Techniques for Mixed-Signal Linearization and Large Signal Handling in Radio-Frequency Receiver Circuits

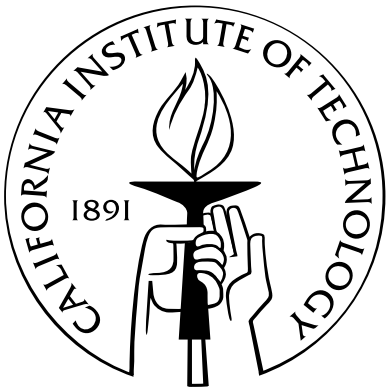
Thesis by

Edward A. Keehr

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy



California Institute of Technology

Pasadena, California

2011

(Defended June 1, 2010)

To my parents.

Acknowledgements

Chief among those whose support was crucial to the completion of this dissertation was my advisor, Professor Ali Hajimiri. In addition to providing the original research objective for the thesis, Ali served as a constant source of inspiration and motivation for moving forward with these projects. Ali's deep love for and knowledge of physics also served to provide solid and intuitive explanations to many problems considered difficult and unwieldy when approached from traditional circuit analysis methodologies.

Perhaps the biggest surprise during the completion of the work in this dissertation was the manual dexterity required to successfully assemble a working test module. Fortunately, Hamdi Mani was present to provide considerable assistance in this regard. In addition to performing incredibly intricate tasks on my behalf during deadline periods, Hamdi patiently taught me a number of essential techniques for handling small components and wires.

During the road to completing graduate studies, it is traditional for the student to pass the candidacy and thesis defense examinations. I would like to thank and acknowledge those professors and scientists for their participation in these endeavors, both for their time and for the valuable advice and questioning that came out of these processes. The individuals so acknowledged include Prof. Azita Emami, Prof. Babak Hassibi, Prof. Dave Rutledge, Prof. P. P. Vaidyanathan, Dr. Sander Weinreb, all of Caltech, and Dr. Larry D'Addario of JPL.

I am also indebted to the members of the Moore lab support staff, especially for their timely assistance during deadline periods. Our group secretary Michelle Chen has been superb with assisting in ordering supplies, travel, and arranging group get-togethers. For computer assistance both for the CHIC group cluster and other miscellaneous questions I am grateful for the help of Naveed Near-Ansari, John Lilley, and Gary Waters. I am also indebted to the assistance of Carol Sosnowski, Linda Dozsa, Lyn Hein, Janet Couch, and Tanya Owen.

As always, I am indebted to the advice and help of my fellow graduate students in the CHIC and MICS laboratories at Caltech. In particular, Matt Loh provided assistance with the tools and

methodologies required for Verilog coding and for accessing digital design libraries. Juhwan Yoo provided help with assembly and chip layout during the times leading up to various deadlines, while Hua Wang, Yu-Jiu Wang, Sanggeun Jeon, Shouhei Kousai, Aydin Babakhani, and Jay Chen provided helpful suggestions and information regarding testing, equipment, and design kits. The frequency divider utilized in the IM3-cancelling receiver was provided as hard layout IP by Florian Bohn. Joe Bardin and Glenn Jones provided LaTeX thesis example templates to help quickly determine the proper formatting commands for the Caltech dissertation document.

Circuit design is not an inexpensive field of study, and for this reason, discounted materials and software are thoroughly appreciated. To this end, I acknowledge UMC for the fabrication of the 90nm die described in Chapter 7, the Rogers Corporation for free dielectric substrates, Terasic for discounts on FPGA boards, Cadence for free software design tools under the Cadence University Program, Zeland Corporation for free licenses of the IE3D software package, DVH Circuits of Van Nuys for discounts in etching and gold plating of dielectric substrates, and the Altera Corporation for providing free licenses for FPGA programming tools. I would also like to thank and acknowledge the NDSEG Fellowship program and the Lee Center for Advanced Networking for funding portions of my stay at Caltech.

Finally, and most importantly, I would like to thank my beloved wife Kim for her patience and understanding during the completion of this work.

Abstract

In this dissertation, two effective linearization schemes for radio-frequency receivers are introduced. The first of these comprises a mixed-signal feedforward path which regenerates third-order intermodulation (IM3) products at radio frequencies, downconverts these products, digitizes them, and then uses them to cancel corruptive IM3 products in the digital baseband portion of a nominally linear receiver path. The combined implemented receiver represents a SAW-less direct-conversion receiver for UMTS FDD Region 1 that achieves an uncorrected out-of-band IIP3 of -7.1dBm under worst-case blocking specifications. Under IM3 equalization, the receiver achieves an effective IIP3 of +5.3dBm and meets the UMTS BER sensitivity requirements with 3.7dB of margin. To enable this mixed-signal feedforward path, a multistage cubic term generator is introduced which uses cascaded nonlinear operations to generate reference IM3 products. The multistage nature of this circuit is considered in the context of the aforementioned linearization scheme and is shown to provide sufficient dynamic range for nearly complete IM3 cancellation while dissipating far less power than the original receiver front end. In particular, the effect of the group delay between stages is analyzed and shown to permit large IM3 cancellation ratios for interstage group delays less than 1ns.

Expanding upon the first effective linearization approach led to the development of a large signal handling receiver with an out-of-band 1-dB desensitization point of +12.5dBm. Enabling this large signal handling capability is a passive mixer downconverter preceded by a novel wide-swing LNTA. With a stacked push-pull class-AB common-gate architecture, the LNTA reduces the magnitude of input-referred distortion by up to 40dB beyond that predicted by an initial slope-of-3 characteristic while at the same time minimally impacting the effective small-signal gain of the receiver. To compensate for intermodulation distortion terms of order greater than 3, IM3 and IM2 products are processed down to digital baseband where they are successively multiplied to generate approximations to higher-order terms. In the case of a +12.4dBm QPSK-modulated signal and a -16.3dBm CW blocker, cancellation improves receiver input-referred error by over 24dB, resulting in an extrapolated IIP3 of +43.5dBm.

Contents

Acknowledgements **iv**

Abstract **vi**

1 Background Material and Introduction **1**

1.1 Problem Definition 1

1.1.1 Radio-Frequency (RF) Systems 1

1.1.2 Error in RF Systems 2

1.1.3 Nonlinear Distortion in RF Systems 3

1.1.4 Metrics of Nonlinear Distortion 7

1.1.5 Out-of-Band Blocking 11

1.2 Inherent Challenges of Designing Linear RF Circuits 11

1.2.1 Insufficient On-Chip Filtering 11

1.2.2 Unfavorable Power/Linearity Tradeoff 12

1.3 Approaches and Challenges to Improving Nonlinearity in RF Circuits 13

1.3.1 Negative Feedback 13

1.3.2 Predistortion 15

1.3.3 Piecewise Linearization 16

1.3.4 Feedforward Cancellation 16

1.4 Adaptive Feedforward Cancellation of Self-Generated Error 17

1.4.1 LMS as Adaptive Equalization Algorithm 19

1.5 Prior Art in Adaptive Feedforward Cancellation of Nonlinear Distortion 20

1.6 Motivation and Challenges of Designing Very Large-Signal Handling RF Receivers . 22

2 Mixed-Signal / Mixed-Domain Equalization of IM3 Products in RF Receivers **26**

2.1 Target Application and Context of Project 26

2.2 General Concept of Proposed Solution Architecture 28

2.2.1 Choice of Feedforward Loop Overcoming Limitations of Prior Art 28

2.2.2 Fixed and Adaptive Equalization 29

2.2.3 Dc Offset in Direct Conversion Receivers 30
2.3 Sources of Error in Alternate Feedforward Path 32
2.3.1 Linear Term Feedthrough 32
2.3.2 IM3-to-Noise Ratio (INR) and IM3-to-Error Ratio (IER) 33
2.3.3 Higher-Order Nonlinear Terms 34
2.3.4 Quadrature Mismatch and LMS Equalization 34
2.4 Translation of Receiver Specification 35
2.4.1 Main Path 36
2.4.2 Alternate Path 38
2.5 Enhanced Degree of Freedom Adaptive Algorithms 38
2.6 Criteria and Techniques for Enabling and Disabling Feedforward Loop 39
2.6.1 Behavior of System in Fading Environment 41

3 Implementation and Testing of Experimental Linearized RF Receiver 42
3.1 Receiver Architecture 42
3.2 Circuit Block Implementations 43
3.2.1 Main Path Circuit Design 43
3.2.1.1 LNA and Balun 43
3.2.1.2 Mixer and LO Buffer 44
3.2.1.3 Analog and Digital Baseband Circuitry 45
3.2.2 Alternate Path Circuit Design 47
3.3 Alternate Path Postfilter Specification Requirements 50
3.4 Experimental Setup 52
3.5 Experimental Results 55
3.5.1 Receiver IIP3 Measurement Results 55
3.5.2 Receiver Sensitivity Measurement Results 58
3.5.3 Alternate Path Measurement Results 59
3.5.4 Additional Measurement Results 60

4 Multistage Cubic Term Generators for RF Receivers 61
4.1 Prior Art in Cubic Term Generators 61
4.2 Proposed Cubic Term Generator 62
4.3 Requirements of Cubic Term Generators for RF Receivers 63
4.4 Usefulness and Sufficiency of Internal Bandwidth Limitations in Reconstructing IM3 Products 63
4.5 Dynamic Range Calculations and Simulation Results 67
4.6 Comparison to Single-Stage Cubic Term Generators 70

4.7 Effect and Quantitative Analysis of Interstage Group Delay 73
4.7.1 Evaluating Term η 78
4.7.2 Evaluating Term κ 79
4.7.3 Evaluating Term ϵ 80
4.7.4 Evaluating Term χ and λ 80
4.7.5 Final Power Spectral Density Expressions 81

5 A Wide-Swing Low Noise Transconductance Amplifier and the Enabling of Large-Signal Handling Direct Conversion Receivers 84

5.1 General Design of a Large-Signal Handling Direct Conversion Receiver 84
5.2 Prior Art in CMOS LNAs and LNTAs 87
5.2.1 Current Sharing 87
5.2.2 IM2 Cancellation 88
5.2.3 IM3 Improvement in Common-Gate/Common-Source Noise Cancelling LNAs 88
5.2.4 Remarks 88
5.3 Evolution of Wide-Swing CMOS Common-Gate LNTA 89
5.3.1 Considering a Common-Source Push-Pull LNTA 89
5.3.2 Choice of Stacked vs. Parallel Topology 91
5.3.3 Device Sizing in a Stacked Topology 91
5.4 General Theory of LNTA Operation 93
5.5 Obtaining Near-Constant LNTA Transconductance 94
5.5.1 General Concept 94
5.5.2 Mobility Degradation Effects 96
5.5.3 Transition into Triode Region and Body Effect 98
5.5.4 Modeling Subthreshold MOS Conduction 100
5.5.5 Intuition of Device Physics Applied Towards General Concept 101
5.5.6 Total IMD Product Error as Function of Input Voltage Magnitude 102
5.5.7 LNTA Compression 105
5.6 Implementation and Simulation Results 106
5.6.1 Static Swept Input Simulations 106
5.6.2 Dynamic Swept Input Simulations 109
5.6.2.1 Compression Simulations 109
5.6.2.2 Noise Simulations 109
5.6.2.3 Quiescent Current Simulations 110
5.6.2.4 Intermodulation QPSS Simulations 110
5.7 Robustness of Constant-Transconductance Behavior 110

5.8 Heuristic Modeling of LNTA Nonlinearity 111
5.9 Biasing OTA Implementation 112

6 Successive Regeneration and Adaptive Cancellation of Higher-Order Intermodulation Products 114

6.1 Need for Cancelling Higher-Order Intermodulation Products 114
6.2 Naïve Approach to Cancelling Higher-Order Intermodulation Products 114
6.3 Proposed Approach to Cancelling Higher-Order Intermodulation Products 116
6.3.1 Primary Concept 116
6.3.2 Cancellation of IMD Products at Higher-Order Frequency Offsets 118
6.3.2.1 Quantitative Intuition of Higher-Order IMD Terms at IM3-Offset . 119
6.3.3 Intuition of Approximation 121
6.3.4 Justification of Approximation 121
6.3.5 Greater-than-Two-Tone Blocking Scenarios 122
6.3.6 Use of Functional Basis Elements 123
6.4 Analysis of Residual Error from Successive Regeneration Approximation 123
6.4.1 General Approach 123
6.4.2 General Case for Odd-Order Intermodulation Distortion 124
6.4.3 Single Polynomial Term Nonlinear Term Generator Case for Odd-Order Intermodulation Distortion 130
6.4.4 Even-Order Intermodulation Distortion 132
6.4.4.1 Remarks 134
6.5 Simulation of Residual Error from Successive Regeneration Approximation 135
6.5.1 General Information 136
6.5.1.1 Evaluated Waveforms 136
6.5.1.2 Computation of Nonlinear Path INR 136
6.5.2 Odd-Order Intermodulation Products 137
6.5.2.1 Term Generator and Frequency Offsets Considered 137
6.5.2.2 Receiver Odd-Order Polynomial Model 137
6.5.2.3 Reference Basis Set 138
6.5.2.4 Mathematical Simulation Setup 139
6.5.2.5 Results and Remarks 139
6.5.3 Even Order Intermodulation Products 140
6.5.3.1 Term Generator and Frequency Offsets Considered 140
6.5.3.2 Receiver Even-Order Polynomial Model 143
6.5.3.3 Reference Basis Set 143

6.5.3.4 Mathematical Simulation Setup 143
6.5.3.5 Results and Remarks 143

7 Implementation and Testing of a Rail-to-Rail Input Receiver with Successive Regeneration and Adaptive Cancellation of Intermodulation Products 144

7.1 Receiver Architecture 144
7.1.1 RF/Analog Front End Architecture 146
7.1.1.1 Linear Path 147
7.1.1.2 Odd Nonlinear Path 147
7.1.1.3 Even Nonlinear Path 148
7.1.1.4 Frequency Generation 148
7.1.1.5 ADCs and Digital Interface 149
7.1.2 Digital Back End Architecture 149
7.2 90nm Die Circuit Blocks 152
7.2.1 Linear Receiver Path 152
7.2.1.1 $\frac{1}{4}$-Phase Passive Mixer 152
7.2.1.2 Common-Gate TIA Buffer 153
7.2.1.3 TIA and Baseband Filter 156
7.2.1.4 Baseband OTAs 156
7.2.2 Nonlinear Receiver Path 156
7.2.2.1 Odd IMD Nonlinear Reference and Common-Gate Buffers 156
7.2.2.2 Even IMD Nonlinear Reference and Buffer 158
7.2.2.3 Passive Mixers and Baseband Circuitry 159
7.2.3 Frequency Generation 159
7.2.3.1 QVCO 159
7.2.3.2 $\frac{1}{4}$-Phase Logic and Multiply-by-2 162
7.2.3.3 Divide-by-2 162
7.2.3.4 Frequency Divider Chain 162
7.3 Experimental Setup 164
7.4 Experimental Results 167
7.4.1 Baseline Receiver Tests 167
7.4.2 Modulated Blocker Receiver Tests 168
7.4.3 Relation of Measured Modulated Blocker Receiver Tests to Theoretical Achievable Cancellation Ratio 170

8 Conclusion 175

A Achievable IER In Presence of Gain and Phase Mismatch for Canonical LMS-Based Algorithms **177**

B Dc Offset Removal Using the LMS Algorithm and its Equivalency to a High Pass Filter **180**

C Abbreviations Used in Text **183**

Bibliography **187**

List of Figures

1.1 Depiction of a typical radio-frequency data link. 2
1.2 Depiction of a typical superheterodyne radio-frequency data link. 2
1.3 Depiction of signal-to-error ratio over a fixed frequency band. 4
1.4 Depictions of error. a) Inherent. b) Self-generated. 4
1.5 Common classifications of nonlinear distortion in circuits. a) Harmonic(HD). b) Intermodulation(IMD). c) Cross-modulation(XMD). 5
1.6 Depiction of distortion. a) Even-order. b) Odd-order. 6
1.7 Depiction of traditional 1-dB compression point test. 7
1.8 Depiction of out-of-band 1-dB desensitization test. 8
1.9 Depiction of traditional second-order intermodulation distortion input-referred intercept point (IIP2) test. 9
1.10 Depiction of traditional third-order intermodulation distortion input-referred intercept point (IIP3) test. 9
1.11 Depiction of insufficiency of traditional definition of IIP3. 10
1.12 Depiction of single-point effective extrapolated IIP3. 10
1.13 Out-of-band blocking specification for UMTS Region 1. 11
1.14 Attenuation provided by on-chip RLC filtering. 12
1.15 Method of trading off power for linearity in receiver circuits while maintaining constant noise figure. 13
1.16 Basic feedback concept. 14
1.17 Basic common-source RF amplifier schematic. 15
1.18 Basic predistortion concept. 16
1.19 Basic piecewise linearization concept. 16
1.20 Piecewise linearization as implemented in a Gilbert multi-tanh circuit. 17
1.21 Basic feedforward concept. 17
1.22 Simplified schematic of modified derivative superposition linearity-enhanced LNA. . . 18
1.23 General adaptive feedforward error cancellation concept. 18
1.24 Least mean squares (LMS) equalization concept. 20

1.25 Problem with baseband nonlinear term generation - inherent low pass filtering at mixer output. . . . 21
1.26 Basic passive mixer-based receiver front end architecture, showing inherent low-pass filtering prior to first baseband voltage. . . . 21
1.27 Signal bands required for digitization of IM3-producing blockers in UMTS Region 1 (after downconversion to baseband) in a purely digital adaptive feedforward distortion cancellation scheme. . . . 22
1.28 a) Depiction of expansion of nonlinear input-output characteristic to one exhibiting strong nonlinearity. b) Weakening of input-output characteristic nonlinearity. . . . 23
1.29 Motivations for large-signal handling receivers. . . . 24
1.30 Dual NMOS/PMOS LNA as depicted in [53]. . . . 24

2.1 Blocking problem in FDD receivers. . . . 27
2.2 a) Nonlinearity problem in typical integrated receivers. b) Use of a SAW filter to alleviate nonlinearity problem. . . . 27
2.3 Adaptive feedforward IM3 distortion cancellation - proposed receiver system. . . . 29
2.4 Reduction of required adaptive filter FIR taps via use of IIR filter to compensate for known path difference. . . . 30
2.5 Dc offset cancellation problem in LMS filters. . . . 31
2.6 Proposed dc offset cancellation scheme showing dc offset transient waveforms. . . . 31
2.7 Exemplary digital dc offset correction circuit. . . . 32
2.8 Insufficiency of complex LMS/NLMS in presence of alternate and main path signal vector mismatches. . . . 34
2.9 IER vs. rotational mismatch and difference in phase mismatch. . . . 36
2.10 Expected UMTS blocker profile at various points in the receiver. . . . 37
2.11 Hardware implementation of complex NLMS algorithm. a) Canonical. b) Proposed architecture with enhanced degree of freedom. . . . 39
2.12 Proposed procedure for determining a turn-on condition for the linearity enhancement. 40
2.13 Proposed procedure for determining a turn-off condition for the linearity enhancement. 40

3.1 Experimental UMTS receiver architecture. . . . 43
3.2 ADC-FPGA bidirectional interface. . . . 44
3.3 Implementation of main path LNA and balun. a) Schematic depiction. b) Balun 3-D CAD representation. . . . 44
3.4 Main path high-IIP2 mixer. . . . 45
3.5 Mixer common-mode feedback circuit. . . . 45
3.6 Main path Cherry-Hooper LO buffer. . . . 46

3.7 Main path 2nd-order biquad. 46
3.8 Main path gain and buffering stage prior to ADC. 47
3.9 Cubic term generator schematic. 48
3.10 Simulated phase response of cubic term generator interstage phase response. 48
3.11 Alternate path mixer schematic. 49
3.12 Alternate path LO buffer schematic. 49
3.13 Alternate path baseband filter and ADC driver. 50
3.14 Depiction of blocker mapping algorithm. 52
3.15 Worst-case undesired IM3 product baseband profile as function of frequency. 52
3.16 Experimental implementation of proposed receiver. 53
3.17 Zoomed-in view of bonded chip mounted on gold-plated Rogers board. 53
3.18 RF front end die photo. 54
3.19 Experimental test setup picture. 54
3.20 Experimental test setup simplified schematic. 55
3.21 Modified two-tone test concept. 56
3.22 Measured results of modified two-tone test. a) For proposed enhanced complex NLMS architecture. b) For canonical complex NLMS architecture. 56
3.23 Measured results of modified two-tone test in which IM3 products contain squared TX leakage. a) Concept. b) Results. 57
3.24 Measured results of modified two-tone test for -26dBm TX leakage, -34dBm CW blocker swept over LO frequency. a) Input-referred error vs. f_{LO}. b) IIP3 vs. f_{LO}. 57
3.25 Measured results of modified two-tone test for -25dBm TX leakage, -33dBm CW blocker swept over LO frequency, input-referred error vs. f_{LO}. 57
3.26 Measured convergence behavior of adaptive equalization algorithm. 58
3.27 Digital spectrum analyzer measurement of power spectral density with and without correction. 58
3.28 Measured receiver DPCH despread SNR and BER under sensitivity, sensitivity / blocking / correction and sensitivity / blocking / no correction, respectively. 59
3.29 Measured INR performance of alternate path. 60

4.1 Schematic of cubic term generator used in Chapter 3. 62
4.2 Implementation of gain between nonlinear operations in multistage cubic term generator. 64
4.3 Effects of baseband frequency response on IM3 products in a multistage cubic term generator for a two-tone test. 65
4.4 Frequency domain depiction of proposed cubic term generator internal bandwidth limitation. 65

4.5 Frequency translation of bandpass noise by multiple blocker signals to desired signal level. . . . 69
4.6 Schematic of cubic term generator used for simulation. . . . 70
4.7 a) Calculated vs. simulated INR due to input stage components swept over V_{OD}. b) Simulated INR, input stage devices (Instage) and all devices (Total), swept over V_{OD}. 71
4.8 a) Calculated vs. simulated INR due to input stage components swept over A_2/A_1 ratio. b) Simulated INR, input stage devices (Instage) and all devices (Total), swept over A_2/A_1 ratio. . . . 71
4.9 Schematic of cubing circuit from [88]. . . . 71
4.10 Depiction of effect of interstage group delay on CW and BPSK blocking signals producing IM3 products. . . . 73
4.11 Baseband PSD (single-ended) of squared QPSK modulated signal for a) I and b) Q channels when Θ=0°. . . . 82
4.12 Baseband PSD (single-ended) of squared QPSK modulated signal group delay-related error for a) I and b) Q channels when Θ=0°. . . . 82
4.13 Calculated IER as a function of group delay for T=Tr. . . . 83
4.14 Calculated IER vs. blocker modulation chip rate for fixed delay. . . . 83

5.1 General CMOS push-pull LNTA topology. . . . 89
5.2 Available signal swing of cascoded common-source and common-gate transconductors. 90
5.3 Biasing strategy for stacked class-AB common-gate LNTA. . . . 92
5.4 Stacked push-pull LNTA large-signal current flows. . . . 93
5.5 Circuit structure utilized for quantitative analysis of LNTA. . . . 94
5.6 Calculated output characteristics of single-ended LNTA for square-law MOSFET model. a) Output current. b) Transconductance. c) Residual after removal of small-signal linear fit. . . . 95
5.7 Intuitive shape of multiplicative term required to effectively linearize LNTA. . . . 96
5.8 Nonlinear effects contributing to the effective linearization of the LNTA for large signals. a) Mobility degradation ($VS_n(x)$, $VS_p(x)$). b) Transition into triode region ($F_{s,n}(x)$, $F_{s,p}(x)$). . . . 97
5.9 Calculated output characteristics of single-ended LNTA for MOSFET model incorporating mobility degradation effects. a) Output current. b) Transconductance. c) Residual after removal of small-signal linear fit. . . . 98
5.10 Calculated output characteristics of differential LNTA for MOSFET model incorporating mobility degradation effects. a) Output current. b) Transconductance. c) Residual after removal of small-signal linear fit. . . . 98

5.11 Calculated output characteristics of differential LNTA for MOSFET model incorporating continuous transition from MOS saturation to triode region. a) Output current. b) Transconductance. c) Residual after removal of small-signal linear fit. 100

5.12 Calculated output characteristics of differential LNTA for MOSFET model incorporating subthreshold conduction effects. a) Output current. b) Transconductance. c) Residual after removal of small-signal linear fit. 101

5.13 Intuitive illustration of generation of transconductance W-shape. 102

5.14 Intuitive illustration of generation of transconductance W-shape - second chain rule term. 103

5.15 Calculated output characteristics of differential LNTA for MOSFET model incorporating subthreshold conduction effects, for wide input range. a) Output current. b) Transconductance. c) Residual after removal of small-signal linear fit. 104

5.16 Calculated input-referred upper-sideband IMD products for two-CW-tone blocking scenario. a) For two equal magnitude CW blockers. b) For asymmetric CW blockers. . . 104

5.17 Calculated input-referred upper-sideband IMD products for two-CW-tone blocking scenario with slope-of-3 labels. a) For two equal magnitude CW blockers. b) For asymmetric CW blockers. 105

5.18 Calculated input-referred LNTA out-of-band desensitization (160MHz offset) for a CW-tone blocking scenario. a) Effective transconductance. b) Relative transconductance. . 106

5.19 Simplified schematic of implemented LNTA. 107

5.20 Static LNTA simulation: transconductance as function of differential input voltage over corner. 107

5.21 Static LNTA simulation: small signal linear fit residual as function of differential input voltage over corner. 108

5.22 Static LNTA simulation: S_{11} as function of differential input voltage over corner. . . . 108

5.23 Miscellaneous dynamic LNTA simulations using QPSS/QPNOISE. a) Effective transconductance gain. b) Noise figure. c) Quiescent current. 109

5.24 Dynamic LNTA odd-order two-tone IMD simulations using QPSS (second tone 29dB less than large tone). a) Zoomed out. b) Annotated. c) Zoomed in. 110

5.25 Cubic-Gaussian basis fit to nonlinear residual curves. a) Fitting to calculated model. b) Fitting to simulated model. 112

5.26 Simplified schematic of implemented center class-AB OTA for LNTA. 112

6.1 Depiction of higher-order intermodulation distortion products. 115

6.2 Naïve approaches to cancelling higher-order intermodulation distortion products. a) Trivial extension of work in Chapter 2. b) Moving adaptive cancellation algorithm from digital into analog baseband. 115
6.3 Polynomial fit to residual I-V characteristic of LNTA after small-signal linear fit is removed. 116
6.4 Successive regeneration and adaptive feedforward cancellation of intermodulation products at baseband implemented in this work. a) Concept. b) Simplified system block diagram. 117
6.5 Depiction of insufficiency of pure cubic term generator for odd nonlinear reference in the presence of higher-order terms in the nominally linear path. 119
6.6 Successive regeneration and adaptive feedforward cancellation of intermodulation products at baseband implemented in this work showing explicitly the higher-order terms present in the odd-order nonlinear term generator. a) Concept. b) Simplified system block diagram. 120
6.7 Intuitive approach as to why baseband successive regeneration is a good approximation for large blocker asymmetries. a) Considering all blockers as a single narrowband signal $s(t)$. b) Considering the envelope of $s(t)$ and the terms immediately surrounding dc. c) Considering higher-order odd IMD products at RF and those surrounding the LO frequency. 122
6.8 Application-based justification for a dominant-blocker approximation. 123
6.9 Successive regeneration and adaptive feedforward cancellation of IMD products at baseband accounting for envelope baseband memory effects. 135
6.10 Comparison of higher-order moments of the amplitudes of several narrowband modulation schemes. a) Zoomed in. b) Zoomed out (Log Scale). 137
6.11 a) 11th-order polynomial best-fit to LNTA input-output characteristic utilized for successive regeneration scheme residual error simulations. b) Simulated input-referred odd-order IM distortion for 11th-order polynomial in a) and "S2" input signal for various IMD offset frequencies. 138
6.12 Concept of higher-order functional reference basis elements for case of impure-odd-order nonlinear term generator. 139
6.13 Simulation and calculation results of approximation error IER for case "S2" as function of $E[a^2]/E[b^2]$ for noiseless nonlinear paths. 140
6.14 Simulation and calculation results of approximation error IER_{Odd} for case "S1T1F1". 141
6.15 Simulation and calculation results of approximation error IER_{Odd} for case "S2T1F1". 141
6.16 Simulation and calculation results of approximation error IER_{Odd} for case "S1T2F1". 141
6.17 Simulation and calculation results of approximation error IER_{Odd} for case "S2T2F1". 142

6.18 Simulation and calculation results of approximation error IER_{Odd} for case "S1T2F2". 142
6.19 Simulation and calculation results of approximation error IER_{Odd} for case "S2T2F2". 142
6.20 Simulation and calculation results of approximation error IER_{Even}. 143

7.1 Depiction of reciprocal mixing with large blocker. 145
7.2 Proposed receiver architecture with detail on RF/Analog front end. 146
7.3 Implementation of PLL loop with off-chip components. 148
7.4 Proposed digital back end architecture. 150
7.5 Hardware implementation of complex LMS equalizer (filter portion) modified to compensate for I/Q mismatch. 151
7.6 Hardware implementation of complex modified-NLMS equalizer (filter portion) modified to compensate for I/Q mismatch. 151
7.7 Schematic of passive mixer and servo biasing. 153
7.8 Simplified schematic of linear path common-gate buffer. 154
7.9 Analysis of optimal g_m of common-gate buffer. 154
7.10 Inverse of input-referred thermal noise power density of CG buffer as function of input device transconductance. 155
7.11 Block diagram of TIA and biquad with dc offset compensation. 156
7.12 Simplified schematic of two-stage OTA used in baseband circuitry. 157
7.13 Simplified schematic of odd nonlinear term generator. 157
7.14 Simplified schematic of even nonlinear term generator. 158
7.15 Depiction of dual-Q ring VCO concept. 160
7.16 Depiction of dual-Q ring VCO after Π-Y transformation. 160
7.17 Simplified schematic of implemented QVCO. 161
7.18 Frequency generation circuitry simulations. a) Time-domain outputs of QVCO. b) Simulated total noise of the oscillator and 1/4 phase logic circuitry. 161
7.19 Simplified schematic of implemented 1/4-phase logic circuitry. 163
7.20 Block diagram of CML divide-by-2 circuit. 163
7.21 Simplified schematics of a) CML D-flip-flops and b) actively-loaded Cherry-Hooper buffers. 164
7.22 Photograph of assembled RF/Analog front end. 165
7.23 Photograph of mounted RF/Analog 90nm die with RF and local baseband connections. 165
7.24 RF/Analog chip die photo. 166
7.25 Photograph of experimental setup. 166
7.26 Schematic of experimental setup. 167

7.27 Two-CW tone IIP3 test: measured input-referred error, effective IIP3, and out-of-band desensitization of standalone RF/Analog die. 168
7.28 Measured baseline receiver performance metrics. 168
7.29 Measured input-referred error with various degrees of cancellation. a) Even- and odd-order cancellation for two-signal blocking. b) Even order cancellation for large QPSK-modulated blocker and phase noise floor (measured with CW blocker). 169
7.30 a) Measured input-referred error as function of modulation bandwidth for +12.4dBm QPSK blocker / -16.3dBm CW blocker. b) Convergence behavior of full adaptive algorithm for +11.4dBm 2MSPS QPSK blocker / -16.3dBm CW blocker. 169
7.31 Measured input-referred even-order IM distortion error as function of modulation bandwidth for +12.4dBm QPSK blocker only. 170
7.32 Cartoons depicting a) Spreading factor concept. b) Noise spreading out of and into channel bandwidth. 171
7.33 PSD of various signals within modified IER simulation and calculation. a) Baseband QPSK large blocker and odd-order IM distortion. b) Signal and noise in nonlinear paths. 171
7.34 Measured in-channel INR of nonlinear paths, with and without spurious content at low frequencies. 172
7.35 Simulation and modified calculation results of approximation error IER_{Odd} for case "S1T1F1". 173
7.36 Simulation and modified calculation results of approximation error IER_{Odd} for case "S1T2F1". 173
7.37 Simulation and modified calculation results of approximation error IER_{Odd} for case "S1T2F2". 173

B.1 LMS with dc offset reference input. 181
B.2 Equivalence of LMS with dc offset removal to high pass filter. 181

Chapter 1

Background Material and Introduction[1]

1.1 Problem Definition

1.1.1 Radio-Frequency (RF) Systems

Radio-frequency (RF) communication systems exchange information wirelessly via the propagation of electromagnetic waves. As depicted in Fig. 1.1, data is sent by an RF transmitter circuit, which multiplies an information-bearing complex baseband signal $a(t) = a_I(t) + a_Q(t)$ with a high-frequency complex exponential $c(t) = e^{j2\pi f_0 t} = \cos(2\pi f_0 t) + j\sin(2\pi f_0 t)$ and then applies the resultant high-frequency narrowband signal to an electromagnetic radiating structure such as an antenna. The radiated signal then travels in free space as an electromagnetic wave. The purpose of an RF receiver circuit in such a system is to convert this electromagnetic signal incident on a receiver antenna back into its complex baseband representation so that relevant information can be extracted from $a(t)$.

It is important to note that the direct application of $c(t)$ to $a(t)$ is denoted direct-conversion RF. While the majority of contemporary commercial RF transmitters and receivers employ direct-conversion architectures, including all of those to be discussed in this dissertation, for many decades the predominant radio architecture was known as a superheterodyne system, depicted in Fig. 1.2, which relied on the application of multiple frequency shifts interspersed with filters. The reasons for the previous dominance of the superheterodyne architecture and the recent shift to direct-conversion topologies have a direct bearing on the topic of this dissertation and will be discussed more fully below.

[1] Portions of this material have been previously published in [1] and [2].

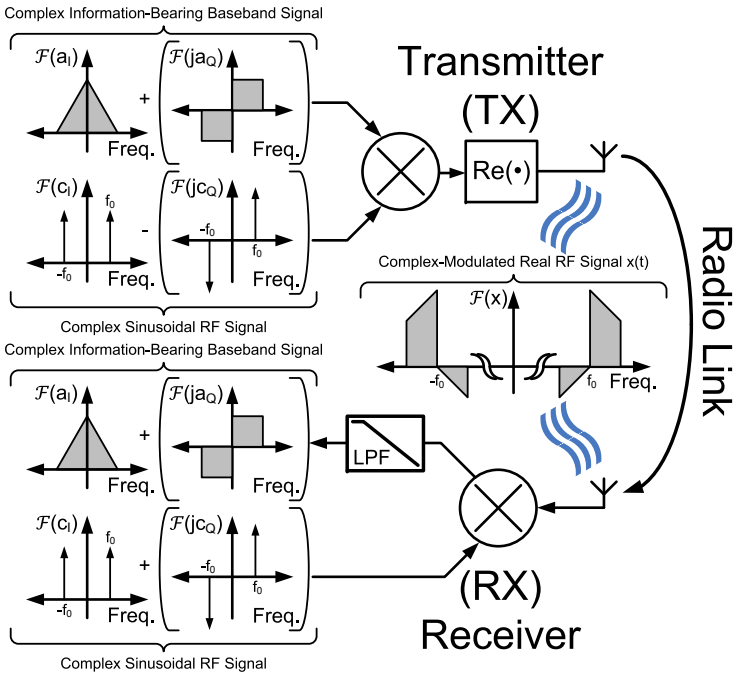


Figure 1.1: Depiction of a typical radio-frequency data link.

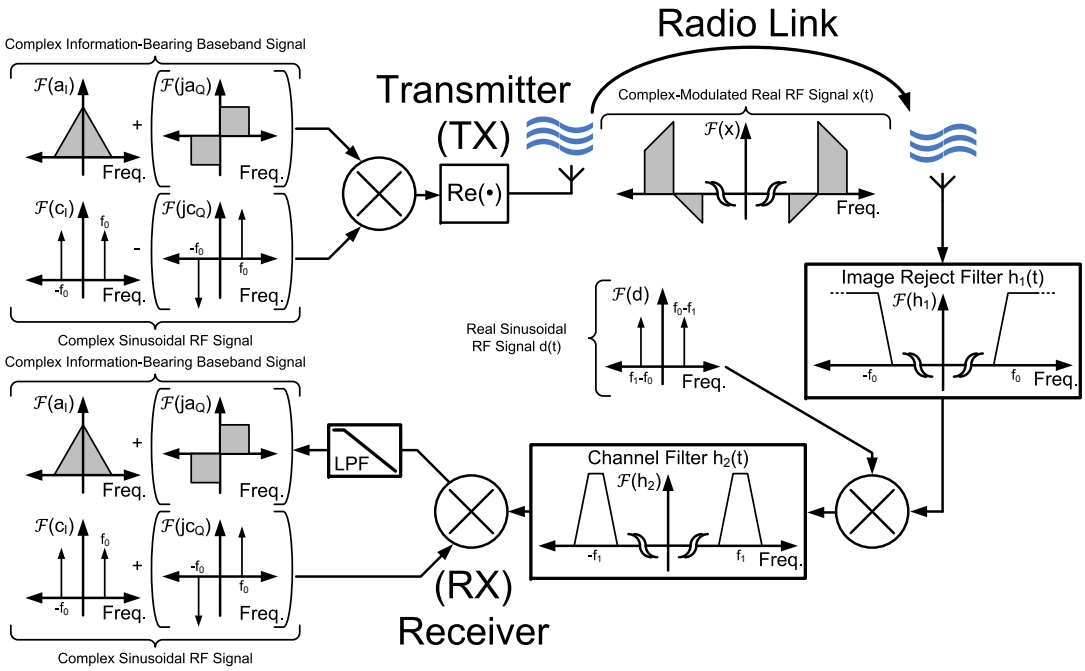


Figure 1.2: Depiction of a typical superheterodyne radio-frequency data link.

1.1.2 Error in RF Systems

The reliability of an RF communication link is often best described by its bit error rate (BER) [3, pg. 23], which is most often specified explicitly under particular conditions by a standards-

Standard	Input Signal Condition	Required Error Rate	Reference
GSM	-102dBm reference input/no undesired signals	BER $< 10^{-4}$	[4]
GSM	-99dBm reference input/large undesired signals	BER $< 10^{-3}$	[4]
cdma2000	-104dBm reference input/no undesired signals	FER < 0.005	[5]
cdma2000	-101dBm reference input/large undesired signals	FER < 0.01	[5]
UMTS	-117dBm data input/-106.7dBm total ref. input /no undes. signals	BER $< 10^{-3}$	[6]
UMTS	-114dBm data input/-103.7dBm total ref. input /large undes. signals	BER $< 10^{-3}$	[6]

Table 1.1: Examples of data error rate specifications for notable communications standards.

Standard	Modulation Scheme	Key Relationship	Reference
GSM	GMSK	BER = 10^{-4} for SER = 10dB	[7, pg. 719]
GSM	GMSK	BER = 10^{-3} for SER = 8dB	[7, pg. 719]
cdma2000	QPSK	FER = 0.01 for SER = 3.6dB	[8]
UMTS	QPSK	BER = 10^{-3} for despread SER = 1dB	[9]
UMTS	QPSK	BER = 10^{-3} for spread (SF=128) SER = -20dB	[9]

Table 1.2: Examples of significant relationships between data error rate specifications and signal-to-error ratio

setting committee. Some notable examples of BER specifications from modern communications standards are summarized in Table 1.1. For the cdma2000 standard, the performance of the radio link is specified by the frame error rate (FER), where a frame is a collection of bits involved in both data communication and communication link control. For each of these conditions, the signal-to-error ratio (SER), depicted in Fig. 1.3, of the received baseband communication signal $a(t)$ must be above a particular level in order to achieve a particular bit error rate. That is, for some constant K, the relationship in (1.1) must hold to obtain a particular BER, where $S(f)$ and $E(f)$ are the power spectral densities of the signal and error, respectively. This level K depends on many factors associated with the communications system, including type of modulation used, type and parameters of convolutional coding used, and any code spreading used. Some commonly encountered relationships between SER and BER are detailed in Table 1.2 where the error is assumed to be Gaussian white noise.

$$\frac{\int_{f_0}^{f_1} S(f)df}{\int_{f_0}^{f_1} E(f)df} > K \tag{1.1}$$

Error in the RF receiver can be classified into two types, as depicted in Fig. 1.4. One is inherent error that is present in the absence of a signal incoming to the receiver. This class includes all types of noise processes and dc offset. The other is error that is self-generated either in response to the desired incoming signal or to undesired incoming signals. This class includes error due to I/Q mismatch residue and nonlinear distortion.

1.1.3 Nonlinear Distortion in RF Systems

Corruptive nonlinear distortion can take the form of either harmonic distortion (HD), intermodulation distortion (IMD), and cross-modulation distortion (XMD). As depicted in Fig. 1.5, HD arises at multiples of the fundamental frequency of a large signal. IMD arises at permutations of high and low frequency shifts induced by two or more signals. Essentially, these distortion compo-

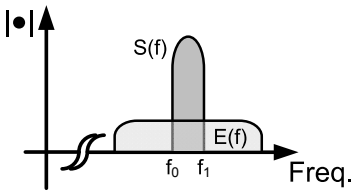


Figure 1.3: Depiction of signal-to-error ratio over a fixed frequency band.

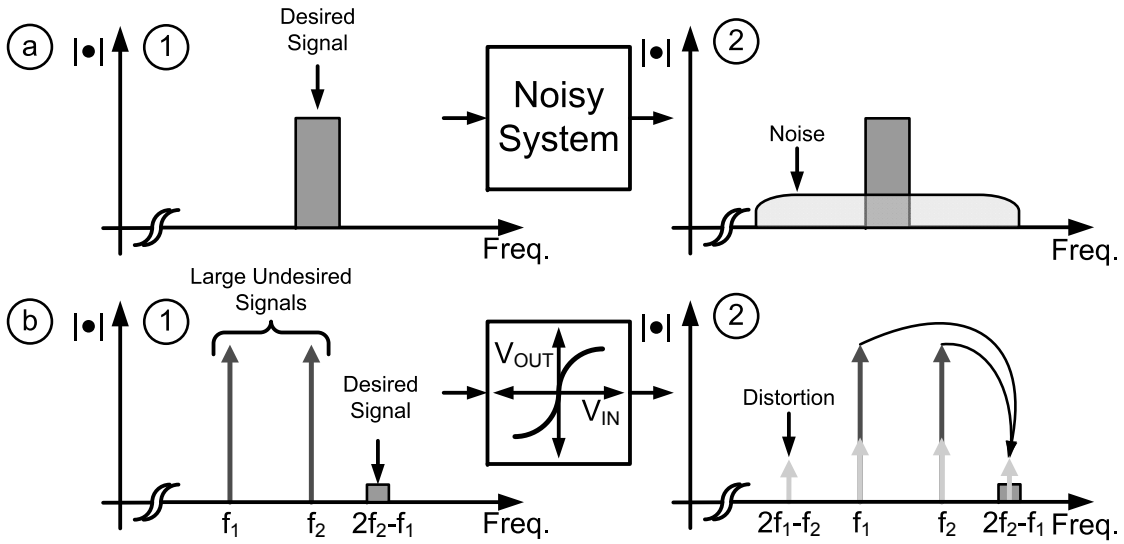


Figure 1.4: Depictions of error. a) Inherent. b) Self-generated.

nents arise due to the signal mixing that occurs when two or more signals are raised to a polynomial power greater than unity [10, pg. 17]. As a trivial example, consider two RF cosines at frequencies f_1 and f_2 such that $a(t) = A_1 \cos(2\pi f_1 t) + A_2 \cos(2\pi f_2 t)$. When passed through a system with static nonlinear transfer function $g(x) = \alpha_3 x^3$, the following third-order distortion products in (1.2) result. Those terms involving more than one frequency component, such as $2f_2 - f_1$, are termed intermodulation distortion (IMD) components.

$$\begin{aligned} g(a(t)) = &\left(\frac{3}{4}\alpha_3 A_1^3 + \frac{3}{2}\alpha_3 A_1 A_2^2\right)\cos(2\pi f_1 t) + \left(\frac{3}{4}\alpha_3 A_2^3 + \frac{3}{2}\alpha_3 A_2 A_1^2\right)\cos(2\pi f_2 t) \\ &+ \frac{1}{4}\alpha_3 A_1^3 \cos(2\pi(3f_1 t)) + \frac{1}{4}\alpha_3 A_2^3 \cos(2\pi(3f_2 t)) \\ &+ \frac{3}{4}\alpha_3 A_1^2 A_2 \cos(2\pi(2f_1 + f_2)t) + \frac{3}{4}\alpha_3 A_2^2 A_1 \cos(2\pi(2f_2 + f_1)t) \\ &+ \frac{3}{4}\alpha_3 A_1^2 A_2 \cos(2\pi(2f_1 - f_2)t) + \frac{3}{4}\alpha_3 A_2^2 A_1 \cos(2\pi(2f_2 - f_1)t) \end{aligned} \quad (1.2)$$

For two large blocker signals near the desired signal with small $|f_2 - f_1|$, on-chip filtering cannot easily attenuate the IMD-producing signals without significantly attenuating the desired signal as well. Hence, any small desired signal falling at a frequency of $2f_2 - f_1$ will be corrupted by IMD

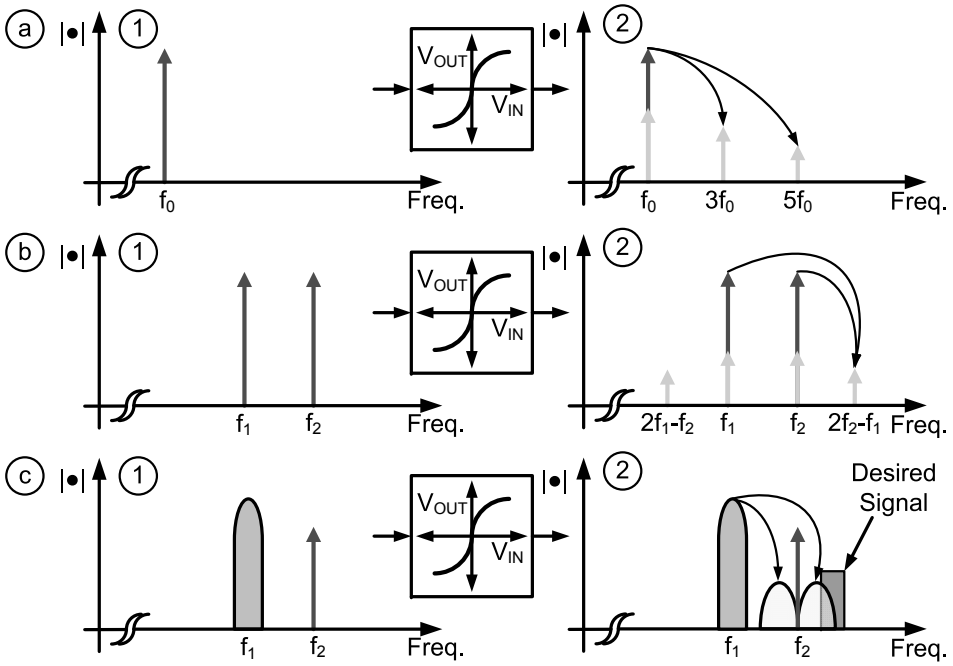


Figure 1.5: Common classifications of nonlinear distortion in circuits. a) Harmonic(HD). b) Intermodulation(IMD). c) Cross-modulation(XMD).

products in the presence of two large blockers at those frequencies. As these IMD products are produced by a third-order nonlinearity, they are termed IM3 products. This represents a much more significant problem than HD, where the distortion terms generated by large blocker signals near the desired signal are easily filtered out at no matter what frequency the desired signal passband exists. Furthermore, for the more general case of complex modulated bandpass signals $a(t) = a_I(t)\cos 2\pi f_1 t + a_Q(t)\sin(2\pi f_1 t)$ and $b(t) = b_I(t)\cos(2\pi f_2 t) + b_Q(t)\sin(2\pi f_2 t)$, the following third-order distortion terms arise in (1.3), where the time arguments have been removed in the complex baseband signals for simplicity:

$$
\begin{aligned}
g(a(t)+b(t)) = \alpha_3 \frac{3}{4}\cdot \\
&\big(((a_I^2+a_Q^2)+2(b_I^2+b_Q^2))a_I\cos(2\pi f_1 t) + ((a_I^2+a_Q^2)+2(b_I^2+b_Q^2))a_Q\sin(2\pi f_1 t) \\
&+((b_I^2+b_Q^2)+2(a_I^2+a_Q^2))b_I\cos(2\pi f_2 t) + ((b_I^2+b_Q^2)+2(a_I^2+a_Q^2))b_Q\sin(2\pi f_2 t) \\
&+(\tfrac{1}{3}a_I^3 - a_I a_Q^3)\cos(2\pi f_1 t) + (\tfrac{-1}{3}a_Q^3 + a_I^2 a_Q)\sin(2\pi f_1 t) \\
&+(\tfrac{1}{3}b_I^3 - b_I b_Q^3)\cos(2\pi f_2 t) + (\tfrac{-1}{3}b_Q^3 + b_I^2 b_Q)\sin(2\pi f_2 t) \\
&+(a_I^2 b_I - 2a_I a_Q b_Q - a_Q^2 b_I)\cos(2\pi(2f_1+f_2)t) + (a_I^2 b_Q + 2a_I a_Q b_I - a_Q^2 b_Q)\sin(2\pi(2f_1+f_2)t) \\
&+(a_I b_I^2 - 2a_Q b_I b_Q - a_I b_Q^2)\cos(2\pi(f_1+2f_2)t) + (a_Q b_I^2 + 2a_I b_I b_Q - a_Q b_Q^2)\sin(2\pi(f_1+2f_2)t) \\
&+(a_I^2 b_I + 2a_I a_Q b_Q - a_Q^2 b_I)\cos(2\pi(2f_1-f_2)t) + (a_I^2 b_Q - 2a_I a_Q b_I - a_Q^2 b_Q)\sin(2\pi(2f_1-f_2)t) \\
&+(a_I b_I^2 + 2a_Q b_I b_Q - a_I b_Q^2)\cos(2\pi(2f_2-f_1)t) + (a_Q b_I^2 - 2a_I b_I b_Q - a_Q b_Q^2)\sin(2\pi(2f_2-f_1)t)\big)
\end{aligned}
\tag{1.3}
$$

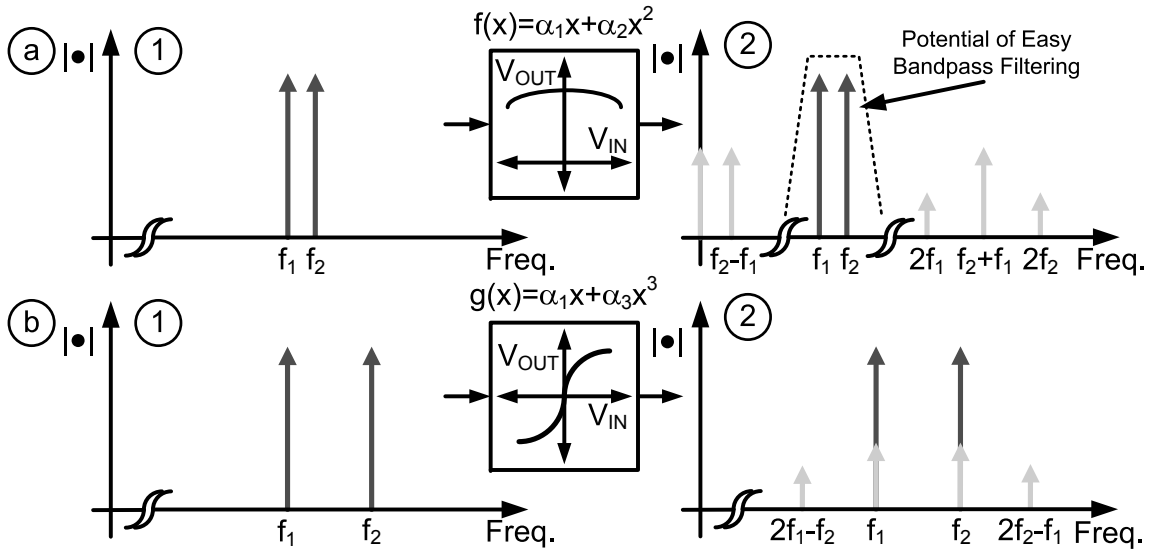


Figure 1.6: Depiction of distortion. a) Even-order. b) Odd-order.

It can be seen in (1.3) that if $a(t)$ is a desired signal at f_1, it itself is corrupted, although to a much lesser extent than the small desired signal in the previous case. For these reasons, large undesired incoming signals that can potentially create self-generated error are denoted as "blockers" or "jammers", as the distortion error terms that result can prevent decoding of the desired signal.

As shown in Fig. 1.6, IMD can be either even-order or odd-order in nature, each constituting a separate set of issues. Odd-order IMD such as in the conditions previously described result in corruptive components close to the error generating signals. Even-order IMD results in corruptive components near dc or near even-order harmonics of the IMD-producing signals. To see this, one can consider the same two RF cosines as before with static nonlinear transfer function $h(x) = \alpha_2 x^2$ in (1.4):

$$h(a(t)) = \tfrac{1}{2}\alpha_2(A_1^2 + A_2^2 + A_1^2\cos(2\pi(2f_1)t) + A_2^2\cos(2\pi(2f_2)t)) \\ + \alpha_2(A_1A_2\cos(2\pi(f_1 - f_2)t) + A_1A_2\cos(2\pi(f_1 + f_2)t)) \quad (1.4)$$

In this case, even-order IMD products are in principle easy to remove using an on-chip bandpass filter, as depicted in Fig. 1.6. However, various mechanisms exist in RF receivers [11] [12] by which even-order IMD terms at baseband can effectively leak through to the baseband portion of the RF receiver where they cannot be filtered out, thereby corrupting the downconverted desired signal. For many years, problems associated with even-order IMD were among several issues that dictated the use of superheterodyne radio receivers, in which filtering after each frequency conversion helped to attenuate both low-frequency IMD and IMD-producing blockers.

XMD is a special case of odd-order IMD in which an incoming signal $a(t)$ at f_2 is corrupted by the envelope of a much larger undesired signal $b(t)$ at f_1 as seen in the first two terms of (1.3). This scenario is commonly illustrated as in Fig. 1.5c where $a(t)$ is a CW signal. Here, it can be seen that

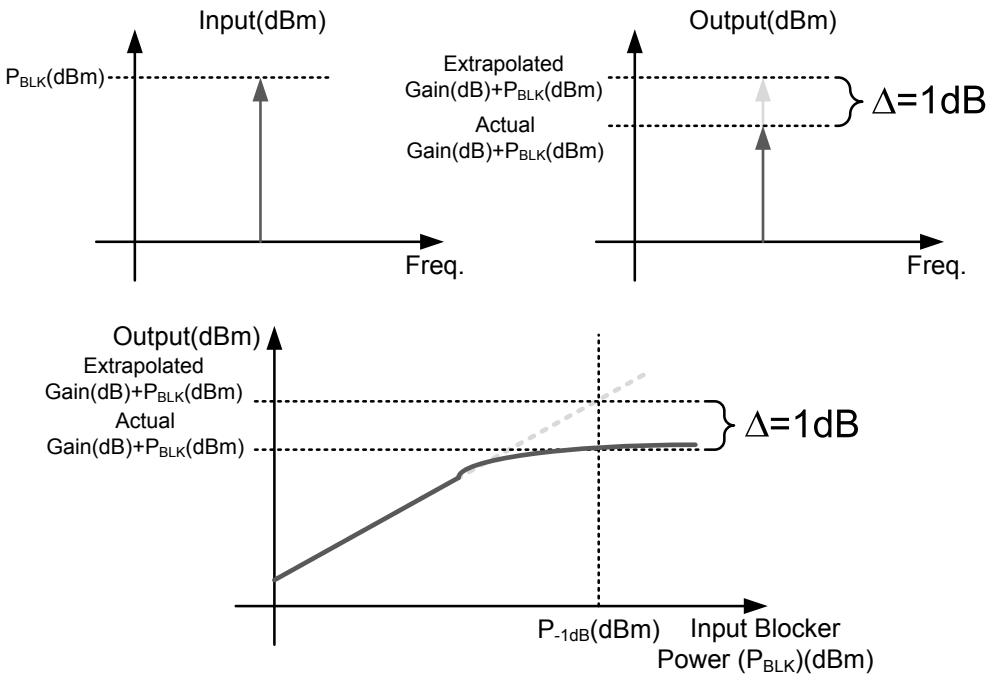


Figure 1.7: Depiction of traditional 1-dB compression point test.

the envelope of $b(t)$ is frequency-shifted to f_2 and is corrupting a nearby desired signal frequency channel. XMD can be problematic for both an undesired and desired $a(t)$, although typically the former of the two cases is limiting when both $a(t)$ and $b(t)$ are considerably larger than the desired signal.

1.1.4 Metrics of Nonlinear Distortion

Like quantities such as noise figure (NF) and sensitivity, nonlinearity metrics most often represent output error referred to the input of the receiver by the receiver small-signal gain. This allows for unbiased comparison between receivers with different small-signal gains, the exact value of which is usually unimportant so long as it is sufficiently large. This also allows for a direct comparison of the input-referred error and the power of the minimum desired signal incident on the antenna, which is usually specified by a given communications standard.

One caveat when input-referring the output error is that the small-signal gain must be computed under each blocking condition. This is due to the fact that the dominant third-order XMD resulting from the amplitude of large blockers mixing with the desired signal typically has the net effect of reducing the amplitude of the desired signal and hence the effective small-signal gain. The reason for this is twofold in that the envelope of large blockers typically has a large dc component and that the third-order polynomial constant α_3 is typically negative [10, pg. 15]. This effect is often captured in a measurement sweep, depicted in Fig. 1.7, in which a large signal is applied and increased in power until the receiver output level is 1dB less than that which would have been obtained in a completely linear system. The input power at which this point occurs is denoted the input-referred 1-dB compression point, alternatively P_{-1dB}. A more relevant way of computing the

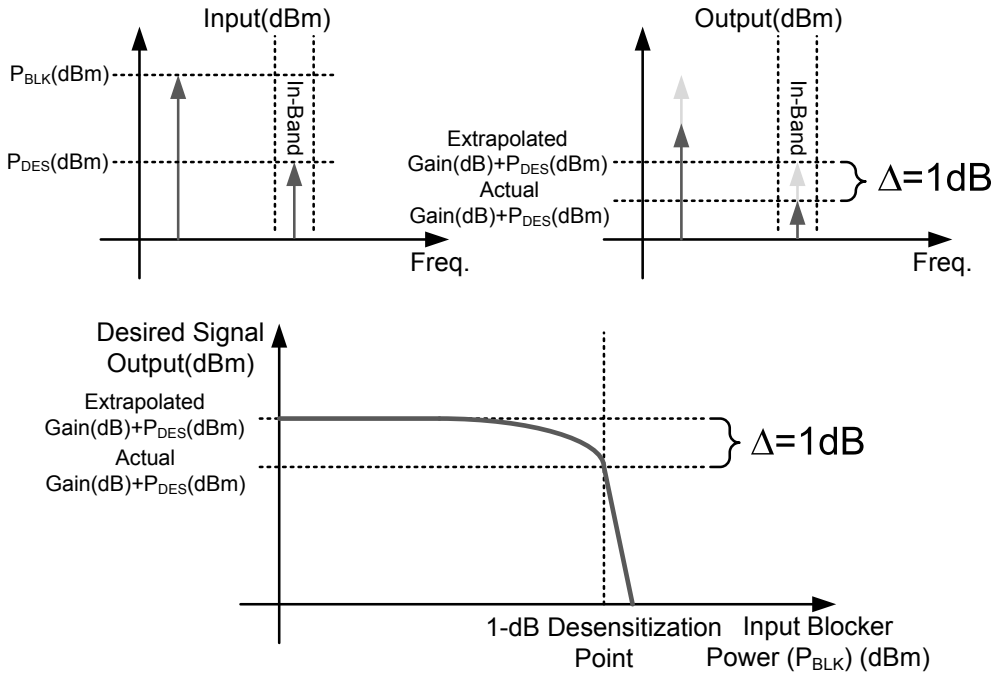


Figure 1.8: Depiction of out-of-band 1-dB desensitization test.

1-dB compression point for receivers beset by out-of-band blocking problems is depicted in Fig. 1.8 and consists of sweeping the power of an undesired blocking signal and then measuring the small-signal gain change of a desired signal near the receiver LO frequency. This particular method is often termed the desensitization test [13] or an outband compression test [14]. For the purposes of this dissertation, this metric will be denoted the out-of-band 1-dB desensitization test. According to [13], this metric is 3dB lower than the P_{-1dB} for a circuit characterized by a purely cubic nonlinearity.

Third-order nonlinear distortion is commonly characterized by a test in which two large CW signals each of power P_{BLK} are swept and the four output tones nearest the original CW signals are measured. P_{BLK} is chosen sufficiently small such that higher-order nonlinear terms and XMD-induced small-signal gain compression are negligible [10, pg. 19]. The output powers at the original and IM3 frequencies, respectively, are then logarithmically plotted as a function of P_{BLK} and extrapolated until they intercept. The value of P_{BLK} at which the two extrapolated lines intercept is denoted the input-referred intercept point for third-order distortion (IIP3), as depicted in Fig. 1.10. Extrapolating the linear and IM3 terms is necessary because IIP3 is typically about 10-20dB larger than P_{-1dB}. The same procedure is carried out for second-order distortion in which the IM2 product at the two-tone beat frequency $f_2 - f_1$ is measured instead of the IM3 product to obtain the input-referred intercept point for second order distortion (IIP2), as depicted in Fig. 1.9.

A serious problem with each of these two tests is the potential for misapplication. Worst-case blocking conditions for some communication standards can easily violate the assumptions of the aforementioned tests. In this case, the IIP3 and IIP2 metrics are not accurate predictors of the total input-referred error in response to the worst-case blocking conditions. For example, as shown in Fig. 1.11, in a hypothetical example in which IM5 products dominate the output for a worst-case

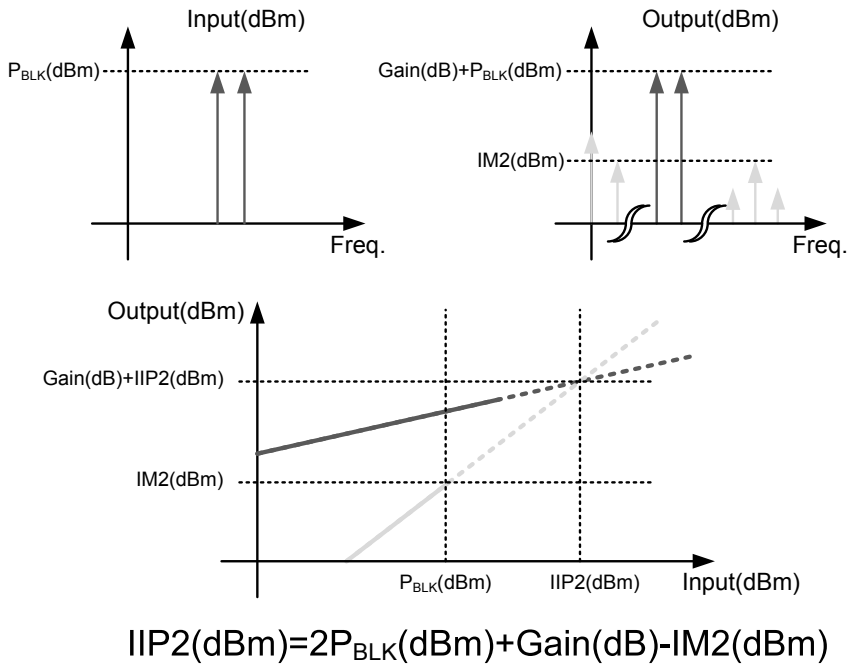


Figure 1.9: Depiction of traditional second-order intermodulation distortion input-referred intercept point (IIP2) test.

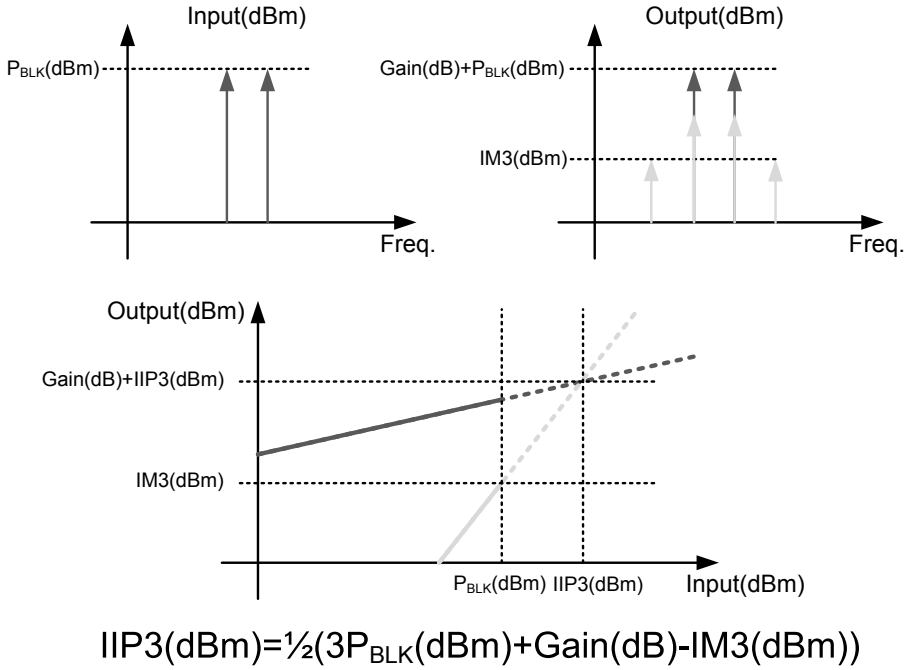


Figure 1.10: Depiction of traditional third-order intermodulation distortion input-referred intercept point (IIP3) test.

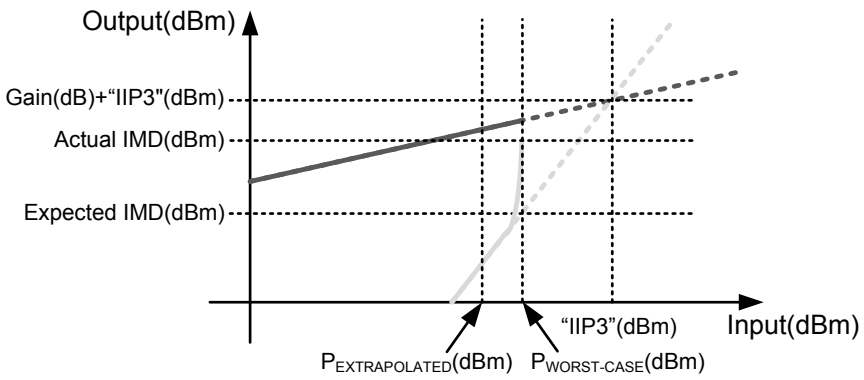


Figure 1.11: Depiction of insufficiency of traditional definition of IIP3.

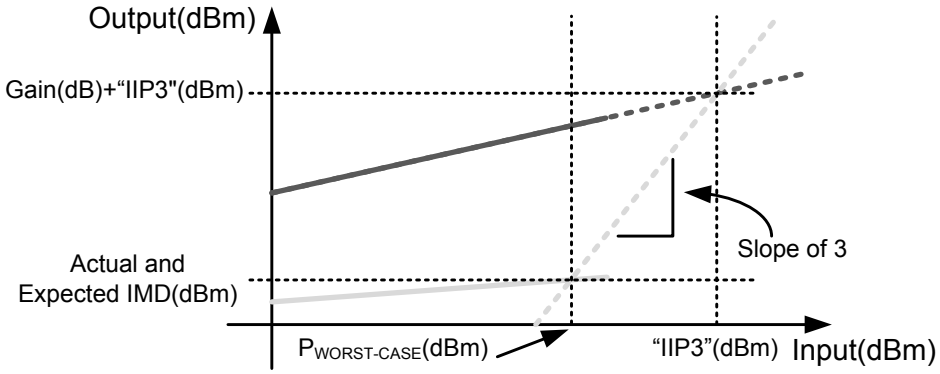


Figure 1.12: Depiction of single-point effective extrapolated IIP3.

input signal, extrapolating the IIP3 from input levels at which higher-order nonlinear terms are negligible at the output leads to a gross underestimate of the worst-case input-referred distortion error. One solution to this problem is to utilize an effective IIP3 or IIP2 obtained from a single-point extrapolation from the input-referred error obtained under worst-case blocking, as depicted in Fig. 1.12 for IIP3. In this case, the modified IIP3 and IIP2 metrics provide accurate predictors of an upper bound on total input-referred error due to odd- and even-order distortion. This permits more meaningful comparisons across different circuit blocks and architectures for a given communications standard. In order to perform the extrapolation for IIP3, (1.5) may be used [15]. In order to perform the extrapolation for IIP2 for a two-tone test with two equal-power tones each of power P_{1T}, (1.6) may be used [16]. Note that in each case, the power of the measured IMD product is referred first to the input of the circuit block or entire receiver (i.e. Gain=0dB).

$$IIP_3 = \tfrac{1}{2}(2P_{BLK,1} + P_{BLK,2} - P_{IM3,(2f_1-f_2)}) \tag{1.5}$$

$$IIP_2 = 2P_{1T} + P_{IM2,(f_2-f_1)} \tag{1.6}$$

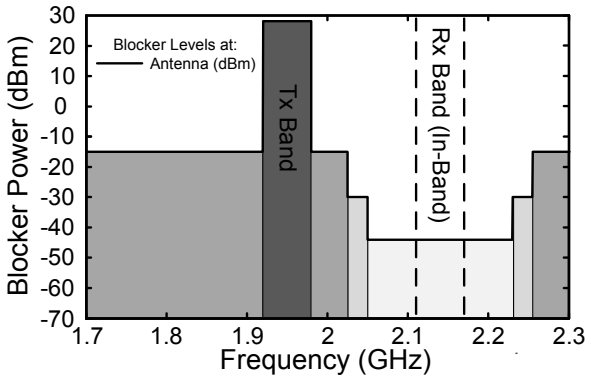


Figure 1.13: Out-of-band blocking specification for UMTS Region 1.

1.1.5 Out-of-Band Blocking

In modern communication schemes such as those described in Section 1.1.2, the standard specifications not only dictate the receiver sensitivity requirements for particular blocker power levels, but also for different blocker frequency offsets from the specified communication frequency band. Typically the blocker power required to be handled at the receiver input increases as the absolute value of the blocker frequency offset. For example, the UMTS Region 1 blocking specification dictates the BER requirement in the bottom entry of Table 1.1 to be met for blocker power levels of -45dBm, -30dBm, and -15dBm at the receiver input for frequency offsets of 0MHz, 65MHz, and 85MHz away from the edge of the receive (RX) communication band as shown in Fig. 1.13. In addition, the frequency-domain-duplexed (FDD) nature of UMTS also places a large implicit blocking specification at a 130MHz-190MHz offset due to the concurrent operation of the mobile station transmitter.

What this implies is that the dominant blocking problem for many important classes of RF receivers is that of out-of-band blocking. Although RF filters can and do mitigate this problem to a certain extent by attenuating the magnitude of such blockers prior to interacting with the active circuitry of the receiver, such filters can be quite expensive both in terms of cost and area footprint.

1.2 Inherent Challenges of Designing Linear RF Circuits

1.2.1 Insufficient On-Chip Filtering

The problem of self-generated distortion error due to undesired signals entering an RF receiver is a considerable problem in maintaining a high SER for a desired signal. In principle, undesired signals can be removed on-chip with RLC bandpass filtering. However, the amount of achievable attenuation is limited by the Q of the on-chip inductors, as depicted in Fig. 1.14. For modern processes, the maximum achievable value of Q is about 10 at 2GHz. For a typical RLC transfer function as (1.7) below, this implies an attenuation of 7dB at a 200MHz frequency offset from the LC center frequency. Compared to the wide dynamic range required by the exemplary sensitivity specifications in Table 1.3, this is a grossly insufficient solution.

Standard	Minimum Desired Signal	Maximum Undesired Signal	Difference	Reference
GSM	-99dBm	+0dBm (max. out of band blocker)	99dB	[4]
cdma2000	-101dBm	+23dBm (TX leakage)	124dB	[5]
UMTS	-114dBm data /-103.7dBm total	+28dBm (TX leakage)	142dB	[6]

Table 1.3: Examples of maximum-to-minimum signal ratios in notable communication standards.

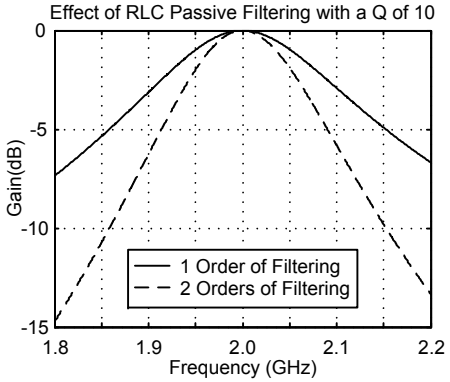


Figure 1.14: Attenuation provided by on-chip RLC filtering.

$$\begin{aligned} Z(f) &= \frac{1}{\frac{1}{R} + j(2\pi fC - \frac{1}{2\pi fL})} \\ |Z(f)|^2 &= \frac{1}{\frac{1}{R^2} + (2\pi fC - \frac{1}{2\pi fL})^2} \end{aligned} \tag{1.7}$$

The solution most frequently employed to circumvent this limitation in the past was the use of off-chip filters with much higher Q. For example, surface acoustic wave (SAW) filters can achieve effective Qs of over 500 [17] with in-band loss of 1-2dB. Bulk acoustic wave (BAW) resonators [18] and film bulk acoustic resonator (FBAR) structures [19] can achieve Q factors of greater than 1000 in the GHz range. However, such filters are expensive and can rival the footprint of the entire integrated RF transceiver. As cellular and other radios converge towards massively multimode solutions, as many as 12 [20] RF front end channels are required, each with their own set of filters. In this case, the off-chip filters can quickly dominate the cost of the entire radio. Therefore, there exists a need to investigate alternative solutions to reducing the impact of large undesired signals aside from simply attenuating them with filters.

1.2.2 Unfavorable Power/Linearity Tradeoff

Power may also be traded off with nonlinearity while keeping the noise of the receiver constant. One way of concretely depicting this tradeoff, shown in Fig. 1.15 for a given nonlinear circuit, is to introduce an attenuation of 3dB prior to the circuit and then place two of the circuits in parallel. In this case, both the net circuit gain (including the attenuation block) and total output noise rise by 3dB, keeping the circuit noise figure (NF) constant. Assuming the unit circuit nonlinearity is memoryless and dictated by $f(x) = \alpha_1 x + \alpha_3 x^3$, the new composite circuit can be modeled as $g(x)$

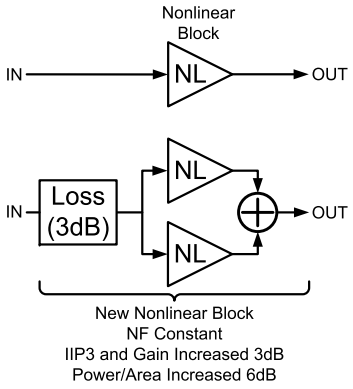


Figure 1.15: Method of trading off power for linearity in receiver circuits while maintaining constant noise figure.

below in (1.8):

$$\begin{aligned} g(x) &= 2f(\tfrac{1}{\sqrt{2}}x) \\ g(x) &= \sqrt{2}\alpha_1 x + \tfrac{1}{\sqrt{2}}\alpha_3 x^3 \\ \tfrac{g(x)}{\sqrt{2}\alpha_1} &= x + \tfrac{1}{2}\tfrac{\alpha_3}{\alpha_1}x^3 \end{aligned} \tag{1.8}$$

It can be seen that the total input-referred IM3 error is reduced by 6dB and IIP3 is improved by 3dB for a doubling of circuit power. To provide an example as to why this solution is even less palatable than adding an external SAW filter prior to the block, consider the results from the receiver in [21] in which a SAW filter was used prior to the downconverter (pre-amplifier plus mixer) to obtain an overall IIP3 of +1.2dBm, which was dominated by the LNA. As the LNA obtains a gain of 14.2dB and the downconverter obtains an IIP3 of -0.7dBm, the overall receiver IIP3 would be approximately -15 dBm without the filter. To re-establish the receiver IIP3 of +1.2dBm requires the downconverter current of 14.4mA to increase by a factor of over $2^5 = 32$ to over 460mA, which is an unacceptable solution for a mobile application.

1.3 Approaches and Challenges to Improving Nonlinearity in RF Circuits

According to Lee [22, pg. 423], there exist four broad classes of techniques to improve the effective linearity of any circuit-based system. It is worthwhile to visit each of these in turn to examine its feasibility for the improvement of the linearity of an RF receiver.

1.3.1 Negative Feedback

The principle of feedback in the context of circuit design is usually attributed to Harold Black of Bell Laboratories in 1927. Black came up with the idea as a possible solution to reducing distortion in repeater amplifiers used for telephone transmission [23]. The essential concept of feedback, as

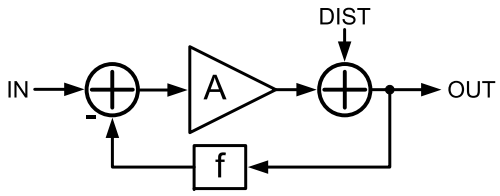


Figure 1.16: Basic feedback concept.

depicted in Fig. 1.16, is to take the output of a nonideal system (usually an amplifier) with open-loop gain A, reduce the output by a scalar factor f, and subtract the resultant signal from the input. The resultant closed-loop system then has a gain of (1.9).

$$\frac{OUT}{IN} = \frac{A}{1+Af} \overset{Af\to\infty}{\longrightarrow} \frac{1}{f} \tag{1.9}$$

And output error sensitivity of (1.10).

$$\frac{OUT}{DIST} = \frac{1}{1+Af} \overset{Af\to\infty}{\longrightarrow} 0 \tag{1.10}$$

Clearly, feedback becomes more effective as A is increased for constant f. As the feedback network f can be made up of inherently linear passive elements, the system input/output relationship can approach perfectly linear behavior as A increases, no matter how nonlinear the nonideal system is.

The initial difficulty Black faced in gaining acceptance for his invention hinged upon the fact that feedback trades off active device gain for linearization (or any error reduction, for that matter). At the time, the principal active device used was the vacuum tube, which had a peak dc gain of about 5 [22, pg. 442]. Sacrificing this already low gain for a moderate improvement in distortion would have undoubtedly negated the purpose of many amplifiers of this time. Developments in vacuum tube technology, and later bipolar and CMOS transistors, soon made large values of dc gain possible. Consequently, feedback became a ubiquitous feature in low-frequency (baseband) circuit design.

However, like the vacuum tubes of yesteryear, transistor gain is scarce at RF frequencies due to the limitation of the drain impedance imposed by parasitic capacitances. These capacitances can be tuned out by on-chip inductors, but the peak gain is still limited by the finite Q of the inductor. To get a quantitative sense of the limitations here, the reader may consider a common-source cascoded LNA with transconductance (g_m) of 60mS and total parasitic output capacitance of 0.5pF as shown in Fig. 1.17. For a 2GHz amplifier, a 12.6nH inductor is required to tune out this capacitance. For a Q of 9 (possible in a thick metal process), the equivalent output impedance is 1.4kΩ. This results in a peak voltage gain of 86 for the amplifier. Simulations in a 90nm process suggest that this peak voltage gain will be limited to about 50, or 34dB, when accounting for the active device channel conductance. To obtain a closed loop dc voltage gain of 15dB[2], only 20dB of open-loop gain may be applied as feedback to reduce 3rd-order output IMD. Compared to the 60-80dB of open-loop gain commonly achievable by baseband operational amplifiers, this 20dB of gain stands out as a small number. Even worse, this analysis neglects the increase in nonlinearity due to increased feedback of second order IMD from the output to the input to the LNA. These products then mix at the

[2]A typical value of LNA voltage gain.

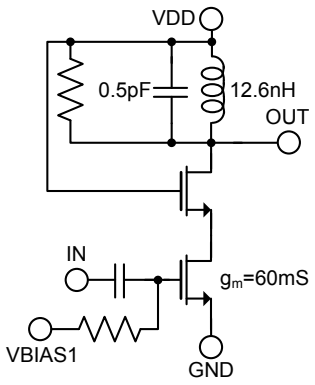


Figure 1.17: Basic common-source RF amplifier schematic.

input of the LNA with the original blocker signals with the very strong 2nd order nonlinearity of the CMOS device, generating additional IM3 products [24]. This analysis also neglects the group delay associated with such a high-Q tank, which is equal to $2Q/w_0$, or about 1ns, for a center frequency w_0 of 2π(2e9)radians/s. The problem with this is that the group delay is twice as long as the desired signal carrier frequency period and is not well-controlled over process variations. Therefore, the feedback will not effectively cancel distortion products and may even lead to instability.

It can be further argued that the above case of linearizing a voltage-domain RF LNA via feedback is the easiest conceivable. In traditional receiver architectures, the mixer constitutes the dominant linearity burden in the system because the input blockers have experienced gain due to the LNA but have not yet been attenuated by baseband filters. However, feedback in a mixer requires another frequency upconversion after downconversion to baseband. Downconversion to baseband typically entails an unavoidable passing of the current through a large time constant and consequent group delay of at least several tens of nanoseconds. Re-upconversion to RF and cancellation after such a large time constant prevents effective feedback. The exact same problem prevents the effective use of feedback in passive-mixer based systems that have become more prevalent in recent years.

1.3.2 Predistortion

The concept of predistortion is to cascade two inverse nonlinear functions in series such that the composite function is nearly linear. Successful application of this technique, depicted in Fig. 1.18, relies on careful matching. In power amplifier linearization systems, the matching problem is often alleviated with adaptive feedback in which the power amplifier output is downconverted and compared with the original non-distorted signal. This information can be used to adjust the weights of an adaptive look-up table [25], or a polynomial predistorter [26]. Implementing such blocks in RF receivers proves difficult. Adaptive look-up tables are digital baseband circuits by nature and cannot be easily implemented in the RF portion of the receiver. Polynomial predistorters could in principle be used but due to the presence of large amounts of active circuitry would seriously degrade the NF of the receiver.

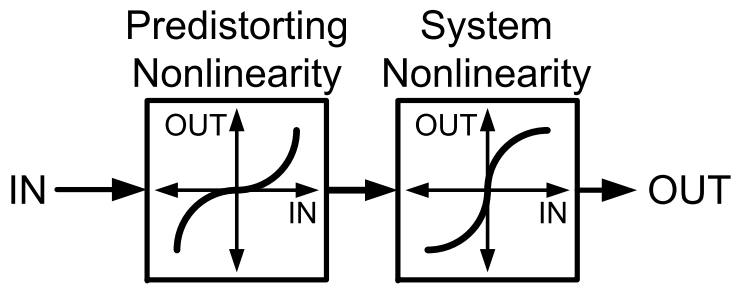


Figure 1.18: Basic predistortion concept.

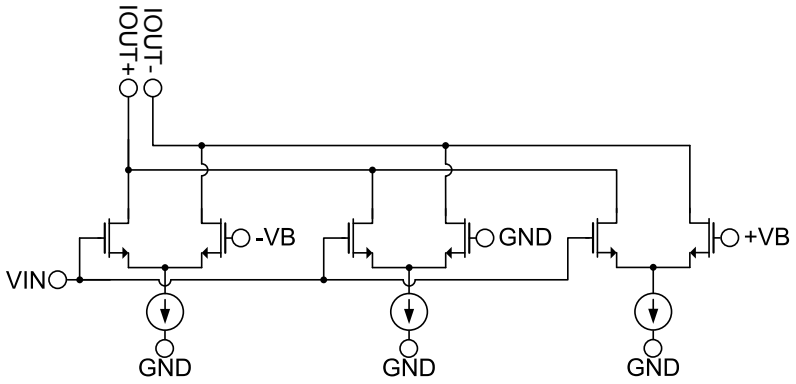


Figure 1.19: Basic piecewise linearization concept.

