TX signal accurately, it is useful to have a higher bandwidth mode up to the highest testing duplex spacing. Furthermore, it was desired to not perform downconversion at all, and instead view the RX signal directly at RF in an extreme bandwidth mode. This could be used, for example, to measure the effect of reciprocal mixing on the leakage signal, via comparison of the TX leakage noise skirts between the RF and baseband outputs. A moderate noise figure of 4.5dB was targetted in order to be dominated by canceller noise for high TX power.

The architecture of the receiver consisted of a complementary CG-CS LNTA, passive mixers, a baseband TIA, and 50Ω pad drivers. The LNTA had a fixed G_m , while the impedance value and bandwidth of the TIA was configurable using resistor and capacitor DACs. Detail for the creation and optimization of this receiver design is provided in [48]. The top level of the receiver is shown in Fig. 3.8.

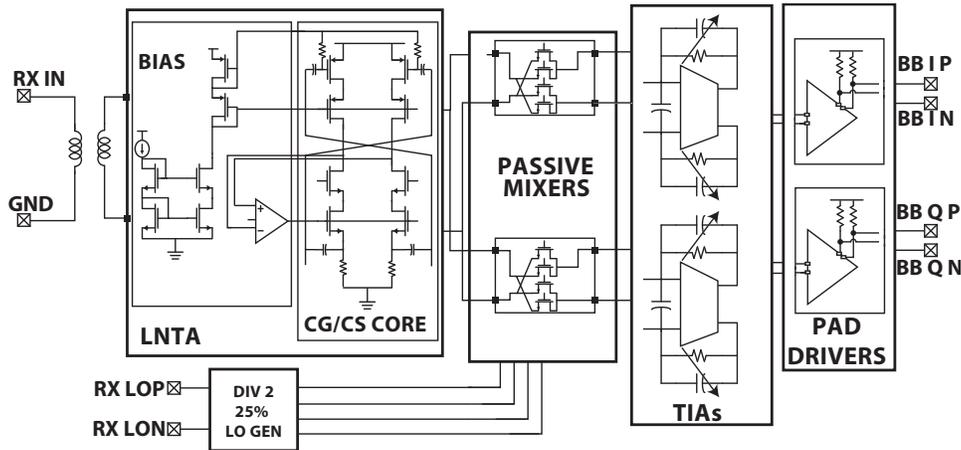


Figure 3.8: RX top level.

3.2 DAC Design

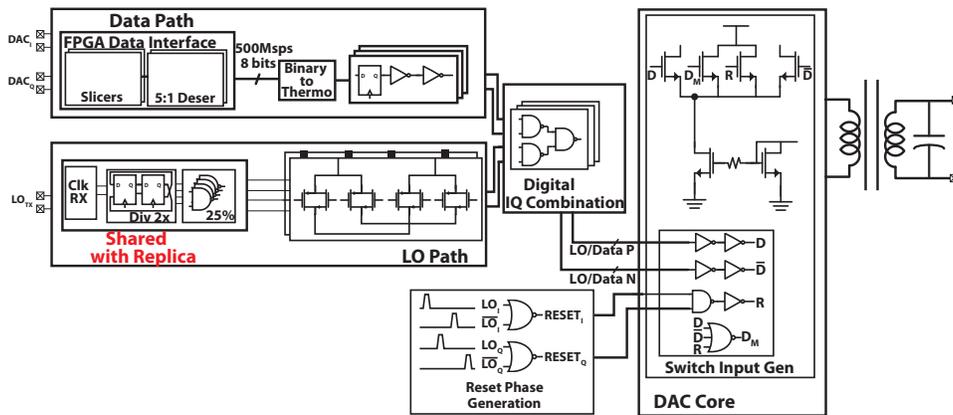
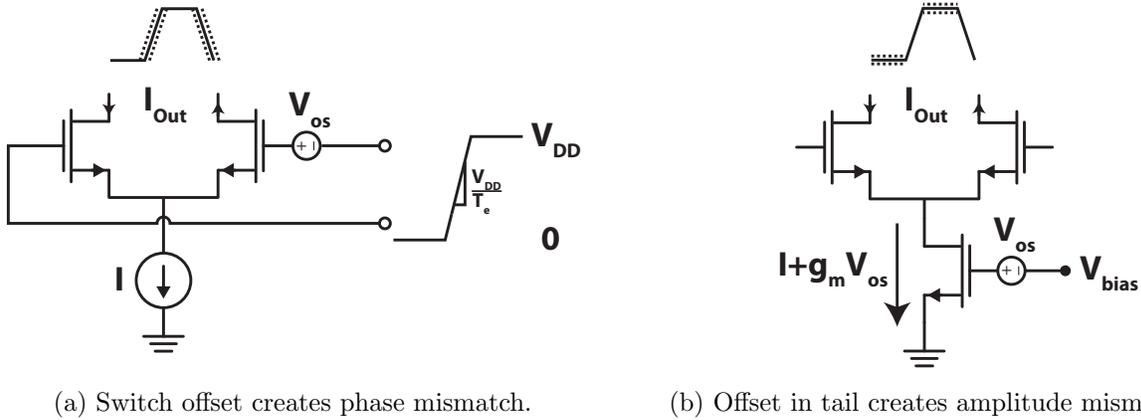


Figure 3.9: Top level of DAC.

In this section, the design of the cancellation DAC is motivated and explained. From the high level calculations in Section 2, to achieve >50dB TX cancellation in the presence of even significant nonlinearities in the DAC, 10 bits in total are desired. With a half-binary, half-thermometer segmentation, 36 total cells are used in this configuration, creating a 6x6 square. To reduce DNL in the presence of linear gradients, a 2 dimensional common centroid configuration is used. To minimize INL in the presence of dishing gradients [53], the ordering of thermometer cells were placed in a spiral pattern shown in Fig. 3.12b.

The DAC consists of a tail device connected to a set of 4 hard-switched transistors which perform the I/Q cell sharing modulation. The main source of DAC mismatch is the tail device. A 1% offset in tail current corresponds directly to a 1% change in output current.



(a) Switch offset creates phase mismatch.

(b) Offset in tail creates amplitude mismatch.

Figure 3.10: Mismatch propagation for switches and tail device.

The switches, on the other hand, do not require nearly as much matching due to the fast risetime of the square wave signal. Consider a mismatched DAC unit cell driven by a square wave of frequency F and edge transition time T_e , shown in Fig. 3.10. To equate mismatches for the tail device and the switch devices, the offset voltage V_{os} for some fractional mismatch in current can be found as:

$$\left(\frac{\Delta I}{I}\right)_{Tail} = 2\frac{V_{os}}{v_{Ov}} \quad (3.3)$$

while for an input-referred voltage offset V_{os} in a hard-switched differential pair with an LO slope defined in Fig. 3.10a, the current output mismatch is:

$$\left(\frac{\Delta I}{I}\right)_{Switch} = 2\pi\frac{V_{os}}{V_{DD}}T_eF \quad (3.4)$$

Substituting Eq. (3.3) into Eq. (3.4) and rearranging gives:

$$\left(\frac{\Delta I}{I}\right)_{Switch} = \pi\frac{v_{Ov}}{V_{DD}}T_eF\left(\frac{\Delta I}{I}\right)_{Tail} \quad (3.5)$$

which, for a 2GHz 1.2V LO with a fanout of 4 inverter chain and a v_{Ov} of 200mV, the same V_{os} which produces 1% current mismatch in the tail produces only 0.015% current mismatch in the switches. Therefore, the effect of switch mismatch can be ignored completely.

To reduce noise from the switching transistors, the tail drain impedance must remain high at the TX LO frequency, necessitating a low routing capacitance from the tail devices to the switching quad. This requires that the DAC be constructed as an array of full unit cells, rather than a tail device array and a switching array. Such a design is seen in [53, 54, 55], while separate current and switching blocks are seen in [56, 57].

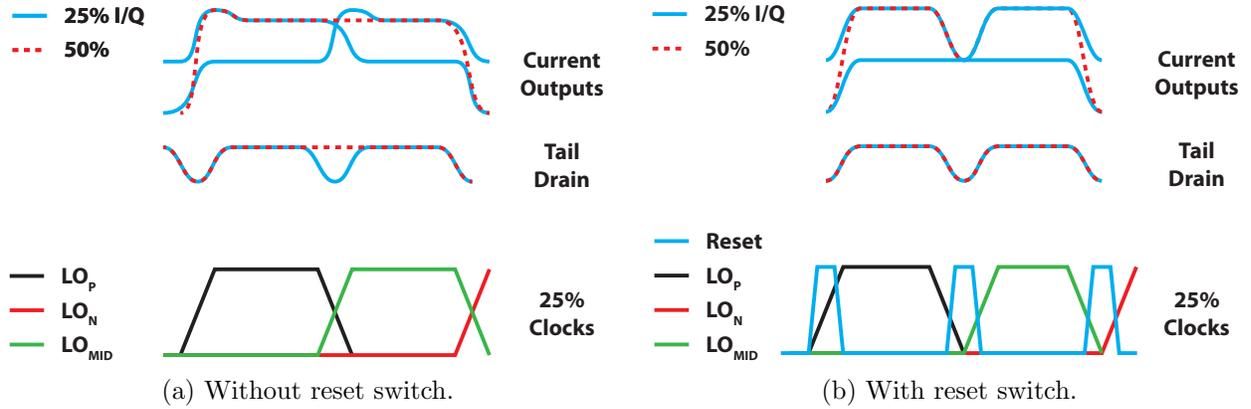


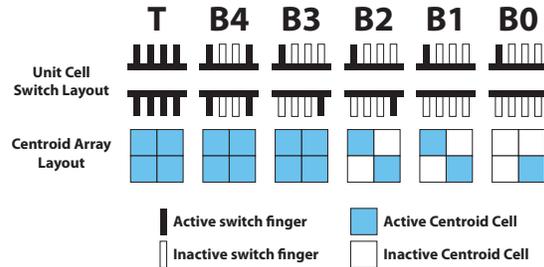
Figure 3.11: DAC output current waveforms with and without reset.

Because the replica current source is an RF DAC, I/Q nonlinearity is an important consideration, as shown in Section 2. In the DAC, this can appear due to various sources. Tail device drain swing is one significant source of nonlinearity, where a 25% signal creates a dip in the drain voltage due to timing mismatches between the LO_{MID} turning on and the LO_P turning off, while there is no corresponding dip for a 50% waveform which has only one transition, shown in Fig. 3.11a. The difference in voltage at the beginning of the transition leads to either a surge or reduction in current at the start, impairing I/Q nonlinearity. To combat this, the duty-cycle of the input DATA & LO waveforms is reduced, and an extra Reset switch is introduced to the switching array. This Reset switch activates at the end of each 25% phase, forcing the same drain voltage swing regardless of the complex code sent to the unit cell, illustrated in Fig. 3.11b.

Another form of nonlinearity, deterministic DNL, requires attention to the binary unit cells to mitigate. If binary cells are naively constructed from the thermometer cells by reducing the size of the output switches, capacitive load mismatch for the unit cell drivers can severely impact DNL, due to differences in driver bandwidth. By making the binary switches the same total size as the thermometer cells, but shunting fractions of the total output current to the center tap, as shown in Fig. 3.12a, all unit cell drivers and switches have the same bandwidth.

B3	T25	T20	T22	T24	B1 _N	B1	T24	T22	T20	T25	B3
T27	T12	T4	T7	T14	T29	T29	T14	T7	T4	T12	T27
T18	T10	T1	T3	T9	T16	T16	T9	T3	T1	T10	T18
T15	T8	T2	T0	T11	T17	T17	T11	T0	T2	T8	T15
T28	T13	T6	T5	B4	T30	T30	B4	T5	T6	T13	T28
B2	T23	T21	T19	T26	B0 _N	B0	T26	T19	T21	T23	B2 _N
B2 _N	T23	T21	T19	T26	B0 _N	B0 _N	T26	T19	T21	T23	B2
T28	T13	T6	T5	B4	T30	T30	B4	T5	T6	T13	T28
T15	T8	T2	T0	T11	T17	T17	T11	T0	T2	T8	T15
T18	T10	T1	T3	T9	T16	T16	T9	T3	T1	T10	T18
T27	T12	T4	T7	T14	T29	T29	T14	T7	T4	T12	T27
B3	T25	T20	T22	T24	B1	B1 _N	T24	T22	T20	T25	B3

(a) DAC centroid configuration.



(b) DAC switch binary division.

Figure 3.12: DAC high level structure.

3.3 DAC Thermal Noise Cancellation

In receivers, thermal noise reduction is routinely performed, creating multiple paths where the noise of a device interferes destructively while the signal interferes constructively [58, 59, 60]. While this is a common trait of state-of-the-art receivers and analog devices requiring very high sensitivity, thermal noise mitigation methods are not employed for transmitters or DACs due to the large signal levels they produce. In DACs especially, the quantization noise floor is far higher than the thermal noise floor. Take, for example, a transmitter capable of outputting +20dBm on a 50Ω load. In order for a noise figure of 10dB to negatively affect the SQNR of a Nyquist DAC by 3dB, 18 bits would be required. Even if an oversampling ratio of 100 were used, 15 bits would still be required. This is far above the requirements for the TX, but a 20dB noise figure due to a cancellation DAC would heavily impact overall system performance because this noise adds directly to the RX noise figure. Note that in the LTE standard, a 15dB noise figure limit is used for FDD [25]. Mentioned in Section 2.10 and expanded upon in Section 3.6, digital cancellation of the DAC and TX quantization noise can be employed to increase the effective bits in the system, but this does not remove non-deterministic sources of noise, such as DAC thermal noise, necessitating a noise cancellation mechanism. In VCOs requiring high spectral purity, feedback techniques have been proposed to reduce tail noise [61, 62], and similar techniques are used in the cancellation DAC.

In Section 2.4.4, a model for the DAC’s noise contribution was introduced, and in Section 3.2 was motivated by the DAC’s architecture. This model consists of a noisy baseband current source which is upconverted to F_{TX} through a noiseless mixer. This upconverted current is present at the RX input port, directly adding to the overall RX NF. In Chapter

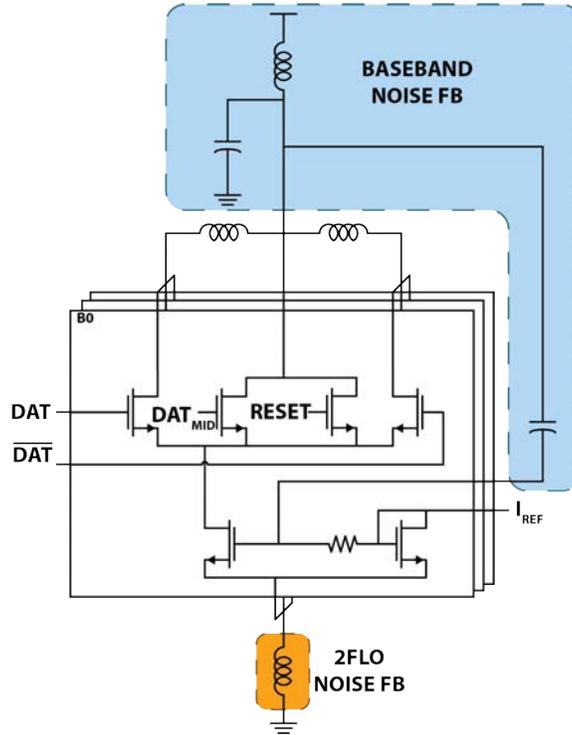


Figure 3.13: DAC with noise feedback highlighted.

2, Eq. (2.12) shows that there are three ways of reducing the current noise output by the DAC: increase tail v_{Ov} , reduce RX transformer turns ratio, and lower mixer A_{conv} . All three of these methods increase DAC power consumption, the first through increasing the DAC supply to compensate for higher headroom requirements, and the last two through increasing the tail current required for the same TX output power at the antenna.

For another method to achieve noise reduction, a feedback technique can be used which does not increase the power consumption of the DAC. This technique takes advantage of the fact there are nodes where tail noise can be sensed independently from the RX or TX signals. This is very important because it allows the noise to be reduced through the feedback loop gain, without reducing the output signal of the DAC. There are two such isolating nodes: the RX balun center tap and the source of the unit cell tail device. Focusing first on the center tap, consider that the Mid switches on the replica DAC enable class-A backoff operation, where disabled unit cells shunt their current directly to the center tap. Ideally, the only fluctuation in the center tap current waveform would come from noise in the tail current devices, regardless of DAC data. If the balun is well-balanced, the RX signal will have no common mode component at the RX input port, keeping the center tap isolated from the RX signal. Corruption from the RX signal and chip switching noise, as well as supply noise from the LDO are irrelevant because the center tap will be used for a baseband feedback technique, on the order of tens of MHz, far from frequencies where these corruption signals

are present. Second, the source of the tail devices also has this same property, where the same value of current flows into the DAC ground independent of DAC data. For this design, the baseband portion of tail noise is sensed at the center tap, while the $2F_{TX} + F_{Duplex}$ portion is sensed at the tail source. The DAC has a dedicated supply to minimize the switching noise present on the center tap node and has a dedicated ground to minimize interference around $2F_{TX}$.

In both cases, the current noise is converted to a voltage through a resonant impedance and is fed back using the DAC aggregate G_m . Because of the high bias current of the DAC, this G_m is very large (300mS), creating a large loop gain for moderate source and center tap impedances. While this DAC G_m is constant, the effective noise reduction is code-dependent because of noise injected into the loop from inactive cells. This effect is most prominent at low codes is not an issue because at lower DAC codes, the RX is the dominant source of desensitization.

In the following section, the thermal noise reduction vs. code will be analyzed, and a testing setup will be introduced for validating both the upconversion mechanism and verifying the center tap resonance frequency.

Fig. 3.14a shows a simplified schematic of the DAC and baseband noise feedback mechanism. The switching quad is represented as a single cascode transistor because there is always a path from the tail to the center tap, independent of data and LO phase. The only distinction between unit cells in this case is which cells are active (the tail outputs to the differential RX input), or inactive (tail shunted to center tap). All active cells are lumped into a single tail transistor of transconductance $G_{m,A}$, and all inactive cells are lumped into a tail with $G_{m,I}$. In this simplification, it is clear that the active and inactive devices are simply diode connected to the center tap node. I_A is the total active tail current signal, i_A^2 and i_I^2 are the active and inactive tail device noise sources, respectively. A further simplification can be made, shown in Fig. 3.14b, where the diode-connected inactive transistor is replaced with a resistor and the inactive current noise is kept in the same position. Intuitively, since both the center tap impedance and the disabled cells are connected to small signal ground, and it doesn't matter what current flows through them, they can be put in parallel. This reduction in the effective center tap impedance implies that, even ignoring the current noise injected by the inactive devices, the current noise rejection is code-dependent.

$$Z = Z_{CT} \parallel \frac{1}{G_{m,I}} \quad (3.6)$$

$$= \frac{Z_{CT}}{1 + G_{m,I}Z_{CT}} \quad (3.7)$$

For the active current noise:

$$i_A = i_{nA} \frac{1 + G_{m,I}Z_{CT}}{1 + G_m Z_{CT}} \quad (3.8)$$

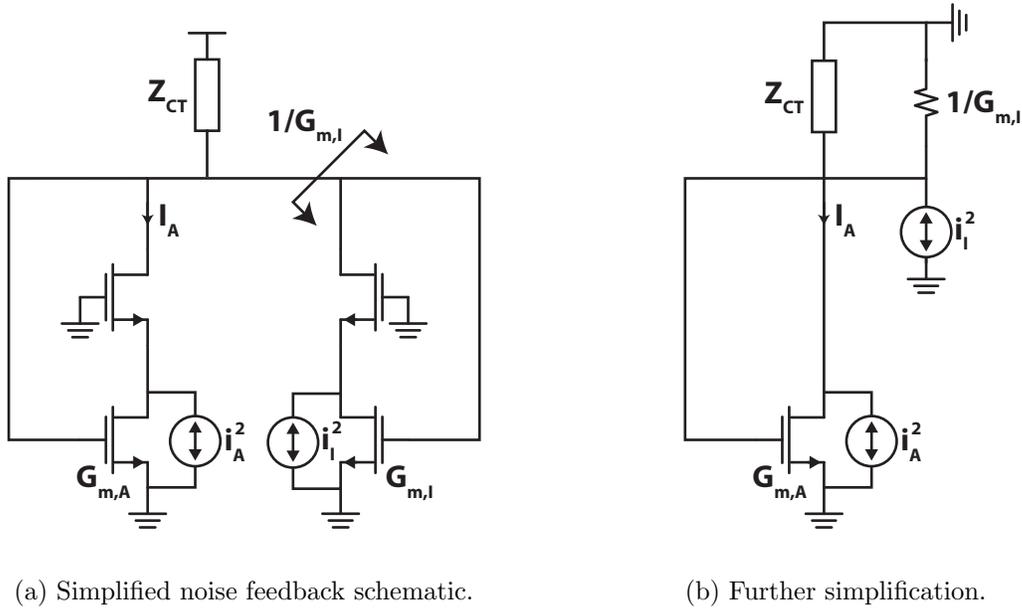


Figure 3.14: Current DAC noise propagation models.

