

**DEVELOPMENT OF ROBUST ANALOG AND MIXED-SIGNAL CIRCUITS IN  
THE PRESENCE OF PROCESS-VOLTAGE-TEMPERATURE VARIATIONS**

A Dissertation

by

MARVIN OLUFEMI ONABAJO

Submitted to the Office of Graduate Studies of  
Texas A&M University  
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

May 2011

Major Subject: Electrical Engineering

Development of Robust Analog and Mixed-Signal Circuits in the Presence of Process-  
Voltage-Temperature Variations

Copyright 2011 Marvin Olufemi Onabajo

**DEVELOPMENT OF ROBUST ANALOG AND MIXED-SIGNAL CIRCUITS IN  
THE PRESENCE OF PROCESS-VOLTAGE-TEMPERATURE VARIATIONS**

A Dissertation

by

MARVIN OLUFEMI ONABAJO

Submitted to the Office of Graduate Studies of  
Texas A&M University  
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Approved by:

Chair of Committee,  
Committee Members,

Head of Department,

Jose Silva-Martinez  
Edgar Sánchez-Sinencio  
Sunil Khatri  
Duncan M. H. Walker  
Costas N. Georghiades

May 2011

Major Subject: Electrical Engineering

**ABSTRACT**

Development of Robust Analog and Mixed-Signal Circuits in the Presence of Process-Voltage-Temperature Variations. (May 2011)

Marvin Olufemi Onabajo,

B.S., The University of Texas at Arlington;

M.S., Texas A&M University

Chair of Advisory Committee: Dr. Jose Silva-Martinez

Continued improvements of transceiver systems-on-a-chip play a key role in the advancement of mobile telecommunication products as well as wireless systems in biomedical and remote sensing applications. This dissertation addresses the problems of escalating CMOS process variability and system complexity that diminish the reliability and testability of integrated systems, especially relating to the analog and mixed-signal blocks. The proposed design techniques and circuit-level attributes are aligned with current built-in testing and self-calibration trends for integrated transceivers. In this work, the main focus is on enhancing the performances of analog and mixed-signal blocks with digitally adjustable elements as well as with automatic analog tuning circuits, which are experimentally applied to conventional blocks in the receiver path in order to demonstrate the concepts.

The use of digitally controllable elements to compensate for variations is exemplified with two circuits. First, a distortion cancellation method for baseband

operational transconductance amplifiers is proposed that enables a third-order intermodulation (IM3) improvement of up to 22dB. Fabricated in a 0.13 $\mu\text{m}$  CMOS process with 1.2V supply, a transconductance-capacitor lowpass filter with the linearized amplifiers has a measured IM3 below -70dB (with 0.2V peak-to-peak input signal) and 54.5dB dynamic range over its 195MHz bandwidth. The second circuit is a 3-bit two-step quantizer with adjustable reference levels, which was designed and fabricated in 0.18 $\mu\text{m}$  CMOS technology as part of a continuous-time  $\Sigma\Delta$  analog-to-digital converter system. With 5mV resolution at a 400MHz sampling frequency, the quantizer's static power dissipation is 24mW and its die area is 0.4mm<sup>2</sup>.

An alternative to electrical power detectors is introduced by outlining a strategy for built-in testing of analog circuits with on-chip temperature sensors. Comparisons of an amplifier's measurement results at 1GHz with the measured DC voltage output of an on-chip temperature sensor show that the amplifier's power dissipation can be monitored and its 1-dB compression point can be estimated with less than 1dB error. The sensor has a tunable sensitivity up to 200mV/mW, a power detection range measured up to 16mW, and it occupies a die area of 0.012mm<sup>2</sup> in standard 0.18 $\mu\text{m}$  CMOS technology.

Finally, an analog calibration technique is discussed to lessen the mismatch between transistors in the differential high-frequency signal path of analog CMOS circuits. The proposed methodology involves auxiliary transistors that sense the existing mismatch as part of a feedback loop for error minimization. It was assessed by performing statistical Monte Carlo simulations of a differential amplifier and a double-balanced mixer designed in CMOS technologies.

## ACKNOWLEDGMENTS

I would like to express my sincere gratitude to my advisor Dr. Jose Silva-Martinez for his support, guidance, and constructive critique over the past years. I also greatly appreciate Dr. Edgar Sánchez-Sinencio's mentorship and assistance related to several research projects and to my graduate studies. Having received valuable advice from Dr. Sunil Khatri and Dr. Duncan Walker, I want to thank them for serving on my dissertation committee.

Various funding sources have made this work possible. I thank the Department of Electrical and Computer Engineering, Texas Instruments, and Broadcom for financial support. I would like to acknowledge the sponsorship of the chip fabrications by Jazz Semiconductor and United Microelectronics Corporation, as well as partial funding of the test cost by grants from TAMU-CONACYT and the National Science Foundation.

It has been a pleasure and great learning experience to collaborate on research projects with several other graduate students at Texas A&M University; namely Xiaohua Fan, Felix Fernandez, Mohamed Mobarak, Cho-Ying Lu, Venkata Gadde, Yung-Chung Lo, Vijay Periasamy, Fabian Silva-Rivas, Hsien-Pu Chen, Hemasundar Mohan Geddada, Chang Joon Park, and Aravind Kumar Padyana.

Many thanks also go out to fellow department members for helpful conversations regarding research and course projects; especially to Raghavendra Kulkarni, Jason Wardlaw, Mohamed El-Nozahi, Heng Zhang, Jusung Kim, John Mincey, Alfredo Perez, Mandar Kulkarni, Nicolas Frank, Casey Wang, Joselyn Torres, Erik Pankratz, Mohammed Mohsen Abdul-Latif, Ramy Saad, Chadi Geha, Sang Wook Park, Chinmaya

Mishra, Manisha Gambhir, Younghoon Song, and Vijay Dhanasekaran. Furthermore, I would like to thank Ella Gallagher for helping to facilitate events and completion of paperwork on many occasions.

I appreciate having had the opportunity to work together with Dr. Josep Altet from the Universitat Politècnica de Catalunya (UPC) in Barcelona, Spain; and thank him for sharing his experience related to on-chip temperature sensing during his stay at Texas A&M University. I also thank Dr. Eduardo Aldrete-Vidrio, Dr. Diego Mateo, and Didac Gómez from UPC for the collaboration related to thermal testing strategies.

In closing the acknowledgments, I am grateful for the encouragement, understanding, as well as support from my parents and brother. They have inspired me in many aspects of life, including education.

## TABLE OF CONTENTS

	Page
ABSTRACT .....	iii
ACKNOWLEDGMENTS.....	v
TABLE OF CONTENTS .....	vii
LIST OF FIGURES.....	x
LIST OF TABLES .....	xvi
<b>I. INTRODUCTION .....</b>	<b>1</b>
I.1. Background and Motivation.....	1
I.2. Research Focus and Dissertation Organization.....	4
I.2.1. Linearization scheme for transconductance amplifiers.....	7
I.2.2. Process variation-aware quantization .....	8
I.2.3. Non-invasive on-chip measurement of thermal gradients and RF power.....	9
I.2.4. Analog calibration for transistor mismatch reduction.....	10
<b>II. PROCESS VARIATION CHALLENGES AND SOLUTIONS APPROACHES.....</b>	<b>12</b>
II.1. Current Trends .....	12
II.1.1. The impact of rising process variations .....	12
II.1.2. Circuit and system design tendencies .....	14
II.2. A System Perspective on Transceiver Built-In Testing and Self-Calibration .....	18
II.2.1. Digital correction and calibration .....	19
II.2.2. Analog measurements and tuning.....	22
II.2.3. Loopback testing.....	26
II.2.4. Digital performance monitoring with analog compensation .....	28
II.2.5. Combined digital monitoring, analog measurements, and tuning .....	30
II.2.6. High-volume manufacturing testing.....	31
<b>III. HIGH-LINEARITY TRANSCONDUCTANCE AMPLIFIERS WITH DIGITAL CORRECTION CAPABILITY .....</b>	<b>34</b>
III.1. Background.....	34
III.2. Attenuation-Predistortion Linearization Methodology .....	37
III.2.1. Single-ended circuits .....	38
III.2.2. Fully-differential circuits.....	40
III.2.3. Scaling of attenuation ratios .....	42
III.2.4. Volterra series analysis.....	44



	Page
III.3. Circuit-Level Design Considerations .....	45
III.3.1. Fully-differential OTA with floating-gate FETs .....	45
III.3.2. Proof-of-concept filter realization and application considerations .....	49
III.4. Compensation for PVT Variations and High-Frequency Effects .....	53
III.5. Prototype Measurement Results .....	57
III.5.1. Standalone OTA .....	57
III.5.2. Second-order lowpass filter .....	62
III.6. Summarizing Remarks .....	68
IV. QUANTIZER DESIGN FOR A CONTINUOUS-TIME SIGMA-DELTA ADC WITH REDUCED DEVICE MATCHING REQUIREMENTS .....	69
IV.1. Background .....	69
IV.1.1. State of the art continuous-time $\Sigma\Delta$ ADCs .....	70
IV.1.2. Quantizer design trends .....	72
IV.1.3. Quantizer design considerations for the $\Sigma\Delta$ modulator architecture .....	77
IV.2. 3-Bit Two-Step Current-Mode Quantizer Architecture .....	83
IV.2.1. Quantizer design .....	83
IV.2.2. Process variations .....	91
IV.2.3. Simulation results and technology scaling .....	97
IV.2.4. ADC chip measurements with embedded quantizer .....	102
IV.3. Summarizing Remarks .....	107
V. AN ON-CHIP TEMPERATURE SENSOR TO MEASURE RF POWER DISSIPATION AND THERMAL GRADIENTS .....	109
V.1. Background .....	109
V.2. Temperature Sensing Approach .....	111
V.2.1. Integration with transceiver calibration techniques .....	111
V.2.2. Modeling of the thermal coupling .....	113
V.2.3. Electro-thermal analysis example: low-noise amplifier .....	117
V.3. CMOS Differential Temperature Sensor Design .....	122
V.3.1. Previous sensors .....	122
V.3.2. Design of the proposed sensor topology .....	123
V.3.3. Adjustment of the sensor's sensitivity .....	130
V.3.4. Sensor design optimization procedure .....	132
V.4. Measurement Results .....	135
V.4.1. Temperature sensor characterization .....	136
V.4.2. RF testing with the on-chip DC temperature sensor .....	141
V.5. Summarizing Remarks .....	146

	Page
VI. MISMATCH REDUCTION FOR TRANSISTORS IN HIGH-FREQUENCY DIFFERENTIAL ANALOG SIGNAL PATHS .....	147
VI.1. Background .....	147
VI.2. A Mismatch Reduction Technique for Differential Pair Transistors .....	148
VI.2.1. Approach .....	148
VI.2.2. Simulation results .....	154
VI.3. Second-Order Nonlinearity Enhancement for Double-Balanced Mixers.....	156
VI.3.1. Introduction .....	156
VI.3.2. Proposed mixer calibration .....	163
VI.3.3. Double-balanced mixer design.....	175
VI.3.4. Simulation results.....	180
VI.4. Summarizing Remarks .....	195
VII. SUMMARY AND CONCLUSIONS .....	197
VII.1. Overall Perspective .....	197
VII.2. Dissertation Projects .....	198
REFERENCES .....	202
APPENDIX A .....	217
APPENDIX B .....	222
APPENDIX C .....	229
APPENDIX D .....	231
VITA .....	235

## LIST OF FIGURES

		Page
Fig. 1.	Smartphone market trend.....	2
Fig. 2.	Single-chip transceiver in a cell phone. ....	4
Fig. 3.	Specification variation impact on the fraction of discarded chips.....	13
Fig. 4.	Process corner-based vs. $3\sigma$ design approaches. ....	15
Fig. 5.	Receiver with digital I/Q mismatch compensation ([14]).....	20
Fig. 6.	Analog I/Q calibration for image-rejection receivers. ....	23
Fig. 7.	BIT with analog instrumentation along the signal path. ....	25
Fig. 8.	Generalized transceiver block diagram with loopback.....	27
Fig. 9.	Transceiver with digital monitoring and tuning of analog blocks. ....	29
Fig. 10.	Transceiver with digital monitoring, analog measurements, and tuning. ....	30
Fig. 11.	Attenuation-predistortion linearization for single-ended circuits. ....	39
Fig. 12.	Attenuation-predistortion linearization for fully-differential circuits. ....	41
Fig. 13.	Low-frequency model for the attenuation-predistortion scheme.....	43
Fig. 14.	Folded-cascode OTA (implements $G_m$ in the main and auxiliary paths). ...	46
Fig. 15.	Error amplifier circuit in the CMFB loop. ....	48
Fig. 16.	$2^{\text{nd}}$ -order lowpass filter diagram and design parameters.....	49
Fig. 17.	Block diagram of the proposed automatic linearity tuning scheme.....	51
Fig. 18.	Simulated AC amplitude at the input of the main OTA (PD <sub>3</sub> in Fig. 17)....	53
Fig. 19.	Sensitivity of $ IM3 $ (in dBc) to component mismatches.....	54
Fig. 20.	Simulated sensitivity to critical component variations and mismatches. ....	56
Fig. 21.	Measured linearity with $0.2V_{p-p}$ input swing from two tones.....	58

	Page
Fig. 22.	IM3 vs. input voltage swing for reference OTA and compensated OTA. ....60
Fig. 23.	Measured IM3 dependence of the compensated OTA on phase shift. ....60
Fig. 24.	Measured filter frequency response and linearity.....63
Fig. 25.	Filter IM3 vs. frequency measured with two tones spaced by 100KHz. ....63
Fig. 26.	IM3 vs. input peak-peak voltage for the linearized filter.....64
Fig. 27.	Measured in-band intercept point curves for the filter.....65
Fig. 28.	Measured out-of-band intercept point curves for the filter. ....66
Fig. 29.	Die micrograph of the OTAs and filter in 0.13 $\mu$ m CMOS technology. ....68
Fig. 30.	Simplified diagram of a continuous-time $\Sigma\Delta$ modulator. ....70
Fig. 31.	Conventional 3-bit flash quantizer.....73
Fig. 32.	The two-step ADC principle.....75
Fig. 33.	Block diagram of the 5 <sup>th</sup> -order continuous-time modulator. ....78
Fig. 34.	Feedback path with 3-bit quantizer and PWM DAC.....80
Fig. 35.	Relative 3-bit DAC linearity error comparison: conventional vs. PWM.....81
Fig. 36.	Single-ended equivalent block diagram of the quantizer.....84
Fig. 37.	Timing of the successive quantization decisions and output code words....84
Fig. 38.	Simplified schematic of the current-mode quantizer core circuitry.....86
Fig. 39.	Simulated example of the quantization timing. ....89
Fig. 40.	Schematic of the latched comparator.....90
Fig. 41.	Latched comparator Monte Carlo simulation without device matching.....93
Fig. 42.	Latched comparator Monte Carlo simulation with device matching.....95

	Page
Fig. 43.	Quantizer core Monte Carlo simulation with device matching. ....96
Fig. 44.	Quantizer layout (0.18 $\mu$ m CMOS technology).....97
Fig. 45.	Output bit transitions with an input ramp from -200mV to 200mV. ....98
Fig. 46.	Quantizer post-layout simulations: (a) DNL (b) INL. ....99
Fig. 47.	Tuning range of the -150mV transition level (schematic simulations).....100
Fig. 48.	Die microphotograph (2.6mm <sup>2</sup> area excluding pads and ESD circuitry). .103
Fig. 49.	Measured output spectrum of the $\Sigma\Delta$ modulator. ....104
Fig. 50.	Measured SNR and SNDR vs. input signal power. ....105
Fig. 51.	Generalized receiver diagram with on-chip thermal sensing.....112
Fig. 52.	RC network model for electro-thermal coupling. ....114
Fig. 53.	Electro-thermal coupling between CUT and sensing device.....116
Fig. 54.	Area of the die with CUT (LNA) and temperature-sensing PNP device...119
Fig. 55.	Simulated average powers at devices in the CUT vs. RF input power.....120
Fig. 56.	Temperature change $T_s$ at the sensing device vs. RF input power. ....121
Fig. 57.	Transient behavior of $T_s$ with -5dBm input power. ....122
Fig. 58.	A differential CMOS temperature sensor with lateral PNP devices.....123
Fig. 59.	Proposed wide dynamic range differential temperature sensor. ....124
Fig. 60.	Simplified small-signal equivalent circuit of the sensor core. ....126
Fig. 61.	Simulated sensor sensitivity ( $\Delta I_{st1}/\Delta T$ ) vs. gain ( $A_v$ ) for amplifier $A_1$ . ....127
Fig. 62.	Amplifier ( $A_1$ ) schematic with annotated width/length dimensions.....128
Fig. 63.	Common-mode feedback (CMFB) circuit schematic. ....129

	Page
Fig. 64. Simulated dynamic range of the sensor core. ....	130
Fig. 65. Assessment of offsets in the sensor core with Monte Carlo simulations...	131
Fig. 66. Simulated $V_{be}$ mismatch of $Q_1/Q_2$ vs. ambient temperature. ....	132
Fig. 67. Combined CUT and sensor simulation. ....	134
Fig. 68. Micrograph of the chip with differential temperature sensor and LNA. ...	135
Fig. 69. Sensor output vs. power dissipation at resistor $R_t$ . ....	136
Fig. 70. Sensor output vs. power of diode-connected MOS transistors $D_{1,2}$ . ....	137
Fig. 71. Sensitivity control to power in $R_t$ and $D_{1,2}$ via $I_{core}$ adjustments. ....	138
Fig. 72. Common-mode sensitivity of the temperature sensor. ....	138
Fig. 73. Offset calibration with currents $I_{cal1}$ and $I_{cal2}$ ( $I_{core} = 500\mu A$ ). ....	139
Fig. 74. Offset calibration range with $I_{cal1}$ ( $I_{cal2} = 0$ , $I_{core} = 500\mu A$ ). ....	141
Fig. 75. Measurement vs. simulation comparison for the CUT characterization....	143
Fig. 76. LNA output power and log-magnitude of the sensor output voltage. ....	144
Fig. 77. The CUT's output spectrum from a two-tone test around 1GHz (case 1)..	145
Fig. 78. The CUT's output spectrum from a two-tone test around 1GHz (case 2)..	145
Fig. 79. An unmatched RF transistor pair. ....	149
Fig. 80. An RF transistor pair with DC mismatch reduction loop. ....	150
Fig. 81. Differential amplifier with transistor mismatch reduction loop. ....	152
Fig. 82. Operational transconductance amplifier (A) in the calibration loop. ....	152
Fig. 83. Monte Carlo simulation results (100 runs at 30°C). ....	156
Fig. 84. Double-balanced mixer. ....	158

	Page
Fig. 85. Mixer with conceptual mismatch reduction for the LO transistors. ....	164
Fig. 86. Mixer with calibration loop components. ....	166
Fig. 87. DC signal flow diagram for one calibration loop with offsets. ....	167
Fig. 88. Common-mode feedback circuit for the main calibration loop. ....	169
Fig. 89. Frequency response of the main CMFB circuit. ....	170
Fig. 90. Schematic of amplifiers $A_1$ - $A_4$ in the calibration loop. ....	171
Fig. 91. Frequency response of the amplifiers in the calibration loop. ....	172
Fig. 92. Open-loop frequency response of the calibration circuit. ....	174
Fig. 93. Detailed double-balanced mixer schematic. ....	177
Fig. 94. Common-mode feedback amplifier at the mixer output. ....	178
Fig. 95. Simulated gain and phase of the CMFB loop at the mixer output. ....	178
Fig. 96. Conversion gain vs. frequency. ....	180
Fig. 97. SSB noise figure vs. frequency. ....	181
Fig. 98. IIP3 curve. ....	181
Fig. 99. 1-dB compression curve. ....	182
Fig. 100. IIP2 curve with 0.5% mismatch between the load resistors ( $R_L$ ). ....	182
Fig. 101. Feedthrough between mixer ports. ....	183
Fig. 102. Transient simulation with a 20MHz IF output signal. ....	184
Fig. 103. Conversion gain vs. LO signal power. ....	185
Fig. 104. SSB Noise figure at IF = 1MHz vs. LO signal power. ....	185
Fig. 105. IIP2 (with 0.5% $R_L$ mismatch) and IIP3 vs. LO signal power. ....	186

	Page
Fig. 106. IIP2 comparison with 100 Monte Carlo runs.....	188
Fig. 107. Mixer with intentional threshold voltage offsets ( $\Delta V_{Th}$ ). .....	189
Fig. 108. $\overline{\Delta I_D}$ (average mismatch of $I_{D1}$ - $I_{D4}$ ) vs. $\Delta V_{Th}$ .....	190
Fig. 109. Transient settling behavior of critical control voltages.....	191
Fig. 110. Transient IF output after settling of the calibration control voltages.....	191
Fig. 111. Conversion gain comparison with 100 Monte Carlo runs. ....	192
Fig. 112. IIP3 comparison with 100 Monte Carlo runs.....	193
Fig. 113. Comparison of the SSB NF at 1MHz with 100 Monte Carlo runs. ....	193
Fig. 114. Nonlinear model for differential attenuation-predistortion cancellation. ..	217
Fig. 115. OTA model with additional nonidealities.....	222
Fig. 116. Single-ended equivalent block diagram of a bandpass biquad. ....	224
Fig. 117. Single-ended diagram of a bandpass biquad with phase compensation. ...	227
Fig. 118. BP filter simulations with different $R_s$ values for phase compensation.....	228



## LIST OF TABLES

	Page
Table I. Intra-die variability (with min. dimensions) vs. CMOS technology node...	12
Table II. Comparison of transceiver built-in testing and calibration techniques.....	33
Table III. Measured main parameters of the reference folded-cascode OTA.....	57
Table IV. Comparison of OTA linearity and noise measurements .....	61
Table V. OTA comparison with prior works .....	62
Table VI. Comparison of wideband $G_m$ -C lowpass filters .....	67
Table VII. Component parameters in the quantizer core (Fig. 38).....	101
Table VIII. Key quantizer performance parameters .....	102
Table IX. Measured $\Sigma\Delta$ ADC performance .....	105
Table X. Comparison with previously reported lowpass $\Sigma\Delta$ ADCs.....	106
Table XI. CUT design parameters and simulation results .....	116
Table XII. Simulated amplifier ( $A_1$ ) specifications.....	129
Table XIII. Measured CUT* performance parameters .....	142
Table XIV. Differential amplifier and calibration loop components.....	153
Table XV. Calibration circuitry components.....	173
Table XVI. Subthreshold mixer components.....	179
Table XVII. Simulated mixer specifications with and without calibration.....	187
Table XVIII. Down-conversion mixer performance comparison .....	194
Table XIX. Simulated comparison: OTA linearization without power increase .....	230

## I. INTRODUCTION

### I.1. Background and Motivation

As rapid progress encompasses the integration of voice, video, and internet connectivity functions into small low-power integrated circuits, portable wireless devices continue to become more prevalent in our lives to the point that many vital situations depend on the reliable operation of the integrated circuits. Consequently, there is an increasing incentive to incorporate self-test and correction features for improved reliability of wireless devices, especially in medical and military applications in which life-saving information is transmitted and received. Even though new technologies allow the design of smaller chips with more functionality, manufacturing process variability and post-production aging effects pose growing challenges for the design, fabrication, and reliability of single-chip mixed-signal systems that are realized with complementary metal-oxide-semiconductor (CMOS) technology in the modern nanometer regime. Consequently, many current research efforts are concentrated on the development of more robust analog and mixed-signal circuits by devising built-in test methodologies that enable digitally-assisted performance tuning.

On the analog circuit level, rising parameter variability is a fundamental contributor to yield and reliability problems. As a result, designing for optimum performance specifications alone is not sufficient anymore. In parallel, it has become critical to improve the on-chip measurement and self-calibration capabilities as well as the

---

This dissertation follows the style and format of the *IEEE Journal of Solid-State Circuits*.

testability of single-chip systems during high volume production testing, all in order to increase product yields and to lower the cost of testing. Both yield and cost improvement have been identified as needs in the International Technology Roadmap for Semiconductors [1], giving the incentive for novel built-in test features and alternative test strategies. Additionally, progressive on-chip self-calibration of wireless devices will help to enhance their reliability and allow full utilization of future CMOS technologies with smaller feature sizes despite of increased parameter variations.

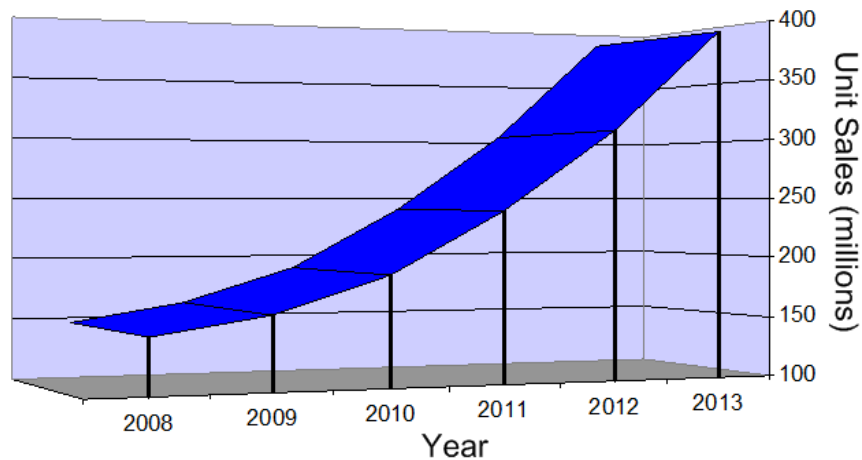


Fig. 1. Smartphone market trend.

Due to high manufacturing volumes, consumer products are a key driving force behind the development of highly integrated chips for wireless communication. For example, the projected global sales of Smartphones is plotted in Fig. 1, which is based on the data provided in [2]. The push towards mobile internet and multimedia features has led to ongoing efforts to incorporate additional functionality. At the same time,

single-chip transceivers have emerged to perform the analog signal reception and transmission operations, as well as much digital signal processing on the same chip as possible. This approach has allowed to reduce product dimensions and production cost. Nowadays, cell phones have fewer chips on the printed circuit board (Fig. 2), but the complexity of those chips causes significant design complications. In the case of integrated transceivers, the demand to support multiple communication standards has created design issues related to more stringent linearity requirements for the broadband radio frequency (RF) front-end circuits, reconfigurability of many blocks along the transmit/receive chains, interference avoidance among circuits, minimization of total power consumption, and other aspects. Within the scope this dissertation is that RF system performance monitoring is becoming significantly more important and difficult with the trend towards increasing integration and power densities in single-chip systems fabricated with modern CMOS technologies. On-chip electrical power detectors are commonly used to monitor and optimize the dynamic range of RF systems through measurements and controlled amplifications in RF front-ends. However, the adverse effects from parasitic input capacitances of electrical detectors become more detrimental at higher frequencies. Non-invasive temperature sensors for RF power detection offer an attractive alternative to conventional power detectors, as shown by the investigations presented in this dissertation.

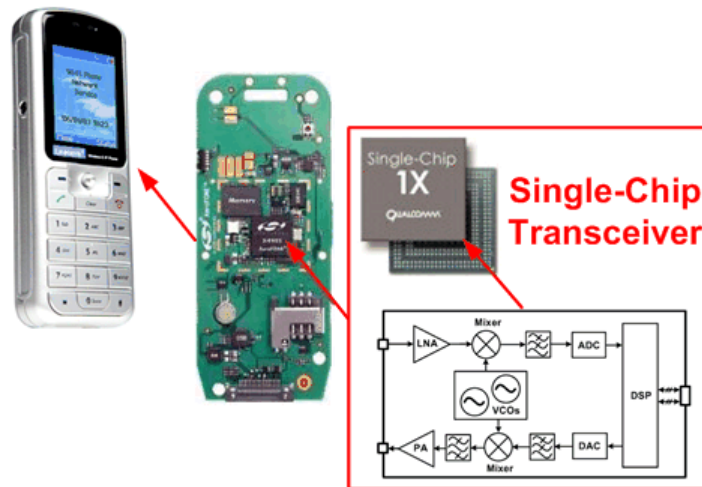


Fig. 2. Single-chip transceiver in a cell phone.

## I.2. Research Focus and Dissertation Organization

Contemporary CMOS technologies have offered progress with respect to circuit properties such as smaller device dimensions, better high-frequency operation, and power efficiency. But, analog designers in particular face various technology-related drawbacks associated with newer technologies, for example signal swing limitations due to decreased supply voltage and gain reduction due to lower transistor output resistance. Other major disadvantages, which are elaborated in Section II, are worsening process variations and intra-die device mismatches. These have a strong impact on the product yield and reliability, translating into manufacturing cost and risk factors in critical medical or military applications. Variations and circuit sensitivity to environmental conditions such as temperature changes and interference from other nearby circuits are becoming more problematic as the complexity of integrated systems increases. In this dissertation, special attention is given to augmentations of analog and mixed-signal

circuits in response to the emerging variability problems and system-level calibration approaches concerning current and future CMOS technologies.

An intricate issue is the high number of possible failure causes for analog circuits as a result of the random nature of process variation, ambient temperature changes, and interference signals. Typically, it is insufficient to monitor a single quantity and extract the necessary information to determine the severity of faults or the actions to be taken for their correction. For instance, measurement of an RF circuit's quiescent current can be helpful to identify gross defects, but has very limited usefulness when the goal is to tune RF metrics such as gain or linearity parameters. This creates a need for continuous expansion of on-chip measurement capabilities, especially because the acceptability of an analog circuit's performance normally depends on many parameters that can take on a continuous range of values. Moreover, the integration of more functionality and transistors into integrated systems leads to higher power densities on the chips, which leads to more pronounced temperature gradients and interference between circuits due to thermal coupling. A temperature sensing strategy is introduced in Section V to provide alternative means for on-chip measurements of RF characteristics and to increase the observability of temperature gradients. The section also contains descriptions of the proposed temperature sensor topology for built-in testing of analog circuits and the simulation methodology for its design.

A digital circuit whose functionality has been verified during the characterization test phase will predominantly be affected by process variation of the transition frequency and threshold voltage, which will have main effects on the maximum speed of operation

and power consumption. This eases the determination of performance limits for digital circuits by verifying their logic outputs or the output of test structures at the mandated speed. As alternative for test cost reductions or performance optimizations, local process monitors can be embedded in the layout design to measure the transition frequency or threshold voltages (as representatives for areas of a partitioned die), and to compensate for variations by adjusting nearby digital circuits through features such as adaptive body bias or supply voltage. Such systematic approaches have become increasingly popular to deal with variability in digital circuits, but they are less effective for analog circuits because their performance depends on more parameters and each analog block has a different dependence on a given parameter. For that reason, the design strategies for robust analog circuits tend to be tailored to the circuit type or even its specific topology.

The approach taken in this dissertation is to present examples of circuits and their features that alleviate the effects of process variations. With adaptations, the presented methodologies can be extended to similar analog circuits. In particular, the use of digitally programmable circuit elements or bias conditions will be emphasized and related to the compatibility of individual blocks with emerging system-level self-calibration strategies. The first example to be discussed in Section III is the linearization of transconductance amplifiers in broadband filter applications. Section IV describes another case study, which is a 3-bit quantizer that was designed for continuous-time  $\Sigma\Delta$  analog-to-digital converters. Section V introduces a strategy to utilize differential temperature sensors as on-chip RF power detectors for built-in testing. Next, a general technique to reduce the mismatch between transistors is proposed in Section VI, in

which it is applied to differential pair transistors of a wide bandwidth amplifier and the switching transistors of a double-balanced mixer. To finish, Section VII summarizes the contributions of this work. The following subsections give a more detailed overview of the focal points in this dissertation.

### *1.2.1. Linearization scheme for transconductance amplifiers*

Operational transconductance amplifiers (OTAs) are elements of transconductance-capacitor ( $G_m$ -C) filters in many wireless receivers and continuous-time  $\Sigma\Delta$  analog-to-digital converters. Thus, OTA performance and dependability improvements manifest themselves in system-level enhancements of communication circuits and sensor signal conditioning circuits. The push towards wider bandwidths in these applications mandates OTA designs with progressively better linearity at higher frequencies. Towards this end, an architectural solution is presented in Section III that can be applied to diverse circuit-level OTA configurations. Effective linearization over a wide frequency range demands a mechanism to correct for high-frequency effects and process variations. Accordingly, digital programmability was realized to ensure high linearity and compatibility with modern CMOS technologies.

The linearization technique utilizes two matched OTAs to cancel output harmonic distortion components, creating a robust architecture. Compensation for process variations and frequency-dependent distortion based on Volterra series analysis is achieved by employing a delay equalization scheme with on-chip programmable resistors. An OTA design with the proposed broadband linearization method has third-order inter-modulation (IM3) distortion better than -74dB up to 350MHz with  $0.2V_{p-p}$



input, 70dB signal-to-noise ratio (SNR) in 1MHz bandwidth, and 5.2mW power consumption. The distortion-cancellation technique enables an IM3 improvement of up to 22dB compared to a commensurate OTA without linearization. A proof-of-concept lowpass filter with the linearized OTAs has a measured IM3 < -70dB and 54.5dB dynamic range over its 195MHz bandwidth. The standalone OTAs and the filter were fabricated on a 0.13 $\mu$ m CMOS test chip with 1.2V supply.

### *1.2.2. Process variation-aware quantization*

Future wireless devices will require extensive connectivity to accommodate several services, which means that the receivers must cover broader frequency bands. Therefore, on-chip analog-to-digital converters (ADCs) in multi-standard receivers not only demand increased signal-to-quantization-noise-ratio, but also more bandwidth for the conversion of the analog signal into the digital domain. Our research group developed a lowpass continuous-time  $\Sigma\Delta$  ADC for next generation broadband receiver applications using a 0.18 $\mu$ m CMOS process. Rather than using multiple signal levels, a multi-bit digital-to-analog converter (DAC) realization based on a feedback signal with time-varying pulse duration was employed. This approach alleviates nonlinearity problems associated with typical multi-bit DACs. Section IV of this dissertation describes the corresponding 3-bit quantizer architecture with multi-phase clocking. The reference levels for the quantizer are adjustable to compensate for process variations after fabrication if the application necessitates fine resolution. Designed with 5mV resolution at a 400MHz sampling frequency, the quantizer power dissipation is 24mW and its die area with auxiliary logic circuitry and routing is 0.4mm<sup>2</sup>. With embedded quantizer, the

5<sup>th</sup>-order  $\Sigma\Delta$  ADC achieves a measured peak SNDR of 67.7dB in 25MHz bandwidth, consumes a total of 48mW with a 1.8V supply, and occupies 2.6mm<sup>2</sup> die area.

### *1.2.3. Non-invasive on-chip measurement of thermal gradients and RF power*

One aspect of designing robust analog and mixed-signal circuits in wireless products is the inclusion of on-chip monitors that can determine whether device performance parameters are within an acceptable range or whether a detrimental shift has occurred due to effects from aging, temperature variations, interfering signals, or other conditions. This information can then be incorporated into self-calibration schemes that tune circuit blocks to restore satisfactory functionality. A part of this dissertation work is directed towards the conception of a practical monitoring strategy employing differential temperature sensors with high sensitivity and accuracy for measuring on-chip temperature gradients over the range of interest. Due to thermal coupling, the temperature in the vicinity of a device depends on its power dissipation, and this relation can be exploited for testing purposes [3].

In Section V, a design methodology is presented which aims at the extraction of RF circuit performance characteristics from the DC output of an on-chip temperature sensor. Any RF input signal can be applied to excite the circuit under examination because only dissipated power levels are measured, which makes this approach attractive for online thermal monitoring and built-in test scenarios. A fully-differential sensor topology is introduced that has been specifically designed for the proposed method by constructing it with a wide dynamic range, programmable sensitivity to DC and RF power dissipation, as well as compatibility with CMOS technology. Furthermore, a procedure

is outlined to model the local electro-thermal coupling between heat sources and the sensor, which is used to define the temperature sensor's specifications as well as to predict the thermal signature of the circuit under test.

A prototype chip with an RF amplifier and temperature sensor was fabricated in a conventional  $0.18\mu\text{m}$  CMOS technology. The proposed concepts were validated by correlating RF measurements at 1GHz with the measured DC voltage output of the on-chip sensor and the simulation results, demonstrating that the RF power dissipation can be monitored and the 1-dB compression point can be estimated with less than 1dB error. The sensor circuitry occupies a die area of  $0.012\text{mm}^2$ , which can be shared when several on-chip locations are observed by placement of multiple  $11\mu\text{m} \times 11\mu\text{m}$  temperature-sensing devices.

#### *1.2.4. Analog calibration for transistor mismatch reduction*

An analog calibration technique is presented to lessen the mismatch between transistors in the differential high-frequency signal path of analog CMOS circuits. It can be applied for offset reduction in high-speed amplifiers and comparators in which short-channel devices are utilized to minimize bandwidth reduction from parasitic capacitances. In general, this approach is suitable for RF applications in which direct matching of the transistors is undesired because sophisticated layout practices would increase the coupling between the high-frequency paths. The proposed methodology involves auxiliary devices that sense the existing mismatch as part of a feedback loop for error minimization. This technique is demonstrated in Section VI with a differential amplifier having a loaded gain and -3dB frequency of 13dB and 2.14GHz. It was

designed in 90nm CMOS technology with a 1.2V supply. Monte Carlo simulations indicate that the 4.17mV standard deviation of the amplifier's anticipated input-referred offset voltage improves to 0.76mV-1.29mV with the mismatch reduction loop, which is contingent on the layout configuration of the mismatch-sensing transistors.

Section VI also provides a second application example for the analog mismatch reduction loop, which is to enhance the matching between the switching transistors in a double-balanced CMOS mixer. Simulation results show that this scheme improves the mixer's IIP2 by 5dB while having negligible impact on other performance parameters with the exception of 30% higher power due to the dissipation in the calibration circuitry. The calibration method helps to compensate for the large process variations of the mixer transistors that are biased with small currents in the subthreshold region. As a result, the power consumption of the presented mixer is still more than six times lower than that of conventional down-conversion mixers using saturation region bias, whereas its specifications are similar to the state of the art.

## II. PROCESS VARIATION CHALLENGES AND SOLUTIONS APPROACHES

### II.1. Current Trends

#### II.1.1. The impact of rising process variations

Most semiconductor product improvements over the past decades are direct or indirect consequences of the perpetual shrinking of devices and circuits, allowing performance enhancements at lower fabrication cost. A paralleling trend is that process variations and intra-die variability increase with each technology node. Since most high-performance analog circuits depend on matched devices and differential signal paths, this trend has begun to diminish yields and reliabilities of chip designs. Fundamentally, the problem is that parameters of devices on the same die show increasing intra-die variations, thereby exhibiting different characteristics. For example, Table I displays the evolution of the typical transistor threshold voltage standard deviation  $\sigma\{V_{Th}\}$  normalized by the threshold voltage ( $V_{Th}$ ) for several technologies, as reported in [4]. Also notice that  $V_{Th}$  exhibits further dependence on gate length variations through the drain-induced-barrier-lowering (DIBL) effect under large drain-source voltage bias conditions, as demonstrated by the characterization in [5] using 65nm technology. Since DIBL worsens as the channel is scaled down, this additional impact on  $V_{Th}$  variations can be assumed to be even stronger beyond the 65nm technology node.

Table I. Intra-die variability (with min. dimensions) vs. CMOS technology node

Technology Node	250nm	180nm	130nm	90nm	65nm	45nm
$\sigma\{V_{Th}\}/V_{Th}$	4.7%	5.8%	8.2%	9.3%	10.7%	16%

A direct consequence of device parameter variations is a decrease in production yields because block-level and system-level parameters will show a corresponding increase in variations. This relationship between variations and yield can be inferred from the visualization in Fig. 3, where the Gaussian distribution of a specification with a standard deviation  $\sigma$  around the mean value  $\mu$  is shown together with the specification limits ( $\pm 3\sigma$  in this example). For standalone analog circuits, parameters such as gain may have an upper and/or lower specification limit, and the samples that exceed the limit(s) during production testing must be discarded. Guardbands are often defined to account for measurement uncertainties by following procedures such as repeating the same test or performing other more comprehensive tests to determine whether the part can be sold to customers, which incurs additional test cost in a manufacturing environment.

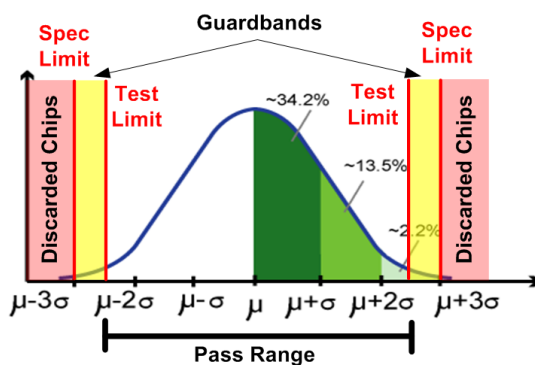


Fig. 3. Specification variation impact on the fraction of discarded chips.

An important observation from Fig. 3 is that an increase of variation ( $\sigma$ ) widens the Gaussian distribution, which leads to a higher percentage of parts that fall within the highlighted ranges that require them to be scrapped or retested. Clearly, there is a direct

relationship between the amount of process variations and production cost due to low yields. In the case of wireless mixed-signal integrated systems, the trend towards increasing integration and complexity has also been paralleled by technical challenges and rising cost of testing, which can amount up to 40-50% of the total manufacturing cost [6], [7]. As a consequence, built-in self-test, design-for-test, and design-for-manufacturability methods for analog and mixed-signal circuits have received growing attention over the past years.

### *II.1.2. Circuit and system design tendencies*

System complexities and process variations raise the importance of considering testability early in the design phase to avoid technical complications and time-to-market delays in the pre-production phase as well as test cost reduction during the production phase. Worst-case process corner models have been used extensively to account for variations during the design of analog circuits. But more recently, a paradigm shift towards the use of statistical models and Monte Carlo simulations has occurred. One of the main reasons for this development is that corner-based design easily results in too pessimistic designs [8], which is evident in Fig. 4. In this figure, the x-axis and y-axis represent the ranges over which two parameters can vary, and the area inside the ellipse indicates the combined range in which the  $3\sigma$  limits are met. This region can be predicted with statistical Monte Carlo simulations for yield estimation. On the other hand, the area outside of the elliptical design space corresponds to design implementations that meet the specifications, but are overdesigned. This means that “investments” of area, power, or trade-offs with other parameters are made in order to

allow acceptable performance despite of increased deviations of the two parameters from their nominal values. The rectangular region between the combination of the four worst corner cases of the two parameters includes overdesign space, implying that it involves costly performance or parameter trade-offs. This economic reason and the availability of more efficient computational tools have created a trend towards statistical yield optimizations rather than corner-based design [8].

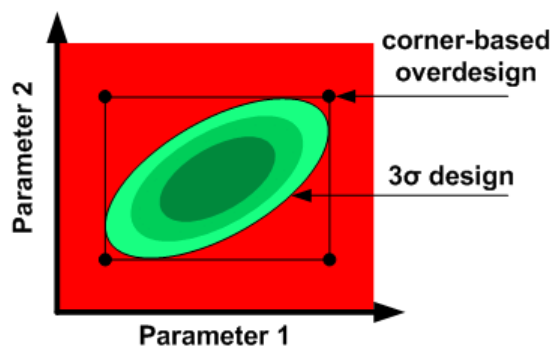


Fig. 4. Process corner-based vs.  $3\sigma$  design approaches.

Defect densities on wafers become worse in newer technologies and production yields decrease with increased chip size [9]. Self-test and self-repair schemes for digital circuits have been routinely incorporated into products for a long time, especially since on-chip verification of logic blocks and repair with redundant circuitry do not require analog instrumentation resources. The inclusion of scan chains gives easy access to internal digital circuitry through a minimal number of pins during production testing. Similarly, the standardized mixed-signal test bus (IEEE Std. 1149.4) has been developed to improve the testability of analog blocks by allowing better observation of internal



nodes. Nowadays, the use of analog test buses within single-chip systems is feasible in the industry, but significant design considerations are required to avoid that the interface circuitry does not affect the integrity of the analog signals or measurements [10].

In addition to the underlying variation and defect issues on the device level, several system-level and technology trends impair the testability and manufacturability of integrated circuits for mobile applications:

#### Support of multiple communication standards and more features on low-power chips

The wireless communication industry has experienced phenomenal growth in the past decade that resulted in low-power handheld devices with multi-purpose functionality such as video, voice, pictures, and internet access. The wireless local-area networks for laptops, desktops, and personal digital assistants (PDAs) include standards like Bluetooth, WiFi, IEEE 802.16, WiMAX, Ultra-Wideband (UWB), and GPS. Most relevant services for handheld devices range from 470 MHz to almost 11GHz. The main technical challenge is the co-existence of wireless devices, which results in signal interference. This can be solved if more linear high-performance analog receiver front-ends are available to tolerate and filter out high-power interfering signals without saturation of the analog blocks due to excessive signal power levels. Further filtering and channel selection can be performed in the digital domain when the signal integrity is maintained by the processing through unsaturated highly-linear analog blocks. Support of multiple communication standards requires chips with more circuitry and complexity, which makes them less testable in the production stage because of limited access to internal nodes, interactions between blocks, and a higher number of test cases to verify

functionality. Systems with more channels are more likely to fail, which is another reason why yields of integrated receivers, transmitters, and transceivers are on the decline. Simultaneously, the processing of broadband signals in their front-ends mandates high-performance analog circuits, which in many cases requires continued circuit-level innovations for on-chip self-calibration to tune for optimum performance.

#### Process technology optimizations for digital circuits create analog design challenges

The main advantages of device scaling with CMOS technology are improved performance at higher frequencies, reduced power consumption, and increased levels of integration. Those benefits are particularly aiding the development of digital circuits and systems. With regards to analog circuits, deep-submicron technology scaling progress comes together with adverse effects such as reduced gains from lower transistor output impedances, design with limited voltage headroom, higher flicker noise levels, and reduced transistor linearity. Larger variability of parameters is caused by physical and fabrication limitations such as under-etching uncertainties, variations of effective transistor dimensions, severe channel length modulation due to higher electric fields, and channel dopant fluctuations. Interestingly, the random dopant fluctuations have reached a severity that can lead to threshold voltage mismatch in neighboring devices at the 65nm node [11]. Additional reliability concerns arise from the restricted power that transistors can supply to the load without exceeding the low breakdown voltage of the deep submicron devices. Furthermore, digital CMOS processes often do not provide high-quality passive devices required for conventional high-performance analog designs. For example, metal-insulator-metal (MIM) capacitors, high-resistivity polysilicon

resistors, or well-characterized inductor models might not be available in a digital process, forcing analog designers to get by with metal-oxide-semiconductor (MOS) capacitors and standard polysilicon resistors. Both of these have higher parasitic capacitance to the substrate than the equal-valued MIM capacitors or high-resistivity polysilicon resistors. Scaling down transistors permits more digital functionality and memory on a single chip, but with less reliability especially for analog signal processing.

## **II.2. A System Perspective on Transceiver Built-In Testing and Self-Calibration**

The concepts and examples presented in this dissertation are all involving circuit blocks which are found in conventional transceivers within mobile wireless devices. While equipping the circuit blocks with built-in test (BIT) and self-calibration features to compensate for variations, it is important to keep their role as part of the system in mind because of the interaction between blocks and the overall goal to optimize system-level performance specifications such as bit error rate (BER) or error vector magnitude (EVM). In general, the self-calibration challenge can be divided into two parts: one is to add tunability and controllability capabilities in the individual blocks, and the other one is to devise comprehensive system-level calibration algorithms in a digital signal processing unit. The former task is the focus of this dissertation, but the existing approaches for the latter task will be briefly discussed in the remainder of this section and when applicable throughout the dissertation.

BIT strategies for transceivers vary tremendously depending on the transceiver architecture, communication standard, available on-chip measurement and computation resources, the production volume, and whether the BIT is designed for production

testing (quality control) or on-line self-calibration (reliability) during the life time of the chip. Consequently, most BITs involve a mix of analog and digital blocks, on-chip and off-chip measurement devices, long calibration routines at start-up, and shorter periodic or on-line calibration. Generally, a trend has emerged to combine techniques for verification of complex mixed-signal transceivers implemented as single chips. Nevertheless, the BIT approaches can be grouped into a few rough high-level categories that represent the different design philosophies in academia and the industry. In the following overview, a few example cases will be discussed to highlight the distinctive characteristics of methods that can be broadly classified into these categories:

- Digital correction and calibration
- Analog measurements and tuning
- Loopback testing
- Combined digital performance monitoring and analog compensation
- Combined digital monitoring, analog measurements, and analog compensation

### *II.2.1. Digital correction and calibration*

Digital BIT approaches involve measurements and compensation techniques that are realized in the digital baseband processor of the transceiver. They are suitable for parameters that are observable and traceable in the digital domain, such as slowly drifting DC offsets or mismatch between the in-phase (I) and quadrature-phase (Q) paths in the front-end. Generally, digital methods have the advantage of high precision when sufficient computational resources are available. They are also very attractive for on-line calibration schemes that run in the background.

Digital I/Q mismatch compensation is a widely used method that involves digital measurement and compensation of the I/Q gain and phase mismatches in the analog front-end circuitry. For example, the work in [12] presents a scheme that runs during start-up or in a dedicated calibration mode to ensure acceptable performance of a low-IF receiver even with up to 10% gain and  $10^\circ$  phase imbalance in the analog front-end. On-line digital I/Q compensation techniques have also been reported, such as [13], in which the training symbols that are standard in orthogonal frequency-division multiplexing (OFDM) transmissions are exploited for background I/Q calibration. It was also demonstrated in [13] how digital I/Q compensation relaxes the overall signal-to-noise ratio (SNR) requirements in the receiver chain because I/Q imbalance directly affects the SNR and thereby degrades the bit error rate (BER). In the OFDM receiver example presented in [13], the digital calibration allowed to improve the tolerance to I/Q imbalances from 1%-gain/ $1^\circ$ -phase to 10%-gain/ $10^\circ$ -phase.

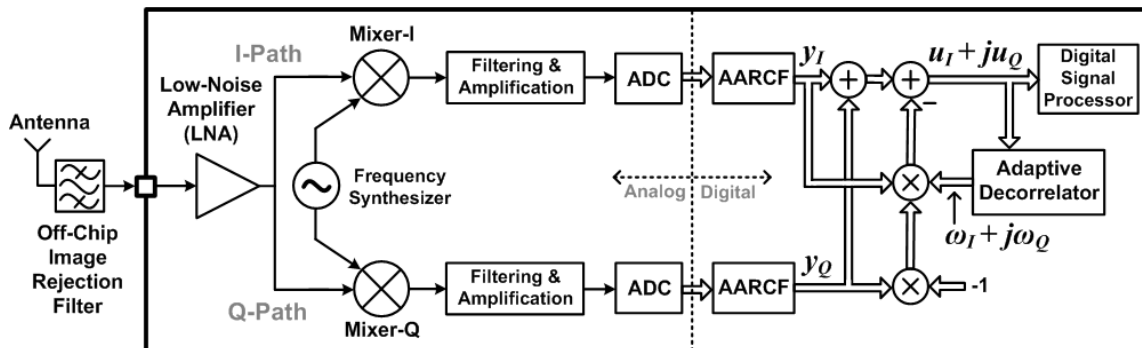


Fig. 5. Receiver with digital I/Q mismatch compensation ([14]).

Digital I/Q calibration is widely used in the industry. An example is the work from Texas Instruments describing a low-IF GSM receiver in 90nm CMOS technology [14]. This receiver utilizes an adaptive filter that obtains the mismatch information from on-line I/Q correlations, for which the modified block diagram from [14] is displayed in Fig. 5. The interesting part of the block diagram is the adaptive decorrelator after the analog-to-digital converter (ADC) and anti-aliasing rate change filter (AARCF). In the digital domain, gain mismatch appears as difference in the auto-correlation between I and Q paths, while phase mismatch appears as nonzero cross-correlation between I and Q. The authors use an algorithm that takes advantage of the aforementioned relationships by implementing an adaptive decorrelator which attempts to minimize the auto-correlation and the cross-correlation between I and Q outputs ( $y_I, y_Q$ ). This is done by adjusting the correction coefficients:

$$\omega_{I(n+1)} = \omega_{I(n)} + \mu \cdot [u_{I(n)} \cdot u_{I(n)} - u_{Q(n)} \cdot u_{Q(n)}] \text{ and } \omega_{Q(n+1)} = \omega_{Q(n)} + 2\mu \cdot u_{I(n)} \cdot u_{Q(n)}, \quad (1)$$

where  $\mu$  is the adaptation step size which is inversely proportional to the signal energy. Thus, periodic training sequences are required with this scheme. Depending on process-voltage-temperature (PVT) variations, 15-30dB image rejection ratio (IRR) improvement has been demonstrated in practice with phase mismatch  $< 1^\circ$  and amplitude mismatch  $< 10\%$  in [14] with a settling time in the range of 3-4 milliseconds. This settling time is lengthy compared to analog tuning approaches that can be as short as a few microseconds [15], which becomes important in production testing situations because any adjustments for different test conditions in the front-end (different gain settings, channel, etc.) would require 3-4ms idle time for digital I/Q calibration before

the BER test can begin. On the other hand, settling times of analog tuning schemes depend on the loop bandwidth, which can be designed in the megahertz range to achieve settling times in the microseconds regime. Hence, analog I/Q tuning approaches would fill the niche of situations that require fast convergence.

The incentive for using a digital BIT technique is high when the circuit under test itself has digital features. An example is the BIT of a transmitter in [16] that includes an all-digital phase-locked loop (ADPLL). In that case, the error signal of the ADPLL is already in the digital domain, allowing to monitor failures and the center frequency drift of the digitally controlled oscillator. Furthermore, the authors of [16] state that digital filtering and spectral estimation can be used to monitor and adjust the phase noise transfer function.

### *II.2.2. Analog measurements and tuning*

The analog equivalent to the digital I/Q imbalance calibration scheme has been proposed and demonstrated for image-reject receiver (IRRX) architectures [17]. A simplified block diagram of such a BIT is displayed in Fig. 6, which is representing the work from [17]. In an IRRX, the down-conversion scheme with two mixing stages and lowpass filters suppresses the image signal at the second intermediate frequency output  $\text{Out}(f_{IF2})$ , which avoids the need for an external image-rejection filter. The quality of the image-rejection is typically expressed with the image-rejection ratio (IRR) that depends on the I/Q amplitude mismatch ( $\Delta A$ ) and phase mismatch ( $\Delta\theta$ ):

$$IRR_{(dB)} \approx 10 \cdot \log\left( (1/4) \cdot [(\Delta\theta)^2 + (\Delta A / A)^2] \right). \quad (2)$$

In practice, the IRR is normally limited to 25dB-40dB due to mismatches, even though almost 60dB are required for acceptable BER performance. In [17], a purely analog calibration scheme was implemented with the auxiliary path shown in Fig. 6. This path contains the duplicate mixing operations as in the main path with the exception that the output signal at the second intermediate frequency ( $f_{IF2}$ ) can be of the form  $\cos(2\pi \cdot f_{IF2} \cdot t)$  or  $\sin(2\pi \cdot f_{IF2} \cdot t)$ , depending on which phases of the two local oscillators ( $LO_1$ ,  $LO_2$ ) are routed to the auxiliary mixers. Finally, mixer<sub>3</sub> correlates the signals from the two paths to extract the I/Q mismatch information contained in the DC component after the lowpass filter (LPF). This analog DC voltage ( $V_{cal}$ ) can be directly used to tune the bias voltages of analog circuits for mismatch compensation, resulting in high IRR (e.g. 57dB in [17]). A similar automatic IRR calibration with analog mixers, variable phase shifter, and gain tuning has been realized in [18], achieving an IRR of 59dB.

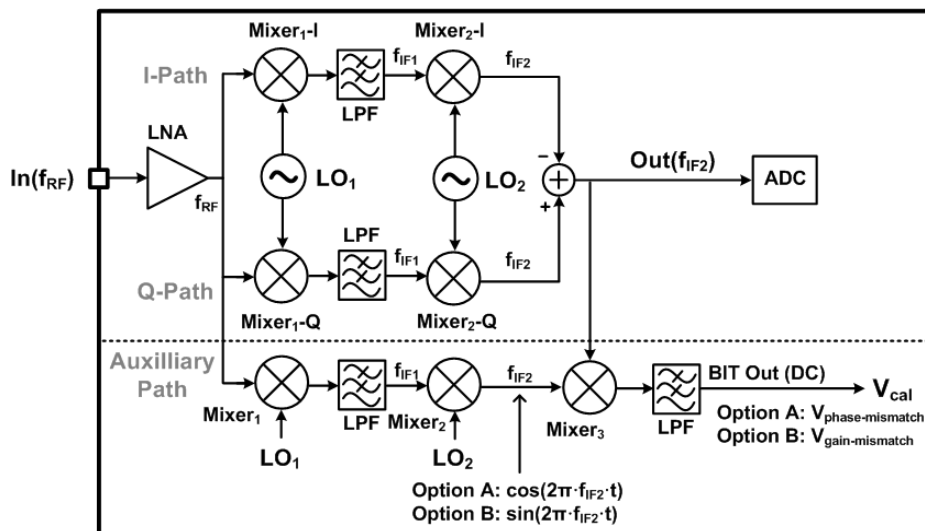


Fig. 6. Analog I/Q calibration for image-rejection receivers.



A benefit with local analog tuning is that the bias conditions of the analog blocks under calibration are controlled and less affected by PVT variations due to the correcting action of the local loops, thereby allowing higher yields as a result of automatic correction in the analog front-end. However, the power and area consumption of the BIT circuitry is the main trade-off. Furthermore, the BIT circuits themselves have to be designed robustly to avoid failures, making the implementation more challenging and invasive than digital schemes. Efforts for the analog approach are generally more justified in transceivers that have limited on-chip digital resources and in scenarios that require fast automatic correction. For example, the IRR calibration in [18] can be used on-line with a settling time that depends on the bandwidth of the analog control loops rather than convergence of digital algorithms that take several milliseconds as in [14]. Another fast analog calibration method with a convergence time in the microseconds regime is described in [15].

Instead of using a system-level test strategy, it has been very popular to extract information from each block in the analog front-end for characterization or tuning of the block, which is visualized in Fig. 7. The circuit under test (CUT) represents a block in the RF front-end or analog baseband that can be connected to a BIT circuit in test mode by closing the two switches  $S_1$  and  $S_2$ . In [19] for instance, a low-noise amplifier (LNA) was tested with a BIT block containing a test amplifier and two power detectors to measure input impedance, gain, noise figure, input return loss, and output SNR. This approach has the advantage that the fault location/cause can be identified clearly and that the DC or digital outputs of the BIT circuits can be used to recover from certain failure

modes. High-frequency RF front-ends have been targeted in particular with dedicated design of BIT circuits because gain, impedance matching, and linearity performances are very sensitive to variations. Also, direct signal digitization is not feasible at high frequencies, eliminating many digital compensation schemes. Hence, several RF block-level measurement approaches involve power or amplitude detectors along the signal path [20]-[23].

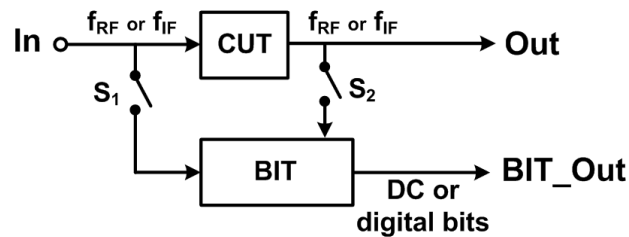


Fig. 7. BIT with analog instrumentation along the signal path.

Self-calibration of impedance matching for an LNA at the input of the receiver chain as done in [24] also requires on-chip analog sensing circuitry, especially to achieve a short calibration time such as the  $30\mu\text{s}$  reported in [24]. An alternative proposition to monitor individual blocks in the signal path was made in [25], in which the transient supply currents of the CUTs are monitored with the BIT circuitry by placing small series resistors in the power supply lines. However, a clear disadvantage with any block-level measurement is that the BIT circuitry is connected to the CUT and therefore must be designed carefully to avoid impact on block or system performance. But, some degradation due to loading effects from BIT circuitry must usually be tolerated.

Furthermore, switches in or along the signal path are undesired due to their losses and signal feedthrough due to finite isolation, particularly at RF frequencies.

Though with less accuracy than off-chip measurement equipment, efforts have also been made to mimic conventional instrumentation such as spectrum analyzers ([26], [27]) on the chip with sufficient accuracy for BIT applications. In [26] for example, the analyzer with a frequency range of 33MHz to 3GHz could cover the entire signal paths of many wireless transceivers in handheld consumer products. A multiplexor could be used to selectively route a test input at a time to one spectrum analyzer, but the on-chip measurement circuitry still takes up large area and significant power that might not be permissible in certain applications. For example the analyzer in [26] consumes  $0.384\text{mm}^2$  and more than 20mW.

### *II.2.3. Loopback testing*

Loopback testing is a system-level BIT technique in which the BER is monitored in the digital baseband [28]. It allows simultaneous verification of the analog and digital transceiver blocks (Fig. 8) with a low-frequency digital input signal applied to the baseband subsection of the transmitter. This up-converted signal is routed from the transmitter (TX) output to the receiver (RX) input via a loopback connection [29]. After down-conversion and digitization in the RX, the received bitstream is analyzed in the digital baseband processor to determine the BER. Attenuation and frequency translation with a mixer are required in the loopback block to maintain signal integrity and to ensure that the power levels during testing are comparable to normal operation. If the communication standard does not require frequency translation between TX and RX,

then only the RF attenuator is required. In any case, the overhead of the BIT circuitry is below 10% of the complete transceiver, which is efficient. However, the loopback BIT cannot be executed on-line; it requires a dedicated test mode during production testing or self-checks during times when the transceiver is idle.

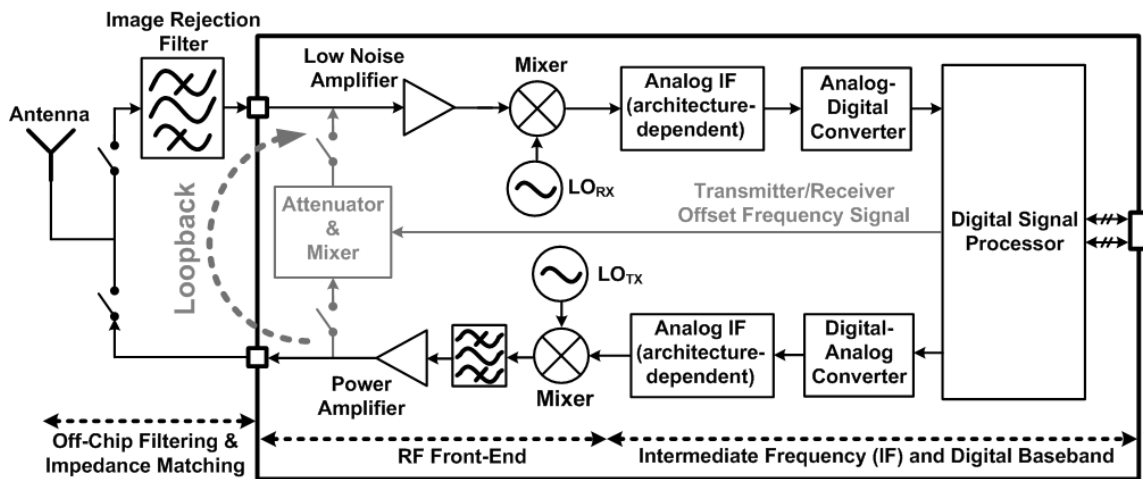


Fig. 8. Generalized transceiver block diagram with loopback.