1.3.3 Piecewise Linearization

Piecewise linearization relies on the fact that many types of circuits are approximately linear over a significant voltage range. If this voltage range can be shifted by an offset, several of these circuits can be placed in parallel to effectively extend this linear range. The canonical example of this concept is the Gilbert multi-tanh circuit [27] shown in Figs. 1.19 and 1.20. The tradeoff here is an increase in power dissipation and input capacitance [22, pg. 427] for each additional branch added to the circuit.

1.3.4 Feedforward Cancellation

Like feedback, feedforward cancellation can be used to substantially reduce many forms of self-generated error arising from the interaction of desired and undesired signals with circuit nonidealities. However, feedforward cancellation is potentially more suited to improving the performance of RF receivers because it avoids the gain loss and group delay issues inherent in feedback. As shown in Fig. 1.21, a feedforward loop employs a rough model of a circuit nonideality to recreate an approximation to the self-generated error. This self-generated error can then be subtracted from the main receiver path. Like predistortion, however, feedforward cancellation is extremely sensitive to a wide array of mismatches that can occur within a practical system, which ultimately limit its performance. For example, to attenuate self-generated error by 40dB, components must match in both gain and group delay to about 1%, which is challenging to achieve with even modern components.

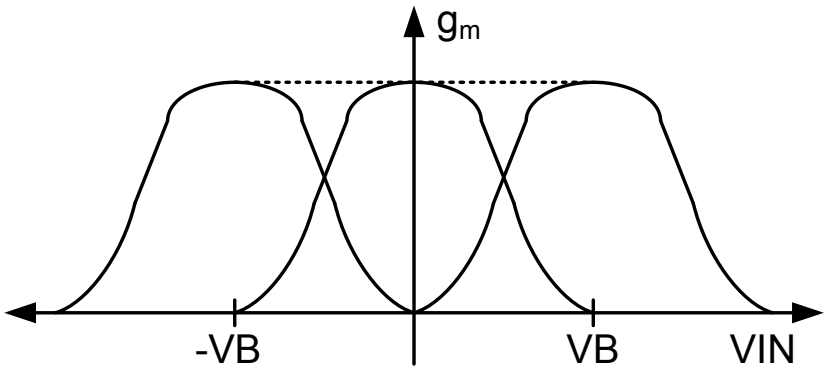


Figure 1.20: Piecewise linearization as implemented in a Gilbert multi-tanh circuit.

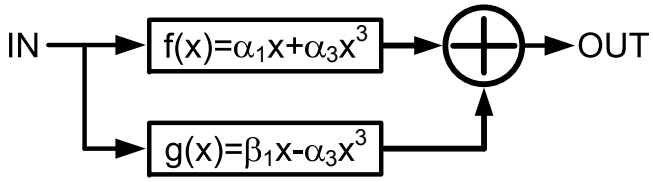


Figure 1.21: Basic feedforward concept.

One common class of local feedforward cancellation of IMD products in RF receivers [24], shown in Fig. 1.22, exploits the fact that the third-order Taylor series coefficient of the MOS device undergoes a sign change as the gate bias voltage is reduced. Also as the gate bias voltage is reduced, the device enters the weak inversion region and the linear gain term is substantially reduced in magnitude. When a device biased in this fashion is placed in parallel with a more typically biased common source or inductively source degenerated MOS device in an LNA, any third-order products are effectively cancelled at the output while only minimally reducing the LNA power gain. This solution has two inherent problems. The first of these is that the amount of improvement varies substantially over process variation and temperature due to the fact that precise cancellation depends on two exact biasing conditions. The second of these is that the fifth-order Taylor series coefficient of the net structure is increased, as can be seen in the swept measurement results of Fig. 10 in [24], which limits the effective IMD improvement to smaller input signal levels. This limitation explains the frequent usage of such techniques in LNAs but almost never in mixers, where the large undesired blocking signals are 10-15 dB larger than those seen at the LNA input.

1.4 Adaptive Feedforward Cancellation of Self-Generated Error

In order to solve the process variation and mismatch problems that limit the performance of feedforward cancellation, adaptive equalizers may be used to perform the subtraction of the self-generated error model signal from the main receiver path, as shown in Fig. 1.23. By feeding back the

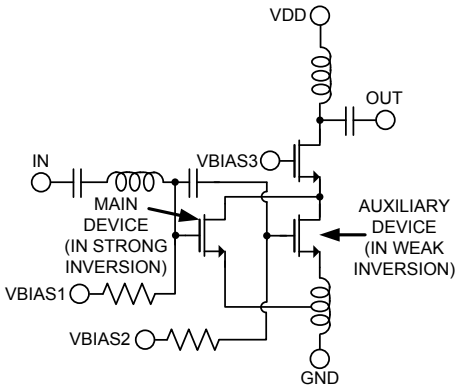


Figure 1.22: Simplified schematic of modified derivative superposition linearity-enhanced LNA.

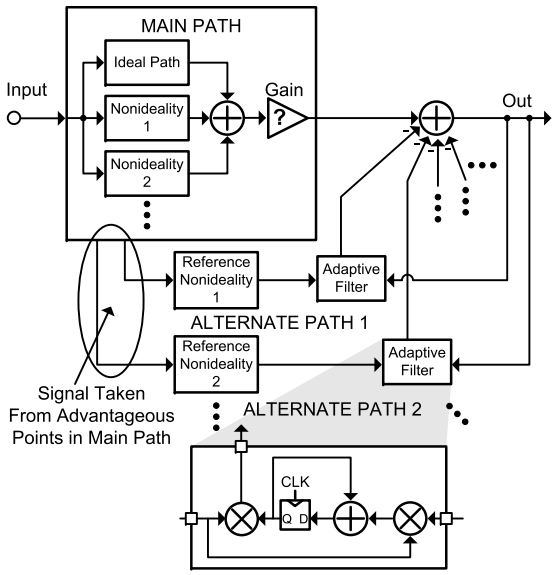


Figure 1.23: General adaptive feedforward error cancellation concept.

residual output from this subtraction process to the adaptive equalization algorithm, the algorithm will converge so as to minimize the amount of error present in the output while leaving the desired signal relatively unaltered. In recent years, this concept has been proposed for use in RF receivers to remove IM2 distortion and dc offset [28], IM3 distortion [29] [30], signal interference resulting from quadrature I-Q mismatch in IF-receiver circuitry [31] [32], and was described as a general concept in [32]. Other examples of the use of adaptive signal processing in receivers include the cancellation of error-producing signals themselves [33] and additional techniques to remove the effects of dc offset and I-Q mismatch in direct-conversion receivers [34] [35].

Although adaptive feedforward error cancellation may in principle be used to locally correct for the error generated by a particular block, cancellation of the error at the end of the physical portion of the receive chain is also possible and makes more sense for two reasons. First, all of a particular class of error may be cancelled at once. For example, many different blocks in a receiver contribute

IM3 distortion error that is qualitatively the same. Cancelling this error at multiple points or at a single point has the same effect, but clearly the latter is the simpler solution. Second, canonical adaptive equalizers can be implemented in digital baseband circuitry without the designer having to account for the effects of process variation and temperature changes in the operation of the equalizer itself.

Such a system-level solution methodology in turn begets even more advantages, not the least of which is its indifference towards the details of the original receiver. For example, the receiver LNA topology may be changed from an inductively degenerated common source topology to a common gate topology and the adaptive feedforward error cancellation scheme need not change at all to have qualitatively the same effect. Furthermore, as this technique focuses only on cancellation of the error itself and not the error producers, it permits a relatively narrowband alternate path through which the recreated error signals pass. That is, although large, undesired, error-producing signals may occur over a wide frequency range (100-500MHz), the relevant error terms that they produce are only restricted to a much smaller bandwidth equal to that of the desired signal (1-5MHz). This implies that the alternate path circuits can be constructed using traditional RF circuit design techniques, as opposed to more recent extremely-wideband architectures used for software-defined radios [36]. The circuits processing the error can operate with a lower dynamic range than their original receiver counterparts, as the corruptive error to noise ratio is in general much less than the error producer to noise ratio in the main receiver path. Also, if the LMS algorithm is used to perform the adaptive equalization, multiple LMS loops in parallel can be used to cancel different error signals as shown in Fig. 1.23 [28]. Furthermore, in the event that the self-generated error of interest is due to large, undesired blocker signals, the power dissipation of an alternate path can be further reduced by only powering it on when needed, as most receivers only need to operate in the presence of strong interferers for a small fraction of the time. Finally, the adaptive nature of the equalizer permits tracking of rapidly changing blocker conditions.

The indifference of the adaptive feedforward error cancellation technique to the details of the original receiver can also be exploited to simplify the design of multimode receiver terminals intended to work with communication standards dominated by different sources of error. For example, a frequency-band-adjustable downconverter could be designed using simple canonical block architectures. If one particular standard then requires a much higher IIP2 than the base receiver can provide, an IM2-error canceling alternate path can be enabled. Similarly, if high IIP3 is required, an IM3-error canceling alternate path can be turned on.

1.4.1 LMS as Adaptive Equalization Algorithm

Least mean squares (LMS)-based adaptive equalizers are common choices in power-constrained applications due to their simplicity and robustness [37, pp. 231, 297]. The concept behind LMS-based adaptive equalization is depicted in Fig. 1.24. Here, a time-varying finite impulse response (FIR) filter is utilized to modify a reference signal and to subtract it from an incoming signal corrupted by a version of the same reference signal. The taps of the FIR filter are adjusted based on the instantaneous correlation estimate between the equalizer output and the reference signal. That is, if there is any signal correlated with the reference signal in the equalizer output, each tap is adjusted on average in a direction so as to reduce the reference signal content in the output. If

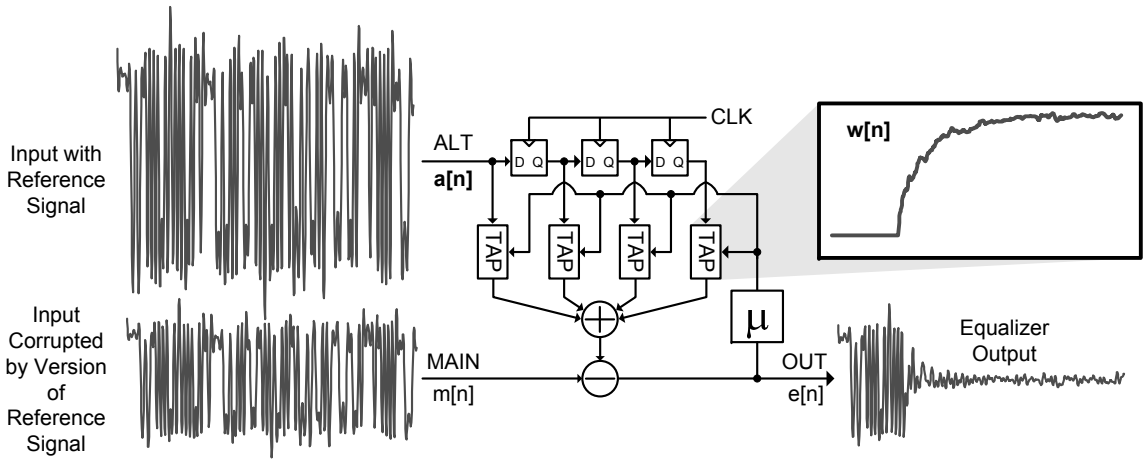


Figure 1.24: Least mean squares (LMS) equalization concept.

the equalizer is designed properly, the filter taps will converge close to a solution that yields the minimum mean squared error at the output, but will exhibit a small excess deviation around that solution. In the case of Fig. 1.23, single-tap LMS adaptive filters are used as part of an exemplary adaptive equalizer which removes multiple types of error.

1.5 Prior Art in Adaptive Feedforward Cancellation of Nonlinear Distortion

In [28], an adaptive error cancellation scheme using postdistortion to generate a reference error signal for IM2 cancellation was proposed and implemented using discrete components. As shown in Figs. 2 and 3 of [28], a squaring circuit was connected to the output of the direct-conversion downconversion mixer. In this case, all downconverted blocker signals are squared along with any other desired signal. Because the desired signal is much smaller than any of the downconverted blockers, its square is much much smaller than those of the downconverted blockers and can be assumed to be negligible in the cancellation process. This scheme relies on the fact that the IM2 products produced by the blockers after downconversion are the same as those produced by the blockers prior to downconversion. However, this scheme does not recognize the inherent difficulty in maintaining the large signal bandwidth after downconversion required to maintain large signal levels for the downconverted out-of-band blockers that dominate blocking problems such as those described in Section 1.1.5. In [28] it is assumed that the downconverted blocker signals can be processed by a squaring circuit prior to the application of any filtering. However, it is common practice to embed the first real pole of the analog baseband filter in the current-to-voltage conversion immediately following the mixer Gilbert cell switching pair, such as in [5] [21] [38] [39] [40] [41] [42] [43], where the first-order RC time constants ranged from 0.1 to 2 MHz. One depiction of this truism is shown in Fig. 1.25. Receiver architectures using a passive mixer scheme as shown in Fig. 1.26 universally embed a low-pass filter in the transimpedance amplifier that follows the mixer to perform the baseband current-to-voltage conversion. Neglecting the effect of this filtering may be acceptable for the 6MHz

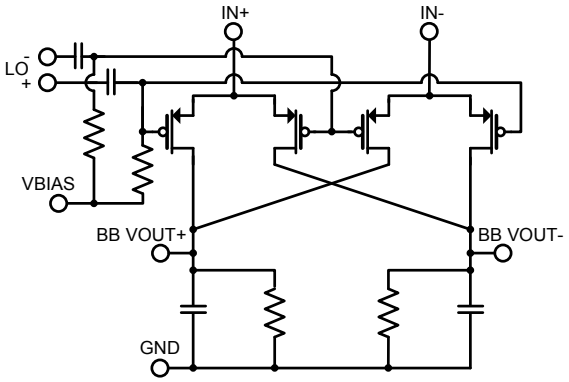


Figure 1.25: Problem with baseband nonlinear term generation - inherent low pass filtering at mixer output.

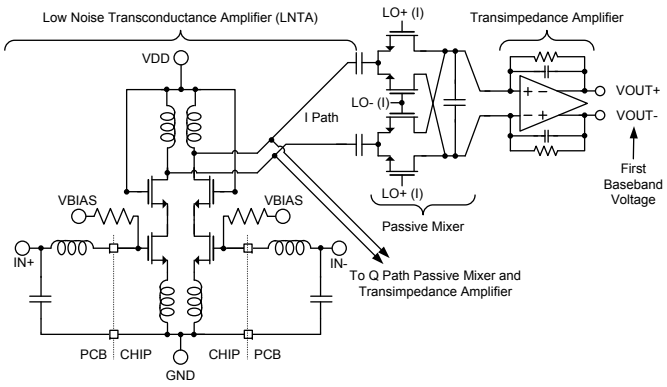


Figure 1.26: Basic passive mixer-based receiver front end architecture, showing inherent low-pass filtering prior to first baseband voltage.

blocker offset used in [28] but, for example, is not for the 250MHz maximum frequency offset between the out-of-band TX leakage blocker and desired RX signal in a UMTS Region 1 receiver. In the latter case, the downconverted TX leakage baseband voltage signal will be attenuated by over 40dB with respect to the desired signal. This means that the regenerated IM2 products will be 80dB lower than they would have been at a low frequency offset and would likely be lower in magnitude than the noise floor of a low-power squaring circuit.

The digital algorithm utilized in [28] is also notable in that it cancels two types of distortion simultaneously. If one considers dc offset to be 0th order distortion, the algorithm removes it by using a digital value of 1 as the reference input to an LMS tap which is placed in parallel with the LMS taps utilized to adaptively cancel IM2 products. It can be shown (Appendix A) that the operation of this LMS tap is the same as a 1st-order digital high-pass filter with the LMS convergence constant μ setting the corner frequency of the filter.

In [29] and [30], postdistortion is used in the digital domain after analog-to-digital conversion to generate reference error signals for adaptive IM3 cancellation. The scheme described in [30] is summarized in Fig. 4 and in the first paragraph of section IV in [30]. Although the complete

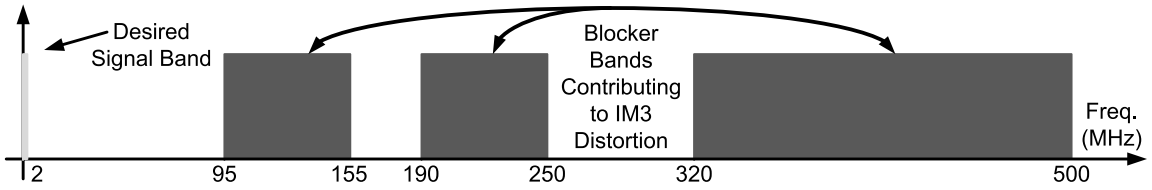


Figure 1.27: Signal bands required for digitization of IM3-producing blockers in UMTS Region 1 (after downconversion to baseband) in a purely digital adaptive feedforward distortion cancellation scheme.

schematic of the receiver discussed in [30] is never shown, Fig. 4 of [30] shows an "I or Q" input prior to a nonlinear component, implying that this nonlinear component exists after downconversion to baseband. The text description at the beginning of section IV of [30] places the analog-to-digital interface between the nonlinear component and band-split filtering of Fig. 4. In the case of either [29] or [30], the issues with generating reference IM3 products at baseband for an out-of-band blocking scenario hold as in [28] except in this case the problem is worse due to the fact that regenerated IM3 products will be nearly 120dB lower than they would have been at a low frequency offset due to the 3rd-order nature of the distortion. Furthermore, for the architectures in [29] and [30] to cancel IM3 products, the ADCs must pass the full spectrum of potential problematic blockers. For a modern communication standard such as UMTS Region 1, this means the frequency bands of 1670MHz-1850MHz, 1920MHz-1980MHz, and 2015-2075MHz must be digitized, as depicted in Fig. 1.27, along with the desired signal bandwidth for comparison. Achieving this would require at least 6 Nyquist ADCs with sampling rates in excess of 60MHz. The ADC outputs would also need to be recombined in such a way that the digitized blockers interact nonlinearly as they would in the continuous time domain. Furthermore, since the frequency bands surrounding those listed are equally likely to have blockers not associated with the generation of IM3 products, the anti-aliasing filters prior to the ADCs would have to be extremely sharp, necessitating a very high filter order, and hence area and power consumption.

1.6 Motivation and Challenges of Designing Very Large-Signal Handling RF Receivers

The compression metrics of the receiver denote the maximum signal levels that can be reached before the effective gain of the receiver drops by 1dB. Table 1.4 shows a representative sample of the P_{-1dB} quantities of recent receiver LNAs and Table 1.5 of the various reported values for compression and desensitization for receiver front ends. The P_{-1dB} plot is commonly provided for LNAs in the literature, with an example provided by Fig. 20 in [44]. What this plot typically shows is that as the input signal power is increased past P_{-1dB}, the circuit small-signal gain drops off very quickly and the output signal power levels off or even drops. These latter characteristics indicate the presence of nonlinear terms of order much greater than 3. In fact, one can consider a signal large enough to move the active amplifying devices into the triode region such that the effective system nonlinear function transitions from a simple 3rd-order nonlinearity to one with a strong compressive

Reference	P_{-1dB} (dBm)	Noise Figure (dB)	Supply Voltage (Vdd)
[45]	-7.0	2.9	1.8
[46]	0.0	3.9	1.2/2.5
[47]	-8.4	1.7	3.0
[48]	-12.0	2.6	1.5
[49]	-18.0	3.6	1.2

Table 1.4: Examples of reported LNA P_{-1dB} performance in the literature.

Reference	1-dB Point (dBm)	1-dB Point Type	Noise Figure (dB)	Supply Voltage (Vdd)
[9]	-13	Unknown	3	3.0
[21]	-7.8	Unknown	5.3	3.0
[13]	-10, 200MHz Blocker Offset	Desensitization	4	1.2
[13]	-6, 400MHz Blocker Offset	Desensitization	4	1.2
[14]	-7.0	Desensitization	5.5	1.2/2.5
[50]	-20	Unknown	6	1.2
[51]	-14	Unknown	3.5	1.2

Table 1.5: Examples of reported receiver 1-dB compression metrics in the literature.

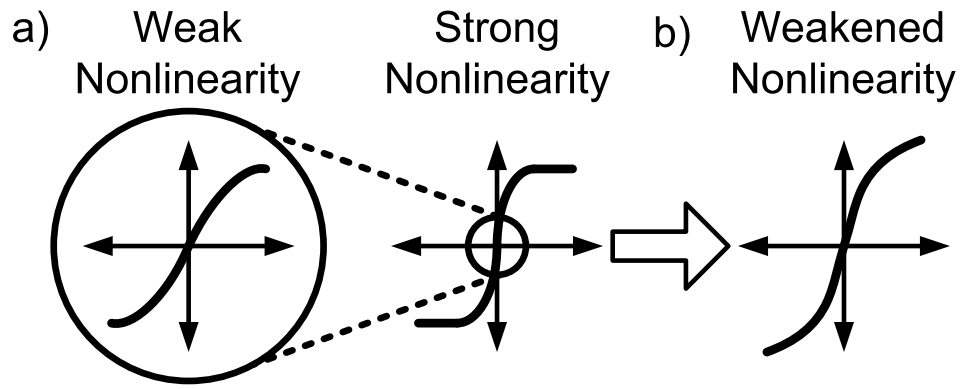


Figure 1.28: a) Depiction of expansion of nonlinear input-output characteristic to one exhibiting strong nonlinearity. b) Weakening of input-output characteristic nonlinearity.

tanh-like nonlinearity, as depicted in Fig. 1.28a.

What this information conveys is that the performance of receivers using adaptive feedforward nonlinear distortion cancellation to extend the system dynamic range will eventually be limited by higher-order intermodulation distortion products and compressive effects (themselves a result of higher-order nonlinear distortion terms). These two problems need to be tackled independently. Compressive effects can reduce the gain of the desired signal to the point where the signal-to-noise ratio is below that required for proper demodulation and decoding. At this point, recovering the desired signal in digital baseband becomes difficult, if not impossible. Therefore, weakening the nonlinearity as in Fig. 1.28b takes precedence as a key design goal. Even after doing so, however, the nonlinear input-output characteristic must be modeled with many higher-order polynomial terms as implied by the kinks in the input-output characteristic depicted in Fig. 1.28b. The discussion in the previous section makes plain that higher-order IMD reference terms cannot be generated strictly in the digital domain due to the large digitization bandwidths required to retain information related to the out-of-band blockers. Hence, novel solutions to cancel higher-order IMD reference terms must

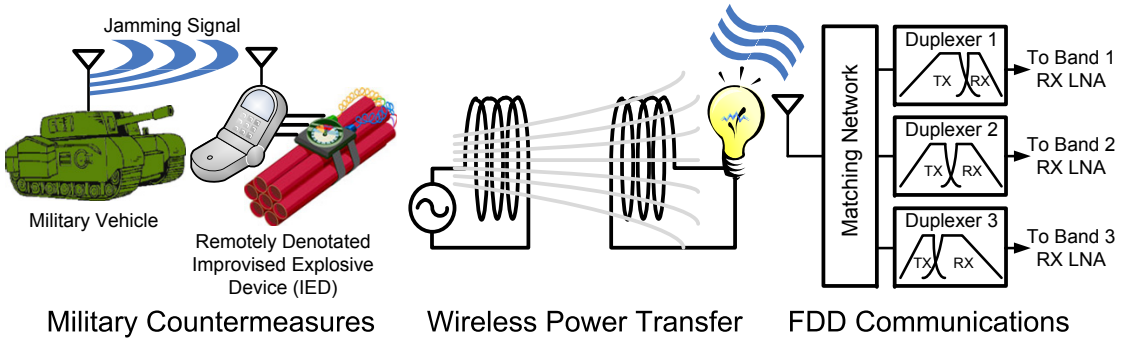


Figure 1.29: Motivations for large-signal handling receivers.

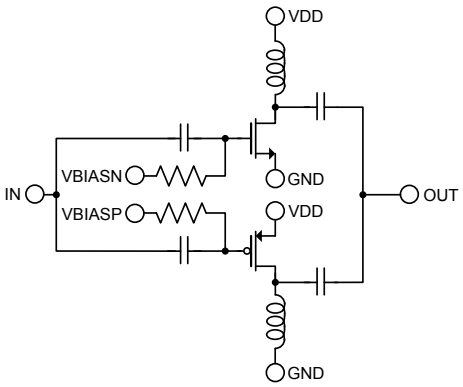


Figure 1.30: Dual NMOS/PMOS LNA as depicted in [53].

also be developed in conjunction with analog solutions to reduce the compressive effects found within the receiver.

While the aforementioned compression performance in Table 1.5 may suffice for contemporary applications, the signal-handling demands on radio receivers are due to increase dramatically in the coming years. As depicted in Fig. 1.29, radar systems and military communications systems need to be able to operate in the presence of very large intentional jamming signals that are used to prevent remote IED detonation and enemy communications. Co-location issues in portable combined jamming/communications radios result in the appearance of very large signals at the receiver input. Wireless power transfer is making increasingly frequent appearances, both in powering gadgets from across a room and in transferring energy to implantable medical sensors and drug delivery devices [52]. In each of these cases, data communications must co-exist with a very large out-of-band undesired blocker signal. Finally, frequency-domain-duplexed (FDD) receivers are under increasing pressure to relieve some of the requirements on up-front duplexers. Because multimode receivers now may be required to handle a dozen or more cellular standards, and because each standard requires its own duplexer, the requirements of these duplexers need to be reduced in order to lower the cost and area burden to the radio. The upshot of doing this is that a much larger TX leakage signal appears at the receiver input. Fortunately, recent milestones hit by CMOS scaling present some hope in overcoming the limited signal-handling capability of receivers in the prior art.

Since the beginning of the widespread use of CMOS in the early 1970s, improvements in the process lithography and oxide growth control have permitted a steady decrease in minimum achievable device dimension and parasitic capacitance. The most notable consequence of this improvement has been termed Moore's law, in which the number of transistors that can be placed inexpensively on an integrated circuit die has doubled about every two years [54]. A corollary to this improvement has been the increase in maximum operating frequency of MOS transistors enabled by the reduction in parasitic capacitance that accompanies finer process lithographies. Although the results of scaling may be most commercially conspicuous in the constant increase in digital computing performance, scaling has also consistently made possible new applications and fields in the analog and RF domains. For example, digitally-assisted analog circuits such as [55] [56] began to make sense when the cost and size of a single transistor became negligible with respect to the total circuit die cost. In this case, massive digital algorithms were justifiable to correct for nonidealities in a few critical analog transistors. In the more recent past, a similar trend took place in domain of RF circuits. To see how, the reader may consider that a rough measure of the maximum frequency at which current gain can be obtained from a transistor is given by (1.11).

$$f_T = \frac{g_m}{2\pi(C_{gs} + C_{gd})} \tag{1.11}$$

Making the common assumption that C_{gd} is negligible and substituting into the formula the physical expressions from a square-law MOSFET I-V characteristic yields:

$$f_T \approx \frac{g_m}{2\pi C_{gs}} \approx \frac{\mu_n C_{ox}(W/L)(V_{GS} - V_{TH})}{\frac{2}{3}WLC_{ox}} = \frac{3}{2}\frac{\mu_n(V_{GS} - V_{TH})}{L^2} \tag{1.12}$$

As minimum device lithography shrunk, lowering the minimum L, the device f_T moved into the 100GHz range, making possible the implementation of RF CMOS circuits in the GHz frequency bands that had been allotted for personal mobile radio communications by various national governments. As NMOS devices have a higher electron mobility than PMOS devices ($\mu_n \approx 3\mu_p$), the first RF CMOS circuit topologies were by necessity all-NMOS. As scaling continued, PMOS transistors in turn also became viable RF devices [53], allowing for far broader range of circuit topologies such as current-reuse common-source LNAs [13], and push-pull circuit topologies capable of linear behavior over a wider input signal range than either NMOS or PMOS alone [53] such as the dual common-source topology shown in Fig. 1.30. Scaling has also enabled the practical use of passive mixer topologies such as the one shown in Fig. 1.26, which have become increasingly common due to the reduced power required to fully transition a MOS device gate voltage from ground to supply.

Chapter 2

Mixed-Signal / Mixed-Domain Equalization of IM3 Products in RF Receivers[1]

2.1 Target Application and Context of Project

In RF receivers, perhaps the most significant contemporaneous self-generated error problem is nonlinear distortion, due to the continued and rising popularity of the UMTS FDD standard [60] for 3G communications. UMTS and similar FDD standards typically possess stringent linearity requirements due to the necessity of having to concurrently handle a large, unwanted blocker signal in the presence of TX leakage through the frequency-domain duplexer, as shown in Fig. 2.1. This situation, defined in the UMTS standard specifications [6], sets up an implicit two-tone test which can yield intermodulation distortion products that corrupt the desired signal. As depicted in Fig. 2.2a, in a typical direct-conversion receiver that uses a voltage-domain LNA output, if standard block design values are used for the LNA and mixer with no additional enhancements, the worst-case IM3 products overwhelm the desired signal at RF. After downconversion, the desired signal is still corrupted. In order to meet the receiver linearity requirements, several reported commercial receivers [9] [21] [61] have resorted to the use of interstage SAW filters to attenuate large blocker signals and hence to relax requirements on the integrated circuit blocks, as depicted in Fig. 2.2b. In this case, the large blocker signals are attenuated prior to the mixer, which dominates the linearity performance of the receiver. As a result, the signal-to-error ratio of the downconverted desired signal is still sufficient for proper decoding.

Several SAW-less UMTS receivers have also been reported, albeit with somewhat lower out of band IIP3 performance [43] [62] [63] than those previously referenced. It should be noted that reported IIP3 specifications and achievements vary widely within the literature, as shown in Table 2.1. This variation is due to the fact that the IIP3 specification depends on the receiver achieved

[1]Portions of this material have been previously published in [57],[58],[1], [59], and [2].

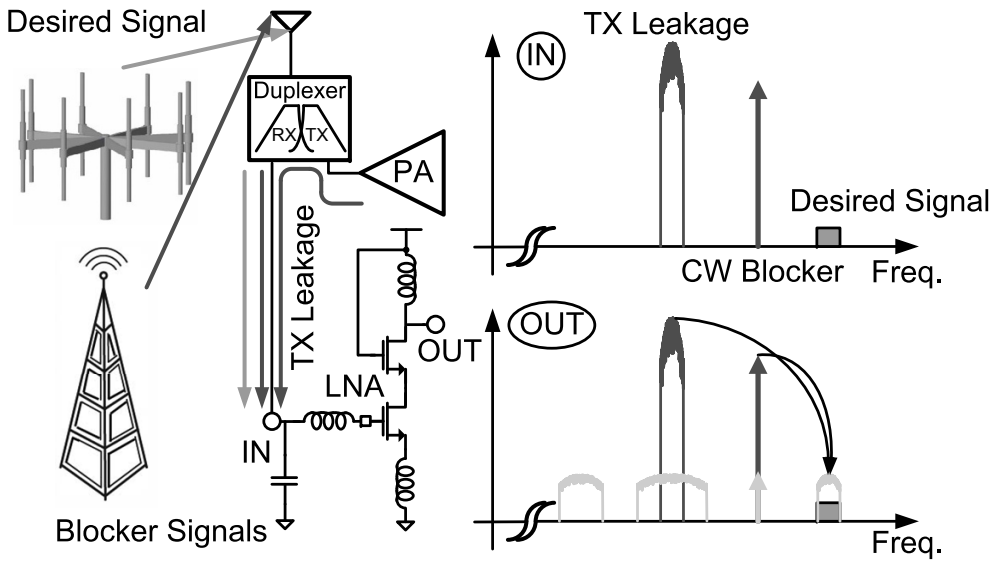


Figure 2.1: Blocking problem in FDD receivers.

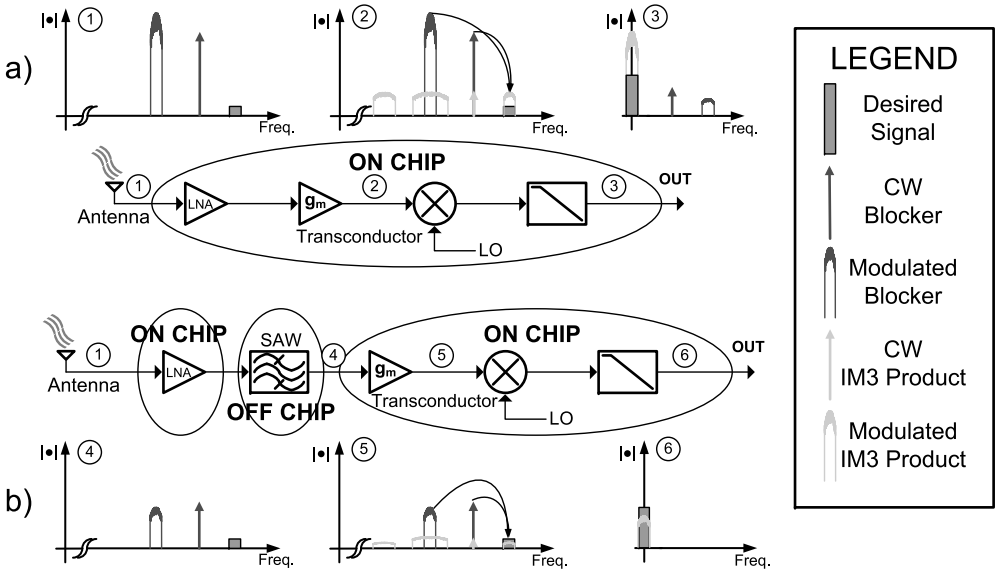


Figure 2.2: a) Nonlinearity problem in typical integrated receivers. b) Use of a SAW filter to alleviate nonlinearity problem.

noise figure, the peak TX power which must be handled, and the characteristics of the particular duplexer used, as discussed in Section 2.4. Although the SAW-less receivers mentioned above meet their respective self-imposed out-of-band IIP3 specifications, it is worthwhile to consider the design of SAW-less receivers with still higher IIP3 in order to reduce overall receiver power dissipation and to permit the use of less stringent, and possibly also less expensive, duplexer blocks.

Reference	IIP3 Specification	IIP3 Performance
[9]	+1.3dBm	+1.6dBm
[61][a]	-6.1dBm	-0.5dBm
[43]	-10dBm	-7.4dBm
[62]	-8dBm	-7dBm
[63]	-5dBm	-2dBm

Table 2.1: Reported IIP3 specification and performance comparison.

[a]Measurement at mixer transconductor input.

2.2 General Concept of Proposed Solution Architecture

The general concept behind the adaptive feedforward IM3 cancellation scheme introduced in this work is shown in Fig. 2.3. In this figure, a nonlinear main receiver path is subject to two large blockers such that the desired signal is overwhelmed by the self-generated IM3 interference products. The figure reflects the fact that the UMTS out-of-band blocking requirement [6] is an implicit two-tone test with one of the "tones" being the modulated TX output leakage through the antenna duplexer. Reference IM3 products are generated at RF by a cubic term generator, while choosing the LO frequency of the alternate feedforward path mixer to be the same as that of the main path guarantees that the proper set of IM3 products are downconverted to baseband frequencies. In this fashion, a reference IM3 signal can be generated for any specified RF blocker that can produce significant IM3 distortion interference in the receiver, regardless of the blocker frequency offset from the RX LO frequency.

Baseband postfiltering attenuates to negligible levels unwanted IM3 products in the alternate feedforward path such that the adaptive equalizer converges based only on the statistics of the IM3 products corrupting the desired signal at baseband. In this fashion, the equalized IM3 products of the alternate feedforward path can be directly subtracted from the main path, leaving only the desired RX signal. Note that this technique is not limited to a two-tone test but also removes IM3 products resulting from a three-tone test. Furthermore, this technique does not require prior knowledge of one or more of the blocker frequencies, as do techniques that rely on the cancellation of TX leakage to meet the UMTS linearity specifications [33] [64]. Hence, this technique is applicable to a far broader class of radio problems than the aforementioned references. Although both IM2 and IM3 products can be concurrently cancelled by parallel LMS adaptive filters, it was decided in this project to leverage recent work [41] [65] in order to perform local IIP2 improvements so as to avoid adding additional ADCs to the system to pass reference IM2 products.

2.2.1 Choice of Feedforward Loop Overcoming Limitations of Prior Art

The adaptive feedforward error cancellation technique described in this paper differs significantly from those presented in [29] and [30] in that reference IM3 products are generated at analog RF rather than at digital baseband. That is, the alternate feedforward path loop is a mixed-mode system that stretches back to almost equal the length of the main receiver path. It can be shown with trigonometric identities that the resultant downconverted IM3 products are the same as those that would have been produced by I and Q baseband cubing circuits. This architectural characteristic

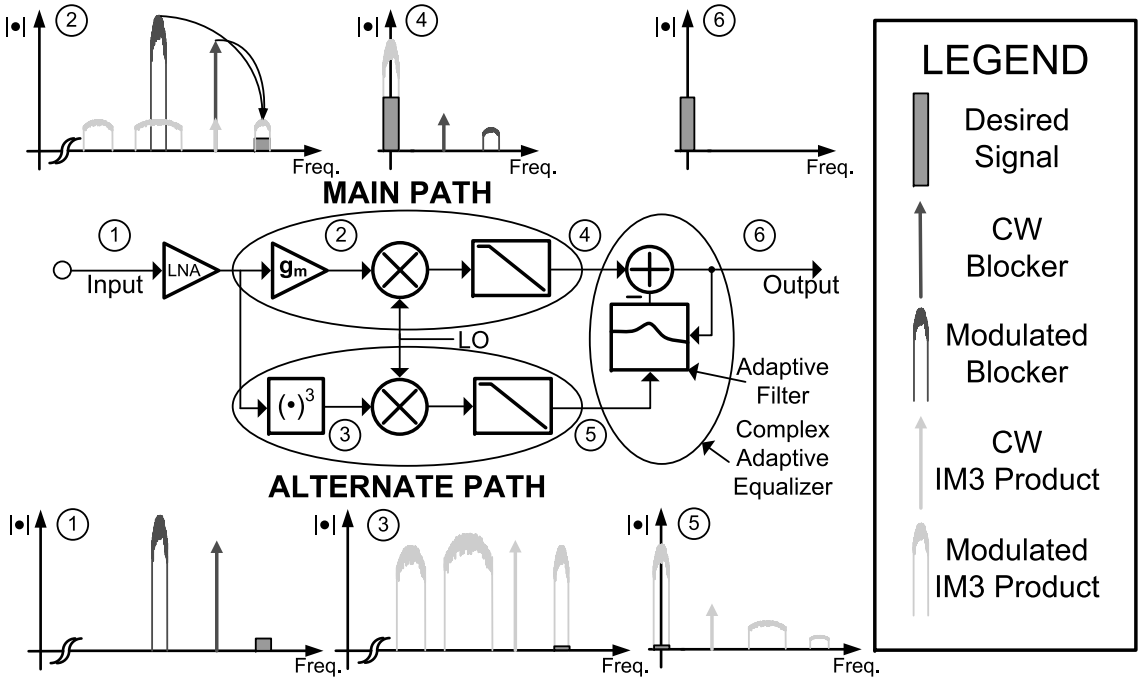


Figure 2.3: Adaptive feedforward IM3 distortion cancellation - proposed receiver system.

overcomes the limitations of the prior art discussed in Section 1.5 with regards to out-of-band blockers in that at analog RF, the blockers have not yet been attenuated by inherent postfiltering after downconversion. IM3 generation in the continuous-time (analog) domain as performed in this architecture is the preferred method in an integrated downconversion receiver since it permits the use of relatively narrowband mixers, postfilters and ADCs such as those based on sigma-delta modulators in the alternate path. Given that IM3 generation is therefore to be performed in the analog domain, it should additionally be performed at RF after the LNA, where the blocker voltage magnitudes are at their largest point in the receiver. Doing so minimizes the required power dissipation of the cubic term generator, as a higher noise floor can then be tolerated.

2.2.2 Fixed and Adaptive Equalization

It can turn out in the design of the alternate feedforward path system that the linear time-invariant (LTI) baseband path difference[2] is known to a large degree of certainty. For example, tuning schemes designed to track RC [66] or g_m-C [67] filter variations are commonly employed to keep time constant variation to only a few percent over process and temperature variation. In cases such as this, the baseband path difference is primarily deterministic. Adaptive equalization of this difference is computationally inefficient, as depicted in Fig. 2.4. The main reason for this is twofold. First, most analog path differences are infinite impulse response (IIR) in nature, while adaptive equalization algorithms such as those in the LMS family are FIR. Secondly, even if the path difference were FIR, the LMS-based adaptive equalizer requires two multipliers per filter tap, as opposed to just one for a fixed FIR filter. Hence, in the finalized version of the receiver described

[2]i.e. the ratio between the main and alternate path effective baseband frequency responses

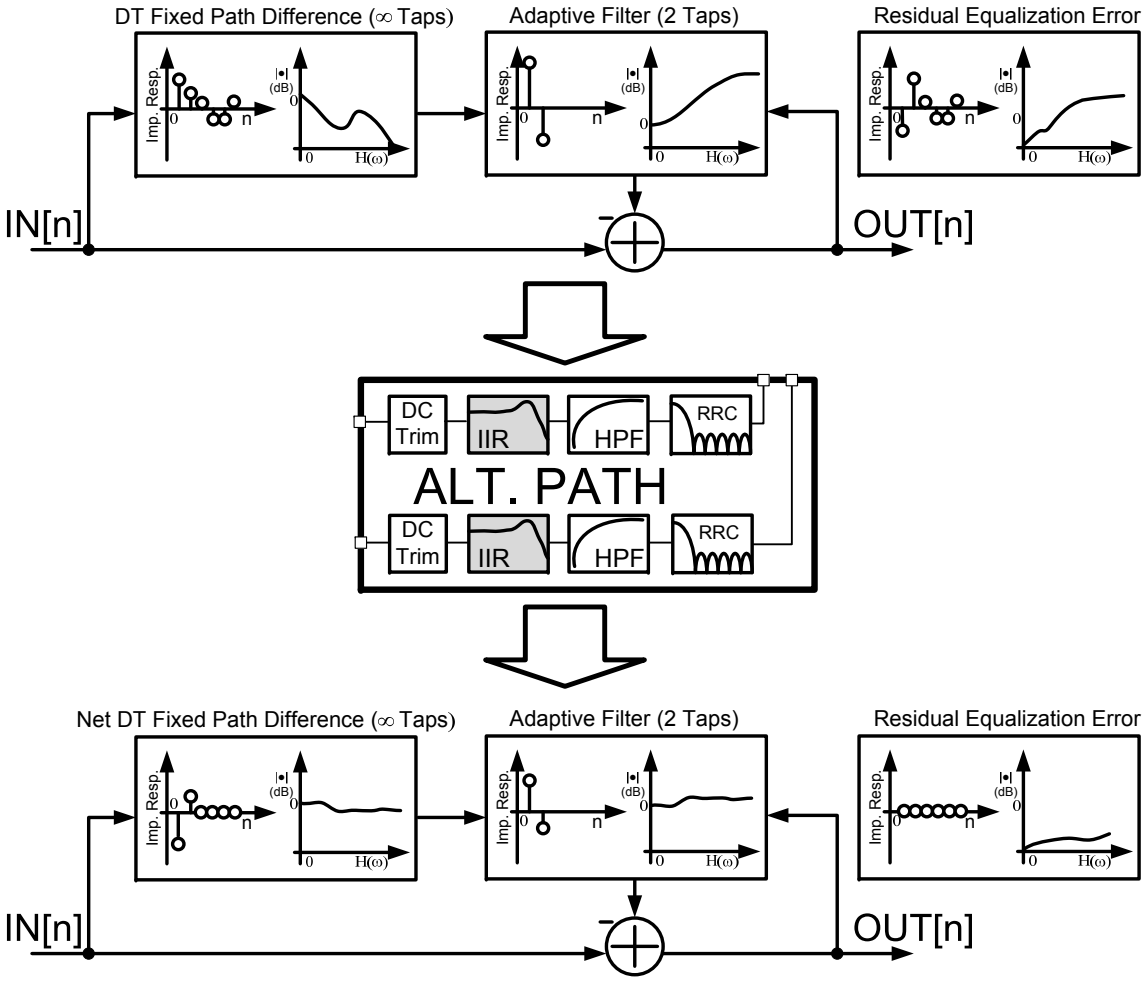


Figure 2.4: Reduction of required adaptive filter FIR taps via use of IIR filter to compensate for known path difference.

in this chapter, the known difference between the main and alternate path transfer functions is equalized by fixed real three-multiplier IIR filters. The remaining difference between the two paths is a complex dc gain and a small random mismatch in the baseband transfer function. This difference is broadband in the frequency domain and by the duality principle will correspond to a small number of taps required in the adaptive equalizer, as depicted in the lower portion of Fig. 2.4.

In this project, the normalized-LMS (NLMS) algorithm is utilized, as its convergence speed is in general superior to that of canonical LMS [68] [69] [70]. Furthermore , the NLMS does not suffer from the gradient noise amplification [37, pg. 320] that occurs in LMS for a large input signal. Functionally, the NLMS algorithm differs from the canonical LMS algorithm in that the tap update variable is normalized by the magnitude of the incoming reference signal.

2.2.3 Dc Offset in Direct Conversion Receivers

As is well-known, direct conversion receivers are susceptible to large dc offsets at baseband [10, pg. 131] [71]. This presents an issue with regards to the adaptive filter in that if both dc offset and

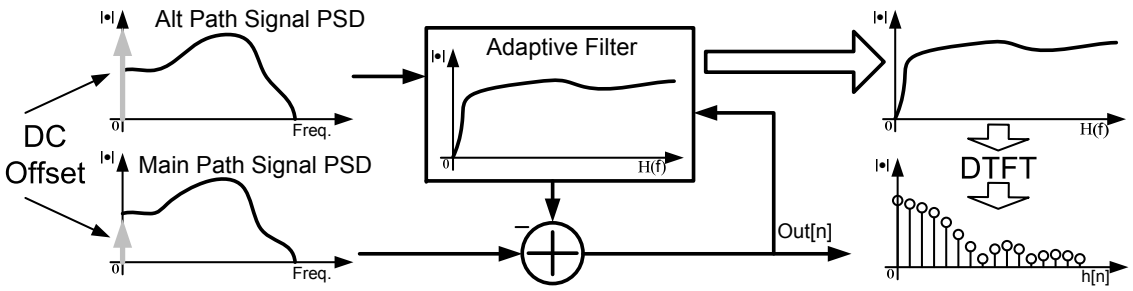


Figure 2.5: Dc offset cancellation problem in LMS filters.

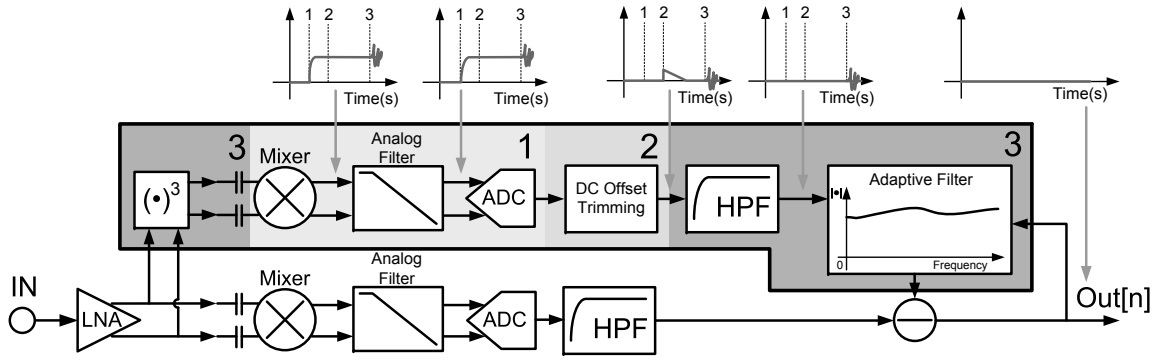


Figure 2.6: Proposed dc offset cancellation scheme showing dc offset transient waveforms.

IM3 signal are present on both main and alternate paths, the adaptive equalizer will attempt to equalize both signals. However, because the dc offset is large and uncorrelated with the baseband path difference at low frequencies, the optimal transfer function of the adaptive equalizer will have a large impulse at dc in the frequency domain. Due to the duality principle, this corresponds to a very large number of adaptive filter taps in the time domain, as depicted in Fig. 2.5, which will consume inordinate amounts of power and degrade the performance of the equalizer.

A common solution to the dc offset issue in UMTS receivers is to use high-pass filtering at baseband [9] [40] [61], with a cutoff frequency no greater than 10kHz [40]. In this work, high-pass filtering at 10kHz is performed in the digital domain for both main and alternate paths to remove the dc offset of the complete analog portion of the receiver, including the ADC. One significant problem with high pass filtering in either domain is that when the alternate path powers on, the dc offset of the mixer appears as a step to the alternate path high pass filter. The resultant step response takes 3-4 HPF time constants to settle below the error floor, effectively preventing convergence of the adaptive filter during this time. This is a problem because the cutoff frequency of the HPF is very low. Another option is to adaptively remove dc offset as part of the equalizer algorithm [28] [72]. However, it is shown in Appendix B that this technique effectively implements a high-pass filter and would have the same settling time issue.

The solution utilized in this work to remove this startup transient is to retain the high-pass filters mentioned earlier, but to also power on the alternate path and measure the dc offset in

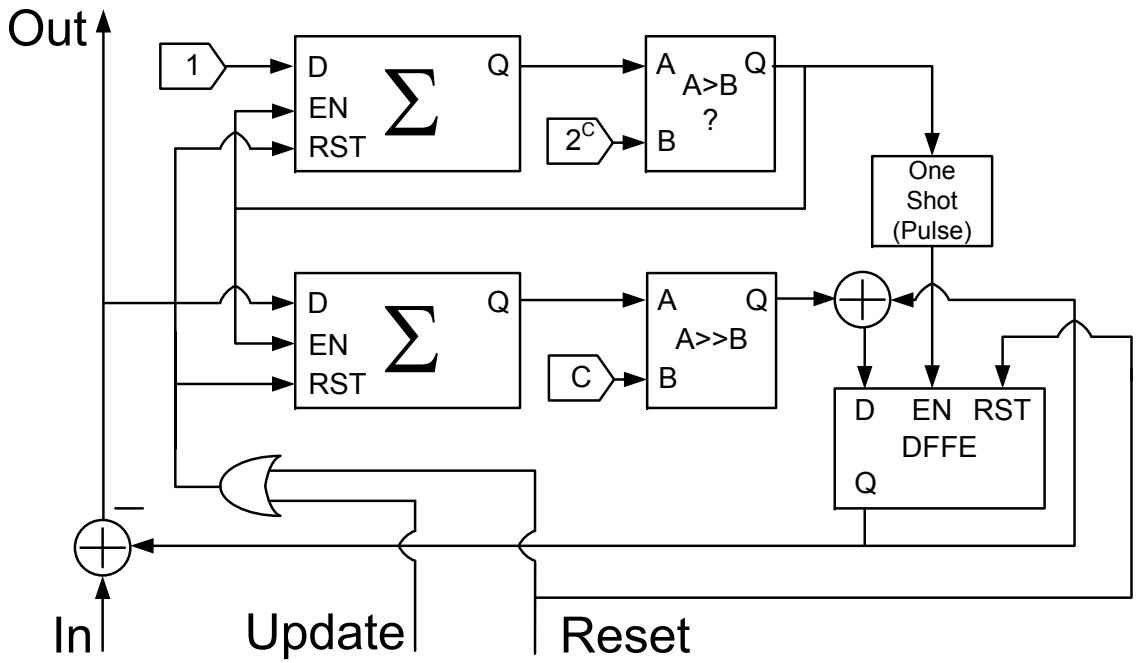


Figure 2.7: Exemplary digital dc offset correction circuit.

the digital domain in the absence of IM3 products being passed through the alternate path. This measurement is then immediately subtracted from the incoming signal to remove the dc offset. Since the only signals present at this time are dc offset and a small degree of noise, these operations can be performed relatively quickly (a few μs) by a simple averaging circuit immediately prior to enabling the full alternate path. The complete dc offset cancelling scheme is depicted in Fig. 2.6. A possible digital dc offset trimming circuit, used in the experimental receiver architecture described in Chap. 3, is shown in Fig. 2.7. Here, the input signal average is computed over 2^C samples and then subtracted from the input signal.

2.3 Sources of Error in Alternate Feedforward Path

The alternate feedforward path of the proposed receiver architecture suffers from some of the same error sources as do traditional receiver designs, but in often different ways and to different degrees.

2.3.1 Linear Term Feedthrough

The first requirement of the alternate path is that it must heavily attenuate the incoming desired signal (i.e. linear term feedthrough) with respect to the IM3 products. This is visually depicted at point 3 of Fig. 2.3. The reason for this is that any desired signal at the reference input of the equalizer will be treated as error by the adaptive algorithm. As the algorithm functions so as to minimize mean squared error, it will attempt to strike a balance between eliminating the IM3 products and desired signal, reducing small-signal gain, IM3 cancellation, or both.

2.3.2 IM3-to-Noise Ratio (INR) and IM3-to-Error Ratio (IER)

Common metrics used to quantify the signal processing quality of a circuit block include signal-to-noise ratio (SNR) and signal-to-noise-and-distortion ratio (SNDR). However, in the proposed receiver architecture a principal concern is the quality of signal processing of nonlinear IM3 distortion products. To avoid confusion between these signals and desired receiver signals, a quantity termed INR is introduced which denotes the ratio of IM3 products to integrated noise within the desired signal channel. Generalizing this metric to other forms of error results in a quantity denoted IM3-to-error ratio, or IER. This latter metric is most useful in that it can be used to quantify the performance impact of the alternate path on the main path receiver. Taking the output of the alternate feedforward path as the output of the adaptive filter of Fig. 2.3, the adaptive equalizer enforces the equality of the IM3 products at the outputs the main and alternate paths (2.1). Although this equality is never perfectly achieved, any such deviation can be thought of as contributing to the alternate path IER.

$$I_{MAIN} = I_{ALT} \tag{2.1}$$

After cancellation of the IM3 products, the equalizer output contains error due to thermal noise, higher order IMD products, and several other effects. As depicted in (2.2), this error can be attributed to separate sources in the main and alternate paths. In (2.2) it is assumed that this error is uncorrelated to good approximation.

$$E_{TOT}^2 = E_{MAIN}^2 + E_{ALT}^2 \tag{2.2}$$

Using (2.1) and (2.2) it can be seen that the relationship of (2.3) holds.

$$E_{TOT}^2 = E_{MAIN}^2 \left(1 + \left(\frac{IER_{MAIN}}{IER_{ALT}}\right)^2\right) \tag{2.3}$$

With the relation (2.3), the total receiver error can be input-referred using the main path small-signal gain. This error can then be compared with the receiver specifications, which are typically also input-referred. Since the total allowed error, main path IM3 products, and other main path error terms are known prior to consideration of the alternate path, the required IER of the alternate path can be determined by (2.3) for a given blocking condition. As the power of the IM3-producing blocker signals are varied, the IER quantities of the two paths roughly track. Using (2.3) it can be seen that this implies that the maximum total error occurs when the main path error is at its maximum. This condition occurs under the peak, or worst-case, blocking condition. Hence, the performance of the alternate path circuitry is specified at this point. Knowledge of the main path error and IER under peak blocking conditions, denoted $IER_{MAIN,PK}$, along with the receiver error specifications sets a requirement on the alternate path IER under peak blocking, denoted $IER_{ALT,PK}$. Provided that this requirement is met, the receiver error requirement will be met for all other blocking conditions, as E_{MAIN}^2 will decrease and $\frac{IER_{MAIN}}{IER_{ALT}}$ will remain roughly constant for smaller blocker levels.

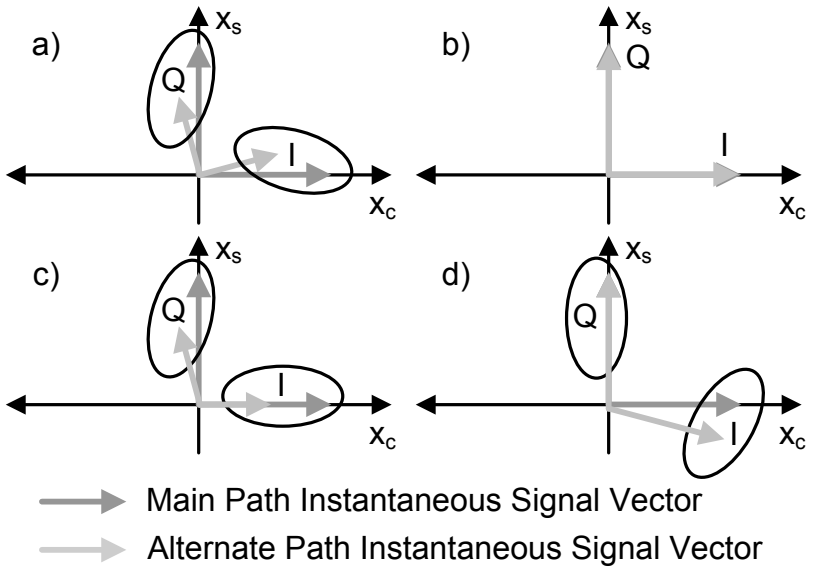


Figure 2.8: Insufficiency of complex LMS/NLMS in presence of alternate and main path signal vector mismatches.

2.3.3 Higher-Order Nonlinear Terms

The large-signal transfer functions of real-world devices may be characterized by Taylor series with an infinite number of terms. Assuming that the Taylor series coefficients in both the main and alternate paths are the same for all possible sets of input blocker frequencies, all higher-order nonlinear interference terms will cancel under equalization. However, this goal is all but impossible to guarantee in design. As the ratio of higher-order nonlinear terms to IM3 products is largest under peak blocking conditions, it is best to guarantee that the IER_{PK} due to these higher-order terms is significantly greater than the total IER_{PK} requirement in both paths. Cutting the IER performance close in this case is a bet on the accuracy of the nonlinear device models, and is not recommended. This requirement can be verified at the block level with a harmonic balance simulation, varying the number of calculated harmonics in order to isolate the magnitude of higher-order terms that fall at the same frequency as the IM3 products.

2.3.4 Quadrature Mismatch and LMS Equalization

Quadrature mismatch is typically not a major issue in meeting the sensitivity requirements of direct-conversion receivers, as the image signal is merely the quadrature component of the desired signal and is hence on the same order of magnitude. In the receiver presented in this work, however, quadrature mismatch holds the same position of importance as it does in image-reject superheterodyne architectures, as the interfering IM3 distortion signal is typically at least an order of magnitude larger than the desired signal. This significance becomes evident after an examination of the complex LMS algorithm.

The operation of the two equalizers in their canonical complex form is described by (2.4), with bold symbols denoting column vector quantities and the italicized portions corresponding to the

NLMS algorithm alone. The variables in (2.4) correspond to the signal variables in Fig. 1.24. For adaptive feedforward error-corrected receivers in which the alternate path extends back past the mixer and in which the multistage cubic term generator proposed in Section 4.2 is used, the complex form of the LMS-based algorithm is required, as it is impossible to guarantee a fixed phase relationship between the main and alternate paths due to the nonconstant nature of the cubic term generator interstage frequency response.

$$e[n] = m[n] - \mathbf{w^H[n]a[n]};\;\; \mathbf{w[n+1]} = \mathbf{w[n]} + \tilde{\mu}e[n]\mathbf{a[n]};\;\; \tilde{\mu} = \frac{\mu}{\|\boldsymbol{a[n]}\|^2} \tag{2.4}$$

When expanded into a physical hardware implementation, (2.4) takes the form of equations (2.5)-(2.9).

$$\tilde{\mu} = \frac{\mu}{\|\boldsymbol{a}_I\boldsymbol{[n]}\|^2 + \|\boldsymbol{a}_Q\boldsymbol{[n]}\|^2} \tag{2.5}$$

$$e_I[n] = m_I[n] - \mathbf{w_I^T[n]a_I[n]} + \mathbf{w_Q^T[n]a_Q[n]} \tag{2.6}$$

$$e_Q[n] = m_Q[n] - \mathbf{w_Q^T[n]a_I[n]} - \mathbf{w_I^T[n]a_Q[n]} \tag{2.7}$$

$$\mathbf{w_I[n+1]} = \mathbf{w_I[n]} + \tilde{\mu}[e_I[n]\mathbf{a_I[n]} + e_Q[n]\mathbf{a_Q[n]}] \tag{2.8}$$

$$\mathbf{w_Q[n+1]} = \mathbf{w_Q[n]} + \tilde{\mu}[e_Q[n]\mathbf{a_I[n]} - e_I[n]\mathbf{a_Q[n]}] \tag{2.9}$$

As can be seen from a one-tap implementation of (2.5)-(2.9), the signals in the main and alternate paths must be related by a Givens rotation (2.10) for a solution to exist to the complex filter tap such that complete equalization is achieved.

$$\begin{bmatrix} m_I[n] \\ m_Q[n] \end{bmatrix} = \begin{bmatrix} w_I[n] & -w_Q[n] \\ w_Q[n] & w_I[n] \end{bmatrix} \begin{bmatrix} a_I[n] \\ a_Q[n] \end{bmatrix} \tag{2.10}$$

This is the case in Fig. 2.8a which shows a vector representation of IM3 products in the main and alternate paths. After the complex equalizer applies the proper Givens rotation to the alternate path signal, subtraction yields complete cancellation, as in Fig. 2.8b. However, if as in Fig. 2.8c phase and rotational mismatch exist between the two paths, then clearly their respective signal vectors are not related by a Givens rotation. In this case, complete cancellation cannot be achieved, limiting the IER of the adaptive equalization. This effect is quantified in Appendix A and Fig. 2.9. It is evident that even in the absence of I/Q gain mismatch, small phase mismatches can severely limit the performance of the adaptive equalization. Note that in Fig. 2.9, ϕ_R is equal to the rotational mismatch between the main and alternate paths, while ϕ_M and ϕ_A are the quadrature phase mismatches in the main and alternate paths, respectively.

2.4 Translation of Receiver Specification

Based on the high-level architectural choices and concepts presented in the previous section, it is now possible to quantitatively approach the block level design of the alternate path. The first step in this task is to determine the performance of the main path, namely, of $IER_{MAIN,PK}$. At this point in the chapter and in the design process, a rough idea of the main receiver path performance

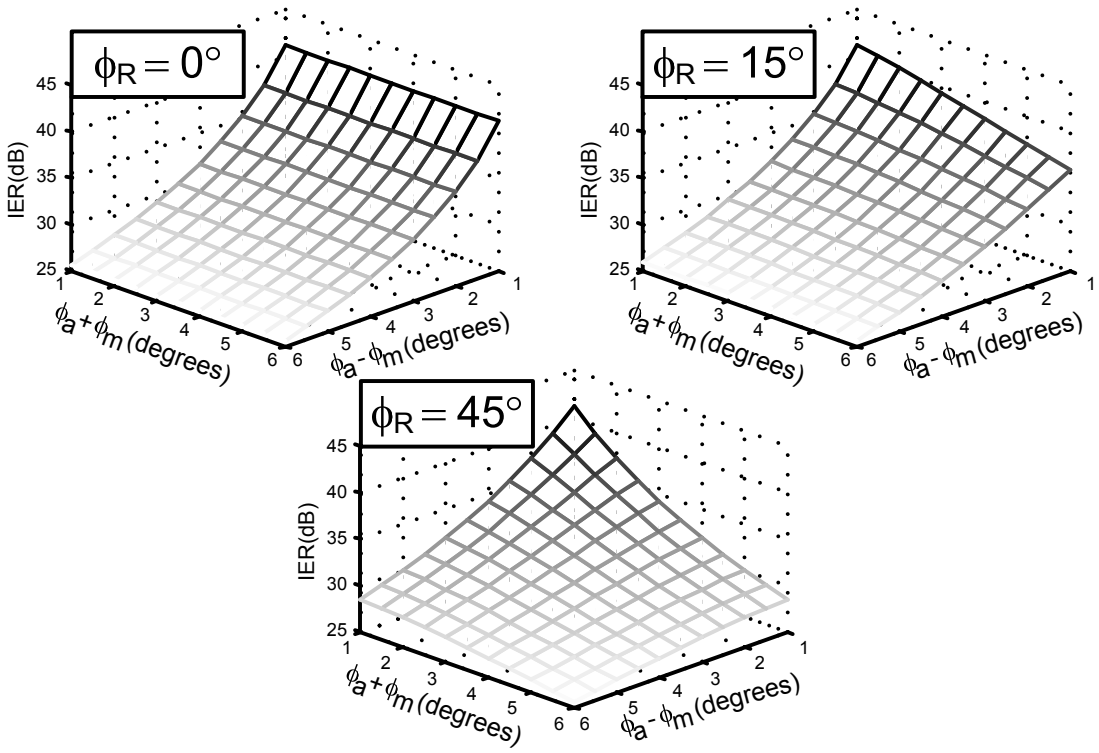


Figure 2.9: IER vs. rotational mismatch and difference in phase mismatch.

is sufficient to move forward.

2.4.1 Main Path

In order to determine the worst-case "two-tone" blocker scenario seen by the receiver circuitry, the UMTS blocker specification [6] must be used in conjunction with the frequency response of the duplexer shown in Fig. 2.1. From Fig. 2.10 it can be seen that for the duplexer described in [73], the largest IM3 products occur when f_{TX}=1.98GHz, f_{CW}=2.05GHz, and f_{RX}=2.12GHz. In this case, the blocker powers are P_{TX} = +28dBm and P_{CW} = -30dBm at the antenna and P_{TX} = -26dBm and P_{CW} = -34dBm at the LNA input. Interestingly, in this case, the worst-case blocking scenario does not occur over the band in which the specified CW blocker is -15dBm.

According to [15], the UMTS specifications impose an analog requirement of $NF_{ANT,MAX}$ = 9dB at the antenna. A more general definition of NF denoted "error figure" (EF) is adopted to encompass other forms of error including distortion products. Although potentially cumbersome at this point, the EF quantity will later help relate more well known receiver specifications to $IER_{ALT,PK}$. The UMTS specification allows $EF = NF$+3dB under blocking conditions. In other words, $EF_{ANT,MAX}$ = 12dB. Given the insertion loss of the duplexer [73] L_{DUP} = 1.8dB and that $NF_{ANT} = L_{DUP}$+NF_{RX}, it is computed that $NF_{RX,MAX}$ = 7.2dB and $EF_{RX,MAX}$ = 10.2dB at the LNA input.

For UMTS, the noise due to the 50Ω source resistance is kTB = -108dBm / 3.84MHz at the LNA input. Denoting all quantities as LNA input-referred, this implies that after removing source

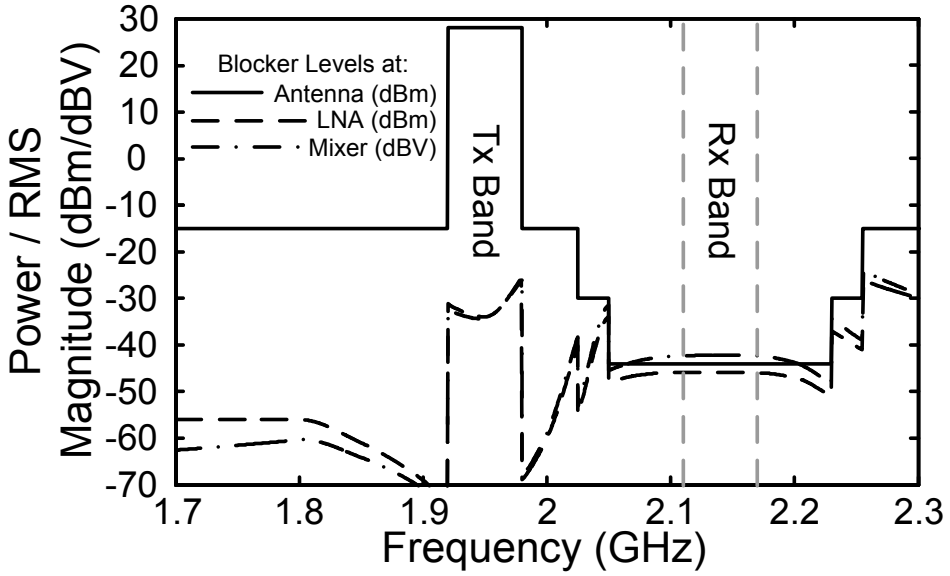


Figure 2.10: Expected UMTS blocker profile at various points in the receiver.

noise, the maximum allowed receiver noise power is $N_{RX,MAX}$ = -101.7dBm / 3.84MHz. Assuming that the error under worst-case blocking is dominated by thermal noise and IM3 distortion, the error figure limit implies that the rms sum of $I_{RX,MAX}$ and $N_{RX,MAX}$ is -98.2dBm / 3.84MHz, where $I_{RX,MAX}$ is the maximum allowed IM3 distortion product power. It follows that $I_{RX,MAX}$ = -100.8dBm / 3.84MHz. Using these values in (2.11) yields $IIP3_{RX,MIN}$ = +3.4dBm.

$$IIP3_{RX,MIN} = \frac{1}{2}[2P_{CW} + P_{TX} - I_{RX,MAX}] \tag{2.11}$$

The problem with this requirement is that it is higher than typical attainable values for SAW-less receivers in the absence of special enhancements (at the time of this research). For example, typical values for $IIP3_{MIXER}$ range from +8 to +12dBm [33]. For the initial design in this work, the simulated values for the LNA gain G_{LNA} and $IIP3_{LNA}$ are 17dB and +6dBm, respectively. Recalling the IIP3 relation (2.12) from [10, pg. 23] as used in [33], such design values yield $IIP3_{RX}$ = -9.1dBm for $IIP3_{MIXER}$=+8dBm.

$$IIP3_{RX} = \left(\frac{1}{IIP3_{LNA}} + \frac{G_{LNA}}{IIP3_{MIXER}}\right)^{-1} \tag{2.12}$$

Clearly a significant discrepancy arises and some sort of additional IIP3 enhancement is required to meet the input-referred error specification. The proposed adaptive feedforward error-cancelling loop is therefore added to the receiver. In order to begin a quantitative design of the alternate path, however, it is necessary to first determine $IER_{MAIN,PK}$. Using (2.11), the main path IM3 product power (input-referred) under peak blocker conditions $I_{MAIN,PK}$ is equal to -75.8dBm / 3.84MHz. Adding 2dB of margin to the NF requirements and assuming that $EF_{RX,MAX} = NF_{RX,MAX}$+2dB (allowing 1dB margin for error due to the alternate path) the maximum error power, including 50Ω source noise but not IM3 products, referred to the main path LNA input is $E_{MAIN,PK}$ = -100.8dBm.

Hence, $IER_{MAIN,PK} = 25$dB. This number can now be used to determine a target design value for $IER_{ALT,PK}$.

2.4.2 Alternate Path

From $IER_{MAIN,PK}$ and the allotted 1dB error margin for the alternate path enhancement, the requirement on $IER_{ALT,PK}$ can be determined as follows. To obtain $EF_{RX,MAX}$, all error is referred to the output of the equalizer in (2.13), where G_{MAIN} represents the small-signal gain of the main path.

$$EF_{RX,MAX}(dB) = 10\log_{10}\left(\frac{E^2_{MAIN,PK} + E^2_{ALT,PK}}{G^2_{MAIN}kTB}\right) \tag{2.13}$$

Recalling that the adaptive equalizer forces $I_{MAIN,PK}$=$I_{ALT,PK}$, substitution yields (2.14).

$$EF_{RX,MAX}(dB) = 10\log_{10}\left(\frac{E^2_{MAIN,PK}}{G^2_{MAIN}kTB}\right) + 10\log_{10}\left(1 + \frac{IER^2_{MAIN,PK}}{IER^2_{ALT,PK}}\right) \tag{2.14}$$

The second term in (2.14) represents the excess error figure due to the operation of the alternate path. Given the design numbers in the previous subsection, $IER_{ALT,PK}$ >31dB. For simplicity, this error requirement is split equally between the cubic term generator and the remainder of the alternate path, yielding $IER_{CUB,PK}$ >34dB.

2.5 Enhanced Degree of Freedom Adaptive Algorithms

In order to overcome the constraint imposed by the Givens rotation relation of canonical complex LMS, an additional degree of freedom must be added to the adaptive algorithm in order to accommodate gain and phase mismatch between the main and alternate paths. In the context of the Givens rotation, this means that the relationship (2.10) should change to (2.15), where the introduction of $x_I[n]$ and $x_Q[n]$ collectively constitute the additional degree of freedom.

$$\begin{bmatrix} m_I[n] \\ m_Q[n] \end{bmatrix} = \begin{bmatrix} w_I[n] & x_I[n] \\ w_Q[n] & x_Q[n] \end{bmatrix} \begin{bmatrix} a_I[n] \\ a_Q[n] \end{bmatrix} \tag{2.15}$$

In this case the new algorithm error subtraction equations become (2.16).

$$\begin{aligned} e_I[n] &= m_I[n] - \mathbf{w_I^T}[\mathbf{n}]\mathbf{a_I}[\mathbf{n}] - \mathbf{x_I^T}[\mathbf{n}]\mathbf{a_Q}[\mathbf{n}] \\ e_Q[n] &= m_Q[n] - \mathbf{w_Q^T}[\mathbf{n}]\mathbf{a_I}[\mathbf{n}] - \mathbf{x_Q^T}[\mathbf{n}]\mathbf{a_Q}[\mathbf{n}] \end{aligned} \tag{2.16}$$

An equivalent solution to the I/Q mismatch problem in a complex adaptive equalizer was suggested in [74][75]. Unlike in [75], however, the equalizer tap update relations are developed in this work for LMS-based adaptive equalizers by viewing the complex LMS algorithm as a completely real implementation and by substituting the relations (2.17) into (2.4). It is then seen that (2.16) is satisfied and that (2.18)-(2.19) realize the tap update equations for the new algorithm.

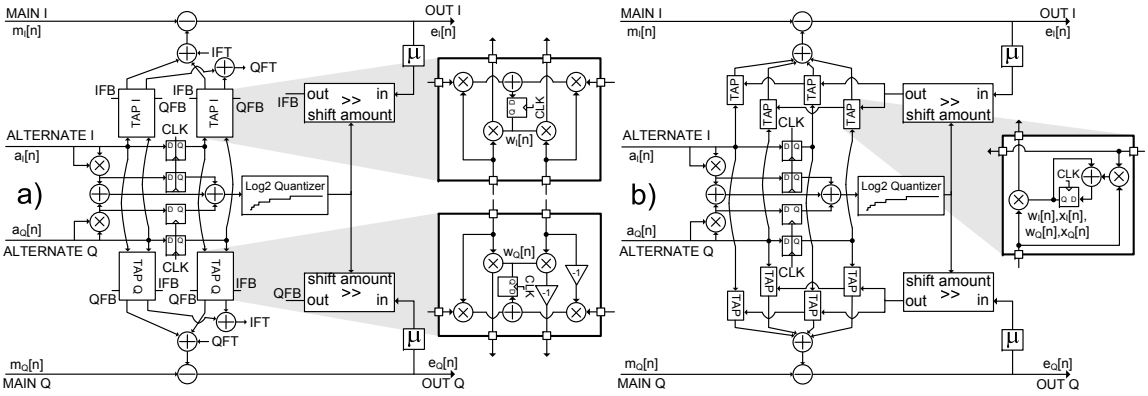


Figure 2.11: Hardware implementation of complex NLMS algorithm. a) Canonical. b) Proposed architecture with enhanced degree of freedom.

$$e[n] = \begin{bmatrix} e_I[n] \\ e_Q[n] \end{bmatrix};\ \mathbf{a[n]} = \begin{bmatrix} a_I[n] \\ a_Q[n] \end{bmatrix}$$
$$m[n] = \begin{bmatrix} m_I[n] \\ m_Q[n] \end{bmatrix};\ \mathbf{w[n]} = \begin{bmatrix} [\mathbf{w_I[n]}\ \ \mathbf{w_Q[n]}]^T \\ [\mathbf{x_I[n]}\ \ \mathbf{x_Q[n]}]^T \end{bmatrix} \tag{2.17}$$

$$\mathbf{w_I[n+1]} = \mathbf{w_I[n]} + \tilde{\mu}[e_I[n]\mathbf{a_I[n]}];\ \ \mathbf{w_Q[n+1]} = \mathbf{w_Q[n]} + \tilde{\mu}[e_Q[n]\mathbf{a_I[n]}] \tag{2.18}$$

$$\mathbf{x_I[n+1]} = \mathbf{x_I[n]} + \tilde{\mu}[e_I[n]\mathbf{a_Q[n]}];\ \ \mathbf{x_Q[n+1]} = \mathbf{x_Q[n]} + \tilde{\mu}[e_Q[n]\mathbf{a_Q[n]}] \tag{2.19}$$

This change is efficient from a hardware perspective, as shown in Fig. 2.11. Both the original NLMS and enhanced-degree-of-freedom NLMS equalizers have the same number of multipliers, which dominate the power and area consumption of the digital implementation. By contrast, an equivalent solution of placing two canonical LMS equalizers in parallel while conjugating the complex input of one of them [76] doubles the digital hardware burden.

2.6 Criteria and Techniques for Enabling and Disabling Feedforward Loop

In order to use the alternate path most efficiently, it must be powered on only when IM3 products corrupt the main path signal. Detecting this condition cannot be done with the use of a simple power detector at RF, as such a circuit cannot discriminate between single and multiple blockers. Using a power detector at main path baseband is similarly ineffective, as it cannot discriminate between large IM3 products and large desired signal.

A superior method of detecting a problematic blocking condition is to use a portion of the alternate path itself, as its baseband output power is proportional to the IM3 products corrupting the main path signal. For example, the RF front end portion of the alternate path can be powered on in the event that a problematic blocking condition is possible and its baseband output monitored

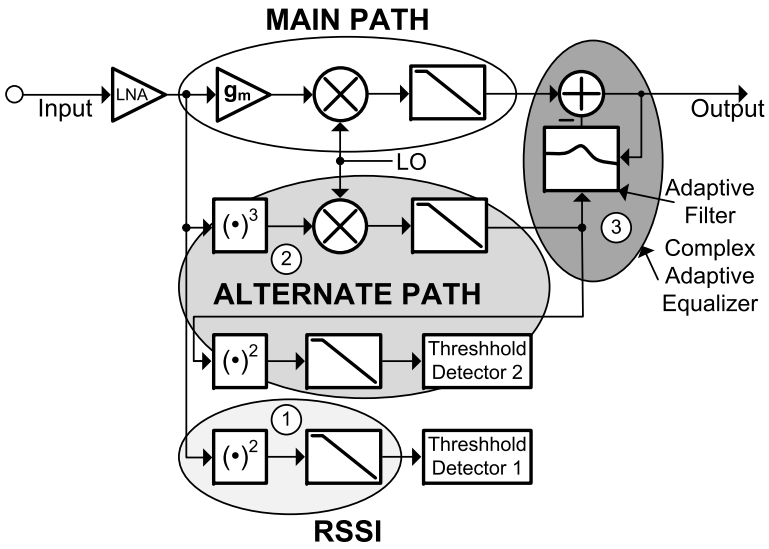


Figure 2.12: Proposed procedure for determining a turn-on condition for the linearity enhancement.

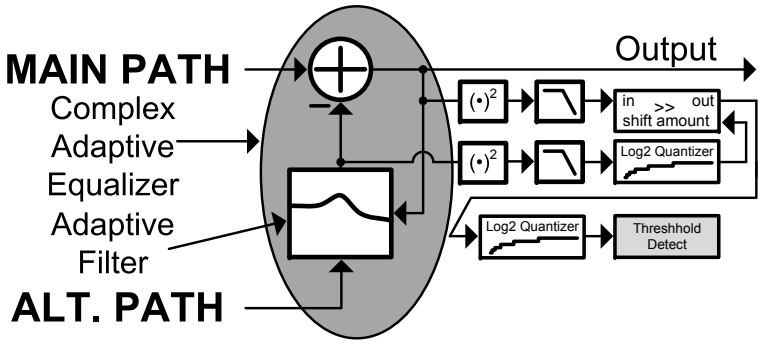


Figure 2.13: Proposed procedure for determining a turn-off condition for the linearity enhancement.

by a simple power detector circuit. Once a given IM3 power threshold is exceeded, the rest of the alternate path is then enabled to perform IM3 cancellation.

The alternate path front end enable condition can be flagged by an additional power detection circuit, such as an RSSI, at RF when the total blocker power exceeds a certain threshold or, as in UMTS or another FDD communications standard, when the TX output power is known to exceed a certain threshold. Use of the former procedure is depicted in Fig. 2.12 in which the output of the RSSI is connected to a threshold detection circuit which triggers the turn-on of the analog portion of the alternate path. The output of the analog portion of the alternate path is in turn rectified and its rms power detected. Once this power exceeds a given threshhold, the digital portion of the alternate path is enabled.

Based on measured results in the exemplary receiver described in Chap. 3, IM3 products large enough to corrupt the desired signal may arise when the TX output power is greater than +10dBm. According to [77], in UMTS this condition occurs 30% of the time in a cell with radius of 1km and 20 users. Thus, the time-averaged power consumption of the RF front end portion of the alternate path

would be reduced by this amount under these conditions. The time-averaged power consumption of the rest of the alternate path would be negligible, as its operation depends on the rare event when another blocker appears with the proper incident power and frequency so as to cause IM3 products.

In any detection procedure such as this, the possibility of a false alarm and its effects on the system must be considered. If this were to occur with the system described in this paper, both main and alternate path inputs to the adaptive equalizer would consist of uncorrelated noise-like signals. Since there is no significant correlated data in the main and alternate paths, the equalizer taps will not converge but will exhibit a small variation around the zero value. This noisy signal then multiplies the thermal noise in the alternate path. Since both signals are small, the resultant noise signal added to the main path is very small, resulting in a negligible increase in the receiver noise figure.

Powering off the alternate path can be done with a modified SNR detection procedure within the adaptive equalizer. The IM3 content of the main path can be easily estimated by considering the total summed signal power at the output of the equalizer taps. Comparing this quantity to the total power of the equalizer output yields an estimate of the desired signal to IM3 product ratio. When this ratio is high enough, the baseband portion of the alternate path can be turned off. This scheme is depicted in Fig. 2.13. Because this SNR estimate can fluctuate under fading conditions, a minimum on-time can be instituted for the alternate path baseband circuitry such that it does not toggle on and off repeatedly within a single blocking incident.

2.6.1 Behavior of System in Fading Environment

The adaptive equalizer in the architecture presented herein exists mainly to compensate for unknown circuit mismatches and the effective rotational phase difference between the main and alternate paths. However, the adaptive equalizer may also need to track slightly to compensate for changing blocker characteristics as a result of fading. Considering that most IM3-producing blockers for UMTS are clustered around 2GHz, and assuming that the maximum speed of the mobile terminal is 250km/h, the minimum coherence time of the blockers is about 400μs [78, pp. 34-40]. The adaptive equalizer needs to be able to converge faster than this amount in order to properly track the changing IM3 products, a requirement that is not difficult to meet.

Chapter 3

Implementation and Testing of Experimental Linearized RF Receiver[1]

3.1 Receiver Architecture

Based on the concepts described in Chapter 2, a radio receiver capable of meeting the UMTS sensitivity requirements under Region 1 specification worst-case blocking was built and tested. A block diagram of this system is shown in Fig. 3.1. The RF front end is implemented in 0.13μm RF CMOS and is mounted on a gold-plated high-frequency laminate substrate to which RF and baseband connections are made. The laminate substrate is in turn mounted on PCB which contains the baseband circuitry of the main and alternate paths, which is composed of low-power commercially available discrete components. The on-board ADCs interface via a bidirectional parallel connection, depicted in Fig. 3.2 to an FPGA platform which conducts the digital signal processing operations of the receiver in real-time. Aside from the bandgap circuitry which utilizes a 2.7V supply voltage (drawing 2.5mA), the remainder of the chip operates under a 1.2V supply voltage. The chip is fully ESD protected.

