

ON-CHIP ANALOG CIRCUIT DESIGN USING BUILT-IN SELF-TEST AND AN
INTEGRATED MULTI-DIMENSIONAL OPTIMIZATION PLATFORM

A Dissertation

by

CONGYIN SHI

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

DOCTOR OF PHILOSOPHY

Chair of Committee, Edgar Sánchez-Sinencio
Committee Members, Kamran Entesari
Peng Li
Halit Üster
Head of Department, Miroslav M. Begovic

August 2017

Major Subject: Electrical Engineering

Copyright 2017 Congyin Shi

ABSTRACT

Nowadays, the rapid development of system-on-chip (SoC) market introduces tremendous complexity into the integrated circuit (IC) design. Meanwhile, the IC fabrication process is scaling down to allow higher density of integration but makes the chips more sensitive to the process-voltage-temperature (PVT) variations. A successful IC product not only imposes great pressure on the IC designers, who have to handle wider variations and enforce more design margins, but also challenges the test procedure, leading to more check points and longer test time. To relax the designers' burden and reduce the cost of testing, it is valuable to make the IC chips able to test and tune itself to some extent.

In this dissertation, a fully integrated in-situ design validation and optimization (VO) hardware for analog circuits is proposed. It implements in-situ built-in self-test (BIST) techniques for analog circuits. Based on the data collected from BIST, the error between the measured and the desired performance of the target circuit is evaluated using a cost function. A digital multi-dimensional optimization engine is implemented to adaptively adjust the analog circuit parameters, seeking the minimum value of the cost function and achieving the desired performance. To verify this concept, study cases of a 2nd/4th active-RC band-pass filter (BPF) and a 2nd order Gm-C BPF, as well as all BIST and optimization blocks, are adopted on-chip.

Apart from the VO system, several improved BIST techniques are also proposed in this dissertation. A single-tone sinusoidal waveform generator based on a finite-impulse-response (FIR) architecture, which utilizes an optimization algorithm to enhance its spur free dynamic range (SFDR), is proposed. It achieves an SFDR of 59 to 70 dBc from 150 to 850 MHz after the optimization procedure. A low-distortion

current-steering two-tone sinusoidal signal synthesizer based on a mixing-FIR architecture is also proposed. The two-tone synthesizer extends the FIR architecture to two stages and implements an up-conversion mixer to generate the two tones, achieving better than -68 dBc IM3 below 480 MHz LO frequency without calibration.

Moreover, an on-chip RF receiver linearity BIST methodology for continuous- and discrete-time hybrid baseband chain is proposed. The proposed receiver chain implements a charge-domain FIR filter to notch the two excitation signals but expose the third order intermodulation (IM3) tones. It simplifies the linearity measurement procedure—using a power detector is enough to analyze the receiver’s linearity.

Finally, a low cost fully digital built-in analog tester for linear-time-invariant (LTI) analog blocks is proposed. It adopts a time-to-digital converter (TDC) to measure the delays corresponded to a ramp excitation signal and is able to estimate the pole or zero locations of a low-pass LTI system.

DEDICATION

To my beloved Mother and Father.

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude towards my advisor, Dr. Edgar Sánchez-Sinencio, for his strong support, enduring help, smart guidance, and constructive critique during the entire course of my doctoral research. I would also like to thank Dr. Jiang Hu for his valuable advice on my research of optimization. I would further appreciate the support from Dr. Kamran Entesari, Dr. Peng Li, and Dr. Halit Üster, who are serving on my dissertation committee. Moreover, I would like to thank Ms. Ella Gallagher, Ms. Tammy Carda, and Ms. Katharine Bryan for facilitating event organizations and paperwork preparations on many occasions.

It has been a great pleasure and study experience to work with the bright minds we have here, in the Department of Electrical and Computer Engineering. Firstly, I would like to express my gratitude to Jiafan Wang, who is Dr. Jiang Hu's PhD student, for his enlightened idea and indispensable work in the analog circuit optimization projects. I would also like to appreciate the help from Hatem Osman, Adriana Sanabria and Sanghoon Lee for their active discussion and assistance in the optimization project. Furthermore, I would like to thank my seniors who have selflessly taught me knowledge in both course works and research projects, namely Jiayi Jin, Joselyn Torres, Salvador Carreon, Jorge Zarate, and Adrian Colli. Many thanks also go out to fellow AMSC members for helpful conversations regarding research projects and paper reviews, especially to Xiaosen Liu, Heng Zhang, and Mohamed Mostafa. I also thank Intel, Texas Instruments, Qualcomm, and Silicon Labs for their financial support. I would also like to acknowledge the sponsorship of chip fabrications by MOSIS and Global Foundries.

In closing the acknowledgments, I am grateful for the encouragement and under-

standing from my family during the years of my academic study. My parents have unconditionally supported me in every which way I needed ever since my very first school day.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supported by a dissertation committee consisting of Dr. Edgar Sánchez-Sinencio, Dr. Kamran Entesari and Dr. Peng Li of the Electrical Engineering at Texas A&M University and Dr. Halit Üster of the Department of Engineering Management, Information, and Systems at Southern Methodist University.

All the work conducted for the dissertation was completed by the student independently.

Funding Sources

Graduate study and research was supported by fellowships from Silicon Labs, Texas Instruments and Qualcomm. Chip fabrication was supported by MOSIS and Global Foundries.

TABLE OF CONTENTS

	Page
ABSTRACT	ii
DEDICATION	iv
ACKNOWLEDGEMENTS	v
CONTRIBUTORS AND FUNDING SOURCES	vii
TABLE OF CONTENTS	viii
LIST OF FIGURES	xii
LIST OF TABLES	xx
1. INTRODUCTION	1
1.1 On-chip Analog Circuit Built-in Self-Test	1
1.1.1 On-chip Spectrum Analyzer	1
1.1.2 On-chip Oscilloscope	3
1.1.3 Linearity Measurement In-situ	4
1.2 On-chip Analog Circuit Optimization	7
1.2.1 Analog Circuit As a Macromodel	7
1.2.2 Optimization Used for Analog Design	8
1.2.3 Validation-Optimization System Architecture	9
1.3 Organization	11
2. IN-SITU EXCITATION SIGNAL GENERATOR, OUTPUT RESPONSE ANALYZER AND STABILITY DETECTION	12
2.1 System Blocks and Circuit Implementation	12
2.1.1 Excitation Signal Generator (ESG)	12
2.1.2 Output Response Analyzer (ORA) and Quantization Block	15
2.1.3 Frequency Response Measurement Procedure	18
2.1.4 Stability Check Procedure	21
2.2 Optimization Procedure and Implementation	22
2.2.1 Cost Function Definition	22
2.2.2 Combined meta-heuristic engine	26

2.3	CUO Study Cases–Active Filters	29
2.3.1	Non-ideal Active-RC Biquad	30
2.3.2	Non-ideal Gm-C Biquad	32
2.4	Precision Analysis	33
2.4.1	Measurement Errors	34
2.4.2	Computational Error	36
2.4.3	System Analysis	39
2.5	Experimental Results	39
2.5.1	2-D Problem: 2 Decision Variables	40
2.5.2	4-D Problem: Four Decision Variables	44
2.6	Conclusion	46
3.	HIGH-LINEARITY SINE-WAVE SYNTHESIZER ARCHITECTURE BASED ON FIR FILTER APPROACH AND SFDR OPTIMIZATION	47
3.1	Motivation	47
3.2	Harmonic Cancellation Technique	52
3.2.1	Waveform, Fourier Series and Harmonics	52
3.2.2	Principles of the Harmonic Cancellation	58
3.2.3	Odd Order Cancellation FIR Filter Approach	61
3.3	Circuit Implementation	67
3.3.1	System Architecture	67
3.3.2	Oscillator and Phase Shifter	69
3.3.3	Weighted Resistor Summing Network	70
3.3.4	LPF and Output Buffer	74
3.3.5	Design Procedure	74
3.4	Iterative SFDR Optimization	81
3.4.1	Error Analysis	81
3.4.2	Min-Max Optimization	83
3.4.3	Iterative Optimization Algorithm	86
3.4.4	Optimization Procedure Simulation	87
3.4.5	Discrete Phase Shifter	89
3.4.6	Temperature Stability Analysis	90
3.4.7	Clock with Jitter	93
3.5	Experimental Results	95
3.6	Conclusions	100
4.	ON-CHIP TWO-TONE SYNTHESIZER BASED ON A MIXING-FIR ARCHITECTURE	102
4.1	Background	102
4.2	Two-tone Generation	106
4.2.1	Two-tone Signal Generation Architecture	106

4.2.2	Linearity Test using Weakly Nonlinear Stimulus	107
4.2.3	Cascade FIR-based Harmonic Cancellation	110
4.3	Circuit Implementation	115
4.3.1	System Architecture	116
4.3.2	Current Mirror Array for FIR Tap Coefficients	118
4.3.3	Clock Divider and Current Combiner for FIR Tap Delays	119
4.3.4	Up-Conversion Mixer	123
4.3.5	3-bit DEM Rotator	126
4.3.6	Design Procedure	128
4.4	Non-ideality Analysis	133
4.4.1	Baseband Current Mismatch and Phase Error	133
4.4.2	3-bit DEM Rotator	137
4.4.3	Nonlinear Up-conversion	139
4.4.4	Aliasing of the Residue Harmonics	141
4.4.5	LO Leakage and Imbalance	143
4.5	Experiment Results	144
4.6	Conclusions	152
5.	CT+DT HYBRID BASEBAND CHAIN USING HARMONIC CANCEL- LATION FOR ON-CHIP LINEARITY TEST	153
5.1	Background	153
5.2	Linearity Measurement of a Hybrid Chain	156
5.2.1	Hybrid Chain System Architecture	156
5.2.2	Harmonic Cancellation for Two-tone Suppression	157
5.2.3	Linearity Measurement Method	158
5.2.4	Circuit Implementation	160
5.3	Evolutionary Charge-domain Filter Design	162
5.3.1	5-phase MA-3 ² Filter	162
5.3.2	8-phase MA-3 ² Filter	165
5.3.3	Stage Compaction and HC-3 ² Filter	167
5.4	Design Procedure	170
5.4.1	Operational Amplifier	170
5.4.2	Programmable CT Filter	175
5.4.3	HC-3 ² DT Filter	176
5.5	Analysis of Measurement Precision	180
5.5.1	Notching Degradation	180
5.5.2	Measurement Precision	182
5.6	Experiment Results	184
5.7	Conclusion	189
6.	ANALOG LTI SYSTEM AC/DC BIST BASED ON A TIME-TO-DIGITAL CONVERTER	192

6.1	Background	192
6.2	Proposed Analog BIST Approach Using Only Digital I/O	194
	6.2.1 System Architecture	194
	6.2.2 Time-domain measurement principle	195
6.3	Block Circuits Design	199
	6.3.1 Ramp Generator	199
	6.3.2 Trigger/Multiplexer	200
	6.3.3 Time-to-digital Converter	201
6.4	Measurement Analysis	202
	6.4.1 Quantization Error	202
	6.4.2 Exponential-term-induced Error	203
	6.4.3 Linearity Analysis	204
6.5	Experimental Results	206
6.6	Conclusions	208
7.	CONCLUSIONS AND FUTURE WORKS	209
	7.1 Conclusions	209
	7.2 Future Works	212
	7.2.1 Strengthen or cancel arbitrary harmonic(s)	212
	7.2.2 Improve the notching of HC-3 ² DTF	214
	7.2.3 Time-domain measurement of LTI system with multiple poles and zeros	215
	REFERENCES	216

LIST OF FIGURES

FIGURE	Page
1.1 The architecture of an on-chip spectrum analyzer.	2
1.2 Delayed sampling mechanism.	3
1.3 Principle of the IM3 measurement using two-tone excitation signals.	5
1.4 On-chip validation-optimization system concept.	10
2.1 ESG schematic: an excitation sine-wave generator with harmonic cancellation technique.	14
2.2 Schematic of the output response analyzer (ORA) and the quantization block details.	16
2.3 Simulated waveform of the proposed ORA	17
2.4 I/Q measurement.	18
2.5 Frequency response measurement procedure: (a) Gain computation data flow and its (b) timing plot.	19
2.6 Principle of the oscillation detection.	21
2.7 Definition of a 2nd-order BPF frequency response matching cost function.	23
2.8 Work flow of the proposed combined meta-heuristic algorithm.	27
2.9 Topology and design parameters of an active-RC BPF biquad.	30
2.10 The actual Q_a contour for (a) ideal $Q_0 = 4$ (active-RC) and (b) ideal $Q_0 = 16$ (Gm-C, $\phi_E = \omega_0/BW_O$, where BW_O is the bandwidth of the OTA).	31
2.11 Schematic of the Gm-C biquad.	33
2.12 Abstracted model for the precision analysis.	34
2.13 Measured SFDR of the excitation sine-wave at 24 MHz.	35

2.14	Simulation results of the computational errors: (a) Simulated SNR_G by sweeping the ADC and the division fractional bit-width in the gain computation stage and (b) simulated SNR_F by sweeping the division and $F(v)$ fractional bit-width in the cost function computation stage.	38
2.15	Chip die photograph of the proposed VO system. (CUO includes only the active-RC BPF. The Gm-C version is almost the same size.)	42
2.16	Experiment results: (a) Measured response impacted by biquad bias current, (b) $F(\mathbf{v})$ 3-D surface and the optimization steps (dots) ($X_1 = R_K = R_Q$ and $X_2 = R_1 = R_2$ for a biquad in Fig. 2.9), (c) $F(\mathbf{v})$ contour and the optimal \mathbf{v}_{best} , (d) fixed response matching with power sizing (active-RC, 2nd order), (e) f_0 shift with power sizing (active-RC, 2nd order), and (f) fixed response matching with capacitance sizing (Gm-C, 2nd order).	43
2.17	4-D optimization experimental results: (a) Optimization procedure of a 4th order Butterworth BPF ($X_1 = R_K = R_Q$ and $X_2 = R_1 = R_2$ for the 1st biquad stage, and $X_3 = R_K = R_Q$ and $X_4 = R_1 = R_2$ for the 2nd biquad stage) and (b) experimental response matching with power sizing (active-RC, 4th order).	45
3.1	Architecture of the sine-wave synthesizers: (a) High-order BPF or LPF, (b) basic harmonic cancellation synthesizer, and (c) direct digital synthesizer.	48
3.2	Sawtooth waveform with infinite and limited number of terms for (a) $k = 1$, (b) $k = 1 \dots 3$, and (c) $k = 1 \dots 5$.	53
3.3	Frequency spectrums for (a) $y(t)$ of Fig. 3.3a, (b) $y(t)$ of Fig. 3.3b, (c) $y(t)$ of Fig. 3.3c, and (d) an ideal sawtooth waveform.	54
3.4	Manipulating waveform shapes by (a) changing a harmonic amplitude or (b) changing a harmonic phase shift.	55
3.5	Manipulation of the sawtooth waveform.	57
3.6	Spectrum of the square waveform.	58
3.7	Categories of the harmonic cancellation implementations.	59
3.8	Sine-wave synthesis architecture proposed in S.W. Park's paper.	60
3.9	Sine-wave synthesis architecture with limited α_i proposed in M.M. Elsayed's paper.	61

3.10	Odd order harmonic filter design: (a) Equivalent architecture of the FIR filter, (b) time-domain half-cosine pulse $y(t)$, and (b) Fourier transform $Y(\omega)$ of the half-cosine pulse.	62
3.11	Odd order cancellation FIR filter design: (a) 2-tap coefficients, (b) $Y_{aliased}(\omega)$ for the 2-tap FIR filter, (c) 5-tap coefficients, and (d) $Y_{aliased}(\omega)$ for the 5-tap FIR filter.	64
3.12	System architecture of the proposed synthesizer.	68
3.13	Schematic of the ring oscillator and the phase shifter.	70
3.14	Schematic of the weighted resistor summing network.	71
3.15	Schematic of the ring oscillator delay cell with the phase shifters.	75
3.16	Simulated oscillation frequency versus bias current.	76
3.17	Layout of the ring oscillator and the phase shifters.	77
3.18	Simulated rising edge delay versus phase shifter control code.	78
3.19	Layout arrangement of the resistors	79
3.20	Layout of the PDC switches and the weighted resistor summing network.	80
3.21	Cost function surface for $M = 5$, $N = 12$, $q = 10$, and sweeping $\Delta\theta_1$ and $\Delta\theta_{10}$ (a) with ideal tap coefficients, (b) with ideal tap coefficients and a fixed $\Delta\theta_6 = -5\% \cdot \frac{2\pi}{N}$, or (c) with a non-ideal tap coefficient $\Delta\alpha_2 = 5\% \cdot \alpha_2$	84
3.22	Simulated F_{cost} distribution with Monte-Carlo simulation before/after the optimization procedure.	88
3.23	Cases of the proposed optimization procedure.	88
3.24	Simulated F_{cost} distribution after the optimization: (a) $f_0=500$ MHz, 3-bit phase shifter, (b) $f_0=100$ MHz, 3-bit phase shifter, and (c) $f_0=500$ MHz, 2-bit phase shifter.	89
3.25	Temperature stability analysis: (a) F_{cost} fluctuation versus temperature, (b) simulated max (F_{cost}) distribution before/after an one-time optimization, and (c) simulated distribution of Δ due to a one-time optimization.	91

3.26	Behavior model for evaluating the impact of the clock jitter.	93
3.27	Simulation results of the jitter impact on the V_O spectrum with (a) 1-ps RMS jitter and (b) 100-ps RMS jitter.	94
3.28	The die photograph of the proposed synthesizer.	95
3.29	Chip experimental results: (a) Measured SFDR @ 150 MHz before the optimization, (b) measured SFDR @ 150 MHz after the optimization, (c) measured SFDR @ 750 MHz before the optimization, and (d) measured SFDR @ 750 MHz after the optimization.	96
3.30	Measured SFDR/-THD vs. output frequency (BO: before the optimization, AO: after the optimization).	97
4.1	Traditional two-tone test configuration using testing equipment.	102
4.2	Two-tone generation architecture concept: (a) Mixing-FIR two-tone generation and (b) output two-tone signal spectrum.	106
4.3	Proposed two-stage cascade FIR harmonic cancellation and tap coefficients for the “baseband” single-tone generation.	111
4.4	Frequency response of the cascade FIR architecture.	112
4.5	The architecture of the single-stage 11-tap FIR approach.	113
4.6	Comparison between two-stage FIR approaches: (a) Proposed two-stage FIR with three 5-tap FIRs and (b) a single 3-tap FIR followed by a single 5-tap FIR.	114
4.7	System architecture of the proposed two-tone synthesizer and the corresponding rearranged FIR path	115
4.8	The proposed two-stage FIR architecture after the path rearrangement.	117
4.9	Current mirror implementation of the two-stage FIR coefficients.	118
4.10	Current steering implementation with shifted clocks and the equivalent flow diagram. CK_b is CK_a delayed by $T/4$, where T is the clock period.	119
4.11	Current combiner topology and clock connections of the proposed two-stage FIR architecture.	120

4.12	Schematic of a simple passive up-conversion mixer.	123
4.13	Passive up-conversion mixer with bootstrapped MOS switches. . . .	124
4.14	IM3 tone comparison between the mixers with and without the bootstrapped switches.	125
4.15	Mismatches existed in the current mirrors.	126
4.16	Current branch rotator with dynamic element matching and the corresponding current branch rotation patterns.	127
4.17	Full display of the current mirror array arrangement and the layout patterns.	130
4.18	Layout of the current branch rotator.	131
4.19	The timing of the 24-phase clock output $\phi_{0...23}$	133
4.20	Schematic of the 24-phase clock generator.	133
4.21	Definition of the phase error $\Delta\theta_i$	134
4.22	Monte-Carlo simulation results: (a) $I_{a1}<0>$ branch current deviation, (b) phase error deviation of ϕ_0 , and (c) distribution of the HD3 calculated from the current and the clock phase mismatches.	135
4.23	Contour of the simulated average HD3 by sweeping the current mismatch and the clock phase error.	136
4.24	Simulated power spectrum density with current mismatches plus (a) DEM OFF or (b) DEM ON.	138
4.25	Simulated HD3 suppression by using the proposed DEM technique. .	139
4.26	Simulated IM3 of the passive mixer ($f_0 = 1\text{MHz}$).	140
4.27	“Fake” IM3 induced by fold-back harmonics.	141
4.28	(a) Die photograph, (b) power distribution, and (c) block area distribution of the proposed two-tone generator.	144
4.29	Test bench configuration for the proposed two-tone generator. . . .	145

4.30	Measurement results: (a) IM3 ($f_{LO}=4.8$ MHz and $f_0=100$ kHz), (b) IM3 ($f_{LO}=480$ MHz and $f_0=1$ MHz), (c) IM3, where f_{LO} is swept from 24 MHz to 1008 MHz ($f_0=1$ MHz), and (d) IM3, where f_{LO} is swept from 2.4 MHz to 153.6 MHz (10 kHz, 100 kHz, and 1 MHz f_0).	146
4.31	Measured IM3 of the two-tone signals ($f_{LO} = 4.8$ MHz, $f_0 = 100$ kHz) (a) without DEM or (b) with DEM.	149
4.32	Measured IM3 improvement by turning on DEM.	150
5.1	DTF architecture with the decimation technique.	154
5.2	Receiver architecture using CT+DT hybrid baseband chain.	156
5.3	Proposed IP3 measurement procedure: (a) In-band IP3 and (b) out-of-band IP3.	159
5.4	Frequency responses with different sampling clock settings: (a) Normal operation mode ($\omega_S = \frac{16}{3}\omega_0$), (b) test tone suppression for in-band IP3 test ($\omega_S = \frac{8}{5}\omega_0$), and (c) test tone suppression for out-of-band IP3 test ($\omega_S = 4\omega_0$).	161
5.5	Schematic of the Tow-Thomas LPF biquad and the Miller-compensated amplifier.	162
5.6	Minimum two-stage MA-3 ² filter with (a) filter architecture and (b) timing diagram for the five-phase MA-3 ² filter.	163
5.7	8-phase MA-3 ² filter before the compaction.	166
5.8	Compacted 8-phase HC-3 ² filter with weighted capacitors ($C_{11} : C_{12} : C_{13} = 1 : 17/12 : 1$).	168
5.9	Timing diagram for DT filters: (a) Eight-phase MA-3 ² filter and (b) compacted 8-phase HC-3 ² filter	169
5.10	Detailed schematic of the the Miller-compensated amplifier.	170
5.11	I_{DS} vs V_{DS} simulation result for N3 cascode transistors.	172
5.12	Simulated transfer function of the miller-compensated amplifier.	173
5.13	Simulated transfer function of the amplifier's DCOC loop.	174
5.14	Configuration of the capacitor and resistor arrays.	175

5.15	Layout of the CT LPF biquad.	177
5.16	Layout of the HC-3 ² DT Filter.	178
5.17	Hybrid baseband chain architecture (a) without the integration stage or (b) with the integration stage.	179
5.18	Comparison of the frequency responses between the baseband chains without and with the integration stage.	179
5.19	Monte-Carlo simulated small-signal two-tone suppression at $3f_0$ and $5f_0$ ($f_S = 8f_0$).	180
5.20	Simulated small-signal tone suppression at $3f_0$ versus the sampling frequency, $f_S = 8f_0$ is swept from 8 MHz to 800 MHz. Four cases were arbitrarily picked from the Monte-Carlo simulation results.	181
5.21	IP3 measurement methodology: (a) Conventional measurement method, (b) PD-based measurement method, and (c) PD-based measurement induced error.	183
5.22	Die photograph of the proposed baseband chain.	184
5.23	Configuration of the chip test bench.	185
5.24	Measured output spectrum of the proposed baseband: (a) Normal operation mode and (b) tone-suppression mode for in-band IIP3 mea- surement.	186
5.25	In-band IIP3 measurement results: (a) IIP3 measured by a commercial spectrum analyzer and (b) IIP3 measured by the proposed PD-based method.	187
5.26	In-band IIP3 measurement results: (a) IIP3 measured by a commercial spectrum analyzer and by the power detector and (b) normalized IIP3 measurement results.	188
5.27	Measured frequency response of the proposed baseband: (a) Normal operation and tone suppression mode for in-band IP3 measurement and (b) normal operation and tone suppression mode for out-of-band IP3 measurement.	190
6.1	Proposed BIST approach: (a) BIST system architecture, (b) DC gain measurement, (c) t_d measurement with error cancellation.	195

6.2	Schematic of the system blocks: (a) Ramp generator, (b) trigger and multiplexer, and (c) time-to-digital convertor.	199
6.3	Measurement analysis: (a) Measure DC gain, $k\delta_{TDC} = 50 \mu\text{V}$, sweep ΔV_{LAT} , (b) measure DC gain, $\Delta V_{LAT} = 10 \text{ mV}$, sweep $k\delta_{TDC}$, (c) ramp-slope-induced error in a 1-pole system.	202
6.4	Measurement analysis: (a) Nonlinearity-induced error sources, (b) periodic-ripple-induced error delay, (c) settling/relaxation-time-induced error delay.	204
6.5	Experimental result: (a) Photo diagram of the chip die, (b) f_{3dB} measurement results of a 1st order low-pass active filter, (c) DC gain measurement results of an inverting amplifier, and (d) quality factor measurement results of a 2nd order low-pass filter.	207
7.1	Expected spectrum of enforcing an arbitrarily selected harmonic. . .	212

LIST OF TABLES

TABLE	Page
2.1 Summary of System Design Parameters	40
2.2 Comparison of the Optimization or Tuning Systems	41
3.1 Normalized coefficients of the odd order cancellation FIR filter	66
3.2 Iterative optimization algorithm	86
3.3 Comparison of sine-wave synthesizer performance	99
4.1 24-phase Clock and Current Branch Connection Pattern (\emptyset represents the dummy connection)	122
4.2 Comparison of Two-/Single-tone Generation Performance	151
5.1 Hardware cost comparison of the DTF architectures	169
5.2 Summary of the Miller-compensated amplifier parameters	174
5.3 Design parameter of the Miller-compensated amplifier	175

1. INTRODUCTION

1.1 On-chip Analog Circuit Built-in Self-Test

Testing an analog integrated circuit (IC) is a traditional research topic accompanied with the analog IC design. However, it received less attention because the resource limitation on the testing is much less than that on the circuit design. Traditionally, the testing of analog ICs is conducted by external instruments that is inherited from testing the circuits with discrete components. This scheme can satisfy most applications till recently. On the one hand, the rapid evolution of system-on-chip (SoC) ICs introduces tremendous complexity, particularly the increased number of integrated analog parts, into the ICs; and thus the corresponding cost of testing is increasing. ITRS report [1] predicts that the test data volume of SoCs will exceed that of the high performance microprocessors at around the year of 2019. On the other hand, the development of built-in self-test (BIST) techniques have set up a series of standard test procedure and methodology for digital circuits [2]. Therefore, as concluded by [1], “Improvements in analog/mixed-signal DFT and BIST are needed” (DFT – Design for Testability, BIST – Built-in Self-test) to reduce test time and cost. Before proposing such improvements in this dissertation, let’s first examine some popular BIST methods for analog/mixed-signal circuits beyond the simplest DC test.

1.1.1 On-chip Spectrum Analyzer

As an alternative of the bulky external spectrum analyzer, researchers have attempted to mimic the analyzer’s behavior on chip. A conventional architecture of on-chip spectrum analyzer is demonstrated in Fig. 1.1. The frequency-variable sine-wave generator generates a high quality sinusoidal waveform as the input excitation

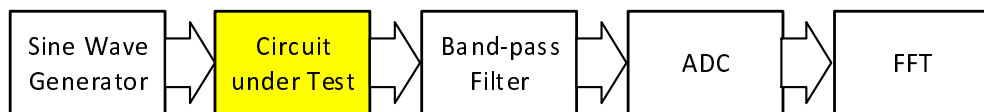


Figure 1.1: The architecture of an on-chip spectrum analyzer.

signal fed into the circuit-under-test (CUT). The output of the CUT is then filtered by a band-pass filter (BPF) with a high Q factor. The amplitude of the BPF's output is picked by an analog-to-digital converter (ADC). By sweeping a range of frequencies, the spectrum can be retrieved via a digital signal processor (DSP) conducting Fast Fourier Transform (FFT). This architecture is first proposed in [3]. It also proposed a switched-capacitor sinusoidal waveform generator utilizing the principle of harmonic cancellation (HC) technique. [4] further improved this traditional architecture. It first up-converts the excitation signal to a higher frequency band and then down-converts the CUT's output to the baseband. This method is able to relax the speed constraint on the design of the spectrum analyzer, although more complexity is introduced and calibration is needed. Further enhancement uses a logarithmic amplifier [5][6] or a lockin amplifier [7] to replace the complicated ADC circuit. Additionally, [8] proposed a Σ - Δ modulator to measure the amplitude and DC level by simply counting the output sequence of the modulator. Although [8] works only for the oscillation-based-test, [9] further extends the test scheme to an on-chip spectrum analyzer. By using the coherent sampling with quadrature clocks. This measurement scheme is insensitive to clock jitters thanks to the long count time. Longer is the Σ - Δ sequence counting, higher is the measurement accuracy. Of course, such long test time is also a disadvantage, plus the requirement of high-speed over-sampling clock.

1.1.2 On-chip Oscilloscope

Analog-to-digital converters (ADCs) is able to conduct the function of an on-chip oscilloscope. However, they're not intentionally designed for test purpose. Low overhead low cost on-chip oscilloscope can be found in some previous research works on the supply noise measurement. Supply noise measurement technique usually adopts delayed sampling and autocorrelation to extend the bandwidth of the measurement while keep a small footprint of the test circuitry [10][11].

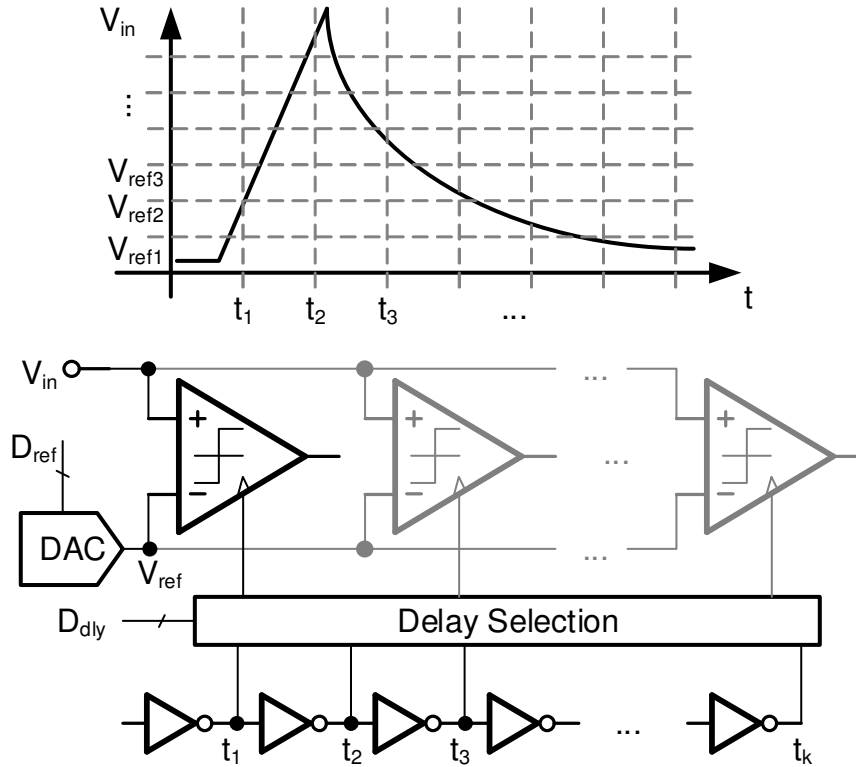


Figure 1.2: Delayed sampling mechanism.

The principle of the delayed sampling is illustrated in Fig. 1.2. Time-shifted sampling clocks drive one or more comparators to sample the input signal V_{in} for a

fixed reference voltage V_{ref} . By changing V_{ref} , the whole waveform could be reconstructed after several measurement cycles. Obviously, the delayed sampling cannot finish the measurement in one time and thus it works only for periodical signal (e.g. the supply noise in clock-driven digital systems). Autocorrelation is used to identify such strongly repeated patterns [10]. [12] implemented a set of clocked comparators to sample the supply noise with the time-shifted versions of sample signal, resulting in an effective sampling rate of 20 Gbps. Then, by changing the reference voltage step by step, the waveform could be reconstructed. [13] reduced the number of comparator to only one. Moreover, it expands the measurement to both of the supply noise and the substrate noise, by using a PMOS-based differential sensor. [14] uses a VCO as the converter and applies the sampled supply voltage directly to the VCO's supply. Moreover, it should be noted that, no extra excitation signal is needed for the supply noise measurement. The disadvantage of the delayed sampling technique is that the multiple time and reference steps lead to slow measurement speed.

1.1.3 Linearity Measurement In-situ

In order to evaluate the linearity of an analog circuit, several metrics can be adopted, such as total harmonic distortion (THD), third-order intercept point (IP3), third-order intermodulation (IM3), 1dB compression point (P1dB), etc. Fortunately, any one of these metrics can be converted to the others [15]. Later in Section 4.2.2, these linearity metrics will be analyzed.

In the instrumentation measurement, the THD are usually used for large signal test at low frequency domain, while IP3 is used for RF/microwave circuits. The former one can be simply read from the spectrum analyzer with a single-tone input sine-wave, although the measurement of IP3 requires more effort, which takes two-tone sinusoidal signals as the excitation and some graphical derivation, which will

be elaborated in Section 5.5.2, is needed to obtain the exact intercept point.

As an on-chip alternative, [16] proposed a two-tone PLL to generate the required excitation. Moreover, it is more difficult to design an ORA for the on-chip linearity measurement.