The main benefit of the loopback technique is that a BER test is the most important metric, which is only low when all components function properly. This property makes loopback very attractive for fast pass/fail production testing and quick self-checks during in-field use, especially when few or no off-chip test resources are available. For example, a loopback test for the on-wafer production test stage was presented in [30].

A drawback of early loopback implementations is the lack of information regarding failure causes and fault locations. In response, one proposed variant [31] involves more computations in the digital baseband processor to determine the spectral content of the received bits and to use the data for estimation of receiver/transmitter nonlinearity

specifications. Alternatively, power detectors could be placed at critical nodes to extract block-level gain and 1dB-compression point measurements. Or, similarly, statistical sampling blocks were placed along the signal path in [32]. These blocks produce digital bitstreams for analysis of fault locations. In general, inclusion of auxiliary circuitry during a loopback test increases the observability of faults, but with the associated trade-offs that have been discussed for on-chip measurement circuitry in Section II.2.2.

#### *II.2.4. Digital performance monitoring with analog compensation*

A BIT approach for complex transceiver chips that has become increasingly popular in recent years is depicted in Fig. 9. It incorporates accurate digital monitoring and I/Q mismatch correction in the baseband processors as well as a few analog observables such as outputs from received signal strength indicators (RSSIs) or DC control voltages of blocks that give some insights into their operating conditions. A significant aspect is that many analog bias voltages for RF front-end and baseband circuits are generated with digital-to-analog converters (DACs). These DACs are utilized for coarse adjustments at start-up in order to compensate for PVT variations. They also reduce DC offsets in the analog circuits to prevent saturation of internal nodes due to large gains in the receiver. Thus, more mismatches can be tolerated because of the capability to counteract them.

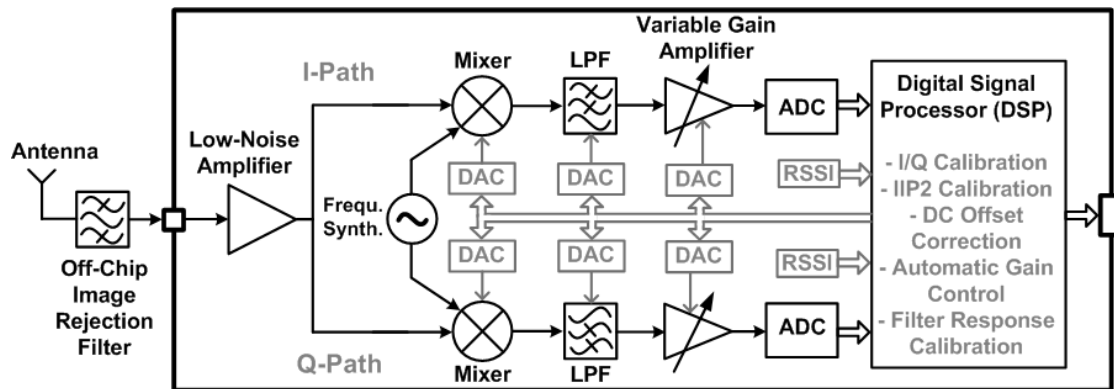


Fig. 9. Transceiver with digital monitoring and tuning of analog blocks.

Combined digital monitoring/calibration with analog compensation DACs has been reported in publications describing industrial transceivers. Some examples are:

- Single-chip GSM/WCDMA transceiver in 90nm CMOS [33], (Freescale, 2009)
  - DC offset, I/Q gain & phase, IIP2 calibration in the digital signal processor
  - 6-bit DACs for analog compensation
- 2.4GHz Bluetooth Radio in 0.35 $\mu$ m CMOS [34], (Broadcom, 2005)
  - Bias networks with digital settings for LNA, mixer, filter
  - Tuning patent (US 7,149,488 B2); RSSIs & digital block-level bias trimming
- 5.15-5.825GHz WLAN transceiver in 0.18 $\mu$ m CMOS [35], (Athena, 2003)
  - Digital I/Q mismatch correction
  - Multiple internal loopback switches for self-calibration in test mode
  - 8-bit DACs for DC offset minimization after mixers and filters
- 2.4GHz WLAN transceiver in 0.25 $\mu$ m CMOS [36], (MuChip, 2005)
  - Baseband I/Q gain and phase calibration
  - Extra analog mixer & peak detector

### II.2.5. Combined digital monitoring, analog measurements, and tuning

The circuit-level research projects discussed in the following sections are based on the hybrid analog/digital approach outlined in the previous subsection. One goal is to improve fault observability and calibration effectiveness by adding more measurement circuitry in the analog segments to provide data that can become part of the system-level calibration routine. Information from measurements can be used for block-level tuning prioritizations and optimizations, leading to shorter start-up routines and convergence times of algorithms. Fig. 10 portrays the envisioned transceiver with enhanced analog measurements, where power detectors (PD) measure gains along the analog chain [20]-[23]. Power gain and linearity measurements through temperature sensing are explored in Section V. In contrast to conventional power detectors, temperature sensors do not physically come in contact with the CUT and thus avoid loading effects.

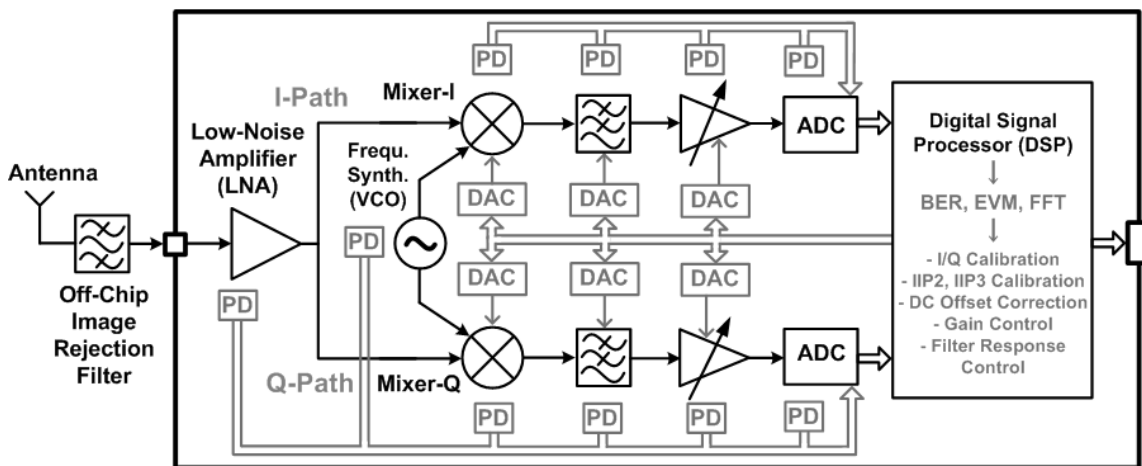


Fig. 10. Transceiver with digital monitoring, analog measurements, and tuning.

Another aspect of comprehensive system-level self-calibration is that the analog circuits must have tunable or programmable elements, meaning that “knobs” to adjust performance parameters must be identified. Progress towards more analog features for detection of process parameter shifts and performance degradations is also beneficial because detection and tuning in the analog domain is often faster than the digital counterpart. Hence, start-up routines could be improved with added analog tuning features. One tool to do so is the analog mismatch reduction scheme presented in Section VI. Current trends show that the conglomerate of analog and digital techniques is crucial for effective built-in tests of complex single-chip systems, motivating the continued development of BITs and digitally controllable analog circuit blocks. Pros and cons of the aforementioned self-test and calibration concepts are recapped in Table II.

#### *II.2.6. High-volume manufacturing testing*

A production test strategy for transceiver systems-on-a-chip has recently been proposed in [37] to address cost savings through the use of soft specification limits based on statistical parameter distributions in combination with a defect-oriented test approach that enables low-cost testing using less accurate equipment or built-in circuitry. Such a test strategy would open doors for positive impact of the circuit-level adjustment features from this research on product yields. Since the suggested approach in [37] involves crude and fast tests around the acceptable minimum and maximum



specification limits for a given parameter, digital programmability in the analog blocks makes retesting with fast on-chip performance tuning possible. Therefore, in reference to Fig. 3, self-calibration leads to narrower parameter distributions and thus higher production yields [37].

The on-chip temperature sensor in this work extracts the gain and linearity information that conventional power detectors ([20]-[23]) for built-in testing provide. Since the sensors also have DC output voltages, they simplify production testing by avoiding RF outputs requiring well-designed impedance-matched interfaces with the automatic test equipment (ATE). Furthermore, RF measurements drive up the production test cost and are undesirable in multi-site (parallel) testing setups due to the limited number of RF channels on the ATE [38]. Since reading out DC voltages with on-chip multiplexors is more practical than routing high-frequency signals, built-in test and calibration typically reduces the number of I/O pads, thereby decreasing die sizes.

Table II. Comparison of transceiver built-in testing and calibration techniques

Approach	Typical Applications	Advantages	Disadvantages
<b>Digital Correction and Calibration</b> (Section II.2.1)	<ul style="list-style-type: none"> <li>I/Q mismatch calibration</li> <li>Digital dynamic offset compensation</li> <li>System-level performance measurements (BER, FFT, EVM) with external test input or training symbols during normal operation</li> </ul>	<ul style="list-style-type: none"> <li>High accuracy</li> <li>No measurement circuitry in the analog front-end that could load the signal path</li> <li>Well-suited for background calibration</li> <li>Digital BIT circuit performance is robust to PVT variations</li> <li>Low area and power overhead (when the DSP is on the chip)</li> </ul>	<ul style="list-style-type: none"> <li>Large variations in the analog front-end gain or linearity cannot be corrected (e.g. saturation of analog stages from DC offset amplification)</li> <li>Convergence times are longer (millisecond range). Convergence times increase with PVT variation severity.</li> <li>Adaptive optimization of analog circuits is not possible because failure cause information is not available.</li> </ul>
<b>Analog Measurements and Tuning</b> (Section II.2.2)	<ul style="list-style-type: none"> <li>I/Q mismatch calibration in image-reject receivers</li> <li>Block-level characterization and tuning</li> <li>Dedicated transceiver front-end chips without on-chip digital resources</li> </ul>	<ul style="list-style-type: none"> <li>Direct correction of analog blocks with control voltages</li> <li>Fast settling times</li> <li>Typically suitable for background calibration</li> <li>The only option when the digital baseband processor is on a different chip</li> <li>Can be applied to high-frequency blocks</li> </ul>	<ul style="list-style-type: none"> <li>Increased power and die area due to analog BIT circuitry</li> <li>BIT circuitry is connected to CUTs and failures can impact the main signal path</li> <li>Intensive design efforts (BIT circuitry implementation is significantly different, depending on transceiver types, applications, and accuracy requirements.)</li> </ul>
<b>Loopback Testing</b> (Section II.2.3)	<ul style="list-style-type: none"> <li>Production testing</li> <li>Quick self-tests when the transceiver is idle</li> </ul>	<ul style="list-style-type: none"> <li>The most important system-level parameter is verified: bit error rate performance</li> <li>Fast verification of all on-chip blocks</li> <li>Low area and power overhead for BIT circuits</li> </ul>	<ul style="list-style-type: none"> <li>No or limited data about fault locations unless combined with analog measurement circuits</li> <li>Not suitable for on-line calibration (transceiver must be idle and in test mode)</li> </ul>
<b>Combined Digital Performance Monitoring and Analog Compensation</b> (Section II.2.4)	<ul style="list-style-type: none"> <li>I/Q mismatch calibration</li> <li>Analog dynamic offset compensation to prevent saturation</li> <li>Coarse start-up calibrations</li> <li>Production testing and on-line calibration</li> </ul>	<ul style="list-style-type: none"> <li>Analog compensation overcomes large PVT variations and reduces design margin requirements</li> <li>Front-end circuitry adjustments for deficiencies that cannot be corrected in the digital domain (transistors in unacceptable operating region due to process variations, low SNR from diminished front-end gain, amplified DC offsets in analog circuits that saturate internal nodes or the ADC input)</li> <li>Well-suited for background calibration</li> </ul>	<ul style="list-style-type: none"> <li>Limited insights into block-level performance</li> <li>Complex calibration algorithms</li> <li>Solutions are developed specific to the transceiver under test</li> <li>Analog circuits must be programmable</li> </ul>
<b>Combined Digital Monitoring, Analog Measurements, and Analog Compensation</b> (Section II.2.5)		<ul style="list-style-type: none"> <li>Highest detection capability of faults and performance shifts on the block-level and system-level</li> <li>Block-level optimization as part of system calibration algorithms</li> <li>Well-suited for background calibration</li> </ul>	<ul style="list-style-type: none"> <li>Area and power overhead for measurement circuitry</li> <li>Complex calibration algorithms</li> <li>Intensive design efforts (BIT circuitry implementation is significantly different depending on transceiver types, applications, and accuracy requirements.)</li> </ul>

### III. HIGH-LINEARITY TRANSCONDUCTANCE AMPLIFIERS WITH DIGITAL CORRECTION CAPABILITY\*

#### III.1. Background

Operational transconductance amplifiers (OTAs) are essential elements of transconductance-capacitor ( $G_m$ -C) filters [39]-[40],  $\Delta\Sigma$  modulators [41], gyrators, variable-gain amplifiers, and negative-resistance elements. Compared to their active-RC counterparts,  $G_m$ -C filters enable low-power operation and tuning of the filter characteristics at higher frequencies, but are less linear. Tunable active-RC filters are suitable for low-frequency applications; however, extending their use to higher frequencies would require significantly more power. On the other hand, OTA-based filters in wireless receivers and continuous-time (CT)  $\Delta\Sigma$  analog-to-digital converters (ADCs) increasingly mandate good linearity at higher frequencies. These applications typically require highly linear OTAs with third-order inter-modulation (IM3) distortion better than -60dB. Further advances in high-frequency  $G_m$ -C filters with SNDRs over

---

\* © 2010 IEEE. Section III is in part reprinted, with permission, from “Attenuation-predistortion linearization of CMOS OTAs with digital correction of process variations in OTA-C filter applications,” M. Mobarak, M. Onabajo, J. Silva-Martinez, and E. Sánchez-Sinencio, *IEEE J. Solid-State Circuits*, vol. 45, no. 2, pp. 351-367, Feb. 2010. This material is included here with permission of the IEEE. Such permission of the IEEE does not in any way imply IEEE endorsement of any of Texas A&M University's products or services. Internal or personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution must be obtained from the IEEE by writing to [pubs-permissions@ieee.org](mailto:pubs-permissions@ieee.org). By choosing to view this material, you agree to all provisions of the copyright laws protecting it.

50dB are also desirable for channel selection/equalization in multi-Gbps portable data communication devices [40], and for possible application in next generation analog-to-information receivers if dynamic range  $> 90\text{dB}$  in 200MHz bandwidth becomes attainable [42].

Viable high-frequency  $G_m$ -C filter solutions were presented in [39] and [43] with 3-dB frequencies at 275MHz and 184MHz, respectively. The topology reported in [39] has low noise, limited linearity, and a pseudo-differential realization prone to low power supply rejection ratio (PSRR). The filter in [43] achieves high linearity with relatively low power but higher noise. Trade-offs between linearity, noise, power, and operating frequency are common and have been incorporated into figures of merit (FOMs) such as in [44] and [45]. Recent works also address alternative filter structures such as the source-follower-based approach [46] and performance improvement of typical OTA topologies [47].

A popular linearization approach is to cross-couple two transconductors, theoretically cancelling certain harmonics at specific bias conditions over a limited frequency range. A typical cross-coupled OTA contains two paths; each having different transconductance and the same amount of harmonic distortion. When cross-coupled, the equal harmonics cancel under ideal conditions and the effective transconductance is the difference between the two paths. The frequency dependence of this approach has been analyzed with a Volterra series in [48] and [49], in which the analytical expressions are correlated with measurement results. Process-voltage-temperature (PVT) variations, high-frequency effects, and device modeling inaccuracies will create unforeseen

mismatches between the two amplifiers. Therefore, precision tuning of bias currents/voltages is typically required. Attenuation and cross-coupling has been combined for the low-noise amplifier in [50], in which distortion cancellation is restricted to third-order nonlinearities with feedforward path and precise off-chip input attenuation.

The proposed methodology is an architectural solution that achieves up to 22dB IM3 improvement over an identical nonlinearized OTA design at frequencies as high as 350MHz. It can be generalized to fully-differential topologies which offer high PSRR and common-mode rejection ratio (CMRR). Since the maximum frequency is mainly limited by process parasitics and OTA performance, the approach shows promise of exceeding 350MHz bandwidth in future nanoscale CMOS processes. Robust linearization over a wide frequency range demands a mechanism to correct for high-frequency effects and PVT variations, for which a digital programmability scheme is proposed. Section III.2 describes the proposed attenuation-predistortion linearization methodology along with the result from Volterra analysis that ensures broadband performance. The corresponding OTA and  $G_m$ -C filter design issues are addressed in Section III.3. Section III.4 presents digital correction requirements based on PVT simulations. Measurement results for a linearized fully-differential OTA and a 2<sup>nd</sup>-order biquadratic  $G_m$ -C lowpass filter in 0.13 $\mu$ m CMOS technology are provided in Section III.5, and conclusions are given in Section III.6.

### III.2. Attenuation-Predistortion Linearization Methodology

Signal attenuation at the OTA input [48] reduces the effective transconductance and decreases the SNR. Alternatively, distortion cancellation by means of cross-coupled differential pairs results in increased power consumption and noise proportional to the transistor parameters in the additional path. Since the extra differential pair normally has less transconductance than the main pair, the effective transconductance is reduced by 10-50%. However, both transistor pairs should have the same third-order nonlinearity, which translates into different transistor sizes and bias currents for each pair. As a result, the cross-coupling technique is sensitive to PVT variations and restricted to narrow frequency ranges. Another common method to linearize a transistor having transconductance  $g_m$  is to add a degeneration resistor  $R_{sd}$  at the source [48], which makes the third-order harmonic distortion proportional to the factor  $1/(1+g_m R_{sd})^3$ . Nonetheless, large degeneration resistance results in higher input-referred noise, lower transconductance, and less voltage headroom. The effective transconductance ( $g_{msd}$ ) and the input-referred noise ( $v_{nsd}^2$ ) with *resistive source degeneration* are given by

$$g_{msd} = \frac{g_m}{1 + g_m R_{sd}}, \quad v_{nsd}^2 \approx \frac{4kT}{g_m} (2/3 + g_m R_{sd}); \quad (3)$$

where the noise coefficient  $\gamma$  was approximated as  $2/3$ . For example, using a degeneration factor  $g_m R_{sd} = 2$  will ideally result in IM3 improvement of approximately 29dB, an input-referred noise power increase by a factor of 4, and a decrease of the transconductance to one third of its original value. But based on simulations of the OTA from this work with  $g_m R_{sd} = 2$ , the expected IM3 improvement would be 25.2dB with an associated noise power increase of more than 9 times.

The proposed attenuation-predistortion method is independent of OTA topology and involves cancellation of all distortion components except those from secondary effects at high frequencies. It can be used in conjunction with other circuit-level linearization techniques internal to the OTA, such as source degeneration or cross-coupling.

### III.2.1. Single-ended circuits

Fig. 11 depicts the single-ended architecture that contains an auxiliary branch with an OTA having identical dimensions, DC bias, and AC common-mode conditions as in the main path to generate the distortion components required for cancellation. An important advantage of identical paths is robustness to PVT variations because of optimal device matching obtainable from proper layout. In this scheme, it is avoided to base the distortion cancellation on branches with different transconductor device dimensions or bias conditions, which would degrade matching accuracy. But even with minimized mismatches, nonlinearities are particularly frequency-dependent at high frequencies and remain sensitive to PVT variations as established in Section III.4. Hence, the proposed linearization method involves variable resistors to tune performance and counteract high-frequency degradation as well as PVT variations. Either a resistive or capacitive divider can form the attenuator at the input of the auxiliary path; however, resistors add more noise.

Distortion cancellation in the single-ended case requires  $G_m \times R = 1$ , which is ascertained by the following analysis. For a certain input voltage amplitude  $V_m$ , the output current can be divided into a linear part  $i_{lin}\{V_m\} = G_m \times V_m$  and a nonlinear part  $i_{non-lin}\{V_m\} = g_{m2} \times V_m^2 + g_{m3} \times V_m^3 + \dots$ , where  $g_{m2}, g_{m3}, \dots$  are Taylor series coefficients

of the transconductance. The differential input of the main OTA is:  $V_{dif} = V_{in} - [V_{in}/2 + i_{non-lin}\{V_{in}/2\}/G_m] = V_{in}/2 - i_{non-lin}\{V_{in}/2\}/G_m$ . Under ideal conditions, the distortion generated in the auxiliary path,  $-i_{non-lin}\{V_{in}/2\}$ , cancels out the distortion in the main voltage-to-current conversion. In practice, distortion caused by nonlinearities at the output of the auxiliary OTA and high-frequency effects introduce some finite uncanceled distortion. Capacitor  $C_o$  represents the lumped output capacitance of the auxiliary OTA, input capacitance of the main OTA, and layout parasitics. Resistor  $R_c$  of the phase shifter and equivalent input capacitance  $C_i$  provide 1<sup>st</sup>-order frequency compensation, creating a pole to equalize the phase shift between the main and auxiliary paths. Compensation is necessary at high frequencies because  $C_o$  at the negative input terminal of the main OTA creates a pole with resistor  $R$  in the auxiliary path.

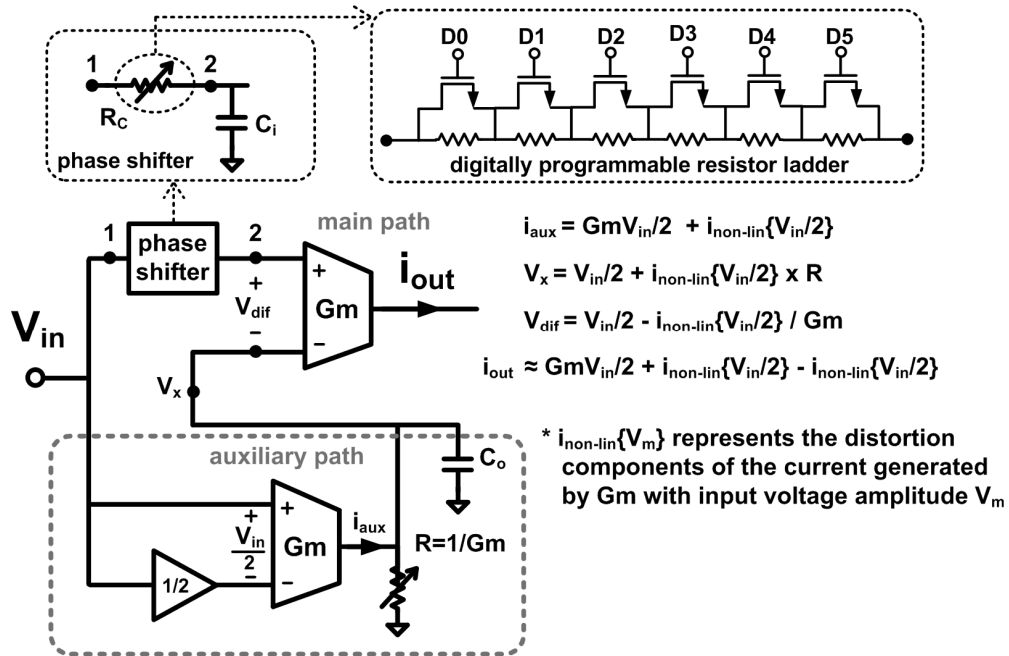


Fig. 11. Attenuation-predistortion linearization for single-ended circuits.



### III.2.2. Fully-differential circuits

A conceptual diagram of the proposed linearization approach for a fully-differential transconductor ( $G_m$ ) is displayed in Fig. 12. In the fully-differential case, attenuation factors at the input of the transconductors are realized with floating-gate devices described in Section III.3.1. As discussed in [48] and [51], the inherent input attenuation with floating-gate stages enhances the OTA linearity. The distortion cancellation principle is the same as in the single-ended case, but different conditions must be satisfied for fully-differential implementation, which are explained in sections III.2.3 and III.3.1 with regards to the attenuation ratios. By selecting an input attenuation ratio of  $1/3$  and voltage gain of  $3$  in the auxiliary branch ( $G_m \times R = 3$ ), the signal amplitude  $V_x$  is equal to  $V_{in}$  plus three times the distortion components caused by the nonlinear current  $i_{non-lin}\{V_{in}/3\}$  from the transconductor with input amplitude of  $V_{in}/3$ . In the main path, the effective differential OTA input signal is:  $V_{dif} = 2V_{in}/3 - V_x/3 = 2V_{in}/3 - [V_{in} + 3 \times i_{non-lin}\{V_{in}/3\}/G_m] / 3 = V_{in}/3 - i_{non-lin}\{V_{in}/3\}/G_m$ . Thus, the differential signal contains the attenuated input signal and the inverse of the distortion generated by the identical  $G_m$  in the auxiliary branch for distortion cancellation during the voltage-to-current conversion in the main path. Ideally, the distortion components are canceled by the equal and opposite terms from the predistortion of the differential input signal except for negligible higher-order components.



### III.2.3. Scaling of attenuation ratios

Depending on application-specific requirements, the design parameters in the attenuation-predistortion linearization approach can be selected to adjust the voltage swings and the effective transconductance. Fig. 13 shows the differential attenuation-predistortion linearization scheme, where frequency compensation and parasitic capacitors have been omitted for simplicity. The following analysis assumes floating-gates as a practical attenuator implementation choice under the constraint that factors  $k_1$  and  $(1-k_1)$  are related as elaborated upon in Section III.3.1, but less restrictive types of attenuators could also be used. The output current  $i_o$  of an OTA due to an input voltage  $V_m$  can be modeled as having a linear and a nonlinear part:  $i_o = G_m V_m + i_{non-lin}\{V_m\}$ . Ignoring high-frequency and secondary effects, the following relation can be written:

$$i_{out} \approx G_m [k_1 - (1-k_1)k_2 G_m R] V_{in} - (1-k_1) G_m R \cdot i_{non-lin}\{k_2 V_{in}\} + i_{non-lin}\{[k_1 - (1-k_1)k_2 G_m R] V_{in}\}; \quad (4)$$

where:  $i_{non-lin}\{k_2 V_{in}\} \cdot R(1-k_1) \ll (k_1 - (1-k_1)k_2 G_m R) V_{in}$  is assumed in the approximation. To cancel the distortion, the following conditions should hold:

- i) The auxiliary and main OTAs should have the same effective input voltage amplitudes such that an identical distortion is created at their respective outputs.
- ii) The gain in the auxiliary path must ensure that the distortion through this signal path reaches the output of the main OTA with a gain of -1.
- iii) The internal signal swings should be bounded, i.e.:

$$k_2 G_m R \leq 1. \quad (5)$$

Applying conditions i) and ii), cancellation of the nonlinear terms in (4) requires:

$$(1 - k_1)G_m R = 1 \quad , \quad k_2 = k_1 / 2 . \quad (6)$$

Consequently, the effective transconductance with linearization is given by

$$G_{meff} = [k_1 - (1 - k_1)k_2 G_m R]G_m = (k_1 / 2)G_m = k_2 G_m . \quad (7)$$

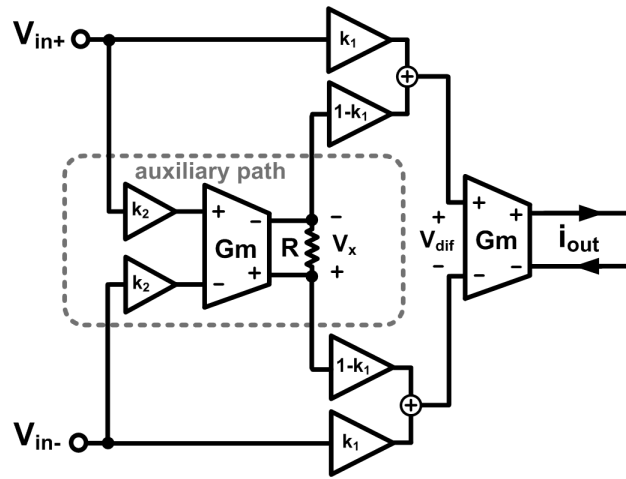


Fig. 13. Low-frequency model for the attenuation-predistortion scheme.

Condition iii) depends on the application and is not always necessary. Cancellation of distortion with the proposed technique requires weakly nonlinear operation in the auxiliary branch, which is ensured by limiting the signal swing with this condition. The example that is presented in Fig. 12 was derived with  $k_2 G_m R = 1$ , ensuring that the signal swing at the output of the auxiliary OTA is the same as at its input. This choice was made to maintain the same maximum input voltage swing as the initial OTA without saturating the OTA in the linearization path. If the specified input signal is  $k_2 G_m R$  times below the OTA saturation level, then  $k_2$  can be increased accordingly to obtain  $k_2 G_m R >$

1 and higher effective transconductance based on (7). But, this choice is only permissible if a reduction of the maximum input swing by  $k_2 G_m R$  can be tolerated, which would imply a reduction in the dynamic range. Typically, choosing  $k_2 G_m R = 1$  is advantageous to maintain the same maximum input voltage swing as the original OTA after the linearization. Selection of  $k_1 = 2/3$  and  $k_2 = 1/3$  results in the highest effective transconductance that can be achieved in (7) based on the above conditions while also satisfying the attenuation factor relationships in the floating-gate devices (Section III.3.1) with identical signal swings at the input and output of the auxiliary OTA ( $k_2 G_m R = 1$ ). Hence,  $G_m R = 3$  under the stated conditions.

#### III.2.4. Volterra series analysis

The preceding expressions are valid at low frequencies and give insights into the conditions to cancel total distortion when secondary effects are negligible. Following the procedure outlined in [52], the 3<sup>rd</sup>-order Volterra series analysis in Appendix A reveals the following requirement for the phase shifter resistor in Fig. 12 to minimize IM3 at high frequencies:

$$\begin{aligned}
 i_{IM3} \approx & g_{m3} \left( \frac{k_1/2}{1+2C_p/C} \right)^3 (3V_{in1}^2 V_{in2}/4) \\
 & \times \left( \frac{1+j\omega_1 C((1-k_1)R - k_1 R_c) + 2j\omega_1 C_o R}{1+j\omega_1 b - c\omega_1^2} \right)^2 \left( \frac{1-j\omega_1 C((1-k_1)R - k_1 R_c) - 2j\omega_1 C_o R}{1-j\omega_1 b - c\omega_1^2} \right) \quad (8) \\
 & - g_{m3} \left( \frac{k_1/2}{1+2C_p/C} \right)^3 (3V_{in1}^2 V_{in2}/4) \frac{1+j\omega_1 C k_1 R_c}{1+j\omega_1 b - c\omega_1^2} \\
 \Rightarrow R_c \approx & \frac{(1-k_1) + 2C_o/C}{2k_1} R \quad \text{for } i_{IM3} \approx 0
 \end{aligned}$$

In the discussed example case with  $k_1 = 2/3$ , the condition to cancel IM3 with the phase shifter block in Fig. 12 is  $R_c = (R/4) \cdot (1+6C_o/C)$ . To ensure high linearity with variations of parasitic capacitances, the programmable range of  $R_c$  is selected based on process corner simulations as described in Section III.4.

### III.3. Circuit-Level Design Considerations

#### III.3.1. Fully-differential OTA with floating-gate FETs

Fig. 14 displays the schematic of the OTAs implemented on the 0.13 $\mu$ m CMOS test chip with a 1.2V supply. Attenuators  $k_1$ ,  $(1-k_1)$ , and  $k_2$  are realized with floating-gate devices for attenuation-predistortion linearization of this fully-differential topology. The gates (G) of the standard NMOS transistors in the OTA core are not resistively biased and are only connected to two conventional metal-insulator-metal (MIM) capacitors. Fig. 14 also visualizes the equivalent capacitive load seen at the  $V_{1+}$  and  $V_{1-}$  inputs, where  $C_{pt}$  represents the effective gate-to-ground<sub>(AC)</sub> capacitance from transistor parasitic capacitances. With this configuration, the gate voltages are:  $V_{G+/-} = (C_{FG1}/C_{total})V_{1+/-} + (C_{FG2}/C_{total})V_{2+/-}$ , where  $C_{total} \approx C_{FG1} + C_{FG2}$  when  $C_{pt}$  is negligible. It follows that the attenuation factors in Fig. 13 are:  $C_{FG1}/C_{total} = k_1$  and  $C_{FG2}/C_{total} = (C_{total}-C_{FG1})/C_{total} = 1-k_1$ . The accuracy of the  $k_1$  and  $(1-k_1)$  factors predominantly depends on the matching of the MIM capacitors  $C_{FG1}$  and  $C_{FG2}$ , which can be achieved within 0.1-1% using proper layout techniques. As assessed in Section III.4, such a matching accuracy is more than sufficient with the 3%-step programmability of resistor  $R$  for gain mismatch compensation in both paths.

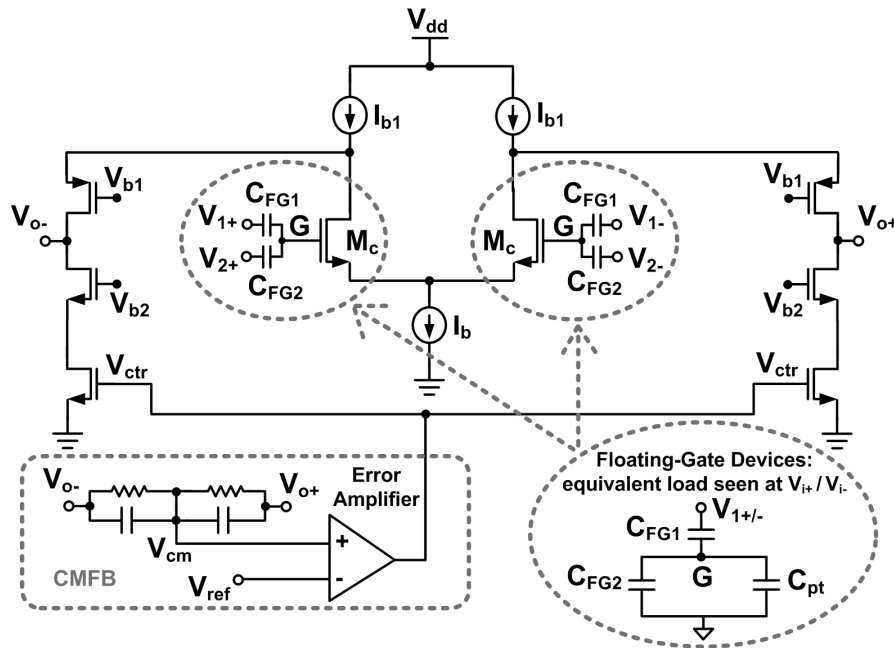


Fig. 14. Folded-cascode OTA (implements  $G_m$  in the main and auxiliary paths).

In the layout, all nodes G at the floating gates in Fig. 14 are connected to the top metal layer using standard poly-metal contacts and metal-metal vias. During fabrication, this connection ensures that any charge stored on the floating gates flows to the substrate because all connections to the top metal are still joined prior to their separation during the last etching step. Thus, no charge is stored on the floating gates when the substrate contacts are also connected to the top metal layer [53], allowing gate discharge into the substrate before the last etching operation. After etching, the top metal extensions of the gates without trapped charge are floating, leaving only the connections to the two MIM capacitors. The floating-gate device design expressions for  $k_l$  and  $(1-k_l)$  above are assuming absence of excess charge on the floating gates, which is a satisfied condition without extra fabrication steps as a consequence of the gate and substrate connections to

the top metal. A special programming technique for non-zero charge on the floating gates was not utilized in this work, but a more sophisticated floating-gate device implementation as presented in [51] could be explored, which promises additional potential for compensation of inherent transistor threshold voltage offsets in the OTA's input differential pair.

The phase shifter in Fig. 12 creates an extra pole within the linearized architecture that the reference OTA does not have. This phase delay is roughly the same as the delay from the pole formed by  $R$  and  $C_o$  in the auxiliary path. In low-loss (high-Q) designs, the additional pole can affect the gain of integrators and the frequency response of biquad sections if  $1/(RC_o)$  is not significantly larger than the operating frequency. A load compensation scheme based on [54] is discussed in Appendix B for such situations.

Identical standalone OTAs are included on the same die to obtain reference linearity measurements. The reference OTA also has a floating-gate input attenuation of  $1/3$  for fair performance comparison. In this way, the linearity benefit from the input attenuation is isolated from the architectural linearization proposed in Fig. 12, and both OTAs have the same effective transconductance ( $G_m/3$  in this case), but the linearization results in doubled power consumption. Since attenuation and feedback linearization techniques have known linearity and effective transconductance trade-offs, the circuit-level comparison in this work is focused on the predistortion linearization scheme relative to a commensurate OTA with equal input attenuation factor. This baseline OTA in Fig. 14 was biased with  $I_b = 0.95\text{mA}$  and  $I_{b1} = 0.85\text{mA}$ , having an effective transconductance of  $510\mu\text{A/V}$ . The linearization does not require any design changes in this core OTA, but



redesign of the OTA is an option if it is required to meet the same power budget after linearization, which is possible as long as OTA bandwidth reduction can be tolerated. Such a linearization under power constraint is disclosed in Appendix C.

Suppression of undesired common-mode signals and noise is vital for linearity at high frequencies. The common-mode feedback (CMFB) circuit should have high gain to accurately control the common-mode output voltage while maintaining a wide bandwidth to reject common-mode noise in the band of interest. The CMFB amplifier is shown in Fig. 15, where  $V_{ctr}$  is the control voltage applied to the OTA in Fig. 14. The addition of the compensation resistor  $R_z$  results in two zeros in the transfer function of the error amplifier, which helps to insure stability of the CMFB loop. The simulated AC response of the CMFB loop has a 51.9dB low-frequency gain and a 424.9MHz unity-gain frequency with  $42.5^\circ$  phase margin.

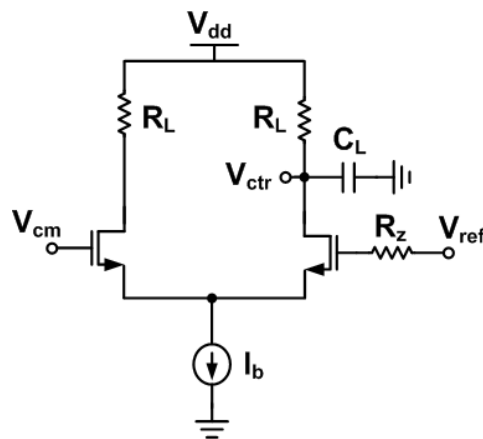
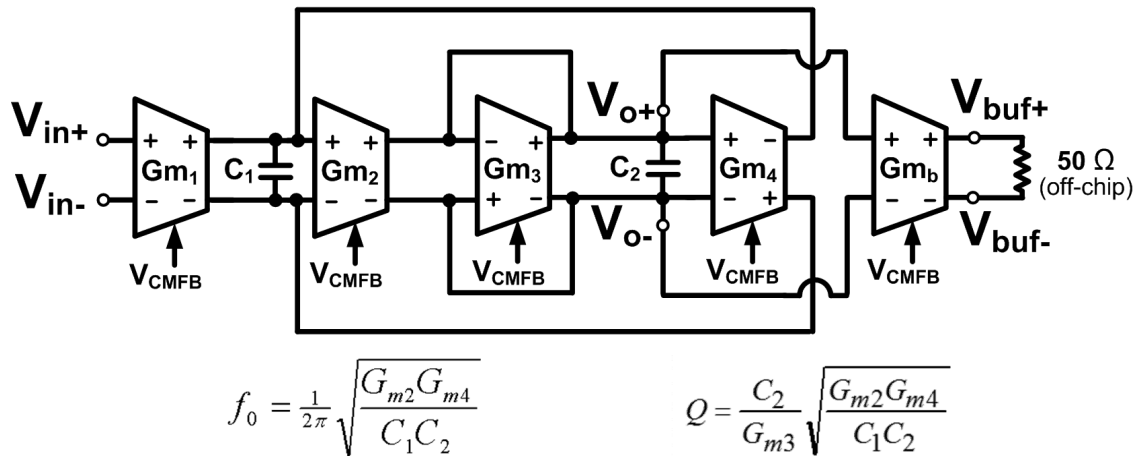


Fig. 15. Error amplifier circuit in the CMFB loop.

### III.3.2. Proof-of-concept filter realization and application considerations

A 2<sup>nd</sup>-order  $G_m$ -C biquad filter was designed with attenuation-predistortion-linearized OTAs to verify that the proposed methodology is suitable for filters with  $G_m$ -C integrator loops. Fig. 16 shows the filter schematic and specifications. The lowpass output of the biquad was measured using another OTA as buffer to drive the 50 $\Omega$  input impedance of the spectrum analyzer.



Parameter	Value
Corner frequency ( $f_{3db}$ )	194.7 MHz
Passband gain	0 dB
Output buffer gain ( $G_{m_b} \times 50 \Omega$ ) & off-chip losses	-34.2 dB
$G_{m_{1,2,3,4}}$	510 $\mu$ A/V

Fig. 16. 2<sup>nd</sup>-order lowpass filter diagram and design parameters.

The primary motivation for digital correction (Section III.4) to enhance linearity performance with severe process variation is the compatibility with digitally-controlled receiver calibration approaches that involve the baseband filter. Practical implementation details for receivers with digital performance monitoring and calibration of analog blocks are described in [33]-[36]. They incorporate accurate digital monitoring and I/Q mismatch correction in the digital signal processor (DSP) as well as a few analog observables that give some insights into the operating conditions, such as outputs from received signal strength indicators or DC control voltages of blocks. The possibility exists to generate and apply test tones at the input of an analog block and extract performance indicators from the output spectrum in the DSP, which contains distortion components. Conversely, calibration could also be performed by monitoring the bit error rate (BER) in the DSP from processing a special test sequence or customary pilot symbols at the beginning of receptions. Since linearity degradation impacts the BER, such a calibration could be computationally more efficient than calculating and analyzing the fast Fourier transform in the DSP. Regardless of the specific digital calibration algorithm, the digitally-controlled correction capability of the proposed linearization scheme can potentially enable filter linearity tuning in integrated receiver applications without the need for extra DACs.

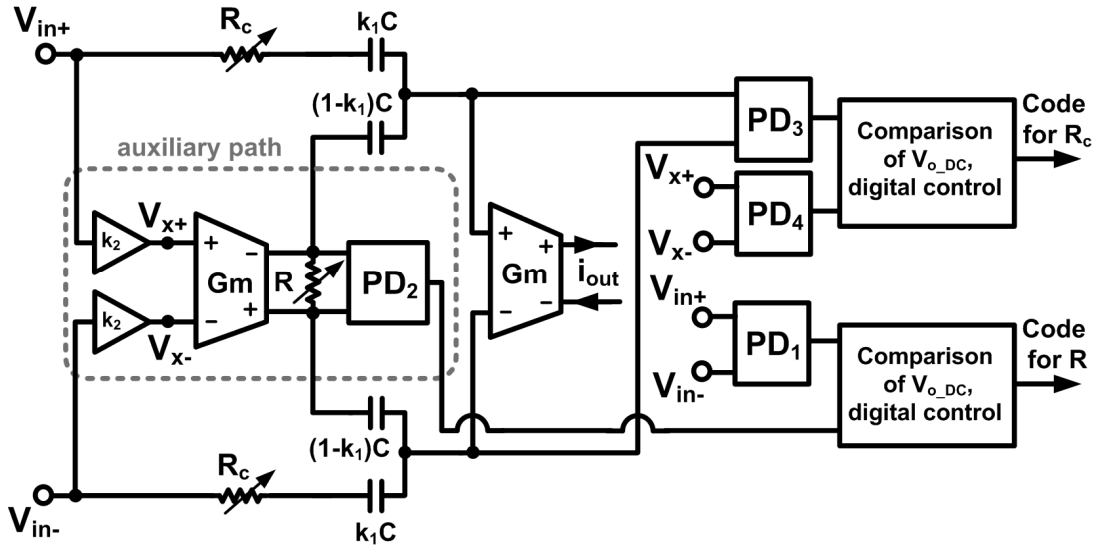


Fig. 17. Block diagram of the proposed automatic linearity tuning scheme.

An alternative automatic calibration that does not involve an on-chip DSP but dedicated analog and simpler digital logic circuitry is displayed in Fig. 17. From the conditions for optimum distortion cancellation described in Section III.2.3, the gain of the auxiliary path must be equal to  $k_2 G_m R$ , which is unity in the discussed design example. This can be ensured by measuring the signal level at the input and output of the auxiliary OTA with power or peak detectors ( $PD_1$ ,  $PD_2$ ), and controlling the digital code of resistor  $R$  until the gain is unity. The simplest control algorithm would be to cycle through the codes that determine the value of  $R$  until the difference in the DC output voltages of  $PD_1$  and  $PD_2$  is minimized, which can be performed digitally by detecting the toggling instance at the output of a single comparator. At higher frequencies, the parasitic pole in the auxiliary path starts to affect the distortion cancellation, causing the signal level at the output of the auxiliary OTA to decrease with increasing frequency.

Hence, the differential input signal to the main OTA at PD<sub>3</sub> increases as a result, which is shown in Fig. 18. By measuring this signal that is ideally equal to  $V_x = k_2 \cdot V_{in}$  with PD<sub>3</sub>, the value of the phase shift resistor  $R_c$  can be adjusted until the outputs of PD<sub>3</sub> and PD<sub>4</sub> are equal. This comparison can be completed with the same logic as for PD<sub>1</sub>/PD<sub>2</sub>, but it has to be done with an input signal at the maximum frequency at which high linearity is desired. The automatic tuning has not been implemented on the circuit level, but simulations with different values of  $R_c$  showed that amplitude detection within 4.6% is required to detect  $R_c$  changes within 5% at 350MHz, which is sufficient for IM3 higher than 70dBc (Section III.4). In differential gain measurements, PVT errors in the detectors are cancelled except for the errors from unavoidable mismatches between the two detectors. Errors from mismatches are less than 5% at 2.4GHz in [55], and more accurate amplitude detection is achievable at lower frequencies. In [23] for example, differential on-chip amplitude measurements were conducted up to 2.4GHz using detectors with 0.031mm<sup>2</sup> die area and negligible loading of the signal path ( $C_{in} < 15\text{fF}$ ).

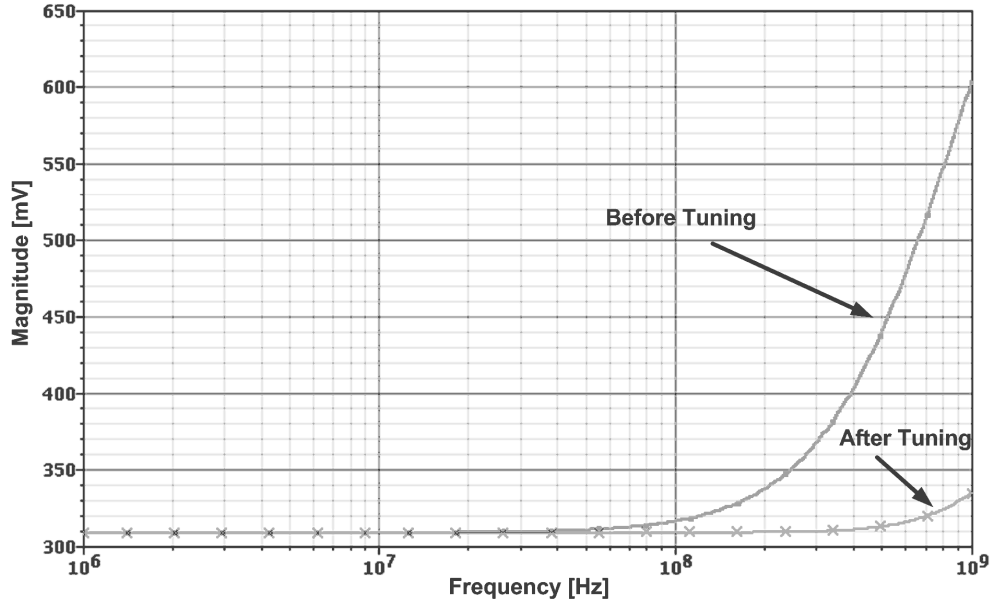


Fig. 18. Simulated AC amplitude at the input of the main OTA (PD<sub>3</sub> in Fig. 17).

#### III.4. Compensation for PVT Variations and High-Frequency Effects

Since the frequency compensation is based on equalization of phase shifts from RC time constants in the main and auxiliary paths, the optimum linearity point is subjected to PVT variations. Resistors  $R$  and  $R_c$  in Fig. 12 can be adjusted digitally to ensure high linearity. When implementing the attenuation ratios with matched capacitors, the variation of the resistors and transconductance mismatch between the auxiliary and main paths become the major sources of IM3 degradation. Fig. 19 illustrates the technique's sensitivity to 20% variation of  $R_c$  and  $G_m$  based on the expression for IM3 in (8). In theory, the  $|IM3|$  (in dBc) without parameter variation is infinite. After introducing a numerical resolution constraint, the peak  $|IM3|$  is limited to around 95dBc. Fig. 19a reveals that  $G_m$ -mismatch results in more degradation than  $R_c$  variation at low

frequencies, but at high frequencies variation of  $R_c$  becomes equally significant as evident from Fig. 19b. In general, less than  $\pm 10\%$  mismatch of  $G_m \times R$  and  $\pm 5\%$  variation of  $R_c$  are required for theoretical  $|IM3|$  higher than 70dBc. Under consideration of the trend towards increasing intra-die variability in modern CMOS processes, programmability of  $R$  and  $R_c$  is necessary to guarantee  $G_m \times R$  gain and  $R_c$  values within these limits. The determination of the appropriate incremental resistor step size is elaborated next.

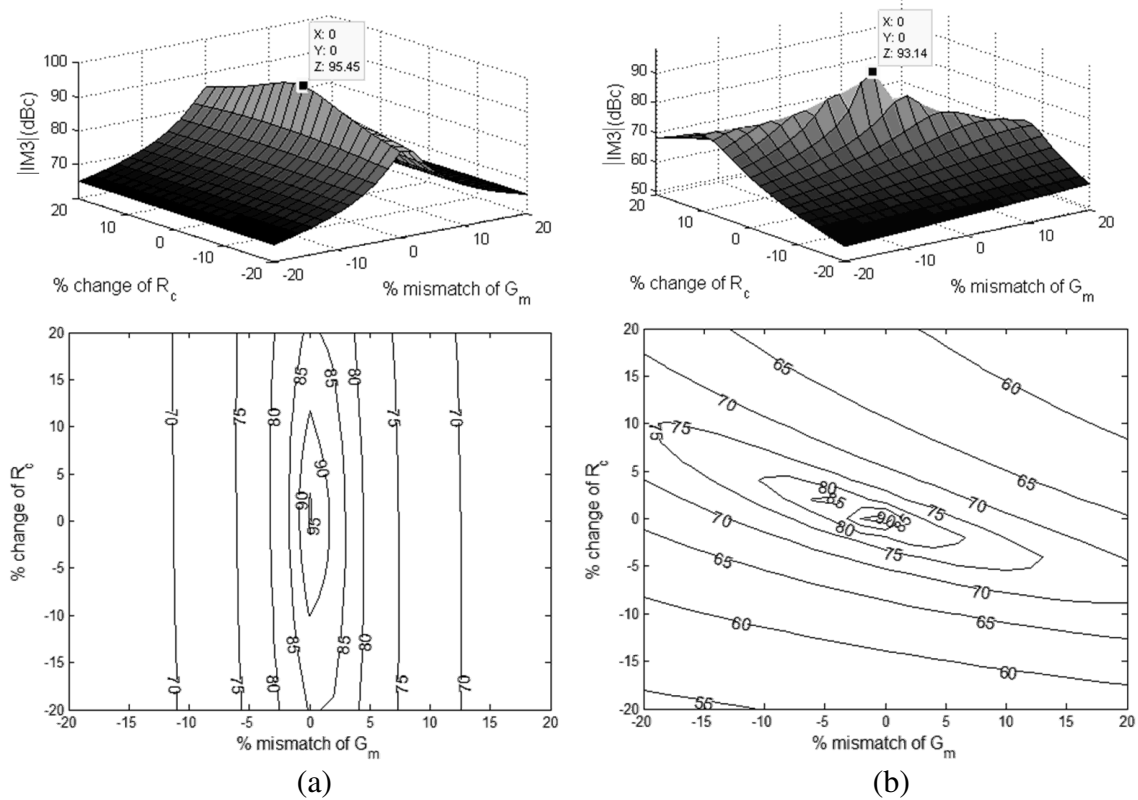
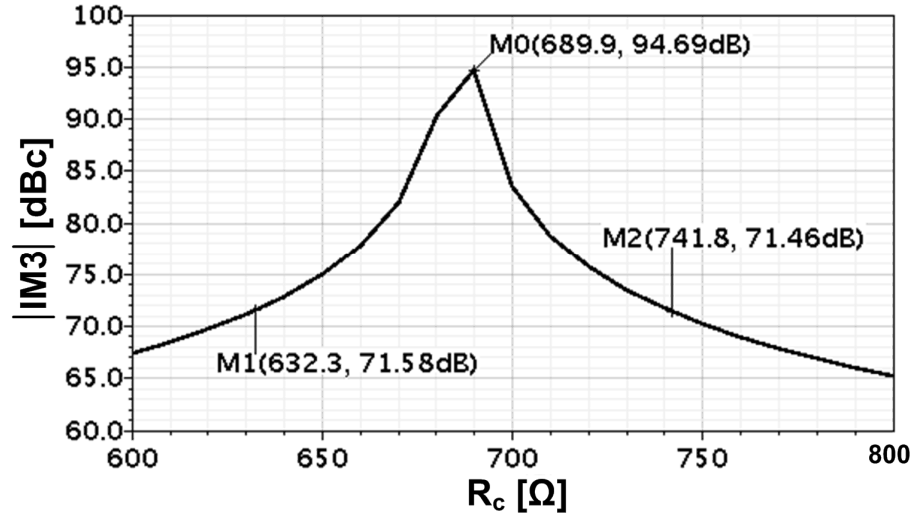


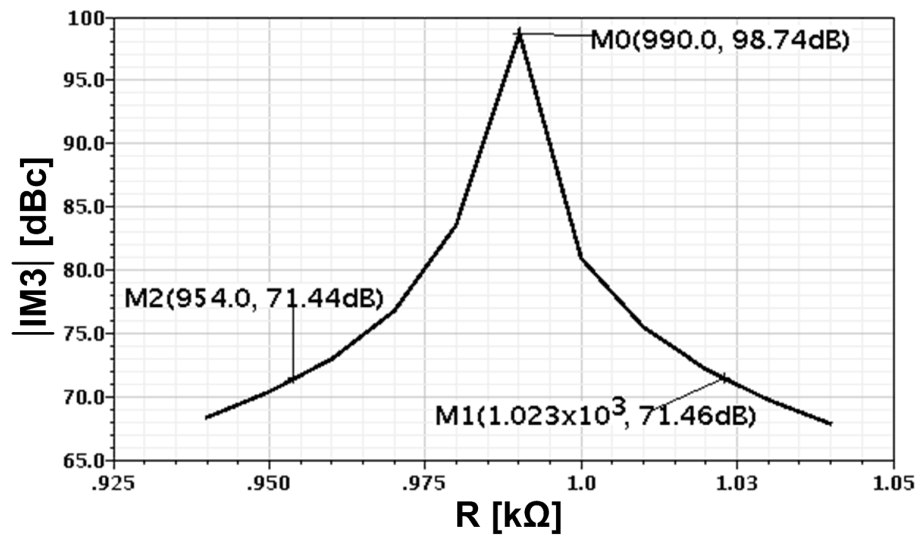
Fig. 19. Sensitivity of  $|IM3|$  (in dBc) to component mismatches. Calculated with equation (8): (a) 10MHz signal frequency, (b) 200MHz signal frequency.

To obtain a practical assessment of the distortion cancellation sensitivity, the compensation resistor value and transconductance mismatch in the two paths were varied in circuit simulations using Spectre. The resulting  $|IM3|$  is plotted vs. deviation from the nominal design parameters in Fig. 20, showing an  $|IM3|$  better than 71dBc for  $\pm 7.5\%$   $R_c$ -variation and  $|IM3|$  better than 71dBc for  $\pm 3.3\%$   $R$ -variation in the presence of 10%  $G_m$ -mismatch. The reference OTA has  $|IM3|$  of 51dBc. It is imperative for effective distortion cancellation to implement the resistor ladders with 3% steps, enabling digital correction of relatively small intra-die mismatches. To account for large absolute variations of parameters, the adequate resistor tuning range should be selected based on simulations under anticipated worst-case conditions. In this work, simulations with process corner models and temperatures ranging from  $-40^\circ\text{C}$  to  $100^\circ\text{C}$  were conducted. Based on these simulation results, a conservative range from  $\sim 30$  to  $2.2\text{k}\Omega$  (approximately 3% - 200% of the nominal value) and 6-bit resolution were chosen for the programmable resistors  $R_c$  and  $R$  (Fig. 12) in this prototype design.





(a)



(b)

Fig. 20. Simulated sensitivity to critical component variations and mismatches. (a)  $|IM3|$  vs. change in  $R_c$  at 350MHz, (b)  $|IM3|$  vs.  $R$  with 10% transconductance mismatch between main OTA and auxiliary OTA at 350MHz.

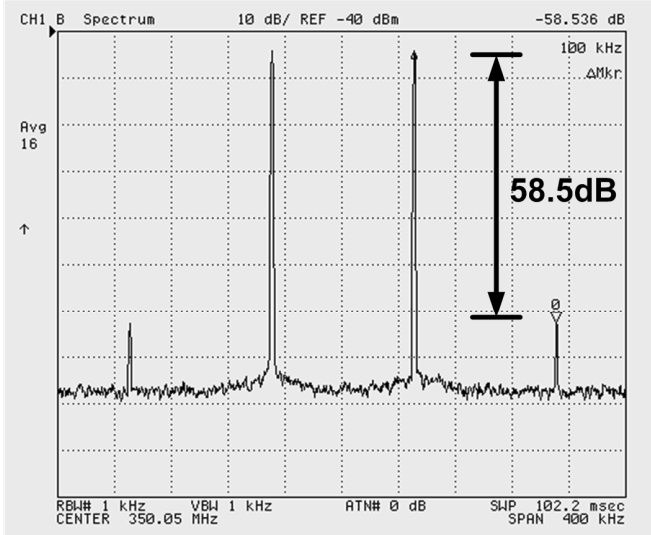
### III.5. Prototype Measurement Results

#### III.5.1. Standalone OTA

Table III summarizes the characterization results for the OTA by itself. Two  $0.1V_{p-p}$  (-16dBm) tones with 100KHz frequency separation and a combined voltage swing of  $0.2V_{p-p}$  were applied during IM3 measurements. The results in Fig. 21 demonstrate IM3 enhancement from  $-58.5\text{dB}$  to  $-74.2\text{dB}$  at 350MHz coupled with a rise in input-referred noise from  $13.3\text{nV}/\sqrt{\text{Hz}}$  to  $21.8\text{nV}/\sqrt{\text{Hz}}$  and twice the power dissipation, while other performance parameters are not affected significantly. The linearization decreased the SNR in 1MHz BW from  $74.5\text{dB}$  to  $70.2\text{dB}$ , but allowed to improve the IM3 by  $15.7\text{dB}$ . Depending on the frequency and switch settings, IM3 enhancement up to  $22\text{dB}$  was achieved with the compensation resistor ladders having 6-bit resolution. If more linearity improvement is required, the resolution of the resistor ladders ( $R$  and  $R_c$ ) in Fig. 12 can be increased by adding more control bits or using a MOS in triode region as one of the elements to obtain a series resistance that is closer to the optimum value for distortion cancellation.

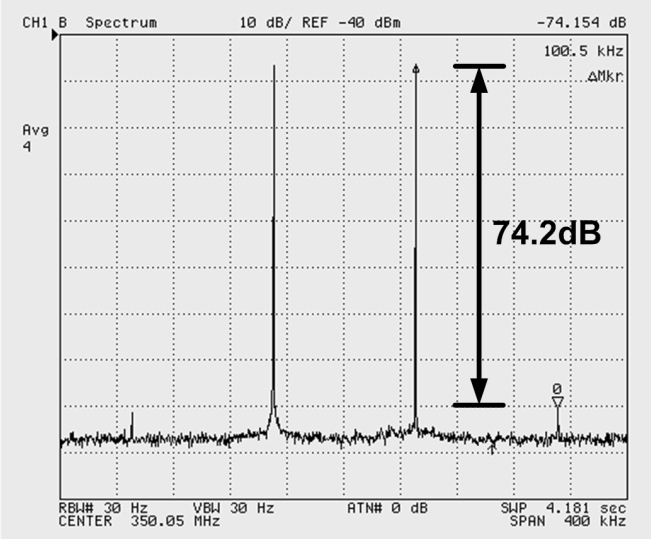
Table III. Measured main parameters of the reference folded-cascode OTA

Parameter	Measurement
Transconductance ( $G_m$ )	$510 \mu\text{A}/\text{V}$
IM3 @ 50MHz ( $V_{in} = 0.2 \text{ V}_{p-p}$ )	$-55.3 \text{ dB}$
Noise (input-referred)	$13.3 \text{ nV}/\sqrt{\text{Hz}}$
Power with CMFB	$2.6 \text{ mW}$
PSRR @ 50MHz	$48.9 \text{ dB}$
Supply	$1.2 \text{ V}$



**Uncompensated OTA IM3**  
**(input: 0.2V<sub>p-p</sub>@350MHz)**

(a)



**Compensated OTA IM3**  
**(input: 0.2V<sub>p-p</sub>@350MHz)**

(b)

Fig. 21. Measured linearity with 0.2V<sub>p-p</sub> input swing from two tones.  
(Each tone: 0.1V<sub>p-p</sub> (-16dBm) on-chip after accounting for off-chip losses at the input).  
Displayed outputs: (a) reference OTA, (b) compensated OTA.