Although it is possible to perform the equalization in analog circuitry, shifting as much of the signal processing to the digital domain affords several advantages. For example, the behavior of digital circuitry is relatively insensitive to process variations, the continued scaling of CMOS processes has rendered baseband digital blocks power-competitive compared to equivalent analog blocks [71], and digital circuits facilitate the implementation of reconfigurable multimode receivers [79]. Implementing the baseband components off-chip has negligible impact on this experiment, as the integrated front end typically dominates the performance of the RF receiver. In an actual implementation, the receiver main path would also include VGA functionality. The experimental receiver described herein is for proof-of-concept only and represents the case when the receiver is attempting to decode

[1]Portions of this material have been previously published in [57],[58],[1], [59], and [2].

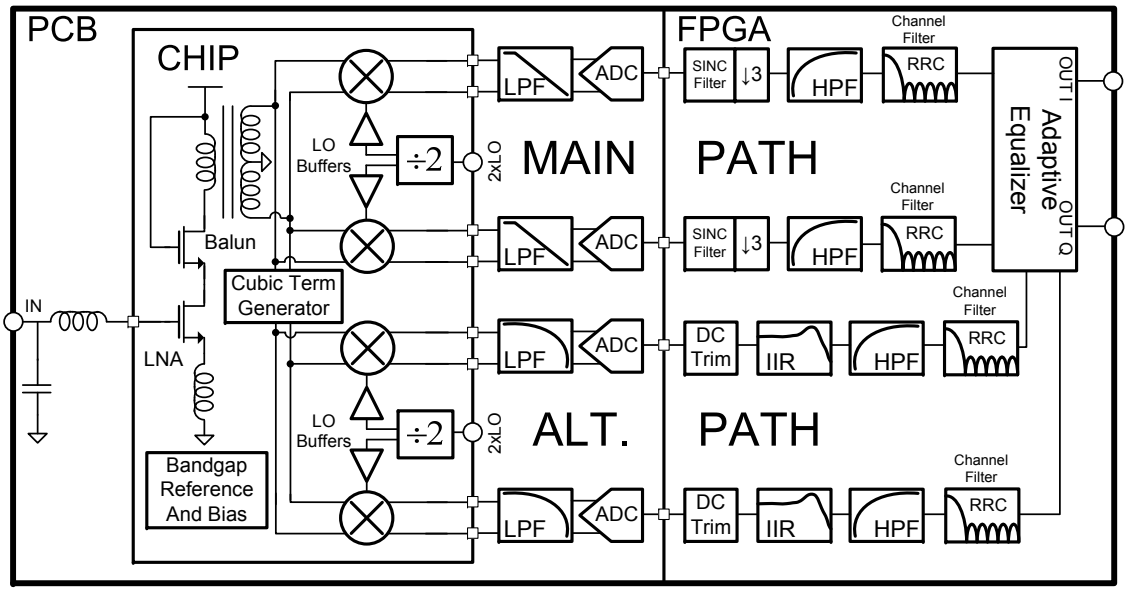


Figure 3.1: Experimental UMTS receiver architecture.

a signal near sensitivity. Were a VGA present in the main path, and were its gain to be changed during alternate path operation, the adaptive nature of the alternate path would allow the equalizer to quickly track the change and maintain IM3 cancellation.

The measurements performed on the receiver with a full equivalent UMTS downlink signal are postprocessed in MATLAB to obtain the DPCH sensitivity results. In the MATLAB code, rate change, synchronizing, derotation, despreading and decoding are performed to recover the original bit stream from the physical receiver output.

3.2 Circuit Block Implementations

3.2.1 Main Path Circuit Design

3.2.1.1 LNA and Balun

Reflecting a typical choice for narrowband receivers, the integrated front end employs an inductively degenerated cascode LNA. As the duplexer [73] has a single ended output, the LNA must have a single ended input. However, as the SAW filter to be removed for this design previously handled the single-ended to differential conversion between the LNA and mixer, provisions for performing this task must now be made on chip. An area-efficient method of accomplishing this goal is to place a secondary inductor winding inside of the LNA load inductor, creating a transformer balun. The secondary coil should be designed to have a maximum number of turns to improve voltage gain. Although this makes the Q of the secondary relatively poor, this has little effect on the Q of the LNA tuned load, as the load impedance seen by the secondary coil is high, and as a result power is not lost to it. The LNA and balun designs are depicted in Fig. 3.3.

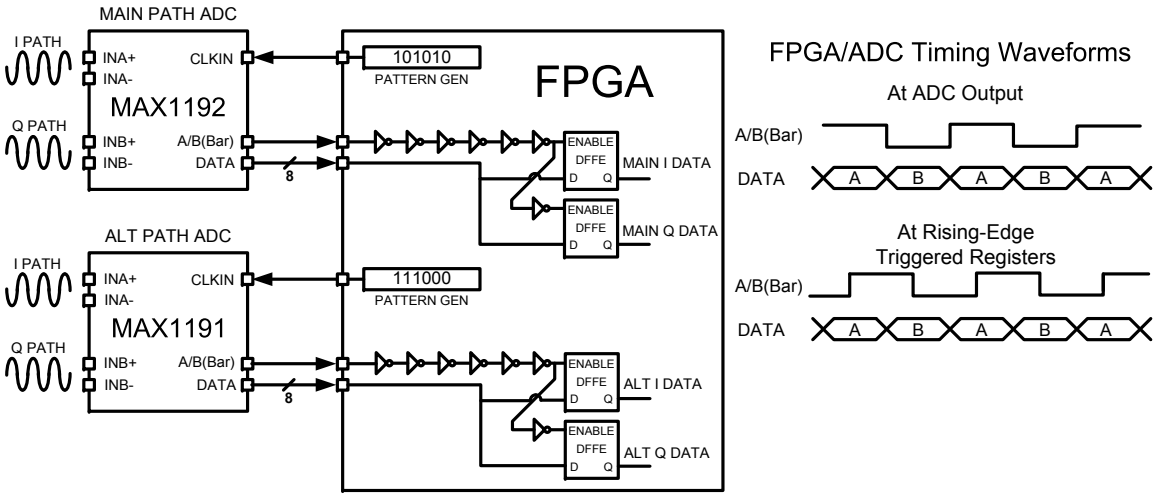


Figure 3.2: ADC-FPGA bidirectional interface.

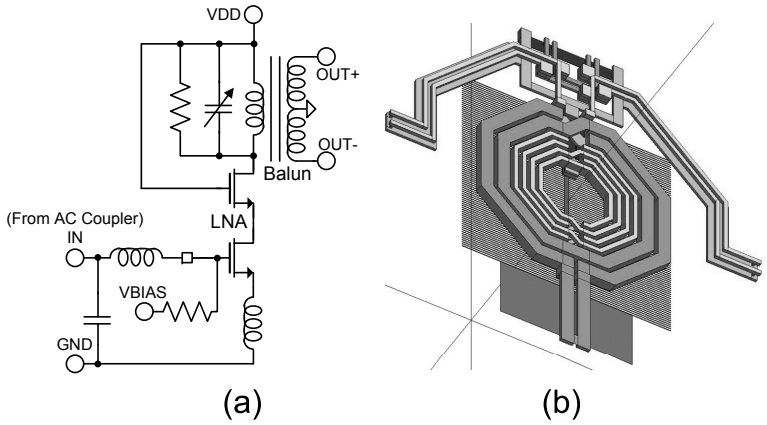


Figure 3.3: Implementation of main path LNA and balun. a) Schematic depiction. b) Balun 3-D CAD representation.

3.2.1.2 Mixer and LO Buffer

As the system proposed in this paper only equalizes IM3 products, the mixer utilizes a folded high-IIP2 mixer in order to obviate any IM2 equalization [41]. The schematic of this mixer as implemented is shown in Fig. 3.4. The common-mode feedback (CMFB) OTA is expanded in Fig. 3.5. In order to drive the large gate capacitances of the mixer switching pair, an actively loaded Cherry-Hooper LO buffer is utilized. The schematic of this block, shown in Fig. 3.6, reflects some biasing and neutralization modifications to the circuit shown in [21], permitting it to function under the low voltage supply headroom. Separate divide-by-two circuits are included immediately adjacent to the LO buffers in order to avoid problems associated with on-chip RF-LO coupling [11].

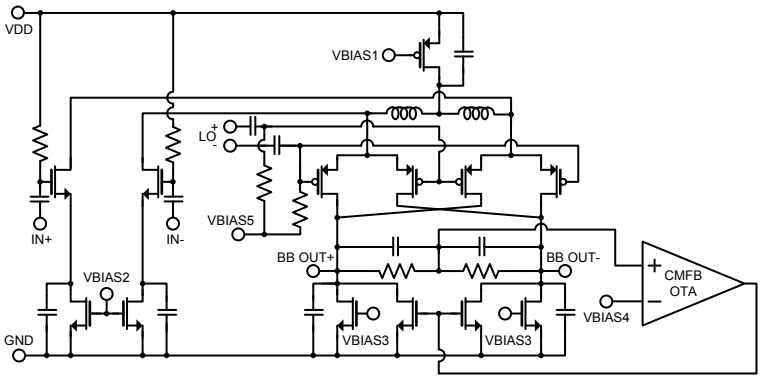


Figure 3.4: Main path high-IIP2 mixer.

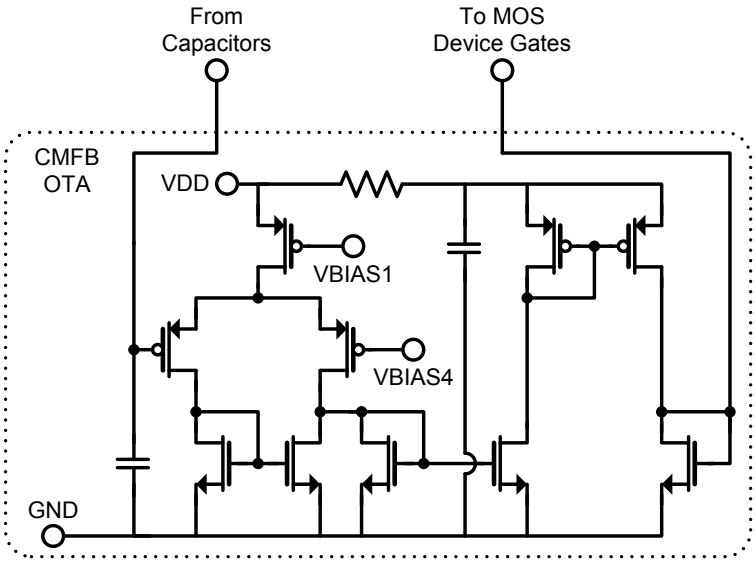


Figure 3.5: Mixer common-mode feedback circuit.

3.2.1.3 Analog and Digital Baseband Circuitry

The required order of the analog postfilter is obtained by considering the worst-case frequency translation to baseband of the blocker profile depicted in Fig. 2.10 after the LNA transfer function is added to the profile. It was found that for an ADC sampling rate of 50MHz, a 3rd-order Chebyshev filter in the analog domain was sufficient to attenuate downconverted out-of-band blocker signals to levels negligible with respect to the thermal noise floor. This filter was partially implemented on PCB with a nominal passband edge frequency of 2.3MHz and nominal amplitude and group delay ripples of 0.5dB and 82ns, respectively. Due to the integration of the analog postfilter with discrete components on PCB substrate, the actual amplitude and group delay ripple values are somewhat worse than these numbers. Nevertheless, these parameters to not constrain the effectiveness of the adaptive equalization. The analog postfilter does not produce attenuation within the desired signal channel due to the location of its passband edge frequency. Rather, close-in adjacent channel filtering

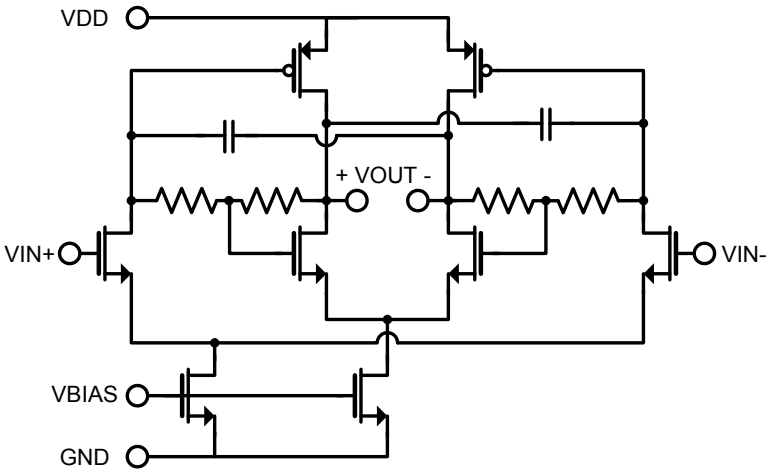


Figure 3.6: Main path Cherry-Hooper LO buffer.

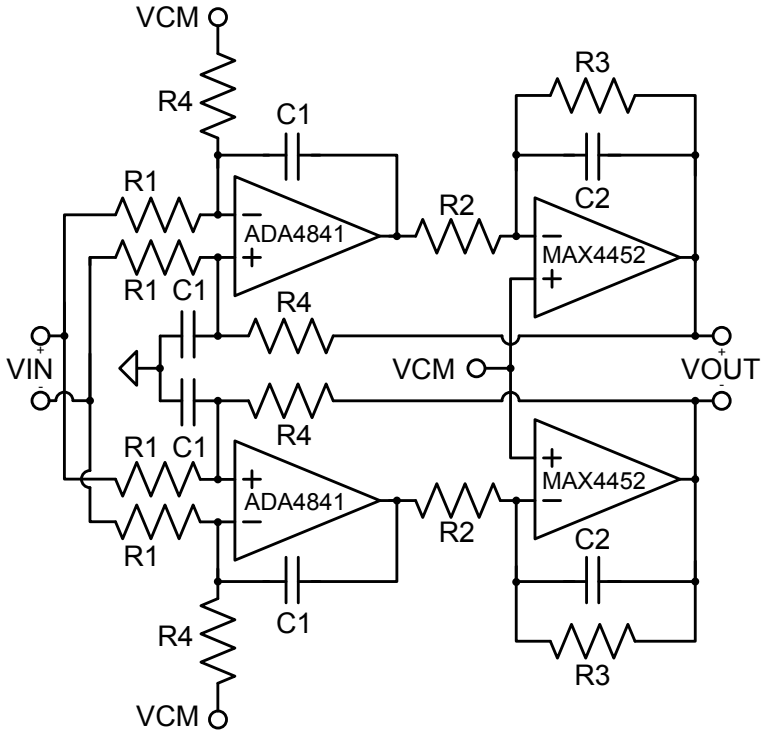


Figure 3.7: Main path 2nd-order biquad.

is provided in the digital domain by a 25-tap root-raised cosine FIR filter running at 16.66MHz. The third-order Chebyshev filter is formed by the combination of the real pole at the output of the mixer and the 2nd-order biquad filter shown in Fig. 3.7, the latter of which is composed of discrete components.

The analog postfilter also includes an additional gain and buffering stage, shown in Fig. 3.8, that drives the ADC which utilizes an 8-bit pipelined architecture running at 50MHz. The ADC

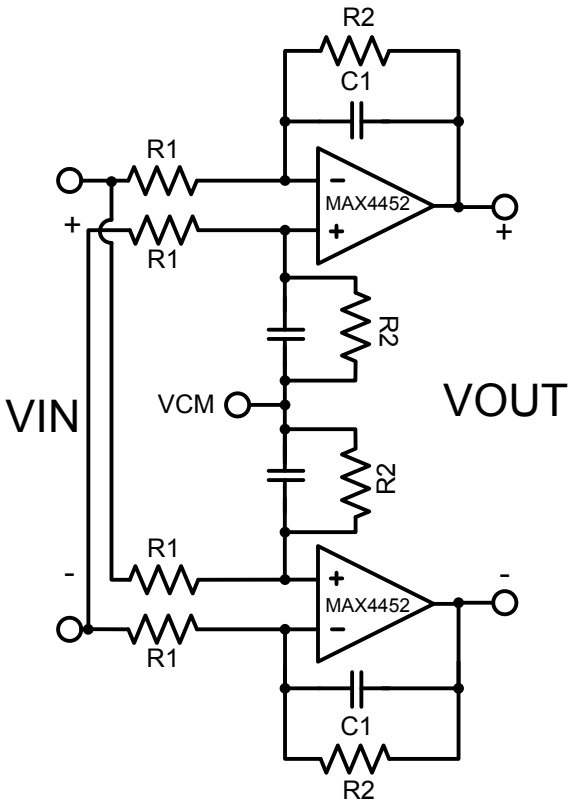


Figure 3.8: Main path gain and buffering stage prior to ADC.

sampling rates were chosen to be not equal to a multiple of the UMTS chip rate in order to facilitate testing with the particular FPGA platform used. However, as the baseband digital signal is oversampled, this choice does not compromise the integrity of the experiments presented herein. Coarse dc offset adjustments prior to the digital HPFs are also implemented to avoid saturating the baseband circuitry.

3.2.2 Alternate Path Circuit Design

The design of the alternate path in this experiment is consistent with the objectives described in Sections 2.3 and 2.4.2. The architecture and design implications of the cubic term generator are significant topics and are discussed fully in Chap. 4. The cubic term generator schematic is shown in Fig. 3.9 and is characterized by its multistage architecture in which third-order IMD products are generated by cascading a squaring, gain, and a final multiplication using a Gilbert cell. One pertinent implication of this multistage architecture is the phase response introduced by the delay of the interstage circuitry, shown in Fig. 3.10. Due to this large phase shift, a complex adaptive equalizer must be used in order to properly cancel IMD products in the main path even in the absence of test setup-induced phase differences. The mixer schematic is shown in Fig. 3.11. It is the same as that in Fig. 3.4 but without the IIP2-enhancing tuning inductor and with the tail current source split between the two differential halves of the circuit. The LO buffers need here be only

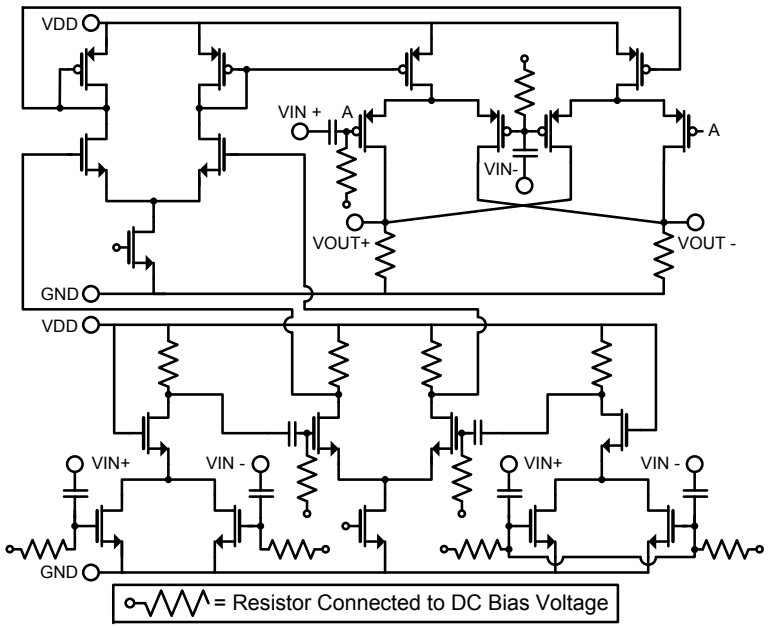


Figure 3.9: Cubic term generator schematic.

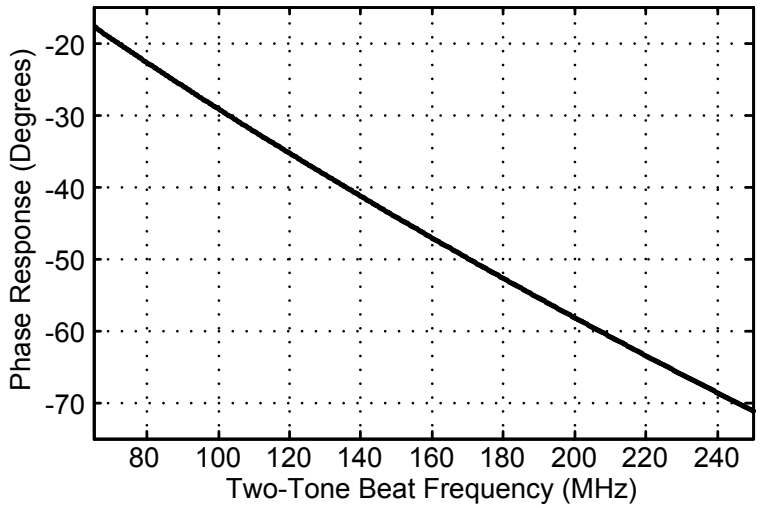


Figure 3.10: Simulated phase response of cubic term generator interstage phase response.

simple resistively loaded differential pairs, as shown in Fig. 3.12, due to the relatively small switching pair capacitance of the mixer. The active analog die circuitry of the alternate path consumes 6.7mA of current from the 1.2V supply and only 0.2mm^2 of die area, making it suitable for monolithic integration.

The outputs of the mixers are buffered by positive-feedback gain amplifiers with first-order filtering composed of discrete components on PCB, as shown in Fig. 3.13. These buffers drive the quadrature ADCs directly. As in the main path, 8-bit pipelined ADCs are used, but sample at

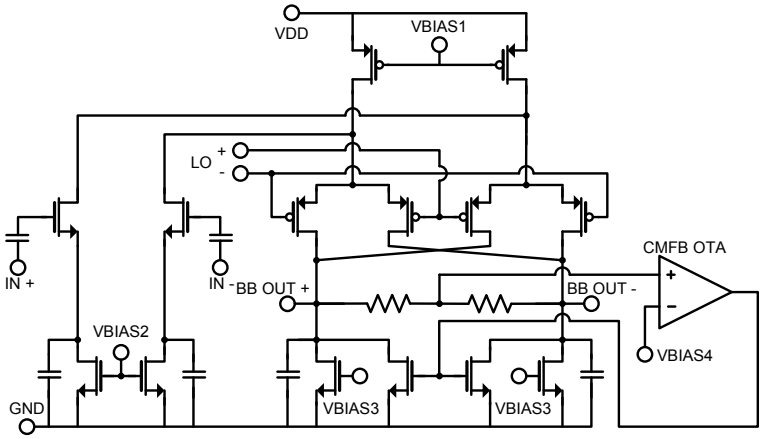


Figure 3.11: Alternate path mixer schematic.

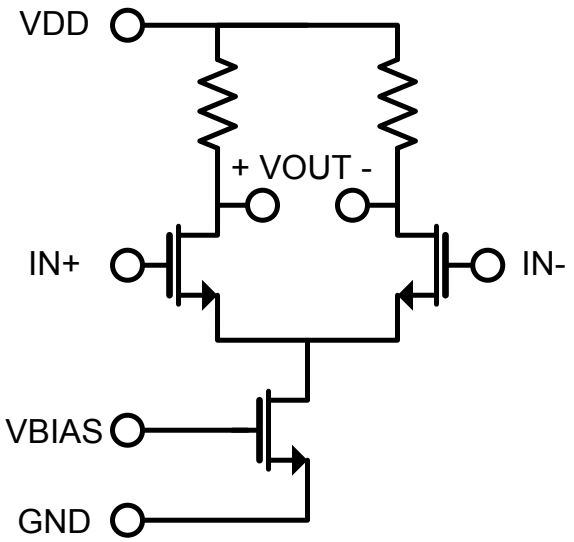


Figure 3.12: Alternate path LO buffer schematic.

16.66MHz. The alternate path baseband circuitry on the PCB, including the postfilters and ADCs, consumes less than 7.6mA under a 2.7V supply voltage. Also in the alternate path digital domain exists an IIR fixed equalization filter which accounts for most of the known LTI path mismatch between the main and alternate path. It is posited that calibration of this digital IIR filter is unnecessary, as the use of automatic calibration in the analog baseband postfilter to maintain a relatively constant frequency response over PVT variation is common practice in cellular receivers [9] [40] [61]. The implemented version of the adaptive equalizer is shown in Fig. 2.11b. Two doubly complex filter taps were used to adjust for perturbations in the baseband group delay over variations in LO frequency, although in practice the baseband group delays will also vary with temperature. The division associated with the NLMS algorithm is log2-quantized here, allowing the use of a simple barrel shifter as a divider [80].

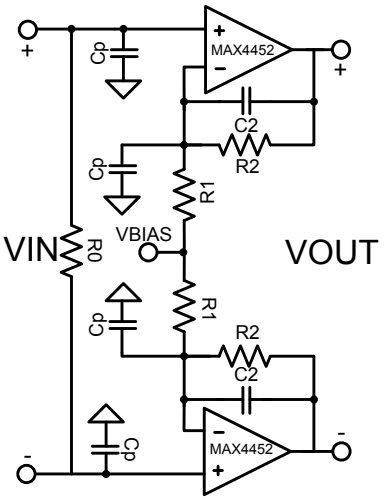


Figure 3.13: Alternate path baseband filter and ADC driver.

3.3 Alternate Path Postfilter Specification Requirements

As in the main path, the objective of the alternate path baseband postfiltering is to attenuate unwanted out-of-band signals to below the noise floor prior to aliasing at the continuous-time to discrete-time conversion. However, the typical approach to specify the main path postfilter must be modified to suit the alternate path for two reasons. First, the out-of-band signals are quantities derived, not taken directly, from the blocker profile in Fig. 2.10. That is, for each set of TX and RX LO frequencies $\{f_{TX}, f_{RX}\}$ there exist two CW blockers at frequencies $\{f_{CW}\}$ that cause IM3 products at the set of frequencies $\{f_{RX}, f_{TX}, f_{CW}, f_O\}$. Hence an additional step is required to compute the unwanted IM3 product magnitudes at these frequencies prior to downconversion. Extra care must be taken to perform downconversion, aliasing, and integration of error within the signal channel for each set $\{f_{RX}, f_{TX}, f_{CW}\}$ due to the fact that some undesired out-of-band IM3 products arising from different subsets of $\{f_{TX}, f_{CW}\}$ will land at the same frequency, causing an overestimate of the error within the signal channel if the subsets are not evaluated separately.

Secondly, the error floor requirement of the alternate path is defined with respect to $I_{ALT,PK}$. Hence, the new postfilter derivation procedure must also relate aliased out-of-band error to this quantity. Ideally, this filter requirement would be cast in terms of the previously defined term $IER_{ALT,PK}$. However, this assumes that error due to this effect is highest under peak blocking conditions, which occur for a specific set of blocker frequencies. While this is true for most error sources, error terms in this case depend on the relationship between blocker frequencies, RX LO frequency, and the ADC sampling rate. Hence, the worst case aliasing error $E^2_{ALTPF,MAX}$ may occur under conditions other than peak blocking and $IER_{ALT,PK}$ may underestimate peak error. Nevertheless, it is useful to relate the error due to this effect to $E_{MAIN,PK}$ through $I_{ALT,PK}$ in order to leverage the analysis in (2.13)-(2.14). This can be done by placing the bound (3.1) on the receiver error figure.

$$EF_{RX} = 10\log_{10}\left(\frac{E_{MAIN}^2 + E_{ALT}^2}{kTB}\right) \leq 10\log_{10}\left(\frac{E_{MAIN,PK}^2 + E_{CUB,PK}^2 + E_{ALTPF,MAX}^2}{kTB}\right) \tag{3.1}$$

Denoting a new quantity $I_{PK}ER_{ALTPF,MIN} = \frac{I_{ALT,PK}}{E_{ALTPF,MAX}}$ and noting that $I_{MAIN,PK} = I_{ALT,PK}$ at the adaptive equalizer subtraction node, the same substitution that was used to obtain (2.14) can be used to obtain the following bound:

$$EF_{RX} < 10\log_{10}\left(\frac{E_{MAIN,PK}^2}{kTB}\right) \leq 10\log_{10}\left(1 + \frac{IER_{MAIN,PK}^2}{IER_{CUB,PK}^2} + \frac{IER_{MAIN,PK}^2}{I_{PK}ER_{ALTPF,MIN}^2}\right) \tag{3.2}$$

Using the conditions described in Section 2.4.2, this implies $I_{PK}ER_{ALTPF,MIN} > 34$dB. Whether or not a postfilter meets this requirement can be determined with a nested sweep on $\{f_{RX}, f_{TX}, f_{CW}\}$. It can be shown for the proposed cubic term generator in Fig. 3.9 that the IM3 product magnitudes at $\{f_{RX}, f_{TX}, f_{CW}\}$, and $\{f_O\}$ when $f_{CW} > f_{TX}$ are proportional to $P_{CW}^2P_{TX}$, $P_{CW}^2P_{TX}$, $P_{CW}P_{TX}^2$, and $P_{CW}P_{TX}^2$, respectively.

The following steps are performed for each nested sweep iteration and are depicted in Fig. 3.14:

1. For each set $\{f_{RX}, f_{TX}, f_{CW}\}$ the blocker profile in Fig. 2.10 is accessed to determine P_{CW}, P_{TX}.

2. The IM3 product rms magnitudes at $\{f_{RX}, f_{TX}, f_{CW}, f_O\}$ are computed as described above.

3. The frequency domain spectrum in this case is downconverted to baseband by $\{f_{RX}\}$ and the proposed postfilter model is applied.

4. The baseband frequency domain spectrum is aliased by the discrete-time nature of the ADC sampling at F_S. The energy falling within the RX channel bandwidth is integrated and is used to divide $I_{PK,ALT}$. This quantity is then compared to a running minimum and if less the running minimum is updated.

The result of this procedure is $I_{PK}ER_{ALTPF,MIN}$. For this project, $F_{S,ADC}$=16.66MHz was chosen to allow the use of a power-efficient commercially available ADC. In this case, if the only postfiltering present were a mixer real pole with $f_{-3dB} = 1.5$MHz, then $I_{PK}ER_{ALTPF,MIN} = 27.18$dB. Adding another first-order pole with $f_{-3dB} = 8$MHz yields $I_{PK}ER_{ALTPF,MIN} = 46.21$dB, which meets the requirements of this design with 12dB of margin. In the event that a higher sampling rate can be used, such as 50MHz, if the only postfiltering present were a mixer pole with $f_{-3dB} = 1.5$MHz, then $I_{PK}ER_{ALTPF,MIN} = 37.43$dB. In this case, the required postfiltering is even simpler, constituting a minimal burden to power, area, and system complexity. To visually depict the worst-case undesired IM3 products relative to $I_{PK,ALT}$, the procedure described above may be terminated at Step 3, the result divided by $I_{PK,ALT}$, and a maximum operator instituted at each frequency bin, resulting in Fig. 3.15.

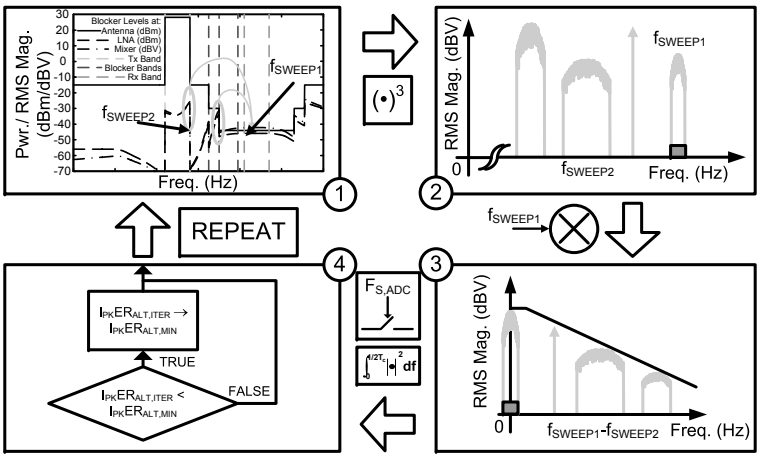


Figure 3.14: Depiction of blocker mapping algorithm.

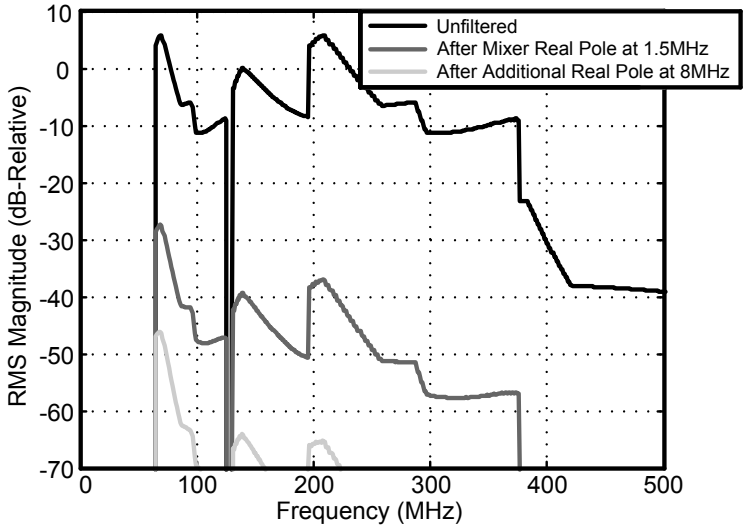


Figure 3.15: Worst-case undesired IM3 product baseband profile as function of frequency.

3.4 Experimental Setup

The complete assembled receiver up to the interface to the FPGA is shown in Fig. 3.16. Calibration is performed up to the leftmost vertical SMA connector which is coupled via 50Ω line to the external matching network at the LNA input. A close-up of the center of the gold-plated Rogers board is shown in Fig. 3.17 with a mounted and wirebonded chip. Immediately below the chip is the passive input matching network, including a Coilcraft 0603CS chip inductor and a ATC 100A series shunt capacitor. The die photo of this chip is shown in Fig. 3.18. A picture of the experimental setup is shown in Fig. 3.19 with the setup schematic shown in Fig. 3.20. In Fig. 3.20 the three input signal sources are seen added together with a Krytar Model 7020265 4-way power divider with the fourth input terminated with a passive 50Ω load.

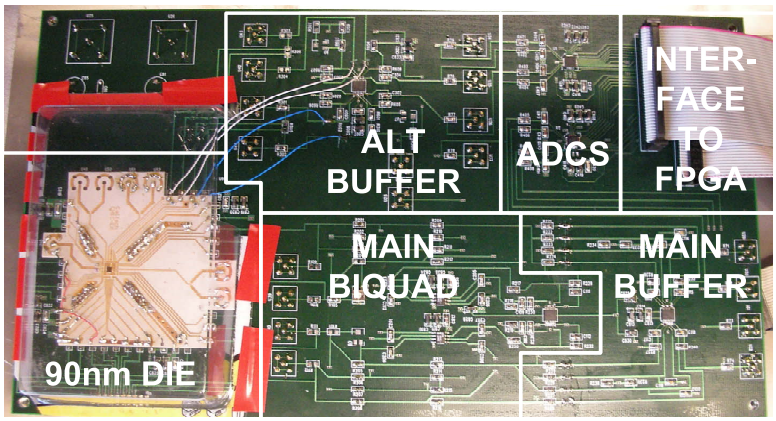


Figure 3.16: Experimental implementation of proposed receiver.

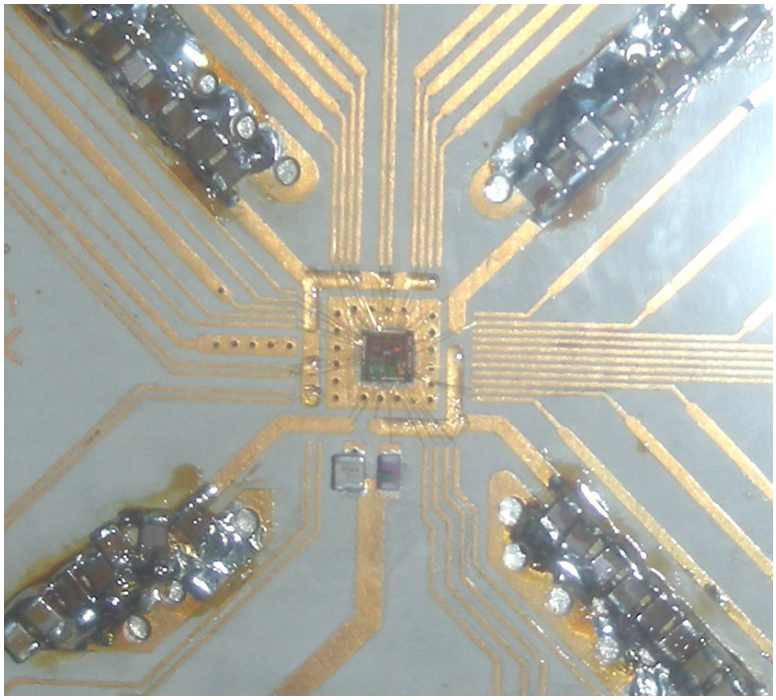


Figure 3.17: Zoomed-in view of bonded chip mounted on gold-plated Rogers board.

The output of the pattern generator supplying the TX leakage model signal is filtered by two Panasonic EFCH1950TCD1 SAW filters in series in order to attenuate the signal generator phase noise at the RX LO frequency. In the case where the signal generator is denoted by two names, the second of the two names denotes the instrument used specifically for the UMTS sensitivity test. The output of the power divider is coupled to a HP 8563E spectrum analyzer via a Krytar model 1851 directional coupler. The spectrum analyzer is used to verify that the proper signal characteristics are present at the receiver input at any given time. The directional coupler is connected to the receiver RF input via a Mini-circuits ZFBT-6GW bias tee. Calibration of the input network is performed by recording key signal power levels at the output of the cable leading from the bias tee to the receiver

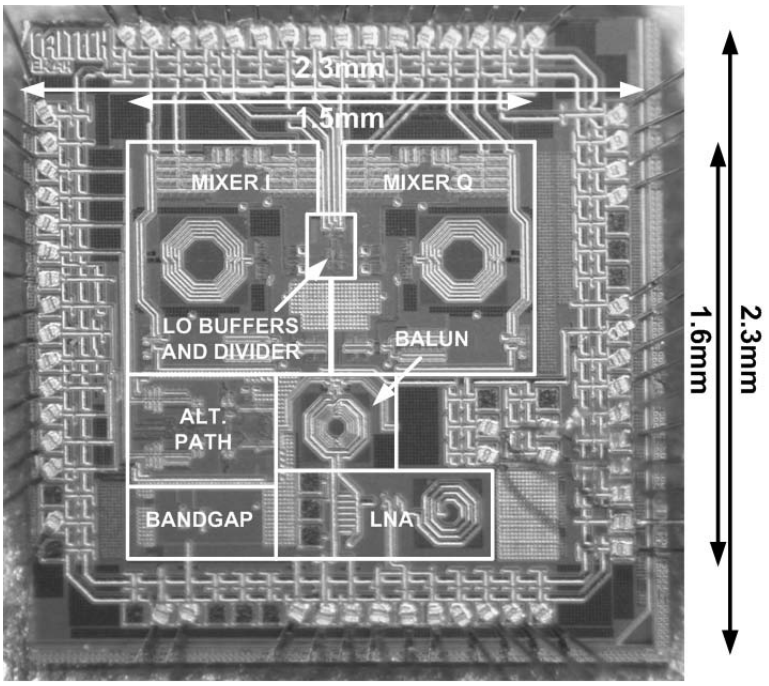


Figure 3.18: RF front end die photo.

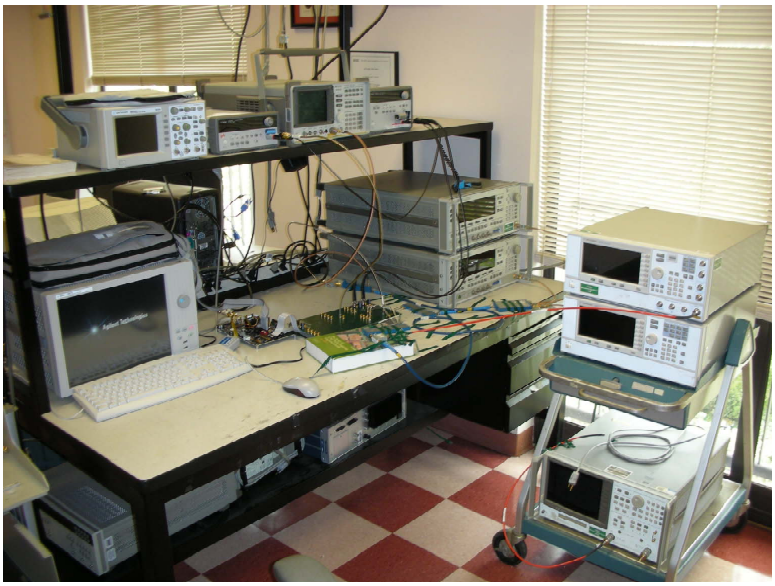


Figure 3.19: Experimental test setup picture.

SMA input with a HP 8487A power sensor and an E4418B Agilent power meter. The receiver noise figure test is performed using the Y-factor method with an HP 346C noise source connected to the input of the bias tee. In this case, the loss of the bias tee is measured and calibrated out of the noise figure and receiver gain measurement.

The 2x LO signal is applied via the use of a Mini-circuits ZFSC-2-10G signal splitter and two

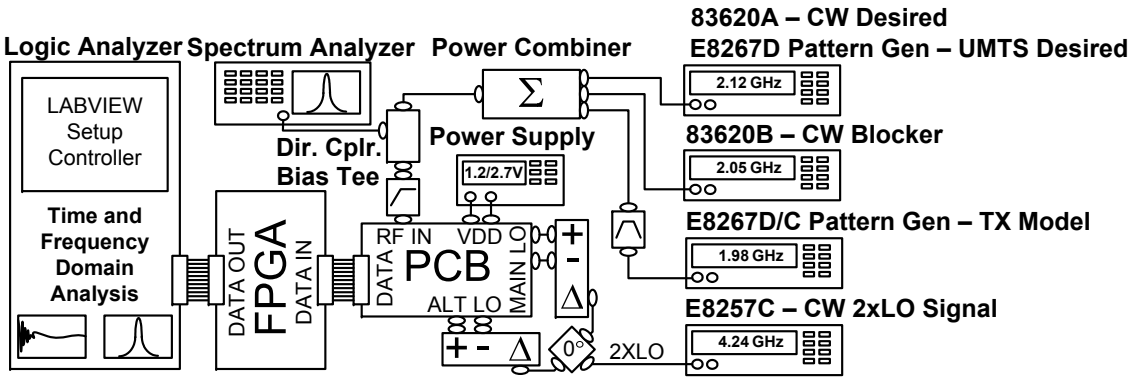


Figure 3.20: Experimental test setup simplified schematic.

Krytar 4010180 180° hybrid couplers. It is important to note that this setup, along with the cabling, introduces a considerable frequency-dependent phase mismatch between the main and alternate path LO signals that must be compensated for by the digital adaptive IM3 cancellation algorithm. The 1.2V/2.7V power supplies are applied directly to the chip and PCB discrete components. The ADC digital power supply is provided by the FPGA board via the bidirectional interface. This ensures that the ADC digital output logic levels are at the appropriate voltages to trigger the FPGA and vice versa.

3.5 Experimental Results

3.5.1 Receiver IIP3 Measurement Results

Fig. 3.21 shows the concept behind the modified two-tone IIP3 test used to evaluate the proposed receiver. The goal is to reproduce the TX leakage and CW blocker signals at the LNA input at several power levels (including the worst-case specified), at all 12 UMTS RX frequencies and to measure the output across the RX channel in each instance. The TX signal is a QPSK-modulated pseudorandom bit sequence at 3.84MSPS that is upsampled, passed through the UMTS-specified channel filter, and upconverted to 1.98GHz. The CW power is fixed at 8dB less than that of the TX. As this test is designed to predict the sensitivity of the receiver for UMTS communication, the RX signal power is set to zero when measuring the receiver output error. This methodology represents a valid proxy for predicting the receiver sensitivity because the total required RX signal level (including pilot channel, etc.) at sensitivity is below the system thermal noise floor. The receiver output error is input-referred by running the modified two-tone IIP3 test along with -90dBm and -97dBm CW RX signal models in order to determine small-signal gain under blocking and blocking+correction conditions, respectively. This is necessary because the receiver small-signal gain will change due to the total blocker input edging closer to the out-of-band 1-dB desensitization level. These CW RX signal models are swept across the RX band and the measured small-signal gain across all of the frequency points is rms-averaged in order to capture any frequency dependence in the effective baseband transfer function.

The resultant steady-state input-referred error over swept TX leakage power is shown in Fig.

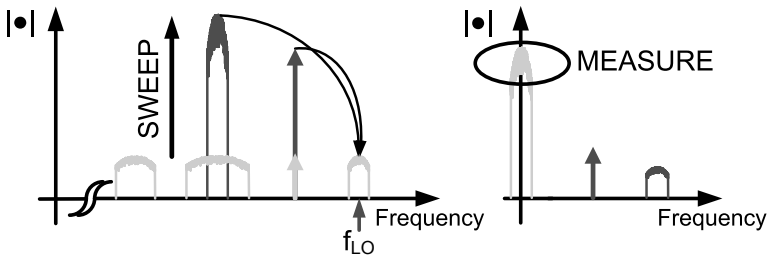


Figure 3.21: Modified two-tone test concept.

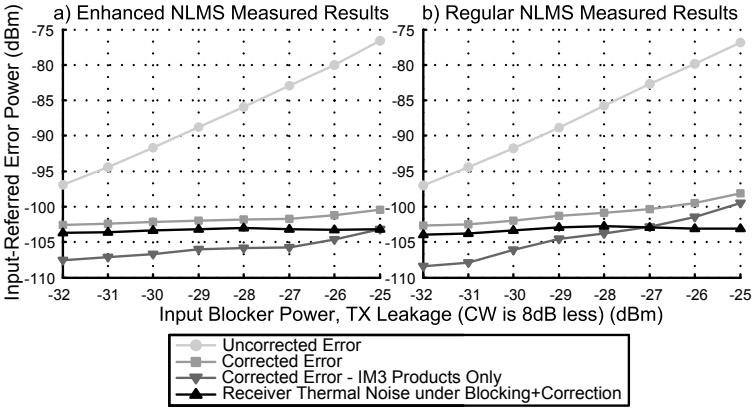


Figure 3.22: Measured results of modified two-tone test. a) For proposed enhanced complex NLMS architecture. b) For canonical complex NLMS architecture.

3.22a. Note that all plots shown and numbers reported depict the I/Q receiver channel with worst-case performance under worst-case specified blocking. The total input-referred error accounts for gain loss, thermal noise, and all IMD products. Removing the effects of main path thermal noise and IM2 products yields a lumped input-referred error quantity consisting of all other error sources. From this quantity, which is treated as residual IM3 error, a slope-of-3 line is extrapolated from the worst-case input blocker power to obtain an effective IIP3 metric. Other measurement results show that the measured IIP3 performance is limited by higher-order distortion products in the main path. Note that 50Ω kTB noise is removed from these plots and that the maximum total input-referred error in this regard, computed in Section 2.4 (-98.2dBm) under the worst-case scenario of -26dBm TX leakage input power, is met with 3dB of margin when correction is applied. The contribution of the baseband circuitry to the uncorrected out-of-band IIP3 has been measured and found to be negligible.

In Fig. 3.22b are the results of this same test using the NLMS algorithm without the enhanced degree of freedom. A phase mismatch of about 3° in the main path along with mismatch in the baseband frequency responses are responsible for the higher input-referred IM3 products. This confirms experimentally that this algorithm enhancement produces a noticeable performance improvement even for the moderate correction ratios required for this design.

Although the case in which the CW blocker frequency is less than the TX frequency does not require alternate path equalization for this duplexer referenced for this study, the CW blocker power

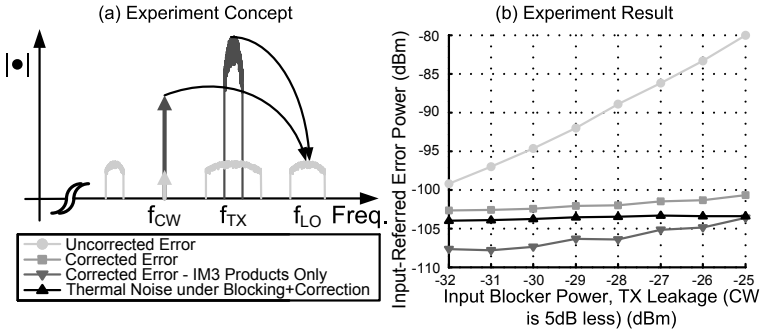


Figure 3.23: Measured results of modified two-tone test in which IM3 products contain squared TX leakage. a) Concept. b) Results.

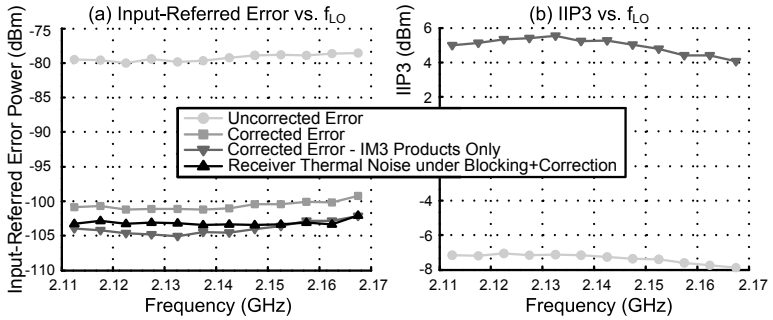


Figure 3.24: Measured results of modified two-tone test for -26dBm TX leakage, -34dBm CW blocker swept over LO frequency. a) Input-referred error vs. f_{LO}. b) IIP3 vs. f_{LO}.

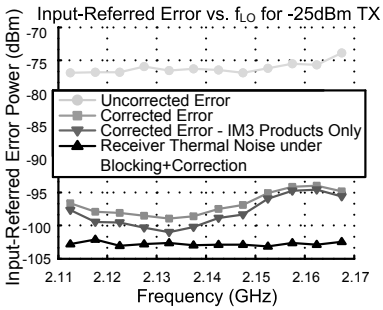


Figure 3.25: Measured results of modified two-tone test for -25dBm TX leakage, -33dBm CW blocker swept over LO frequency, input-referred error vs. f_{LO}.

was exaggerated far above specification in order to show that this case is covered by the proposed architecture as well. As described more fully in Chapter 4, this case generates an IM3 product in the main path consisting of a frequency-translated version of the squared modulated TX leakage. It is worthwhile to measure these results, as this condition is subject to an additional error term in the alternate path stemming from the fact that the "squared" TX leakage in this path is now TX leakage multiplied by a delayed version of itself. The results of this test are shown in Fig. 3.23 and show similar correction ratios to those seen in Fig. 3.22a.

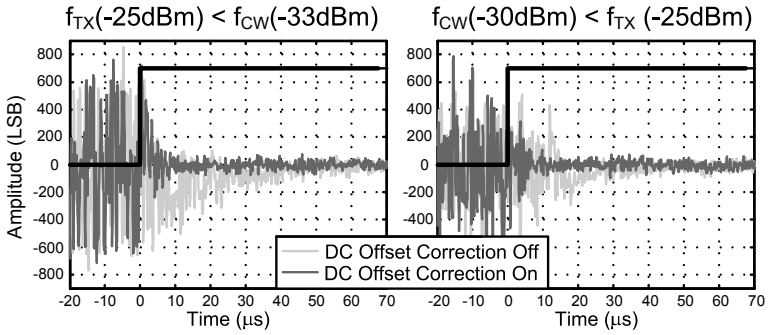


Figure 3.26: Measured convergence behavior of adaptive equalization algorithm.

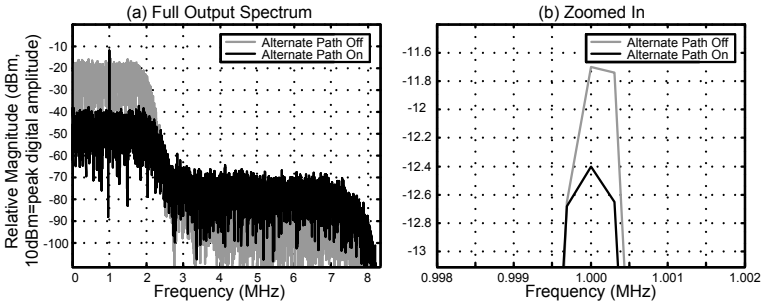


Figure 3.27: Digital spectrum analyzer measurement of power spectral density with and without correction.

The results of the experiment in Fig. 3.22a repeated across the UMTS RX band are shown in Fig. 3.24. The TX frequency is kept at 1.98GHz and the CW frequency adjusted such that the IM3 products fall within the RX channel around the LO frequency. This experiment is also performed for the case in which the adaptive equalizer is the canonical NLMS algorithm. It can be seen in Fig. 3.25 that the achieved IM3 cancellation varies greatly as a function of LO frequency and that the performance at about 2.16GHz is far worse than at 2.12GHz, where the results from Fig. 3.22b are obtained. This is due to the fact that the rotational mismatch ϕ_R varies as a function of LO frequency due to the fixed time delays through the LO cabling assembly. Convergence behavior of the proposed adaptive equalizer is shown in Fig. 3.26 for the case where $f_{LO} = 2.1225$GHz and the IM3 products land directly on the RX channel. It is seen that if dc offset correction is not applied prior to the enabling of the alternate path digital back end, the effective convergence time is dramatically extended, as expected. The frequency-domain measurement of the digital receiver output in the presence of a moderately large tonal desired signal is shown in Fig. 3.27 both with and without correction.

3.5.2 Receiver Sensitivity Measurement Results

Although the IIP3 test provides insight as to how nonlinear terms contribute to the input-referred error of the receiver, the actual performance specification that must be met is that of the sensitivity test. In this work, such a test is performed using a specification-equivalent UMTS 12.2kbps downlink

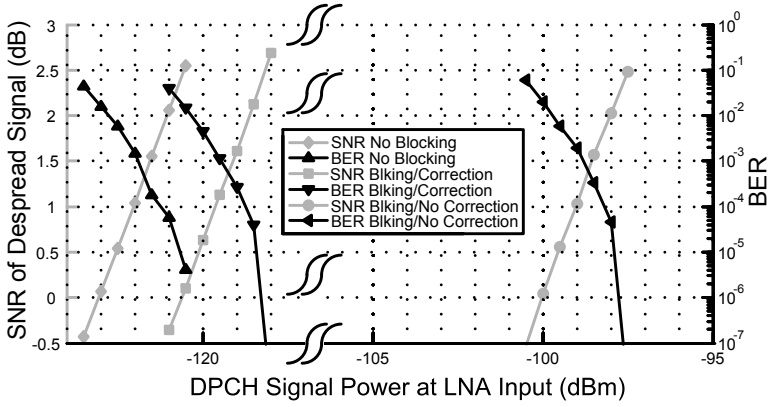


Figure 3.28: Measured receiver DPCH despread SNR and BER under sensitivity, sensitivity / blocking / correction and sensitivity / blocking / no correction, respectively.

reference measurement channel [6] with both I and Q channels active, and the results comparable to those in [9]. The theoretical relation (3.3) described in [9] relates the receiver sensitivity to noise figure, where S represents the receiver sensitivity in dBm and NF the receiver noise figure in dB.

$$SNR_{Despread} = S - (10\log_{10}(kTB) + 30) - NF + G - IL \tag{3.3}$$

This relation also holds for the error figure quantity introduced earlier in this chapter. In this test, the UMTS spreading gain G=21.1dB and the back end implementation loss IL is approximately 0dB. With L_{DUP}=1.8dB the receiver must achieve BER=10^{-3} for DPCH_E_C=-118.8dBm under typical conditions and DPCH_E_C=-115.8dBm under blocking. The results of the test for f_{LO}=2.1225GHz are shown in Fig. 3.28. The fact that BER=10^{-3} occurs with despread SNR≈1dB indicates that the MATLAB postprocessing of the physical receiver output was done correctly [9]. Each point in Fig. 3.28 represents the average of 4.88×10^5 bits (2000 data frames), which is sufficient to accurately resolve BER down to 10^{-4} [81]. The baseline sensitivity is -121.9dBm, 0.5dB greater than predicted by noise figure, with the discrepancy accounted for by unfiltered noise at frequencies greater than 1.92MHz. The sensitivity of the receiver under worst-case blocking and correction is -119.5dBm, 0.9dB greater than predicted by total input-referred error, with 0.6dB of this difference accounted for by noise at frequencies greater than 1.92MHz. This shows that the aforementioned effective IIP3 test is an accurate predictor of the actual sensitivity performance. Without correction, sensitivity significantly exceeds specification at -98.8dBm under worst-case blocking.

3.5.3 Alternate Path Measurement Results

The INR performance of the alternate path is measured and shown in Fig. 3.29. Performance under worst-case specified blocking conditions at f_{LO}=2.1225GHz is 31dB. Additional measurements suggest that higher order distortion products lower IER from INR by less than 1dB. Alternate path linear term feedthrough is also measured and referred to the main path input. The attenuation referred to the main path input is found to be greater than 46dB over all LO frequencies, indicating that the effect of these terms in the equalization process is negligible.

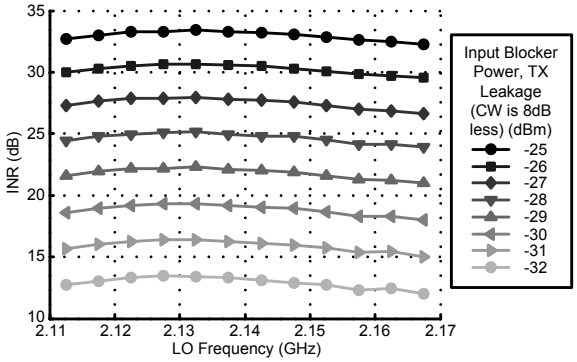


Figure 3.29: Measured INR performance of alternate path.

Parameter Measured at f_{LO}=2.1225GHz	Result
Active Analog Die Area	1.6mm x 1.5mm
Active Analog Die Alternate Path Area	0.5mm x 0.4mm
Analog Die Technology Node	130nm RF CMOS
Analog Die Supply Voltage	1.2V / 2.7V
Estimated Alternate Path Digital Die Area	0.42mm x 0.42mm
Digital Die Technology Node	90nm Bulk CMOS
Digital Die Supply Voltage	1.0V
Analog Die LNA + Main Path Current	28mA (1.2V)
Analog Die Alternate Path Current	6.7mA (1.2V)
Estimated Digital Alternate Path Current	5.6mA (1.0V)
Analog Die LNA + Main Path Dc Gain	30.5dB
Complete Main Path System Dc Gain to ADC Input	70.2dB
Input Return Loss (S11) 2.11 GHz-2.17 GHz	<-13dB
IIP2@1.98 GHz	+58dBm
Uncorrected IIP3 @ 1.98 GHz/2.05125 GHz	-7.1dBm
Effective IIP3 @ 1.98 GHz/2.05125 GHz	+5.3dBm
Out-of-band 1-dB desensitization point@1.98GHz	-19dBm
Analog Die LNA + Main Path NF	5.0dB
Complete LNA + Main Path System NF	5.5dB
Baseline DPCH Sensitivity	-121.9dBm
DPCH Sensitivity Under Blocking/Correction Off	-98.8dBm
DPCH Sensitivity Under Blocking/Correction On	-119.5dBm
Baseband Signal Measurement Bandwidth	10kHz-1.92MHz

Table 3.1: Receiver performance summary.

3.5.4 Additional Measurement Results

The performance summary for the system and front end is shown in Table 3.1. The power consumption estimate of the alternate path digital back end circuitry was obtained from switching statistics of a gate-level Verilog simulation referencing a 90nm CMOS process standard cell library.

Chapter 4

Multistage Cubic Term Generators for RF Receivers[1]

4.1 Prior Art in Cubic Term Generators

Cubic term generators have been extensively used in the past for the predistortion of nonlinear wireline, radio, and laser power amplifiers. Many early predistortion circuits utilized the nonlinear impedance characteristic of back-to-back diodes [83] [84] [85], while an active implementation of such a circuit with higher cubing gain using BJT devices was reported in [86]. Such architectures were not considered for the receiver in Chapter 3 as they are not easily implemented in CMOS processes. CMOS cubic term generators have also been presented that utilize the third order Taylor series coefficients of the MOSFET [87] [88] [89]. However, as the MOSFET is a square-law device the third-order Taylor series coefficient is relatively weak. Furthermore, the cubing operation heavily attenuates the IM3 output signal with respect to the noise of the IM3-producing devices. As will be discussed in Section 4.3 the noise of the cubic term generator is of prime importance in the receiver of Chap. 3 and renders the use of these architectures unattractive in such an application, though they may represent competitive design options when noise is not an issue.

Cubic term generators that realize polynomial function generation by using a cascade of multiplier operations represent a superior approach to cubic term generator design in this regard. Predistortion circuits using cascaded Gilbert cell multipliers have been implemented with discrete components [90], and in BiCMOS [91] and CMOS [92] processes. The authors of [93] propose the use of multipliers modified from [94] that are based on a sum-and-difference squaring technique which utilizes the square-law dependence of the MOS transconductance, but have not reported on a complete predistortion circuit to the knowledge of the authors of this dissertation. Architectures of this class potentially reap a noise benefit relative to those using explicit third-order nonlinearities, as the initial nonlinear products are only attenuated with respect to the noise of the nonlinear devices as the square of the input signal, not the cube.

[1] Portions of this material have been previously published in [82] and copyright is owned by IEEE.

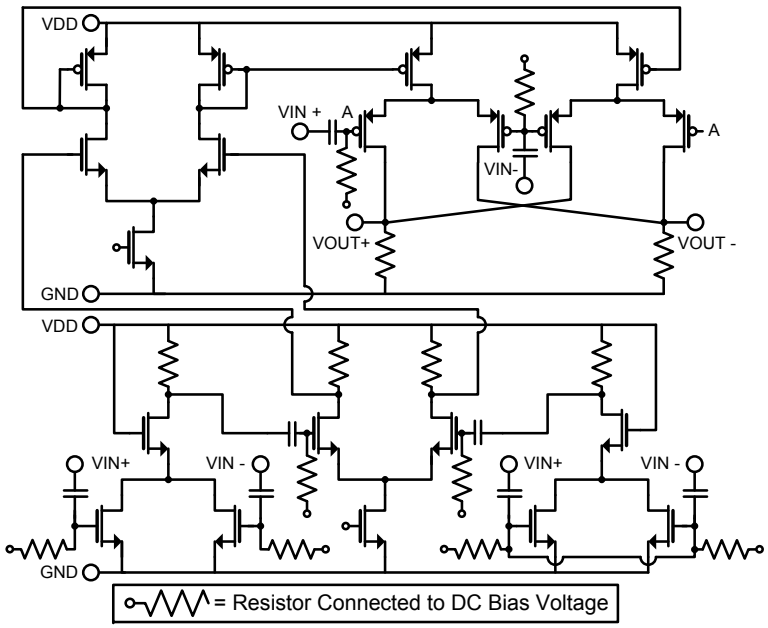


Figure 4.1: Schematic of cubic term generator used in Chapter 3.

4.2 Proposed Cubic Term Generator

The cubic term generator shown in Fig. 4.1 was implemented for the receiver in Chap. 3 and will be the focus of this chapter. Like [90]-[92] it utilizes multiple nonlinear operations to generate third-order distortion and hence much of the analysis in this chapter applies to these architectures as well. The first nonlinear operation in this case is performed by the simple MOS squaring transconductor in the lower-left hand corner of the schematic. This choice is made in order to avoid the generation of higher-order intermodulation products associated with the nonlinearity of the Gilbert cell current commutating devices. It also represents a simpler approach than that described in [94], using fewer noise-producing transistors (2 vs. 12).

As the MOS squaring circuit produces a single-ended output, it must be followed by an active balun to recast the signal differentially. One potential issue with this scheme is that the squaring circuit directly passes common-mode signal. If the balun negative terminal were grounded, the common-mode signal would be recast differentially as well and would propagate through the remainder of the circuit. In order to provide some measure of common-mode rejection, a dummy squaring circuit is added to the negative terminal of the balun. With the gate terminals of the dummy squaring circuit tied together, this circuit only passes common-mode signal. Hence, the common-mode signal is rejected by the CMRR of the balun and subsequent gain circuits. The use of a balun does not constitute an extra burden on the circuit as it also functions as a gain stage to suppress the effective noise contributions of subsequent circuitry.

The final multiplication of the cubic term generator is performed by a Gilbert cell multiplier. In this case, the nonlinearity of the current commutating devices can be improved at the expense of gain by increasing the multiplying device overdrive voltages or by attenuating the RF signal at this point via capacitive division. The lost gain can then be made up earlier in the circuit. The circuit

as implemented in Fig. 4.1 is somewhat power-inefficient due to the voltage output at RF. This was done for testing purposes, but in a commercial implementation the IM3 reference signal would be passed to a mixer switching pair in the current domain.

4.3 Requirements of Cubic Term Generators for RF Receivers

The requirements discussed in [83]-[92] are notable in that noise is not considered as a relevant design parameter. This is because the large signals processed by the predistorter are in fact desired signals and are well above the noise floor. This is not the case in the system presented in Chapter 3 in which the IM3 producing signals are undesired and the desired signal is buried underneath IM3 products over an order of magnitude greater. In this case, any noise present in the cubic term generator, even if it is an order of magnitude less than the IM3 products, can add significantly to the noise floor of the receiver when the alternate path is enabled. Likewise, higher-order distortion terms should be taken into account. Although they are typically much smaller than the IM3 products, they may still render the performance of the enhanced receiver inadequate to meet specification. Linear term feedthrough of desired signal must also be suppressed, as desired signal at the reference input of the equalizer will be treated as error by the adaptive algorithm. As the algorithm functions so as to minimize mean squared error, it will attempt to strike a balance between eliminating the IM3 products and desired signal, reducing small-signal gain, IM3 cancellation, or both.

The relationship between signal and noise is typically quantified using an SNR metric. Since IM3 products are not desired signal, however, we denote the ratio of these terms to the noise integrated across the desired signal channel as INR. Generalizing this metric to other forms of error, including noise and other distortion terms, results in a quantity termed IER. Quantifying the performance impact of the alternate path on the complete receiver is easily done using these metrics, as shown previously in Section 2.3.

4.4 Usefulness and Sufficiency of Internal Bandwidth Limitations in Reconstructing IM3 Products

It was mentioned in Section 4.1 that cubic term generator architectures made of multiple nonlinear operations represent a potential noise advantage over circuits which generate IM3 products using a single cubic nonlinearity due to the fact that the initial nonlinear products are only attenuated with respect to the noise floor as the square of the input signal rather than the cube. The reason for this is that the noise contributions of subsequent multiplier circuits can be made negligible by adding gain immediately after the initial squaring, as depicted in Fig. 4.2.

Adding significant amounts of gain in between nonlinear operations can be problematic if all IM2 products after the initial squaring were required to be faithfully retained. For example, most of the relevant blockers in UMTS are clustered around 2GHz, as shown in Chap. 3, and broadband circuitry operating up to 4GHz is required to retain all second-order distortion products. Fortunately, this is not the case. Figure 4.3 depicts the response of a multistage 3rd-order distortion generator with different interstage filtering schemes, using a two-tone test to represent a more arbitrary modulated signal. It can be seen in the second column that if the high-frequency second-order distortion

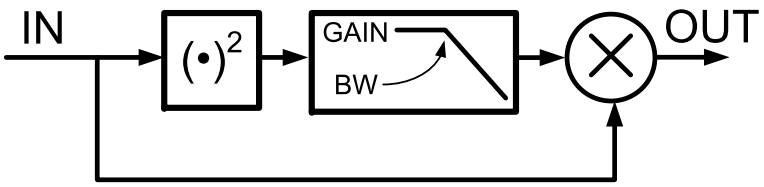


Figure 4.2: Implementation of gain between nonlinear operations in multistage cubic term generator.

products are removed prior to the final multiplication, the relationships between the resultant third-order IMD products are the same as if the high-frequency second-order distortion products were still present. In order to remove these higher-frequency terms without adding memory in the form of an interstage filter, previous polynomial predistorters generated only the low-frequency second-order IMD products by using polyphase splitters [91] and separate I and Q squaring circuits [92] for the initial nonlinear operation.

In the receiver of Chap. 3, it is difficult to separate the I and Q terms of a broadband set of blockers from the incoming RF signal in order to utilize separate I and Q squaring circuits. However, using an interstage filter does not cause significant problems in the way of memory effects due to the fact that the double-sided bandwidth of the QPSK blocker is only about 4MHz while the high-frequency second-order harmonics lie at 4GHz. In this case, the frequency response can be made roughly constant at low frequencies with a cutoff pole still low enough to effect substantial attenuation at high frequencies. Although the filtered high-frequency second-order harmonics may still be significant enough to contribute a small amount of signal energy to the final IM3 products, the narrowband approximation will hold, and the effects of a nonconstant frequency response on these terms prior to the final multiplication will be negligible.

It also turns out that unlike in transmit predistortion circuits, retaining the IM2 products that fall around dc is not required once higher-frequency IM2 products and harmonics are strongly attenuated. This is because in the receiver, the precise relation of all of the third-order distortion components is unimportant; the only requirement is that the proper IM3 products occur around f_{LO}. This difference is depicted in the third column of Fig. 4.3. In the second column, all of the third-order distortion products have the proper relative amplitude and can be used to cancel all of the terms arising from third-order nonlinearities in a predistortion scheme. In the third column, the inner and outer IM3 products no longer have the proper relationship to allow concurrent cancellation of all terms. Although such a reference signal could function in order to cancel only the outer IM3 products in a predistorter two-tone proxy linearity test, for a practical modulated signal all of the third-order products are important and an improper relation between the terms would manifest itself negatively in ACPR and EVM tests.

Therefore, only IM2 products around the beat-frequency in a two-tone test need to be retained, as depicted in Fig. 4.4. This is a significant advantage in the circuit of Fig. 4.1, as the peak IM2 products around dc can be larger than the peak IM2 products around the beat frequency, dictating the compression point of the cubic term generator without contributing to the desired IM3 output products. Furthermore, the reduction in amplitude of the two center IM3 products relaxes the requirements on the anti-aliasing filter in the alternate path, as the initial amplitudes of these undesired IM3 products are less than they would have been without the interstage high-pass filter.

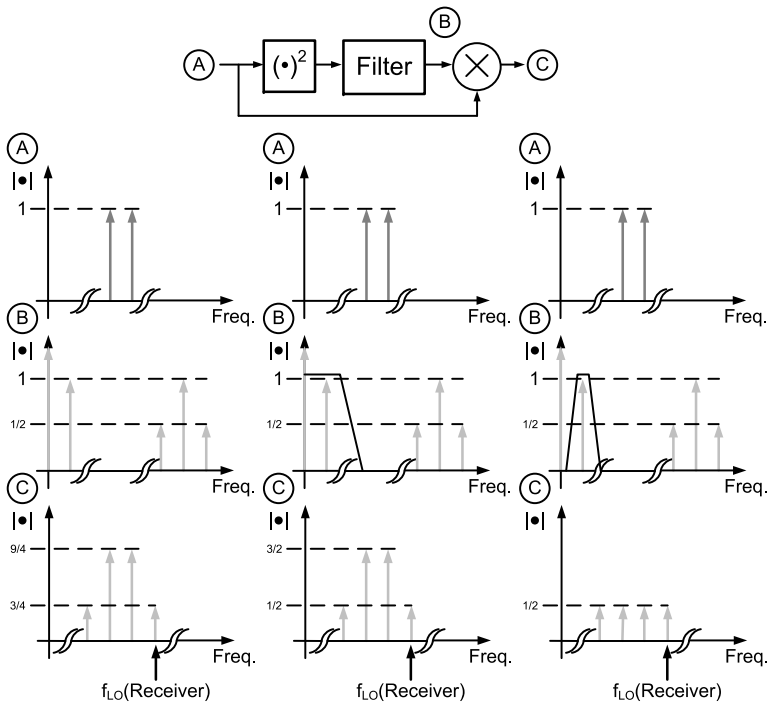


Figure 4.3: Effects of baseband frequency response on IM3 products in a multistage cubic term generator for a two-tone test.

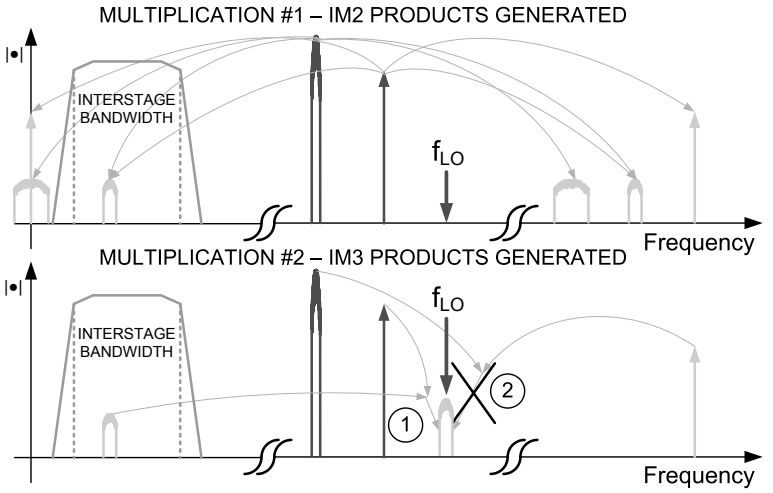


Figure 4.4: Frequency domain depiction of proposed cubic term generator internal bandwidth limitation.

For FDD UMTS Region 1, it can be calculated that the required interstage bandwidth, denoted f_{IS}, is 185MHz centered around 157.5MHz. As shown in Fig. 4.3, the relationship between the four output IM3 products in this case is not the same as that in the output of a true cubing circuit. As a result, the circuit of Fig. 4.1 must be described as a "cubic term generator".

Although the linearity requirements of the FDD UMTS Region 1 specification [6] are dictated

by nonlinear interactions between only two signals (TX leakage and a CW blocker), the cubic term generator should also be able to handle the more general case of problematic IM3 products arising as a result of three arbitrary bandpass signals. A quick proof is presented below that shows that in such a situation, signal around only one of two beat frequencies needs to be retained in order to reproduce the proper IM3 products. The proof sketch is as follows:

1. Consider 3 modulated signals, each at different frequencies, such that an IM3 product falls at $f_{LO} = f_A + f_B - f_C$.
2. Signal content at only 3 out of the 10 possible IM2 frequencies might possibly need to be retained: $f_A + f_B$, $|f_A - f_C|$, and $|f_B - f_C|$.
3. The goal is to show that the IM3 products produced by multiplying each of these IM2 components by the original set of three modulated signals are the same. In this case, signal content near only 1 of these 3 relevant IM2 frequencies needs to be retained to properly reproduce the IM3 products at f_{LO}.
4. In order to do this, one of the two beat frequencies is chosen and the IM2 products at that frequency are computed.
5. Next, the signal content at this frequency is multiplied by the signal not involved in the aforementioned IM2 products and the relevant IM3 signal content at f_{LO} is retained.
6. Next, the IM2 products at $f_A + f_B$ are computed and multiplied by the signal at f_C. The resultant terms at f_{LO} should equal the terms computed in step 5 if these high-frequency terms can be discarded.
7. Finally, if only one of the two beat frequency terms is necessary, then the letters A and B can be interchanged in the expression obtained in steps 5 and 6 with no change in the expression.
8. Note that the analysis shown below is indifferent as to whether difference terms such as $f_A - f_C$ are greater than or less than zero. Hence, this analysis is general for the various permutations of relative blocker frequency locations.

Elaboration on this sketch commences at step 4 and the IM2 terms at $f_A - f_C$ are computed. Recall that the bandpass signals may be expressed as:

$$s_A(t) = s_{AI}(t)\cos(2\pi f_A t) + s_{AQ}(t)\sin(2\pi f_A t) \tag{4.1}$$

$$s_C(t) = s_{CI}(t)\cos(2\pi f_C t) + s_{CQ}(t)\sin(2\pi f_C t) \tag{4.2}$$

Multiplying these two signals together and applying the relevant trigonometric identities yields four separate terms:

$$\begin{aligned} &s_A(t)s_C(t) = \\ &\tfrac{1}{2}\left[s_{AI}(t)s_{CI}(t) + s_{AQ}(t)s_{CQ}(t)\right]\cos(2\pi(f_A - f_C)t) + \\ &\tfrac{1}{2}\left[s_{AQ}(t)s_{CI}(t) - s_{AI}(t)s_{CQ}(t)\right]\sin(2\pi(f_A - f_C)t) + \\ &\tfrac{1}{2}\left[s_{AI}(t)s_{CI}(t) - s_{AQ}(t)s_{CQ}(t)\right]\cos(2\pi(f_A + f_C)t) + \\ &\tfrac{1}{2}\left[s_{AI}(t)s_{CQ}(t) + s_{AQ}(t)s_{CI}(t)\right]\sin(2\pi(f_A + f_C)t) \end{aligned} \tag{4.3}$$

Taking the top two terms and multiplying by $s_B(t)$ yields a set of IM3 products at f_{LO}:

$$\begin{aligned}
&\Big[s_B(t)\left[s_A(t)s_C(t)\right]_{f_A-f_C}\Big]_{f_A+f_B-f_C} = \\
&\tfrac{1}{4}\left[s_{AI}(t)s_{CI}(t)+s_{AQ}(t)s_{CQ}(t)\right]s_{BI}(t)\cos\left(2\pi f_{LO}t\right)+ \\
&\tfrac{1}{4}\left[s_{AI}(t)s_{CQ}(t)-s_{AQ}(t)s_{CI}(t)\right]s_{BQ}(t)\cos\left(2\pi f_{LO}t\right)+ \\
&\tfrac{1}{4}\left[s_{AI}(t)s_{CI}(t)+s_{AQ}(t)s_{CQ}(t)\right]s_{BQ}(t)\sin\left(2\pi f_{LO}t\right)+ \\
&\tfrac{1}{4}\left[s_{AQ}(t)s_{CI}(t)-s_{AI}(t)s_{CQ}(t)\right]s_{BI}(t)\sin\left(2\pi f_{LO}t\right)
\end{aligned} \tag{4.4}$$

Moving on to step 6, the last two terms of (4.3) are retained and C is replaced with B to obtain the relevant high-frequency IM2 products:

$$\begin{aligned}
&\left[s_A(t)s_B(t)\right]_{f_A+f_B} = \\
&\tfrac{1}{2}\left[s_{AI}(t)s_{BI}(t)-s_{AQ}(t)s_{BQ}(t)\right]\cos\left(2\pi(f_A+f_B)t\right)+ \\
&\tfrac{1}{2}\left[s_{AI}(t)s_{BQ}(t)+s_{AQ}(t)s_{BI}(t)\right]\sin\left(2\pi(f_A+f_B)t\right)
\end{aligned} \tag{4.5}$$

Multiplying by $s_C(t)$ and taking the terms at f_{LO} yields another set of IM3 products at f_{LO}:

$$\begin{aligned}
&\Big[s_C(t)\left[s_A(t)s_B(t)\right]_{f_A+f_B}\Big]_{f_A+f_B-f_C} = \\
&\tfrac{1}{4}\left[s_{AI}(t)s_{BI}(t)-s_{AQ}(t)s_{BQ}(t)\right]s_{CI}(t)\cos\left(2\pi f_{LO}t\right)+ \\
&\tfrac{1}{4}\left[s_{AI}(t)s_{BQ}(t)+s_{AQ}(t)s_{BI}(t)\right]s_{CQ}(t)\cos\left(2\pi f_{LO}t\right)+ \\
&\tfrac{1}{4}\left[s_{AQ}(t)s_{BQ}(t)-s_{AI}(t)s_{BI}(t)\right]s_{CQ}(t)\sin\left(2\pi f_{LO}t\right)+ \\
&\tfrac{1}{4}\left[s_{AQ}(t)s_{BI}(t)+s_{AI}(t)s_{BQ}(t)\right]s_{CI}(t)\sin\left(2\pi f_{LO}t\right)
\end{aligned} \tag{4.6}$$

Rearranging (4.4) and (4.6) shows that they are the same expression:

$$\begin{aligned}
&\left[s_A(t)s_B(t)s_C(t)\right]_{f_A+f_B-f_C} = \\
&\tfrac{1}{4}\begin{bmatrix} s_{AI}(t)s_{BI}(t)s_{CI}(t)-s_{AQ}(t)s_{BQ}(t)s_{CI}(t)+ \\ s_{AI}(t)s_{BQ}(t)s_{CQ}(t)+s_{AQ}(t)s_{BI}(t)s_{CQ}(t) \end{bmatrix}\cos\left(2\pi f_{LO}t\right)+ \\
&\tfrac{1}{4}\begin{bmatrix} s_{AQ}(t)s_{BQ}(t)s_{CQ}(t)-s_{AI}(t)s_{BI}(t)s_{CQ}(t)+ \\ s_{AQ}(t)s_{BI}(t)s_{CI}(t)+s_{AI}(t)s_{BQ}(t)s_{CI}(t) \end{bmatrix}\sin\left(2\pi f_{LO}t\right)
\end{aligned} \tag{4.7}$$

Since the letters A and B can be interchanged in (4.7) with no resultant change to the expression, it does not matter which of the two beat frequency terms is originally retained. This satisfies the requirements of step 7 and hence the proof is complete.

4.5 Dynamic Range Calculations and Simulation Results

As mentioned in Sec. 4.3, the target value of IER_{CUB} under peak blocking is 34dB for a UMTS receiver with an uncorrected IIP3 of -9dBm. In a conservative design, the INR should dominate the IER, as accurately predicting higher-order nonlinear terms requires precise device modeling at high frequencies, which cannot always be guaranteed. To determine the INR of the proposed cubic term generator architecture shown in Fig. 4.1, the INR of the initial squaring circuit is analyzed

for the effective two-tone test imposed by the UMTS Region 1 out-of-band blocking requirement by referring both signal and noise quantities to the input of the active balun. The key assumption here is that the noise of these input devices dominates that of the entire circuit under the conditions of interest, namely under peak blocking.