For inactive current noise:

$$i_A = -\frac{G_{m,A}Z_{CT}}{1 + G_m Z_{CT}} i_{n,I} \quad (3.9)$$

Combining Eq. (3.8) and Eq. (3.9) and simplifying:

$$i_A^2 = i_{n,A}^2 \frac{|1 + G_{m,I}Z_{CT}|^2 + G_{m,A}G_{m,I}|Z_{CT}|^2}{|1 + G_m Z_{CT}|^2} \quad (3.10)$$

The feedback noise current normalized by non-feedback current as a function of replica DAC code is plotted in Fig. 3.15, assuming a fully thermometer DAC and $G_m Z_{CT} = 2$. In this plot, the output noise is normalized by the noise without any feedback, assuming that all cells are active. Shown here, the maximum noise variance with feedback is less than one third of the maximum without, and after approximately half the cells are active, the output noise reduces with increasing code because the attenuation rises with n^2 , while the noise power rises with n .

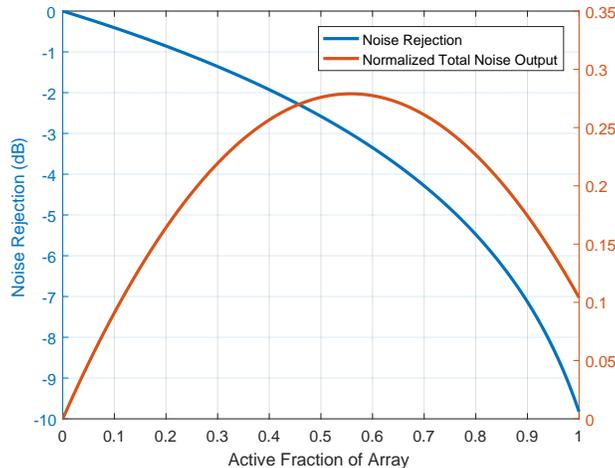


Figure 3.15: Feedback rejection and normalized output noise.

3.4 TX Efficiency Degradation Mechanisms

Because this work creates a novel interface between the TX and RX, any effective reduction in TX efficiency must also be considered. Here, two main sources of degradation to the effective TX efficiency exist: RX winding loss and the power consumption of the canceller power and digital predistortion (DPD). The first source of degradation comes from resistive losses within the RX which cause a short on the balun RX input side to be inadequately translated to the antenna side. While the current flowing into the RX input side is almost entirely circulated by the DAC, creating minimal voltage excursions at the RX, the voltage swing on the antenna side is not as sharply reduced. In simulation, taking into account all on-chip passives in the TX/RX interface for the chip, the power gain of the TX has been plotted with and without the RX and cancellation DAC in Fig. 3.16. Across the full operating frequency range, the power loss due to the RX winding is approximately 0.35dB.

The second form of efficiency degradation is due to the increased system power consumption due to the power draw of the canceller as well as that of the DPD and filtering schemes which are used to improve cancellation of the TX signal at the RX input. A conservative estimate of requirements for the digital filter is 8 taps for 200MS/s with 10 bit coefficients. According to [50], the power consumption for this filter is 10mW in a 65nm process. Digital predistortion is achieved through a lookup table, estimated to cost 15mW in power.

There is additional power consumed to run the adaptation algorithms to change these filters and lookup tables as the network or other nonidealities change, but because the dynamics of the channel are far slower than the datarate, the power consumption of these digital algorithms can be amortized over a very large operation time, making their average power consumption negligible. Additionally, the power consumption of the DAC is a dominant source of power consumption. As stated before, the current required from the DAC supply is proportional to $\sqrt{P_{TX}}$, though power at backoff for modulated data depends also on the

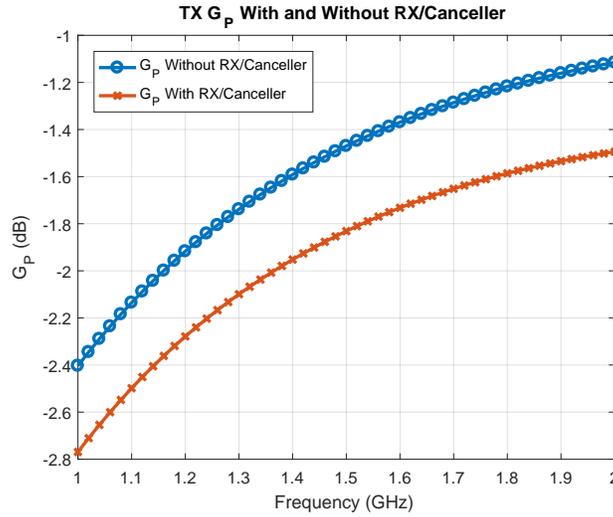


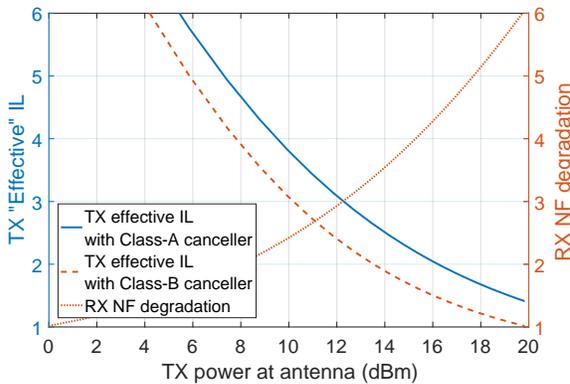
Figure 3.16: Simulated TX insertion loss due to RX transformer.

TX loss parameters.

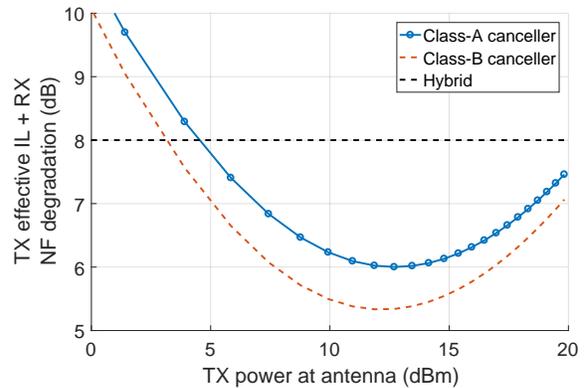
TX Average Backoff (dB)	6
TX Average PAE	25%
Digital Filter Power (mW)	10
Canceller DPD Power (mW)	15
IL From RX Winding (dB)	-0.35
DAC Supply Voltage (V)	1

RX NF Parameters.

RX NF (dB)	2.5
RX XFMR IL (dB)	1
RX XFMR N_{Turns}	2
$R_{TX}(\omega)$	7
DAC v_{Ov} (mV)	800
Uncorrelated Phase Noise (dBc/Hz)	-190



(a) TX/RX degradation versus output power.



(b) Total degradation comparison with hybrid.

Figure 3.17: Total insertion loss and desensitization for architecture.

type of backoff available for the DAC. In Fig. 3.17a, degradation of the TX efficiency due to these mechanisms, for a modulated data signal with 6dB PAPR, is plotted with class-A and

class-B DAC backoff.

Combining the effects of TX efficiency degradation and RX noise figure degradation, a fair comparison can be made with the hybrid technique. RX noise figure degradation can be cleanly converted to RX insertion loss, and TX efficiency degradation to effective TX insertion loss. In practical hybrids, TX and RX insertion losses are typically both 4dB. In Fig. 3.17b, the combined effective TX and RX insertion losses for the aggressive design point specified in Fig. 2.18 are combined and compared with an 8dB combined loss from the hybrid. As shown, from approximately +5dBm to above +20dBm, the replica DAC gives better performance than the hybrid. Additionally, the tunability of the replica canceller over a wide range of antenna VSWR makes this a more attractive solution than the hybrid even for power levels where the effective combined insertion loss is close to or exceeds that of the hybrid.

3.5 Measurement Results

The system was taped out in the TSMC 65nm process, shown in Fig. 3.18b and tested using the PCB in Fig. 3.18a. At a high level, the measurements for this chip consisted of characterizing the TX, RX, and cancellation DAC in isolation, then together as a system. The most important system level measurements were level of rejection, the RX 1dB compression point due to the TX with both single tone and modulated data, cancellation in the presentation of nonidealities such as VSWR, and the cancellation system noise figure. Additional measurements were made of TX phase noise propagation and feedforward cancellation in the system.

3.5.1 Isolated Measurements

In isolation, the TX performance is summarized in Fig. 3.19. It achieved a maximum transmit power of +19dBm at 1.2GHz and has an extremely linear output vs. code response due to the switched-capacitor architecture. The simulated output impedance is 5Ω .

The receiver performance is shown in Fig. 3.20a and Fig. 3.20b. Gain and bandwidth were tunable from 6-18dB and 15-140MHz respectively. The matching network in the first version had a 3dB bandwidth from 1.2-2GHz and a nominal loss of 2.9dB. The noise figure of the RX alone was found to be 4.7dB, whereas the full system noise figure was 7.6 due to matching network loss. It should be noted that the RX itself produces a very wideband response, but the passives in the antenna interface determine the system bandwidth.

The DAC's location directly in front of the LNA made testing it in isolation very difficult. The large voltage swing generated by the DAC differential current is significantly different from normal operation, where the TX current would cancel this voltage swing, making it impossible to measure the DAC constellation accurately for codes higher than the first few bits. Additionally, measuring the phase of the complex baseband is difficult because of phase drift in the measurement instruments, even when a 10MHz reference is shared.

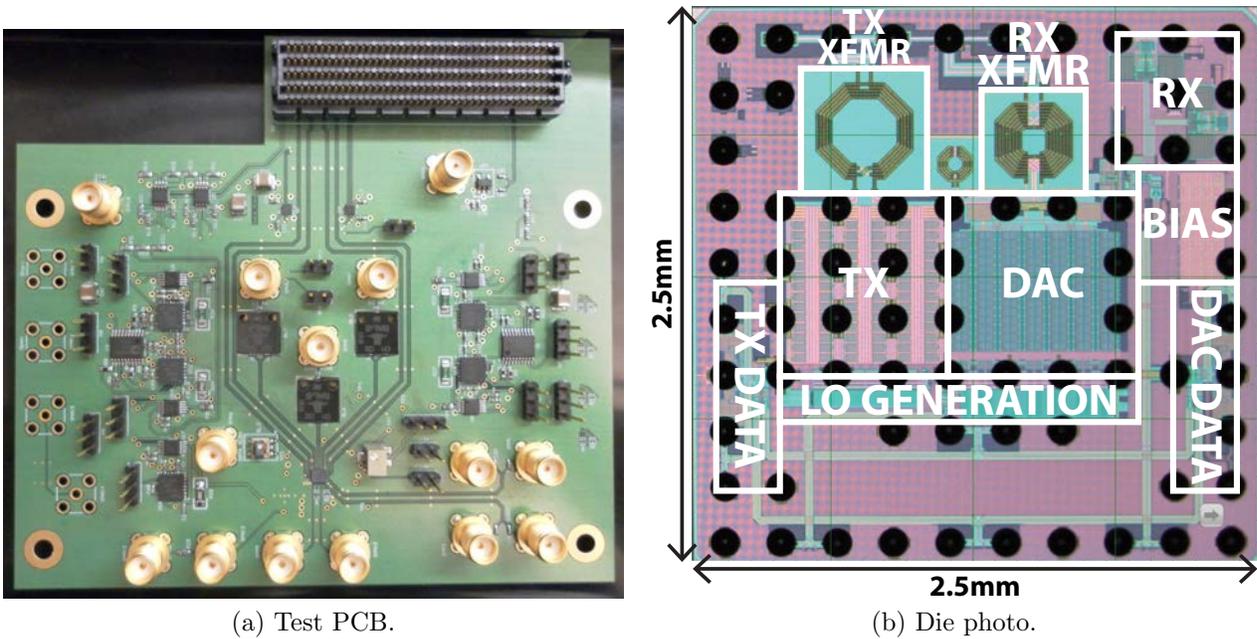


Figure 3.18: System test setup.

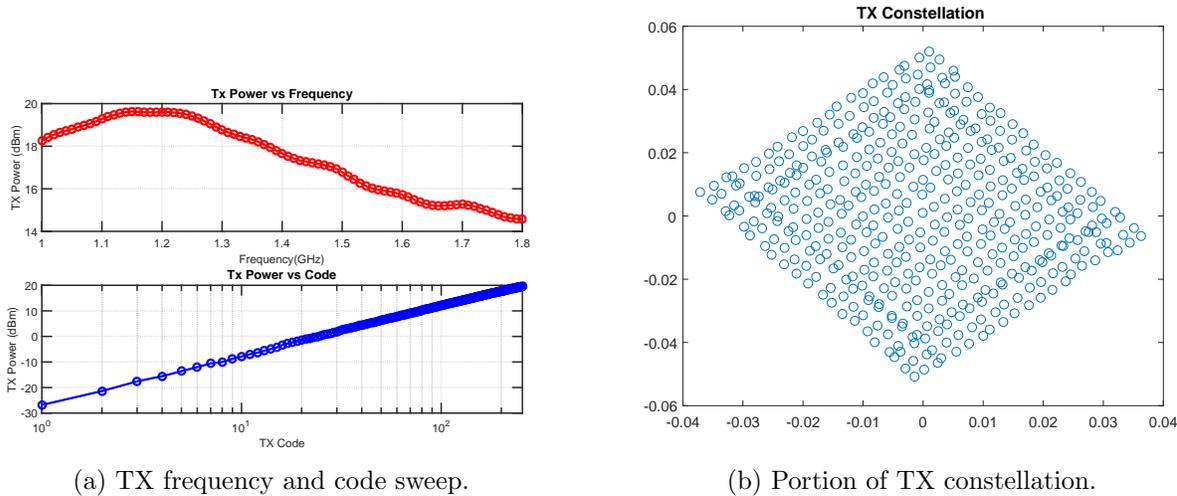


Figure 3.19: TX measurements.

Much of the slow-drift phase noise can be mitigated by using direct conversion of the DAC signal, but in some cases measurements at an offset may be required. Two architecture advantages present in the DAC were used to reconstruct the DAC constellation in the presence of phase drift. In the case where a subset of the total codes was used, unit cells

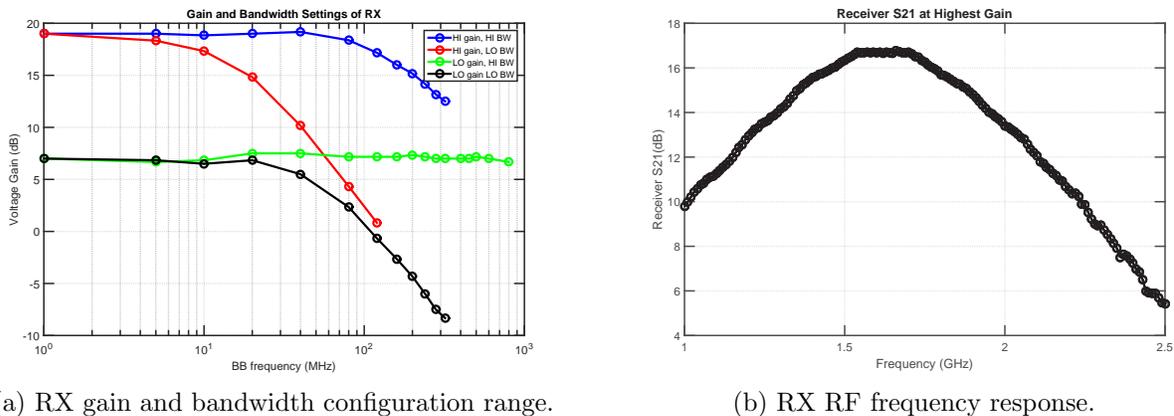


Figure 3.20: RX measurements.

are shown to add highly linearly in Section 3.6. The second architecture advantage is that the distribution of the I and Q LOs is such that there is very little mismatch between unit cells, and that there is little mismatch due to the output routing. This is due to the common centroid configuration of the unit cells. Therefore, it can be assumed that the phase of a unit cell outputting I only has the same phase regardless of the amplitude of the output.

Putting these together, there is a method for extrapolating phase measurements from purely amplitude measurements, even in the presence of I/Q nonlinearity and DNL. The objective is to find both the amplitude and phase of an I only signal, a Q only signal, and those two signals added together. The I and Q signals should be selected such that their binary representations do not overlap ($\text{AND}(C_I, C_Q) = 0$). Measurements of $(C_I, 0)$, $(0, C_Q)$, (C_I, C_Q) are then made. The amplitude measurements can be expressed as a dot product via $|I + Q|^2 = \langle I + Q, I + Q \rangle$, where I and Q are vectors. Thus, θ , the angle between I and Q can be found:

$$\theta = \arccos \frac{|I + Q|^2 - |I|^2 - |Q|^2}{2|I||Q|} \quad (3.11)$$

The same technique can be used to find the angle between I only and $I = Q$. This is possible to perform fully with binary cells and can then be bootstrapped for thermometer cells, such that all constellation points can be accurately measured using only amplitude.

A measurement of the DAC constellation using the extrapolation method detailed above is shown in Fig. 3.21. While only a small subset of the total codes have been exercised, the I/Q nonlinearity should be independent of output amplitude because it is largely an LO generation effect. It can be seen that there is I/Q nonlinearity, both in Q LO mismatch relative to I , as well as I/Q summation nonlinearity. Q mismatch in this measurement is 3.3, while I/Q summation magnitude mismatch is 0.5% and I/Q summation phase mismatch is 4.3. From the discussion of digital predistortion and its relaxation of DAC linearity requirements, this DAC performance is well within the range to achieve good signal cancellation.

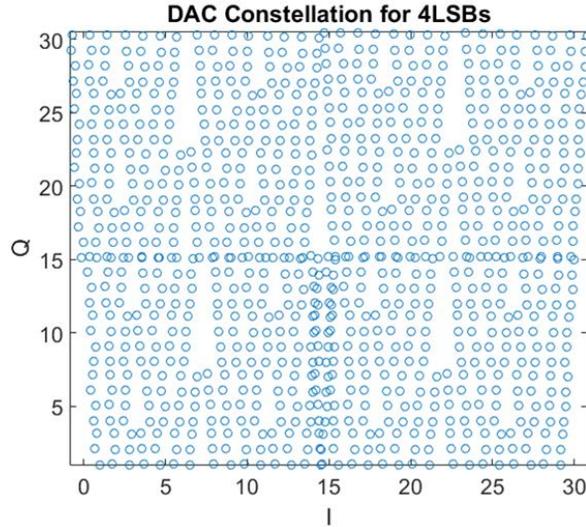


Figure 3.21: DAC quadrant I constellation up to 4 bits.

3.5.2 System Measurements

Connecting the TX and RX baluns together on chip, the full system can be measured. A single tone power sweep of the TX is performed while adapting the DAC code for cancellation and the residual is plotted relative to output referred TX power as the blue curve in Fig. 3.23a. Across TX output power, the residual remains constant, confirming that the LSB of the DAC sets the TX residual at the RX input. This means that the TX isolation is proportional to the TX output power, unlike conventional active cancellation or filtering approaches, which offer a fixed cancellation across TX output power. Moreover, the high cancellation power of this architecture is shown by the ability to cancel a +12.6dBm single-tone signal by >50dB, and order of magnitude greater power handling than prior active cancellation approaches.

Furthermore, cancellation measurements were performed using 20MHz modulated data, where the residual is plotted in black on Fig. 3.23a. Here, measurements were conducted in a two stage fashion. First, the corresponding DAC code for each TX symbol in the sequence is found using the same adaptation procedure as the single-tone case. Next, using this DAC sequence as a starting point, the sequence is adapted point by point to minimize the integrated energy in the 20MHz TX band. This second stage effectively accounts for the taps of the leakage channel as well as any dynamic nonlinearities present on the TX and DAC. The maximum modulated data signal handled before 1dB compression of the RX was a +12.6dBm peak, 6dB PAPR sequence. It should be noted that while the average modulated data residual is approximately the same as the average single tone residual, this is simply a coincidence, as the residual is a function of both the number of exercised bits, as well as the oversampling ratio of the data. In measurement, the oversampling ratio on the DAC and TX data is increased from 3.125x to 6.25x to 12.5x, and each time a 3dB decrease is seen in the residual power in the 20MHz band, apparent in Fig. 3.22. Utilizing 9 bits

of DAC dynamic range and 12.5x data oversampling, a maximum cancellation of 64dB was measured. This matches up very well with the expected cancellation of

$$6 \frac{\text{dB}}{\text{bit}} \times 9\text{bits} + 10 \log_{10} 12.5 = 65\text{dB} \quad (3.12)$$

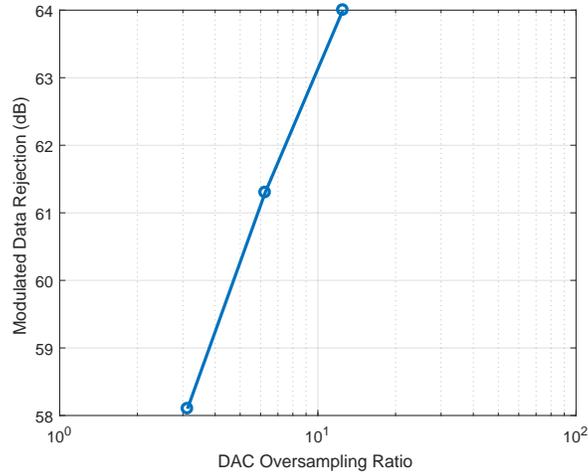


Figure 3.22: TX rejection vs. DAC oversampling ratio.