The IM3 from the two-tone tests of the reference and linearized OTAs around 350MHz is plotted versus input peak-to-peak voltage in Fig. 22. This comparison demonstrates that the IM3 enhancement from the linearization scheme requires weakly nonlinear operation. Even though the linearization effectiveness decreases with increasing input signal swing, the IM3 improvement is still 11dB with  $0.8V_{p-p}$  differential signal swing for this design with 1.2V supply. Since the distortion cancellation exhibits the highest sensitivity to phase shifts at high frequencies, the control code of the phase shift resistor  $R_c$  in Fig. 12 has been changed from its optimum value. The resulting effect on the IM3 of the linearized OTA at 350MHz is plotted in Fig. 23, which validates that variable phase compensation is in fact required for optimum linearity performance. Two resistor ladder settings satisfy that the IM3 attenuation is more than 74dB, hence the selected 3% step for the least significant digital bit in this design was appropriate. Together with the plot obtained by sweeping resistor  $R_c$  in simulations (Fig. 20a), the measurements indicate that the amount of IM3 improvement predominantly depends on the step size of the programmable resistor ladder, which promises even better distortion cancellation with finer resolution.

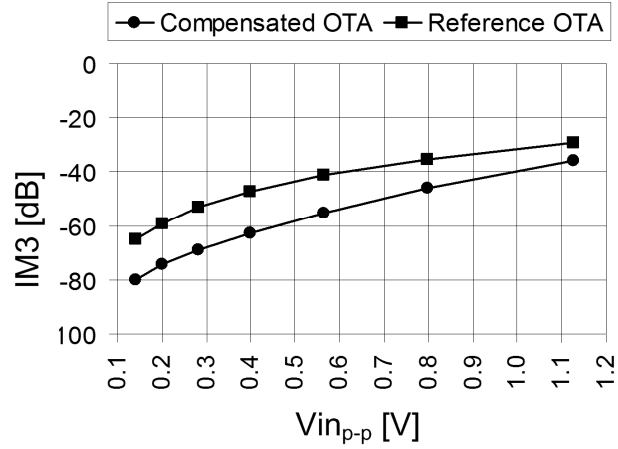


Fig. 22. IM3 vs. input voltage swing for reference OTA and compensated OTA. Obtained with two tones having 100KHz separation around 350MHz.

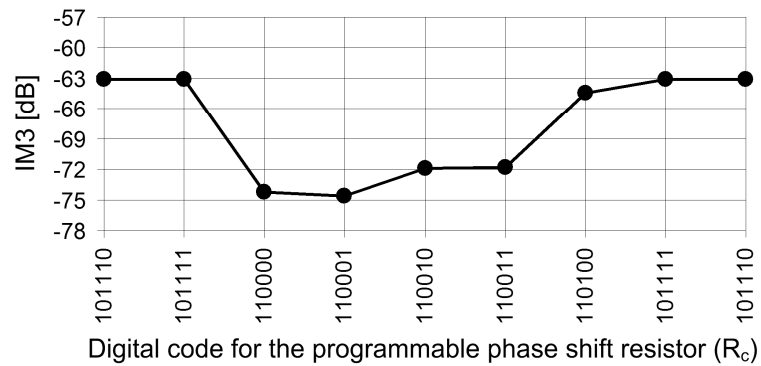


Fig. 23. Measured IM3 dependence of the compensated OTA on phase shift. Obtained with two test tones having 100KHz separation around 350MHz. The least significant bit of the digital control code changes  $R_c$  by  $\sim 3\%$ .

Table IV. Comparison of OTA linearity and noise measurements

OTA type	Input-Referred Noise	Power Consumption	IM3 ( $V_{in} = 0.2 V_{p-p}$ )			Normalized  FOM * (at 350 MHz)
			50 MHz	150 MHz	350 MHz	
<b>Reference</b> (input attenuation = 1/3)	13.3nV/ $\sqrt{\text{Hz}}$	2.6mW	-55.3dB	-60.0dB	-58.5dB	56.7
<b>Linearized</b> (attenuation = 1/3 & compensation)	21.8nV/ $\sqrt{\text{Hz}}$	5.2mW	-77.3dB	-77.7dB	-74.2dB	64.3

\* See Table V for details.

Table IV includes noise and IM3 measurement results at various frequencies, demonstrating the effectiveness of the broadband linearization scheme with the associated input-referred noise. Performance trade-offs can be assessed with the figure of merit from [44]:  $FOM = NSNR + 10\log(f/1\text{MHz})$ , where  $NSNR = SNR_{(dB)} + 10\log[(IM3_N / IM3)(BW / BW_N)(P_N / P_{dis})]$  from [45], the SNR is integrated over 1MHz BW, IM3 is normalized with  $IM3_N = 1\%$ , bandwidth is normalized with  $BW_N = 1\text{Hz}$ , and power consumption is normalized with  $P_N = 1\text{mW}$ . Experimental results are compared with previously reported architectures in Table V. The OTA linearized with input attenuation-predistortion shows a competitive performance with respect to the state of the art. High linearity at high frequencies is realized in this design example, showing the potential of the technique.

Table V. OTA comparison with prior works

	[39]*	[46]*	[48]	[47]	[43]*	This Work
<b>IM3</b>	-	-47dB	-70dB	-60dB	-	-74.2dB
<b>IIP3</b>	-12.5dBV	-	-	-	7dBV	7.6dBV
<b>f</b>	275MHz	10MHz	20MHz	40MHz	184MHz	350MHz
<b>Input Voltage</b>	-	0.2V <sub>p-p</sub>	1.0V <sub>p-p</sub>	0.9V <sub>p-p</sub>	-	0.2V <sub>p-p</sub>
<b>Power / Transconductor</b>	4.5mW	1.0mW	4mW	9.5mW	1.26mW	5.2mW
<b>Input-Referred Noise</b>	7.8nV/ $\sqrt{\text{Hz}}$	7.5nV/ $\sqrt{\text{Hz}}$	70.0nV/ $\sqrt{\text{Hz}}$	23.0nV/ $\sqrt{\text{Hz}}$	53.7nV/ $\sqrt{\text{Hz}}$	21.8nV/ $\sqrt{\text{Hz}}$
<b>Supply Voltage</b>	1.2V	1.8V	3.3V	1.5V	1.8V	1.2V
<b>Technology</b>	65nm CMOS	0.18 $\mu\text{m}$ CMOS	0.5 $\mu\text{m}$ CMOS	0.18 $\mu\text{m}$ CMOS	0.18 $\mu\text{m}$ CMOS	0.13 $\mu\text{m}$ CMOS
<b>FOM<sub>(dB)</sub>**</b>	87.5	92.9	96.1	99.1	100	105.6
<b>Normalized  FOM ***</b>	1.0	3.4	7.1	14.3	17.8	64.3

\* Power/transconductor calculated from filter power. Individual OTA characterization results not reported in full.

\*\* FOM<sub>(dB)</sub> = 10log( $f / 1\text{MHz}$ ) + NSNR from [44] ; NSNR = SNR<sub>(dB)</sub> + 10log[( IM3<sub>N</sub> / IM3 ) ( BW / BW<sub>N</sub> ) ( P<sub>N</sub> / P<sub>dis</sub> )] from [45].

( SNR integrated over 1MHz BW, normalization: IM3<sub>N</sub> = 1%, BW<sub>N</sub> = 1Hz , P<sub>N</sub> = 1mW )

( IM3 in FOM for [39] and [43] was calculated with: IM3<sub>(dB)</sub> = 2 x [ Pin<sub>(dBm)</sub> - IIP3<sub>(dB)</sub> ]. )

\*\*\* Normalized FOM magnitude relative to [39]: Normalized |FOM| = 10<sup>(FOM<sub>(dB)</sub>/10)</sup> / ( 10<sup>(FOM<sub>(dB)</sub>/10)</sup> of [39] )

### III.5.2. Second-order lowpass filter

Fig. 24 shows the filter frequency response for the proof-of-concept biquad design in Fig. 16, and its linearity performance is plotted against frequency in Fig. 25. The IM3 of the filter is up to 8dB worse than that of the standalone OTA. However, the measured filter IM3 includes approximately 2-3dB degradation due to the nonlinearity of the output buffer, which was not de-embedded from the measurement results. By adjusting the resistor ladders with digital controls that are common for all OTAs, the filter achieves IM3  $\approx$  -70dB up to 150MHz for a 0.2V<sub>p-p</sub> two-tone input. At 200MHz, which is

above the 194.7MHz filter cutoff frequency, the IM3 is -66.1dB, demonstrating the effectiveness of the broadband linearization due to compensation with the phase shifter.

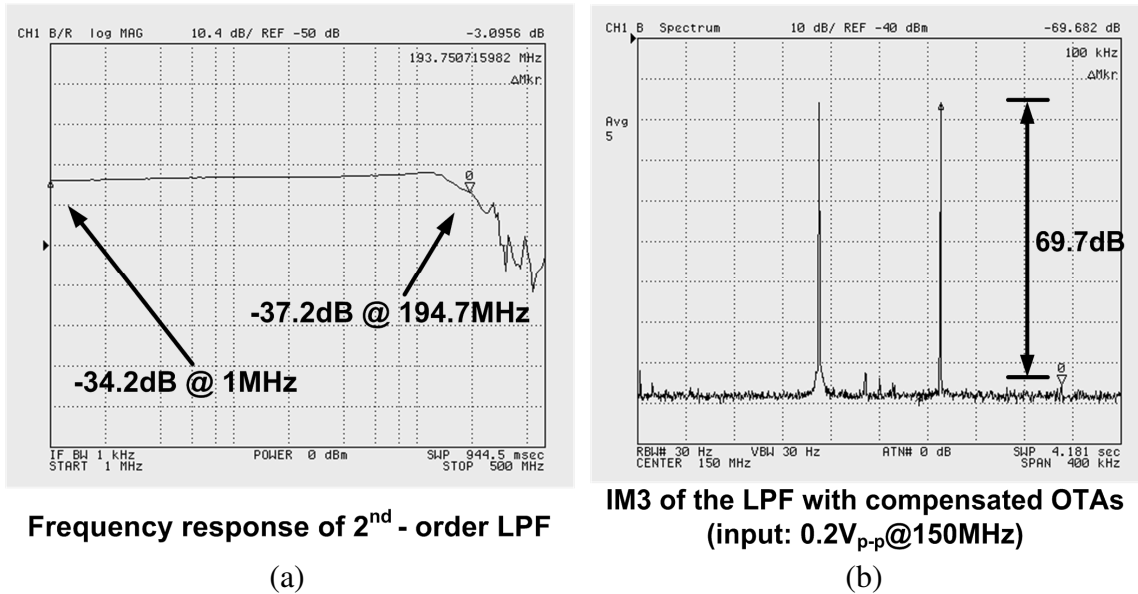


Fig. 24. Measured filter frequency response and linearity.  
 (a) Transfer function with ~34dB total losses (input loss and output buffer attenuation).  
 (b) IM3 with 0.2V<sub>p-p</sub> input swing from two tones, each 0.1V<sub>p-p</sub> (-16dBm) on-chip after accounting for off-chip input losses.

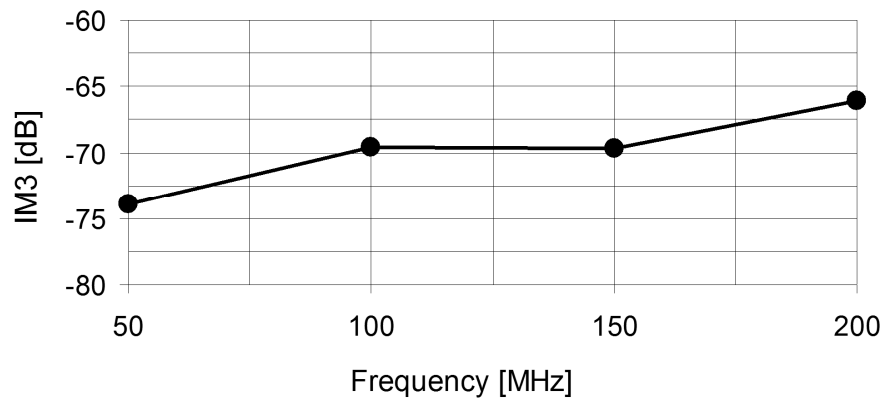


Fig. 25. Filter IM3 vs. frequency measured with two tones spaced by 100KHz.



Fig. 26 visualizes the measured IM3 with increasing input voltage up to 1.13V peak-peak differential swing, which follows the expected trend. At 150MHz, an IM3 of approximately -31dB occurs with an input signal of  $0.75V_{p-p}$ .

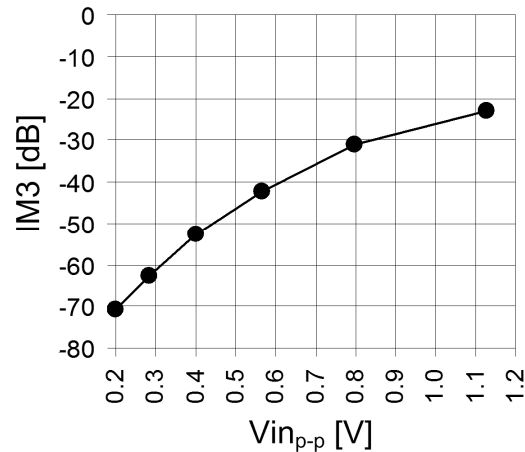


Fig. 26. IM3 vs. input peak-peak voltage for the linearized filter. Measured with two test tones separated by 100KHz around 150MHz.

Fig. 27 illustrates the in-band third-order intermodulation intercept point (IIP3 = 14.0dBm) and second-order intermodulation intercept point (IIP2 = 33.7dBm) curves measured with two tones separated by 100KHz around 150MHz and 2MHz, respectively. In broadband receiver applications with limited filtering in the RF front-end, the presence of numerous out-of-band interference signals results in intermodulation components within the desired signal band. Thus, high out-of-band linearity is desirable in addition to the baseband filter attenuation in order to minimize in-band distortion. This is one of the main motivations to employ OTAs with high linearity at high frequencies even for baseband filters with low cutoff frequencies. The

out-of-band IIP3 plot in Fig. 28a confirms that the linearization scheme's effectiveness is preserved beyond the cutoff frequency. The slight degradation of the out-of-band IIP3 to 12.4dBm is most likely due to the different phase shifts experienced by the 275MHz and 375MHz test tones from the input to node 2 in the auxiliary path (Fig. 12). The digital control code for the phase shift resistor  $R_c$  of the OTAs in the filter was set to optimize linearity in the 195MHz bandwidth, hence the linearity degradation due to the frequency difference of the out-of-band tones. The out-of-band IIP2 (Fig. 28b) is 30.4dBm, which is 3.3dB lower than the in-band IIP2 due to suboptimum phase shifts at 375MHz. Despite of that, the use of OTAs with high out-of-band linearity helps to reduce in-band distortion from out-of-band interferers in broadband scenarios.

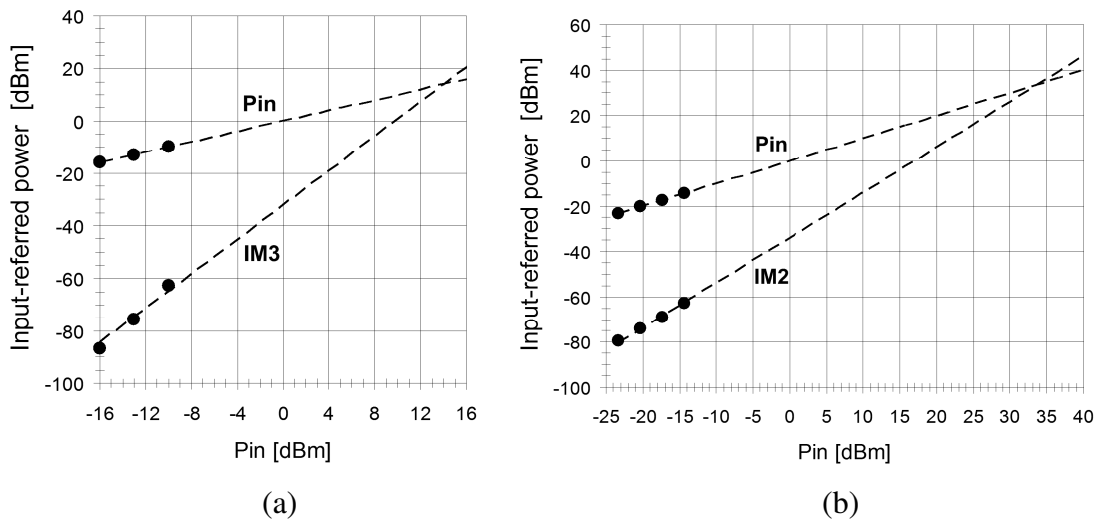


Fig. 27. Measured in-band intercept point curves for the filter.  
 (a) IIP3 [two tones,  $\Delta f = 100\text{KHz}$  around 150MHz],  
 (b) IIP2 [two tones,  $\Delta f = 100\text{KHz}$  around 2MHz].

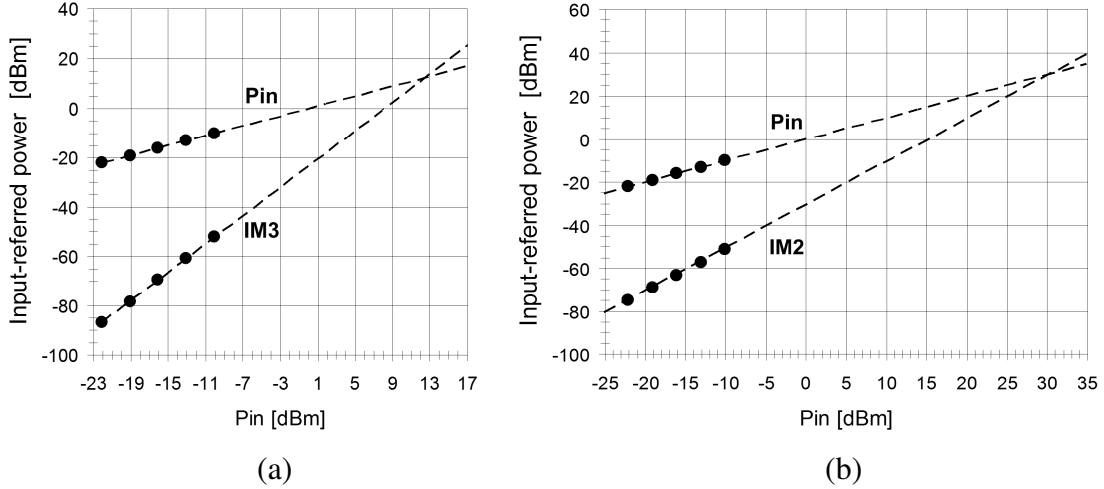


Fig. 28. Measured out-of-band intercept point curves for the filter.  
 (a) IIP3 [ $f_1 = 275\text{MHz}$ ,  $f_2 = 375\text{MHz}$ ,  $f_{\text{IM3}} = 100\text{MHz}$ ],  
 (b) IIP2 [ $f_1 = 375\text{MHz}$ ,  $f_2 = 375.1\text{MHz}$ ,  $f_{\text{IM2}} = 100\text{KHz}$ ].

Table VI summarizes the filter's key performance parameters in contrast to other wideband lowpass filters. The 54.5dB dynamic range integrated over the 195MHz noise bandwidth is competitive with prior works having similar power consumption per pole, most of which were implemented under less voltage headroom constraints than with the 1.2V supply in this design. The proposed linearization is independent of OTA topology, but the proof-of-concept design is comprised of a restrictive fully-differential OTA core in order to demonstrate the concept with a conventional topology. The last two columns in Table VI indicate that the proposed linearization allows almost similar filter linearity performance (in-band IIP3 = 14.0dBm with 1.2V supply) by means of fully-differential OTAs as with the pseudo-differential OTAs in [60], in which an in-band IIP3 of 16.9dBm was recently achieved with 1.8V supply. Apart from linearity considerations, the optimizations involving power consumption, input-referred noise, power supply

noise rejection, and CMRR depend on the application-specific constraints. According to the FOM comparison with the reference OTA in Table IV, the proposed linearization methods improves OTA linearity with justifiable power and noise trade-offs. Furthermore, the most best dynamic range improvement with the proposed technique can be achieved in bandpass designs, in which the noise is integrated over a narrow passband and the linearity improvement significantly reduces the power of the in-band distortion. The filter area on the die (Fig. 29) is  $\sim 0.5\text{mm}^2$  including the output buffer.

Table VI. Comparison of wideband  $G_m$ -C lowpass filters

	[39]	[43]	[56]	[57]	[58]	[59]	[60]	This work
<b>Filter Order</b>	5	5	8	4	7	5	3	2
<b><math>f_c</math> (max.)</b>	275MHz	184MHz	120MHz	200MHz	200MHz	500MHz	300MHz	200MHz
<b>Signal Swing</b>	-	$0.30V_{p-p}$	$0.20V_{p-p}$	$0.88V_{p-p}$	$0.80V_{p-p}$	$0.50V_{p-p}$	-	$0.75V_{p-p}$
<b>Linearity with max. <math>V_{in-p-p}</math></b>	-	HD3, HD5: $< -45\text{dB}$	THD: $-50\text{dB}$ @ 120MHz	THD: $-40\text{dB}$ @ 20MHz	THD: $-42\text{dB}$ @ 200MHz	THD: $< -40\text{dB}$ @ 70MHz	-	IM3: $-31\text{dB}$ **** @ 150MHz
<b>In-Band IIP3</b>	$-12.5\text{dBV}$ (0.5dBm)	$7\text{dBV}$ (20dBm)	-	-	-	-	$3.9\text{dBV}$ (16.9dBm)	$1.0\text{dBV}$ (14.0dBm)
<b>In-Band IIP2</b>	-	-	-	-	-	-	$19\text{dBV}$ (32dBm)	$20.7\text{dBV}$ (33.7dBm)
<b>Out-of-Band IIP3</b>	$-8\text{dBV}$ (5dBm)	-	-	-	-	-	-	$-0.6\text{dBV}$ (12.4dBm)
<b>Out-of-Band IIP2</b>	$15\text{dBV}$ (28dBm)	-	-	-	-	-	-	$17.4\text{dBV}$ (30.4dBm)
<b>Power</b>	36mW	12.6mW	120mW	48mW	210mW	100mW	72mW	20.8mW
<b>Power per Pole</b>	7.2mW	2.5mW	15mW	12mW	30mW	20mW	24mW	10.4mW
<b>Input-Referred Noise</b>	$7.8\text{nV}/\sqrt{\text{Hz}}$	$53.7\text{nV}/\sqrt{\text{Hz}}$ **	-	-	-	-	$5\text{nV}/\sqrt{\text{Hz}}$	$35.4\text{nV}/\sqrt{\text{Hz}}$
<b>Dynamic Range</b>	44dB*	43.3dB***	45dB	58dB	-	52dB	-	54.5dB***
<b>Supply Voltage</b>	1.2V	1.8V	2.5V	2V	3V	3.3V	1.8V	1.2V
<b>Technology</b>	65nm CMOS	$0.18\mu\text{m}$ CMOS	$0.25\mu\text{m}$ CMOS	$0.35\mu\text{m}$ CMOS	$0.25\mu\text{m}$ CMOS	$0.35\mu\text{m}$ CMOS	$0.18\mu\text{m}$ CMOS	$0.13\mu\text{m}$ CMOS

\* Reported spurious-free dynamic range. \*\* Calculated from  $9.3\mu\text{V}_{\text{RMS}}$  in 30KHz BW. \*\*\* Calculated from max.  $V_{p-p}$ ,  $f_c$ , and input-referred noise density. \*\*\*\* IM3 of  $-31\text{dB}$  measured close to  $f_c$  ensures THD  $< -40\text{dB}$ .

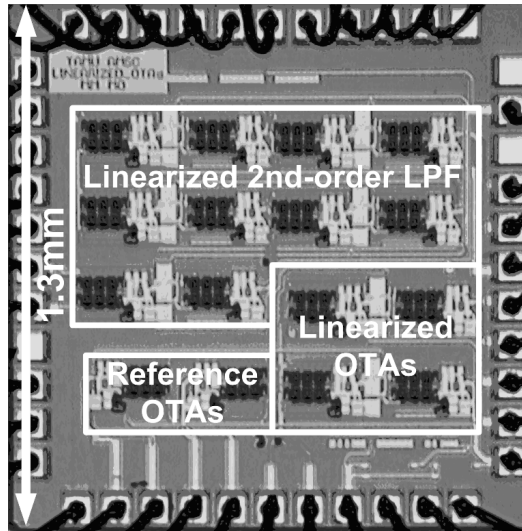


Fig. 29. Die micrograph of the OTAs and filter in  $0.13\mu\text{m}$  CMOS technology. Reference OTA area:  $0.033\text{mm}^2$ , linearized OTA area:  $0.090\text{mm}^2$ .

### III.6. Summarizing Remarks

An attenuation-predistortion technique was described to linearize transconductance amplifiers in  $G_m$ -C filter applications over a wide frequency range and across PVT variations. The high-frequency linearity enhancement is based on Volterra series analysis. Experimental results confirm the efficacy of the OTA linearization at high frequencies to obtain IM3 as low as  $-74\text{dB}$  with  $0.2V_{\text{in,p-p}}$  at  $350\text{MHz}$ . Measurements of a biquad demonstrated that the linearization methodology is suitable for  $G_m$ -C filter applications requiring an overall  $\text{IM3} \leq -70\text{dB}$  up to the cutoff frequency. The proposed linearization approach is independent of the OTA architecture and robust due to the use of matched OTAs to cancel output distortion, resulting in an IM3 improvement of up to  $22\text{dB}$ . Compensation for PVT variations and high-frequency effects is based on digital adjustment of resistors without changing the bias conditions, which would affect other design parameters. Hence, the main OTA can be optimized for its target application.

## IV. QUANTIZER DESIGN FOR A CONTINUOUS-TIME SIGMA-DELTA ADC WITH REDUCED DEVICE MATCHING REQUIREMENTS\*

### IV.1. Background

The quantizer under investigation was specifically designed as part of a continuous-time  $\Sigma\Delta$  modulator, which is an analog-to-digital converter (ADC) that uses oversampling and filtering to achieve quantization noise-shaping to obtain an effective number of bits (signal-to-quantization-noise ratio) significantly higher than the quantizer in the loop (e. g. a 12-bit ADC with a 3-bit quantizer). Such an ADC is visualized in Fig. 30 just to show the quantizer's location in the loop, where the most conventional quantizer is a flash ADC. Details regarding the operation and design of typical continuous-time  $\Sigma\Delta$  modulators are outside of the scope of this dissertation, but they can be found in [61].

---

\* © 2010 IEEE. Excerpts from Section IV are in part reprinted, with permission, from "A 25MHz bandwidth 5th-order continuous-time lowpass sigma-delta modulator with 67.7dB SNDR using time-domain quantization and feedback," C.-Y. Lu, M. Onabajo, V. Gadde, Y.-C. Lo, H.-P. Chen, V. Periasamy, and J. Silva-Martinez, *IEEE J. Solid-State Circuits*, vol. 45, no. 9, pp. 1795-1808, Sept. 2010.

This material is included here with permission of the IEEE. Such permission of the IEEE does not in any way imply IEEE endorsement of any of Texas A&M University's products or services. Internal or personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution must be obtained from the IEEE by writing to [pubs-permissions@ieee.org](mailto:pubs-permissions@ieee.org). By choosing to view this material, you agree to all provisions of the copyright laws protecting it.

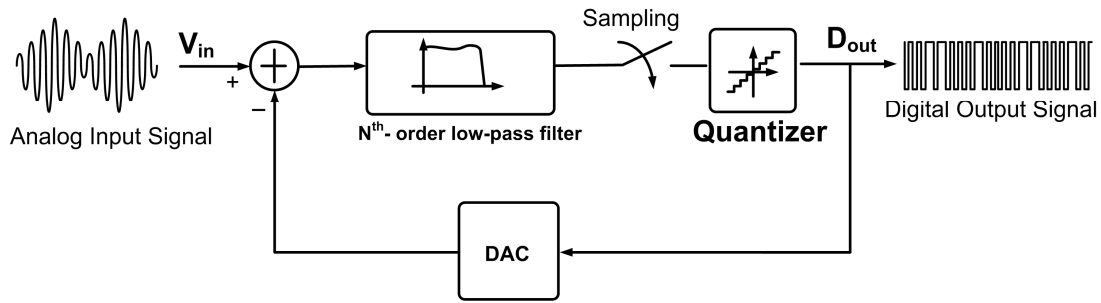


Fig. 30. Simplified diagram of a continuous-time  $\Sigma\Delta$  modulator.

#### IV.1.1. State of the art continuous-time $\Sigma\Delta$ ADCs

Various wireless standards such as WiMAX have been developed in recent years due to the high demand for faster data rate in portable wireless communications, which has pushed baseband bandwidths up to a few tens of megahertz. When high-resolution lowpass  $\Sigma\Delta$  ADC architectures are selected for emerging products because of their efficiency, a wide bandwidth is essential in multi-standard applications to accommodate receiver bandwidth requirements. A popular way to improve the signal-to-quantization-noise ratio (SNDR) over wide bandwidth without increasing the sampling frequency is to use a multi-bit quantizer and a multi-bit feedback digital-to-analog converter (DAC) [62]. With this approach, the noise-shaping gain required in the loop filter can be relaxed due to the reduced quantization noise associated with the multi-bit quantizer. Even though multi-bit architectures have been successfully utilized in multi-MHz bandwidth designs, the “digital friendly” advantages of the 1-bit architecture are typically compromised with the multi-bit solution. In particular, the feedback DAC nonlinearity significantly affects the ADC performance because it directly adds error to the filter input signal and it is not noise-shaped. Dynamic element matching (DEM) and data

weighted averaging (DWA) techniques have been proposed to tackle this problem [63]-[65]. However, the additional power and complexity of DEM methods is not permissible in some applications. In a more recent work [66], the thresholds of the comparators in a 9-level quantizer were shuffled rather than performing DAC element rotations, which shortens the delay for the mismatch-shaping realization. The feasibility of this method has been demonstrated in a modulator having 82dB SNDR over 10MHz bandwidth with a 5<sup>th</sup>-order loop filter. In general, the shaping of the mismatch error provided by DEM/DWA techniques is less effective for designs with low oversampling ratio (OSR) and high conversion speeds. On the contrary, the line of attack in this work is to prevent DAC element matching issues altogether by using a multi-bit single-element DAC. This strategy, on the other hand, necessitates accurate digital timing circuitry, which is a trade-off whose attractiveness parallels technology scaling.

Recent practical works have incorporated a digital-intensive time-based multi-bit quantizer [67] and quantizer/DAC combination [68] in the modulator architecture, achieving 72dB SNDR and 60dB SNDR over 10MHz and 20MHz bandwidths, respectively. Since scaling of CMOS process technologies provides an advantageous environment for high-speed digital timing control but perilous conditions for analog device matching in the DAC/quantizer, the time-based approaches and pulse-width modulation (PWM) feedback DACs are promising solutions for future technologies. The recent simulation results for the designs in [69] provide further insights into the effectiveness of this design methodology. In anticipation of increasing process variations, the approach taken in this work involves a 3-bit quantizer and a single-



element DAC that realizes 3-bit feedback via time-based operation (generation of a PWM waveform). Hence, the need for DAC unit element matching or DEM/DWA techniques is eliminated. However, time-based approaches require strict control over the timing signals and clock jitter to attain high SNDR. The main trade-off is that the DAC linearity depends on the mismatches between the clock phases for the PWM waveform generation rather than unit element mismatches as in conventional multi-bit DACs.

#### *IV.1.2. Quantizer design trends*

Fig. 31 displays a typical 3-bit flash quantizer, in which the input signal  $V_{in}$  is compared to seven reference voltage levels (obtained with a resistor ladder) using seven comparators ( $C_1$ - $C_7$ ). For high-speed operation, the comparators are often comprised of preamplifiers followed by latches. The quantization occurs in one clock cycle, yielding thermometer code as output that can be converted to the desired digital output code with an encoder. With regards to PVT variations, a relevant condition is that the resistors ( $R$ ) must be matched in the layout to avoid shifts in the reference voltage levels. Similarly, the input-offset voltages of the comparators are subjected to PVT variations, in particular through the worsening threshold voltage variations (Table I on page 12). Compensating for these variations and the resulting offsets that cause ADC nonlinearity errors is an ongoing research topic to which many solutions have been proposed over the past decades. Similar to transceiver system calibration approaches (Section II.2.5), recently proposed methods involve calibration control in the digital domain in combination with programmable circuit element through the use of switches. In [70] for example, additional resistors are included in the reference voltage ladder to generate extra voltage

levels between the ideal references. The best combination of references is selected with switches and a digital control scheme in order to compensate for offsets from process variations, improving the effective number of bits of the flash ADC from 3 to 5.6. Another recently proposed digital calibration technique ([71]) employs programmable load resistors in the differential preamplification stage within the comparators in order to make adjustments that counteract random offsets. To maintain compatibility with such digital calibration methods, the quantizer architecture introduced in this dissertation has been designed to allow reference voltage tuning without affecting components that are directly in the signal path.

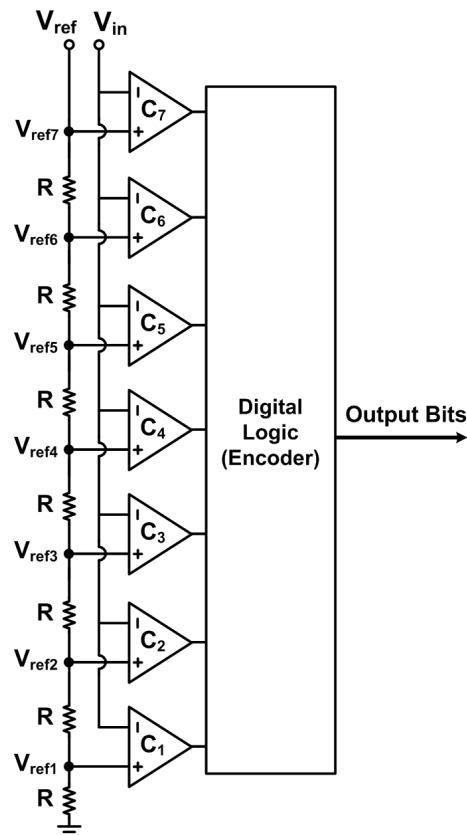


Fig. 31. Conventional 3-bit flash quantizer.

Traditional two-step flash analog-to-digital converter (ADC) architectures are a subset of subranging ADCs that typically consist of a sample-and-hold (S/H), a most-significant bit(s) (MSB) ADC, a digital-to-analog converter (DAC), a gain block, and a least-significant bit(s) (LSB) ADC [72]. As an example, the adapted block diagram of the two-step ADC described in [72] is displayed in Fig. 32, which utilizes two DACs and does not require amplifiers. Conceptually, the operation is as follows: After the input signal is sampled by the S/H circuit, the MSBs are resolved using a fixed reference voltage range ( $V_{ref1}$ ). Next, DAC<sub>1</sub> generates the upper reference voltage ( $V_{ref2a}$ ) for the decision with the LSB ADC by incrementing the quantized MSB value by one in the digital domain. The lower range for the LSB decision is set with DAC<sub>2</sub>, which directly converts the quantized MSB into an analog voltage ( $V_{ref2b}$ ). With the selected reference voltage subrange, the LSB ADC performs a fine quantization of the sampled input voltage. Such a two-step flash approach has the advantage that the output bits from two low-resolution ADCs can be combined to obtain more precision, reducing the number of comparators that a conventional flash ADC would require for the same resolution. Hence, multi-step quantization can be used to lower area and power consumption when a delay of multiple clock cycles or clock phases can be tolerated.

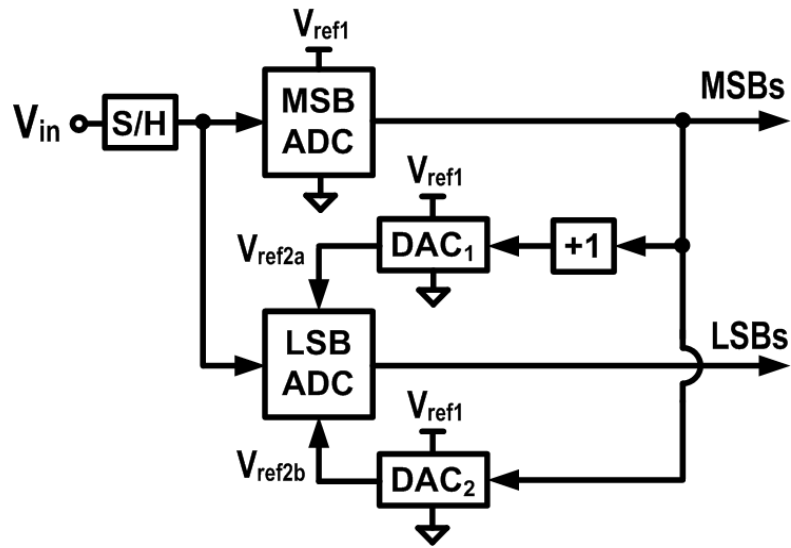


Fig. 32. The two-step ADC principle.

In the past years, several alternative quantizer architectures have been proposed to optimize the operation by taking advantage of technology scaling for enhanced performance at higher conversion speeds; reducing power consumption, and improving compatibility with digital CMOS processes. However, design challenges also arise from adverse effects in deep-submicron technologies such as reduced gains from lower transistor output impedances, design with limited voltage headroom, reduced transistor linearity, and increased PVT variations as well as intra-die variability. As a result, recent works involved quantizer design trade-offs that exploit the advantage of modern CMOS processes while avoiding the drawbacks. For instance, the folding flash ADC in [73] is comprised of 16 instead of 31 (conventional flash) comparators for 5-bit resolution to decrease the power consumption. In addition, the folding topology in [73] circumvents the use of amplifiers in 90nm CMOS technology, which increases its attractiveness with regards to scaling and integration. With the availability of fast-switching devices,

successive approximation ADCs are not constrained to low-speed operation anymore as demonstrated by realizations with low- to medium-resolution at medium- to high-speed [74]-[76]. The 6-bit 600MS/s ADC in [74] exemplifies how asynchronous processing can be utilized to shorten the comparison cycles when employing multiple comparisons to resolve the bits from MSB to LSB sequentially. In [74], the asynchronous successive approximations are performed with a single comparator by weighing the input against a reference that is dynamically changed with a switchable capacitor array before each comparison. With similar operation, a two-step 7-bit ADC having a 150MS/s conversion rate is described in [75], where the MSB is quantized first and the remaining bits are determined with an asynchronous binary-search procedure. In [76], the successive approximations with the sampled input are made via charge-sharing that occurs while cycling through a binary-scaled capacitor array. With the comparator being the only active block, power consumption below 0.7mW was achieved with the 9-bit ADC at conversion rates up to 50MS/s. When a multi-bit lowpass  $\Sigma\Delta$  modulators is designed with a high oversampling ratio, then the sample-to-sample voltage changes of the slow-varying input signal are small. Therefore, only a small number of comparators connected to the reference voltages above and below the current signal level are required in consecutive conversions with a conventional flash architecture. This characteristic can be exploited to reduce the number of comparators by either shifting the references associated with the reduced number of comparators or by shifting the input signal prior to the comparison. In [77] the lowpass  $\Sigma\Delta$  modulator with 104MHz sampling frequency and 2MHz bandwidth for instance, a tracking ADC with 3 comparators was used in lieu

of a conventional 4-bit flash quantizer that would require 15 comparators, which shows how quantizer operation can be optimized for its application in a specific  $\Sigma\Delta$  modulator architecture.

#### *IV.1.3. Quantizer design considerations for the $\Sigma\Delta$ modulator architecture*

When designing high-resolution lowpass multi-bit  $\Sigma\Delta$  ADCs in modern CMOS technologies with rising process variations, the linearity performance of the feedback DAC at the ADC input becomes a limiting factor for the overall performance because its nonlinearity errors are not noise-shaped by the loop dynamics. The quantizer in this work has been designed as part of a group project in which an alternative multi-bit feedback approach was explored by constructing an architecture that does not rely on unit element matching in the front-end DAC. Instead, it employs an inherently linear single-element PWM DAC that is controlled via multi-phase clock signals. The general aim of this approach is to circumvent analog device matching requirements by relying more on well-timed digital operations. Fig. 33 depicts the fully-differential 5<sup>th</sup>-order lowpass  $\Sigma\Delta$  modulator with a sampling frequency of 400MHz for 25MHz signal bandwidth. A 5<sup>th</sup>-order quasi-linear phase inverse Chebyshev lowpass filter with 49dB pass-band gain is employed, which consists of two cascaded active-RC 2<sup>nd</sup>-order lowpass sections and a lossy integrator with sufficient linearity. The summing amplifier ( $\Sigma$ ) couples all feedforward paths of the filter to the quantizer input. A level-to-PWM converter translates the multi-bit signal into a time-domain digital PWM signal such that only a 1-bit current-steering DAC is required for global feedback with 3-bit equivalence. This realization avoids performance degradation originating from current mismatch



input of the loop filter are generated by adding the outputs of the appropriate number of unit element current sources for the quantizer output code. Device mismatches from process variations generate out-band noise that folds into the frequency range of interest as well as in-band harmonic distortion components that degrade the modulator's SNDR. Solutions such as noise-shaping dynamic element matching (DEM) [63], tree-structure DEM [64], and the data weighted averaging technique [65] were proposed in the past to reduce the DAC linearity degradation from mismatch. However, improvements in wideband ADCs are usually limited due to restrictions on loop delay and increased noise levels from the randomization procedure. In this work, a single-element DAC having an output waveform with variable pulse width per sampling period generates a 3-bit charge injection feedback as shown in Fig. 34. Since only one inherently linear single-bit DAC produces different feedback charge levels at the loop filter input, the current mismatch problem of multi-amplitude DACs is avoided. A level-to-PWM converter is implemented in the feedback path to convert the digital codes from the 3-bit quantizer to time-domain PWM signals compatible with the 1-bit DAC having time-varying output pulses of current amplitude  $\pm I$ . The PWM DAC output pulse shapes are arranged as symmetric as possible within a clock period to minimize the power of potential aliasing tones [69]. These pseudo-symmetric high and low amplitude levels of the single-element DAC during one clock period are also visualized in Fig. 34 together with their binary equivalent codes.



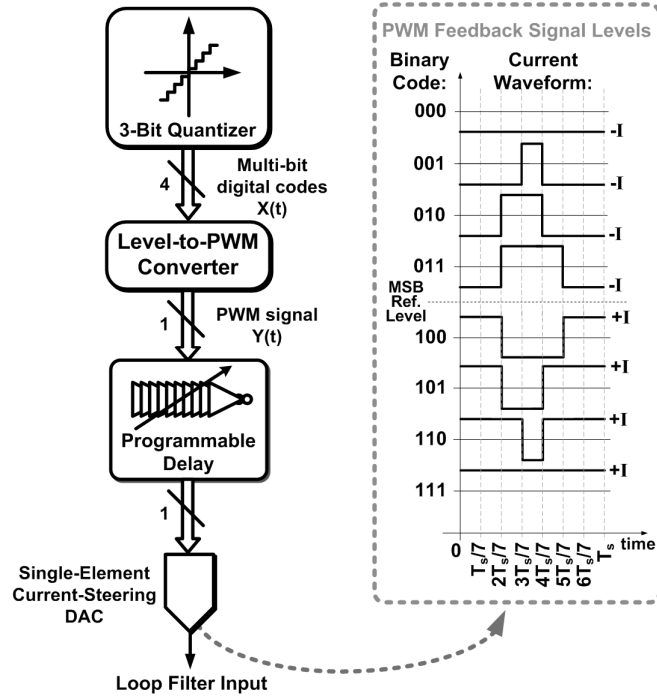


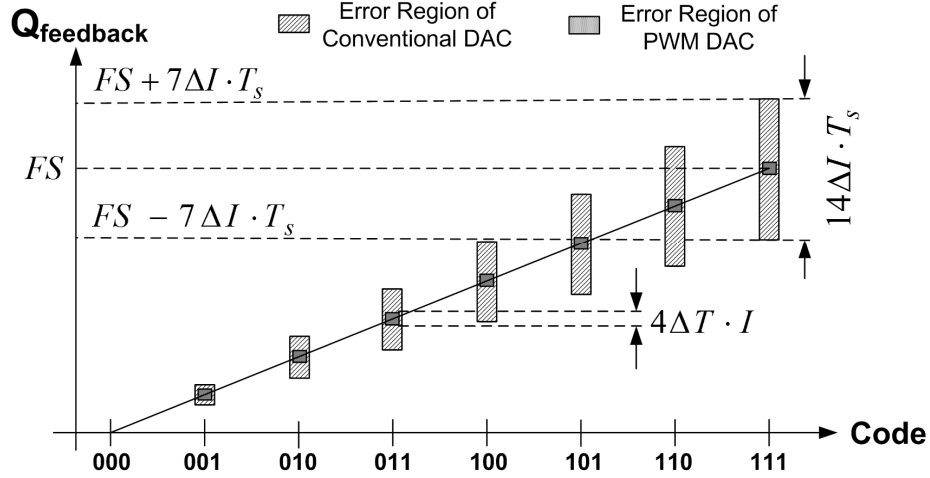
Fig. 34. Feedback path with 3-bit quantizer and PWM DAC.

The main drawback of employing multi-phase time-domain signals is increased sensitivity to jitter noise because of larger and more frequent DAC output transitions compared to a conventional 3-bit NRZ DAC. In general, the maximum signal-to-jitter-noise ratio (SJNR) of the modulator can be analytically estimated for any feedback pulse shape with [61]:

$$SJNR_{peak} = 10 \cdot \log_{10} \left( \frac{T_s^2 \cdot OSR}{2 \cdot \sigma_y^2 \cdot \sigma_\beta^2} \right), \quad (9)$$

where  $OSR = 1/(2 \cdot BW \cdot T_s)$ ,  $\sigma_\beta$  is the clock jitter standard deviation, and  $\sigma_y$  is the standard deviation of  $[y(n) - y(n-1)]$ ; with  $y(n)$  being the  $n^{\text{th}}$  combined digital output of the modulator. The SJNR of the modulator with level-to-PWM converter was evaluated in

comparison to a conventional 3-bit modulator [79], showing that the simulated SJNR limit of the PWM DAC with  $\sigma_\beta \approx 0.5\text{ps}$  is 5dB lower than that of a conventional 3-bit NRZ DAC at 400MHz. Furthermore, the worst-case clock jitter requirement for SNDR > 68dB with the proposed modulator is  $\sigma_\beta < 0.54\text{ps}$  [79].



$ \Delta Q $	000	001	010	011	100	101	110	111
Conventional DAC	0	$\Delta I_1 \cdot T_s$	$\sum_{i=1}^2 \Delta I_i \cdot T_s$	$\sum_{i=1}^3 \Delta I_i \cdot T_s$	$\sum_{i=1}^4 \Delta I_i \cdot T_s$	$\sum_{i=1}^5 \Delta I_i \cdot T_s$	$\sum_{i=1}^6 \Delta I_i \cdot T_s$	$\sum_{i=1}^7 \Delta I_i \cdot T_s$
PWM DAC	0	$(\Delta T_4 + \Delta T_5) \cdot I$	$(\Delta T_3 + \Delta T_5) \cdot I$	$(\Delta T_3 + \Delta T_6) \cdot I$	$(\Delta T_3 + \Delta T_6) \cdot I$	$(\Delta T_3 + \Delta T_5) \cdot I$	$(\Delta T_4 + \Delta T_5) \cdot I$	$2\Delta T_1 \cdot I$

Fig. 35. Relative 3-bit DAC linearity error comparison: conventional vs. PWM.

The nonlinearity of the PWM DAC due to static timing mismatches can be assessed from a feedback charge error comparison relative to the conventional 3-bit DAC. Fig. 35 visualizes the worst-case peak-to-peak charge errors for each code, which are resultants of static mismatch  $\Delta I_i$  for each current  $\cdot T_s$  cell in the conventional DAC and static timing error  $\Delta T_j$  of clock phase  $\Phi_j$  in the PWM DAC.  $\Delta T_j$  originates from static CILFD

mismatches and unequal propagation delays due to routing parasitics, but it does not accumulate in the inverter chain because each stage is locked to the VCO signal. The ideal feedback charge per code is identical for both DACs. Notice that the errors depend on mismatches in up to seven unit elements of the conventional DAC, but only up to two timing phases with the PWM scheme in which two edges define the area under the pulse regardless of the deviations that the phases in between have. Assuming equal mismatches ( $\Delta I_i = \Delta I$ ,  $\Delta T_j = \Delta T$ ) yields worst-case errors of  $\pm 7\Delta I \cdot T_s$  and  $\pm 2\Delta T \cdot I$  for conventional and PWM DACs, respectively. Letting  $\delta_{\%I} = \Delta I / (I/7)$  and  $\delta_{\%T} = \Delta T / (T_s/7)$  be the percent standard deviations of the mismatches in each case, the worst-case accumulated errors are  $\Delta Q_{conv.-worst} = \pm 7\delta_{\%I} \cdot (I/7) \cdot T_s$  and  $\Delta Q_{PWM-worst} = \pm 2\delta_{\%T} \cdot I \cdot (T_s/7)$ . Monte Carlo simulations including delay mismatches in all clock phases showed that  $\delta_{\%T} = 0.16\%$  as a result of the synchronizing effect from the injection-locking. Since  $\delta_{\%I}$  is typically 0.5% with good layout practices for a standard DAC, the anticipated worst-case linearity error of the PWM DAC is favorably lower. Assuming that two timing mismatches are accumulated in the case of the PWM-based ADC, all mismatches in the conventional realization are accumulated, and errors are un-correlated in both cases; the induced third-order harmonic distortion (HD3) ratio can be estimated as derived in [79]:

$$\frac{HD3_{PWM}}{HD3_{conventional}} \cong \left( \sqrt{\frac{2}{N}} \right) \left( \frac{\delta_{\%T}}{\delta_{\%I}} \right), \quad (10)$$

where  $N$  is the number of DAC levels. For  $N = 7$  and the aforementioned distributions, the linearity of the proposed PWM DAC theoretically outperforms the conventional DAC by 15.3dB according to (10). It is important to note that this estimated

improvement is based on the timing mismatch prediction from Monte Carlo simulations of this particular clock generation circuitry, and that nonidealities such as supply noise and ground bounce should be minimized to avoid PWM DAC linearity degradations due to timing errors in the digital circuitry.

## **IV.2. 3-Bit Two-Step Current-Mode Quantizer Architecture**

### *IV.2.1. Quantizer design*

As illustrated in Fig. 36, the quantizer utilizes the seven on-chip clock phases to control four sequential comparison instances ( $\tau_1$ - $\tau_4$ ), which cuts the number of comparators from seven to four with respect to a typical 3-bit flash ADC. The two-step process makes the MSB available after the first step, creating timing margin for the digital control logic that sets up the PWM DAC. Successive approximations during the second step resolve the remaining bits that are processed by the level-to-PWM converter. As a result, and similar to the combination of the PWM generator and TDC in [68], the 1-bit DAC is driven by a PWM waveform. However, in the approach presented here, successive approximations are employed for comparison with the input signal rather than generation of a continuous ramp. Since this successive algorithm only has one MSB and three LSB quantization steps, the comparison to discrete reference levels is a simple alternative that also gives the option to calibrate each level individually if necessary.

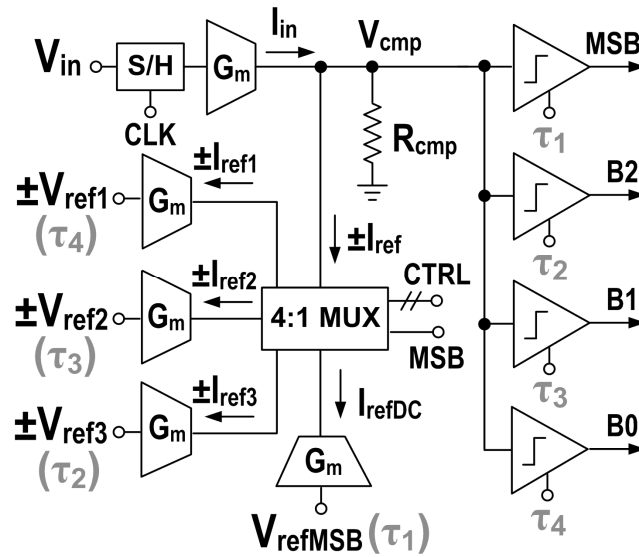


Fig. 36. Single-ended equivalent block diagram of the quantizer.

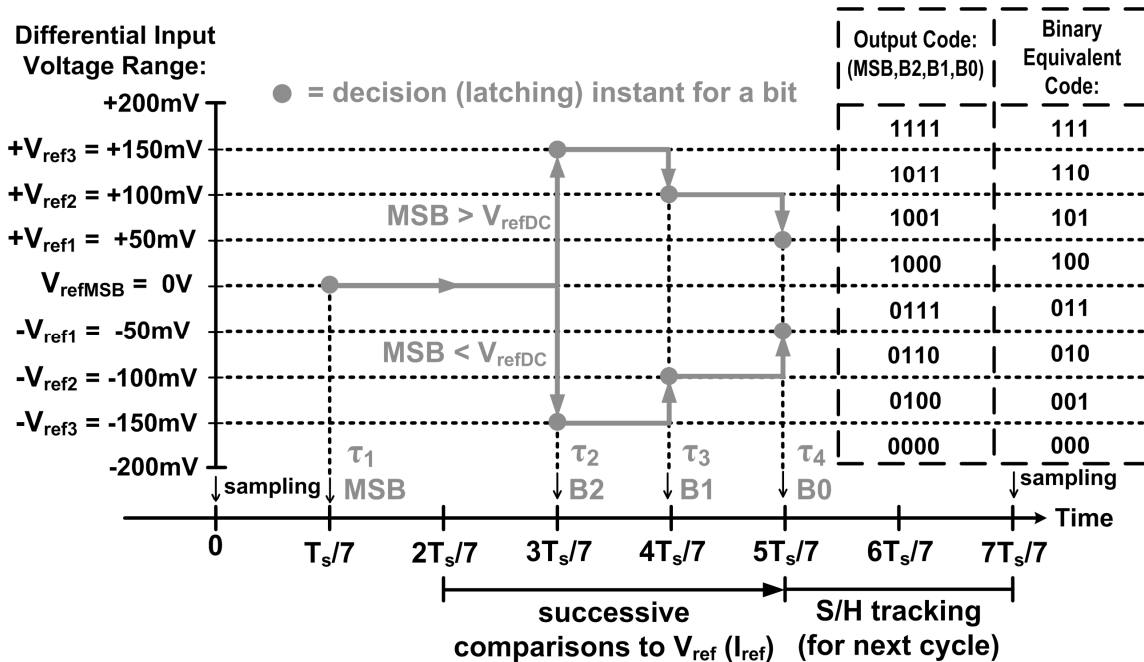


Fig. 37. Timing of the successive quantization decisions and output code words. The arrows show the two possible sequences based on the MSB value.

### Decision timing

The quantizer operates as follows with regards to the topology in Fig. 36 and corresponding timing diagram in Fig. 37. The differential input signal  $V_{in}$  is sampled with a S/H circuit by the 400MHz master clock having a period  $T_s$ , and then it is converted to current  $I_{in}$  via a transconductance stage ( $G_m$ ). First, the MSB is resolved after  $\tau_1$  seconds by comparing  $I_{in}$  to the current from  $V_{refMSB}$  applied to an identical  $G_m$  stage. Depending on the timing control bits (CTRL) and the MSB decision, a multiplexing configuration (MUX) is utilized to compare  $I_{in}$  to current  $I_{ref}$  derived from the appropriate differential reference voltage ( $\pm V_{ref1} \dots \pm V_{ref3}$ ) during each subsequent instant ( $\tau_2$ - $\tau_4$ ). The order of the subranging comparisons and output bits was chosen based on the timing needs in the multi-phase DAC control circuitry because larger signal magnitudes require DAC feedback pulse changes early in the next clock cycle. Comparison resistor ( $R_{cmp}$ ) converts the difference in currents into a positive or negative voltage. A binary result of the current-mode comparison is stored using a latched comparator for each of the four decisions. The tabular inset in Fig. 37 lists the output codes corresponding to the input ranges.

### Circuit-level design considerations

Fig. 38 displays the schematic of the quantizer core in which the current-mode comparisons are made. All devices with the same names are equal-sized and matched in the layout. The simplified S/H circuit represents a transistor-level implementation with gate-bootstrapping [80], and the AND gates effectively function as time-controlled MUX. After the S/H operation, the differential input voltage is converted to current by

the transistor pair ( $M_n$ ) and mirrored 1:1 by pair  $M_p$ . The other  $M_n$  transistors convert the differential reference voltages to currents for successive comparisons, where the difference current flows through the load resistors  $R_{cmp}$  to generate  $V_{cmp} = V_{cmp+} - V_{cmp-}$ . In this fully-differential circuit,  $V_{refMSB} = 0V$  (MSB decision) level is obtained by applying the DC voltage  $V_{refCM}$  that is equivalent to the 1.1V common-mode level at the input of the quantizer to both transistors in one of the branches for comparison with the input signal. The other differential reference voltages listed below Fig. 38 were selected to span the  $400mV_{p-p}$  full-scale swing at the quantizer input. For each reference current step, the polarity of this differential voltage is resolved by the latched comparator.

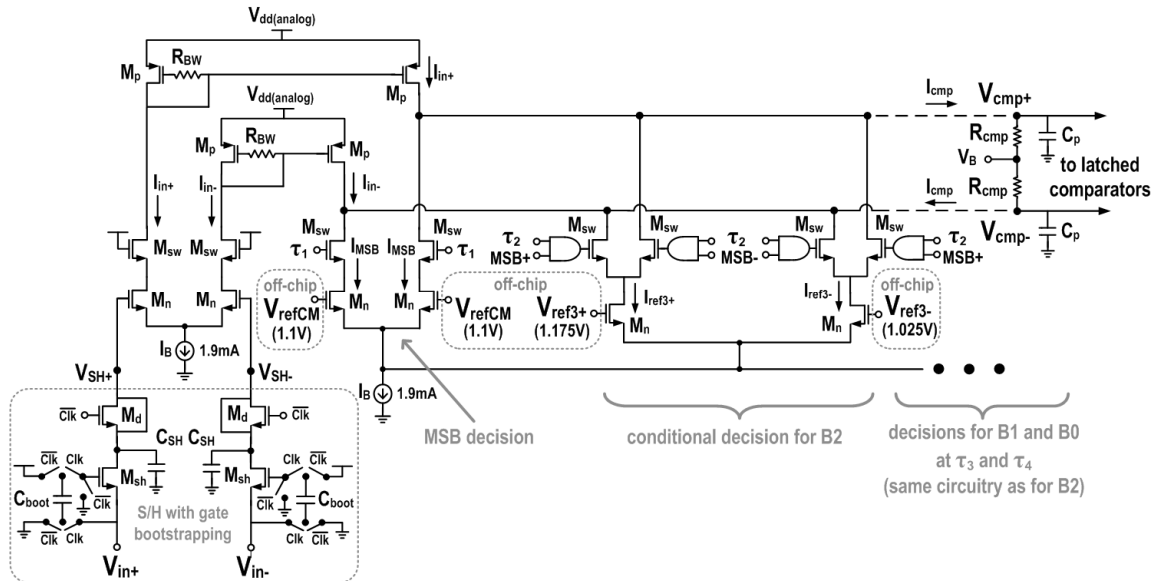


Fig. 38. Simplified schematic of the current-mode quantizer core circuitry. Reference voltages  $\pm V_{ref3} = \pm 150mV = \pm(V_{ref3+} - V_{ref3-}) = \pm(1.175V - 1.025V)$  and  $\pm V_{refMSB} = 0V = V_{refCM} - V_{refCM} = 1.1V - 1.1V$  are shown. The other references are:  $\pm V_{ref2} = \pm 100mV = \pm(1.15V - 1.05V)$ ,  $\pm V_{ref1} = \pm 50mV = \pm(1.125V - 1.075V)$ .

Polysilicon resistors ( $R_{BW}$ ) in Fig. 36 extend the bandwidth of the current mirrors [81] for high-frequency operation according to:

$$BW_{mirror} = \frac{1}{2\pi} \sqrt{g_{mp} / (R_{BW} \cdot C_{gsp}^2)} , \quad (11)$$

where  $g_{mp}$  and  $C_{gsp}$  are the transconductance and gate-source capacitance of  $M_p$ , correspondingly. With  $R_{BW} = 330\Omega$ , the simulated 3-dB bandwidth of the current mirrors is 3.36GHz, which is sufficiently high to prevent it from becoming the factor that limits the comparison speed. More critical is that speed performance is ensured by selecting the value of resistors  $R_{cmp}$  such that the RC time constant formed with parasitic capacitance  $C_p$  at the comparison nodes ( $V_{cmp+}$ ,  $V_{cmp-}$ ) does not impose limitations. After switch  $M_{sw}$  closes to compare the current from the input signal with the corresponding reference in each comparison cycle, the difference current  $I_{cmp} = I_{cmp+} - I_{cmp-}$  will cause a step response at the input of the latches ( $V_{cmp} = V_{cmp+} - V_{cmp-}$ ). With a first-order model, this step response can be expressed as

$$V_{cmp}(s) = \frac{2 \cdot R_{cmp}}{1 + s \cdot R_{cmp} \cdot C_p} \times \frac{I_{cmp}}{s} , \quad (12)$$

where  $s = j\omega$  and  $C_p$  is the cumulative parasitic capacitance at the comparison node from transistors  $M_p$ ,  $M_{sw}$ , input devices of the four latches, as well as routing parasitics. Taking the inverse Laplace transform of (12) gives the transient response during each comparison phase:

$$V_{cmp}(t) = 2 \cdot I_{cmp} \cdot \left( R_{cmp} - R_{cmp} \cdot e^{-t/(R_{cmp} \cdot C_p)} \right) . \quad (13)$$



Fig. 39 displays one sampling clock cycle of the simulated transient behavior at the comparison node, where the polarity (delineated by marker A on the  $V_{cmp+} - V_{cmp-} = 0V$  line) of the differential voltage is latched on the falling edge of the shown timing signals that correspond to  $\tau_1$ - $\tau_4$  in Fig. 37. The latching instants are labeled with arrows, resulting in an output code of (MSB, B2, B1, B0) = (1, 0, 0, 0) for this example quantization cycle. Note,  $V_{cmp(t)}$  settles within 5% of its final value after approximately three  $R_{cmp}C_p$  time constants. In this design,  $R_{cmp}$  is  $405\Omega$  and  $C_p$  is approximately 250fF, resulting in a theoretical time constant of 100ps. Nevertheless, it is only critical that  $V_{cmp}$  is larger than the resolution of the latch that resolves whether  $V_{cmp}$  is positive or negative. This zero-crossing event must occur sufficiently early to allow pre-charging of the nodes inside the activated latch by its preamplifier within the  $T_s/7$  comparison time window of the LSBs. If the aforementioned zero-crossing is delayed due to the large parasitic capacitance ( $C_p$ ) or insufficient preamplification prior to latching, then false decisions could occur. Hence, the timing and signal amplitude at this comparison node is the most significant factor affecting the quantizer resolution. Note, other factors such as the switch turn-on delay (of  $M_{sw}$  in Fig. 38), finite rise/fall times of the control signals, delay variations of the control signals, clock jitter, and kickback from the latches also impact the decision accuracy and cause the deviation of the  $V_{cmp}$  signal waveform in Fig. 39 from the ideal sequence of step responses.