Modeling the transconductances of the MOS devices as a Taylor series (4.8), the beat frequency IM2 product voltage at the balun input can be expressed as (4.9) where A_X represents the differential amplitude of sinusoidal blocker #X and R represents the load resistor of the circuit input stage. Here, the UMTS TX leakage is modeled as sinusoidal blocker #2.

$$I_d(v_{in}) = a_0 + a_1 v_{in} + a_2 v_{in}^2 + a_3 v_{in}^3 + O(4) \tag{4.8}$$

$$IM2_{RMS}^2 = \tfrac{1}{8}(a_2 A_1 A_2 R)^2 \tag{4.9}$$

The value of a_2 can be found via Taylor series expansion of the MOS short-channel current equation (4.10) [95, pg. 589]. To simplify notation, the equalities in (4.11) have been substituted.

$$I_D(x) = KB(V_{OD} + x)^2 (V_{OD} + x + B)^{-1} \tag{4.10}$$

$$V_{OD} = V_{GS} - V_{TH};\;\; K = \tfrac{1}{2}\tfrac{W}{L}\mu_n C_{ox};\;\; B = \frac{1}{\theta + \frac{\mu_0}{2L_{eff} v_{sat}}} \tag{4.11}$$

Taking repeated derivatives of (4.10) yields (4.12) and (4.13).

$$g_m = a_1 = \left.\frac{\partial I_d}{\partial x}\right|_{x=0} = \frac{I_D}{V_{OD}} \cdot \frac{V_{OD} + 2B}{V_{OD} + B} \tag{4.12}$$

$$\begin{aligned} a_2 &= \frac{1}{2!}\left.\frac{\partial^2 I_d}{\partial x^2}\right|_{x=0} = \frac{KB}{(V_{OD}+B)}\left(1 - \frac{V_{OD}}{V_{OD}+B}\right)^2 \\ &= \frac{KB^3}{(V_{OD}+B)^3} = \frac{I_D B^2}{V_{OD}^2 (V_{OD}+B)^2} \end{aligned} \tag{4.13}$$

The PSD of the thermal noise of the input devices is evaluated using the MOS long-channel noise model. This was done as the simulation models available for this work permitted a more realistic and constant value of the noise γ parameter over V_{OD} than did the short-channel noise model. In this case, the PSD at the input of the balun is given by (4.14), where g_m is the transconductance of one of the input devices and k is Boltzmann's constant. Note that the dummy input devices contribute to this quantity as well.

$$N_{RMS}^2(\Delta f) = 4kTR(4\gamma g_m R + 2)\Delta f \tag{4.14}$$

As shown in Fig. 4.5, the thermal noise of the input devices referred to the input of the Gilbert multiplier has a bandpass characteristic due to the frequency response discussed in the previous section. Clearly noise around the desired beat frequency IM2 products is upconverted to the IM3 products at RF by one of the blockers. In the scenario described at the beginning of this analysis,

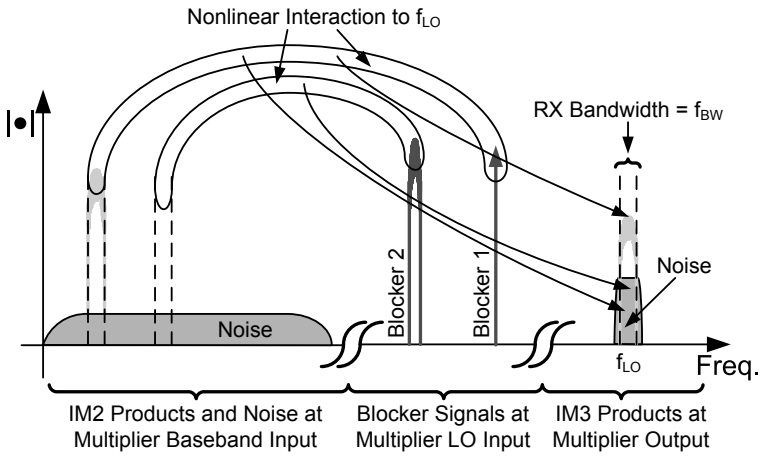


Figure 4.5: Frequency translation of bandpass noise by multiple blocker signals to desired signal level.

this is done by blocker #1. However, any blocker signals falling within $f_{LO} - f_{IS}, f_{LO} + f_{IS}$ will also contribute bandpass noise to the final INR quantity, even if they do not participate in the generation of IM3 products. To refer this noise back to the input of the balun stage, it suffices to realize that, with local bandpass gain G_X, the noise PSD upconverted and contributed by blocker X is given by $G_X^2 A_X^2 N_{RMS}^2(\Delta f)$. Integrating over the double-sided desired signal bandwidth f_{BW}=3.84MHz yields the total noise power. To refer back to the input of the balun, each noise contribution must be divided by $G_1^2 A_1^2$. In this case, the INR is given by (4.15), where M represents the total number of blockers in the relevant frequency range.

$$INR = \frac{I_D B^4 A_1^4 A_2^2}{64kTf_{BW}\left(\sum_{i=1}^{M}\left(\frac{G_i}{G_1}A_i\right)^2\right)V_{OD}^3(V_{OD}+B)^3} \cdot \frac{1}{2\gamma(V_{OD}+2B)+\frac{1}{R}\frac{V_{OD}(V_{OD}+B)}{I_D}} \tag{4.15}$$

Although the summation in (4.15) seems to suggest that INR can be arbitrarily low depending on how many blockers exist, in practice the total blocker power is bounded and for a real-world deployment would have to be estimated by field measurements in the relevant frequency range. For the UMTS specification tests of interest in Chap. 3, however, only the TX leakage and a CW blocker are present at any given time and M=2. From (4.15), it would also seem that the optimal design strategy would be to set V_{OD} as low as possible by grounding the input device gates and increasing device width until enough signal is obtained. However, it can be shown [96] that as V_{OD} approaches 0V well into the weak inversion region that the ratio of IM2/IM4 reaches a maximum after which it rapidly falls off. This quantity represents a practical lower bound on V_{OD}.

The only remaining task is to find the relationship between the current of each input device and the current consumed by the remainder of the circuit. A simple way to do this is to render the current draw of each gain, balun, and Gilbert cell stage equal. To adequately suppress the noise of the Gilbert cell in this case requires multiple gain stages, as shown in the circuit of Fig. 4.6, which is used to compare simulation vs. calculation. For simulations, two gain stages in addition to the active balun were utilized to adequately suppress the noise of the balun and Gilbert cell devices,

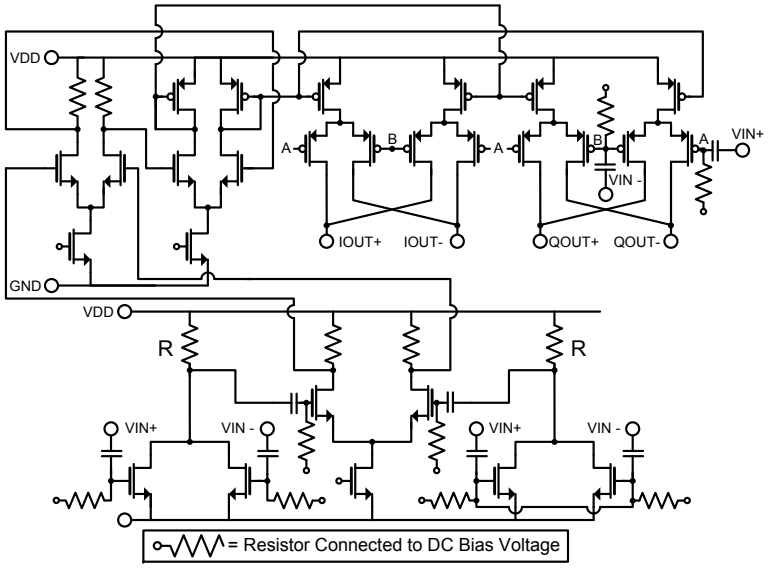


Figure 4.6: Schematic of cubic term generator used for simulation.

resulting in $I_{D,TOT} = 24I_D$.

To confirm the calculations, they are compared with simulation in Fig. 4.7 for L=0.6μm, $I_{D,TOT}$=1mA, T=298K and swept V_{OD} in a deep submicron process. A relatively large value of L is targeted in order to improve matching and to minimize linear term feedthrough. Values for γ and B were extracted from simulated device models and found to be 0.55 and 1.08V, respectively. The value of γ is consistent with other recent reports of submicron CMOS devices [97] [98], but a bit lower, likely due to device operation near weak inversion and with relatively low V_{DS}. G_2/G_1 was simulated and found to be 0.8. Fig. 4.8 shows the results of simulation vs. calculation for V_{OD}=132mV, $I_{D,TOT}$=1mA, T=298K, and for swept A_2/A_1. In both Figs. 4.7b. and 4.8b. the simulated INR of both the input stage and total circuit is shown. The data points at A_2=-26dBV (TX leakage model) in Fig. 4.7 represent peak blocking for the receiver of Chap. 3 and it can be seen here that the input stage indeed dominates the INR performance of the circuit. Furthermore, for most presented values of V_{OD}, the circuit meets the target INR value of 34dB for 1mA of current. As this current draw is much less than the 28mA consumed by the original receiver path presented in Chap. 3, it represents a power-efficient design option for a portion of the alternate path enhancement circuitry.

4.6 Comparison to Single-Stage Cubic Term Generators

Although it has been qualitatively argued in Sec. 4.5 that for purposes of IM3-to-noise ratio a multistage cubic term generator is superior to a single-stage cubic term generator, it is still instructive to consider a quantitative example to appreciate the extent to which this claim is true under reasonable design constraints. Consider the circuit described in [88], shown in Fig. 4.9. It can be shown via Taylor series analysis that the total output of the single-ended composite transconductor in Fig. 4.9 is given by (4.16) where a_0, a_2, etc are the Taylor series coefficients of the MOS saturation

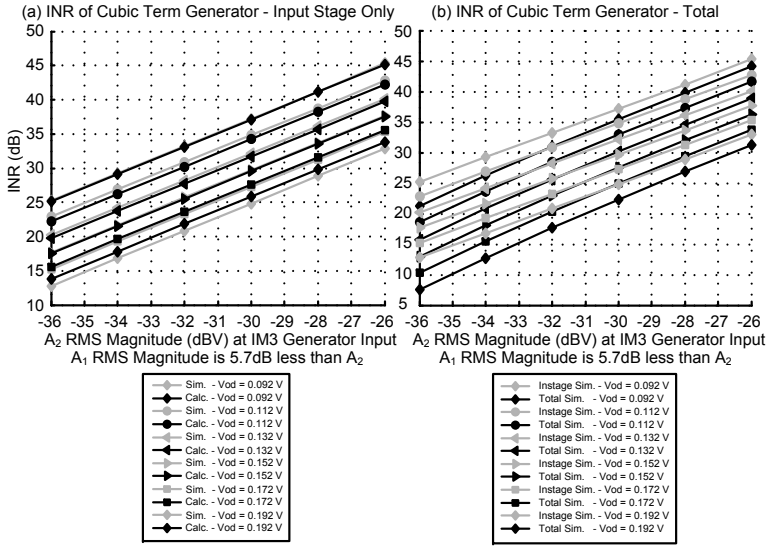


Figure 4.7: a) Calculated vs. simulated INR due to input stage components swept over V_{OD}. b) Simulated INR, input stage devices (Instage) and all devices (Total), swept over V_{OD}.

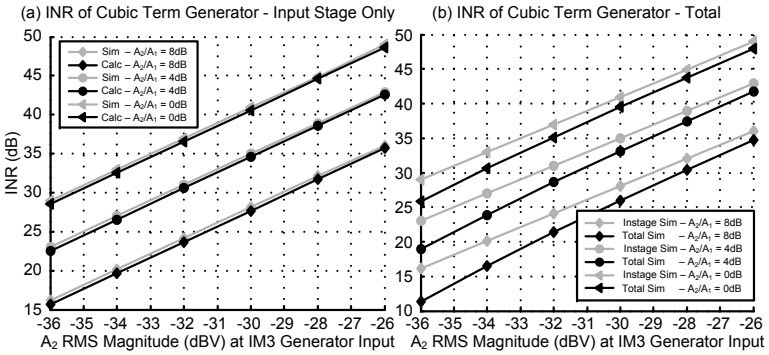


Figure 4.8: a) Calculated vs. simulated INR due to input stage components swept over A_2/A_1 ratio. b) Simulated INR, input stage devices (Instage) and all devices (Total), swept over A_2/A_1 ratio.

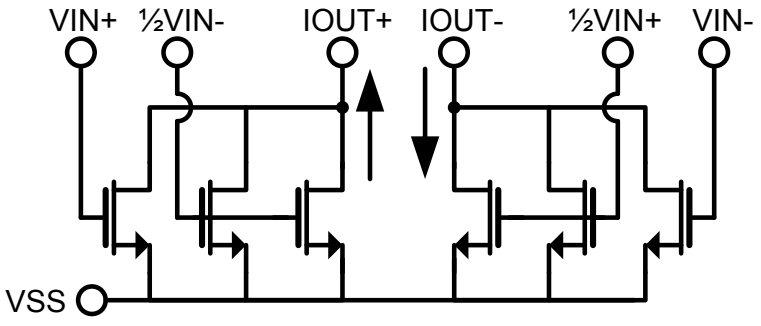


Figure 4.9: Schematic of cubing circuit from [88].

V-I relation.

$$I_{OUT+}(v_{in,SE}) = 3a_0 + \tfrac{3}{2}a_2 v_{in,SE}^2 + \tfrac{3}{4}a_3 v_{in,SE}^3 + O(4) \tag{4.16}$$

The resultant differential output in the absence of differential mismatch is:

$$\begin{aligned} I_{OUT}(v_{in,SE}) &= \tfrac{6}{4}a_3 v_{in,SE}^3 + O(5) \\ I_{OUT}(v_{in,DE}) &= \tfrac{3}{16}a_3 v_{in,DE}^3 + O(5) \end{aligned} \tag{4.17}$$

Where the third-order Taylor series coefficient is given by:

$$\begin{aligned} a_3 = \frac{1}{3!}\frac{\partial^3 I_D}{\partial x^3}\bigg|_{x=0} &= \frac{-KB}{(V_{OD}+B)^2}\left(1 - \frac{V_{OD}}{V_{OD}+B}\right)^2 \\ = \frac{-KB^3}{(V_{OD}+B)^4} &= \frac{-I_D B^2}{V_{OD}^2(V_{OD}+B)^3} \end{aligned} \tag{4.18}$$

In this case, the output squared RMS IM3 current at f_{LO}=2f_1-f_2 in response to the same two-tone blocking scenario described in Section 4.5 is given by:

$$\begin{aligned} IM3_{RMS}^2 &= \left(\tfrac{3}{16}a_3 \tfrac{3}{4}A_1^2 A_2 \tfrac{1}{\sqrt{2}}\right)^2 \\ &\approx \left(0.1 a_3 A_1^2 A_2\right)^2 \end{aligned} \tag{4.19}$$

It can likewise be shown that the RMS noise current power at the output of this cubing circuit is given by:

$$N_{RMS}^2 = 24kT\gamma g_m f_{BW} = 24kT\gamma f_{BW}\frac{I_D}{V_{OD}}\frac{V_{OD}+2B}{V_{OD}+B} \tag{4.20}$$

For isolation purposes, one such circuit is implemented for each of a pair of I,Q passive down-conversion mixers. In this case:

$$I_D = \tfrac{1}{12}I_{D,TOT} \tag{4.21}$$

Combining (4.18), (4.19), (4.20), and (4.21) yields an expression for the INR of the circuit:

$$INR_{SingleStage} \approx \frac{I_{D,TOT}B^4 A_1^4 A_2^2}{12 \cdot 2400kT\gamma f_{BW} V_{OD}^3 (V_{OD}+B)^5 (V_{OD}+2B)} \tag{4.22}$$

Neglecting the terms due to resistor noise, considering only two blocker signals, and invoking the earlier design choice of $I_{D,TOT} = 24I_D$ in (4.15) yields the INR of the multistage cubic term generator:

$$INR_{Multistage} \approx \frac{I_{D,TOT}B^4 A_1^4 A_2^2}{3072kT f_{BW}\left(\left(\frac{G_2}{G_1}\right)^2 A_2^2 + A_1^2\right) V_{OD}^3 (V_{OD}+B)^3 (V_{OD}+2B)} \tag{4.23}$$

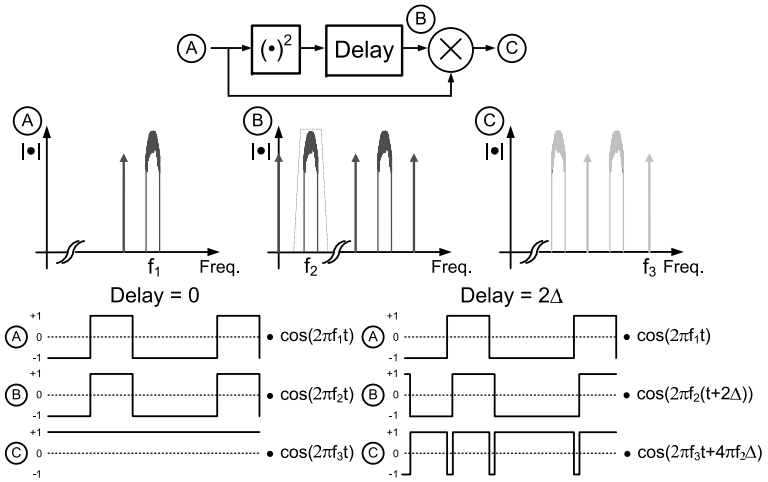


Figure 4.10: Depiction of effect of interstage group delay on CW and BPSK blocking signals producing IM3 products.

Comparing the ratio of the two INR quantities yields the approximate INR advantage obtained by the use of the multistage cubing circuit, despite the increased complexity:

$$\frac{INR_{Multistage}}{INR_{SingleStage}} = \frac{12 \cdot 2400(V_{OD}+B)^2}{3072\left(\left(\frac{G_2}{G_1}\right)^2 A_2^2 + A_1^2\right)} \tag{4.24}$$

From Sec. 4.5, we have that V_{OD}=0.132V and B=1.08V. We also have that the peak blocking condition has A_1=0.037V and A_2=0.071V for G_2/G_1=0.8. In this case, the INR power advantage of using a multistage cubic term generator is about 35dB. Although this number is obtained from somewhat flexible design choices and a specific blocking condition, the magnitude of the advantage is enough to make a general statement regarding the superiority of using multistage cubic term generators over single-stage cubic term generators when attempting to maximize the circuit INR for blocker input levels about an order of magnitude less than 1V.

4.7 Effect and Quantitative Analysis of Interstage Group Delay

Group delay between the multiplications in the IM3 term generator may result in a modulated signal multiplying a delayed version of itself or another modulated signal. This can be seen by considering an unfiltered BPSK-modulated signal, as shown in Fig. 4.10. In this case, the IM3 products consist of a frequency-translated version of the squared BPSK signal. Error due to the delay occurs in the time interval over which adjacent symbols both differ and overlap, denoted by the dashed lines. The instance of this effect which occurs under the FDD Region 1 UMTS specification [6] is an IM3 product arising from the squared QPSK TX leakage and a CW tone ranging from 1.67GHz-1.85GHz.

The power spectral density of a squared complex digitally modulated signal has been studied

in the past [99]; however, it is vital to perform a similar analysis for QPSK here for two reasons. First, unlike the XMD evaluated in [99] which was composed of a CW-shifted O-QPSK baseband low-frequency IM2 envelope, in this work the relevant distortion is the IMD effectively resulting from a CW-shifting of the high-frequency terms resulting from the squaring of a QPSK signal. In the latter case, the IMD I and Q components have completely different PSD expressions. Secondly, the analysis needs to be performed taking into account the delay between multiplications in a multistage cubic term generator in order to compute the error due to this effect. The authors of [91] correctly noted that in a predistortion system with only gain and phase adaptive adjustments, the achievable cancellation of IM3 products is given by a bandpass first-order noise shaping function of the delay time. In the receiver, however, half the value of this delay is assumed to be reversed by the multi-tap adaptive equalizer along with any gain and phase mismatch in the process of minimizing mean squared error. In this case the error integrated across the desired signal channel will be lower than that given by [91].

The statistical analysis begins by recognizing that a bandpass complex modulated signal can be represented as (4.25), where A_m is the modulated signal carrier amplitude and where f_c is the modulated signal carrier frequency.

$$x(t) = A_m v(t)\cos(2\pi f_c t) + A_m w(t)\sin(2\pi f_c t) \tag{4.25}$$

As in [99], the TX channel filter is approximated with a brick wall frequency response with a cutoff frequency of 1/2T, where T equals the chip rate of the interfering modulation. Hence, for QPSK modulation the signals are expressed as in (4.26), where α is a random time shift uniformly distributed in (0,T).

$$\begin{array}{l} v(t) = \sum_{n=-\infty}^{\infty} i_n \mathrm{sinc}\left(\frac{t-nT+\alpha}{T}\right) \\ w(t) = \sum_{n=-\infty}^{\infty} q_n \mathrm{sinc}\left(\frac{t-nT+\alpha}{T}\right) \end{array} \tag{4.26}$$

For the sake of this analysis, i_n and q_n are i.i.d. discrete random variables realizing one of $\{1,-1\}$. To begin computation of the IM3 component of interest, the bandpass modulated signal must be squared and its high-frequency components retained, yielding (4.27).

$$y(t) = A_m^2\left[\tfrac{1}{2}v^2(t)\cos(4\pi f_c t) + v(t)w(t)\sin(4\pi f_c t) - \tfrac{1}{2}w^2(t)\cos(4\pi f_c t)\right] \tag{4.27}$$

To generate the relevant IM3 products at baseband, the expression in (4.27) is multiplied by two CW signals, a blocker and an LO signal. This constitutes an effective multiplication by a single sinusoidal signal with an arbitrary phase Θ with respect to the carrier terms of the modulated signal $s(t) = A_{CW}\cos(4\pi f_c t + \Theta)$ and yields after low-pass filtering at baseband:

$$z(t) = \tfrac{1}{4}A_m^2 A_{CW}\left[\left[v^2(t) - w^2(t)\right]\cos(\Theta) - 2v(t)w(t)\sin(\Theta)\right] \tag{4.28}$$

Note that this signal appears at the I channel baseband. In order to find the signal that appears at the Q channel baseband, it is merely sufficient to substitute Θ=Θ+90° into the final expression. The order of multiplications above is taken to simplify the analysis and is not the same as that performed in the circuit in Fig. 4.1. However, the proof in Sec. 4.3 can be used to show that the result is the same in either case.

In the alternate path, (4.28) is altered by the presence of the interstage group delay, which is denoted by 2Δ, to yield (4.29).

$$\begin{aligned} z'(t) = \tfrac{1}{4}A_m^2 A_{CW} \left[[v(t+\Delta)v(t-\Delta) - w(t+\Delta)w(t-\Delta)] \cos(\Theta) \right. \\ \left. - [v(t+\Delta)w(t-\Delta) + w(t+\Delta)v(t-\Delta)] \sin(\Theta) \right] \end{aligned} \tag{4.29}$$

Expressing (4.29) in this fashion makes an implicit assumption that the group delay is constant across the entire range of frequencies spanned by the IM2 beat frequency products at any given time. This assumption is reasonable, at least for the conditions directly of interest → simulation results of the cubic term generator in Chap. 3 show that the average group delay deviation within the interstage frequency band over the 3.84MHz TX leakage bandwidth is 0.67%. In this case, the error due to the delay is given by (4.30). Starting from (4.30) the time parameter is removed from the arguments of v and w, while Δ is relegated to a subscript to condense the expressions.

$$\begin{aligned} z(t) - z'(t) = e(t) = \tfrac{1}{4}A_m^2 A_{CW} \left[\left[(v^2 - v_\Delta v_{-\Delta}) - (w^2 - w_\Delta w_{-\Delta}) \right] \cos(\Theta) \right. \\ \left. - [2vw - v_\Delta w_{-\Delta} - w_\Delta v_{-\Delta}] \sin(\Theta) \right] \end{aligned} \tag{4.30}$$

The autocorrelation of the error is therefore given by (4.31), where the delay offset τ is also placed within the subscript of v and w where it exists.

$$\begin{aligned} &R_{ee}(t+\tau;t) = \tfrac{1}{16}A_m^4 A_{CW}^2 \cdot \\ &E[[[(v_\tau^2 - v_{\tau+\Delta}v_{\tau-\Delta}) - (w_\tau^2 - w_{\tau+\Delta}w_{\tau-\Delta})] \cos(\Theta) - [2v_\tau w_\tau - v_{\tau+\Delta}w_{\tau-\Delta} - w_{\tau+\Delta}v_{\tau-\Delta}] \sin(\Theta) \cdot \\ &[(v^2 - v_\Delta v_{-\Delta}) - (w^2 - w_\Delta w_{-\Delta})] \cos(\Theta) - [2vw - v_\Delta w_{-\Delta} - w_\Delta v_{-\Delta}] \sin(\Theta)]] \end{aligned} \tag{4.31}$$

Note that terms such as $E[\bullet]\sin(\Theta)\cos(\Theta)$ are equal to zero due to the fact that one out of the four terms is a zero-mean bit sequence that is independent of the other three bit sequences. Multiplying out only the terms resulting in $E[\bullet](\cos^2(\Theta))$ and $E[\bullet](\sin^2(\Theta))$ yields, respectively:

$$\begin{aligned} &E\begin{bmatrix} v_\tau^2 v^2 - v_\tau^2 v_\Delta v_{-\Delta} - v_\tau^2 w^2 + v_\tau^2 w_\Delta w_{-\Delta} - v_{\tau+\Delta}v_{\tau-\Delta}v^2 \\ +v_{\tau+\Delta}v_{\tau-\Delta}v_\Delta v_{-\Delta} + v_{\tau+\Delta}v_{\tau-\Delta}w^2 - v_{\tau+\Delta}v_{\tau-\Delta}w_\Delta w_{-\Delta} \\ -w_\tau^2 v^2 + w_\tau^2 v_\Delta v_{-\Delta} + w_\tau^2 w^2 - w_\tau^2 w_\Delta w_{-\Delta} + w_{\tau+\Delta}w_{\tau-\Delta}v^2 \\ -w_{\tau+\Delta}w_{\tau-\Delta}v_\Delta v_{-\Delta} - w_{\tau+\Delta}w_{\tau-\Delta}w^2 + w_{\tau+\Delta}w_{\tau-\Delta}w_\Delta w_{-\Delta} \end{bmatrix} \cdot \cos^2(\Theta) = \\ &\begin{bmatrix} C_0 - A - D_0 + B \\ -A_{FR} + C + B_{FR} - D \\ -D_0 + B + C_0 - A \\ +B_{FR} - D - A_{FR} + C \end{bmatrix} \left(\tfrac{1}{2} + \tfrac{1}{2}\cos(2\Theta)\right) \end{aligned} \tag{4.32}$$

$$E\begin{bmatrix} 4v_\tau w_\tau vw - 2v_\tau w_\tau v_\Delta w_{-\Delta} - 2v_\tau w_\tau w_\Delta v_{-\Delta} \\ -2v_{\tau+\Delta}w_{\tau-\Delta}vw + v_{\tau+\Delta}w_{\tau-\Delta}v_\Delta w_{-\Delta} + v_{\tau+\Delta}w_{\tau-\Delta}w_\Delta v_{-\Delta} \\ -2w_{\tau+\Delta}v_{\tau-\Delta}vw + w_{\tau+\Delta}v_{\tau-\Delta}v_\Delta w_{-\Delta} + w_{\tau+\Delta}v_{\tau-\Delta}w_\Delta v_{-\Delta} \end{bmatrix} \sin^2(\Theta) = \begin{bmatrix} 4F_0 - 2E - 2E \\ -2E_{FR} + F + G \\ -2E_{FR} + G + F \end{bmatrix} \left(\tfrac{1}{2} - \tfrac{1}{2}\cos(2\Theta)\right) \tag{4.33}$$

A brief discussion on the notation in (4.32) and (4.33) is in order. Terms with the same letter can be shown to have the same PSD. The terms subscripted with 0 denote the terms that constitute the original IM3 signal. The PSD of the original IM3 products can therefore be obtained by setting Δ=0 in the PSD of C, D, and F. The terms denoted with FR can be shown to have PSDs equivalent to their respective terms, with the exception that the PSD is frequency-reversed.

The computation continues by evaluating each of these terms in turn to obtain subexpressions that can be made time-independent such that their Fourier transforms can be taken. Term C is evaluated first in (4.34).

$$C(t+\tau;t) = \sum_{n,m,p,r=-\infty}^{\infty} E[i_n i_m i_p i_r] \cdot \frac{1}{T}\int_0^T \text{sinc}(\tfrac{t+\tau-nT+\Delta+\alpha}{T})\text{sinc}(\tfrac{t+\tau-mT-\Delta+\alpha}{T}) \cdot \text{sinc}(\tfrac{t-pT+\Delta+\alpha}{T})\text{sinc}(\tfrac{t-rT-\Delta+\alpha}{T})d\alpha \tag{4.34}$$

The term in the expected value operator evaluates to unity rather than zero under only the following conditions:

(a) $n = m = p = r$

(b) $n = m = k$ and $p = r = l \neq n$

(c) $n = p = k$ and $m = r = l \neq n$

(d) $n = r = k$ and $m = p = l \neq n$

These four cases yield the terms in (4.35), in respective order. In order to condense the expressions, "sinc" is replaced with "s".

$$
\begin{aligned}
C(t+\tau;t) = & \\
& -\frac{2}{T}\sum_{n=-\infty}^{\infty}\int_0^T s(\tfrac{t+\tau-nT+\Delta+\alpha}{T})s(\tfrac{t+\tau-nT-\Delta+\alpha}{T})s(\tfrac{t-nT+\Delta+\alpha}{T})s(\tfrac{t-nT-\Delta+\alpha}{T})d\alpha \\
& +\frac{1}{T}\sum_{k,l=-\infty}^{\infty}\int_0^T s(\tfrac{t+\tau-kT+\Delta+\alpha}{T})s(\tfrac{t+\tau-kT-\Delta+\alpha}{T})s(\tfrac{t-lT+\Delta+\alpha}{T})s(\tfrac{t-lT-\Delta+\alpha}{T})d\alpha \\
& +\frac{1}{T}\sum_{k,l=-\infty}^{\infty}\int_0^T s(\tfrac{t+\tau-kT+\Delta+\alpha}{T})s(\tfrac{t-kT+\Delta+\alpha}{T})s(\tfrac{t+\tau-lT-\Delta+\alpha}{T})s(\tfrac{t-lT-\Delta+\alpha}{T})d\alpha \\
& +\frac{1}{T}\sum_{k,l=-\infty}^{\infty}\int_0^T s(\tfrac{t+\tau-kT+\Delta+\alpha}{T})s(\tfrac{t-kT-\Delta+\alpha}{T})s(\tfrac{t+\tau-lT-\Delta+\alpha}{T})s(\tfrac{t-lT+\Delta+\alpha}{T})d\alpha \\
& = \epsilon+\gamma+\eta+\kappa
\end{aligned}
\tag{4.35}
$$

Note that the first term has a negative sign, to account for the fact that three such terms need to be removed to avoid repeat counting of the same expression. The term A can be evaluated in a similar manner, resulting in the following expression:

$$
\begin{aligned}
& A(t+\tau;t) = \\
& \sum_{n,m,p,r=-\infty}^{\infty} E[i_n i_m i_p i_r]\cdot\frac{1}{T}\int_0^T s(\tfrac{t+\tau-nT+\alpha}{T})s(\tfrac{t+\tau-mT+\alpha}{T})s(\tfrac{t-pT+\Delta+\alpha}{T})s(\tfrac{t-rT-\Delta+\alpha}{T})d\alpha = \\
& -\frac{2}{T}\sum_{n=-\infty}^{\infty}\int_0^T s(\tfrac{t+\tau-nT+\alpha}{T})s(\tfrac{t+\tau-nT+\alpha}{T})s(\tfrac{t-nT+\Delta+\alpha}{T})s(\tfrac{t-nT-\Delta+\alpha}{T})d\alpha \\
& +\frac{1}{T}\sum_{k,l=-\infty}^{\infty}\int_0^T s(\tfrac{t+\tau-kT+\alpha}{T})s(\tfrac{t+\tau-kT+\alpha}{T})s(\tfrac{t-lT+\Delta+\alpha}{T})s(\tfrac{t-lT-\Delta+\alpha}{T})d\alpha \\
& +\frac{1}{T}\sum_{k,l=-\infty}^{\infty}\int_0^T s(\tfrac{t+\tau-kT+\alpha}{T})s(\tfrac{t-kT+\Delta+\alpha}{T})s(\tfrac{t+\tau-lT+\alpha}{T})s(\tfrac{t-lT-\Delta+\alpha}{T})d\alpha \\
& +\frac{1}{T}\sum_{k,l=-\infty}^{\infty}\int_0^T s(\tfrac{t+\tau-kT+\alpha}{T})s(\tfrac{t-kT-\Delta+\alpha}{T})s(\tfrac{t+\tau-lT+\alpha}{T})s(\tfrac{t-lT+\Delta+\alpha}{T})d\alpha \\
& = \lambda+\beta+2\chi
\end{aligned}
\tag{4.36}
$$

It can similarly be shown that:

$$
\begin{aligned}
B=\beta, D=\gamma, \\
E=\chi, F=\eta, \\
G=\kappa
\end{aligned}
\tag{4.37}
$$

The complete expression for the autocorrelation of the error signal can therefore be given by:

$$
R_{ee}(\tau) = \tfrac{1}{16}A_m^4 A_{CW}^2\cdot\big(2[\epsilon_0-(\lambda+\lambda_{TR})+\epsilon][\tfrac{1}{2}+\tfrac{1}{2}\cos(2\Theta)]+2[2\eta_0-2(\chi+\chi_{TR})+\eta+\kappa]\big) \tag{4.38}
$$

The complete expression for the autocorrelation of the original squared modulated signal can be

given by:

$$R_{zz}(\tau) = \tfrac{1}{16} A_m^4 A_{CW}^2 [\epsilon_0[1 + \cos(2\Theta)] + 4\eta_0] \tag{4.39}$$

The general procedure for taking the Fourier transform of each of the five subterms can be shown in detail for term η. As many of the steps are repeated for each of the terms, they will only be performed for term η, while the unique steps for each term will be performed in detail.

4.7.1 Evaluating Term η

The first manipulation is to make the substitution that l=k+s, where k and s are both integers.

$$\eta = \frac{1}{T} \sum_{k,s=-\infty}^{\infty} \int_0^T s(\tfrac{t+\tau-kT+\Delta+\alpha}{T}) s(\tfrac{t+\tau-kT-sT-\Delta+\alpha}{T}) s(\tfrac{t-kT+\Delta+\alpha}{T}) s(\tfrac{t-kT-sT-\Delta+\alpha}{T}) d\alpha \tag{4.40}$$

Next, the substitutions that t-kT+α=u, du=dα are made and the appropriate adjustments to the limits of the integral are made.

$$\eta = \frac{1}{T} \sum_{s=-\infty}^{\infty} \int_{-\infty}^{\infty} s(\tfrac{u+\tau+\Delta}{T}) s(\tfrac{u+\tau-sT-\Delta}{T}) s(\tfrac{u+\Delta}{T}) s(\tfrac{u-sT-\Delta}{T}) du \tag{4.41}$$

Taking the Fourier transform of this expression yields:

$$H(f) = \frac{1}{T} \int_{-\infty}^{\infty} \sum_{s=-\infty}^{\infty} \int_{-\infty}^{\infty} s(\tfrac{u+\tau+\Delta}{T}) s(\tfrac{u+\tau-sT-\Delta}{T}) s(\tfrac{u+\Delta}{T}) s(\tfrac{u-sT-\Delta}{T}) e^{-j2\pi f\tau} du d\tau \tag{4.42}$$

Because it is known *a priori* that the power spectral density of the complete expression is a finite quantity at each frequency, integration and summation can be interchanged at will. Doing so yields the following steps:

$$H(f) = \frac{1}{T} \sum_{s=-\infty}^{\infty} \int_{-\infty}^{\infty} s(\tfrac{u+\tau+\Delta}{T}) s(\tfrac{u+\tau-sT-\Delta}{T}) e^{-j2\pi f\tau} d\tau \cdot \int_{-\infty}^{\infty} s(\tfrac{u+\Delta}{T}) s(\tfrac{u-sT-\Delta}{T}) du \tag{4.43}$$

Considering the two inner integrals separately, if the variable change x=u+τ is made, the following expression can be obtained:

$$H(f) = \frac{1}{T} \sum_{s=-\infty}^{\infty} \int_{-\infty}^{\infty} s(\tfrac{x+\Delta}{T}) s(\tfrac{x-sT-\Delta}{T}) e^{-j2\pi fx} dx \cdot \int_{-\infty}^{\infty} s(\tfrac{u+\Delta}{T}) s(\tfrac{u-sT-\Delta}{T}) e^{-j2\pi fu} du \tag{4.44}$$

Note that the Fourier transform of this sinc function is a rectangle function of total width 1/T in frequency and of height T, denoted $\Pi_T(f)$. Therefore, the Fourier transform of a sinc multiplied by a delayed sinc is a rectangle function convolved by a rectangle function which has been multiplied by a complex exponential. Performing this operation and successively interchanging integration and summation yields:

$$H(f) = \frac{1}{T}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\sum_{s=-\infty}^{\infty} e^{j2\pi(\varphi-\phi)sT}\Pi_T(f-\phi)\Pi_T(\phi)e^{-j4\pi\phi\Delta}d\phi\Pi_T(f-\varphi)\Pi_T(\varphi)e^{j4\pi\varphi\Delta}d\varphi \quad (4.45)$$

Recalling that the Fourier series of an impulse train is a sum of periodically spaced complex exponentials [100, pg. 208], the following change can be made to (4.45):

$$H(f) = \frac{1}{T^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\sum_{s=-\infty}^{\infty} \delta(\varphi-\phi-\tfrac{s}{T})\Pi_T(f-\phi)\Pi_T(\phi)\Pi_T(f-\varphi)\Pi_T(\varphi)e^{j4\pi(\varphi-\phi)\Delta}d\varphi d\phi \quad (4.46)$$

Integration and summation can again be exchanged and the inner integral evaluated using the sifting property of the Dirac delta function.

$$H(f) = \frac{1}{T^2}\int_{-\infty}^{\infty}\sum_{s=-\infty}^{\infty} \Pi_T(f-\phi)\Pi_T(\phi)\Pi_T(f-\phi-\tfrac{s}{T})\Pi_T(\phi+\tfrac{s}{T})e^{j4\pi s\Delta/T}d\phi \quad (4.47)$$

Due to the strict bandlimited nature of the rectangle function, (4.47) only evaluates to a nonzero value for s=0. The summation and complex exponential thus disappear in (4.48).

$$H(f) = \frac{1}{T^2}\int_{-\infty}^{\infty}\Pi_T^2(f-\phi)\Pi_T^2(\phi)d\phi = \int_{-\infty}^{\infty}\Pi_T(f-\phi)\Pi_T(\phi)d\phi = T\cdot Tr(f) \quad (4.48)$$

It can be seen that $H(f) = H_0(f)$ and that the triangle function Tr(f) represents:

$$Tr(f) = \begin{cases} |1-fT| & 0 \le |f| \le \frac{1}{T} \\ 0 & otherwise \end{cases} \quad (4.49)$$

4.7.2 Evaluating Term κ

Taking the Fourier transform of κ and applying the steps outlined earlier results in:

$$\begin{aligned} &K(f) = \\ &\frac{1}{T}\sum_{s=-\infty}^{\infty}\int_{-\infty}^{\infty}\Pi_T(f-\phi)e^{j2\pi(f-\phi)\Delta}\Pi_T(\phi)e^{-j2\pi\phi sT}e^{-j2\pi\phi\Delta}d\phi\cdot \\ &\int_{-\infty}^{\infty}\Pi_T(f-\varphi)e^{j2\pi(f-\varphi)\Delta}\Pi_T(\varphi)e^{j2\pi\varphi sT}e^{-j2\pi\varphi\Delta}d\varphi \end{aligned} \quad (4.50)$$

$$K(f) = \frac{1}{T^2}\int_{-\infty}^{\infty}\sum_{s=-\infty}^{\infty}\Pi_T(f-\phi)\Pi_T(\phi)e^{-j4\pi\phi\Delta}e^{j4\pi(f-\phi-\frac{s}{T})\Delta}\cdot\Pi_T(f-\phi-\tfrac{s}{T})\Pi_T(\phi+\tfrac{s}{T})d\phi \quad (4.51)$$

Noting again that the above expression is nonzero only when s=0, the expression simplifies to:

$$K(f) = \frac{1}{T^2}\int_{-\infty}^{\infty}\Pi_T^2(f-\phi)e^{j4\pi(f-\phi)\Delta}\Pi_T^2(\phi)e^{-j4\pi\phi\Delta}d\phi = \int_{-\infty}^{\infty}\Pi_T(f-\phi)e^{j4\pi(f-\phi)\Delta}\Pi_T(\phi)e^{-j4\pi\phi\Delta}d\phi \tag{4.52}$$

Expression (4.52) represents the convolution of a function with its conjugate. Hence, it can be re-expressed by rearranging the real and imaginary components as in (4.53).

$$\begin{aligned} K(f) = & \\ & \int_{-\infty}^{\infty}\Pi_T(f-\phi)\cos(4\pi(f-\phi)\Delta)\Pi_T(\phi)\cos(4\pi\phi\Delta)d\phi \\ & + \int_{-\infty}^{\infty}\Pi_T(f-\phi)\sin(4\pi(f-\phi)\Delta)\Pi_T(\phi)\sin(4\pi\phi\Delta)d\phi \end{aligned} \tag{4.53}$$

Using the linearity property of the integral operator and a trigonometric identity, the expression in (4.53) can be condensed.

$$K(f) = \Psi(4, f, \Delta) \tag{4.54}$$

Where the following function is defined:

$$\Psi(x, f, \Delta) = \int_{-\infty}^{\infty}\Pi_T(f-\phi)\Pi_T(\phi)\cos(x\pi(f-2\phi)\Delta)d\phi \tag{4.55}$$

4.7.3 Evaluating Term ϵ

Using the steps in the previous section, a preliminary expression for the Fourier transform of ϵ can be derived.

$$E(f) = -\frac{2}{T}\int_{-\infty}^{\infty}s(\tfrac{x+\Delta}{T})s(\tfrac{x-\Delta}{T})e^{-j2\pi fx}dx \cdot \int_{-\infty}^{\infty}s(\tfrac{u+\Delta}{T})s(\tfrac{u-\Delta}{T})e^{j2\pi fu}du \tag{4.56}$$

Further manipulation yields:

$$E(f) = -\frac{2}{T}\int_{-\infty}^{\infty}\Pi_T(f-\phi)e^{j2\pi(f-\phi)\Delta}\Pi_T(\phi)e^{-j2\pi\phi\Delta}d\phi \cdot \int_{-\infty}^{\infty}\Pi_T(f-\varphi)e^{-j2\pi(f-\varphi)\Delta}\Pi_T(\varphi)e^{j2\pi\varphi\Delta}d\varphi \tag{4.57}$$

$$E(f) = -\frac{2}{T}\Psi^2(2, f, \Delta) \tag{4.58}$$

4.7.4 Evaluating Term χ and λ

The Fourier transform of the expression obtained for χ is given by:

$$X(f) = \frac{1}{T}\sum_{s=-\infty}^{\infty}\int_{-\infty}^{\infty}s(\tfrac{x}{T})s(\tfrac{x-sT}{T})e^{-j2\pi fx}dx \cdot \int_{-\infty}^{\infty}s(\tfrac{u-\Delta}{T})s(\tfrac{u-sT+\Delta}{T})e^{j2\pi fu}du \tag{4.59}$$

Further manipulations in the vein of those previously described yield the following expression:

$$X(f) = \int_{-\infty}^{\infty} \Pi_T(f-\phi) e^{j2\pi(f-\phi)\Delta} \Pi_T(\phi) e^{-j2\pi\phi\Delta} d\phi = \Psi(2, f, \Delta) \tag{4.60}$$

Because the function $\Psi(x, f)$ is symmetric with respect to frequency, it also holds that $X(f) = X_{TR}(f)$. Evaluating λ follows the exact same set of steps and results in (4.61).

$$\Lambda(f) + \Lambda_{TR}(f) = -4Tr(f) \cdot \Psi(2, f, \Delta) \tag{4.61}$$

4.7.5 Final Power Spectral Density Expressions

It can be shown through Taylor series analysis that:

$$\Psi(x, f, \Delta) = \frac{T^2}{x\pi\Delta} \sin\left(\frac{x\pi\Delta}{T} Tr(f)\right) \tag{4.62}$$

L'Hôpital's rule yields that for the case of Δ=0, (4.62) simplifies to T·Tr(f).

Combining the final expressions obtained above yields expressions for the power spectral densities of the IM3 products and the error due to the interstage delay at the I channel baseband output. Recall that to obtain the Q channel baseband output it is only necessary to substitute Θ=Θ+90°.

$$S_{zz}(f) = \tfrac{1}{4} A_m^4 A_{CW}^2 T \cdot Tr(f)[1 - Tr(f)(\tfrac{1}{2} + \tfrac{1}{2}\cos(2\Theta))] \tag{4.63}$$

$$S_{ee}(f) = \tfrac{1}{16} A_m^4 A_{CW}^2 [\Omega(f) + \Phi(f)(\tfrac{1}{2} + \tfrac{1}{2}\cos(2\Theta))] \tag{4.64}$$

$$\Omega(f) = 6T \cdot Tr(f) + 2\Psi(4, f, \Delta) - 8\Psi(2, f, \Delta) \tag{4.65}$$

Further evaluation of (4.65) requires a Taylor series expansion of the Ψ function. Retaining only the dominant terms yields:

$$\Omega(f) \approx \tfrac{16}{5T^3} \pi^4 \Delta^4 Tr^5(f) \tag{4.66}$$

Proceeding in the same manner yields:

$$\Phi(f) = -4T \cdot Tr^2(f) + 8Tr(f)\Psi(2, f, \Delta) - \tfrac{4}{T}\Psi^2(2, f, \Delta) \tag{4.67}$$

Again retaining only the dominant terms of the Taylor series expansion yields (4.68).

$$\Phi(f) \approx -\tfrac{16}{9T^3} \pi^4 \Delta^4 Tr^6(f) \tag{4.68}$$

The final result for the error PSD is then given by (4.69).

$$S_{ee}(f) \approx A_m^4 A_{CW}^2 \frac{\pi^4 \Delta^4}{T^3} Tr^5(f)[\tfrac{1}{5} - \tfrac{1}{9} Tr(f)(\tfrac{1}{2} + \tfrac{1}{2}\cos(2\Theta))] \tag{4.69}$$

The correctness of these calculations and the validity of the Taylor series truncation can be

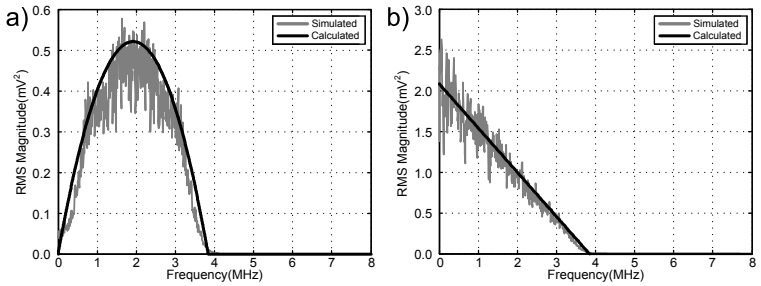


Figure 4.11: Baseband PSD (single-ended) of squared QPSK modulated signal for a) I and b) Q channels when Θ=0°.

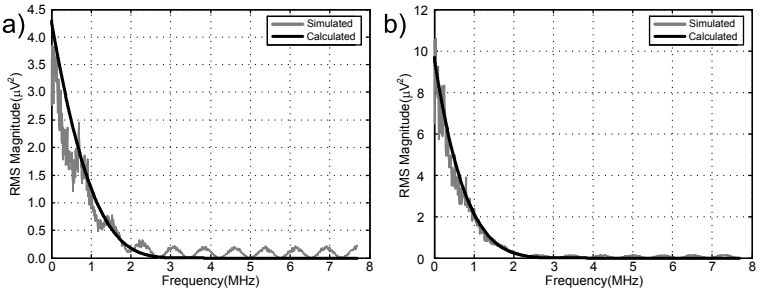


Figure 4.12: Baseband PSD (single-ended) of squared QPSK modulated signal group delay-related error for a) I and b) Q channels when Θ=0°.

confirmed by considering Figs. 4.11 and 4.12. In this case, the results were obtained using a 641-tap sinc pulse-shaping channel filter in a discrete-time simulation construct where Fs=245.76MHz. These figures represent the case where A_m=2, A_{CW}=1, and Δ=4ns.

Using these results, one can predict the IER due to this error mechanism. In Figs. 4.13 and 4.14, the "I" represents total downconverted baseband IM3 power in order to provide a relevant measure of total blocker power. "E" represents error quantities falling within the RX channel bandwidth 1/2Tr, where Tr is the chip rate of the received modulated signal. In Fig. 4.13, Δ is swept and T=Tr=260.4ns, as in the UMTS specification test. In Fig. 4.14, the single-sided RX channel bandwidth and delay Δ are fixed at 1.92MHz and 4ns, respectively, while T is varied. Note that in the work presented in Chap. 3, Δ_{MAX} was simulated at 0.5ns and the alternate path IER target was 31dB for a receiver with uncorrected IIP3 of -9dB. In this case, the effect due to the interstage group delay is negligible, as Fig. 4.13 shows. However, for more aggressive receiver designs requiring greater IM3 cancellation, this effect eventually becomes a limiting factor.

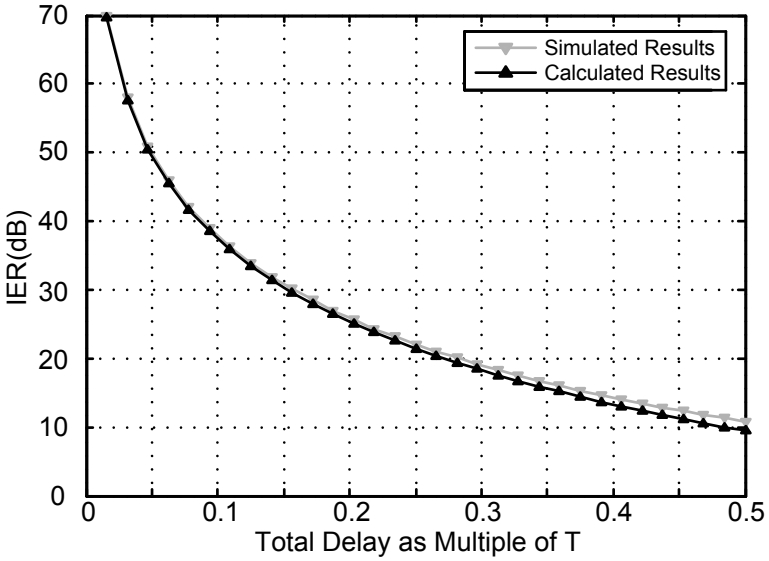


Figure 4.13: Calculated IER as a function of group delay for T=Tr.

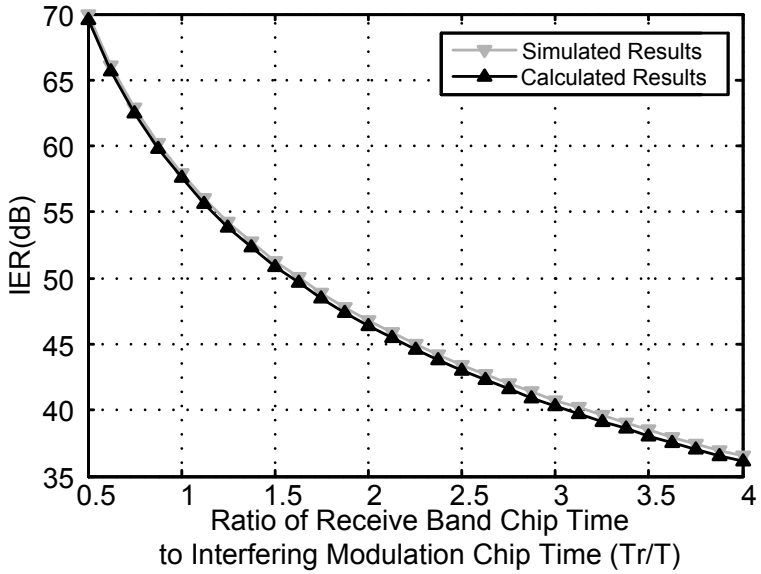


Figure 4.14: Calculated IER vs. blocker modulation chip rate for fixed delay.

Chapter 5

A Wide-Swing Low Noise Transconductance Amplifier and the Enabling of Large-Signal Handling Direct Conversion Receivers

5.1 General Design of a Large-Signal Handling Direct Conversion Receiver

Large-signal handling receivers achieve appreciable sensitivity to a small desired signal in the presence of out-of-band interferers whose voltage amplitude approaches that of the supply rail. As discussed in Section 1.6, such receivers would find use in frequency-domain duplexed communications transceivers, in implantable medical devices using separate power and data links, and in radars robust against hostile jamming interference. Once it is determined that building such a receiver is a worthwhile goal, the natural question becomes how to generally approach the architectural design of the receiver.

Inevitably, the effective small-signal gain provided by a semiconductor device is a function of electric potential, not flow. For example, the boundaries of the three regions of operation of an ideal square-law MOS device are thresholds dictated by the terminal voltages, not currents. In principle, a large enough MOS device can pass any finite amount of current, but no MOS device can maintain appreciable small-signal current gain per unit quiescent current when the gate-to-source V_{GS} and drain-to-source V_{DS} voltages drop to zero.

As the majority of radio receivers must decode desired signals in the μV range incident on the antenna, and as the analog-to-digital converter input-referred noise floor is typically near the mV range, and given that the analog-to-digital converter is typically designed such that it is not the dominant burden of the receiver noise figure (NF), the receiver front end must employ nearly three orders of magnitude of in-band voltage gain. However, for an out-of-band blocker that is already nearly as large as the supply rails, *any* voltage gain would result in gain desensitization of the

receiver. Although on-chip filtering at RF may be employed to reduce the magnitude of out-of-band blockers at the current-to-voltage conversion, this filtering is typically composed of a single RLC tank with which only 3-7dB of attenuation at a 100-200MHz frequency offset can be attained as shown in Fig. 1.14.

To provide a quantitative example of why on-chip filtering at RF with RLC tanks cannot solve this problem in a power- or area-efficient manner, a plausible scenario is assumed in which a near rail-to-rail out-of-band blocker 110dB above the desired signal exists at a 100MHz frequency offset. Because the downconverted blocker can possibly alias into the desired signal channel at the analog-to-digital conversion point, this out-of-band blocker needs to be attenuated at least to the level of the desired signal, if not more. Examining Fig. 1.14 shows that at least 37 RLC tanks are required to achieve this goal. In order to achieve the high Q depicted in Fig. 1.14 the inductors need to be set to the maximum radius allowed by the design kit, typically 250μm square. In order to avoid coupling and parasitic magnetic feedback loops, the inductors also need to be set some distance apart from each other, perhaps 250μm. In this case, the total area consumed by the filtering is 18.5mm^2. Furthermore, the small signal gain of the first few stages must be set on the order of 0dB so as not to rail the output of each amplifier. In this case, the noise of these first few stages and some immediately thereafter will approximately rms-add. For the typical square-law MOS model, the input-referred noise power spectral density of a common-source gain stage is given by (5.1).

$$P_{N,Stage}(\Delta f) = \frac{\frac{8}{3}kT\Delta f}{g_m} = \frac{\frac{4}{3}kTV_{OD}\Delta f}{I_D} \tag{5.1}$$

Choosing V_{OD}=$V_{GS} - V_{TH}$=200mV to maintain reasonable linearity and to avoid operation in the weak inversion region, the input-referred noise power contribution of each stage per Hertz is $\frac{4}{15}kT/I_D$. In this case, for an I_D=266mA, each stage contributes a noise power kT[1] to the original input referred noise power of kT[2]. This yields an approximate NF of 7dB counting just the first five stages of such a scheme. From this rough calculation, the combined power and area of this proposed solution is clearly unjustifiable, especially in mobile communications application.

Filtering of out-of-band blockers is performed far more efficiently at baseband. One way of seeing this is to note that selectivity in such circuits is accomplished through frequency-selective impedance division between resistive and capacitive elements. For a given order filter, the maximum achievable attenuation is limited by the Q of the capacitor, which is normally several orders of magnitude higher than that of the inductor required for filtering at RF. The noise of a baseband filter is typically dictated by its equivalent input resistance. Lowering this resistance requires a proportional increase in capacitance to maintain the same filter time constant and cutoff frequency. As opposed to inductors, the lithography and fine oxide growth capabilities in modern planar semiconductor processes make high-density flux containment in integrated capacitors easily attainable. Hence, the baseband filter capacitors do not constitute the same serious area burden as do integrated inductors.

Based on these realizations, the design of an optimal large-signal handling direct conversion receiver would operate primarily in the current domain at RF, with as few current-to-voltage conversions as possible. The principal current-to-voltage conversion should occur at baseband where a large filtering capacitor can result in a substantial attenuation of an out-of-band blocker. For ex-

[1]Per unit Hertz.
[2]Per unit Hertz.

Reference	Process	Best Reported IIP3	Noise Figure	Conditions
[102]	180nm	-1dBm	4.4dB	Tone Spacing < 10MHz
[51]	130nm	-2dBm	3.5dB	In-Band, 0.2MHz < Tone Spacing < 5MHz
[103]	130nm	+14dBm	10.5dB	Tone Spacing = 10MHz
[104]	180nm	+9dBm	8dB	Unknown
[105]	180nm	+10.9dBm	3.4dB	1 Tone: 1MHz, 2 Close Tones: 45MHz
[13]	65nm	+16dBm	4dB	Tone Spacing = 800MHz

Table 5.1: Best-case reported IIP3 from current-domain passive-mixer architectures.

ample, for a 2MHz signal bandwidth, an out-of-band blocker at a 100MHz offset may be attenuated by over 30dB with respect to the low-frequency downconverted desired signal with a single real-pole time constant.

Indeed, these sorts of current-mode RF techniques have recently enjoyed a renaissance in the form of passive mixer receiver architectures. Such architectures utilize a low-noise transconductance amplifier (LNTA) to convert the receiver input voltage (or perhaps a subsequent RF voltage) into a current which is then commutated by a set of MOS pass transistors that perform the downconversion to baseband frequencies. An initial study [101] investigated the 1/f noise properties of passive mixers and provided some practical design guidelines. A typical passive-mixer based receiver architecture such as the one shown in Fig. 1.26 has no dc current flowing through the downconverter switching pair. Although conventional theory predicts that due to the absence of dc current, the 1/f noise contribution of these downconverter devices is negligible, the authors of [101] found that 1/f noise is still present. However, passive mixer-based current-mode direct conversion receivers reported in [102] and [51] showed results in which low 1/f noise corners were obtained along with high IIP3 performances of -1dBm and -2dBm, respectively. It was realized in [102] that the current-mode downconversion architecture dramatically improved the linearity of the mixer to the extent that the receiver linearity depended ideally solely on the LNA. The work in [36] was the first to note the advantages of current-domain mixer operation and baseband-only filtering as advantageous to software-defined radio operation. The concepts of impedance translation by the passive mixer and filtering at baseband prior to the first I-to-V conversion to improve mixer linearity and 1-dB compression point for out-of-band blockers both make their first appearance in [103]. Although the work in [103] presents itself as merely a demodulator, were the transconductor input impedance made to be 50Ω over the band of interest, the work would very well qualify as a complete receiver front end. Subsequently, several other current-mode passive mixer architectures were reported with the sole filtering at baseband [104] [105]. Finally, the work in [13] offers a thorough and concise presentation of the concept of operating a current-mode passive mixer receiver such that the first I-V conversion occurs at baseband after filtering explicitly for the purpose of improving out-of-band linearity.

It is worth pondering briefly why current mode passive mixer based systems have gained increasing prominence over the last decade. Furthermore, it is interesting to note that the best IIP3 numbers reported for such systems (some reports do not distinguish between in-band and out-of-band) have been steadily increasing, as can be seen in Table 5.1. One reason is that scaling has considerably reduced the power required to fully transition a MOS device gate voltage from ground to supply. As the minimum device length shrank, the maximum achievable W/L ratio increased for a given mixer driver power dissipation, permitting the handling of larger currents before a given

distortion threshold was reached. This effect can be captured in (5.2) in which a passive mixer figure of merit (FOM) is proposed which is composed of the product of the power required to drive the passive mixer and the passive mixer on resistance. Clearly, a lower FOM is superior than a higher FOM in this case. It can be seen that the result is inversely proportional to (1.12), showing that as the minimum device dimension decreases by a factor K, the maximum operating frequency of the passive mixer for a given power dissipation increases by a factor of K^2.

$$FOM_{PassiveMixer} = P \cdot R = \frac{WLC_{ox}V_{DD}^2 f}{\frac{W}{L}\mu_n C_{ox}(V_{DD} - V_{TH})} = \frac{L^2 V_{DD}^2 f}{\mu_n (V_{DD} - V_{TH})} \tag{5.2}$$

In the context of a large-signal handling receiver, however, all of the aforementioned architectures remain limited in their signal handling capability by the initial V-I conversion in the receiver. The authors of [101] proposed the use of an inductively degenerated common-source LNA but did not build it. Later, the architecture proposed in [101] was implemented, both with a stacked NMOS/PMOS inductively degenerated LNA [102] and with a cascoded NMOS inductively degenerated LNA [51]. In the case of [102], large signal operation was not reported. In [51] input P_{-1dB} was reported to be -14dBm but it is unclear if this number corresponds to in-band signals. Similarly, out-of-band large-signal operation is not reported for the other receivers with the exception of [13]. The receiver in [13] reports an out-of-band desensitization of -24dBm at a 400MHz offset frequency, with the performance improving to -10dBm by 500MHz, -6dBm by 800MHz and about 0dBm by 2.4GHz offset. This 0dBm number may be benefiting from the reported 6GHz RF bandwidth of the receiver in which the input signals are attenuated prior to effecting any desensitization. As the LO frequency is 400MHz in this case, the blocker frequency is at 2.8GHz and likely achieves some filter attenuation benefit at RF in addition to baseband. Regardless, it appears that the LNA or LNTA is limiting the desensitization performance of the receiver to well below a magnitude commensurate with that of the supply rail. Hence, there exists a need to design an LNA or LNTA with which rail-to-rail input operation of the receiver may be achieved.

5.2 Prior Art in CMOS LNAs and LNTAs

5.2.1 Current Sharing

As conveyed in Section 1.6, scaling has also permitted the utilization of PMOS/NMOS RF low noise amplifiers in recent years. As such, there has been no shortage of activity in this area. Like PMOS/NMOS baseband "rail-to-rail" opamp architectures, such LNAs and LNTAs have the potential to operate over a wider input signal range than do LNAs or LNTAs utilizing only one type of MOS device.

Many PMOS/NMOS LNA or LNTA architectures present in the literature exploit the availability of both types of device to take advantage of the current sharing enabled by stacking two single-stage amplifiers operating in parallel from a small-signal perspective. For example, [106] [103] [102] [13] all stack NMOS and PMOS common-source (CS) amplifiers, while [107] stacks NMOS and PMOS common-gate (CG) amplifiers. The work in [49] exploits current sharing between parallel CS and CG stages, while the transconductor in [104] exploits current sharing within a cross-coupled CG/CS stage.

5.2.2 IM2 Cancellation

Cable television tuners must process an effectively ultra-wideband signal with widely varying signal strength across the band at any given time. Because of this, not only is odd-order intermodulation distortion important, but even-order IMD and HD can show up almost anywhere within the effective signal band. For example, IM2 distortion from strong cable TV channels at low frequencies can corrupt the signal of weaker channels at approximately double the strong channel frequency. Although differential circuits can improve IIP2 performance of RF amplifiers, the single-ended cabling present in these systems necessitates at least some single-ended amplifiers within the system. One solution to this problem is to place NMOS and PMOS common-source stages in parallel without current sharing [108] [109] [110]. In this case, the parallel NMOS and PMOS stages act effectively as two halves of a differential circuit assuming that they can be matched. This concept has also recently been extended [111] to confer the benefit of IM2 cancellation on a common-gate, common-source amplifier utilizing the noise cancellation concept described in [112].

5.2.3 IM3 Improvement in Common-Gate/Common-Source Noise Cancelling LNAs

Noise-cancelling LNAs that combine the wideband input matching properties of the common-gate LNA with the high transconductance values associated with a common-source LNA have been reported extensively since the work performed in [112]. Combined common-gate/common-source LNA architectures can also be engineered such that nonlinear distortion components arising from the common-gate amplifier can be made to cancel [113]. The architecture of [48] exploits the fact that the derivative of the common-gate transconductance g'_m of NMOS and PMOS common-gate devices is different, while g_m and g''_m remain the same. This results in an extra degree of freedom by which additional suppression of third-order distortion may be achieved. However, this IM3 cancellation scheme is relevant only for blocker levels as large as -20dBm, as higher-order distortion terms are not canceled.

5.2.4 Remarks

Despite the profusion of CMOS LNA and LNTA architectures, none purport to exhibit any wide-swing or large-signal handling capabilities. Although the CMOS push-pull LNAs that provide IM2 cancellation ameliorate the abrupt turn-off behavior of the MOS device, no evidence of odd-order IMD cancellation or large-signal handling is demonstrated or claimed. However, it seems as if the input-output characteristic stitching behavior inherent in class-AB push-pull topologies is indeed a promising path towards enabling a large voltage at the receiver input. In order to enable an even larger input, the push-pull LNTA should comprise a differential topology, as depicted in Fig. 5.1. Missing in Fig. 5.1 are the actual transconductance (TC) elements to be used, which is the topic to be considered next.

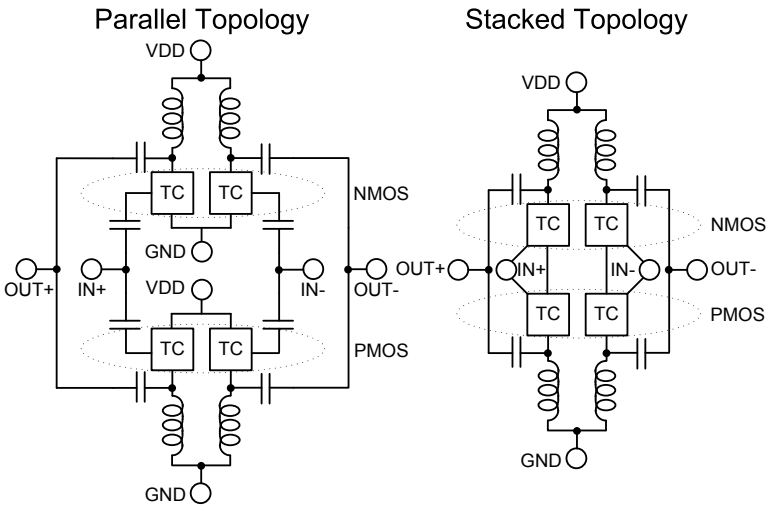


Figure 5.1: General CMOS push-pull LNTA topology.

5.3 Evolution of Wide-Swing CMOS Common-Gate LNTA

5.3.1 Considering a Common-Source Push-Pull LNTA

In addition to effecting a voltage-to-current conversion at the input, the LNTA should also provide a measure of buffering between the passive mixer and the antenna. Without such buffering, the LO signal may couple through to the antenna, possibly resulting in detectable LO radiation that violates the specified emissions mask of the transmitter with which the receiver is typically integrated. This radiation may also reflect off an external antenna and couple back into the antenna, resulting in time-varying LO self-mixing that results in large and difficult-to-cancel baseband signal offsets [11]. Proper isolation between the output and input ports of an LNA or LNTA is typically ensured via the use of a cascode device at the drain of the transconductor.

When handling a large input signal, the finite resistance of a passive mixer's switches results in a voltage appearing at the input that is a function of both the input and LO signals. For an optimally designed passive mixer with switches just large enough to pass the maximum input signal, this voltage may be appreciable. If present at the drain of a short-channel input transconductor, this signal may modulate the transconductor device channel length and operating region, resulting in excessive distortion and unintended signal mixing. Using a cascode device to separate the transconductor from the passive mixer dramatically reduces the extent to which such undesirable interactions occur.

Unfortunately, the use of this important technique is not consistent with the practical realization of a wide-swing common-source push-pull LNTA. To see why, the reader may consider as in Fig. 5.2a a simple cascoded common-source amplifier. In the typical case, the cascode device is sized to be the same as the input device in order to minimize the parasitic capacitance between the two nodes [114]. Neglecting the body effect, as the transconductor gate input voltage is increased, the drain voltage decreases by the same amount. The requirement to maintain the transconductor operation in the saturation region is given by (5.3):

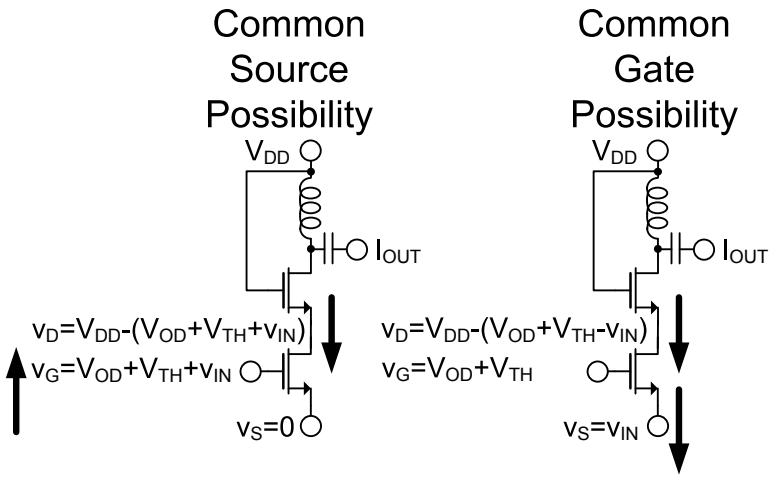


Figure 5.2: Available signal swing of cascoded common-source and common-gate transconductors.

$$v_{DS} > v_{GS} - V_{TH} \tag{5.3}$$

As a result, the maximum input voltage allowable before the device nominally enters the triode region of operation is given by (5.5):

$$V_{DD} - (V_{OD} + V_{TH} + v_{IN}) > (V_{OD} + V_{TH} + v_{IN}) - V_{TH} \tag{5.4}$$

$$\frac{V_{DD} - 2V_{OD} - V_{TH}}{2} > v_{IN} \tag{5.5}$$

The minimum input voltage allowable before the device enters the cut-off region is $v_{IN} = -V_{OD}$ for a total input voltage span of $(V_{DD} - V_{TH})/2$.

Of course, the device will begin transitioning into the triode region well before this point and experience an attendant loss in transconductance. Although the cascode device may be made larger to mitigate this effect, doing so for the PMOS device quickly adds enough parasitic capacitance to result in noticeable signal current losses to ac ground. More capacitance at this node also reduces the source degeneration impedance whose feedback attenuates the noise contribution of the cascode device.

By contrast to the common-source case, the transconductor drain voltage of the cascoded common-gate amplifier tracks with the input in the absence of the body effect, as depicted in Fig. 5.2b. In this case, the device drain-source voltage is roughly constant and the requirement to maintain transconductor operation in the saturation region is still given by (5.3). As a result, the minimum input voltage allowable before the device nominally enters the triode region of operation is given by (5.6):

$$\begin{aligned} V_{DD} - (V_{OD} + V_{TH} - v_{IN}) - v_{IN} &> V_{OD} + V_{TH} - v_{IN} \\ v_{IN} &> 2V_{OD} + V_{TH} - V_{DD} \end{aligned} \tag{5.6}$$

The maximum voltage allowable before the device enters the cut-off region is $v_{IN} = V_{OD}$ for a total input voltage span of $(V_{DD} - V_{TH} - V_{OD})$. For a small V_{TH} and V_{OD}, the common-gate amplifier maintains nearly twice the available signal swing as does the common-source amplifier, making it more suitable for a large signal handling LNTA.

One disadvantage of the common-gate LNTA is the relatively high noise figure of $1+\gamma$. However, in the CMOS process utilized in the foregoing experiments, $\gamma = 0.5$ for NF=1.8dB in simulation, which is well below the required NF of all major cellular communications standards. It is expected that the use of the wide-swing LNTA will also relax filtering requirements prior to the LNTA in many applications. Relaxing the filtering requirements leads to lower loss in the initial filtering, improving the composite noise figure of the LNTA and up-front filter, assuming such an up-front filter is even still required. Another disadvantage of the common-gate LNTA is that the small signal transconductance is fixed by the 50Ω real impedance at the input that is required for optimal power matching to the driving source. The resultant 20mS transconductance is 3-4 times less than what might be achieved in an inductively degenerated common-source amplifier. Choosing the common-gate LNTA topology thus places an atypically large noise burden on the passive mixer and baseband. However, for the purposes of this project, the capability of handling a large input signal is paramount and the noise burden of the other receiver circuit blocks must be improved on a case-by-case basis.

5.3.2 Choice of Stacked vs. Parallel Topology

Given that a CMOS common-gate LNTA is most desirable for large input signal handling, the question remains whether to implement it in a parallel or stacked topology, as depicted in Fig. 5.1. The parallel topology clearly gives a wider signal swing, but the stacked topology results in quiescent current and inductor area savings, the latter of which comes about due to one less required input tuning inductor. One method to obtain the best of both worlds is to increase the power supply rail voltage to allow for greater signal swing and to stack the transistors. In this case, the peak input signal swing would be dictated by maintaining a safe margin away from the device oxide breakdown voltage. In a deep submicron CMOS process, the NMOS oxide breakdown voltage is slightly over 2V. By adopting a supply voltage rail of 1.5V and constraining the maximum v_{GS} to be about 1.2V, the risk of long-term oxide degradation is substantially mitigated. Were a supply rail of 1.5V adopted and a parallel topology adopted, the maximum v_{GS} would be much higher at about 1.9V, but would be risky in terms of potential oxide failure. Hence, a stacked topology is chosen.

5.3.3 Device Sizing in a Stacked Topology

Once the decision to design with a stacked topology is made, the next design choice is to determine the approximate sizing and quiescent current consumption of the stacked structure. One obvious constraint is the requirement of a composite differential transconductance equal to about $g_m = 1/(50\Omega)$. The required single-ended, single-device transconductance is the same. The quiescent current drawn by each half of the structure is roughly given by (5.7) assuming for design purposes a square-law MOSFET I-V characteristic.

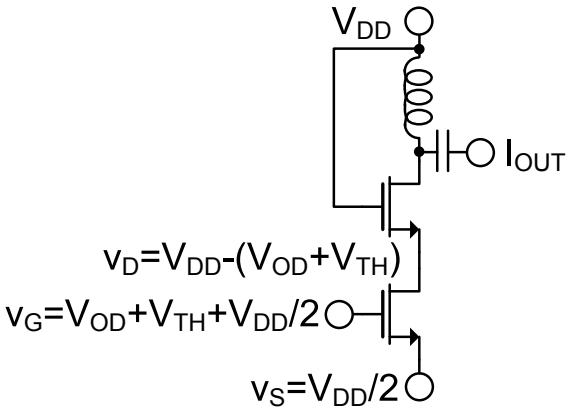


Figure 5.3: Biasing strategy for stacked class-AB common-gate LNTA.

$$\begin{aligned} I_D &= \tfrac{1}{2} g_m V_{OD} \\ V_{OD} &= V_{GS} - V_{TH} \end{aligned} \tag{5.7}$$

Operating the input devices of the LNTA in the saturation region guarantees that the noise figure of the LNTA is minimized under small-input signal operation. Doing so is vital given that the achievable noise figure of the common-gate LNTA is relatively large to begin with. One reason operating the input devices in saturation minimizes noise figure of the LNTA is that the isolating cascode devices are degenerated by a high impedance, driving their effective noise contribution to the output current to zero. Operating the input devices of the LNTA in the saturation region also maximizes effective input transconductance as a function of quiescent current.

In order to ensure robust operation in the saturation region over PVT corners, a voltage margin V_{MG}=0.2V is typically utilized such that $V_{DS} >= V_{GS} - V_{TH} + V_{MG}$. Having chosen the LNTA structure and considering the known biasing conditions for the LNTA upper half, as shown in Fig. 5.3, the input device overdrive voltage is determined by (5.10) if a square-law MOSFET model is assumed. Combining (5.10) with (5.7) yields the quiescent current of each branch of the LNTA.

$$V_{DD} - V_{OD} - V_{TH} - \tfrac{1}{2} V_{DD} = V_{OD} + V_{MG} \tag{5.8}$$

$$\tfrac{V_{DD}}{2} - V_{MG} - V_{TH} = 2V_{OD} \tag{5.9}$$

$$\tfrac{1}{2}(\tfrac{V_{DD}}{2} - V_{MG} - V_{TH}) = V_{OD} \tag{5.10}$$

Using low threshold voltage MOS devices with V_{TH}=250mV and for V_{DD}=1.5V, the approximate required overdrive voltage V_{OD} is 150mV. Along with the g_m requirement, this dictates the device width sizing and sets a quiescent current per LNTA half of about 1.5mA. Although there exists a degree of freedom in sizing the cascode device, in general the cascode device is sized to be the same as the input device in order to trade off the impedance seen by the input device drain and to minimize parasitic capacitance at the interdevice node.

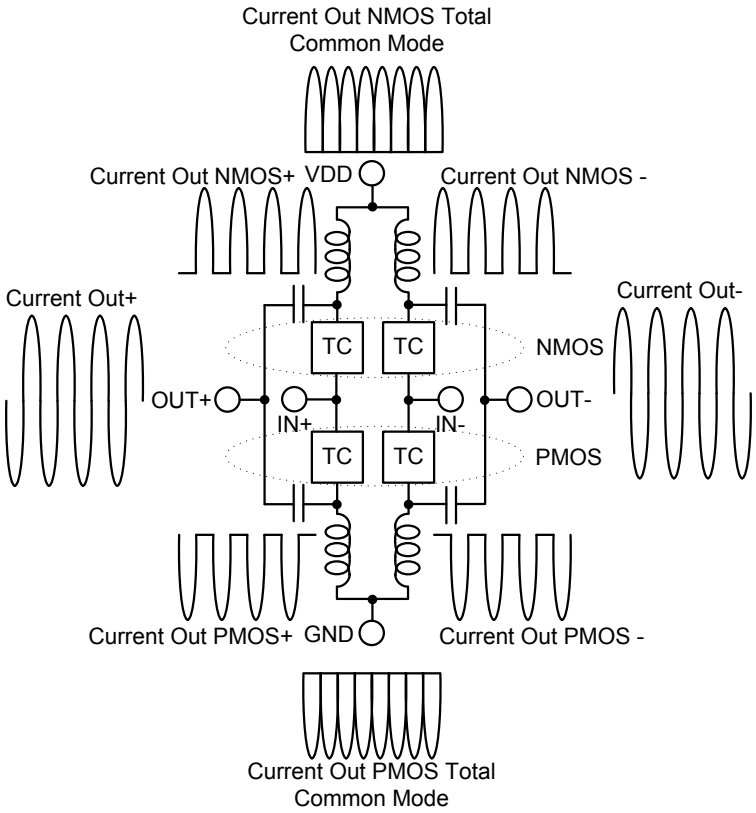


Figure 5.4: Stacked push-pull LNTA large-signal current flows.

5.4 General Theory of LNTA Operation

Visually considering the operation of a stacked push-pull LNTA of general transconductor as in Fig. 5.4 shows that each of the transconductors operates for slightly over one-half of a sinusoidal input cycle. The complete sinusoidal output is obtained by splicing together the ac output waveforms via capacitive coupling. When the NMOS and PMOS half-cycle currents are combined, the large even-order harmonics resulting from the turn-off of the transconductor devices are largely canceled. Although slightly imbalanced NMOS and PMOS devices result in some remaining even-order distortion products on each of the LNTA single-ended outputs, they are attenuated again by about an order of magnitude in the differential-mode output current due to their common-mode nature. Low-frequency even-order IMD product currents are blocked by the output capacitors and instead flow through the large inductors to the supplies.

Whether implemented with common-source or common-gate devices, one important implication of utilizing a differential push-pull LNTA is the generation of low-frequency (including dc) current in the common-mode output, as depicted in Fig. 5.4. Both NMOS and PMOS transconductors conduct current primarily in one direction, effecting a strong second-order nonlinearity which gives rise to a dc component in the output current of each device. This current passes through the inductors to the supply rails where it effects a static power dissipation proportional to the strength of the input signal. In a more accurate sense it is the low-frequency envelope which is generated by the second-order

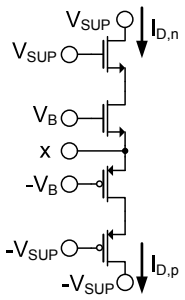


Figure 5.5: Circuit structure utilized for quantitative analysis of LNTA.

transconductor nonlinearity and is shunted to the supply rails by the frequency-selective impedance division of the large inductor and output capacitor. Although matching the characteristics of NMOS and PMOS devices at the quiescent point is possible in principle via the use of separate NMOS and PMOS constant-g_m biasing, maintaining exactly the same operation over the entire signal swing is in general not possible due to the different nonlinear characteristics of the device types. For example, mobility degradation effects in NMOS devices are noticeably greater than in PMOS devices because the inherent mobility of electrons is greater than that of holes. In this case, the low-frequency large-signal envelope current generated by the NMOS and PMOS halves of the circuit may differ by 20-30% even if the quiescent currents are matched.

In order to regulate the common-mode input voltage, this current must be sourced by a common-mode feedback OTA itself class-AB in nature, as its output current should only be large when the LNTA input signal is large. To properly source this envelope current, the unity-gain bandwidth of the OTA open-loop response should be greater than the maximum frequency at which significant components of the input signal envelope are present. Large OTA loop gain values attenuate low-frequency envelope signal appearing at the input common-mode voltage. When present at a significant power level, this signal then adds to the input signal and nonlinear interaction within the active devices results in successive powers of cross-modulation distortion appearing on the output signal current and any on distortion products generated by the input signals. Although this distortion may be removed using the scheme shown in Fig. 6.9, the presence of frequency-dependent attenuation of the baseband envelope at the LNTA input common-mode voltage serves to increase the complexity of the distortion cancellation scheme.

5.5 Obtaining Near-Constant LNTA Transconductance

5.5.1 General Concept

Assuming that the MOSFET devices of the LNTA can be well-modeled by the long-channel square-law equation (5.11) and denoting the single-ended input voltage of the LNTA as the variable x, the I-V relation of one half of the LNTA can be expressed as in (5.12)-(5.13). This expression assumes a split supply in order to exploit the symmetry of the stacked LNTA in the analysis. In this case as illustrated in Fig. 5.5, $x = 0$ in the quiescent condition, while V_{SUP}=$V_{DD}/2$ and V_B is the difference between the bias voltage on the gate of the input MOS device and the quiescent input voltage. Both

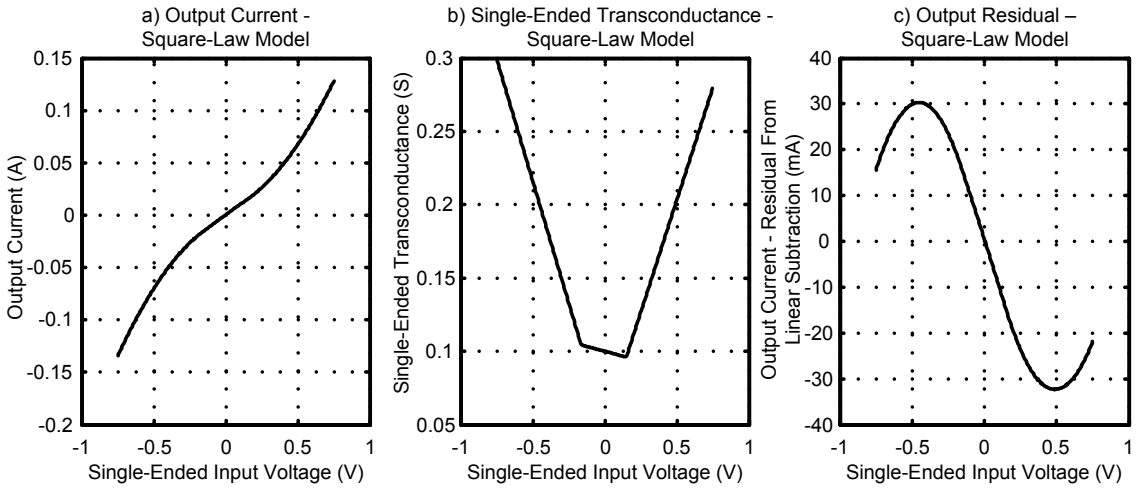


Figure 5.6: Calculated output characteristics of single-ended LNTA for square-law MOSFET model. a) Output current. b) Transconductance. c) Residual after removal of small-signal linear fit.

the I-V and g_m-V relations are plotted in Fig. 5.6, showing a roughly linear dependence on the input variable x when both devices are operating in the saturation region and a quadratic dependence of the input variable x when one device is operating in the saturation region and the other device is operating in the cutoff region. In this case, the drain-source voltage V_{DS} of the input device is assumed to be constant, as implied by the qualitative description of the circuit in Section 5.3.1. For illustrative purposes, the g_m-V relation for the single-ended output current is computed in (5.14) when the NMOS and PMOS device widths are sized to have the same transconductance as a function of voltage. In this case, there exists a small region of constant transconductance bounded by regions of transconductance as a linear function of input voltage. In Fig. 5.6 this constant region is tilted slightly due to the mismatched characteristics of the NMOS and PMOS devices.

$$I_{sq} = \tfrac{1}{2}\tfrac{W}{L}\mu_0 C_{ox}(v_{OD})^2(1+\lambda V_{DS}) \tag{5.11}$$

$$\begin{aligned} v_{OD} &= V_G - v_S - V_{TH} \\ v_{OD,n} &= \max(V_B - V_{TH,n} - x, 0) \\ v_{OD,p} &= \max(V_B - V_{TH,p} + x, 0) \end{aligned} \tag{5.12}$$

$$\begin{aligned} I_{n,sq}(x) &= \tfrac{1}{2}\tfrac{W_n}{L}\mu_{n,0}C_{ox}(v_{OD,n})^2(1+\lambda_n V_{DS}) \\ I_{p,sq}(x) &= -\tfrac{1}{2}\tfrac{W_p}{L}\mu_{p,0}C_{ox}(v_{OD,p})^2(1+\lambda_p V_{DS}) \\ I_{tot,sq}(x) &= I_{p,sq}(x) + I_{n,sq}(x) \end{aligned} \tag{5.13}$$

$$g_{m,tot}(x) = \begin{cases} K(V_B - V_{TH} - x) & x < -(V_B - V_{TH}) \\ 2K(V_B - V_{TH}) & -(V_B - V_{TH}) < x < V_B - V_{TH} \\ K(V_B - V_{TH} + x) & x > V_B - V_{TH} \end{cases} \tag{5.14}$$

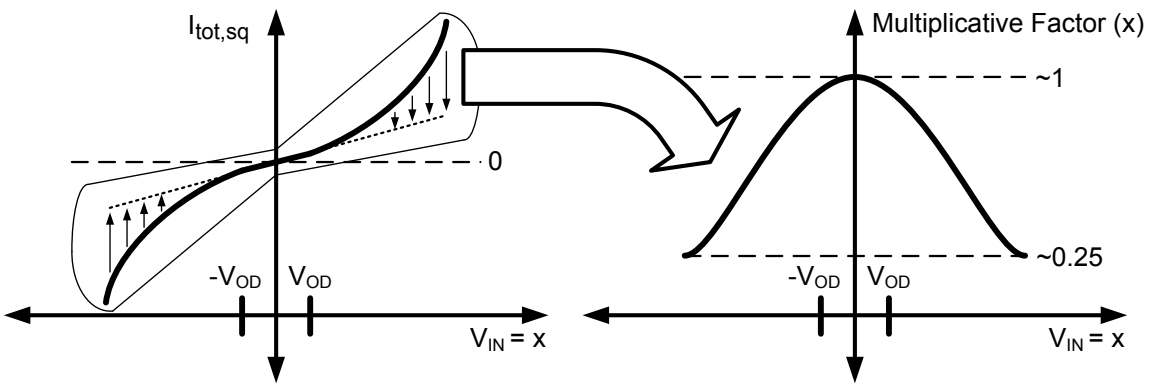


Figure 5.7: Intuitive shape of multiplicative term required to effectively linearize LNTA.