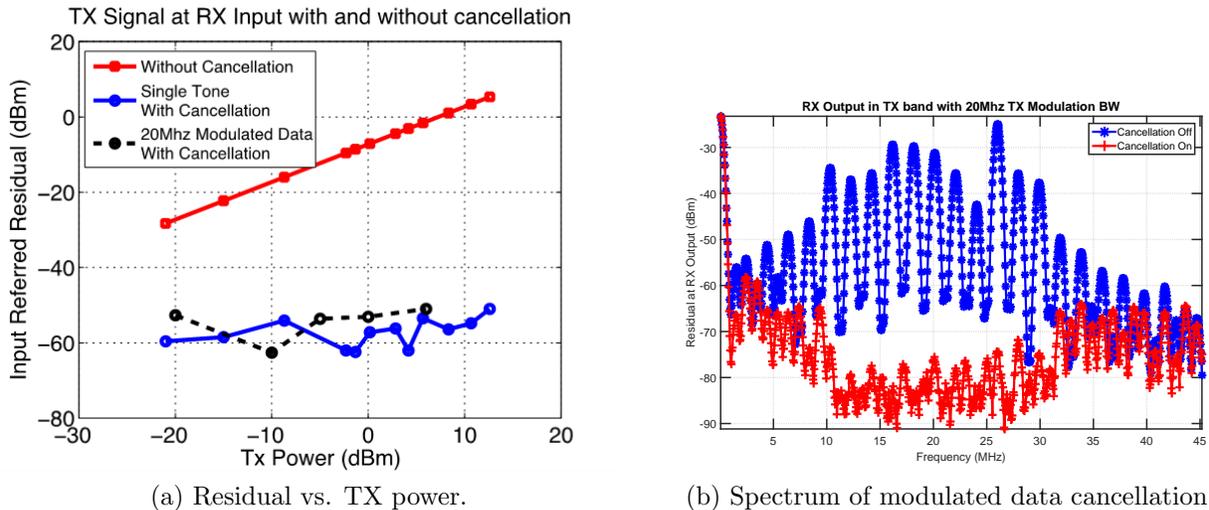


Figure 3.23: Initial TX cancellation measurements.

In the measurement shown in Fig. 3.23b, a sequence consisting of 10 tones across a 20MHz bandwidth was sent through the TX and adapted with the DAC. Cancellation of all 10 tones to similar residual power levels was achieved, showing that signal cancellation takes place equally across the TX bandwidth. The format of 10 tones vs. a general wideband signal was

simply due to limitations in the testing setup, where the FPGA memory was set to only 128 symbols at 250MS/s.

To determine the bandwidth over which cancellation is effective, a sweep of the TX/RX LO at a fixed output power of 0dBm was performed, where at each frequency point, the DAC code was re-adapted. Results are shown in Fig. 3.24. The residual power remained constant over the TX frequency sweep from 1-1.8GHz, due to the high bandwidth of the virtual ground current subtraction node. The limited bandwidth of the RX matching network does not affect the bandwidth of cancellation; instead it simply changes the amplitude/phase shift of the TX current. This difference can be compensated for by re-adaptation.

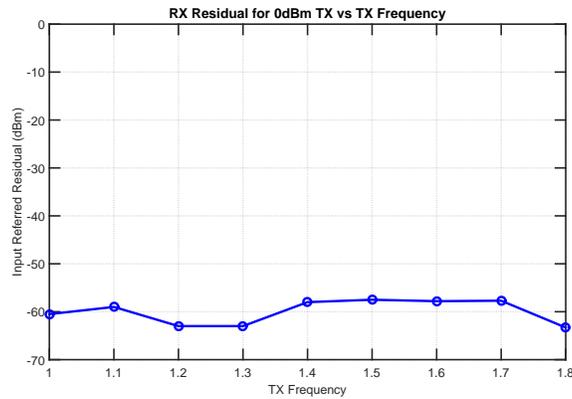


Figure 3.24: Residual vs. TX frequency.

A major advantage of the cancellation DAC over purely analog cancellation architectures is the ability to handle large ranges of leakage channels while maintaining high rejection levels. This strength is emphasized by the measurement of cancellation over antenna VSWR. In this measurement, with setup shown in Fig. 3.25a, an antenna tuner, in the form of a sliding short in parallel with a 50Ω calibration standard, was used to vary the antenna impedance up to 5:1 VSWR. For all points along the sweep range, >50dB of cancellation was achieved after re-adapting, shown in Fig. 3.25b.

Another important metric for the self-interference cancelling transceivers is the RX NF degradation due to the thermal and phase noise of the DAC and TX. Single tone cancellation of the TX signal is performed in a power sweep up to +10.6dBm and the associated RX NF degradation is plotted against this sweep in Fig. 3.26. Because of the 1dB compression at +12.6dBm, the TX power was backed off for a more fair comparison of noise degradation. This measurement can be fit to constant, thermal, and phase noise curves in order to infer the more detailed noise performance. The detailed noise performance summary is that of panel a) of Fig. 2.18. This measured curve suggests an uncorrelated phase noise profile of -180dBc/Hz at 40MHz offset.

Further measurements were performed to showcase feedforward cancellation of the transmitter phase noise. Phase noise of varying bandwidth was injected into the TX LO input by using an external noise source with filter. Power combining this noise signal with the LO

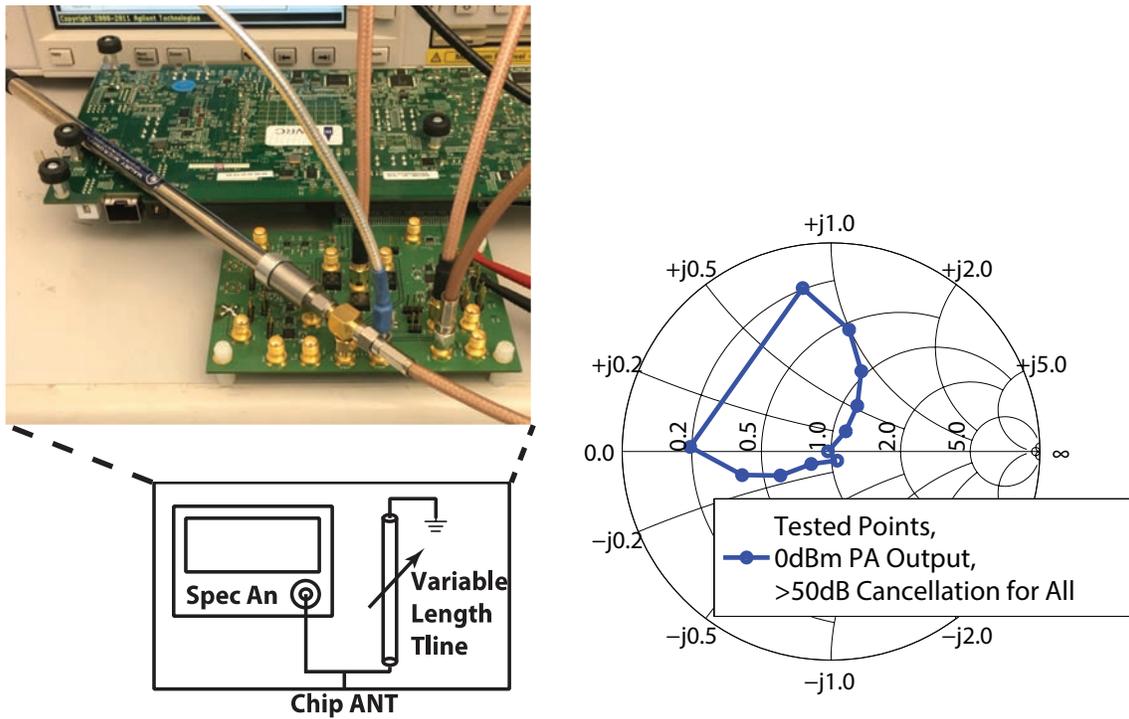


Figure 3.25: VSWR test setup and results.

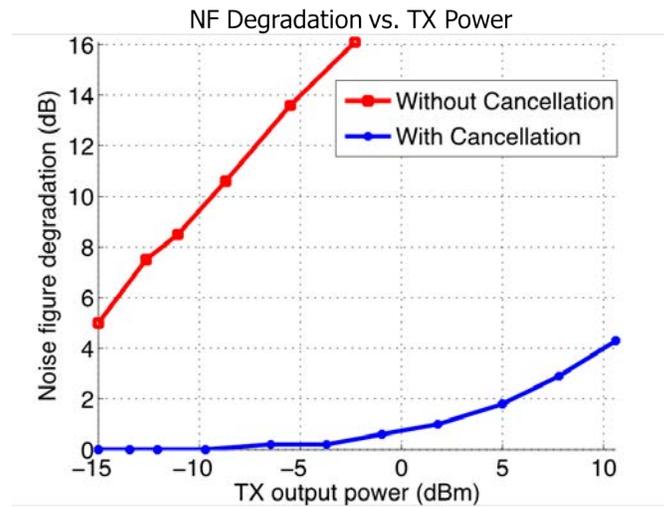


Figure 3.26: NF degradation vs. TX power.

and feeding it into the limiting buffer chain creates an LO with phase noise only because the amplitude is unchanged at the output of the limiter [63]. This injected noise is a much higher power than the intrinsic noise of the system, allowing accurate verification of phase noise

cancellation, diagrammed in Fig. 3.27. First, white noise was filtered to 100MHz bandwidth with a sharp rolloff, preventing any noise at higher harmonics from folding back down. The result, shown in red in Fig. 3.28, follows the same contour of the cancellation of single tone spurs added to the LO. For a duplex spacing of 40MHz, the closest of the LTE FDD bands, 20dB of phase noise cancellation is observed. Next, the TX and RX were isolated from one another and a meter long cable was connected between the two, emulating a channel with large group delay. The results, shown in Fig. 3.29, display significant reduction in phase noise cancellation bandwidth due to the large frequency-dependent phase shift of the cable. Additionally, it can be seen that past 40MHz, the “cancelled” phase noise is larger than the phase noise without cancellation, due to the fact that the correlated phase noise between the TX and DAC can add constructively, rather than destructively with enough phase shift.

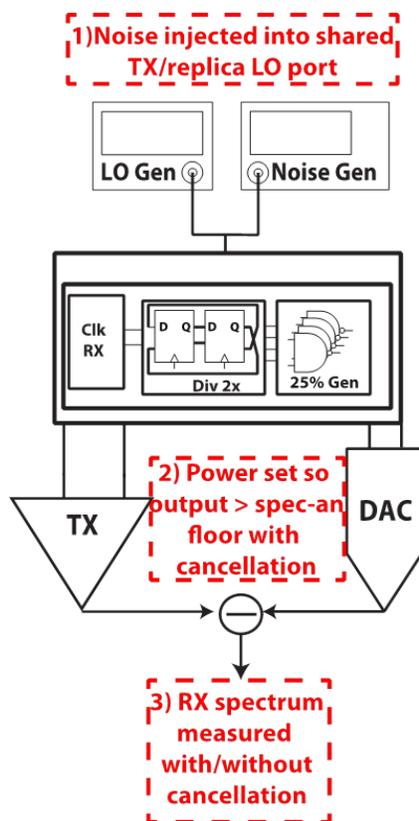


Figure 3.27: Test setup for phase noise cancellation measurement.

If wideband noise is injected into the TX LO, shown in black in Fig. 3.28, cancellation falls off and levels out to a lower value due to noise at higher harmonics of the LO folding down with a different phase shift. The mechanism causing this phase noise cancellation limitation is detailed in [48], but this effect can be condensed into a measurement showing that tones injected to the 2x LO input have different phase relationships than 90° between I and Q outputs. In Fig. 3.30, a tone is injected into the LO at an offset from some harmonic

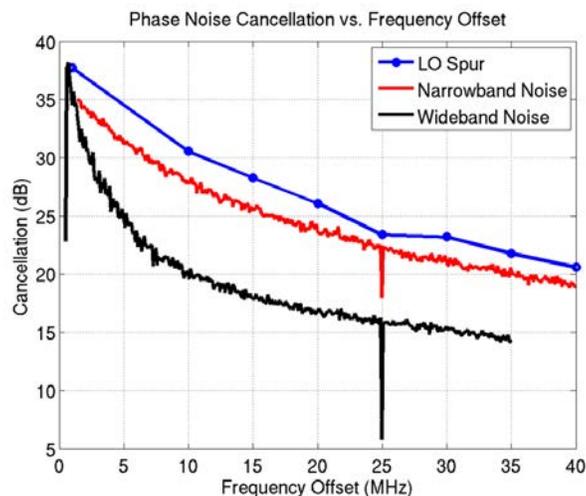


Figure 3.28: Phase noise cancellation measurement with single-tone, narrowband, and wide-band injection.

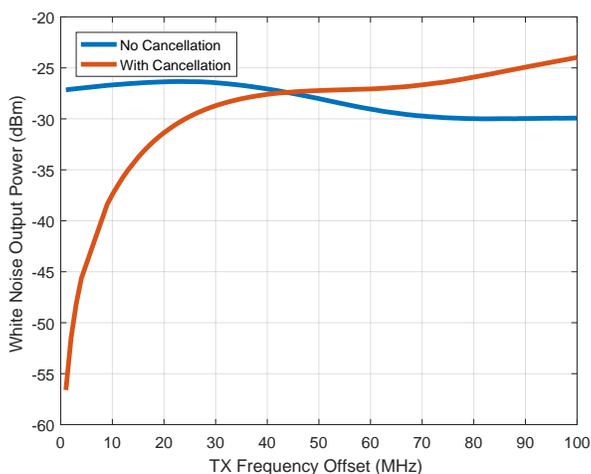


Figure 3.29: Phase noise cancellation with 1m cable.

of the LO, then the output of the PA is measured with an I only symbol and a Q only symbol, and the relative phase of the spurs, as well as their conversion gain, is checked. Shown in Fig. 3.31, spurs around $2F_{TX}$, $6F_{TX}$ fold with the correct $+90^\circ$ phase shift, while spurs at $4F_{TX}$ folds with a -90° shift. The conversion gain of the $4F_{TX}$ spur is -20dB , which when constructively rather than destructively combined with the DAC, can lead to the 15dB cancellation floor seen in Fig. 3.28. Note that the higher levels of cancellation close to the TX frequency is due to the Lorentzian phase noise spectrum of the source LO, where close-in phase noise dominates over the far-out white phase noise, obscuring cancellation limits due to folding at small frequency offsets.

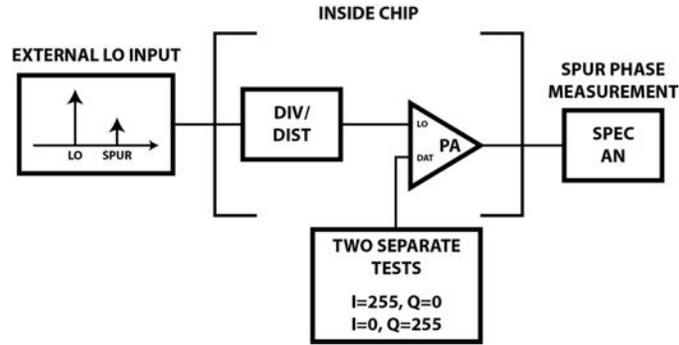


Figure 3.30: Test setup for phase noise folding measurement.

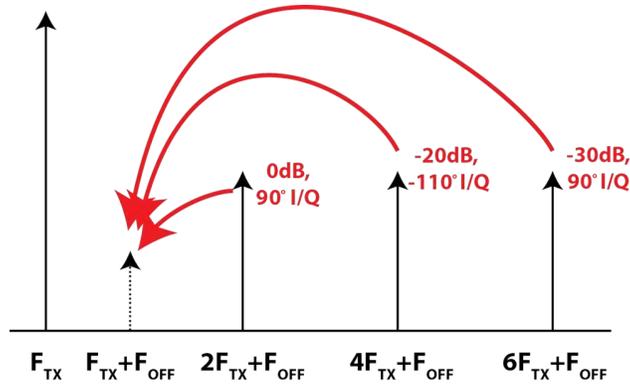


Figure 3.31: Measurement results for LO spur injection relative phase and amplitude.

Compared with prior art at the time of publication in Table 3.1, this work significantly outperforms all but one prior art for maximum power handling capability. Compared with [51], which can handle up to +14dBm TX power before compressing, this work achieves 25dB higher cancellation of single-tone and modulated data, and has far lower noise figure at low TX power levels. It is also worth noting that while some works either are incapable of performing modulated data cancellation [41] or achieve significantly lower cancellation for modulated data [40], this work can actually achieve higher average cancellation for modulated data vs. single-tone due to the quantization noise limited regime that the canceller operates in.

3.5.3 DAC Thermal Noise Cancellation

To measure DAC thermal noise cancellation, the power of thermal noise was required to overcome phase noise. To reduce the power of the phase noise, a full-duplex scheme was implemented, where the TX/DAC and RX clocks were shared through a splitter. The DAC, TX, and RX LOs are all generated from the source LO in the same manner, meaning that the source noise propagation is identical, with the only difference being delay between the

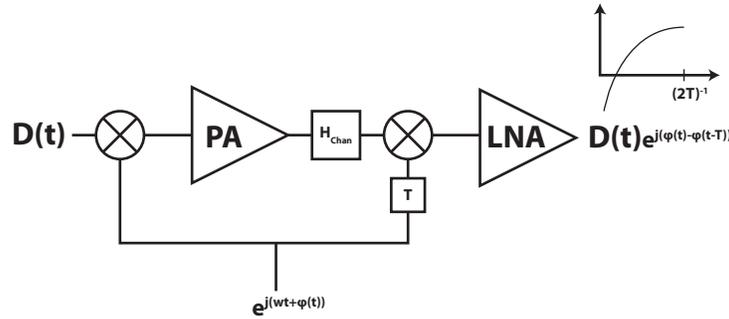
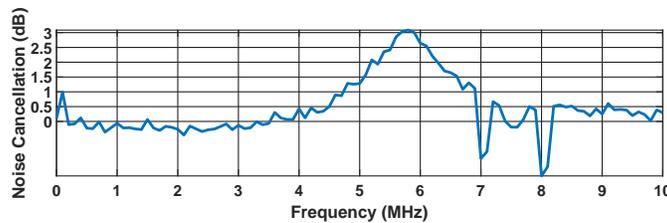


Figure 3.32: Source phase noise cancelled through sharing LO between TX/RX.

separate paths. Mixing this LO with the TX/DAC residual, the phase noise at low offset is cancelled, illustrated in Fig. 3.32. Accordingly, only phase noise from unshared buffers on the different LO paths will contribute, which is a much lower level than that of the source noise, allowing thermal noise to dominate. Next, not only must thermal noise dominate, but the DAC code must be high enough such that a large portion of the noise current at the center tap is due to active unit cells. To ensure this, the current in DAC unit cells was lowered considerably, such that at moderate TX powers high DAC codes are used for cancellation.



(a) Thermal noise cancellation (6MHz TX offset).



(b) Thermal noise cancellation spectrum.

Figure 3.33: DAC thermal noise cancellation measurements.

In measurement, shown in Fig. 3.33, up to 3dB DAC thermal noise reduction was found, with a 1dB bandwidth of 2MHz. Due to the requirements of low-offset measurements, the bandwidth is smaller than it would be for the same Q at 40MHz offset, where the bandwidth would be 15MHz.

	This Work	Columbia, ISSCC 2014 [41]	Columbia, ISSCC 2015 [40]	Twente, ISSCC 2015 [42]	Columbia, ISSCC 2016 [20]	Cornell, RFIC 2016 [51]
Technology	65nm	65nm	65nm	65nm	65nm	65nm
Frequency (GHz)	1.0-1.8	0.3-1.7	0.8-1.4	0.15-3.5	0.6-0.8	0.3-1.6
Max TX power leakage (dBm)	+12.6	+2	+8	+1.5	-6	+14 ¹
Cancellation at max TX power (dB)	>50	>30	33	>27	42	>25
Cancellation, 20MHz BW (dB)	>60	-	20	27	42 ²	>25
RX NF (dB)	7.6	4.2	7.5 ³	6.3	5.0	8.0
NF degradation at +2dBm TX, 40MHz offset	1.1	0.8 ⁴	0.9 ⁴	4.0	5.9	4.0
Fully integrated TX/RX	Yes	No	No	Yes	No	Yes
Single antenna, no external isolation	Yes	No	No	No	Yes	Yes
Canceller power (mW)	60	13-72	44-182	Not reported	89	*

¹ With RX-band TX degeneration

² 12MHz modulation bandwidth

³ With LC duplexer loss included

⁴ TX power not reported

Table 3.1: Comparison with prior art.

3.6 Digital Cancellation Measurements

Having measured the analog performance of the chip, it was also necessary to test the potential for using digital cancellation to further reduce the TX interference level, as motivated in Section 2.10. In this section, the modeling process for the DAC and TX alone is introduced, and a series of refinements to the leakage network are presented, ending with results of digital cancellation for removing the TX and DAC quantization residual when they operate simultaneously to provide analog cancellation.

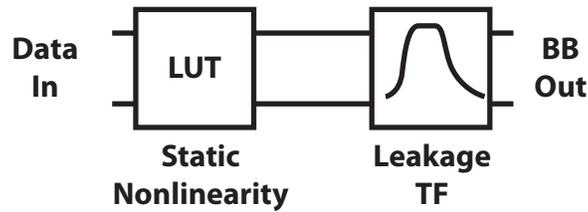


Figure 3.34: DAC model.