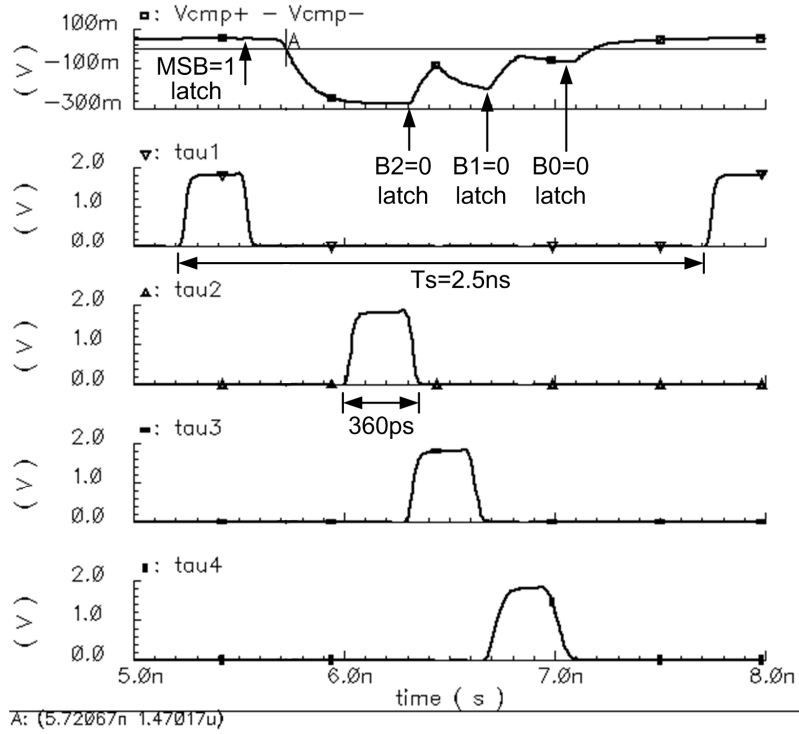


Fig. 39. Simulated example of the quantization timing. From top to bottom: transient voltage at the comparison node ( $V_{cmp} = V_{cmp+} - V_{cmp-}$ ), signals  $\tau_1$ - $\tau_4$  that trigger latching on the falling edge.

The clocked comparators connected to  $V_{cmp}$  in Fig. 36 are implemented with the fully-differential circuit shown in Fig. 40. In the tracking phase,  $\Phi_{LA}$  is low and bias current  $I_B$  is steered into the preamplifier stage consisting of input transistor  $M_1$  and load resistor  $R_{L1}$ . To save power, the bias current is reused in the latch phase (high  $\Phi_{LA}$ ) when it flows into  $M_{LA1}$ . Devices  $M_2$ ,  $R_{L2}$ ,  $M_{LA2}$  form a second preamplification and latch stage, but this stage is controlled by the phase-reversed latch signal to hold the decision for almost one clock period ( $T_s$ ). Transistors  $M_7$ - $M_{10}$  form a self-biased differential amplifier [82] which creates a rail-to-rail output during the long latch phase to drive the subsequent CMOS inverter ( $M_P$ ,  $M_N$ ).

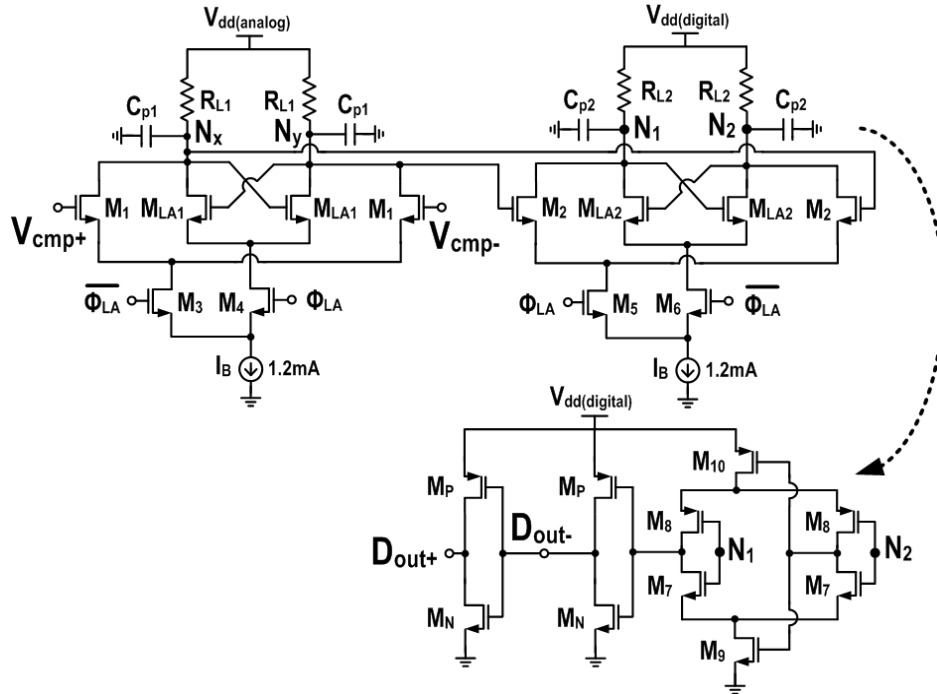


Fig. 40. Schematic of the latched comparator.

The preamplifier and first latch stage also play an important role in the quantizer operation, impacting the overall resolution and speed that can be achieved. First of all, the input transistor  $M_1$  in Fig. 40 should be as small as permissible to avoid introduction of excessive capacitance at the output node of the current-mode comparator core. The associated trade-off with small dimensions is increased input offset, which should be assessed via statistical simulations. Secondly, the bandwidth of the preamplifier must be high to avoid delay. In this design, its first pole is around 3.5GHz with  $R_{L1} = 570\Omega$  and  $C_{p1} = 80\text{fF}$  including routing parasitics. With sufficient preamplifier bandwidth margin, the most critical timing constraint is the propagation delay  $t_{LA1}$  of the first latch, which can be estimated with the expression below obtained by substituting the preamplifier gain ( $g_{m1}R_{L1}$ ) into the equation from [83].

$$t_{LA1} = \frac{C_{p1}}{g_{mLA1}} \cdot \ln \left( \frac{V_{OH} - V_{OL}}{2 \cdot g_{m1} R_{L1} (V_{cmp+} - V_{cmp-})} \right) \quad (14)$$

In (14),  $g_{m1}$  and  $g_{mLA1}$  are the transconductances of  $M_1$  and  $M_{LA1}$  in Fig. 40; and  $V_{OH} - V_{OL}$  is the output voltage difference at nodes  $N_x$  and  $N_y$  between high and low logic levels after latching, which is 1.4V in this design.

#### IV.2.2. Process variations

##### Mismatch analysis

Since transistor dimensions should be small for optimum speed, the input offset of the first latch stage (Fig. 40) must be assessed carefully in the design. Neglecting the charge injection errors, this input offset can be expressed for the latch under investigation by utilizing the general expression for a latched comparator from [84]:

$$V_{off} = V_{off1} + V_{off2} / (g_{m1} R_{LA1}) . \quad (15)$$

The offset  $V_{off1}$  in (15) is the offset from the input differential pair  $M_1$ , which is [85]:

$$V_{off1} = \Delta V_{T1} + (1/2) \cdot (V_{gs1} - V_{T1}) \cdot \left( \frac{\Delta R_{L1}}{R_{L1}} + \frac{\Delta \beta_1}{\beta_1} \right) , \quad (16)$$

where  $\Delta V_{T1}$  is the threshold voltage mismatch,  $V_{gs1}$  is the gate-source voltage,  $\Delta \beta_1$  is the W/L mismatch of  $M_1$ , and  $\Delta R_{L1}$  is the preamplifier load resistor mismatch.

From [86], the latch offset  $V_{off2}$  in (15) also depends on its threshold voltage ( $\Delta V_{TLA1}$ ) variation, device dimensions, and gate-source voltage overdrive ( $V_{gsLA1}$ ):

$$V_{off2} = \Delta V_{TLA1} + \frac{V_{gsLA1} - V_{TLA1}}{2} \left( \frac{\Delta W_{LA1}}{W_{LA1}} - \frac{\Delta L_{LA1}}{L_{LA1}} \right) + \frac{\Delta Q}{C_p} , \quad (17)$$

where  $W_{LA1}$  and  $L_{LA1}$  are the width and length of  $M_{LA1}$ , and  $\Delta Q$  is the charge injection error. Charge injection from control signals should be minimized by using small-sized switching devices because it can cause decision errors. A comparator reset or compensation technique might be required if the application mandates better resolution. In this analysis, charge injection error is omitted for simplicity and to maintain a focus on the expressions that show how transistor sizes and bias conditions can be optimized for enhanced resolution with timing constraints and device mismatches, which both have more severe impact on the performance of the proposed quantizer topology. Based on the analysis in [87], the following equations can be used as guidelines during the design of the first latch stage in Fig. 40 in order to minimize the variances ( $\sigma^2$ ) corresponding to the above offset voltages:

$$\sigma_{off}^2 = \sigma_{off1}^2 + \left( \frac{1}{g_{m1} R_{LA1}} \right)^2 \cdot \sigma_{off2}^2, \quad (18)$$

$$\sigma_{off1}^2 = \frac{A_{VT1}^2}{W_1 L_1} + \frac{(V_{gs1} - V_{T1})^2}{4} \cdot \left( \frac{A_{RL1}^2}{W_{RL1} L_{RL1}} + \frac{A_{\beta M1}^2}{W_1 L_1} \right), \quad (19)$$

$$\sigma_{off2}^2 = \frac{A_{VTLA1}^2}{W_{LA1} L_{LA1}} + \frac{(V_{gsLA1} - V_{TLA1})^2}{4} \cdot \frac{A_{\beta LA1}^2}{W_{LA1} L_{LA1}}; \quad (20)$$

where  $A_x$  represents the process-dependent mismatch constant for parameter  $x$  with units of: (units of  $x$ ) $\times\mu\text{m}$ . The above expressions reveal the trade-off between input offset voltage and speed because offset reduction requires large devices with minimal  $V_{gs}$ , which increases the parasitic capacitances and reduces the effective transconductances of the transistors at high frequencies.

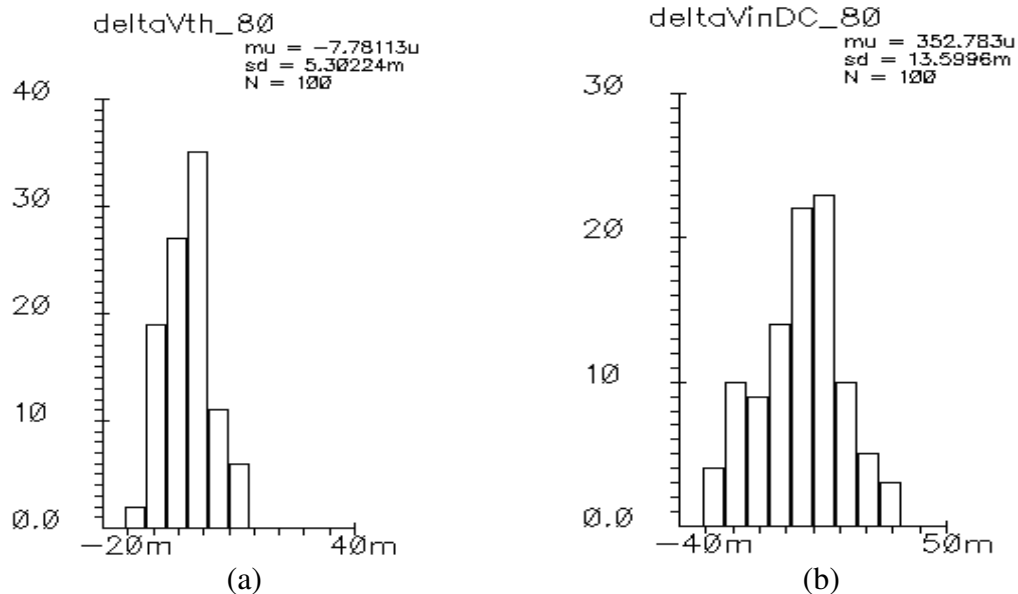


Fig. 41. Latched comparator Monte Carlo simulation without device matching. Histograms (100 runs) for critical offsets in the first comparator stage (Fig. 40):  
 (a)  $\Delta V_{T1}$  (threshold voltage difference of transistor pair M<sub>1</sub>),  
 (b) input offset voltage (at gates of transistor pair M<sub>1</sub>).

Monte Carlo simulations were performed to verify that the static offset voltages of the latched comparator and current-mode core are expected to cause errors less than 10% of the 50mV quantization step, which are noise-shaped by the  $\Sigma\Delta$  modulator. Fig. 41 displays the histograms from 100 Monte Carlo runs at 80°C assuming that none of the devices are matched in layout. The threshold voltage mismatch ( $\Delta V_{T1}$ ) of transistor pair M<sub>1</sub> and the overall input offset (at  $V_{cmp+/-}$  in Fig. 40) have standard deviations of 5.3mV and 13.6mV, respectively. In this simulation result from the complete quantizer circuit, the overall input offset at  $V_{cmp+/-}$  is affected by the mismatches of the circuitry that impact the DC voltage at  $V_{cmp+/-}$  in Fig. 38, including the comparison resistors ( $R_{cmp}$ ). To determine the impact of this offset on the quantizer resolution,  $V_{cmp} = V_{cmp+} - V_{cmp-}$  (Fig.

38, Fig. 40) has to be related to the measurable difference of  $V_{in} - V_{refx}$ , where  $V_{in} = V_{in+} - V_{in-}$  (assuming negligible sampling errors) and  $V_{refx} = V_{refx+} - V_{refx-}$  is an arbitrary differential reference voltage in Fig. 38. The input current and subtracted current at the comparison node depend on  $V_{in}$ ,  $V_{ref}$  and the transconductance  $g_{mn}$  of  $M_n$ . Since, this difference current flows into  $R_{cmp}$  to generate  $V_{cmp}$ , it can be shown that the following expression relates  $V_{cmp}$  to  $V_{in} - V_{refx}$ :

$$V_{in} - V_{refx} = \frac{V_{cmp}}{g_{mn} R_{cmp}}, \quad (21)$$

which is  $V_{in} - V_{refx} = V_{cmp} / 2.63$  in this design since  $g_{mn} = 6.5\text{mA/V}$  and  $R_{cmp} = 405\Omega$ . Using equation (21) to refer the 13.6mV input offset of the latched comparator to the quantizer input results in 5.2mV. Such an input offset contribution from the latched comparator alone would be too high for the intended application, which is why the devices in Fig. 40 with identical labels were matched in the layout. Hence, the Monte Carlo simulations were repeated with correlation coefficients of 0.95 for the matched transistors and of 0.97 for the matched polysilicon resistors. The results in Fig. 42 show that the standard deviations  $\Delta V_{T1}$  of transistor pair  $M_1$  and the input offset voltage of the latched comparator reduce to 1.2mV and 3.6mV, respectively. After referring the latched comparator's input offset to the quantizer input based on (21) as before, the estimated input offset standard deviation becomes  $3.6\text{mV}/2.63 = 1.37\text{mV}$  with device matching. *Thus, about 95% of the chips are expected to have an input-referred offset voltage below 2.7mV (within two standard deviations assuming a Gaussian distribution) due to latched comparator mismatches.*

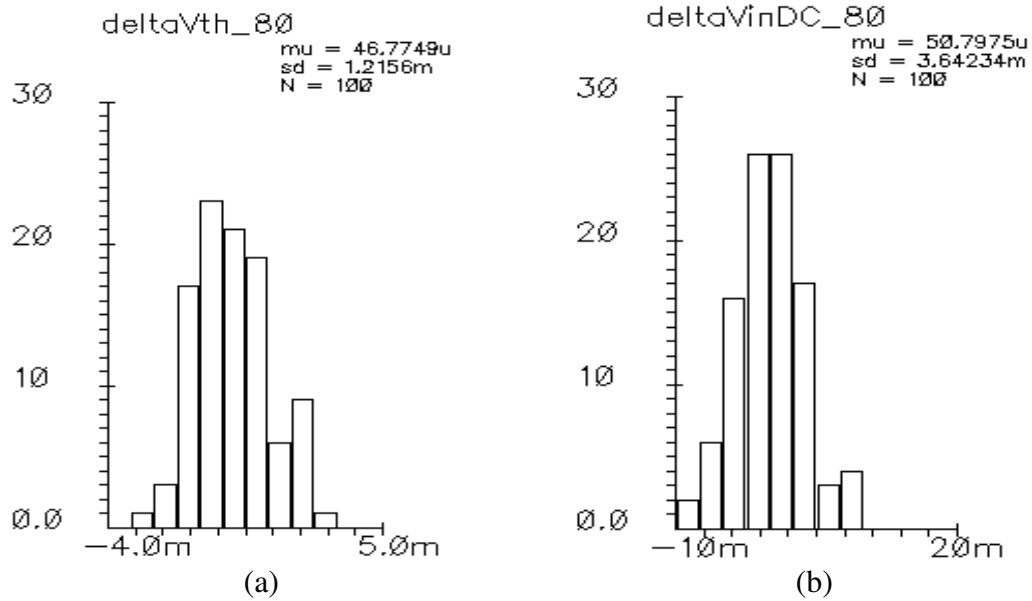


Fig. 42. Latched comparator Monte Carlo simulation with device matching. Histograms (100 runs) for critical offsets in the first comparator stage (Fig. 40):  
 (a)  $\Delta V_{T1}$  (threshold voltage difference of transistor pair M<sub>1</sub>),  
 (b) input offset voltage (at gates of transistor pair M<sub>1</sub>).

An input offset variation evaluation was also conducted for the differential pairs M<sub>n</sub> in the current-mode comparator core (Fig. 38). All transistors with identical names in the core are also matched with a common-centroid layout, and Fig. 43 shows the histogram of the threshold voltage difference obtained from 100 Monte Carlo runs using the same correlations as defined in the latched comparator simulations. The estimated standard deviation is 0.97mV for each differential pair M<sub>n</sub>, which is the approximate input offset under the assumption that the errors from the matched current mirrors (M<sub>p</sub> in Fig. 38) are not significant. Since the output currents of two differential pairs are compared in this circuit, the effective input offset voltage is found by combining the variances:

$$V_{off(core)} = \sqrt{\sigma_{Mn(input)}^2 + \sigma_{Mn(reference)}^2} = \sqrt{2 \cdot V_{off\_Mn}^2} \quad (22)$$



where  $\sigma_{M_n(input)}$  is the standard deviation of the input offset voltage of the differential pair  $M_n$  in the input signal path and  $\sigma_{M_n(reference)}$  is the standard deviation of the input offset voltage of the equal-sized reference differential pair by which the comparison current is generated. From Fig. 43 and equation (22), the estimated combined input-referred offset voltage during a comparison in the current-mode core is 1.4mV. Hence, about 95% of the chips are expected to have an input-referred current-mode comparator core offset below 2.8mV, which is additional static error because this offset is directly at the input.

In summary, the latched comparator and current-mode comparator core input offsets are expected to create a combined static input-referred inaccuracy of less than 5.5mV with likelihood of 95%. However, this error can be compensated by tuning the reference voltages in Fig. 38 as demonstrated by a simulation in the next subsection.

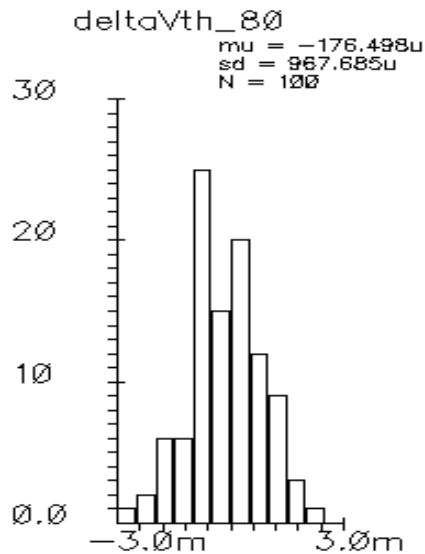


Fig. 43. Quantizer core Monte Carlo simulation with device matching. Histogram (100 runs) for the threshold voltage difference of pairs  $M_n$  in Fig. 38.

### IV.2.3. Simulation results and technology scaling

#### Post-layout simulations

Fig. 44 shows the layout of the quantizer, which was designed in 0.18 $\mu\text{m}$  CMOS technology and embedded in the  $\Sigma\Delta$  modulator. The quantizer's 0.39mm<sup>2</sup> die area includes the bias and timing generation circuitry that generates the control signal from the seven clock phases provided by the on-chip complementary injection-locked frequency divider.

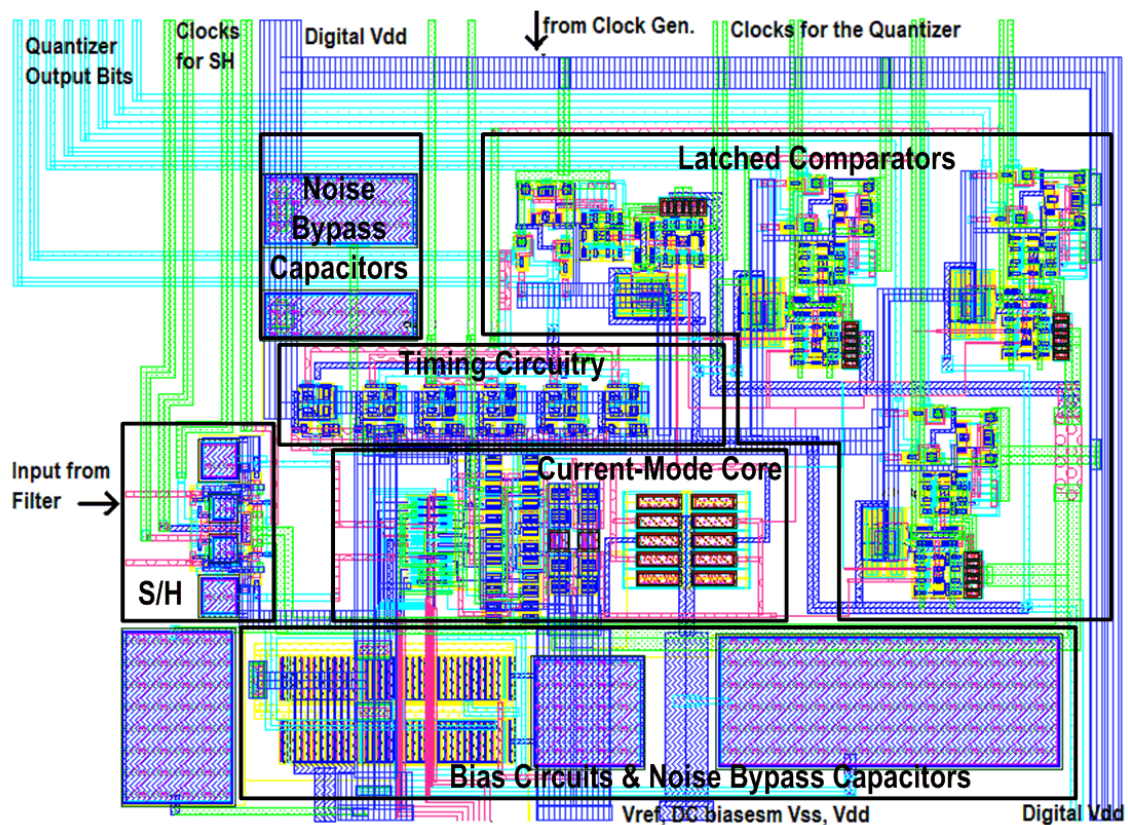


Fig. 44. Quantizer layout (0.18 $\mu\text{m}$  CMOS technology).  
Area of quantizer & timing circuitry: 750 $\mu\text{m}$   $\times$  520 $\mu\text{m}$ .

The simulation testbench included models for pad parasitics, bonding wire inductances, and 100fF capacitance as rough estimate for the effect of the package. The output bit transitions during a ramp test with an input between -200mV and 200mV is shown in Fig. 45. From it, the transition levels with typical device models were verified to be within approximately  $\pm 5\text{mV}$  of the ideal values, which is 10% of the 50mV quantization step size. One level deviates by 7.6mV from the ideal 150mV. From system-level simulations of the continuous-time  $\Sigma\Delta$  ADC in Matlab, it was determined that up to  $\pm 10\text{mV}$  reference level shifts are permissible to achieve a signal-to-quantization noise ratio better than 72dB.

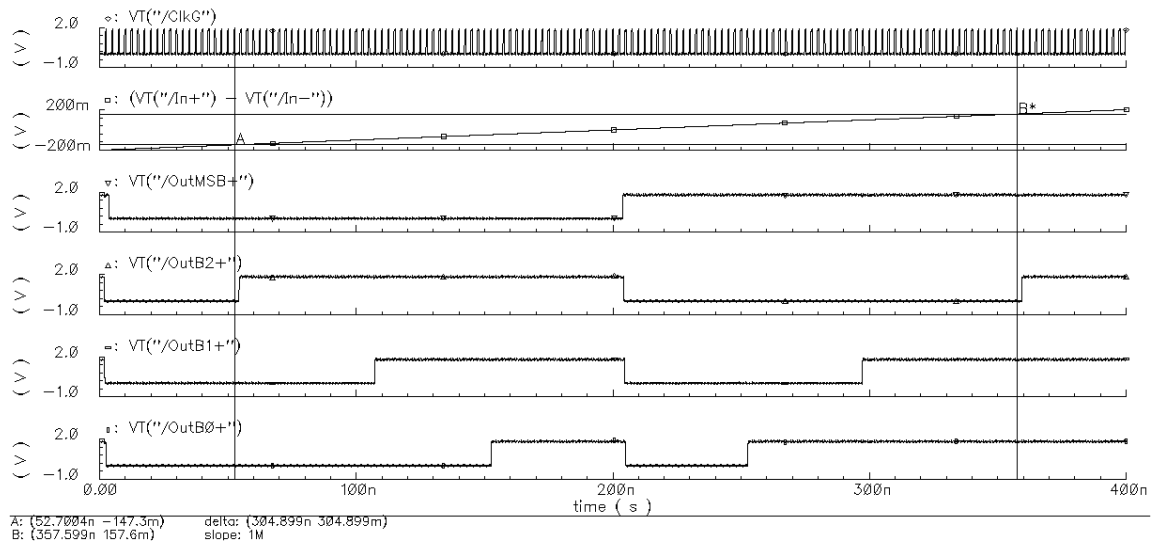


Fig. 45. Output bit transitions with an input ramp from -200mV to 200mV.  
 Top-to-bottom: clock signal, input ramp, bits from Fig. 37: MSB, B2...B0.  
 Quantization transition levels: -147.3mV, -94.8mV, -49.9mV, 2.7mV,  
 50.2mV, 95.1mV, 157.6mV.

Fig. 46 displays the differential nonlinearity (DNL) and the integral nonlinearity (INL) corresponding to the transition levels from the post-layout simulation. The adjustable reference voltages in Fig. 38 offer a way to alleviate the effects of PVT variations. As an example, Fig. 47 visualizes this feature for the -150mV transition level of the 0.18 $\mu$ m design, which can be shifted +/-30mV by adjusting  $V_{ref3}$ .

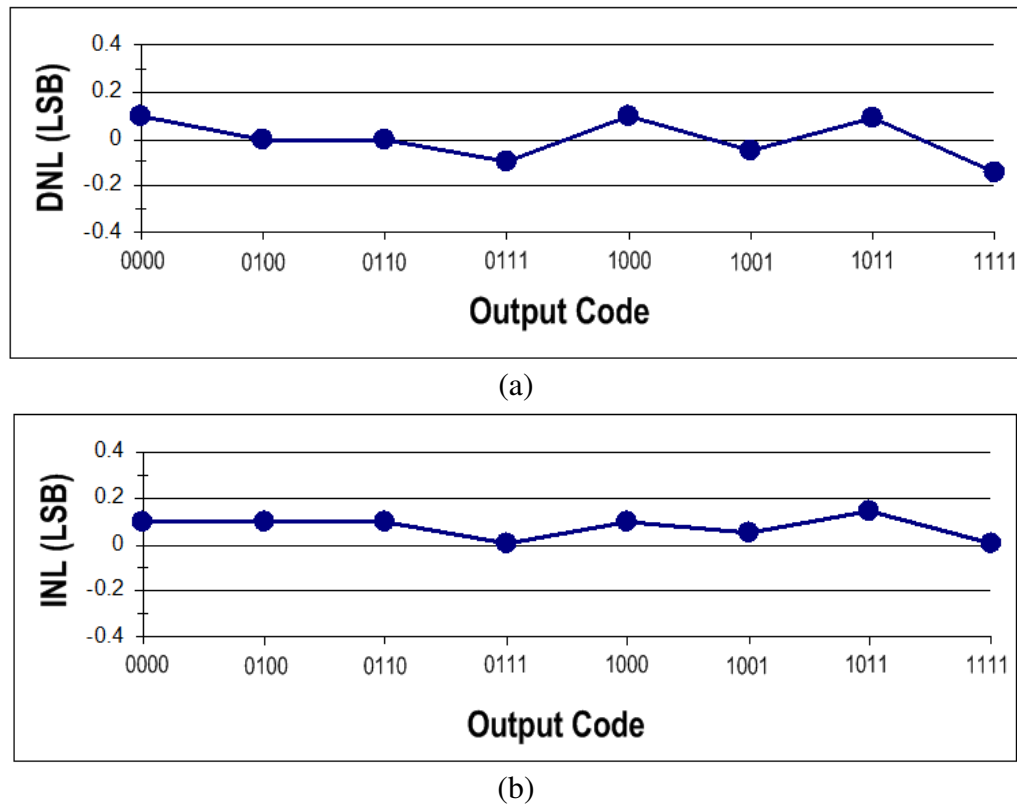


Fig. 46. Quantizer post-layout simulations: (a) DNL (b) INL.

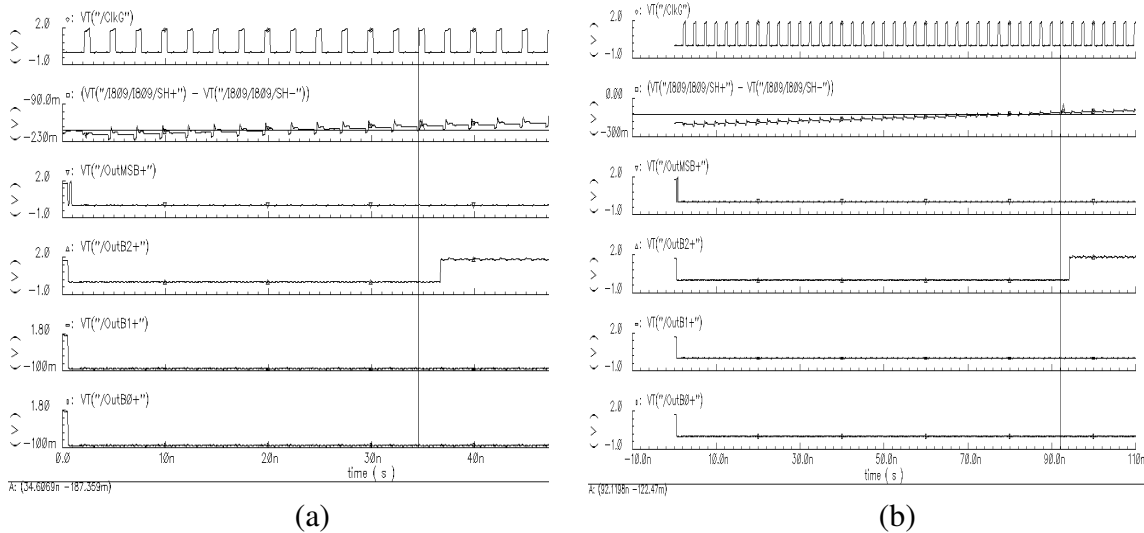


Fig. 47. Tuning range of the -150mV transition level (schematic simulations).  
 (a) Bit transition at -187.4mV with  $V_{\text{ref}3+} - V_{\text{ref}3-} = 180\text{mV}$ ,  
 (b) bit transition at -122.5mV with  $V_{\text{ref}3+} - V_{\text{ref}3-} = 120\text{mV}$ .  
 Top-to-bottom: clock signal, input (after S/H), bits from Fig. 37: MSB, B2...B0.

### Technology scaling

The operation of the proposed current-mode quantizer architecture relies heavily on switching transistors and digital auxiliary circuitry. Hence, performance improvements can be expected in technologies with devices that have a high unity gain frequency (high- $f_T$ ). To verify this hypothesis, the quantizer and control circuitry were re-designed with UMC 90nm CMOS technology and 1V supply voltage, and then simulated with the identical setup as the 0.18 $\mu\text{m}$  design. The dimensions of the components in the quantizer core (Fig. 38) are given in Table VII for both designs, which shows that the active area with UMC 90nm technology was reduced by more than four times. But, over half of the quantizer layout area (Fig. 44) in 0.18 $\mu\text{m}$  technology consists of routing and capacitors to filter out noise at critical components. Since the requirements for routing and passives

do not change significantly, an area reduction of up to 25% in the 90nm process is a more reasonable estimate.

Table VIII provides a comparison of the most important quantizer properties, showing that the resolution in 90nm is only reduced by 2mV. This reduction is partially due to the limited voltage headroom with a 1V supply, which causes inaccuracies with large input signals because the devices operate at the edge of their intended regions of operations. Nevertheless, design optimizations through the use of non-minimum device dimensions could be explored to improve the resolution of the 90nm design at the expense of an increase in power consumption.

Table VII. Component parameters in the quantizer core (Fig. 38)

Device	Jazz 0.18 $\mu$ m CMOS Design (W/L Dimensions or Parameter)	UMC 90nm CMOS Design (W/L Dimensions or Parameter)
$M_n$	28 $\mu$ m / 0.2 $\mu$ m	10 $\mu$ m / 0.36 $\mu$ m
$M_p$	56 $\mu$ m / 0.18 $\mu$ m	20 $\mu$ m / 80nm
$M_{sw}$	21 $\mu$ m / 0.18 $\mu$ m	16 $\mu$ m / 80nm
$M_{B1}, M_{B2}$ for $I_B$ (current mirror)	800 $\mu$ m / 1 $\mu$ m	110 $\mu$ m / 1 $\mu$ m
$R_{cmp}$	405 $\Omega$	633 $\Omega$
$R_{BW}$	333 $\Omega$	1.4k $\Omega$
$I_B$	1.9mA	0.25mA
$V_{refMSB}$	1.1V	0.5V
$V_{ref1+} / V_{ref1-}$	1.125V / 1.075V	0.525V / 0.475V
$V_{ref2+} / V_{ref2-}$	1.150V / 1.050V	0.550V / 0.450V
$V_{ref3+} / V_{ref3-}$	1.175V / 1.025V	0.575V / 0.425V
Supply Voltage	1.8V	1V

Table VIII. Key quantizer performance parameters

	<b>Jazz 0.18<math>\mu</math>m CMOS</b>	<b>UMC 90nm CMOS</b>
<b>Resolution</b>	+/-5mV	+/-7mV
<b>Static Power: Quantizer Core</b>	6.8mW	0.5mW
<b>Static Power: Latched Comparators</b>	4 $\times$ 4.3mW	4 $\times$ 0.3mW
<b>Layout Area</b>	750 $\mu$ m $\times$ 520 $\mu$ m (actual area for core, logic, routing)	estimate: ~500 $\mu$ m $\times$ 500 $\mu$ m (~1/4 of active area, similar passives /routing )
<b>Clock Frequency</b>	400MHz	400MHz

A significant reduction in power was possible for the 90nm design, in which the quantizer core consumes only 0.5mW. On the contrary, this core consumed 6.8mW in the initial 0.18 $\mu$ m design. The power savings were enabled by the facts that the quantizer operation mainly depends on switching speeds and on the amount of parasitic capacitance from all devices connected to nodes  $V_{cmp+}$  and  $V_{cmp-}$  in Fig. 38. At these nodes, the parasitic capacitances form RC time constants with resistors ( $R_{cmp}$ ) that limit the speed of the comparison. On the whole, less current is required to perform the comparisons with a 400MHz clock rate due to the smaller dimensions and higher ratio of transconductance to parasitic capacitance (i. e. higher  $f_T$ ) in 90nm technology.

#### IV.2.4. ADC chip measurements with embedded quantizer

As mentioned earlier, the two-step current-mode quantizer has been designed for a  $\Sigma\Delta$  modulator chip that was fabricated by our research group. Due to the complexity of the system, the test chip and printed circuit board were not equipped with sufficient inputs and outputs to characterize the individual blocks. A brief overview of system-

level measurements is presented in this subsection to demonstrate the 3-bit quantizer's functionality and that the block-level requirements have been met to achieve the targeted system performance. Fig. 48 displays the die microphotograph of the multi-phase continuous-time 5<sup>th</sup>-order lowpass  $\Sigma\Delta$  modulator fabricated in Jazz Semiconductor 0.18 $\mu\text{m}$  1P6M CMOS technology, which was assembled in a QFN-80 package. It occupies a total area of 2.6mm<sup>2</sup>, including the VCO and CILFD but excluding pads and electrostatic discharge (ESD) protection circuitry. The four output bit streams of the 3-bit quantizer were captured with a 4-channel oscilloscope synchronized at 400Msamples/s prior to post-processing in Matlab.

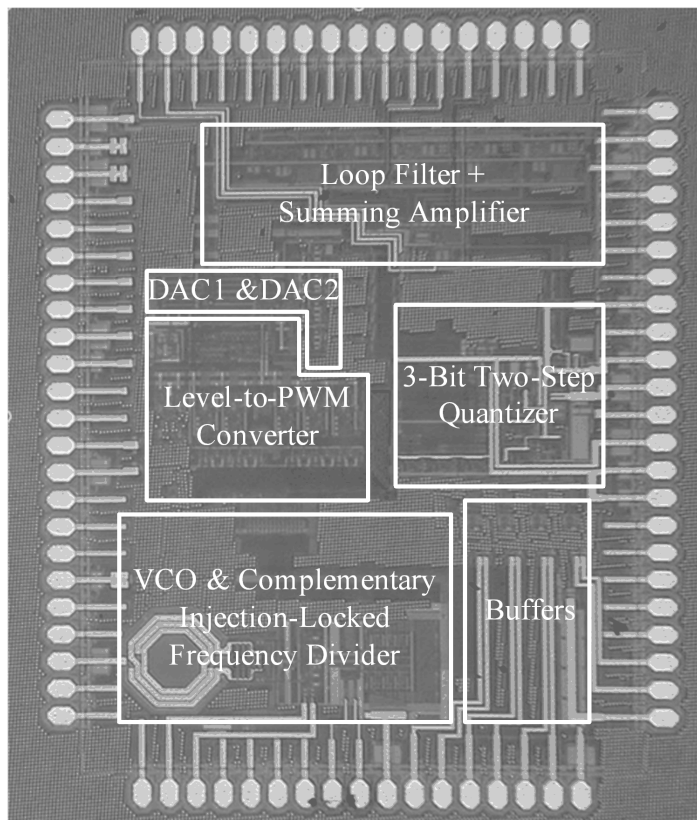


Fig. 48. Die microphotograph (2.6mm<sup>2</sup> area excluding pads and ESD circuitry).



Fig. 49 shows the output spectrum of the modulator with an input of -2.2dBFS at 5MHz. Based on the noise bandwidth of 6.1KHz during the measurement, the average noise floor is around -145dBFS/Hz and the peak SNR is 68.5dB in 25MHz bandwidth. The third-order harmonic distortion (HD3) in this case is 78dB below the test tone, which demonstrates the high linearity properties of both the loop filter and the PWM DAC/quantizer feedback scheme. The peak SNDR including the harmonic tones in the 25MHz bandwidth is 67.7dB. The measured SNR and SNDR for different input signal powers are plotted in Fig. 50, in which the 69dB dynamic range (DR) is annotated.

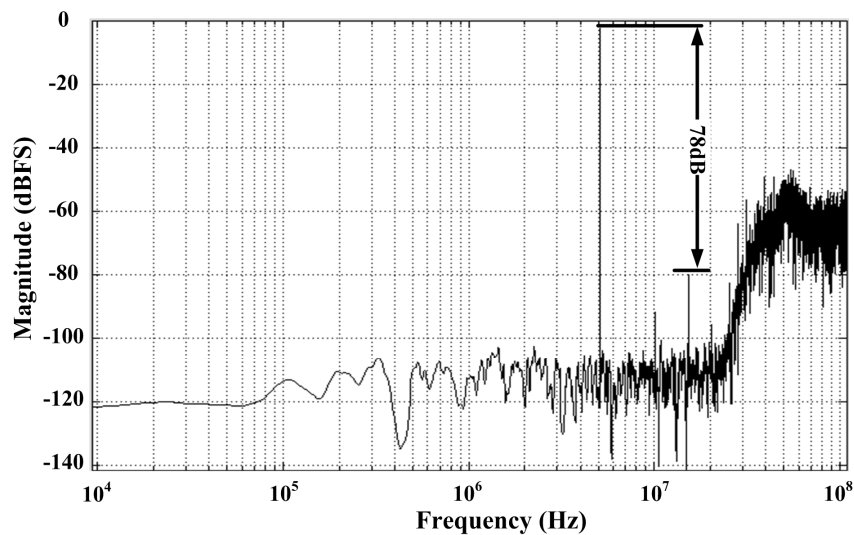


Fig. 49. Measured output spectrum of the  $\Sigma\Delta$  modulator.  
A -2.2dBFS input tone was applied at 5.08MHz.

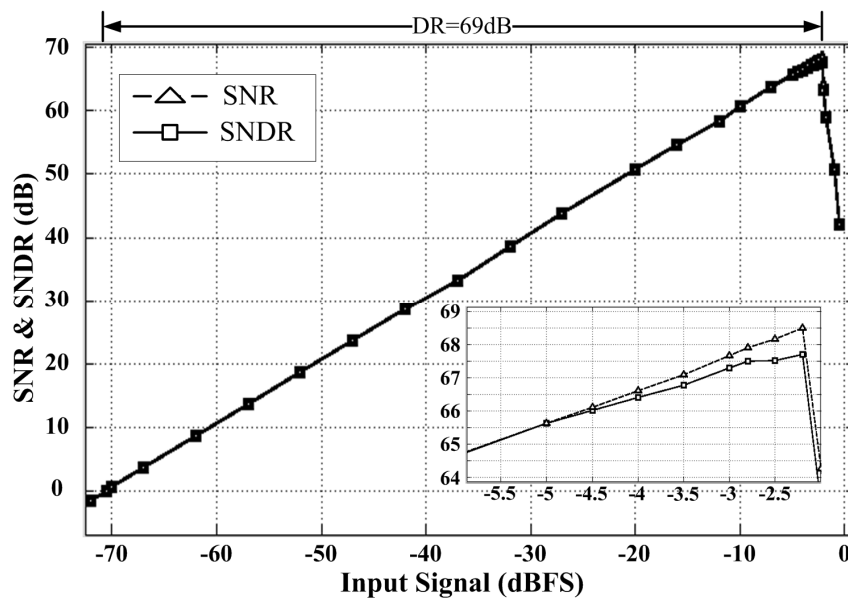


Fig. 50. Measured SNR and SNDR vs. input signal power.

Table IX. Measured  $\Sigma\Delta$  ADC performance

Technology	Jazz 0.18 $\mu$ m CMOS
Power Supply	1.8V
Clock Frequency	400MHz
Bandwidth	25MHz
Peak SNR / SNDR* @ 25MHz Bandwidth	68.5dB / 67.7dB
SFDR	78dB
IM3 (-5dBFS per tone)	< -72dB
Dynamic Range	69dB
Power Consumption	48mW
Area without pads &ESD protection	2.6mm <sup>2</sup>

\* Includes total in-band distortion power and noise.

Table IX provides a summary of the modulator specifications. The linearity performance (IM3) was characterized by injecting two tones with 2MHz separation,

each having a power of -5dBFS. Excluding the VCO, the power budget is 44mW for the modulator core, 2.5mW for the locked ring oscillator, and 1.5mW due to clock buffers. Table X shows a comparison between the proposed modulator architecture and recently reported modulators based on the following figure-of-merit ( $FoM$ ):

$$FoM = \frac{Power}{2^{ENOB} \cdot (2 \cdot BW)}, \quad (23)$$

where  $ENOB$  is the effective number of bits and  $BW$  is the bandwidth. Although fabricated in an economical technology, the achieved 444fJ/bit  $FoM$  of the proposed modulator core is competitive with the current state of the art. In addition, a  $FoM$  improvement is anticipated if the solution is exported to deep submicron technologies, which would lower the quantizer power (see Table VIII) and level-to-PWM converter power as a result of more efficient switching circuitry.

Table X. Comparison with previously reported lowpass  $\Sigma\Delta$  ADCs

Reference	Technology	$f_s$	BW	Filter Order	Peak SNDR	Power	FoM (fJ/bit)
[66] ISSCC 2008	180nm CMOS	640MHz	10MHz	5	82dB	100mW <sup>†</sup>	487
[67] JSSC 2008	130nm CMOS	950MHz	10MHz	2	72dB	40mW*	500
[68] ISSCC 2009	65nm CMOS	250MHz	20MHz	3	60dB	10.5mW <sup>†</sup>	319
[88] ISSCC 2007	90nm CMOS	340MHz	20MHz	4	69dB	56mW <sup>#</sup>	608
[89] ISSCC 2008	90nm CMOS	420MHz	20MHz	4	70dB	28mW <sup>†</sup>	271 <sup>Δ</sup>
[90] JSSC 2006	130nm CMOS	640MHz	20MHz	3	74dB	20mW <sup>†</sup>	122
[91] ISSCC 2009	130nm CMOS	900MHz	20MHz	3 (+1 digital)	78.1dB	87mW*	330
<b>This Work</b>	180nm CMOS	400MHz	25MHz	5	67.7dB	48mW* (44mW <sup>†</sup> )	484* (444 <sup>†</sup> )

\* Includes clock generation circuitry. <sup>†</sup> For modulator circuitry only. <sup>#</sup> Includes digital calibration of RC spread & noise cancellation filter. <sup>Δ</sup> Discrete-time modulator (would require anti-aliasing filter for comparable blocker rejection).

### IV.3. Summarizing Remarks

A two-step current-mode quantizer was described in this section. The architecture was constructed for application within a  $\Sigma\Delta$  modulator loop, and it incorporates characteristics that are aligned with present-day quantizer design trends. First, successive approximations controlled by multiple clock phases are used to reduce the number of required comparators in comparison to the classical flash quantizer architecture. Since switching operations become more efficient as technology scaling progresses, the discussed successive comparison scheme in the quantizer core helps to take advantage of the speed benefits in modern CMOS technologies. Second, the quantizer has easily adjustable reference voltage levels, allowing it to be part of a system-level calibration technique as discussed in Section II.2.5. In such a scenario, the on-chip voltage references at the high-impedance input gates in the quantizer core (Fig. 38) can be generated with a low-power on-chip DAC.

With regards to the  $\Sigma\Delta$  modulator application for which the quantizer was designed, the utilization of time-based processing methods within the continuous-time  $\Sigma\Delta$  modulator shifts more operations into the digital realm, improving the system's robustness, scalability, and potential for power savings. A 5<sup>th</sup>-order continuous-time lowpass  $\Sigma\Delta$  modulator using 3-bit time-domain quantization and feedback has been demonstrated in a 0.18 $\mu\text{m}$  CMOS process. Nonlinearities from element mismatch of traditional multi-level DACs are circumvented because the 3-bit PWM feedback is realized with an inherently linear single-element DAC. Since low-jitter clocks are essential in time-based continuous-time  $\Sigma\Delta$  modulators, the required jitter performance

is accomplished by means of an injected-locked clock generation technique which provides 400MHz clock signals with seven phases. The measured peak SNDR of the modulator with 25MHz bandwidth is 67.7dB, while the SFDR and DR are 78dB and 69dB, respectively. Its power consumption is 48mW from a 1.8V supply. Approximately 56% of this power is dissipated in the quantizer and the level-to-PMW converter, which mainly contain circuits based on high-frequency switching. Technology scaling is expected to significantly enhance the efficiency of the proposed modulator architecture via power reduction in the digital circuitry, especially in the quantizer.

## V. AN ON-CHIP TEMPERATURE SENSOR TO MEASURE RF POWER DISSIPATION AND THERMAL GRADIENTS \*

### V.1. Background

Monitoring performances of individual blocks that constitute a single-chip RF receiver chain is beneficial for identification of faulty devices and self-calibration. In conventional built-in test (BIT) strategies, electrical detectors are placed along the signal path for power measurements [20]-[23] or extraction of input impedance matching conditions in the RF front-end [19], [24], [92]. Although small, the input impedance of the electrical detectors degrades performance; and the impact of parasitic capacitances from detectors worsens with increasing operating frequencies.

Thermal coupling through the semiconductor substrate generates a rise in temperature in the vicinity of a circuit/device that depends on the device's power dissipation. This thermal coupling can be modeled in the DC domain [93] or with complex small-signal parameters [94]. Moreover, it can be utilized for IC testing

---

\* © 2011 IEEE. Section V is in part reprinted, with permission, from "Electro-thermal design procedure to observe RF circuit power and linearity characteristics with a homodyne differential temperature sensor," M. Onabajo, J. Altet, E. Aldrete-Vidrio, D. Mateo, and J. Silva-Martinez, accepted for publication in *IEEE Trans. Circuits and Systems I: Regular Papers*.

This material is included here with permission of the IEEE. Such permission of the IEEE does not in any way imply IEEE endorsement of any of Texas A&M University's products or services. Internal or personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution must be obtained from the IEEE by writing to [pubs-permissions@ieee.org](mailto:pubs-permissions@ieee.org). By choosing to view this material, you agree to all provisions of the copyright laws protecting it.

purposes [95]. Using on-chip temperature gradients as test observables to measure power dissipation is advantageous because the sensors do not load the circuit under test (CUT) as electrical detectors do. Instead, the small temperature-sensing devices are placed near the CUT, making the technique non-invasive. Furthermore, temperature gradients become more critical to both analog and digital system performance as the integration levels of modern single-chip systems increase, creating incentives to improve diagnosis and compensation techniques. For example, the sensitivity of a direct-conversion receiver in [96] was degraded by 2-4dB from transient heating effects.

Thermal gradients on a silicon die can be detected with embedded differential temperature sensors [95]. Temperature measurements are usually conducted up to 10KHz because thermal coupling has low-pass characteristics [94]. But, the multiplication of voltages and currents of different frequencies creates electrical power components at DC and various frequencies [97]. In heterodyne measurement strategies [98], two RF tones at frequencies  $f_1$  and  $f_2$  are applied to the CUT in order to measure the low-frequency power dissipation at  $\Delta f = f_2 - f_1$  ( $<10\text{KHz}$ ) with a temperature sensor. While this approach enables indirect power measurement without interference from on-chip DC temperature gradients, it also necessitates the use of a spectrum analyzer or lock-in amplifier. It is highly desirable to perform measurements at DC to reduce the complexity of the measurement setup and to provide a step towards BIT integration. The RF signal power detected in the thermal DC regime is a result from mixing voltage and current signals at the same frequency, which is why this strategy is referred to as the homodyne method. Since the generated DC temperature gradients are also strongly

influenced by the power dissipation in bias circuitry, sensing the RF power requires an on-chip sensor with a wide dynamic range.

This research effort concentrated on the development of a differential temperature sensor feasible for a homodyne BIT strategy. To ensure CMOS compatibility, the sensing devices are formed with parasitic vertical bipolar (PNP) transistors. Section V.2 provides an overview of the proposed BIT methodology and the application to low-noise amplifier (LNA) characterization is presented as an example. The proposed differential temperature sensor design and tuning features are discussed in Section V.3, for which the measurement results are presented in Section V.4 together with experimental verification of the LNA BIT. Finally, Section V.5 provides conclusions from the work.

## **V.2. Temperature Sensing Approach**

### *V.2.1. Integration with transceiver calibration techniques*

A temperature sensing strategy is appealing for BIT applications where the goal is to: i) identify gross failures that affect the power dissipation in bias circuitry; ii) measure the signal power along processing paths; iii) design self-calibration schemes that can adapt to temporary thermal hot spots occurring near a sensitive circuit. The envisioned purpose of a homodyne sensing scheme is illustrated in Fig. 51, in which several small temperature-sensing devices ( $S_i$ , where  $i$  ranges from 1 to 6 in Fig. 51) are located at various test points within analog blocks of an RF receiver and at one reference location ( $S_{\text{ref}}$ ). In a system-on-a-chip, the temperature gradients between the sensing devices  $S_i$  and  $S_j$  ( $i \neq j$ ) or  $S_i$  and  $S_{\text{ref}}$  can be acquired through processing the sensor core output signals. This larger sensor core contains the necessary bias and amplification circuits to



provide a DC output to an on-chip analog-to-digital converter (ADC). If the on-chip ADC is not available for reuse, then a dedicated 8-12 bit low-power ( $< 50\mu\text{W}$ ) ADC with  $0.05\text{-}0.7\text{mm}^2$  die area would be sufficient for online digitization of the DC sensor output at a low sampling rate (e.g.  $100\text{KHz}$  as in [99], [100]). In such a case, the total area overhead of the sensor core, 20 sensing devices, the ADC, and  $0.75\text{mm}^2$  room for reference voltage and bias current generation circuitry would be between 2% and 15% for a  $10\text{-}25\text{mm}^2$  receiver chip. Finally, the comparisons of the differential measurements conclude in the digital signal processor (DSP), allowing DC temperature gradients and the signal power (i.e. gain) along the analog receiver chain to be monitored. As a step towards realizing such a system-level BIT, the focus in this work is on the measurement of the RF power dissipation and 1-dB compression point of an LNA. In brief, the goal is to design a practical sensor circuit that can be employed as on-chip detector near analog blocks for system-level calibration methods as those described in Section II.2. Another potential use of the sensors is to monitor the average power dissipating in digital blocks for the detection of faults.

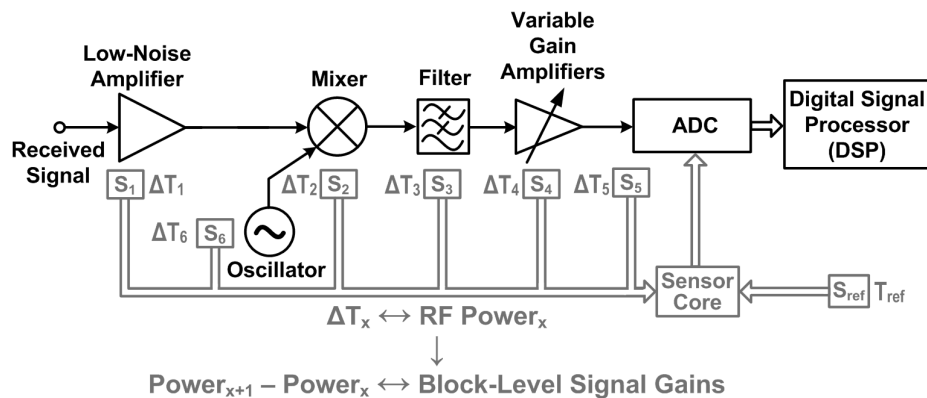


Fig. 51. Generalized receiver diagram with on-chip thermal sensing.

The proposed approach could be used for low-cost pass/fail screening in a high-volume manufacturing test environment or for online monitoring of parameter drift during normal operation. As temperature linearly depends on dissipated power, specification variations and faults that cause a change in power dissipation are detectable; e.g. variations of either  $S_{11}$  in the front-end or gains (output differences between two detectors). Compared to conventional electrical power detectors, a major advantage is that the temperature-sensing devices do not load the signal paths because they are not electrically connected to the input or output of the CUTs, leaving only the coupling path through the common substrate. The discussed approach can also be extended to an individual die in a stacked-die assembly, but each die should include its own reference sensing device ( $S_{ref}$  in Fig. 51) and the differential power gain comparisons should only be made for test points on the same die because each die has its own common-mode temperature.

#### *V.2.2. Modeling of the thermal coupling*

Various modeling ([93]-[95], [101]-[102]) and simulation strategies ([103]-[104]) exist to account for the static and dynamic effects of thermal coupling on the performance of electrical devices on the same die. In this BIT application, the primary interest lies in estimating the temperature increase from power dissipation in the CUT at the location of the sensing device. Hence, the silicon substrate has been modeled with an RC network in order to allow coupled analysis with the electrical behavior of the CUT and temperature sensor using the Spectre simulator in Cadence.



selected because it is comparable with sensing device dimensions, which implies that the devices are approximated as having the same size as the unit grid. With this model, the electrical voltage at each node in the network is equivalent to a temperature change in degrees (Kelvin or Celsius) relative to the ambient die temperature during the electro-thermal co-simulation, and any injected electrical current is equivalent to power dissipation of a device located at the node. Capacitors in the network can be omitted if only DC temperature analysis is needed, but they are included in this work to predict settling times and to maintain a generic model that accounts for frequency-dependence. Points M, C, and R in Fig. 52 represent the locations of devices (Table XI) from which dissipated power (in Watts) is injected into the network modeled as current (in Amperes). As shown in Fig. 53, these current sources are connected to the equivalent points M, C, and R in Fig. 52 based on the layout locations of the devices. The local temperature change is measured with a parasitic vertical PNP device at point S having spacing in the layout of  $7\mu\text{m}$  and  $10\mu\text{m}$  from points C and M, respectively. The temperature ( $T_s$ ) change of the temperature transducer in the sensor is obtained by coupling the voltage at node S to the PNP device through an ideal voltage-controlled voltage source with gain of  $k = -1.8\text{mV/K}$  to modulate the base-emitter voltage ( $V_{be}$ ) of the PNP transistor according to its temperature sensitivity [105]. Here, the temperature-dependence is assumed to be linear over the range of interest.



In the RC network model (Fig. 52), the less critical layers 2 through 5 have z-direction lengths of 20 $\mu\text{m}$ , 40 $\mu\text{m}$ , 80 $\mu\text{m}$ , and 160 $\mu\text{m}$  to model 310 $\mu\text{m}$  of the 330 $\mu\text{m}$  thick substrate. To reduce the complexity and simulation time, the fine-resolution grid (shown in 3D) was only extended by 10 $\mu\text{m}$  around points M, S, C, and R; while low-resolution unit volumes with the following dimensions were employed at the sides and corners to expand the grid by 450 $\mu\text{m}$  into the horizontal directions (only shown in the top view): 10 $\mu\text{m} \times 150\mu\text{m} \times z_u$ , 150 $\mu\text{m} \times 10\mu\text{m} \times z_u$ , and 150 $\mu\text{m} \times 150\mu\text{m} \times z_u$ . Finally, the lateral edges are terminated with infinite impedances and the bottom of the 5<sup>th</sup> layer is grounded, i.e. the thermal boundary conditions are assumed adiabatic and isothermal, respectively. Each discretized capacitance and the directional node resistances in Fig. 52 are calculated as follows [96]:

$$C = \rho \cdot c \cdot x_u y_u z_u , \quad (24)$$

$$R_x = x_u / (\kappa \cdot y_u z_u) , \quad (25)$$

$$R_y = y_u / (\kappa \cdot x_u z_u) , \quad (26)$$

$$R_z = z_u / (\kappa \cdot x_u y_u) ; \quad (27)$$

where the mass density ( $\rho$ ), specific heat capacity ( $c$ ), and thermal conductivity for silicon ( $\kappa$ ) are 2.3 $\cdot 10^6 \text{g/m}^3$ , 0.7J/(g $\cdot$ K), and 120W/(m $\cdot$ K) at 75 $^\circ\text{C}$  , respectively [96].

### V.2.3. Electro-thermal analysis example: low-noise amplifier

Fig. 53 depicts the main devices of the CUT, the PNP sensing device, and how the RC network couples both circuits. The CUT is a typical broadband LNA with resistive load for which design details can be found in Table XI and [106]. Next, it will be shown

that circuit-level power and linearity characteristics of blocks can be extracted using temperature sensors even with a single test tone. However, in system-level testing strategies, multi-tone tests or a frequency sweep of a single test tone typically enhance the fault coverage. Assuming a sinusoidal signal with voltage amplitude  $A$  at  $v_{in}$  in Fig. 53 and combining the DC analysis with the small-signal analysis, simplified expressions for the average power dissipation of the devices can be derived in terms of the transconductances ( $g_{mM}$ ,  $g_{mC}$ ) and DC drain-source voltages ( $V_{dsM}$ ,  $V_{dsC}$ ) of the transistors ( $M_M$ ,  $M_C$ ), load resistor  $R_L$ , and DC current  $I_{DC}$ :

$$R_L: \quad P_r = R_L \cdot I_{DC}^2 + \frac{1}{2}(g_{mM} \cdot A)^2 \cdot R_L, \quad (28)$$

$$M_M: \quad P_m = V_{dsM} \cdot I_{DC} - \frac{1}{2}(g_{mM} \cdot A)^2 / g_{mC}, \quad (29)$$

$$M_C: \quad P_c = V_{dsC} \cdot I_{DC} - \frac{1}{2}(g_{mM} \cdot A)^2 \cdot (R_L - 1 / g_{mC}), \quad (30)$$

Here, the energy conservation principle holds since the AC amplitude-dependent terms sum up to zero and  $P_r + P_m + P_c = V_{DD} \cdot I_{DC}$ . The above expressions show that the average power from the RF signal adds to the DC power at the load resistor but subtracts from the DC power at the active devices acting as RF power sources. This property implies that the ideal placement of the temperature-sensing PNP device in the layout is either on the side of the load resistor that does not face the MOS transistors, or between the two transistors where their temperature effects add. The latter location was selected as shown in Fig. 54. Resistor  $R_L$  was placed more than  $50\mu\text{m}$  away from the sensor to reduce thermal interference, which can be assessed by injecting the power of  $R_L$  at a point R on the RC network in Fig. 53 during the simulations.