In order to create a I-V characteristic that is roughly linear over the entire input span, it is required to reduce both the current (and equivalently, transconductance) as an increasing function of $\max(|x - V_{OD}|, 0)$. As depicted in Fig. 5.7, the current and transconductance can be multiplied by a function decreasing in magnitude from a factor of 1 to a factor of 0.25 to bring the overall I-V characteristic to a condition in which the total mean squared deviation from the linear condition is significantly reduced.

It might appear at first glance that subtraction of the nonlinear portion of the characteristic would be a more sensical approach than multiplication. One reason that this is not the case is that additional subtracting devices placed in parallel with the existing input devices would invariably add more noise to the system. A second reason is that nonlinear effects operating on the MOS devices in the LNTA are already effectively multiplying the input current by a function decreasing roughly as $\max(|x - V_{OD}|, 0)$, as described in the following section.

5.5.2 Mobility Degradation Effects

The current-voltage characteristic of the short-channel MOS devices utilized in this work is already linearized to a certain extent by the presence of mobility degradation effects. As described in [95, pg. 589], both the horizontal and vertical electric fields in the drain-to-source channel reduce the effective mobility of the transistor as a function of gate-to-source voltage (5.15), where v_{sat} is the saturation velocity of carriers within the channel. For the purposes of the analysis here, this effect is consolidated into multiplicative factors $VS_n(x)$ and $VS_p(x)$ that are used to multiply each of the two device type currents in (5.13).

$$\mu_{eff} = \frac{\mu_0}{1 + (\theta + C)v_{OD}} = \mu_0 VS(x) \tag{5.15}$$

$$C = \frac{\mu_0}{2L_{eff}v_{sat}} \tag{5.16}$$

Due to the different scattering effects of holes and electrons, both μ_0 and θ differ considerably between NMOS and PMOS devices. As a result, VS_n and VS_p can differ considerably, lending an asymmetry to the large-signal operation of the LNTA half-circuit. This asymmetry, alluded to

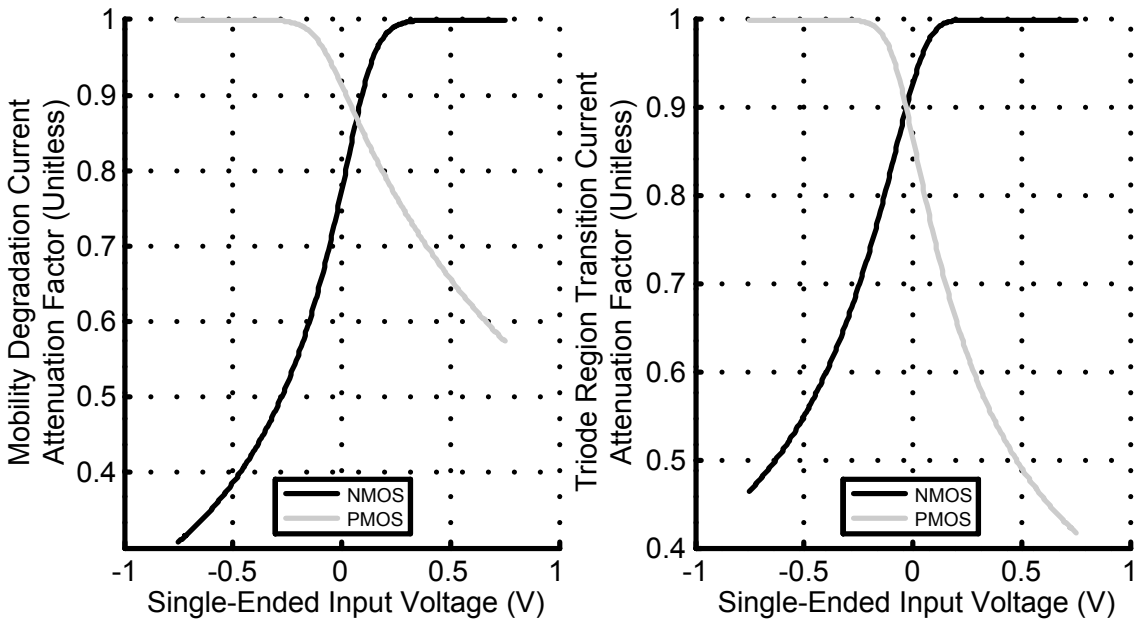


Figure 5.8: Nonlinear effects contributing to the effective linearization of the LNTA for large signals. a) Mobility degradation ($VS_n(x)$, $VS_p(x)$). b) Transition into triode region ($F_{s,n}(x)$, $F_{s,p}(x)$).

earlier in Section 5.4, results in the generation of even-order distortion products that are isolated from subsequent stages of the receiver by the ac-coupling capacitor (for low-frequency IMD products) or by frequency separation (for high-frequency harmonics and IMD products). The asymmetry also results in different large-signal dc and low-frequency IM2 currents that must be sourced by the class-AB OTA at the center of the LNTA. For L_{eff}=70nm and $v_{sat} = 1 \times 10^5$ m/s and for extracted values of θ_p, θ_n, $\mu_{0,n}$, and $\mu_{0,p}$, $VS_n(x)$ and $VS_p(x)$ are plotted in Fig. 5.8a.

Incorporating these effects into the MOS I-V relation yields (5.17)

$$I_{tot,md}(x) = I_{p,sq}(x)VS_p(x) - I_{n,sq}(x)VS_n(x) \tag{5.17}$$

Plotting (5.17) as a function of input voltage yields the asymmetric curve in Fig. 5.9. Although the asymmetry of Fig. 5.9 obscures the full extent of the improvement in linearity, taking the differential output current using the relation in (5.18) (which neglect the effects of differential mismatch) yields a visually superior result. As even-order distortion products are of minimal concern in the current output of the LNTA, the relation in (5.18) is the one of immediate relevance to the design, as odd-order IMD products end up dominating the in-band output error.

$$I_{diff,tot,md}(x) = I_{tot,md}(x) - I_{tot,md}(-x) \tag{5.18}$$

Taking the derivative of (5.18) with respect to the input voltage yields the differential transconductance as a function of input voltage. Plotting both as a function of the input voltage as in Fig. 5.10 makes evident the reduction in integrated mean squared error along each of the curves. The result shown in Fig. 5.10 suggests that if another function similar to that resulting from mobility

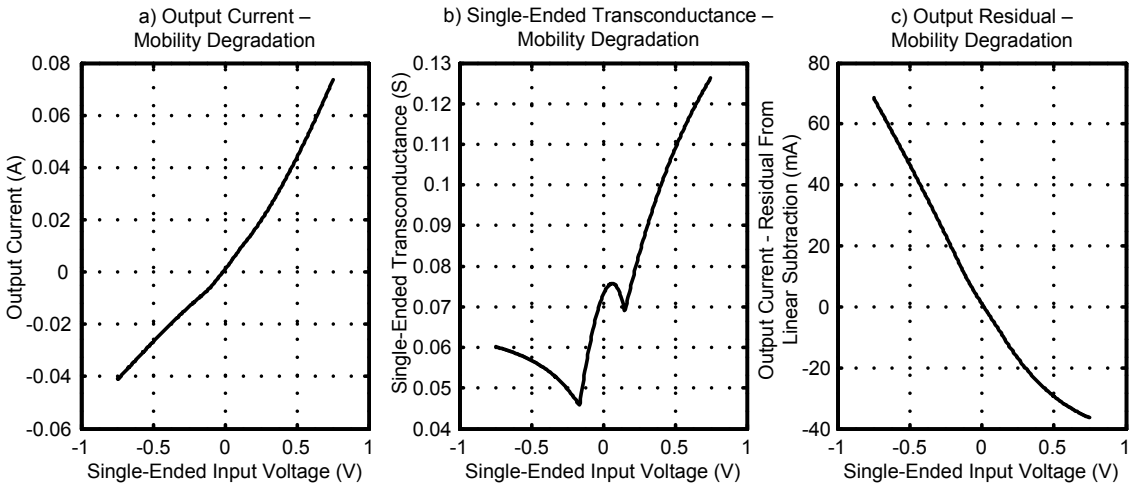


Figure 5.9: Calculated output characteristics of single-ended LNTA for MOSFET model incorporating mobility degradation effects. a) Output current. b) Transconductance. c) Residual after removal of small-signal linear fit.

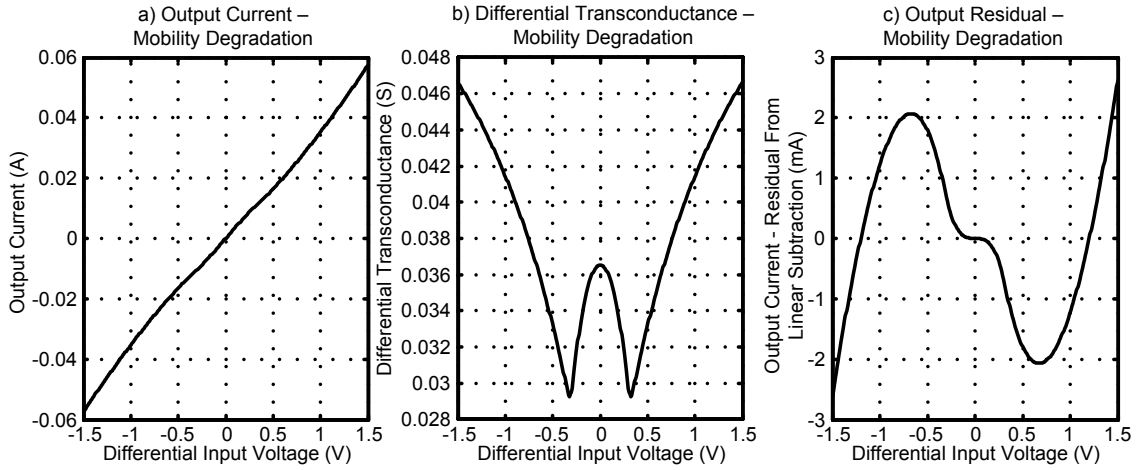


Figure 5.10: Calculated output characteristics of differential LNTA for MOSFET model incorporating mobility degradation effects. a) Output current. b) Transconductance. c) Residual after removal of small-signal linear fit.

degradation effects were present, the resultant I-V characteristic would be nearly linear and the g_m-V characteristic nearly constant.

5.5.3 Transition into Triode Region and Body Effect

In addition to providing a significant measure of isolation between the receiver input and the passive mixer, the presence of the cascode devices also maintains a relatively constant drain-to-source voltage across the input devices when the input devices are in strong inversion, as described in Section 5.3.2. As the magnitude of the input voltage increases, increasing the gate-to-source voltage of the

input device, the input device eventually transitions into the triode region of operation. Multiplying the current-voltage relation by a "saturation function" F_s [115] can account for the effect of this operating region transition on the drain current of the input transistor. Given by (5.19), F_s is seen to approach unity as the drain-source voltage $V_{DS}^{'} >> V_{D,SAT}$, to equal $\frac{1}{2}$ when $V_{DS}^{'} = V_{D,SAT}$ and to approach zero as $V_{DS}^{'} \rightarrow 0$, where $V_{DS}^{'} = V_{DS} + I_D(R_S + R_D) \approx V_{DS}$.

$$F_s = \frac{\frac{V_{DS}^{'}}{V_{D,SAT}}}{(1 + (\frac{V_{DS}^{'}}{V_{D,SAT}})^{\beta})^{1/\beta}} \tag{5.19}$$

For the purposes of this analysis, β_p is taken to be 1.4 for PMOS devices and β_n to be 1.8 for NMOS devices as in [115]. Although in principle $V_{D,SAT} = \frac{v_{OD}}{1+C(v_{OD}/(1+\theta v_{OD}))}$ to account for mobility degradation effects, it was found during the course of this study that modeling $V_{D,SAT} = v_{OD}$ produced a more faithful representation of the qualitative results obtained in simulation.

Due to this smooth transition into the triode region, the equality of V_{DS} and $V_{D,SAT}$ does not happen as quickly as predicted by the idealized model presented in Section 5.3.2. As the input voltage x increases in magnitude, the active input device begins entering the linear region, decreasing the magnitude of $\partial I/\partial x$ from what it would have been in the saturation region. As the active cascode device remains in the saturation region, the term $\partial v_{GS,casc}/\partial x$ also decreases in magnitude from what it would have been if the active input device had remained well within the saturation region. Therefore, the drain-source voltage of the active input device actually increases a small amount as the input voltage magnitude approaches its maximum, slowing the transition into the triode region.

Furthermore, because the source voltage of the cascode device is larger in magnitude than the source voltage of the input device, the change in source voltage of the cascode device as a function of input current is smaller than the change in source voltage of the input device as a result of the body effect. Due to this effect, the drain-to-source voltage of the active input device increases further still as the input voltage magnitude approaches its maximum.

It was found in simulation that incorporating this effect into the analytical model roughly entailed modifying (5.19) to (5.20), where K=0.2. Replacement of V_{DS} with $V_{DS,Init} \pm Kx$ was also performed in the modeling of the channel-length modulation. Here, $V_{DS,Init}$ is the value of V_{DS} when $x = 0$.

$$\begin{aligned} F_{s,n} &= \frac{\frac{V_{DS,Init}-Kx}{V_{D,SAT}}}{(1 + (\frac{V_{DS,Init}-Kx}{V_{D,SAT}})^{\beta_n})^{1/\beta_n}} \\ F_{s,p} &= \frac{\frac{V_{DS,Init}+Kx}{V_{D,SAT}}}{(1 + (\frac{V_{DS,Init}+Kx}{V_{D,SAT}})^{\beta_p})^{1/\beta_p}} \end{aligned} \tag{5.20}$$

Although the terms in (5.20) are of a different form than those in (5.15), when plotted as a function of x in Fig. 5.8b, they have roughly the same qualitative behavior. Adding them to the total current equation yields (5.21).

$$I_{tot,lintrans}(x) = I_{p,sq}(x)VS_p(x)F_{s,p}(x) - I_{n,sq}(x)VS_n(x)F_{s,n}(x) \tag{5.21}$$

For even more accurate modeling, the body effect is added to the overall current equation through the substitutions in (5.22).

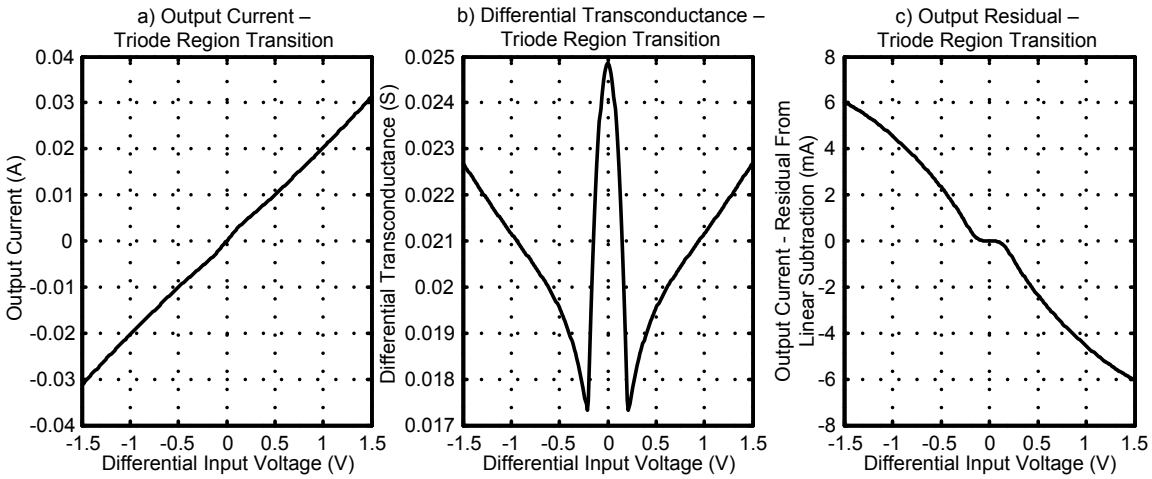


Figure 5.11: Calculated output characteristics of differential LNTA for MOSFET model incorporating continuous transition from MOS saturation to triode region. a) Output current. b) Transconductance. c) Residual after removal of small-signal linear fit.

$$\begin{aligned} V_{TH,p} &= V_{TH,0,p} + \gamma_p[\sqrt{2\Phi_F + V_{SUP} - x} - \sqrt{2\Phi_F}] \\ V_{TH,n} &= V_{TH,0,n} + \gamma_n[\sqrt{2\Phi_F + V_{SUP} + x} - \sqrt{2\Phi_F}] \end{aligned} \tag{5.22}$$

Manipulating (5.21) in the fashion of (5.18) to obtain the differential I-V relation after these modifications to the LNTA model yields the I-V and g_m-V characteristics in Fig. 5.11.

It is important to note that while the effects due to mobility degradation in the MOS devices are mostly predetermined due to the constraint that the device lengths (especially those of the PMOS) must be near the minimum allowable by the process lithography, the properties of $F_{s,p}$ and $F_{s,n}$ may be engineered in order to achieve quasi-linear results qualitatively similar to those depicted in Fig. 5.11 by altering the abruptness of the transition of the triode region. This can be done by changing the size or gate bias of the cascode device, the precise values of which are not critical to the small-signal operation of the LNTA.

5.5.4 Modeling Subthreshold MOS Conduction

Infinite derivatives of the g_m-V curve as seen in Fig. 5.11b cannot exist in a physical device. Properly modeling this section of the curve requires a smooth transition of the MOS region of operation from strong to weak inversion. This can be done by replacing v_{OD} with the expression in (5.24), where the value of "n" represents the subthreshold parameter and can be modeled as $1 + \frac{\gamma}{2\sqrt{2\Phi_F - V_{BS}}}$ [116]. For the low threshold voltage transistors used in the LNTA design, $n \approx 1.05$ and is only a very weak function of the value of x.

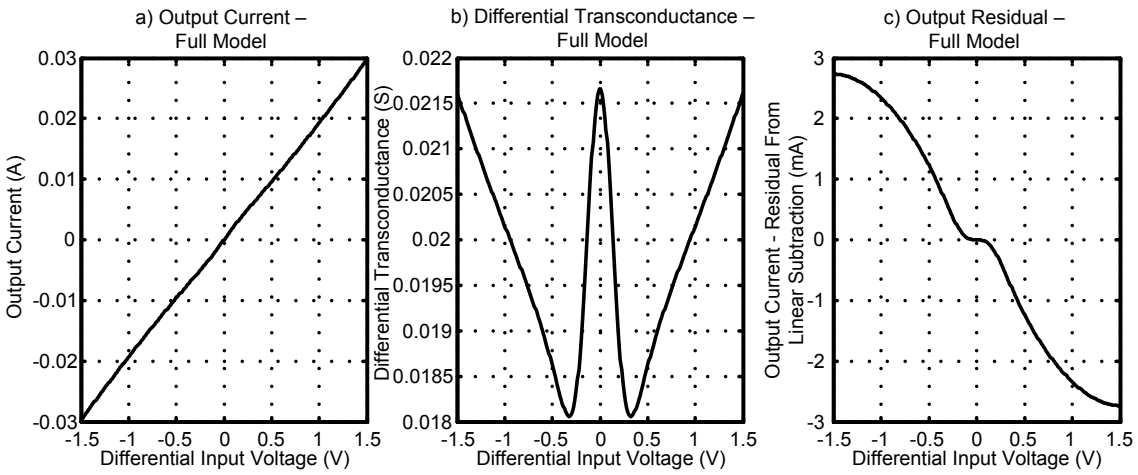


Figure 5.12: Calculated output characteristics of differential LNTA for MOSFET model incorporating subthreshold conduction effects. a) Output current. b) Transconductance. c) Residual after removal of small-signal linear fit.

$$v_{OD,n} = 2nV_t \ln(1 + \exp((V_B - V_{TH,p} + x)/(2nV_t))) \tag{5.23}$$

$$v_{OD,p} = 2nV_t \ln(1 + \exp((V_B - V_{TH,n} - x)/(2nV_t))) \tag{5.24}$$

$$V_t = \frac{kT}{q} \tag{5.25}$$

With these changes made to the model, the I-V and g_m-V curves appear as in Fig. 5.12. In this case, the transconductance is constant to within +/- 10% over the entire input range and deviates no more than 20% from the value at the quiescent point.

5.5.5 Intuition of Device Physics Applied Towards General Concept

In summary, the twin effects of the mobility degradation and the smooth transition of the input device operating region from saturation to linear result in a steady attenuation of the output current and net transconductance as a function of input voltage. Taken together, the composite attenuation functions applied to the NMOS and PMOS devices approximate a Gaussian-like function, as depicted in Fig. 5.7b. As such a function is repeatedly applied (Fig. 5.13), the sides of the g_m-V plot are progressively curved inwards to form the W-like shape shown in Fig. 5.12b. Because the initial simple-square-law model featured a discontinuity when one of the devices enters the cut-off region, the negative dips in transconductance persist as progressive refinements to the MOS modeling are applied. For this reason, the final curve has a W-like shape that can not qualitatively be improved upon.

Although the foregoing description is intuitively accurate, from a mathematical standpoint it is only half correct, considering that the composite I-V curve can be decomposed into two functions. One, $I_{tot,sqr}(x)$, represents the original square-law current relation while also accounting for sub-

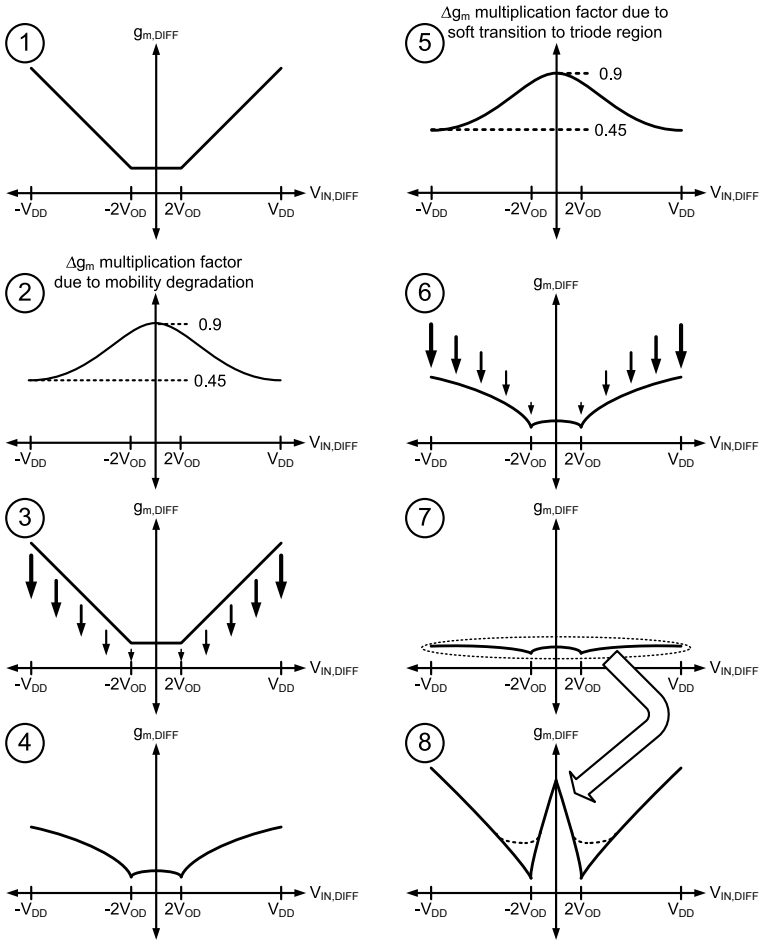


Figure 5.13: Intuitive illustration of generation of transconductance W-shape.

threshold and body effects. The other, $F_{atten}(x)$, represents the effects of the mobility degradation. The transconductance is given by (5.26). Using the chain rule it can be seen that the above intuition explains only the first half of the transconductance relation.

$$\begin{aligned} I_{tot}(x) &= I_{tot,sqr}(x)F_{atten}(x) \\ g_m(x) &= \frac{\partial I_{tot}(x)}{\partial x} = \frac{\partial I_{tot,sqr}(x)}{\partial x}F_{atten}(x) + I_{tot,sqr}(x)\frac{\partial F_{atten}(x)}{\partial x} \end{aligned} \tag{5.26}$$

Taking the derivative of a Gaussian function and multiplying it by the square-law current relation generates another W-shaped curve which is then added to the original as depicted in Fig. 5.14, reinforcing the final characteristic W-shape of the g_m-V curve.

5.5.6 Total IMD Product Error as Function of Input Voltage Magnitude

Maintaining a relatively constant transconductance in the LNTA in the manner described above substantially reduces the magnitude of the composite IMD products generated when large signals are

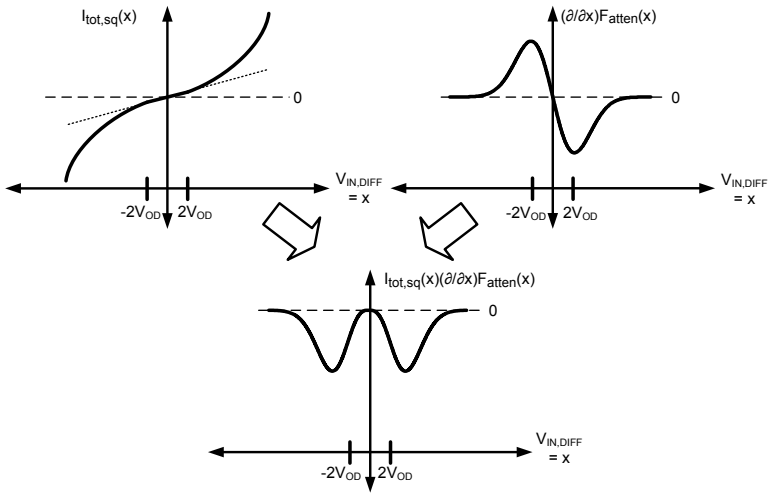


Figure 5.14: Intuitive illustration of generation of transconductance W-shape - second chain rule term.

applied. One way to think about this is to consider that the shape of the LNTA g_m-V curve starts out very similar to a pure 2nd-order function (implying a pure third-order nonlinearity). Therefore, in a swept-amplitude IIP3 simulation using CW tones, the input-referred error will begin rising with a slope of 3 in a manner similar to an ordinary RF circuit. As the amplitude of the input blockers approach the minima of the g_m-V curve, the rate of increase in IMD product amplitude will decrease such that the input-referred IMD error curve begins to flatten out. Since the W-shape of the g_m-V curve implies a sinusoid-like effective nonlinearity (because as the g_m-V curve is integrated to obtain the I-V curve, a sinusoid integrates to a co-sinusoid with a 90° phase shift), the terms of the Taylor series which approximate the nonlinearity alternate in sign, providing for a degree of cancellation of correlated IMD terms of higher order. As the higher-order IMD terms begin to cancel each other as the inflection point is passed, the slope of the input-referred error as a function of blocker magnitude may even turn negative for some period.

Eventually as the input signal voltage magnitude increases, it traverses the outer tails of the W-shaped g_m-V characteristic. When zoomed out, as in Fig. 5.15, the W-shape again appears to be roughly a pure second-order function (implying a pure third-order nonlinearity). Therefore it is expected that the input-referred error rises again at a slope of 3 for very large input blocking signals.

In fact, this predicted behavior is observed when a two-CW tone blocking condition is applied to the I-V transfer characteristic. In Fig. 5.16a, two equal CW blockers are applied to the I-V model in Fig. 5.15. The input-referred IMD products at first rise with a slope of 3. Gradually the higher-order IMD products begin to cancel each other out until near-perfect cancellation is achieved. Beyond this point, the input-referred IMD products begin to rise with a slope of 3 again. However, at this point it can be seen in Fig. 5.17 that the input-referred IMD products are about 40dB less than they would have been had the IMD products continued to rise at a slope of 3 for the entire sweep.

Another very important type of blocking condition to be discussed in subsequent chapters is that of asymmetric blocking. In this case, one of the blocker signals is much larger than all of the

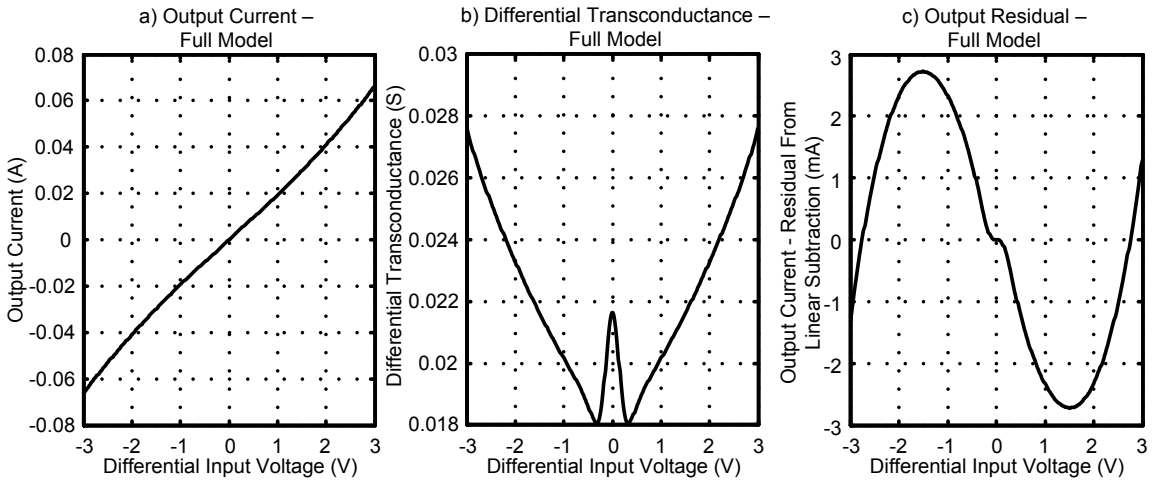


Figure 5.15: Calculated output characteristics of differential LNTA for MOSFET model incorporating subthreshold conduction effects, for wide input range. a) Output current. b) Transconductance. c) Residual after removal of small-signal linear fit.

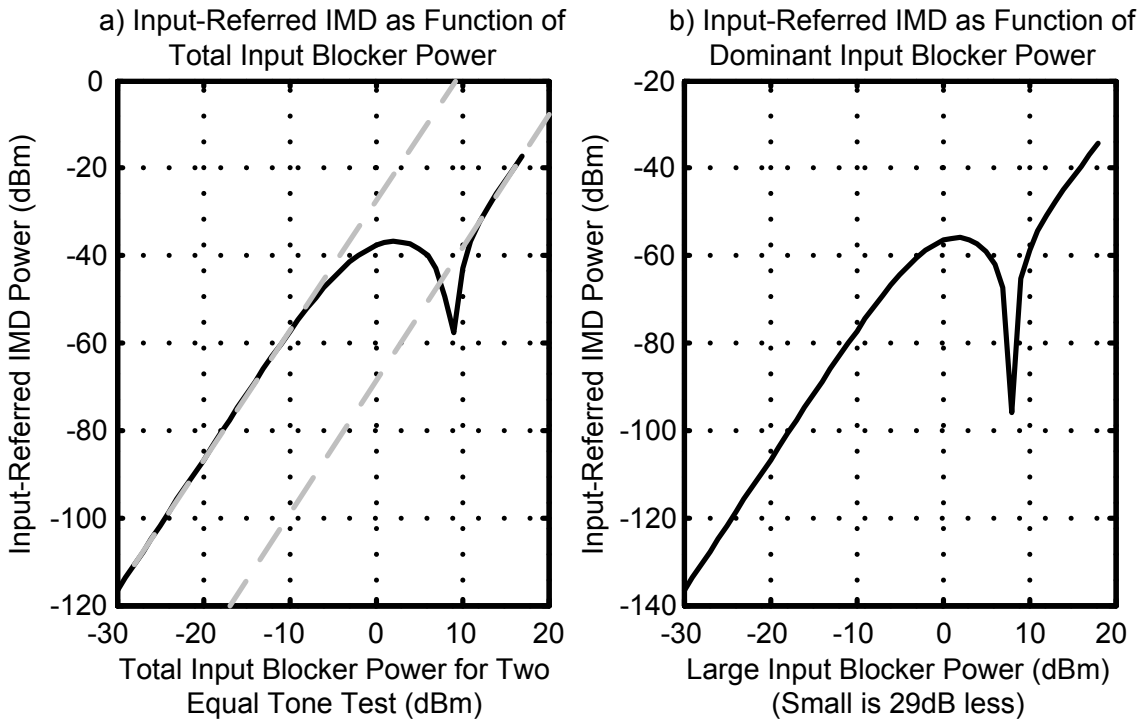


Figure 5.16: Calculated input-referred upper-sideband IMD products for two-CW-tone blocking scenario. a) For two equal magnitude CW blockers. b) For asymmetric CW blockers.

rest. This situation wherein the blocking scenario is dominated by one very large jamming signal arises frequently in applications such as FDD communications (where the TX leakage dominates), in implantable circuits requiring both power and data wireless links (where the power link dominates),

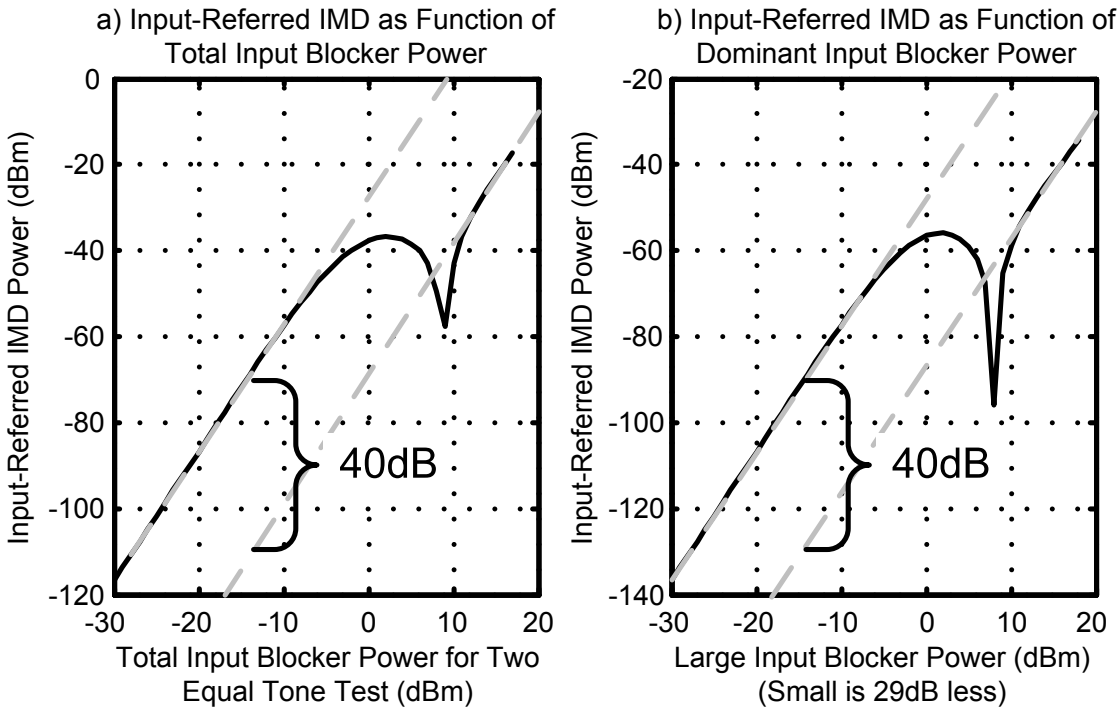


Figure 5.17: Calculated input-referred upper-sideband IMD products for two-CW-tone blocking scenario with slope-of-3 labels. a) For two equal magnitude CW blockers. b) For asymmetric CW blockers.

and in hostile military environments where large jammers are used to prevent enemy communications. In this case, a very large CW blocker is applied to the transfer function in Fig. 5.15 along with a smaller CW blocker of magnitude 29dB lower. We see in Fig. 5.16b that the conclusions of the symmetric blocking condition hold here as well.

5.5.7 LNTA Compression

A relatively constant transconductance in the LNTA also directly implies that the small signal compression resulting from the LNTA is also bounded. As a result, the out-of-band 1-dB desensitization point of the LNTA is very high, and may not even exist at all. To see why, the reader may consider a blocking scenario consisting of one large blocker and one small desired signal. In this case, the Taylor series of the nonlinear I-V curve may be taken with respect to the small desired signal and the output expressed as (5.27).

$$i_{OUT}(t) \approx g_m(v_{BLOCKER}(t))v_{DESIRED}(t) \tag{5.27}$$

The small signal gain change as a function of input voltage can be computed as (5.28).

$$\Delta_{Gain} = E[g_m(v_{BLOCKER}(t))] - g_m(0) \tag{5.28}$$

Plotting Δ_{Gain} for the g_m-V relation responsible for the curve in Fig. 5.15b and a large sinusoidal

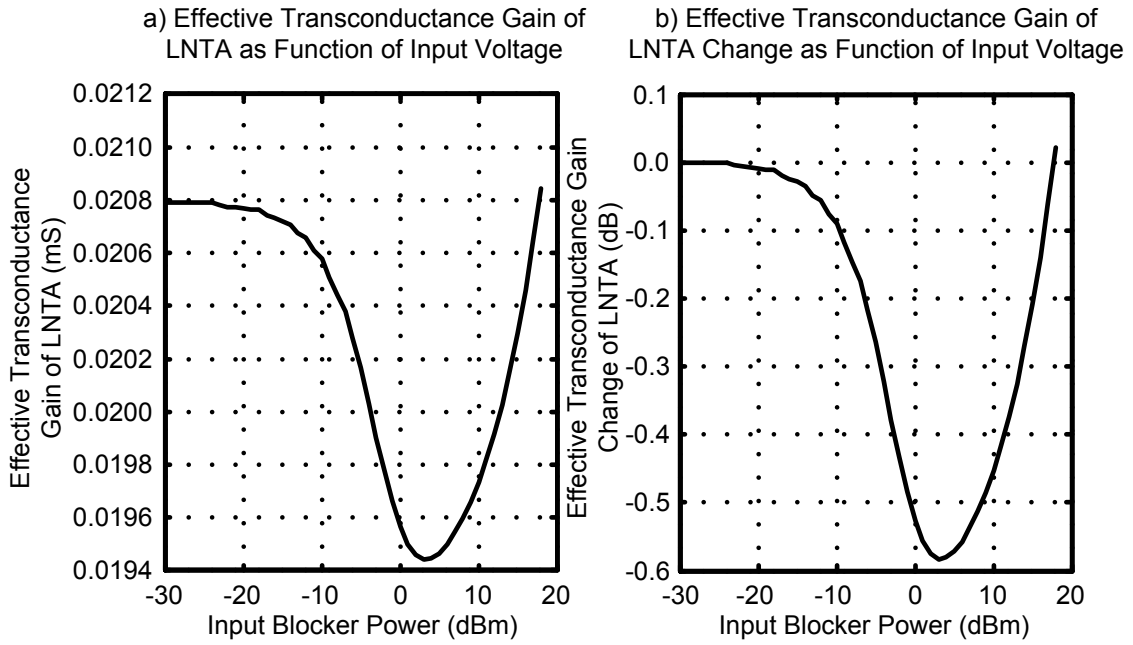


Figure 5.18: Calculated input-referred LNTA out-of-band desensitization (160MHz offset) for a CW-tone blocking scenario. a) Effective transconductance. b) Relative transconductance.

blocker results in the curve shown in Fig. 5.18. Like the output IMD products, the nonmonotonicity of this curve follows directly from the nonmonotonicity of the g_m-V characteristic. In this case, the 1-dB desensitization point *does not exist*. In reality, the 1-dB desensitization point of the LNTA will be limited by the current-shunting action of the input device parasitic source-to-bulk diodes once the diode turn-on voltage is exceeded.

5.6 Implementation and Simulation Results

Based on the foregoing calculations and intuition, a stacked push-pull common-gate LNTA was designed using low-threshold voltage devices in a 90-nm RF CMOS process. Shown using a simplified schematic in Fig. 5.19, the gate voltages of the input devices are defined with replica bias loops that ensure that the quiescent current flowing through the structure is proportional to the input bias current. The source voltages of the input devices are defined at $V_{DD}/2$ by the OTA with the class-AB architecture dictated by the requirements imposed by the large-signal mismatch between PMOS and NMOS devices.

5.6.1 Static Swept Input Simulations

The static current and transconductance of the LNTA are obtained by removing the ac-coupling capacitors from the output terminals and replacing them with ideal voltage sources that monitor the output current. Although memory effects are not taken into account in this simulation, it is found that the results gleaned from this simulation correspond well to the results involving the application of high-frequency signals to the LNTA while producing a reassuring correspondence

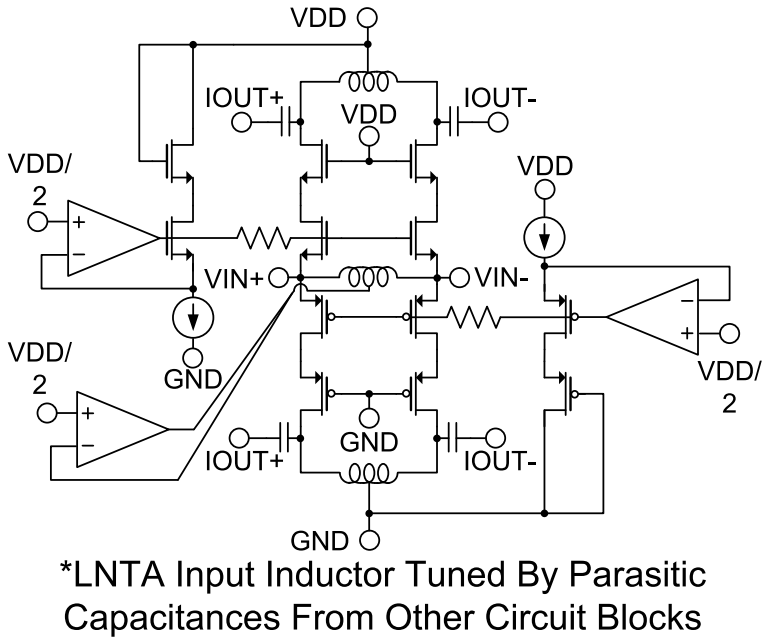


Figure 5.19: Simplified schematic of implemented LNTA.

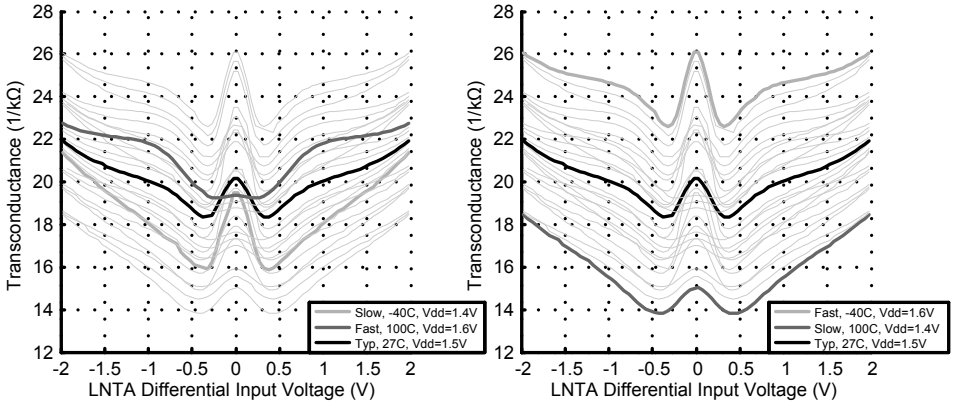


Figure 5.20: Static LNTA simulation: transconductance as function of differential input voltage over corner.

with the memoryless analytical results presented in the foregoing sections. Furthermore, ideally terminating the LNTA output reflects the fact that the frequency-translated baseband impedance seen at the output of the LNTA by large out-of-band blockers in a real receiver implementation is very small so long as large passive mixer switches and a large baseband filtering capacitor is used to generate a low impedance at the blocker baseband offset frequency.

Subtracting the currents from the two halves of the LNTA yields the differential output current and removes the second-order nonlinear terms from the I-V characteristic. Taking the derivative of the output I-V characteristic yields the g_m-V characteristic, which is plotted in Fig. 5.20. Removing the linear fit at the quiescent point from the output I-V characteristic also yields the residual-V

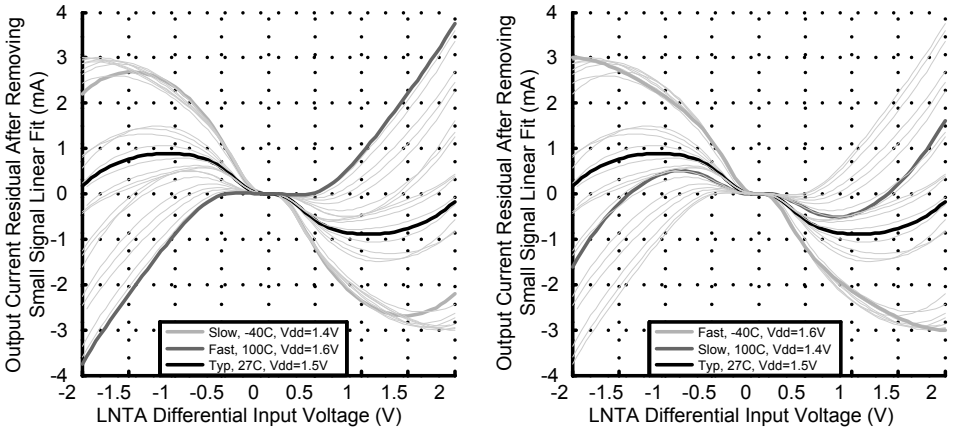


Figure 5.21: Static LNTA simulation: small signal linear fit residual as function of differential input voltage over corner.

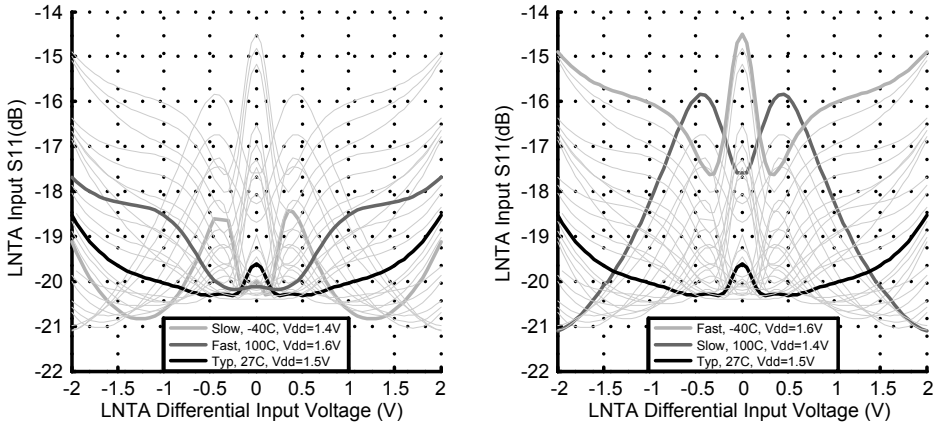


Figure 5.22: Static LNTA simulation: S_{11} as function of differential input voltage over corner.

characteristic, plotted in Fig. 5.21. These simulation results are obtained over 3-σ MOS device corner, temperature, and supply voltage variation. Interesting extrema of the corner sweeps are highlighted in Figs. 5.20 and 5.21.

It is important to note that the constant-transconductance behavior of the LNTA also guarantees a good 50Ω match to the antenna over input signal swing. This is tested during the static sweep by performing an ac analysis at each point of the dc sweep. The results of this test shown in Fig. 5.22 reveal that the return loss S_{11} varies only a few dB from a minimum of -20dB over the typical case rail-to-rail signal swing and never exceeds -13dB.

The foregoing simulation results were obtained via constant-current biasing over MOS device corner, temperature, and supply voltage variation. Even if all three parameters are simultaneously varied to their extremes, the defining characteristics of the LNTA are still present, namely a relatively constant g_m-V curve and a nonlinear current residual that is on the order of a magnitude less than the peak output current of 30mA.

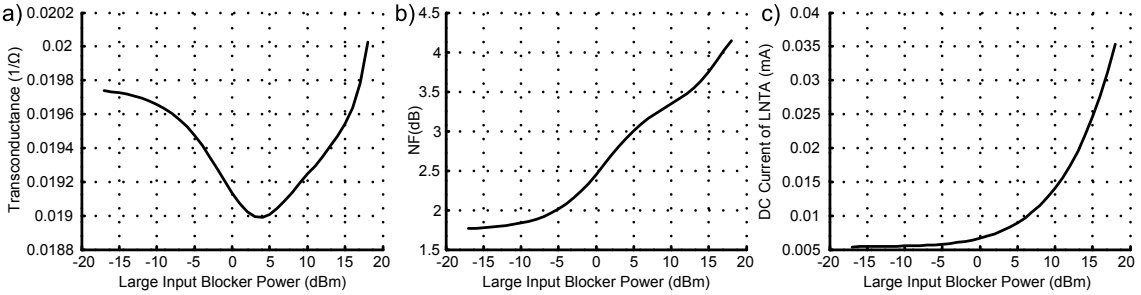


Figure 5.23: Miscellaneous dynamic LNTA simulations using QPSS/QPNOISE. a) Effective transconductance gain. b) Noise figure. c) Quiescent current.

5.6.2 Dynamic Swept Input Simulations

In order to confirm the link between the static swept input simulation and actual operation, several QPSS simulations were performed to examine the dynamic performance of the LNTA under typical conditions. Due to the fact that in this case, the peak-to-peak transconductance and nonlinear current residual variation are about half that of the calculated model, it is expected that compression effects and input-referred IMD products will be about half of that predicted using the calculated model. In this case, the ESD typically present at the input of the LNTA is removed, although an ideal capacitor modeling the total capacitive load at the LNTA/chip input is utilized.

5.6.2.1 Compression Simulations

Using a QPSS simulation, the small signal transconductance gain of the LNTA is determined by applying a small desired CW signal at 2.14GHz along with a large blocker CW signal at 1.98GHz (a 160MHz offset). It is seen that over the course of the input sweep, the LNTA gain drops by at most only 0.3dB as shown in Fig. 5.23a. If the input blocker power were increased past the maximum value shown on this plot, the transconductance curve would eventually reach a point where it drops sharply, corresponding to the input power at which the parasitic source-to-bulk diodes of the input devices turn on.

5.6.2.2 Noise Simulations

As the input blocker signal increases in magnitude, the output noise increases due to the reduced source degeneration available to the cascode devices. The input devices only approach the edge of the triode region and hence the noise figure due to these devices remains the same at $1 + \gamma$, which is independent of the device transconductance. In any event, the transconductance changes little. Performing a QPNOISE analysis while sweeping the input blocker as a function of blocker amplitude results in the curve shown in Fig. 5.23b. It is seen that the noise figure rises from about 1.8dB to 4.2dB. Although this is a relatively large increase, it should be taken into the context that most wireless communications standards effectively allow for a 3dB increase in input-referred error power under blocking conditions.

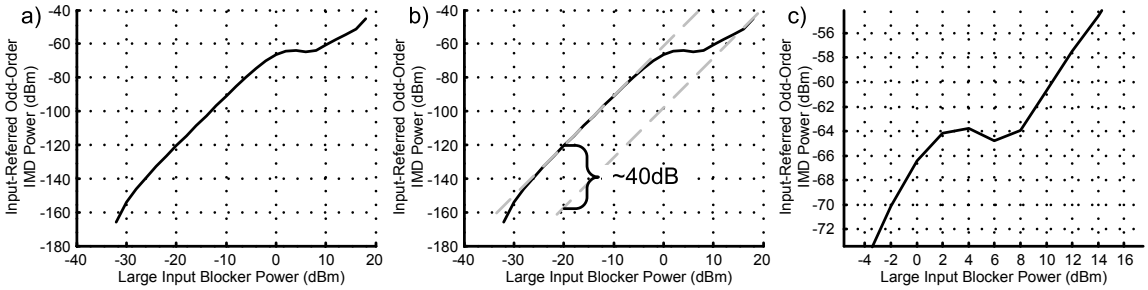


Figure 5.24: Dynamic LNTA odd-order two-tone IMD simulations using QPSS (second tone 29dB less than large tone). a) Zoomed out. b) Annotated. c) Zoomed in.

5.6.2.3 Quiescent Current Simulations

As described in Section 5.4, a large input signal is rectified by the second-order nonlinearities present in the LNTA devices and its dc component shunted to the supply rails through the output inductors. This effect is quantified in Fig. 5.23c, where the quiescent current of 5.4mA (including biasing) increases to 18.5mA for a +12.5dBm CW input blocker and to +35.3mA for a +18dBm CW blocker.

5.6.2.4 Intermodulation QPSS Simulations

In order to confirm that the reduction in IMD at large signal amplitudes still holds for high-frequency inputs, an asymmetric two-tone blocking input was applied to the LNTA. The larger CW blocker was placed at a 160MHz LO offset, while the smaller CW blocker was placed at a 320MHz LO offset. Although the deep notch present in Fig. 5.16 has been noticeably attenuated, the reduction in IMD at large signal amplitudes is still about 40dB over what would be predicted from extrapolating the slope-of-3 characteristic from smaller blocker magnitudes. It is likely that the presence of memory effects within the LNTA do not permit the precise cancellation of IMD products required to effect the deep notch seen in the memoryless analytical result.

It is also interesting to note that the simulated input-referred IMD is about 10dB less than the results shown in Fig. 5.16. Accounting for 6dB of this discrepancy is the fact that the peak-to-peak deviation of the static simulation transconductance and residual in Figs. 5.20 and 5.21 are about half those in Fig. 5.12. Another 3dB is accounted for by the fact that the features of the curves present in Fig. 5.12 are compressed by 1dB with respect to the features present in Figs. 5.20 and 5.21. This effective compressive scaling of the x-axis by 1dB results in a 3dB increase in the odd IM3 products that dominate the slope-of-3 portions of Figs. 5.16 and 5.24.

5.7 Robustness of Constant-Transconductance Behavior

This physical intuition behind the W-like shape of the g_m-V curve also explains its robustness in the face of PVT variation. To understand why, the reader may consider the concepts developed in each of the subsections of Section 5.4.

1. The shape of the initial square-law I-V and g_m-V relations are relatively independent of PVT

variation - only the mobility and threshold voltages change, altering the magnitude and break-points of the curve, respectively.

2. In a deep submicron process, the mobility degradation is largely a result of normal-field effects that depend on the well-controlled thickness of the MOS oxide layer [117]. Although the intrinsic carrier mobilities may vary +/- 50% with doping and temperature, this will only partially effect one of the two linearization mechanisms.

3. The transition into the triode region depends on the supply voltage, device threshold voltages, and device overdrive voltages V_{OD}. Using the relation (5.20) as a proxy, we see that its initial behavior is determined by $V_{DS}/V_{D,SAT}$. Relating $V_{DS}/V_{D,SAT}$ to the aforementioned quantities results in (5.29).

$$\frac{V_{DS}}{V_{D,SAT}} \approx \frac{V_{DD} - V_{GS} - V_{DD}/2}{V_{GS} - V_{TH}} = \frac{V_{DD}/2 - V_{TH}}{V_{OD}} - 1 \tag{5.29}$$

In general, the supply voltage is well-regulated and does not vary much. The value of $V_{OD} = V_{GS} - V_{TH}$ for large input signals is dominated by the input signal itself (since the magnitude of V_S will be in general larger than $V_G - V_{TH}$ when the device is on. The value of V_{OD} when the input is small is determined by the device mobility. Given that for small V_{OD} the MOS relation approximates the canonical square-law model, for a constant-g_m biasing, it can be shown that the overdrive voltage V_{OD} is inversely proportional to the device intrinsic carrier mobility. For constant-current biasing it can be shown that V_{OD} is inversely proportional to the square root of the device intrinsic carrier mobility. In the latter case, a +/- 50% change in mobility results in only a +/- 22% change in V_{OD}. Based on the values used for the models above, this results in F_s varying by about +/- 10%, which is a dramatic attenuation of effect from the initial +/- 50% change in mobility. Taken together, in the presence of PVT variation, the shape of the multiplication functional representing the transition into the triode region in the absence of change in V_{DS} is relatively robust.

5.8 Heuristic Modeling of LNTA Nonlinearity

When fully expanded, the equation used to model the I-V characteristic of the LNTA is rather unwieldy and can be cumbersome to apply. This equation may not precisely reflect simulated and/or measured data (as in the case above) and a fitted model to the macroscopic behavior to the LNTA is desirable for proper behavioral simulations. A clue as to a possible compact basis set for the LNTA transfer characteristic can be seen from the curves shown in Fig. 5.13. In this case, the original I-V characteristic is largely cubic in nature. This cubic relation is then multiplied by two relations of form similar to (5.15). Multiplying the two linearizing relations together yields a curve that looks very similar to a Gaussian function. Given that approximating a function with a basis set of functions that "look like" the original functions makes intuitive sense, it is proposed here to model the LNTA nonlinearity with a set of functions of the form (5.30).

$$f(x, a, k) = ax^3 e^{-|k|x^2} \tag{5.30}$$

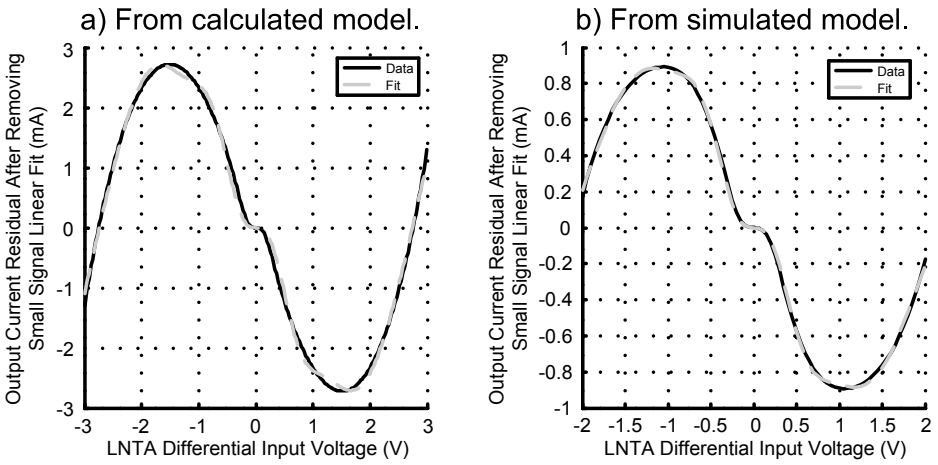


Figure 5.25: Cubic-Gaussian basis fit to nonlinear residual curves. a) Fitting to calculated model. b) Fitting to simulated model.

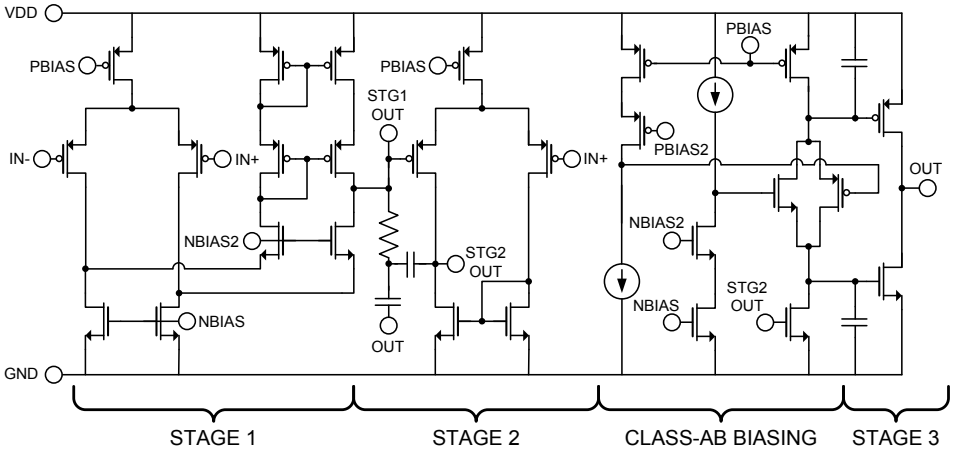


Figure 5.26: Simplified schematic of implemented center class-AB OTA for LNTA.

For example, the I-V curve (in units of mA and V) of the analytical model nonlinear residual in Fig. 5.15 can be well-represented for a differential voltage input by the summation of the functions f(x,0.08434,0), f(x,-2.199,0.435), and f(x,-12.59,2.5), as shown in Fig. 5.25a. The I-V curve of the simulated static nonlinear residual can be well-represented by the set of functions f(x,0.03545,0), f(x,-1.736,0.836), and f(x,-8.867,4.05), as shown in Fig. 5.25b.

5.9 Biasing OTA Implementation

Central to maintaining the proper input common-mode voltage of the LNTA is a class-AB biasing OTA that consumes little quiescent current but that can source enough current to correct for the large-signal imbalance between the NMOS and PMOS devices. Assuming that the maximum input power required is +12.5dBm, the maximum current to be sourced is 2mA. In order to maintain

a degree of margin, the OTA was designed to both sink and source up to 8mA. Because the real part of the load impedance to be driven by the OTA is nominally 12.5Ω (due to the two input terminals of the LNTA effectively in parallel), to obtain an appreciable degree of loop gain for a small quiescent power draw, a multi-stage OTA topology is required.

In order to satisfy these requirements, a three-stage reverse-nested-miller-compensated (RNMC) OTA topology [118] was chosen to set the LNTA common-mode voltage. The OTA schematic is shown in Fig. 5.26. It consumes a quiescent current draw of 900μA while achieving in the typical corner a minimum dc loop gain of 56dB, minimum unity-gain bandwidth of 4.7MHz, minimum phase margin of 74.6°, and minimum gain margin of 12.8dB over the output current swing range of +/- 8mA for a 1nF LNTA input blocking capacitor separating the LNTA from the driving source.

Chapter 6

Successive Regeneration and Adaptive Cancellation of Higher-Order Intermodulation Products[1]

6.1 Need for Cancelling Higher-Order Intermodulation Products

Any practical design of an RF receiver faces strict area and power requirements, creating pressure to operate the receiver with as little performance margin as possible. At the same time, compelling applications exist to motivate the use of a receiver with a rail-to-rail input signal. Taken together, these goals all but guarantee the presence of higher-order odd and even intermodulation distortion products within the receiver that must be cancelled, as depicted in Fig. 6.1.

6.2 Naïve Approach to Cancelling Higher-Order Intermodulation Products

One approach to cancelling the higher-order intermodulation products is to trivially extend the scheme presented in Chapter 2. As depicted in Fig. 6.2a, a reference branch for each required polynomial term may be designed each with a nonlinear term generator of an appropriate order followed by a downconversion mixer for the odd-order terms. Each reference branch would require a separate analog-to-digital converter followed by an adaptive filter whose weights are updated by a correlation operation on the output of the complete equalizer. In this case, the composite LMS-based adaptive equalizer (i.e. all of the adaptive filters combined) utilizes a basis of reference signals that is not only polynomial in the discrete time z-domain (assuming multiple transversal taps on the adaptive filters) but that is also polynomial in the signal amplitude domain.

The issue with using such an architecture is the large number of required amplitude polynomial terms. For example, to generate a fit for the simulated LNTA residual of Section 5.6 over the range

[1] Portions of this material have been previously published in [119] and copyright is owned by IEEE.

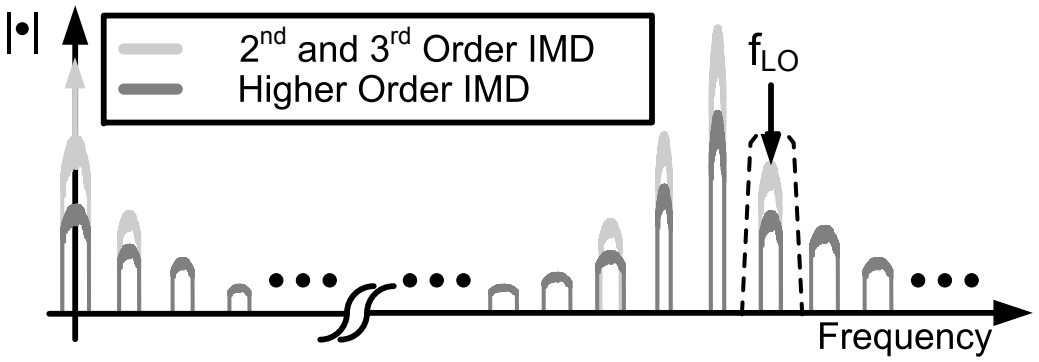


Figure 6.1: Depiction of higher-order intermodulation distortion products.

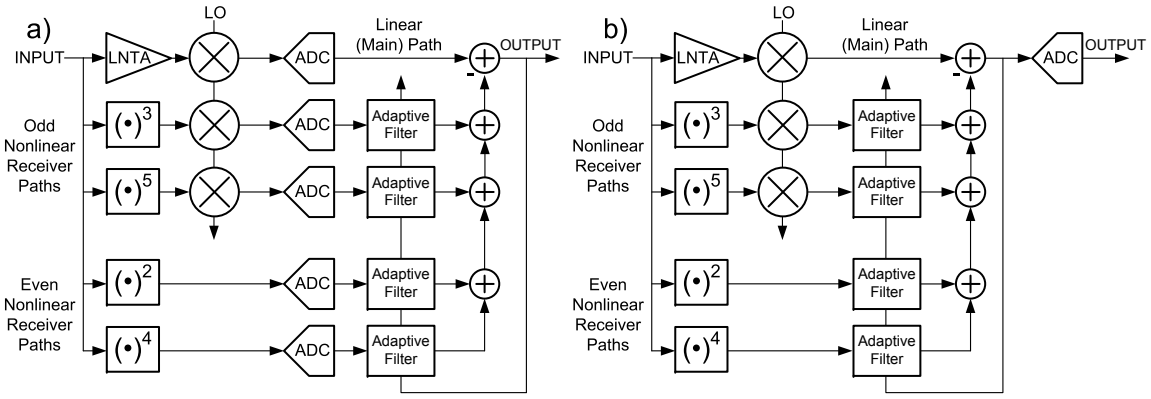


Figure 6.2: Naïve approaches to cancelling higher-order intermodulation distortion products. a) Trivial extension of work in Chapter 2. b) Moving adaptive cancellation algorithm from digital into analog baseband.

(-2V,2V) with rms error of 5.5%, a 15th-order polynomial is required, as seen in Fig. 6.3. For the odd-order reference basis terms, this results in 14 analog-to-digital converters alone.

A solution to the problem of having many analog-to-digital converters is to move the adaptive equalizers to the analog domain, as depicted in Fig. 6.2b. Although these circuits are better controlled and have historically been implemented in the digital domain, work performed in [33] has shown that accurate implementations can be had in the analog domain as well. However, in order to ensure that the adjacent channel signal in the main receiver path does not interfere with the correlation process of the adaptive algorithm, it must be completely filtered out prior to the adaptive equalizers. Meeting this requirement compels the use of fifth- to seventh-order low-pass filters in telecommunications standards such as UMTS and CDMA [61] [40] [62]. Combined with the analog adaptive algorithms, these filtering requirements would prove to be a major design, power, and area burden for the analog die. For both of these trivial solutions, creating seven odd-order intermodulation distortion reference circuits results in a similar burden for the analog die.

One method to cut down on the number of polynomial basis elements required to attain a particular degree of error cancellation is to develop an analog functional basis of elements similar

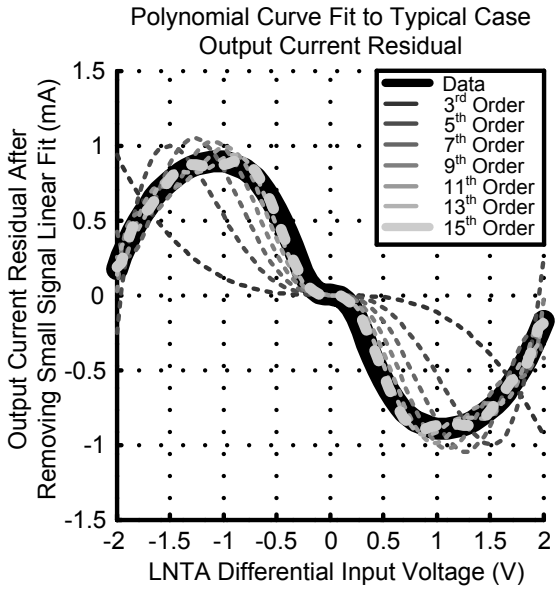


Figure 6.3: Polynomial fit to residual I-V characteristic of LNTA after small-signal linear fit is removed.

to that described in Sec. 5.8. In this case, the proposed reference basis of three elements results in a best-fit rms error of 2.3%. However, it is unclear how to design analog circuitry to generate such a basis without also generating a large linear term, that when applied to the adaptive algorithm, would result in the partial cancellation of the desired signal.

6.3 Proposed Approach to Cancelling Higher-Order Intermodulation Products

6.3.1 Primary Concept

From the foregoing discussion it appears that it would be best if higher-order intermodulation terms could be generated, but not in the analog domain. Ruling out the analog domain leaves only the digital domain. In this case, one can entertain the notion of multiplying together baseband IM2 and IM3 products together in the digital domain to generate higher-order terms, as conceptually depicted in Fig. 6.4a. A large-signal receiver architecture that employs such a scheme is shown in Fig. 6.4b. Here, IM2 and IM3 terms are multiplied together to create approximations to IM5 products, IM2 terms are multiplied by IM2 and IM3 terms to generate approximations to IM7 terms, and so on.

Not shown in Fig. 6.4b is the anti-alias filtering required in the nonlinear paths immediately prior to the analog-to-digital interface. Although these filters are not shown in Fig. 6.4b for simplicity, they may be seen in Fig. 7.2, the block diagram of an implemented receiver front end. Designing these anti-aliasing filters should be done in such a way that they are effectively memoryless across the bandwidth of the downconverted even-order reference products. Although a roughly constant group delay may still exist, this memory effect is trivially compensated for in the digital back end by adding

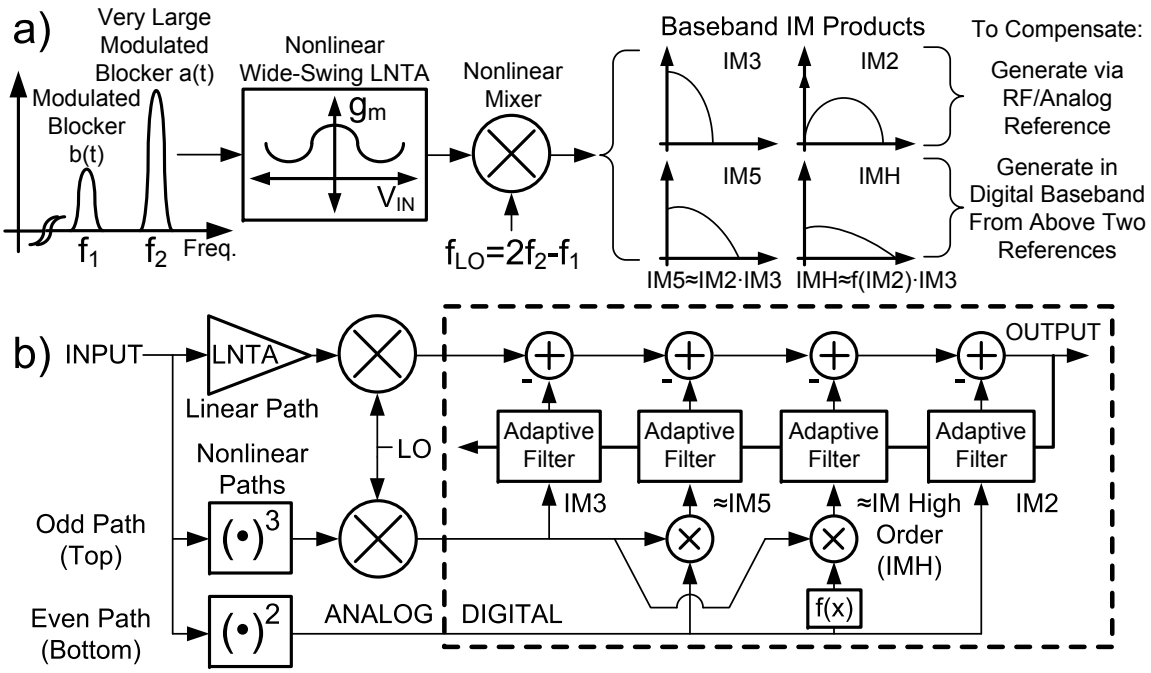


Figure 6.4: Successive regeneration and adaptive feedforward cancellation of intermodulation products at baseband implemented in this work. a) Concept. b) Simplified system block diagram.

compensatory delay elements where needed (hence the nomenclature "effectively memoryless"). The reason for this design constraint is that the memory effects present at baseband have the potential to deviate far more dramatically as a function of frequency than do those at RF. Significant frequency-dependent amplitude and group delay ripple in between multiplications cannot be compensated for later by multi-tap LMS or other filter types because the multiplication and filtering operations are not commutative. Keeping the anti-aliasing filters effectively memoryless across the bandwidth of the downconverted even-order reference products (as opposed to the channel bandwidth of the receiver) is required because the even-order reference IMD products will spread out-of-channel odd-order reference IMD products back in-channel. Hence, it may be important to maintain the accuracy of these signals even outside of the channel bandwidth. Obtaining this "effectively memoryless" characteristic for most conceivable scenarios entails designing the cutoff frequency of the nonlinear path anti-alias filter to be several times greater than that of the nominally linear path, with the exact difference depending on the expected modulation rates of the blocker signals.

Although this approach is reminiscent of the polynomial predistorters used in RF transmitters and referenced in Section 4.1 in that even-order baseband IMD products are successively multiplied, it is fundamentally different in several aspects. First, polynomial predistorters derive the exact envelope of the incoming signal, successively multiply this envelope, and then apply the resultant higher-even-order terms to the incoming signal at RF. In the proposed approach, all successive multiplication is performed at digital baseband after anti-alias filtering and the original IMD-generating signals are not available. This is a key point - the large IMD generating signals are not available because their frequencies are not known and, more importantly, because their frequency spacing may be too large to be accommodated by any auxiliary path analog-to-digital converters if it were

decided to downconvert and digitize the entire frequency band over which large undesired signals may lie. Rather, a derivative of the large input signals is utilized, namely the IM3 products, since their frequency is known. And, as described in Section 2, this frequency is known to be the receiver LO frequency when the baseband output is known to be corrupted.

Second, the polynomial predistorters listed in Section 4.1 reside only in the analog domain at RF. In principle they may also be implemented in baseband circuitry, either analog or digital. However it is neither required nor desirable to place a frequency-translating element (e.g. an LO-driven mixer) within the polynomial predistortion structure. By contrast, a frequency-translating element must be used in the proposed successive regeneration scheme proposed in Fig. 6.4 since IM3 terms must be generated at analog RF but higher-order IMD terms must be generated at digital baseband for reasons already discussed.

Third, the approach proposed here to regenerate nonlinear reference terms is neither one of predistortion or postdistortion. Rather, feedforward distortion cancellation as implemented in Fig. 6.4 is a parallel-distortion technique in which the composite nonlinear receiver branch is effectively matched to the nominally linear receiver branch via adaptive equalizers.

Fourth, unlike the polynomial predistorter that need only process directly relevant IMD terms, the proposed nonlinear reference generation scheme must deal with odd-order IMD products unrelated to the ones corrupting the signal band around the LO frequency of the receiver. When downconverted, these signals must be filtered out in order to avoid unintended aliasing to baseband during the digital sampler process. Hence, the placement and design of postfiltering within the analog IM3 and IM2 branches is critical in this case, as opposed to the case of a polynomial predistorter.

6.3.2 Cancellation of IMD Products at Higher-Order Frequency Offsets

If taken literally, the scheme depicted in Fig. 6.4 is fundamentally limited by the fact that the odd nonlinear term generator is depicted as a pure cubing circuit. To see why, it suffices to consider that the nominally linear (main) receiver path is capable of generating nonlinear terms of order much greater than 3. What this means, for example, is that there will exist in the nominally linear (main) path at the higher-order intermodulation frequency offsets of $3f_1 - 2f_2$ and $3f_2 - 2f_1$ IM5, IM7, and higher-order IMD terms. At the frequency offsets of $4f_1 - 3f_2$ and $4f_2 - 3f_1$ there will similarly exist in the linear main path IM7, IM9, IM11, and higher order terms. However, a purely cubic odd nonlinear path generates reference intermodulation distortion products in response to a two-tone blocking scenario at the frequencies $2f_2 - f_1$ and $2f_1 - f_2$ only. In this case, the successive regeneration and cancellation of nonlinear terms can only proceed if $f_{LO} = 2f_1 - f_2$ or $f_{LO} = 2f_2 - f_1$. This situation is depicted in Fig. 6.5 in which odd-order intermodulation products exist at the LO frequency in the linear path but not in the nonlinear path.

To circumvent this problem, the solution proposed in Fig. 6.6 exploits the fact that the nonlinear reference in the nonlinear odd path does not need to be (and likely cannot be made to be) a pure cubing circuit or cubic term generator. In this case, both the linear path and the nonlinear reference each generate a single linear combination of polynomial terms that decreases in absolute magnitude as the order of the polynomial term increases. The problem of the successive regeneration and adaptive cancellation scheme then turns to expanding the nonlinear path odd-order polynomial

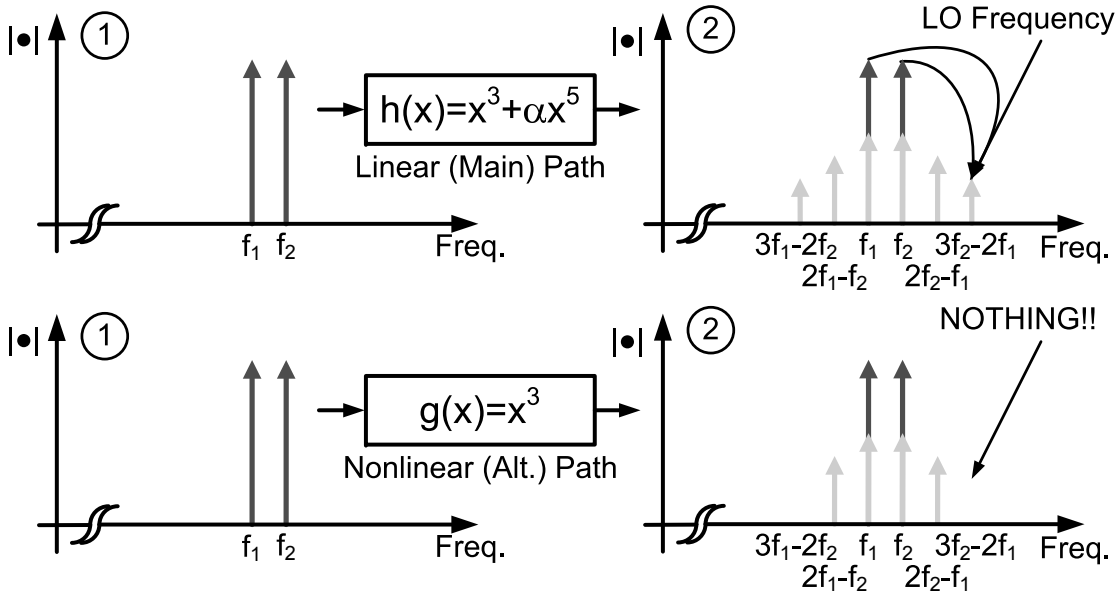


Figure 6.5: Depiction of insufficiency of pure cubic term generator for odd nonlinear reference in the presence of higher-order terms in the nominally linear path.

into a new basis set that can be used to more completely span the possible polynomials realizable by the nonlinearities in the nominally linear path. This new basis set is depicted in Fig. 6.6 as $\{p_{Odd,3}, p_{Odd,5}, p_{Odd,7}\}$.

Denoting the leading coefficients of the polynomial terms in the odd nonlinear term generator as χ_n with magnitudes $|\chi_3| > |\chi_5| > |\chi_7| > \ldots$, the nonlinear path will produce odd-order reference terms at higher-order intermodulation frequency offsets such as $3f_1 - 2f_2$ and $4f_1 - 3f_2$. These terms will be lower in magnitude than those at the IM3 offset of $2f_1 - f_2$ and $2f_2 - f_1$ but as long as the coefficient relationship of $|\chi_3| > |\chi_5| > |\chi_7| > \ldots$ is similar to that of the linear (main) path, the IMD-to-noise ratio (INR) of the nonlinear paths will track the INR of the nominally linear receiver path as the frequency offset is incremented, and thus the effectiveness of the cancellation will not be significantly compromised, if at all.

6.3.2.1 Quantitative Intuition of Higher-Order IMD Terms at IM3-Offset

With such a nonlinear reference generator, attention now turns to the case of $f_{LO} = 2f_1 - f_2$ and $2f_2 - f_1$. Since the nonlinear reference is of impure order (i.e. it is not a pure cubing circuit), it may seem that the achievable cancellation of nonlinear terms in the nominally linear path would be limited if under some conditions the nonlinearity in the main path were to return to a pure third-order term.