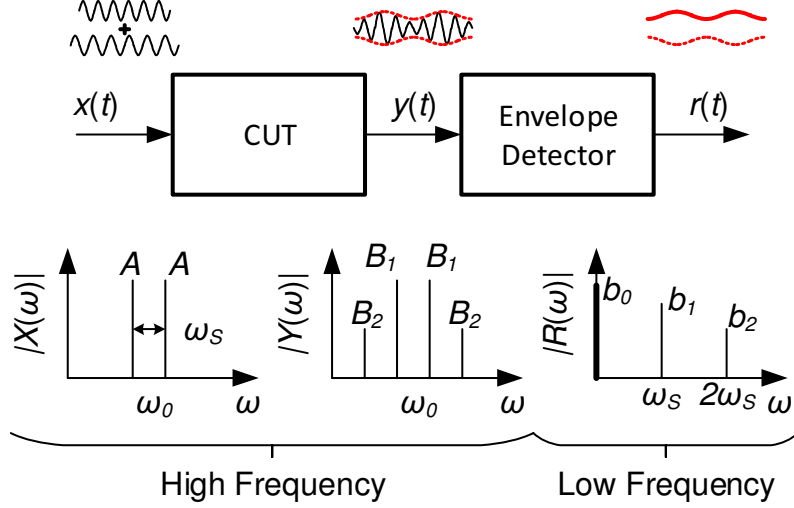


Figure 1.3: Principle of the IM3 measurement using two-tone excitation signals.

Fig. 1.3 shows the principle of the IM3 measurement which is popular among on-chip solutions. Two-tone excitation ($x(t)$) signals at frequencies of $\omega_0 \pm \frac{1}{2}\omega_s$ are fed into the CUT, which generates the output signal $y(t)$. The envelope of $y(t)$ can be derived via

$$r(t) = \left| 2 \cdot B_1 \sin\left(\frac{1}{2}\omega_s t\right) + 2 \cdot B_2 \sin\left(\frac{3}{2}\omega_s t\right) \right| \quad (1.1)$$

where B_1 is the fundamental magnitudes of $y(t)$, and B_2 is the third-order nonlinear

components. $r(t)$ turns out to be a periodical signal [17], whose Fourier series is

$$r(t) = b_0 + \sum_{k=1}^{\infty} b_k \cos(k\omega_s t) \quad (1.2)$$

We have $b_0 = 4/\pi (B_1 - B_2/3)$ and

$$b_k = \frac{8}{\pi} \left(\frac{B_1(-1)^{k+1}}{4k^2 - 1} + \frac{3B_2(-1)^k}{4k^2 - 9} \right) \quad k = 1, 2, \dots \quad (1.3)$$

Therefore, the IM3 component can be extracted by measuring any two b_k . For instance, b_1 and b_2 are selected,

$$\begin{aligned} B_1 &= 0.9204b_1 - 1.2885b_2 \\ B_2 &= 0.1432b_1 + 0.7159b_2 \end{aligned} \quad (1.4)$$

Solving Eq. 1.4 can get the IM3 components of $x(t)$, $\text{IM3} = B_2/B_1$.

It should be noted that, on the one hand, b_k is obtained from the spectrum of $r(t)$, which is a slow signal after the envelope detector. On the other hand, the frequency spectrum can also be used to evaluate the THD of a circuit working at the low frequency range. Therefore, it is possible to share some part of the linearity ORA for both low frequency and RF/microwave circuits theoretically.

The amplitude of b_k can be obtained in the frequency domain, although an intensive computation of fast Fourier transform (FFT) should be implemented. Prior research works focuses on reduce the complexity of digital FFT computation. For two-tone test, [18] directly sampled the CUT's output signal, and indicated the coherent sampling could be used to avoid spectrum leakage and simplify the design of the FFT engine if the criteria $f_{\text{samp}C} = \Delta f \cdot N_{\text{FFT}}/N_{\text{cycle}}$ is met, where $f_{\text{samp}C}$ is the coherent sampling frequency, Δf is the spacing between the two test tones,

N_{FFT} is the number of the FFT points, and N_{cycle} is the chosen number of cycles of the input signal. [18] further simulated the impact on the computation accuracy caused by limited bit-length and other non-ideality. It concludes that, the proposed 16-point FFT engine combined with a 10-bit ADC, could achieve IM3 extraction error within 1.5dB for IM3 components ≤ 50 dBc. Moreover, the synthesized fully digital 16-point FFT engine occupies around 0.073 mm^2 in 45 nm CMOS technology.

Based on [18], [19] further optimized the FFT engine design. It implemented the test architecture in Fig. 1.3 and sampled the output of the envelope detector. As explained above, not all FFT outputs are needed to calculate IM3, and thus [19] further removed the computational resources in the butterfly stages for computing non-use FFT outputs. [19] synthesized a 512-point FFT engine, which takes only 0.036 mm^2 in $0.13 \text{ }\mu\text{m}$ CMOS technology.

This dissertation also focuses on the on-chip linearity test. An on-chip two-tone sine-wave synthesizer is proposed in Section 4, and a low cost linearity evaluation methodology is further proposed in Section 5 for hybrid baseband chain. A combination of these two proposals will allow designers to integrate the linearity testability fully on-chip for analog circuits.

1.2 On-chip Analog Circuit Optimization

1.2.1 Analog Circuit As a Macromodel

Analog IC designers traditionally rely on device-level models to construct analog systems from the bottom up. In order to counteract with the process-voltage-temperature (PVT) variations and device aging, design centering [20] is emphasized in most analog design procedures. However, due to the rapid scaling down of silicon-based and other emerging technologies [21], such as FinFET, FD-SOI, SiGe, etc., the design complexity is increasing tremendously not only because of the wider dis-

tribution of PVT variations [22], but also the more complicated transistor physical behaviors. To accelerate the analog design procedure, macromodels with higher level abstractions were developed in [23] to replace a group of trivial device models. Later, [24] proposed a macromodel validation procedure via stimulating the circuit and measuring its outcome; thus, it is not critical to know the detailed structure or the operating principle of a circuit. In [25], a parameterized macromodel was further proposed to predict the circuit behavior as a function of its design parameters, leading to a tractable system-level synthesis process. From the perspective of macromodel, designing a circuit is to find a combination of all design parameters that will make the circuit achieve the desired performance.

1.2.2 Optimization Used for Analog Design

Optimization algorithms have been well researched such as the steepest descent algorithm [26][27], the simulated annealing (SA) algorithm [28][29], the genetic algorithm (GA) [30], the ant colony (AC) optimization [31] and the particle swarm (PS) optimization [32]. Some other methods, such as the neural network [33] and the support vector machines [34] were implemented to automatically predict the macromodel of an analog circuit. Among all these algorithms, the SA belongs to the category of meta-heuristic, which depends on the analog circuit's behavior but not on its analytic model. In other words, the meta-heuristic methodology is a model-free approach.

The post-fabrication tuning of analog circuits receives increasing attention when a cell array structure was implemented as in [35], which provided a possible way to calibrate each analog component towards its design performance. A fully integrated implementation was reported in [36] for matching the capacitor array in a pipeline ADC, yet the proposed random search algorithm has limited efficiency in general.

Later in [37], the complex tuning process of an RF circuit, guided by the gradient search or genetic optimization, was executed by an external digital signal processor (DSP). Furthermore, to make the built-in self-optimization method applicable to general analog circuits, recent work [38] presented the use of a neural network as an on-chip classifier which, however, requires an external training process and large memory consumption.

The optimization algorithm has been well adopted to solve this type of design problems. However, currently, this high-level macromodeling and synthesis only works as auxiliary CAD tools for analog design.

1.2.3 Validation-Optimization System Architecture

The full concept of the self-contained model-free integrated validation-optimization (VO) system is proposed in Fig. 1.4. The complete system consists of an analog self-validation path and a digital optimization path. The self-validation path contains the circuit under optimization (CUO) with k design variables, whose performance can be changed by the k -dimension design vector \mathbf{v} . Apart from the CUO, an excitation signal generator (ESG) and an output response analyzer (ORA) block are implemented to stimulate the CUO and measure its behaviors. Generally, for different types of CUO, the ESG can be designed to generate signals, such as DC, pulses, single- and two-tone sinusoidal, ramp, etc. While for different output measurement types, ORA can accomplish the measurement of DC level, frequency response ($H(s)$), transient waveform ($h(t)$), linearity, noise, etc. The output of the ORA, which contains the CUO's performance metrics, is then digitalized by a quantization block and passed to the optimization path. Note that the validation overhead for multiple CUOs is fixed. In the all digital optimization path, a cost function, $F(\mathbf{v})$, is used to evaluate the error between the measured actual performance y_{actual} and the desired performance

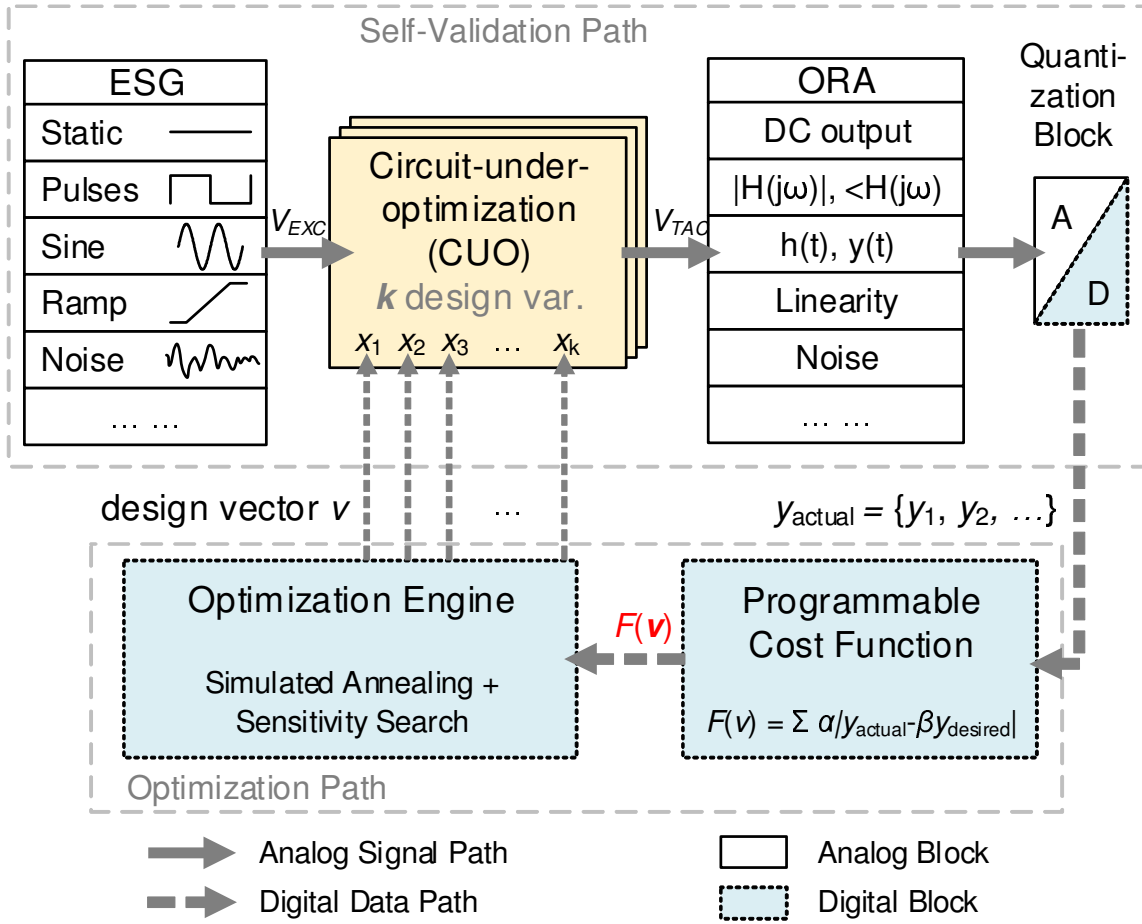


Figure 1.4: On-chip validation-optimization system concept.

$y_{desired}$, where \mathbf{v} is the vector of design variables. Integrating the optimization engine on-chip is a significant challenge, because most algorithms require tremendous computation and memory resources. A general optimization engine, which adopts a combined algorithm based on simulated annealing [28] and sensitivity search [26], is proposed to find the minimum value of $F(\mathbf{v})$ by changing \mathbf{v} , aiming at balancing the searching efficiency, the algorithm complexity and the area overhead. The whole digital optimization path is fully integrated on chip and can assist different analog circuit types in real time.

1.3 Organization

This dissertation is organized as follows. Section 2 introduces the details of the proposed self-validation and optimization system using a digital multi-dimensional optimization engine. The excitation sinusoidal signal synthesizer of the proposed system is further elaborated in Section 3. Section 4 extends the sine-wave synthesizer to generate two-tone signals for on-chip linearity test. The following Section 5 proposes an on-chip RF receiver linearity built-in test methodology for hybrid baseband chain. Furthermore, Section 6 proposes a low cost built-in analog tester with fully digital input/output for linear-time-invariant (LTI) analog blocks. Finally, in Section 7, this dissertation concludes all designs and discusses potential futures works.

2. IN-SITU EXCITATION SIGNAL GENERATOR, OUTPUT RESPONSE ANALYZER AND STABILITY DETECTION

2.1 System Blocks and Circuit Implementation

2.1.1 Excitation Signal Generator (ESG)

2.1.1.1 ESG of the VO system

In the test using an external equipment, an arbitrary waveform generator (AWG), such as a Keysight 33600A [39], is used to generate the excitation signal. Similarly, the excitation signal generator (ESG) is introduced into the VO system proposed in Fig. 1.4 to emulate the function of an AWG. The ESG design should first consider the hardware overhead, because the ESG is used solely for test purpose. Some possible ESG implementations are listed below,

- A static DC voltage or current bias can be implemented by a simple resistor-string or current-steering digital-to-analog converter (DAC).
- Pulse or square waveforms can be obtained from a digital clock, an oscillator or a frequency synthesizer [40]. This type of excitation signals is usually adopted for the measurement of the time-domain response [41], the slew rate, etc.
- A ramp signal, or a triangular waveform, can be generated using a charge-pump. It can be utilized for measuring the RC time constant [42]. Additionally, Section 6 in this dissertation will introduce a ramp-based methodology for the measurement of poles and zeros in a linear-time-invariant (LTI) system.
- The sinusoidal waveform is a widely adopted excitation signal in the test of analog/RF circuits. A sine wave with a variable frequency is usually associated with the characterization of the frequency response, such as the test shown in [4]. The noise transfer function can also be characterized indirectly via the

sinusoidal excitation [43]. Apart from the single-tone sinusoidal waveform, two-tone signals are popular among the linearity tests, as described previously in Section 1.1.3. In this dissertation, Section 3 and Section 4 will introduce different implementations of sinusoidal ESGs and propose new low-cost highly linear single-tone and two-tone generation techniques, respectively.

- The noise ESG can be a simple resistor. However, the fully integrated noise test remains an unexplored area, because the direct noise test requires the measurement circuit produces less noise than that of the circuit-under-optimization (CUO). Such strict requirement may render the built-in solution expensive or even not practical.
- Arbitrary waveforms can always be generated by a digital-to-analog converter. Nevertheless, the harmonic cancellation technique analyzed in Section 3 has the potential of synthesizing any signal more economically. This technique worths a further research effort.

2.1.1.2 ESG design for study cases

The proposed ESG is a compact sine-wave generator using finite impulse response (FIR) filter approach [44], which generates a high-linear sinusoidal waveform by combining five square waves, $P_0 \cdots P_4$. These five clocks have a voltage magnitude ratio of $\frac{1}{2} : \frac{13}{15} : 1 : \frac{13}{15} : \frac{1}{2}$ and are shifted by $0^\circ, 30^\circ, 60^\circ, 90^\circ$ and 120° , so as to cancel the 3rd, 5th, 7th and 9th harmonics of the output waveform V_{EXC} . The detailed computation of the magnitudes and the phase shifts will be further discussed in Section 3. In addition, a 50% duty-cycle is applied to eliminate the even order harmonics. Fig. 2.1 demonstrates the detailed ESG implementation. A frequency synthesizer with a 1 MHz reference clock is constructed to output CLK_{HF} , which is able to sweep from 128 to 508 MHz. It is further divided by 12 to generate phases

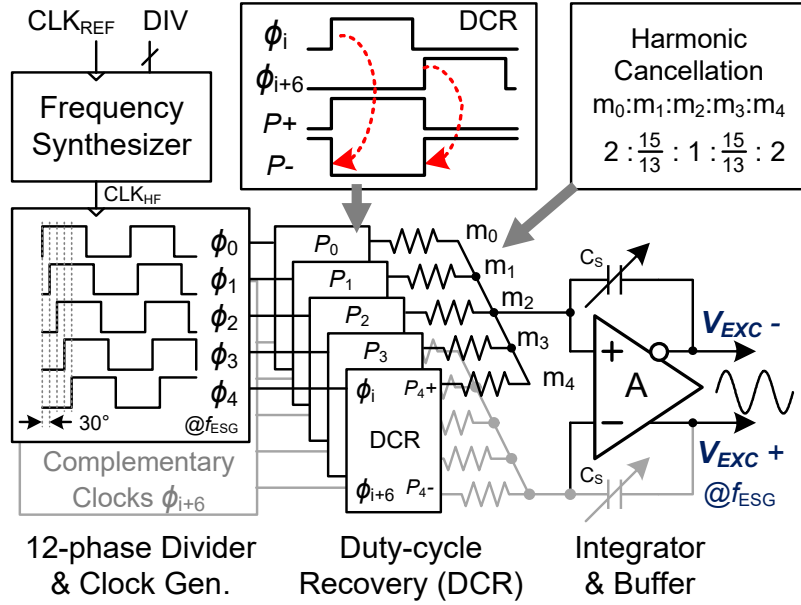


Figure 2.1: ESG schematic: an excitation sine-wave generator with harmonic cancellation technique.

ϕ_i at a frequency of f_{ESG} (10.7 to 42.3 MHz), which has a 30° phase shift between each other. $\phi_0 \dots \phi_4$ and their complementary clocks $\phi_6 \dots \phi_{10}$ are used to align the edges of the five designated clocks $P_0 \dots P_4$ and achieve 50% duty-cycle differential outputs. Note that ϕ_5 and ϕ_{11} are not used. The following resistive network converts $P_0 \dots P_4$ to currents, which are summed at capacitor C_S . C_S is a capacitor array with 4-bit control words, which is used to control the amplitude of V_{EXC} . It should be mentioned that, $\frac{1}{2} : \frac{13}{15} : 1 : \frac{13}{15} : \frac{1}{2}$ is the magnitude ratios of current branches when they're summed on C_S . To convert the voltage pulses, $P_0 \dots P_4$, to branch currents, reciprocal ratio should be applied to the resistors, as shown in Fig. 2.1, $m_0 : m_1 : m_2 : m_3 : m_4 = 2 : \frac{15}{13} : 1 : \frac{15}{13} : 2$.

2.1.2 Output Response Analyzer (ORA) and Quantization Block

2.1.2.1 ORA of the VO system

The output response analyzer is an interface stage between the CUO's output and the quantization ADC block. It is not always necessary for all measurement types. In some cases, the present of the ORA can help relax the design constraint of the ADC.

- An ORA is not needed for DC measurement. The DC voltage/current can be easily sampled by an low-speed low-power ADC.
- A direct sampling on the CUO's output with a high sampling frequency is able to retrieve $|H(j\omega)|$ and $\angle H(j\omega)$, which are the frequency-domain responses, with the assistance of the Fast Fourier Transform (FFT). However, a high-speed ADC design is a significant overhead if the ADC is only used for the test purpose. In this circumstance, a down-conversion ORA can be adopted [4]. A detailed design of this type of ORA will be discussed later in Section 2.1.2.2.
- Similar to the frequency-domain responses, the time-domain responses, $h(t)$, can use a delay line as the ORA. Its usage has been discussed in Section 1.1.2.
- Few papers have discussed the measurement circuit design for linearity test. This dissertation will propose a linearity test methodology using a power detector ORA in Section 4. The proposed power detector outputs a static DC voltage and thus the ADC speed constraint can be relaxed.
- An external noise test is conducted by using a noise figure analyzer (NFA), such as a Keysight N8973B [45]. During the test, noise generated by a resistive noise source is fed into the CUO, and the NFA will sample and compare the input and output noise to compute the noise figure. The major concern of the noise test is that the noise produced by the measurement circuit should

be much less than that of the CUO, which makes the fully integration not practical. However, alternative solutions of noise test exist, like measuring the noise transfer function [43], measuring the noise induced timing uncertainty [46], etc.

- Other electrical or physical parameters can also be collected by the VO system (Fig. 1.4) to help optimize the CUO, such as the power, the temperature [47], the stress [48], etc. Sensor techniques should be used for these parameters, although this dissertation is not going to elaborate on this topic.

2.1.2.2 ORA design for the study cases

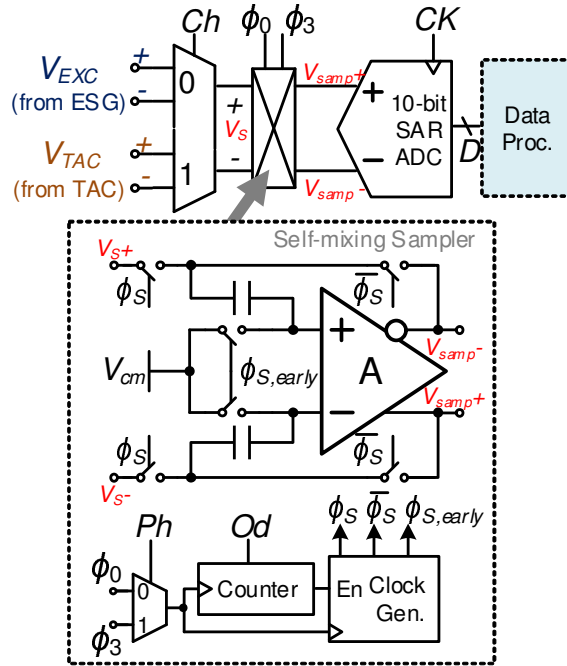


Figure 2.2: Schematic of the output response analyzer (ORA) and the quantization block details.

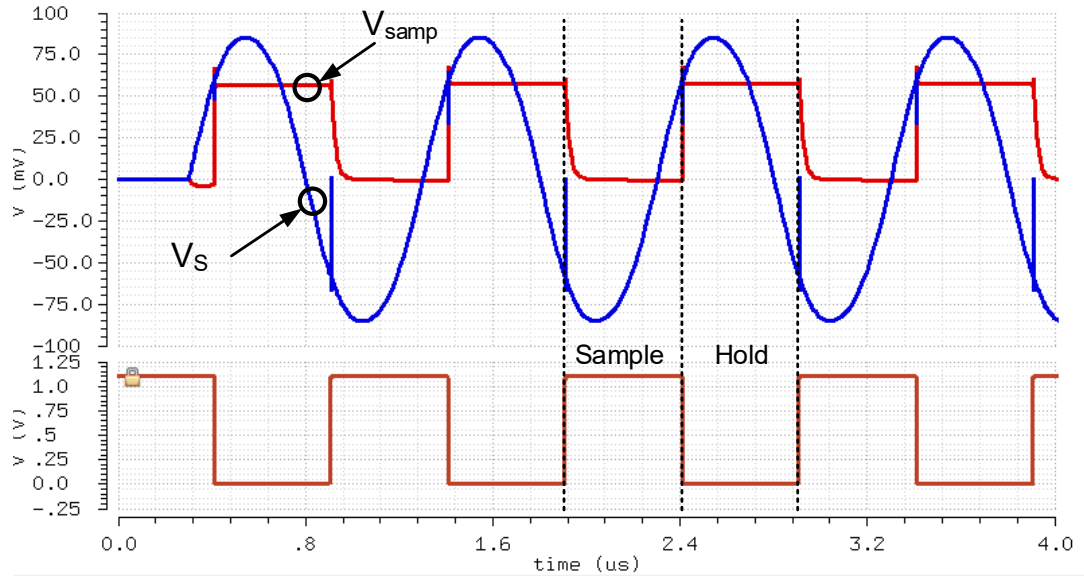


Figure 2.3: Simulated waveform of the proposed ORA

Fig. 2.2 shows the ORA, which is a down-conversion self-mixing sampler. Fig. 2.3 further demonstrates the simulated sample-and-hold behavior of the proposed ORA. The channel selection signal, Ch , selects the CUO's differential excitation input V_{EXC} or differential output V_{CUO} to be sampled by an active sample-and-hold (S/H) circuit. The non-overlapping sampling clocks $\phi_{S,early}$ and $\bar{\phi}_S$ are derived from ϕ_0 or ϕ_3 , which has a 90° phase difference between each other and are selected by Ph . ϕ_0 or ϕ_3 are from the ESG block. Therefore, the proposed S/H circuit acts as a self-mixing mixer, down-converting the sampled periodic signal to a DC level V_{samp} . By switching Ph , the in-phase and quadrature (I/Q) detection technique [49] can be used to retrieve the amplitude $\sqrt{I^2 + Q^2}$ and phase $\arctan(Q/I)$ of the sampled signal, as shown in Fig. 2.4. A counter is further introduced to enable the sampling clocks. $Od = 0$ will keep on generating clocks of N cycles, while $Od = 1$ outputs $N - 1$ cycles. On the one hand, multiple clock cycles guarantee V_{samp} stable after the

CUO reconfiguration. On the other hand, different sampling cycles help to detect the stability failure of the CUO, which will be discussed in Section 2.1.4. Moreover, the proposed down-converting mechanism relaxes the speed constraint of the ADC. In this design, a 100-kS/s power efficient 10-bit SAR ADC within a small footprint [50] is adopted with a digital output D , which is handled by the data processing circuit discussed next. And the details about the bootstrapped sampling switch refer to Fig. 4.13, which will be discussed later in Section 4.

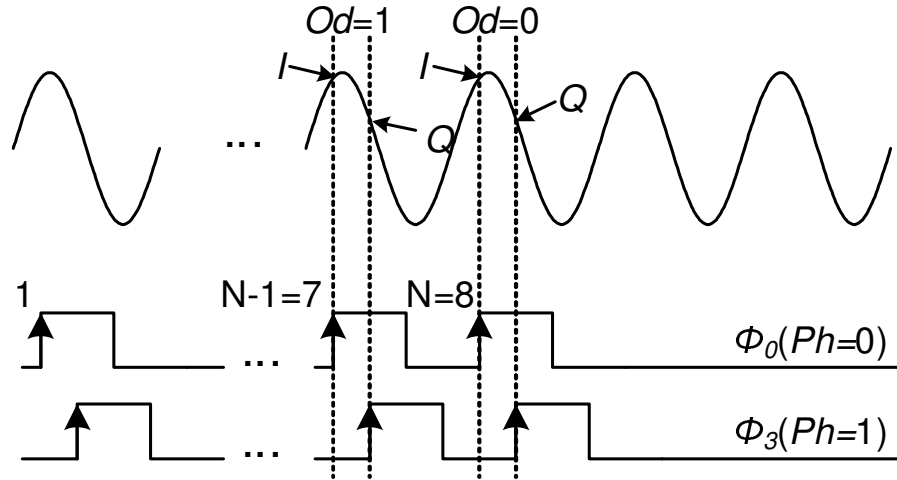


Figure 2.4: I/Q measurement.

2.1.3 Frequency Response Measurement Procedure

Instead of the envelope detector adopted in [51], I/Q detection is used in this design; hence, the amplitudes of the CUO input V_{EXC} and CUO output V_{CUO} are calculated by a digital arithmetic circuit. The pipelined data processing block, shown in Fig. 2.2, is described in detail in Fig. 2.5a. In order to reduce the complexity of the digital circuit, the arithmetic unit is multiplexed and the register resources are

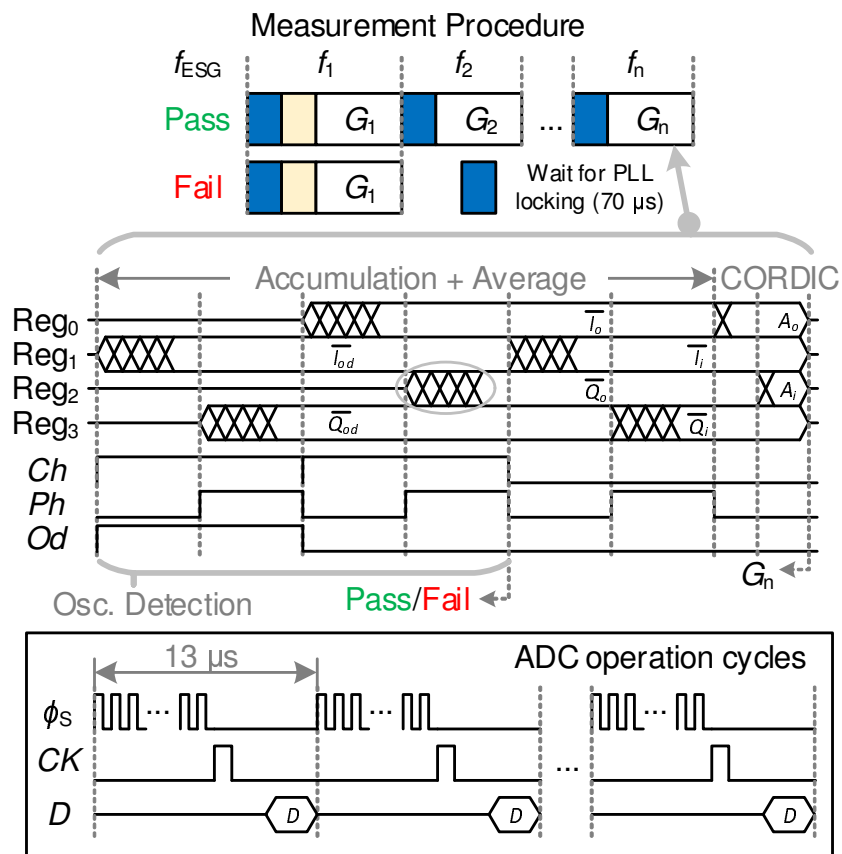
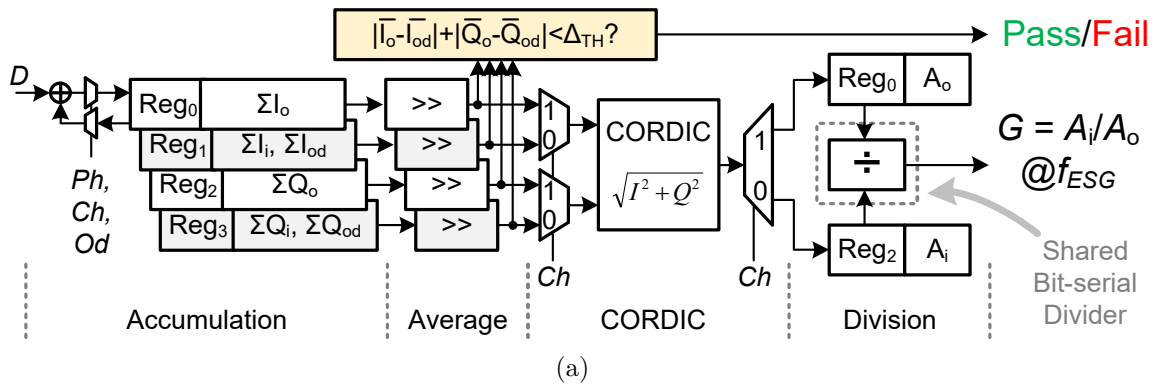


Figure 2.5: Frequency response measurement procedure: (a) Gain computation data flow and its (b) timing plot.

reused. The temporal measurement procedure is depicted in Fig. 2.5b. The complete measurement consists of multi-cycle ADC operations. In each cycle, the ϕ_S in the ORA waits till the CUO is stable, and then samples the sinusoidal waveform at the N -th or $(N - 1)$ -th rising edge. $N = 8$ is chosen in this design for the CUO settling. After the down-conversion is done, a CK pulse triggers the ADC to digitize the DC level of V_{samp} , outputting D . One ADC operation cycle takes about $13 \mu s$. It is executed multiple times and D is accumulated in a selected register. A combination of signals Ch , Ph , and Od defines the meaning of the measured values. The sampled I/Q values of the output ($Ch = 1$) I_o and Q_o are stored in Reg_0 and Reg_2 , respectively. While for the input ($Ch = 0$), I_i and Q_i are accumulated in Reg_1 and Reg_3 . After the accumulation, the data in the register is averaged. In this design, I/Q values are accumulated eight times and then right-shifted by 3 bits. Following the averaging computation, a CORDIC algorithm [52], is implemented to retrieve the amplitude. The outputs of the CORDIC algorithm are the CUO input and output amplitudes, A_i and A_o , and they reuse the spaces of Reg_0 and Reg_2 . Later, the gain (G) of the CUO can be produced by a bit serial divider.

The complete frequency response measurement contains the quantization and magnitude gain computation of $G_1 \cdots G_n$ at n frequency characteristic points $f_1 \cdots f_n$. Before the I/Q detection at each frequency, it takes about $70 \mu s$ to wait until the frequency synthesizer is locked. Then, an oscillation/instability detection is carried out. It will tell whether the CUO is “PASS” or “FAIL.” The measurement will stop immediately if an oscillation or instability condition is detected. It should be noted that, the offset of the ORA and the ADC must be corrected before any measurement. Simply connecting the ORA’s input to the common mode voltage and running an ADC operation cycle could obtain the digitized offset value D_{OFF} . Later in all measurement steps, D_{OFF} but not 512 (for the 10-bit ADC) are used as the differential

zero reference.