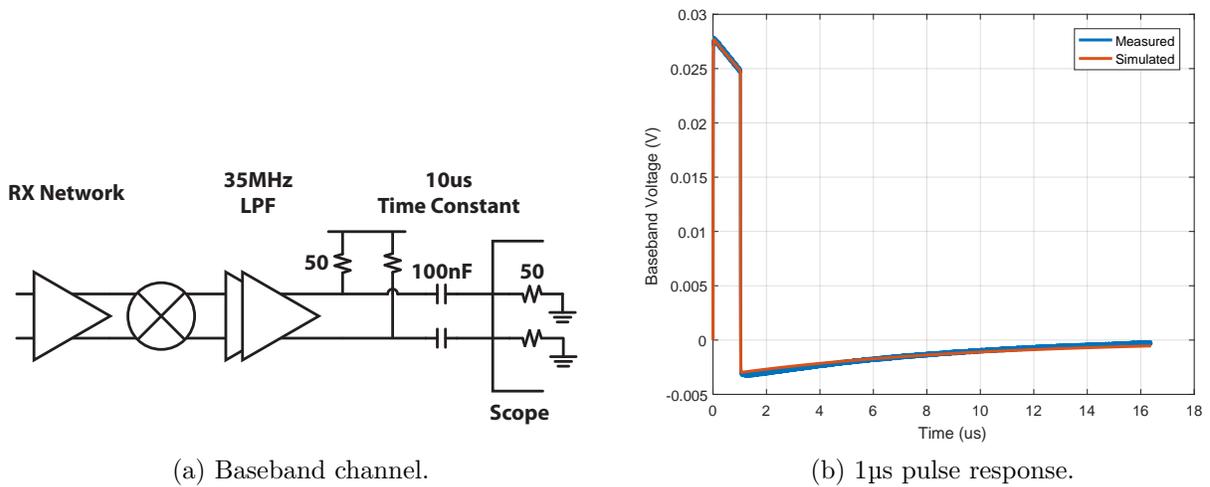


Figure 3.35: Baseband channel and effective baseband pulse response.

In the beginning of the modeling process, a very simple two-stage digital model was used, diagrammed in Fig. 3.34, where the first block is a lookup table accounting for static nonlinearity in the TX/DAC, and the second block is a complex equivalent baseband transfer function, representing the leakage channel. First, the channel response measurement will be discussed, as it is the less complex of the two. The DAC/TX output pulses of either current/voltage, which are later sampled by the RX ADC. Therefore, sampling the pulse response of the leakage channel for both is sufficient to provide all the channel information required in this simple model. It is initially seen that the pulse response has a long tail, which in fact is due to the baseband processing network, most notably the AC coupling capacitor at the baseband outputs, diagrammed in Fig. 3.35a, rather than the RF leakage channel itself, as shown by the comparison in Fig. 3.35b. This simulation assumes a 35MHz baseband amplifier bandwidth, as well as a 10 μ s time constant high-pass filter due to the AC coupling capacitor. This very long tail significantly constellation measurements for 64ns held data, as shown in Fig. 3.37a, where the sampled values of an RX data packet are graphed.

Due to the highpass behavior of the RX baseband, it is not possible to measure each constellation point statically, so a method to calibrate out the channel effect on a packet of data

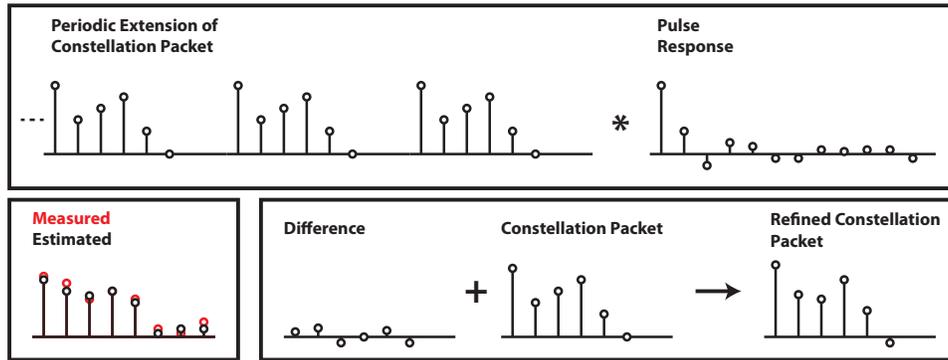
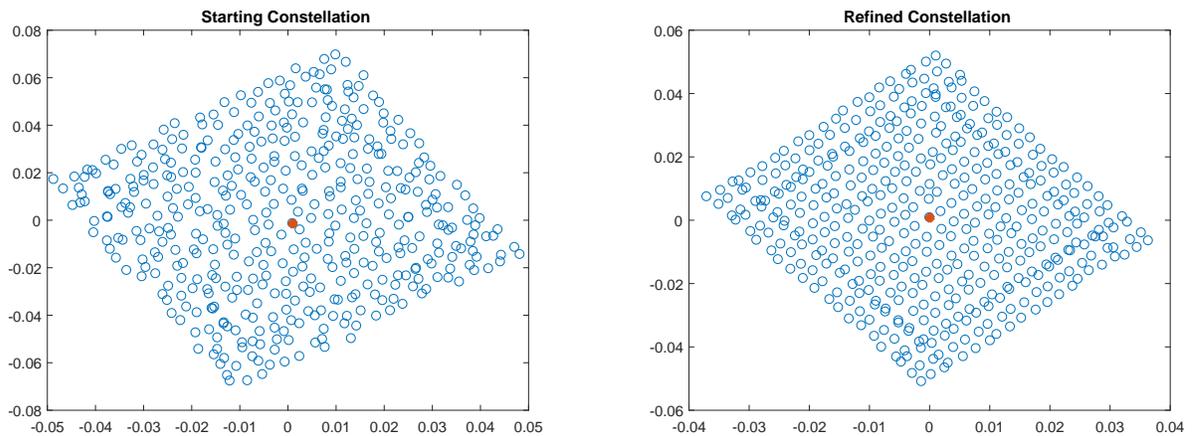


Figure 3.36: Constellation refinement procedure.



(a) Constellation before refinement.

(b) Constellation after refinement.

Figure 3.37: Comparison of measured constellations before after channel de-convolution.

is needed. Later, the AC coupling cap is removed to improve the channel, but a packet-based constellation measurement system is still preferred because it allows the constellation points to be measured with little measurement time overhead. The constellation measurement and calibration procedure is detailed in Fig. 3.36, and consists of transmitting a data packet containing every point in the desired constellation, then using the measured pulse response of the channel to subtract the ISI from the measured constellation sequence. After multiple iterations, the pulse response is de-convolved from the measured sequence, leading to a refined constellation in Fig. 3.37b, which has symmetric characteristics and clearly shows some amount of DNL, which are both indicative of a correct calibration procedure.

For generating a static nonlinearity lookup table equal in size to the full range of the DAC/TX, measuring each constellation point individually is a long process; measuring every constellation point in the DAC would require 17ms for a 4ns symbol period. Instead of this laborious process, the DAC and TX may be significantly subsampled because their unit

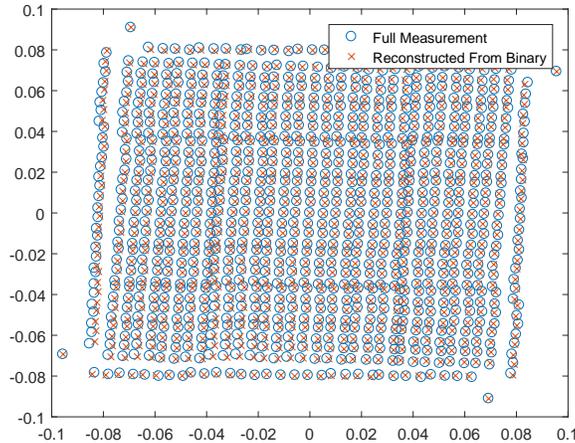


Figure 3.38: Comparison of full measured constellation with a reconstruction using 2% of the total measured points.

cells add together with high linearity. As stated earlier in Section 2.7, the major sources of static nonlinearity come from within the DAC unit cells, while their current summation at the RX input can be thought of as highly linear. The same argument can be made for the switched-capacitor PA in this architecture. By measuring the different unit cell output states for the top two quadrants (i.e. $1, j, -1, 1 + j, -1 + j$ for each cell), the rest of the constellation can be recovered with high precision, shown in Fig. 3.38. Here, a portion of the PA constellation is reconstructed with $>40\text{dB}$ precision, corresponding to 3 bits below the LSB. Similar reconstruction accuracy was found for the replica DAC. For the 10 bit DAC, only 180 points would need to be measured, less than 0.1% of the total number of constellation points. Since this would take less than $1\mu\text{s}$ with a 4ns symbol period, more points or processing could be added to improve the accuracy of this reconstruction by adapting a model for summation nonlinearity.

Next, a set of oversampled QPSK symbols were sent through the DAC and this sequence was compared with a simulated sequence using the calibrated static constellation, as well as the pulse response data. Using a reduced constellation region ($|I|, |Q| \leq 16$), this refined constellation and pulse response can give 47dB matching of a DAC sequence, corresponding to 4 bits below the DAC LSB Fig. 3.39.

PA Modelling

This simple mode works well for all codes of the DAC, due to its current-mode operation, but is only appropriate for the PA low power and begins to degrade as a higher average power is used, due to a nonlinearity inherent to the PA construction. This additional source of nonlinearity within the PA can be seen by considering its voltage-mode architecture. A switched capacitor power amplifier, chosen for its low and code-independent output impedance, can be

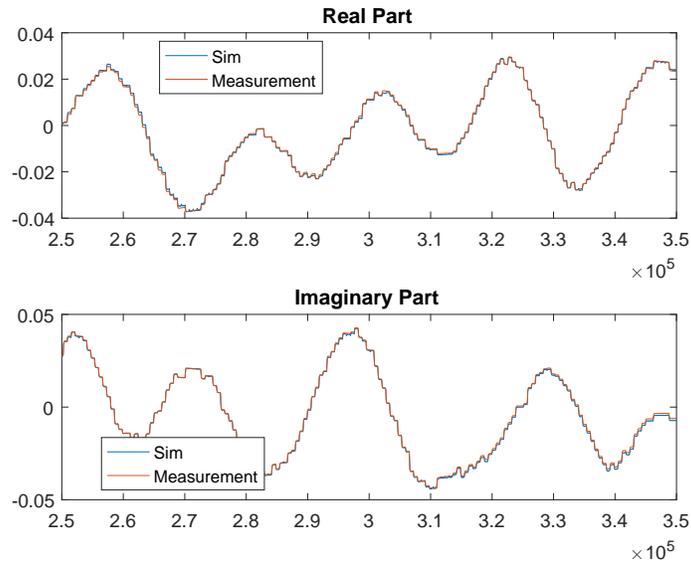


Figure 3.39: Matching of 15MHz data.

modeled as a tunable capacitive voltage divider connected to the PA supply. When connected to a supply with some ripple, $VDD(t)$, the ripple directly modulates the output waveform $O_{PA}(t) = \frac{VDD(t)}{VDD_{nominal}} O_{PA,ideal}(t)$. An illustration of this effect is shown in Fig. 3.40.

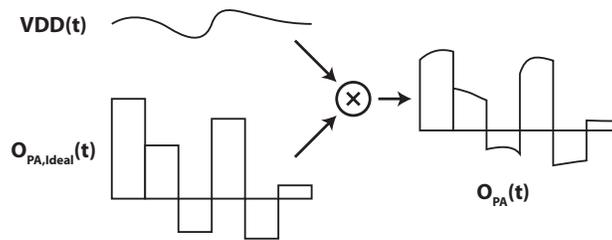


Figure 3.40: TX passes supply ripple to output.

This modulation effect means that for a 10% ripple amplitude on the supply, there will be a 10% inaccuracy in the PA output compared to no ripple. If the ripple were constant in amplitude with code, then there should be no difference in digital cancellation accuracy for high codes vs. low codes, since rejection is a relative measurement. In the case of the 10% ripple, the digital cancellation would be limited to 20dB for all codes.

In the case of a PA with an isolated supply, the ripple amplitude is a strong function of the PA output power due to series resistance in the supply network. In the presence of a series resistance on the supply, the voltage ripple is a linear function of the current draw. In

the case of a switched-capacitor PA, the current draw is proportional to $\sqrt{P_{TX}}$, or $|C_{TX}|$, where $C_{TX} = I_{TX} + jQ_{TX}$. This leads to a memoryless nonlinearity:

$$\begin{aligned} O_{PA}(C) &= \frac{C}{|C_{MAX}|} (VDD - R\alpha |C|) \\ &= \frac{C}{|C_{MAX}|} VDD - \frac{R\alpha}{|C_{MAX}|} C |C| \end{aligned}$$

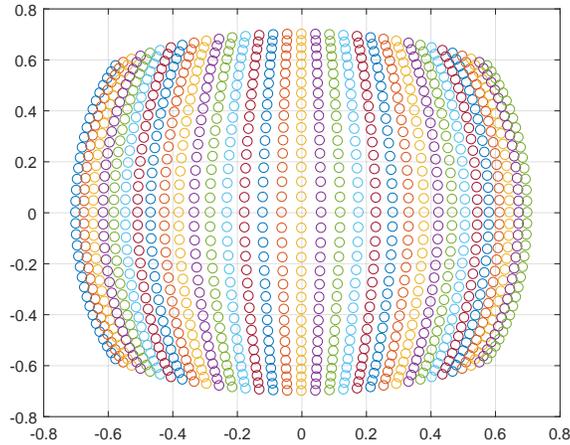


Figure 3.41: Simulated constellation with memoryless PA supply nonlinearity.

An exaggerated example of this memoryless nonlinearity is shown in Fig. 3.41. The voltage supply ripple is a function of the baseband PA code and also directly modulates the output waveform, so this effect is independent of the frequency of the PA. Realistically, there is significant memory on the power supply node. The presence of memory in the supply impedance does not prevent a purely baseband analysis of this nonlinearity, but it can no longer be simply modeled by a constellation distortion. To observe the effect of memory on the supply node, a 64 μ s is step applied in Fig. 3.42, where it is clear that there is a multiplicative effect of the supply ripple on the unit step waveform, and that the supply ripple has a large time constant, approximately 10 μ s, affecting hundreds of symbols before settling. This very long time constant is due to the large decoupling capacitors on the PA supply. While removing them would make the supply ripple shorter, making this nonlinearity closer to memoryless, it would add significant supply noise to the output of the PA, adding wideband interference to the RX band. Therefore, this memory effect should be modelled rather than removed.

To complicate the supply ripple further, for different unit step sizes in Fig. 3.43a, not only does the magnitude of the supply ripple increase with code, but the Q of the ripple changes

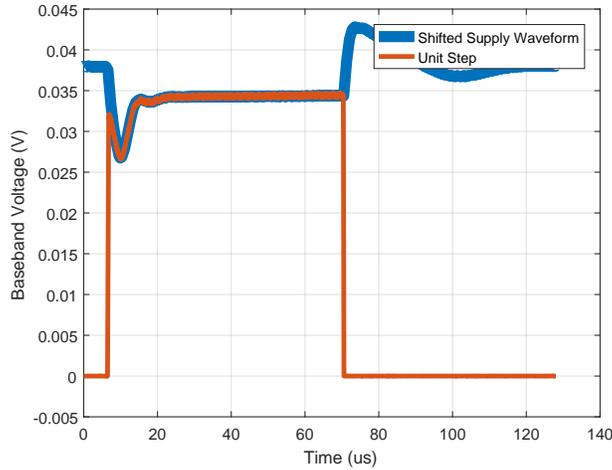
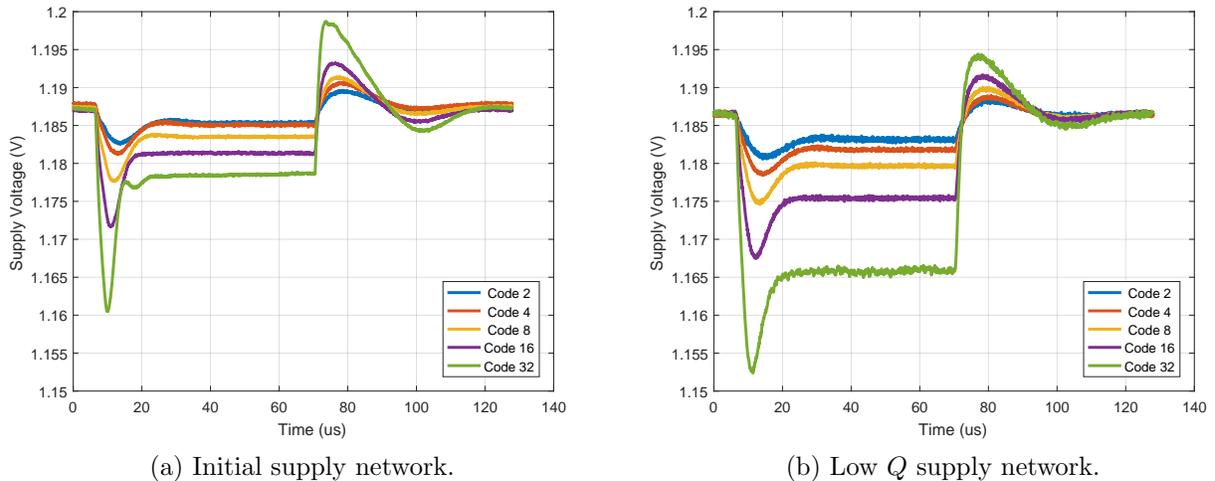


Figure 3.42: TX unit step compared with supply ripple.



(a) Initial supply network.

(b) Low Q supply network.

Figure 3.43: Effect of lowering PA supply network Q .

as well. This is due to the nonlinear output impedance of the LDO used in the supply network, which is difficult to model for digital cancellation. To simplify the supply network, a resistor is added to the supply, reducing the significance of the LDO output impedance, leading to a much more linear set of curves shown in Fig. 3.43b. While long unit steps are extremely rare in data-carrying transmissions, the unit step response looks similar to the supply response for when a burst of data is sent through the PA, as shown in Fig. 3.44. This is because the average current of the PA is an effective unit step response when modulated data begins transmitting, and the time constant of the supply network is long enough that spurious changes in code due to TX PAPR do not affect the supply voltage significantly.

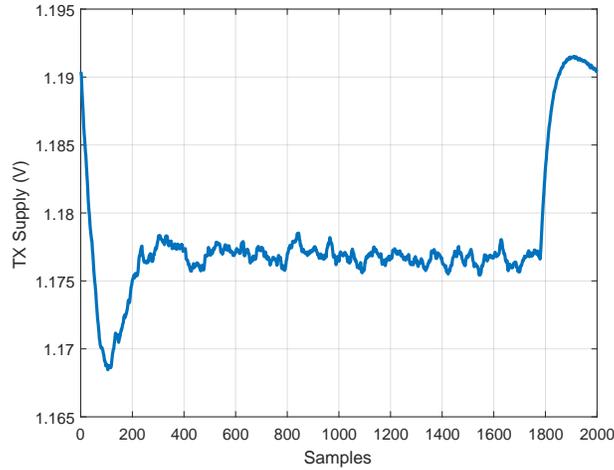


Figure 3.44: TX supply, arbitrary sequence with deadtime.

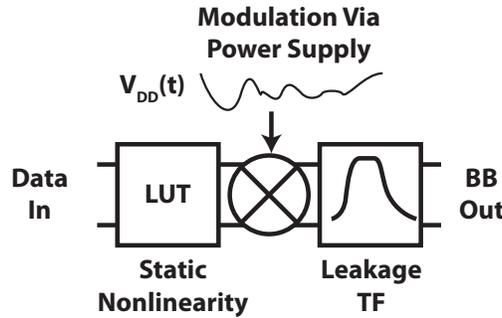


Figure 3.45: PA model including supply modulation.

The PA baseband model is modified in Fig. 3.45. To estimate the supply ripple affecting the PA, there are two approaches. First, a model can be made of the off-chip power supply network (at the level of precision required, there is no benefit to considering the on-chip passive network, which has far higher bandwidth), which could include both passive devices and nonlinear models of active devices, such as the LDO supplying the PA. This model, coupled with a model of the current draw of the PA given code, would provide an estimate of the power supply ripple as a function of the data sequence.

Another approach is to simply measure the power supply ripple. This can be done off-chip by probing the PA power supply pin, or a dedicated low-bandwidth power supply measurement device could be implemented on-chip. Measuring the power supply ripple allows for far lower model complexity, a strong advantage. One disadvantage is the fact that the ripple cannot be predicted, therefore this could not be used for predistortion unless an iterative adaptation loop is used.

In this work, measuring the supply ripple off-chip was used. The same procedure for

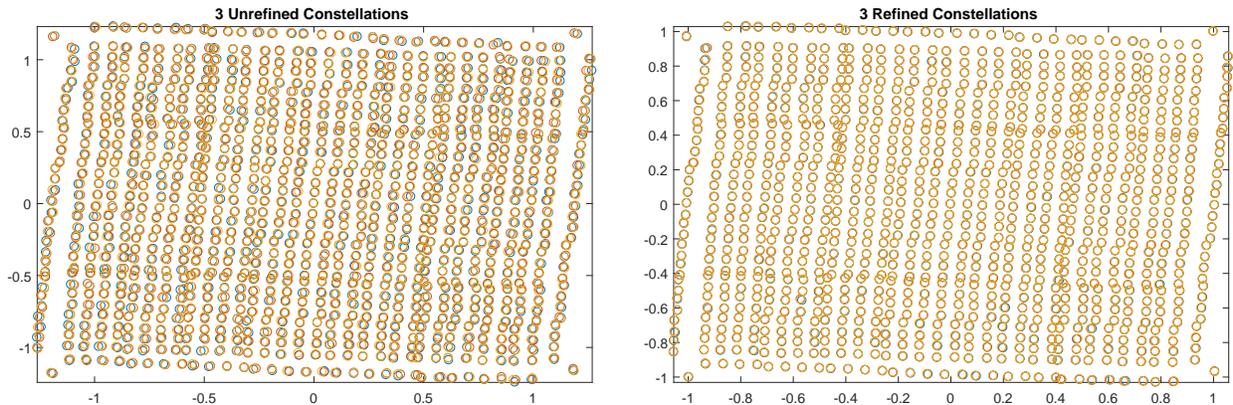


Figure 3.46: Constellation comparison using supply modulation model.

refining the constellation, Fig. 3.36, is performed, but the supply ripple is recorded while the constellation packet is output. Iteratively using this method with the new model provides far better matching in the constellations at higher codes, as shown in Fig. 3.46. Here, the AC coupling capacitor has been removed to significantly reduce the ISI in measurements. The small variation in initial constellation points is almost entirely due to the supply ripple differences. Using this new technique, 55dB of matching is achieved for 15MHz sequences from $(|I|, |Q| \leq 16)$ constellation, corresponding to 5 bits below the PA LSB.