However, this is not the case. Although the initial nonlinear reference polynomial would have multiple terms, for example $p_{Odd,3} = \chi_3 x^3 + \chi_5 x^5 + \chi_7 x^7$, the successively regenerated nonlinear reference polynomials would be approximately proportional to $p_{Odd,5} \approx \chi_3 x^5 + \chi_5 x^7 + \chi_7 x^9$, $p_{Odd,7} \approx \chi_3 x^7 + \chi_5 x^9 + \chi_7 x^{11}$, and so on. Provided that the nonlinear reference generator has been designed such that $|\chi_3| > |\chi_5| > |\chi_7| > \ldots$, the n adaptive filters connecting the odd nonlinear reference

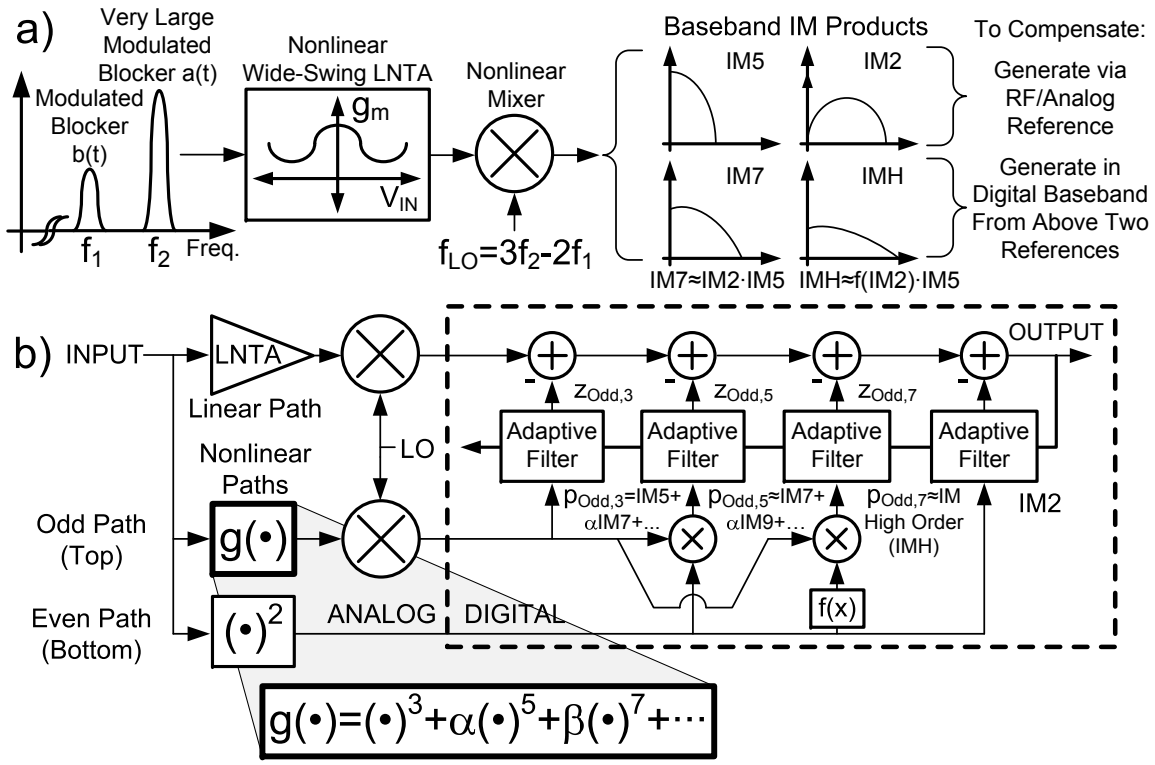


Figure 6.6: Successive regeneration and adaptive feedforward cancellation of intermodulation products at baseband implemented in this work showing explicitly the higher-order terms present in the odd-order nonlinear term generator. a) Concept. b) Simplified system block diagram.

branches to the linear receiver path would choose the dc weights β_n in (6.1) such that the odd nonlinear reference polynomials $z_{Odd,n}(x)$ generated by the adaptive equalization algorithms to cancel the IM3 terms in the main path are given by (6.2).

$$\begin{aligned} \beta_3 &: 1 \\ \beta_5 &: -\frac{\chi_5}{\chi_3} \\ \beta_7 &: \frac{\chi_5^2}{\chi_3^2} - \frac{\chi_7}{\chi_3} \end{aligned} \tag{6.1}$$

$$\begin{aligned} z_{Odd,3} &: \chi_3 x^3 + \chi_5 x^5 + \chi_7 x^7 \\ z_{Odd,5} &: -\chi_5 x^5 - \frac{\chi_5^2}{\chi_3} x^7 - \frac{\chi_5 \chi_7}{\chi_3} x^9 \\ z_{Odd,7} &: +\left(\frac{\chi_5^2}{\chi_3} - \chi_7\right) x^7 + \left(\frac{\chi_5^3}{\chi_3^2} - \frac{\chi_5 \chi_7}{\chi_3}\right) x^9 + \left(\frac{\chi_5^2 \chi_7}{\chi_3^2} - \frac{\chi_7^2}{\chi_3}\right) x^{11} \\ z_{Odd,Tot} &: \chi_3 x^3 + \left(\frac{\chi_5^3}{\chi_3^2} - \frac{\chi_5 \chi_7}{\chi_3}\right) x^9 + \left(\frac{\chi_5^2 \chi_7}{\chi_3^2} - \frac{\chi_7^2}{\chi_3}\right) x^{11} \end{aligned} \tag{6.2}$$

Now since $|\chi_5|$ and $|\chi_7|$ were small and less than $|\chi_3|$ to begin with, their successive products will be much less than $|\chi_3|$. Hence, the residual IM9 and IM11 terms from the operation described above will be much smaller in magnitude than the original IM5 and IM7 products. This procedure may be continued until the residual terms are below the noise floor or below the cancellation floor required by the receiver.

It is evident that for the case of linear path IMD products falling at an LO offset of $lf_2-(l-1)f_1$ or $lf_1-(l-1)f_2$, the leading polynomial term of the nonlinear path output is the $(2l-1)$th-order term representing the IM$(2l-1)$ distortion (e.g. IM3 for $l=2$). Replacing the above IM$(2l-1)$ terms above with IM$(2l+1)$ shows that the above argument also applies to for any possible intermodulation frequency offset. The principle of superposition then implies that for any linear path polynomial, only a finite number of higher-order distortion terms are required to be regenerated in order to guarantee that the uncancelled residual terms are negligible with respect to the noise floor.

6.3.3 Intuition of Approximation

One way to see intuitively why this general procedure results in approximations to higher-order IMD terms is via the depiction in Fig. 6.7. In part a), it suffices to consider all blockers as one composite signal $s(t)$. The blocking signal $s(t)$ may consist of two or more subsignals, but its defining characteristic is that the blocking profile is dominated by only one of them. For any single signal $s(t)$, $s(t)$ multiplied by successive powers of its envelope result in higher-order odd IMD products around the original frequency of $s(t)$.

However, due to the bandlimited nature of the analog to digital converters, only the portion of the envelope around dc can be digitized. Other IM2 terms that lie far away from dc due to the out-of-band nature of the blockers cannot be digitized. It can be seen in Fig. 6.7b that in the case of one dominant blocker, the dc portion of the envelope contains the bulk of the information in the envelope signal and for this reason the approximations to the higher-order terms improve as one blocker dominates the input spectrum.

Now to effectively regenerate higher-order terms it is required to obtain signal information around the LO frequency, of which there is none in the out-of-band blocking scenario of $s(t)$. However, in the spectrum of $s(t)|s(t)|^{2(l-1)}$, namely the IM$(2l-1)$ spectrum, there does exist signal content around the LO frequency. Since $s(t)|s(t)|^{2(l-1)}$ is a linear combination of all of the narrowband IM$(2l-1)$ products resulting from the original input blocking profile $s(t)$, multiplying the envelope approximation $|s'(t)|^2$ by each of these sub-terms results in an approximation to the IM$(2l+1)$ products that would appear around each of the narrowband carrier frequencies.

Finally, due to the invariance of performing the multiplication by an amplitude-modulated signal such as the envelope at RF or at baseband, this multiplication can be performed at baseband after downconverting odd-order terms to generate approximations to IM$(2l+1)$, IM$(2l+3)$, and higher-order terms.

6.3.4 Justification of Approximation

Although the proposed receiver only achieves large cancellation performance for large ratios of $|a(t)|/|b(t)|$, it is important to note that this condition is precisely the same as that of many important nonlinear blocking problems, as shown in Fig. 6.8. For example, in FDD communications

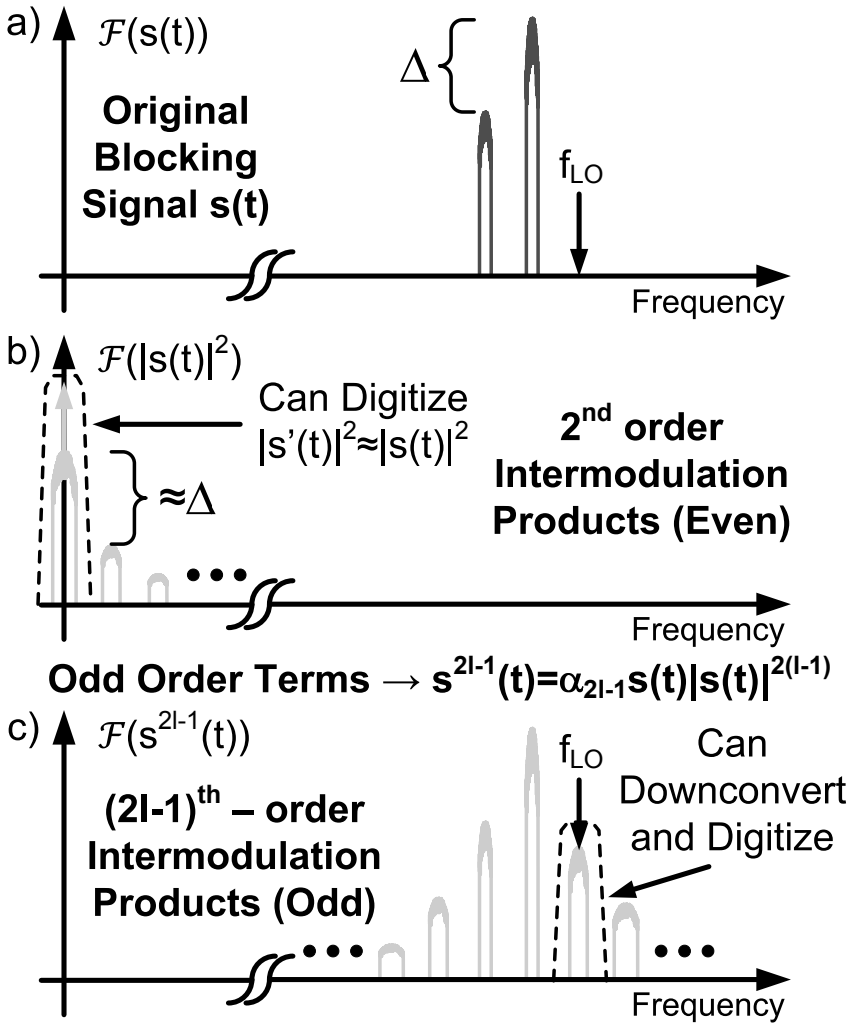


Figure 6.7: Intuitive approach as to why baseband successive regeneration is a good approximation for large blocker asymmetries. a) Considering all blockers as a single narrowband signal $s(t)$. b) Considering the envelope of $s(t)$ and the terms immediately surrounding dc. c) Considering higher-order odd IMD products at RF and those surrounding the LO frequency.

systems with relaxed PA/LNA isolation, the TX leakage appears as the dominant blocker to the receiver. Implantable medical sensors that receive power wirelessly may also need to demodulate a small data signal in the presences of a dominant power transfer signal [52]. Finally, radar systems and military communications systems need to be able to operate in the presence of very large dominant hostile jammer signal.

6.3.5 Greater-than-Two-Tone Blocking Scenarios

From the intuitive viewpoint of the approximation made when moderate-frequency baseband envelope signal content is discarded, the approximation is valid no matter how many tones are involved in IM3 generation, so long as only one of them dominates. Furthermore, if there are

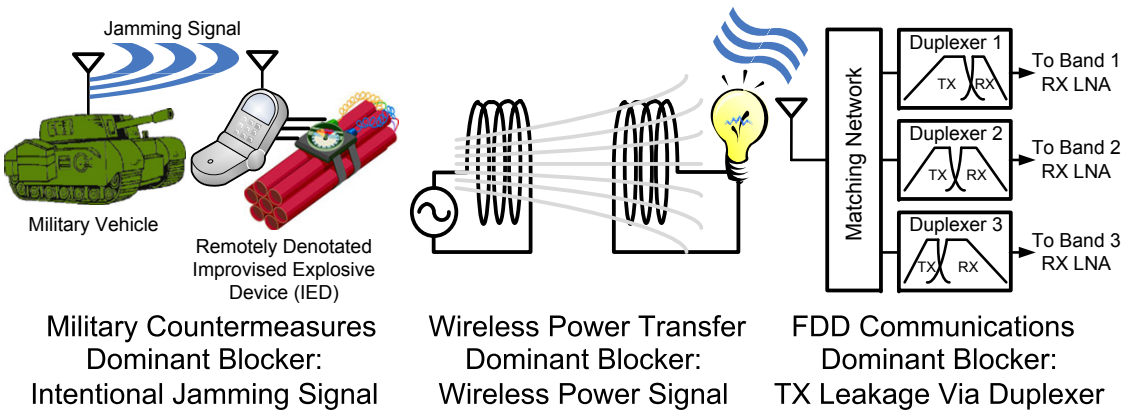


Figure 6.8: Application-based justification for a dominant-blocker approximation.

multiple tones present that are not involved in IM3 generation, they will appear as part of higher-order IMD products due to their contribution to the portion of the envelope at dc. In this case, the approximation still holds.

6.3.6 Use of Functional Basis Elements

In order to well-approximate nonlinearities in the main receiver path such as the residual of the LNTA, it is useful to use a compact functional representation such as the one developed in Section 5.8. In this case, the basis of Gaussian functions lends itself well to the block diagram of Fig. 6.4 in that it is clear that an exponential function need only be raised to a function of the IM2 signal and multiplied by the IM3 signals to recreate approximations to the terms generated in the actual LNTA.

However, this approach clearly works for general functions as well. As any function may be approximated by a polynomial, any odd-order functional basis element may be expressed as a polynomial $f(x) = ax^3 + bx^5 + cx^7 + \ldots$. Performing a polynomial expansion of a functional basis element $f(x^2)$ yields $f(x^2) = a + bx^2 + cx^4 + dx^6 + \ldots$. which can approximate any even-order IM distortion. Multiplying this by IM3 terms yields $x^3 f(x^2) = ax^3 + bx^5 + cx^7 + dx^9 + \ldots$ which can yield any odd-order distortion.

6.4 Analysis of Residual Error from Successive Regeneration Approximation

6.4.1 General Approach

In order to quantify the error in the proposed IMD successive regeneration scheme, two statistically independent blocker signals are represented using the form of (4.1-18) in [120]. That is, the two blocker signals $s_1(t)$ and $s_2(t)$ are modeled as in (6.3) where $a(t)$ and $b(t)$ are the time-varying signal envelopes and $\Theta_2(t)$ and $\Theta_1(t)$ the time-varying phases.

$$\begin{aligned} s_1(t) &= a(t)\cos\left(2\pi f_2 t + \Theta_2(t)\right) \\ s_2(t) &= b(t)\cos\left(2\pi f_1 t + \Theta_1(t)\right) \\ x(t) &= s_1(t) + s_2(t) \end{aligned} \tag{6.3}$$

In order to compact the following analysis the time arguments are removed and the substitutions shown in (6.4) are made.

$$\begin{aligned} a(t) &\rightarrow a \\ b(t) &\rightarrow b \\ a^2 + b^2 &\rightarrow IM2 \\ 2\pi f_2 t + \Theta_2(t) &\rightarrow \Theta \\ 2\pi f_1 t + \Theta_1(t) &\rightarrow \Psi \end{aligned} \tag{6.4}$$

With these representations, the input to the receiver is given by (6.5).

$$x = a\cos(\Theta) + b\cos(\Psi) \tag{6.5}$$

6.4.2 General Case for Odd-Order Intermodulation Distortion

Assuming that the receiver odd-order nonlinearity may be expressed as a memoryless polynomial, the receiver odd-order output y may be expressed as in (6.6).

$$y = \sum_{\substack{n=1 \\ n\in\mathbb{Z}_{Odd}}}^{\infty} \alpha_n x^n = \sum_{\substack{n=1 \\ n\in\mathbb{Z}_{Odd}}}^{\infty} \alpha_n (a\cos(\Theta) + b\cos(\Psi))^n \tag{6.6}$$

Expanding this relation using the binomial theorem yields (6.7).

$$y = \sum_{\substack{n=1 \\ n\in\mathbb{Z}_{Odd}}}^{\infty} \sum_{k=0}^{n} \alpha_n \binom{n}{k} a^{n-k} b^k \cos^{n-k}(\Theta)\cos^k(\Psi) \tag{6.7}$$

At this point, it is desired to isolate the IMD products falling at the frequency $f_{LO} = (l)f_2 - (l-1)f_1$ where $2l - 1 \leq n$. This generality covers the case in which IMD products fall at frequency offsets allowed by higher-order nonlinear terms. The first step in doing so is to utilize the power reduction formula for sinusoids, given below in (6.8).

$$\cos^p(\Theta) = \begin{cases} \frac{2}{2^p} \sum_{q=0}^{\frac{p-1}{2}} \binom{p}{q} \cos\left((p-2q)\Theta\right) & p \in \mathbb{Z}_{Odd} \\ \frac{1}{2^p}\binom{p}{p/2} + \frac{2}{2^p} \sum_{q=0}^{\frac{p-1}{2}} \binom{p}{q} \cos\left((p-2q)\Theta\right) & p \in \mathbb{Z}_{Even} \end{cases} \tag{6.8}$$

The second step in isolating the IMD products falling at f_{LO} is to substitute the terms in (6.8) in which $p - 2q = l$ and $p - 2q = l - 1$ into the terms $\cos^{n-k}(\Theta)$ and $\cos^k(\Psi)$, respectively. Proper substitution in this regard entails making the change of variables shown in (6.9), (6.10), and (6.11).

$$\begin{aligned} p-2q=l &\rightarrow \tfrac{p-l}{2}=q \\ p-2q=l-1 &\rightarrow \tfrac{p-l+1}{2}=q \end{aligned} \tag{6.9}$$

$$\begin{aligned} \Theta &: \frac{2}{2^p}\binom{p}{\frac{p-l}{2}}\cos\left((l)\Theta\right) \\ \Psi &: \frac{2}{2^p}\binom{p}{\frac{p-l+1}{2}}\cos\left((l-1)\Psi\right) \end{aligned} \tag{6.10}$$

$$\begin{aligned} \Theta : p=n-k &\rightarrow \frac{2}{2^{n-k}}\binom{n-k}{\frac{n-k-l}{2}}\cos\left((l)\Theta\right) \\ \Psi : p=k &\rightarrow \frac{2}{2^k}\binom{k}{\frac{k-l+1}{2}}\cos\left((l-1)\Psi\right) \end{aligned} \tag{6.11}$$

The final substitution is shown in (6.12). In this case, only intermodulation distortion powers $n \geq 2l-1$ and $l-1 \leq k \leq n-l, k-l \in \mathbb{Z}_{Odd}$ are required for consideration.

$$y_{Intermediate} = \sum_{\substack{n=2l-1 \\ n\in\mathbb{Z}_{Odd}}}^{\infty} \sum_{\substack{k=l-1 \\ k-l\in\mathbb{Z}_{Odd}}}^{n-l} \alpha_n \binom{n}{k} a^{n-k} b^k \tfrac{1}{2^{n-2}} \binom{n-k}{\frac{n-k-l}{2}} \binom{k}{\frac{k-l+1}{2}} \cos\left((l)\Theta\right)\cos\left((l-1)\Psi\right) \tag{6.12}$$

Multiplying the cosine terms together and consolidating terms yields the expression in (6.13)-(6.14). Here, the phase-dependent cosine expression is consolidated into the term c_o.

$$y_{Odd} = \sum_{\substack{n=2l-1 \\ n\in\mathbb{Z}_{Odd}}}^{\infty} \sum_{\substack{k=l-1 \\ k-l\in\mathbb{Z}_{Odd}}}^{n-l} \alpha_n \tfrac{1}{2^{n-1}} a^{n-k} b^k \binom{n}{k} \binom{n-k}{\frac{n-k-l}{2}} \binom{k}{\frac{k-l+1}{2}} \cos\left((l)\Theta-(l-1)\Psi\right) \tag{6.13}$$

$$y_{Odd} = \sum_{\substack{n=2l-1 \\ n\in\mathbb{Z}_{Odd}}}^{\infty} \sum_{\substack{k=l-1 \\ k-l\in\mathbb{Z}_{Odd}}}^{n-l} \alpha_n \gamma_{n,k,l} a^{n-k} b^k c_o \tag{6.14}$$

Expanding the n-choose-r terms of $\gamma_{n,k,l}$ yields (6.15).

$$\begin{aligned} \gamma_{n,k,l} &= \tfrac{1}{2^{n-1}} \frac{n!}{(n-k)!k!} \cdot \frac{(n-k)!}{(\frac{n-k+l}{2})!(\frac{n-k-l}{2})!} \cdot \frac{k!}{(\frac{k+l-1}{2})!(\frac{k-l+1}{2})!} \\ \gamma_{n,k,l} &= \tfrac{n!}{2^{n-1}} \frac{1}{(\frac{n-k+l}{2})!(\frac{n-k-l}{2})!(\frac{k+l-1}{2})!(\frac{k-l+1}{2})!} \end{aligned} \tag{6.15}$$

The above terms in (6.14) represent the true odd-order IMD products generated by the receiver linear path nonidealities. The reference odd-order IMD products in principle have the same phase as those in the linear path after adaptive equalization and are generated from the two noisy nonlinear

reference path signals r_{Even} and r_{Odd}, defined in (6.16). More precisely, the two noisy nonlinear reference path signals are proportional to those in (6.16). This does not matter for the purposes of this analysis, however, as the proportionality constant implicity referred to is compensated for in the adaptive equalizers between the nonlinear and linear paths.

To begin the treatment of the odd nonlinear path reference, the general case of an impure-polynomial odd nonlinear term generator with coefficients χ_m is assumed. In addition to terminology previously defined, n_{Ev} and n_{Od} represent the noise signals on the raw even and odd nonlinear reference paths, respectively. The time-varying phase of the odd nonlinear path noise is denoted by a compacted cosine term c_d.

$$\begin{aligned} r_{Even} &= a^2 + b^2 + n_{Ev} \\ r_{Odd} &= \sum_{\substack{m=2l-1 \\ m\in\mathbb{Z}_{Odd}}}^{\infty} \sum_{\substack{i=l-1 \\ i-l\in\mathbb{Z}_{Odd}}}^{m-l} (\chi_m \gamma_{m,i,l} a^{m-i} b^i c_o) + n_{Od} c_d \end{aligned} \tag{6.16}$$

In order to create the composite odd-order reference z_{Odd}, the raw even-order reference r_{Even} is first successively multiplied to generate higher orders of even-order intermodulation products. These higher even-order terms are then multiplied by the raw odd-order reference r_{Odd} and then weighted by the adaptive algorithm to produce z_{Odd}. Formally, the first of these steps are performed in (6.17)-(6.18).

$$z_{Odd} = r_{Odd} \sum_{\substack{n=2l-1 \\ n\in\mathbb{Z}_{Odd}}}^{\infty} \beta_n r_{Even}^{\left(\frac{n+1}{2}-l\right)} \tag{6.17}$$

$$\begin{aligned} z_{Odd} = &\sum_{\substack{n=2l-1 \\ n\in\mathbb{Z}_{Odd}}}^{\infty} \sum_{\substack{m=2l-1 \\ m\in\mathbb{Z}_{Odd}}}^{\infty} \sum_{\substack{i=l-1 \\ i-l\in\mathbb{Z}_{Odd}}}^{m-l} \beta_n (\chi_m \gamma_{m,i,l} a^{m-i} b^i c_o)(a^2 + b^2 + n_{Ev})^{\left(\frac{n+1}{2}-l\right)} + \\ &\sum_{\substack{n=2l-1 \\ n\in\mathbb{Z}_{Odd}}}^{\infty} n_{Od} c_d (a^2 + b^2 + n_{Ev})^{\left(\frac{n+1}{2}-l\right)} \end{aligned} \tag{6.18}$$

Expanding the r_{Even} term using the trinomial theorem yields (6.19).

$$\begin{aligned} (a^2 + b^2 + n_{Ev})^{\left(\frac{n+1}{2}-l\right)} &= \sum_{\substack{k=l-1 \\ k-l\in\mathbb{Z}_{Odd}}}^{n-l} \sum_{j=0}^{\frac{k-l+1}{2}} \frac{\left(\frac{n+1}{2}-l\right)!}{\left(\frac{n-k-l}{2}\right)!\left(\frac{k-l+1}{2}-j\right)!j!} a^{n-k-l} b^{k-l+1-2j} n_{Ev}^j \\ (a^2 + b^2 + n_{Ev})^{\left(\frac{n+1}{2}-l\right)} &= \sum_{\substack{k=l-1 \\ k-l\in\mathbb{Z}_{Odd}}}^{n-l} \sum_{j=0}^{\frac{k-l+1}{2}} \lambda_{n,k,j,l} a^{n-k-l} b^{k-l+1-2j} n_{Ev}^j \end{aligned} \tag{6.19}$$

Incorporating this expansion into (6.18) yields (6.20).

$$z_{Odd} = \sum_{n=2l-1} \sum_{k=l-1} \sum_{j=0} \sum_{m=2l-1} \sum_{i=l-1} \beta_n \lambda_{n,k,j,l} \chi_m \gamma_{m,i,l} a^{n+m-k-i-l} b^{k+i-l+1-2j} n_{Ev}^j + \left(\sum_{n=2l-1} \sum_{k=l-1} \sum_{j=0} \beta_n \lambda_{n,k,j,l} a^{n-k-l} b^{k-l+1-2j} n_{Ev}^j \right) n_{Od} c_d \tag{6.20}$$

Continuing to assume a memoryless effective nonlinearity for both the linear and nonlinear receiver paths, and assuming that the LMS adaptive equalizer adjusts the values β_n so as to minimize the mean squared error, we find that the final LMS tap weight vector $\vec{\beta}_{opt}$, whose ∞ elements are indexed by n, is provided by (6.21), which follows from the mean squared error orthogonality principle.

$$\begin{aligned} \vec{\beta}_{opt} &= \bar{\bar{R}}^{-1} E[\vec{v} y_{Odd}] \\ \bar{\bar{R}} &= E[\vec{v}\vec{v}^T] \end{aligned} \tag{6.21}$$

In (6.21), the vector $\vec{v}$ consists of ∞ elements and is indexed by the variable n. Each nth element of $\vec{v}$ is given by (6.22).

$$v_n = \sum_{k=l-1} \sum_{j=0} \sum_{m=2l-1} \sum_{i=l-1} \lambda_{n,k,j,l} \chi_m \gamma_{m,i,l} a^{n+m-k-i-l} b^{k+i-l+1-2j} n_{Ev}^j c_o + \left(\sum_{k=l-1} \sum_{j=0} \lambda_{n,k,j,l} a^{n-k-l} b^{k-l+1-2j} n_{Ev}^j \right) n_{Od} c_d \tag{6.22}$$

Performing the computations of (6.21) requires knowledge of the characteristics of the nondominant blocker and the phases of both the incoming IMD products and the odd nonlinear path noise. This information may not be known *a priori*. In the event that only the characteristics of the dominant blocker are known *a priori*, the coefficients β_n may be determined by finding the MMSE solution of the simplified relation (6.23).

$$\sum_{n=2l-1} \alpha_n \gamma_{n,l-1,l} a^{n-l+1} = \sum_{n=2l-1} \sum_{m=2l-1} \beta_n \lambda_{n,l-1,0,l} \chi_m \gamma_{m,l-1,l} a^{n+m-3l+2} \tag{6.23}$$

Given that a closed-form solution is not possible for β_n in either case, the analysis is continued without substituting into this symbolic placeholder. However, after the more complete analysis is concluded, the special case in which $r_{Odd} = a^l b^{l-1} c_o + n_{Od} c_d$ will be considered which does have a closed-form solution to β_n. Continuing with the general analysis yields (6.24) with the three components of the error residual given by $e_{Odd,1}$, $e_{Odd,2}$, and $e_{Odd,3}$.

$$
\begin{aligned}
z_{Odd} = & \sum_{n=2l-1} \sum_{k=l-1} \sum_{m=2l-1} \sum_{i=l-1} \beta_n \lambda_{n,k,l,0} \chi_m \gamma_{m,i,l} a^{n+m-i-k-l} b^{k+i-l+1} c_o + \\
& + \sum_{n=2l+1} \sum_{k=l+1} \sum_{j=1} \sum_{m} \sum_{i} \beta_n \lambda_{n,k,l,j} \chi_m \gamma_{m,i,l} a^{n+m-i-k-l} b^{k+i-l+1-2j} n_{Ev}^j c_o + \\
& + \sum_{n=2l-1} \sum_{k=l-1} \sum_{j=0} \beta_n \lambda_{n,k,l,j} a^{n-k-l} b^{k-l+1-2j} n_{Ev}^j n_{Od} c_d \\
= & \, y_{Odd} + e_{Odd,1} + e_{Odd,2} + e_{Odd,3}
\end{aligned}
\tag{6.24}
$$

The first, or dominant, error term represents the residual left over from cancelling signal-only terms and is shown explicitly in (6.25). By the operations performed in (6.23) and in the MMSE computation of the β_n, this term is equal to 0 for the dominant signal term, i.e. that for which $k, i = l-1$, provided that $\chi_m = 0, m > 2l-1$. The next most dominant residual term occurs when $k = l+1, i = l-1$ and will be referred to hereafter as the nondominant error term. Neglecting other terms, a more compact approximation to $e_{Odd,1}$ may be developed in (6.26). Here, the maximum of the dominant and nondominant error is taken, a step that is useful for accurately predicting error when $\chi_m \neq 0, m > 2l-1$.

$$
e_{Odd,1} = \sum_{n,k} \alpha_n \gamma_{n,k,l} a^{n-k} b^k c_o - \sum_{n,k,m,i} \beta_n \lambda_{n,k,0,l} \chi_m \gamma_{m,i,l} a^{n+m-i-k-l} b^{k+i-l+1} c_o \tag{6.25}
$$

$$
\begin{aligned}
\delta_{n,l} &= \alpha_n \gamma_{n,l-1,l} \\
\epsilon_{n,m,l} &= \beta_n \chi_m \lambda_{n,l-1,0,l} \gamma_{m,l-1,l} \\
\delta'_{n,l} &= \alpha_n \gamma_{n,l+1,l} \\
\epsilon'_{n,m,l} &= \beta_n \chi_m \lambda_{n,l+1,0,l} \gamma_{m,l-1,l} \\
e_{Odd,1} &\approx \max \left(\sum_{n=2l-1} (\delta_{n,l} a^{n-l+1} - \sum_{m=2l-1} \epsilon_{n,m,l} a^{n+m-3l+2}) b^{l-1} c_o, \right. \\
&\qquad \left. \sum_{n=2l+1} (\delta'_{n,l} a^{n-l-1} - \sum_{m=2l-1} \epsilon'_{n,m,l} a^{n+m-3l}) b^{l+1} c_o \right)
\end{aligned}
\tag{6.26}
$$

One important thing to note at this point is that not all of the signal energy in the nondominant term of (6.26) contributes to error. In fact, as terms containing b^{l-1} are readily cancelled by minor adjustments to the dominant odd-order reference terms, a b^2 term may be factored out of b^{l+1} and the expected value removed to leave the terms that truly contribute to uncorrectable error as in (6.27). For cases in which l is odd, additional terms may be factored out, for example if $l = 3$ then b^{l-1} may be replaced by $b^2 - E[b^2]$ in (6.27) below.

$$e_{Odd,1} \approx \max\left(\sum_{n=2l-1}(\delta_{n,l}a^{n-l+1} - \sum_{m=2l-1}\epsilon_{n,m,l}a^{n+m-3l+2})b^{l-1}c_o, \right. \\ \left. \sum_{n=2l+1}(\delta'_{n,l}a^{n-l-1} - \sum_{m=2l-1}\epsilon'_{n,m,l}a^{n+m-3l})b^{l-1}(b^2 - E[b^2])c_o\right) \tag{6.27}$$

From (6.27), the error power due to this quantity may be derived as in (6.28), assuming that the phase of the IM3 products is independent of the amplitude of either blocker signal.

$$E[e^2_{Odd,1}] \approx \max\left(\left(\sum_{n=2l-1}\sum_{p=2l-1}\left(\delta_{n,l}\delta_{p,l}E[a^{n+p-2l+2}] - 2\sum_{m=2l-1}\delta_{p,l}\epsilon_{n,m,l}E[a^{n+m+p-4l+3}] \right.\right.\right. \\ \left.\left.\left. + \sum_{m=2l-1}\sum_{q=2l-1}\epsilon_{n,m,l}\epsilon_{p,q,l}E[a^{n+m+p+q-6l+4}]\right)\right)\cdot E[b^{2l-2}]E[c_o^2], \\ \left(\sum_{n=2l+1}\sum_{p=2l+1}\left(\delta'_{n,l}\delta'_{p,l}E[a^{n+p-2l-2}] - 2\sum_{m=2l-1}\delta'_{p,l}\epsilon'_{n,m,l}E[a^{n+m+p-4l-1}] \right.\right. \\ \left.\left. + \sum_{m=2l-1}\sum_{q=2l-1}\epsilon'_{n,m,l}\epsilon'_{p,q,l}E[a^{n+m+p+q-6l}]\right)\right)\cdot\left(E[b^{2l+2}] - 2E[b^{2l}]E[b^2] + E[b^{2l-2}]E[b^2]^2\right)E[c_o^2]\Bigg) \tag{6.28}$$

Likewise with the quantity $e_{Odd,2}$, a simplified approximation, assuming that $|a| \gg |b|$, can be made by only considering the dominant term where $k = l+1, i = l-1$, as in (6.29).

$$E[e^2_{Odd,2}] \approx \\ \sum_{n,p=2l+1,m,q=2l-1}\beta_n\beta_p\chi_m\chi_q\lambda_{n,l+1,1,l}\lambda_{p,l+1,1,l}\gamma_{m,l-1,l}\gamma_{q,l-1,l}E[a^{n+m+p+q-6l}]E[b^{2l-2}]E[n^2_{Ev}]E[c_o^2] \tag{6.29}$$

Similarly, for $e_{Odd,3}$, the dominant term occurs for $k = l-1$, $j = 0$, resulting in an error power given by (6.30) In this case $E[c_d^2] = 1/2$ due to the fact that the noise is a bandpass Gaussian signal.

$$E[e^2_{Odd,3}] \approx \frac{1}{2}\sum_{n,p=2l-1}\beta_n\beta_p\lambda_{n,l-1,0,l}\lambda_{p,l-1,0,l}E[a^{n+p-4l+2}]E[n^2_{Od}] \tag{6.30}$$

Finally, the power of the total odd-order intermodulation distortion is given by (6.31). Also in this case, the approximation is made to consider the dominant terms (i.e. in which $k = l-1$).

$$E[y^2_{Odd}] \approx \sum_{n,m=2l-1}\alpha_n\alpha_m\gamma_{n,l-1,l}\gamma_{m,l-1,l}E[a^{n+m-2l+2}]E[b^{2l-2}]E[c_o^2] \tag{6.31}$$

The definition of signal to noise ratio (power) in a traditional context is given by (6.32).

$$SNR = \frac{E[s^2(t)]}{E[n^2(t)]} \tag{6.32}$$

Applying this definition to the expressions in (6.28), (6.29), and (6.30) results in an approximation to the IER resulting from the proposed successive IMD regeneration procedure for odd-order IMD products, shown in (6.33).

$$IER_{Odd} \approx \frac{E[y_{Odd}^2]}{E[e_{Odd,1}^2] + E[e_{Odd,2}^2] + E[e_{Odd,3}^2]} \tag{6.33}$$

6.4.3 Single Polynomial Term Nonlinear Term Generator Case for Odd-Order Intermodulation Distortion

As mentioned previously, a somewhat more explicit representation of the error terms may be obtained if it is assumed that the odd nonlinear analog reference circuit produces only an output proportional to $a^l b^{l-1} c_o$. This result may be obtained beginning with setting $m = 2l-1$ and $i = l-1$ in (6.16). In this case, the reference quantities are as given below in (6.34).

$$\begin{aligned}
r_{Odd} &= \chi_{2l-1}\gamma_{2l-1,l-1,l}a^l b^{l-1}c_o + n_{Od}c_d \\
z_{Odd} &= r_{Odd}\sum_{n=2l-1}\beta_n(a^2+b^2+n_{Od})^{\left(\frac{n+1}{2}-l\right)} \\
z_{Odd} &= \sum_{n=2l-1,k=l-1,j=0}\beta_n\chi_{2l-1}\gamma_{2l-1,l-1,l}\lambda_{n,k,j,l}a^{n-k}b^{k-2j}n_{Ev}^j c_o + \\
&\quad \sum_{n=2l-1,k=l-1,j=0}\beta_n\lambda_{n,k,j,l}a^{n-k-l}b^{k-l+1-2j}n_{Ev}^j n_{Od}c_d
\end{aligned} \tag{6.34}$$

Solving for a nearly optimal set of β_n can be done explicitly in this case by equating the dominant signal terms (i.e. by setting $k = l-1$, $j = 0$) in z_{Odd} to those of y_{Odd}. Cancelling common terms results in the relation (6.35).

$$\begin{aligned}
\alpha_n\gamma_{n,l-1,l} &= \beta_n\chi_{2l-1}\gamma_{2l-1,l-1,l}\lambda_{n,l-1,0,l} \\
\alpha_n\frac{n!}{2^{n-1}}\frac{1}{(\frac{n+1}{2})!(\frac{n-2l+1}{2})!(l-1)!0!} &= \beta_n\chi_{2l-1}\frac{1}{l!(l-1)!}\frac{(\frac{n-2l+1}{2})!}{(\frac{n-2l+1}{2})!0!}\frac{(2l-1)!}{2^{2l-2}} \\
\alpha_n\frac{n!}{2^{n-2l+1}}\frac{1}{\chi_{2l-1}}\frac{l!}{(2l-1)!}\frac{1}{(\frac{n+1}{2})!(\frac{n-2l+1}{2})!} &= \beta_n
\end{aligned} \tag{6.35}$$

Substituting this result into the expressions for e_{Odd} in the previous section in addition to setting $\chi_n = 0, n > 0$ yields (6.36), (6.37), and (6.38) below. Note that since in this case the odd nonlinear term generator is of pure order, the dominant IMD products (i.e. $k = l-1$) are completely cancelled.

$$
\begin{aligned}
e_{Odd,1} &\approx \sum_{n=2l+1} (\alpha_n \gamma_{n,l+1,l} - \beta_n \chi_{2l-1} \lambda_{n,l+1,0,l} \gamma_{2l-1,l-1,l}) \cdot \\
&\quad a^{n-l-1} b^{l-1} (b^2 - E[b^2]) c_o \\
e_{Odd,1} &\approx \sum_{n=2l+1} \left(\alpha_n \frac{n!}{2^{n-1}} \frac{1}{(\frac{n-2l-1}{2})!(\frac{n-1}{2})! l! 1!} - \right. \\
&\quad \left. \alpha_n \frac{n!}{2^{n-2l+1}} \frac{l!}{(2l-1)!} \frac{1}{(\frac{n+1}{2})!(\frac{n-2l+1}{2})!} \left(\frac{(\frac{n+1}{2}-l)!}{(\frac{n-2l-1}{2})! 0!} \frac{1}{l!(l-1)!} \frac{(2l-1)!}{2^{2l-2}} \right) \right) \cdot \\
&\quad a^{n-l-1} b^{l-1} (b^2 - E[b^2]) c_o \\
e_{Odd,1} &\approx \sum_{n=2l+1} \alpha_n \frac{n!}{2^{n-1}} \left(\frac{\frac{n+1}{2} - l}{l!(\frac{n+1}{2})!(\frac{n-2l-1}{2})!} \right) a^{n-l-1} b^{l-1} (b^2 - E[b^2]) c_o \\
e_{Odd,1} &\approx \sum_{n=2l+1} \alpha_n \gamma_{n,l+1,l} (1 - \tfrac{2l}{n+1}) a^{n-l-1} b^{l-1} (b^2 - E[b^2]) c_o \\
E[e_{Odd,1}^2] &\approx \sum_{n=2l+1} \sum_{m=2l+1} \alpha_n \alpha_m \gamma_{n,l+1,l} \gamma_{m,l+1,l} (1 - \tfrac{2l}{n+1})(1 - \tfrac{2l}{m+1}) E[a^{n+m-2l-2}] \cdot \\
&\quad \left(E[b^{2l+2}] - 2E[b^{2l}]E[b^2] + E[b^{2l-2}]E[b^2]^2 \right) E[c_o^2]
\end{aligned}
\tag{6.36}
$$

$$
\begin{aligned}
e_{Odd,2} &\approx \sum_{n=2l+1} \beta_n \chi_{2l-1} \lambda_{n,l+1,1,l} \gamma_{2l-1,l-1,l} a^{n-l+1} b^{l-1} n_{Ev} c_o \\
e_{Odd,2} &\approx \sum_{n=2l+1} \alpha_n \frac{n!}{2^{n-2l+1}} \frac{1}{\chi_{2l-1}} \frac{l!}{(2l-1)!} \frac{1}{(\frac{n+1}{2})!(\frac{n-2l+1}{2})!} \chi_{2l-1} \frac{1}{l!(l-1)!} \frac{(2l-1)!}{2^{2l-2}} \frac{(\frac{n+1}{2}-l)!}{(\frac{n-2l-1}{2})! 0!} \cdot \\
&\quad a^{n-l+1} b^{l-1} n_{Ev} c_o \\
e_{Odd,2} &\approx \sum_{n=2l+1} \alpha_n \frac{n!}{2^{n-1}} \frac{1}{(\frac{n+1}{2})!(\frac{n-2l-1}{2})!(l-1)!} a^{n-l+1} b^{l-1} n_{Ev} c_o \\
e_{Odd,2} &\approx \sum_{n=2l+1} \alpha_n \gamma_{n,l+1,l} \tfrac{2l}{n+1} a^{n-l+1} b^{l-1} n_{Ev} c_o \\
E[e_{Odd,2}^2] &\approx \sum_{n=2l+1} \sum_{m=2l+1} \alpha_n \alpha_m \gamma_{n,l+1,l} \gamma_{m,l+1,l} \tfrac{4l^2}{(n+1)(m+1)} E[a^{n+m-2l+2}] E[b^{2l-2}] E[n_{Ev}^2] E[c_o^2]
\end{aligned}
\tag{6.37}
$$

$$
\begin{aligned}
e_{Odd,3} &\approx \sum_{n=2l-1} \beta_n \lambda_{n,l-1,0,l} a^{n-2l+1} n_{Od} c_d \\
e_{Odd,3} &\approx \sum_{n=2l-1} \alpha_n \frac{n!}{2^{n-2l+1}} \frac{1}{\chi_{2l-1}} \frac{l!}{(2l-1)!} \frac{1}{(\frac{n+1}{2})!(\frac{n-2l+1}{2})!} a^{n-2l+1} n_{Od} c_d \\
e_{Odd,3} &\approx \sum_{n=2l-1} \alpha_n \xi_{n,l} a^{n-2l+1} n_{Od} c_d \\
E[e_{Odd,3}^2] &\approx \frac{1}{2} \sum_{n=2l-1} \sum_{m=2l-1} \alpha_n \alpha_m \xi_{n,l} \xi_{m,l} E[a^{n+m-4l+2}] E[n_{Od}^2]
\end{aligned}
\tag{6.38}
$$

Where $\xi_{n,l}$ is defined in place in (6.38).

6.4.4 Even-Order Intermodulation Distortion

A similar set of procedures can be instituted to analyze the error due to the approximation used in the even-order reference term generation. It is important to consider only the non-dc portion of the even-order intermodulation products due to the fact that the dc portions will be removed via dc offset-cancellation circuitry and high-pass filters. Here, it is assumed that the even nonlinear path reference is a very good approximation to a squaring circuit (which it is in practice).

In this case, the even-order signal (including dc terms) produced by the nominally linear path is again given by the binomial theorem in (6.39).

$$y = \sum_{\substack{n=2 \\ n\in\mathbb{Z}_{Even}}}^{\infty} \sum_{\substack{k=0 \\ n\in\mathbb{Z}_{Even}}}^{n} \alpha_n \binom{n}{k} a^{n-k} b^k \cos^{n-k}(\Theta)\cos^{k}(\Psi) \tag{6.39}$$

Noting that only even-order IMD products near dc are of interest and that k and $n-k$ are always even integers, the substitution dictated by the power reduction formula in (6.40) is performed, eventually resulting in (6.41) after simplification.

$$\cos^{p}(\Theta) \rightarrow \frac{1}{2^p}\binom{p}{p/2} \tag{6.40}$$

$$y_{Even} = \sum_{\substack{n=2 \\ n\in\mathbb{Z}_{Even}}}^{\infty} \sum_{\substack{k=0 \\ n\in\mathbb{Z}_{Even}}}^{n} \alpha_n \frac{n!}{2^n}\left[\frac{1}{(\frac{n-k}{2})!\,(\frac{k}{2})!}\right]^2 a^{n-k}b^k \tag{6.41}$$

Removing dc terms from (6.41) yields (6.42)

$$y_{Even} = \sum_{\substack{n=2 \\ n\in\mathbb{Z}_{Even}}}^{\infty} \sum_{\substack{k=0 \\ n\in\mathbb{Z}_{Even}}}^{n} \alpha_n \frac{n!}{2^n}\left[\frac{1}{(\frac{n-k}{2})!\,(\frac{k}{2})!}\right]^2 (a^{n-k}b^k - E[a^{n-k}]E[b^k]) \tag{6.42}$$

Retaining only the dominant terms of y_{Even} in which $k = 0$ and computing the power yields (6.43)

$$E[y_{Even}^2] = \sum_{n,m=2} \alpha_n\alpha_m \frac{n!m!}{2^{n+m}}\left[\frac{1}{(\frac{n}{2})!}\right]^2\left[\frac{1}{(\frac{m}{2})!}\right]^2 (E[a^{n+m}] - E[a^n]E[a^m]) \tag{6.43}$$

Noting again that the raw even nonlinear reference can be represented by $r_{Even} = a^2 + b^2 + n_{Ev}$, the higher even-order intermodulation terms generated by successive multiplication of r_{Even} by itself are given in (6.44) from the trinomial theorem. In order to match the zero-dc characteristic of y_{Even}, the dc portions of each of the terms comprising the composite reference signal z_{Even} are removed.

$$z_{Even} = \sum_{\substack{n=2 \\ n\in\mathbb{Z}_{Even}}}^{\infty} \beta_n((a^2 + b^2 + n_{Ev})^{n/2} - E[(a^2 + b^2 + n_{Ev})^{n/2}]) \tag{6.44}$$

$$z_{Even} = \sum_{\substack{n=2 \\ n\in\mathbb{Z}_{Even}}}^{\infty} \sum_{\substack{k=0 \\ k\in\mathbb{Z}_{Even}}}^{n} \sum_{j=0}^{k/2} \beta_n \frac{(\frac{n}{2})!}{(\frac{n-k}{2})!(\frac{k}{2}-j)!j!}(a^{n-k}b^{k-2j}n_{Ev}^j - E[a^{n-k}]E[b^{k-2j}]E[n_{Ev}^j]) \tag{6.45}$$

Again, a compact symbolic expression for the MMSE solution for the β_n is not possible and the best-fit solution here is approximated to be that in which y_{Even} and z_{Even} are equal for $k = 0$ and $j = 0$. In this case, β_n is given by (6.46).

$$\beta_n = \alpha_n \frac{n!}{2^n} \left(\frac{1}{(\frac{n}{2})!} \right)^2 \tag{6.46}$$

Substituting this result into (6.45) yields (6.47).

$$z_{Even} = \sum_{n,k,j} \alpha_n \frac{n!}{2^n} \frac{1}{(\frac{n}{2})!(\frac{n-k}{2})!(\frac{k}{2}-j)!j!} (a^{n-k} b^{k-2j} n_{Ev}^j - E[a^{n-k}]E[b^{k-2j}]E[n_{Ev}^j]) \tag{6.47}$$

Subtracting z_{Even} from y_{Even} yields two error components, the first of which is related to uncancelled even-order IMD products, while the second is related to the noise present in the even nonlinear reference.

$$e_{Even,1} = \sum_{n,k=2} \alpha_n \frac{n!}{2^n} \left[\frac{1}{(\frac{n-k}{2})!(\frac{k}{2})!} \right]^2 \left(1 - \frac{(\frac{n-k}{2})!(\frac{k}{2})!}{(\frac{n}{2})!} \right) (a^{n-k} b^k - E[a^{n-k}]E[b^k]) \tag{6.48}$$

$$e_{Even,2} = \sum_{n,k,j=1} \alpha_n \frac{n!}{2^n} \frac{1}{(\frac{n}{2})!(\frac{n-k}{2})!(\frac{k}{2}-j)!j!} a^{n-k} b^{k-2j} n_{Ev}^j \tag{6.49}$$

Approximating the two error terms by again assuming that $|a| \gg |b|$ and retaining only the dominant terms for which $k = 2$ and $j = 1$ yields (6.50) and (6.51) below.

$$e_{Even,1} \approx \sum_{n=2} \alpha_n \frac{n!}{2^n} \left[\frac{1}{(\frac{n-2}{2})!} \right]^2 \left(1 - \frac{2}{n} \right) (a^{n-2} b^2 - E[a^{n-2}]E[b^2]) \tag{6.50}$$

$$e_{Even,2} \approx \sum_{n=2} \alpha_n \frac{n!}{2^n} \left[\frac{1}{(\frac{n-2}{2})!} \right]^2 \frac{2}{n} a^{n-2} n_{Ev} \tag{6.51}$$

As in the analysis for the odd-order cancellation residual, not all of the terms in (6.50) actually contribute to error. To see why, one can consider that the term $a^{n-2}b^2$ can be expanded as in (6.52).

$$a^{n-2}b^2 = (a^{n-2} - E[a^{n-2}])(b^2 - E[b^2]) + (a^{n-2} - E[a^{n-2}])E[b^2] + E[a^{n-2}](b^2 - E[b^2]) + E[a^{n-2}]E[b^2] \tag{6.52}$$

The second term of (6.52) does not contribute to non-cancellable error, while the fourth term cancels the dc term in (6.50). Removing the second term and cancelling the dc error results in the expression (6.53).

$$e_{Even,1} \approx \sum_{n=2} \alpha_n \frac{n!}{2^n} \left[\frac{1}{(\frac{n-2}{2})!} \right]^2 \left(1 - \frac{2}{n} \right) (a^{n-2}(b^2 - E[b^2])) \tag{6.53}$$

The respective powers of the two error components are given by (6.54) and (6.55).

$$E[e_{Even,1}^2] \approx$$
$$\sum_{n,m=2} \alpha_n \alpha_m \frac{n!m!}{2^{n+m}} \left[\frac{1}{(\frac{n-2}{2})!(\frac{m-2}{2})!} \right]^2 \left(1 - \frac{2}{n}\right)\left(1 - \frac{2}{m}\right) E[a^{n+m-4}](E[b^4] - E[b^2]^2) \quad (6.54)$$

$$E[e_{Even,2}^2] \approx \sum_{n,m=2} \alpha_n \alpha_m \frac{(n-1)!(m-1)!}{2^{n+m-2}} \left[\frac{1}{(\frac{n-2}{2})!(\frac{m-2}{2})!} \right]^2 E[a^{n+m-4}] E[n_{Ev}^2] \quad (6.55)$$

The IER resulting from this cancellation procedure is given by (6.56).

$$IER_{Even} \approx \frac{E[y_{Even}^2]}{E[e_{Even,1}^2] + E[e_{Even,2}^2]} \quad (6.56)$$

6.4.4.1 Remarks

It is seen that the error between the true IMD products and the regenerated IMD products decreases as one of the two blockers dominates the other. Hence, the proposed successive intermodulation distortion regeneration scheme results in approximations to the true baseband IMD products in the nominally linear receiver path that are increasingly valid as one of the blockers dominates the other.

This analysis assumes that the nonlinearity in the nominally linear (main) receiver path occurs at RF and is memoryless. However, given that the frequency responses seen in the receiver RF circuitry at the blocker frequencies vary little in amplitude and phase over the bandwidth of the blocker signals, the narrowband approximation can be invoked, making the IMD products susceptible only to single-valued phase and amplitude considerations as a function of frequency at given frequencies of interest. In the context of the nonlinear transfer function approach to the analysis of weakly nonlinear circuits [121, pp. 81-90], this implies that the presence of memory in the nonlinear system will imbue a separate amplitude and phase rotation to each of the polynomial terms falling at a given intermodulation offset frequency. Given that the proposed adaptive cancellation scheme allows for a different phase rotation for each nonlinear reference basis element, these individual phase rotations for each of the polynomial terms will not constitute a limiting factor in the process of cancelling the baseband IMD products. It is important to note that despite the large number of polynomial terms required to model it to a high degree of accuracy, the sinusoid-like nonlinearity of the LNTA is still considered to be weak due to the fact that the transfer characteristic in Fig. 5.12 does not cut off abruptly, the criterion for a strong nonlinearity [121, pp. 1-4].

Depending on the precise implementation of the RF/analog circuit blocks, the case may also occur in which the envelope of the large blocker is implicitly downconverted to baseband frequencies and coupled into the biasing network. Here at low frequencies, the downconverted blocker envelope may experience considerable frequency-dependent memory effects. Once present in the biasing circuitry, this filtered even-order distortion then interacts with the input signal and the odd-order polynomial terms of the RF circuit blocks to effect higher-order odd IMD products. This possibility likewise need not constitute a limiting scenario. In principle, compensatory reference IMD products for these non-memoryless terms may be generated by passing one or more successive powers of the even

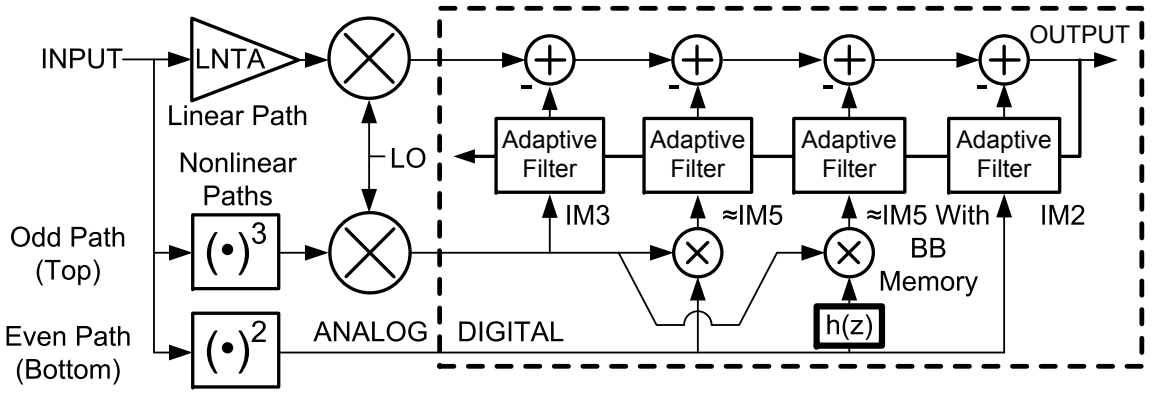


Figure 6.9: Successive regeneration and adaptive feedforward cancellation of IMD products at baseband accounting for envelope baseband memory effects.

nonlinear reference signal through discrete-time models of the analog baseband transfer function prior to multiplying it by the odd-order reference IMD basis elements. This concept is depicted in Fig. 6.9 in which an IM5 reference term is generated with a discrete-time model $\mathbf{h(z)}$ of the baseband envelope filtering.

Computing the predicted IER of the proposed successive IMD regeneration and cancellation scheme is seen to require knowledge of the higher-order moments of the dominant blocker signal. Knowledge of the second- and fourth-order moments of the nondominant blocker signal is also required. In many cases, including those discussed in Section 6.3.4, the statistical properties of the dominant blocker are known *a priori*, and therefore a closely-approximated analysis may be carried out in the early stages of the receiver design. In the event that this is not possible, one option is to note that if the dominant blocker is expected to be a digitally-modulated signal, the ratios of its higher-order central moments to that of its variance are bounded by those of a Gaussian-modulated signal, as shown in the next section.

6.5 Simulation of Residual Error from Successive Regeneration Approximation

In order to arrive at quantifiable conclusions regarding the relationship between the ratio $E[a^2]/E[b^2]$ and the INR of the two nonlinear paths, MATLAB simulations are performed in which the odd- and even-order IMD products in the nominally linear receiver path are produced along with the odd- and even-order nonlinear reference signals. Subtraction of the two signals takes place after MMSE weighting of the elements in the reference basis and the residual error power then compared to the original signal power. At the same time, the simplified expressions developed in the previous section are evaluated numerically and the results placed alongside the simulation results for comparison.

6.5.1 General Information

6.5.1.1 Evaluated Waveforms

In order to maximize the signal power passing through the LNTA while not exceeding its breakdown voltage, QPSK was chosen to be the modulation scheme of the dominant blocker a. In the simulations and numerical calculations below, the input-referred power of a is not swept but rather fixed at +12.5dBm. The nondominant blocker b is modeled as either a CW tone or a Gaussian signal. These two cases are denoted in the plots below by the shorthand "S1" and"S2", respectively. The CW tone is evaluated due to the fact that many communications standards specify performance under CW blocking and due to the fact that modeling blocker scenarios with a CW tone simplifies the testing apparatus. However, in these specifications this CW tone is often merely a proxy for a bandpass modulated signal. Furthermore, in the expressions from the previous section it is seen that $e_{Odd,1}$ and $e_{Even,1}$ evaluate to zero in the presence of a CW nondominant blocker signal since $b^2 = E[b^2]$.

Given this, it makes sense to evaluate the residual error $e_{Odd,1}$ and $e_{Even,1}$ for the case in which the kurtosis of the nondominant blocker is at a relatively large value. Although in principle the kurtosis of the nondominant blocker can approach infinity in the case of a broadband pulse-based signal, in this event the blocker bandwidth would be extremely wide and would include the receiver LO frequency. Consequently in such a case, the interference energy within the receiver is dominated by that of the first-order (i.e. linear) Taylor-series term while the IMD products would contribute much smaller amounts of interference energy. Whether considering the receiver nonlinearity or not, therefore, such a problematic situation can be dealt with utilizing solutions already described in the literature, including the use of median prefilters [122] and maximum-likelihood RAKE receivers [123].

Discarding the requirement of handling such broadband signals, the objective then turns to finding a narrowband signal with a kurtosis exceeding that of most of the common modulation schemes that may serve as one of the nondominant blockers. A bandpass Gaussian-modulated (AWGN) signal satisfies this objective, as shown in Fig. 6.10. Furthermore, in Fig. 6.10 it is seen that all of the higher-order moments of the Gaussian signal exceed those of a number of common modulation schemes.

For each of the waveforms evaluated herein, the band-limited nature of the signals is enforced with the application of root-raised cosine filters similar to those used in wireless communications systems. The chip rate of the modulated signals is 2MSPS while the sampling rate of the simulation is 32MHz.

6.5.1.2 Computation of Nonlinear Path INR

In order to simplify the flow of the simulation code, the INR of the two nonlinear paths is computed in an *a posteriori* fashion. For this reason, the INR of the two nonlinear paths will vary amongst the various test cases considered below. Furthermore, the INR is maintained roughly constant over each $E[a^2]/E[b^2]$ sweep by pegging the noise rms power to the power of the nondominant blocker b. As in the even-order IMD computations, the INR of the even nonlinear path is computed utilizing only the non-dc portion of the baseband even-order IMD products due to the fact that the dc portion does not contribute to the successive regeneration of higher-order terms.

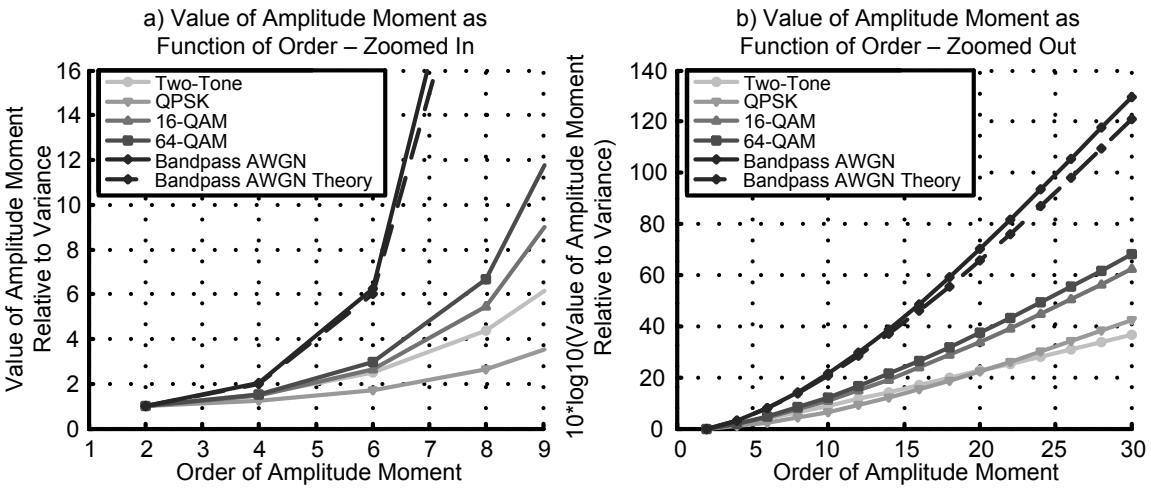


Figure 6.10: Comparison of higher-order moments of the amplitudes of several narrowband modulation schemes. a) Zoomed in. b) Zoomed out (Log Scale).

6.5.2 Odd-Order Intermodulation Products

6.5.2.1 Term Generator and Frequency Offsets Considered

Three conditions are considered while evaluating the successive regeneration scheme. First, the case of a pure third-order odd nonlinear term generator is considered while evaluating IMD products at the third-order IMD offset frequency (i.e. in the analysis above, $l = 2$ and $\chi_3 = 1; \chi_m = 0, m > 3$). In the plots below, this case is denoted by the shorthand"T1F1". Second, the case of a odd nonlinear term generator characterized by $\chi_3 = 1$, $\chi_5 = 0.3$, and $\chi_m = 0, m > 5$ is considered. Here, the successive regeneration scheme can generate compensatory IMD products for $l = 2$ and $l = 3$, where these two cases are handled separately using the shorthand terminology "T2F1" and "T2F2", respectively.

6.5.2.2 Receiver Odd-Order Polynomial Model

Because to isolate the IMD products that fall around the LO frequency out of the many other frequencies at which IMD products occur requires a polynomial expansion of the effective receiver nonlinearity in the analysis of Section 6.4, a polynomial model of the receiver nonlinearity was utilized for this exercise. In order to speed up the computations and to minimize effects due to numerical instability, an 11th-order polynomial best-fit (conducted over the range (-2.7V, 2.7V)) model of the simulated LNTA was utilized. The result of the best-fit process is shown in Fig. 6.11a and represents the polynomial shown in (6.57) where the input x is in volts and the output y is in mA. Furthermore, the third order polynomial term is decremented by 0.02667 (i.e. $\alpha_{3,Tot} = -2.036$) in order to represent a passive mixer with a receiver-input-referred IIP3 of +40dBm.

$$\begin{aligned} y_{LNTA}(mA) &= \alpha_3 x^3 + \alpha_5 x^5 + \alpha_7 x^7 + \alpha_9 x^9 + \alpha_{11} x^{11} \\ y_{LNTA}(mA) &= -2.010x^3 + 1.506x^5 - 0.433x^7 + 0.056x^9 - 0.002645x^{11} \end{aligned} \tag{6.57}$$

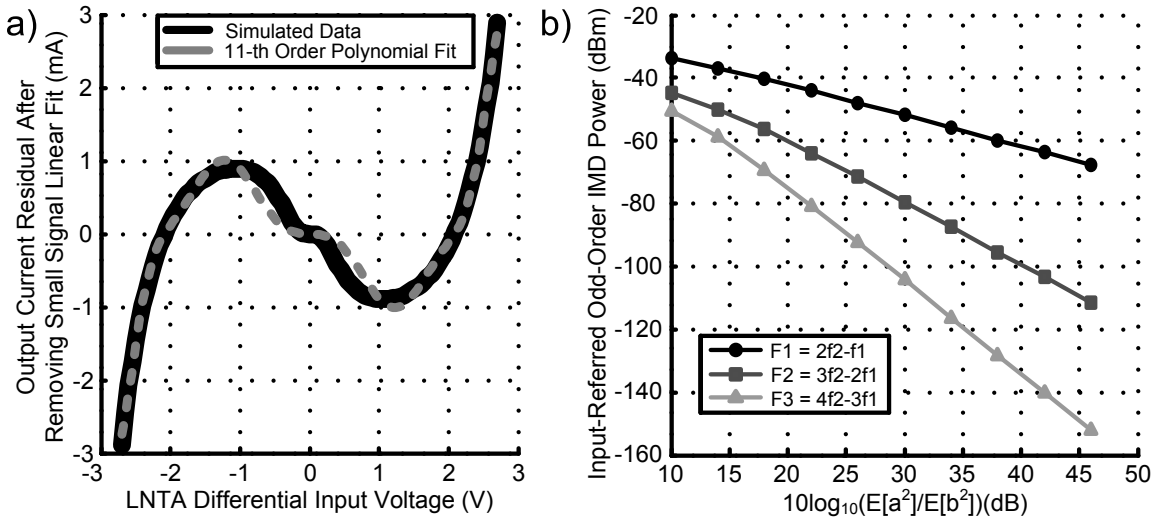


Figure 6.11: a) 11th-order polynomial best-fit to LNTA input-output characteristic utilized for successive regeneration scheme residual error simulations. b) Simulated input-referred odd-order IM distortion for 11th-order polynomial in a) and "S2" input signal for various IMD offset frequencies.

Figure 6.11b shows the input-referred odd-order IM distortion error as a function of the IMD frequency offset. It is seen that for the 7th-order IMD offset of $4f_2 - 3f_1$, the input-referred error is below or near the required sensitivity levels of the common communications standards detailed in Table 1.1 when $10\log_{10}(E[a^2]/E[b^2])$ >30 dB. Depending on the exact standard, little or no IMD cancellation would be required under this condition. Higher-order LNTA polynomial models were also examined and yielded similar input-referred error results to those from the 11th-order polynomial.

6.5.2.3 Reference Basis Set

As postulated in Section 6.3.6, in addition to a three-element polynomial odd-order reference basis set (3rd-, 5th-, and 7th-order), a functional basis element is added to explicitly model the LNTA nonlinearity. In this case, the exact same polynomial model as (6.57) is utilized, as the purpose of this quantitative exercise is to determine the IER due to the successive regeneration dominant-blocker approximation and nonlinear path noise, not to determine the IER due to nonidealities within the reference basis set. Should only one of these functional basis elements be used, however, in the "T2F1" case there will exist significant uncancelled residual due to higher-order IMD terms regenerated in the impure odd-order nonlinear term generator. In order to reduce this residual, three functional basis elements are utilized, each based on the polynomial of (6.57). That is, the basis elements are constructed as shown in (6.58) and in Fig. 6.12.

$$z_{Odd,Fxn1} = r_{Odd}(\alpha_3\gamma_{3,l-1,l} + \alpha_5\gamma_{5,l-1,l}r_{Even} + \alpha_7\gamma_{7,l-1,l}r_{Even}^2 + \alpha_9\gamma_{9,l-1,l}r_{Even}^3 + \alpha_{11}\gamma_{11,l-1,l}r_{Even}^4) \quad (6.58)$$
$$z_{Odd,Fxn2} = z_{Odd,Fxn1}r_{Even}; z_{Odd,Fxn3} = z_{Odd,Fxn2}r_{Even}$$

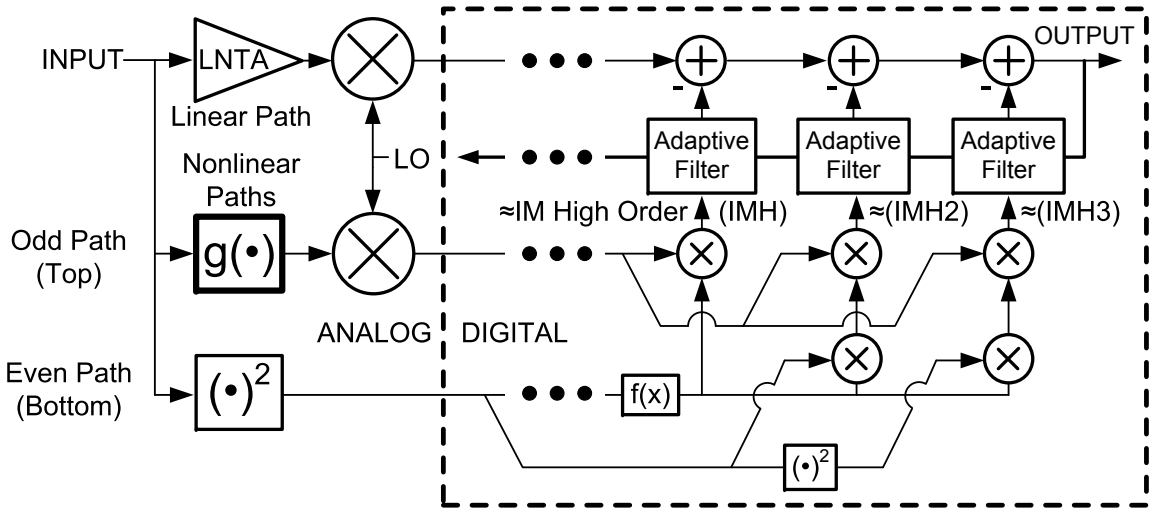


Figure 6.12: Concept of higher-order functional reference basis elements for case of impure-odd-order nonlinear term generator.

6.5.2.4 Mathematical Simulation Setup

The simulations reported below operate completely at baseband to minimize the required computational power. That is, the simulated y_{Odd} and z_{Odd} are provided by (6.14) and (6.17), respectively, in addition to a downconversion operation (here modeled by setting $2\pi(2f_2 - f_1)t = \pi/4$). Applying this phase offset of $\pi/4$ accommodates the situation in which the I and Q components of the odd-order IMD have different in-band power, such as that described in Section 4.7, by giving equal weighting to each of these components in a unified analytical expression.

Also in the simulated case, the MMSE relation is used to determine the proper weightings β_n in (6.17). In the calculated case, the weightings β_n are determined by applying the MMSE relation to (6.23) while the IER is determined by (6.28)-(6.33). Furthermore, for the case of $l = 3$ the substitution of $(b^2 - E[b^2])$ is made into b^{l-1} as mentioned previously in Section 6.4.2.

6.5.2.5 Results and Remarks

Performing the numerical computations for the case of "S2" and in which the nonlinear reference paths are noiseless result in the curves shown in Fig. 6.13. It can be seen that calculation and simulation match up well and that the IER is roughly equal to $20\log_{10}(E[a^2]/E[b^2]) - 10$dB for large $E[a^2]/E[b^2]$. To put this in the context of a common telecommunication standard, using UMTS as an example, a rail-to-rail input signal due to the TX leakage would be +12.5dBm after some up-front filtering, while the nondominant blocker might be about 30dB less than this after being attenuated somewhat from its specified -15dBm input-referred value. In this case, the achievable IER is close to 50dB, meaning that the input-referred IMD error power of slightly less than -50dBm seen at the input of the LNTA could be attenuated to over -100dBm under this cancellation scheme, a level which is not substantially greater than the thermal noise floor of the input source (-108dBm in UMTS).

For small and decreasing values of $E[a^2]/E[b^2]$ it is seen that the simulated IER begins leveling

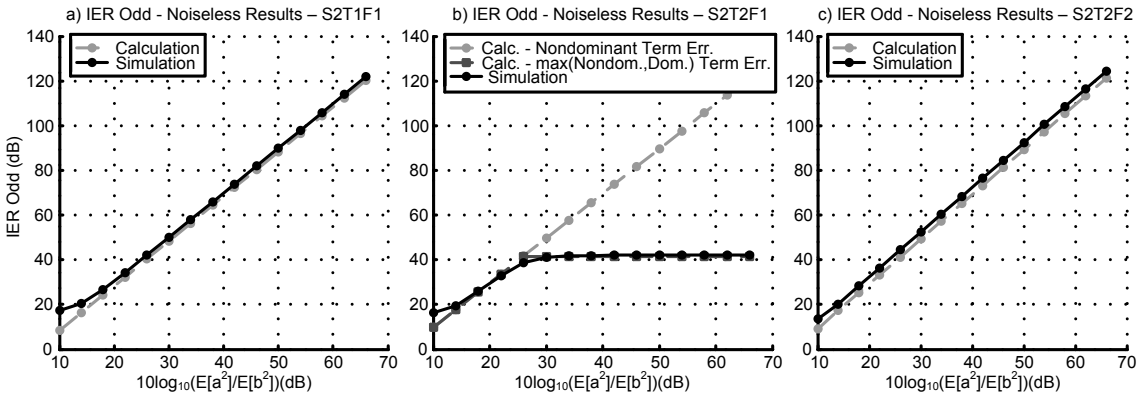


Figure 6.13: Simulation and calculation results of approximation error IER for case "S2" as function of $E[a^2]/E[b^2]$ for noiseless nonlinear paths.

off while the calculated results maintain the same trajectory. This is due to the fact that the calculated IER depends solely on the statistics of a and scales proportionally with the power of b. In the simulated case, as b becomes larger in amplitude, the smaller residual error terms neglected in the simplified calculations become significant. As a result of the MMSE condition, these smaller residual terms partially cancel larger residual error terms, resulting in an improvement in the IER not predicted by the calculations.

Furthermore, it can be seen in the case of "S2T2F1" that taking the maximum of the dominant and nondominant error terms accurately predicts the achievable IER. Here, the achievable IER ceiling is at 41dB and can be improved upon by adding functional basis elements as was done in Fig. 6.12.

Adding noise to the nonlinear paths results in the curves shown in Figs. 6.14-6.19. Simulated results are shown in bold curves, while the results from the simplified calculations are shown in dashed curves. In general, simulation and calculation match well. It can be seen in the case "S2T1F1" that in order to achieve a particular IER, the even path target INR should be about equal to the target IER while the odd path target INR should be about 5-10dB greater than the target IER.

One notable feature of many of the plots shown in Figs. 6.14-6.19 is that as $E[a^2]/E[b^2]$ decreases, the simulated IER improves with respect to the calculated IER, even when the IER is dominated by nonlinear reference path thermal noise. This is due to nondominant noise error terms neglected in the calculation becoming more significant and partially cancelling the dominant noise terms.

6.5.3 Even Order Intermodulation Products

6.5.3.1 Term Generator and Frequency Offsets Considered

For the even nonlinear path, there only exists one frequency offset, that of the frequencies immediately surrounding dc. This fact, combined with the ease of designing an effectively pure square-law even-order term generator, obviates the need to handle multiple test cases here.

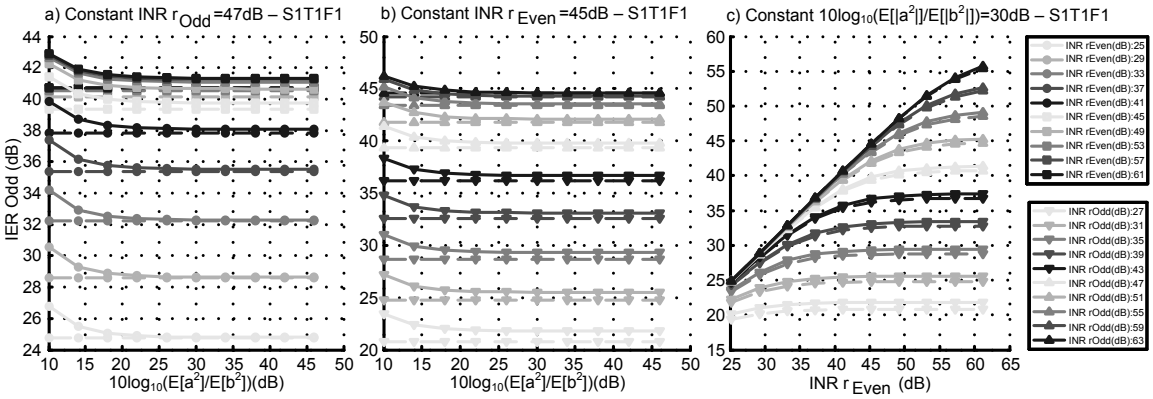


Figure 6.14: Simulation and calculation results of approximation error IER_{Odd} for case "S1T1F1".

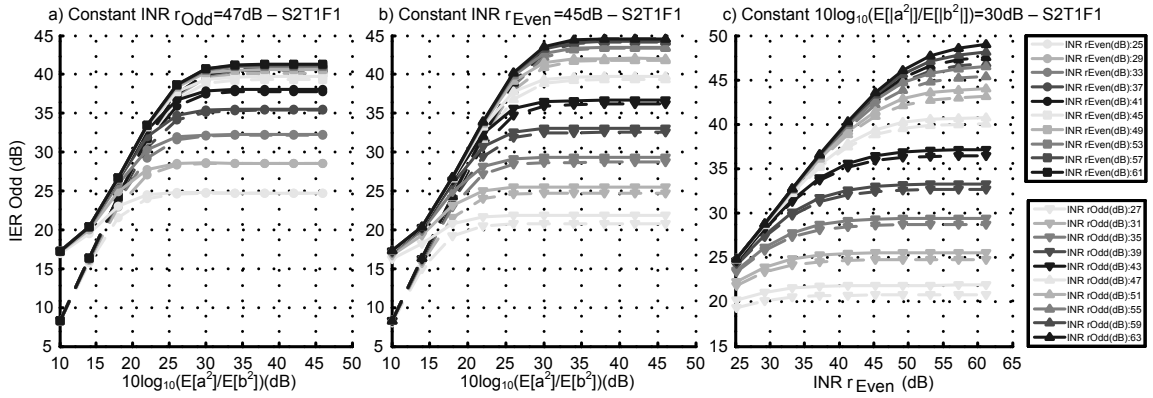


Figure 6.15: Simulation and calculation results of approximation error IER_{Odd} for case "S2T1F1".

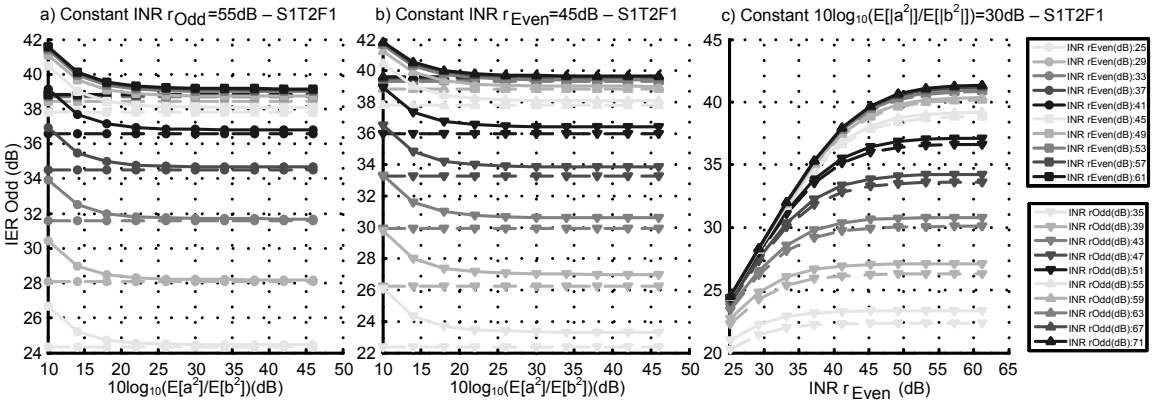


Figure 6.16: Simulation and calculation results of approximation error IER_{Odd} for case "S1T2F1".

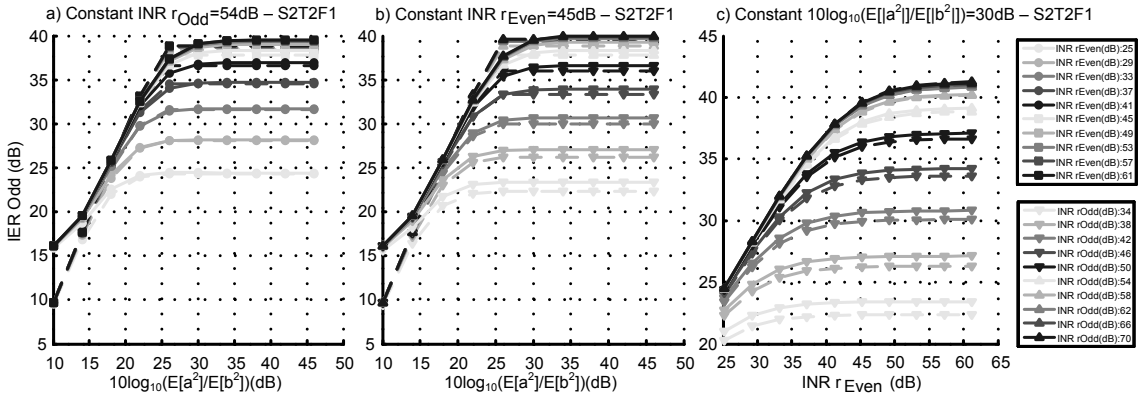


Figure 6.17: Simulation and calculation results of approximation error IER_{Odd} for case "S2T2F1".

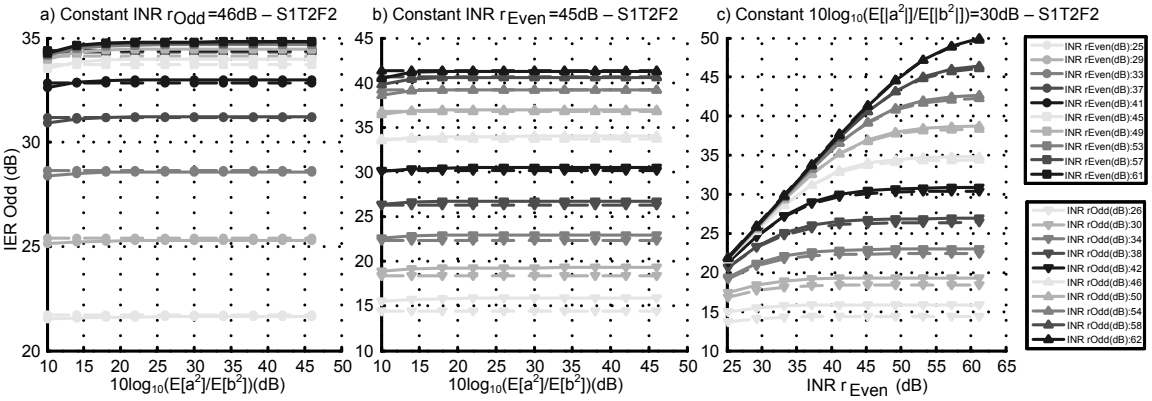


Figure 6.18: Simulation and calculation results of approximation error IER_{Odd} for case "S1T2F2".

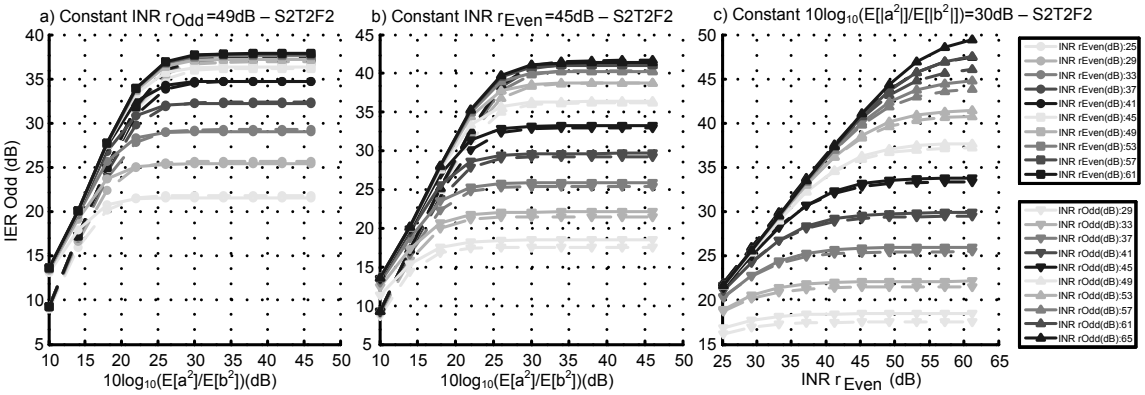


Figure 6.19: Simulation and calculation results of approximation error IER_{Odd} for case "S2T2F2".