2.1.4 Stability Check Procedure

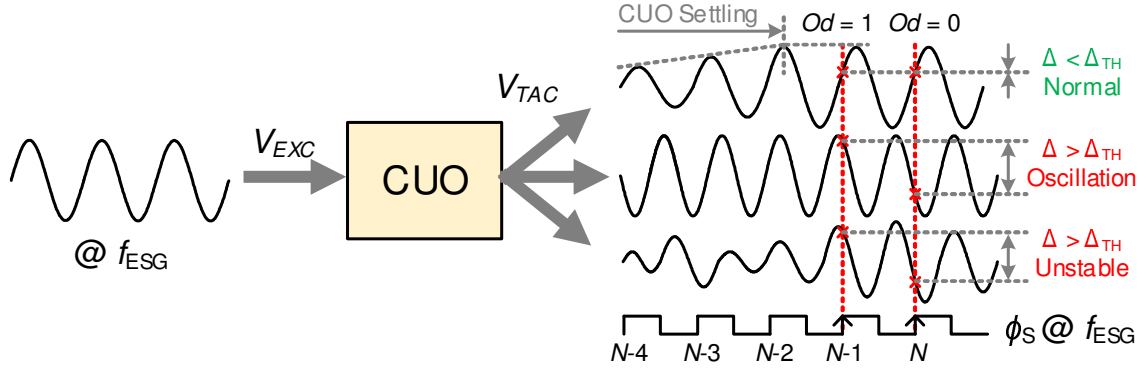


Figure 2.6: Principle of the oscillation detection.

An early warning mechanism is adopted to detect the oscillation/instability state of the CUO, as demonstrated in Fig. 2.6. This feature allows extreme designs while keeping the circuit stable. If, for instance, the CUO is working in a normal condition and is excited by a sinusoidal signal at a frequency of f_{ESG} , after the settling time, the two samples at the N -th and $(N - 1)$ -th rising edges of ϕ_S will lead to two identical DC levels, whose difference is much smaller than a threshold value Δ_{TH} . To the contrary, the CUO may self-oscillate at a frequency different from f_{ESG} or work in an unstable state. In these circumstances, two consecutive samples will lead to a difference value bigger than Δ_{TH} . The threshold is set to around 35 mV in this design. Fig. 2.5b shows that the oscillation/instability detection procedure adds two extra measurements of I_{od} and Q_{od} , which reuse Reg_0 and Reg_2 for storage when $O_d = 1$.

2.2 Optimization Procedure and Implementation

2.2.1 Cost Function Definition

As previously mentioned in Fig. 1.4, the cost (error) function $F(\mathbf{v})$ is defined as the difference between the actual performance y_{actual} and the desired specification $y_{desired}$,

$$F(\mathbf{v}) = \sum_{i=1}^m \alpha_i \|y_{actual,i} - \beta_i \cdot y_{desired,i}\|_{s_i} + P \quad (2.1)$$

where \mathbf{v} is the multi-dimensional design vector, m is the number of performance data points, $\|\cdot\|_{s_i}$ represents the s_i -norm operation, α_i is the i -th sizing factor of the normalized term, β_i is the i -th sizing factor of the desired specification, and P is a penalty function. The penalty function can be used to penalize $F(\mathbf{v})$ for some performance that can not be quantified (e.g. stability). With the help of the stability check procedure proposed in Section 2.1.4, (2.3) will demonstrate such an example. When designing a cost function for a specific CUO, one important rule should be followed. That is $F(\mathbf{v})$ would achieve its minimal value $F(\mathbf{v}_{best})$, ideally zero, if and only if the desired specification is completely satisfied under the design vector \mathbf{v}_{best} .

Consider a 2nd-order BPF frequency response as an example. An intuitive definition of the cost function for matching the desired frequency response is the least mean square (LMS) error. On the one hand, implementing such a cost function in digital domain requires the computation of 2-norm, which means costly floating point operations of square and square root. On the other hand, a detailed frequency sweep is required to evaluate the shape of the frequency response, leading to long measurement time. Noting that we only care about the minimum value of $F(\mathbf{v})$, absolute operation (1-norm), which consumes less computation resources, can be used. And instead of fitting the whole frequency spectrum, a limited number of key characteristic frequency points can be extracted, not only reducing the data amount

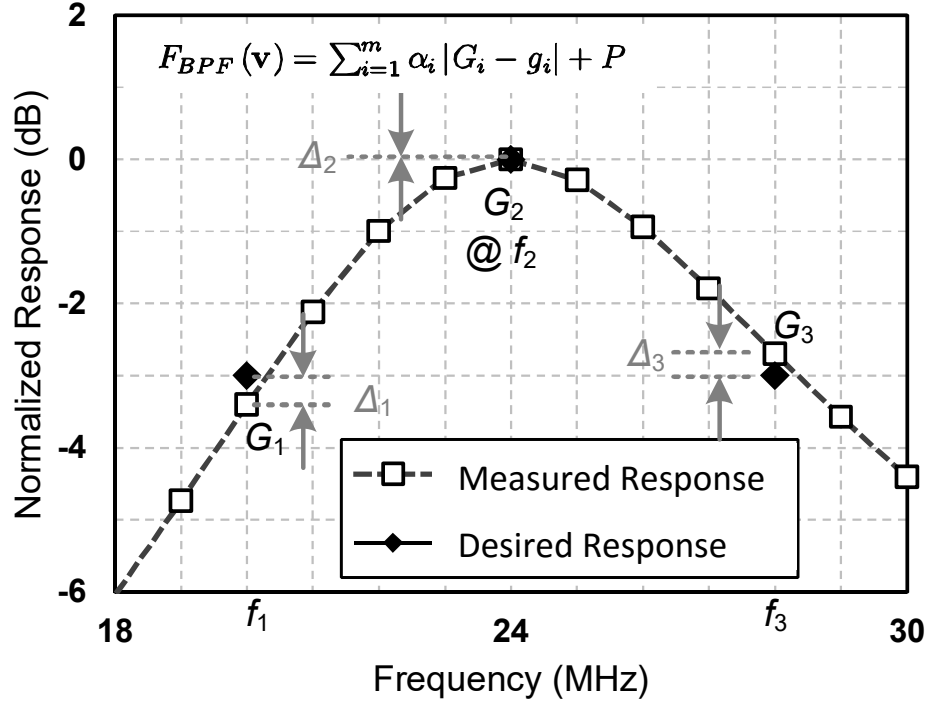


Figure 2.7: Definition of a 2nd-order BPF frequency response matching cost function.

but also accelerating the computing procedure. Therefore, a general cost function $F_{BPF}(\mathbf{v})$ for frequency response matching is shown in Fig. 2.7. where G_i (a.k.a. $y_{actual,i}$) is the measured gain at the frequency of f_i , g_i (a.k.a. $y_{desired,i}$) is the desired gain at f_i , and P is a penalty term, which will be discussed later. Particularly, α_i is the coefficient of each absolute value term, and it can be used to emphasize some specific frequency points. $F_{BPF}(\mathbf{v})$ has only addition and subtraction, reducing the arithmetic operation complexity.

2.2.1.1 $F(\mathbf{v})$ for a 2nd order biquad

In practice, we have to make some modification to the general $F_{BPF}(\mathbf{v})$. Considering that Q enhancement may boost the gain of the BPF while the gain is easy to set by changing R_K , the absolute gain matching can be relaxed. It can then be

rewritten as

$$\begin{aligned}
 F_{2nd}(\mathbf{v}) &= \alpha_1 |G_{i1} - \sqrt{2}G_{b1}| + \alpha_2 |G_{i1} - \sqrt{2}G_{b2}| \\
 &\quad + \alpha_3 |G_{b1} - G_{b2}| + P_{2nd}
 \end{aligned} \tag{2.2}$$

where the gains (G) at three frequency points, f_{b1} , f_{b2} (“b” for the 3dB frequencies), and the central frequency f_{i1} (“i” for “in-band”) are used. Note that $Q_a = f_{i1}/(f_{b2}-f_{b1})$. In (2.2), the first two terms emphasize the 3dB roll-off, and the third term imposes the symmetry on the transfer function. In the optimization theory, penalty function is a technique for transforming a constrained optimization problem into an unconstrained one. Here the penalty term is derived from the oscillation/instability detection criteria and the gain matching

$$P_{2nd} = \begin{cases} \infty & \text{if unstable,} \\ 1 - G_{i1} & \text{if } G_{i1} < 1, \\ 0 & \text{otherwise.} \end{cases} \tag{2.3}$$

Experimental results in Section 2.5 show the feasibility of such a cost function.

2.2.1.2 $F(\mathbf{v})$ for a 4th and higher order BPF

Similar to (2.2), we can further design the cost function for the 4th or higher order. Apart from emphasizing the bandwidth and the symmetry, in-band flatness and out-of-band suppression are also considered. The cost function can be written

as

$$\begin{aligned}
F_{4th}(\mathbf{v}) &= \alpha_1 |G_{i3} - \sqrt{2}G_{b1}| + \alpha_2 |G_{i1} - \sqrt{2}G_{b2}| \\
&\quad + \alpha_3 |G_{i1} - G_{i2}| + \alpha_4 |G_{i2} - G_{i3}| \\
&\quad + \alpha_5 |G_{i1} - G_{i3}| + P_{4th}
\end{aligned} \tag{2.4}$$

where the subscripts of G represent the frequency points. Among them, f_{i1} , f_{i2} and f_{i3} define the left the center, and the right corner of the BPF's pass band, and f_{b1} and f_{b2} are the left and the right 3dB bandwidth frequencies. Each term in (2.4) can further be emphasized by the weight α_i . However, relying only on (2.4) may mislead the optimization procedure to such a case where only one of the two biquads works, and thus the response of the whole filter is exactly a 2nd order BPF. This unpractical case is because (2.4) focuses on matching the in-band response but ignores the out-of-band roll-off. To avoid such a dilemma, we should penalize the cost function if it doesn't behave as a 4th order filter

$$P_{4th} = P_{2nd} + \begin{cases} \beta_1 (G_{o1} - g_o) & \text{if } G_{o1} > g_o, \\ \beta_2 (G_{o2} - g_o) & \text{if } G_{o2} > g_o, \\ \dots & \dots, \end{cases} \tag{2.5}$$

where G_{o1}, G_{o2}, \dots are out-of-band (annotated by “o”) characteristic frequency points, and g_o is a threshold gain for the frequency response in the out-of-band region. If the suppression is not enough ($G_o > g_o$), $F_{4th}(\mathbf{v})$ will be penalized. This guarantees the filter order matching.

To sum up, the design of the cost function is to select some characteristic frequency points which may reflect key factors (f_0 , 3dB BW, in-band flatness, out-of-

band roll-off, etc.) in the BPF design, and then combine them into a sum of errors and penalties.

2.2.2 Combined meta-heuristic engine

Another key part of the optimization procedure is the digital meta-heuristic engine. Its first approach was adopted in [53]. In this design, a hybrid algorithm combining simulated annealing (SA) [29] and sensitivity search (SS) [27] is proposed to improve multi-dimensional searching ability. Fig. 2.8 illustrates the full optimization flow, which consists of MAX_SA iterations of SA and MAX_SS iterations of SS. In addition, SS and SA are chosen because of their low computation complexity and limited dependence on historical information, which means less memory consumption in the hardware implementation.

The SA iteration starts just after the initialization step, after which the iteration counter is reset. In Fig. 2.8, the gray block named “2-D variable selection” within the simulated annealing section in Fig. 2.8, is utilized to decompose the multi-dimensional problem into several two dimensional (2-D) problems. It groups two variables in the design vector \mathbf{v} to be changed by the optimization engine, while the other variables remain fixed if the problem dimension $k > 2$. Such a grouping pattern includes two variables out of k and is maintained during the next M SA iterations. After that, a new grouping pattern is generated arbitrarily and kept for another group of M SA iterations, and so on so forth. In this way, the dimension of the large scale problem is reduced to two at the time, thereby limiting the complexity of the optimization hardware implementation. To be more specific, in each SA iteration, the two selected variables x_1 and x_2 of \mathbf{v} are changed into x'_1 and x'_2 , where $(x'_1, x'_2) = (x_1, x_2) + (a, b)$ and $|a| + |b|$ is defined equal to the temperature T . The signs of a and b define the direction, and their magnitudes denote the step size.

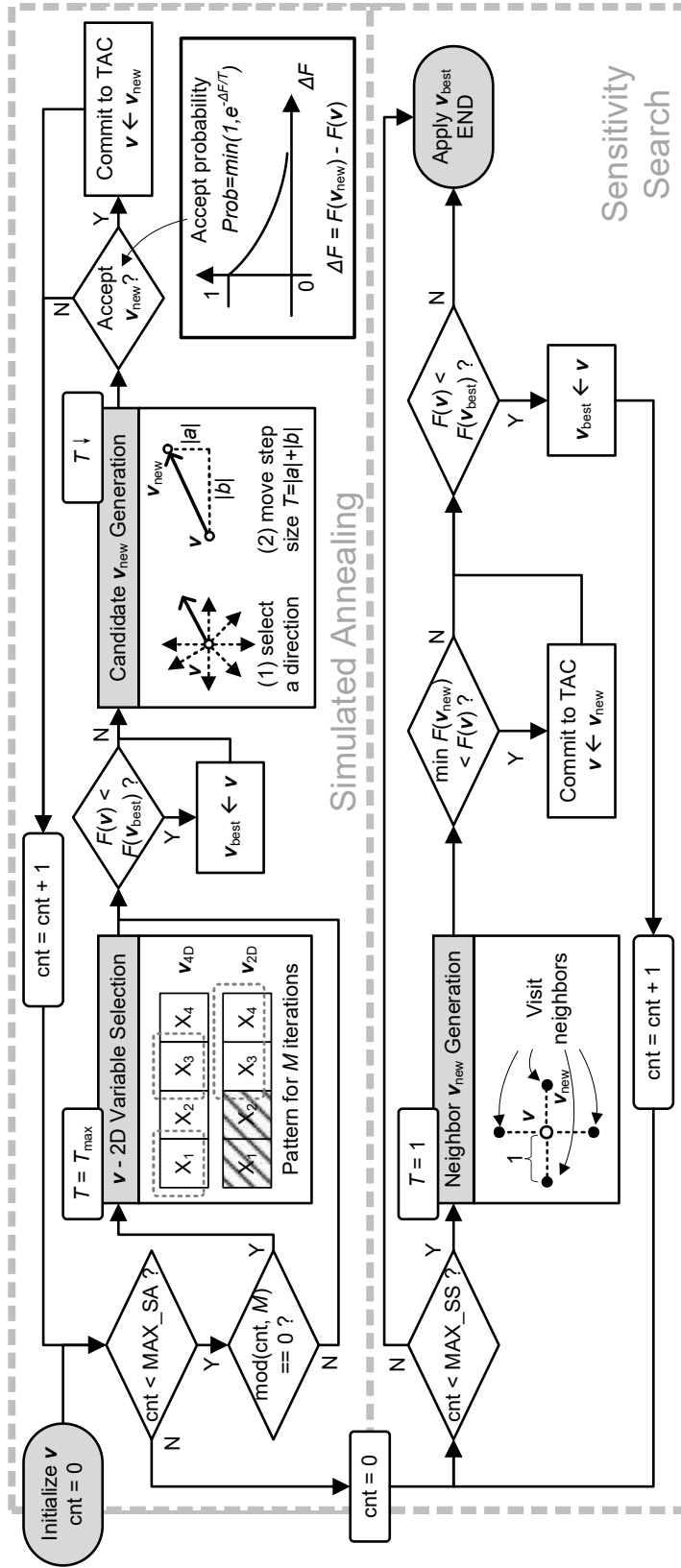


Figure 2.8: Work flow of the proposed combined meta-heuristic algorithm.

A new vector \mathbf{v}_{new} is obtained in the candidate generation. Then, \mathbf{v}_{new} is submitted to CUO as design variables and substituted for the design vector \mathbf{v} according to a probability which is a function of $F(\mathbf{v})$ change and T . Additionally, T is reset to T_{max} at the beginning of each 2-D variable selection, and decreases gradually after every SA iteration. This process simulates the annealing in the physical world. In other words, the random acceptance of \mathbf{v}_{new} and the changing temperature T give the algorithm a chance to escape the local minima and approach the global optima. Although this means a better performance in searching global optima, SA shows a slower convergence rate compared to the SS algorithm mentioned next. Following the SA iterations, the sensitivity search starts to work on the basis of \mathbf{v}_{best} found by SA. Within every SS iteration, all the neighbors of \mathbf{v} (for instance, four neighbors in the 2-D problem, $(x_1 \pm 1, x_2)$ and $(x_1, x_2 \pm 1)$), are evaluated and the one that leads to the minimum cost is further checked against $F(\mathbf{v})$. Only if $F(\mathbf{v}_{new}) < F(\mathbf{v})$ will \mathbf{v}_{new} be updated to \mathbf{v} and a new round of searches start. Otherwise, the algorithm terminates. In contrast to the SA algorithm, the stand-alone SS tends to converge fast but it is prone to be trapped into local optima. However, by taking the advantage of SA's result, SS improves the performance of a fast and global optima search. During the execution of SA and SS algorithms, the best design vector \mathbf{v}_{best} is updated by \mathbf{v} whenever a smaller $F(\mathbf{v})$ is available. Finally, when both algorithms are done, \mathbf{v}_{best} is applied to the CUO.

By employing SA as a global explorer and SS as a detailed explorer, the combined approach reaches a balance between solution quality and search speed. This combined algorithm is completely implemented in digital circuits. Moreover, the art of the meta-heuristic engine design should be emphasized here – the same engine can be reused for different cost functions, different design vectors and significantly different CUOs.

2.3 CUO Study Cases—Active Filters

To show the feasibility of the conceptual system proposed in Fig. 1.4, which conducts a self-contained in-situ analog circuit design based on a black-box validation and optimization, two study cases, an active-RC BPF with reduced GB values and a Gm-C BPF are illustrated in this section. However, this system-level concept can be further applied to any analog/RF circuit, and future works will address noise and linearity constraints.

Traditionally, filter tuning techniques can achieve accurate response calibration with low overhead hardware as in [42] and [54], or the master-slave approach [55]. Nevertheless, [42] and [54] assume ideal amplifier implementation, while [55] requires a replica filter cell whose condition may be different from the original circuit due to mismatches and PVT variations. Moreover, all these tuning methods handle only a single design variable, such as the capacitance in [42], the resistor array in [54], and the Gm in [55]. Different from these classical methods, the proposed VO concept is based on a black box approach rather than a transistor-level model. It first digitizes the input and output signals of the circuit under optimization (CUO) in-situ, taking all non-ideality into consideration. Based on the measured data, the error between the desired and actual circuit performance is computed using a cost function, which combines and emphasizes different design specifications. Particularly, instability detection is proposed as a penalty of the cost function. It is able to keep the CUO stable and work under extreme conditions. An optimization engine is further implemented to minimize the error (cost function), yielding the optimal design parameters. Multiple independent parameters can be adjusted on demand as needed. In the proposed filter study cases below where no assumption on ideal amplifiers is made, several RC time constants of the active-RC biquad with small

GB s and transconductance values of the Gm-C biquad are tuned respectively within 2-D or 4-D optimization. Furthermore, a digital-to-analog conversion is required in the proposed approach at the expense of some silicon areas. This overhead can be leveraged at the system level by reusing the same validation circuit for different CUOs. In addition, the intensive digital implementation of the cost function and the optimization engine will benefit from the area shrink in small-size IC technologies.

2.3.1 Non-ideal Active-RC Biquad

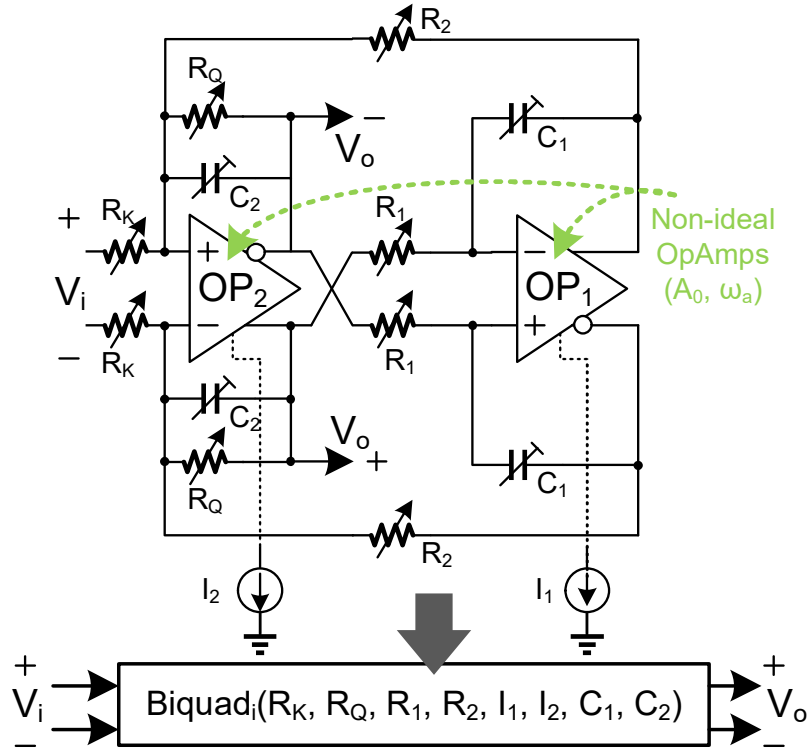


Figure 2.9: Topology and design parameters of an active-RC BPF biquad.

Ideal OpAmps are usually assumed in the design of active-RC BPF, whose topology is shown in Fig. 2.9. We can set the central frequency $\omega_0 = 1/\sqrt{R_1 R_2 C_1 C_2}$ and the

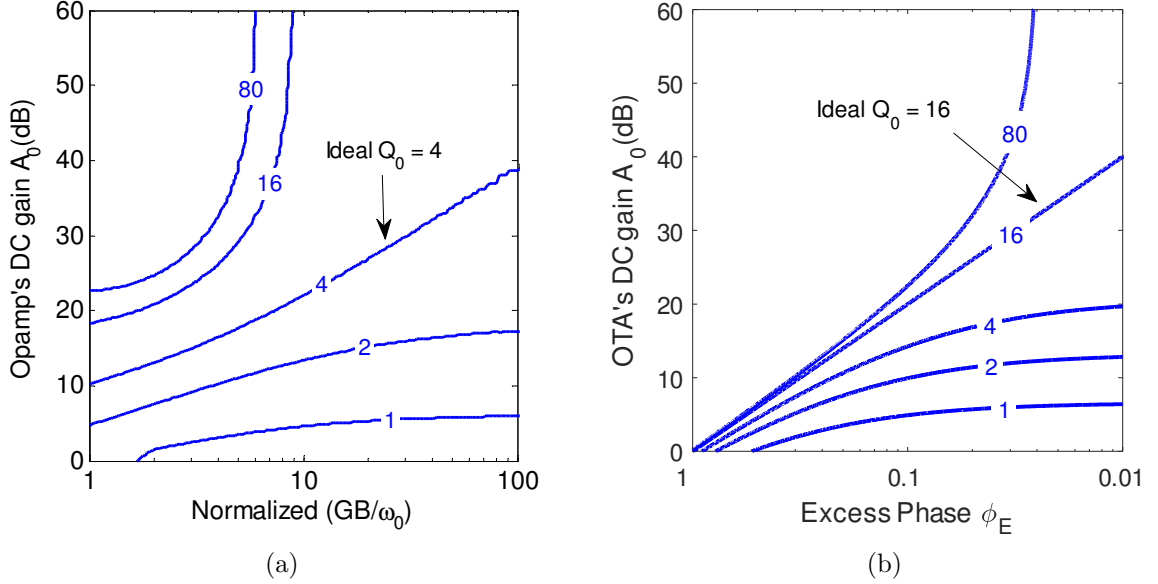


Figure 2.10: The actual Q_a contour for (a) ideal $Q_0 = 4$ (active-RC) and (b) ideal $Q_0 = 16$ (Gm-C, $\phi_E = \omega_0/BW_O$, where BW_O is the bandwidth of the OTA).

nominal Q factor $Q_0 = \sqrt{\frac{R_Q^2 C_1}{R_1 R_2 C_2}}$. The BPF transfer function deviates from the ideal case if the real OpAmp has limited DC gain A_0 and bandwidth ω_a . The actual Q factor of a biquad [56] can be approximated as

$$Q_a = Q_0 \cdot \left(1 + \frac{2Q_0}{A_0} - 4Q_0 \cdot \frac{\omega_0}{GB}\right)^{-1} \quad (2.6)$$

where the gain-bandwidth product $GB = A_0 \cdot \omega_a$. The numerical simulation results in Fig. 2.10a for $Q_0 = 4$ show the contours of Q_a for different A_0 and normalized GB/ω_0 . It should be noted that OpAmps with a relatively low DC gain but extended GB [57] may degrade the actual Q_a . To the contrary, if the GB is too small, Q_a will rise sharply and soon the circuit becomes unstable. Therefore, a proper GB should be found to avoid instability or excessive power consumption. As a result, classical

tuning methods, such as the one considering only passive components [42], or one that assumes a good match between the actual and the theoretical phase shift [54], are not applicable. Instead, the proposed black-box optimization solution can tackle such a non-ideal case. It heuristically explores the CUO’s ability to meet the desired performance at a proper cost; thus, it is possible to scale the power for different performance requirements even if the amplifier is far from ideal.

In this design, OP_1 and OP_2 intentionally implement the Miller-compensated two-stage OpAmps [58] with a low DC gain around 30 dB. Bias currents I_1 and I_2 are used to affect the OpAmps’ A_0 and ω_a . Moreover, resistor arrays (R_K , R_Q , R_1 , and R_2), and capacitor arrays (C_1 and C_2) are implemented as design variables. The resistance value is defined as $R_{base} + K_R \cdot R_u$, where K_R is the 5-bit control word. Similarly, the capacitor array adopts $C_{base} + K_C \cdot C_u$. Design values are summarized in Table 2.1. Particularly, classical filter design usually sets all resistors proportional to a fixed unit resistance [42]. In this design, R_u are different between R_K, R_Q and R_1, R_2 so as to demonstrate the flexibility of the proposed optimization algorithm. Section 2.5 will verify the design parameter flexibility meeting specifications and trade-offs.

2.3.2 Non-ideal Gm-C Biquad

The Gm-C BPF transfer function is also complicated by non-ideal operational transconductance amplifiers (OTA). Similar to (2.6), the actual Q factor [59] can be described as

$$Q_a = Q_0 \cdot \left[1 + 2Q_0 \left(\frac{1}{A_0} - \phi_E \right) \right]^{-1} \quad (2.7)$$

where A_0 and ϕ_E are the finite DC gain and the excess phase introduced by the OTA, while the nominal $Q_0 = g_{m2}/g_{m1}$. The contour of (2.7) when $Q_0 = 16$ is illustrated

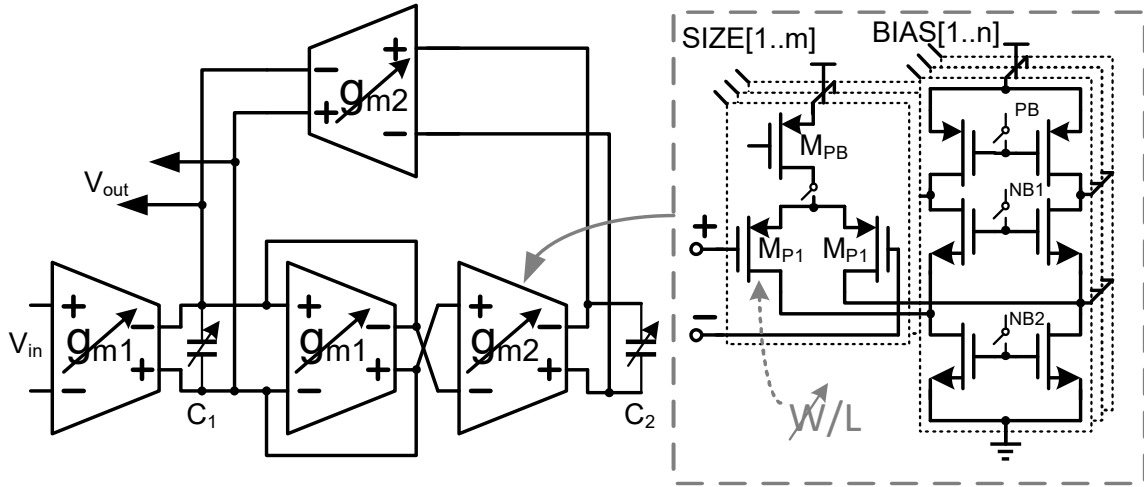


Figure 2.11: Schematic of the Gm-C biquad.

in Fig. 2.10b. As the OTA non-ideality leads to performance deviation and risk of instability, a Gm-C biquad is also implemented as a study case. Folded-cascode OTAs, which are shown in Fig. 2.11, are adopted with both the programmable differential pair and the bias branches. On one hand, multiple folded differential pair transistors are connected to or disconnected from the cascode branches, controlled by the signal *SIZE*. This is equivalent to a tunable width for a single differential pair. On the other hand, signal *BIAS* enables the cascode branches, which equivalently changes the bias current. The combination of *SIZE* and *BIAS* results in a wide programmable g_m in the Gm-C biquad. Furthermore, programmable C_1 and C_2 are implemented.

2.4 Precision Analysis

As mentioned above, self-validation technique is implemented to conduct measurements of the circuit performance metrics, and then a cost function is calculated based on these metrics. Fig. 2.12 demonstrates the abstraction of the signal path from the ESG to the cost function. Because the cost function is used as the baseline

for evaluating the circuit performance, it is necessary to analyze different factors that impact the precision in each measurement/computation steps. To obtain the systematic signal-to-noise ratio (SNR_{SYS}), the measurement error (SNR_{MEAS}) and the digital computation error (SNR_{DIG}) will be discussed individually.

2.4.1 Measurement Errors

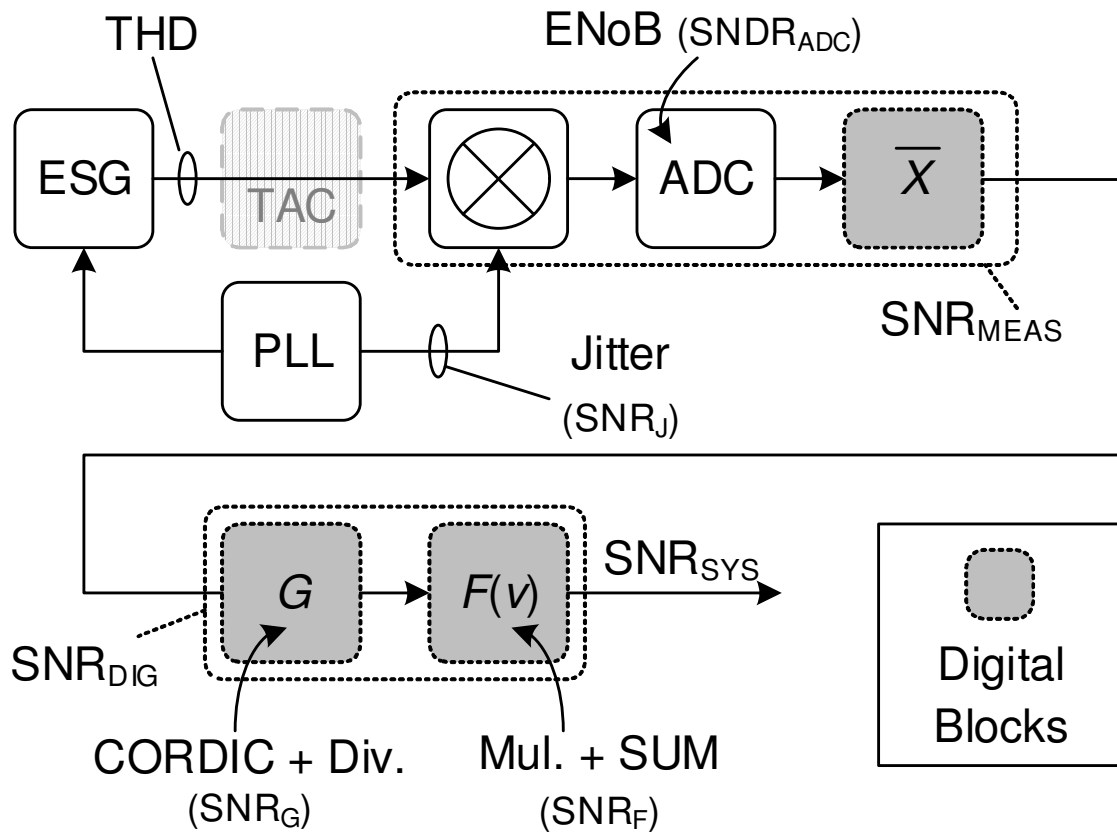


Figure 2.12: Abstracted model for the precision analysis.

The first error source of the self-validation path comes from the ESG. As the excitation sinusoidal waveform is synthesized from square waves, its distortion will contribute to the measurement error. However, measurement result (Fig. 2.13) shows

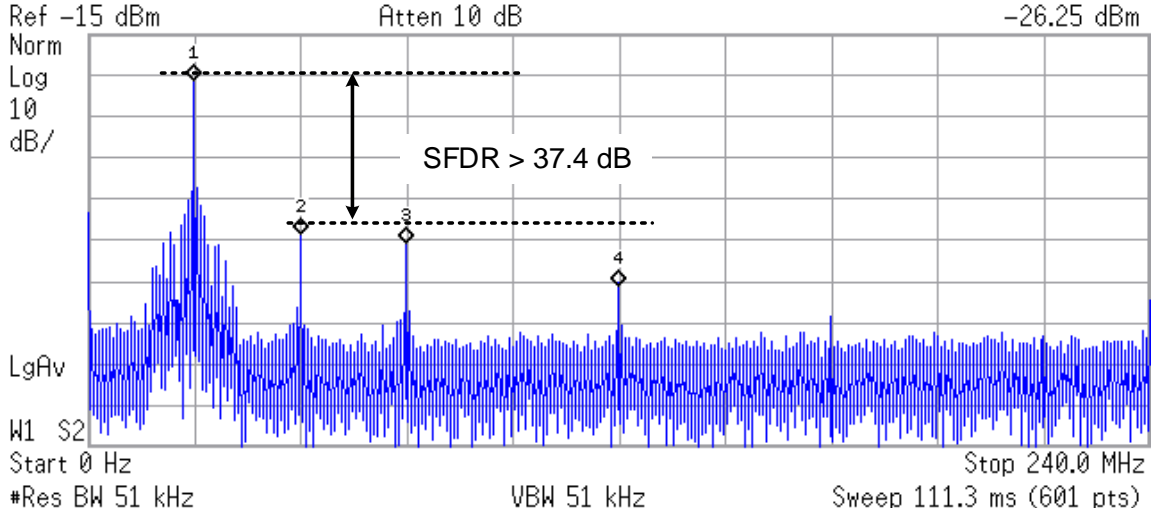


Figure 2.13: Measured SFDR of the excitation sine-wave at 24 MHz.

better than 37.4 dB spur free dynamic range (SFDR) is obtained by the ESG, leading to less than 2% total harmonic distortion (THD). Although noise at all harmonics will be folded back to the DC due to the self-mixing mechanism, the extra noise is not dominant thanks to the high linearity.