TX+DAC Residual Modelling

In addition to exploring prediction of arbitrary data for the PA and DAC alone, prediction of the DAC/TX residual through direct measurement of the “residual constellation” was performed. This allowed prediction of much higher TX power levels because the linearity of the receiver was less of an issue due to cancellation of the fundamental. Here, constellation sequences were measured the same as before, except that both the TX and DAC were given nonzero codes, where the DAC sequence was chosen to cancel the TX sequence. For each TX code, a corresponding DAC code was found which canceled the TX code, assuming zero ISI. This dictionary was then used to cancel the TX to a level low enough such that the RX did not compress. A measurement of multiple residual constellations is superimposed in Fig. 3.47, and Fig. 3.48 shows a similarity of 25-30dB, corresponding to 4-5 bits of extra matching on top of the analog cancellation.

Given the complexity of the model used for the TX alone, it may seem surprising that simple measurement of the residual symbols from a constellation packet were enough to provide good cancellation, especially considering that 7 out of 8 bits of the PA were exercised. The major change in model complexity came from the fact that the residual sequences were measured without any deadtime between packets, which prevented the TX supply from rising back up to its nominal value. As can be seen from Fig. 3.44, after the initial supply

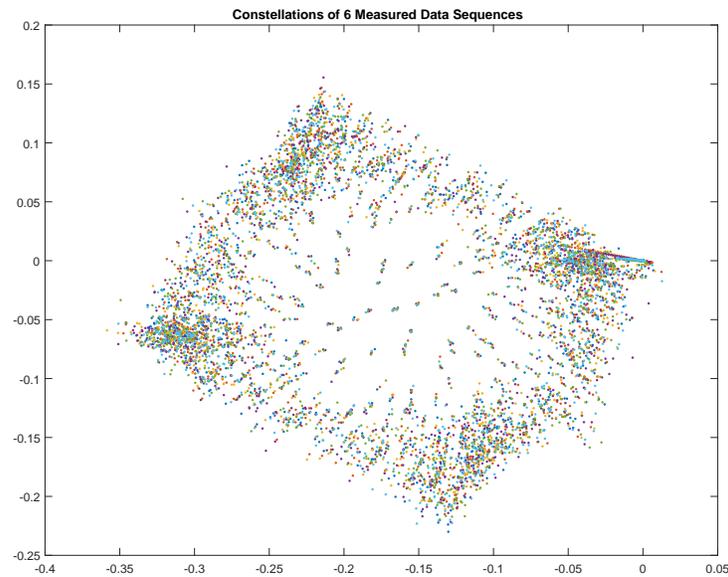


Figure 3.47: Comparison of multiple constellation packets.

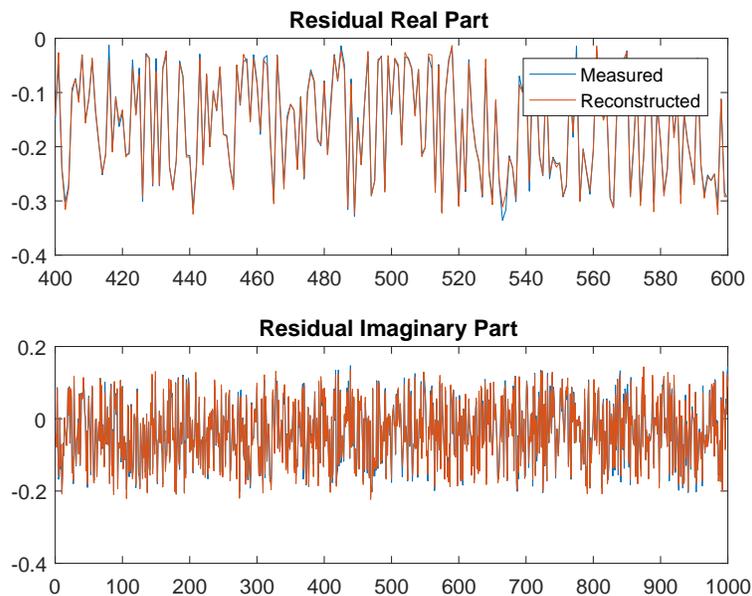


Figure 3.48: Comparison of arbitrary data measurement with reconstruction.

reduction, which spans many symbols, the supply ripple reduces considerably in amplitude. If more than 5 bits of digital matching are required, or in the instance where arbitrary

combinations of TX and DAC are required, then the supply swing dependence of the TX should be considered again.

Chapter 4

A Transceiver with >64 dB TX Signal Cancellation and Thermal/Phase Noise Rejection

4.1 Overview

The test results in the previous section were used to inform improvements to the transceiver design. First and foremost, a new design for the RX could leverage the cancellation architecture for lower overall noise figure. Secondly, the efficiency of the matching network and its bandwidth can be improved by reducing the DAC output capacitance as well as adding tunability to the available gain and matching impedance of the network. Finally, while the fundamental of the TX interference was cancelled to a low and constant residual, there was still significant harmonic content which was uncanceled due to differing transfer functions between F_{TX} and $3F_{TX}$, necessitating an RX and passive network which can mitigate these higher harmonics.

In the sections below, each point of improvement is detailed and measurement results follow.

4.2 Passive Mixer First RX

A passive mixer first receiver was designed in order to significantly improve the nominal noise figure, improve transceiver linearity, and improve the S_{21} bandwidth. A detailed account of the design tradeoffs and topology motivations are given in [48], but a short summary of some aspects of the RX chain are given below.

As depicted in Fig. 4.1, the passive mixer first architecture utilizes a TIA as the first stage, which both provides gain and the desired input impedance. By using a voltage-mode amplifier in the feedback loop, rather than a transconductance stage, the input matching constraint is independent of the fundamental noise figure. This is because for a voltage-mode

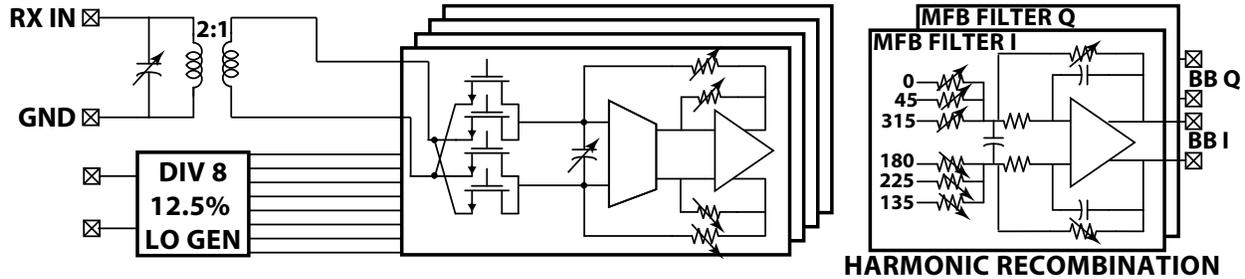


Figure 4.1: Top level of passive mixer first RX architecture.

shunt-shunt-feedback stage, $Z_{In} = R / (1 + T)$, while the noise is equal to $4kT \frac{R_{In}}{A} \Delta f$, where the feedback resistor noise dominates. This is contrasted with a transconductance stage, where the input-referred voltage noise is equal to $4kTR_{In} \Delta f$, such that the noise figure is limited to 3dB.

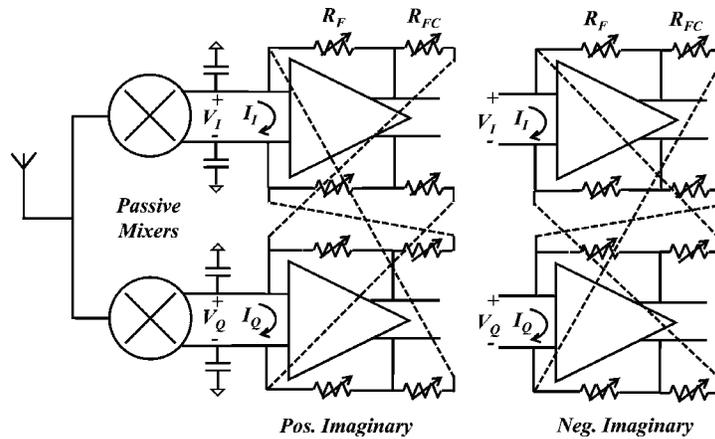


Figure 4.2: Baseband cross-coupling for imaginary input impedance synthesis [64].

Using a 2-stage design with a class-AB output stage, the mixer-first design can achieve <2dB NF in simulation. The class-AB stage also significantly improves the linearity of the stage, getting .

Explained in detail in [48, 64], the passive mixer first architecture has an additional benefit in that it can match to complex impedances, offsetting the impact of RX input node capacitance from the replica DAC. This ability is gained through cross-coupling the I and Q baseband paths, as shown in Fig. 4.2. Considering the I path, baseband current from the Q path is injected into the baseband I input. Due to being mixed by a quadrature clock, this current has a 90° phase shift, which is then upconverted to RF as an imaginary admittance component.

The second baseband stage uses a harmonic recombination architecture to cancel the 3rd and 5th harmonics of the TX/DAC. Using an 8-phase mixing architecture, the phases 0° ,

+45°, -45° are added together, with the 0 phase having $\sqrt{2}$ larger weight than the other two, producing an output which nulls the residual downconverted by the 3rd and 5th harmonics.

4.3 TX/RX Passive Network

The improvement to the intrinsic noise figure of the receiver chain makes an improvement to the passive network loss highly desirable. There are two main points for improvement: best-case G_A and bandwidth. In the first version, the matching network was optimized by assuming an input impedance to the RX and tuning the network based on no other considerations. For the second version, a co-optimization of the matching network and RX was performed, due to the passive-mixer first RX's ability to match to complex impedances. This ability, explained in the previous section, is highly desirable for a network with significant passive parasitics, such as this cancellation network, though it is not perfect. Matching the RX to an imaginary part has significant penalties for the intrinsic noise figure, as shown in Fig. 4.3. It is therefore desirable to minimize the imaginary part of the matching impedance across the receiver bandwidth. A capacitor DAC can be used to tune the imaginary part in the network.

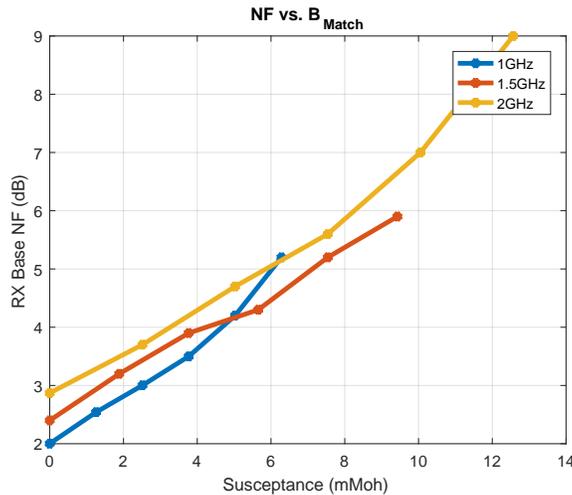


Figure 4.3: NF vs. susceptance for $R_{Match} = 200\Omega$.

Before sizing the DAC, the ideal system location for the device must be determined. There are two logical choices for the location, which are detailed in Fig. 4.4: the primary or secondary of the RX balun. The total system noise figure can be approximated by an addition of the RX intrinsic noise figure with the available gain of the matching network:

$$NF_{Total} = NF_{Intrinsic} - G_A \tag{4.1}$$

The intrinsic noise figure of the RX can be simplified by assuming it is roughly independent of required matching resistance. This is because the receiver chain is designed such that

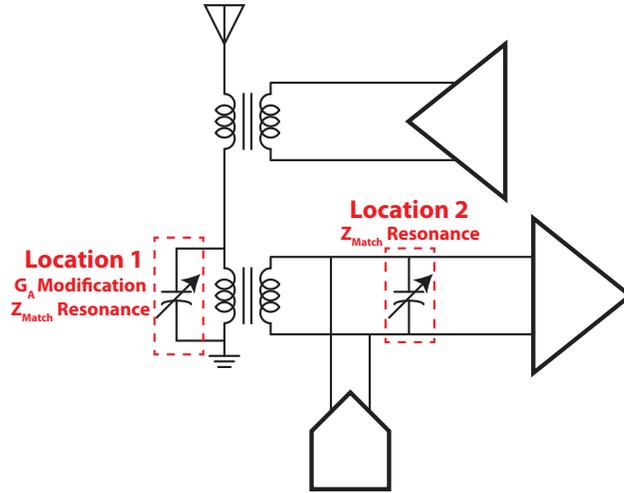


Figure 4.4: Possible locations for capacitor DAC.

the noise from the feedback resistor dominates, and in order to maintain current consumption and linearity while tuning the impedance, only the feedback resistance is changed. If the matching impedance is halved, then the resistor is halved, producing twice the current noise variance, but this is compared with a halved antenna resistance referred to the input of the RX, so there is no change in noise figure.

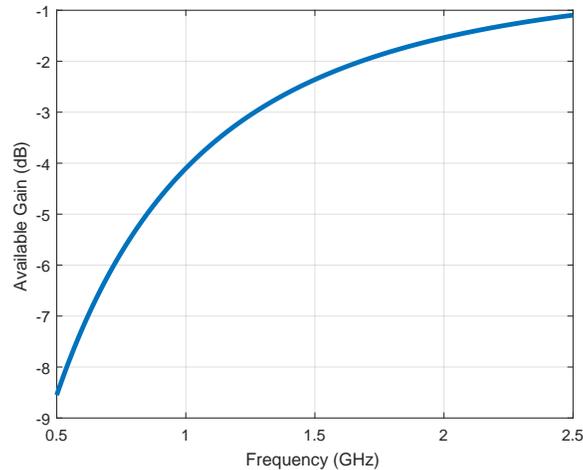


Figure 4.5: Representative 1:2 transformer available gain ($Q_P = Q_S = 10, k = 0.9$).

It is then the capacitor DAC's effect on G_A which matters most. If the capacitor DAC is placed on the secondary of the RX transformer, it can resonate with the inductance of the RX balun in order to create a purely real matching impedance, but G_A is unchanged

because the available gain is not a function of the impedance at the output port. Using a lumped transformer model, its G_A falls off as Fig. 4.5, which does not assure low noise figure across a large bandwidth. The second case, where the capacitor DAC is on the primary side, both resonates with the inductive component of the balun impedance, but also modifies G_A , allowing for a large bandwidth of operation. In Fig. 4.6, the system noise figures for these two locations are shown, where the capacitor DAC is assumed to have infinite Q . The capacitor DAC on the primary side, while requiring approximately 4x the area, produces a noise figure >0.5 dB lower than the case of a secondary capacitor DAC across a wide bandwidth.

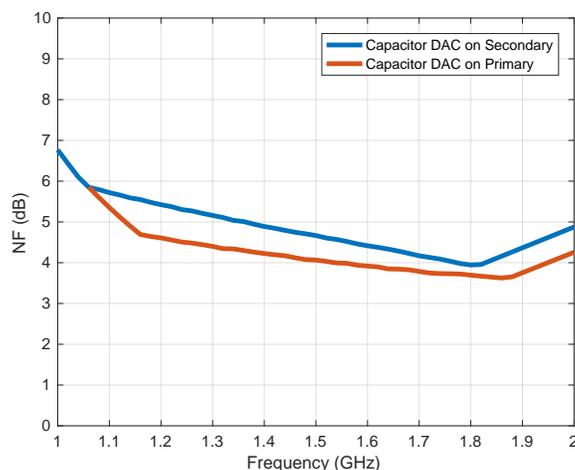


Figure 4.6: Total NF vs. capacitor DAC location.

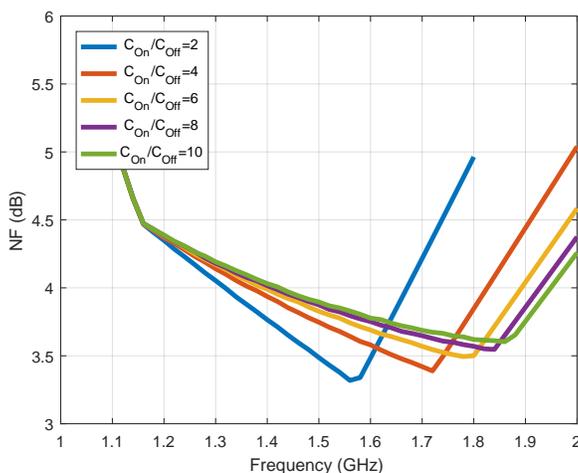


Figure 4.7: Total NF vs. C_{On}/C_{Off} .

The ratio between C_{On} and C_{Off} simply sets the bounds for low system noise figure. It can be seen from Fig. 4.7 that the range over which the system noise figure is low is proportional to C_{On}/C_{Off} . A value of 10 for the ratio provides acceptable bandwidth. Furthermore, the number of bits in the capacitor DAC sets the range of variation for the system noise figure, as shown in Fig. 4.8. This makes intuitive sense because in the frequency gap between resonance at one DAC code and the next, the imaginary part of the matching admittance will become nonzero, degrading the noise figure of the RX when it compensates for this effect. The lower the bits, the larger the gap, and therefore the larger the increase in noise figure. For 4 bits in the capacitor DAC, there is $<0.15\text{dB}$ variation in noise figure, an acceptable range. Finally, the effect of finite Q in the capacitor DAC is shown in Fig. 4.9. Past the point of a Q of 8, there is not a significant improvement in system noise figure.

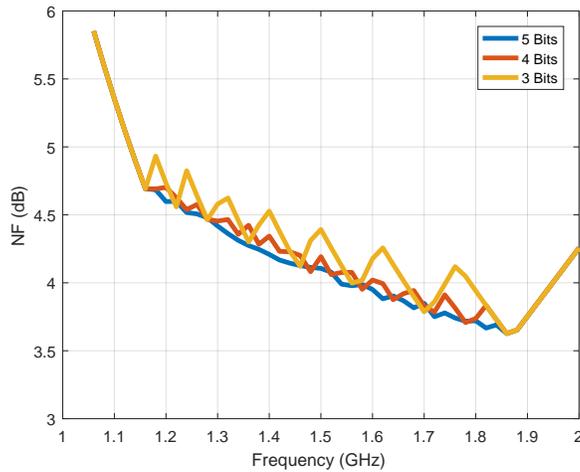


Figure 4.8: Total NF vs. capacitor DAC bits.

In addition to tuning the RX available gain transfer function over a wide bandwidth, the capacitor DAC can be used to create a tunable harmonic trap for the TX third harmonic. A small on-chip inductance is placed in series with the capacitor DAC to resonate around the third harmonic, while not significantly affecting the RX noise figure. Given the choice of balun inductance, a 200pH series inductor is required to resonate at 3x the fundamental while still maximizing G_A and minimizing the system noise figure. Shown in Fig. 4.10, assuming a conservative Q of 4 at 1.5GHz, there is not a significant increase in the total noise figure due to the introduction of this inductance.

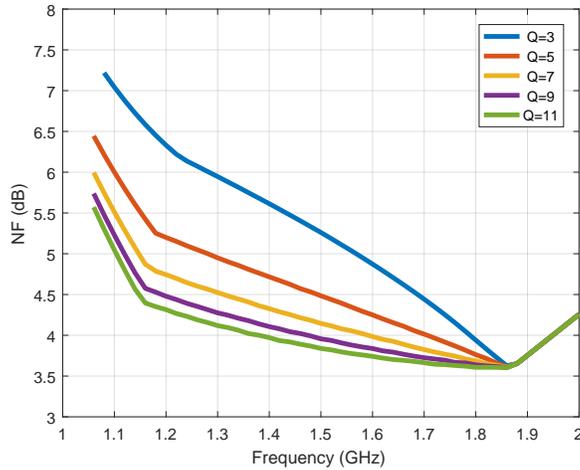


Figure 4.9: Total NF vs. capacitor DAC Q .

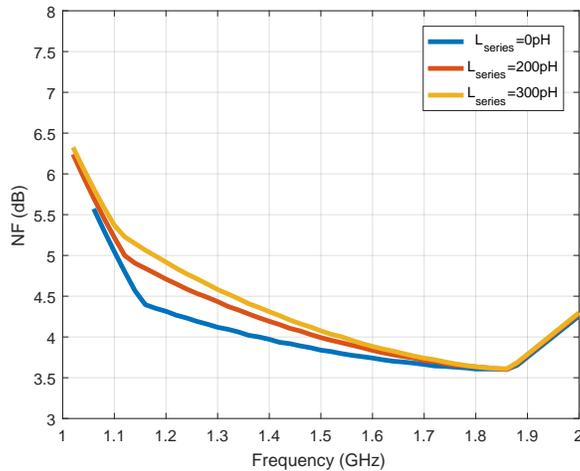


Figure 4.10: Total NF vs. harmonic trap series inductance.