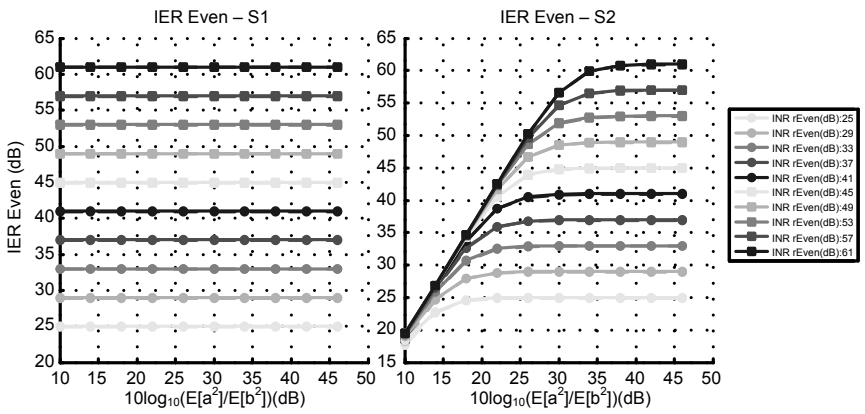


Figure 6.20: Simulation and calculation results of approximation error IER_{Even}.

6.5.3.2 Receiver Even-Order Polynomial Model

The even-order polynomial model used to evaluate the successive regeneration approximation error is developed from the measured even-order intermodulation distortion reported in Fig. 7.29b. The coefficients α_2, α_4 were found to be 0.026 and -0.004 respectively, while α_6 was chosen somewhat arbitrarily to be 0.00015 in order to provide a non-negligible contribution to the total IMD power.

6.5.3.3 Reference Basis Set

Because the even-order coefficients of the measured receiver drop off in magnitude very quickly, a functional basis element was not required here. Hence, a three-element polynomial basis consisting of IM2, IM4, and IM6 reference terms was utilized.

6.5.3.4 Mathematical Simulation Setup

The simulated waveforms were generated with the use of (6.42) and (6.44). The coefficients β_n were determined via the MMSE relation and the reference terms subtracted from the original IMD terms, with the power of the resultant error compared to the power of the original signal to obtain the IER. The equations utilized to compute the simplified calculation IER are given by (6.43), (6.54), (6.55), and (6.56).

6.5.3.5 Results and Remarks

Performing the numerical computations specified above result in the curves shown in Fig. 6.20. Although as for the odd-order case calculated results are plotted with dashed lines, in this case the agreement with simulation is close enough to obscure the view of this data. Clearly the curves in the left hand plot correspond to the case where a CW nondominant blocker signal is applied to the system and there does not exist any degradation of IER as the ratio $E[a^2]/E[b^2]$ is reduced. However for $10\log_{10}(E[a^2]/E[b^2])$ >30 dB the results are nearly identical irrespective of whether a CW or bandpass Gaussian nondominant blocker is applied to the nonlinearity model.

Chapter 7

Implementation and Testing of a Rail-to-Rail Input Receiver with Successive Regeneration and Adaptive Cancellation of Intermodulation Products[1]

7.1 Receiver Architecture

Having introduced the wide-swing LNTA in Chapter 5 and the higher-order IMD-cancellation approaches in Chapter 6, the tools now exist to build a receiver capable of withstanding large input signals on the same order of magnitude as the receiver supply voltage. Since the LNTA and nonlinear reference blocks must interface monolithically with the rest of the circuitry in a passive-mixer based receiver, a completely new RF/Analog die must be designed and fabricated in order to validate the proposed concepts.

In order to guarantee repeatable performance of the receiver, the design was approached with the goal of handling a differential rail-to-rail input signal. Although in principle the wide-swing LNTA is useful for input signals larger than this, oxide breakdown of the input devices may occur and render the performance of the receiver unreliable.

The recommended supply voltage of a 90nm process is 1.0V. However, if the constituent circuit blocks are designed in such a way that the devices are stacked, the supply voltage may be increased without concern for device degradation. To ensure that device degradation mechanisms do not come into play over large-signal operation, the circuits may be simulated and their internal voltages monitored to check that the MOSFET terminal voltages do not exceed a particular value representing a voltage with some margin below the MOS oxide breakdown voltage.

In order to accommodate an input signal of appreciable magnitude, a supply voltage of 1.5V was chosen in order to comfortably accommodate the four-transistor stack required of the wide-swing LNTA in Chapter 5. In this case, a differential rail-to-rail input signal is on the order of +12.5dBm to +13.5dBm. As this large signal is filtered after downconversion to baseband, it must first pass

[1] Portions of this material have been previously published in [119] and copyright is owned by IEEE.

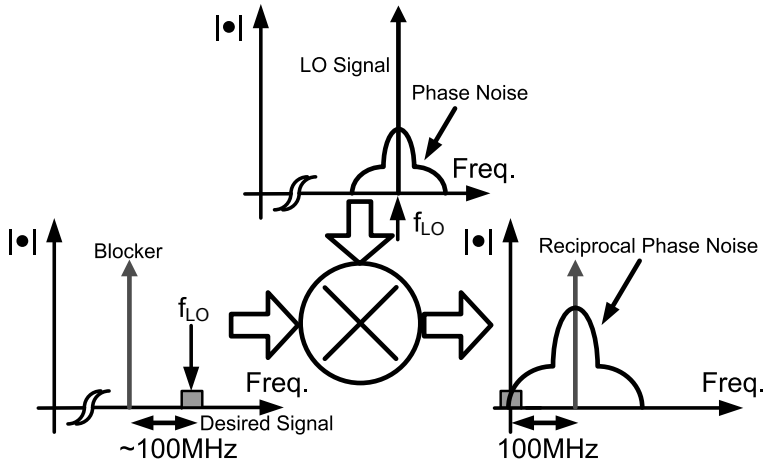


Figure 7.1: Depiction of reciprocal mixing with large blocker.

through the passive mixer where it will downconvert LO phase noise into the desired signal band via reciprocal mixing, as depicted in Fig. 7.1. A quick calculation in (7.1) shows that the phase noise of the LO at the offset of the large blocker must be less than -180dBc/Hz to obtain a noise figure of 6dB due to this effect alone. In (7.1), $P_C(dBm)$ is the LO carrier power, $P_B(dBm)$ is the power of the blocker, $P_{N,\phi}(dBm/Hz)$ is the absolute phase noise power per unit bandwidth, and $P_{N,IR}(dBm/Hz)$ is the effective input-referred noise power. Surveying the phase noise performance of several top-end Agilent and Anritsu signal generators reveal a typical thermal phase noise floor of around -150dBc/Hz [124] [125] and -160dBc/Hz [126], which is much higher than desirable for the goals of this project. Due to the insufficiency of these potential LO sources, an on-chip frequency generation solution must be added to the RF/Analog die. At minimum, the VCO along with a divider chain is required to be integrated on-chip, while the baseband portion of a PLL may be implemented off-chip.

$$\begin{aligned} P_{N,IR}(dBm/Hz) &= NF(dB) + 10\log_{10}(kT) + 30dB \\ P_{N,IR}(dBm/Hz) &= P_{N,\phi}(dBm/Hz) + P_B(dBm) - P_C(dBm) \\ P_{N,IR}(dBm/Hz) - P_B(dBm) &= (P_{N,\phi} - P_C)(dBc/Hz) \\ NF(dB) + 10\log_{10}(kT) + 30dB - P_B(dBm) &= (P_{N,\phi} - P_C)(dBc/Hz) \\ NF(dB) - 186.3dB &= (P_{N,\phi} - P_C)(dBc/Hz) \end{aligned} \tag{7.1}$$

For the purposes of this academic exercise, the analog-to-digital converters and digital back end are implemented off-chip in discrete, commercially available ICs, and in software, respectively. As the design of analog-to-digital converters for radio receivers is a well-established art and as the large blockers are proposed to be filtered out within the RF/Analog front end, integrating ADCs onto the RF/Analog die would yield very little in the way of improving the proof-of-concept of this receiver. Similarly, while practical challenges of integrating analog and digital hardware on the same die always exist, they are typically handled on a case-by-case basis and are usually surmountable. Hence, implementing and integrating the digital back end algorithm would similarly add little to

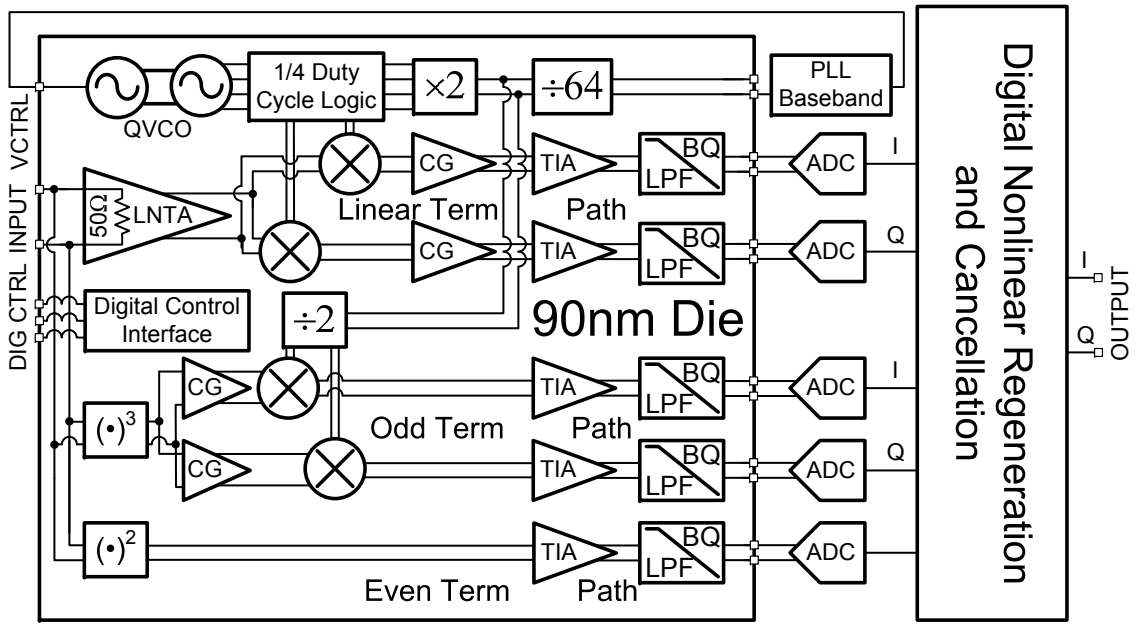


Figure 7.2: Proposed receiver architecture with detail on RF/Analog front end.

the proof of concept.

7.1.1 RF/Analog Front End Architecture

Due to the current-domain approach of the large-signal receiver, there do not exist any internal voltage-domain nodes at which the nonlinear paths may branch off from the nominally linear receiver path as in the architecture described in Chapter 3. As shown in Fig. 7.2, all three receiver paths access the chip input, which is the only RF node with appreciable signal voltage amplitude. Although the voltage at this node is slightly distorted due to the nonlinear impedance division between the source impedance and the LNTA input impedance, the magnitude of the distortion relative to the large input blockers is small enough such that explicit nonlinear interaction between these distortion terms and the much larger blockers is negligible relative to nonlinear interaction between the blockers themselves. In other words, unintentional higher-order distortion terms in the nonlinear paths due to the nonlinearity of the LNTA input impedance are negligible with respect to the magnitude of the intentionally generated higher-order reference distortion terms.

To see why, the reader may consider from a quantitative perspective a peak asymmetric blocking condition of incident signals with powers of +12.5dBm and -16.5dBm. Based on the simulated results of Fig. 6.11b, the IMD products at the LO frequency are expected to be slightly less than -50dBm at the input of the receiver. Any nonlinear reference IMD terms generated in the nonlinear term paths from the interaction between this signal and the smaller blocker are at least 43.5dB smaller in magnitude than the IMD resulting from the interaction of the smaller blocker with itself. Even if IMD cancellation greater than 40dB is desired it is important to note that the interaction of these IMD products at the LO frequency with the original blockers results in higher-order IMD terms in the nonlinear reference paths that can be canceled using the successive regeneration scheme

described in Chapter 6. Hence, placement of the nonlinear reference circuitry at the input of the LNTA does not constitute a significant limitation to the design at hand.

7.1.1.1 Linear Path

Processing the desired communication signal is the responsibility of the linear path of the receiver and as such it is composed of nominally linear circuit blocks. The LNTA interfaces to the quadrature passive mixers in a capacitive fashion via the capacitors in Fig. 5.19. In reality, each of the four capacitors shown in Fig. 5.19 constitutes two parallel capacitors, each of which goes to one of the two quadrature mixers. In order to isolate the I and Q downconversion chains with minimal voltage swing at the LNTA output, a 1/4-phase passive mixer scheme was used. The noise generated by the transimpedance amplifier (TIA) in a passive mixer system is a well-known problem in cases such as this, where the impedance looking back up into the passive mixer is low [51]. In order to provide a high source impedance to the TIA, it is preceded by a common-gate (CG) buffer, thereby lowering its effective noise contribution. This technique was previously shown in [127] although the details of the CG buffer in this case were not shown. When placed in parallel with a very large (335pF) differential capacitor, the input impedance of the CG buffer also aids in attenuating the amount of large downconverted blocker that is passed to the remainder of the receiver chain. A second-order active RC biquad (BQ) is utilized to both buffer the TIA and to complete a 3rd-order Chebyshev low-pass anti-aliasing filter. The biquad outputs are designed to drive the discrete ADCs through the ESD network and remain stable over process corner even when loaded with 20pF of capacitance. The 3-dB cutoff frequency of the composite (CG-TIA-BQ) filter is approximately 2.3MHz so as to avoid introducing substantial group delay distortion for desired signals occupying double-sided bandwidths up to 4MHz. Coarse dc offset cancellation is provided by adding a differential static current to the virtual ground nodes of the first biquad OTAs. This allows the receiver to process large baseband IMD products even in the presence of large dc offset.

7.1.1.2 Odd Nonlinear Path

Utilizing a distributed cubic term generator with an architecture similar to the one described in Chapter 4, the odd nonlinear path principally generates 3rd-order IMD products. The output of the cubic term generator is a current-mode signal which is buffered by CG amplifiers that provide separation between the I and Q passive mixers that downconvert the reference odd-order IMD products to baseband. Because the odd-order IMD reference current is small relative to the current that must be handled by the linear path, the I/Q buffering may be done in such a way that places a large impedance in series with the signal path. In this case, a more traditional passive mixer scheme may be employed as opposed to a 1/4-phase scheme that requires greater power due to the sharper rise and fall times of the mixer drive waveforms. Since the ratio between the magnitudes of the desired odd-order IMD products and undesired odd-order IMD products (namely those appearing at frequencies other than that of the LO) is much smaller than that of the undesired blocker and the desired signal, the required order and out-of-band attenuation of the baseband anti-alias filtering in the nonlinear paths is much less than in the linear path. Hence, the large filtering capacitor in the linear path is not required here. Similarly, the large output impedance presented by the CG buffers, even at RF, prevents the input-referred TIA noise from being substantially amplified, although care

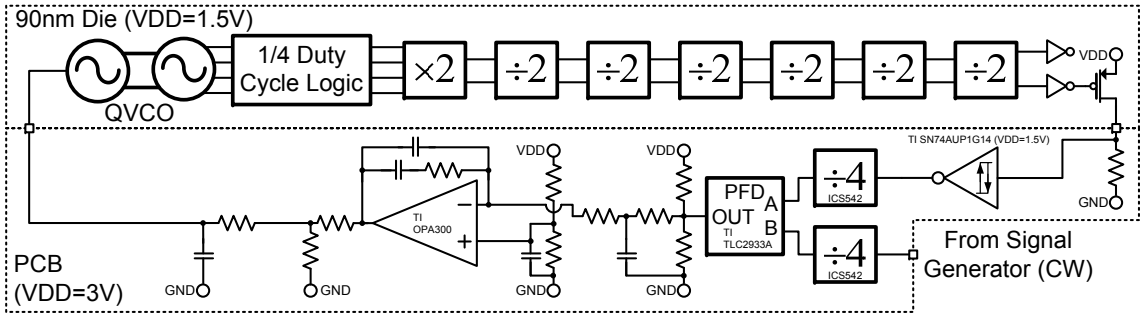


Figure 7.3: Implementation of PLL loop with off-chip components.

must be taken to avoid the introduction of substantial parasitic capacitance at this node. Like the linear path, 2nd-order biquads are utilized to buffer the TIA, to complete a 3rd-order Chebyshev filter, and to drive the discrete ADC inputs. In this case, the 1-dB and 3-dB cutoff frequencies of the filters are 6.5MHz and 7.2MHz, respectively, and the in-band ripple is minimized in order to approximate the effectively memoryless (aside from a constant group delay) analog signal path required by the successive regeneration scheme to obtain large IMD cancellation ratios for blocker signals of double-sided bandwidth in the MHz range. The larger cutoff frequency of the Chebyshev baseband filters relative to those of the linear path is made possible by the reduced amplitude discrepancy between desired and undesired signals relative to the linear path as described in Section 3.3. In this case, less attenuation is required in the anti-alias filtering than in the linear path to achieve the desired IER. Coarse dc offset cancellation circuitry is present in the nonlinear path biquads as well but was not utilized during testing.

7.1.1.3 Even Nonlinear Path

Fronted by a canonical MOS squaring circuit, the even nonlinear path incorporates the same baseband building blocks as the odd nonlinear path, saving design time. Since the desired even-order (principally IM2) products already exist at baseband, there is no need for any frequency conversion.

7.1.1.4 Frequency Generation

Accurately setting the reference frequency of the experimental receiver requires a complete PLL. Although the PLL is not required to be active for IMD cancellation to properly occur, enabling the PLL is required to obtain reliable measurements of complex signals passing through the receiver. The principal purpose of this PLL is to prevent large-scale low-frequency drift of the LO frequency and to maintain the out-of-band phase noise performance of the QVCO rather than to achieve a particular in-band phase noise floor. Although the IMD products around the LO frequency can be large, the phase noise imbued onto these signals when downconverted is the same in both the linear and nonlinear paths. In principle, therefore, the in-band phase noise of the PLL does not prevent cancellation of IMD products, although it may slightly degrade the EVM of an incident desired signal.

The VCO oscillates at the LO frequency in order to minimize the out-of-band phase noise floor for a given power dissipation. Since the out-of-band phase noise floor is dictated by the thermal

noise of the devices at high frequencies, driven circuits such as frequency dividers will contribute substantially in this frequency range unless they are of inordinate power dissipation. Not running the VCO at 2x the LO frequency is an atypical choice due to the fact that it promotes LO-RF and RF-LO coupling, increasing dc offset and decreasing IIP2, respectively. However, in this architecture, dc offset is compensated by adding a differential static current to the first OTA virtual ground in the biquads. IM2 products are ultimately cancelled using the scheme described in Chapter 6, improving the effective IIP2 performance.

The PLL loop, shown in Fig. 7.3, begins off-chip with discrete frequency dividers which permit the use of low-frequency reference oscillator inputs and phase-frequency detectors. The outputs of the frequency dividers are compared using a TI TLC 2933A discrete phase frequency detector/charge pump circuit. As the "charge" pump of the TI TLC 2933A consists merely of low-impedance switches to each of the power supplies, another mechanism to regulate the current flowing into the loop filter must be implemented. This is done by first using a resistive voltage divider to establish a common-mode voltage of $V_{DD}/2$ at the charge pump output and then by low-pass filtering the resultant perturbations in voltage produced by the charge pump switching. Secondly, the output voltage of the low-pass filter and the virtual ground node of an active loop filter are applied across a resistor to establish the current to be integrated onto the loop filter capacitor. Because the discrete PLL components operate under a supply voltage of twice that of the 90nm RF/Analog die, a passive resistive divider is used to reduce the output of the active loop filter in order to avoid overvoltage damage to the chip. This output is in turn low-pass filtered in order to remove high-frequency noise on the reference line prior to re-entering the chip to act as the control voltage of the QVCO.

7.1.1.5 ADCs and Digital Interface

The analog outputs of the receiver are captured by 12-bit ADI AD9235 ADCs running at 25MHz. The digital outputs of the ADCs are received by an FPGA platform that serializes the data so that it can be sent to a logic analyzer/digital spectrum analyzer for data acquisition and real-time measurement of the receiver baseline metrics.

7.1.2 Digital Back End Architecture

For the proof-of-concept demonstration of the successive generation and adaptive cancellation of IMD products, the digital back end (DBE) is implemented in a fixed-point software model, with its architecture shown in Fig. 7.4. The nonlinear path inputs are upsampled and filtered prior to successive nonlinear reference generation to ensure that unwanted higher order nonlinear terms do not alias into the signal band. This filtering also compensates for the small amount of group delay distortion present in the nonlinear path analog baseband filters. After this process is complete, an approximate digital model of the analog linear path baseband filter removes undesired residue from these operations and helps to better match the known difference between the frequency responses of the linear and nonlinear paths.

The remaining frequency-domain difference between the linear and nonlinear path transfer functions is fine-tuned via LMS adaptive equalizers modified to compensate for I/Q mismatch, shown in Fig. 7.5. Quantized-NLMS adaptive equalizers modified to divide by the square root of the norm were placed on the IM2 and IM4 lines to reduce gradient noise amplification for large signal levels.

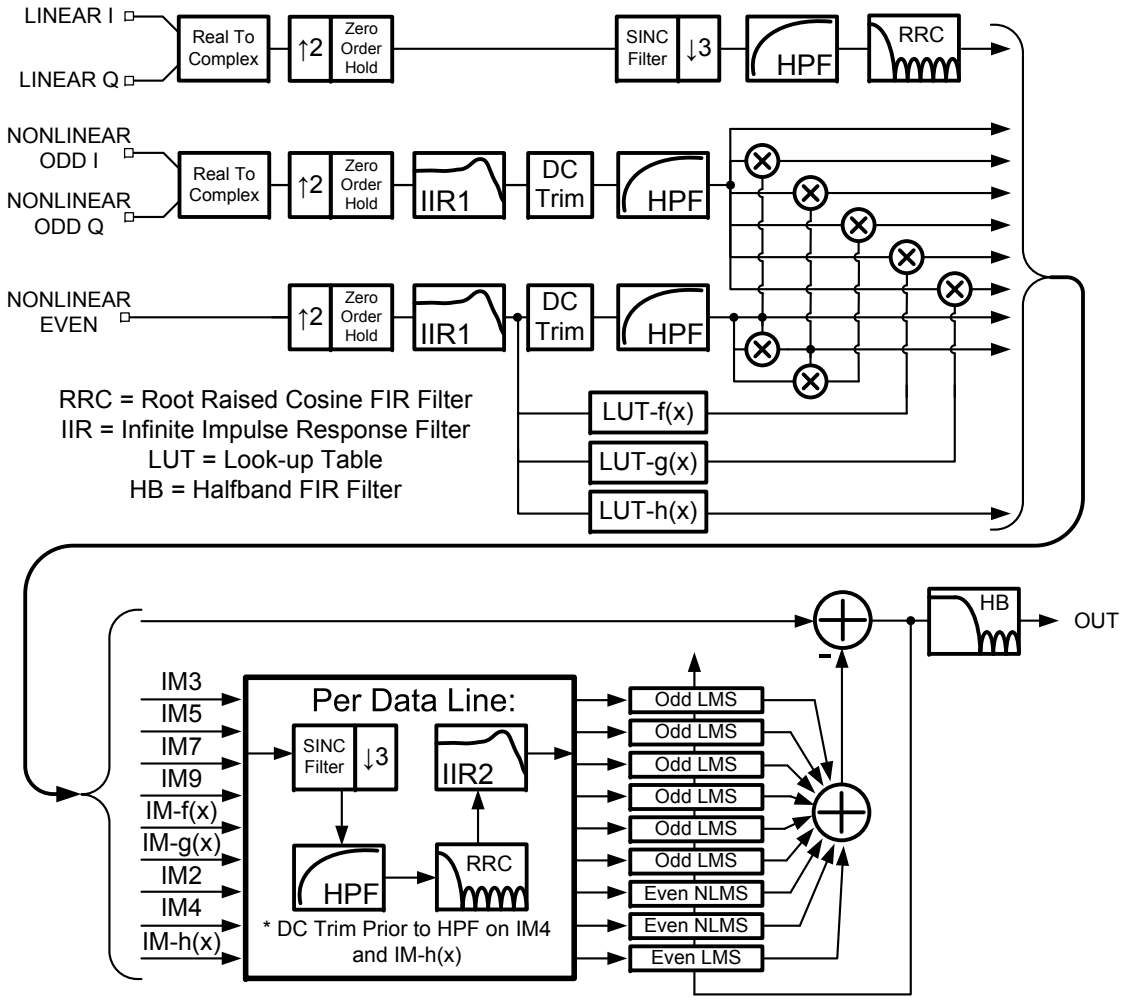


Figure 7.4: Proposed digital back end architecture.

Modifying to divide by the square root of the norm merely involves a bit shift, as shown in Fig. 7.6, and does not constitute an undue hardware burden.

High-pass filters remove dc offsets and 1/f noise from the incoming signals. Dc trimming circuits are utilized to cancel dc offsets prior to the high-pass filters and can be periodically turned on to measure the analog die output offsets at regular intervals when the IMD cancellation scheme is not in use. This is done so that the combination of dc offsets along with the high-pass filters do not result in large settling times during alternate path turn-on that would delay convergence of the adaptive algorithm. Dc trim circuits are also utilized prior to the high-pass filters on the higher-order even IMD lines, as dc signal is regenerated by even-order nonlinear operations.

Root-raised cosine filters provide a large degree of adjacent channel rejection and have a 3-dB cutoff frequency of 2.0MHz. It is important to completely remove large adjacent channel signals prior to adaptive cancellation as these signals will interfere in the LMS correlation process, generating a significant amount of tap noise. These filters are required in typical communications systems in any event in order to complete a raised cosine filter (the other root-raised cosine filter is used to

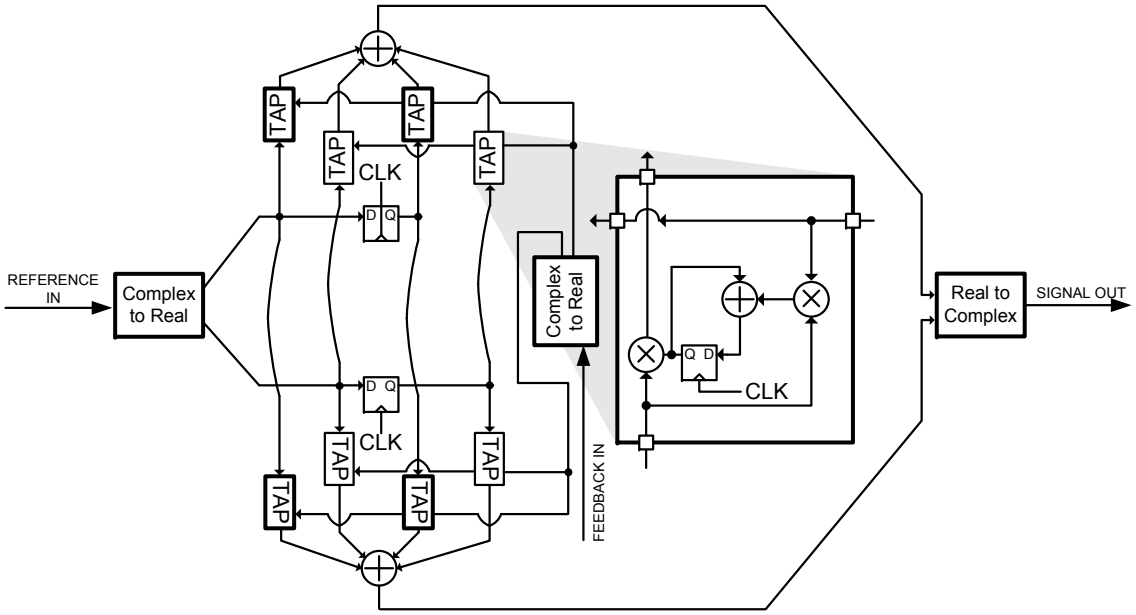


Figure 7.5: Hardware implementation of complex LMS equalizer (filter portion) modified to compensate for I/Q mismatch.

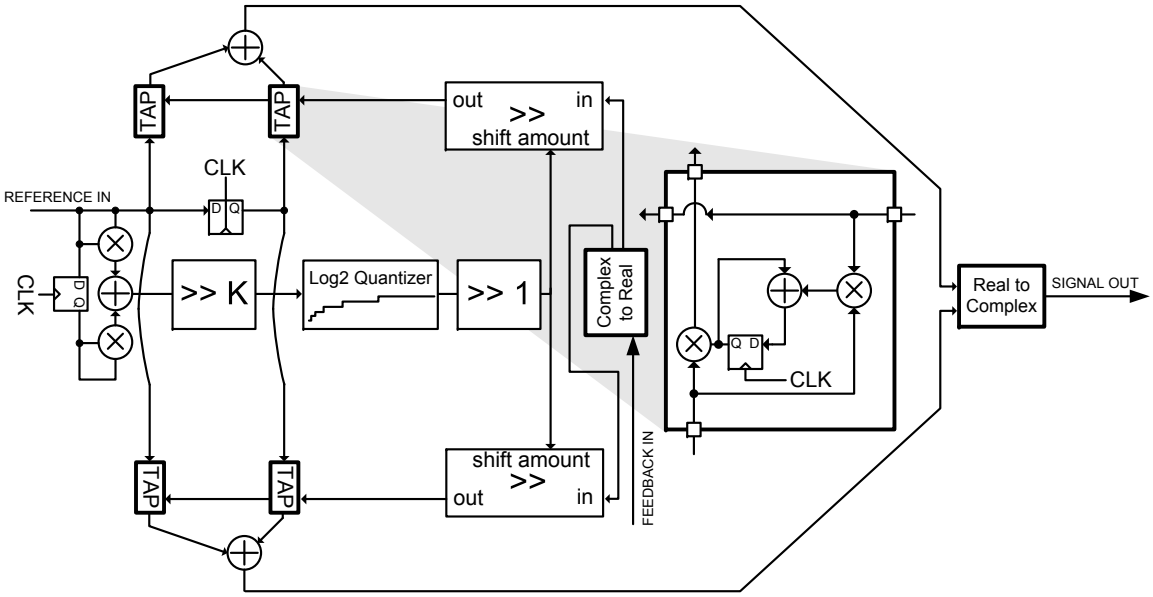


Figure 7.6: Hardware implementation of complex modified-NLMS equalizer (filter portion) modified to compensate for I/Q mismatch.

pulse-shape the output data of the transmitter) so that intersymbol interference is minimized.

The basis elements required by the LNTA from Section 5.8 are implemented as the 256-element lookup tables f(x) and g(x) in Fig. 7.4. Both f(x) and g(x) realize functions of the form $e^{-|k|x}$ where a squaring term is not needed due to the fact that the IM2 products have already experienced a

squaring. The performance of the cancellation scheme is not extremely sensitive to the value of k used - the value of k may change ± 20% without affecting input-referred cancellation more than 0.1dB in the measured worst-case blocking condition. A function $\tanh(k \cdot x)$ is utilized in the even-order distortion cancellation. It was found that adding this functional basis helps to reduce the effects of low-frequency corruptive signals on the odd nonlinear path outputs during application of a large signal to the receiver input. Over short time periods, this corruptive signal appears as dc on the odd nonlinear path baseband input. When this dc signal multiplies the outputs of the f(x) and g(x) blocks, it creates yet another spurious signal which is partially canceled by the $\tanh(k \cdot x)$ basis.

Finally, a half-band filter was added to the output of the algorithm to suppress glitching during the peak blocking condition. As the glitch impulses are composed mostly of high-frequency content, they are dramatically attenuated by the half-band filter. This glitching occurs because the IMD products in the linear path baseband occasionally rail the output of the OTAs, generating higher-order harmonics that are not well-modeled by the basis set shown in Fig. 7.4. Although in principle a basis element can be developed to explicitly model this source of error, the half-band filter is a less multiplier-intensive solution to the problem.

The complete digital back end circuitry used in the nonlinear path, including the adaptive equalizers, utilizes 39 and 342 16-bit multipliers running at 50MHz and 16.66MHz, respectively. Based on the results of [128] and assuming that the multipliers dominate the power consumption, the extra digital circuitry and adaptive equalizers would consume about 12mA and 41.5mA under a 1.3V supply for the even and odd nonlinear paths, respectively. In practice, these quantities would be dramatically reduced by time-averaging, as correction is only required under infrequent blocking conditions.

7.2 90nm Die Circuit Blocks

7.2.1 Linear Receiver Path

7.2.1.1 $\frac{1}{4}$-Phase Passive Mixer

Immediately following the LNTA described in Chapter 5, the passive mixer shown in Fig. 7.7 modulates the input signal current with the 1/4-phase LO waveforms, downconverting it to baseband. As specified by the literature [51] [103], the passive mixer switches are biased such that the dc gate-source voltage is only slightly larger than the threshold voltage of the device to prevent the flow of dc current. This is accomplished by the servo biasing circuit shown in the left half of Fig. 7.7 which forces a small leakage current through a dummy device and then forces the source voltage to equal the mixer output common-mode voltage. A similar biasing scheme was reported in [103] that used a fixed bias voltage rather than the passive mixer baseband output common-mode voltage.

The passive mixer gates are relatively large at 120μm by 80nm, resulting in an on-resistance of about 3.3Ω at the typical process corner. For a rail-to-rail input signal that produces a peak current of 30mA, this results in a voltage drop of 100mV across the switch.

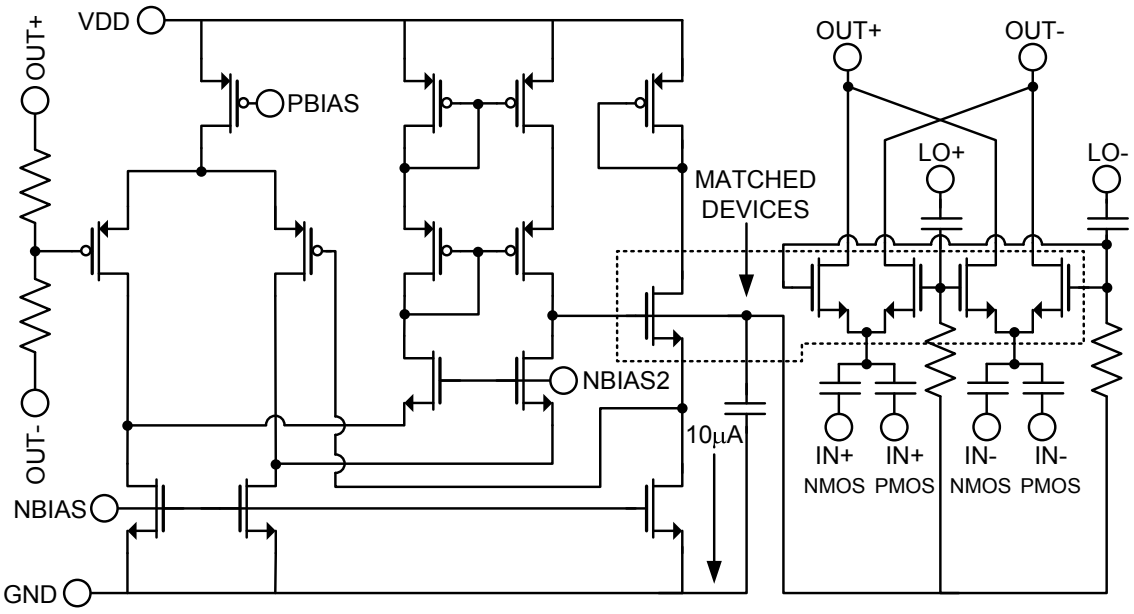


Figure 7.7: Schematic of passive mixer and servo biasing.

7.2.1.2 Common-Gate TIA Buffer

This common-mode feedback circuit is vital to the operation of the passive mixer, as its baseband output voltage is specified in an open-loop fashion by the input devices of the common-gate buffer. Shown in Fig. 7.8, the common-gate buffer is implemented by a g_m-doubling circuit which places a common-source PMOS device in parallel with the input common-gate devices while sharing the bias current, similar to the design of the input transconductor in [104]. Unlike the design in [104], this design is at baseband rather than at RF and is not as easily biased by ac-coupling techniques that separate the biasing (dc) and signal (ac) domains. Here, the output common-mode voltage is established with an OTA-based loop that uses split-pole compensation to obtain 37+dB of dc gain and 13+MHz of bandwidth over MOS and resistor corner. The PMOS devices are thick-oxide devices to take advantage of the higher threshold voltage offered. This permits a direct dc biasing of the PMOS devices that is compatible with maintaining all of the other devices in the saturation region over MOS and resistor corner.

The transconductance of the CG buffer devices is determined first by noting that each of the devices must be designed with a narrowly defined overdrive voltage $V_{OD} = V_{GS} - V_{TH}$ in order to maintain device operation in the saturation region over corner variation. In other words, the transconductance and sizing relationships between all of the devices are roughly fixed, and the only practical design variable is that of the quiescent current, which sets the absolute device sizing.

In order to choose the optimal g_m, the analytical setup in Fig. 7.9 is utilized in which the effective LNTA output impedance Z_S is frequency-translated to baseband. The RF input current at the LNTA output node can also be frequency-translated to baseband. At this point, the signal current sees an impedance division between the LNTA output impedance and the input impedance of the CG-buffer. With this setup, the output signal current power is given by (7.2), where the

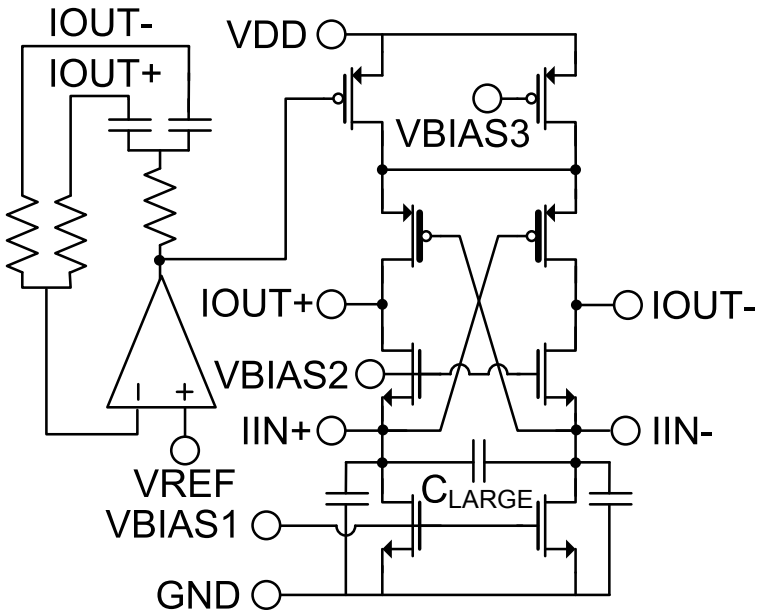


Figure 7.8: Simplified schematic of linear path common-gate buffer.

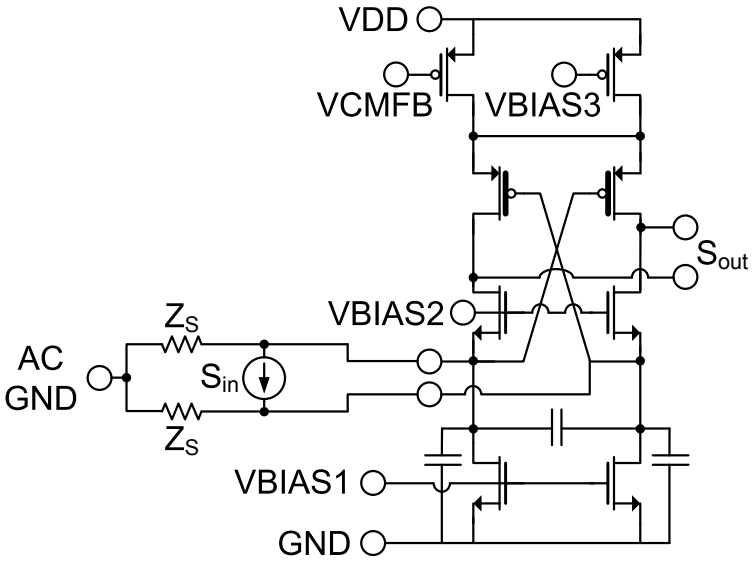


Figure 7.9: Analysis of optimal g_m of common-gate buffer.

transconductances are denoted as $g_{m,n}$ for the NMOS input device, $g_{m,nc}$ for the NMOS current source device, and $g_{m,p}$ for the PMOS common-source device.

$$S^2_{out,rms} = S^2_{in,rms}\left((g_{m,n}+g_{m,p})\cdot\left(\frac{Z_S}{1+g_{m,n}Z_S}\right)\right)^2 \tag{7.2}$$

The current thermal noise power density at the output of the CG-buffer is given by (7.3):

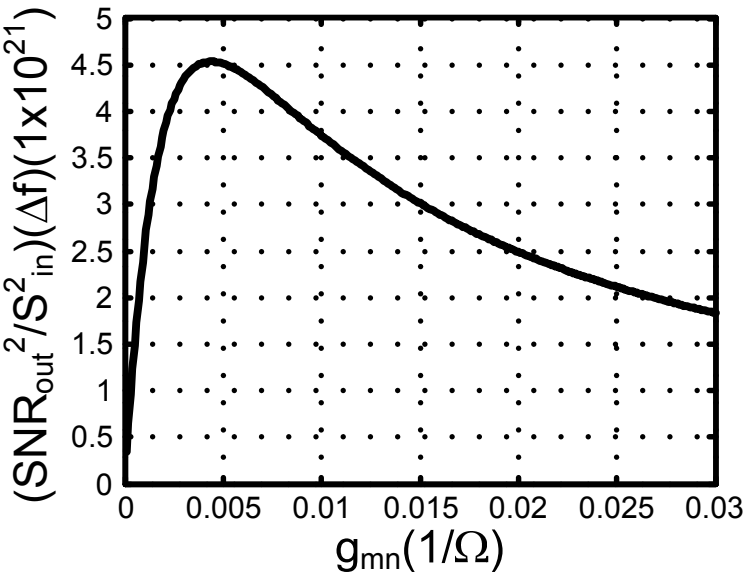


Figure 7.10: Inverse of input-referred thermal noise power density of CG buffer as function of input device transconductance.

Element	Value
R0	3kΩ
R1	6kΩ
R2	6kΩ
R3	60kΩ
R4	60kΩ
C0	40pF
C1	7pF
C2	2pF

Table 7.1: Passive values in linear path baseband filtering.

$$N_{out,rms}^2(\Delta f) = 4kTR\Delta f \left(g_{m,p} + g_{m,nc} \cdot \left(\frac{Z_S}{1 + g_{m,n} Z_S} (g_{m,n} + g_{m,p}) \right)^2 + g_{m,n} \left(\frac{1 + g_{m,p} Z_S}{1 + g_{m,n} Z_S} \right)^2 \right) \tag{7.3}$$

Given that Z_S is determined to be 181Ω in simulation, the inverse of the input-referred thermal noise power density is plotted in Fig. 7.10 versus $g_{m,n}$ when $g_{m,p} = 0.76 g_{m,n}$ and $g_{m,p} = 0.465 g_{m,nc}$. In this case, a soft optimum is obtained for $g_{m,n}$=4.5mS. For design purposes, $g_{m,n}$ was set to 6mS for better handling of the attenuated large out-of-band blocker that still makes it through the CG buffer while losing very little SNR from the optimum-g_m point.

Also included in Fig. 7.8 is the large differential input capacitor used to attenuate the majority of the large downconverted blocker current. At 335pF it is clearly visible in the chip die photograph. Along with the impedance $Z_S \parallel 1/g_{m,n}$ it sets a first-order pole at about 3MHz that by 100MHz attenuates the large signal out of band blocker current continuing on to the TIA by about 30dB.

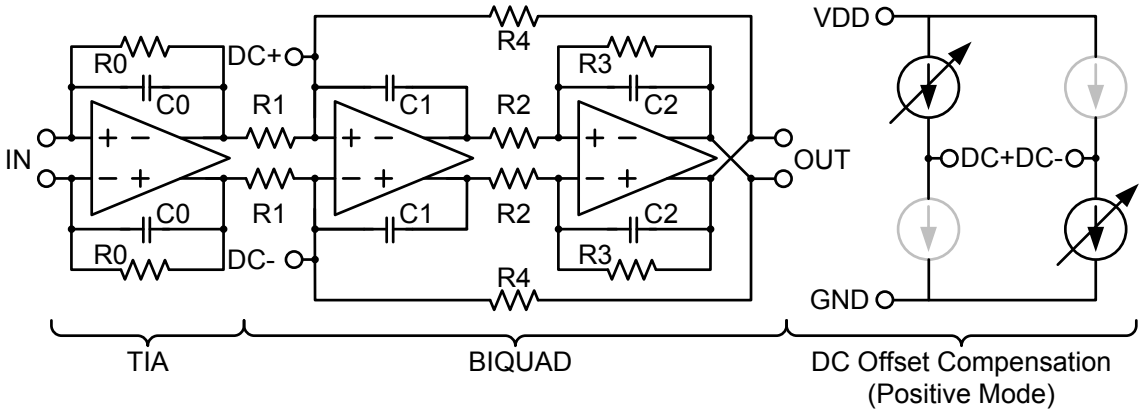


Figure 7.11: Block diagram of TIA and biquad with dc offset compensation.

7.2.1.3 TIA and Baseband Filter

Using a standard TIA/Biquad architecture, shown in Fig. 7.11, the large out-of-band signals are filtered to an even greater extent prior to sampling by the ADC. The resistor and capacitor sizes, listed in Table 7.1, are chosen to trade-off in-band noise and distortion. The somewhat large resistance values do slightly impact the noise figure of the entire receiver and might be reduced in a future implementation. Dc offset compensation is provided by a set of programmable current sources, shown in the right hand side of Fig. 7.11, tied to the virtual ground nodes of the first OTA in the biquad. For positive offsets, the pair of highlighted current sources is enabled. For negative offsets, the highlighted current sources are disabled and the greyed-out sources are enabled.

7.2.1.4 Baseband OTAs

Standard two-stage Miller-compensated OTAs, shown in Fig. 7.12, provide the loop gain necessary for accurate realization of the biquad filter transfer functions. In order to save design time, all of the OTAs throughout the 90nm die are essentially the same and draw 0.75mA of quiescent current, except for that of the linear path TIA, which is a doubled version of the others and draws 1.5mA.

7.2.2 Nonlinear Receiver Path

7.2.2.1 Odd IMD Nonlinear Reference and Common-Gate Buffers

Like the receiver described in Chapter 3, the odd IMD nonlinear reference is comprised of a multi-stage cubic term generator such as the one described in Chapter 4. In order to improve the dynamic range of the cubic term generator, stacked NMOS/PMOS squaring circuits are used to generate the initial second-order IM2 terms as shown in Fig. 7.13. Like the circuit described in Chapter 4, half of the squaring circuits actively generate IM2 products, while the other half generate currents only in response to common-mode inputs. More specifically, the lower left NMOS and upper right PMOS squaring circuits generate IM2 products, while the upper left PMOS and lower right NMOS with shorted inputs generate signals only in response to common-mode inputs.

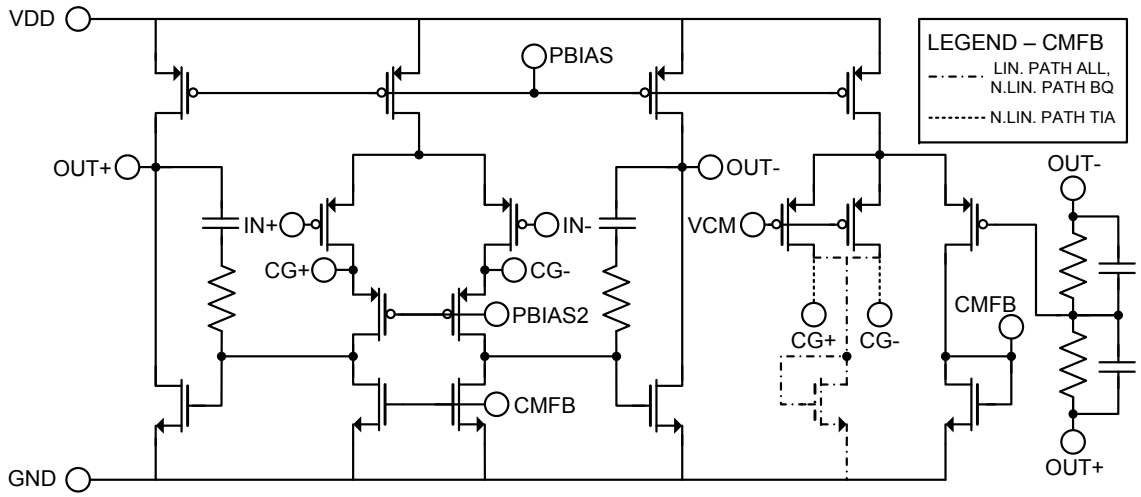


Figure 7.12: Simplified schematic of two-stage OTA used in baseband circuitry.

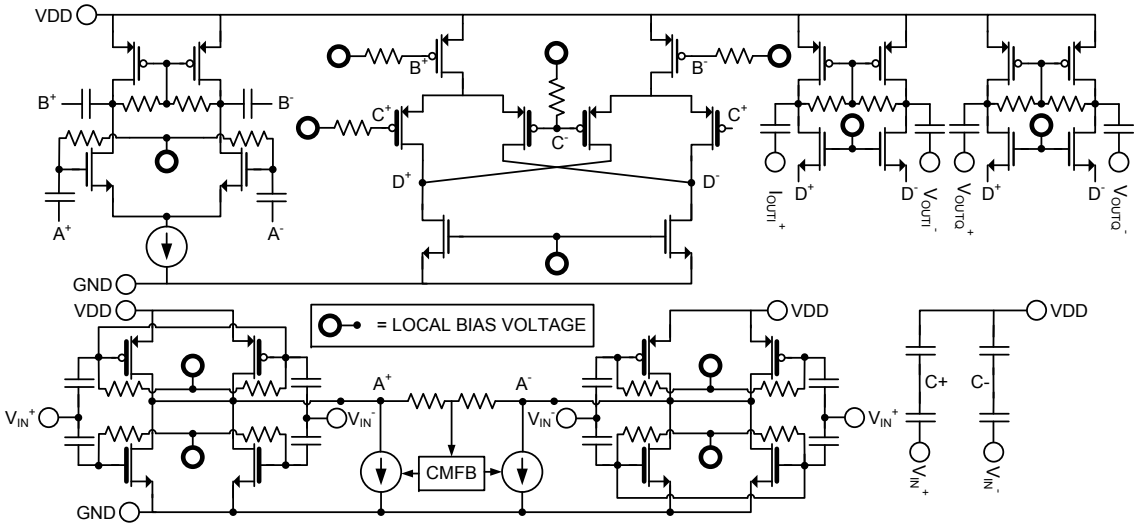


Figure 7.13: Simplified schematic of odd nonlinear term generator.

In order to more accurately achieve a square-law I-V characteristic, the input squaring circuits are comprised of thick oxide devices with channel lengths of 1μm. Even having done this does not permit accurate square-law behavior when a large rail-to-rail signal is applied to the input of the circuit. For this reason, the input RF signal is capacitively divided down between the ac coupling capacitance and the large squaring device input capacitance.

Chapter 4 argues that one advantage of a multistage cubic term generator with a bandpass interstage characteristic is that IM2 terms directly surrounding dc are removed so that they do not saturate the interstage circuitry while contributing nothing to the usable output IMD products. This advantage is nowhere greater than in the case of an asymmetric blocking condition dominated by a large signal. In this case, the dc signal content is greater than the beat-frequency IM2 content by

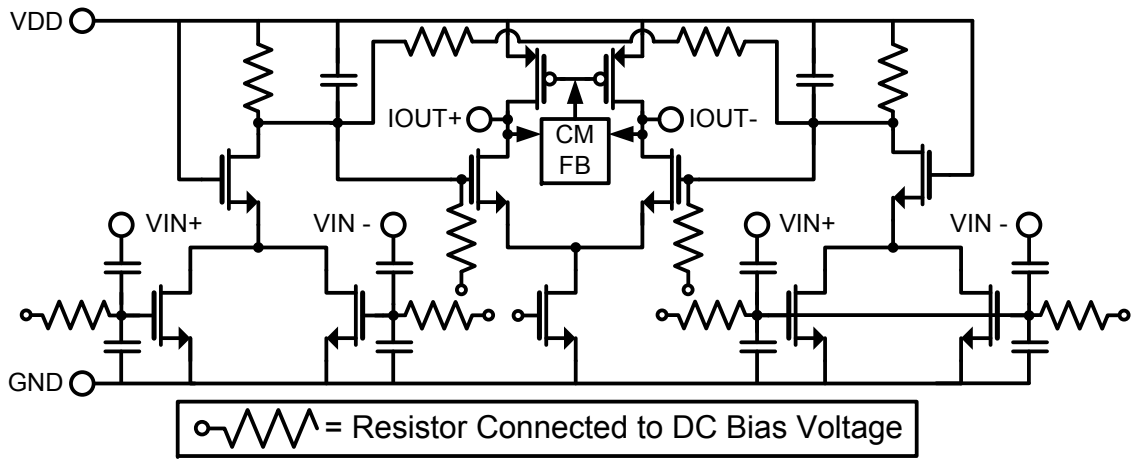


Figure 7.14: Simplified schematic of even nonlinear term generator.

an amount equal to the difference in power between the large blocker and smaller blocker and may approach 20-30dB depending on the precise application. Removing the dc IM2 content also relaxes requirements on the nonlinear path baseband postfiltering. Were an explicit cubing used to generate the odd-order IMD (principally IM3) products, the resultant output at the large out-of-band blocker frequency would be much larger than the odd-order IMD products at the LO frequency due to the fact that it consists primarily of a cubed version of the original large signal. Without the dc IM2 content, however, the output at the large signal frequency is a much smaller odd-order IMD term proportional to the square of the small blocker multiplied by the large blocker.

In principle, it is desired that IMD products be canceled to the thermal noise floor limit of the linear receiver path. In this case, it would be desirable to cancel IMD products such that the input-referred distortion error power is reduced by 40dB or more. Given this, it is prudent to ask whether the interstage group delay constitutes a limitation to achievable cancellation. Returning to Fig. 4.13 it is seen that 50dB IER is achievable when the interstage group delay is 1/20th of the chip time. As a 1ns interstage group delay is easily achievable, the results of Section 4.7 imply that the multistage cubic term generator is suitable for generating reference IM3 products for modulated blockers with sampling rates up to 50MHz.

In order to isolate the I and Q passive mixers from one another, two CG buffers are utilized at the output of the odd IMD nonlinear reference, resulting in an even split of the output reference current.

7.2.2.2 Even IMD Nonlinear Reference and Buffer

The even-order IMD nonlinear reference shown in Fig. 7.14 utilizes the same input stage as the cubic term generator from Chapter 4. Because the desired IM2 products in this case exist at baseband, a direct dc connection is made between the squaring transconductor outputs and the subsequent buffering stage. Like the odd-order IMD reference circuit from the previous paragraph, the input is capacitively divided down so as not to overwhelm the input devices.

Element	Value
R0	3kΩ
R1	6kΩ
R2	12kΩ
R3	6kΩ
R4	24kΩ
C0	13.2pF
C1	6.6pF
C2	1.2pF

Table 7.2: Passive values in nonlinear path baseband filtering.

7.2.2.3 Passive Mixers and Baseband Circuitry

The odd nonlinear path utilizes the passive mixer architecture shown in Fig. 7.7. However, the mixer here utilizes much smaller switches and takes as an input a 50%-duty cycle LO waveform from the nonlinear path frequency divide-by-2 circuit. Both nonlinear paths employ the TIA/Biquad architecture shown in Fig. 7.11 with the addition of a 5pF differential capacitor placed across the TIA OTA input terminals to suppress upconverted high-frequency signal. In order to avoid any distortion at baseband due to the saturation of the output OTAs, the gain of the biquad is designed to be less than that of the linear path (6dB vs. 20dB). The OTAs used are the same as those in the linear path, with the exception that the common-mode feedback is modified as in [129, pg. 146] with the connections made using the dashed lines in Fig. 7.12. Because the OTA inputs see an infinite impedance at dc, without the dashed-line connections, the OTA common-mode voltage has two stable solutions, one of which does not permit rail-to-rail output operation.

Coarsely matched odd and even baseband filters are required for the successive regeneration of nonlinear terms at digital baseband so that the odd-order and even-order reference IMD terms experience the same group delay prior to multiplication. Mismatch between the group delays does not constitute a limitation for the proposed successive regeneration scheme, however, as different IMD reference basis elements may be created with staggered digital delays between the odd and even references prior to multiplication. In this case, the adaptive algorithm will effectively perform a system identification on the true relationship between the analog group delays and in the process will strongly cancel the higher-order IMD terms.

7.2.3 Frequency Generation

7.2.3.1 QVCO

Out-of-band VCO noise is typically dictated by the thermal noise of the oscillator active devices and the filtering attenuation provided by the L-C tank. For this reason, the Q of the L-C tank must be maximized with generous metal lines comprising the inductor and all interconnect. Given the aggressive requirements of the large-signal handling receiver, it is desired to reduce the out-of-band noise even further. One way to accomplish this is by using a ring oscillator with capacitively coupled dual L-C tanks [130] whereby the dual nature of the tank increases its effective Q. As depicted in Fig. 7.15, capacitive coupling between the L-C tanks also results in a 90° phase shift as the voltage from one tank induces a current through the capacitor to then form a voltage on the next tank. Utilizing two of these 90° phase shifts in conjunction with active devices and a 180° transition implemented by a differential signal swap results in a structure that satisfies the Barkhausen criterion at some

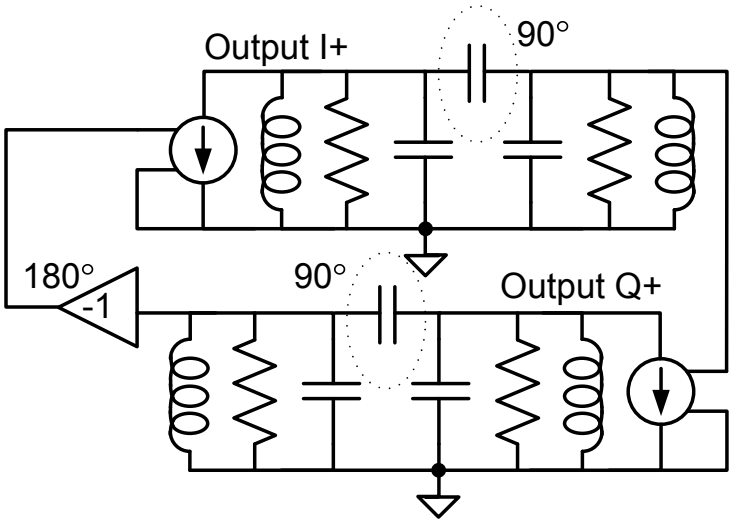


Figure 7.15: Depiction of dual-Q ring VCO concept.

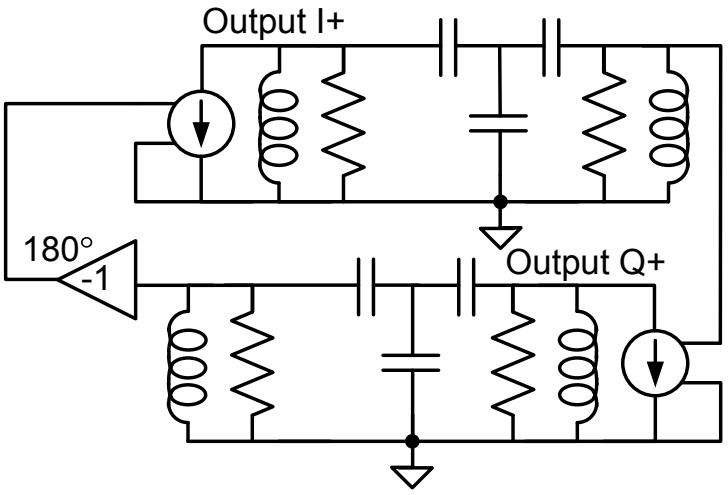


Figure 7.16: Depiction of dual-Q ring VCO after Π-Y transformation.

frequency dictated by the values of the passive elements in the structure. As a side benefit, the two halves of the circuit output signals 90° out of phase with each other, creating a QVCO.

As also determined by the work in [130] and depicted in Fig. 7.16, the capacitive coupling network in the RLC tanks can be manipulated using a Π-Y transformation. According to [130] this results in lower loss in the network and has the added advantage that when implementing the capacitors as tunable PIN diodes, the control voltage at the center node of the Y-network is easily accessible at low frequencies via a resistive connection.

These concepts were implemented in a working QVCO, whose schematic is shown in Fig. 7.17 and that is modified from the one described in [130]. PMOS devices were used for lower phase noise [131] while utilizing the thick-oxide option permitted large signal swings to increase the carrier-

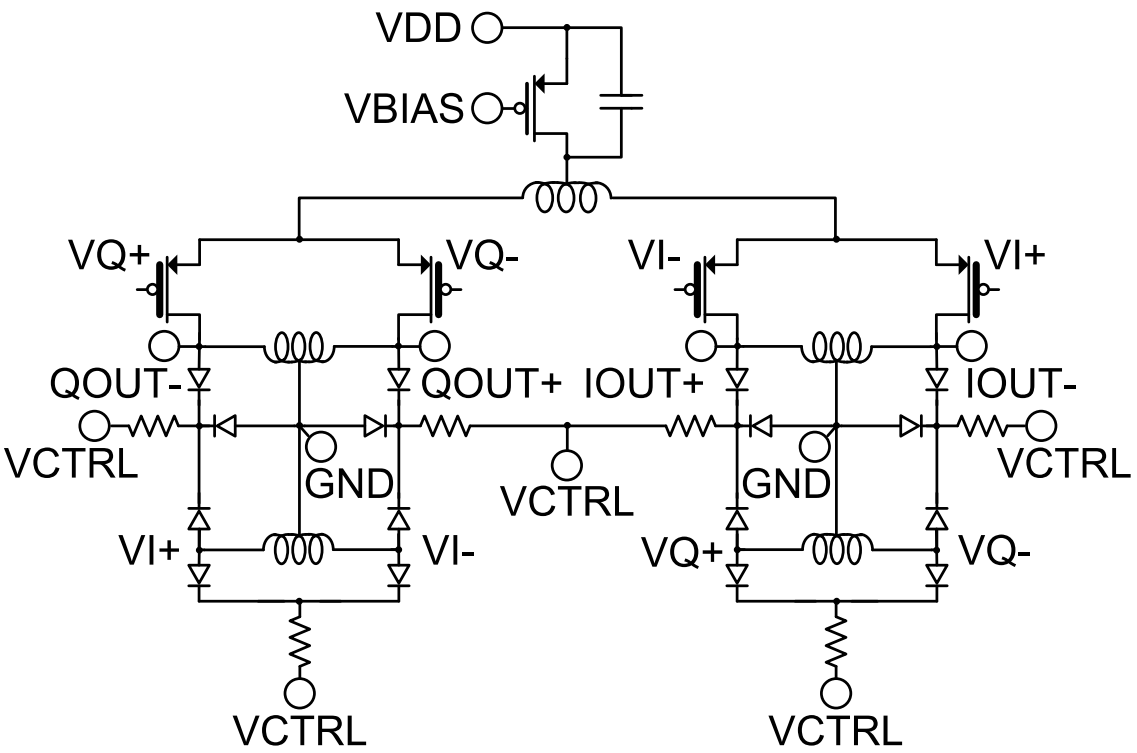


Figure 7.17: Simplified schematic of implemented QVCO.

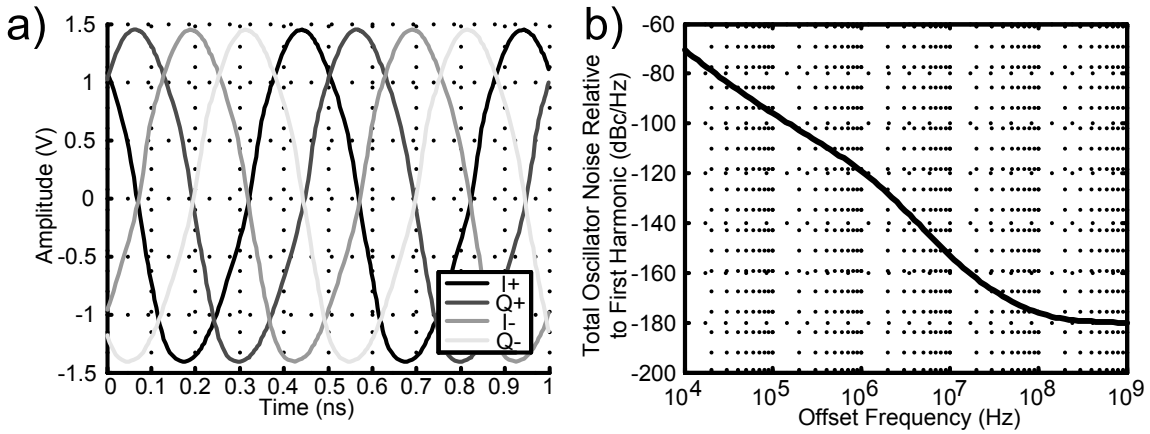


Figure 7.18: Frequency generation circuitry simulations. a) Time-domain outputs of QVCO. b) Simulated total noise of the oscillator and 1/4 phase logic circuitry.

to-noise ratio. Adding a differential inductor from the tail current source to the active devices of the oscillator reinforces the quadrature relationship between the two halves of the QVCO [132], increases the output signal swing (and hence carrier-to-noise ratio) [132] and reduces the phase-noise contribution of the oscillator bias circuitry [133]. The simulated output waveform of the oscillator with extracted capacitances and when loaded with the $\frac{1}{4}$-phase logic is shown in Fig. 7.18a. These

waveforms are subsequently attenuated by the ac-coupling network at the input of the 1/4-phase logic such that they do not break down the oxide of the CMOS NAND device gates.

7.2.3.2 $\frac{1}{4}$-Phase Logic and Multiply-by-2

Directly supplying waveforms to the passive mixer gates is performed by the 1/4-phase generation logic circuitry shown in Fig. 7.19. With the outputs of the QVCO, CMOS NAND gates generate a set of nonoverlapping 1/4-phase pulses which are then buffered by CMOS inverters that drive the passive mixer gates. The common-mode voltages of the NAND gate inputs are set via a reference buffer and ac-biasing. Because the relationship between the four input waveforms is known and periodic, only one of the two PMOS transistors typically present in a NAND gate is required. Since the PMOS device is much larger than the NMOS gate, removing one results in a reduction in dynamic power dissipation due to the decrease in capacitive load. The simulated total noise of the oscillator and the $\frac{1}{4}$-phase logic from a PNOISE simulation is plotted in Fig. 7.18b. Note that due to the equipartition theorem of statistical mechanics, half of the noise power is amplitude noise power and half is phase noise power. Hence, the relevant noise metric is 3dB less than shown in Fig. 7.18b. At a 100MHz offset frequency, the noise is dominated by various components of the QVCO, while the 1/4-phase logic contributes just 23% of the noise power.

Sending the LO phase information to the nonlinear path may entail the use of a substantial routing length, which increases the likelihood that the LO signal will couple to a sensitive node in the circuit. Because of this, the LO frequency is multiplied by 2 before routing it out to the nonlinear path so that any coupling to one of the signal paths is relatively benign. Performing this multiplication is facilitated by the 1/4-phase logic in that when the I and Q 1/4-phase outputs of the NAND gates are again applied to NAND gates, the output signal period is half that of the input signals. Applying this circuitry to both I and Q 1/4 phase outputs generates a differential signal that incorporates information from all of the signals controlling the linear path mixer. Because this 2xLO signal most closely reflects the totality of the waveforms actually supplying the mixer, it is used as the input to the frequency divider chain and the rest of the PLL.

7.2.3.3 Divide-by-2

As shown in Fig. 7.20, the divide-by-2 circuit used both in the nonlinear path and in the frequency divider chain is comprised of two current-mode logic (CML) D-flip-flop (DFF) latches buffered by actively loaded Cherry-Hooper amplifiers, both shown in Fig. 7.21.

7.2.3.4 Frequency Divider Chain

The frequency divider chain and its connections within the PLL are shown in Fig. 7.3. It consists of six divide-by-2 circuits cascaded for a net division of 64. As each divide-by-2 circuit has one differential input and two differential outputs, the I output of each frequency divider is used to connect to the next one while the Q output is kept floating. As the frequency divider chain was implemented as part of a last-minute notification to tape out the chips, its termination circuitry relied on pre-laid out blocks capable of driving the off-chip capacitive load. In this case, CMOS inverters were used to buffer the divide-by-2 output to drive an on-chip PMOS-based common-source amplifier. The output of this amplifier is connected to an off-chip discrete Schmitt-trigger

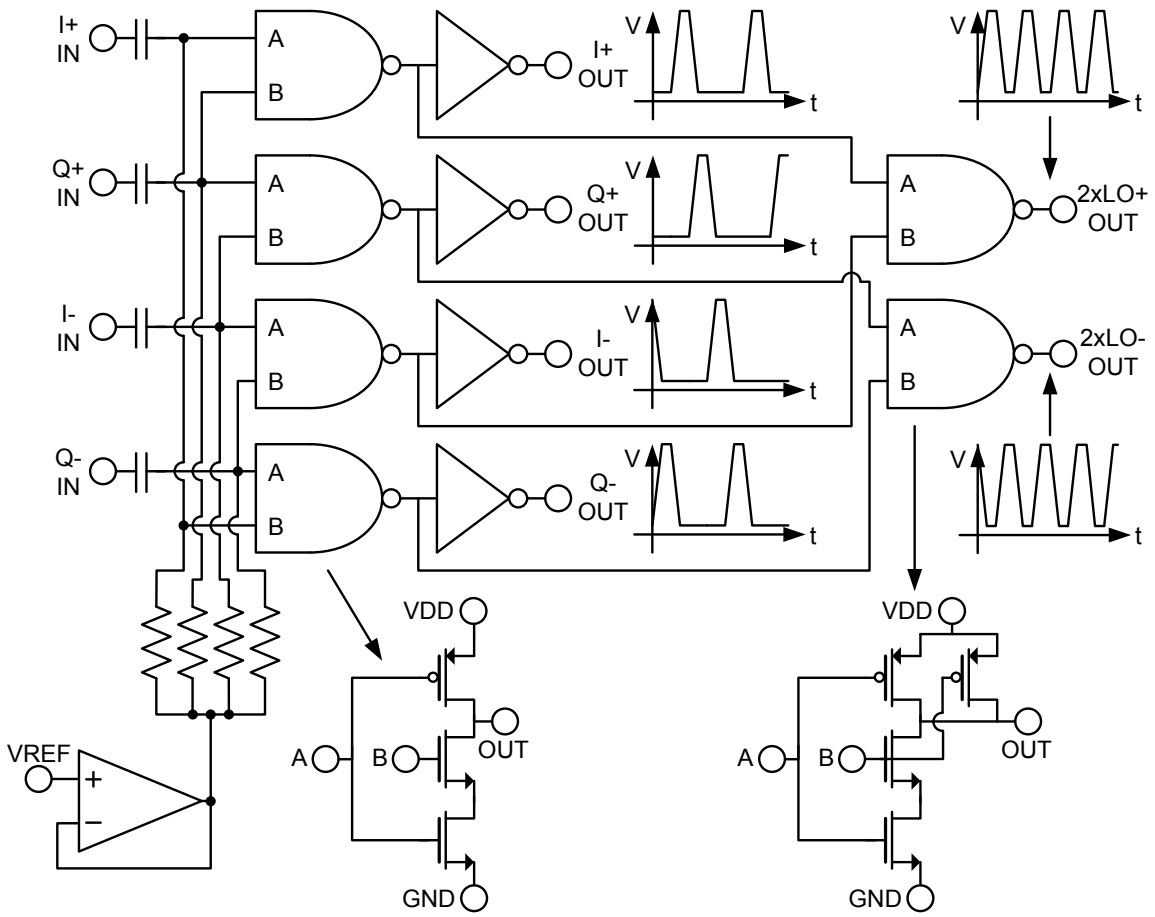


Figure 7.19: Simplified schematic of implemented 1/4-phase logic circuitry.

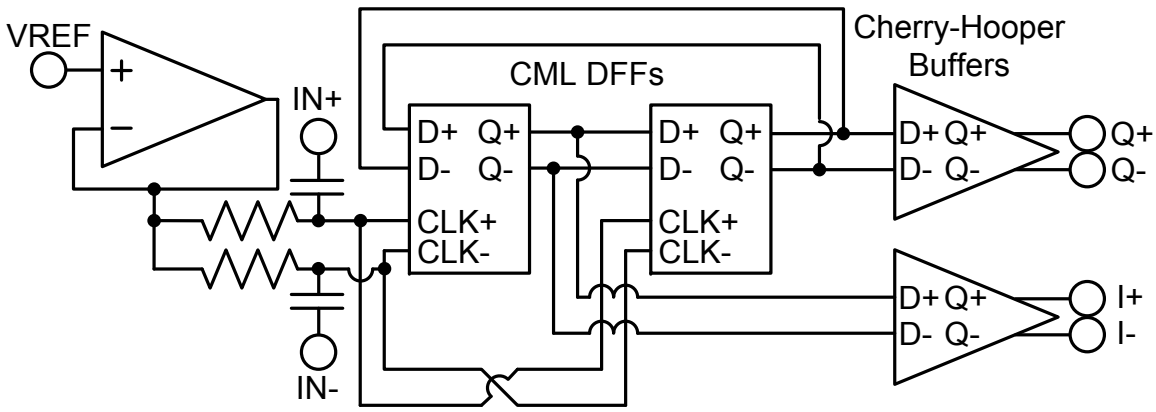


Figure 7.20: Block diagram of CML divide-by-2 circuit.

which provides buffering prior to driving a long trace on the printed circuit board.

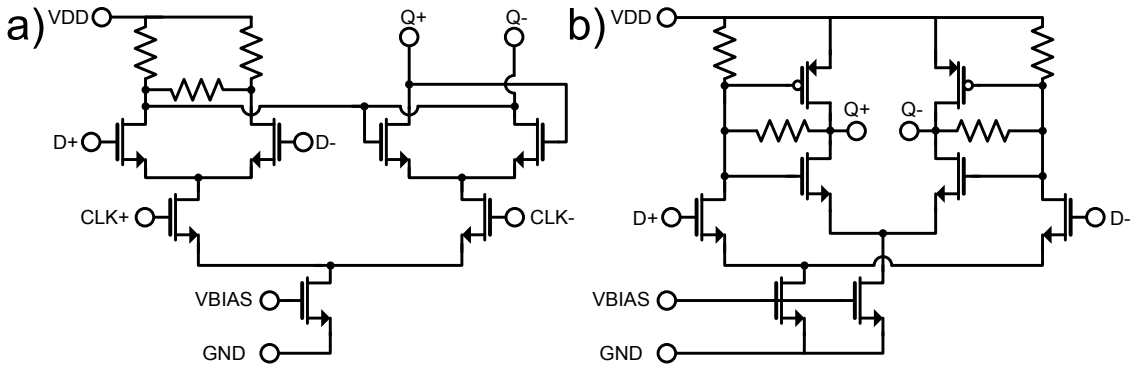


Figure 7.21: Simplified schematics of a) CML D-flip-flops and b) actively-loaded Cherry-Hooper buffers.

7.3 Experimental Setup

The complete assembled receiver up to the connection to the FPGA board is shown in Fig. 7.22. A close-up of the gold-plated Rogers board is shown in Fig. 7.23, while the die photo of the chip is shown in Fig. 7.24. Included on the Rogers board are an external 1:1 balun and ac coupling capacitors to the input of the LNTA. Although a balun is not strictly required in the event that a differential antenna is utilized with the receiver, it is required to interface the differential 50Ω LNTA with the single-ended 50Ω test equipment. Calibration of the input sources is performed up to the leftmost (RF input) SMA connector of the receiver. The combined losses of the SMA connector, balun, routing, and ac coupling capacitors are taken into account by shorting the exposed terminals of the ac coupling capacitors, which sit 180 μm apart, and by measuring the S_{11} scattering parameter. The result of this experiment in dB divided by two yields the combined loss of the input network up to the LNTA. As the final connection to the LNTA occurs in free space, the wavelength of a 1.9GHz input signal at this point is 160mm. Since the shorted connection length is three orders of magnitude less than the signal wavelength, this shorted connection is a very good approximation to an ideal short, validating this measurement procedure. The measured losses of this input network are then de-embedded from the raw measurement data of the receiver itself, setting the calibration plane at the chip input.

A picture of the experimental setup is shown in Fig. 7.25 with the setup schematic shown in Fig. 7.26. The large modulated blocker is supplied by an E8267 vector signal generator followed by three Panasonic EFCH1842TCD1 SAW filters in series in order to attenuate the signal generator phase noise at the RX LO frequency. In order to overcome subsequent losses, the large modulated signal is amplified by a Mini-Circuits ZHL-4240 power amplifier and then followed by three additional EFCH1842TCD1 SAW filters in order to remove more noise from the E8267 and the power amplifier. An E8257C signal generator is used to model a CW desired signal, while an 83620B signal generator was used to generate the smaller of the two blockers. These two signals are combined using a Mini-Circuits ZFSC-2-2500 signal combiner/splitter. The resultant output is then combined with the output of the final large-signal SAW filters using a Mini-Circuits ZAPDQ-2 power splitter/combiner. According to the ZAPDQ-2 data sheet, its typical port isolation is 25dB at 1.8GHz, helping to protect the two smaller-signal generators from the power amplifier output. The output of the ZAPDQ-2

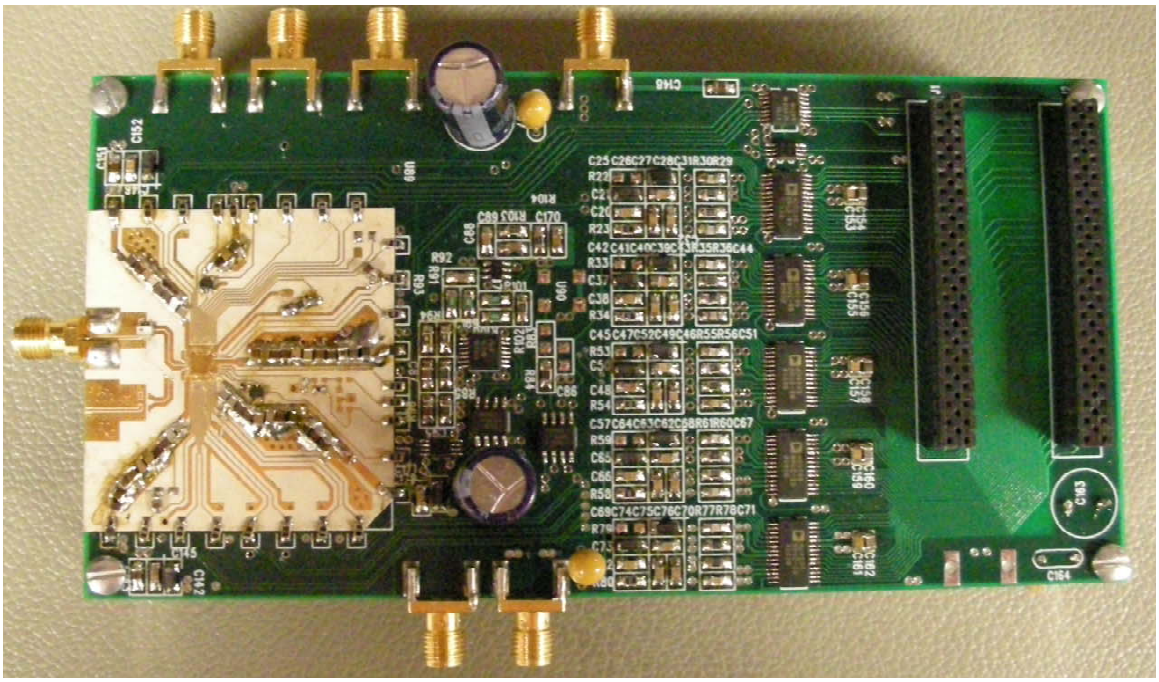


Figure 7.22: Photograph of assembled RF/Analog front end.

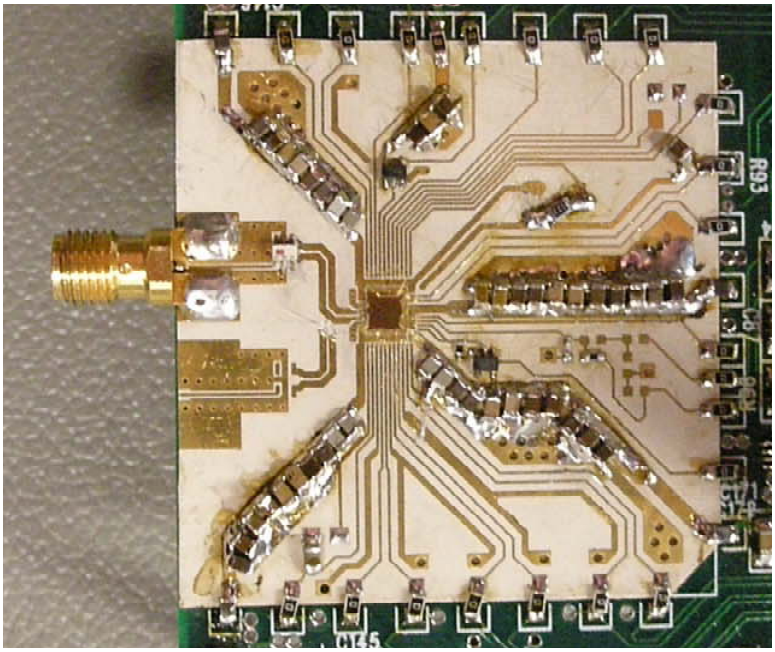


Figure 7.23: Photograph of mounted RF/Analog 90nm die with RF and local baseband connections.

is coupled to a HP 8563E spectrum analyzer via a Krytar model 1851 directional coupler. The spectrum analyzer is used to verify that the proper signal characteristics are applied to the receiver input at any given time. The directional coupler is connected to the receiver RF input via a Mini-

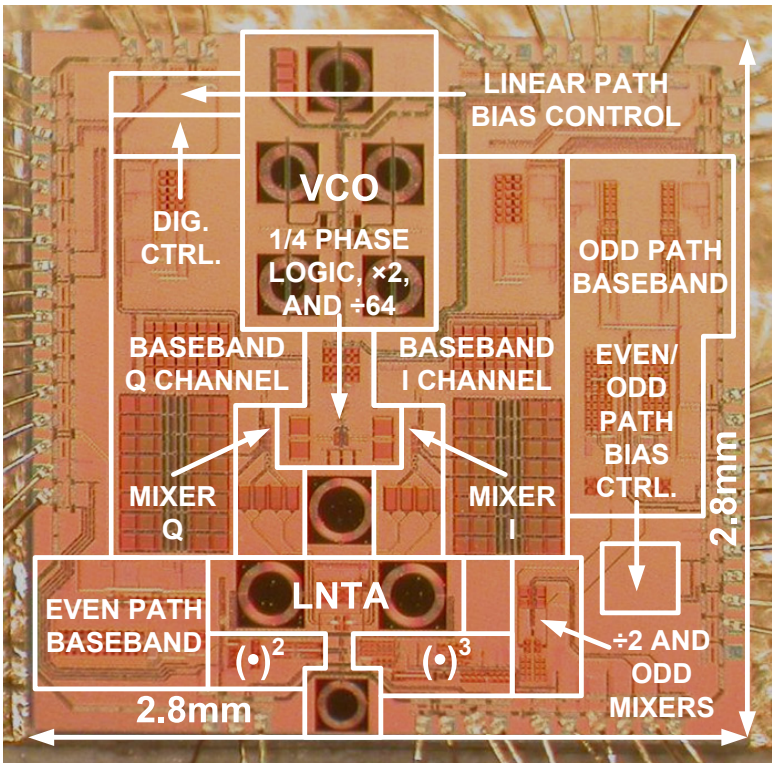


Figure 7.24: RF/Analog chip die photo.