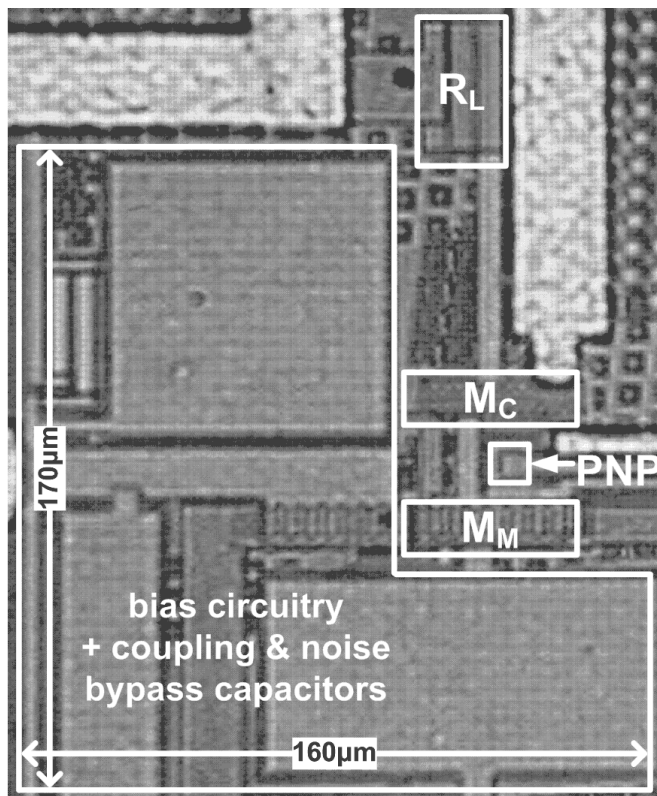


Fig. 54. Area of the die with CUT (LNA) and temperature-sensing PNP device.

The broadband LNA used as the CUT in Fig. 53 was designed with 11dB gain for on-die probing [106]. Table XI lists the key design and performance parameters from simulations of this LNA with estimated parasitics for the packaged prototype chip. The graphs in Fig. 55 were obtained by sweeping the RF power of a single-tone input to the CUT and plotting the average power for each device. As expected from (28)-(30), the DC component of the dissipated power due to RF signal processing adds to the DC bias power at the resistor and subtracts from the DC bias power at the MOS transistors. The analysis in Appendix D explains how the nonlinearities of the MOS transistors cause their DC power curves ( $P_m$ ,  $P_c$ ) to have minima. Notice that the DC component of the



power due to RF circuit activity is significantly less than the DC bias power dissipation of each device, which translates into a high dynamic range requirement when the same sensor should be capable to measure the effects of DC bias as well as of RF signal processing via temperature changes. In addition, the sensor must at least have sufficient sensitivity to detect a change in the dissipated DC power from  $20\mu\text{W}$  to  $200\mu\text{W}$  associated with the  $-10\text{dBm}$  to  $0\text{dBm}$  electrical signal input power levels.

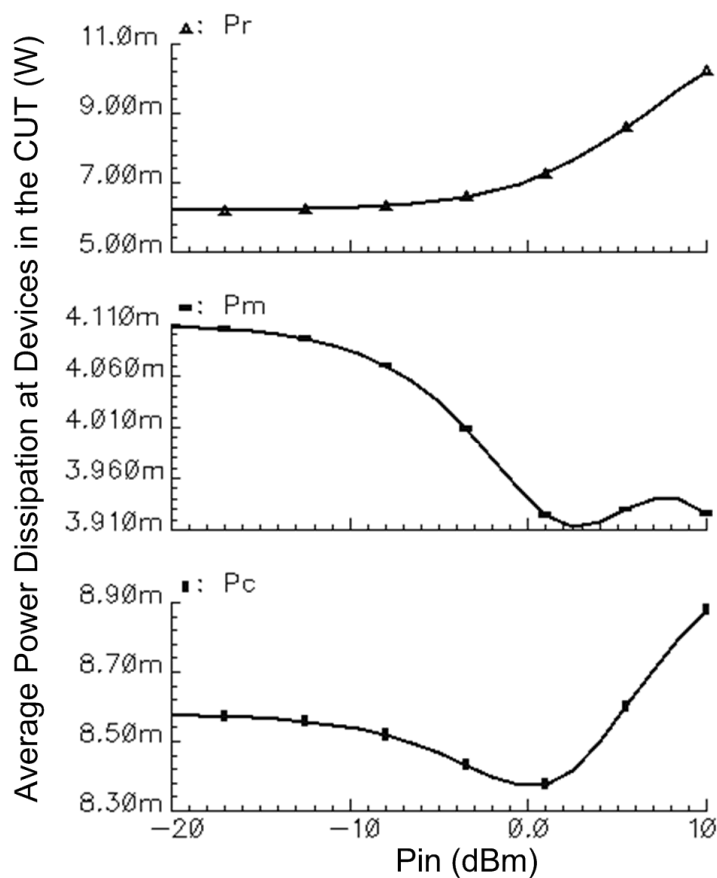


Fig. 55. Simulated average powers at devices in the CUT vs. RF input power. Top:  $P_r$  at  $R_L$ , middle:  $P_m$  at  $M_M$ , bottom:  $P_c$  at  $M_C$ .

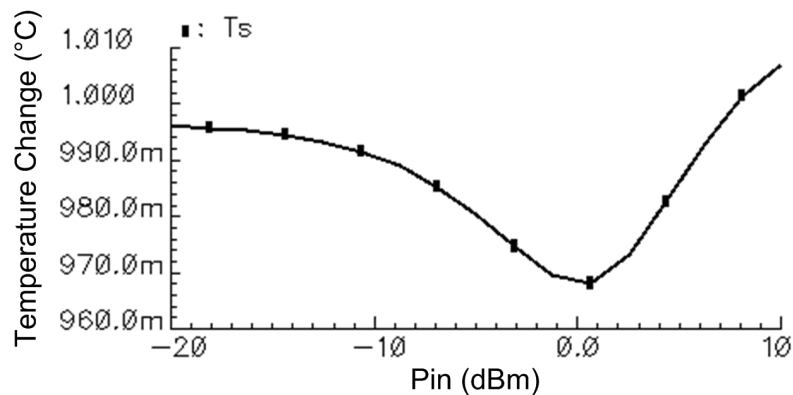


Fig. 56. Temperature change  $T_s$  at the sensing device vs. RF input power.

Fig. 56 visualizes the simulated local temperature change  $T_s$  at node S shown in Fig. 52 and Fig. 53. The DC bias of the CUT creates static 0.996°C change of  $T_s$  with respect to the ambient temperature. As the amplitude of the electrical signal applied to the CUT input increases, the local temperature changes as a result of the superimposed thermal coupling from the power dissipations (Fig. 55) in devices  $M_M$ ,  $M_C$ , and  $R_L$ . The DC power/temperature reaches a minimum that can be related to the 1-dB compression point with a shift on the x-axis (Appendix D). The simulation result in Fig. 56 also indicates that the sensor sensitivity should be high enough to detect 5m°C to 30m°C changes in the -15dBm to 0dBm range of interest. The CUT and electro-thermal network were simulated with -5dBm input power to assess the transient response of the temperature change. Fig. 57 reveals that the settling time is approximately 8 $\mu$ s, which is adequately short for production testing.

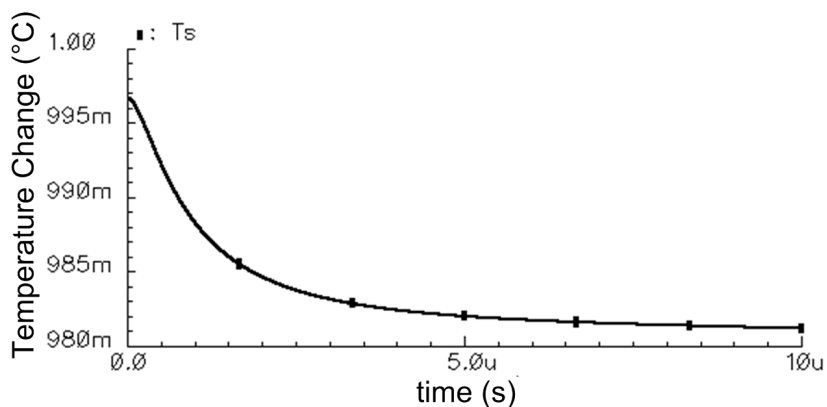


Fig. 57. Transient behavior of  $T_s$  with -5dBm input power.

### V.3. CMOS Differential Temperature Sensor Design

#### V.3.1. Previous sensors

Various passive and active sensors for on-chip differential temperature measurements are experimentally compared in [107], and a schematic representation of a previously presented CMOS-compatible fully-differential sensor is shown in Fig. 58. Conceptually, the two temperature-sensing parasitic PNP devices ( $Q_1$ ,  $Q_2$ ) are placed as a differential pair within an operational transconductance amplifier (OTA) configuration. The collector current difference between  $Q_1$  and  $Q_2$  due to temperature difference  $\Delta T = T_1 - T_2$  is amplified by current mirrors within the OTA before flowing into the high impedance nodes at the output. Currents  $I_{cal1}/I_{cal2}$  can be adjusted to compensate for electrical and thermal offsets. This sensor has a high sensitivity of up to  $\sim 400\text{mV/mW}$  when the CUT that dissipates power is placed at  $20\mu\text{m}$  distance from  $Q_1$  (or  $Q_2$ ) and there is a spacing of  $400\mu\text{m}$  between  $Q_1$  and  $Q_2$ . A drawback of this topology is its limited dynamic range of less than  $1.5\text{mW}$  with this sensitivity. Generally, such

differential sensors with high sensitivity are optimal for the heterodyne approach ([97], [98]) and the AC setup at low frequencies in [107] with external lock-in amplifier or spectrum analyzer. Since the heterodyne measurements of two RF tones at the  $\Delta f$  frequency are free of interference from DC temperature gradients, the previous sensors are well-suited to sense and amplify the low-power mixing product at  $\Delta f$  without saturating the sensor.

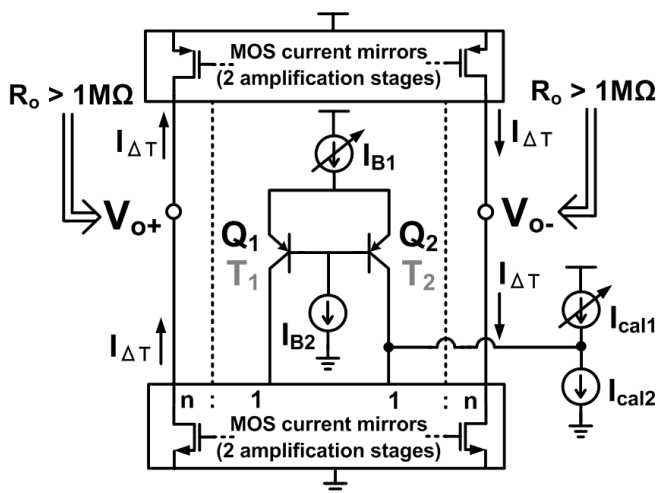


Fig. 58. A differential CMOS temperature sensor with lateral PNP devices. (This circuit was proposed in [107].)

### V.3.2. Design of the proposed sensor topology

In this dissertation, the focus is on the homodyne measurement approach and the development of a sensor core optimized for application to RF BIT measurements at DC without relying on any external equipment. Hence, the sensor must have a wide dynamic range to enable concurrent DC and RF power measurements. Additionally, differential temperature sensors are often comprised of lateral parasitic PNP devices, but some

CMOS processes only model vertical PNP devices which are more restrictive because the collector (p-type substrate) is typically grounded. Parasitic vertical PNP devices are popular temperature sensors because they offer high precision and repeatability; e.g.  $\pm 0.1^\circ\text{C}$  absolute error from  $-50^\circ\text{C}$  to  $130^\circ\text{C}$  in [108], where the error can be treated as DC offset and  $V_{be}$  temperature sensitivity spread due to process variations is limited to below 2% depending on the technology.

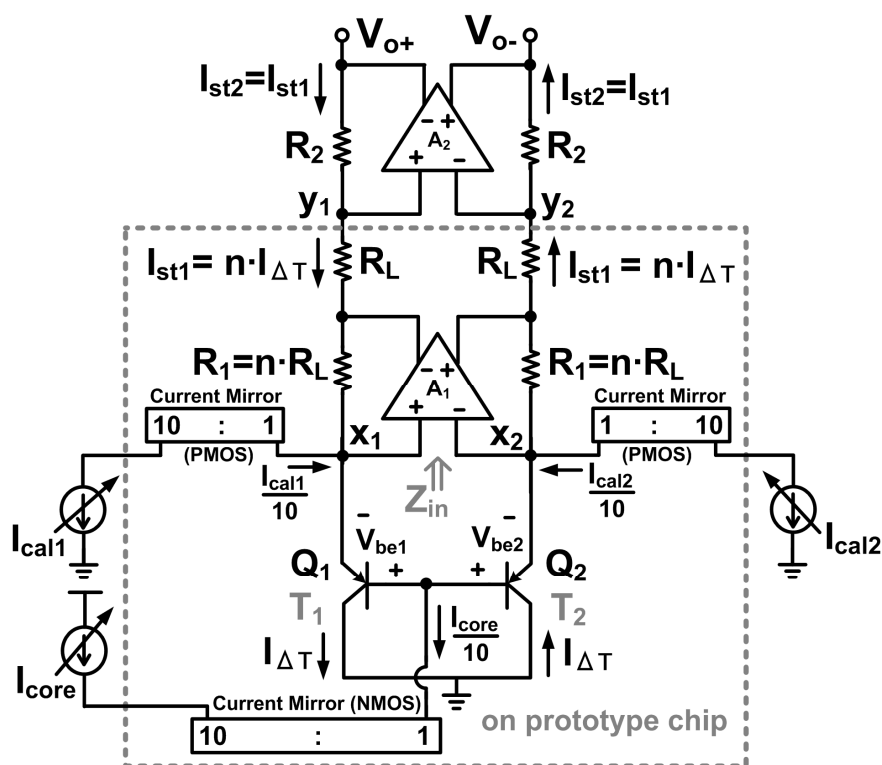


Fig. 59. Proposed wide dynamic range differential temperature sensor.  
(The devices  $Q_1$  and  $Q_2$  are vertical parasitic PNP transistors in a CMOS process.)

Fig. 59 displays the proposed sensor topology that was constructed with vertical PNP devices. Sensing transistors  $Q_1$  and  $Q_2$  are biased with the same operating point,

having common base and collector terminals. These two devices can be either the sensing or reference points ( $S_i$  or  $S_{ref}$ ) in Fig. 51. The DC emitter voltages are also forced to be identical due to the virtual ground created by the feedback from the first amplifier ( $A_1$ ). Notice that the collector current difference of  $Q_1$  and  $Q_2$  under this DC bias ideally only depends on the temperature difference ( $\Delta T = T_1 - T_2$ ) between their respective locations. In practice, device mismatches and thermal gradients cause offsets that can be compensated with currents  $I_{cal1}$  and  $I_{cal2}$ . The temperature-dependent differential current ( $I_{\Delta T}$ ) is amplified with a cascade of a transimpedance amplifier (TIA) stage ( $A_1, R_1$ ) and resistive load  $R_L = R_1/n$  connected to a virtual ground from a subsequent TIA stage ( $A_2, R_2$ ). Consequently, the current amplification ( $I_{st1} = n \cdot I_{\Delta T}$ ) depends on reliable resistive matching to minimize sensitivity variations. Moreover, the sensitivity can be changed with the base current  $I_{core}$  [95] to allow reuse of the same sensor near low- and high-power devices on the chip and to compensate for any process-dependent gain variations. As a proof-of-concept, the sensor core and first amplification stage with  $R_L$  were implemented on the prototype chip, while stage 2 was realized with an off-chip amplifier for simplified external DC voltage measurements. In a BIT application, the output current of stage 1 could be digitized directly or the second amplification stage could be included on the chip.

The dynamic range improvement with the proposed sensor topology comes from the virtual ground at nodes  $x_{1,2}$  in Fig. 59, which furnishes a low impedance at the emitters of  $Q_{1,2}$ . It also avoids that  $I_{\Delta T}$  is converted to a voltage difference at the emitters, which

would cause imbalance of the bias conditions of the sensing PNP devices. Instead, the current is processed by a low-gain TIA stage with a controlled amplification ratio.

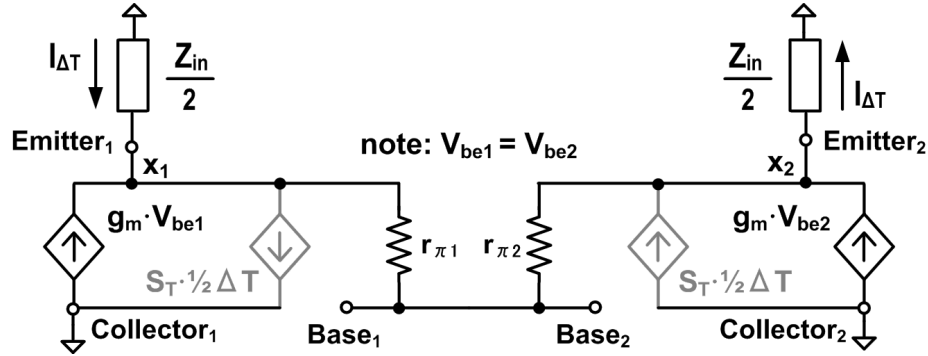


Fig. 60. Simplified small-signal equivalent circuit of the sensor core. (The PNP devices are represented with the hybrid- $\pi$  model.)

A simplified small-signal model of the PNP pair in the sensor core is shown in Fig. 60. The temperature difference causes a change of the emitter current with a sensitivity that can be roughly estimated as  $S_T \approx k \cdot g_{mQ}$ ; where  $k$  and  $g_{mQ}$  are the temperature sensitivity of the base-emitter voltage and the transconductance of  $Q_{1,2}$ , assuming  $Z_{in} \ll 1/(k \cdot g_{mQ})$ . Part of this temperature-dependent current will not be amplified by the TIA because it will flow through resistance  $r_{\pi}$ , resulting in unavoidable sensitivity loss. It is important to minimize the effective load impedances at the emitters presented by the input impedance ( $Z_{in}$ ) of the first TIA stage. A high amplifier gain improves the overall sensitivity by lowering  $Z_{in}$  in Fig. 59 according to the following approximation:

$$Z_{in} \approx R_1 / [1 + A_1 (\frac{R_L \parallel R_1}{R_L \parallel R_1 + r_o})] = \frac{R_1}{1 + A_v}, \quad (31)$$

where  $r_o$  and  $A_v$  are the output resistance and loaded voltage gain of amplifier  $A_1$ . To determine the appropriate gain prior to the design of amplifier  $A_1$ , the sensor core was simulated with an ideal amplifier model having a variable gain. The sensitivity ( $\Delta I_{st1}/\Delta T$ ) vs.  $A_v$  is plotted in Fig. 61 and a target value of  $A_v \approx 32 \approx 30\text{dB}$  was selected from these simulations to avoid major efficiency degradation in the sensor core. Additionally, matched polysilicon resistors of  $R_I = 8\text{k}\Omega$  and  $R_L = 1\text{k}\Omega$  were selected for a robust current amplification ratio of  $n = 8$ . To ease testing of this prototype design,  $R_2$  was an off-chip  $100\text{k}\Omega$  resistor and  $A_2$  was an off-the-shelf operational amplifier (NJM4580D) with  $110\text{dB}$  DC gain.

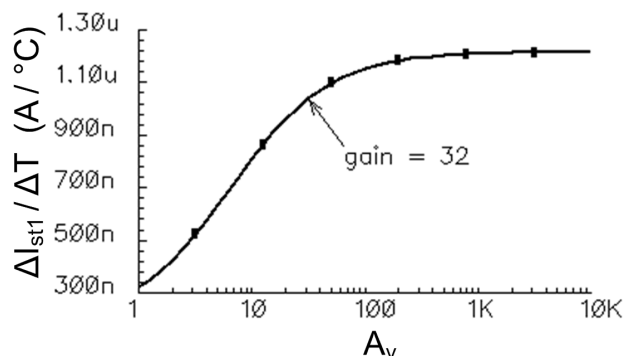


Fig. 61. Simulated sensor sensitivity ( $\Delta I_{st1}/\Delta T$ ) vs. gain ( $A_v$ ) for amplifier  $A_1$ . ( $R_I = 8\text{k}\Omega$ ,  $R_L = 1\text{k}\Omega$ , and  $I_{core} = 100\mu\text{A}$ .)

Fig. 62 shows the schematic of amplifier  $A_1$ . It consists of a simple differential pair ( $M_2$ ) loaded by transistors ( $M_3$ ) in saturation region and a PMOS source follower output stage ( $M_4, M_5$ ). The amplifier's input DC level depends on the bias conditions of  $Q_{1,2}$  in the sensor core (Fig. 59), which is why nodes  $n_{1,2}$  are regulated by the common-mode





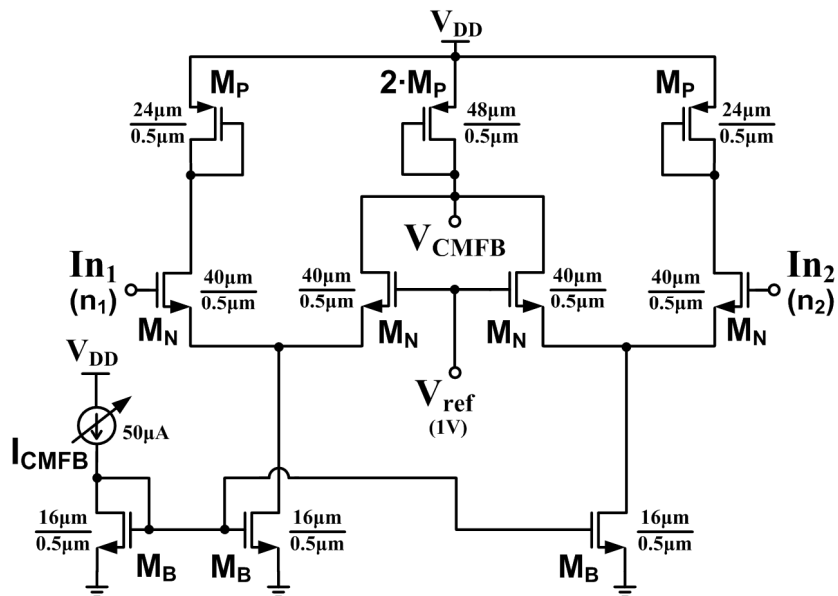


Fig. 63. Common-mode feedback (CMFB) circuit schematic.

Table XII. Simulated amplifier ( $A_I$ ) specifications

Parameter	Value
DC Gain	30.2dB
$f_{3dB}$	1.74MHz
Unity Gain Frequency ( $f_u$ )	56.9MHz
Phase Margin	89.7°
Integrated Input-Referred Noise (DC - $f_u$ )	55.1μV
Common-Mode Rejection Ratio*	75.5dB at 10KHz
Power Supply Rejection Ratio*	36.4dB at 10KHz
Output Resistance	270Ω
5% Settling Time (1mV step input, unloaded)	264ns
CMFB Loop: DC Gain / Phase Margin	35.1dB / 74.4°
Input Offset Voltage (standard deviation)	1.5mV
Technology / $V_{DD}$	0.18μm CMOS / 1.8V
Power Dissipation (with CMFB)	1.05mW

\* For a single output. The fully-differential processing in the sensor topology improves the noise rejection.

### V.3.3. Adjustment of the sensor's sensitivity

DC simulations of the standalone sensor circuit can be performed by sweeping the SPICE parameter  $Trise$  of one PNP device to emulate its temperature increase above the ambient temperature due to local heating from the CUT. For example, the plots in Fig. 64 were generated this way in order to evaluate the dynamic range based on the output current  $I_{st1}$  of the first amplification stage in Fig. 59. The results show that the linear range is  $\pm 4.7^\circ\text{C}$  with  $2.94\mu\text{A}/^\circ\text{C}$  sensitivity and  $\pm 13.4^\circ\text{C}$  with  $0.99\mu\text{A}/^\circ\text{C}$  sensitivity for  $I_{core} = 1\text{mA}$  and  $I_{core} = 100\mu\text{A}$ , respectively. The sensor core's wide dynamic range with adjustable sensitivity is sufficient to monitor devices with power beyond 50mW. Large differential output currents cause a large voltage drop across  $R_I$  in Fig. 59, which forces  $M_{4,5}$  in the amplifier (Fig. 62) out of the saturation region.

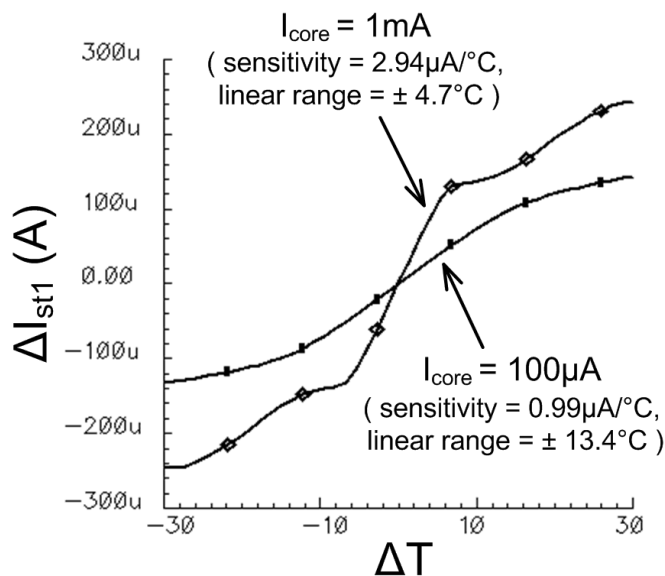


Fig. 64. Simulated dynamic range of the sensor core.

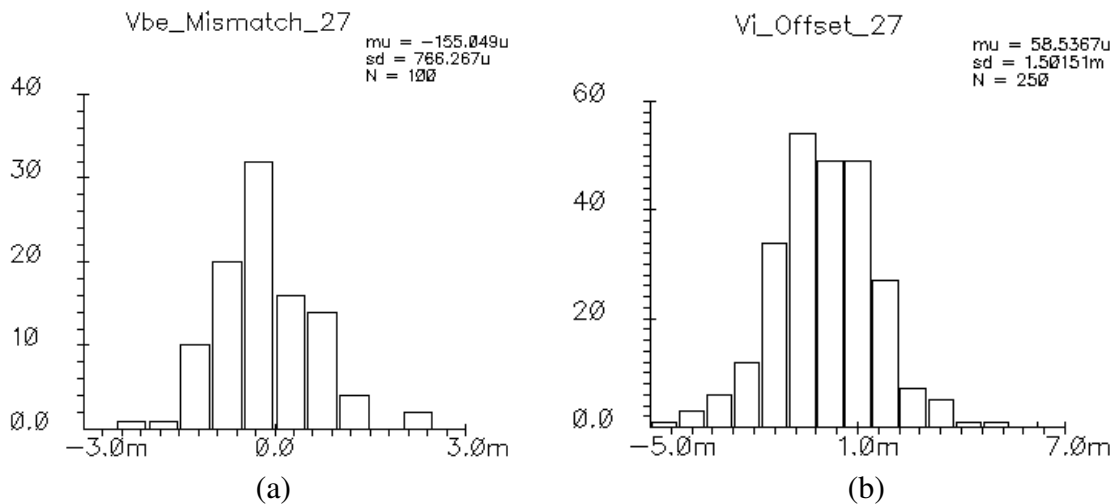


Fig. 65. Assessment of offsets in the sensor core with Monte Carlo simulations. (a)  $V_{be}$  mismatch of  $Q_1/Q_2$ , (b) input offset voltage of amplifier  $A_I$ .

Currents  $I_{cal1}$  and  $I_{cal2}$  (Fig. 59) permit the compensation of DC temperature gradients as well as electrical offsets from mismatches in the cascaded amplifier stages. The appropriate calibration current ranges can be determined with DC simulations that include anticipated electrical device mismatches while modeling the heat sources of the CUT or any other nearby circuits in the simulation based on Fig. 53. For example,  $I_{cal1} = 100\mu A$  compensates for an equivalent thermal offset at the sensing device location of approximately  $8^\circ C$  ( $0.99\mu A/^\circ C$  sensitivity setting). Offset voltages are also calibrated out. Based on the Monte Carlo simulation results in Fig. 65, the  $V_{be}$  mismatch of the PNP pair and the input offset of amplifier  $A_I$  have standard deviations of  $0.8mV$  and  $1.5mV$ , respectively; and the simulated  $V_{be}$  mismatch due to absolute temperature changes from  $-50^\circ C$  to  $130^\circ C$  is less than  $0.2mV$  (Fig. 66). In the calibration step preceding a measurement, the sensor can be balanced by adjusting  $I_{cal1}$  and  $I_{cal2}$  under monitoring of the differential output until it is close to  $0V$ . This was done manually in

the experimental characterization (Section V.4.1), but could be performed with the same on-chip ADC that resolves the sensor output in a system-level BIT scenario (Fig. 51).

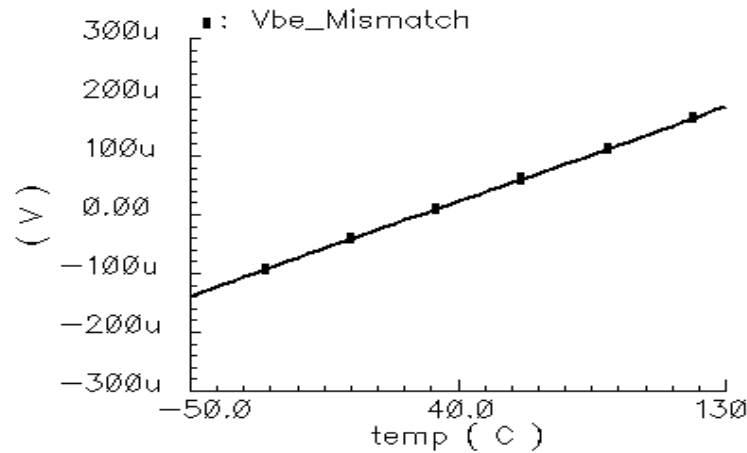


Fig. 66. Simulated  $V_{be}$  mismatch of  $Q_1/Q_2$  vs. ambient temperature.

#### V.3.4. Sensor design optimization procedure

To perform co-simulations of the CUT and appropriate sensor circuit it is advisable to follow these steps:

- 1) Construct the electro-thermal coupling network described in Section V.2.2 based on the actual or anticipated layout locations of the devices in the CUT. The capacitors can be removed if only DC analysis is to be performed.
- 2) Select a suitable layout location to place a single parasitic PNP transistor near the device(s) to be monitored, and perform the simulation in Section V.2.3 which will reveal the temperature change at the related node in the grid. Select a suitable location for the reference parasitic transistor that will be used to process the thermal gradient. In this example,  $Q_2$  is located at a distance of

420 $\mu\text{m}$  where the simulated temperature change is about two orders of magnitude lower than at the sensing device  $Q_1$ .

- 3) Determine the required dynamic range and temperature sensitivity for the sensor from the results in 2). In the previously discussed example, average power dissipations of 4.1mW and 8.6mW at  $M_M$  and  $M_C$  caused almost a 1 $^\circ\text{C}$  imbalance between the PNP transistors. A wide dynamic range is desirable to monitor low- and high-power devices on a chip. On the other hand, a sensitivity around 5m $^\circ\text{C}$  is needed to detect the RF signal power at the LNA. Hence, the sensor circuitry must have sufficient gain to achieve this resolution. Notice that, if this technique is utilized to characterize other blocks, then the higher power levels of the signals processed in the receiver chain makes it easier to sense the temperature changes.
- 4) Design a differential temperature sensor circuit consisting of the PNP transistor pair in step 2) as well as bias and amplification circuitry to meet the specifications in step 3). Nodes in the extended RC network allow assessing that the temperature change at the reference PNP device ( $Q_2$ ) is significantly smaller than at  $Q_1$  near the CUT. In the presented case, the DC temperature changes are 0.996 $^\circ\text{C}$  at  $Q_1$  and 96m $^\circ\text{C}$ , 20m $^\circ\text{C}$ , 13m $^\circ\text{C}$  at 150 $\mu\text{m}$ , 300 $\mu\text{m}$ , 450 $\mu\text{m}$  away from  $Q_1$  respectively. In an integrated system, effects from circuits further than 150 $\mu\text{m}$  away from the CUT are attenuated by more than one order of magnitude, but their impacts can be accounted for by injecting their power dissipations as currents into the extended RC grid.

- 5) Simulate the CUT, electro-thermal network, and complete sensor circuit by coupling the schematics as in Fig. 53. Optimize the sensing device placement as well as the sensor circuit's gain, dynamic range, and transient response based on the simulated electro-thermal coupling.

As example, the plot in Fig. 67 was obtained with a CUT/sensor co-simulation to assess the 0.5dBm 1-dB compression point identification capability, showing that the sensor output reaches a -79.0mV minimum with 0.63dBm input power. Based on the analysis in Appendix D, the simulated relative input power shift (0.13dB) should be subtracted from the minimum power point to predict the 1-dB compression point.

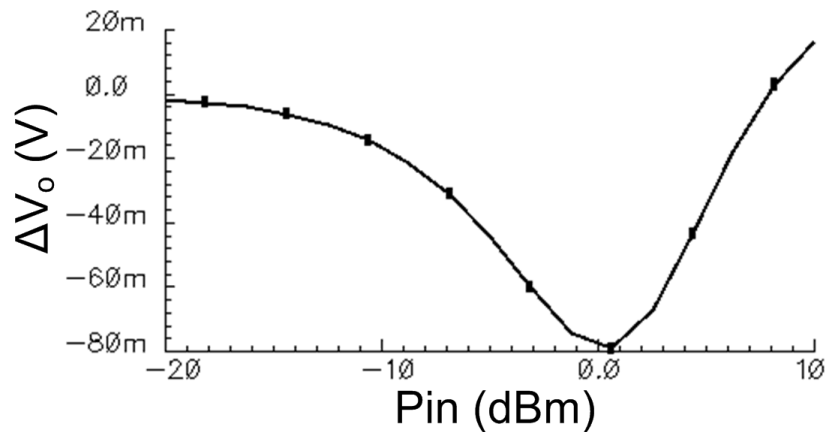


Fig. 67. Combined CUT and sensor simulation.

The plot shows the differential sensor output voltage after settling vs. average RF input power applied to the CUT having a 1-dB compression point of 0.5dBm.

#### V.4. Measurement Results

Fig. 68 displays the microphotograph of the chip fabricated in Jazz Semiconductor 0.18 $\mu\text{m}$  1P6M CMOS technology. Sensing device  $Q_2$  ( $11\mu\text{m} \times 11\mu\text{m}$ ) is located at a reference point that is separated from active devices of the sensor core by  $150\mu\text{m}$ . Additional diode-connected MOS transistors ( $D_{1,2}$  with  $W/L = 60\mu\text{m}/0.18\mu\text{m}$ ) and a  $50\Omega$  polysilicon resistor  $R_t$  ( $5\mu\text{m} \times 33.8\mu\text{m}$ ) are placed  $4\mu\text{m}$  away from the sensing devices as extra test heat sources. Standard multimeters were used for the measurement of voltage drops and currents to determine the DC power at these heat sources.

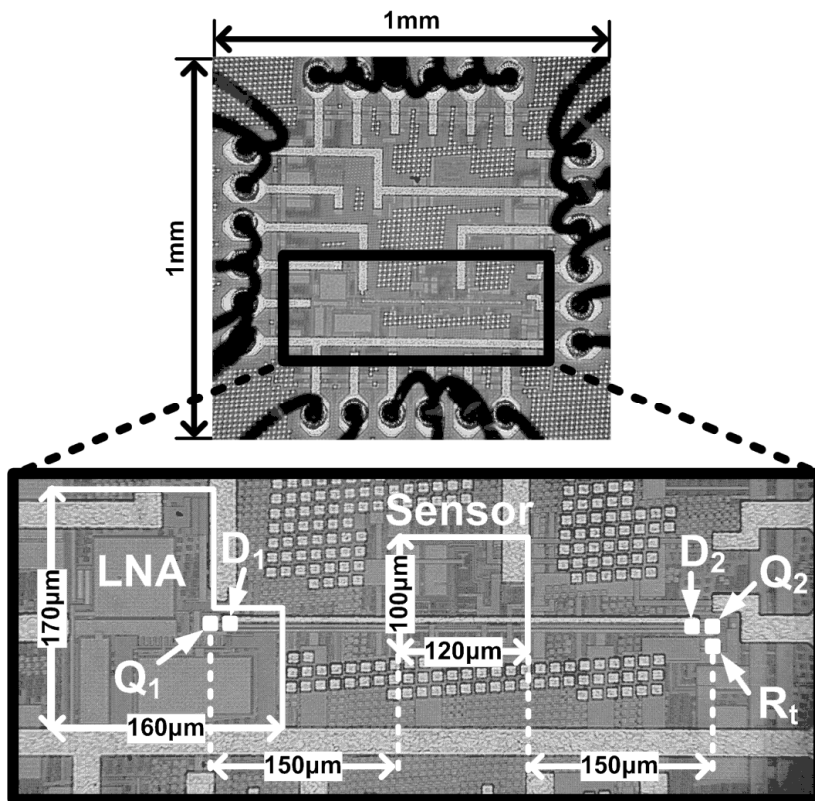


Fig. 68. Micrograph of the chip with differential temperature sensor and LNA. Emitter area of  $Q_{1,2}$ :  $11\mu\text{m} \times 11\mu\text{m}$ . Area of sensor core:  $0.012\text{mm}^2$  (reusable with additional  $Q_x$  devices to monitor multiple locations on a die).



#### V.4.1. Temperature sensor characterization

Fig. 69 shows the measured differential output voltage in response to the DC power dissipation at resistor  $R_t$ , which was kept below 16mW to prevent damage based on the process-specific recommendations for the device and interconnect dimensions. The plots show that the linear range with 199.6mV/mW sensitivity is slightly above 12mW, but it extends beyond 16mW in the 41.7mV/mW sensitivity setting. Although 16mW dynamic range is adequate to monitor conventional high-power devices, the simulations (Fig. 64) indicate that the range is more than 30mW with  $I_{\text{core}} = 100\mu\text{A}$  (41.7mV/mW sensitivity).

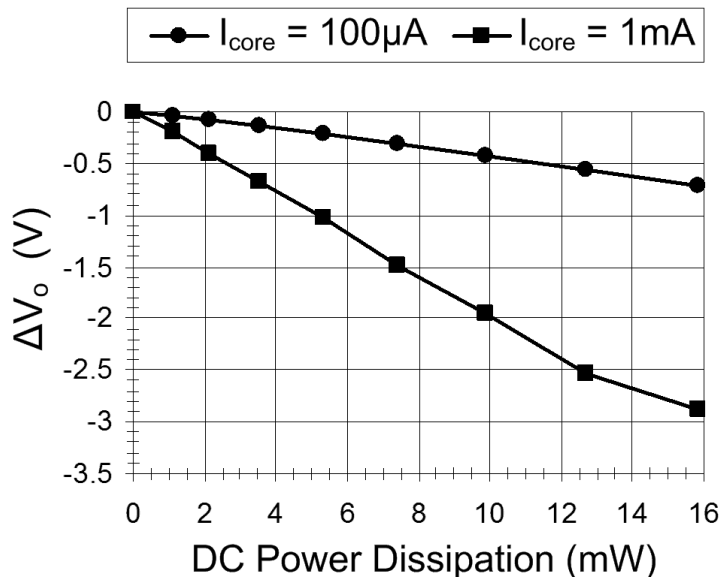


Fig. 69. Sensor output vs. power dissipation at resistor  $R_t$ . The measurements were performed with  $I_{\text{core}} = 100\mu\text{A}$  (sensitivity = 41.7mV/mW) and  $I_{\text{core}} = 1\text{mA}$  (sensitivity = 199.6mV/mW). Distance between  $R_t$  and  $Q_2$ : 4  $\mu\text{m}$ .

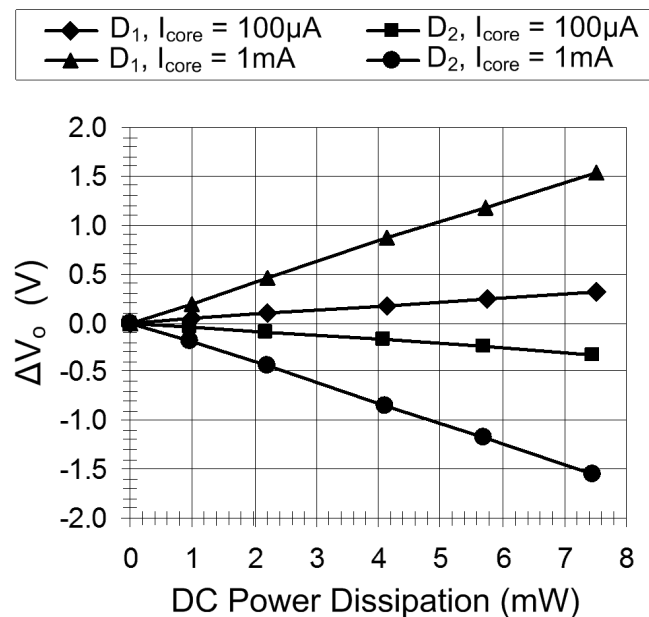


Fig. 70. Sensor output vs. power of diode-connected MOS transistors  $D_{1,2}$ . The measurements were performed with  $I_{core} = 100 \mu A$  (sensitivity =  $42.0 \text{ mV/mW}$ ) and  $I_{core} = 1 \text{ mA}$  (sensitivity =  $207.9 \text{ mV/mW}$ ). Distance between  $D_{1,2}$  and  $Q_{1,2}$ :  $4 \mu \text{m}$ .

Fig. 70 displays the plots from the sensor characterization measurements in which the DC power in the diode-connected transistors  $D_{1,2}$  near each sensing device ( $Q_1, Q_2$ ) was swept individually up to the safe limits for the particular device layouts. The results reveal the symmetric nature of the fully-differential circuitry and that the sensitivity to power in the MOS devices is approximately the same as for the resistor within the sensor's linear range, which can also be observed from the sensitivity vs.  $I_{core}$  plots in Fig. 71.

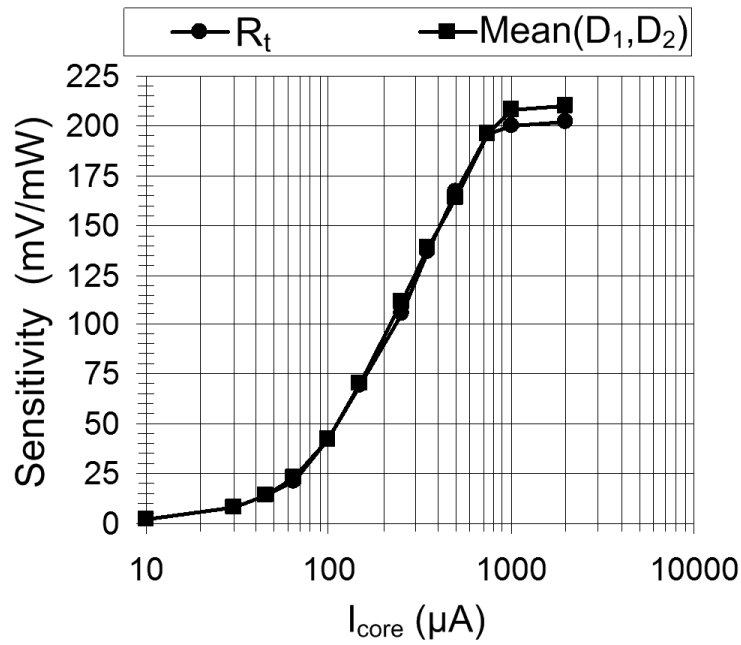


Fig. 71. Sensitivity control to power in  $R_t$  and  $D_{1,2}$  via  $I_{\text{core}}$  adjustments.

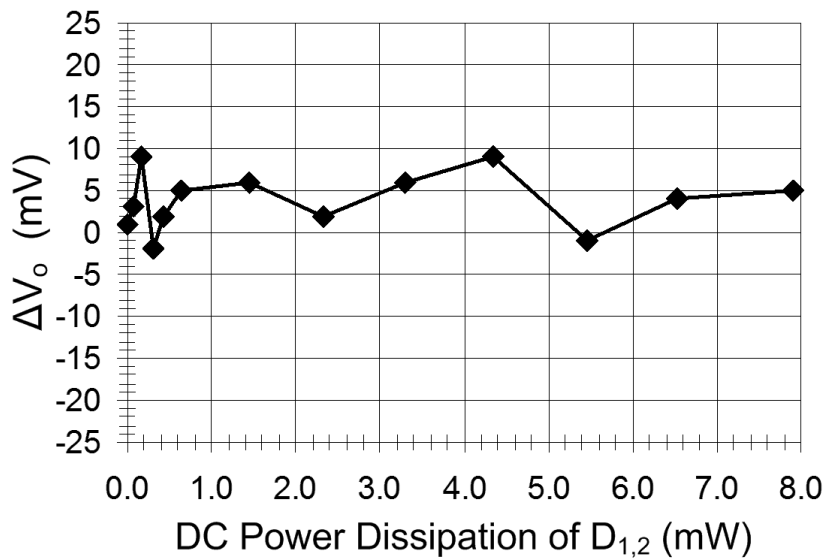


Fig. 72. Common-mode sensitivity of the temperature sensor.  
(The common-mode sensitivity was measured by sweeping the power dissipation in  $D_1$  and  $D_2$  simultaneously with  $I_{\text{core}} = 500\mu\text{A}$ .)

To verify that the sensor has a high rejection to ambient temperature changes,  $D_1$  and  $D_2$  were excited concurrently by injecting DC currents and adjusting the currents such that the measured DC power in both devices is identical for each data point in Fig. 72. Even though the sensitivity to common-mode power is below 10mV/mW, the fluctuations suggest that the sensor calibration step should precede the CUT measurement if the ambient temperature is expected to have changed significantly since the last measurement.

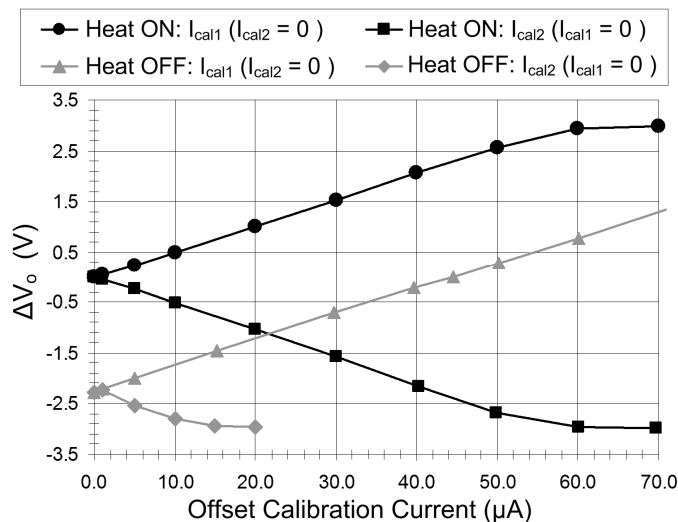


Fig. 73. Offset calibration with currents  $I_{cal1}$  and  $I_{cal2}$  ( $I_{core} = 500\mu\text{A}$ ).

The offset calibration range was evaluated under three conditions: i) when the test heat sources ( $D_1$ ,  $D_2$ ,  $R_t$ ) do not dissipate power and with a deactivated LNA (named “Heat OFF”), ii) with an activated LNA and 3.9mW additional power dissipation in  $D_1$  (named “Heat ON”) to achieve  $\Delta V_o \approx 0\text{V}$  with  $I_{cal1} = I_{cal2} = 0$ , iii) when  $R_t$  alone dissipates 15.9mW. Case i) gives insight into the ability to recover from the sensor’s

inherent electrical offsets due to component mismatches without interference from the LNA's DC bias. As shown in Fig. 73, the differential output voltage has a linear dependence on the calibration currents as long as the electrical amplification stages in the sensor are not saturated, and  $I_{\text{cal1}} = 44.6\mu\text{A}$  is required to compensate for on-chip and off-chip component variations of this prototype design. Case ii) makes it evident that heat sources can also be used to balance the sensor, which in this case requires 3.9mW power in  $D_1$  in addition to the DC bias power of the LNA to achieve  $\Delta V_o \approx 0\text{V}$ . Furthermore, the plot in Fig. 73 under the "Heat ON" condition shows the symmetry of the output voltage dependence on  $I_{\text{cal1}}$  and  $I_{\text{cal2}}$ . In case iii), the 15.9mW power dissipation at  $R_t$  without activation of other heat sources creates an extreme imbalance in the operating conditions of the two bipolar transistors due to both the offset from process variations and the extra temperature gradient. The measured sensor output voltage for this case is plotted versus  $I_{\text{cal1}}$  in Fig. 74, demonstrating that  $I_{\text{cal1}} = 95.6\mu\text{A}$  establishes a balanced output and that the offset compensation capability spans the linear range of the sensor circuitry. The offset calibration currents were adjusted to compensate for DC temperature gradients and electrical offsets by obtaining  $\Delta V_o \approx 0\text{V}$  prior to each set of measurements under certain bias conditions, which requires adjustments in the micro-ampere range. In practice, the ADC and digital post-processing will limit the test time because the settling times of the temperature change (Fig. 57) and amplifier (Table XII) are below 10 $\mu\text{s}$  and 500ns, respectively. However, up to 18 clock cycles could be required for the calibration phase (assuming 6-bit programmability for the calibration test sources and a binary search algorithm until  $\Delta V_o \approx 0\text{V}$ ), averaging of several sensor

output measurements (might be required in a noisy system-on-chip environment), and test control operations. At a 100KS/s rate, this would imply 0.18ms per test point. The test time could be even shorter with the availability of a faster on-chip ADC or off-chip test resources in a production test environment.

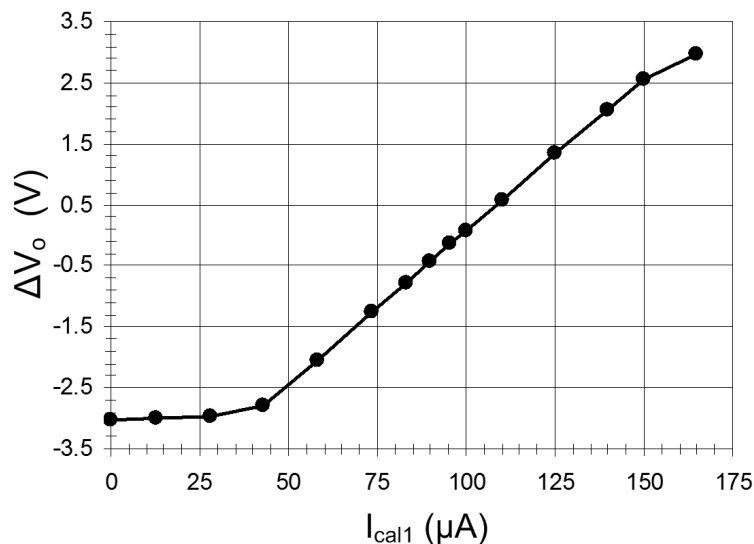


Fig. 74. Offset calibration range with  $I_{cal1}$  ( $I_{cal2} = 0$ ,  $I_{core} = 500\mu A$ ).

#### V.4.2. RF testing with the on-chip DC temperature sensor

Table XIII gives an overview of the CUT parameters that are relevant to the correlation of its RF output and the temperature sensor output. The RF measurements were taken around 1GHz because the parasitics of the QFN package and PCB assembly degraded  $S_{11}$  to worse than -6.3dB at higher frequencies. Losses from cables, power combiner, bias-T, and impedance mismatches were characterized and de-embedded from the measurements reported below. A spectrum analyzer was used to measure the CUT output while simultaneously reading the differential sensor output with a DC voltmeter

in order to experimentally verify that the CUT's RF performance can be extracted with temperature sensor measurements. To correlate measurements with simulations, Fig. 75 contains plots of the CUT and sensor outputs from a sweep of the RF input power applied to the CUT with a single tone at 1GHz. Offsets on the y-axes are caused by the ~3dB CUT gain difference between simulations and measurements with extra losses. The curves show that input power levels above -15dBm can be monitored at the output of this DC sensor, which is sufficient when a signal with more power than a typical LNA input signal is applied during testing. Online testing with input signals below -15dBm would require sensor sensitivity improvements. Options that can be explored are designing the sensor with more amplification or implementing  $Q_{1,2}$  with PNP devices that are electrically connected in Darlington configuration to boost the gain and to increase the coupling to the CUT surrounded by two nearby PNP devices ( $Q_1$ ).

Table XIII. Measured CUT\* performance parameters

Parameter	Value at 1GHz
Gain ( $S_{21}$ )	-2.3dB**
1-dB Compression Point	0.5dBm
Third-Order Intercept Point (IIP3)	12.0dBm
$S_{11}$	-6.3dB
$S_{22}$	-12.7dB
$I_{DC}$	8.7mA
Technology / $V_{DD}$	0.18 $\mu$ m CMOS / 2.4V

\* LNA loaded (without buffer) by a 50 $\Omega$  analyzer impedance.

\*\* Reduced due to the external 50 $\Omega$  load in addition to the on-chip load resistor ( $R_L$ ) and due to  $S_{11}$  degradation from packaging/PCB parasitics at 1GHz;  $S_{21} \approx 0$ dB up to 500MHz.

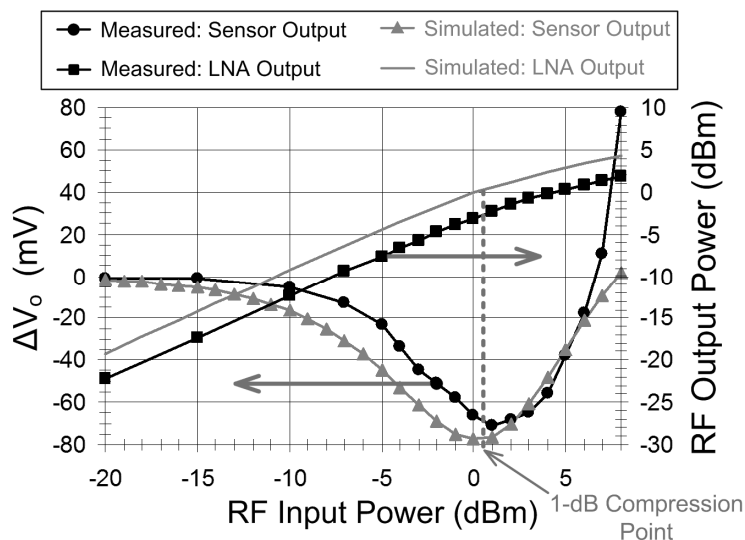


Fig. 75. Measurement vs. simulation comparison for the CUT characterization. The plots show the LNA's 1-dB compression point curve and the DC output voltage of the sensor with  $I_{\text{core}} = 500\mu\text{A}$  ( $167\text{mV/mW}$  sensitivity).

In Fig. 75, the minimum of the temperature sensor's  $\Delta V_o$  curve is  $-71\text{mV}$  with  $1\text{dBm}$  input power. Subtracting the fixed  $0.13\text{dB}$  shift according to the simulations results in Section V.3.4, the estimated 1-dB compression point is  $0.87\text{dBm}$ . This value approximates the electrically measured 1-dB compression point with an error of  $0.37\text{dB}$ , which is comparable to standard RF power detectors in BIT applications. As described in Appendix D, estimation inaccuracies create further uncertainty of  $\pm 0.6\text{dB}$ , yielding up to  $1\text{dB}$  error for the 1-dB compression point prediction.

Compared to the simulated plot in Fig. 75, it can be observed that the measured minimum is about 10% higher due to electro-thermal modeling inaccuracies. This discrepancy is acceptable since the sensitivity of the sensor can be adjusted with  $I_{\text{core}}$  over a tuning range of roughly a decade (Fig. 71). A log-magnitude plot of the measured sensor output voltage vs. CUT input power is displayed in Fig. 76 to visualize how the



1-dB compression point corresponds to the vicinity of the peak log-magnitude of the sensor output voltage.

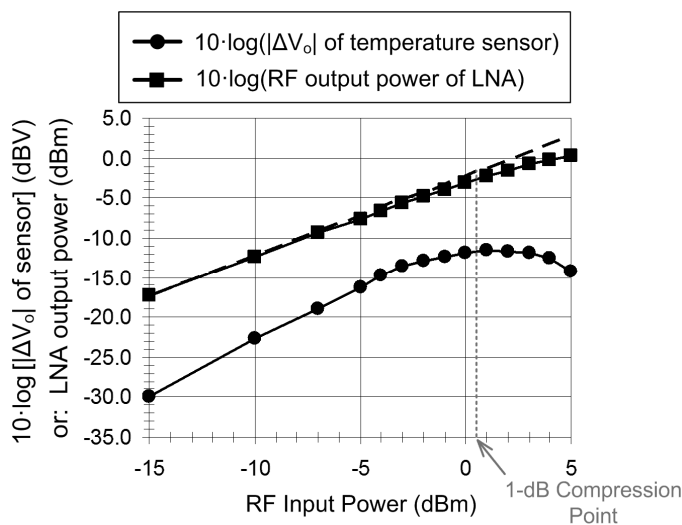


Fig. 76. LNA output power and log-magnitude of the sensor output voltage.  $I_{\text{core}}$  was  $500\mu\text{A}$  ( $167\text{mV/mW}$  sensitivity) during these measurements.

Fig. 77 displays the CUT's output spectrum around 1GHz that was obtained with two  $-22.2\text{dBm}$  test tones having a separation of  $200\text{KHz}$ . As reference, the third-order intermodulation (IM3) of  $-67.4\text{dB}$  is annotated for this linear operating condition. The  $1\text{dBm}$  input power level was identified as critical nonlinear point based on the temperature sensor output measurements. For comparison, Fig. 78 shows the output spectrum with two  $-2.2\text{dBm}$  test tones that have a combined power of  $1.2\text{dBm}$ . The resulting IM3 is  $-29.9\text{dB}$ , which demonstrates the usefulness of this point as indicator for nonlinear operation. Since the DC temperature sensor characterization of the CUT circumvents the use of RF measurement equipment, it provides a viable alternative to

monitor RF signal levels and linearity performance in BIT applications and pass/fail production testing in which a 1dB error is permissible.

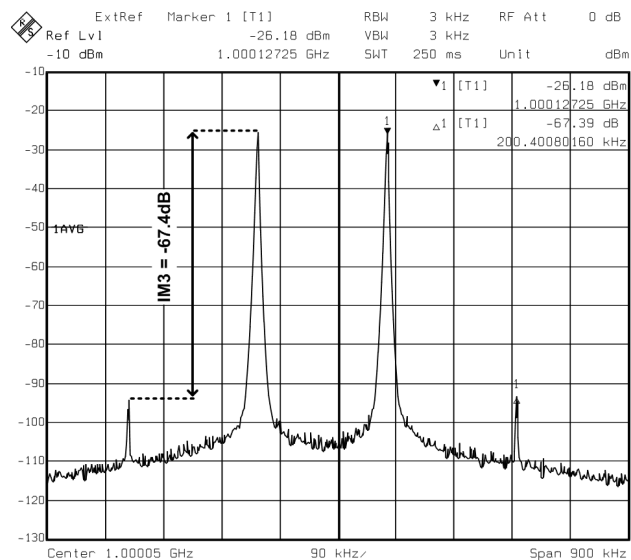


Fig. 77. The CUT's output spectrum from a two-tone test around 1GHz (case 1). Measured with: 200KHz tone spacing, -22.2dBm per tone (-19.2dBm combined).

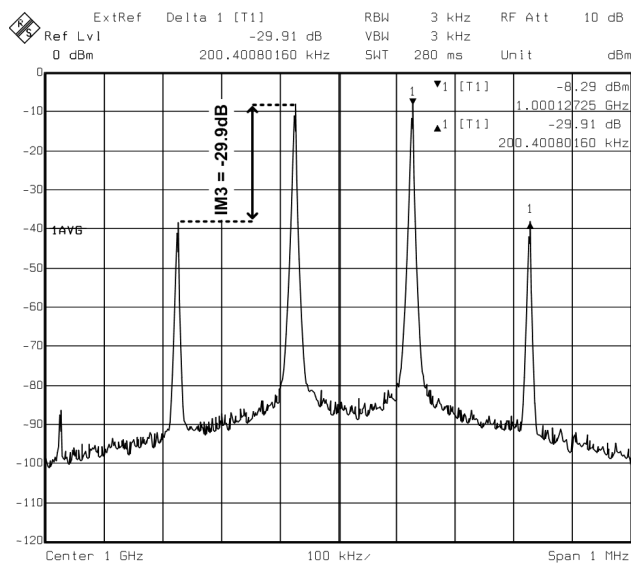


Fig. 78. The CUT's output spectrum from a two-tone test around 1GHz (case 2). Measured with: 200KHz tone spacing, -2.2dBm per tone (1.2dBm combined).

## V.5. Summarizing Remarks

A sensing methodology was proposed that exploits the intrinsic down-conversion of circuit performance information from the RF domain to the DC domain with the homodyne temperature measurement approach. It was shown that this property is useful for application in built-in testing and monitoring of on-chip thermal gradients that can impact system performance. Since this alternative technique does not require a connection to the circuit under test or the signal path, it provides a non-influential method for monitoring variations. The presented CMOS-compatible sensor architecture has been developed for the wide dynamic range and programmability requirements as built-in power detector based on the homodyne approach.

Furthermore, an electro-thermal design procedure for differential temperature sensors has been experimentally validated. Coupling at low frequencies could impact the CUT's operation, which can be evaluated with electro-thermal simulations. Measurement results obtained with an RF amplifier and a  $0.012\text{mm}^2$  built-in temperature sensor on a  $0.18\mu\text{m}$  CMOS test chip revealed that the same sensor can detect the DC and RF power dissipation, and that the 1-dB compression point can be predicted from the sensor's output with an error below 1dB without RF measurement equipment.

## **VI. MISMATCH REDUCTION FOR TRANSISTORS IN HIGH-FREQUENCY DIFFERENTIAL ANALOG SIGNAL PATHS**

### **VI.1. Background**

Until now, the approaches discussed in this dissertation are mostly aimed at making analog and mixed-signal circuits more robust by either circumventing their dependence on mismatches or by introducing digitally programmable elements for post-fabrication adjustments. An alternative approach to deal with rising variability is to decrease the mismatches of analog circuits by lessening them in a statistical sense. The approach discussed in this section is targeting the static mismatch between critical transistors in particular, where the goal is to decrease the standard deviation of the parameter variations by employing an automatic analog calibration loop.