Nevertheless, another path, the sampling clock ϕ_S in the PMQ block is the dominant source of the measurement error, because of the clock jitter. As analyzed in [60], the noise induced by PLL's jitter in sampling could be evaluated as

$$SNR_J = -20 \log \left[\frac{f_{in}}{f_S} \sqrt{\int_{-\infty}^{+\infty} 10^{L(f-f_{in})/10} df} \right] \quad (2.8)$$

where f_{in} is the frequency of the input sinusoidal waveform, and f_S is the sampling frequency. $L(f)$ represents the single-side-band (SSB) phase noise of the sampling clock generated by a PLL. About -83 dBc/Hz in-band phase is measured from the PLL at 24 MHz f_{ESG} , and the loop BW of the PLL is set to 100 kHz. Thus, we have a typical $SNR_J = 28$ dB, which is the major error source in the whole system.

To reduce this jitter-induced noise, several steps could be taken into consideration. The most effective way is to reduce the in-band phase noise of the PLL, which may increase the power consumption of the charge pump and the phase frequency detector (PFD). Reducing loop BW can also improve the SNR_J , however, larger area and longer PLL locking time is the tradeoff. Additionally, digital averaging filter (\bar{X}) is a simple solution.

Following the down-conversion block, a 10-bit SAR ADC is implemented. As reported in [50], the ADC's effective number of bits (ENoB) is 9.2 bits, which represents a signal-to-noise and distortion ratio $SNDR_{ADC} = 57.1 \text{ dB}$. A digital averaging filter, \bar{X} , is further adopted to accumulate and average the measured data, which is previously illustrated in Fig. 2.5a. This averaging filter could improve the SNR by N times, where N is the number of samples taken for generating one average number. Consider a combination of the down-conversion S/H circuit, the SAR ADC and the averaging filter, its SNR could be derived as

$$SNR_{MEAS} = -10 \log \left(10^{\frac{-SNR_J}{10}} + 10^{\frac{-SNDR_{ADC}}{10}} \right) + 10 \log N \quad (2.9)$$

Therefore, we have $SNR_{MEAS} = 37 \text{ dB}$ for 8x averaging, while $SNR_{MEAS} = 40 \text{ dB}$ for 16x averaging.

2.4.2 Computational Error

As shown in Fig. 2.12, digital fixed-point computation procedures, which consist of a gain computation stage (G) and a cost function ($F(\mathbf{v})$) evaluation block, are implemented to process the ADC-measured data. The gain computation (G) is introduced in the right part of Fig. 2.5a, including the CORDIC and the division. For simplicity, the following analysis is based on Eq. 2.2.

Error occurs during these computation procedures. [61] introduces the fixed-

point error analysis, which will be applied in the proposed design. Similar to the definition in the analog block, the digital fixed-point error could also be presented by the SNR, although it is slightly different. The fixed point error SNR is defined as $SNR = -10 \log(P_{signal}/P_{error})$, where P_{signal} is the signal power and P_{error} is the noise power. Both of them could be further formulated as

$$P_{signal} = \frac{1}{n} \sum_{i=1}^n Y_i^2, P_{error} = \frac{1}{n} \sum_{i=1}^n (\hat{Y}_i - Y_i)^2 \quad (2.10)$$

In the equations above, uniform distributed input number is quantized, goes through the fixed point computation process and then produces the fixed-point values \hat{Y}_i . Correspondingly, the ideal numerical result, Y_i , is obtained by applying Eq. 2.2.

It can be found in Fig. 2.5a, the gain computation is divided into several steps, thereby SNR_G is affected by the bit-width configurations among these steps. We denote the ADC bit-width (which is also the bit-width of register Reg_i), CORDIC and division modules as $\{IB_A, FB_A\}$, $\{IB_C, FB_C\}$ and $\{IB_D, FB_D\}$, respectively. IB represents the integer bit-width and FB is the fraction bit-width. For simplicity, IB is assumed to be fixed. This assumption holds if the full scale of the ADC output is normalized to 1, which means $IB_A = 0$. Since the root of sum of squares process accomplished by CORDIC module will at most enlarge the ADC measured-data by $\sqrt{2}$ times, thus we have $IB_C = 1$. Moreover, IB_D limits the maximum gain that can be evaluated. Considering the relaxed gain constraint in Eq. 2.3, we set $IB_D = 2$, and thus a gain higher than four will be truncated. Therefore, the discussion below focuses only on the fractional bit-width. In addition, simulation result shows that FB_C is not necessary to be larger than FB_A , because wider FB_C has low impact on SNR_G . Thus, in the fixed point simulation, $FB_A = FB_C$ is assumed.

Fixed-point simulation of the gain computation stage is carried out in Matlab.

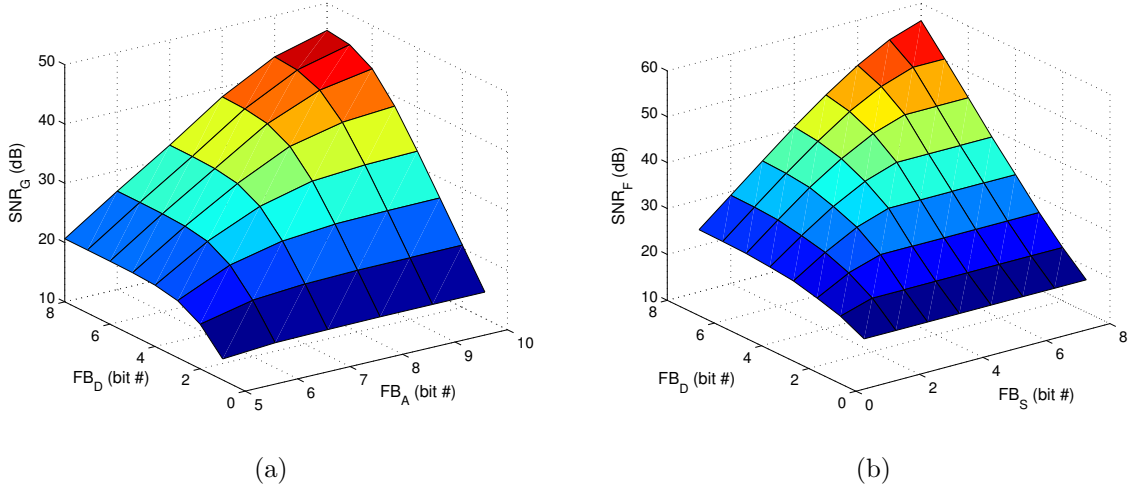


Figure 2.14: Simulation results of the computational errors: (a) Simulated SNR_G by sweeping the ADC and the division fractional bit-width in the gain computation stage and (b) simulated SNR_F by sweeping the division and $F(v)$ fractional bit-width in the cost function computation stage.

By sweeping FB_A from 5 to 10 and FB_D from 1 to 8, Fig. 2.14a demonstrates the change of SNR_G , which benefits from the increasing bit-width of both FB_A and FB_D . Furthermore, it could be found that, for a fixed FB_A , the SNR may saturate after some FB_D . Selecting FB_D at these corner points can help to avoid over design of the divider module, which occupies much area in the digital circuit. In the proposed design, $FB_A = 10$ and $FB_D = 7$ are implemented, leading to $SNR_G = 44.9$ dB.

SNR_F is determined by the following $F(\mathbf{v})$ stage. The input of $F(\mathbf{v})$ inherits the divider's output from the gain computation stage. We further examine the bit-width of the final output, $\{IB_S, FB_S\}$. On the one hand, IB_S will be extended due to the summation operation in $F(\mathbf{v})$. However, because all the weights are constants, no additional degree of freedom is introduced to the integer part. In other words, IB_S is still considered fixed. On the other hand, SNR is simulated by sweeping FB_D and

FB_S from 1 to 8, and the plot is shown in Fig. 2.14b. Similarly, the corner position could be found, where $FB_S = FB_D$ roughly. $FB_D = FB_S = 7$ is chosen in the proposed design, which achieves 50.4 dB SNR_F . Additionally, it should be noted that only the minimum value of $F(\mathbf{v})$ matters for the optimization engine, and thus we don't need a full bit-width of $F(\mathbf{v})$. In fact, only the eight least significant digits are enough, making $IB_S = 1$ and $FB_S = 7$.

In short, the bit-width choices in the proposed design are $\{0,10\}$, $\{2,7\}$, $\{1,7\}$ for the ADC measured data, the division and $F(\mathbf{v})$ outcome, respectively. Thus, $SNR_{DIG} = 42.2$ dB is obtained for the whole digital optimization path. It can be found that SNR_G is 6 dB lower than SNR_F , which means the gain computation stage dominates the digital performance. A further improvement could be made to save the area overhead, although the proposed parameters are relatively optimal.

2.4.3 System Analysis

Finally, the performance of the whole VO system can be defined as

$$SNR_{SYS} = -10 \log \left(10^{\frac{-SNR_{MEAS}}{10}} + 10^{\frac{-SNR_{DIG}}{10}} \right) \quad (2.11)$$

Based on the proposed design parameters, the VO system could achieve 36 dB (8x averaging) or 38 dB (16x averaging) total SNR. A summary of all system parameters is listed in Table 2.1.

2.5 Experimental Results

The proposed validation-optimization system is fabricated in 0.18 μm standard CMOS technology. Its die micrograph is demonstrated in Fig. 2.15. Analysis shows it achieves a total of 36 dB system-level SNR, which is enough for the optimization procedure. Moreover, a comparison is summarized in Table 2.2.

Table 2.1: Summary of System Design Parameters

Excitation Signal Generator (ESG)	
Reference Clock CLK_{REF}	1 MHz
PLL Output Clock CLK_{HF}	128 to 508 MHz
Sine-wave Frequency f_{ESG}	10.67 to 42.33 MHz
f_{ESG} Sweep Resolution	333 kHz
Output Response Analyzer (ORA)	
Sampling Frequency	f_{ESG}
CUO Settling Cycles (N)	7 or 8
Quantization Block	
ADC Bitwidth	10
Quantization Averaging #	8
CORDIC Bitwidth	10
Division Bitwidth	9
System SNR (dB)	36
Circuit-under-optimization: Active-RC	
R_K, R_Q	$865 \Omega + K_R \times 865 \Omega$
R_1, R_2	$1.23 k\Omega + K_R \times 445 \Omega$
C_1, C_2	$50 fF + K_C \times 112 fF$

2.5.1 2-D Problem: 2 Decision Variables

Fig. 2.16a shows the measured frequency response of an active-RC BPF biquad using non-ideal OpAmp parameters on purpose. When the bias current (I_1 and I_2 in Fig. 2.9) is reduced, f_0 is shifted, and the Q_a factor arises, see (2.6). A further power decrement results in a distorted response curve, which may make the biquad unstable. The oscillation/instability detection mechanism and the emphasis of symmetry in $F(\mathbf{v})$ definition can avoid these circumstances. By setting $R_K = R_Q = X_1$ and $R_1 = R_2 = X_2$ in Fig. 2.9, and fixing C_1 and C_2 , a sweep of the design vector $\mathbf{v} = (X_1, X_2)$ leads to the 3D topography of $F(\mathbf{v})$, which is illustrated in Fig. 2.16b. Darker region indicates smaller values of $F(\mathbf{v})$. Black spots represent the candidate design vectors searched by the optimization engine, noting that around $\mathbf{v}_{best} = (20, 23)$,

Table 2.2: Comparison of the Optimization or Tuning Systems

Ref.	Application	Tuning Method/Algorithm	Control Mode	Dimension	Design Variables	Self-contained Stimulus
This work	Arbitrary specification matching of active filter	Sensitivity Search + Simulated Annealing	Digital	Multi-dimensional	Resistor/GBW* (active-RC) G_m array (Gm-C)	Yes
[36]	ADC capacitor matching improvement	Random Search	Digital	Multi-dimensional	Capacitor array	No
[37]	LNA performance enhancement **	Multiple Starting Points Gradient Search	Digital	Multi-dimensional	Capacitor array	No
[42]	Active filter tuning ***	Time Constant Matching	Mixed-Signal	1-D	Capacitor array	No
[54]	Active filter tuning ***	Phase Shift Matching	Mixed-Signal	1-D	Voltage-controlled resistor	No
[55]	Active filter tuning	Master-Slave Tuning	Mixed-Signal	1-D	Voltage-controlled OTA G_m	No

* OpAmps' tunable GBW provides an extra degree of freedom for power reduction.

** An external DSP is used.

*** These approaches assume the ideal amplifier implementation.

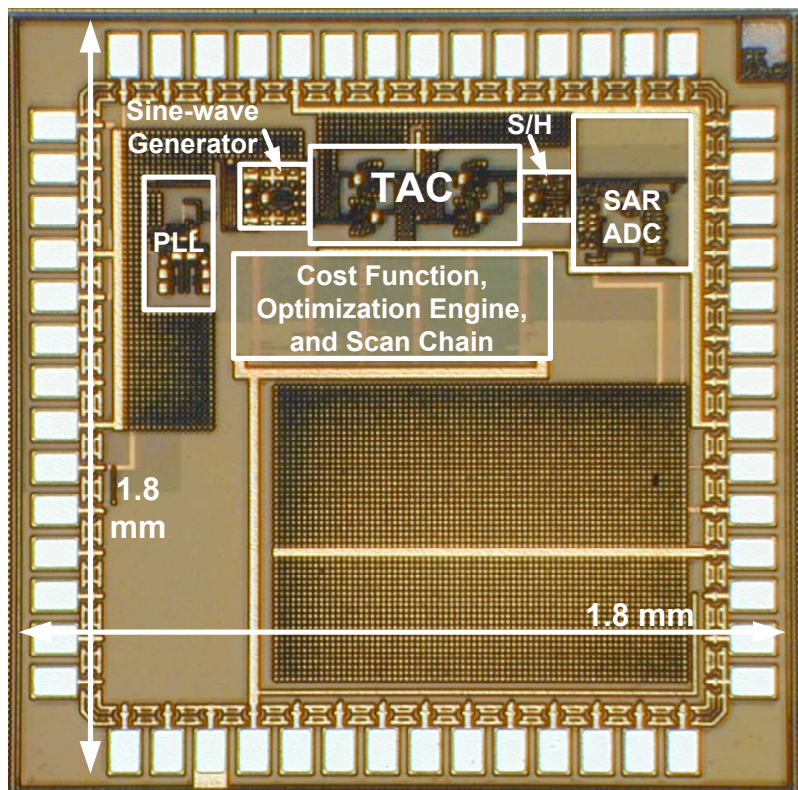


Figure 2.15: Chip die photograph of the proposed VO system. (CUO includes only the active-RC BPF. The Gm-C version is almost the same size.)

intensive search is applied due to the sensitivity search algorithm. The contour of $F(\mathbf{v})$ is also illustrated in Fig. 2.16c, along with the optimal \mathbf{v}_{best} . This value means $R_K = R_Q = 18.17 \text{ k}\Omega$ and $R_1 = R_2 = 11.47 \text{ k}\Omega$. Different from the brute-force full sweep, the optimization engine works in a smarter way that tries a limited number of solutions and concludes with the optimal value. Furthermore, this \mathbf{v}_{best} is obtained for the BPF design target $f_0 = 24 \text{ MHz}$, $BW = 8 \text{ MHz}$, and when the biquad consumes relatively large power, i.e. 2.1 mW. Simulation shows the OpAmps' GB to be around 1 GHz with this power. If we squeeze the power consumption by reducing the OpAmps' bias current, optimal matching is achieved at $\mathbf{v}_{best} = (14, 22)$ for 0.9 mW biquad power ($GB = 610 \text{ MHz}$), and $\mathbf{v}_{best} = (11, 19)$ for 0.6 mW ($GB =$

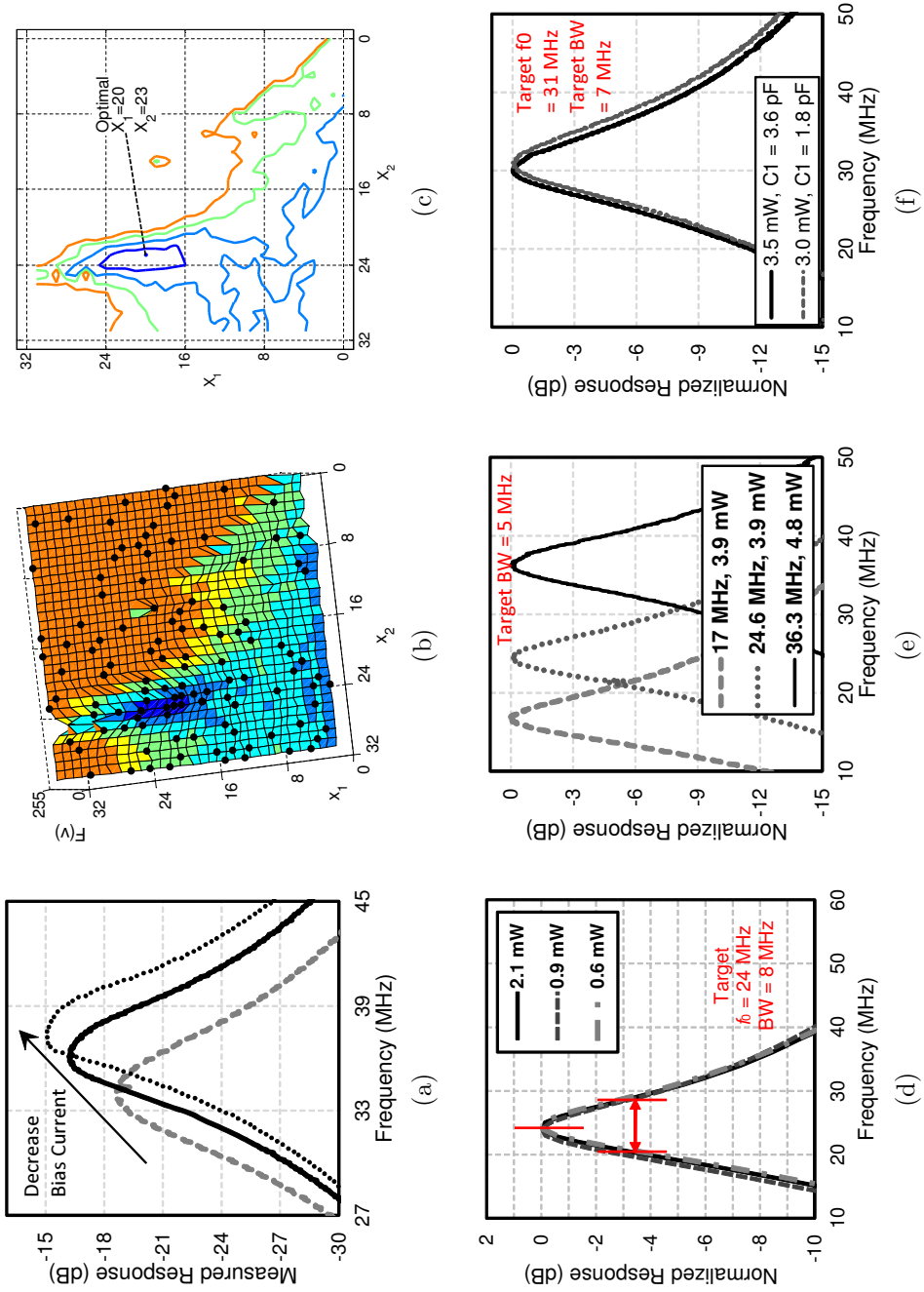
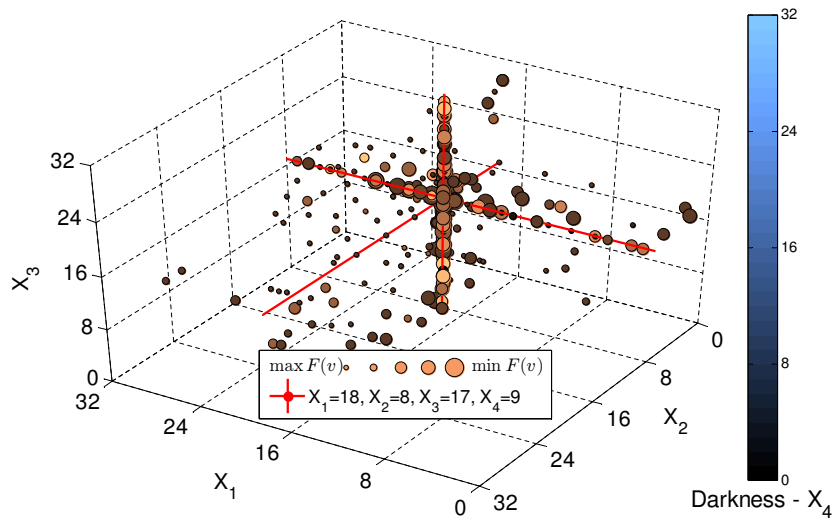


Figure 2.16: Experiment results: (a) Measured response impacted by biquad bias current, (b) $F(\mathbf{v})$ 3-D surface and the optimization steps (dots) ($X_1 = R_Q$ and $X_2 = R_K = R_1 = R_2$ for a biquad in Fig. 2.9), (c) $F(\mathbf{v})$ contour and the optimal \mathbf{v}_{best} , (d) fixed response matching with power sizing (active-RC, 2nd order), (e) f_0 shift with power sizing (active-RC, 2nd order), and (f) fixed response matching with capacitance sizing (Gm-C, 2nd order).

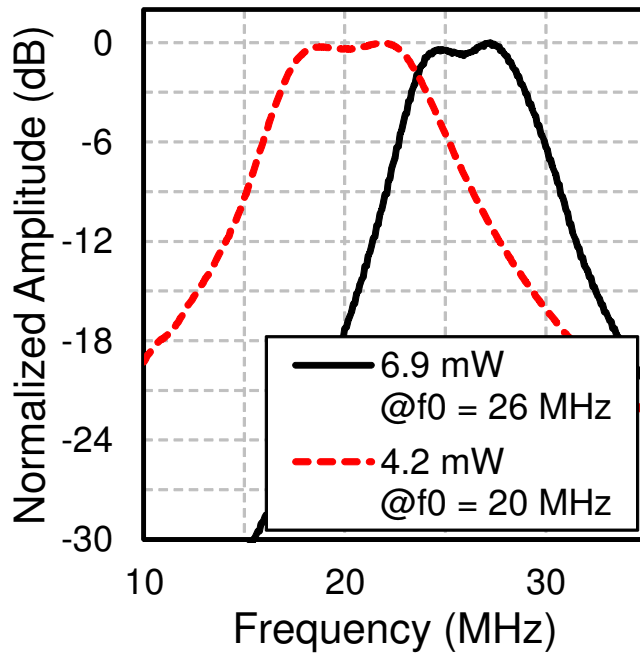
450 MHz). The algorithm will fail to find \mathbf{v}_{best} if the power is further reduced, which indicates a total failure of the CUO. The successful hits of the matching are shown in Fig. 2.16d for different power configurations, while providing less than 5% and 1.3% errors for the BW and f_0 . In addition, simulation results show that the input-referred IP3 (IIP3) of one biquad changes from +29 dBm to +23 dBm due to the power reduction. Another experiment is carried out with the fixed $BW = 5$ MHz target. The desired central frequency f_0 is shifted to 17.0, 24.6 and 36.3 MHz, as demonstrated in Fig. 2.16e. For all cases, the optimization engine can find the optimal design vectors that match the responses. The biquad power is also sized, 3.9 mW for f_0 of 17.0 and 24.6 MHz, which corresponds to about 1.33 GHz GB of OpAmps. But later it should be increased to 4.8 mW (1.41 GHz GB) and thus the optimization engine could push f_0 to 36.3 MHz. The identical VO system is also applied to a Gm-C biquad CUO. Design variables X_1 and X_2 are used to control g_{m1} and g_{m2} in Fig. 2.11 by switching variables $BIAS$ and $SIZE$. The frequency response matching result is shown in Fig. 2.16f for the target of 31 MHz f_0 and 7 MHz BW . For different load capacitances C_1 , the power consumption can be sized to obtain the correct response.

2.5.2 4-D Problem: Four Decision Variables

The optimization engine can further be applied to a 4-D problem, where two active-RC biquads are cascaded. Define $R_K = R_Q = X_1$ and $R_1 = R_2 = X_2$ for the first biquad stage, and $R_K = R_Q = X_3$ and $R_1 = R_2 = X_4$ for the second, thus we have the new design vector $\mathbf{v} = (X_1, X_2, X_3, X_4)$. The 4-D cost function (2.4) is currently implemented by an FPGA for the proof of concept, while the on-chip optimization engine is reused. An example searching procedure is shown in Fig. 2.17a. The value of $F(\mathbf{v})$ is described by the size of the circle, where a larger



(a)



(b)

Figure 2.17: 4-D optimization experimental results: (a) Optimization procedure of a 4th order Butterworth BPF ($X_1 = R_K = R_Q$ and $X_2 = R_1 = R_2$ for the 1st biquad stage, and $X_3 = R_K = R_Q$ and $X_4 = R_1 = R_2$ for the 2nd biquad stage) and (b) experimental response matching with power sizing (active-RC, 4th order).

circle represents a smaller $F(\mathbf{v})$ value. This procedure helps the optimization engine find a Butterworth filter response for $f_0 = 20 \text{ MHz}$ and $BW = 7 \text{ MHz}$ with a total 4.2 mW power consumption. Meanwhile a higher power, 6.9 mW, can help the engine find a solution for $f_0 = 26 \text{ MHz}$ and $BW = 6 \text{ MHz}$. The matched response is demonstrated in Fig. 2.17b.

2.6 Conclusion

A proof of the self-contained validation-optimization (VO) system concept has been presented. It makes the analog design robust against PVT variations, aging effects and even lack of transistor models by implementing a digital optimization engine as well as built-in self-validation circuits. The robust self-validation path consists of a digital based sine-wave generator, and the output signal is also converted to digital-domain. The proposed system illustrates that, not only the traditional design variables (time constants, trans-conductance, etc.) but also the unconventional design parameters, such as the GB of the OpAmp in active-RC biquads, can be incorporated to yield power reduction while meeting the design specifications. To conclude, the proposed VO system concept does not involve the internal operating principle of the target circuit, and thus, the system is applicable to other types of analog circuits, while the cost function should be redefined according to the performance desired.

3. HIGH-LINEARITY SINE-WAVE SYNTHESIZER ARCHITECTURE BASED ON FIR FILTER APPROACH AND SFDR OPTIMIZATION *

3.1 Motivation

High-linear sinusoidal signal generation is critical in some scenarios of IC design, such as the mixed-signal circuit testing and the electrochemical impedance spectroscopy (EIS) [62]. The test of a mixed-signal circuit relies on a sine-wave stimulation with variable frequency to measure the frequency response [51], examine the supply noise tolerance [43], evaluate an analog-to-digital converter (ADC) [63], or even characterize a whole RF receiver [64]. For the sine-wave generation, current research works focus on the trade-off among spectral purity, bandwidth, area, and power efficiency. Currently, wide-bandwidth active filters and high-speed ADCs continue to meet the evolving demands for more powerful broadband communication systems. For example, a low-power 6th-order 240 MHz-to-500 MHz active-RC low-pass filter (LPF) is reported in [57], and a 10-bit 800 MHz CMOS ADC has been reported in [65]. This section will propose a high frequency compact sine-wave synthesizer solution which covers the sub-1 GHz frequency range and can work as a building block of the test architecture for emerging broadband circuits.

To generate a high-linearity sinusoidal waveform, the synthesizer should suppress any higher order harmonics and ideally leave only the fundamental tone. A straightforward solution is to implement a high-order LPF or a high-selectivity bandpass filter (BPF) as illustrated in Fig. 3.1a. However, on the one hand, a high order filter significantly increases the whole design complexity. For instance, [62] reports

*Part of this chapter is reprinted from “150-850 MHz high-linearity sine-wave synthesizer architecture based on FIR filter approach and SFDR optimization” by C. Shi and E. Sanchez-Sinencio, IEEE Transactions on Circuits and Systems I: Regular Papers, vol. 62, pp. 2227–2237, copyright 2015 by IEEE.

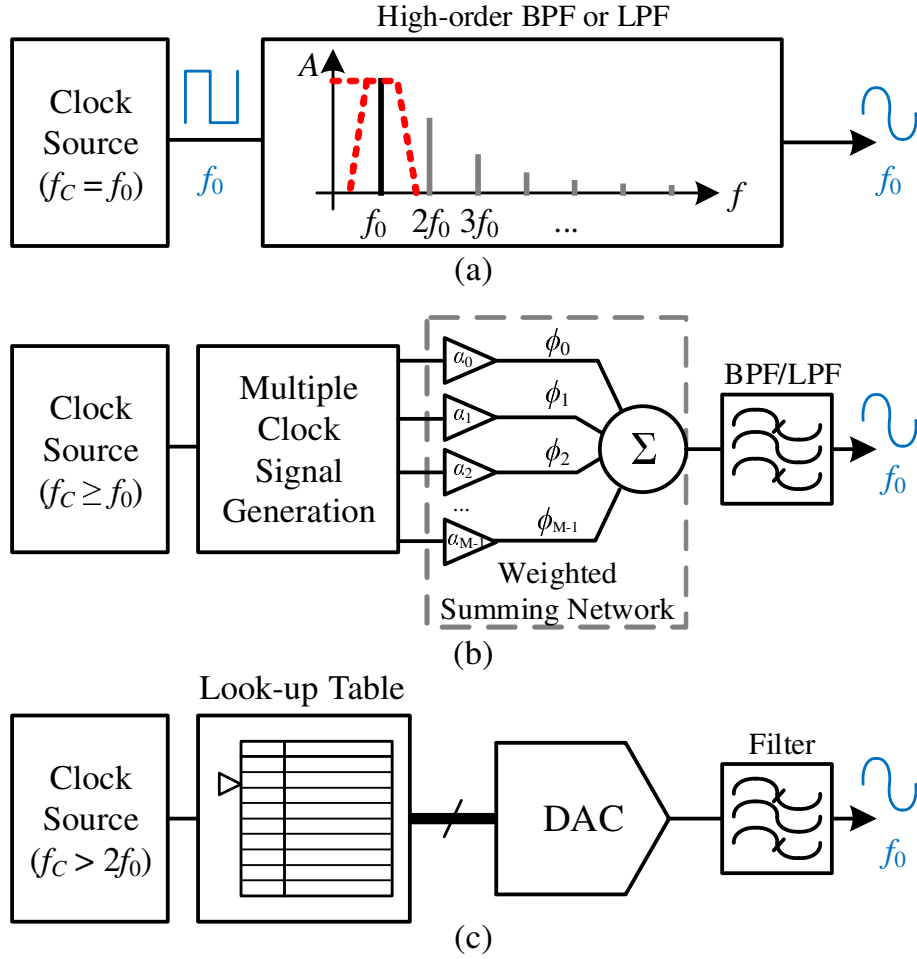


Figure 3.1: Architecture of the sine-wave synthesizers: (a) High-order BPF or LPF, (b) basic harmonic cancellation synthesizer, and (c) direct digital synthesizer.

a sine-wave synthesizer implementing five operational amplifiers (OpAmps) to construct a fifth order switched capacitor (SC) ladder filter. On the other hand, a low order filter has limited attenuation on the close-in harmonics, such as second and the third order harmonics. Therefore, harmonic cancellation (HC) techniques are proposed to economically enhance linearity [43, 66, 67, 68, 69, 70, 71]. These techniques are able to cancel some specific close-in harmonics by manipulating the phases and amplitudes of multiple arbitrary periodic waveforms and adding them together. In

practice, square-wave clock signals are chosen instead of arbitrary periodic waves, because they're fully compatible with digital circuits. As illustrated in Fig. 3.1b, the HC sine-wave synthesizer mainly consists of a multiple clock signal generation block, a summing network, and a low-order output filter.

The clock signal generation block produces multiple phases of the square waveforms $\phi_{0..M-1}$. There are basically two different schemes for this block. [66, 67, 68, 70, 71] use conditional clocks, which means each clock ϕ_i has a different phase and duty-cycle. This conditional-clock-based scheme requires an auxiliary control logic, like the delay lines in [67] or the counter/divider-based logic in [66], [68], [70] and [71]. For the counter/divider-based approaches, it is inevitable to use a source clock with much higher frequency than the desired sinusoidal frequency ($f_C > f_0$). Particularly in [66], a division ratio of 116 was adopted to generate a 10 MHz output from a 1.16 GHz input. Such high input frequency prevents the output frequency from being further raised. To the contrary, [43] and [69] implement another scheme—the multi-phase clocks. Multi-phase clocks have the identical magnitude and the same frequency ($f_C = f_0$) but different phases. Compared to the previous scheme, the control logic is removed, and clock signals can be generated by a N -stage ring oscillator (RO) or a delay locked loop (DLL). The M clocks $\phi_{0..M-1}$ are picked from the N clocks ($M \leq N$) generated in the RO or the DLL, which have equally-distributed initial phases. This scheme is more suitable for high frequency application. The synthesizer in [43] has pushed the output frequency to 220 MHz. In addition, [69] analyzes HC conditions using 3 to 10 phases. It derived the amplitude for each phase to achieve the best cancellation.