4.4 RX Capacitor DAC Design

As shown in Section 4.3, lower noise figure is the main advantage of locating the capacitor DAC on the antenna side of the RX balun. The main disadvantage of the antenna side is that harmonics of the TX/DAC can cause large voltage swings across the RX balun which would otherwise not appear across the secondary side. In simulation, a $+20\text{dBm}$ TX output can cause a 1.7V amplitude swing the top of the capacitive DAC. The switches for enabled capacitor DAC cells do not experience this voltage swing, as a Q of 10 (chosen in Section 4.3 for low noise figure) ensures that the voltage swing across the resistive portion of the series RC is $<200\text{mV}$ in amplitude, well within the acceptable range of a 1.2V device. It is possible

to damage the disabled cells though, since the high impedance of the disabled switch causes it to see the full voltage swing Fig. 4.11.

Additionally, due to the RX negative node being pinned out to the testing PCB for ease of testing the RX in isolation, approximately 200pH of inductance is present between this node and ground. The 1.7V amplitude swing is not relative to the ground node but instead relative to this swinging node. Due to higher harmonics, this node swings 500mV amplitude relative to ground. Across the off switches, 2.4V peak-to-peak is present, and the bottom node is capable of going 500mV below ground.

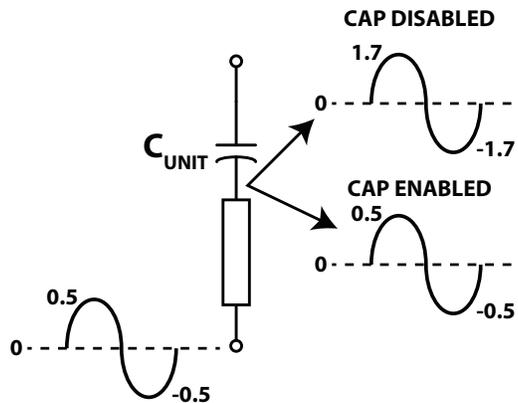


Figure 4.11: Capacitor DAC unit cell voltage swing for $+20\text{dBm}$.

To evaluate the extent that this voltage swing damages the off switches, the mechanisms by which a transistor can be damaged should be taken into account. In [65], a comprehensive study of these mechanisms is performed, with the main cause of damage being hot carrier injection into gate oxide. Electrons can tunnel into the gate oxide with a very low probability, but high V_{GS} causes a very high density of electrons to appear in an NMOS channel, making tunneling occurrences happen more frequently. These tunneled electrons serve to lower the potential barrier and further increase the likelihood of tunneling events. Eventually, the oxide is damaged enough that a conductive junction is created between the gate and source, permanently shorting the transistor. An important point is that if no channel is present (transistor in cutoff or accumulation mode), then the transistor is far more resilient to high voltages than if a channel were present. If the transistor can be kept in cutoff or accumulation over the entirety of the voltage cycle, it is possible to use a single transistor in the capacitor DAC, improving C_{On}/C_{Off} for a given Q . While this option is attractive, the cost of mis-design is far higher in a one transistor configuration, rather than two, so a two transistor stack was chosen for this chip.

In the two stack configuration, the total swing can be split between the two devices and brought to safe levels. First, both off switches are given a large explicit capacitance between gate and source and are biased with high resistance voltage sources in order to keep V_{GS}

below threshold for the entire excursion of the large signal. Secondly, the bodies of both transistors are floated to prevent forward-biasing of the source/body junction due to the -500mV worst-case source voltage. With gates and sources swinging in unison, the final consideration is the drain node. The top of the capacitor can be thought to swing from 1.2V to -1.2V, and the drains can be biased and the swings can be distributed accordingly such that damage is prevented. If both transistors are sized identically and laid out in a symmetric fashion, their drain voltages in the off state can be assumed to split close to evenly across the 2.4 peak-to-peak voltage. By biasing the top transistor drain at 1.2V, and the bottom transistor drain at 0.6V, the two transistors' V_{DS} each go from 0 to 1.2V. Including the pad inductance and its swing, Fig. 4.12 shows the voltage excursions of each node.

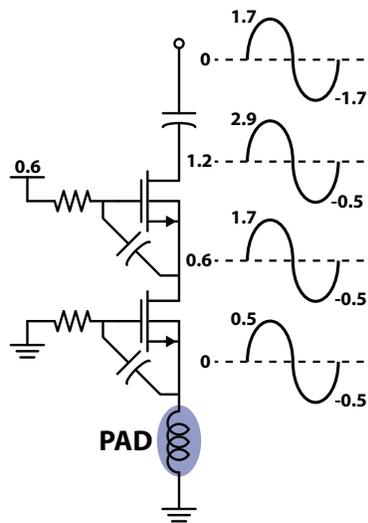


Figure 4.12: Stacked capacitor DAC, disabled unit cell.

While this simple configuration shown in Fig. 4.13 prevents overvoltage in the presence of large PA swings, there is a tradeoff between noise and bias accuracy due to the resistors used to bias the drains. An AC noise simulation was performed on the RX matching network with capacitor DAC under two bias resistor values, shown in Fig. 4.14. Here, as described before, a sweep of capacitor DAC codes was performed and the lowest total noise figure was chosen for each frequency point. It is clear that at lower frequencies (higher capacitor DAC codes and more unit cells enabled), a higher biasing resistance is desired. Unfortunately, given leakage from the devices, a larger resistance on the off cells leads to inaccuracies in the drain voltages, which could lead to overvoltage of the capacitor DAC off transistors. Therefore, PMOS switches were added to the top and bottom switch drain biasing such that off switches see low resistance and on switches see high resistance.

With this addition of another set of switches, there is another chance of overvoltage of devices. When the unit cell is disabled, the drain of the top PMOS sees the full 1.7V

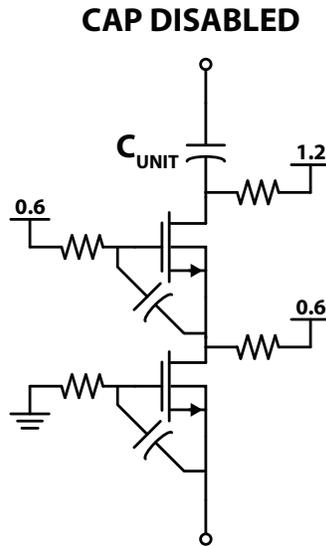


Figure 4.13: Simple biasing of unit cell.

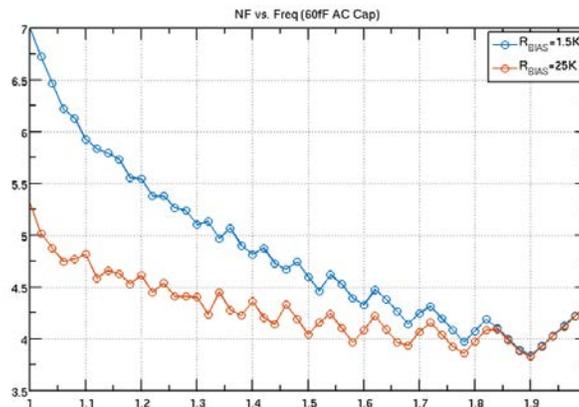


Figure 4.14: Effect of bias resistor size on system noise figure.

amplitude voltage swing, meaning that pinning the PMOS source at 1.2 will cause damage. The series resistance is divided into two portions and the PMOS switch is placed between the two such that the drain and source voltages never swing below ground. The drain is capacitively coupled to the gate since in both cell enabled and disabled cases, the drain swings, while the source only swings when the unit cell is disabled. This difference between the source and drain also determines which side the body should be tied to. When the unit cell is enabled, the drain has a 0.5V amplitude swing centered around 0V, meaning in the worst case it may weakly forward bias the NWELL/PSUB diode. By floating the body based off the source, this problem is not an issue, as shown in Fig. 4.15.

To verify the lumped simulation of the harmonic trap noise figure from Section 4.3, a

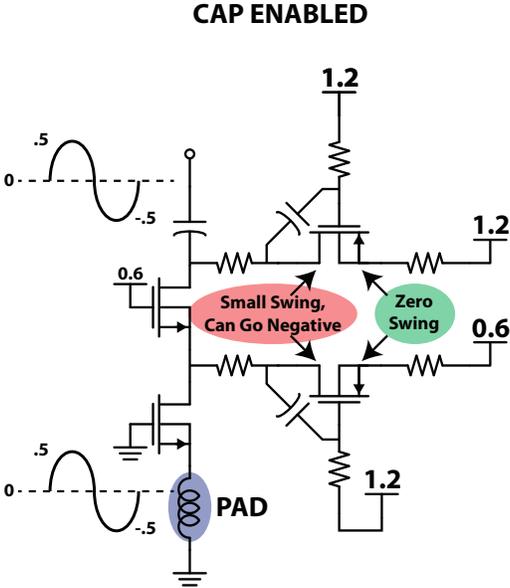


Figure 4.15: PMOS switch voltage swings.

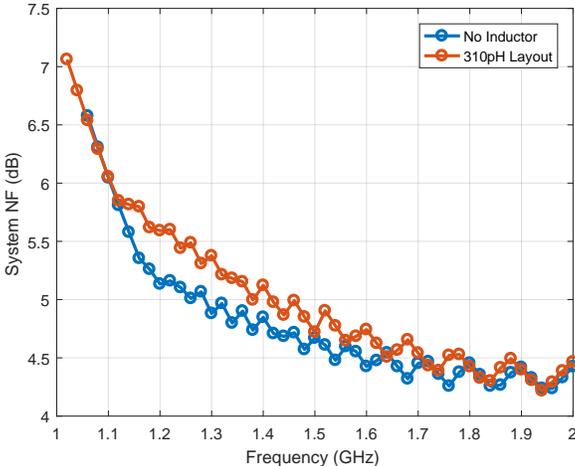


Figure 4.16: System noise figure with series harmonic trap included.

simulation was performed with an extracted layout of the capacitor DAC along with the harmonic trap inductor in series in Fig. 4.16. Including an electromagnetic simulation of the harmonic trap inductor layout, the system noise figure is negligibly impacted by the trap over a large range of frequencies.

4.5 Column Cascodes

The capacitive load on a transformer strongly affects the efficiency of a network. In simulation, the capacitive load of the RX balun (set mainly by the DAC output capacitance) is swept and a matching network is optimized for maximum G_A at a single frequency, shown in Fig. 4.17. The output impedance of the DAC in the first chip was 1.5pF. Lowering the output capacitance of the DAC to 0.5pF yields a 1.5dB improvement in matching network G_A . As stated in the previous section, available gain increase directly translates to system noise figure reduction when using a passive mixer first receiver.

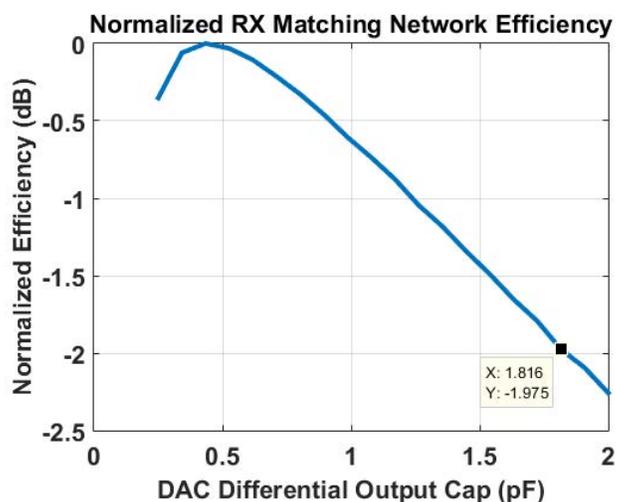


Figure 4.17: Matching network efficiency vs. DAC cap.

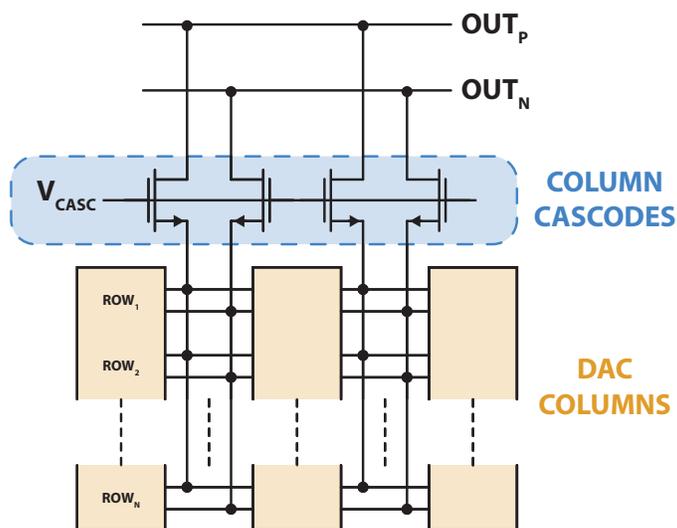


Figure 4.18: DAC columns with cascodes.

Accounting for all capacitances at the RX input, nearly 40% of the total capacitance of the DAC is due to the output routing lines whose width and spacing is set by space constraints. Dummy switch devices in the unit cell contribute another 25-30% of the total output capacitance. The capacitance of the active devices, whose widths are set by current handling and linearity requirements, only account for approximately 30% of the total RX loading capacitance. Summing the currents within a column together at a low impedance node, shown in Fig. 4.18, then buffering this current to the output reduces the routing capacitance seen by the RX to a small fraction of what it was before and heavily minimizes the relative contribution of dummy devices due to the large number of active fingers in these cascodes. Additionally, because of the low $1/g_m$ impedance of the cascodes, the DAC intermediate node can maintain a high bandwidth, not causing large changes to the code to current relationship of the current DAC. This has a twofold effect, where the first advantage is that predistortion becomes simpler, as well as digital cancellation models for the DAC. Both these advantages make an integrated version of predistortion and digital cancellation consume less power and area.

4.5.1 Column Cascode Nonlinearity

The dominant distortion mechanism due to the column cascode is a code-dependent rotation of the constellation. This is because the $1/g_m$ source impedance is, to first order, inversely proportional to $\sqrt{I_{Column}}$. The current flowing into the cascode source is a single-pole low-pass filter, whose bandwidth is dependent on the current draw, or current DAC code, to first order:

$$A_{Column}(s) = \frac{1}{1 + skC_{Column}\sqrt{I_{Column}}^{-1}} \quad (4.2)$$

In an extreme case, where the risetime of the column cascode takes a significant portion of a quarter-period of the LO, the constellation is distorted as shown in Fig. 4.19b. Intuitively, this can be seen in the time domain, illustrated in Fig. 4.19a, where low codes have a larger delay due to the low bandwidth of the node, significantly rotating the low magnitude codes in the constellation. Then as the codes become higher, the incremental reduction in delay becomes less and less, leveling the rotation off.

As detailed in Section 2.7, with the use of digital predistortion, the only nonlinearities which the cancellation system is sensitive to are those which create discontinuities in constellation spacing. The column cascode spiral constellation nonlinearity is a smooth one with no discontinuities, so it only weakly limits cancellation levels. Managing the complexity of the DAC model is advantageous though, because it allows simpler, lower power consumption, predistortion blocks to be created. It is therefore useful to mitigate this effect by reducing the cascode transconductance code dependence. Adding “bleeder” currents to each column cascode, shown in Fig. 4.20, which forces a minimum cascode current, has a two-fold effect on reducing cascode nonlinearity. First, it sets a lower-bound on the bandwidth of the cascode node, meaning that less delay change with code will occur. Second, because the g_m has an

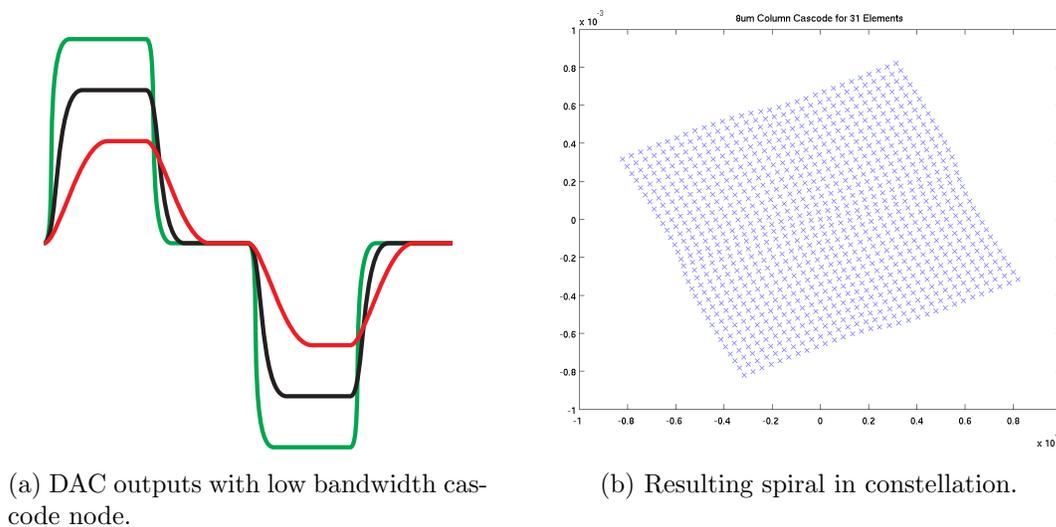


Figure 4.19: Cascode code-dependent bandwidth.

approximate square-root dependence on current, the incremental change in g_m as a function of current is less at higher currents.

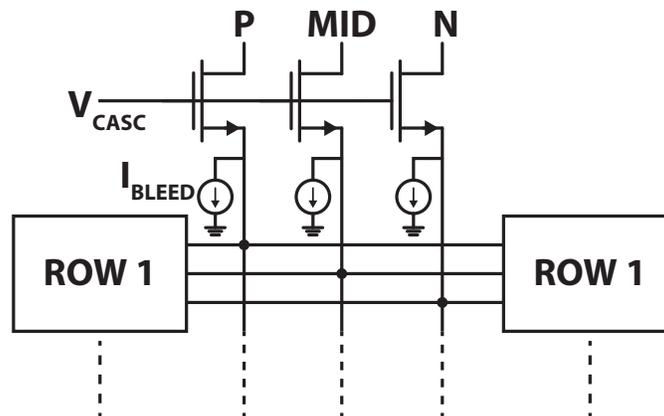
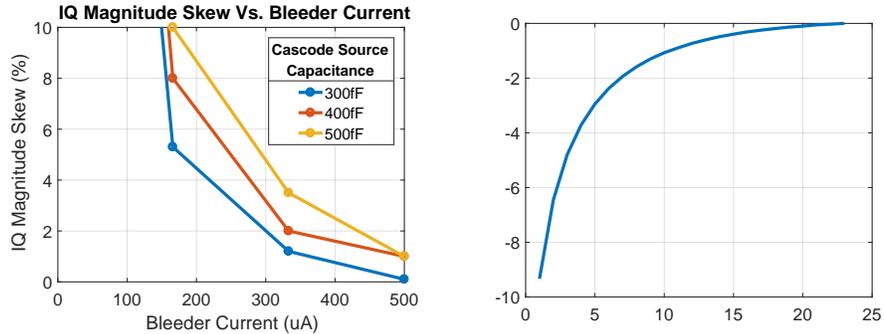


Figure 4.20: Bleeder currents for column cascodes.

A series of simulations was performed where a cascode source capacitance is assumed worst-case I/Q nonlinearity, as well as AM-PM distortion is measured, shown in Fig. 4.21. With 500fF differential capacitance in a column, sizing the 3 current bleeders at 300uA each, in total 10% of the current for a column, creates acceptable nonlinearity.

A simulation of DNL at this nominal bias point was performed, where a single column's current was ramped from zero to maximum, with a 50% duty cycle. In this case, the DNL maintained at approximately 0.1LSB, as seen in Fig. 4.22.



(a) Column cascode I/Q nonlinearity. (b) Column cascode AM-PM distortion.

Figure 4.21: Simulations of column cascode-induced nonlinearities.

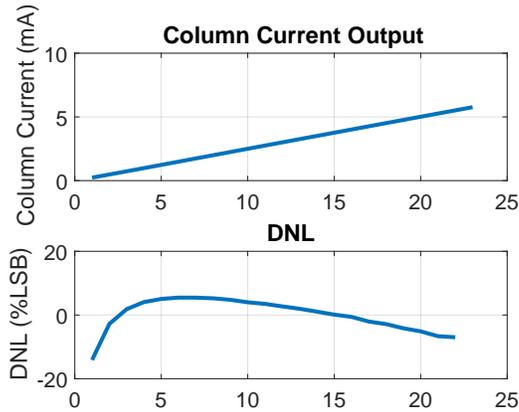


Figure 4.22: DNL due to column cascode (500fF source cap, 300uA bleeder).