Device mismatches become more severe as technology scaling continues, especially when minimum transistor dimensions are used to optimize for high-speed operation or to bias with high overdrive voltage for yield enhancement [109]. In addition to higher percent errors for small fabrication dimensions, the threshold voltage mismatch worsens even for neighboring transistors due to the increasing effect of dopant fluctuations in modern CMOS processes [11]. The resulting offsets degrade the performance of analog circuits that rely on device matching. For example, the second-order intermodulation intercept point (IIP2) of mixers strongly depends on matching of transistors, for which a digital mismatch reduction scheme was proposed in [110] to adjust gate bias voltages separately for each switching transistor.

Another issue in RF circuit design is that designers might place transistors next to each other with a safe distance instead of elaborately matching them in the layout. Even though the use of non-minimum dimensions can reduce process variations, devices with large area (i. e., large parasitic capacitances) in the signal path are often not feasible since they imply increased power consumption and/or performance degradation, which is the case in high-speed amplifiers and comparators [111]. Similarly, layout matching techniques such as interleaved or common-centroid styles create more high-frequency coupling through parasitic capacitances of crossing metal lines or leakage through the substrate due to the proximity of the devices. An alternative design technique towards the goal of alleviating transistor mismatches is proposed in this section. The method involves an analog calibration loop in which device mismatches are indirectly detected and reduced through layout-based parameter correlations rather than directly measuring characteristics of the circuit. This calibration loop continuously operates in the background without requiring digital resources or switches in the signal path. Its short convergence time below  $10\mu\text{s}$  prevents excessive start-up calibration time for time-critical situations such as during production testing.

## **VI.2. A Mismatch Reduction Technique for Differential Pair Transistors**

### *VI.2.1. Approach*

In RF applications, designers may choose to place transistors next to each other with a safe distance as shown in Fig. 79 instead of matching them in the layout. The advantage with such a configuration is that the physical separation of the devices provides isolation against RF signal leakage that leads to crosstalk between the

differential signal paths. Often, each RF transistor is surrounded by a guard ring for enhanced isolation and by deep trenches (if available). A drawback in this scenario is that the unmatched devices have significant parameter mismatches which are observable through the static drain current difference.

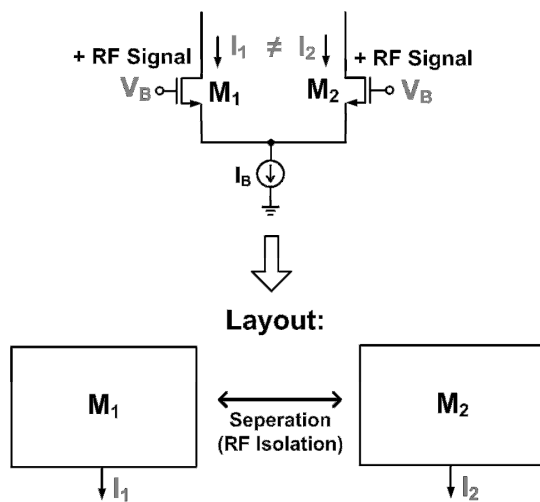


Fig. 79. An unmatched RF transistor pair.

To alleviate the mismatch problem, the alternative approach visualized in Fig. 80 is proposed here. Instead of matching the RF transistors  $M_1$  and  $M_2$  to each other, they are individually matched to mismatch-sensing transistors  $M_{1S}$  and  $M_{2S}$  in a DC calibration loop. Thus, the currents  $I_{1S}$  and  $I_{2S}$  of the mismatch-sensing transistors are correlated to  $I_1$  and  $I_2$  of the main transistor pair, respectively. Even though it is optimal to use the same dimensions and number of fingers for  $M_{1S}$  and  $M_{2S}$  as for  $M_1$  and  $M_2$ , they do not have to be identical. However, their electrical device parameters must be correlated to  $M_1$  and  $M_2$  through layout matching techniques. The feedback action in the loop

compares  $I_{1S}$  to  $I_{2S}$  and adjusts the separate gate bias voltages  $V_{B1}$  and  $V_{B2}$  of the mismatch-sensing transistors until the currents are approximately equal to each other. Consequently, the drain current difference in the main transistor pair is also reduced due to the parameter correlations between the matched transistors and the shared gate bias voltages. In this way, the mismatches are lessened while the RF isolation between the main transistors is maintained. Additionally, low-pass filter nodes within the calibration loop suppress any RF signal that might couple into it through layout parasitics.

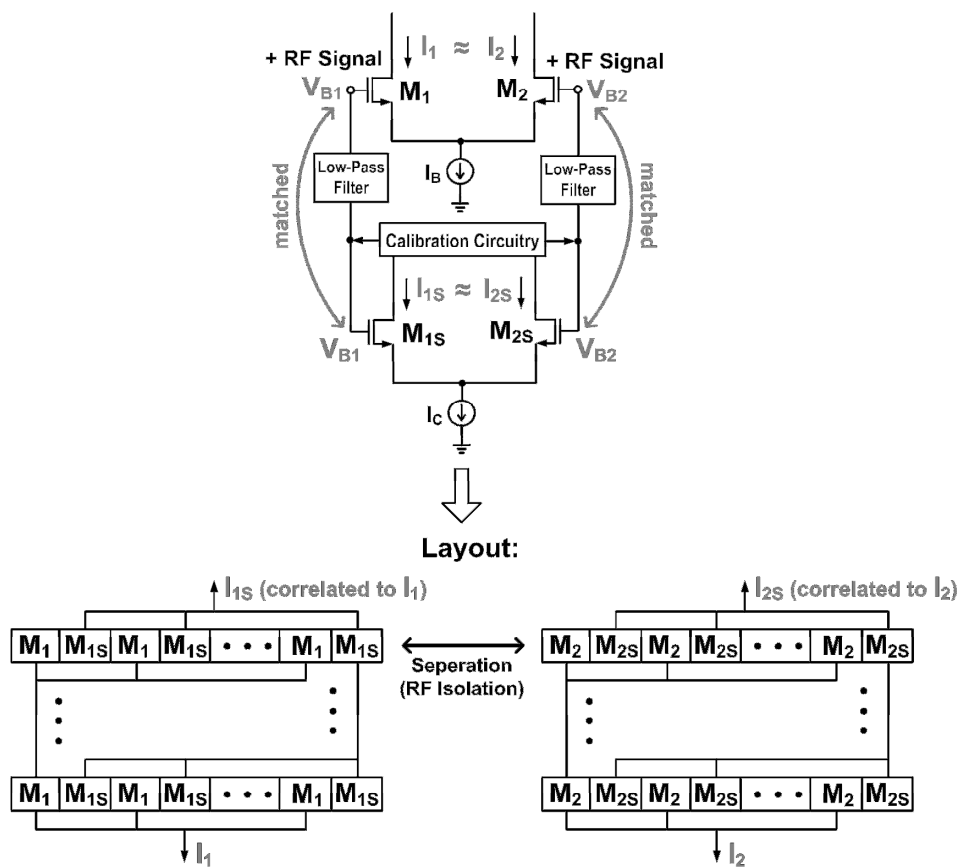


Fig. 80. An RF transistor pair with DC mismatch reduction loop.

To demonstrate the abovementioned concept, Fig. 81 depicts a differential amplifier consisting of a transistor pair ( $M_1, M_2$ ) with polysilicon resistor loads ( $R_L$ ), where the resistor dimensions were selected large enough to ensure that the input-referred offset voltage is dominated by  $M_1$  and  $M_2$ . Table XIV lists the device dimensions for the circuit. The characteristics of  $M_1$  and  $M_2$  are ideally equal, but considerable deviations occur when they are not matched in the layout through interleaved, common-centroid, or similar configurations. Hence, crosstalk between the differential signal paths is avoided by physically separating them, while parameter variations of  $M_1$  and  $M_2$  should be treated as uncorrelated. However,  $M_1$  and  $M_2$  can be laid out with  $N$  ( $=20$  in this example) subdevices, and matched to sensing-transistors  $M_{1S}$  and  $M_{2S}$  respectively. In this configuration,  $M_{1S}$  and  $M_{2S}$  are part of the DC calibration loop that detects a mismatch between currents  $I_1$  and  $I_2$ , and that generates bias voltages  $V_{B1}$  and  $V_{B2}$  individually for each branch. If the drain currents of  $M_{1S}$  and  $M_{2S}$  are forced to be equal in the absence of mismatches, then their gate-source voltage overdrives must be equal [11], which only occurs when  $V_{C1} = V_{C2}$  in Fig. 81. Here,  $M_{1S}$  and  $M_{2S}$  are placed in a differential amplifier configuration with a tail-current source ( $I_B/10$ ) and active loads ( $M_3, M_4$ ) for high gain with self-regulation via feedback resistors ( $R_{cm}$ ). Capacitors ( $C_{st}$ ) stabilize the loop by creating a dominant pole at nodes  $V_{C1}$  and  $V_{C2}$ . If  $I_1 \neq I_2$  in the presence of device mismatches, then the resulting imbalance of  $V_{C1} - V_{C2}$  is amplified by the amplifier (A). The feedback action differentially adjusts  $V_{B1}$  and  $V_{B2}$  until  $V_{C1} \approx V_{C2}$  to minimize mismatches without requiring on-chip digital resources. Capacitors ( $C_{filt}$ ) are included to filter out high-frequency noise. Amplifier A, whose schematic is





Table XIV. Differential amplifier and calibration loop components

Component	Dimensions / Value
$M_1, M_2, M_{1S}, M_{2S}$	$W/L = 90\text{nm} \times 20 \text{ fingers} / 90\text{nm}$
$M_3, M_4$	$W/L = 6.25\mu\text{m} \times 8 \text{ fingers} / 3.7\mu\text{m}$
$R_L$	$1.12\text{k}\Omega$ ( $L/W = 9\mu\text{m} / 2\mu\text{m}$ )
$C_L$	0.1pF
$R_B$	100k $\Omega$
$C_{\text{filt}}$	1pF
$C_c$	5pF
$C_{\text{st}}$	10pF
$R_{\text{cm}}$	100k $\Omega$ ( $L/W = 20 \times 10\mu\text{m} / 1\mu\text{m}$ )
$I_B$	1mA
Technology / Supply Voltage	90nm CMOS / 1.2V
<b>Operational Transconductance Amplifier (A)</b>	
$M_N$	$W/L = 8\mu\text{m} \times 4 \text{ fingers} / 4\mu\text{m}$
$M_P$	$W/L = 5\mu\text{m} \times 2 \text{ fingers} / 1.55\mu\text{m}$
$M_B$	$W/L = 3\mu\text{m} \times 4 \text{ fingers} / 1\mu\text{m}$
$R_{\text{fb}}$	38k $\Omega$ ( $L/W = 8 \times 19\mu\text{m} / 1\mu\text{m}$ )
$I_{\text{bias}}$	50 $\mu\text{A}$
DC Gain: Amplifier, Calibration Loop	18.3dB, 38.6dB

This scheme exploits that the parameters of  $M_1/M_{1S}$  (and  $M_2/M_{2S}$ ) are highly correlated so that the mismatch can be continuously extracted in the background to compensate for drifts from temperature changes as well as process variations. Since the calibration loop has several low-pass filtering nodes, the differential signal integrity is not jeopardized by coupling between  $M_1$  and  $M_2$  through the loop. Instead, coupling to  $M_{1S}/M_{2S}$  via layout parasitics and substrate leakage due to the matching only create small signal losses. The large bias resistors ( $R_B$ ) prevent that the input capacitances

looking into the gates of  $M_{1S}$  and  $M_{2S}$  cause any significant loading effects at the RF inputs (In+, In-).

The accuracy of the proposed method relies on the matching between  $M_1/M_{1S}$  and  $M_2/M_{2S}$ , which depends on their number of subdevices [112], [113]. Let  $\sigma_{\Delta V_{th}}$  be the standard deviation of the threshold voltage difference for an unmatched transistor pair. In a matched pair with  $N$  fingers and the same effective dimensions in a stripe pair structure, this standard deviation decreases to [113]:

$$\sigma_{\Delta V_{th}(m)} = \sigma_{\Delta V_{th}} / \sqrt{N} . \quad (32)$$

More complex common-centroid configurations are expected to improve the spread reduction, but the aforementioned relationship will be used as plausible worst-case estimate. Being outside of the signal path, the parasitic capacitances of the matched transistors in the core of the calibration loop do not affect the RF performance. Hence, their dimensions can be increased to ensure that their offsets are negligible. Therefore, non-minimum transistor lengths (L) and widths (W) were selected (Table XIV, Fig. 81, Fig. 82) for matched pairs  $M_3/M_4$ ,  $M_B$ ,  $M_N$ ,  $M_P$ ,  $I_B$  (NMOS current mirror), and  $I_B/10$  based on the inverse proportionality of  $\sigma_{\Delta V_{th}}$  to  $\sqrt{W \cdot L}$  [114]. Likewise, the polysilicon resistors  $R_{cm}$  and  $R_{fb}$  were sized sufficiently large with the help of statistical device models and Monte Carlo simulations.

### VI.2.2. Simulation results

The test circuit (Fig. 81, Fig. 82) was designed using UMC 90nm CMOS technology with a 1.2V supply, and simulations were performed with the foundry's

statistical device models. The loaded differential amplifier under calibration has a gain of 13dB with a -3dB bandwidth of 2.14GHz. Its minimum AC input impedance magnitude within the passband is 1.77k $\Omega$ , which changes less than 1% when the loop is added and activated. The amplifier (A) has a loaded DC gain of 18.3dB, resulting in an overall gain of 38.6dB in the loop starting at  $V_{B1}/V_{B2}$  and traversing through  $V_{C1}/V_{C2}$ .

Device matching was taken into account during the Monte Carlo analysis with the Cadence Spectre simulator (process and mismatch variations enabled) by calculating the expected spread reduction in equation (32) based on the number of fingers in each matched pair. According to this reduction, the corresponding correlation coefficient ( $C_m$ ) was specified from the relation given in [115]:

$$1/\sqrt{N} = \sqrt{1 - C_m} \quad . \quad (33)$$

For example, the  $N = 20$  fingers ( $C_m = 0.95$ ) of the matched pairs  $M_1/M_{1S}$  and  $M_2/M_{2S}$  leads to an expected spread reduction of 4.47 with the proposed scheme when other offsets in the loop are negligible. Fig. 83 displays the histograms of the input-referred offset voltage of the amplifier obtained with 100 Monte Carlo runs at 30°C, showing that its standard deviation decreases from 4.17mV to 1.29mV when the calibration loop is added. Notice that an input offset decrease from 4.17mV to 1.29mV corresponds to a drain current difference reduction from 3.1% to 1.0% for  $M_1$  and  $M_2$ ). At -40°C and 100°C, 100 Monte Carlo runs revealed that the predicted offset decreases from 4.10mV to 1.22mV and from 4.25mV to 1.40mV, respectively. With the large-sized devices in the calibration circuit, the accuracy improvement mainly depends on the correlation of the parameters between  $M_1/M_{1S}$  (and  $M_2/M_{2S}$ ). For instance, using  $C_m =$

0.99 in the simulation instead of the previous worst-case assumption, the input offset with calibration reduces to 0.76mV (0.6%  $M_1/M_2$  drain current difference); provided that 20 subdevices can be appropriately matched with a common-centroid layout.

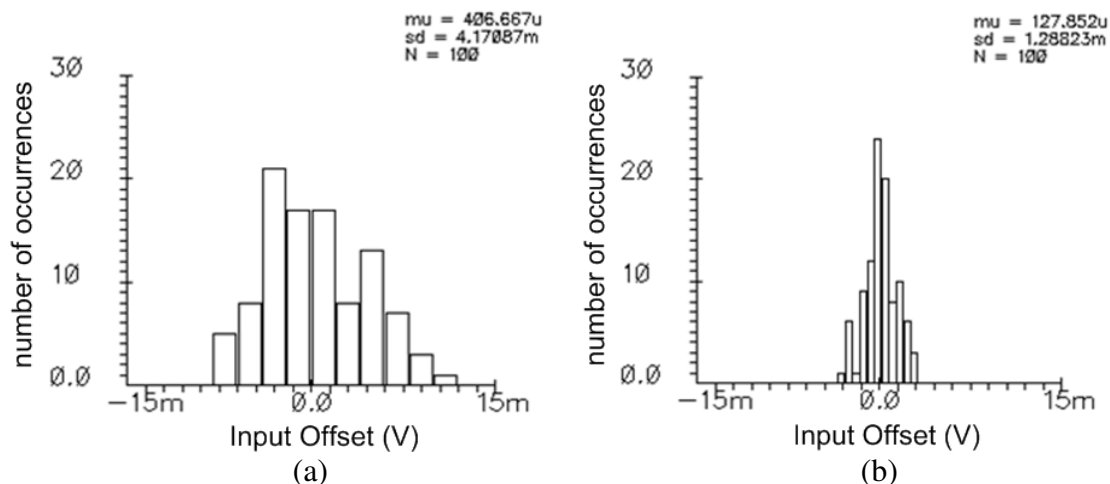


Fig. 83. Monte Carlo simulation results (100 runs at 30°C).

Input-referred offset voltage of the differential amplifier:

- (a) without mismatch calibration (ideal bias voltages:  $V_{B1} = V_{B2} = V_B$ ),  
 (b) with activated mismatch calibration loop.

### VI.3. Second-Order Nonlinearity Enhancement for Double-Balanced Mixers

#### VI.3.1. Introduction

Second-order nonlinearity of the down-conversion mixer is typically the bottleneck for the overall achievable second-order intermodulation intercept point (IIP2) performance with direct-conversion and low-IF receiver architectures ([116], [117]), which are appealing low-power architectures for low-cost portable wireless devices. Thus, stringent IIP2 specification demands are imposed on mixers in these systems,

especially with the tendency towards wider bandwidths that leads to increased interference signals at the RF front-end. For instance, a minimum mixer IIP2 requirement of 60dBm has been identified for the UMTS receiver design budget in [118]. Similarly, the down-conversion mixer IIP2 for WCDMA systems has been specified as 59dBm in [119], whereas the IIP2 target for the WCDMA/CDMA2000 mixer in [120] was 50dBm. Even though the IIP2 mixer requirement depends on the given communication standard and system-level design, 50dBm can be regarded as the minimum tolerable mixer IIP2 for direct-conversion receivers based on findings in the literature. A general approach to derive this mixer specification is given in [121], where even the need for mixers with IIP2 > 70dBm has been outlined.

#### IIP2 degradation mechanisms with ideal switching transistors in the core

The schematic of a double-balanced mixer ([122]) is displayed in Fig. 84, in which the bias circuitry is omitted for simplicity. Transistors labeled  $M_{RF}$  are the input transconductors to which the RF signal is applied. Assuming a hard-switching local oscillator (LO) signal and the corresponding square-wave approximation, it has been shown in [116] that the IIP2 can be estimated with the following equation when the switching core transistors ( $M_{SW}$ ) are considered ideal:

$$IIP2 \approx \frac{\sqrt{2}}{\pi\eta_{nom}} \frac{a_2}{\alpha_2} \times \frac{4}{2 \cdot \Delta\eta(\Delta g_m + \Delta A_{RF}) + \Delta R_L(1 + \Delta g_m)(1 + \Delta A_{RF})} \quad (34)$$

Parameters  $g_m$ ,  $a_2$ , and  $\Delta g_m$  in (34) are the nominal transconductance, second-order non-linearity coefficient, and transconductance deviation of the two transistors  $M_{RF}$ .  $\Delta A_{RF}$  is the amplitude difference at the RF+ and RF- inputs, and  $\Delta R_L$  is the discrepancy between



that 5-bit programmability in the mixer load resistors leads to receiver IIP2 improvements in the 23-26dB range. Analogously, when the mixer load contains current sources, an additional feedback loop can be added to reduce the common-mode output impedance mismatch for approximately 20dB IIP2 enhancement [124].

Revisiting equation (34), another observation is that any of the parameters in the second denominator can be tuned to minimize this mismatch-dependent denominator. Various IIP2 improvement schemes involve tuning of parameters other than the load mismatch. In [125] for example, an LO buffer with tunable phase for one of the differential outputs was used to change the duty cycle term ( $\Delta\eta$ ) in order to maximize IIP2. Alternatively, LO duty cycle modification is also possible by adjusting the gate bias voltages of the individual LO transistors, which affects the turn on/off time instants of the switches [126]. However, notice that such an approach will impact the maximum achievable IIP2 limit under consideration of mismatches in the LO transistors because the LO bias conditions are altered as discussed in the next subsection. It was also shown in [125] that programmable bias circuitry for one of the  $M_{RF}$  transistors can be employed to vary the transconductance mismatch ( $\Delta g_m$ ) until a maximum IIP2 is reached based on (34). The effectiveness of these abovementioned tuning methods depends on the resolution of the programmable elements or the accuracy of the calibration loop, generally providing 20-30dB higher IIP2 after tuning.

#### IIP2 degradation mechanisms with non-ideal switching transistors in the core

The results with the methods summarized in the previous subsection demonstrate the capabilities of IIP2 tuning based on the ideal hard-switching LO model with



negligible mismatches in the switching transistors. However, the intrinsic IIP2 limit depends primarily on the mismatches in the switching transistors [117] for fully-differential double-balanced mixers (e.g., the mixer in Fig. 84 with a shared tail current source added at the sources of the  $M_{RF}$  transistors). In the pseudo-differential case (e.g., the mixer in Fig. 84 without any modifications), the intrinsic IIP2 limit depends predominantly on the input transconductor as well as on the switching transistor mismatch, where a common-mode feedback circuit at the IF output can be used to suppress the input transconductor's contribution to the IIP2 [127]. This makes the mismatch of the LO switching transistors critical for the achievable best-case IIP2.

Let  $L$  be the low-frequency leakage parameter due to mismatches between the  $M_{SW}$  transistors in Fig. 84. A detailed expression for  $L$  can be found in [117], but it is important to point out here that this parameter is zero for perfectly matched  $M_{SW}$  transistors, and that it is directly proportional to the relative offset voltages of non-ideal  $M_{SW}$  transistors. Thus,  $L$  is a statistically-varying mismatch parameter. Its impact on the RMS voltage of the IIP2 is evident from the following equation [117]:

$$\sigma_{IIP2} = \frac{(2/\pi) \cdot g_m}{\sqrt{L^2 \cdot [(\alpha_2^{dif})^2 + (\alpha_2^{cm})^2] + [(\Delta R_L / R_L) \cdot \alpha_2^{cm}]^2}}, \quad (35)$$

where  $R_L$  and  $\Delta R_L$  are the load resistors in Fig. 84 and their mismatch, respectively. As before,  $g_m$  is the transconductance of the RF input transistor  $M_{RF}$ , whose second-order nonlinearity has a differential component  $\alpha_2^{dif}$  and a common-mode component  $\alpha_2^{cm}$ . Equation (35) reflects that load resistor mismatch only degrades IIP2 in the presence of  $\alpha_2^{cm}$ , which is alleviated when a common-mode feedback is added [127] or when fully-

differential input transconductors with high common-mode rejection at low frequencies (within IF bandwidth) are employed [118]. On the other hand, the mismatches of the LO switching transistors limit the achievable IIP2 through parameter  $L$  and the combined input transistor nonlinearities. Even if  $\alpha_2^{cm}$  is made negligible by designing with high common-mode rejection, the differential second-order nonlinearity  $\alpha_2^{dif}$  will deteriorate IIP2 with non-perfectly matched LO transistors. The approach presented in [110] aims at cancelling the offset between the LO transistors by using separate digitally programmable gate bias voltages. With regards to equation (35), this means a reduction of parameter  $L$  by setting the switches to the exact combination that gives minimal offsets between the transistors, resulting in simulated (theoretical) IIP2 improvements up to roughly 40dB with 6-bit resolution of the bias adjustment voltage. The mixer calibration technique proposed in Section VI.3.2 applies the automatic analog calibration scheme from Section VI.2 for reduction of the LO transistor mismatches in order to boost the intrinsic IIP2 limit based on equation (35).

#### IIP2 calibration with digital control

Regardless of which mechanisms degrade the IIP2, a DC offset can be dynamically injected at the output of the mixer to improve the IIP2. A system-level IIP2 calibration technique has been demonstrated in [128] by injecting an offset current at the mixer output with a digitally controllable current source having 6-bit resolution. Such a scheme is aligned with the system-level calibration approach discussed in Section II.2.4. The ADC output in the receiver is analyzed in the digital signal processor to control the offset current sources based on the digitally measured static and dynamic DC offsets.

Similarly, the calibration in [129] involves an auxiliary second-order intermodulation (IM2) generator that cancels the IM2 in the mixer. The IM2 generator contains a programmable scaling unit that can be adjusted for optimum IIP2 performance when IIP2 monitoring capabilities exist on the chip. Another digital calibration method utilizes a least-mean-square (LMS) algorithm operating on the digitized output of a common-mode detector at the mixer output and the baseband filter's output to tune the IIP2 by injecting a DC current [130]. Even though digital approaches are effective and allow calibration control through the DSP, they typically involve significantly longer convergence times compared to analog control loops. Additionally, they rely on DSP resources for the measurement of performance degradation and the corresponding corrective actions, which might not be available on the chip with the RF front-end circuitry.

#### Autonomous IIP2 reduction/cancellation

The benefits and trade-offs of digital and analog circuit-level calibrations have been discussed in the subsections of Section II.2. Instead of using digitally programmable elements to tune IIP2, automatic analog feedback loops can be employed as well. The work in [131] is a representative paradigm for analog IIP2 calibration, which involves an IM2 generator whose output determines how much current is injected into the mixer core to cancel the IM2. With such a scheme, the amount of IIP2 improvement (e.g., 22dB from simulations in [131]) depends on the gain in the feedback loop. In theory, the IM2 component with calibration is given by

$$IM2_{cal} = \frac{IM2_i}{1+A_L}, \quad (36)$$

where  $IM2_i$  is the IM2 without calibration and  $A_L$  is the loop gain. In practice, the calibration circuitry must be designed with care to avoid that component offsets and mismatches degrade its effectiveness. Since the calibration loop bandwidth is typically in the range of the IF signal bandwidth, the required frequency response is usually achievable using non-minimum device dimensions to lessen mismatches.

Another reported IIP2 improvement method involves cancellation of the input transistor's second-order nonlinearity parameter  $\alpha_2$  in equation (34) with a modified bias network that serves as IM2 generator [132]. Simulations of this alternative approach indicate that 20-40dB IIP2 improvement is achievable with this method even though it does not involve a feedback loop.

### VI.3.2. Proposed mixer calibration

In this work, the objective is to improve the intrinsic IIP2 of a double-balanced down-conversion mixer by reducing the mismatches of the LO switching transistors that proportionally increase the leakage parameter  $L$  in equation (35). It is intended for applications in which limited on-chip digital computational resources are available or in which a fast analog IIP2 tuning at start-up helps to reduce the convergence time and required range of a digital system-level calibration algorithm.

Fig. 85 gives an overview of the proposed calibration for a double-balanced mixer based on the mismatch reduction loop discussed in Section VI.2. Here, the goal is to force equal currents in the calibration branches ( $I_{D(M1S)} \approx I_{D(M2S)} \approx I_{D(M3S)} \approx I_{D(M4S)}$ ),

minimizing their mismatches and the corresponding mismatches in the transistors of the mixer that are switched by the LO signal.

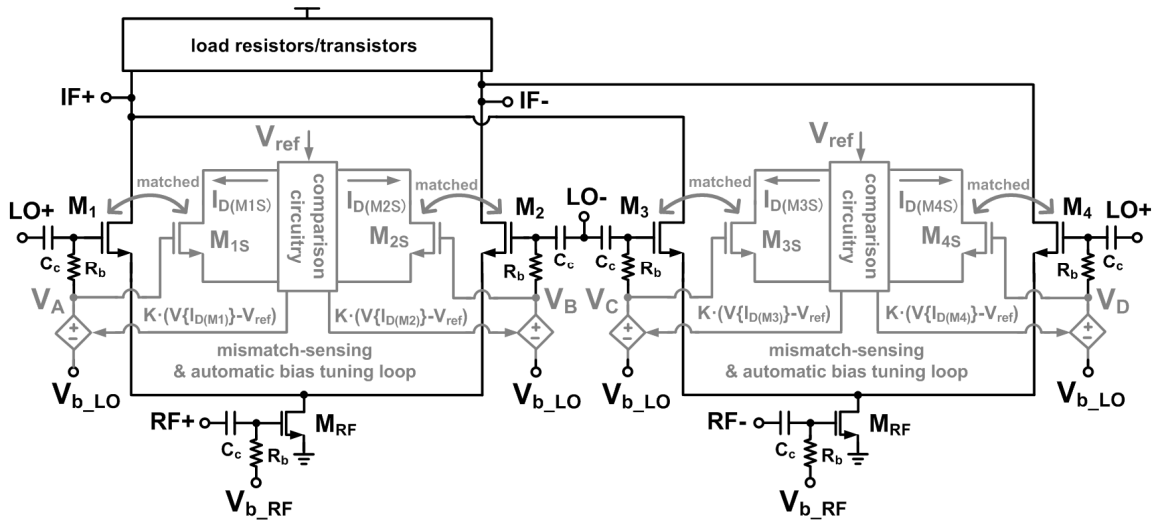


Fig. 85. Mixer with conceptual mismatch reduction for the LO transistors.

The comparison circuitry in Fig. 85 utilizes the same mechanism to accomplish the mismatch reduction as the calibration loop described in Section VI.2. In this circuit, the LO transistors  $M_1$ - $M_4$  are assumed to be matched to the associated mismatch-sensing transistors  $M_{1S}$ - $M_{4S}$  in the layout, which results in the parameter correlations described in Section VI.2. Within the comparison circuitry, all currents from the sensing transistors are converted to a voltage  $V\{I_{D(Mx)}\}$  which is then compared to a common reference  $V_{ref}$ . The difference is amplified by a factor  $K$  within the control loops for the individual bias voltages  $V_A$ - $V_D$ . These gate bias voltages are shared by each LO transistor and its mismatch-sensing transistor, and they are controlled around the gate bias voltage  $V_{b\_LO}$  with which the mixer is designed. Notice that bias resistors ( $R_b$ ) and coupling capacitors

( $C_c$ ) form high-pass filters that allow the RF signals to pass, whereas the DC mismatch calibration circuitry contains low-pass filters (not shown). The high-valued resistors ( $R_b$ ) further isolate the calibration circuitry from the LO signal. It is also worth mentioning that the gate bias voltage  $V_{b\_RF}$  for the input transconductor in Fig. 85 is independent of the calibration loop and available for tuning. In receivers with I/Q paths, this gate bias voltage of the transconductor  $M_{RF}$  can be adjusted for I/Q amplitude matching of the mixer outputs in both paths [123].

The key building blocks of the calibration scheme are displayed in Fig. 86. All mismatch-sensing transistors have a shared tail current source  $I_C$ . Without mismatches, the currents in all four sensing branches are identical. The voltages  $V_1$ - $V_4$  are also equal in the absence of mismatches since they are derived from comparisons of the drain currents of  $M_{1S}$ - $M_{4S}$  with the same current  $I_P$  from well-matched current sources with large transistor dimensions. Notice that the current  $I_P$  is controlled by a common-mode feedback ( $CMFB_{cal}$ ) loop that regulates the high-impedance nodes at the drains of the sensing-transistors to maintain the average of  $V_1$ - $V_4$  equal to  $V_{cal}$ . As in Section VI.2, the capacitors  $C_{st}$  and  $C_{filt}$  serve to stabilize the loop and to filter out high-frequency signal components that might leak into the calibration circuitry. At steady state, the errors between the currents  $I_{D(M1S)}$ - $I_{D(M4S)}$  become very small due to the high loop gain.

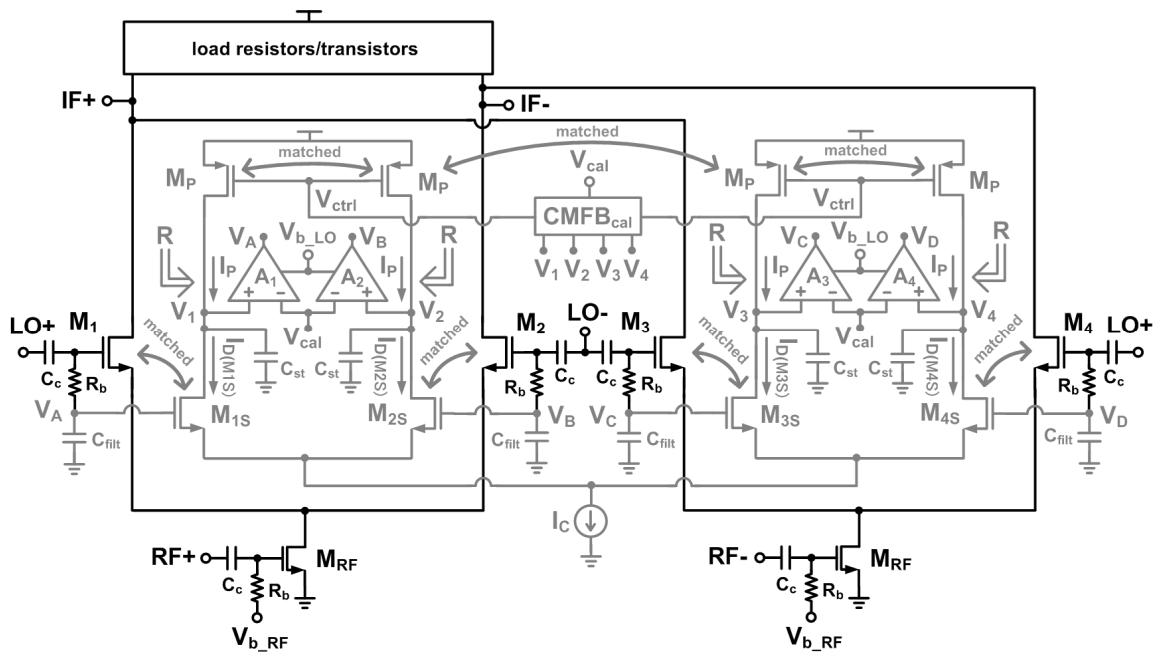


Fig. 86. Mixer with calibration loop components.

With mismatched transistors  $M_1$ - $M_4$  in Fig. 86, the different correlated currents  $I_{D(M1S)}$ - $I_{D(M4S)}$  of the sensing-transistors will be converted to distinct voltages  $V_1$ - $V_4$ . These voltages are compared to the common-mode voltage  $V_{cal}$  by amplifiers  $A_1$ - $A_4$  in each branch for further amplification and automatic adjustment of the individual bias voltages  $V_A$ - $V_D$  around the set bias  $V_{b\_LO}$  for the switching transistors. For example, if  $I_{D(M1S)}$  is relatively low compared to the other currents due to parameter mismatches, then  $V_1$  will be higher than  $V_{cal}$ . Consequently, the output voltage  $V_A$  of amplifier  $A_1$  will rise above  $V_{b\_LO}$ , and the increase of the gate bias voltage in this branch will increase  $I_{D(M1S)}$  until it is equal to the currents in the other branches.

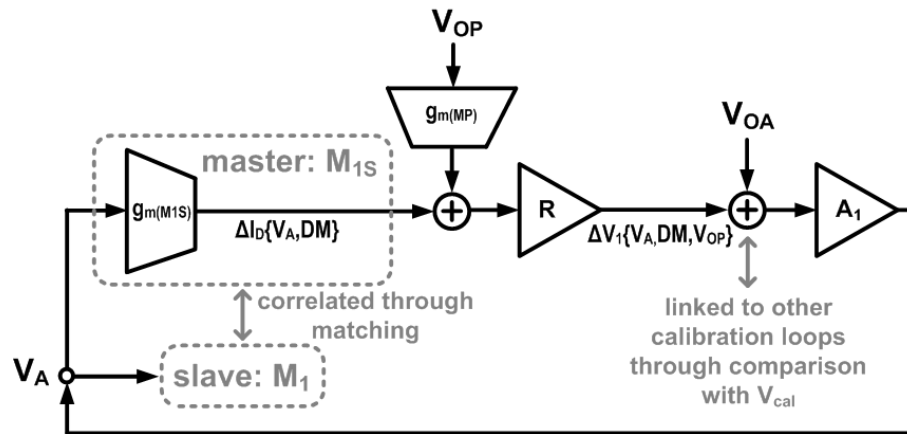


Fig. 87. DC signal flow diagram for one calibration loop with offsets.

An equivalent diagram for the DC calibration loop containing  $M_{1S}$  is portrayed in Fig. 87, which includes the offsets that affect the scheme's accuracy. It can be considered a master/slave configuration, in which  $M_{1S}$  is in the master loop and the shared gate bias voltage  $V_A$  is controlling the slave element  $M_1$ . The transconductors  $g_{m(M1S)}$  and  $g_{m(MP)}$  are representing the transconductance parameters of  $M_{1S}$  and  $M_P$  in Fig. 86.  $V_{OP}$  is the gate-referred offset voltage of  $M_P$ . The current  $\Delta I_D\{V_A, DM\}$  is the difference of the sensing transistor's drain-source current relative to the mean of the same current in all branches, which depends on  $V_A$  and the device mismatches ( $DM$ ) under correction. The block labeled “ $R$ ” in Fig. 87 represents the equivalent resistance looking into the node at which the drains of  $M_{1S}$  and  $M_P$  are connected together. At this node, the voltage  $\Delta V_I$  (the divergence of  $V_1$  from the mean of  $V_1 - V_4$ ) is a function of  $V_A$ ,  $V_{OP}$ , and  $DM$ . Furthermore, the input-referred offset voltage  $V_{OA}$  of the amplifier  $A_1$  adds at the same node. This node is significant because it links the calibration loop for  $M_1$  to



the other branches by comparison of  $V_1$ - $V_4$  with  $V_{cal}$  at the inputs of the amplifiers (Fig. 86).

As explained in Section VI.2, the intrinsic limit of the calibration loop's ability to reduce the standard deviation of the parameter mismatches between the slave transistors in the main circuit depends on their layout-dependent correlation to the mismatch-sensing transistors. For optimum effectiveness, the offsets associated with devices in the loop relative to their counterparts in the other branches must be minimized as well. From Fig. 87, two conditions can be identified by inspection:

$$V_{OP} \ll \frac{\Delta I_D \{V_A, DM\}}{g_{m(MP)}} , \quad (37)$$

$$V_{OA} \ll \Delta I_D \{V_A, DM\} \times R . \quad (38)$$

Since the offset voltages are inversely proportional to the device dimensions ([114]) of the current sources  $M_P$  and in amplifier  $A_1$ , the strategy to meet the criteria in equations (37) and (38) is to increase these dimensions until the simulated offsets are negligible. This is feasible because the parasitic capacitances from the large devices are not critical in this DC loop.

It is also insightful to assess the input-referred offset voltage of the calibration loop. Since  $V_A$  links the master and slave elements, it is preferred to maximize the sensitivity to  $\Delta I_D \{V_A, DM\}$  by minimizing the impact of offsets at that node. Referring to Fig. 87 again, it can be derived that the offset of  $V_A$  (from  $V_{b\_LO}$  in Fig. 86) is:

$$V_A = \frac{\Delta I_D \{V_A, DM\}}{g_{m(M1S)}} + \frac{g_{m(MP)} \cdot V_{OP} + V_{OA} / R}{g_{m(M1S)}} . \quad (39)$$

Apart from the need to minimize offset voltages  $V_{OP}$  and  $V_{OA}$ , the above expression reveals the importance of maximizing the gain in the first amplification stage by designing  $R$  to be large. This suggests the use of a small current  $I_C$  in combination with non-minimum transistor lengths for  $M_P$  to increase the resistance looking into the node at the drains of  $M_P$  and  $M_{IS}$  in Fig. 86.

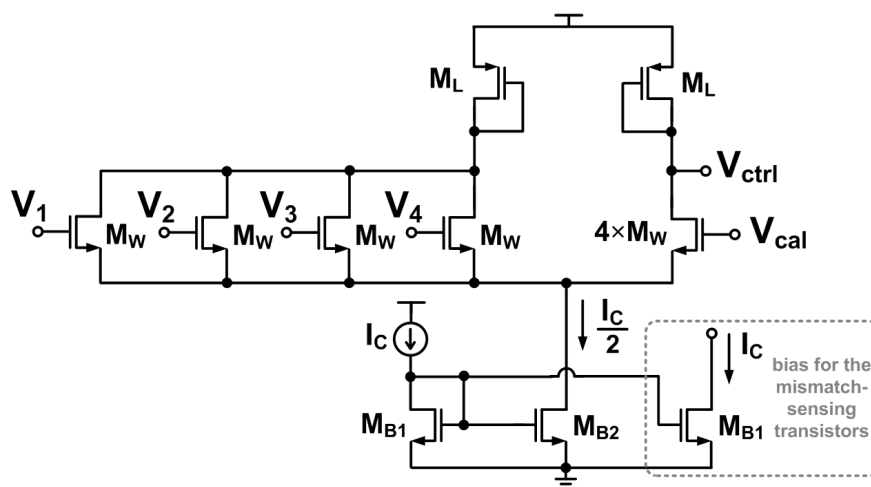


Fig. 88. Common-mode feedback circuit for the main calibration loop.

Since the nodes labeled  $V_1$ - $V_4$  in Fig. 86 are high-impedance nodes, common-mode control circuitry is necessary to ensure that the positive inputs of amplifiers  $A_1$ - $A_4$  are maintained close to the calibration reference  $V_{cal}$  at the negative inputs. Fig. 88 shows the schematic of the CMFB circuit in the calibration loop, which weighs voltages  $V_1$ - $V_4$  equally and compares their averaged value to the reference voltage  $V_{cal}$ . For convenience, the current mirror to bias the CMFB circuit also provides the current  $I_C$  that is routed to the sources of the mismatch-sensing transistors in the main loop. The stability of the CMFB loop is strongly related to the main calibration loop due to the

shared dominant pole at  $V_1$ - $V_4$ . Hence, a large value can be selected for  $C_{st}$  in Fig. 86 in order to stabilize both loops. A mixer calibration design example will be described with more details in the remainder of this section. The simulated gain and phase responses of the CMFB loop in this design are displayed in Fig. 89. It has a low-frequency gain of 14.4dB and a phase margin of 91.0°.

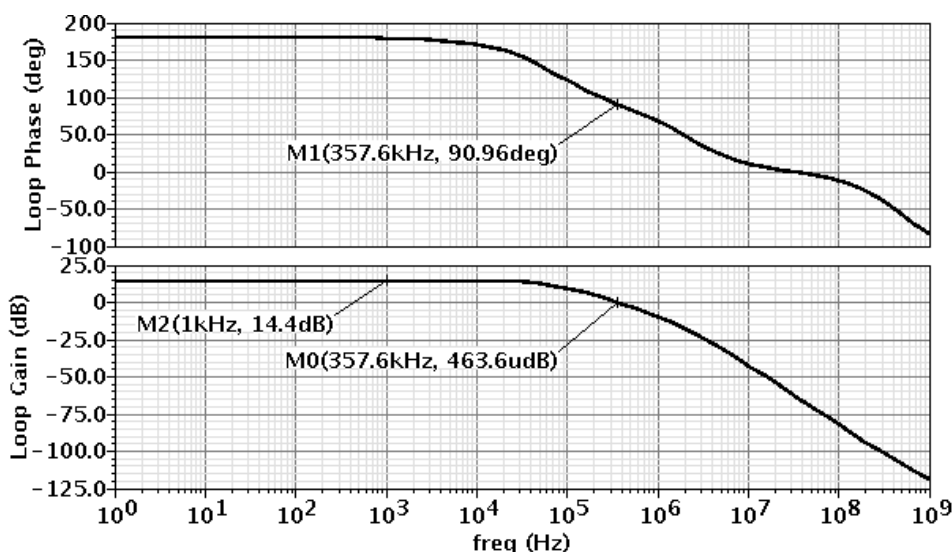


Fig. 89. Frequency response of the main CMFB circuit.

The schematic of the amplifiers  $A_1$ - $A_4$  is displayed in Fig. 90. It consists of a simple differential pair ( $M_A$ ) loaded by resistors ( $R_{CM}$ ) and controlled current sources ( $M_{CTR}$ ). The resistors serve as common-mode detectors for the CMFB amplifier ( $M_{CM1}$ ,  $M_{CM2}$ ) that is connected to the gates of  $M_{CTR}$  to regulate the output of the main amplifier. When the mismatches are sensed ( $In+ - In- \neq 0$ ), the voltage at the output terminal (Out) can move freely to counteract the sensed difference as part of the mismatch calibration loop, but the CMFB of the amplifier ensures that this change occurs around the required gate

bias voltage level  $V_{b\_LO}$  of the switching transistors in the mixer. Besides its role in the common-mode detection, the internal node  $N_{int}$  is not utilized as output. However, the same capacitor  $C_{filt}$  as present at the amplifier output (Fig. 86) is connected to  $N_{int}$  for loading symmetry.

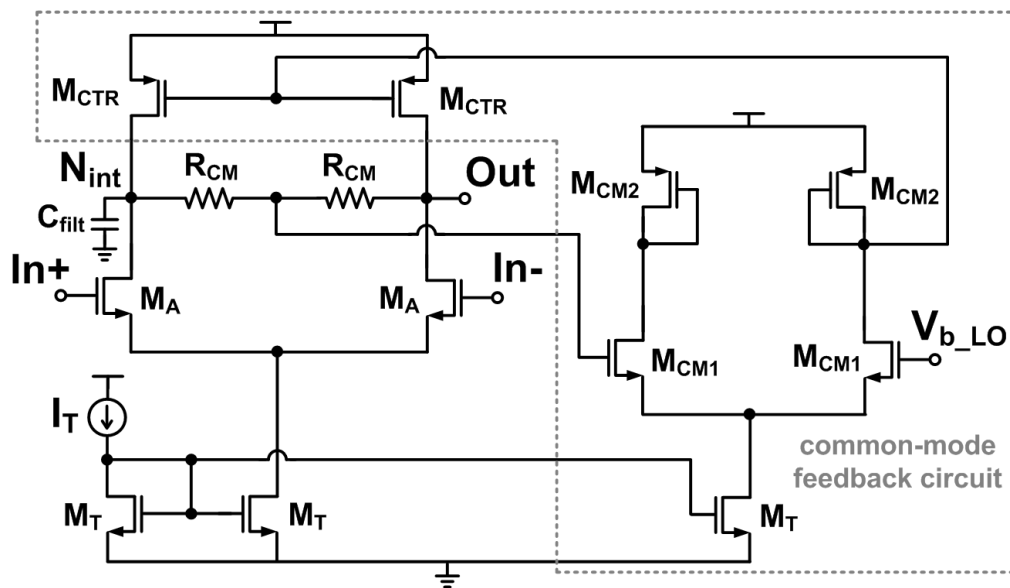


Fig. 90. Schematic of amplifiers  $A_1$ - $A_4$  in the calibration loop.

In the following design example, the amplifier in Fig. 90 was designed with a DC gain of 21.5dB from the differential input to the single-ended output. Apart from the stability considerations, its frequency response (Fig. 91) is not critical in the DC calibration loop for static mismatch reduction. Nonetheless, the bandwidth of the amplifier and overall calibration loop can be optimized when fast settling is desired for test time reduction.

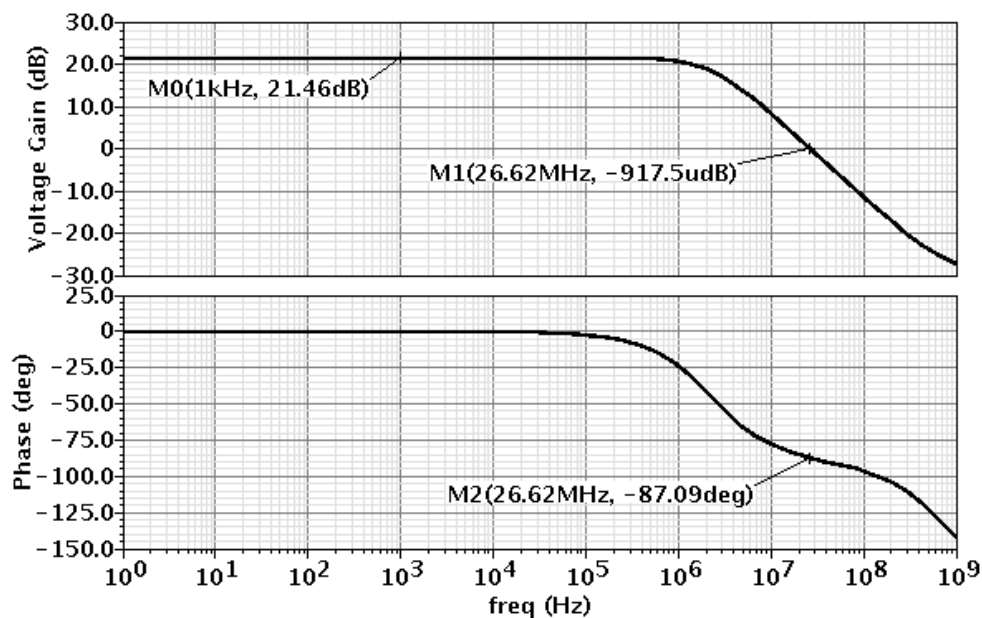


Fig. 91. Frequency response of the amplifiers in the calibration loop.

To demonstrate the calibration method, a double-balanced mixer was designed (see Section VI.3.3 for details) with the auxiliary circuitry described above. Table XV lists the component parameters of the design in TSMC 0.13 $\mu$ m CMOS technology using a 1.2V supply. Only the mismatch-sensing transistors  $M_{1S}$ - $M_{4S}$  have minimum transistor lengths. Their dimensions were selected identical to those of the switching transistors in the mixer under calibration, and they have the same number of fingers for improved parameter correlations according to equations (32) and (33). As explained previously, all other transistors in the mismatch calibration loop have non-minimum dimensions to decrease mismatches and offset voltages.

Table XV. Calibration circuitry components  
(0.13 $\mu\text{m}$  CMOS technology with 1.2V supply)

Component	Dimensions / Value
<b>Mismatch-sensing and first amplification stage (Fig. 86):</b>	
$M_{1S}, M_{2S}, M_{3S}, M_{4S}$	W/L = 2 $\mu\text{m}$ $\times$ 40 fingers / 0.13 $\mu\text{m}$
$M_P$	W/L = 6.9 $\mu\text{m}$ $\times$ 12 fingers / 5 $\mu\text{m}$
$C_{st}$	55pF
$C_{filt}$	0.5pF
$V_{cal}$	0.8V
$I_C$	50 $\mu\text{A}$
$V_{b\_LO}$	0.665V
$C_c$	1pF
$R_b$	100k $\Omega$ (L/W = 6 $\times$ 15.8 $\mu\text{m}$ / 1 $\mu\text{m}$ )
<b>Common-mode feedback circuit (Fig. 88):</b>	
$M_W$	W/L = 3 $\mu\text{m}$ $\times$ 4 fingers / 0.3 $\mu\text{m}$
$M_L$	W/L = 3.3 $\mu\text{m}$ $\times$ 2 fingers / 0.3 $\mu\text{m}$
$M_{B1}$	W/L = 2.5 $\mu\text{m}$ $\times$ 8 fingers / 0.5 $\mu\text{m}$
$M_{B2}$	W/L = 2.5 $\mu\text{m}$ $\times$ 4 fingers / 0.5 $\mu\text{m}$
<b>Amplifiers <math>A_1</math>-<math>A_4</math> (Fig. 90):</b>	
$M_A$	W/L = 6 $\mu\text{m}$ $\times$ 14 fingers / 4 $\mu\text{m}$
$M_{CTR}$	W/L = 5.2 $\mu\text{m}$ $\times$ 8 fingers / 3 $\mu\text{m}$
$M_{CM1}$	W/L = 1.8 $\mu\text{m}$ $\times$ 2 fingers / 0.3 $\mu\text{m}$
$M_{CM2}$	W/L = 2.8 $\mu\text{m}$ $\times$ 2 fingers / 0.3 $\mu\text{m}$
$M_T$	W/L = 2 $\mu\text{m}$ $\times$ 8 fingers / 1 $\mu\text{m}$
$I_T$	20 $\mu\text{A}$
$R_{CM}$	128k $\Omega$ (L/W = 20 $\times$ 6 $\mu\text{m}$ / 1 $\mu\text{m}$ )

With the design parameters in Table XV, the DC gain from the gate to the drain of each sensing transistor ( $M_{1S}$ - $M_{4S}$ ) is 20.5dB. Considering the 21.5dB amplifier gain ( $A_1$ - $A_4$ ), the total DC loop gain per branch is 42dB. When assessing the stability, it is important to keep in mind that the loops interact through the shared sources of  $M_{1S}$ - $M_{4S}$  and the common-mode feedback circuit (CMFB<sub>cal</sub>). Simulations were performed to

determine the appropriate capacitor values of  $C_{st}$  and  $C_{filt}$  for stability by inserting a probe at the gate of one mismatch-sensing transistor in Fig. 86 and plotting the loop's frequency response, which is also influenced by the  $CMFB_{cal}$  circuit. This assessment is to assure tolerance to any perturbation that could occur from high-frequency noise in one branch. The gain of the differential comparison involving the mismatch currents in each branch is very high, which is also evident from the evaluation of the mismatch current reduction that follows on page 189. Nevertheless, the response to an AC disturbance in an individual loop has a lower gain when only one of the voltage inputs ( $V_1$ - $V_4$ ) of the  $CMFB_{cal}$  block changes because the common-mode feedback action lowers the single-ended equivalent impedance seen at nodes  $V_1$ - $V_4$ . As shown in Fig. 92, this combined loop response for a single branch has an effective DC gain of 11.4dB and phase margin of  $47.7^\circ$  at the 3.8MHz unity gain frequency.

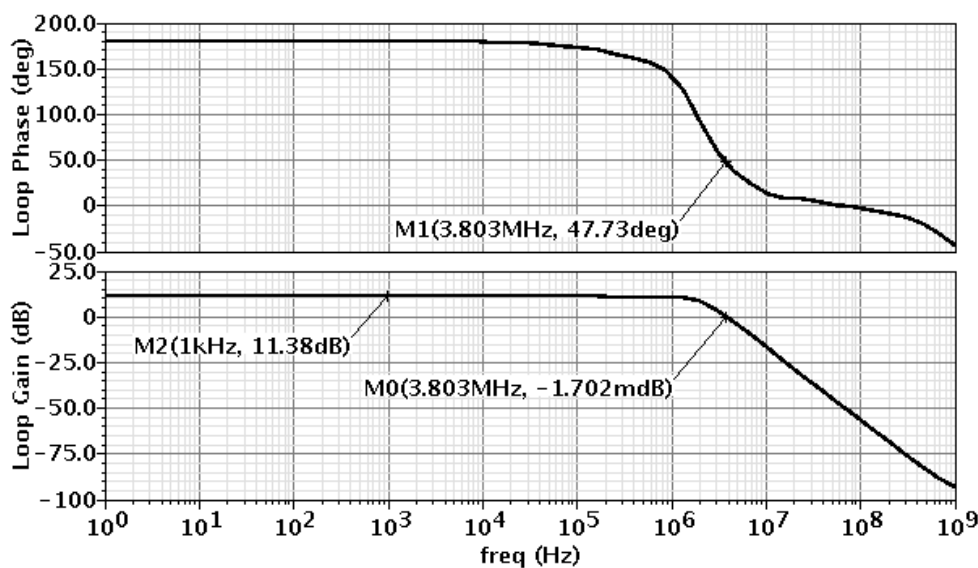


Fig. 92. Open-loop frequency response of the calibration circuit.  
(The simulation for a single branch was performed with the  $CMFB_{cal}$  block activated.)

### *VI.3.3. Double-balanced mixer design*

Since the transition frequency ( $f_T$ ) of devices in CMOS technologies continues to increase, several recent works have taken advantage of this trend by designing RF mixers with devices operating in the subthreshold region [133]-[136]. Even though the  $f_T$  of a device is much lower in the subthreshold (weak inversion) region than in the saturation (strong inversion) region, the technology improvements make up for  $f_T$  deficiencies that existed in the past. The primary benefit of designing mixers with devices in subthreshold region is that significant power savings can be achieved, as demonstrated in [133] with a 2.4GHz down-conversion mixer consuming only 0.5mW. Additionally, the LO signal can have a smaller swing for hard-switching of the transistors with reduced gate-source overdrive voltage, which translates into more power savings in the LO signal generation circuitry. With less DC currents in the mixer branches, subthreshold designs also have the tendency to allow for more voltage headroom. Thus, the possibility exists to use larger load resistors in order to increase the conversion gain. On the contrary, the main trade-offs are reduced linearity, higher device noise levels, and increased die area to obtain comparable transconductance values. Furthermore, subthreshold designs are generally more susceptible to PVT variations. For example, the results in [11] and [137] show how the percent mismatch of the drain-source current for MOS transistors increases drastically as the gate-source voltage is decreased.

Although the IIP2 calibration technique presented in the previous subsection can be applied to any double-balanced mixer, it is demonstrated here for a subthreshold mixer



in order to simultaneously explore this promising design methodology further. Fig. 93 shows the mixer schematic from before with more details. The approach taken here is to optimize the subthreshold mixer for linearity and noise performance that approximates state of the art mixers in saturation region as much as possible for typical conversion gain. This requires transistors with high W/L ratios to obtain the appropriate transconductances in subthreshold region. However, the use of large devices increases the total parasitic capacitances at the drains of the LO transistors ( $M_{SW}$ ), causing IIP2 degradation. As explained in [118], the inductors ( $L_S$ ) resonate with these parasitic capacitances to improve the IIP2 performance. In addition, the mismatch reduction method for the LO transistors is utilized for further IIP2 enhancement. While the LO transistor bias voltages  $V_A$ - $V_D$  are generated with the previously described loop, the RF input transconductors are biased with a simple current mirror to produce the DC current  $I_{DC}$  on each side of the mixer. If the transconductance mismatch of the  $M_{RF}$  transistors becomes detrimental, then the same mismatch reduction loop as for the LO transistors can be employed to generate the RF bias voltages individually. However, IIP2 is typically more sensitive to LO transistor mismatches as described in Section VI.3.1. To achieve sufficient transconductance in this subthreshold mixer design, the RF input transistors  $M_{RF}$  are five times larger than the LO transistors  $M_1$ - $M_4$ , which makes it even less important to calibrate the  $M_{RF}$  transistors.





the switching transistors  $M_1$ - $M_4$  are exactly the same as the sensing transistors  $M_{1S}$ - $M_{4S}$  in the mismatch reduction loop.

Table XVI. Subthreshold mixer components  
(0.13 $\mu$ m CMOS technology with 1.2V supply)

Component	Dimensions / Value
<b>Main mixer components (Fig. 93):</b>	
$M_1, M_2, M_3, M_4$	W/L = 2 $\mu$ m $\times$ 40 fingers / 0.13 $\mu$ m
$M_{RF}$	W/L = 10 $\mu$ m $\times$ 40 fingers / 0.13 $\mu$ m
$M_{ctrl}$	W/L = 1.2 $\mu$ m $\times$ 26 fingers / 0.25 $\mu$ m
$R_L$	3k $\Omega$ (L/W = 10 $\times$ 8.87 $\mu$ m / 8 $\mu$ m)
$C_L$	0.15pF
$L_S$	7nH
$C_c$	1pF
$R_b$	100k $\Omega$ (L/W = 6 $\times$ 15.8 $\mu$ m / 1 $\mu$ m)
$V_{b\_LO}$ (nominal values of $V_A, V_B, V_C, V_D$ )	0.665V
$V_{refL}$	0.565V
$I_{DC}$	200 $\mu$ A
<b>Common-mode feedback amplifier <math>A_{CM}</math> (Fig. 94):</b>	
$M_{CP}$	W/L = 1.5 $\mu$ m $\times$ 4 fingers / 0.13 $\mu$ m
$M_{LCM}$	W/L = 1.5 $\mu$ m $\times$ 4 fingers / 0.13 $\mu$ m
$M_{BI}$	W/L = 2.5 $\mu$ m $\times$ 8 fingers / 0.5 $\mu$ m
$M_{BT}$	W/L = 2.5 $\mu$ m $\times$ 18 fingers / 0.5 $\mu$ m
$R_{LCM}$	3.9k $\Omega$ (L/W = 6 $\times$ 4.5 $\mu$ m / 2 $\mu$ m)
$I_{BT} / I_C$	110 $\mu$ A / 50 $\mu$ A
<b>Mismatch reduction loop (Fig. 86):</b>	
$M_{1S}, M_{2S}, M_{3S}, M_{4S}$ , comparison circuitry	listed in Table XV

### VI.3.4. Simulation results

#### Characterization of the subthreshold mixer design

Unless noted otherwise, the simulation results for the subthreshold mixer design described in Section VI.3.3 were obtained with a 1.988GHz sinusoidal LO signal having a power of -1dBm. As seen in Fig. 96, this mixer has a conversion gain of  $11.5\text{dB} \pm 0.5\text{dB}$  for RF input signals located up to 125MHz away from the LO frequency. It has been demonstrated that designing active mixers in the subthreshold region allows high gain (e.g., 32dB in [136]) with low power consumption from the use of small bias currents, which also leaves voltage headroom for large load resistors. However, the mixer in this dissertation was optimized to achieve high linearity for broadband applications. This required a conversion gain trade-off that resulted in 11.5dB gain, which is comparable to conventional double-balanced active mixers designed in the saturation region.

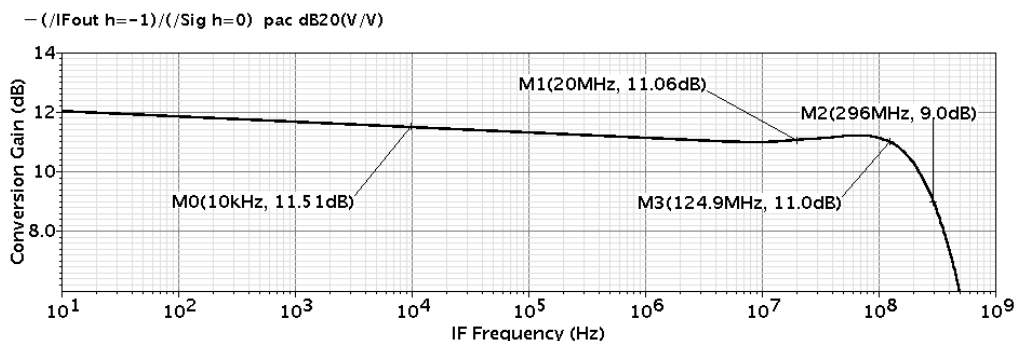


Fig. 96. Conversion gain vs. frequency.

Fig. 97 shows that a reasonable noise figure (NF) can be attained in the subthreshold region by using large RF input transistors to ensure that they have

sufficient transconductance. In this case, the single-sideband (SSB) NF is 16.2dB with a flicker noise corner at 266KHz. The corresponding double-sideband (DSB) NF is normally 3dB lower than the SSB NF [122].