Following the clock signal generation block, the weighted summing network sizes amplitudes of all generated clocks ($\alpha_{0..M-1}$ in Fig. 3.1b) and adds them together. The sinusoidal oscillator in [67] adopts a digital-to-analog converter (DAC) as a

multilevel hard limiter. It shapes the output waveform with multiple voltage levels and outputs the summation result immediately. Different from [67], most other approaches first convert voltage signals into the current domain or the charge domain and then conduct the summation. [68] and [71] adopt a charge adder using SC circuits with different capacitance. Inverting amplifiers are implemented in [69] and [70], sizing the amplitudes with different resistors and summing the currents. However, the operation speed of these amplifier-based approaches will be limited by the gain and bandwidth of the amplifier design. Furthermore, the synthesizer in [43] uses current-steering architecture, which implements seven groups of current mirrors whose transistor sizes define the current amplitudes. [66] proposes a digitally compatible solution—summing network with only inverters and resistors. These two architectures have lower complexity and thus are more efficient for higher frequencies.

After the waveform summation, specific close-in harmonics are eliminated, but the high order frequency components still exist and degrade the total distortion. Only 45 dBc spur-free dynamic range (SFDR) is measured in [43] for a sine-wave synthesizer without an output filter. Thus, it is necessary to use an output filter to further smooth the waveform. [67] adopts a continuous-time BPF (CT-BPF), [68] and [71] use SC-BPFs, and [70] implements a CT-LPF. These active filters are good for preserving the wave swing and driving the following circuit stages. Nevertheless, they also introduce additional distortion, power consumption and parasitic components. The highest achievable output frequency is limited by the bandwidth of the filter’s amplifiers. Furthermore, a passive filter can also be used as the output filter. [66] has shown how a synthesizer with a third order passive RC LPF is able to achieve the total harmonic distortion (THD) of -72 dBc for 10 MHz sinusoidal output. In addition, the LC-VCO, which contains a passive LC BPF, is also considered to be a good sine-wave generator, although the on-chip inductor occupies a tremendous area

for low RF frequency below 1 GHz [40].

Direct digital frequency synthesizer (DDFS) is another powerful tool for sine-wave synthesis as shown in Fig. 3.1c. It integrates a memory-based look-up table and a DAC to store and build the sinusoidal waveform [72, 73, 74]. High linearity of the output waveform can be achieved with high resolution DAC and precise sine-wave encoding. However, huge hardware overhead and power consumption are the major drawback, making it unsuitable for practical integrated applications.

This section proposes a high frequency fully digital sinusoidal wave synthesizer which is competitive for the demanding scaling down of IC technology. Aiming at 150 to 850 MHz frequency range, the proposed compact architecture implements a 5-phase 3-amplitude harmonic cancellation technique by implementing a multi-phase clock generation block, a weighted resistor summing network, and a passive low-pass output filter. The proposed circuit architecture can operate at low supply voltage and is robust against the process-voltage-temperature (PVT) variations with the help of an iterative SFDR optimization algorithm. This section is organized as follows. The harmonic cancellation technique is discussed in Section 3.2, and the circuit architecture is introduced in Section 3.3. Section 3.4 proposes the application of the optimization algorithm. Section 3.5 shows the measurement results followed by conclusions in Section 3.6.

3.2 Harmonic Cancellation Technique

3.2.1 Waveform, Fourier Series and Harmonics

3.2.1.1 Fourier series

As well known [75], an arbitrary periodic waveform $\phi(t)$ can be presented as the Fourier series,

$$\begin{aligned}\phi(t) &= \frac{A_0}{2} + \sum_{k=1}^{+\infty} [A_k \cos(k\omega_0 t) + B_k \sin(k\omega_0 t)] \\ &= \frac{A_0}{2} + \sum_{k=1}^{+\infty} C_k \sin(k\omega_0 t + \varphi_k)\end{aligned}\quad (3.1)$$

where A_0 is the DC component, ω_0 is the fundamental angular frequency, and A_k and B_k are the Fourier coefficients. In addition, we have $C_k = \sqrt{A_k^2 + B_k^2}$ and $\varphi_k = \arctan(A_k/B_k)$. It means that any arbitrary periodic waveform is a combination of multiple sine waves. These sine wave components are called harmonics. The k -th order harmonic has an amplitude of C_k , a phase shift of φ_k and is located at a frequency of $k\omega_0$. Usually, the first harmonic ($k = 1$) is named as the fundamental.

For instance, a sawtooth waveform can be synthesized in such a way

$$\phi_{st}(t) = \frac{2}{\pi} \sum_{k=1}^{\infty} \frac{1}{k} \sin \left[k\omega_0 t + (-1)^{k+1} \frac{\pi}{2} - \frac{\pi}{2} \right] \quad (3.2)$$

where the number of terms goes infinite, $C_k = \frac{1}{k}$, and $\varphi_k = (-1)^{k+1} \frac{\pi}{2} - \frac{\pi}{2}$. For simplicity, $-\sin(k\omega_0 t)$ is used in the equations below to represent a 180° phase shift, and we use $\omega_0 = 1$. An ideal sawtooth waveform is approximated gradually by increasing the number of harmonics k . Fig. 3.2 demonstrates the synthesized waveforms using only one, three or five harmonics. Fig. 3.3 shows the corresponding spectrums in the frequency domain, including that of the ideal sawtooth wave. It

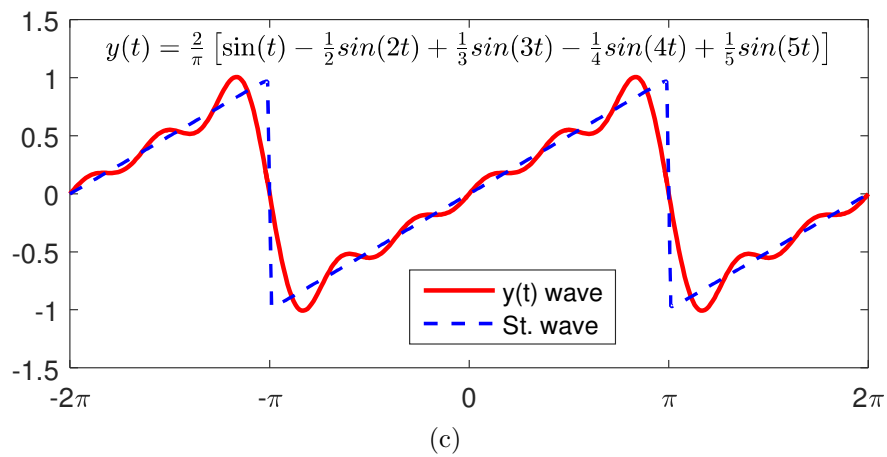
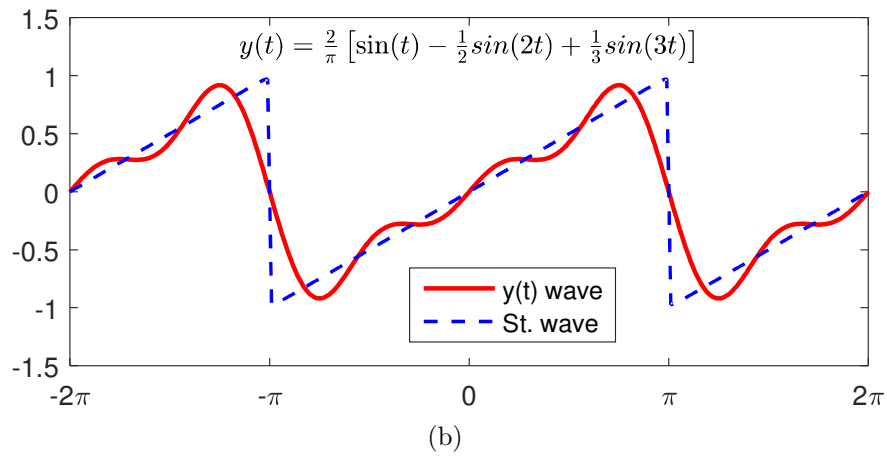
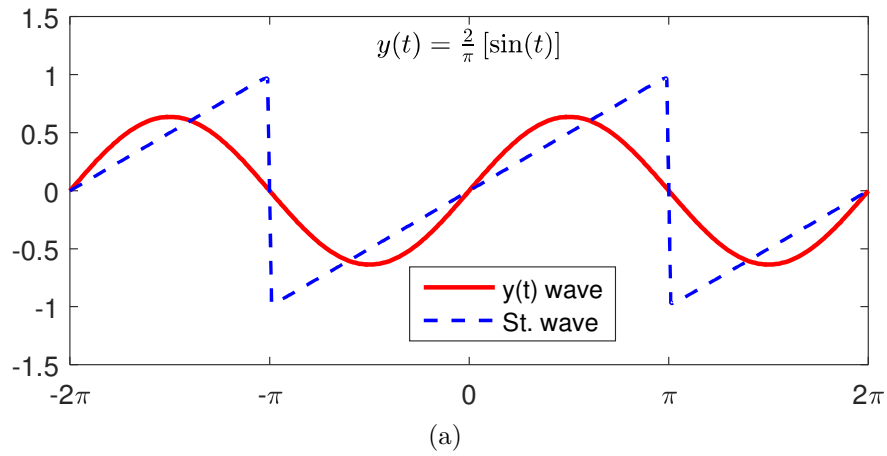
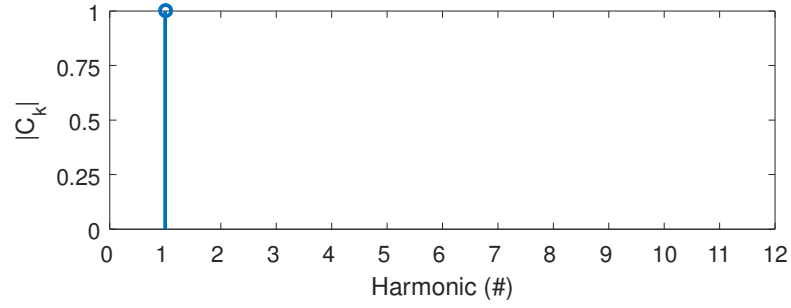
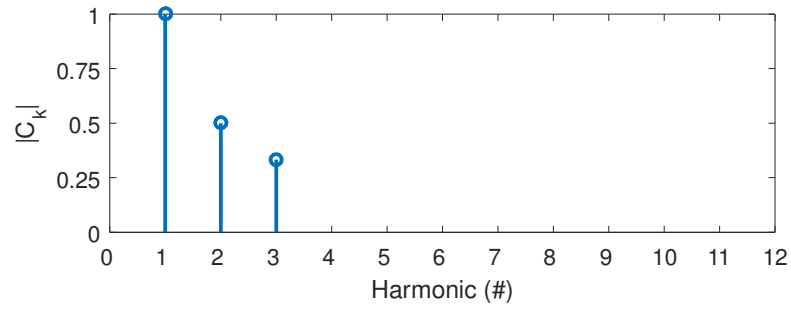


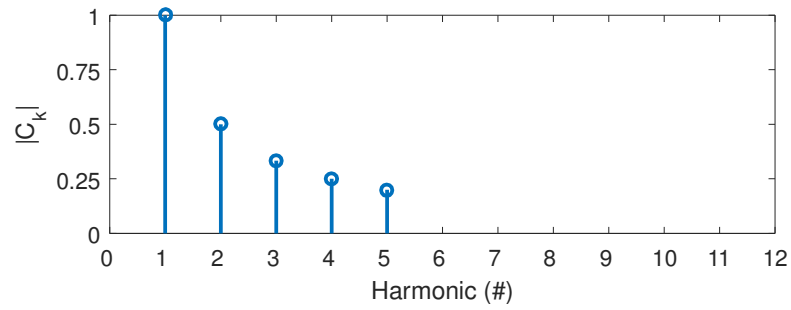
Figure 3.2: Sawtooth waveform with infinite and limited number of terms for (a) $k = 1$, (b) $k = 1 \dots 3$, and (c) $k = 1 \dots 5$.



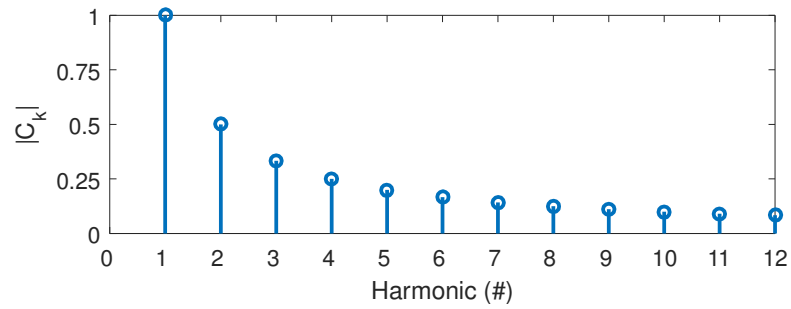
(a)



(b)



(c)



(d)

Figure 3.3: Frequency spectrums for (a) $y(t)$ of Fig. 3.3a, (b) $y(t)$ of Fig. 3.3b, (c) $y(t)$ of Fig. 3.3c, and (d) an ideal sawtooth waveform.

should be noted that, the spectrum contains only the absolute value of C_k , and thus it cannot reveal the phase shift of each harmonic. In fact, both of the amplitudes and phase shifts of the harmonics can be manipulated to construct a different waveform.

3.2.1.2 Sizing and Phase Shift of Harmonics

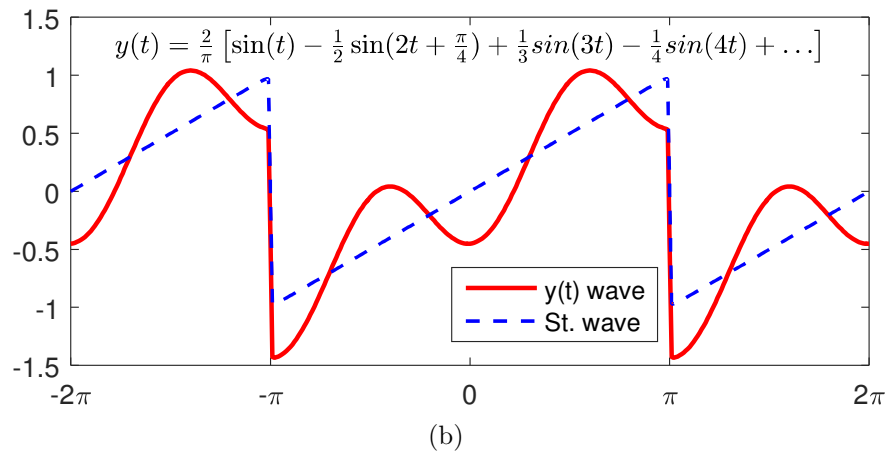
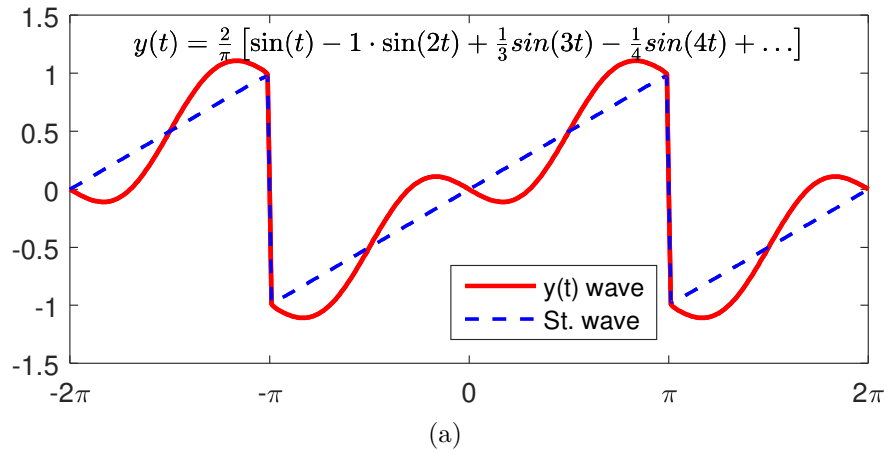


Figure 3.4: Manipulating waveform shapes by (a) changing a harmonic amplitude or (b) changing a harmonic phase shift.

By manipulating each harmonic, we can change the shape of the waveform. For

example, if one harmonic of the sawtooth (3.2) is modified, the shape will become different. Fig. 3.4a shows the change of one harmonic amplitude, while Fig. 3.4b shows the change of one harmonic phase shift.

Moreover, a phase shift θ can also be applied to the arbitrary waveform,

$$\phi\left(t + \frac{\theta}{\omega_0}\right) = \frac{A_0}{2} + \sum_{k=1}^{+\infty} C_k \sin(k\omega_0 t + \varphi_k + k\theta) \quad (3.3)$$

Because $\phi(t)$ is a periodic signal, a phase shift is equivalent to a time delay of $\frac{\theta}{\omega_0}$. Using the sawtooth waveform as an example, after the phase shift, it becomes

$$\phi_{st}\left(t + \frac{\theta}{\omega_0}\right) = \sum_{k=1}^{\infty} \frac{1}{k} \sin\left[k\omega_0 t + (-1)^{k+1} \frac{\pi}{2} - \frac{\pi}{2} + k\theta\right] \quad (3.4)$$

We can find that the phase shift θ of the waveform results in $k\theta$ phase shift on its k -th order harmonic. Note that because $\phi(t)$ is a periodic signal, we have $0 \leq \theta < 2\pi$. Waveform phase shift is so important that we can use it to strengthen or cancel some harmonics when we combine multiple shifted arbitrary waveforms.

Here gives some examples of the harmonic manipulation. Considering the shifted sawtooth waveform in (3.4), if we apply $\theta = \pi$, new waveforms can be generated from addition and subtraction. Results are shown in Fig. 3.5. Particularly, the subtraction gives

$$\frac{1}{2} \left(\phi_{st}(t) - \phi_{st}\left(t + \frac{\pi}{\omega_0}\right) \right) = \sum_{k=1}^{\infty} \frac{1}{2k-1} \sin[(2k-1)\omega_0 t] = \phi_{sq}(t) \quad (3.5)$$

This is a square waveform and all of its even order harmonics are eliminated. Its spectrum is illustrated in Fig. 3.6.

Additionally, it should be emphasized that, the frequency spectrum diagram does not contain the phase information. For instance, $\phi_{st}(t)$ and $\phi_{st}\left(t + \frac{\pi}{\omega_0}\right)$ share the same spectrum diagram, which is shown in Fig. 3.3, although their time-domain

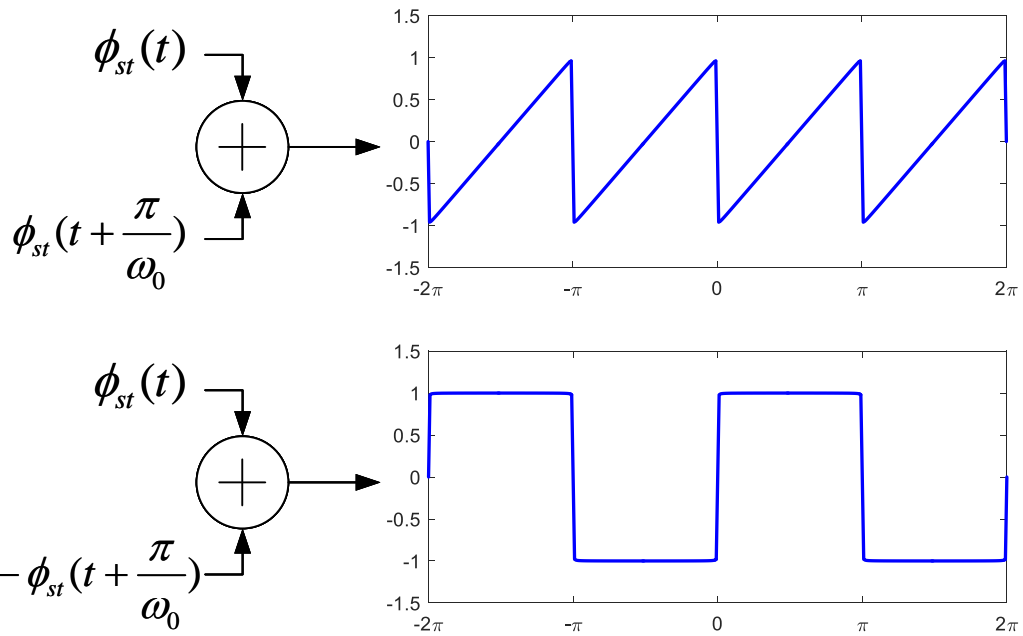
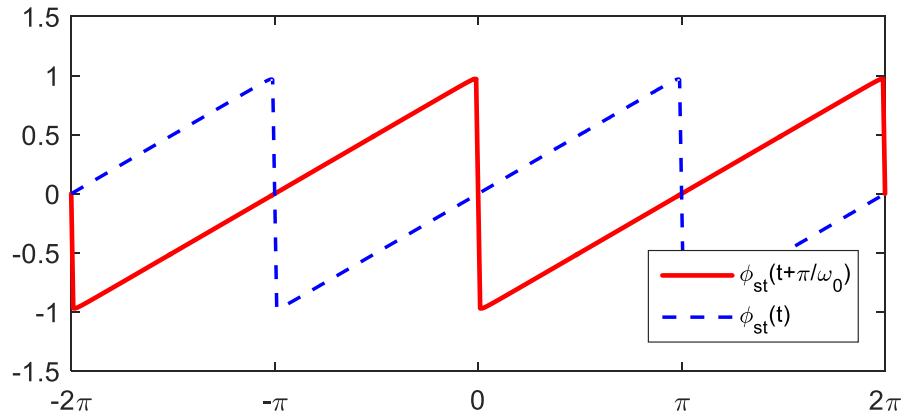


Figure 3.5: Manipulation of the sawtooth waveform.

waveform are different as illustrated in Fig. 3.5.

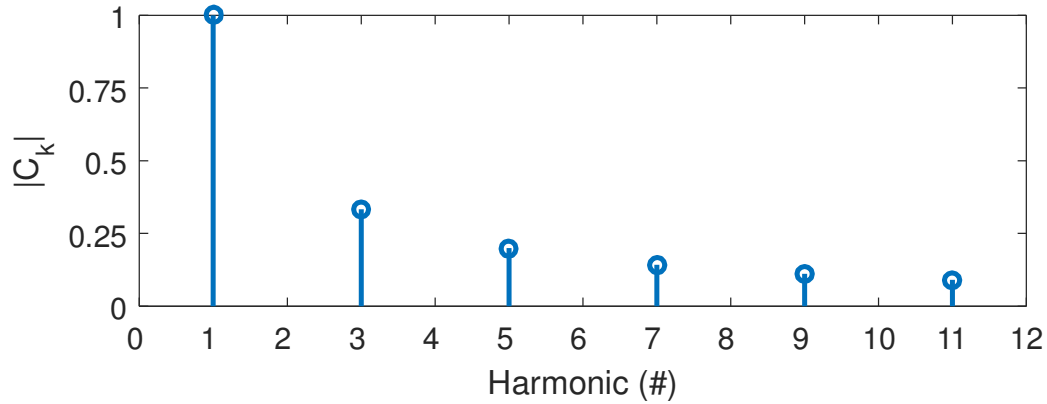


Figure 3.6: Spectrum of the square waveform.

3.2.2 Principles of the Harmonic Cancellation

3.2.2.1 Definition

If we ignore A_0 and combine multiple $\phi(t)$ waveforms from (3.1),

$$\begin{aligned}
 F(t) &= \sum_{i=0}^{M-1} \alpha_i \phi\left(t + \frac{\theta_i}{\omega_0}\right) \\
 &= \sum_{k=1}^{+\infty} X_k \cos(k\omega_0 t) + Y_k \sin(k\omega_0 t) \\
 &= \sum_{k=1}^{+\infty} D_k \sin(k\omega_0 t + \gamma_k)
 \end{aligned} \tag{3.6}$$

$$X_k = \sum_{i=0}^{M-1} \alpha_i [A_k \cos(k\theta_i) + B_k \sin(k\theta_i)] \tag{3.7}$$

$$Y_k = \sum_{i=0}^{M-1} \alpha_i [B_k \cos(k\theta_i) - A_k \sin(k\theta_i)] \tag{3.8}$$

$$D_k = \sqrt{X_k^2 + Y_k^2}, \quad \gamma_k = \arctan\left(\frac{X_k}{Y_k}\right) \tag{3.9}$$

where α_i are the magnitude sizing factors (note that X_k , Y_k , D_k and γ_k are parameters related to the combined waveform $F(t)$, while A_k , B_k , C_k and φ_k belong to the single

waveform $\phi(t)$ of (3.1)), and θ_i are the initial phase of the i -th single waveform. The target of the harmonic cancellation is to make $X_k = 0$ and $Y_k = 0$ ($D_k = 0$) for some specific k or all $k \geq 2$. Furthermore, we can find in (3.6) that there are two degrees of freedom for waveform manipulation— α_i and θ_i .

3.2.2.2 Design Methodology

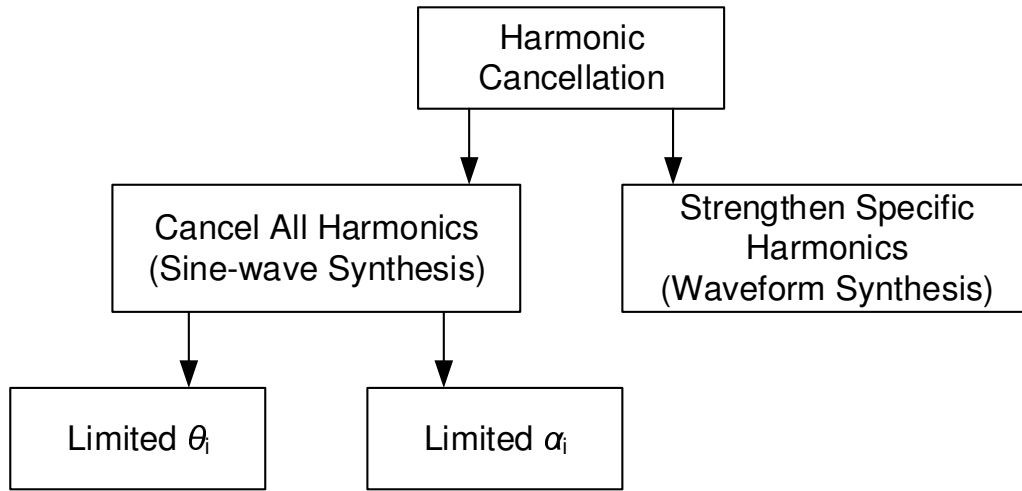


Figure 3.7: Categories of the harmonic cancellation implementations.

For different purposes, the harmonic cancellation implementations can be categorized as illustrated in Fig. 3.7.

There exists two different paths for designing the sine-wave synthesizer, emphasizing either α_i (limited θ_i) or θ_i (limited α_i). Implementing the harmonic cancellation with limited θ_i is the main-stream method for synthesizing a sinusoidal waveform. S. W. Park’s work [68] is an example of this method. As shown in Fig. 3.8, three delayed square waveforms are amplified and summed together to generate the output, a quasi-sinusoidal waveform. Limited θ_i means that a fixed value T_0/N is chosen as

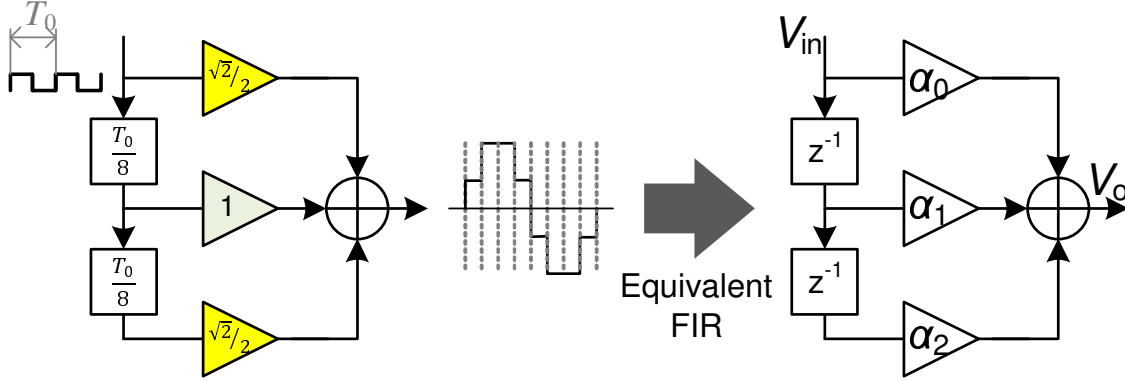


Figure 3.8: Sine-wave synthesis architecture proposed in S.W. Park’s paper.

the delay time (phase shift) for all delay units. Designers further rely on different α_i factors to achieve the desired harmonic cancellation. This architecture is equivalent to a finite-impulse-response (FIR) filter. Later in Section 3.2.3, the intuitive FIR approach and general design equations will be introduced. For limited θ_i , the design complexity mainly comes from the different α_i factors. Particularly, α_i may be an irrational number that is difficult to implement in hardware. Section 3.2.3 will introduce the effort that can approximate the irrational number, and Section 4.2.3 will discuss a two-stage FIR architecture aiming at further reducing the complexity.

Compared to the limited θ_i method, limited α_i design is more complicated because each square-wave clock needs a fine tuning on its delay. M. M. Elsayed introduced this type of harmonic cancellation in [66], and its main idea is shown in Fig. 3.9. All α_i are fixed to 1, and thus different delays are adopted. [66] further proposed a search algorithm to find the proper delays to achieve the best cancellation. There’s not a general equation for limited α_i method. However, it is able to relax the constraint on the hardware implementing α_i .

One more step of the harmonic cancellation technique is to cancel or enforce arbitrary harmonics. [76] demonstrated a digital harmonic synthesis block (DHSB),

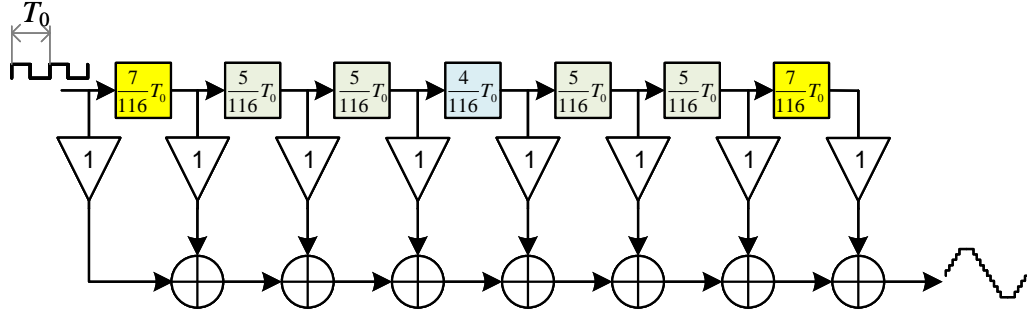


Figure 3.9: Sine-wave synthesis architecture with limited α_i proposed in M.M. Elsayed's paper.

which extends the aforementioned limited α_i concept to enforce a higher order harmonic but cancel the fundamental tone. Nevertheless, the unwanted harmonics near the desired one remain relatively high. [76] further adopts a LC BPF to purify the output spectrum, leading to large hardware overhead. Therefore, enforcing desired harmonic but eliminating unwanted harmonics with only operations of θ_i and α_i is still an open problem which worths further research effort.

3.2.3 Odd Order Cancellation FIR Filter Approach

As discussed above, $\phi(t)$ can be a square wave clock $\phi_{sq}(t)$. For this special case, we can handle the even and the odd order harmonics separately. For the even order harmonic cancellation, consider a square waveform with 50% duty cycle where $A_k = B_k = 0$ for $k = 2, 4, 6, \dots$. In other words, it has only the odd order harmonic components. However, impacted by the PVT variations, it is difficult to produce an exact 50% duty cycle in a real circuit. A differential signal path, a phase-to-duty-cycle converter and an optimization algorithm were adopted to further improve the symmetry. All these techniques will be introduced later.

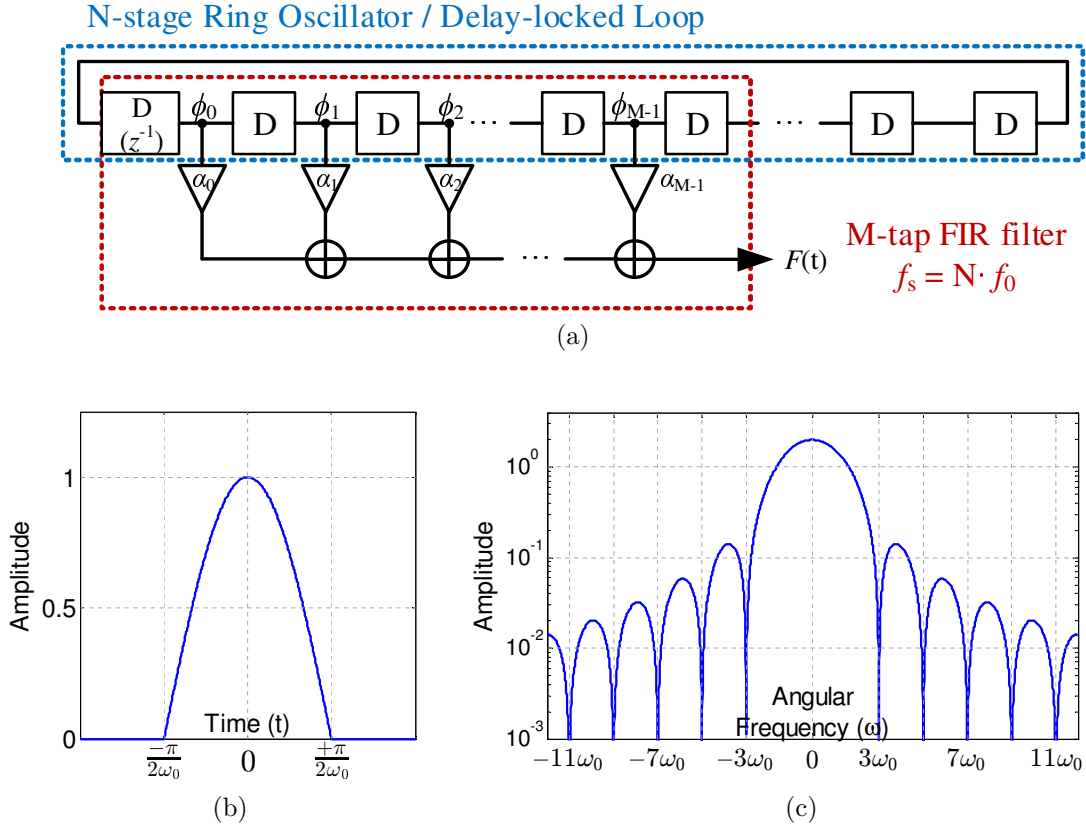


Figure 3.10: Odd order harmonic filter design: (a) Equivalent architecture of the FIR filter, (b) time-domain half-cosine pulse $y(t)$, and (c) Fourier transform $Y(\omega)$ of the half-cosine pulse.