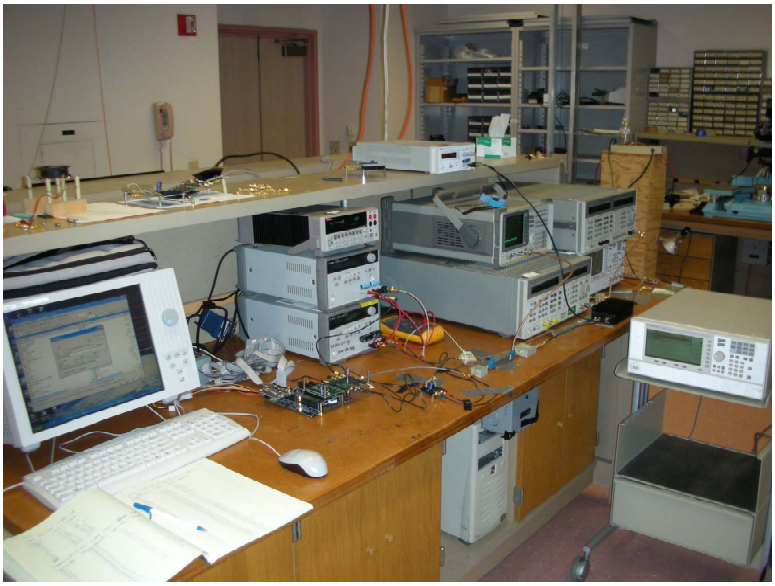


Figure 7.25: Photograph of experimental setup.

circuits ZFBT-6GW bias tee. Calibration of the input source network is performed by recording key signal power levels at the output of the cable connecting to the assembled receiver with a HP 8487A power sensor and an E4418B Agilent power meter. The receiver noise figure test is performed using

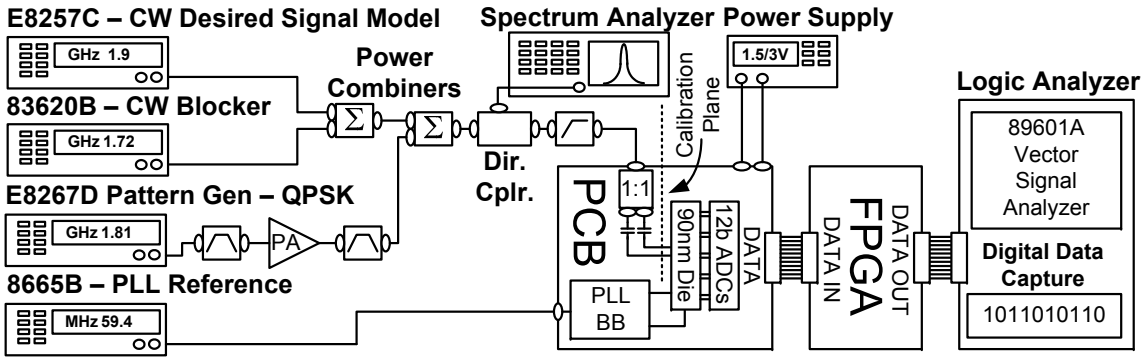


Figure 7.26: Schematic of experimental setup.

the Y-factor method with an HP 346C noise source connected to the input of the aforementioned bias tee. In this latter test, the losses of the bias tee and cable to the receiver are measured and calibrated out of the receiver noise figure and gain measurements.

The PLL reference is applied using an HP 8665B sinusoidal signal generator to ensure good spectral purity of the reference signal in series with a Mini-Circuits ZFBT-6GW bias tee externally biased to 1.5V (half the 3V power supply of the PCB-based portion of the PLL).

The digital ADC outputs interface to the Terasic DE3 FPGA board via a GPIO-HSMC card. The FPGA board provides the ADC digital power supply, supplies the ADC clock signals, and acquires and multiplexes the different ADC channels to the Agilent 16901 logic analyzer, which performs data acquisition and spectral measurements using an 89601A digital vector signal analyzer/spectrum analyzer. Baseline measurements of the RF/Analog portion of the receiver were made using the 89601A digital spectrum analyzer. Measurements involving large modulated blockers were made by acquiring data streams from the FPGA which were then exported to the software digital back end model for IMD cancellation.

7.4 Experimental Results

7.4.1 Baseline Receiver Tests

The non-monotonic nature of the LNTA nonlinearity is apparent in the results of the two-tone measurement shown in Fig. 7.27. In this case, a large CW blocker at 92MHz LO frequency offset is swept with a smaller blocker at 185MHz offset while the LO is set to run at 1.9GHz. As predicted in Chapter 5, the magnitude of the IMD products remains roughly constant over the 0 to +12.5dBm range. Also like the results presented in Chapter 5, the large-signal IMD products are about 40dB less than what would be predicted if a slope-of-3 line were extended from the smaller-signal IMD products. By extrapolating at each point of the input-referred error curve of Fig. 7.27, an effective IIP3 metric can also be obtained, reaching a peak of +33.5dBm for a rail-to-rail blocker at the input. At the same time, the receiver small-signal gain only drops by slightly less than 1dB for a +12.5dBm CW blocker input. This out-of-band 1-dB desensitization point result is worse than that predicted for an ideally loaded LNTA in Chapter 6 but is corroborated by simulations of the LNTA loaded by both passive mixers. The performance summary of the baseline receiver is shown in Fig.

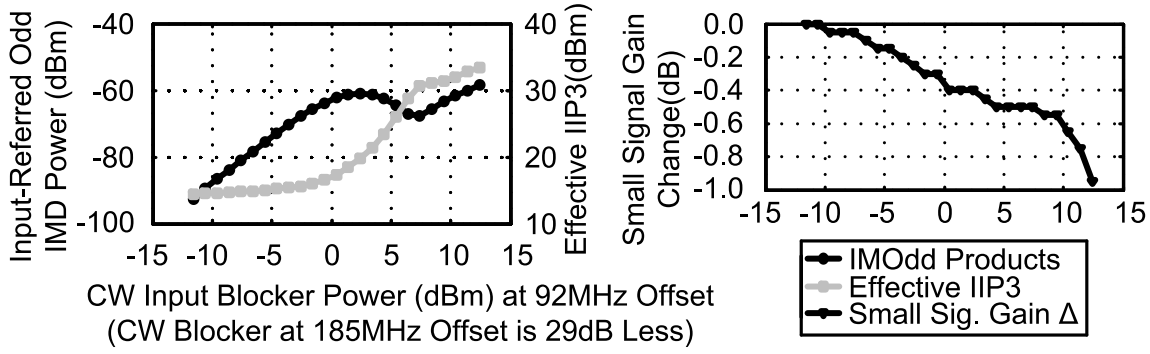


Figure 7.27: Two-CW tone IIP3 test: measured input-referred error, effective IIP3, and out-of-band desensitization of standalone RF/Analog die.

Measurement at f_{LO}=1.9GHz	Result
RF/Analog Die Area	(2.8mm)^2
RF/Analog Die External Supply	1.5V
RF/Analog Die Process	90nm CMOS
Receiver Linear Path Voltage Gain	50.3dB
Sim. DC Gain of Lin. Path Biquad	20.0dB
Receiver Linear Path Noise Figure	10.7dB
Peak Effective Two-Tone IIP3 (Uncorrected) @1.81GHz/1.72GHz	+33.5dBm
Two-Tone IIP2 (Uncorrected) @ 1.81GHz	+64dBm
Return Loss (S11) 1.6GHz-2.0GHz	<-16dB
1-dB Desensitization @1.81GHz	+12.5dBm
Linear Path Quiescent Current	14mA
On-chip LO Generation Current	46.2mA
Even/Odd Path Quiescent Current	3.5/14.3mA
Total RF/Ana. Die Quies. Current	84.8mA
Baseband Signal Meas. Bandwidth	0.01-1.92MHz

Figure 7.28: Measured baseline receiver performance metrics.

7.28. Notable measured results include the 50.3dB linear path voltage gain, 10.7dB noise figure and +64dBm two-tone IIP2.

7.4.2 Modulated Blocker Receiver Tests

Fig. 3.23a shows the concept behind the modified two-tone test used to evaluate the proposed receiver with IMD cancellation scheme. Using this methodology, the linearity performance of the receiver was also tested under modulated blocking conditions by applying a large QPSK signal at a modulation rate of 2MSPS along with a smaller CW blocker at LO offsets of 92MHz and 184MHz, respectively. The use of a CW non-dominant blocker in this case is justified by the fact that the use of CW blockers is commonly used in specification tests for communication standards such as UMTS. Furthermore, the error due to the successive regeneration approximation for a modulated nondominant blocker for the blocker powers considered would be below the receiver thermal noise

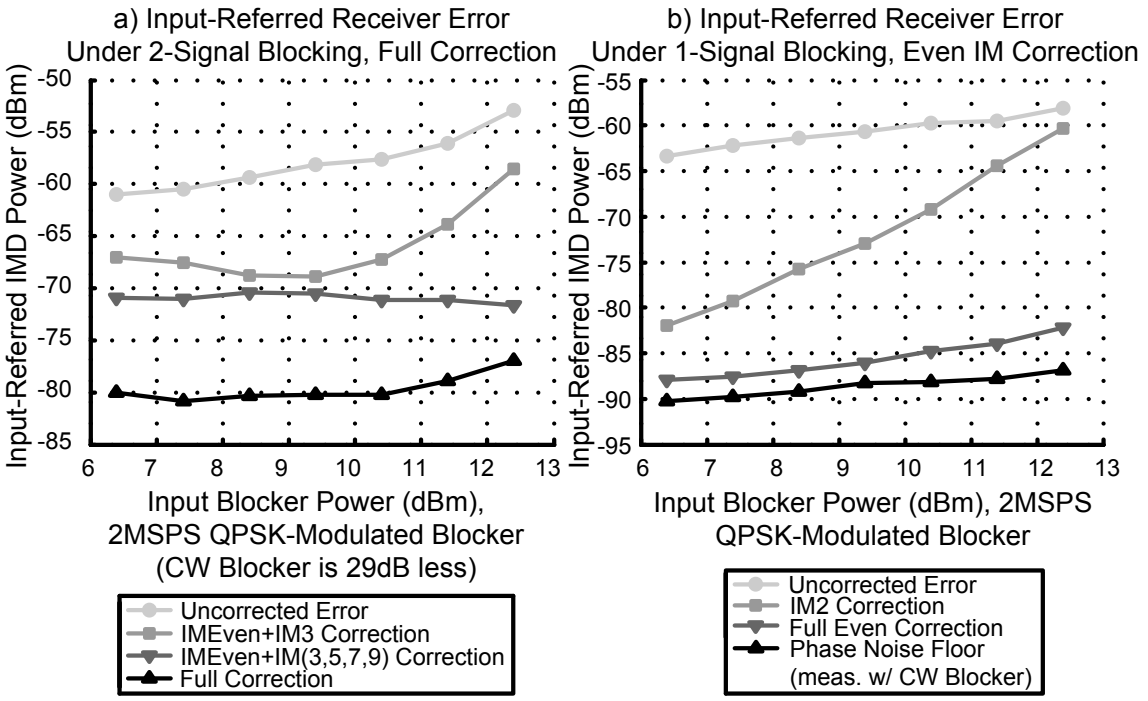


Figure 7.29: Measured input-referred error with various degrees of cancellation. a) Even- and odd-order cancellation for two-signal blocking. b) Even order cancellation for large QPSK-modulated blocker and phase noise floor (measured with CW blocker).

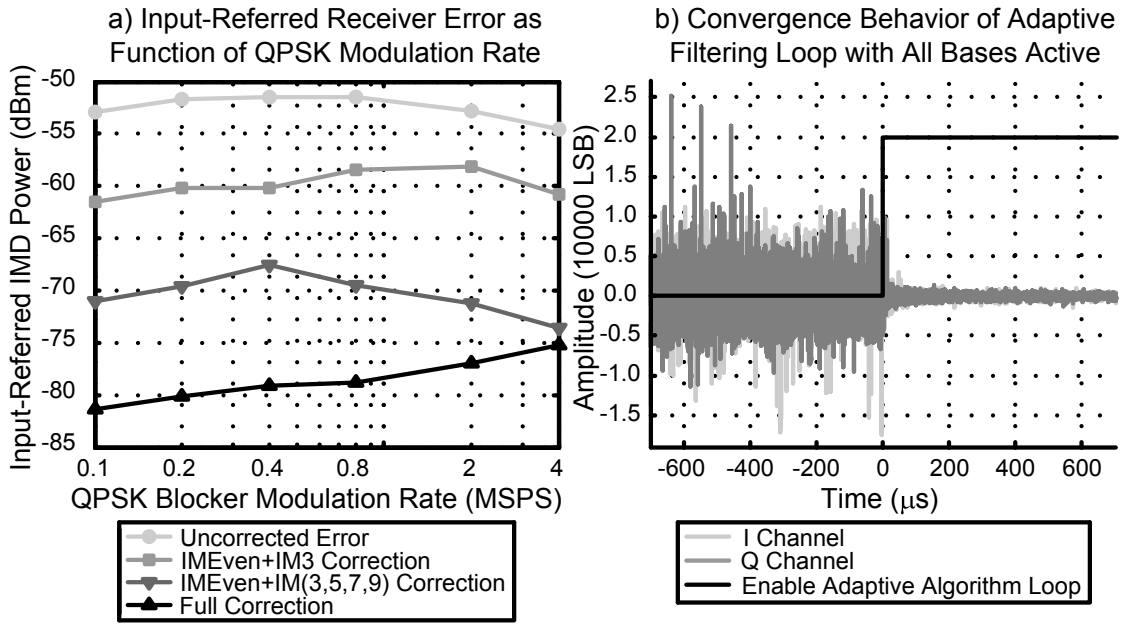


Figure 7.30: a) Measured input-referred error as function of modulation bandwidth for +12.4dBm QPSK blocker / -16.3dBm CW blocker. b) Convergence behavior of full adaptive algorithm for +11.4dBm 2MSPS QPSK blocker / -16.3dBm CW blocker.

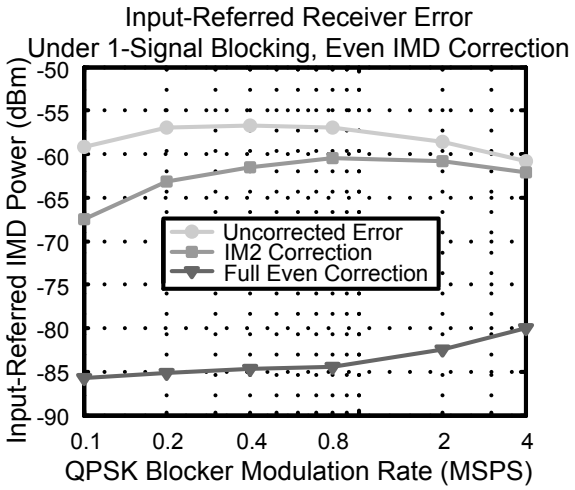


Figure 7.31: Measured input-referred even-order IM distortion error as function of modulation bandwidth for +12.4dBm QPSK blocker only.

floor, as predicted in Section 6.5. The measured cancellation performance for different levels of applied correction is shown in Fig. 7.29. At the worst-case full-correction value in Fig. 7.29a, an extrapolated IIP3 metric of +43.5dBm[2] is obtained after de-embedding residual even-order products and phase noise. The correction performance for peak blocking as a function of QPSK modulation bandwidth is shown in Fig. 7.30a, while the convergence behavior of the adaptive algorithm is shown in Fig. 7.30b. Note that convergence behavior shown in Fig. 7.30b is just the convergence of the adaptive algorithm and not the settling behavior of the high-pass filters. The correction performance for peak blocking as a function of QPSK modulation bandwidth for even-order IM distortion is shown in Fig. 7.31. De-embedding the results of Fig. 7.31 from Fig. 7.30a yield an extrapolated IIP3 metric of +44.4dBm for a QPSK blocker modulation rate of 0.8MHz when full correction is applied. As these tests are designed to predict the sensitivity of the receiver in a modern spread-spectrum communications system in which the magnitude of the desired signal at sensitivity is below that of the error floor, the RX signal model is not applied during the determination of the output-referred error. In order to input-refer the error, the modulated blocking tests were re-run in the presence of -66dBm and -86dBm CW desired signal models in order to determine small-signal gain changes of the system under the conditions for the top two curves in Figs. 7.29, 7.30a, and 7.31 and for the bottom curves representing modulated tests in Figs. 7.29, 7.30a, and 7.31, respectively.

7.4.3 Relation of Measured Modulated Blocker Receiver Tests to Theoretical Achievable Cancellation Ratio

Predicting the achievable intermodulation distortion cancellation ratio CR in the absence of other error using the results of Section 6.5 requires a minor computation involving the calculated IER and the IM distortion power of the odd- and even-order terms, respectively (7.4).

[2]This value is corrected from [119]

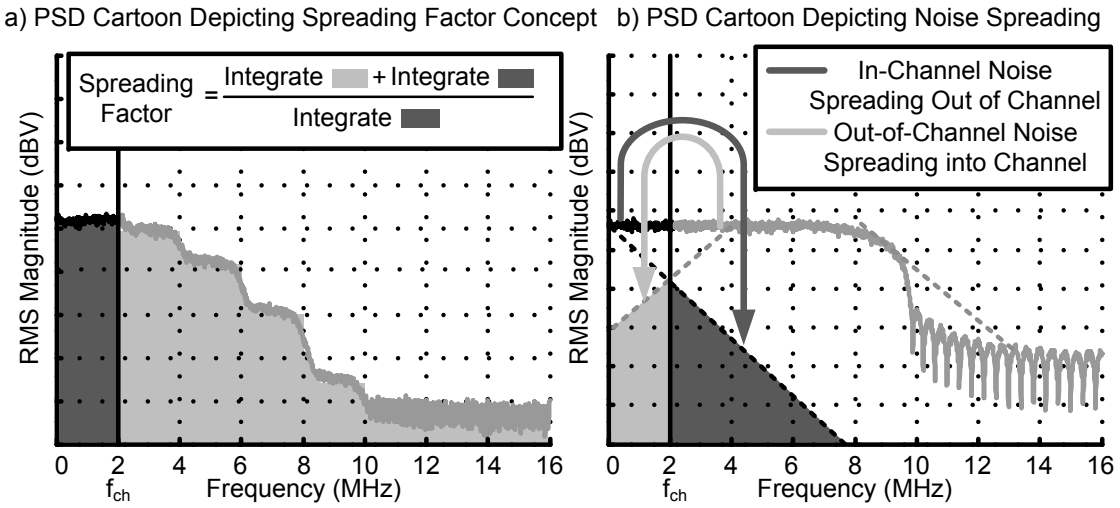


Figure 7.32: Cartoons depicting a) Spreading factor concept. b) Noise spreading out of and into channel bandwidth.

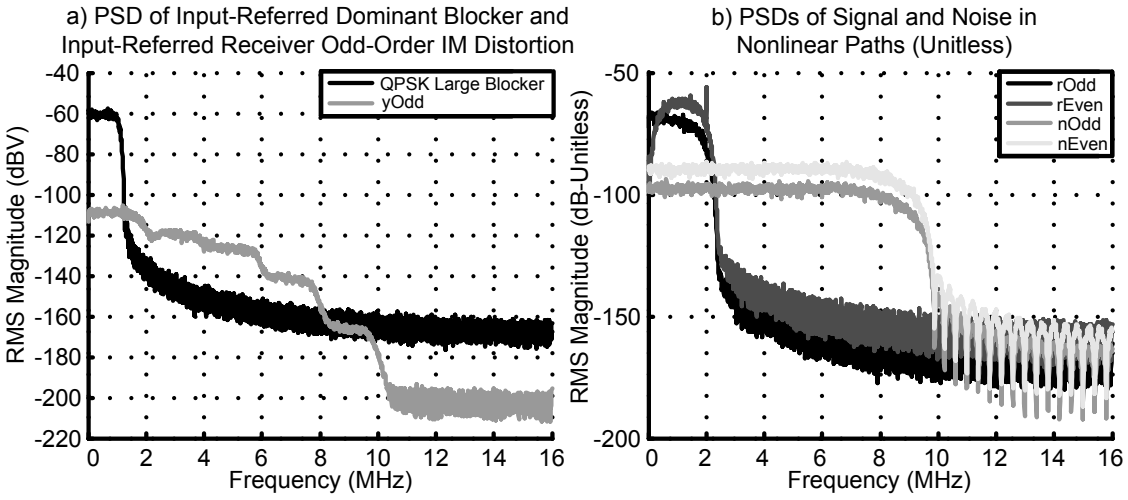


Figure 7.33: PSD of various signals within modified IER simulation and calculation. a) Baseband QPSK large blocker and odd-order IM distortion. b) Signal and noise in nonlinear paths.

$$CR = \frac{P_{IM,Even} + P_{IM,Odd}}{\frac{P_{IM,Even}}{10^{IER_{Even}(dB)/10}} + \frac{P_{IM,Odd}}{10^{IER_{Odd}(dB)/10}}} \tag{7.4}$$

Given that the original analysis presented in Section 6.4 neglected consideration of the signal and noise bandwidth, a small modification is in order before directly applying it to the receiver at hand. This original analysis computed the IER for the total IM distortion energy and for the total residual error energy. However, some of this energy can be spread outside of the channel bandwidth f_{CH} by the receiver nonlinearity, especially in the case of the wide-swing LNTA which effectively contains many high-order polynomial terms. This out-of-channel IMD energy is lumped into a quantity termed a spreading factor (SF) that denotes the ratio of total IMD energy to the

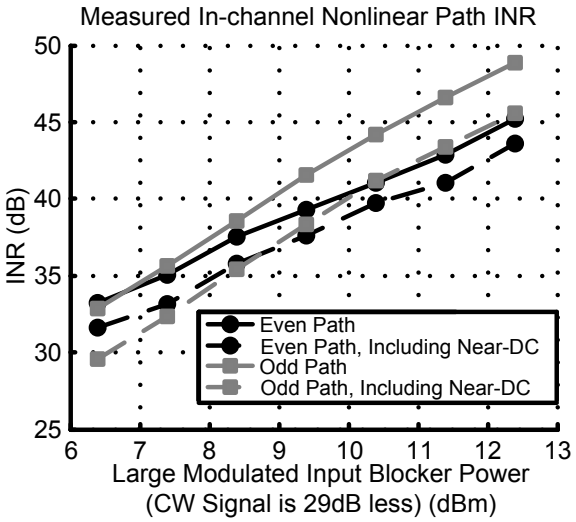


Figure 7.34: Measured in-channel INR of nonlinear paths, with and without spurious content at low frequencies.

amount of IMD energy falling in-channel, depicted in Fig. 7.32a. In order to account for this in the IER computation, the IMD error $E[y_{Odd}^2]$ is divided by the spreading factor SF, which is computed on an ad-hoc basis for each of the simulation/calculation scenarios of Section 6.5.

Similarly, it is assumed that the nondominant successive regeneration approximation error $E[e_{Odd,1}^2]$ is equally spread out of channel and must be divided by this quantity as well. The dominant successive regeneration error in the "S1T2F1" case due to uncancelled higher-order IMD terms is in general spread by a much higher factor. The SF in this case was determined in simulation and also incorporated into the calculation. Although the nonlinear path noise is also spread, because the bandwidth of the noise is greater than the reference IMD bandwidth in the nonlinear paths, a roughly equal amount of out-of-channel noise is spread back into the channel as is spread out of it, as depicted in Fig. 7.32b. Because of this, the noise error quantities of Section 6.5 $E[e_{Odd,2}^2]$ and $E[e_{Odd,3}^2]$ need not be altered in the modified analysis. However, the *a posteriori* nonlinear path INR computation is modified to only reflect the in-channel noise in order to provide continuity with the original simulation and calculation results.

In order to better reflect the implemented receiver, the simulation in this case is modified to include nonlinear path noise bandwidths of about 8MHz while sharp digital FIR filters with cutoff frequencies of about 2MHz are applied to both the y_{Odd} distortion signal and z_{Odd} reference distortion prior to MMSE computation of the weighting β_n and distortion cancellation. The results of these modifications for the case of a 2MSPS modulated QPSK dominant blocker and CW nondominant blocker are seen in Figures 7.35, 7.36, 7.37 and in this case result in only small differences from the plots shown in Section 6.5.

The measured in-channel INR of the two nonlinear paths for a 2MSPS QPSK dominant blocker and CW nondominant blocker is shown in Fig. 7.34. The even-order IMD at dc is removed from this calculation, as it is high-pass filtered out elsewhere in the DBE. The INR is reported for two cases: one in which the noise integration extends to dc and one stopping at about 15kHz to exclude

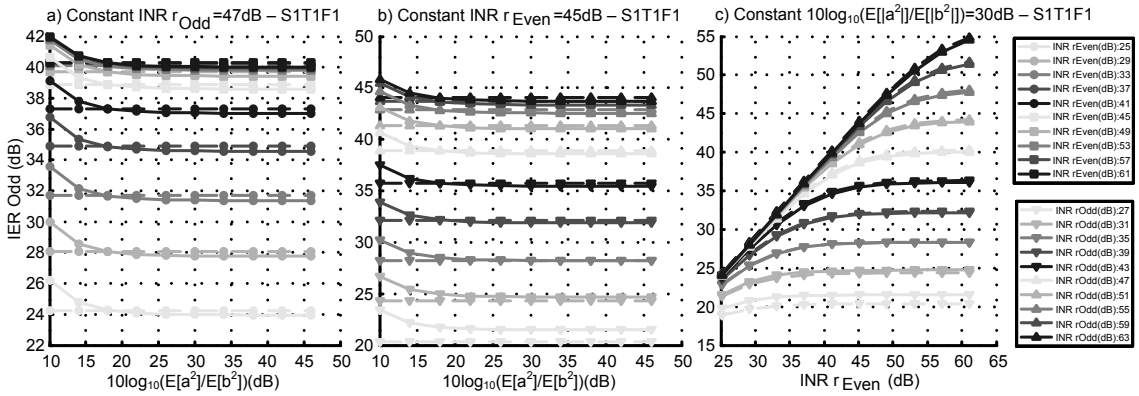


Figure 7.35: Simulation and modified calculation results of approximation error IER_{Odd} for case "S1T1F1".

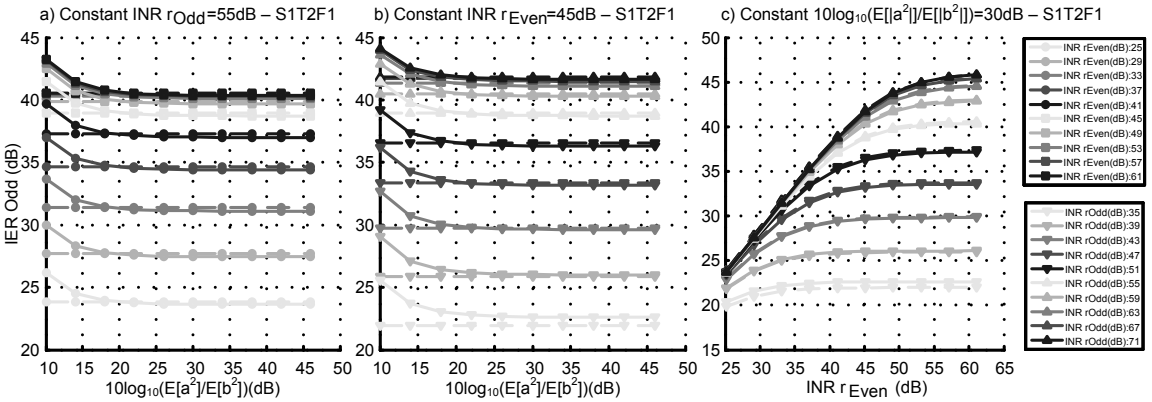


Figure 7.36: Simulation and modified calculation results of approximation error IER_{Odd} for case "S1T2F1".

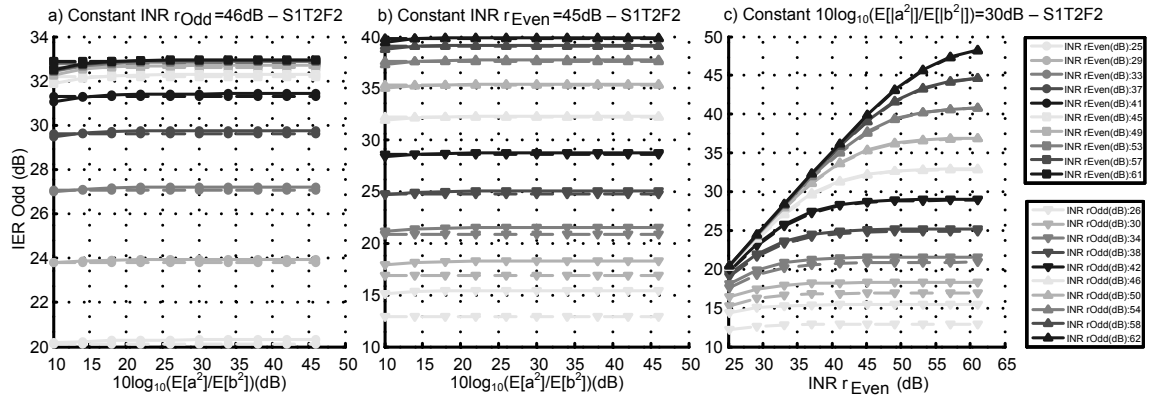


Figure 7.37: Simulation and modified calculation results of approximation error IER_{Odd} for case "S1T2F2".

significant low-frequency spurious signal content. Although this signal content is high-pass filtered in the IM3 reference basis element, its energy is spread in the other reference basis elements and is not filtered out. Hence, the complete measured in-channel INR is the relevant quantity with which to interpret the IER simulation and calculation results.

For the peak blocking scenario, it is seen that INR_{Even} ≈43dB while $INR_{Odd} \approx$ 45dB. In Fig. 7.35 this corresponds to an achievable IER_{Odd} of 36dB and IER_{Even} of 43dB. As seen in Fig. 7.29, the measured correction ratio is lower than either of these two quantities at 24dB. Part of the discrepancy here relates to decreased cancellation of error due to memory effects as seen in Fig. 7.30. A correction ratio of 27dB is achieved for lower QPSK modulation rates, however, in this case the spreading factor is unity and the results of Fig. 7.35 predict an odd-order IMD correction ratio of 37dB. As can be seen in Fig. 7.31, the even-order distortion cancellation ratio is fundamentally limited by the residual phase noise in the linear receiver path. After de-embedding the even-order distortion from the total input-referred error plot at low modulation rates, the input-referred odd-order error power is about -53dBm prior to cancellation and -80dBm afterwards for a cancellation ratio of 27dB, which is still somewhat less than 37dB. A portion of this discrepancy is clearly due to the cancellation limit set by the phase noise in the linear receiver path but the remainder requires further examination and could be due to the lack of another key functional basis or due to interaction between the adaptive filters processing the many different correlated odd-order basis elements.

Chapter 8

Conclusion

In this dissertation, a novel class of mixed-signal system-level techniques was introduced to mitigate the problem of corruptive intermodulation distortion in RF receivers introduced by large out-of-band undesired signals. A first system cancelling only third-order intermodulation distortion products was implemented and designed to the specifications of the Universal Mobile Telephone Service (UMTS) as an initial proof-of-concept. This system was then extended to provide for the cancellation of arbitrary intermodulation distortion products both in polynomial and functional form.

Novel circuitry was introduced to enable the nonlinear feedforward intermodulation distortion reference generation and large signal handling of the receiver. A multistage cubic term generator concept was introduced that showed good dynamic range performance for a given power dissipation. By high-pass filtering out low-frequency IM2 terms, the multistage cubic term generator both mitigated compressive effects of unnecessary signals and minimized the analog postfiltering requirements of the nonlinear reference receiver paths. Implications regarding the limitations of the interstage group delay inherent in the multistage architecture were quantitatively analyzed and deemed to pose little obstacle to the development of feedforward intermodulation distortion cancellation schemes capable of large cancellation ratios.

A novel wide-swing low-noise transconductance amplifier was developed with a near-constant-transconductance property that reduced the total magnitude of intermodulation distortion products needed to be processed by the baseband portion of the receiver. Due to its common-gate nature, this LNTA also presents a near-constant impedance to its source for optimal power matching. The LNTA was shown to be well-represented by a cubic-Gaussian functional basis which could be implemented in the proposed higher-order IMD cancellation scheme.

Techniques introduced in this dissertation take a strong step towards the possibility of implementing RF receivers without any external passive components whatsoever. By mitigating the pernicious effects of large signals rather than the signals themselves, circuit designers can develop low-cost, low-area receiver solutions without concern for high-quality frequency-selective elements at radio frequencies. Enabling these solutions can in turn enable new products, such as watch-sized cellular radios that work for any wireless standard in the world, or wireless sensor network radios both robust to out-of-band interference and small enough to fit on a fingertip.

To reach these exciting possibilities, however, more follow-up research is required. Although a rail-to-rail input receiver was introduced, in many cases it is desired for the receiver input to with-

stand even larger signals, as in the case of duplexerless FDD communications. Moving towards this goal requires the development of planar-process-compatible filtering structures that trade isolation for cost and die area. The rail-to-rail input receiver introduced achieved a high noise figure due to the low transconductance achievable by a common-gate LNTA. Further research in the design of higher-transconductance wide-swing and constant-transconductance LNTAs is desired to move the baseline performance of such systems closer to the requirements of the cellular communications marketplace. It was also shown that increasing the effective impedance looking up into the passive mixer improves the noise of the rail-to-rail input receiver. Additional work is required in order to determine how to maintain that high impedance while at the same time maintaining the capability of sinking over 100mA of undesired signal current without causing compression or breakdown of the receiver devices.

Appendix A

Achievable IER In Presence of Gain and Phase Mismatch for Canonical LMS-Based Algorithms

In this appendix, the maximum available IER_{ALT} allowed by a canonical complex LMS equalizer in the presence of arbitrary dc gain and phase mismatch on the main and alternate paths is derived. Rotational mismatch between I and Q paths and the phase of the incident IM3 products at RF are also factored into account. For the sake of simplicity, the main and alternate path baseband signals are assumed to be time-aligned at the input of a one-tap equalizer. In reality, the two taps of the implemented equalizer perform the delay adjustment to good approximation.

First, the RF gains of the two complex paths are lumped into effective baseband gains. Second, the phase delay produced by the interstage circuitry in the alternate path IM3 term generator is lumped into the alternate path LO rotational mismatch. These manipulations result in the equality of the effective RF IM3 products at both the main and alternate path mixer switching pair inputs. These products are given by (A.1), where $x_c(t)$ and $x_s(t)$ are assumed to be independent, zero mean, unit variance random processes.

$$x(t) = x_c(t)\cos\omega t + x_s(t)\sin\omega t \tag{A.1}$$

The four effective LO signals are denoted by (A.2)-(A.3)

$$l_{MI}(t) = \cos(\omega t + \phi'_M + \phi'_R + \phi'_D);\;\; l_{MQ}(t) = \sin(\omega t - \phi'_M + \phi'_R + \phi'_D) \tag{A.2}$$

$$l_{AI}(t) = \cos(\omega t + \phi'_A - \phi'_R + \phi'_D);\;\; l_{AQ}(t) = \sin(\omega t - \phi'_A - \phi'_R + \phi'_D) \tag{A.3}$$

Here, ϕ'_M denotes half of the I-Q phase mismatch in the main path, ϕ'_A denotes half of the I-Q phase mismatch in the alternate path, ϕ'_R denotes half of the rotational mismatch between the main and alternate paths, and ϕ'_D denotes the phase difference between the incoming IM3 terms and the nominal LO phase. In this case, the IM3 products appearing at baseband in each of the four paths are given by (A.4)-(A.7) where M_I, M_Q, A_I, and A_Q denote the effective lumped dc gains of each of the paths. The baseband sampling period is denoted as T and it is assumed that no information

is lost in the sampling process.

$$\begin{aligned} m_I[n] &= M_I \cos(\phi'_M + \phi'_R + \phi'_D)x_c(nT) - M_I \sin(\phi'_M + \phi'_R + \phi'_D)x_s(nT) \\ &= \alpha_c x_c(nT) - \alpha_s x_s(nT) \end{aligned} \tag{A.4}$$

$$\begin{aligned} m_Q[n] &= -M_Q \sin(\phi'_M - \phi'_R - \phi'_D)x_c(nT) + M_Q \cos(\phi'_M - \phi'_R - \phi'_D)x_s(nT) \\ &= -\beta_s x_c(nT) + \beta_c x_s(nT) \end{aligned} \tag{A.5}$$

$$\begin{aligned} a_I[n] &= A_I \cos(\phi'_A - \phi'_R + \phi'_D)x_c(nT) - A_I \sin(\phi'_A - \phi'_R + \phi'_D)x_s(nT) \\ &= \gamma_c x_c(nT) - \gamma_s x_s(nT) \end{aligned} \tag{A.6}$$

$$\begin{aligned} a_Q[n] &= -A_Q \sin(\phi'_A + \phi'_R - \phi'_D)x_c(nT) + A_Q \cos(\phi'_A + \phi'_R - \phi'_D)x_s(nT) \\ &= -\delta_s x_c(nT) + \delta_c x_s(nT) \end{aligned} \tag{A.7}$$

The LMS algorithm attempts to minimize the total mean squared error at the output of the equalizer. In the absence of thermal noise and as the convergence constant $\mu \to 0$ the equalizer is successful. For a one-tap adaptive complex LMS adaptive equalizer, with complex tap $w = w_I + jw_Q$, it can be shown that the remaining mean squared error is given by (A.8):

$$E^2 = (\alpha_c - w_I\gamma_c - w_Q\delta_s)^2 + (-\alpha_s + w_I\gamma_s + w_Q\delta_c)^2 + (-\beta_s + w_I\delta_s - w_Q\gamma_c)^2 + (\beta_c - w_I\delta_c + w_Q\gamma_s)^2 \tag{A.8}$$

Minimizing this expression entails setting the complex tap gradient equal to zero.

$$\frac{\partial E^2}{\partial w_I} = 2(\gamma_c^2 + \gamma_s^2 + \delta_c^2 + \delta_s^2)w_I - 2(\gamma_c\alpha_c + \gamma_s\alpha_s + \delta_c\beta_c + \delta_s\beta_s) + 2(\gamma_c\delta_s + \gamma_s\delta_c - \gamma_c\delta_s - \gamma_s\delta_c)w_Q = 0 \tag{A.9}$$

$$\frac{\partial E^2}{\partial w_Q} = 2(\gamma_c^2 + \gamma_s^2 + \delta_c^2 + \delta_s^2)w_Q - 2(\delta_c\alpha_s + \delta_s\alpha_c - \gamma_c\beta_s - \gamma_s\beta_c) + 2(\gamma_c\delta_s + \gamma_s\delta_c - \gamma_c\delta_s - \gamma_s\delta_c)w_I = 0 \tag{A.10}$$

Solving these equations and performing a series of trigonometric identities yields:

$$w_I = \frac{M_I A_I + M_Q A_Q}{A_I^2 + A_Q^2} \cos(\phi'_A - \phi'_M)\cos(2\phi'_R) + \frac{M_I A_I - M_Q A_Q}{A_I^2 + A_Q^2} \sin(\phi'_A - \phi'_M)\sin(2\phi'_R) \tag{A.11}$$

$$w_Q = \frac{M_I A_Q + M_Q A_I}{A_I^2 + A_Q^2} \cos(\phi'_A + \phi'_M)\sin(2\phi'_R) + \frac{M_I A_Q - M_Q A_I}{A_I^2 + A_Q^2} \sin(\phi'_A + \phi'_M)\cos(2\phi'_R) \tag{A.12}$$

For the case in which all baseband gains are equal, the expression has a particularly pleasing form.

$$w_I = \cos(\phi'_A - \phi'_M)\cos(2\phi'_R); \quad w_Q = \cos(\phi'_A + \phi'_M)\sin(2\phi'_R) \tag{A.13}$$

With these optimal tap weights realized, the orthogonality principle [134, pg. 439] can be utilized to obtain the residual expected squared error in both channels.

$$E^2 = \alpha_c^2 + \alpha_s^2 + \beta_c^2 + \beta_s^2 - (\alpha_c\gamma_c + \alpha_s\gamma_s + \beta_c\delta_c + \beta_s\delta_s)w_I - (\alpha_c\delta_s + \alpha_s\delta_c - \beta_c\gamma_s - \beta_s\gamma_c)w_Q \tag{A.14}$$

Again for the case in which all baseband gains are equal, the resultant expression for the effective

alternate path IER due to this effect alone is given by:

$$IER_{ALT}(dB) = -10\log_{10}\left[1 - \cos^2(\phi'_A - \phi'_M)\cos^2(2\phi'_R) - \cos^2(\phi'_A + \phi'_M)\sin^2(2\phi'_R)\right] \quad \text{(A.15)}$$

For the case in which baseband dc gains are not equal, expansion of (A.14) yields (A.16).

$$IER_{ALT}(dB) = -10\log_{10}\left[1 - \frac{\left(\begin{array}{c}(M_I A_I + M_Q A_Q)\cos(\phi'_A - \phi'_M)\cos(2\phi'_R) \\ +(M_I A_I - M_Q A_Q)\sin(\phi'_A - \phi'_M)\sin(2\phi'_R)\end{array}\right)^2}{(A_I^2 + A_Q^2)(M_I^2 + M_Q^2)} - \frac{\left(\begin{array}{c}(M_I A_Q + M_Q A_I)\cos(\phi'_A + \phi'_M)\sin(2\phi'_R) \\ +(M_I A_Q - M_Q A_I)\sin(\phi'_A + \phi'_M)\cos(2\phi'_R)\end{array}\right)^2}{(A_I^2 + A_Q^2)(M_I^2 + M_Q^2)}\right] \quad \text{(A.16)}$$

Substituting in for the actual phase error yields for (A.14) and (A.16) the following:

$$IER_{ALT}(dB) = -10\log_{10}\left[1 - \cos^2(\tfrac{1}{2}(\phi_A - \phi_M))\cos^2(\phi_R) - \cos^2(\tfrac{1}{2}(\phi_A + \phi_M))\sin^2(\phi_R)\right] \quad \text{(A.17)}$$

$$IER_{ALT}(dB) = -10\log_{10}\left[1 - \frac{\left(\begin{array}{c}(M_I A_I + M_Q A_Q)\cos(\frac{1}{2}(\phi_A - \phi_M))\cos(\phi_R) \\ +(M_I A_I - M_Q A_Q)\sin(\frac{1}{2}(\phi_A - \phi_M))\sin(\phi_R)\end{array}\right)^2}{(A_I^2 + A_Q^2)(M_I^2 + M_Q^2)} - \frac{\left(\begin{array}{c}(M_I A_Q + M_Q A_I)\cos(\frac{1}{2}(\phi_A + \phi_M))\sin(\phi_R) \\ +(M_I A_Q - M_Q A_I)\sin(\frac{1}{2}(\phi_A + \phi_M))\cos(\phi_R)\end{array}\right)^2}{(A_I^2 + A_Q^2)(M_I^2 + M_Q^2)}\right] \quad \text{(A.18)}$$

Appendix B

Dc Offset Removal Using the LMS Algorithm and its Equivalency to a High Pass Filter

It has been mentioned in the literature that dc offset removal can be added to the LMS algorithm by making the modification shown in Fig. B.1 [28] [72]. In this appendix, it is shown that incorporating this structure is tantamount to adding a high-pass filter in series with the equalizer. Hence, the convergence constant of the dc offset removal portion must be set such that the filter pole frequency is low enough to not cut out a significant portion of the signal band, effectively placing a lower bound on the convergence time of the entire equalizer.

First, it is noted that the dc offset removal portion alone exactly realizes a high-pass filter. This approach was taken in analog form in [40] in order to provide high-pass filtering in the analog domain for a UMTS baseband filter. The equivalence is depicted in Fig. B.2. It can be shown that the LTI transfer function of the structure in Fig. B.2 is given by the expression in (B.1), which is that of a discrete-time high pass filter.

$$\frac{E(z)}{M(z)} = \frac{1 - z^{-1}}{1 - (1 - \mu_{HP})z^{-1}} \tag{B.1}$$

In order to evaluate the structure in Fig. B.1, it is assumed that $m[n] = x[n] + a[n]$ where $x[n]$ and $a[n]$ are nonzero mean independent random processes each with i.i.d. samples. The alternate path input is also expanded as $a[n] = b[n] + a_{dc}$ where $b[n]$ is a zero-mean i.i.d. random process. This implies that $b[n]$ is independent of past values of $e[n]$. For the sake of convenience, it is assumed that there is no path mismatch aside from the dc offset. The output of the equalizer is then given by (B.2).

$$\begin{aligned} e[n] = x[n] + b[n] + a_{dc} - \left[\sum_{k=-\infty}^{n-1} \mu_{AP} e[k] a_{dc}\right] [b[n] + a_{dc}] \\ - \left[\sum_{k=-\infty}^{n-1} \mu_{AP} e[k] b[k]\right] [b[n] + a_{dc}] - \sum_{k=-\infty}^{n-1} \mu_{HP} e[k] \end{aligned} \tag{B.2}$$

Assuming that the tap update constants are set such that the adaptive equalizer converges, the

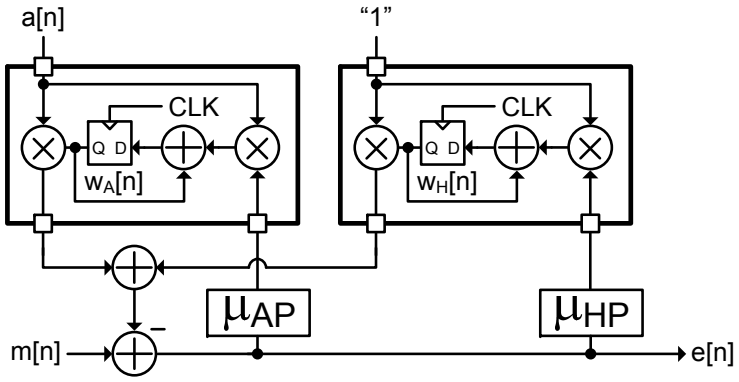


Figure B.1: LMS with dc offset reference input.

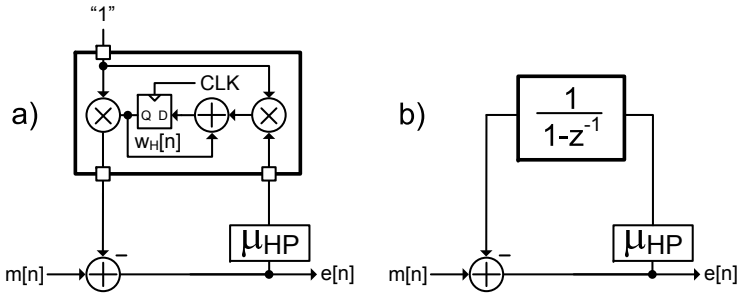


Figure B.2: Equivalence of LMS with dc offset removal to high pass filter.

following relation will hold:

$$\left[\sum_{k=-\infty}^{n-1} \mu_{AP} e[k]b[k]\right][b[n]+a_{dc}] = b[n]+a_{dc}+p[n]; \quad \left[\sum_{k=-\infty}^{n-1} \mu_{AP} e[k]a_{dc}\right] b[n] = q[n] \tag{B.3}$$

The quantities $p[n]$ and $q[n]$ are excess noise terms uncorrelated with $x[n]$ due to the i.i.d. assumption mentioned earlier. Hence, the expression for the output of the equalizer becomes:

$$e[n] = x[n] - p[n] - q[n] - \sum_{k=-\infty}^{n-1} (\mu_{HP} + \mu_{AP} a_{dc}^2) e[k] \tag{B.4}$$

It can be shown that the transfer function seen by the desired input signal $x[n]$ is given by B.5.

$$\frac{E(z)}{X(z)} = \frac{1-z^{-1}}{1-(1-\mu_{HP}-\mu_{AP}a_{dc}^2)z^{-1}} \tag{B.5}$$

In this case the equalizer still functions as a high-pass filter with respect to the desired input signal $x[n]$, albeit with the additional problem that the cutoff frequency is no longer known a priori. Assuming that $\mu_{AP}a_{dc,MAX}^2$ is much smaller than μ_{HP}, it is still true that μ_{HP} must be set so that the filter cutoff frequency is less than 10 kHz with an associated step response of greater than 40-60μs.

If the adaptive equalizer were to start up with the optimal alternate path equalizer tap value $w_{A,OPT}[n]$, the system would still take 40-60 μs to converge due to the dc offset correction portion. In an actual startup situation, the alternate path equalizer tap value varies as it converges, changing the dc content of the equalizer output. This prevents the dc offset correction portion from settling until after its original, minimal, settling time. Hence, the convergence time of such an equalizer will have a lower bound set by the pole frequency of the dc offset correction portion.

Appendix C

Abbreviations Used in Text

>>	Right bit shift
ac	Alternating current
ACPR	Adjacent channel power ratio
ADC	Analog-to-digital converter
ADI	Analog Devices, Incorporated
ALT	Alternate
ATC	American Technical Ceramics
AWGN	Additive white Gaussian noise
BER	Bit error rate
BiCMOS	A monolithic transistor technology incorporating both BJT and CMOS devices.
BJT	Bipolar junction transistor
BPSK	Binary phase-shift keying
BQ	Biquadratic filter (Biquad)
CDMA	Code-division multiple access
CG	Common-gate
CHIC	Caltech High-speed Integrated Circuits research group
CLK	Clock
CMFB	Common-mode feedback
CML	Current-mode logic
CMOS	Complementary metal-oxide-semiconductor
CR	Correction ratio
CS	Common-source
CW	Continuous wave
DAC	Digital-to-analog converter
DBE	Digital back end
dc	Direct current
DFF	D-flip-flop
DFFE	DFF with enable pin
DPCH	Dedicated physical channel
DTFT	Discrete-time Fourier transform
EN	Enable

ESD	Electrostatic discharge
EVM	Error vector magnitude
FBAR	Film bulk acoustic resonator
FDD	Frequency-domain-duplexed
FER	Frame error rate
FIR	Finite impulse response
FOM	Figure of merit
FPGA	Field programmable gate array
FR	Frequency-reversed
g_m	Small-signal transconductance
GMSK	Gaussian minimum-shift keying
GND	Electrical voltage ground
GPIO-HSMC	General-purpose input-output to high-speed mezzanine card
GSM	Global System for Mobile Communications
HB	Halfband
HD	Harmonic distortion
HP	Hewlett-Packard Development Company
HPF	High-pass filter
I	In-phase
IED	Improvised explosive device
IEEE	Institute of Electrical and Electronics Engineers
IER	Intermodulation distortion-to-error ratio
i.i.d.	Independent and identically distributed
IIP2	Second-order intermodulation distortion input-referred intercept point
IIP3	Third-order intermodulation distortion input-referred intercept point
IIR	Infinite impulse response
IM	Intermodulation
IM2	Second-order intermodulation distortion
IM3	Third-order intermodulation distortion
IMD	Intermodulation distortion
INR	Intermodulation distortion-to-noise ratio
IP	Intellectual property
I-V	Current-voltage
JPL	Jet Propulsion Laboratory
kTB	(Boltzmann's constant)·(Temperature)·(Bandwidth)
LMS	Least mean squares
LNA	Low-noise amplifier
LNTA	Low-noise transconductance amplifier
LO	Local oscillator
LPF	Low-pass filter
LTI	Linear time-invariant
LUT	Look-up table
MICS	Mixed-mode Integrated Circuits and Systems group at Caltech.

MMSE	Minimum mean squared error
MOS	Metal-oxide-semiconductor
MOSFET	Metal-oxide-semiconductor field-effect transistor
MSPS	Millions of samples per second
NAND	Not-and
NF	Noise figure
NLMS	Normalized-least mean squares
NMOS	N-channel MOSFET
O-QPSK	Offset quadrature phase-shift keying
OTA	Operational transconductance amplifier
PA	Power amplifier
PCB	Printed circuit board
PFD	Phase / frequency detector
PIN	P-type / Intrinsic / N-type
PLL	Phase-locked loop
PMOS	P-channel MOSFET
PNOISE	Periodic noise analysis
PSD	Power spectral density
PVT	Process-voltage-temperature
Q	Quadrature
QAM	Quadrature amplitude modulation
QPNOISE	Quasi-periodic noise analysis
QPSK	Quadrature phase-shift keying
QPSS	Quasi-periodic steady-state analysis
QVCO	Quadrature VCO
RAKE	A radio receiver designed to counter the effects of multipath fading.
RF	Radio-frequency
RLC	Resistor-inductor-capacitor
RMS	Root mean squared
RRC	Root-raised cosine
RSSI	Received strength signal indicator
RST	Reset
RX	Receive
SAW	Surface acoustic wave
SER	Signal-to-error ratio
SF	Spreading factor
SMA	Subminiature version A
SNR	Signal-to-noise ratio
TAP	Filter tap
TC	Transconductance
TIA	Transimpedance amplifier
TX	Transmit
UMC	United Microelectronics Corporation

UMTS	Universal Mobile Telecommunications System
VCM	Common-mode voltage
VCO	Voltage-controlled oscillator
VDD	Electrical voltage supply
XMD	Cross-modulation distortion

Bibliography

[1] E. A. Keehr and A. Hajimiri, "Equalization of third-order intermodulation products in wideband direct-conversion receivers," *IEEE J. Solid-State Circuits*, vol. 43, pp. 2853–2867, Dec. 2008.

[2] ——, "Digitally enhanced alternate path linearization of RF receivers," in *Multi-Mode/Multi-Band RF Transceivers for Wireless Communications*. Wiley Interscience, 2010, pp. 309–345.

[3] S. Haykin, *Digital Communications*. Singapore: John Wiley and Sons (Asia), 2003.

[4] P. Orsatti, F. Piazza, and Q. Huang, "A 20-mA-receiver, 55-mA-transmit, single-chip GSM transceiver in 0.25-μm CMOS," *IEEE J. Solid-State Circuits*, vol. 34, pp. 1869–1880, Dec. 1999.

[5] M. Hafizi, S. Feng, T. Fu, K. Schulze *et al.*, "RF front-end of direct conversion receiver RFIC for CDMA-2000," *IEEE Journal of Solid-State Circuits*, vol. 39, pp. 1622–1632, Oct. 2004.

[6] "UE Radio Transmission and Reception (FDD)," Tech. Specification Group, 3GPP, (TSG) RAN WG4, TS 25.101, v8.1.0, Dec. 2007.

[7] R. Steele, *Mobile Radio Communications* . New York: IEEE Press, 1992.

[8] S. S. Tsai and Y. C. Yoon, "SNR estimation for power control of cdma2000 forward fundamental channel," in *IEEE Global Telecommunications Conference*, vol. 2, Dec. 2003, pp. 1008–1012.

[9] S. K. Reynolds, B. A. Floyd, T. J. Beukema, T. Zwick, and U. R. Pfeiffer, "Design and compliance testing of a SiGe WCDMA receiver IC with integrated analog baseband," *Proceedings of the IEEE*, vol. 93, pp. 1624–1636, Sep. 2005.

[10] B. Razavi, *RF Microelectronics*. Upper Saddle River, NJ: Prentice Hall, 1998.

[11] R. Magoon, A. Molnar, J. Zachan, G. Hatcher, and W. Rhee, "A single-chip quad-band (850 / 900 / 1800 / 1900 MHz) direct conversion GSM/GPRS RF transceiver with integrated VCOs and fractional-N synthesizer," *IEEE J. Solid-State Circuits*, vol. 37, pp. 1710–1720, Dec. 2002.

[12] D. Manstretta, M. Brandolini, and F. Svelto, "Second-order intermodulation distortion mechanisms in CMOS downconverters," *IEEE J. Solid-State Circuits*, vol. 38, pp. 394–406, Mar. 2003.

[13] Z. Ru, N. A. Moseley, E. A. M. Klumperink, and B. Nauta, "Digitally enhanced software defined radio receiver robust to out-of-band inteference," *IEEE J. Solid State Circuits*, vol. 44, pp. 3359–3375, Dec. 2009.

[14] C. Andrews and A. C. Molnar, "A passive-mixer-first receiver with baseband-controlled RF impedance matching, <6dB NF, and >27dBm wideband IIP3," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2010, pp. 46–47,47a.

[15] A. Springer, L. Maurer, and R. Weigel, "RF system concepts for highly integrated RFICs for W-CDMA mobile radio terminals," *IEEE Trans. Microwave Theory and Techniques*, vol. 50, pp. 254–267, Jan. 2002.

[16] W. Y. Ali-Ahmad, "Effective IM2 estimation for two-tone and WCDMA modulated blocker in zero-IF," *RF Design*, pp. 32–40, Apr. 2004.

[17] R. R. Hay, "Digitally-tunable surface acoustic wave resonator," Ph.D. dissertation, Boise State University, 2009.

[18] D. Petit, E. Cesar, P. Bar *et al.*, "Thermally stable oscillator at 2.5GHz using temperature compensated BAW resonator and its integrated temperature sensor," in *Proceedings of the 2008 International Ultrasonics Symposium*, Nov. 2008, pp. 895–898.

[19] A. Muller, D. Neculoiu *et al.*, "6.3-GHz film bulk acoustic resonator structures based on a gallium nitride/silicon thin membrane," *IEEE Electron Device Letters*, vol. 30, pp. 799–801, Aug. 2009.

[20] A. Hadjichristos, M. Cassia *et al.*, "Single-chip RF CMOS UMTS/EGSM transceiver with integrated receive diversity and GPS," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2009, pp. 118–119.

[21] B. A. Floyd, S. K. Reynolds, T. Zwick, L. Khuon, T. Beukema, and U. R. Pfeiffer, "WCDMA direct-conversion receiver front-end comparison in RF-CMOS and SiGe BiCMOS," *IEEE Trans. Microwave Theory and Techniques*, vol. 53, pp. 1181–1188, Apr. 2005.

[22] T. H. Lee, *The design of CMOS radio-frequency integrated circuits*, 2nd ed. Cambridge, UK: Cambridge University Press.

[23] H. S. Black, "Stabilized feedback amplifiers," *Elec. Eng.*, vol. 53, pp. 114–120, Jan. 1934.

[24] V. Aparin and L. E. Larson, "Modified derivative superposition method for linearizing FET low-noise amplifiers," *IEEE Trans. Microwave Theory and Techniques*, vol. 53, pp. 571–581, Feb. 2005.

[25] A. A. M. Saleh and J. Salz, "Adaptive linearization of power amplifiers in digital radio systems," *The Bell System Technical Journal*, vol. 62, pp. 1019–1033, Apr. 1983.

[26] S. P. Stapleton and F. C. Costescu, "An adaptive predistorter for power amplifier based on adjacent channel emissions," *IEEE Trans. Vehicular Technology*, vol. 41, pp. 49–56, Feb. 1992.

[27] B.Gilbert, "The multi-tanh principle: a tutorial overview," *IEEE J. Solid-State Circuits*, vol. 33, pp. 2–17, Jan. 1998.

[28] M. Faulkner, "DC offset and IM2 removal in direct conversion receivers," *IEE Proceedings Communications*, vol. 149, pp. 179–184, Jun. 2002.

[29] V. H. Estrick and R. T. Siddoway, "Receiver distortion circuit and method," U.S. Patent 5,237,332, filed Feb. 25 1992, granted Aug. 17 1993.

[30] M. Valkama, A. S. H. Ghadam, L. Antilla, and M. Renfors, "Advanced digital signal processing techniques for compensation of nonlinear distortion in wideband multicarrier radio receivers," *IEEE Trans. Microwave Theory and Techniques*, vol. 54, pp. 2356–2366, June 2006.

[31] L. Yu and M. Snelgrove, "A novel adaptive mismatch cancellation scheme for quadrature IF radio receivers," *IEEE Trans. Circuits and Systems - II*, vol. 46, pp. 789–801, June 2006.

[32] ——, "Signal processor for reducing undesirable signal content," U.S. Patent 6,804,359, filed Aug. 3, 1998, granted Oct. 12, 2004.

[33] V. Aparin, G. J. Ballantyne, C. J. Persico, and A. Cicalini, "An integrated LMS adaptive filter of TX leakage for CDMA receiver front ends," *IEEE J. Solid-State Circuits*, vol. 41, pp. 1171–1182, May 2006.

[34] Z. S. Ebadi and R. Saleh, "Adaptive compensation of RF front-end nonidealities in direct conversion receivers," *IEEE Trans. Circuits and Systems - II*, vol. 55, pp. 354–358, Mar. 2008.

[35] Y. Zheng, M. Cao, E. K. H. Teo, and H. K. Garg, "An adaptive filtering algorithm for direct-conversion receivers: architecture and performance analysis," *IEEE Trans. Circuits and Systems - I*, vol. 46, pp. 1141–1148, May 2008.

[36] R. Bagheri, A. Mirzaei, S. Chehrazi, M. E. Heidari, M. Lee, M. Mikhemar, W. Tang, and A. A. Abidi, "An 800-MHz-6-GHz software-defined wireless receiver in 90nm CMOS," *IEEE J. Solid-State Circuits*, vol. 41, pp. 2860–2876, Dec. 2006.

[37] S. Haykin, *Adaptive Filter Theory*, 4th ed. Upper Saddle River, NJ: Prentice Hall, 1998.

[38] F. Gatta, D. Manstretta, P. Rossi, and F. Svelto, "A fully integrated 0.18-μm CMOS direct conversion receiver front-end with on-chip LO for UMTS," *IEEE J. Solid-State Circuits*, vol. 39, pp. 15–23, Jan. 2004.

[39] V. Aparin, N. Kim, G. Brown, Y. Wu, A. Cicalini, S. Kwok, and C. Persico, "A fully-integrated highly linear zero-IF CMOS cellular CDMA receiver," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2009, pp. 324–601.

[40] J. Rogin, I. Kouchev, G. Brenna, D. Tschopp, and Q. Huang, "A 1.5-V 45-mW direct conversion WCDMA receiver IC in 0.13μm CMOS," *IEEE J. Solid-State Circuits*, vol. 38, pp. 2239–2248, Dec. 2003.

[41] A. Liscidini, M. Brandolini, D. Sanzogni, and R. Castello, "A 0.13μm CMOS front-end for DCS1800 / UMTS / 802.11b-g with multiband positive feedback low-noise amplifier," *IEEE J. Solid State Circuits*, vol. 41, pp. 981–989, Apr. 2006.

[42] P. Sivonen, J. Tervaluoto, N. Mikkola, and A. Parsinnen, "A 1.2-V RF front-end with on-chip VCO for PCS 1900 direct conversion receiver in 0.13-μm CMOS," *IEEE J. Solid-State Circuits*, vol. 41, pp. 384–394, Feb. 2006.

[43] M. Tamura, T. Nakayama, Y. Hino *et al.*, "A low voltage (1.8V) operation triple band WCDMA transceiver IC," in *IEEE RFIC Symp. Dig. Tech. Papers*, June 2005, pp. 269–272.

[44] D. K. Shaeffer and T. H. Lee, "A 1.5-V, 1.5GHz CMOS low noise amplifier," *IEEE Journal of Solid-State Circuits*, vol. 32, pp. 745–759, May 1997.

[45] R. Xu, L. Sun, and J. Wen, "A highly linear wideband CMOS LNA adopting current amplification and distortion cancellation," in *The 9th International Conferernce on Solid-State and Integrated-Circuit Technology*, Oct. 2008, pp. 1512–1515.

[46] H. Darabi, "A blocker filtering technique for SAW-less wireless receivers," *IEEE J. Solid-State Circuits*, vol. 42, pp. 2766–2773, Dec. 2007.

[47] B. A. Floyd and D. Ozis, "Low-noise amplifier comparison at 2GHz in 0.25-μm and 0.18-μm RF-CMOS and SiGe BiCMOS," in *IEEE RFIC Symp. Dig. Tech. Papers*, Jun. 2004, pp. 185–188.

[48] W.-H. Chen, G. Liu, B. Zdravko, and A. M. Niknejad, "A highly linear broadband CMOS LNA employing noise and distortion cancellation," in *IEEE RFIC Symp. Dig. Tech. Papers*, Jun. 2007, pp. 61–64.

[49] S. B. T. Wang, A. M. Niknejad, and R. W. Brodersen, "Design of a sub-mW 960-MHz UWB CMOS LNA," *IEEE J. Solid-State Circuits*, vol. 41, pp. 2449–2456, Nov. 2006.

[50] M. Camus, B. Butaye, L. Garcia, M. Sie, B. Pellat, and T. Parra, "A 5.4mW/0.07 mm^2 2.4GHz front-end receiver in 90nm CMOS for IEEE 802.15.4 WPAN standard," *IEEE J. Solid-State Circuits*, vol. 43, pp. 1372–1383, Jun. 2008.

[51] M. Valla, G. Montagna, R. Castello, R. Tonietto, and I. Bietti, "A 72-mW CMOS 802.11a direct conversion front-end with 3.5-dB NF and 200kHz 1/f Noise Corner," *IEEE J. Solid-State Circuits*, vol. 40, pp. 970–977, Apr. 2005.

[52] W. Liu, M. Sivaprakasam, G. Wang, M. Zhou, J. Granacki, J. Lacoss, and J. Wills, "Implantable biomimetic microeletronic systems design," *IEEE Engineering in Medicine and Biology Magazine*, pp. 66–74, Sep. 2005.

[53] K. Lee, I. Nam, I. Kwon *et al.*, "The impact of semiconductor technology scaling on CMOS RF and digital circuits for wireless application," *IEEE Trans. Electron Devices*, vol. 52, pp. 1415–1422, Jul. 2005.

[54] G. E. Moore, "Progress in digital integrated electronics," in *1975 International Electron Devices Meeting*, vol. 21, Dec. 1975, pp. 11–13.

[55] H. S. Lee, D. A. Hodges, and P. R. Gray, "A self-calibrating 15 bit CMOS A/D converter," *IEEE J. Solid State Circuits*, vol. 19, pp. 813–819, Dec. 1984.

[56] E. Fogleman, J. Welz, and I. Galton, "An audio ADC delta-sigma modulator with 100-dB peak SINAD and 102-dB DR using a second-order mismatch-shaping DAC," *IEEE J. Solid-State Circuits*, vol. 36, pp. 339–348, Mar. 2001.

[57] E. Keehr and A. Hajimiri, "Equalization of IM3 products in wideband direct-conversion receivers," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2008, pp. 204–205.

[58] E. A. Keehr and A. Hajimiri, "Digitally-assisted linearization of wideband direct conversion receivers," in *European Microwave Integrated Circuit Conference*, Oct. 2008, pp. 159–162.

[59] ——, "Digitally-assisted equalization of third-order intermodulation products in wideband direct conversion receivers," *Intl. Journal of Microwave and Wireless Technologies*, vol. 1, pp. 377–385, Aug. 2009.

[60] [Online]. Available: http://www.umts-forum.org/content/view/2315/110

[61] D. Kaczman *et al.*, "A single-chip tri-band (2100, 1900, 850/800 MHz) WCDMA / HSDPA cellular tranceiver IC," *IEEE J. Solid-State Circuits*, vol. 41, pp. 1122–1132, May 2006.

[62] N. K. Yanduru, D. Griffith, S. Bhagavatheeswaran *et al.*, "A WCDMA, GSM/GPRS/EDGE receiver front end without interstage SAW filter," in *IEEE RFIC Symp. Dig. Tech. Papers*, June 2006, pp. 11–13.

[63] B. Tenbroek, J. Strange, D. Nalbantis *et al.*, "Single-chip tri-band WCMDA/HSDPA transceiver without external SAW filtering and with integrated TX power control," in *IEEE Int. Solid-State Circuit Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2008, pp. 202–203.

[64] A. Safarian, A. Shameli, A. Rofougaran, M. Rofougaran, and F. de Flaviis, "Integrated blocker filtering RF front ends," in *IEEE RFIC Symp. Dig. Tech. Papers*, Jun. 2007, pp. 13–16.

[65] M. Brandolini, P. Rossi, D. Sanzogni, and F. Svelto, "A +78dBm IIP2 CMOS direct down-conversion mixer for fully integrated UMTS receivers," *IEEE J. Solid-State Circuits*, vol. 41, pp. 552–559, Mar. 2006.

[66] T. Oshima, K. Maio, W. Hioe, and Y. Shibahara, "Novel automatic tuning method of RC filters using a digital-DLL technique," *IEEE J. Solid-State Circuits*, vol. 39, pp. 2052–2054, Nov. 2004.

[67] Y.-H. Kim and H.-K. Yu, "Automatic tuning circuit for Gm-C filters," in *Electronics, Circuits and Systems, 2005. ICECS 2005. 12th IEEE International Conference on*, Dec. 2005, pp. 1–4.

[68] M. Tarrab and A. Feuer, "Convergence and performance analysis of the normalized LMS algorithm with uncorrelated gaussian data," *IEEE Trans. on Information Theory*, vol. 34, pp. 680–691, Jul. 1998.

[69] D. T. M. Slock, "On the convergence behavior of the LMS and normalized LMS algorithms," *IEEE Trans. Signal Processing*, vol. 41, pp. 2811–2825, Sep. 1993.

[70] V. H. Nascimento, "A simple model for the effect of normalization on the convergence rate of adaptive filters," in *IEEE. Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP) Dig. Tech. Papers*, vol. 2, May 2004, pp. 453–6.

[71] H. J. Bergveld *et al.*, "A low-power highly digitized receiver for 2.4-GHz-Band GFSK applications," *IEEE Trans. Microwave Theory and Techniques*, vol. 53, pp. 453–461, Feb. 2005.

[72] L. Der and B. Razavi, "A 2-GHz CMOS image-reject receiver with LMS calibration," *IEEE J. Solid-State Circuits*, vol. 38, pp. 167–175, Feb. 2003.

[73] muRata Corp. Part Number DFYK61G95LBJCA. http://www.murata.com Data sheet available at http://smartdata.usbid.com/datasheets/usbid/dsid/103495.pdf.

[74] K. Gerlach, "The effect of I,Q mismatch errors on adaptive cancellation," *IEEE Trans. Aerospace and Electronic Systems*, pp. 729–740, Jul. 1992.

[75] K. Gerlach and M. J. Steiner, "An adaptive matched filter that compensates for I, Q mismatch errors," *IEEE Trans. Signal Processing*, vol. 45, pp. 3104–3107, Dec. 1997.

[76] N. A. Moseley, Z. Ru, E. A. M. Klumperink, and B. Nauta, "A 400-to-900 MHz receiver with dual-domain harmonic rejection exploiting adaptive interference cancellation," in *IEEE Int. Solid-State Circuit Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2009, pp. 232–233.

[77] M. Schwab and P. Seidenberg, "Analysis of mobile-originated interference in coexisting UMTS networks," in *Vehicular Technology Conference*, vol. 4, May 2002, pp. 1636–1639.

[78] A. Springer and R. Weigel, *UMTS: The Physical Layer of the Universal Mobile Telecommunications System.* Berlin: Springer-Verlag, 2002.

[79] B. J. Minnis and P. A. Moore, "A highly digitized multimode receiver architecture for 3G mobiles," *IEEE Trans. Vehicular Technology*, vol. 52, pp. 637–653, May 2003.

[80] H. Oba, M. Kim, and H. Arai, "FPGA implementation of LMS and N-LMS processor for adaptive array applications," in *Int. Symp. on Intelligent Signal Processing and Communications (ISPACS) Dig. Tech. Papers*, Dec. 2006, pp. 485–488.

[81] M. Jeruchim, "Techniques for estimating the bit error rate in the simulation of digital communication systems," *IEEE J. Selected Areas Communications*, vol. SAC-2, pp. 153–170, Jan. 1984.

[82] E. Keehr and A. Hajimiri, "Analysis of internally bandlimited multistage cubic term generators for RF receivers," *IEEE Trans. Circuits and Systems-I*, vol. 56, pp. 1758–1771, Aug. 2009.

[83] T. Nojima and N. Konno, "Cuber predistortion linearizer for relay equipment in 800 MHz band land mobile telephone system," *IEEE Trans. Vehicular Technology*, vol. VT-34, pp. 169–177, Nov. 1985.

[84] N. Imai, T. Nojima, and T. Murase, "Novel linearizer using balanced circulators and its application to multilevel digital radio systems," *IEEE Trans. Microwave Theory and Techniques*, vol. 37, pp. 1237–1243, Aug. 1989.

[85] W. Huang and R. E. Saad, "Novel third-order distortion generator with residual IM2 suppression capabilities," *IEEE Trans. Microwave Theory and Techniques*, vol. 46, pp. 2372–2382, Dec. 1998.

[86] L. Roselli, V. Borgioni, V. Palassari, and F. Alimenti, "An active cuber circuit for power amplifier analog predistortion," in *33rd European Microwave Conference*, vol. 3, Oct. 2003, pp. 1219–1222.

[87] R. Sadhwani and B. Jalali, "Adaptive CMOS predistortion linearizer for fiber optic links," *IEEE J. Lightwave Tech.*, vol. 21, pp. 3180–3193, Dec. 2003.

[88] F. Shearer and L. MacEachern, "A precision CMOS analog cubing circuit," in *IEEE NEWCAS*, June 2004, pp. 281–284.

[89] F. Shearer, "CMOS analog cubing circuits for radio-over-fiber predistortion," Master's thesis, Carleton Univ., Ottawa, Canada, 2005.

[90] T. Nesimoglu, C. N. Canagarajah, and J. P. McGeehan, "A broadband polynomial predistorter for reconfigurable radio," in *Vehicular Technology Conference*, vol. 3, May 2001, pp. 1968–1972.

[91] T. Rahkonen *et al.*, "Performance of an integrated 2.1 GHz analog predistorter," in *2006 International Workshop on Integrated Nonlinear Microwave and Millimeter-Wave Circuits*, Jan. 2006, pp. 34–37.

[92] E. Westesson and L. Sundstrom, "A complex polynomial predistorter chip in CMOS for baseband or IF linearization of RF power amplifiers," in *Proc. of IEEE Int. Symp. on Circuits and Systems (ISCAS 99)*, vol. 1, Jun. 1999, pp. 206–209.

[93] T. S. Nielsen, S. Lindfors, S. Tawfik, and T. Larsen, "0.25mm CMOS analog multiplier for polynomial predistorter," in *Proc. Norchip Conf.*, Nov. 2004, pp. 191–194.

[94] H.-J. Song and C.-K. Kim, "A MOS four-quadrant analog multiplier using simple two-input squaring circuits with source followers," *IEEE J. Solid-State Circuits*, vol. 25, pp. 841–848, Jun. 1990.

[95] B. Razavi, *Design of Analog CMOS Integrated Circuits.* New York, NY: McGraw-Hill, 2003.

[96] P. Langlois and A. Demosthenous, "Sweet spots in moderate inversion for MOSFET squarer transconductors," *IEEE Trans. Circuits and Systems-II*, vol. 54, pp. 479–483, Jun. 2007.

[97] K. Han, J. Gil, S.-S.Song *et al.*, "Complete high-frequency thermal noise modeling of short-channel MOSFETs and design of 5.2-GHz low noise amplifier," *IEEE J. Solid-State Circuits*, vol. 40, pp. 726–735, Mar. 2005.

[98] M. T. Yang, C. W. Kuo, P. P. C. Ho *et al.*, "CR018 wideband noise model for AMS/RF CMOS simulation," in *IEEE RFIC Symp. Dig. Tech. Papers*, Jun. 2007, pp. 643–646.

[99] V. Aparin and L. E. Larson, "Analysis and reduction of cross-modulation distortion in CDMA receivers," *IEEE Trans. Microwave Theory and Techniques*, vol. 51, pp. 1591–1602, May 2003.

[100] A. V. Oppenheim, A. S. Willsky, and S. H. Nawab, *Signals and systems*, 2nd ed. Upper Saddle River, NJ: Prentice Hall.

[101] D. M. W. Leenaerts and W. Redman-White, "1/f noise in passive CMOS mixer for low and zero IF integrated receivers," in *Proc. 27th Eur. Solid-State Circuits Conf.*, Sep. 2001, pp. 103–107.

[102] E. Sacchi, I. Bietti, S. Erba, L. Tee, P. Vilmercati, and R. Castello, "A 15mW, 70kHz 1/f corner direct conversion CMOS receiver," in *Proc. IEEE Custom Integrated Circuits Conference*, Sep. 2003, pp. 459–462.

[103] N. Poobuapheun, W.-H. Chen, Z. Boos, and A. M. Niknejad, "A 1.5-V 0.7-2.5-GHz CMOS quadrature demodulator for multiband direct-conversion receivers," *IEEE J. Solid-State Circuits*, vol. 42, pp. 1669–1677, Aug. 2007.

[104] N. Kim, V. Aparin, and L. E. Larson, "A resistively degenerated wide-band passive mixer with low noise figure and +60dBm IIP2 in 0.18μm CMOS," in *Proc. of IEEE 2008 RFIC Symposium*, Jun. 2008, pp. 185–188.

[105] N. Kim, L. E. Larson, and V. Aparin, "A highly linear SAW-less CMOS receiver using a mixer with embedded TX filtering for CDMA," in *Proc. of IEEE 2008 Cust. Integ. Circuits Conference*, Sep. 2008, pp. 729–732.

[106] P. Litmanen, P. Ikalainen, and K. Halonen, "A 2.0-GHz submicron CMOS LNA and a down-conversion mixer," in *Proc. of 1998 IEEE Int'l. Symp., Circuit and Systems*, vol. 4, May 1998, pp. 357–359.

[107] A. A. Youssef and J. W. Haslett, "Low power interference-robust UWB low noise amplifier in 0.18-μm CMOS technology," in *50th Midwest Symposium on Circuits and Systems*, Aug. 2007, pp. 1006–1009.

[108] I. Nam, B. Kim, and K. Lee, "Single-ended differential amplifier and mixer circuits utilizing complementary RF characteristics of both NMOS and PMOS," in *IEEE RFIC Symp. Dig. Tech. Papers*, Jun. 2003, pp. 631–634.

[109] ——, "CMOS RF amplifier and mixer circuits utilizing complementary characteristics of parallel combined NMOS and PMOS devices," *IEEE Trans. on Microwave Theory and Techniques*, vol. 53, pp. 1662–1671, May 2005.

[110] K. Lee, "The impact of semiconductor technology scaling on CMOS radio," in *The 11th IEEE Int'l. Symp. on Electron Devices for Microwave and Optoelectronic Applications*, Nov. 2003, pp. 35–38.

[111] D. Im, H.-T. Kim, and K. Lee, "A wideband CMOS low-noise amplifier employing noise and IM2 distortion cancellation for a digital TV tuner," *IEEE J. Solid-State Circuits*, vol. 44, pp. 686–698, Mar. 2009.

[112] F. Bruccoleri, E. A. M. Klumperink, and B. Nauta, "Wide-band CMOS low-noise amplifier exploiting thermal noise canceling," *IEEE J. Solid-State Circuits*, vol. 39, pp. 275–282, Feb. 2004.

[113] S. C. Blaakmeer, E. A. M. Klumperink, D. M. W. Leenaerts, and B. Nauta, "Wideband balun-LNA with simultaneous output balancing, noise-canceling and distortion-canceling," *IEEE J. Solid-State Circuits*, vol. 43, pp. 1341–1350, Jun. 2008.

[114] P. Sivonen and A. Parssinen, "Analysis and optimization of packaged inductively degenerated common-source low-noise amplifiers with ESD protection," *IEEE J. Solid-State Circuits*, vol. 53, pp. 1304–1313, Apr. 2005.

[115] A. Khakifirooz, O. M. Nayfeh, and D. Antoniadis, "A simple semiempirical short-channel MOSFET current-voltage model continuous across all regions of operation and employing only physical parameters," *IEEE Trans. Electron Devices*, vol. 56, pp. 1674–1680, Aug. 2009.

[116] Y. Tsividis, K. Suyama, and K. Vavelidis, "A simple 'reconciliation' MOSFET model valid in all regions," *Electronics Letters*, vol. 31, pp. 506–508, Mar. 1995.

[117] S. Voinigescu, T. Dickson, M. Gordon *et al.*, "RF and Millimeter-Wave IC Design in the Nano-(Bi)CMOS Era," in *Si- based Semiconductor Components for Radio-Frequency Integrated Circuits (RF IC)*. Transworld Research Network, 2006, pp. 33–62.

[118] R. Mita, G. Palumbo, and S. Pennisi, "Design guidelines for reverse nested miller compensation in three-stage amplifiers," *Trans. on Circuits and Systems - II*, vol. 50, pp. 227–233, May 2003.

[119] E. Keehr and A. Hajimiri, "A rail-to-rail input receiver employing successive regeneration and adaptive cancellation of interference products," in *IEEE RFIC Symp. Dig. Tech. Papers*, May 2010, pp. 47–50.

[120] J. G. Proakis, *Digital Communications*, 4th ed. New York: McGraw Hill, 2001.

[121] D. D. Weiner and J. F. Spina, *Sinusoidal Analysis and Modeling of Weakly Nonlinear Circuits*. New York: Van Nostrand Reinhold, 1980.

[122] Y. Lee and S. Kim, "Median-prefiltering-based robust acquisition of direct-sequence spread-spectrum signals in wide-band pulse jamming," *IEEE Transactions on Vehicular Technology*, vol. 51, pp. 171–179, Jan. 2002.

[123] K. Kowalske and R. C. Robertson, "Performance of a noncoherent RAKE receiver and convolutional coding with ricean fading and pulse-noise jamming," in *Electronics, Circuits and Systems, 2005. ICECS 2005. 12th IEEE International Conference on*, Dec. 2005, pp. 1–4.

[124] Agilent Corp., "Agilent E8257D PSG Microwave Analog Signal Generator Data Sheet," 2009.

[125] Anritsu Corp., "MG3681 Digital Modulation Signal Generator 250kHz to 3GHz Data Sheet," 2004.

[126] ——, "MG 3690B RF/Microwave Signal Generator, 0.1Hz to 70 GHz/325 GHz Data Sheet," 2008.

[127] A. Hadjichristos, M. Cassia *et al.*, "Single-chip RF CMOS UMTS/EGSM transceiver with integrated receive diversity and GPS," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) - Slides of Presentation 6.4*, Feb. 2009.

[128] S. Hsu, S. Mathew, M. A. Anders, B. R. Zeydel, V. G. Oklobdzija, R. K. Krishnamurthy, and S. Y. Borkar, "A 110 GOPS/W 16-bit multiplier and reconfigurable PLA loop in 90-nm CMOS," *IEEE J. Solid-State Circuits*, vol. 41, pp. 256–264, Jan. 2006.

[129] R. Gregorian, *Introduction to CMOS op-amps and comparators.* New York: Wiley Interscience, 1999.

[130] A. M. ElSayed and M. I. Elmasry, "Low-phase-noise LC quadrature VCO using coupled tank resonators in a ring structure," *IEEE J. Solid-State Circuits*, vol. 36, pp. 701–705, Apr. 2001.

[131] A. Jerng and C. G. Sodini, "The impact of device type and sizing on phase noise mechanisms," *IEEE J. Solid-State Circuits*, vol. 40, pp. 360–369, Feb. 2005.

[132] S. L. J. Gierkink, S. Levantino, R. C. Frye, C. Samori, and V. Boccuzzi, "A low-phase-noise 5-GHz CMOS quadrature VCO using superharmonic coupling," *IEEE J. Solid-State Circuits*, vol. 38, pp. 1148–1154, Jul. 2003.

[133] E. Hegazi, H. Sjoland, and A. A. Abidi, "A filtering technique to lower LC oscillator phase noise," *IEEE J. Solid-State Circuits*, vol. 36, pp. 1921–1930, Dec. 2001.

[134] A. Leon-Garcia, *Probabilty and random processes for electrical engineering.* Reading, MA: Addison-Wesley, 1994.