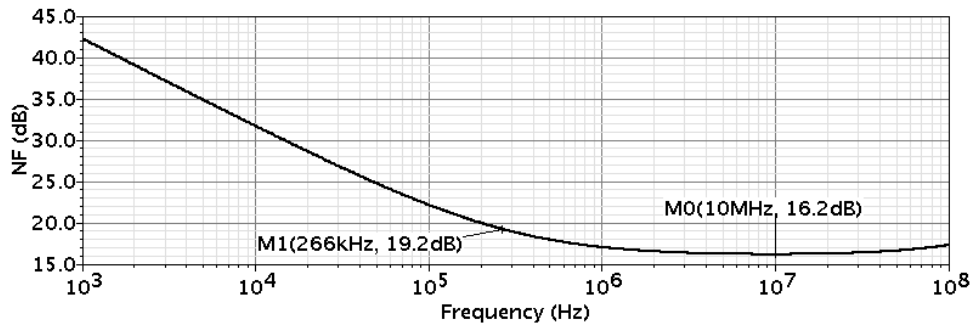


Fig. 97. SSB noise figure vs. frequency.

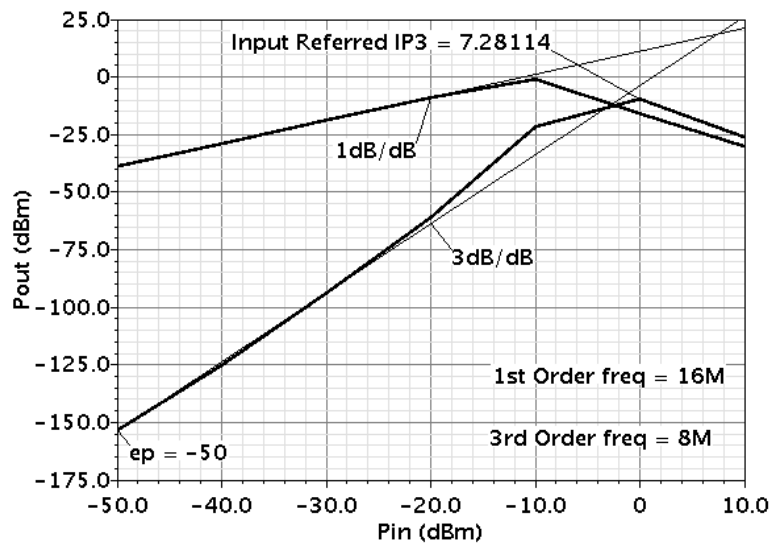


Fig. 98. IIP3 curve.

LO frequency: 1.988GHz, RF test tones: 2GHz, 2.004GHz, IM3 frequency: 8MHz.

Linearity characteristics were assessed within a 20MHz band under consideration that the mixer is intended for broadband wireless target application such as WiMAX.

The simulated IIP3 of 7.3dBm in Fig. 98 was obtained with two tones located at 2GHz and 2.004GHz (12MHz and 16MHz away from the 1.988GHz LO frequency). Fig. 99 shows that the mixer has a simulated 1-dB compression point of -7.7dBm, which was determined by sweeping the power of a single 2GHz RF input tone.

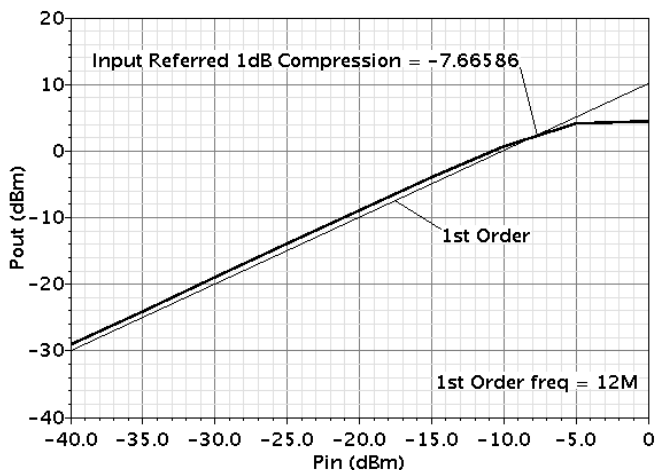


Fig. 99. 1-dB compression curve.

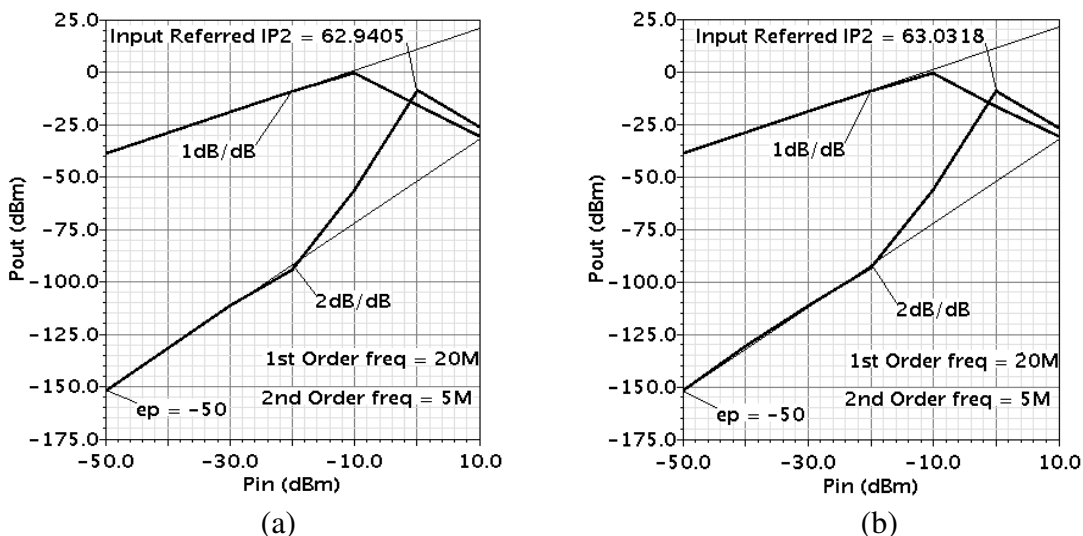


Fig. 100. IIP2 curve with 0.5% mismatch between the load resistors ( $R_L$ ).  
LO frequency: 1.985GHz, RF test tones: 2GHz, 2.005GHz, IM2 frequency: 5MHz;  
(a) without calibration circuitry, (b) with calibration circuitry.

To give first insights into the IIP2 characteristics, the simulated IIP2 curves with 0.5% load resistor mismatches are plotted in Fig. 100 for the mixer without and with calibration circuitry. This assessment condition was selected because the load mismatch leads to common-mode to differential-mode conversion of the IM2 components according to equation (35). Without any other mismatches in the circuits, the results in Fig. 100 reveal that the calibration circuitry has negligible impact. IIP2 characterizations with Monte Carlo simulations using statistical device models provided by the foundry are discussed later in this section to present an estimate for the IIP2 improvement from the calibration circuit in the presence of realistic device mismatches in the mixer and calibration circuit itself.

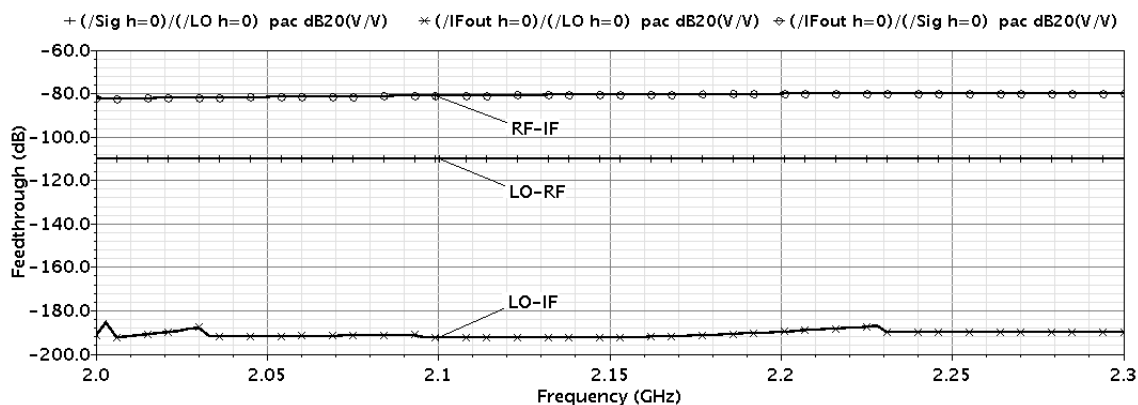


Fig. 101. Feedthrough between mixer ports.

Fig. 101 displays the simulated port-port feedthroughs, showing that the port-port isolation is 80dB or more. This isolation is credited to the fact that minimum lengths are used for the LO switching transistors and RF input transistors, which is particularly



important to minimize the parasitic capacitances when designing in the subthreshold region with high W/L ratios. As for conventional mixers, the measured isolation will be strongly affected by substrate leakage and layout parasitics, as well as package and PCB design choices. As explained in Section VI.2, one of the motivations behind the use of the DC calibration loop with low-pass filter nodes is to avoid RF coupling and substrate leakage due to the proximity of transistors in typical layout matching techniques.

Fig. 102 shows the transient signals from a simulation of the mixer with a -30dBm differential RF input signal at 2.005GHz and a -1dBm differential LO at 1.985GHz. As expected, the differential IF output signal (IF+ – IF-) has a frequency of 20MHz and an amplitude of 38.8mV, indicating a conversion gain of 11.8dB relative to the 10mV RF input amplitude.

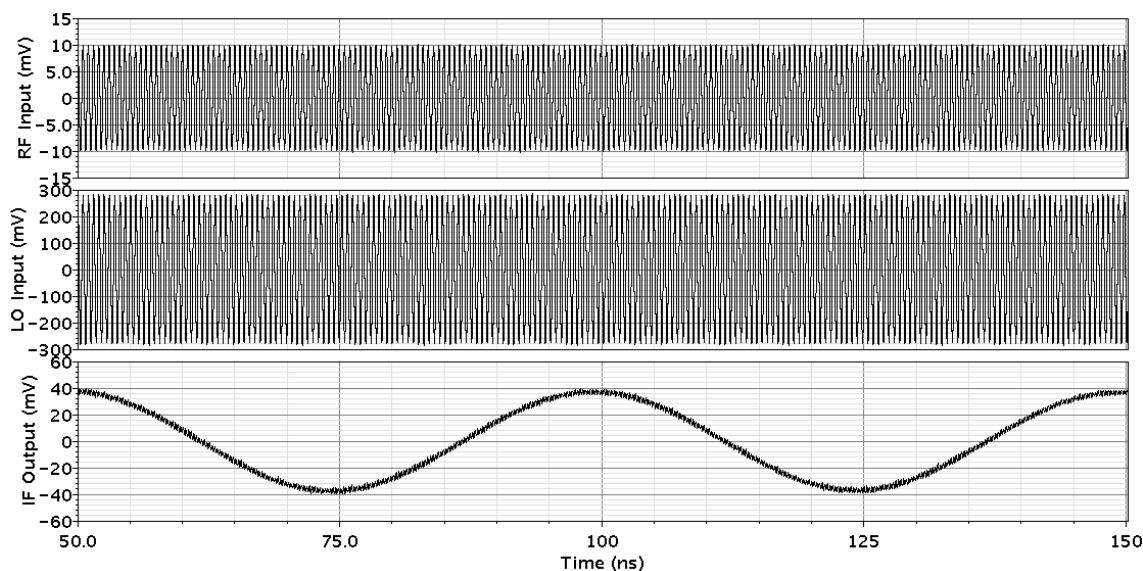


Fig. 102. Transient simulation with a 20MHz IF output signal.  
(LO frequency: 1.985GHz, RF input signal: -30dBm at 2.005GHz.)

Since the mixer is designed in the subthreshold region instead of the saturation region, a smaller LO amplitude is needed to induce hard-switching of the LO transistors due to the reduced gate-source overdrive voltage. The progression of the simulated gain, NF, IIP2, and IIP3 for a sweep of the LO signal power can be observed in Fig. 103 – Fig. 105. Based on the specification trade-offs in these plots, the LO power of -1dBm was selected for this subthreshold mixer design.

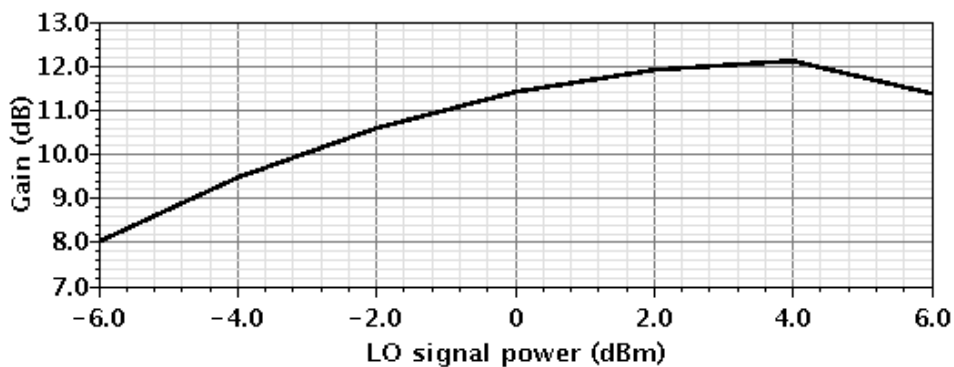


Fig. 103. Conversion gain vs. LO signal power.  
(frequencies: LO = 1.985GHz, RF 2.005GHz, IF = 20MHz.)

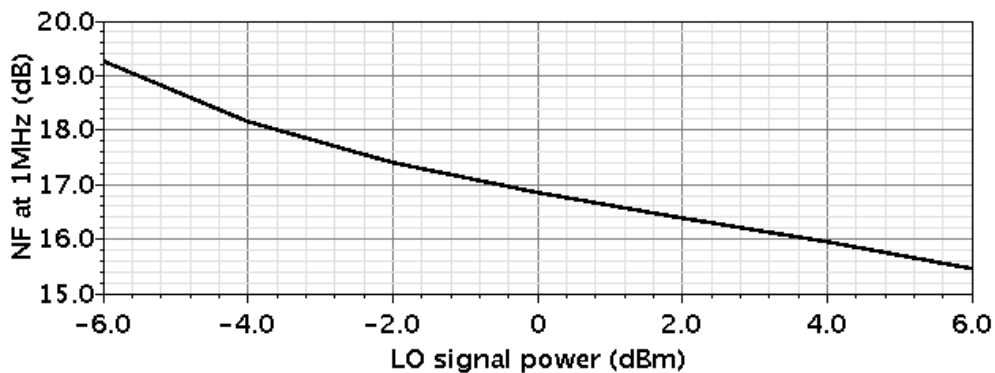


Fig. 104. SSB Noise figure at IF = 1MHz vs. LO signal power.

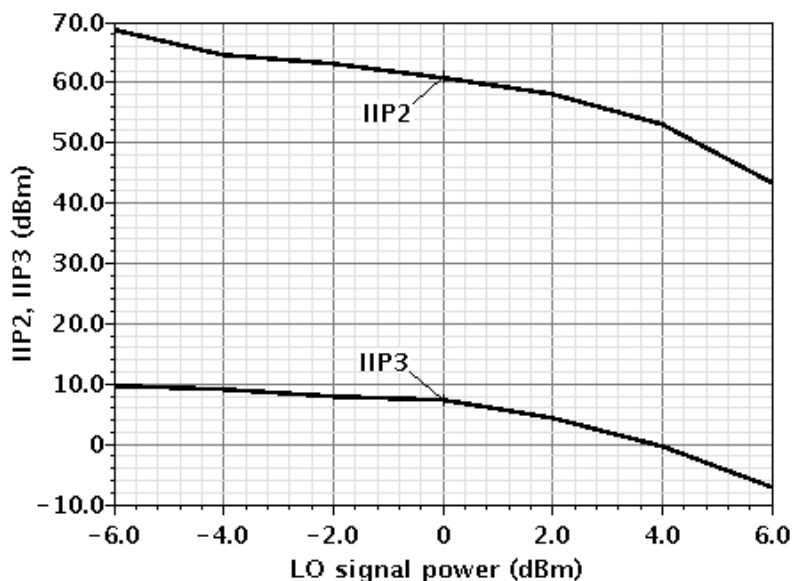


Fig. 105. IIP2 (with 0.5%  $R_L$  mismatch) and IIP3 vs. LO signal power.  
(frequencies: LO = 1.985GHz, RF = 2GHz/2.005GHz, IF = 15MHz/20MHz,  
IM2 = 5MHz, IM3 = 10MHz.)

A summary of the subthreshold mixer performance specifications is provided in Table XVII to compare the simulation results before and after adding the calibration circuitry. The outcomes show that none of the mixer specifications is affected significantly by the DC calibration loops outside of the signal path. A notable difference is the minimum IIP2 observed after Monte Carlo simulations, which will be discussed in the remainder of this section. In general, the impact of the mixer's auxiliary calibration circuits is limited to its ability to compensate for device variations and mismatches as discussed in sections VI.1-VI.2. However, the drawbacks are the increase of the total power consumption from 0.68mW to 0.97mW as well as the die area required for the calibration circuitry.

Table XVII. Simulated mixer specifications with and without calibration  
(0.13 $\mu$ m CMOS technology with 1.2V supply)

	Without Calibration Circuitry	With Calibration Circuitry
<b>RF Frequency</b>	2GHz	2GHz
<b>IF Bandwidth</b>	< 124.9MHz	< 124.3MHz
<b>Conversion Gain</b>	11.5dB	11.5dB
<b>IIP3</b>	7.3dBm	7.3dBm
<b>1-dB Compression Point</b>	-7.7dBm	-7.8dBm
<b>IIP2 (With 0.5% <math>R_L</math> Mismatch)</b>	62.9dBm	63.0dBm
<b>Avg. IIP2* (100 Monte Carlo runs)</b>	58.9dBm	64.2dBm
<b>Yield** (for IIP2 &gt; 54dBm)</b>	75%	91%
<b>DSB Noise Figure</b>	13.2dB	13.2dB
<b>Flicker Noise Corner</b>	266KHz	274KHz
<b>LO-RF Isolation (2-2.3GHz)</b>	> 110dB	> 110dB
<b>LO-IF Isolation (2-2.3GHz)</b>	> 185dB	> 182dB
<b>RF-IF Isolation (2-2.3GHz)</b>	> 80dB	> 79dB
<b>Power (with auxiliary circuits)</b>	0.68mW	0.97mW

\* With foundry-supplied statistical models (process & mismatch) for all devices in the mixer and calibration circuits.

\*\* Defined as the percentage of the Monte Carlo simulation outcomes that meet the IIP2 target.

### IIP2 evaluation before and after the addition of the calibration circuitry

The IIP2 performance was investigated with statistical Monte Carlo simulations using device models provided by the foundry to account for process and mismatch variability. All active and passive devices in the mixer and calibration circuit were simulated with these statistical models, and correlations between matched devices were defined based on equations (32) and (33) as described in sections VI.2.1 and VI.2.2. In the mixer, correlations based on the number of fingers or resistor segments were set only for the load devices  $R_L$  and  $M_{ctrl}$  in Fig. 93 as well as the devices with identical names in the CMFB circuit in Fig. 94. This was done under the assumptions that these will be laid out with matching techniques. On the contrary, correlations were not specified for the

devices that process RF signals ( $M_1$ - $M_4$  and  $M_{RF}$ ), so that these can be placed as individual devices to minimize substrate leakage due to placement proximity and crosstalk via routing parasitics. Since parasitic capacitances in the low-frequency calibration circuits are not critical, they can be laid out with matching techniques. Hence, correlations were defined based on the number of fingers or resistor segments for  $M_{1S}$ - $M_{4S}$  and  $M_P$  in Fig. 86 as well as for the transistors and resistors with equal labels in the  $CMFB_{cal}$  (Fig. 88) and amplifier circuits (Fig. 90).

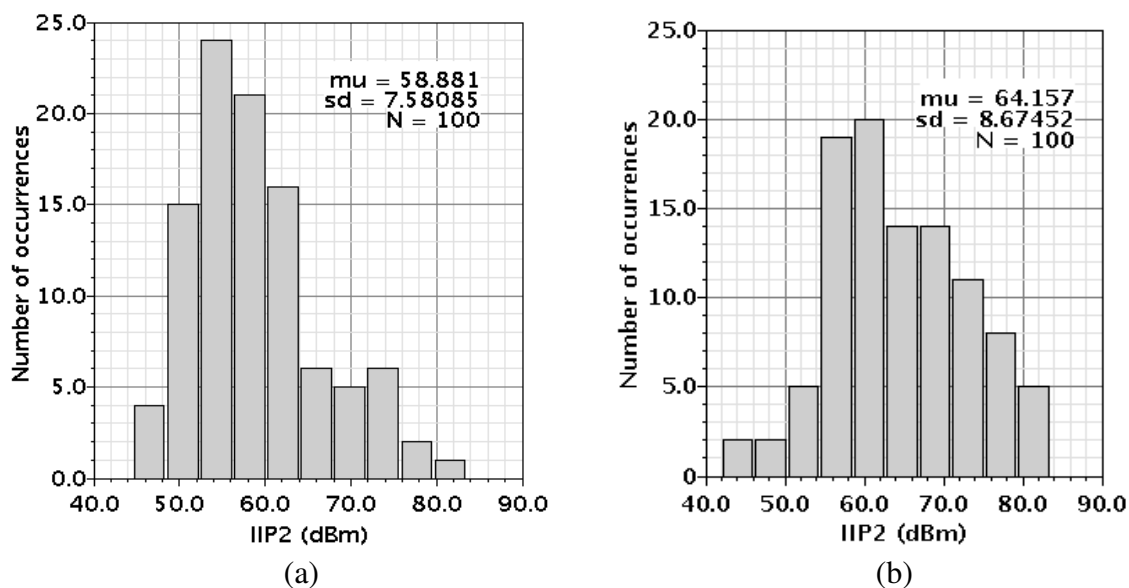


Fig. 106. IIP2 comparison with 100 Monte Carlo runs.  
LO frequency: 1.985GHz, RF test tones: 2GHz, 2.005GHz, IM2 frequency: 5MHz;  
(a) without calibration circuitry, (b) with calibration circuitry.

Fig. 106 displays the histograms of the IIP2 from Monte Carlo simulations (process and mismatch variations enabled) with 100 runs before and after the addition of the calibration circuitry. Without calibration, the IIP2 mean is 58.9dBm (with 7.6dbm

standard deviation), which improved to 64.2dBm (with 8.7dBm standard deviation) due to the calibration. With a target IIP2 of 54dBm for example, this would correspond to a yield increase from 75% to 91% as a result of the calibration.

### Mismatch reduction with the calibration loops

The mismatch in the mixer core can be assessed by purposely introducing offset voltages at the gates of the LO transistors to emulate threshold voltage mismatches as visualized in Fig. 107. In this test setup, a positive DC offset voltage source ( $\Delta V_{Th}$ ) was inserted at the gates of  $M_2$  and to its corresponding matched sensing-transistor  $M_{2S}$ , while the same offset voltage with negative polarity was included at the gates of  $M_4$  and  $M_{4S}$ . The ultimate mismatch indicator is the difference of the LO transistor DC drain currents  $I_{D1}-I_{D4}$ . Here, this average mismatch current is defined using  $I_{D1}$  as reference:

$$\overline{\Delta I_D} = \text{mean} \{ |I_{D2} - I_{D1}| + |I_{D3} - I_{D1}| + |I_{D4} - I_{D1}| \}. \quad (40)$$

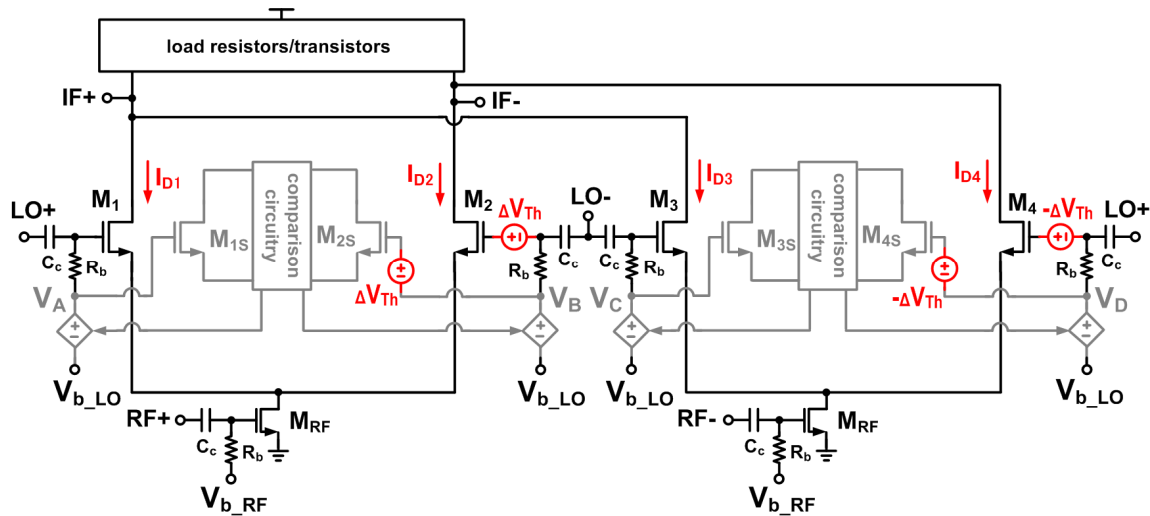


Fig. 107. Mixer with intentional threshold voltage offsets ( $\Delta V_{Th}$ ).

A comparison of the average mismatch current  $\overline{\Delta I_D}$  with and without calibration circuitry is plotted in Fig. 108 for a sweep of  $\Delta V_{Th}$  from 1mV to 30mV, showing a mismatch current reduction by more than two orders of magnitude. This property of the calibrated mixer is the fundamental mechanism behind the IIP2 improvement observed in the Monte Carlo simulation results.

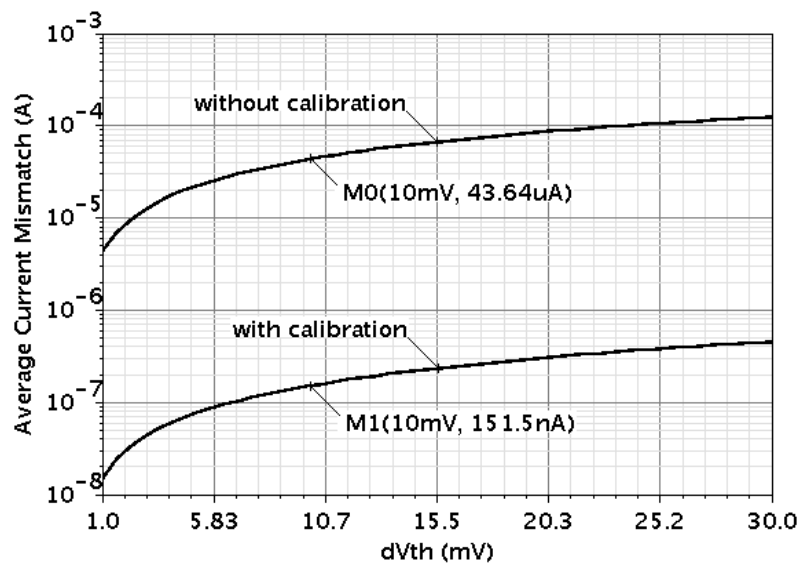


Fig. 108.  $\overline{\Delta I_D}$  (average mismatch of  $I_{D1}-I_{D4}$ ) vs.  $\Delta V_{Th}$ .

### Transient behavior of the calibration loops

Fig. 109 shows the settling of voltages  $V_A-V_D$  and  $V_{ctrl}$  (Fig. 86) from a transient simulation with 1.985GHz LO frequency and a -30dBm RF input signal at 2.005GHz. In this simulation, the offset voltages at the gates of ( $M_2$ ,  $M_3$ ,  $M_4$ ) changed from 0V to (30mV, -15mV, -30mV) at time = 0s. Fig. 110 displays the corresponding transient waveform of the down-converted 20MHz signal at the IF output after settling of the

control voltages. The short settling times below  $4\mu\text{s}$  of the control voltages in this background calibration scheme make it suitable for quick calibrations at system start-up as well as for in built-in test routines during manufacturing testing.

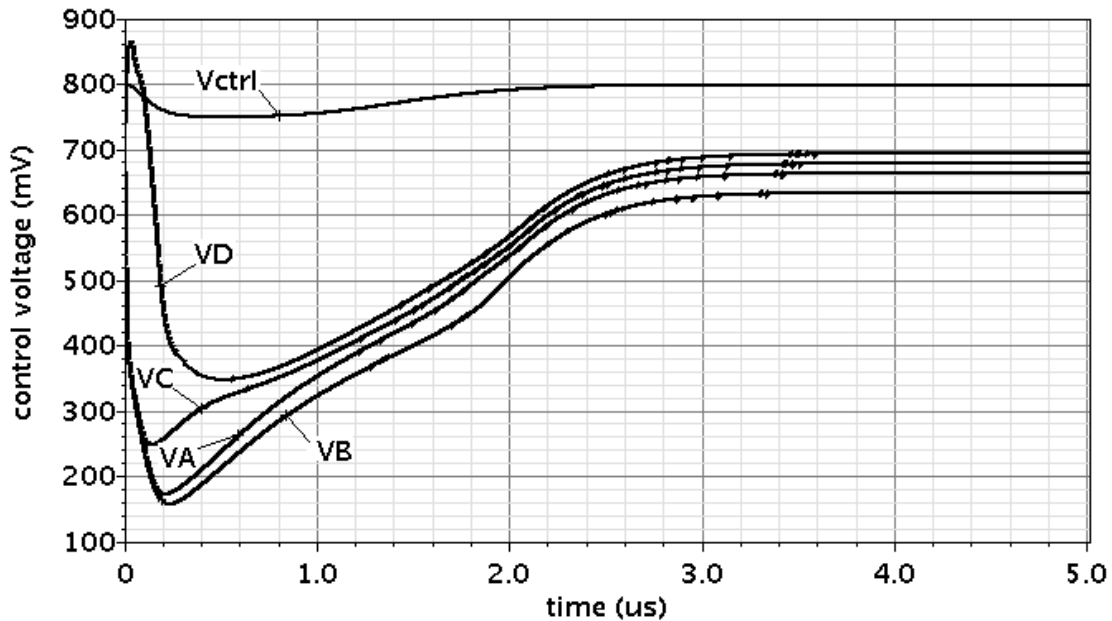


Fig. 109. Transient settling behavior of critical control voltages.

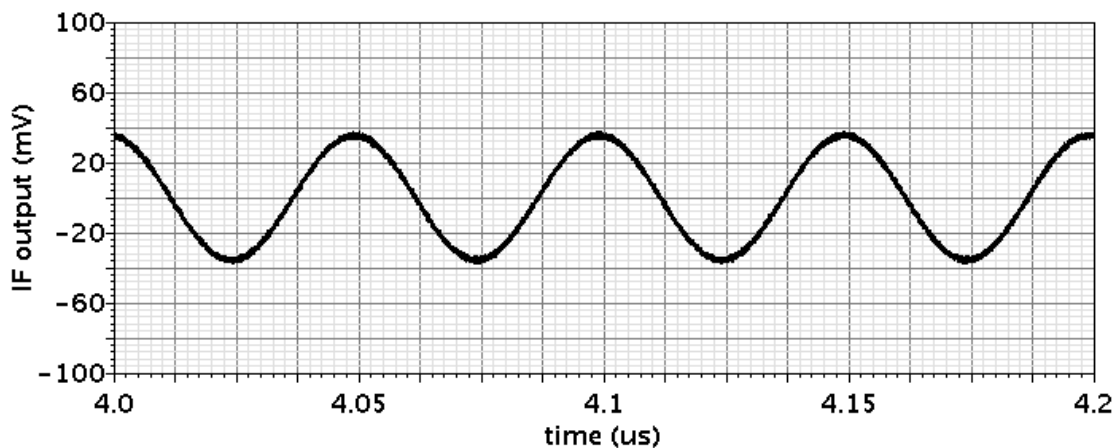


Fig. 110. Transient IF output after settling of the calibration control voltages.  
(LO frequency: 1.985GHz, RF input signal: -30dBm at 2.005GHz.)



### Variations of other mixer parameters

Monte Carlo simulations with statistical device models and the aforementioned correlation definitions were also performed to determine the calibration circuitry's impact on other key mixer specifications. Fig. 111 – Fig. 113 show the histograms of the conversion gain, IIP3, and noise figure after Monte Carlo simulations with 100 runs. By comparing the results, it can be seen that the calibration circuitry has little impact on the mean values and standard deviations of these specifications. However, activation of the calibration slightly increases the IIP3 and its standard deviations by 1.6dBm and 2.5dBm, respectively.

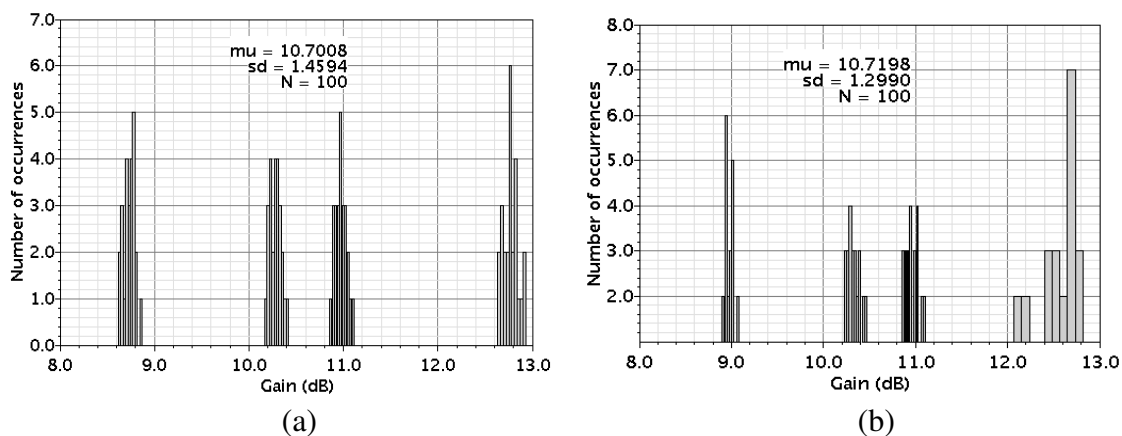


Fig. 111. Conversion gain comparison with 100 Monte Carlo runs.  
 LO frequency: 1.98GHz, RF test tone: 2GHz, IF frequency: 20MHz;  
 (a) without calibration circuitry, (b) with calibration circuitry.

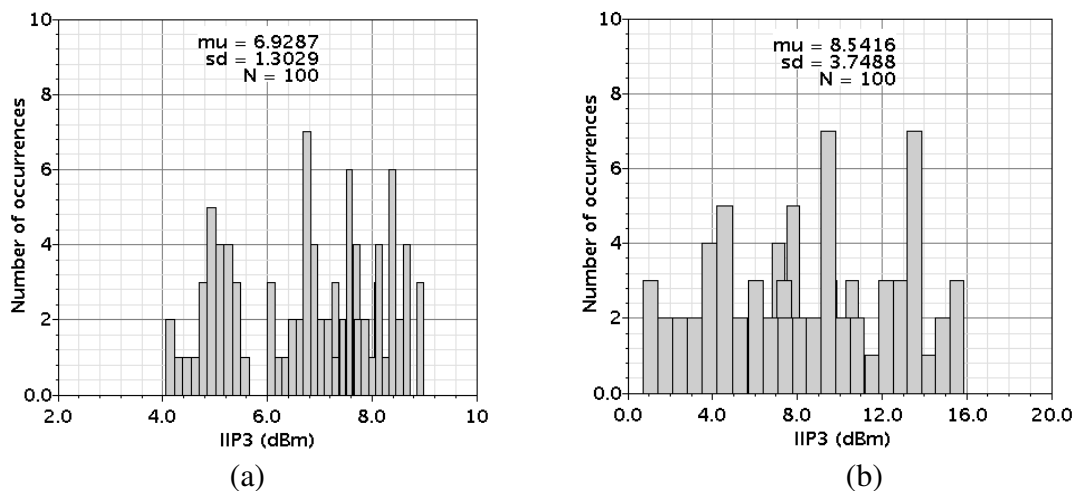


Fig. 112. IIP3 comparison with 100 Monte Carlo runs.  
 LO frequency: 1.98GHz, RF test tones: 2.01GHz, 2.02GHz, IM3 frequency: 20MHz;  
 (a) without calibration circuitry, (b) with calibration circuitry.

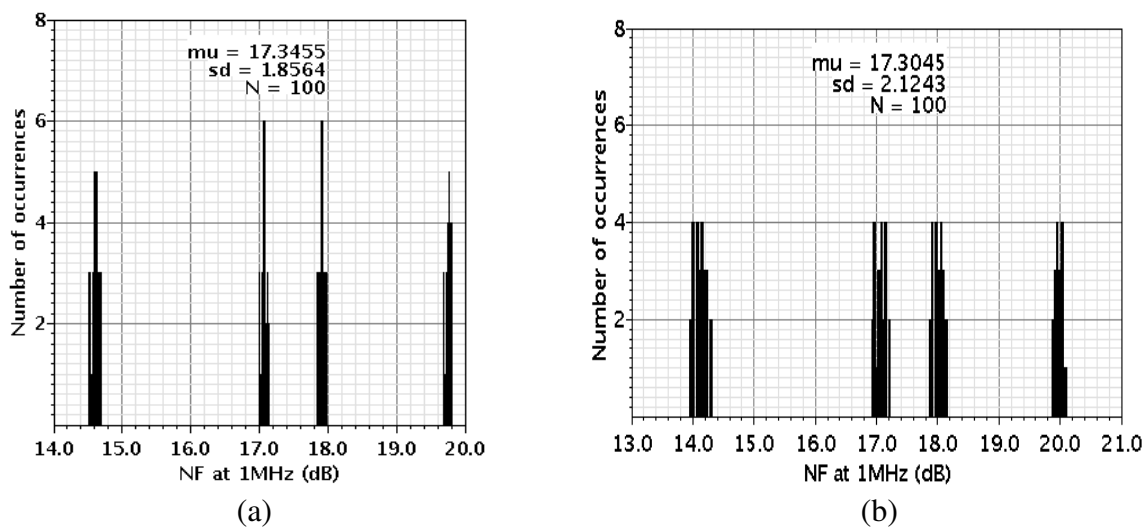


Fig. 113. Comparison of the SSB NF at 1MHz with 100 Monte Carlo runs.  
 The shown cases are: (a) without calibration circuitry, (b) with calibration circuitry.

Assessment with respect to the state of the art

Table XVIII. Down-conversion mixer performance comparison

	1	2	3	4	5	6	7	8	9	10	11
<b>Reference</b>	[110] <sup>Δ*</sup>	[118] <sup>Δ†</sup>	[120] <sup>†</sup>	[121] <sup>Δ†</sup>	[130] <sup>Δ†</sup>	[131] <sup>Δ*</sup>	[133] <sup>#†</sup>	[134] <sup>#†</sup>	[135] <sup>#†</sup>	[136] <sup>#†</sup>	<b>This Work<sup>#Δ*</sup></b>
<b>CMOS Technology</b>	0.18μm	0.18μm	90nm	0.35μm	0.13μm	65nm	0.13μm	0.18μm	0.13μm	0.18μm	0.13μm
<b>RF Freq. (GHz)</b>	3.5	2.1	2.1	0.815	2	2.1	2.4	2.4	3.1 – 10.6	2.4	2
<b>IF Freq. (MHz)</b>	-	< 4.5	< 1.2	< 10	< 1.5	< 10	60	10	264	30	< 124
<b>Conversion Gain (dB)</b>	10	16	9	14.5	53	8	15.7	9	9.8 – 14.0	32	11.5
<b>Noise Meas. or DSB NF (dB)</b>	$4.5 \frac{nV}{\sqrt{Hz}}$ §	$4 \frac{nV}{\sqrt{Hz}}$ §	9.4	12	$3.5 \frac{nV}{\sqrt{Hz}}$ §	16	18.3	11.8	14.5 – 19.6	8.5	13.2
<b>IIP3 (dBm)</b>	8	9	8.9	2.4	12	12	-9	-	-11	-14.5	7.3
<b>1-dB Comp. Point (dBm)</b>	-	-	-	-	4	-	-28	-	-24 – -19	-	-7.8
<b>IIP2 (dBm)</b>	> 65	> 78	> 55.1	> 66	~85	> 75	-	-	-	-	> 54 <sup>x</sup>
<b>Supply (V)</b>	1.8	1.8	1	2.7	1.5	1	1	1.2	1.2	1.8	1.2
<b>Power (mW)</b>	-	7.2	6.25 <sup>θ</sup>	10.8	72	8.5	0.5	0.18	1.85	1	0.97

\* Simulation results. † Measurement results. # Subthreshold design. Δ With IIP2 enhancement circuitry.  
<sup>θ</sup> Reported with LO buffer. § Reported as input-referred noise. <sup>x</sup> With 91% yield.

Table XVIII contains summaries of specifications reported for CMOS down-conversion mixers with similar operating frequencies. The presented subthreshold mixer

in the last column has lower IIP2 than the mixers in columns 1-6 that are designed with transistors biased in saturation region. However, when the IIP2 target is 50dBm as in [120], the IIP2 improvement from the calibration makes it possible to achieve such a target with this subthreshold design. Apart from mixer design optimizations for scenarios with higher IIP2 requirement, it can be explored to make the load resistors of the mixer programmable for further IIP2 tuning through digital trimming as proposed in [123]. Most of the mixers in columns 1-6 of Table XVIII contain auxiliary circuitry for IIP2 enhancements. Notice that they exhibit overall comparable performances but consume at least six times as much power as the proposed subthreshold mixer with calibration. On the other hand, the subthreshold mixer designs in columns 7-10 have similar performances and power consumptions compared to this work, but with the tendency that they have lower IIP3 and 1dB compression point specifications; whereas IIP2 characterization results were not reported for these designs. In general, the presented subthreshold mixer with calibration has competitive performance relative to saturation region mixers, but with significantly lower power dissipation in the same range as other reported subthreshold mixers. The simulation results suggest that the proposed calibration loop effectively improves the second-order linearity and makes the subthreshold design more robust to mismatch variations.

#### **VI.4. Summarizing Remarks**

Alternatively to matching transistors within the RF signal path or increasing their dimensions, a methodology has been proposed to reduce the mismatch between a pair of transistors by indirectly matching them through a DC calibration loop. Monte Carlo

simulation results demonstrated that the input offset standard deviation of the differential amplifier under investigation is expected to reduce from 4.17mV to 1.29mV or 0.76mV, which depends on the layout-based quality of the matching between the RF and mismatch-sensing transistors. The trade-offs with the scheme are an approximately 15% power increase and the die area overhead for the calibration circuitry.

Applied to an example mixer design, it was shown that the proposed calibration scheme improves the IIP2 specification. Monte Carlo simulations revealed that the mean of the IIP2 increased from 58.9dBm to 64.2dBm. While the background calibration loops did not noticeably impact other mixer specifications, the main trade-off was a 30% increase in the power consumption. If the mixer under calibration is designed with saturation region bias conditions using higher currents, then the power overhead could be as low as 10-20% because the bias currents in the amplifiers of the calibration loop can be maintained small. The other investment with this IIP2 enhancement method is the die area required for the calibration circuitry. Depending on the layout style, the mixer area with calibration could be up to twice the area of the mixer without calibration. There is a direct trade-off between the layout area and the IIP2 improvement from better matching between devices. But unlike with conventional matching techniques, the devices with non-minimum lengths in the calibration loops are outside of the signal path and therefore their parasitic capacitances do not degrade the mixer's frequency response.

## VII. SUMMARY AND CONCLUSIONS

### VII.1. Overall Perspective

Contemporary CMOS technologies make it possible to design highly integrated multi-functional chips. On the other hand, the current research and product development trends are associated with several challenges in the quality assurance and reliability of the manufactured devices. As described in Section I, many problems are fundamentally caused by worsening process parameter variations, interactions between individual blocks through coupling effects on the same chip, system complexity, high on-chip power densities, and the increasing number of functions to be verified. In the case of wireless systems, an additional issue is that more and more devices are designed to transmit/receive signals from multiple communication standards, leading to interference problems. A survey of the existing and emerging on-chip built-in test and calibration techniques for single-chip wireless transceivers was presented. Since it embodies various design philosophies in academia and the industry, the overview exposed the diversity among the approaches to solve the current testability and reliability challenges. In general, it can be observed that a tendency exists to combine system-level test and calibration techniques with digitally adaptable circuits within the analog sections of the transceivers, where the digital processor monitors system parameters and controls corrective actions.

Supplemental measurements or calibration loops on the analog circuit level are beneficial to quickly detect and correct gross variations at start-up in order to reduce the computational overhead and time requirements in the digital processor. On-chip built-in

test circuitry also aids the identification of fault location to determine appropriate adjustments. Moreover, certain faults are extremely difficult to observe in the digital baseband of receivers, particularly defects and variations in the RF front-end section such as those related to impedance matching. Hence, many on-chip built-in test and calibration techniques involve analog measurement circuitry. The emphasis in this dissertation was on the exploration of design strategies to make analog circuits more robust to PVT variations. Since this task is very specific to the type of circuit being designed, several examples with different analog and mixed-signal circuits in wireless receivers were discussed. In general, it can be concluded that variation-aware analog design itself is not sufficient to guarantee the required performance in demanding applications. For this reason, it is advisable to equip the analog blocks with features for performance tuning during production testing or even during normal operation of the devices. Most of the alterations proposed for the example circuits in this dissertation encompass digitally programmable elements for compatibility with the system-level calibration approaches that were addressed in Section II.

## **VII.2. Dissertation Projects**

The first example involved the design task to increase the linearity of operational transconductance amplifiers (OTAs) in lowpass filters with wide bandwidth. In Section III, an architectural solution was proposed which is based on cancellation of the main amplifier's nonlinearities with an identical auxiliary OTA. With regards to resilience to PVT variations, the motivation for this approach is that two amplifiers with the same component dimensions and bias conditions exhibit minimal mismatches. This

characteristic is particularly important to arrive at an effective broadband linearization method because it ensures minimal deviations of the high-frequency responses in the main and auxiliary signal paths. Nevertheless, the analysis of the problem and experimental results have revealed that high linearity at high frequencies requires the ability to compensate for PVT variations. To do so, digitally programmable resistor ladders were utilized to perform the necessary post-fabrication gain and phase equalizations for optimum cancellation of nonlinearities. Measurements obtained with a  $0.13\mu\text{m}$  CMOS test chip demonstrated that the nonlinearity cancellation technique improves the IM3 of the designed OTA by up to 22dB at frequencies up to 350MHz. Consuming 5.2mW from a 1.2V supply, the linearized OTA with a  $0.2V_{p-p}$  input signal has an IM3 better than -74dB up to 350MHz and a 70dB signal-to-noise ratio (SNR) in 1MHz bandwidth. The linearization scheme was also tested with multiple OTAs embedded into a lowpass filter having a 195MHz bandwidth. This filter has a measured in-band IIP3 of 14.0dBm and a 54.5dB dynamic range.

In the second presented circuit example, the quantizer topology in Section IV was developed as part of a continuous-time  $\Sigma\Delta$  modulator architecture with 3-bit pulse-width modulation in the feedback path in order to circumvent the nonlinearity problems caused by unit element mismatches in multi-bit feedback circuitry. Besides robustness to process variations, the other incentives for using this  $\Sigma\Delta$  modulator architecture are the scalability and the potential for power savings with state of the art CMOS technology. However, low-jitter clocks are required for this time-based architecture, which is why the 7-phase 400MHz clock signal is provided by an injected-locked clock generator. A



two-step current-mode quantizer was proposed for the  $\Sigma\Delta$  modulator. This 3-bit quantizer utilizes the available clock phases for analog-to-digital conversion with successive approximations. If applications require tuning for finer resolution, the high-impedance of the reference voltage inputs allow them to be generated with low-power on-chip digital-to-analog converters as those used in many system-level calibration schemes. The quantizer functionality was verified through the measurements of the 5<sup>th</sup>-order continuous-time  $\Sigma\Delta$  modulator chip with the embedded quantizer, which was fabricated in a 0.18 $\mu\text{m}$  CMOS process.

Better observability of faults and variations usually improves the accuracy or execution time of test and calibration routines, for which electrical detectors and process monitoring circuits are utilized. Towards this end, a temperature sensing approach has been assessed in Section V. Since this alternative technique does not require a connection to the circuit under test or signal path, it provides a non-influential method for monitoring variations. A design procedure with electro-thermal co-simulation was outlined to evaluate RF circuit performance metrics from the DC output of an on-chip temperature sensor. The proposed fully-differential sensor circuit for this application has been designed with a wide dynamic range, programmable sensitivity to DC and RF power dissipation, and compatibility with CMOS technology. Using an LNA as prototype, measurements obtained with a 0.18 $\mu\text{m}$  CMOS technology test chip showed that RF power dissipation can be observed with the on-chip temperature sensor. Furthermore, the 1-dB compression point can be estimated with less than 1dB error. The sensor circuitry with 0.012 $\text{mm}^2$  die area can be shared when several on-chip test points

are monitored by placement of multiple temperature-sensing parasitic bipolar devices having an emitter area of  $11\mu\text{m} \times 11\mu\text{m}$ .

Finally, an alternative approach to alleviating the effects of process parameter variations was proposed in Section VI. Rather than employing digitally adjustable elements, the mismatch reduction scheme employs an automatic analog calibration loop to improve the matching of transistors in the high-frequency differential signal path. The method is intended for analog circuits in which short-channel devices are used to minimize bandwidth reduction from parasitic capacitances, and in which transistors are not directly matched to reduce high-frequency coupling through layout parasitics and substrate leakage. Monte Carlo simulations were performed to evaluate the approach for two example circuits designed with 90nm and  $0.13\mu\text{m}$  CMOS technology. In the first case, the application of the mismatch reduction loop to a differential amplifier with 13dB gain and a -3dB frequency of 2.14GHz lowered the simulated standard deviation of the input-referred offset voltage from 4.17mV to 0.76mV-1.29mV, depending on the assumed layout of the sensing-transistors. In the second case, the mismatch reduction loop was used to boost the simulated IIP2 of a double-balanced mixer by 5dB via improvement of the matching between the switching transistors. Based on the results, it can be concluded that this mismatch reduction scheme is suitable for fast coarse calibration at start-up because the loop's settling time can be kept in the range of a few microseconds. If further calibration accuracy is needed and on-chip digital resources are available, then it could be explored to merge the analog loop with digitally-controlled elements within the mixer for system-level calibration with longer convergence.

## REFERENCES

- [1] *International Technology Roadmap for Semiconductors, Test and Test Equipment*, 2009 edition. Available: <http://public.itrs.net/reports.html>
- [2] S. Menon and C. L. Horney, "Smartphone & Chip Market Opportunities," Market research report no. 9010, Forward Concepts Co., Feb. 5 2009. Available: <http://fwdconcepts.com/Smartphones>
- [3] J. Altet, E. Aldrete-Vidrio, D. Mateo, A. Salhi, S. Grauby, W. Caléis, S. Dilhaire, X. Perpiñà, and X. Jordà, "Heterodyne lock-in thermal coupling measurements in integrated circuits: applications to test and characterization," *Review of Scientific Instruments*, vol. 80, no. 2, pp. 026101 – 026101-3, Feb. 2009.
- [4] C. Chiang and J. Kawa, *Design for Manufacturability and Yield for Nano-scale CMOS*, Dordrecht, The Netherlands: Springer, 2007, pp. 14-15.
- [5] W. Zhao, Y. Cao, F. Liu, K. Agarwal, D. Acharyya, S. Nassif, and K. Nowka, "Rigorous extraction of process variations for 65nm CMOS design," in *Proc. European Solid-State Device Research Conf. (ESSDERC)*, Sept. 2007, pp. 89-92.
- [6] G.W. Roberts and B. Dufort, "Making complex mixed-signal telecommunication integrated circuits testable," *IEEE Communications Magazine*, pp. 90-96, June 1999.
- [7] A. Zjajo and J. P. de Gyvez, "Evaluation of signature-based testing of RF/analog circuits," in *Proc. European Test Symp.*, May 2005, pp. 62-67.
- [8] G. G. E. Gielen, "Design methodologies and tools for circuit design in CMOS nanometer technologies," in *Proc. European Solid-State Device Research Conf. (ESSDERC)*, Sept. 2006, pp. 21-32.
- [9] H. Masuda, M. Tsunozaki, T. Tsutsui, H. Nunogami, A. Uchida, and K. Tsunokuni, "A novel wafer-yield PDF model and verification with 90–180-nm SOC chips," *IEEE Trans. Semiconductor Manufacturing*, vol. 21, no. 4, pp. 585-591, Nov. 2008.
- [10] V. A. Zivkovic, F. van der Heyden, G. Gronthoud, and F. de Jong, "Analog test bus infrastructure for RF/AMS modules in core-based design," in *Proc. 13th European Test Symp.*, May 2008, pp. 27-32.

- [11] K. Agarwal, J. Hayes, and S. Nassif, "Fast characterization of threshold voltage fluctuation in MOS devices," *IEEE Trans. Semiconductor Manufacturing*, vol. 21, no. 4, pp. 526-533, Nov. 2008.
- [12] J. P. F. Glas, "Digital I/Q imbalance compensation in a low-IF receiver," in *Proc. IEEE Global Telecommunications Conf. (GLOBECOM)*, vol. 3, Nov. 1998, pp. 1461-1466.
- [13] W. Eberle, J. Tubbax, B. Come, S. Donnay, H. De Man, and G. Gielen, "OFDM-WLAN receiver performance improvement using digital compensation techniques," in *Proc. IEEE Radio and Wireless Conf. (RAWCON)*, Aug. 2002, pp. 111-114.
- [14] I. Elahi, K. Muhammad, and P. T. Balsara, "I/Q mismatch compensation using adaptive decorrelation in a low-IF receiver in 90-nm CMOS process," *IEEE J. Solid-State Circuits*, vol. 41, no. 2, pp. 395-404, Feb. 2006.
- [15] B. Shi and Y. W. Chia "An analog mismatch calibration system for image-reject receivers," in *Proc. Eur. Conf. on Wireless Technology*, Oct. 2005, pp. 225-228.
- [16] R. B. Staszewski, I. Bashir, and O. Eliezer, "RF Built-in self test of a wireless transmitter," *IEEE Trans. Circuits and Systems II: Express Briefs*, vol. 54, pp. 186-190, Feb. 2007.
- [17] R. Montemayor and B. Razavi, "A self-calibrating 900-MHz CMOS image-reject receiver," in *Proc. Eur. Solid-State Circuits Conf. (ESSCIRC)*, Sept. 2000, pp. 320-323.
- [18] M. A. I. Elmala and S. H. K. Embabi, "Calibration of phase and gain mismatches in Weaver image-reject receiver," *IEEE J. of Solid-State Circuits*, vol. 39, no. 2, pp. 283-289, Feb. 2004.
- [19] J.-Y. Ryu, B. C. Kim, and I. Sylla, "A new low-cost RF built-in self-test measurement for system-on-chip transceivers," *IEEE Trans. Instrumentation and Measurement*, vol. 55, no. 2, pp. 381-388, April 2006.
- [20] Q. Yin, W. R. Eisenstadt, R. M. Fox, and T. Zhang, "A translinear RMS detector for embedded test of RF ICs," *IEEE Trans. Instrumentation and Measurement*, vol. 54, no. 5, pp. 1708-1714, Oct. 2005.

- [21] S. Bhattacharya and A. Chatterjee, "Use of embedded sensors for built-in-test RF circuits," in *Proc. IEEE Intl. Test Conf. (ITC)*, Oct. 2004, pp. 801-809.
- [22] Q. Wang and M. Soma, "RF front-end system gain and linearity built-in test," in *Proc. 24th IEEE VLSI Test Symp.*, May 2006, pp. 228-233.
- [23] A. Valdes-Garcia, R. Venkatasubramanian, J. Silva-Martinez, and E. Sánchez-Sinencio, "A broadband CMOS amplitude detector for on-chip RF measurements," *IEEE Trans. Instrumentation and Measurement*, vol. 57, no. 7, pp. 1470-1477, July 2008.
- [24] T. Das, A. Gopalan, C. Washburn, and P. R. Mukund, "Self-calibration of input-match in RF front-end circuitry," *IEEE Trans. Circuits and Systems II: Express Briefs*, vol. 52, no. 12, pp. 821-825, Dec. 2005.
- [25] V. Stopjakova, H. Manhaeve, and M. Sidiropulos, "On-chip transient current monitor for testing of low-voltage CMOS IC," in *Proc. Design, Automation and Test in Europe Conf. and Exhib.*, March 1999, pp. 538-542.
- [26] A. P. Jose, K. A. Jenkins, and S. K. Reynolds, "On-chip spectrum analyzer for analog built-in self test," in *Proc. IEEE VLSI Test Symp.*, May 2005, pp. 131-136.
- [27] A. Valdes-Garcia, F. A.-L. Hussien, J. Silva-Martinez, and E. Sánchez-Sinencio, "An integrated frequency response characterization system with a digital interface for analog testing," *IEEE J. Solid-State Circuits*, vol. 41, no. 10, pp. 2301-2313, October 2006.
- [28] J. J. Dabrowski and R. M. Ramzan, "Built-in loopback test for IC RF transceivers," *IEEE Trans. Very Large Scale Integration (VLSI) Systems*, vol. 18, no. 6, pp. 933-946, June 2010.
- [29] M. Onabajo, J. Silva-Martinez, F. Fernandez, and E. Sánchez-Sinencio, "An on-chip loopback block for RF transceiver built-in test," *IEEE Trans. Circuits and Systems II: Express Briefs*, vol. 56, no. 6, pp. 444-448, June 2009.
- [30] G. Srinivasan, A. Chatterjee, and F. Taenzler, "Alternate loop-back diagnostic tests for wafer-level diagnosis of modern wireless transceivers using spectral signatures," in *Proc. 24th VLSI Test Symp.*, May 2006, pp. 222-227.

- [31] A. Haider, S. Bhattacharya, G. Srinivasan, and A. Chatterjee, "A system-level alternate test approach for specification test of RF transceivers in loopback mode," in *Proc. 18th Intl. Conf. on VLSI Design*, Jan. 2005, pp. 289-294.
- [32] M. Negreiros, L. Carro, and A. A. Susin, "An improved RF loopback for test time reduction," in *Proc. Design, Automation, and Test in Europe Conf. and Exhib.*, Mar. 2006, pp. 646-651.
- [33] D. Kaczman, M. Shah, M. Alam, M. Rachedine, D. Cashen, L. Han, and A. Raghavan, "A single-chip 10-band WCDMA/HSDPA 4-band GSM/EDGE SAW-less CMOS receiver with DigRF 3G interface and +90 dBm IIP2," *IEEE J. Solid-State Circuits*, vol. 44, no. 3, pp. 718-739, March 2009.
- [34] H. Darabi, J. Chiu, S. Khorram, H. J. Kim, Z. Zhou, H.-M. Chien, B. Ibrahim, E. Geronaga, L. H. Tran, and A. Rofougaran, "A dual-mode 802.11b/Bluetooth radio in 0.35- $\mu$ m CMOS," *IEEE J. Solid-State Circuits*, vol. 40, no. 3, pp. 698-706, March 2005.
- [35] I. Vassiliou, K. Vavelidis, T. Georgantas, S. Plevridis, N. Haralabidis, G. Kamoulakos, C. Kapnistis, S. Kavadias, Y. Kokolakis, P. Merakos, J. C. Rudell, A. Yamanaka, S. Bouras, and I. Bouras, "A single-chip digitally calibrated 5.15-5.825-GHz 0.18- $\mu$ m CMOS transceiver for 802.11a wireless LAN," *IEEE J. Solid-State Circuits*, vol. 38, no. 12, pp. 2221-2231, Dec. 2003.
- [36] Y.-H. Hsieh, W.-Y. Hu, S.-M. Lin, C.-L. Chen, W.-K. Li, S.-J. Chen, and D. J. Chen, "An auto-I/Q calibrated CMOS transceiver for 802.11g," *IEEE J. Solid-State Circuits*, vol. 40, no. 11, pp. 2187-2192, Nov. 2005.
- [37] O. Eliezer, R. B. Staszewski, and D. Mannath, "A statistical approach for design and testing of analog circuitry in low-cost SoCs," in *Proc. IEEE Intl. Midwest Symposium on Circuits and Systems (MWSCAS)*, Aug. 2010, pp. 461-464.
- [38] M. Onabajo, F. Fernandez, J. Silva-Martinez, and E. Sánchez-Sinencio, "Strategic test cost reduction with on-chip measurement circuitry for RF transceiver front-ends – an overview," in *Proc. IEEE Intl. Midwest Symp. on Circuits and Systems (MWSCAS)*, Aug. 2006, vol. 2, pp. 643-647.

- [39] V. Saari, M. Kaltiokallio, S. Lindfors, J. Ryynänen, and K. A. I. Halonen, "A 240-MHz low-pass filter with variable gain in 65-nm CMOS for a UWB radio receiver," *IEEE Trans. Circuits and Systems I: Regular Papers*, vol. 56, no. 7, pp. 1488-1499, July 2009.
- [40] M. Gambhir, V. Dhanasekaran, J. Silva-Martinez, and E. Sánchez-Sinencio, "A low power 1.3GHz dual-path current mode Gm-C filter," in *Proc. IEEE Custom Integrated Circuits Conf. (CICC)*, Sept. 2008, pp.703-706.
- [41] R. Schoofs, M. S. J. Steyaert, and W. M. C. Sansen, "A design-optimized continuous-time delta-sigma ADC for WLAN applications," *IEEE Trans. Circuits and Systems I: Regular Papers*, vol. 54, no. 1, pp. 209-217, Jan. 2007.
- [42] D. Healy, *Analog-to-Information (A-to-I) Receiver Development Program*, BAA 08-03 Announcement, Defense Advanced Research Projects Agency (DARPA), Microsystems Technology Office (MTO), Nov. 2007.
- [43] J. C. Rudell, O. E. Erdogan, D. G. Yee, R. Brockenbrough, C. S. G. Conroy, and B. Kim, "A 5th-order continuous-time harmonic-rejection GmC filter with in-situ calibration for use in transmitter applications," in *IEEE Intl. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2005, pp. 322-323.
- [44] A. Lewinski and J. Silva-Martinez, "A high-frequency transconductor using a robust nonlinearity cancellation," *IEEE Trans. Circuits and Systems II: Express Briefs*, vol. 53, no. 9, pp. 896-900, Sept. 2006.
- [45] E. A. M. Klumperink and B. Nauta, "Systematic comparison of HF CMOS transconductors," *IEEE Trans. Circuits and Systems II: Express Briefs*, vol. 50, no. 10, pp. 728-741, Oct. 2003.
- [46] S. D'Amico, M. Conta, and A. Baschiroto, "A 4.1-mW 10-MHz fourth-order source-follower-based continuous-time filter with 79-dB DR," *IEEE J. Solid-State Circuits*, vol. 41, no. 12, pp. 2713-2719, Dec. 2006.
- [47] T. Y. Lo and C.-C. Hung, "A 40-MHz double differential-pair CMOS OTA with -60dB IM3," *IEEE Trans. Circuits and Systems I: Regular Papers*, vol.55, no.1, pp. 258-265, Feb. 2008.