For the odd order harmonic cancellation, a finite impulse response (FIR) approach was implemented. This research will show how the multi-phase clock signal generation, which was discussed in the introduction (Section 3.1), can lead to a better understanding of harmonic cancellation and duplicate the function of an FIR filter. The FIR filter is a filter whose impulse response has finite duration. In an FIR filter, the input signal is delayed for limited times. The output of each delay is called a “tap”. The output of the filter is a sum of all taps multiplied by their tap coefficients. For instance, as shown in Fig. 3.10a, “D” is a delay cell in the RO, it can also be treated as a z^{-1} operator in the Z-domain. The weighted summing net-

work is the same as the sum of weighted taps. In the N -stage RO whose oscillation frequency is f_0 , each delay cell has a delay of $(N \cdot f_0)^{-1}$. If we treat this filter as an equivalent discrete-time (DT) filter, we can use the traditional DT-FIR filter design method. In such an equivalent DT-FIR filter, all z^{-1} cells (D-flipflops) are driven by the same sampling clock. To achieve the identical $(N \cdot f_0)^{-1}$ delay in each cell, the sampling frequency f_s should be $N \cdot f_0$. Moreover, in Fig. 3.10a, assume $\phi[n]$ is the input and $F[n]$ is the output, where $[n]$ is used to indicate the most recent sample, $[n - 1]$ presents the previous one, and so on. We have $\phi_0 = \phi[n]$, $\phi_1 = \phi_1[n - 1]$, ... , $\phi_{M-1} = \phi_0[n - (M - 1)]$. The output sequence of this equivalent discrete-time FIR filter is defined as

$$F[n] \triangleq \sum_{i=0}^{M-1} \alpha_i \phi[n - i] \quad (3.10)$$

This is a $(M - 1)$ -th order DT-FIR filter.

Let's consider a time-domain half-cosine pulse (Fig. 3.10b)

$$y(t) = \begin{cases} \cos(\omega_0 t) & -\frac{\pi}{2\omega_0} \leq t \leq \frac{\pi}{2\omega_0} \\ 0 & \textit{otherwise} \end{cases} \quad (3.11)$$

where $\omega_0 = 2\pi f_0$. Its Fourier transform is

$$Y(\omega) = \mathcal{F}[y(t)] = \frac{2}{\omega_0} \frac{\cos\left(\frac{\pi}{2} \frac{\omega}{\omega_0}\right)}{1 - \left(\frac{\omega}{\omega_0}\right)^2} \quad (3.12)$$

Eq. 3.12 represents a filter suppressing the odd order harmonics. Its frequency-domain spectrum is shown in Fig. 3.10c, which has $Y(\omega) = 0$ for all $\omega = \pm 3\omega_0, \pm 5\omega_0, \dots$. One of the FIR filter design method, the impulse response truncation [77], was chosen for this design. Traditionally, this method is not accurate as it simply truncates the infinite impulse response of a desired transfer function. However, this is not the

case for this design because the target response itself, a half-cosine pulse (Eq. 3.11), is finite. Hence, the FIR filter does not suffer from inaccuracy in the truncation. To design the FIR filter, sampling at the time-domain half-cosine pulse gives the FIR tap coefficients.

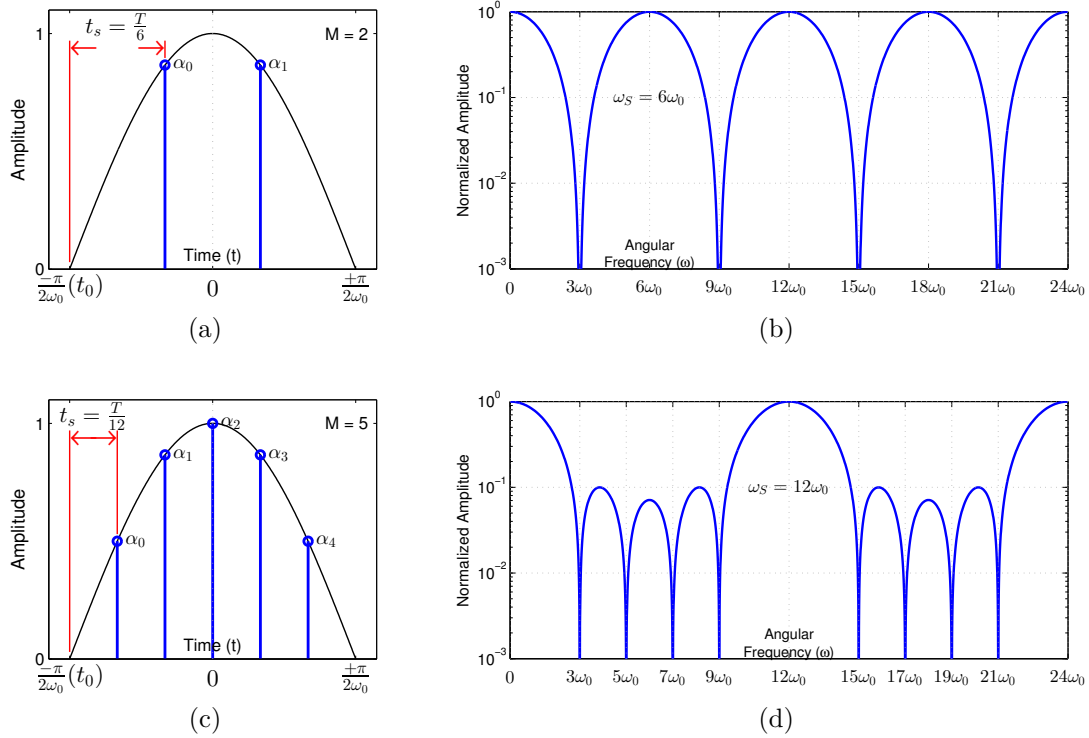


Figure 3.11: Odd order cancellation FIR filter design: (a) 2-tap coefficients, (b) $Y_{alias(ed)}(\omega)$ for the 2-tap FIR filter, (c) 5-tap coefficients, and (d) $Y_{alias(ed)}(\omega)$ for the 5-tap FIR filter.

As demonstrated in Fig. 3.11a and Fig. 3.11c, the sampling frequency is f_s and we start sampling at $t_0 = -\pi/2\omega_0$. The advantage of this start time point is a symmetric coefficients distribution, which reduces the number of different coefficients. The sampling interval is $t_s = 2\pi/N\omega_0 = T/N$, where $T = f_0^{-1}$ is the period of the output

sine wave, and N is the number of delay stages in Fig. 3.10a. N is given by

$$N = 2(M + 1) \quad (3.13)$$

where M is the number of FIR filter taps. Such a discrete sampling in the time domain leads to the aliasing in the frequency domain. The aliased frequency response of the FIR filter can be derived from Eq. 3.12,

$$Y_{aliased}(\omega) = \sum_{k=-\infty}^{+\infty} Y(\omega - kN\omega_0) \quad (3.14)$$

Fig. 3.11b and Fig. 3.11d demonstrate $Y_{aliased}(\omega)$ for a 2-tap and a 5-tap odd order cancellation filter. The aliasing leads to some noncancellable odd order harmonics, such as the fifth for $M = 2$ and the 11th for $M = 5$, assuming the even order harmonics have been canceled by the 50% duty cycle. To mitigate the aliasing problem, we can implement more taps to push the first noncancellable order farther and adopt an output filter to attenuate it.

Table 3.1 summarizes the tap coefficients for the tap numbers from 2 to 7. In this table, the maximum value is normalized to 1, and only 4 digits are retained for the fractional part. More coefficients for Eq. 3.6 can be intuitively obtained from Eq. 3.11 and Fig. 3.10a

$$\begin{aligned} \alpha_i &= y((i + 1)t_s + t_0) \\ &= \cos\left(\frac{(i + 1)\pi}{M + 1} - \frac{\pi}{2}\right) \\ \theta_i &= i \cdot \frac{2\pi}{N} \\ i &= 0, 1, \dots, M - 1 \end{aligned} \quad (3.15)$$

Table 3.1: Normalized coefficients of the odd order cancellation FIR filter

M	N	θ_i	α_0	α_1	α_2	α_3	α_4	α_5	α_6	1st noncancelable harmonic
2	6	$i \cdot \frac{\pi}{3}$	1.0000	1.0000						$5\omega_0$
3	8	$i \cdot \frac{\pi}{4}$	$0.7071(\frac{1}{\sqrt{2}})$	1.0000	$0.7071(\frac{1}{\sqrt{2}})$					$7\omega_0$
4	10	$i \cdot \frac{\pi}{5}$	0.6180	1.0000	1.0000	0.6180				$9\omega_0$
5	12	$i \cdot \frac{\pi}{6}$	$0.5000(\frac{1}{2})$	$0.8660(\frac{\sqrt{3}}{2})$	1.0000	$0.8660(\frac{\sqrt{3}}{2})$	$0.5000(\frac{1}{2})$			$11\omega_0$
6	14	$i \cdot \frac{\pi}{7}$	0.4450	0.8019	1.0000	1.0000	0.8019	0.4450		$13\omega_0$
7	16	$i \cdot \frac{\pi}{8}$	0.3827	$0.7071(\frac{1}{\sqrt{2}})$	0.9239	1.0000	0.9239	$0.7071(\frac{1}{\sqrt{2}})$	0.3827	$15\omega_0$

The number of taps M directly determines the number of stages N in the RO/DLL. Thus, trade-offs should be made between the number of delay stages, which indirectly affects the output sine-wave frequency, and the odd order cancellation ability. In this design, a 5-tap ($M = 5$) odd order cancellation FIR filter is adopted to demonstrate the proposed HC technique. On the one hand, a 5-tap odd order cancellation FIR filter suppresses third, fifth, seventh and ninth order close-in harmonics, while 50% duty cycle kills the even orders. Moreover, given a square wave $\phi(t)$ in Eq. 3.1, we have $A_k = 0$ for all k , $B_k = 0$ for even k and $B_k = 4/k\pi$ for odd k . Apply Eq. 3.6 with the proposed coefficients to the square wave. It can be found that the normalized magnitudes of higher order harmonics above the 10th is below 0.01 (-20 dBc). Simultaneously, with the help of a first order output filter and the first order integration mechanism in the proposed weighted summing network, they further yield at least -40 dB attenuation. Therefore, the total SFDR is expected to be around 60 dBc or higher, which is comparable to the other state-of-the-art sine-wave synthesizers. On the other hand, the 5-tap FIR filter has three different coefficients, $1/2$, $\sqrt{3}/2$ and 1. Among them, only the $\sqrt{3}/2$ is irrational. However, we can find the equivalent fractional number, $13/15 \approx 0.86667$, to be the approximate value. It is only 0.064% larger than the original number. Hence, we can use the tap coefficients, $1/2$, $13/15$ and 1. Note that it is not necessary to use the absolute coefficient values. Instead, maintaining the proportional relationships between coefficients is enough to achieve the proposed harmonic cancellation.

3.3 Circuit Implementation

3.3.1 System Architecture

Fig. 3.12 shows the system architecture of the proposed sine-wave synthesizer. A ring oscillator, a weighted resistor summing network, a programmable first order RC

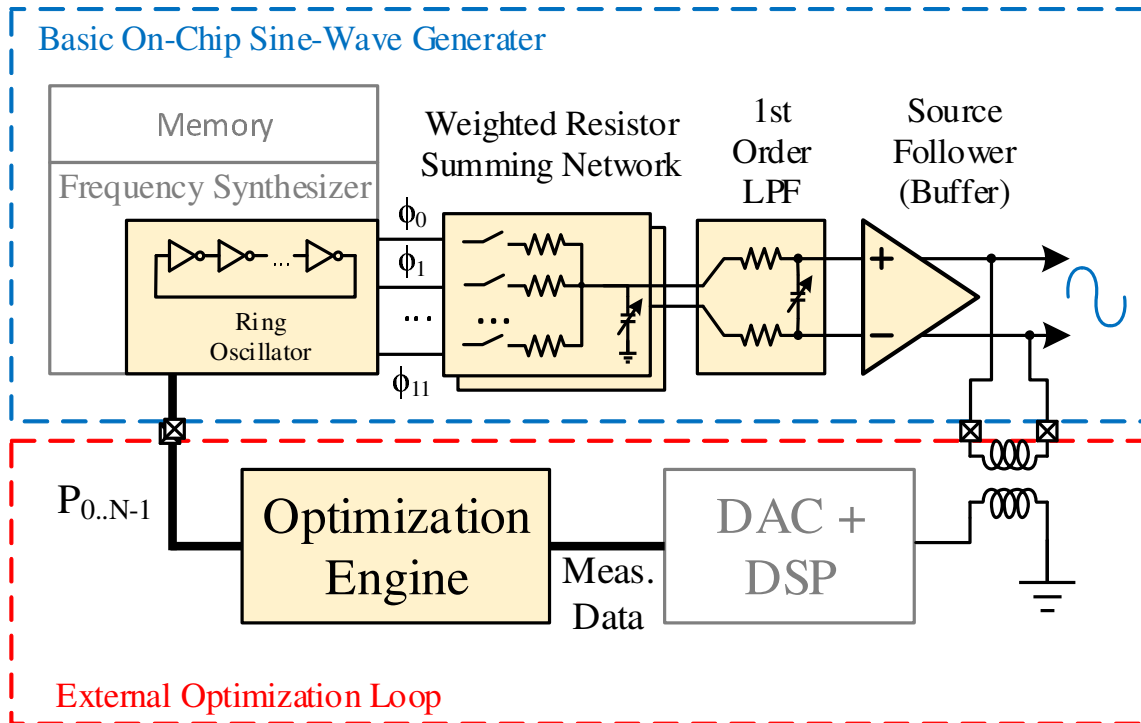


Figure 3.12: System architecture of the proposed synthesizer.

LPF and a buffer are integrated on-chip. A 6-stage differential current-controlled ring oscillator (ICRO) produces square waveforms $\phi_0.. \phi_{11}$. These generated clocks are then fed into the weighted resistor summing network, where nonlinear waveforms are summed together and specific harmonics are canceled. The following LPF further smooths the output waveform, which is buffered and delivered as the sinusoidal signal. Moreover, the on-chip analog path adopts differential blocks so as to suppress the even order harmonics and reduces the sensitivity to the supply noise. The proposed circuit blocks, plus a frequency synthesizer, which stabilizes the output frequency, constructs a basic sine-wave generator, as shown in the blue dashed box in Fig. 3.12. This on-chip generator has a medium harmonic suppression level and is qualified for harmonic-insensitive tasks, such as plotting the filter’s frequency response [51] or testing the supply noise tolerance [43].

Although the HC technique is implemented on-chip, PVT variations may introduce mismatches between clock phases and thus degrade the cancellation effect. To compensate the errors, an external optimization loop is further proposed. As demonstrated in the red dashed box in Fig. 3.12, this auxiliary loop contains an iterative SFDR optimization engine and a spectrum analyzer, which is based on a DAC and a DSP. By using the optimization loop to do a one-time optimization after the chip fabrication, the linearity of the on-chip generator could be further improved. In this case, the one-time optimization procedure should be executed for different output frequencies, and a memory device is necessary to preserve the optimized control words. Therefore, the on-chip sine-wave generator can be used for some applications that have stricter linearity requirements, such as verifying an ADC [63]. In addition, the optimization algorithm and its temperature stability will be analyzed later in Section 3.4.

Furthermore, by permanently enabling the optimization loop, the proposed sine-wave generator can work as a generic signal generator and achieve the best linearity time to time. However, this configuration is not a fully integrated solution. The on-chip spectrum analyzer requires further research effort. In this section, we mainly focused on the implementation of shaded blocks in Fig. 3.12, including the core generator circuits and the optimization engine.

3.3.2 Oscillator and Phase Shifter

The 12 clock signals $\phi_i (i = 0, 1, \dots, 11)$ were generated by the ICRO, which is demonstrated in Fig. 3.13. Ideally, these 12 clocks are identical square waves with 50% duty cycle except that a 30° phase difference exists between the rising edges of every two adjacent clocks ϕ_i and ϕ_{i+1} . Nevertheless, PVT variations will impose different phases and duty cycles on these clocks. Therefore, in the ICRO, each ϕ_i is

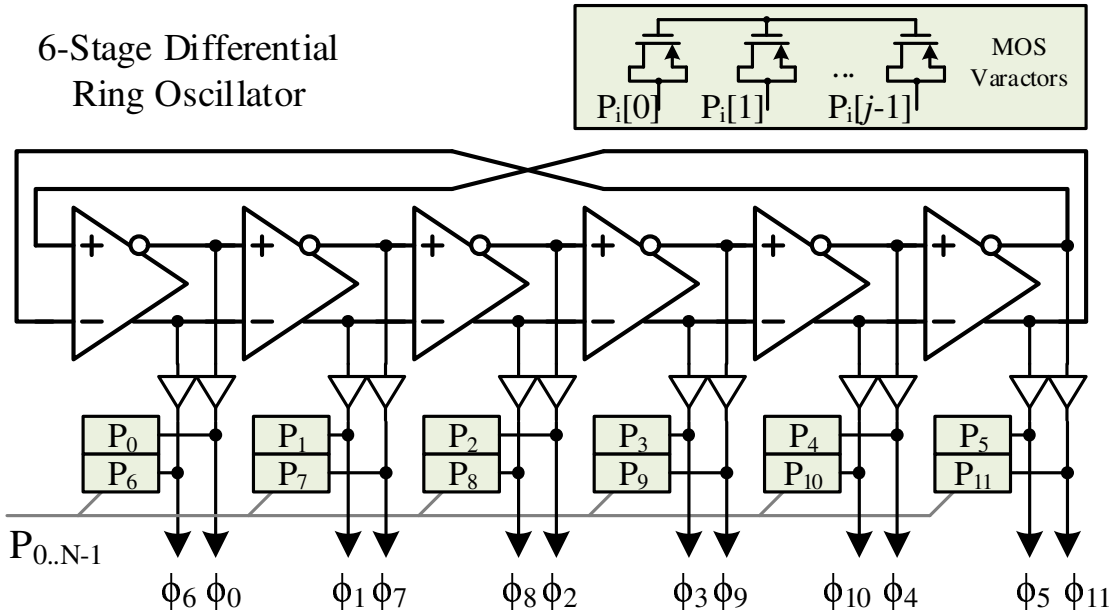


Figure 3.13: Schematic of the ring oscillator and the phase shifter.

buffered and a digitally-controlled phase shifter was implemented as the clock buffer load. The phase shifter is a binary-weighted MOS varactor array, terminated by control signals P_i . By applying different P_i , the external optimization algorithm can tune the rising time of each ϕ_i , and thus change the phase slightly. In this design, each P_i is a 3-bit ($j = 3$) control word, and each phase shifter has a tuning range from 0 to 20 ps. Finally, the 12 clocks were fed into a weighted resistor summing network.

3.3.3 Weighted Resistor Summing Network

To implement the FIR approach ($M = 5$) described in Fig. 3.10a, a weighted resistor summing network was adopted, which is illustrated in Fig. 3.14. The shifted square wave ϕ_i with phase shift θ_i came from the ICRO, and each FIR tap was made up of a switch block pair (SW_i and SW_i^*) and the following resistors. Tap coefficients α_i were defined by the resistance values. The summation operation was

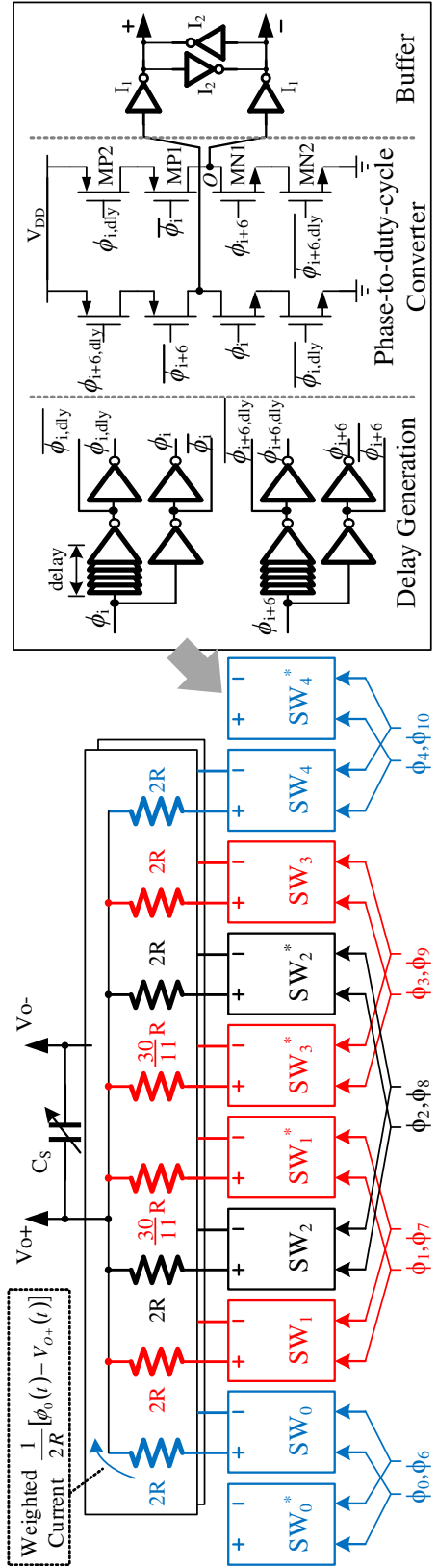


Figure 3.14: Schematic of the weighted resistor summing network.

finally achieved by an integration capacitor C_S . Additionally, all building blocks were differential mode circuits.

In the circuit design, switch blocks were inserted to isolate the resistor summing network from the ICRO's output ϕ_i . As discussed in [66], if simple inverters are used as buffers, they will suffer from the duty-cycle error because of the NMOS & PMOS threshold voltage mismatch. Therefore, slightly different from the conceptual structure in Fig. 3.10a, not only the first five phases $\phi_{0..4}$, but also their complementary phases $\phi_{6..10}$ were used. The inverter was replaced by a phase-to-duty-cycle (PDC) converter. The PDC converter triggered the rising edge and the falling edge of the output separately from clock pair ϕ_i and ϕ_{i+6} . For instance, Fig. 3.14 shows node O is charged to the high voltage (rising edge) when ϕ_i became high and turned on MP1. The charging path will later be cut off by MP2 when delayed $\phi_{i,dly}$ also becomes high. Similarly, ϕ_{i+6} turns on MN1 and discharges the node O (falling edge), and $\phi_{i+6,dly}$ stops the discharging. In this procedure, the proposed PDC converter makes the duty cycle error, which is introduced by the unpredictable and uncontrollable threshold voltage mismatch, compensable by tuning the phase shifts of ϕ_i and ϕ_{i+6} . Particularly, the PDC delay time should be carefully designed because it limits the highest PDC switching frequency, if it was not limited by the ring oscillator. On the one hand, if the delay time is longer than $(2f_0)^{-1}$, both of the MP1-MP2 and the MN1-MN2 paths may be turned on simultaneously, damaging the PDC behavior. On the other hand, the delay time should be long enough to guarantee a full charge/discharge of the node O . In the proposed design, a 500 ps delay is used. Furthermore, a buffer stage was implemented to drive the resistor network. As analyzed in [66], the on-resistance mismatch of the CMOS inverter-transistors will also impact the output waveform's linearity. Particularly, for high frequency synthesizer, this issue is severer, because the selected small resistance R in the resistor network

is comparable to the on-resistance of the CMOS inverter-transistors. To address this issue, a cross-coupled inverter pair I_2 was added to the buffer stage. I_2 played the role of a negative impedance in parallel with the on-resistance of I_1 , and thus reduced the impact of the on-resistance to a certain extent [78].

Each switch block output a square wave with full swing on one terminal of each resistor, converting the voltage signal to a current. Three different tap coefficients, $1/2$, $13/15$ and 1 , are proposed in Section 3.2. In order to achieve the cancellation, the amount of currents flowing through the summing network should keep the same proportional relationship as these coefficients. Thus, the relative resistance values can be obtained from the reciprocal numbers of the coefficients— $2R$, $\frac{15}{13}R$ and R , where R is a unit resistance value. To further reduce the change of resistance values, R can be replaced by a pair of $2R$ resistors in parallel, and $\frac{15}{13}R$ is a $2R$ resistor in parallel with a $\frac{30}{11}R$ resistor. As a result, only two different resistance values, $2R$ and $\frac{30}{11}R$, are needed. Switch block pairs are used to drive the two parallel resistors separately ($SW_{1,2,3}$ and $SW_{1,2,3}^*$) or balance the clock buffer loads ($SW_{0,4}$ and dummy $SW_{0,4}^*$). Moreover, in the layout, the switches and the resistors are arranged as shown in Fig. 3.14 to improve the symmetry among different branches. In this design, we choose $2R = 7.5 \text{ K}\Omega$, a 4-segment poly resistor. For each segment, the length is $4.4 \mu\text{m}$ and the width is 700 nm . Another resistor is $\frac{30}{11}R = 10.2 \text{ K}\Omega$, whose segment length is changed to $6 \mu\text{m}$.

As shown in Fig. 3.14, weighted currents are summed together on the common node of the resistor network and converted back to voltage through the integration capacitor C_S . On the one hand, we can treat the resistor network and C_S as an equivalent LPF. The C_S value should be high enough to make the bandwidth of this LPF much smaller than the output frequency, $(2\pi RC_S)^{-1} \ll f_0$. Only in this way, C_S can actually perform the integration (summation) at f_0 . On the other hand,

the swing of the output sinusoidal waveform is also controlled by C_S . A larger C_S results in a smaller swing. Therefore, a wide-range programmable capacitor, C_S , is implemented, from 35 fF to 2.2 pF for a flexible selection of output frequency and voltage swing.

3.3.4 LPF and Output Buffer

A first order passive-RC LPF was adopted after the weighted resistor summing network. The bandwidth of this LPF is the same as the output sinusoidal frequency f_0 so as to suppress higher order harmonics. Furthermore, two PMOS source followers were adopted as the output buffer to drive the external spectrum analyzer. A simple common drain structure was used to avoid further linearity degradation.

3.3.5 Design Procedure

3.3.5.1 Ring oscillator

The detailed schematic of a delay cell in the ring oscillator (Fig. 3.13) is shown in Fig. 3.15, where VBP and VBN is the bias voltage of the P/NMOS current mirror transistors. The delay cell design follows this procedure,

1. The delay cell is equivalent to a amplifier. According to the Buckhausen criterion, to obtain a stable oscillation, the amplifier's gain should be larger than 0 dB at the oscillation frequency. Therefore, the gain-bandwidth product should satisfy $GBW \geq 1$ GHz, considering 1 GHz is the upper frequency limit with some margin.
2. Set the current budget to 200 μ A at 1 GHz for one delay cell from a 1.8 V supply.
3. For the IBM 180 nm process, the mobility of an NMOSFET is $\mu_N = 490$ $\text{cm}^2/\text{V} \cdot \text{s}$, and the mobility of a PMOSFET is $\mu_P = 98$ $\text{cm}^2/\text{V} \cdot \text{s}$, which is approximately 1/4 of that of the NMOSFET. Moreover, NMOS has $C_{ox,N} =$

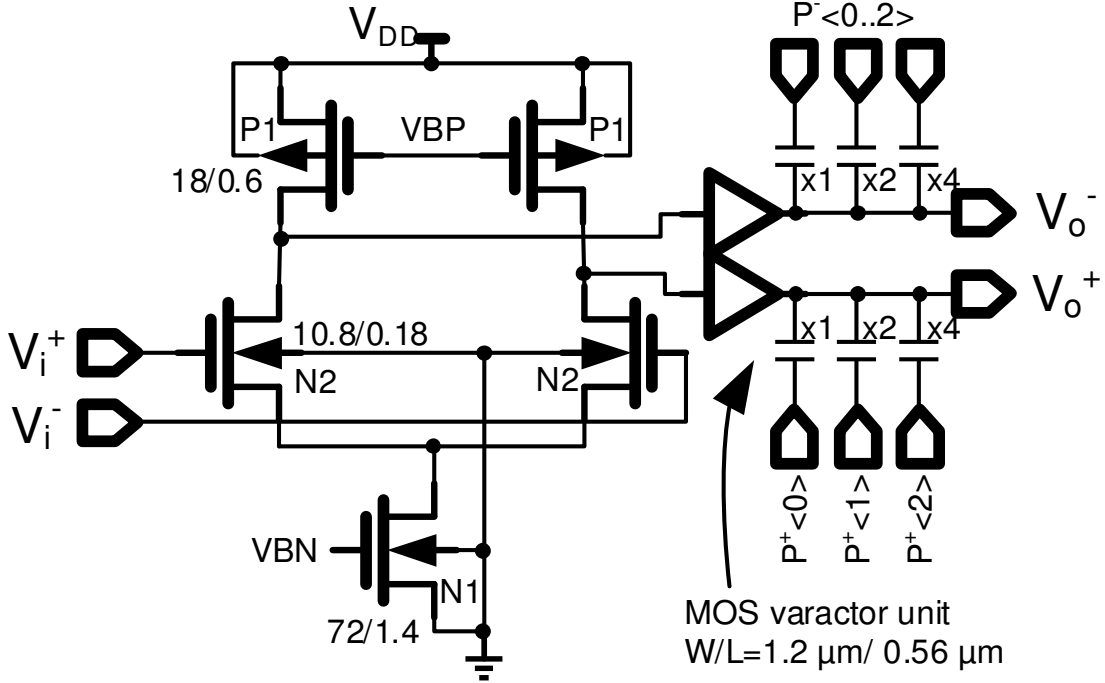


Figure 3.15: Schematic of the ring oscillator delay cell with the phase shifters.

$\epsilon_{SiO_2}/t_{ox,N} = 7.76 \times 10^{-3} \text{ F/m}^2$, and PMOS has $C_{ox,P} = \epsilon_{SiO_2}/t_{ox,P} = 7.51 \times 10^{-3} \text{ F/m}^2$.

4. For P_1 , $I_{P_1} = \frac{1}{2}\mu_P C_{ox,P}(V_{GS,P_1} - V_{TH,P})^2 = 100 \mu\text{A}$. Consider an overdrive voltage $V_{GS,P_1} - V_{TH,P} = 0.2 \text{ V}$. Setting $L = 0.6$ for all current mirror transistors, we can derive that $(W/L)_{P_1} = 68 = 40.8 \mu\text{m}/0.6 \mu\text{m}$.
5. Similarly, we have $I_{N_1} = \frac{1}{2}\mu_N C_{ox,N}(V_{GS,N_1} - V_{TH,N})^2 = 200 \mu\text{A}$, and thus $(W/L)_{N_1} = 26 = 15.6 \mu\text{m}/0.6 \mu\text{m}$.
6. Simulation shows that the parasitic capacitance seen from the drain of P_1 is around 40 fF. Thus, an estimated load capacitance C_L is set to around 120 fF, including the parasitic capacitance from the drain of P_1 , the gate of N_1 in the next delay cell and the gate of the output buffer.
7. The GBW satisfies $\text{GBW} \cdot 2\pi = g_{m,N_2}/C_L = 1 \text{ GHz}$. We have $g_{m,N_2} \approx 7.54 \times$

$$10^{-4}\Omega^{-1}.$$

8. The size of N_2 can be obtained from $g_{m,N2} = \sqrt{2\mu_N C_{ox,N}(W/L)_{N2}I_{P1}}$. As a result, $(W/L) \approx 7.5 \approx 2.7 \mu\text{m}/0.36 \mu\text{m}$.

Using the parameters obtained above as a start point and executing iterative design optimization procedures, the final component parameters are shown in Fig. 3.15. In addition, MOSFET varactors are adopted for the phase shifter, because relatively small capacitance (maximum 5.8 fF for each unit) can be achieved. The corresponding layout design is illustrated in Fig. 3.17. A simulation of the ring oscillator's output frequency versus the bias current is conducted and the result is plotted in Fig. 3.16. It covers the desired frequency range from 150 to 850 MHz. The phase shifter is also simulated. A different control code will apply a different delay on the rising edge of the output square wave, as depicted in Fig. 3.18. The delay step is about 2.25 ps.

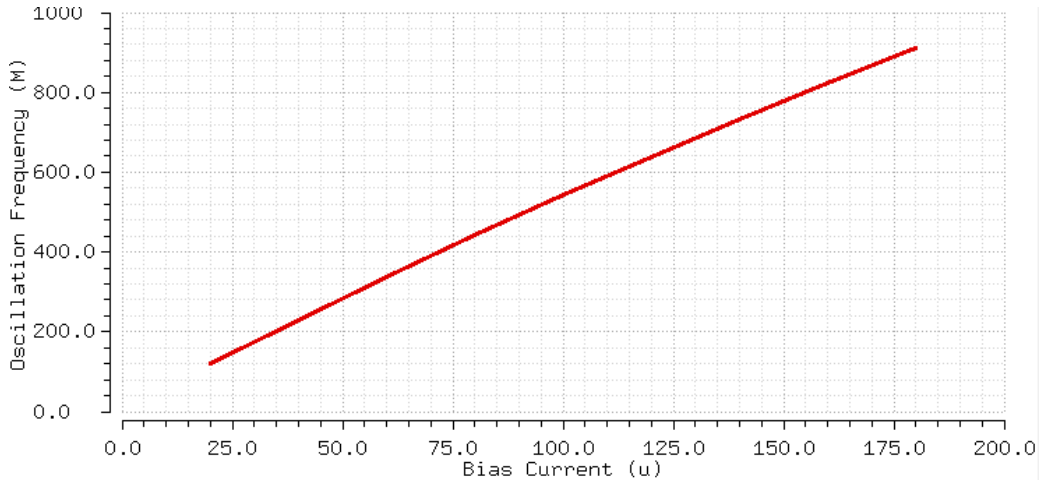


Figure 3.16: Simulated oscillation frequency versus bias current.