4.5.2 Noise

While nearly all the noise from the unit cell switches circulates due to the high bandwidth of the tail drain node, the column cascodes do not receive the same benefit because the cascode sources inherently have low bandwidth. The current-current transfer function for cascode noise is:

$$A_I(s) = \frac{1}{1 + g_m R} \frac{(1 + sRC)}{1 + s \frac{RC}{1 + g_m R}} \quad (4.3)$$

This transfer function reduces the column cascode current by more than 20dB at low frequencies, but the zero in the transfer function is located at $\omega = (RC)^{-1}$, removing this filtering effect at higher frequencies. While this limited bandwidth is not ideal, one mitigating factor is that the cascode noise does not offset the noise reduction from thermal noise feedback. This is because the resonant impedance of the center tap is high only at low

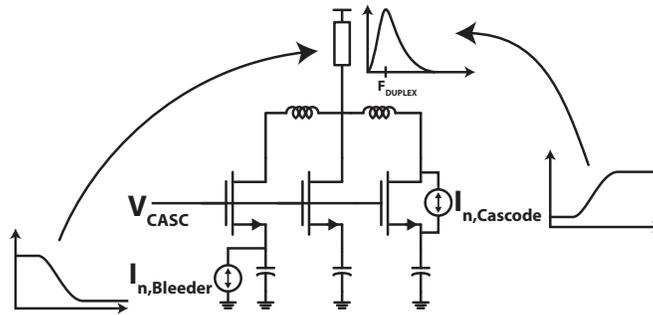


Figure 4.23: Illustration of cascode and bleeder noise current injection.

frequencies, precisely where the column cascodes achieve good thermal noise circulation, as shown in Fig. 4.23.

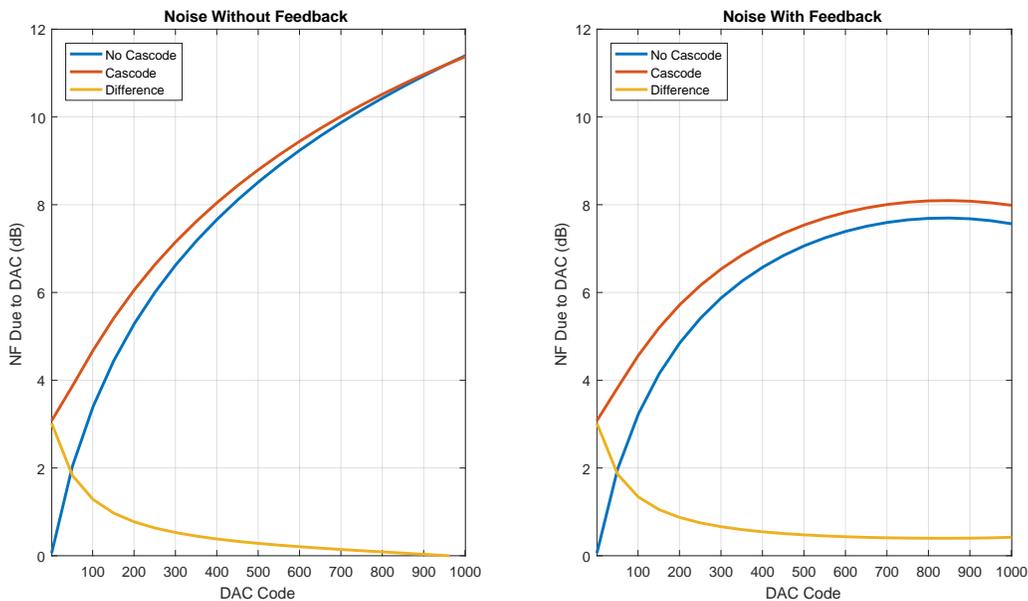


Figure 4.24: NF with and without feedback for column cascode.

On the other hand, the bleeder current sources have the opposite frequency response to the cascode noise sources. At high frequencies, the low impedance of the source capacitance reduces its contribution to the system noise figure, but at low frequencies, nearly all the bleeder noise current is buffered through the cascodes, going directly to the center tap to be fed back, shown in Fig. 4.23. Fundamentally, there is a tradeoff between column cascode linearity and noise from the bleeders, as a higher bias current leads to a lower noise resistance. In this design, resistors are chosen due to their higher noise resistance as compared with transistor current sources Eq. (4.4), where for a 1.5V cascode gate supply and a 400mV

threshold voltage, the resistor produces 3.5x lower noise variance than the current source. In practice, the noise due to the cascodes and bias resistors alone creates approximately a 2dB NF, degrading the total noise figure by <1dB, shown in Fig. 4.24.

$$\frac{R_{n,TransistorMax}}{R_{n,Resistor}} = \frac{2}{1 - \frac{V_T}{V_{Source}}} \quad (4.4)$$

4.6 Measurement Results

The improved system was taped out in the TSMC 65nm process, shown in Fig. 4.25. The main points of improvement for the second version were noise figure (through the improved passive network and RX) and TX power handling before compression.

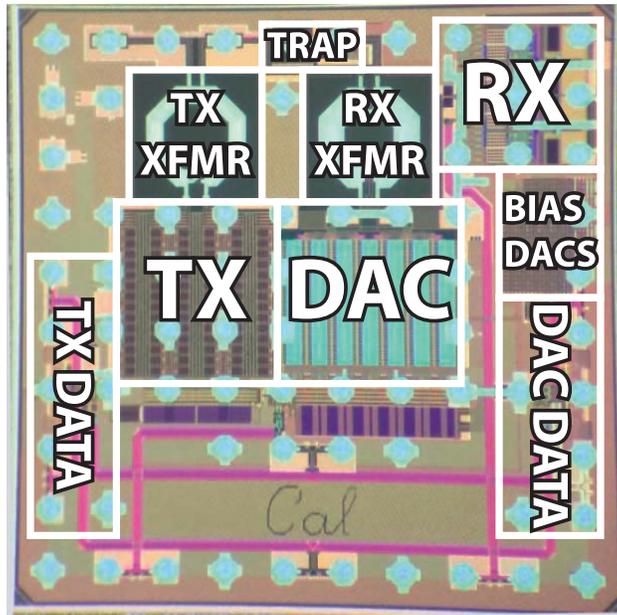


Figure 4.25: Die photo of second version.

A series of measurements were made of the TX/RX isolated board, where the capacitor DAC code was swept and the S-parameters were measured. A comparison between measured and simulated S11 with the RX off is shown in Fig. 4.26. The DAC codes shown (0, half, max) match extremely well to simulated data when small changes from the design point are made. The “on” capacitance of the DAC in this model is 10% higher than designed (715fF vs. 650fF), which could be due to unaccounted fringing or bottom plate capacitance in the devices. The RX balun’s primary inductance is 13% higher than designed for, which was later found to be due to the inductance of the long routing lines connecting the RX balun to the output pads.

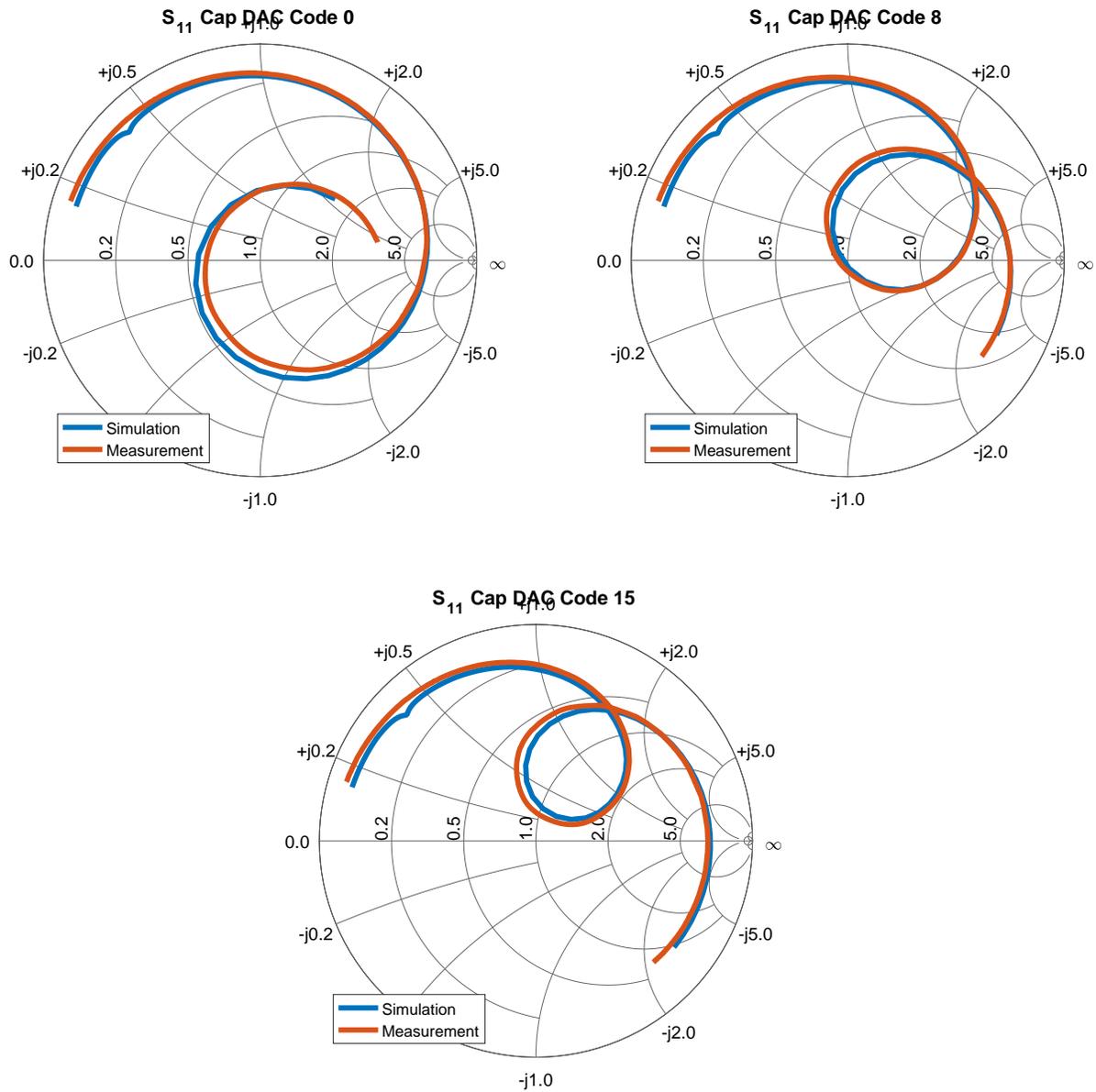


Figure 4.26: S_{11} measurement vs. simulation over capacitor DAC code sweep.

Next, the resonance point of the matching network was found as a function of capacitor DAC code. Here, the resonance point is defined as the frequency for which the S_{11} minimum coincided with the RX LO frequency using a fully real matching admittance. Accordingly, for a fixed capacitor DAC code, the RX LO frequency is swept and the minimum is S_{11} at each F_{RX} is recorded. A plot of the resonance points is shown and compared to simulation in

Fig. 4.27. In design, the capacitor DAC at the lowest code should have given a resonance of 2GHz, but the differences in the network mentioned earlier caused the measurement results to lower.

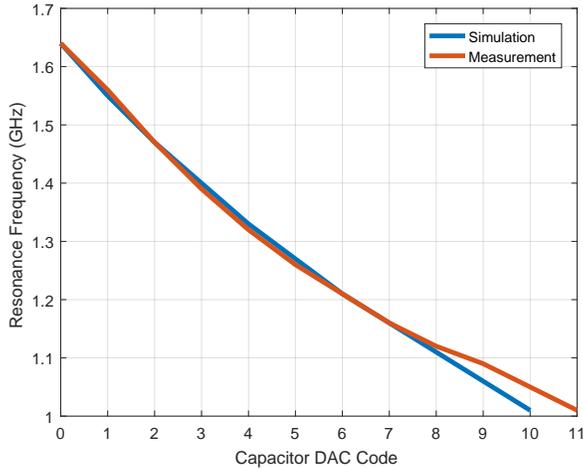


Figure 4.27: S_{11} resonance frequency vs. capacitor DAC code.

The S_{11} of the receiver was measured across different capacitor DAC codes, as shown in Fig. 4.28. Using codes from 0 to 8 on the capacitor DAC, as well as tuning the receiver’s imaginary matching admittance, the RX can achieve better than -20dB S_{11} from 1-1.9GHz. The highly linear design of the first and second class-AB stages lead to a +25dBm OOB IIP3 and +5dBm OOB P1dB, shown in Fig. 4.29a and Fig. 4.29b respectively. For IIP3 measurement, two separate tests were done such that the IM3 product would always fall into the RX passband: one tone is placed at the edge of the passband, and the other is ramped from F_{RX} to the edge of the passband, then both tones are moved, with one moving twice as fast as the other such that the IM3 tone stays at the edge of the passband. This two-phase method accounts for the discontinuity during the switching point.

4.6.1 Phase Noise

A copy from the PCB of Chapter 3, the interface between the chip and the testing PCB consisted of three pads: TX, RX, and the middle pin between the two baluns. The second testing PCB was created in two different ways, where one connected the middle pin directly to the PCB ground plane, while the other did not bond to that pin at all. This is in contrast to the first PCB which used an 0201 zero-ohm resistor to shunt the middle pin. The advantage with this method is two-fold: the strongly grounded middle pin provides significant isolation between the TX and RX, while the reduction in parasitics in the open allows for a more wideband leakage network. This is evidenced by the measurement of narrowband noise injected into the TX LO in Fig. 4.30, where even past 100MHz offset from F_{TX} , the cancellation of phase noise is >20dB.

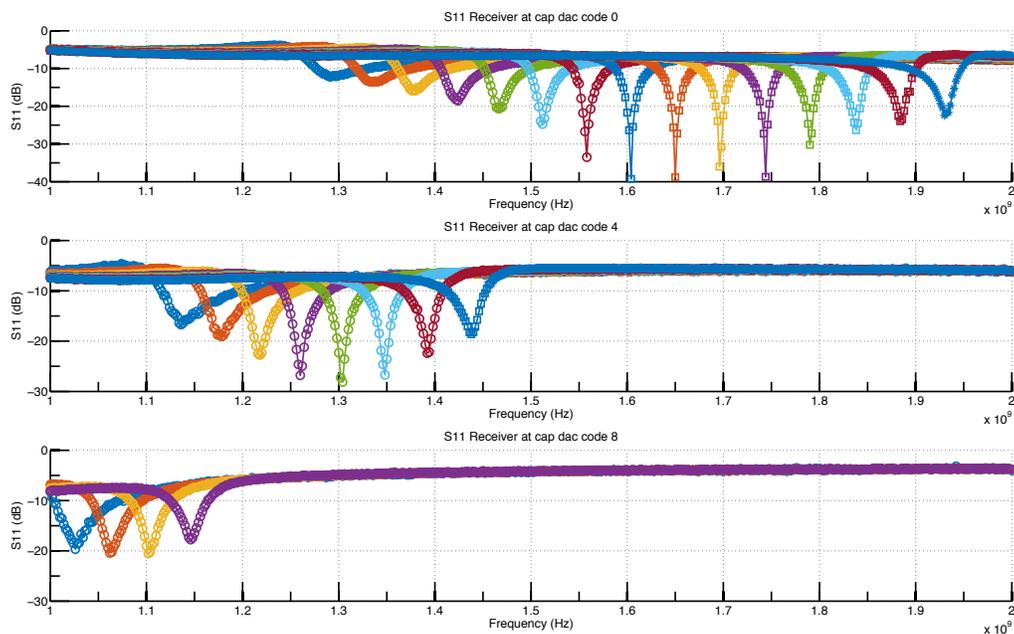
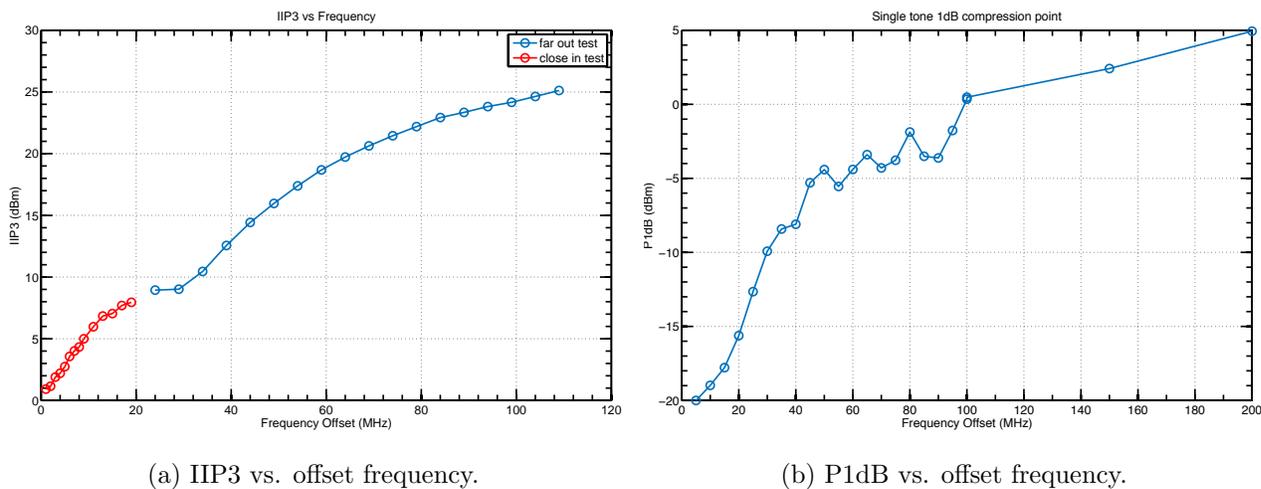


Figure 4.28: Measured S_{11} versus frequency with capacitor DAC tuning.



(a) IIP3 vs. offset frequency.

(b) P1dB vs. offset frequency.

Figure 4.29: RX linearity metrics.

4.6.2 Noise With Cancellation

In measurement, this chip had a 15dB noise figure for +16dBm TX output power cancelled for 40MHz duplex spacing, shown in Fig. 4.31.

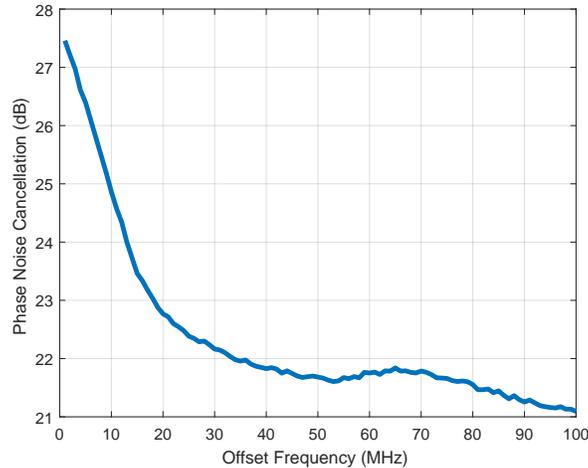


Figure 4.30: Improved phase noise cancellation bandwidth.

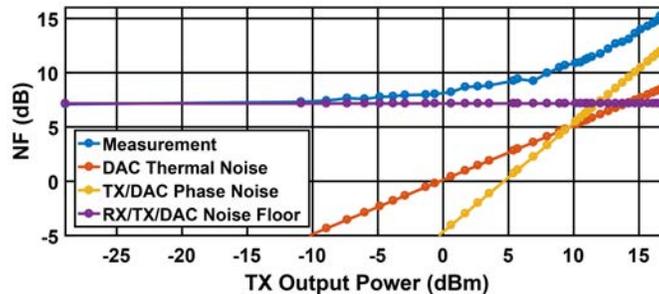


Figure 4.31: Breakdown of noise vs. cancellation power.

4.6.3 RX Compression With Cancellation

Gain compression was measured for 4 different frequency offsets: in-band, 40MHz, 80MHz, and 120MHz spacing. The results are given in Fig. 4.32. Compared with the previous chip, approximately 5dB higher TX power is able to be handled, owing to the large linearity increase, as well as harmonic rejection, of the new receiver. A separate measurement was done of the compression for in-band full-duplex using a circulator, where +13.5dBm TX power was handled before compressing. The circulator’s isolation was 20dB in the TX band, but the higher harmonics not being filtered is likely the reason for compression at only 10dB higher than the in-band measurement without the circulator.

Gain compression due to harmonics is a plausible explanation for the limited linearity of this system. In Fig. 4.33, a capacitor DAC code was fixed and the TX LO frequency was swept around $F_{0,LC}$, where this frequency is the resonance of the capacitor DAC and the series inductor, which form a harmonic trap. As shown, the third harmonic is suppressed around these frequencies, but the fifth remains high. At the signal levels recorded for 1dB compression, the third and fifth can plausibly be the main sources of compression. To further

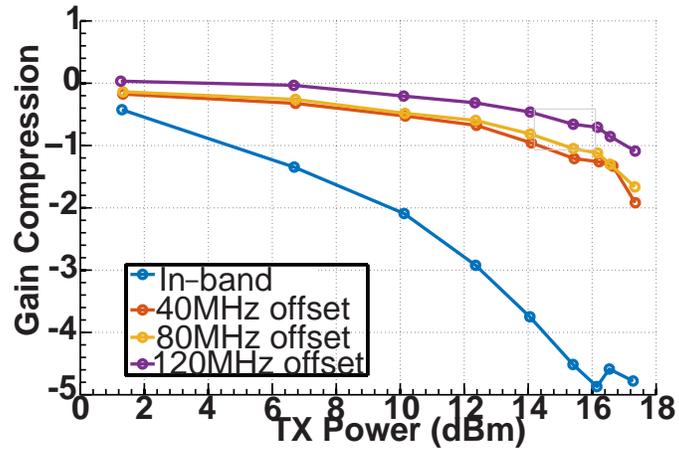


Figure 4.32: Gain compression vs. TX power.

demonstrate this, at a 1MHz duplex spacing (such that third and fifth harmonics would fall in-band), the TX was cancelled up to the signal level for 1dB compression. Then, the third and fifth harmonic powers at the output of the RX were measured, shown in Fig. 4.34. Finally, the RX signal gain was measured without any extra signals added and compared with the case where third and fifth harmonic signals were injected at the antenna, with a power level to match the output referred levels with the TX on. This also produced 1dB gain compression, proving that it is the uncanceled harmonics which cause gain compression.