- [48] J. Chen, E. Sánchez-Sinencio, and J. Silva-Martinez, "Frequency-dependent harmonic-distortion analysis of a linearized cross-coupled CMOS OTA and its application to OTA-C filters," *IEEE Trans. Circuits and Systems I: Regular Papers*, vol. 53, no. 3, pp. 499-510, March 2006.
- [49] W. Huang and E. Sánchez-Sinencio, "Robust highly linear high-frequency CMOS OTA with IM3 below -70 dB at 26MHz," *IEEE Trans. Circuits and Systems I: Regular Papers*, vol.53, no.7, pp. 1433-1447, July 2006.
- [50] D. Yongwang and R. Harjani, "A +18 dBm IIP3 LNA in 0.35 $\mu$ m CMOS," in *IEEE Intl. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2001, pp.162-163.
- [51] R. Chawla, F. Adil, G. Serrano, and P. E. Hasler, "Programmable Gm-C filters using floating-gate operational transconductance amplifiers," *IEEE Trans. Circuits and Systems I: Regular Papers*, vol. 54, no. 3, pp. 481-491, March 2007.
- [52] S. Maas, *Nonlinear Microwave and RF Circuits*. Boston, MA: Artech House, 2003.
- [53] E. Rodriguez-Villegas and H. Barnes, "Solution to trapped charge in FG MOS transistors," *Electronics Letters*, vol. 39, no. 19, pp. 1416-1417, Sept. 2003.
- [54] A. P. Nedungadi and R. L. Geiger, "High-frequency voltage-controlled continuous time lowpass filter using linearized CMOS integrators," *Electronic Letters*, vol. 22, pp. 729-731, July 1986.
- [55] A. Valdes-Garcia, R. Venkatasubramanian, R. Srinivasan, J. Silva-Martinez, and E. Sánchez-Sinencio, "A CMOS RF RMS detector for built-in testing of wireless transceivers," in *Proc. IEEE VLSI Test Symp.*, May 2005, pp. 249-254.
- [56] G. Bollati, S. Marchese, M. Demicheli, and R. Castello, "An eighth-order CMOS low-pass filter with 30-120 MHz tuning range and programmable boost," *IEEE J. Solid-State Circuits*, vol. 36, no. 7, pp. 1056-1066, July 2001.
- [57] A. Otin, S. Celma, and C. Aldea, "A 40-200 MHz programmable 4th-order Gm-C filter with auto-tuning system," in *Proc. 33rd Eur. Solid-State Circuits Conf. (ESSCIRC)*, Sept. 2007, pp. 214-217.



- [58] S. Dosho, T. Morie, and H. Fujiyama, "A 200-MHz seventh-order equiripple continuous-time filter by design of nonlinearity suppression in 0.25- $\mu\text{m}$  CMOS process," *IEEE J. Solid-State Circuits*, vol. 37, no. 5, pp. 559-565, May 2002.
- [59] S. Pavan and T. Laxminidhi, "A 70-500MHz programmable CMOS filter compensated for MOS nonquasistatic effects," in *Proc. 32nd Eur. Solid-State Circuits Conf. (ESSCIRC)*, Sept. 2006, pp. 328-331.
- [60] K. Kwon, H.-T. Kim, and K. Lee, "A 50–300-MHz highly linear and low-noise CMOS Gm-C filter adopting multiple gated transistors for digital TV tuner ICs," *IEEE Trans. Microwave Theory and Techniques*, vol. 57, no. 2, pp. 306-313, Feb. 2009.
- [61] J. A. Cherry and W. M. Snelgrove, *Continuous-Time Delta-Sigma Modulators for High-Speed A/D Conversion*. Boston, MA: Kluwer, 2000.
- [62] K. Matsukawa, Y. Mitani, M. Takayama, K. Obata, S. Dosho, and A. Matsuzawa, "A fifth-order continuous-time delta-sigma modulator with single-opamp resonator", *IEEE J. Solid-State Circuits*, vol. 45, no. 4, pp. 697-706, April 2010.
- [63] A. Yasuda, H. Tanimoto, and T. Iida, "A third-order  $\Delta$ - $\Sigma$  modulator using second-order noise-shaping dynamic element matching," *IEEE J. Solid-State Circuits*, vol. 33, no. 12, pp. 1879-1886, Dec. 1998.
- [64] E. N. Aghdam and P. Benabes, "Higher order dynamic element matching by shortened tree-structure in delta-sigma modulators," in *Proc. European Conf. Circuit Theory and Design*, vol. 1, Sept. 2005, pp. 201-204.
- [65] R. T. Baird and T. S. Fiez, "Improved  $\Delta\Sigma$  DAC linearity using data weighted averaging," in *Proc. IEEE Intl. Symp. Circuits and Systems (ISCAS)*, May 1995, vol. 1, pp. 13-16.
- [66] W. Yang, W. Schofield, H. Shibata, S. Korrapati, A. Shaikh, N. Abaskharoun, and D. Ribner, "A 100mW 10MHz-BW CT  $\Delta\Sigma$  modulator with 87dB DR and 91dBc IMD," in *IEEE Intl. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2008, pp. 498-631.

- [67] M. Z. Straayer and M. H. Perrott, "A 12-bit, 10-MHz bandwidth, continuous-time  $\Sigma\Delta$  ADC with a 5-bit, 950-MS/s VCO-based quantizer," *IEEE J. Solid-State Circuits*, vol. 43, no. 4, pp. 805-814, Apr. 2008.
- [68] V. Dhanasekaran, M. Gambhir, M. M. Elsayed, E. Sanchez-Sinencio, J. Silva-Martinez, C. Mishra, L. Chen, and E. Pankratz, "A 20MHz BW 68dB DR CT  $\Delta\Sigma$  ADC based on a multi-bit time-domain quantizer and feedback element," in *IEEE Intl. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2009, pp. 174-175.
- [69] F. Colodro and A. Torralba, "New continuous-time multibit sigma-delta modulators with low sensitivity to clock jitter," *IEEE Trans. Circuits and Systems I: Regular Papers*, vol. 56, no. 1, pp. 74-83, Jan. 2009.
- [70] C.-Y. Chen, M. Q. Le, and K. Y. Kim, "A low power 6-bit flash ADC with reference voltage and common-mode calibration," *IEEE J. Solid-State Circuits*, vol. 44, no. 4, pp. 1041-1046, April 2009.
- [71] Y.-Z. Lin, C.-W. Lin, and S.-J. Chang, "A 5-bit 3.2-GS/s flash ADC with a digital offset calibration scheme," *IEEE Trans. Very Large Scale Integration (VLSI) Systems*, vol. 18, no. 3, pp. 509-513, March 2010.
- [72] J. Doernberg, P. R. Gray, and D. A. Hodges, "A 10-bit 5-Msample/s CMOS two-step flash ADC," *IEEE J. Solid-State Circuits*, vol. 24, no. 2, pp. 241-249, April 1989.
- [73] B. Verbruggen, J. Craninckx, M. Kuijk, P. Wambacq, and G. Van der Plas, "A 2.2 mW 1.75 GS/s 5 bit folding flash ADC in 90 nm digital CMOS," *IEEE J. Solid-State Circuits*, vol. 44, no. 3, pp.874-882, March 2009.
- [74] S.-W. Chen and R. W. Brodersen, "A 6-bit 600-MS/s 5.3-mW asynchronous ADC in 0.13- $\mu$ m CMOS," *IEEE J. Solid-State Circuits*, vol. 41, no. 12, pp. 2669-2680, Dec. 2006.
- [75] G. Van der Plas and B. Verbruggen, "A 150MS/s 133 $\mu$ W 7b ADC in 90nm digital CMOS using a comparator-based asynchronous binary-search sub-ADC," in *IEEE Intl. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2008, pp. 242-243, 610.

- [76] J. Craninckx and G. Van der Plas, "A 65fJ/conversion-step 0-to-50MS/s 0-to-0.7mW 9b charge-sharing SAR ADC in 90nm digital CMOS," in *IEEE Intl. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2007, pp. 246-247, 600.
- [77] L. Dorrer, F. Kuttner, P. Greco, P. Torta, and T. Hartig, "A 3-mW 74-dB SNR 2-MHz continuous-time delta-sigma ADC with a tracking ADC quantizer in 0.13- $\mu$ m CMOS," *IEEE J. Solid-State Circuits*, vol. 40, no. 12, pp. 2416- 2427, Dec. 2005.
- [78] Y.-C. Lo, H.-P. Chen, J. Silva-Martinez, and S. Hoyos, "A 1.8V, sub-mW, over 100% locking range, divide-by-3 and 7 complementary-injection-locked 4 GHz frequency divider," in *Proc. IEEE Custom Integrated Circuits Conf. (CICC)*, Sept. 2009, pp. 259-262.
- [79] C.-Y. Lu, "Calibrated continuous-time sigma-delta modulators," Ph.D. dissertation, Dept. of Electrical and Computer Eng., Texas A&M University, 2010.
- [80] M. Dessouky and A. Kaiser, "Input switch configuration suitable for rail-to-rail operation of switched op amp circuits," *Electronics Letters*, vol. 35, no. 1, pp. 8-10, Jan. 1999.
- [81] T. Voo and C. Toumazou, "High-speed current mirror resistive compensation technique," *Electronics Letters*, vol. 31, no. 4, pp. 248-250, Feb. 1995.
- [82] M. Bazes, "Two novel fully complementary self-biased CMOS differential amplifiers," *IEEE J. Solid-State Circuits*, vol. 26, no. 2, pp. 165-168, Feb. 1991.
- [83] P. E. Allen and D. R. Holberg, *CMOS Analog Circuit Design*, 2nd ed., London, U.K.: Oxford Univ. Press, 2002, pp. 477-480.
- [84] G. M. Yin, F. O. Eynde, and W. Sansen, "A high-speed CMOS comparator with 8-b resolution," *IEEE J. Solid-State Circuits*, vol. 27, no. 2, pp. 208-211, Feb. 1992.
- [85] L. Sumanen, M. Waltari, V. Hakkarainen, and K. Halonen, "CMOS dynamic comparators for pipeline A/D converters," in *Proc. IEEE Intl. Conf. Circuits and Systems (ISCAS)*, May 2002, vol. 5, pp. V-157 – V-160.
- [86] B. Razavi and B. A. Wooley, "Design techniques for high-speed, high-resolution comparators," *IEEE J. Solid-State Circuits*, vol. 27, no. 12, pp. 1916-1926, Dec. 1992.

- [87] P. Amaral, J. Goes, N. Paulino, and A. Steiger-Garcao, "An improved low-voltage low-power CMOS comparator to be used in high-speed pipeline ADCs," in *Proc. IEEE Intl. Conf. Circuits and Systems (ISCAS)*, May 2002, vol. 5, pp. V-141 – V-144.
- [88] L. J. Breems, R. Rutten, R. van Veldhoven, G. van der Weide, and H. Termeer, "A 56mW CT quadrature cascaded  $\Sigma\Delta$  modulator with 77dB DR in a near zero-IF 20MHz band," in *IEEE Intl. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2007, pp. 238-239.
- [89] P. Malla, H. Lakdawala, K. Kornegay, and K. Soumyanath, "A 28mW spectrum-sensing reconfigurable 20MHz 72dB-SNR 70dB-SNDR DT  $\Delta\Sigma$  ADC for 802.11n/WiMAX Receivers," in *IEEE Intl. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2008, pp. 496-497.
- [90] G. Mitteregger, C. Ebner, S. Mechnig, T. Blon, C. Holuigue, and E. Romani, "A 20-mW 640-MHz CMOS continuous-time  $\Sigma\Delta$  ADC with 20-MHz signal bandwidth, 80-dB dynamic range and 12-bit ENOB," *IEEE J. Solid-State Circuits*, vol. 41, no. 12, pp. 2641-2649, Dec. 2006.
- [91] M. Park and M. Perrott, "A 0.13 $\mu$ m CMOS 78dB SNDR 87mW 20MHz BW CT  $\Delta\Sigma$  ADC with VCO-based integrator and quantizer," in *IEEE Intl. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2009, pp. 170-171, 171a.
- [92] X. Fan, M. Onabajo, F. O. Fernández-Rodríguez, J. Silva-Martinez, and E. Sánchez-Sinencio, "A current injection built-in test technique for RF low-noise amplifiers," *IEEE Trans. Circuits and Systems I: Regular Papers*, vol. 55, no. 7, pp. 1794-1804, Aug. 2008.
- [93] D. J. Walkey, T. S. Smy, R. G. Dickson, J. S. Brodsky, D. T. Zweidinger, and R. M. Fox, "Equivalent circuit modeling of static substrate thermal coupling using VCVS representation," *IEEE J. Solid-State Circuits*, vol. 37, no. 9, pp. 1198-1205, Sept. 2002.
- [94] N. Nenadovic, S. Mijalkovic, L. K. Nanver, L. K. J. Vandamme, V. d'Alessandro, H. Schellevis, and J. W. Slotboom, "Extraction and modeling of self-heating and mutual

- thermal coupling impedance of bipolar transistors," *IEEE J. Solid-State Circuits*, vol. 39, no. 10, pp. 1764-1772, Oct. 2004.
- [95] J. Altet, A. Rubio, E. Schaub, S. Dilahire, and W. Claeys, "Thermal coupling in integrated circuits: application to thermal testing," *IEEE J. Solid-State Circuits*, vol. 36, no. 1, pp. 81-91, Jan. 2001.
- [96] S. Mattisson, H. Hagberg, and P. Andreani, "Sensitivity degradation in a tri-band GSM BiCMOS direct-conversion receiver caused by transient substrate heating," *IEEE J. Solid-State Circuits*, vol. 43, no. 2, pp. 486-496, Feb. 2008.
- [97] D. Mateo, J. Altet, E. Aldrete-Vidrio, and J. L. Gonzalez, "Frequency characterization of a 2.4 GHz CMOS LNA by thermal measurements," in *Proc. IEEE Radio Frequency Integrated Circuits (RFIC) Symp.*, June 2006, pp. 565-568.
- [98] J. Altet, E. Aldrete-Vidrio, D. Mateo, X. Perpiñà, X. Jordà, M. Vellvehi, J. Millán, A. Salhi, S. Grauby, W. Claeys and S. Dilhaire, "A heterodyne method for the thermal observation of the electrical behavior of high-frequency integrated circuits," *Measurement Science and Technology*, vol. 19, no. 11, pp. 115704 (8pp), Nov. 2008.
- [99] Michael D. Scott, Bernhard E. Boser, and Kristofer S. J. Pister, "An ultralow-energy ADC for smart dust," *IEEE J. Solid-State Circuits*, vol. 38, no. 7, pp. 1123-1129, July 2003.
- [100] N. Verma and A. P. Chandrakasan, "An ultra low energy 12-bit rate-resolution scalable SAR ADC for wireless sensor nodes," *IEEE J. Solid-State Circuits*, vol. 42, no. 6, pp. 1196-1205, June 2007.
- [101] L. Codecasa, D. D'Amore, and P. Maffezzoni, "Modeling the thermal response of semiconductor devices through equivalent electrical networks," *IEEE Trans. Circuits and Systems I: Fundamental Theory and Applications*, vol. 49, no. 8, pp. 1187-1197, Aug. 2002.
- [102] V. Szekely, "On the representation of infinite-length distributed RC one-ports," *IEEE Trans. Circuits and Systems*, vol. 38, no. 7, pp. 711-719, July 1991.
- [103] S.-S. Lee and D. J. Allstot, "Electrothermal simulations of integrated circuits," *IEEE J. Solid-State Circuits*, vol. 28, no. 12, pp. 1283-1293, Dec. 1993.

- [104] W. Van Petegem, B. Geeraerts, W. Sansen, and B. Graindourze, "Electrothermal simulation and design of integrated circuits," *IEEE J. Solid-State Circuits*, vol. 29, no. 2, pp. 143-146, Feb. 1994.
- [105] J. Michejda and S. K. Kim, "A precision CMOS bandgap reference," *IEEE J. Solid-State Circuits*, vol. 19, no. 6, pp. 1014-1021, Dec. 1984.
- [106] H. M. Geddada, J. W. Park, and J. Silva-Martinez, "Robust derivative superposition method for linearising broadband LNAs," *Electronics Letters*, vol. 45, no. 9, pp. 435-436, April 2009.
- [107] E. Aldrete-Vidrio, D. Mateo, and J. Altet, "Differential temperature sensors fully compatible with a 0.35- $\mu\text{m}$  CMOS process," *IEEE Trans. Components and Packaging Technologies*, vol. 30, no. 4, pp. 618-626, Dec. 2007.
- [108] M. A. P. Pertijs, G. C. M. Meijer, and J. H. Huijsing, "Precision temperature measurement using CMOS substrate pnp transistors," *IEEE Sensors Journal*, vol. 4, no. 3, pp. 294-300, June 2004.
- [109] D. Gomez and D. Mateo, "Exploiting CMOS short-channel effects for yield enhancement in analogue/RF design," *Electronics Letters*, vol. 46, no. 8, pp. 559-561, April 2010.
- [110] S. Rodriguez, A. Rusu, L.-R. Zheng, and M. Ismail, "CMOS RF mixer with digitally enhanced IIP2," *Electronics Letters*, vol. 44, no. 2, pp. 121-122, Jan. 2008.
- [111] Gupta, V., and Rincon-Mora, G.A., "Achieving less than 2% 3- $\sigma$  mismatch with minimum channel-length CMOS devices," *IEEE Trans. Circuits and Systems II: Express Briefs*, vol. 54, no. 3, pp. 232-236, March 2007.
- [112] M. Conti, P. Crippa, S. Orcioni, and C. Turchetti, "Layout-based statistical modeling for the prediction of the matching properties of MOS transistors," *IEEE Trans. Circuits and Systems I: Fundamental Theory and Applications*, vol. 49, no. 5, pp. 680-685, May 2002.
- [113] T.-H. Yeh, J.C.H. Lin, S.-C. Wong, H. Huang, and J.Y.C. Sun, "Mis-match characterization of 1.8V and 3.3V devices in 0.18 $\mu\text{m}$  mixed signal CMOS

- technology," in *Proc. IEEE Intl. Conf. Microelectronic Test Structures (ICMTS)*, March 2001, pp. 77-82.
- [114] M. J. M. Pelgrom, A. C. J. Duinmaijer, and A. P. G. Welbers, "Matching properties of MOS transistors," *IEEE J. Solid-State Circuits*, vol. 24, no. 5, pp. 1433-1439, Oct. 1989.
- [115] Cadence Design Systems, "Recommended Monte Carlo modeling methodology for Virtuoso Spectre circuit simulator application note," pp. 13-18, Nov. 2003. Available: [http://www.cdnusers.org/community/virtuoso/resources/spectre\\_mcm modelingAN.pdf](http://www.cdnusers.org/community/virtuoso/resources/spectre_mcm modelingAN.pdf)
- [116] K. Kivekas, A. Parssinen, and K. A. I. Halonen, "Characterization of IIP2 and DC-offsets in transconductance mixers," *IEEE Trans. Circuits and Systems II: Analog and Digital Signal Processing*, vol. 48, no. 11, pp. 1028-1038, Nov. 2001.
- [117] D. Manstretta, M. Brandolini, and F. Svelto, "Second-order intermodulation mechanisms in CMOS downconverters," *IEEE J. Solid-State Circuits*, vol. 38, no. 3, pp. 394-406, March 2003.
- [118] M. Brandolini, P. Rossi, D. Sanzogni, and F. Svelto, "A +78 dBm IIP2 CMOS direct downconversion mixer for fully integrated UMTS receivers," *IEEE J. Solid-State Circuits*, vol. 41, no. 3, pp. 552-559, March 2006.
- [119] S. Rodriguez, S. Tao, M. Ismail, and A. Rusu, "An IIP2 digital calibration technique for passive CMOS down-converters," in *Proc. IEEE Intl. Symp. Circuits and Systems (ISCAS)*, May 2010, pp. 825-828.
- [120] S. Peng, C.-C. Chen, and A. Bellaouar, "A wide-band mixer for WCDMA/CDMA2000 in 90nm digital CMOS process," in *Proc. IEEE Radio Frequency Integrated Circuits (RFIC) Symp.*, June 2005, pp. 179-182.
- [121] E. E. Bautista, B. Bastani, and J. Heck, "A high IIP2 downconversion mixer using dynamic matching," *IEEE J. Solid-State Circuits*, vol. 35, no. 12, pp. 1934-1941, Dec. 2000.
- [122] T. H. Lee, *The Design of CMOS Radio-Frequency Integrated Circuits*, Cambridge, U.K.: Cambridge University Press, 1998.

- [123] K. Kivekas, A. Parssinen, J. Ryynanen, J. Jussila, and K. Halonen, "Calibration techniques of active BiCMOS mixers," *IEEE J. Solid-State Circuits*, vol. 37, no. 6, pp. 766-769, Jun 2002.
- [124] W. Kim, S.-G. Yang, J. Yu, H. Shin, W. Choo, and B.-H. Park, "A direct conversion receiver with an IP2 calibrator for CDMA/PCS/GPS/AMPS applications," *IEEE J. Solid-State Circuits*, vol. 41, no. 7, pp. 1535-1541, July 2006.
- [125] M. Hotti, J. Ryynanen, and K. Halonen, "IIP2 calibration methods for current output mixer in direct-conversion receivers," in *Proc. IEEE Intl. Symp. Circuits and Systems (ISCAS)*, May 2005, vol. 5, pp. 5059- 5062.
- [126] K. Dufrene and R. Weigel, "A novel IP2 calibration method for low-voltage downconversion mixers," in *Proc. IEEE Radio Frequency Integrated Circuits (RFIC) Symp.*, June 2006, pp. 327-330.
- [127] M. Brandolini, M. Sosio, and F. Svelto, "A 750 mV fully integrated direct conversion receiver front-end for GSM in 90-nm CMOS," *IEEE J. Solid-State Circuits*, vol. 42, no. 6, pp. 1310-1317, June 2007.
- [128] H. Darabi, H. J. Kim, J. Chiu, B. Ibrahim, and L. Serrano, "An IP2 improvement technique for zero-IF down-converters," in *IEEE Intl. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2006, pp. 1860-1869.
- [129] M. Chen, Y. Wu, and M. F. Chang, "Active 2nd-order intermodulation calibration for direct-conversion receivers," in *IEEE Intl. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2006, pp. 1830-1839.
- [130] K. Dufrene, Z. Boos, and R. Weigel, "Digital adaptive IIP2 calibration scheme for CMOS downconversion mixers," *IEEE J. Solid-State Circuits*, vol. 43, no. 11, pp. 2434-2445, Nov. 2008.
- [131] M. B. Vahidfar and O. Shoaie, "A New IIP2 enhancement technique for CMOS down-converter mixers," *IEEE Trans. Circuits and Systems II: Express Briefs*, vol. 54, no. 12, pp. 1062-1066, Dec. 2007.
- [132] P. Sivonen, A. Vilander, and A. Parssinen, "Cancellation of second-order intermodulation distortion and enhancement of IIP2 in common-source and common-



- emitter RF transconductors," *IEEE Trans. Circuits and Systems I: Regular Papers*, vol. 52, no. 2, pp. 305-317, Feb. 2005.
- [133] H. Lee and S. Mohammadi, "A 500 $\mu$ W 2.4GHz CMOS subthreshold mixer for ultra low power applications," in *Proc. IEEE Radio Frequency Integrated Circuits (RFIC) Symp.*, June 2007, pp. 325-328.
- [134] B. G. Perumana, R. Mukhopadhyay, S. Chakraborty, C.-H. Lee, and J. Laskar, "A low-power fully monolithic subthreshold CMOS receiver with integrated LO generation for 2.4 GHz wireless PAN applications," *IEEE J. Solid-State Circuits*, vol. 43, no. 10, pp. 2229-2238, Oct. 2008.
- [135] J.-B. Seo, J.-H. Kim, H. Sun, and T.-Y. Yun, "A low-power and high-gain mixer for UWB systems," *IEEE Microwave and Wireless Components Letters*, vol. 18, no. 12, pp. 803-805, Dec. 2008.
- [136] A. V. Do, C. C. Boon, M. A. Do, K. S. Yeo, and A. Cabuk, "A weak-inversion low-power active mixer for 2.4 GHz ISM band applications," *IEEE Microwave and Wireless Components Letters*, vol. 19, no. 11, pp. 719-721, Nov. 2009.
- [137] P. Andricciola and H. P. Tuinhout, "The temperature dependence of mismatch in deep-submicrometer bulk MOSFETs," *IEEE Electron Device Letters*, vol. 30, no. 6, pp. 690-692, June 2009.

## APPENDIX A

## OTA LINEARIZATION: VOLTERRA SERIES ANALYSIS

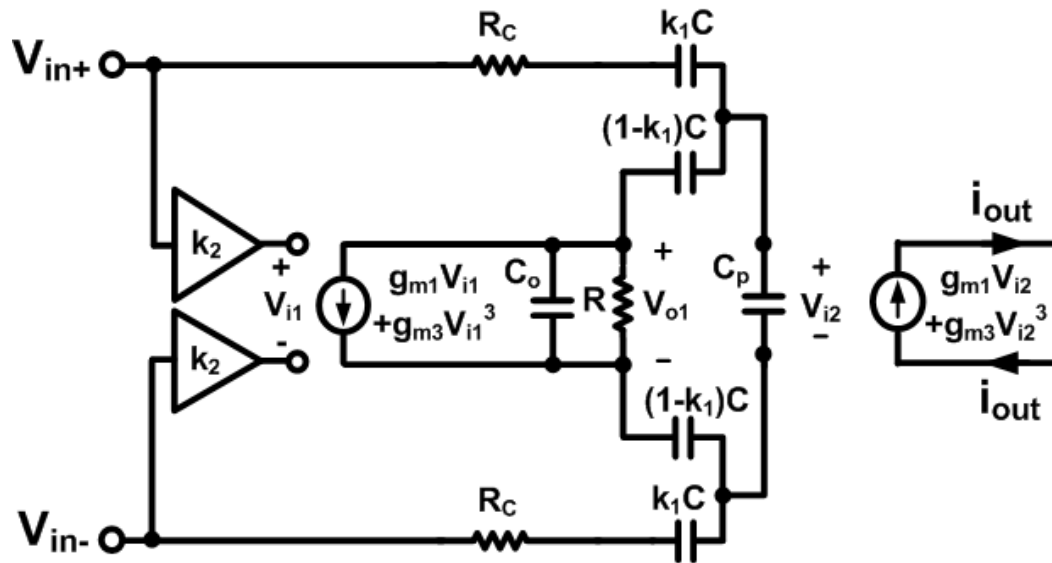


Fig. 114. Nonlinear model for differential attenuation-predistortion cancellation.

In this appendix, the optimum compensation resistor value for linearization at high frequencies is derived with Volterra series analysis [52]. Employing a 3<sup>rd</sup>-order model of transconductor nonlinearity, the simplified model of the proposed attenuation-predistortion linearization technique is shown in Fig. 114. In this analysis,  $g_{m1}$  represents the linear transconductance and  $g_{m3}$  the third-order component. Resistor ( $R_c$ ) compensates for high-frequency linearity degradation by equalizing the delays in the main and auxiliary paths. The differential voltage  $V_{i2}(t)$  at the input of the main OTA is given by

$$V_{i2}(t) = -\left(g_{m1}k_2V_{in}(t) + g_{m3}[k_2V_{in}(t)]^3\right) \cdot \frac{R \cdot (1-k_1)}{1+2C_p/C} \cdot \frac{1+j\omega Ck_1R_c}{1+j\omega b-c\omega^2} \\ + V_{in}(t) \cdot \frac{k_1}{1+2C_p/C} \cdot \frac{1+j\omega C(1-k_1)R/2+j\omega C_oR}{1+j\omega b-c\omega^2},$$

where: (41)

$$b = \frac{C(k_1/2)(1-k_1)(R+2R_c) + 2k_1C_pR_c + (1-k_1)C_pR}{1+2C_p/C} + C_oR \\ c = \frac{k_1(1-k_1)CC_pRR_c + CC_o k_1(1-k_1)RR_c + 2k_1C_pC_oR_cR}{1+2C_p/C}.$$

Following the same analysis as in Section III.2.3 but taking the parasitic capacitances  $C_p$  and  $C_o$  into account, the conditions for distortion cancellation at low frequencies are:

$$\frac{g_{m1} \cdot R \cdot (1-k_1)}{1+2C_p/C} = 1, \quad k_2 = \frac{k_1/2}{1+2C_p/C}.$$
(42)

With the above provisions, the output current of the main OTA after algebraic simplifications is:

$$i_{out}(t) = g_{m1}V_{i2}(t) + g_{m3}[V_{i2}(t)]^3 \\ \approx g_{m1}V_{in}(t) \cdot \frac{(k_1/2)}{1+2C_p/C} \cdot \frac{1+j\omega C((1-k_1)R-k_1R_c)+j2\omega C_oR}{1+j\omega b-c\omega^2} \\ - g_{m3} \left( \frac{k_1V_{in}(t)/2}{1+2C_p/C} \right)^3 \cdot \frac{1+j\omega Ck_1R_c}{1+j\omega b-c\omega^2} \\ + g_{m3} \left( \frac{k_1V_{in}(t)/2}{1+2C_p/C} \cdot \frac{1+j\omega C((1-k_1)R-k_1R_c)+j2\omega C_oR}{1+j\omega b-c\omega^2} \right)^3.$$
(43)

Assuming weakly nonlinear operation based on condition iii) in Section III.2.3 and that the signal can be expressed as a sum of sinusoids with incommensurate frequencies,

the harmonic input method can be applied to calculate the Volterra series coefficients [52] and theoretically demonstrate the nonlinearity cancellation with the proposed scheme. Taking a single input  $V_{in}(t) = e^{j\omega t}$  and substituting into (43) to express the linear transfer function  $H_1$ :

$$H_1 = g_{m1} \cdot \frac{(k_1 / 2)}{1 + 2C_p / C} \cdot \frac{1 + j\omega C((1 - k_1)R - k_1 R_c) + 2j\omega C_o R}{1 + j\omega b - c\omega^2} \quad (44)$$

Selecting  $V_{in}(t) = e^{j\omega_1 t} + e^{j\omega_2 t} + e^{j\omega_3 t}$  and making the appropriate substitutions for calculation of the third-order transfer function ( $H_3$ ) yields the following equality after expansion and omission of all terms that do not contain the  $\exp(j\omega_1 t + j\omega_2 t + j\omega_3 t)$  factor relevant to  $H_3$ :

$$\begin{aligned} H_3(\omega_1, \omega_2, \omega_3) = & \\ & g_{m3} \left( \frac{k_1 / 2}{1 + 2C_p / C} \right)^3 \left( \frac{1 + j\omega_1 C((1 - k_1)R - k_1 R_c) + 2j\omega_1 C_o R}{1 + j\omega_1 b - c\omega_1^2} \right) \\ \times & \left( \frac{1 + j\omega_2 C((1 - k_1)R - k_1 R_c) + 2j\omega_2 C_o R}{1 + j\omega_2 b - c\omega_2^2} \right) \left( \frac{1 + j\omega_3 C((1 - k_1)R - k_1 R_c) + 2j\omega_3 C_o R}{1 + j\omega_3 b - c\omega_3^2} \right) \quad (45) \\ - & g_{m3} \left( \frac{k_1 / 2}{1 + 2C_p / C} \right)^3 \frac{1 + j(\omega_1 + \omega_2 + \omega_3)Ck_1 R_c}{1 + j(\omega_1 + \omega_2 + \omega_3)b - c(\omega_1 + \omega_2 + \omega_3)^2} \end{aligned}$$

The amplitude of the third harmonic distortion (HD3) current due to a sinusoidal input signal  $V_{in} \sin(\omega t)$  is given by

$$\begin{aligned} i_{o3} = \frac{1}{4} V_{in}^3 H_3(\omega, \omega, \omega) = & \\ \frac{1}{4} g_{m3} \left( \frac{V_{in} k_1 / 2}{1 + 2C_p / C} \right)^3 \left( \frac{1 + j\omega C((1 - k_1)R - k_1 R_c) + 2j\omega C_o R}{1 + j\omega b - c\omega^2} \right)^3 & \quad (46) \\ - \frac{1}{4} g_{m3} \left( \frac{V_{in} k_1 / 2}{1 + 2C_p / C} \right)^3 \frac{1 + j3\omega Ck_1 R_c}{1 + j3\omega b - 9c\omega^2} & \end{aligned}$$

Elimination of HD3 requires that  $i_{o3} = 0$ , hence

$$\frac{1 + j\omega C((1 - k_1)R - k_1 R_c) + 2j\omega C_o R}{1 + j\omega b - c\omega^2} = \sqrt[3]{\frac{1 + j3\omega C k_1 R_c}{1 + j3\omega b - 9c\omega^2}} \quad (47)$$

The cubic root in (47) can be approximated with  $\sqrt[3]{1+x} \approx 1+x/3$  for  $x \ll 1$ . Thus,

$$\begin{aligned} \frac{1 + j\omega C((1 - k_1)R - k_1 R_c) + 2j\omega C_o R}{1 + j\omega b - c\omega^2} &\approx \frac{1 + j\omega C k_1 R_c}{1 + j\omega b - 3c\omega^2} \\ \Rightarrow R_c &\approx \frac{(1 - k_1) + 2C_o / C}{2k_1} R \quad \text{to cancel HD3} \end{aligned} \quad (48)$$

For a two-tone input signal of the form  $V_{in1} \sin(\omega_1 t) + V_{in2} \sin(\omega_2 t)$ , the IM3 current can be determined with Volterra series [52] according to the following equation:

$$\begin{aligned} i_{IM3} &= \frac{3}{4} V_{in1}^2 V_{in2} H_3(\omega_1, \omega_1, -\omega_2) = \\ &g_{m3} \left( \frac{k_1 / 2}{1 + 2C_p / C} \right)^3 \left( 3V_{in1}^2 V_{in2} / 4 \right) \left( \frac{1 + j\omega_1 C((1 - k_1)R - k_1 R_c) + 2j\omega_1 C_o R}{1 + j\omega_1 b - c\omega_1^2} \right)^2 \\ &\times \left( \frac{1 - j\omega_2 C((1 - k_1)R - k_1 R_c) - 2j\omega_2 C_o R}{1 - j\omega_2 b - c\omega_2^2} \right) \\ &- g_{m3} \left( \frac{k_1 / 2}{1 + 2C_p / C} \right)^3 \left( 3V_{in1}^2 V_{in2} / 4 \right) \frac{1 + j(2\omega_1 - \omega_2) C k_1 R_c}{1 + j(2\omega_1 - \omega_2) b - c(2\omega_1 - \omega_2)^2} \end{aligned} \quad (49)$$

Simplifying  $i_{IM3}$  for two intermodulation tones that are close together ( $\omega_1 \approx \omega_2 \approx 2\omega_1 - \omega_2$ ) yields:

$$\begin{aligned}
i_{IM3} &\approx g_{m3} \left( \frac{k_1/2}{1+2C_p/C} \right)^3 \left( 3V_{in1}^2 V_{in2} / 4 \right) \left( \frac{1+j\omega_1 C((1-k_1)R - k_1 R_c) + 2j\omega_1 C_o R}{1+j\omega_1 b - c\omega_1^2} \right)^2 \\
&\times \left( \frac{1-j\omega_1 C((1-k_1)R - k_1 R_c) - 2j\omega_1 C_o R}{1-j\omega_1 b - c\omega_1^2} \right) \\
&- g_{m3} \left( \frac{k_1/2}{1+2C_p/C} \right)^3 \left( 3V_{in1}^2 V_{in2} / 4 \right) \frac{1+j\omega_1 C k_1 R_c}{1+j\omega_1 b - c\omega_1^2} \\
\Rightarrow R_c &\approx \frac{(1-k_1) + 2C_o/C}{2k_1} R \quad \text{for } i_{IM3} \approx 0
\end{aligned} \tag{50}$$

## APPENDIX B

## OTA LINEARIZATION: ADVANCED PHASE COMPENSATION

Fig. 115a depicts a model for an OTA in integrator configuration where  $r_o$  represents the OTA output impedance and  $Gm_{(j\omega)}$  the transconductance that changes with frequency due to internal parasitic poles. Both nonidealities cause deviations from ideal integration on the load capacitor  $C$ . The following analysis shows that the linearization introduces an additional pole, which can be cancelled by adding resistor  $R_s$  in series with the load capacitor as in the conventional case [54].

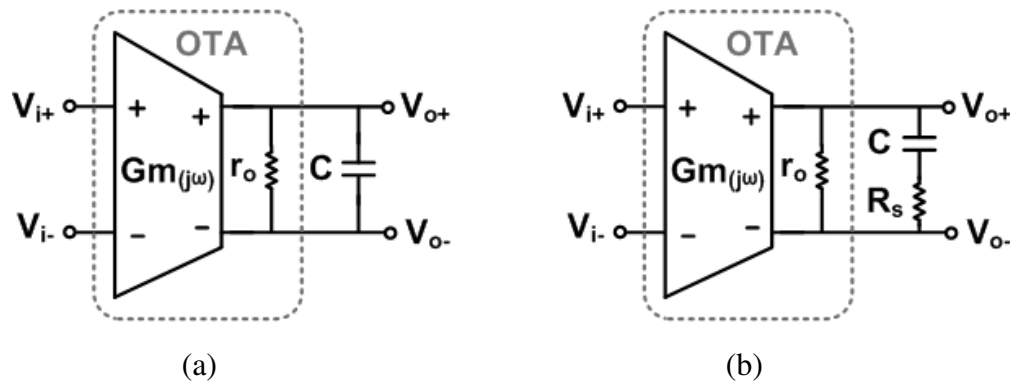


Fig. 115. OTA model with additional nonidealities.

(a) Standard configuration, (b) configuration to compensate for the internal pole from the attenuation-predistortion linearization; where  $Gm_{(\omega)}$ ,  $r_o$ , and  $C$  are the frequency-dependent transconductance, finite output impedance, and load capacitor, respectively.

Let  $\omega_o = 1 / (r_o C)$  be the dominant pole of the integrator configuration and  $\omega_l$  be the internal parasitic pole of the OTA with the lowest frequency. If  $\omega_l \gg \omega_o$ , then the transfer function of the configuration in Fig. 115a is:

$$\frac{V_{o+} - V_{o-}}{V_{i+} - V_{i-}} = \frac{Gm_{(0)} \cdot r_o}{1 + s \cdot r_o C} \approx \frac{Gm_{(0)}}{s \cdot C}, \quad (51)$$

where  $s = j\omega$  and the approximation implies:  $\omega_o \ll \omega \ll \omega_l$ . When using attenuation-predistortion linearization at high frequencies, the additional pole  $\omega_c$  formed by  $R_c$  and  $C_p$  in Fig. 12 is not negligible in all designs. Hence, the integrator has the following transfer function:

$$\frac{V_{o+} - V_{o-}}{V_{i+} - V_{i-}} = \frac{Gm_{(0)} \cdot r_o}{1 + s \cdot r_o C} \cdot \frac{1}{1 + s / \omega_c}. \quad (52)$$

To avoid impact of  $\omega_c$ , a series resistor  $R_s$  can be added to the load capacitor  $C$  as visualized in Fig. 115b, resulting in the new expression for the transfer function:

$$\frac{V_{o+} - V_{o-}}{V_{i+} - V_{i-}} = \frac{Gm_{(0)} \cdot r_o (1 + s \cdot R_s C)}{(1 + sC[r_o + R_s]) \cdot (1 + s / \omega_c)} \approx \frac{Gm_{(0)} \cdot r_o (1 + s \cdot R_s C)}{(1 + s \cdot r_o C) \cdot (1 + s / \omega_c)}, \quad (53)$$

where  $R_s \ll r_o$  is assumed in the approximation. In the range  $\omega_o \ll \omega \ll \omega_l$ , the following condition to nullify the impact of the linearization can be identified by comparing (53) and (51):

$$R_s = \frac{1}{C \cdot \omega_c}, \quad (54)$$

The effect of  $\omega_c$  from the linearization on the key parameters of a biquad section can be assessed by examining the bandpass (BP) case (Fig. 116). The center frequency ( $\omega_{oi}$ ), bandwidth ( $BW_i$ ), and quality factor ( $Q_i$ ) with ideal OTAs are:

$$\omega_{oi} = \sqrt{\frac{Gm_1 Gm_2}{C_A C_B}}, \quad (55)$$



$$BW_i = \frac{Gm_3}{C_B}, \quad (56)$$

$$Q_i = \sqrt{\frac{Gm_1 Gm_2}{(Gm_3)^2} \cdot \frac{C_B}{C_A}}. \quad (57)$$

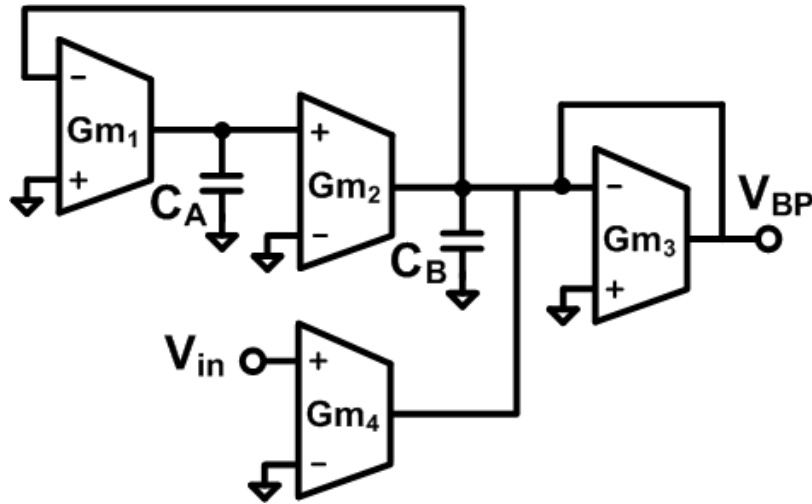


Fig. 116. Single-ended equivalent block diagram of a bandpass biquad.

Substituting  $Gm(s) = Gm / (1 + s/\omega_c)$  for each Gm in the BP transfer function yields the following equation for a linearized BP section:

$$H_{BP}(s) = \frac{V_{BP}}{V_{in}} = \frac{N(s)}{D(s)} = \frac{\frac{Gm_4}{C_B} \cdot \frac{1}{1+s/\omega_{c4}} \cdot s}{s^2 + \frac{Gm_3}{C_B} \cdot \frac{1}{1+s/\omega_{c3}} \cdot s + \frac{Gm_1 Gm_2}{C_A C_B} \cdot \frac{1}{1+s/\omega_{c1}} \cdot \frac{1}{1+s/\omega_{c2}}}. \quad (58)$$

Letting  $Gm = Gm_1 = Gm_2$  and  $\omega_c = \omega_{c1} = \omega_{c2}$  for simplicity and given that  $\omega_{oi} < \omega_c$ , it can be shown that the center frequency ( $\omega_{on}$ ) of the linearized BP biquad can be approximated as:

$$\omega_{on} \approx \omega_{oi} . \quad (59)$$

The denominator of the linearized BP transfer function in (58) can be approximated as follows:

$$\begin{aligned} D(s) &\approx s^2 + \frac{Gm_3}{C_B} \cdot (1 - s / \omega_{c3}) \cdot s + \frac{Gm^2}{C_A C_B} \cdot (1 - 2s / \omega_c) \\ &\approx s^2 + \left( \frac{Gm_3}{C_B} - \frac{2Gm^2}{\omega_c C_A C_B} \right) \cdot s + \frac{Gm^2}{C_A C_B} , \end{aligned} \quad (60)$$

where the second approximation is valid when  $\omega \ll \omega_{c3}$ . From (60),  $BW_n$  and the quality factor ( $Q_n$ ) with linearized OTAs can be written in terms of the above ideal expressions as follows:

$$BW_n \approx \frac{Gm_3}{C_B} - \frac{2Gm^2}{\omega_c C_A C_B} = BW_i - \frac{2\omega_{oi}^2}{\omega_c} = BW_i \left( 1 - \frac{2\omega_{oi}^2}{\omega_c \cdot BW_i} \right) , \quad (61)$$

$$Q_n = \frac{\omega_{on}}{BW_n} \approx \frac{\omega_{oi}}{BW_i \left( 1 - \frac{2\omega_{oi}^2}{\omega_c \cdot BW_i} \right)} . \quad (62)$$

Equation (62) shows that the quality factor error from linearization increases with the ratio of  $\omega_{oi}^2 / (\omega_c \cdot BW_i)$ , where:  $\omega_{oi} \approx \omega_{on}$ . Furthermore, stability requires:

$$\frac{2\omega_{oi}^2}{\omega_c \cdot BW_i} < 1 . \quad (63)$$

The parameter changes in (61)-(62) can be incorporated into the design of linearized biquads by altering the transconductance and capacitor values accordingly. Alternatively, the effects of the linearization can be canceled as described next.

Using series resistors  $R_{sA}$  and  $R_{sB}$  with  $C_A$  and  $C_B$  to compensate for the phase shift from the linearization as described above and shown in Fig. 117, the corresponding zeros are introduced in the denominator:

$$D(s) = s^2 + \frac{Gm_3}{C_B} \cdot \frac{1+s/\omega_{zB}}{1+s/\omega_{c3}} \cdot s + \frac{Gm^2}{C_A C_B} \cdot \frac{(1+s/\omega_{zA})(1+s/\omega_{zB})}{(1+s/\omega_c)^2}, \quad (64)$$

where  $\omega_{zA} = 1/(R_{sA}C_A) = \omega_{zB} = 1/(R_{sB}C_B) = \omega_z = \omega_c$ . Using the same approximations as in (59)-(63), the compensated center frequency ( $\omega_{cn}$ ) and bandwidth ( $BW_{cn}$ ) become:

$$\omega_{cn} \approx \sqrt{\frac{Gm^2}{C_A C_B} \cdot (1 - 2s/\omega_c) \cdot (1 + 2s/\omega_z)} \approx \sqrt{\frac{Gm^2}{C_A C_B} \cdot (1 - \frac{4s^2}{\omega_c \cdot \omega_z})} \approx \sqrt{\frac{Gm^2}{C_A C_B}}, \quad (65)$$

which is equivalent to  $\omega_i$  as a result of the last simplification step ( $4 \cdot \omega^2 \ll \omega_c \cdot \omega_z$ );

$$BW_{cn} \approx \frac{Gm_3}{C_B} \cdot [1 + (\frac{1}{\omega_z} - \frac{1}{\omega_{c3}}) \cdot s - \frac{s^2}{\omega_z \cdot \omega_{c3}}] \approx \frac{Gm_3}{C_B} \cdot [1 + j\omega \cdot (\frac{1}{\omega_z} - \frac{1}{\omega_{c3}})]. \quad (66)$$

Note, a small bandwidth error remains after compensation due to the difference between  $\omega_z$  and  $\omega_{c3}$  because  $\omega_{zA}(R_{sA}, C_A)$  and  $\omega_{zB}(R_{sB}, C_B)$  are optimized to cancel  $\omega_{c1}$  and  $\omega_{c2}$  of  $Gm_1$  and  $Gm_2$ , respectively. Thus, the pole  $\omega_{c3}$  is only partially cancelled if  $Gm_1 \neq Gm_3$ . Nevertheless, the second term in (66) has a small effect in the typical case and  $BW_{cn} \approx BW_i$  since  $\omega \ll |1/\omega_z - 1/\omega_{c3}|^{-1}$ .

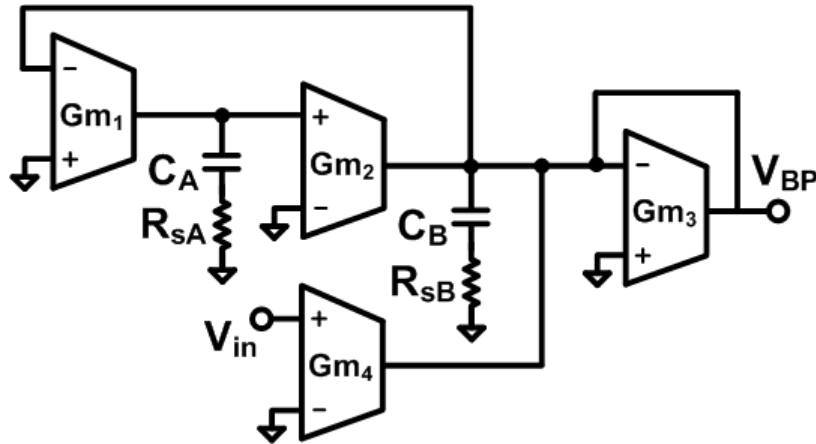


Fig. 117. Single-ended diagram of a bandpass biquad with phase compensation.

The linearized OTAs described in Section III.3.1 were employed in a BP filter (Fig. 117) with  $f_0 = 100\text{MHz}$ ,  $Gm_3 = Gm_4 = Gm/2$ , and  $Gm = Gm_1 = Gm_2$  for simplicity (implying  $\omega_c = \omega_{c1} = \omega_{c2}$ ). Series resistors  $R_{sA}$  and  $R_{sB}$  with  $C_A$  and  $C_B$  compensate for the phase shift from the linearization by creating zeros  $\omega_{zA}$  and  $\omega_{zB}$ :  $\omega_{zA} = 1/(R_{sA}C_A) = \omega_{zB} = 1/(R_{sB}C_B) = \omega_z = \omega_c$ . A small BW error remains after compensation due to the difference between  $\omega_z$  and  $\omega_{c3}$  of  $Gm_3$  because  $\omega_{zA}(R_{sA}, C_A)$  and  $\omega_{zB}(R_{sB}, C_B)$  are optimized to cancel  $\omega_{c1}$  and  $\omega_{c2}$  of  $Gm_1$  and  $Gm_2$ , respectively. Thus, the pole  $\omega_{c3}$  is only partially cancelled since  $Gm_1 \neq Gm_3$ . Nevertheless, the effect is small in the typical case ( $\omega \ll \omega_{c3}$ ). This BP filter achieves simulated IM3 of  $-72.0\text{dB}$  evaluated after an additional output buffer ( $Gm$ ). Fig. 118 contains simulated plots of the frequency responses for different values of  $R_s$  from this example BP filter design. The plots show how the adjustment of  $R_s = R_{sA} = R_{sB} \cdot (C_B/C_A)$  during the design allows tuning of the quality factor to  $\sim 4$  with  $R_s = 7\Omega$  in this case, while  $f_0$  does not change significantly.

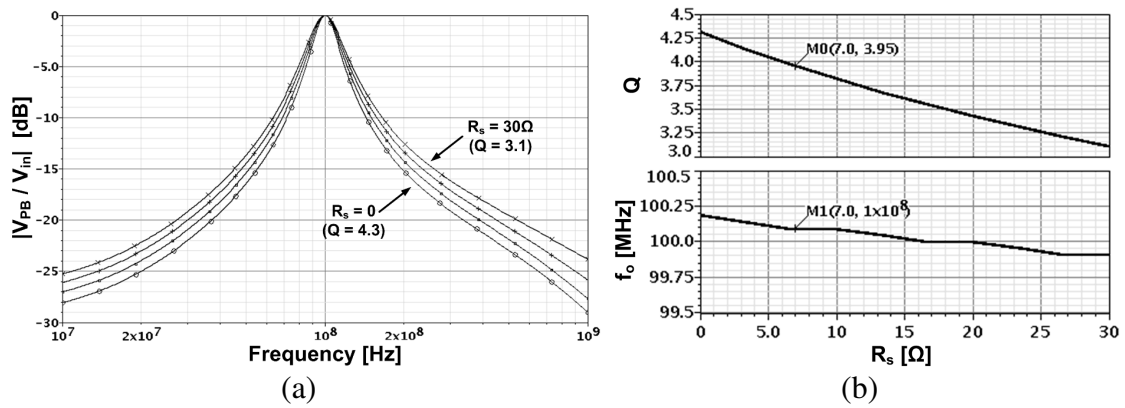


Fig. 118. BP filter simulations with different  $R_s$  values for phase compensation.  
 (a) Frequency responses, (b) quality factor and center frequency;  
 where  $R_s = R_{sA} = R_{sB} \cdot (C_B / C_A)$ .

## APPENDIX C

### OTA LINEARIZATION WITHOUT POWER BUDGET INCREASE

Attenuation-predistortion linearization offers the means to improve the linearity of a given OTA while preserving its AC characteristics without design changes in the OTA core, which is achieved at the expense of increased power, noise, and layout area. Another option is to redesign the two OTAs in the linearization scheme using half of the power in order to meet the same power budget as the original OTA. But, that approach is associated with a reduction of the OTA bandwidth as delineated in this appendix.

To accomplish linearization with equal power budget, the currents  $I_b$  and  $I_{b1}$  in Fig. 14 can be reduced by 50%, which requires increasing the W/L ratios of the transistors in the core ( $M_c$ ) to obtain the same transconductance as before. Thus, the saturation voltage  $V_{DSAT}$  of  $M_c$  becomes approximately half of the initial value. Furthermore, the ratio of transconductance to parasitic capacitance (i.e.  $f_T$ ) of both OTAs in the linearization scheme reduces due to the bias current decrease and width increase for  $M_c$ . Gain vs. frequency simulations of the linearized OTA (50% power reduction in each path) and the reference OTA revealed that the linearization with equal power reduces the effective 3dB bandwidth from 2.49GHz to 1.09GHz with 50 $\Omega$  load. Table XIX summarizes the key results from simulating the linearized OTA in comparison to the reference OTA with identical total power. High linearity through distortion cancellation ( $IM3 \approx -77dB$ ) is achievable, but limited to lower frequencies. Despite of this, the results indicate that higher FOM (see Table V on page 62) can be achieved with low-frequency linearization compared to the linearization with doubled power consumption.

Table XIX. Simulated comparison: OTA linearization without power increase

OTA Type	$V_{DSAT}$ of Input Differential Pair ( $M_c$ )	$f_{3db}$ with $50\Omega$ Load	Input-Referred Noise	Power	IM3 ( $V_{in} = 0.2 V_{p-p}$ )	Normalized  FOM * (at $f_{max}$ )
<b>Reference</b> (input attenuation = 1/3)	90mV	2.49GHz	9.7nV/ $\sqrt{Hz}$	2.6mW	-53.1dB at $f_{max} = 350MHz$ (-53.2 dB at 100MHz)	57.2
<b>Linearized</b> (attenuation = 1/3 & compensation)	54mV	1.09GHz	14.3nV/ $\sqrt{Hz}$	2.6mW	-77.1dB at $f_{max} = 100MHz$	119.2

\* See Table V for details.

**APPENDIX D**

**TEMPERATURE SENSING ANALYSIS:**

**RELATIONSHIP BETWEEN CIRCUIT NONLINEARITIES**

**AND DC TEMPERATURE**

The main purpose of the analysis in this appendix is to show that a minimum temperature point is sensed near a MOS device as the RF power of an applied signal is swept. When a sinusoidal input voltage  $x(t)$  with amplitude  $X$  at frequency  $\omega$  excites a weakly nonlinear MOS device and creates a current  $y(t)$  that can be expressed by a power series with coefficients  $\alpha_0, \alpha_1, \alpha_2, \dots$ ; then the signals can be written as

$$x(t) = X \cos \omega t , \quad (67)$$

$$y(t) = \alpha_0 + \alpha_1 x(t) + \alpha_2 x^2(t) + \alpha_3 x^3(t) + \dots . \quad (68)$$

The effect of the bias current  $\alpha_0$  is removed via calibration before the application of the signal, which avoids interference with the 1-dB compression point characterization. Thus, the signal-dependent current without  $\alpha_0$  can be expressed as

$$y_{sig}(t) = y_{sig}(t)|_{DC} + y_{sig}(t)|_{AC} ; \quad (69)$$

where:

$$y_{sig}(t)|_{DC} \approx \frac{\alpha_2}{2} X^2 , \quad (70)$$

$$\begin{aligned} y_{sig}(t)|_{AC} &= (\alpha_1 X + \frac{3\alpha_3}{4} X^3) \cos \omega t \\ &+ \frac{\alpha_2}{2} X^2 \cos 2\omega t + \frac{\alpha_3}{4} X^3 \cos 3\omega t + \dots . \end{aligned} \quad (71)$$



A conventional 1-dB compression point characterization is a measure of the third-order distortion due to  $\alpha_3$  at frequency  $\omega$ , for which the input amplitude approximation is given by

$$X_{1dB} = \sqrt{\left| (4/3) \cdot (10^{-1/20} - 1) \frac{\alpha_1}{\alpha_3} \right|}. \quad (72)$$

With the homodyne temperature sensing approach, the linearity is assessed from indirect measurement of the DC power, giving rise to the implications analyzed below.

When a signal is applied, the AC amplitude and the signal-dependent part of the drain-source voltage's DC component resulting from  $y_{sig}(t)|_{DC}$  scale proportionally to the RMS drain-source voltage change. Let  $K$  represent this load-dependent proportionality factor. In the transistors of the CUT, the AC drain-source voltage is  $180^\circ$  out of phase with the drain current  $y_{sig}(t)|_{AC}$ . Thus, a simplified approximation for the signal-dependent drain-source voltage around the 1-dB compression point is:

$$\begin{aligned} v_{sig[1dB]}(t) &= K \cdot y_{sig}(t)|_{DC[1dB]} - K \cdot y_{sig}(t)|_{AC[1dB]} \\ &\approx K_{DC} - K_{AC} \cdot \cos(\omega t) \end{aligned} \quad ; \quad (73)$$

where:

$$K_{DC} = K \cdot \frac{\alpha_2}{2} X_{1dB}^2, \quad (74)$$

$$K_{AC} = K \cdot (\alpha_1 X_{1dB} + (3/4) \cdot \alpha_3 X_{1dB}^3). \quad (75)$$

Here, the analysis is simplified by the omission of load-dependent nonlinearities and by disregarding components at  $2\omega$ ,  $3\omega$ , and higher harmonics. More complex expressions and incorporation of electro-thermal coupling would be required for more accurate analytical estimates. Nevertheless, the approximations under the assumed

conditions give insights into the key characteristics of the power that causes the temperature change:

$$p(t) = v_{sig[1dB]}(t) \cdot y_{sig}(t). \quad (76)$$

Notice that (76) only represents the scaled signal-dependent power components. Without calibration, the DC bias current ( $\alpha_0$ ) would have to be included and multiplied with a different factor (unrelated to  $K$ ). But,  $\alpha_0$  was dropped from (68) because its contribution is nullified by the calibration step. After substituting (69)-(71) and (73)-(75) into (76), using the trigonometric identity  $\cos^2(x) = \frac{1}{2}[1 + \cos(2x)]$ , and dropping all remaining AC terms based on the low-pass filter characteristics of the thermal coupling (under condition:  $\omega \gg 2\pi \cdot 10\text{KHz}$ ); the DC power component that causes the measured DC temperature change is obtained as follows:

$$P_{DC \rightarrow \Delta T} \approx K_{DC} \cdot \left(\frac{\alpha_2}{2} X^2\right) - (1/2) \cdot K_{AC} \cdot (\alpha_1 X - (3/4) \cdot |\alpha_3| X^3). \quad (77)$$

The approximation in (77) assumes weakly nonlinear operation, negligible higher-order distortion components, and the typical case in which  $\alpha_0$ - $\alpha_2$  are positive but  $\alpha_3$  is negative. From (77), it can be observed that second-order nonlinearity creates a measurement offset and that the DC component reaches a minimum as  $X$  is swept to evaluate the 1-dB compression property. This theoretical minimum can be derived by taking the derivative of (77), equating the resulting expression to zero, and solving for the amplitude:

$$X_{\min} = \frac{-\alpha_2 (K_{DC} / K_{AC}) + \sqrt{\alpha_2^2 (K_{DC} / K_{AC})^2 + (9/4) \cdot \alpha_1 |\alpha_3|}}{(9/4) \cdot |\alpha_3|}. \quad (78)$$

Equation (78) gives insights into the minimum temperature point characteristics, but it is important to note that it is only a rough approximation due to the aforementioned assumptions. In the absence of thermal coupling to other devices, a relative comparison of (78) and (72) allows to estimate the fixed input power shift (in decibels) between the minimum DC power/temperature point and the 1dB-compression point:

$$shift_{\min[1dB]} = 10 \cdot \log(X_{\min}^2 / X_{1dB}^2). \quad (79)$$

The above equations show that the 1-dB compression point can be inferred from the DC power dissipation monitored by the temperature sensor as long as the second-order nonlinearity is accounted for during simulations. For the nonlinearity coefficients of the example CUT, (79) predicts a 4.73dB shift. Based on this shift with respect to the simulated 0.5dBm 1-dB compression point, the minimum point is expected with 5.23dBm input power. However,  $P_m$  in Fig. 55 has a minimum at 2.6dBm, where the error is mainly caused by the aforementioned idealizations and by deviations from the weak nonlinearity assumption that causes approximately 15% error in (72) alone. Furthermore, the thermal coupling of devices in the CUT affects the minimum temperature point on the x-axis, which follows the superposition principle (e.g. the power of all devices in Fig. 55 results in the combined temperature curve ( $T_s$ ) at the sensing PNP device in Fig. 56). Therefore, the electro-thermal simulation method presented in this dissertation provides a more reliable estimate for the shift, which was around 0.1dB in simulations and 0.5dB in measurements. The difference is affected by process variations as well as electro-thermal modeling inaccuracies, which could cause up to  $\pm 0.6$ dB uncertainty for this CUT that was added to the measurement error.

## VITA

Marvin Olufemi Onabajo was born in Lengerich, Germany in 1982. He received a B.S. degree (*summa cum laude*) in electrical engineering from The University of Texas at Arlington in 2003; as well as the M.S. and Ph.D. degrees in electrical engineering from Texas A&M University in 2007 and 2011, respectively.

During his final year at UT Arlington he worked in the Analog and Mixed-Signal IC group in affiliation with the National Science Foundation's Research Experiences for Undergraduates program. From 2004 to 2005, he was Electrical Test/Product Engineer at Intel Corp. in Hillsboro, Oregon. He joined the Analog and Mixed-Signal Center at Texas A&M University in 2005, where he was engaged in research projects involving analog built-in testing, data converters, and on-chip temperature sensors for thermal monitoring. In the spring 2011 semester, he worked as a Design Engineering Intern in the Broadband RF/Tuner Development group at Broadcom Corp. in Irvine, California. He can be contacted through the Department of Electrical and Computer Engineering, Attn: Jose Silva-Martinez, Texas A&M University, 214 Zachry Engineering Center, TAMU 3128, College Station, TX 77843-3128.