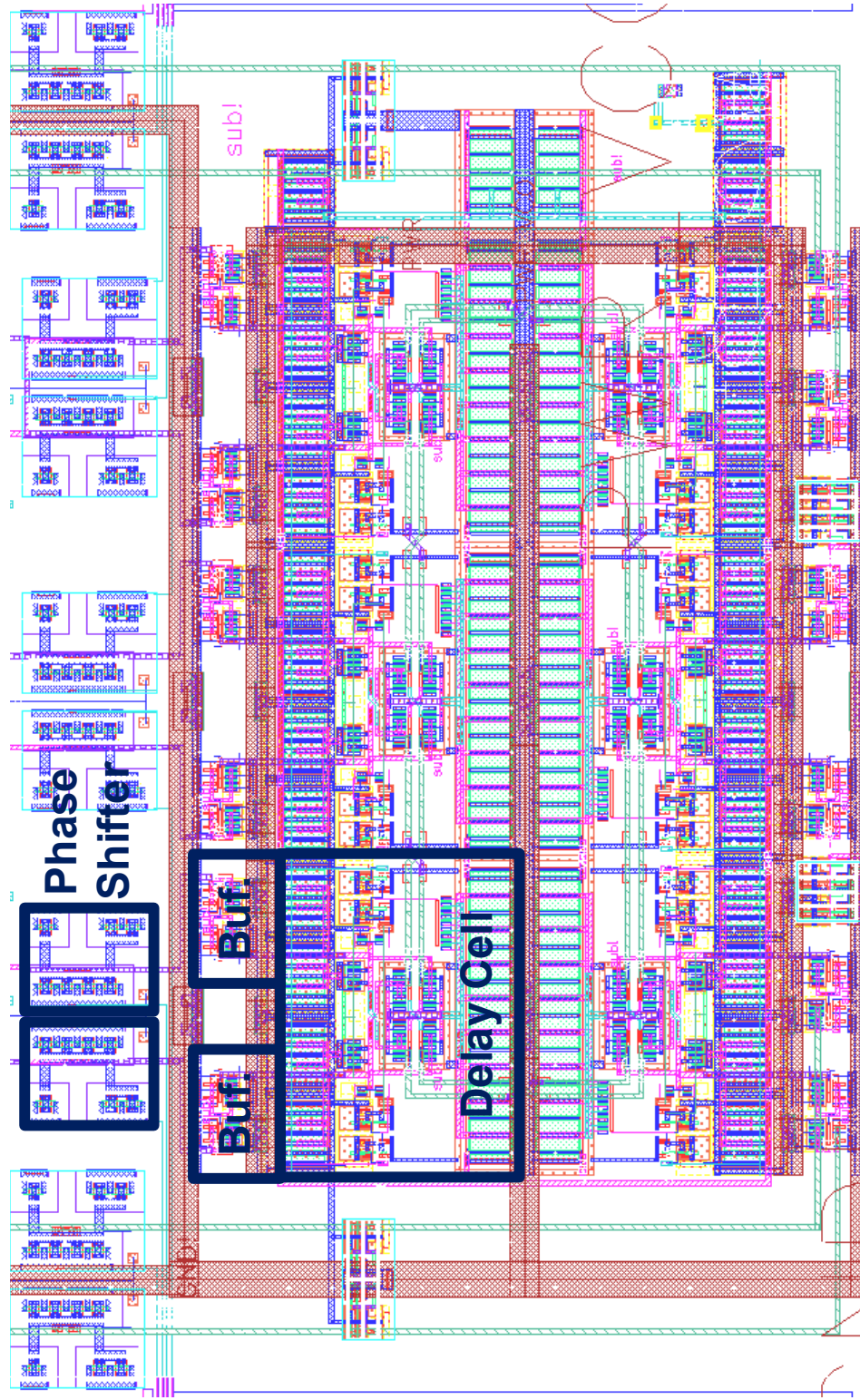


Figure 3.17: Layout of the ring oscillator and the phase shifters.

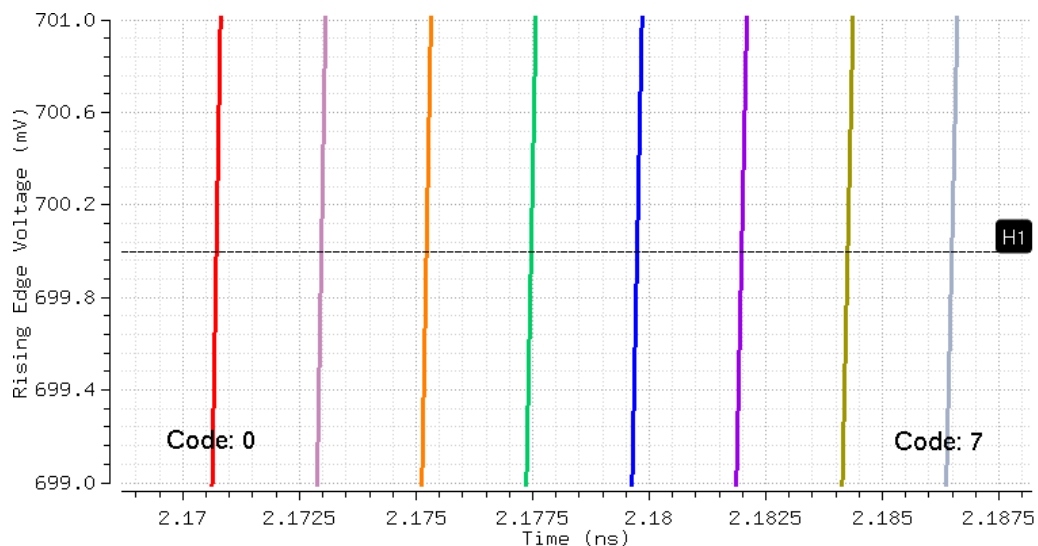


Figure 3.18: Simulated rising edge delay versus phase shifter control code.

3.3.5.2 Resistor summing network

The detailed layout design of the weighted resistor summing network of Fig. 3.14 is illustrated in Fig. 3.19. Partial common centroid layout is used to reduce the mismatches between different resistor segments. It should be mentioned that, two different resistor lengths are used, $4.4 \mu\text{m}$ and $5.9 \mu\text{m}$, so as to achieve the two different weights $2R$ and $\frac{30}{11}R$ in Fig. 3.14. And one leg of resistor consists of total 4 segments as demonstrated in the figure.

3.3.5.3 PDC switch

The design procedure of PDC switches shown in Fig. 3.14 is described below

1. Build the PDC switch array as shown in Fig. 3.14, using the minimum W/L for all transistors.
2. Connect the PDC switch array with the resistor summing networking and the ring oscillator (Fig. 3.13).
3. Simulate the quasi-sinusoidal waveform generated by the proposed synthesizer

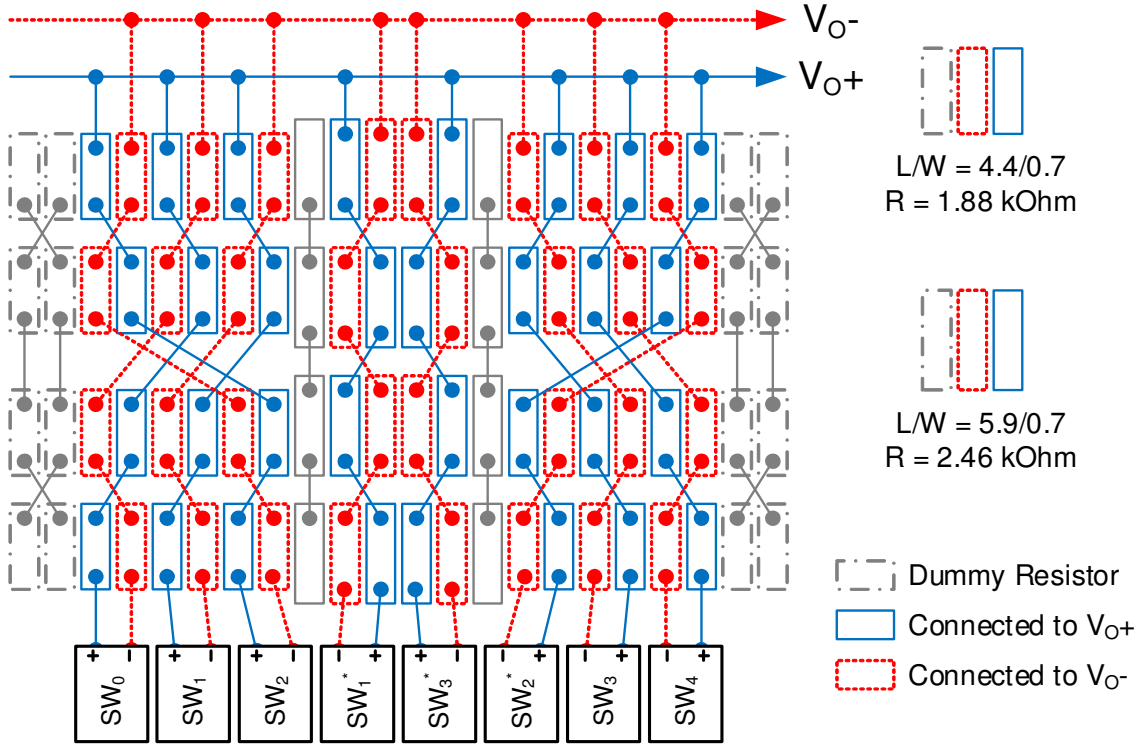


Figure 3.19: Layout arrangement of the resistors

and evaluate its SFDR.

4. Increase the width of the N/PMOS transistors in the buffer stage (I_1 and I_2) and the phase-to-duty-cycle converter (MN1, MN2, MP1 and MP2) until the simulated linearity of the output signal meets the target, $\text{SFDR} < -60 \text{ dBc}$ at $f_0 = 500 \text{ MHz}$.

It should be mentioned that the sizing procedure of step 4 should follow the optimal inverter chain sizing factor [79],

$$\frac{1}{2} \left(\frac{W}{L} \right)_{MN1} : \left(\frac{W}{L} \right)_{NMOS-of-I1} = 1 : 2.7 \quad (3.16)$$

This ratio is also applied to PMOS transistors. Furthermore, the layout of the PDC

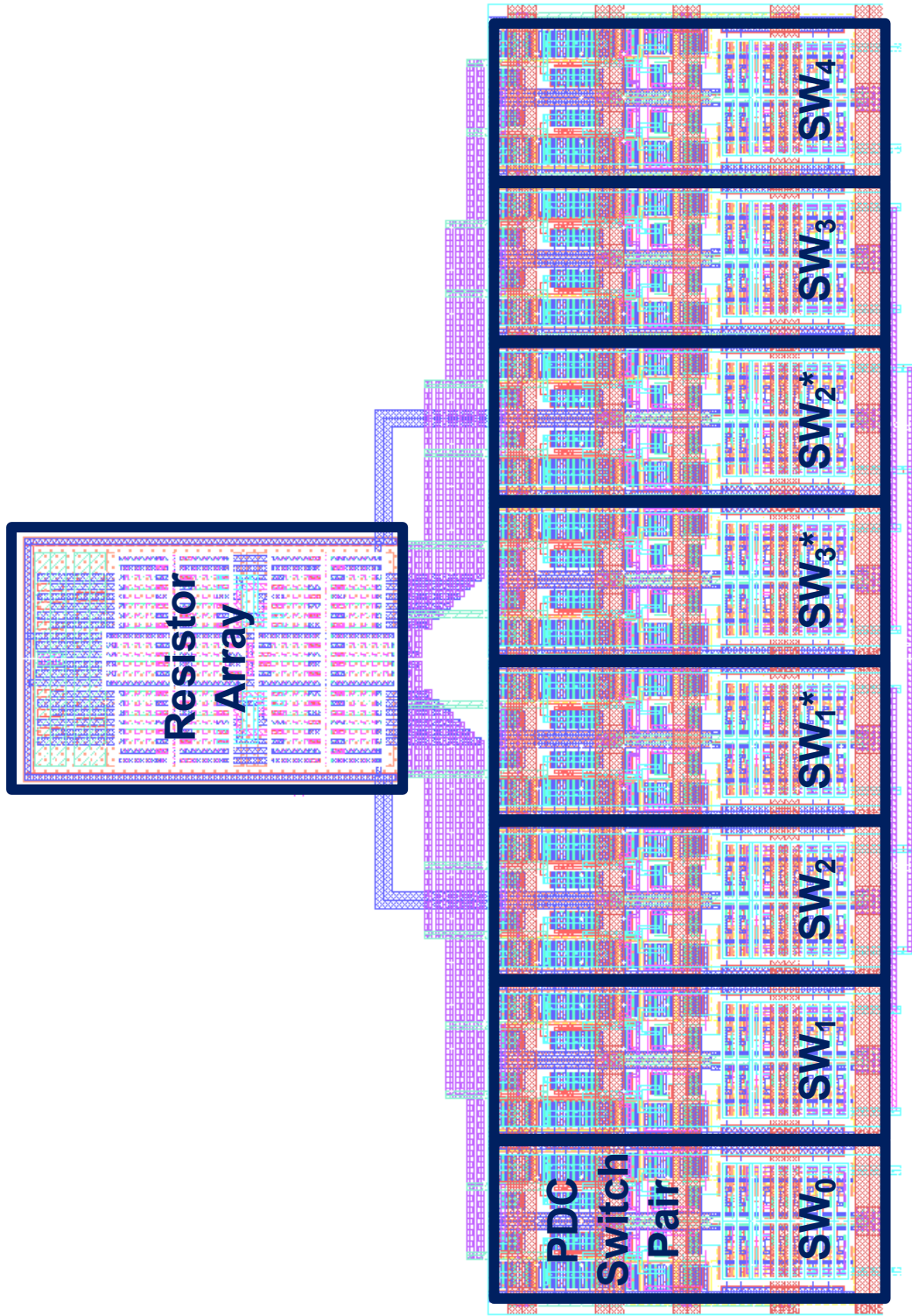


Figure 3.20: Layout of the PDC switches and the weighted resistor summing network.

switches and the resistor summing network is shown in Fig. 3.20. PDC switch pairs (SW_i) are indicated.

3.4 Iterative SFDR Optimization

3.4.1 Error Analysis

Section 3.3 introduces the synthesizer implementation. Particularly, a PDC converter was implemented to make the duty cycle of the clocks controlled by the phases. Therefore, this design reduces the number of error sources to two, the phase (time) error and the amplitude error, compared to the three sources in [66]. The phase error is defined as $\Delta\theta_i$ for the i -th phase generated in the clock generation block, which is caused by uneven delay stages in the RO or the DLL. The amplitude error is expressed as $\Delta\alpha_i$ for each FIR filter tap. The amplitude error is attributed to the nonuniform resistance values in the summing network imposed by the PVT variations.

Consider the complementary clock phases, θ_i and $\theta_{i+\frac{N}{2}}$, the shifted clock signal $\phi_i(t)$ can be generated by the switch pair introduced above. Its normalized form $s_i(\theta)|_{\theta=\omega_0 t}$ without the DC component is

$$s_i(\theta) = \begin{cases} 1 & \theta_i + \Delta\theta_i + 2k\pi \leq \theta \leq \theta_{i+\frac{N}{2}} + \Delta\theta_{i+\frac{N}{2}} + 2k\pi \\ -1 & \textit{otherwise} \end{cases} \quad (3.17)$$

Noting that $\theta_{i+\frac{N}{2}} = \theta_i + \pi$, the Fourier coefficients of the i -th clock $\phi_i(t)$ can be derived as

$$\begin{aligned} A_{k,i} &= \frac{1}{\pi} \int_0^{2\pi} s(\theta) \cos(k\theta) d\theta \\ &= -\frac{2}{k\pi} [C_{k,i} \sin(k\theta_i) + D_{k,i} \cos(k\theta_i)] \end{aligned} \quad (3.18)$$

$$\begin{aligned}
B_{k,i} &= \frac{1}{\pi} \int_0^{2\pi} s(\theta) \sin(k\theta) d\theta \\
&= \frac{2}{k\pi} [C_{k,i} \cos(k\theta_i) - D_{k,i} \sin(k\theta_i)]
\end{aligned} \tag{3.19}$$

where

$$C_{k,i} = \cos(k \cdot \Delta\theta_i) - (-1)^k \cos\left(k \cdot \Delta\theta_{i+\frac{N}{2}}\right) \tag{3.20}$$

$$D_{k,i} = \sin(k \cdot \Delta\theta_i) - (-1)^k \sin\left(k \cdot \Delta\theta_{i+\frac{N}{2}}\right) \tag{3.21}$$

Here, $\phi_i(t)$ is equivalent to $\phi(t + \theta_i)$ for Eq. 3.6. Now, we can rewrite Eq. 3.7 and Eq. 3.8 as

$$X_k = -\frac{2}{k\pi} \sum_{i=0}^{M-1} (\alpha_i + \Delta\alpha_i) [C_{k,i} \sin(k\theta_i) + D_{k,i} \cos(k\theta_i)] \tag{3.22}$$

$$Y_k = \frac{2}{k\pi} \sum_{i=0}^{M-1} (\alpha_i + \Delta\alpha_i) [C_{k,i} \cos(k\theta_i) - D_{k,i} \sin(k\theta_i)] \tag{3.23}$$

Consequently, the combined waveform $F(t)$ in Eq. 3.6 is a function of both the magnitude error $\Delta\alpha_i$ and the phase error $\Delta\theta_i$. The i -th order harmonic distortion is defined as

$$HD_k = 20 \log \left(\frac{H(k)}{H(1)} \cdot \frac{\sqrt{X_k^2 + Y_k^2}}{\sqrt{X_1^2 + Y_1^2}} \right) \quad (k = 2, 3, 4, \dots) \tag{3.24}$$

where $H(k)$ is the attenuation introduced to the k -th order harmonic by the output filter. Furthermore, we can take the non-linearity of the output buffer into consideration. The output waveform of the buffer, which accepts $F(t)$ as its input, can be expressed as,

$$F'(t) = \sum_{i=1}^m b_i [F(t)]^i \tag{3.25}$$

where coefficients b_1, b_2, \dots, b_m model the source follower's non-linearity (up to the m -th order). The final expression of the i -th order harmonic distortion is complicated. However, it's still a function of $\Delta\alpha_i$ and $\Delta\theta_i$. It was determined that there exists a set of $\Delta\theta_i$ which minimizes each HD_k , and thus, achieves the maximum spur-free dynamic range (SFDR).

3.4.2 Min-Max Optimization

To find the maximum SFDR is to solve an optimization problem whose cost function is related to the SFDR definition. Let's consider only the harmonic distortion up to the q -th order wherein the cost function can be defined as

$$\begin{aligned} F_{cost}(\Delta\theta) &= \max\{HD_2, HD_3, \dots, HD_q\} \\ \Delta\theta &= \{\Delta\theta_0, \Delta\theta_1, \dots, \Delta\theta_{N-1}\} \end{aligned} \quad (3.26)$$

Note that $F_{cost}(\Delta\theta)$ is -SFDR, which only takes up to the q -th order into consideration. Based on this cost function, we can define a multidimensional min-max optimization problem[80], solving

$$\begin{aligned} \min \quad & F_{cost}(\Delta\theta) \\ \text{subject to} \quad & \theta_{lower} \leq \Delta\theta < \theta_{upper} \end{aligned} \quad (3.27)$$

where θ_{lower} and θ_{upper} are the lower and upper bounds of the control variables. In addition, not only the SFDR, but also the total harmonic distortion (THD) can be used as the cost function. The THD-based (count up to the q -th order) cost function is defined as $F_{cost,thd}(\Delta\theta) = \sqrt{HD_2^2 + \dots + HD_q^2}$. Simulation results show similar performance for the SFDR-based and the THD-based cost functions. For simplicity, only the SFDR-based cost function is discussed.

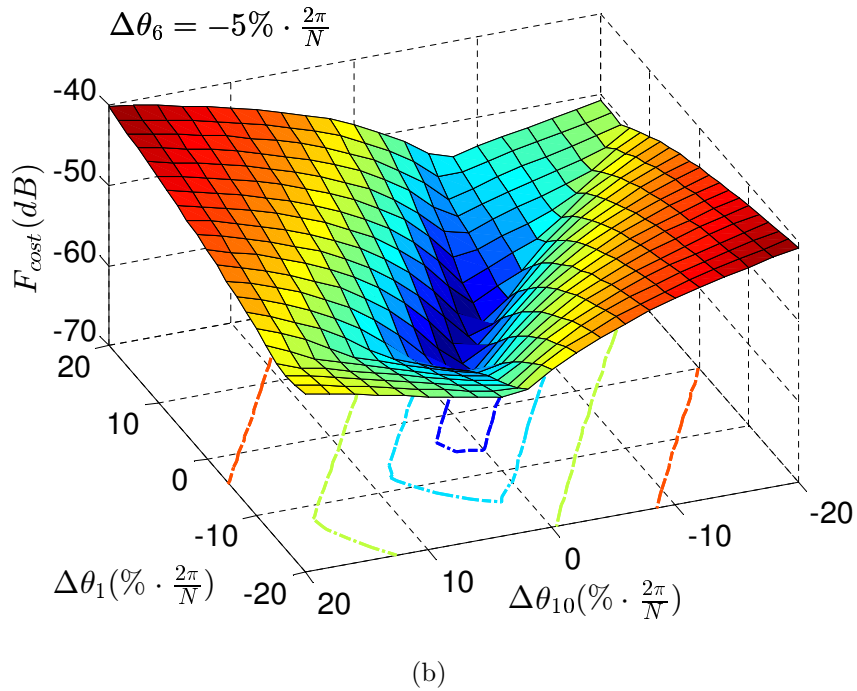
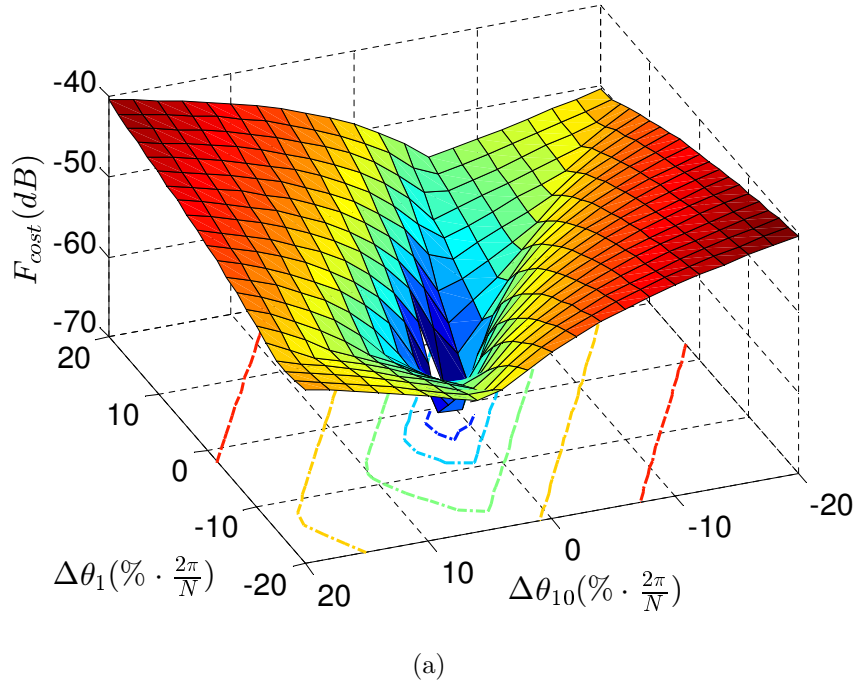


Figure 3.21: Cost function surface for $M = 5$, $N = 12$, $q = 10$, and sweeping $\Delta\theta_1$ and $\Delta\theta_{10}$ (a) with ideal tap coefficients, (b) with ideal tap coefficients and a fixed $\Delta\theta_6 = -5\% \cdot \frac{2\pi}{N}$, or (c) with a non-ideal tap coefficient $\Delta\alpha_2 = 5\% \cdot \alpha_2$.

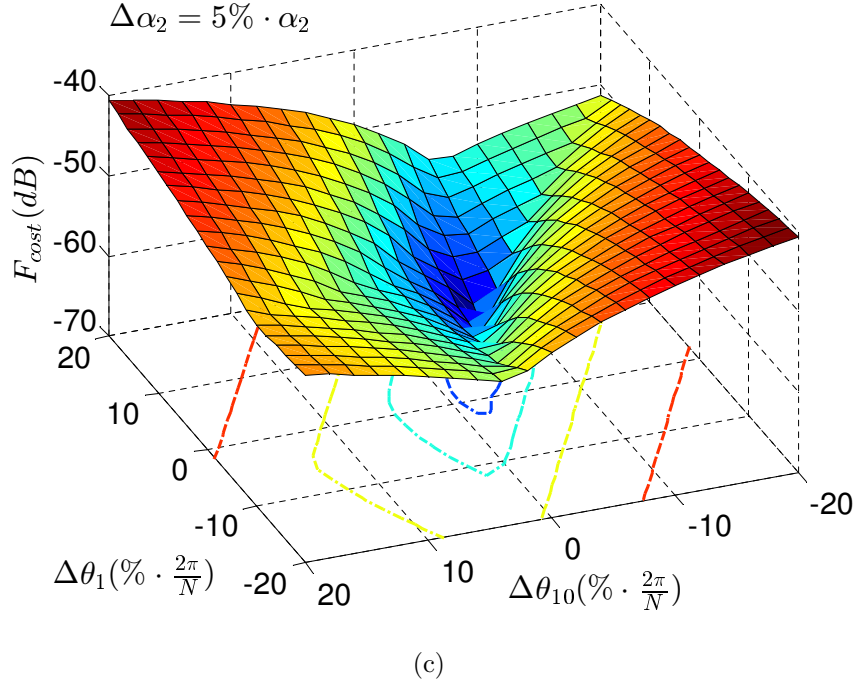


Figure 3.21: Continued.

Let's investigate the surface of the cost function with different variables. Using the proposed circuit implementation as an example, we have $M = 5$ and $N = 12$. Consider up to the 10th order harmonic, which means $q = 10$. In addition, the output filter is a first order LPF, plus the integration capacitor C_S in the summing network plays the role of another first order filtering. Hence, we have $H(k) = (1 + k^2)^{-1}$ for Eq. 3.24. Fig. 3.21a demonstrates the impact of phase errors with ideal tap coefficients. Errors are imposed on the second and the 11th phases, $\Delta\theta_1$ and $\Delta\theta_{10}$, sweeping from $-20\% \cdot \frac{2\pi}{N}$ to $20\% \cdot \frac{2\pi}{N}$. We can find that the minimum cost of $-\infty$ is achieved when $\Delta\theta_1 = \Delta\theta_{10} = 0$, which means no errors at all. Fig. 3.21b further adds a fixed phase error, $\Delta\theta_6 = -5\% \cdot \frac{2\pi}{N}$, to Fig. 3.21a's condition. The minimum cost is -62.65 dB and locates at $\Delta\theta_1 = -\Delta\theta_{11} = 4\% \cdot \frac{2\pi}{N}$. This proves that the error introduced by one element in the set of $\Delta\theta$ can be compensated by changing the others. On

the other hand, Fig. 3.21c adds a non-ideal tap coefficient (amplitude error) instead. The new coefficient for the 3rd tap is $\alpha'_2 = 1.05$ while the other taps keep unchanged. Sweeping results give the minimum cost of -64 dB when $\Delta\theta_1 = -\Delta\theta_{11} = 4\% \times \frac{2\pi}{N}$. Furthermore, this reveals that the amplitude error can also be corrected by changing $\Delta\theta$. To conclude, we can apply an optimization algorithm to change $\Delta\theta$, which maximizes the SFDR of the output sinusoidal waveform.

3.4.3 Iterative Optimization Algorithm

Table 3.2: Iterative optimization algorithm

```

for i = 0 to N - 1
    Pi = Initial Values

DO

    for i = 0 to M - 1
        {Pi, Pi+6} =
            better (
                {Pi + 1, Pi+6 + 1},
                {Pi - 1, Pi+6 - 1},
                {Pi, Pi+6}
            )

    for i = 0 to M - 1
        Pi+6 = better(Pi+6 + 1, Pi+6 - 1, Pi+6)

    for i = 0 to M - 1
        Pi = better(Pi + 1, Pi - 1, Pi)

UNTIL the max iteration number is achieved or no Pi changes.

```

This design proposes the use of an iterative optimization algorithm based on the Gradient Descent algorithm. In the real circuit, the phase shifts are tuned by

knobs. As described in Section 3.3, these knobs are the digitally controlled varactor array. Thus, the cost function and the problem description should be converted to the discrete version. The phase error set $\Delta\theta$ becomes $\{P_0, P_1, \dots, P_{N-1}\}$ in Eq. 3.26. And for Eq. 3.27, P_i should be integer numbers, and we have $0 \leq P_i < 2^j$, where j is the number of the binary-weighted varactors.

The main algorithm is listed in Table 3.2. The concept of the proposed optimization procedure is to reduce the multi-variable problem to multiple 1-dimensional (1-D) problems. In each 1-D solving step, select only one control variable, P_i, P_{i+6} , or $\{P_i, P_{i+6}\}$ pair to be increased or decreased by 1. This is called one move. The algorithm keeps the move which improves the cost. To make the move, a subfunction *better*(A, B, C) is defined, where A, B and C are three different moves. This subfunction will verify each move against the boundary condition and compare the cost function outcomes among all three moves. The best move with the minimum cost will update the selected control variable. The algorithm iteratively tunes the phases ($\{P_i, P_{i+6}\}$ moves) or the duty cycle (P_i or P_{i+6} moves) of the clock signal for each tap until the maximum iteration number is achieved or all P_i are kept unchanged in one full *DO..UNTIL* iteration as shown in Table 3.2.

3.4.4 Optimization Procedure Simulation

A transistor-level Monte-Carlo simulation is carried out to verify the effectiveness of the iterative optimization algorithm. The simulation introduces mismatches to the transistors in the ring oscillator and the resistors in the summing network, which are the major sources of the phase errors and the amplitude errors. A total of 200 Monte-Carlo error data groups were collected. And the phase shifts, which are controlled by different P_i , were also simulated. The optimization procedures were conducted and analyzed in MATLAB. Fig. 3.22 gives the F_{cost} distribution before and after

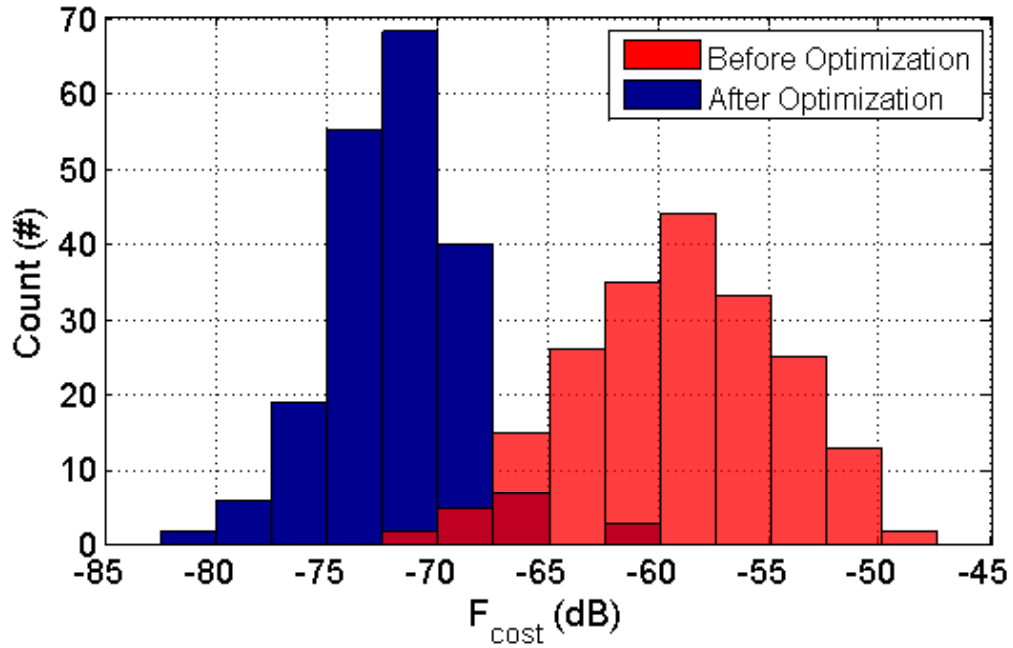


Figure 3.22: Simulated F_{cost} distribution with Monte-Carlo simulation before/after the optimization procedure.