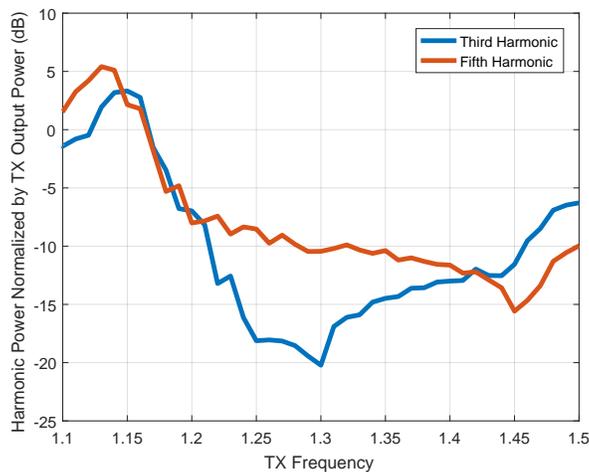
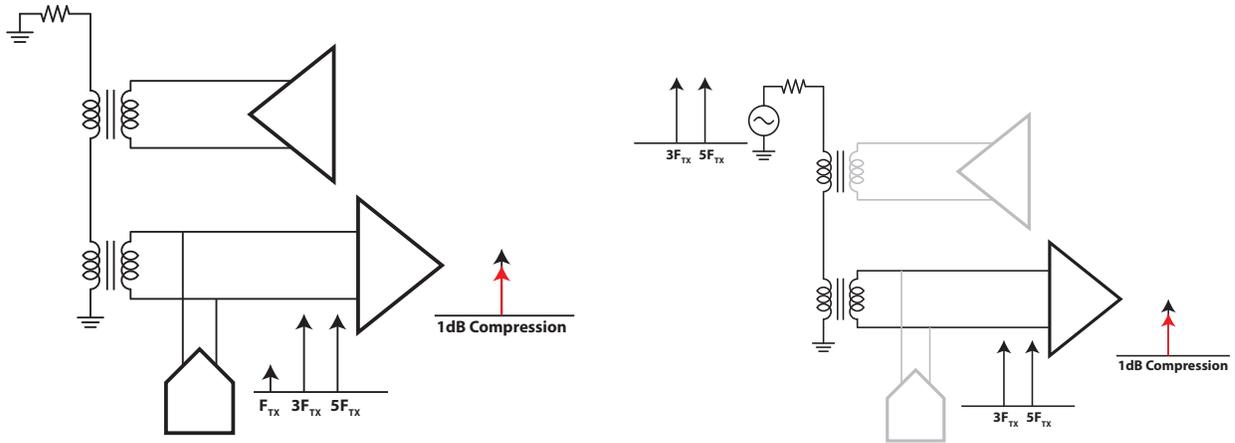


Figure 4.33: RX input harmonic power normalized to fundamental.



(a) RX harmonic measurement.

(b) RX harmonic injection.

Figure 4.34: Harmonic compression testing setup.

4.6.4 Comparison With Prior Art

Comparing this work with prior art with similar power handling capabilities, this system achieves >20dB modulated data frontend cancellation, and a competitive NF with cancellation and without any external isolation. Furthermore, 25-30dB of backend digital cancellation brings the potential system cancellation to >90dB.

	This Work	Yang, RFIC 2016 [51]	Zhou, ISSCC 2015 [40]
Technology	65nm	65nm	65nm
Frequency (GHz)	1.0-2.0	0.3-1.6	0.8-1.5
Frontend Cancellation (dB)	64	25	40
Backend Cancellation (dB)	25-30	-	-
P_{TX} , 1dB compression (dBm)	+17¹	+14 ¹	+15 ¹
RX NF (dB)	7-15.4²	11-16 ¹	4.8-6.3 ¹
External Cancellation (dB)	0	0	35

¹ $F_{Duplex} = 115\text{MHz}$

² $F_{Duplex} = 40\text{MHz}$

Table 4.1: Comparison with prior art.

Chapter 5

Conclusion

5.1 Thesis Summary

It is no longer sustainable for a phone to work in a small region of the world. Instead, phones which can achieve high LTE data rates across the globe are highly desired. This is made difficult due to the large number of global LTE FDD bands, which necessitate a high degree of isolation between the TX and RX at closely spaced bands. The current state of the art in phones requires the use of many discrete filters for each of these bands, which are costly and require extra space on the circuit board.

An additional source of interference is the large number of separate communication protocols which use the same frequency ranges, and the push towards multi-mode transceiver chips. Here, multiple standards could communicate at the same time at very closely spaced frequencies, just like in the case of FDD, once again requiring large self-interference suppression due to the close proximity between the transceivers.

This work, along with [48], proposes a frequency-flexible self-interference cancellation architecture which can cancel record high TX power without compressing the receiver. This architecture exploits a series connection between the transmitter and receiver, with a replica current source shunting the current normally flowing through the receiver. This architecture provides low insertion loss for both the TX and RX, due to this TX signal virtual ground. The use of a current-steering RF-DAC additionally supports a wide range of leakage channels, where nonlinear predistortion may be used to account for nonidealities in the TX or DAC, arbitrarily long FIR filters may be applied to suppress TX signals over a wide range of VSWR and reflection conditions.

Furthermore, it is not simply enough to cancel the TX signal leaking to the RX such that the receiver does not compress; the RX-band interference due to the TX and DAC can be a strong desensitization mechanism if not accounted for. In this work, the main sources of desensitization are TX phase noise, DAC thermal noise, and TX/DAC quantization noise. Phase noise, proportional to P_{TX} , is the most significant non-deterministic interference mechanism, but this work has shown that wideband feedforward cancellation of TX phase

noise using the DAC is possible if care is taken in creating the LO distribution networks for the TX and DAC such that most of the phase noise is shared between them. Thermal noise, proportional to $\sqrt{P_{TX}}$, is not as strong as phase noise, but may still be mitigated by the use of a feedback cancellation technique. Quantization noise is the most significant source of interference when transmitting modulated data, but it is a deterministic effect and can be significantly lowered with a digital predictive model for quantization noise propagation through the transceiver.

5.2 Future Directions

As shown in this thesis, as well as [48], a current-steering RF-DAC is an attractive solution for high-power TX cancellation, but there is room for improvement. Most notably, further harmonic traps may be used to reduce the uncanceled fifth and higher harmonics to push the linearity of the transceiver to higher TX powers. An additional option is the use of a harmonic cancelling architecture for both the TX and DAC, similar to those of [66, 67, 44], in order to cancel the harmonics at the output of each device, obviating the need for resonant traps. Furthermore, the digital backend cancellation detailed in this work shows a large amount of potential for further research. Finally, due to the lack of need for high dynamic linearity to achieve good analog cancellation, one could adopt a class-B cancellation DAC to achieve power savings with modulated data.

Bibliography

- [1] Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 20162021. Tech. rep. Cisco, 2016.
- [2] Techcrunch'. *Facebook Video Pays Off*. URL: <https://techcrunch.com/2016/11/03/facetube/>.
- [3] Mediakix'. *The Snapchat Statistics Every Marketer Needs to Know*. URL: <http://mediakix.com/2016/01/snapchat-statistics-2016-marketers-need-to-know/>.
- [4] Expanded Ramblings. *160 Amazing YouTube Statistics*. URL: <http://expandedramblings.com/index.php/youtube-statistics/>.
- [5] Techcrunch'. *One in five Facebook videos is Live as it seizes the verb*. URL: <https://techcrunch.com/2017/04/06/live-video/>.
- [6] K. Mayama. URL: <http://pictures.reuters.com/archive/POPE-RP6DRMRPZFAB.html>.
- [7] Associated Press. URL: <http://static1.businessinsider.com/image/514205d4eab8ea8b3800004346-3260/vatican-papal-conclave-2013.jpg>.
- [8] 3GPP'. *Releases*. URL: <http://www.3gpp.org/specifications/67-releases>.
- [9] Qorvo. *Redefining Filter Performance, Advanced NoDrift and LowDrift Solutions from Qorvo*. URL: <http://www.qorvo.com/-/media/files/qorvopublic/brochures/qorvo-advanced-filtering-solutions-brochure-2017.pdf>.
- [10] TriQuint. *SAW and BAW Technology Comparison*. URL: https://www.sec.gov/Archives/edgar/data/911160/000119312514092588/g691156page_23.jpg.
- [11] U. C. Fernando. "Non-adjacent carrier aggregation architecture". Pat. US20120294299A1. 2012.
- [12] N. Khlal. "Carrier aggregation radio system". Pat. US20130051284A1. 2013.
- [13] TeleGeography. *SK Telecom trials five-band carrier aggregation with Samsung*. 2017. URL: <https://www.telegeography.com/products/commsupdate/articles/2017/02/27/sk-telecom-trials-five-band-carrier-aggregation-with-samsung/>.
- [14] *Addressing Carrier Aggregation Challenges Using Multiplexer Solutions*. Tech. rep. Qorvo, 2016.

- [15] R. Chen and H. Hashemi. “19.3 Reconfigurable SDR receiver with enhanced front-end frequency selectivity suitable for intra-band and inter-band carrier aggregation”. In: *2015 IEEE International Solid-State Circuits Conference - (ISSCC) Digest of Technical Papers*. 2015, pp. 1–3. DOI: 10.1109/ISSCC.2015.7063068.
- [16] 3GPP. *UMTS UE Specifications*. URL: http://www.etsi.org/deliver/etsi_ts/125100_125199/125101/06_19_00_60/ts_125101v061900p.pdf.
- [17] B. D. Smith. “Analysis of Commutated Networks”. In: *Transactions of the IRE Professional Group on Aeronautical and Navigational Electronics PGAE-10* (1953), pp. 21–26. ISSN: 2168-0167. DOI: 10.1109/TPGAE.1953.5062331.
- [18] B. W. Cook et al. “Low-Power 2.4-GHz Transceiver With Passive RX Front-End and 400-mV Supply”. In: *IEEE Journal of Solid-State Circuits* 41.12 (2006), pp. 2757–2766. ISSN: 0018-9200. DOI: 10.1109/JSSC.2006.884801.
- [19] H. Darabi, A. Mirzaei, and M. Mikhemar. “Highly Integrated and Tunable RF Front Ends for Reconfigurable Multiband Transceivers: A Tutorial”. In: *TCAS* 58.9 (2011), pp. 2038–2050.
- [20] J. Zhou, N. Reiskarimian, and H. Krishnaswamy. “9.8 Receiver with integrated magnetic-free N-path-filter-based non-reciprocal circulator and baseband self-interference cancellation for full-duplex wireless”. In: *2016 IEEE International Solid-State Circuits Conference (ISSCC)*. 2016, pp. 178–180. DOI: 10.1109/ISSCC.2016.7417965.
- [21] 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.
- [22] N. 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.
- [23] D. Murphy et al. “A blocker-tolerant wideband noise-cancelling receiver with a 2dB noise figure”. In: *IEEE International Solid-State Circuits Conference* (2012).
- [24] J. 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).
- [25] 3GPP. *LTE Specifications*. URL: <http://www.3gpp.org/specifications>.
- [26] L. Weinberg. *Network analysis and synthesis*. RE Krieger Publishing Company, 1975.
- [27] A. Mirzaei and H. Darabi. “A Low-Power WCDMA Transmitter With an Integrated Notch Filter”. In: *JSSC* 43.12 (2008), pp. 2868–2881.
- [28] A. Safarian et al. “Integrated Blocker Filtering RF Front Ends”. In: *2007 IEEE Radio Frequency Integrated Circuits (RFIC) Symposium*. 2007, pp. 13–16. DOI: 10.1109/RFIC.2007.380822.

- [29] C. Izquierdo et al. “Reconfigurable wide-band receiver with positive feed-back translational loop”. In: *2011 IEEE Radio Frequency Integrated Circuits Symposium*. 2011, pp. 1–4. DOI: 10.1109/RFIC.2011.5940616.
- [30] C. Izquierdo et al. “Wide-band receiver architecture with flexible blocker filtering techniques”. In: *2010 17th IEEE International Conference on Electronics, Circuits and Systems*. Dec. 2010, pp. 894–897. DOI: 10.1109/ICECS.2010.5724656.
- [31] T. D. Werth, C. Schmits, and S. Heinen. “Active feedback interference cancellation in RF receiver front-ends”. In: *2009 IEEE Radio Frequency Integrated Circuits Symposium*. 2009, pp. 379–382. DOI: 10.1109/RFIC.2009.5135562.
- [32] T. D. Werth et al. “An Active Feedback Interference Cancellation Technique for Blocker Filtering in RF Receiver Front-Ends”. In: *IEEE Journal of Solid-State Circuits* 45.5 (May 2010), pp. 989–997. ISSN: 0018-9200. DOI: 10.1109/JSSC.2010.2041405.
- [33] M. Mikhemar, H. Darabi, and A. Abidi. “A Tunable Integrated Duplexer with 50dB Isolation in 40nm CMOS”. In: *ISSCC* (2009).
- [34] M. Mikhemar, H. Darabi, and A. Abidi. “An On-Chip Wideband and Low-Loss Duplexer for 3G/4G CMOS Radios”. In: *VLSI Symposium* (2010).
- [35] J. G. Kim et al. “Balanced topology to cancel Tx leakage in CW radar”. In: *IEEE Microwave and Wireless Components Letters* 14 (2004).
- [36] T. Zhang et al. “An integrated CMOS passive transmitter leakage suppression technique for FDD radios”. In: *Radio Frequency Integrated Circuits Symposium* (2014).
- [37] S. H. Abdelhalem, P. S. Gudem, and L. 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.
- [38] B. van Liempd et al. “A +70dBm IIP3 single-ended electrical-balance duplexer in 0.18 μ m SOI CMOS”. In: *IEEE International Solid-State Circuits Conference* (2015).
- [39] M. Mikhemar. “Interference Cancellation in Software-Defined CMOS Receivers”. PhD thesis. University of California, Los Angeles, 2010.
- [40] J. Zhou et al. “19.1 Receiver with 20MHz bandwidth self-interference cancellation suitable for FDD, co-existence and full-duplex applications”. In: *2015 IEEE International Solid-State Circuits Conference - (ISSCC) Digest of Technical Papers*. 2015, pp. 1–3. DOI: 10.1109/ISSCC.2015.7063066.
- [41] J. Zhou, P. R. Kinget, and H. Krishnaswamy. “20.6 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: *2014 IEEE International Solid-State Circuits Conference Digest of Technical Papers (ISSCC)*. 2014, pp. 352–353. DOI: 10.1109/ISSCC.2014.6757466.

- [42] D. J. van den Broek, E. A. M. Klumperink, and B. Nauta. “19.2 A self-interference-cancelling receiver for in-band full-duplex wireless with low distortion under cancellation of strong TX leakage”. In: *2015 IEEE International Solid-State Circuits Conference - (ISSCC) Digest of Technical Papers*. 2015, pp. 1–3. DOI: 10.1109/ISSCC.2015.7063067.
- [43] T. Lee and B. Razavi. “A 125-MHz mixed-signal echo canceller for Gigabit Ethernet on copper wire”. In: *IEEE Journal of Solid-State Circuits* 36.3 (2001), pp. 366–373. ISSN: 0018-9200. DOI: 10.1109/4.910475.
- [44] B. Yang et al. “A 65nm CMOS I/Q RF DAC with Harmonic Cancellation and Mixed-Signal Filtering”. In: *VLSI Symposium* (2017).
- [45] V. Vorapipat, C. Levy, and P. Asbeck. “A wideband voltage mode Doherty power amplifier”. In: *2016 IEEE Radio Frequency Integrated Circuits Symposium (RFIC)*. 2016, pp. 266–269. DOI: 10.1109/RFIC.2016.7508302.
- [46] J. B. Johnson. “Thermal Agitation of Electricity in Conductors”. In: *Phys. Rev.* 32 (1 July 1928), pp. 97–109. DOI: 10.1103/PhysRev.32.97. URL: <http://link.aps.org/doi/10.1103/PhysRev.32.97>.
- [47] 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: *2015 IEEE Radio Frequency Integrated Circuits Symposium (RFIC)*. 2015, pp. 75–78. DOI: 10.1109/RFIC.2015.7337708.
- [48] S. Ramakrishnan. “Design of Integrated Full-Duplex Wireless Transceivers”. PhD thesis. University of California, Berkeley, 2016.
- [49] D. Chowdhury. “Efficient Transmitters for Wireless Communications in Nanoscale CMOS Technology”. PhD thesis. University of California, Berkeley, 2010.
- [50] F. Sheikh. “Power-Performance Tradeoffs in ASICs for Next Generation Wireless Datapaths”. PhD thesis. University of California, Berkeley, 2008.
- [51] D. Yang et al. “A fully integrated Software-Defined FDD transceiver tunable from 0.3-to-1.6 GHz”. In: *2016 IEEE Radio Frequency Integrated Circuits Symposium (RFIC)*. 2016, pp. 334–337. DOI: 10.1109/RFIC.2016.7508320.
- [52] S. Yoo et al. “A Switched-Capacitor RF Power Amplifier”. In: *JSSC* 46.12 (2011), pp. 2977–2987.
- [53] Y. Nakamura et al. “A 10-b 70-MS/s CMOS D/A converter”. In: *IEEE Journal of Solid-State Circuits* 26.4 (Apr. 1991), pp. 637–642. ISSN: 0018-9200. DOI: 10.1109/4.75066.
- [54] J. Savoj et al. “A 12-GS/s Phase-Calibrated CMOS Digital-to-Analog Converter for Backplane Communications”. In: *IEEE Journal of Solid-State Circuits* 43.5 (May 2008), pp. 1207–1216. ISSN: 0018-9200. DOI: 10.1109/JSSC.2008.920319.

- [55] C. H. Lin et al. “A 12 bit 2.9 GS/s DAC With IM3 \ll - 60 dBc Beyond 1 GHz in 65 nm CMOS”. In: *IEEE Journal of Solid-State Circuits* 44.12 (Dec. 2009), pp. 3285–3293. ISSN: 0018-9200. DOI: 10.1109/JSSC.2009.2032624.
- [56] A. van den Bosch et al. “A 10-bit 1-GSample/s Nyquist current-steering CMOS D/A converter”. In: *IEEE Journal of Solid-State Circuits* 36.3 (Mar. 2001), pp. 315–324. ISSN: 0018-9200. DOI: 10.1109/4.910469.
- [57] A. Van den Bosch et al. “A 12 bit 200 MHz low glitch CMOS D/A converter”. In: *Proceedings of the IEEE 1998 Custom Integrated Circuits Conference (Cat. No.98CH36143)*. May 1998, pp. 249–252. DOI: 10.1109/CICC.1998.694974.
- [58] D. Murphy et al. “A Blocker-Tolerant, Noise-Cancelling Receiver Suitable for Wide-band Wireless Applications”. In: *IEEE Journal of Solid-State Circuits* 47.12 (Dec. 2012), pp. 2943–2963. ISSN: 0018-9200. DOI: 10.1109/JSSC.2012.2217832.
- [59] F. Bruccoleri, E. A. M. Klumperink, and B. Nauta. “Wide-band CMOS low-noise amplifier exploiting thermal noise canceling”. In: *IEEE Journal of Solid-State Circuits* 39.2 (Feb. 2004), pp. 275–282. ISSN: 0018-9200. DOI: 10.1109/JSSC.2003.821786.
- [60] S. C. Blaakmeer et al. “The Blixer, a Wideband Balun-LNA-I/Q-Mixer Topology”. In: *IEEE Journal of Solid-State Circuits* 43.12 (Dec. 2008), pp. 2706–2715. ISSN: 0018-9200. DOI: 10.1109/JSSC.2008.2004866.
- [61] E. Hegazi, H. Sjolund, and A. A. Abidi. “A filtering technique to lower LC oscillator phase noise”. In: *IEEE Journal of Solid-State Circuits* 36.12 (Dec. 2001), pp. 1921–1930. ISSN: 0018-9200. DOI: 10.1109/4.972142.
- [62] D. Murphy, H. Darabi, and H. Wu. “25.3 A VCO with implicit common-mode resonance”. In: *2015 IEEE International Solid-State Circuits Conference - (ISSCC) Digest of Technical Papers*. Feb. 2015, pp. 1–3. DOI: 10.1109/ISSCC.2015.7063116.
- [63] A. Hajimiri and T. H. Lee. “A general theory of phase noise in electrical oscillators”. In: *IEEE Journal of Solid-State Circuits* 33.2 (1998), pp. 179–194. ISSN: 0018-9200. DOI: 10.1109/4.658619.
- [64] 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 on Circuits and Systems I: Regular Papers* 57.12 (2010), pp. 3092–3103. ISSN: 1549-8328. DOI: 10.1109/TCSI.2010.2052513.
- [65] G. T. Sasse. “Reliability Engineering in RF CMOS”. PhD thesis. University of Twente, 2008.
- [66] Y. H. Chen et al. “9.7 An LTE SAW-less transmitter using 33% duty-cycle LO signals for harmonic suppression”. In: *2015 IEEE International Solid-State Circuits Conference - (ISSCC) Digest of Technical Papers*. 2015, pp. 1–3. DOI: 10.1109/ISSCC.2015.7062981.

- [67] C. Huang et al. “A 40nm CMOS single-ended switch-capacitor harmonic-rejection power amplifier for ZigBee applications”. In: *2016 IEEE Radio Frequency Integrated Circuits Symposium (RFIC)*. 2016, pp. 214–217. DOI: 10.1109/RFIC.2016.7508289.