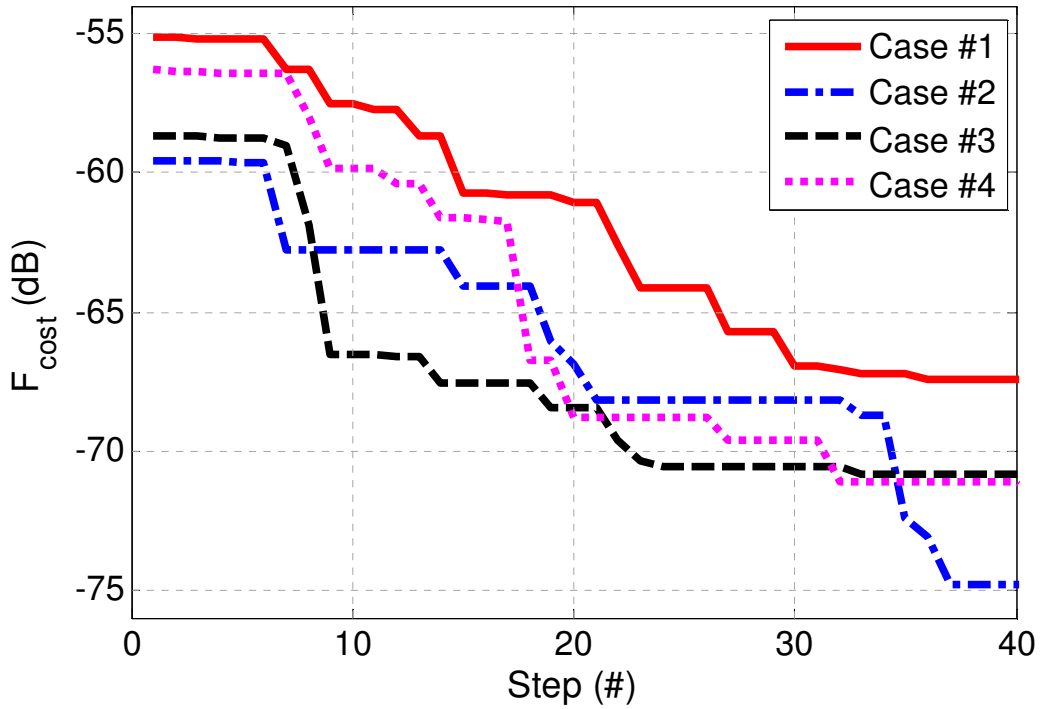


Figure 3.23: Cases of the proposed optimization procedure.

the optimization procedure. A comparison between these two groups of data shows a significant SFDR enhancement from an average below 60 dBc to 72 dBc, and the spread range also narrowed. Fig. 3.23 further demonstrates the execution steps of the algorithm. Each move takes one step. A minimum value of F_{cost} can be found after several iterations. In addition, this simulation is based on the output sine-wave frequency of 500 MHz. The proposed algorithm has proven capable of improving the linearity of the proposed sine-wave synthesizer. It will be applied after the chip fabrication.

3.4.5 Discrete Phase Shifter

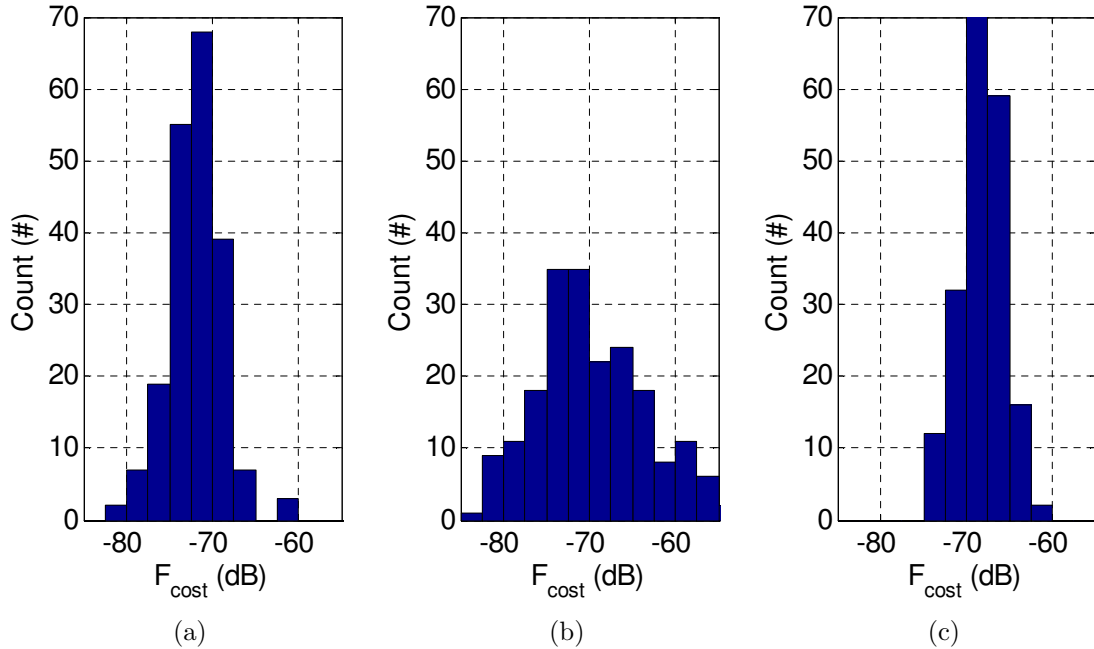


Figure 3.24: Simulated F_{cost} distribution after the optimization: (a) $f_0=500$ MHz, 3-bit phase shifter, (b) $f_0=100$ MHz, 3-bit phase shifter, and (c) $f_0=500$ MHz, 2-bit phase shifter.

As demonstrated in Fig. 3.13, the phase shifter is a discrete capacitor array, thus its tuning range and resolution will affect the ability of the proposed optimization procedure. On the one hand, when the oscillator’s output frequency decreases, the delay time in each delay cell gets longer, and so as the delay (phase) mismatch. However, the tuning range of the phase shifters will not change because they are isolated from the oscillator. If one delay cell’s mismatch that needs to be compensated was larger than the phase shifter’s range, a full compensation would be impossible. Fig. 3.24 shows such an impact. Fig. 3.24a is the distribution of F_{cost} after the optimization obtained from Fig. 3.22. If the ring oscillator’s output frequency was reduced from 500 MHz to 100 MHz, the average of F_{cost} would also reduce from 72 dBc to 69.8 dBc, as shown in Fig. 3.24b. The solution is to increase the number of capacitors in the phase shifter and thus widen its tuning range. On the other hand, the resolution of the phase shifter is another concern. Because the linearity is a very sensitive parameter, finer phase resolution can help to find a better result. By removing the smallest capacitor from the capacitor array of the proposed phase shifter, we can get a 2-bit approach. Its simulated F_{cost} distribution after the optimization is illustrated in Fig. 3.24c. The average is 68.3 dBc, lower than that of the previous 3-bit approach. Therefore, for the discrete phase shifter, a tradeoff should be made among the linearity, the frequency range and the tuning resolution.

3.4.6 Temperature Stability Analysis

As discussed in Section 3.3, the proposed optimization loop relies on a spectrum analyzer to obtain F_{cost} , which is not practical for fully integration. A possible solution is to make a one-time optimization after the chip fabrication. The one-time optimization is able to mitigate the error introduced by the process variation, which is fixed after the production, and the supply voltage, which can be fixed by a power

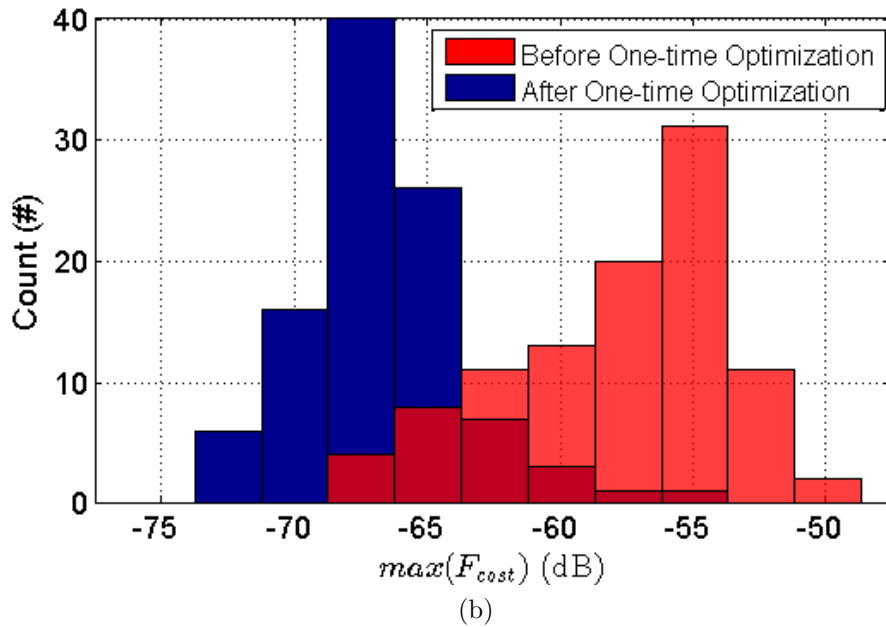
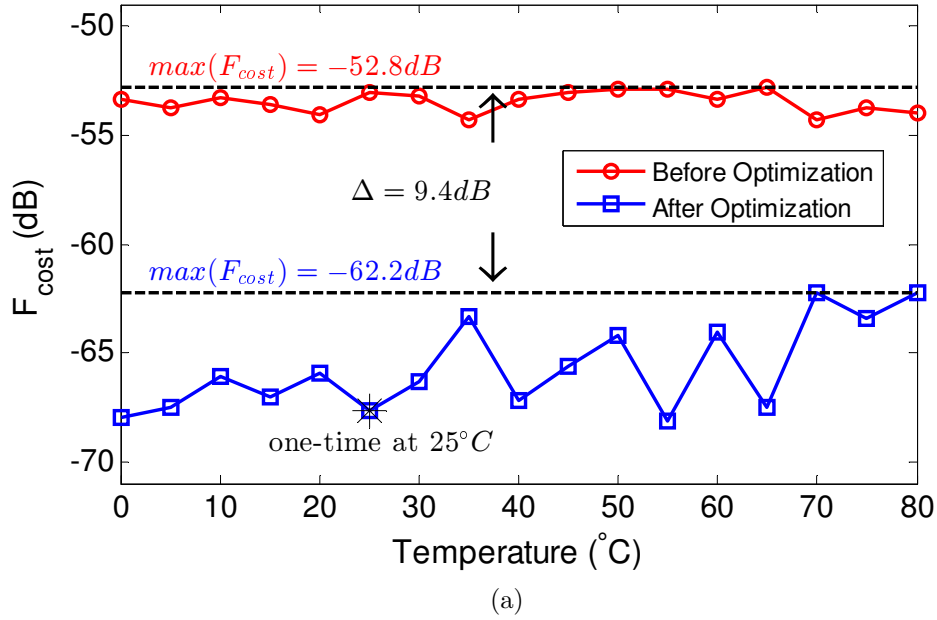


Figure 3.25: Temperature stability analysis: (a) F_{cost} fluctuation versus temperature, (b) simulated $\max(F_{cost})$ distribution before/after an one-time optimization, and (c) simulated distribution of Δ due to a one-time optimization.

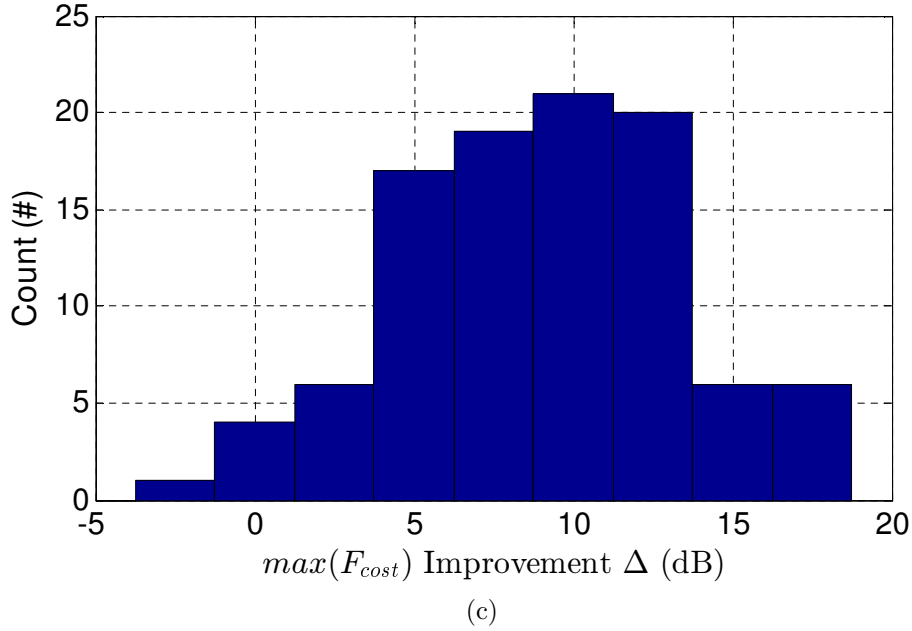


Figure 3.25: Continued.

management circuit. However, temperature is difficult to be stabilized when the chip is working. Therefore, the temperature stability of the proposed synthesizer should be further evaluated.

100 runs of Monte-Carlo simulation, plus a temperature sweep from 0 to 80 °C, are carried out to examine the circuit performance. 500MHz sine-wave output is selected. Fig. 3.25a demonstrates one case of the F_{cost} -versus-temperature fluctuation. No clear link can be found between the linearity and temperature. Even a 5 °C temperature drift can result in several dBs' improvement or degradation on F_{cost} . This phenomenon may contribute to the non-uniform mobility of transistors in the ring oscillator. Therefore, temperature has an uneven impact on each delay cell, and thus, the linearity change is difficult to predict. To deal with such situation, only the worst case of the linearity, $\max(F_{cost})$, will be considered. When talking about the worst case, we can find the optimization procedure is still able to improve

its performance. Fig. 3.25a shows F_{cost} before and after a one-time optimization at 25 °C. It means that the optimization procedure is used only once to obtain an optimized control word at 25 °C, and this new word is applied to the synthesizer across the whole temperature range. Compared to the synthesizer using the default control word, the optimized one has an improvement Δ of 9.4 dB for $\max(F_{cost})$. Moreover, the distribution of $\max(F_{cost})$ before and after the one-time optimization is summarized in Fig. 3.25b. The mean values are -57 dB and -66 dB separately. The distribution of the improvement Δ is reported in Fig. 3.25c with an average value of 9dB. To conclude, although the temperature change may degrade the linearity of the output sine wave, the proposed optimization procedure is still necessary to further suppress the harmonics. Only a one-time optimization is required, and then the proposed synthesizer can work at different temperatures without an optimization loop.

3.4.7 Clock with Jitter

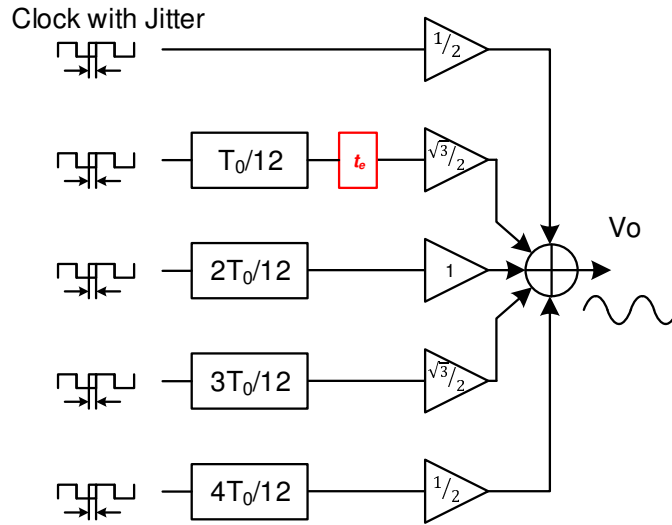
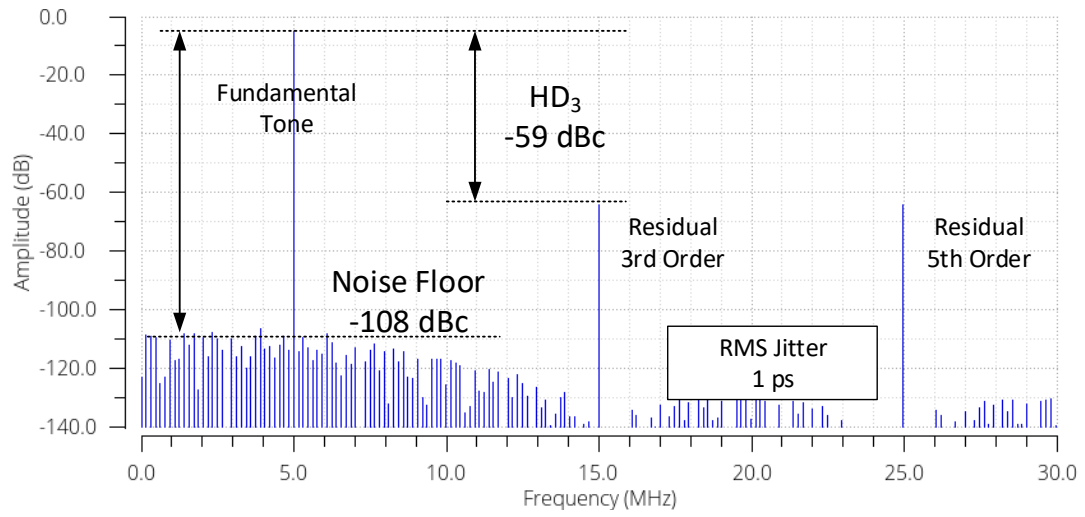
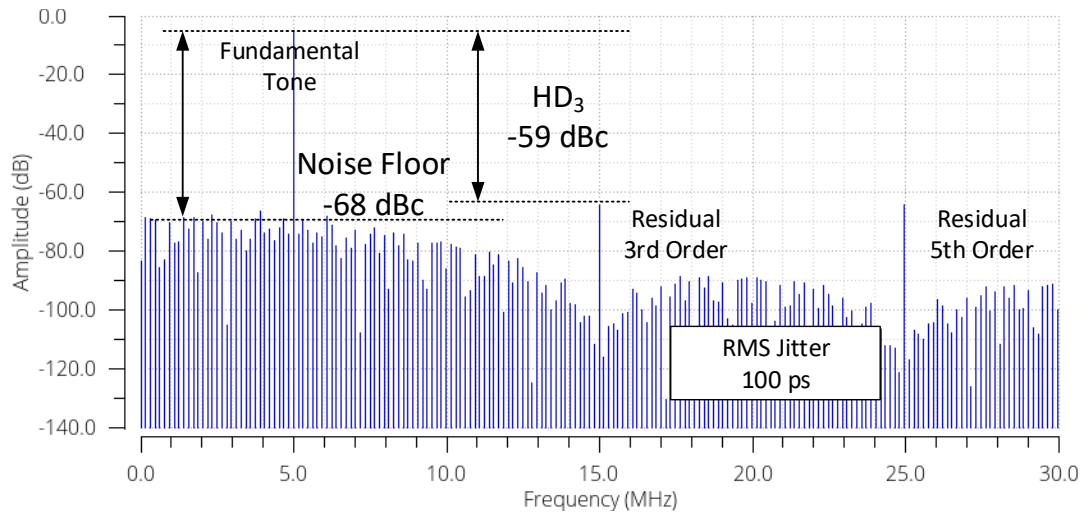


Figure 3.26: Behavior model for evaluating the impact of the clock jitter.



(a)



(b)

Figure 3.27: Simulation results of the jitter impact on the V_O spectrum with (a) 1-ps RMS jitter and (b) 100-ps RMS jitter.

A ring oscillator (Fig. 3.13) will generate clocks with jitters [81]. Fig. 3.26 shows the behavior model that is used to evaluate the impact of the clock jitter on the harmonic cancellation. Consider the worst case, in which jitters generated by every

delay stage in the ring oscillator are not correlated to each other. Therefore, five independent clock sources are adopted. The configuration of these sources with the same frequency (5 MHz) and jitter settings are briefly introduced in [82]. A fixed delay (phase) error, $t_e = 80$ ps, is also inserted to induce residual harmonics (limited SFDR) as analyzed in Section 3.4.1, Thus, we can observe whether the jitter will change the SFDR. Simulation results are illustrated in Fig. 3.27a and Fig. 3.27b for the 1-ps and 100-ps root mean square (RMS) jitter, respectively. To conclude, the clock jitter will not change the SFDR induced by the phase error. However, it changes the noise floor level—the bigger the clock jitter is, the higher the noise floor is. The increment rate of the noise floor is around 20 dB/dec.

3.5 Experimental Results

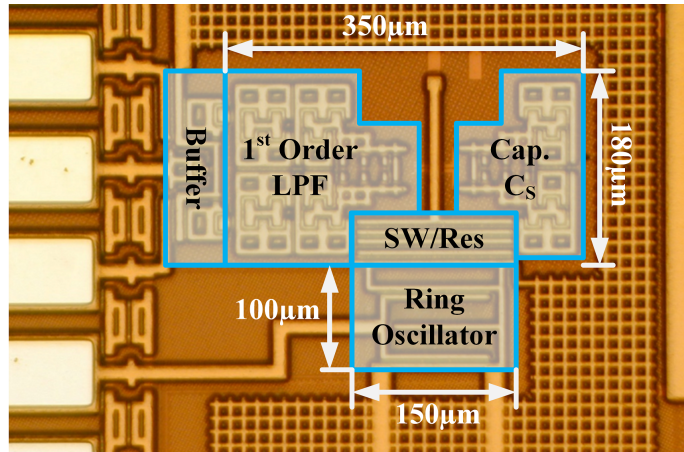


Figure 3.28: The die photograph of the proposed synthesizer.

The proposed sine-wave synthesizer is fabricated in 0.18 μm standard CMOS technology. The chip die micrograph is demonstrated in Fig. 3.28. The weighted resistor summing network, the integration capacitor C_S , and the programmable first

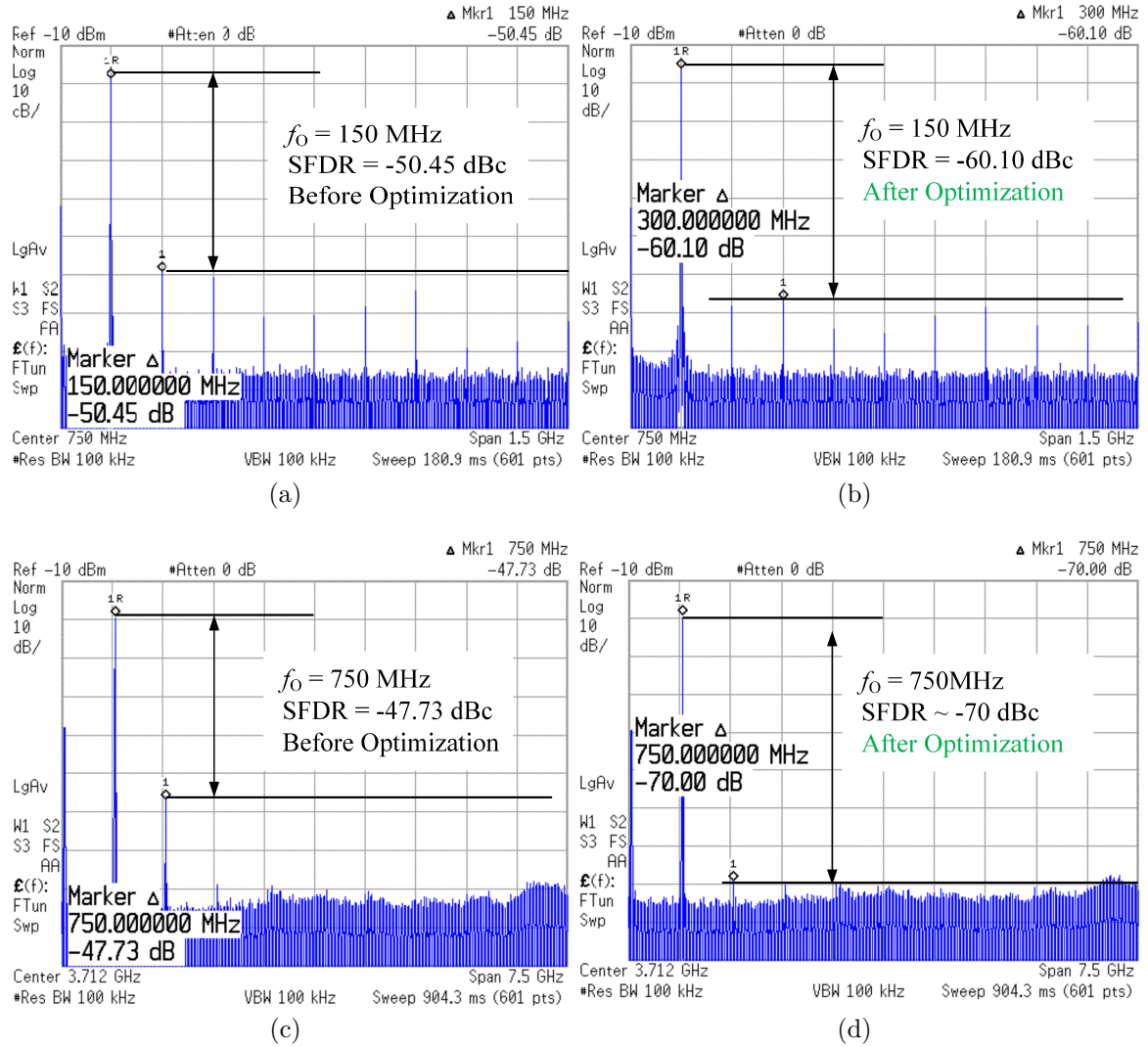


Figure 3.29: Chip experimental results: (a) Measured SFDR @ 150 MHz before the optimization, (b) measured SFDR @ 150 MHz after the optimization, (c) measured SFDR @ 750 MHz before the optimization, and (d) measured SFDR @ 750 MHz after the optimization.

order differential LPF occupies an area of $350 \mu\text{m} \times 180 \mu\text{m}$, and the ring oscillator takes about $100 \mu\text{m} \times 150 \mu\text{m}$. The whole synthesizer occupies 0.08 mm^2 silicon without the output buffer. All synthesizer blocks are working under the same supply voltage, although the supply can change from 1.0 V to 1.8 V. The measured lowest

power consumption (without the output buffer) is 9.11 mW under 1.0 V supply voltage, when the output frequency is the lowest at 150 MHz. For the highest 850 MHz sinusoidal wave, the highest power of 57.2 mW is achieved (without the output buffer) under the supply of 1.8 V. The ring oscillator draws about 33% of the total power, while the other two thirds are consumed by the summing network and the RC LPF. Power in the second part is mostly dissipated to drive the resistive part of the summing network. Because of the different phases, the output nodes of some switch blocks may be connected to the supply, while the others' are grounded. Therefore, direct paths exist between switch blocks and currents flow through the resistor loads in the summing network.

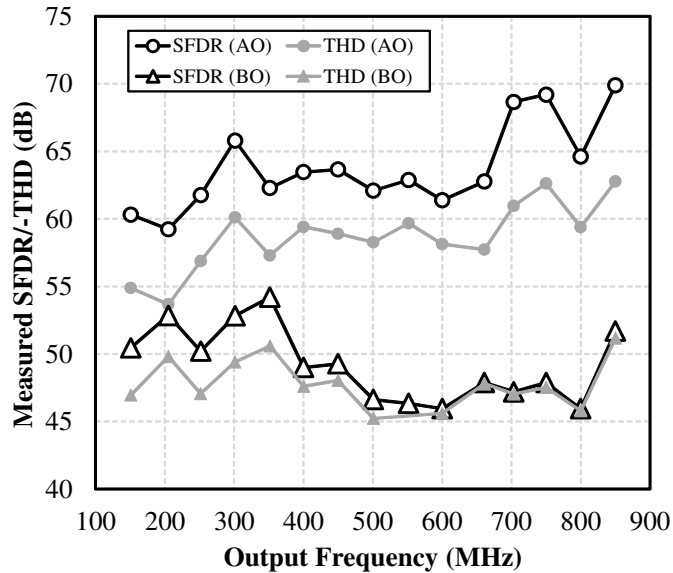


Figure 3.30: Measured SFDR/-THD vs. output frequency (BO: before the optimization, AO: after the optimization).

The proposed iterative SFDR optimization procedure is also tested. A 9.7 dB SFDR improvement is measured at 150 MHz after the optimization (Fig. 3.29a and

Fig. 3.29b), and a 22.3 dB improvement was obtained for a 750 MHz output (Fig. 3.29c and 3.29d). Particularly, both of the odd and even order harmonic cancellations relied on the clock phase matching in the ring oscillator, which becomes worse at higher frequencies. However, the odd order harmonics receive further suppression from the FIR architecture. Therefore, even order harmonics may dominate the linearity degradation at high frequency as shown in Fig. 3.29c. This result also proves the necessity and the effectiveness of the optimization procedure, which can fix the matching errors and improve the overall linearity.

Moreover, Fig. 3.30 compares the SFDRs and THDs (count up to the ninth order) measured before and after the optimization procedure. The measured improvements well match the predicted values in Fig. 3.25c. Even without one-time optimization, the SFDR is still above 45 dBc across the whole frequency range. This means the proposed synthesizer is capable of some applications that have low linearity requirements, such as that in [43]. After the optimization procedure, the minimum increase of the SFDR is 6.4 dB, while the maximum is nearly 22 dB. For higher frequency, the SFDR measured before the optimization is lower because the PVT variations impose more mismatches on clock phases. However, the effect of optimization is more significant. This is because, for higher frequency, the phase tuning range is wider compared to the clock period, and thus the optimization algorithm may find a better solution. To conclude, the weighted resistor summing architecture, the phase programmability and the iterative SFDR optimization algorithm make the synthesizer's linearity performance robust to the PVT variations.

Table 3.3 compares the performance of the proposed synthesizer to the other state-of-the-art works. SFDR and THD values are listed and the negative THD value are indicated by using a marker “*”. For those sine-wave synthesizers, which adopt the HC technique, Table 3.3 further summarizes the types of clock generation

Table 3.3: Comparison of sine-wave synthesizer performance

Ref.	F_{out}	Power (mW)	Area ($\times 10^{-3} mm^2$)	Process (nm)	Supply (V)	SFDR /-THD* (dBc)	Circuit Type	Clock Type	Summing Network	Filter Type	FoM
This work	150 ~ 850 MHz	9.1 ~ 57	80	180	1.0 ~ 1.8	60.3/54.9* @ 150 MHz, 62.6/57.7* @ 650 MHz, 70.0/62.6* @ 750 MHz	HC	Multi-phase Clock	Resistor Network	Passive LPF	170052
[43]	20 ~ 220 MHz	N/A	45.5	90	N/A	45 @ 100 MHz	HC	Multi-phase Clock	Current Mirror	N/A	N/A
[72]	DC ~ 1 GHz	130	100	55	N/A	55.2 @ 660 MHz	DDFS	N/A	N/A	N/A	45241
[74]	DC ~ 650 MHz	350	2000	90	1.2, 1.5	52 @ 612 MHz	DDFS	N/A	N/A	N/A	377
[73]	DC ~ 2500 MHz	460	2100	350 (SiGe)	3.3	45.7 @ 2.49 GHz	DDFS	N/A	N/A	N/A	508
[66]	5 ~ 11 MHz	3.34	186	130	1.2	72* @ 10 MHz	HC	Conditional Clock	Resistor Network	Passive LPF	72526
[71]	1.11 MHz	3.24	40	180	1.8	77* @ 1.11 MHz	HC	Conditional Clock	SC	SC-BPF	62508
[68]	10 MHz	20.1	200	350	3.3	54.8* @ 10 MHz	HC	Conditional Clock	SC	SC-BPF	1397
[67]	10.7 MHz	132	3150	350	3.3	53* @ 10.7 MHz	HC	Conditional Clock	DAC	CT-BPF	12
[62]	40 mHz ~ 40 KHz	N/A	640	350	3.3	63 @ 10 KHz	Filter	N/A	N/A	N/A	N/A
[70]	100 Hz ~ 100 KHz	N/A	N/A	N/A	N/A	51.9 @ 1 KHz	HC	Conditional Clock	Inverting Amplifier	CT-LPF	N/A
[69]	2 KHz ~ 180 KHz	N/A	N/A	N/A	30	82* @ 29 KHz	HC	Multi-phase Clock	Inverting Amplifier	Ideal	N/A

blocks, summing networks and output filters in detail. The Figure of Merit (FoM) for each design, except those using discrete components, is given by

$$FoM = \frac{f_o (MHz) \cdot 2^{SFDR_{worst}(dB)/6}}{P_{total} (mW) \cdot A (mm^2)} \quad (3.28)$$

where f_o is the maximum output frequency, $SFDR_{worst}$ is the measured worst SFDR, P_{total} is the maximum power consumption, and A is the active area.

To sum up, this proposed synthesizer performs better than those works in a similar frequency range. Additionally, references operating at much lower or higher frequencies are also included because of their architecture and good linearity performance. We can find that the proposed approach consumes much less power than that of the DDFS, and is highly area efficient. It pushes the sine-wave synthesizer to sub-1 GHz but still keeps the high linearity. The synthesizer's architecture, which is a combination of the multi-phase clock generation, the weighted resistor summing network, and the passive LPF, has been proven suitable for high frequency application.

3.6 Conclusions

A sine-wave synthesizer generating low distortion high frequency sinusoidal signals is proposed. It only involves square wave clocks of five phases, which have a 30° difference between every two adjacent phases, and three separate amplitudes are adopted to achieve the cancellation of the third, fifth, seventh and ninth order harmonics. 50% duty cycle and differential mode circuit architecture were introduced to reduce the even order harmonics, and an LPF was used to further smooth the output waveform. Thanks to the phase shifters in the ring oscillator and the PDC converters in the weighted resistor summing network, this design simultaneously enables the ability of correcting phase, amplitude and duty cycle errors. Moreover, an

iterative optimization algorithm was incorporated to compensate for the mismatches induced by the PVT variations and improve the linearity of the output waveform. The proposed synthesizer shows low distortion, wide bandwidth, high flexibility, and a robustness for broadband application. Moreover, the compact design, which has only logic gates and passive components, is shrinkable for future advanced IC process. Additionally, this approach represents one step in the direction of built-in optimization for integrated circuit design.

4. ON-CHIP TWO-TONE SYNTHESIZER BASED ON A MIXING-FIR ARCHITECTURE *

4.1 Background

Third order intermodulation distortion (IM3), as well as the IM3 interception point (IP3), give a direct relationship between the input signal level and the linearity. They are very important figures of merits (FoM) in the measurement and characterization of analog/RF circuits or systems. These parameters characterize the linearity of the device-under-test (DUT). To obtain these key metrics, the two-tone test method is the industry standard. It can be used to measure the linearity of a wide-bandwidth active filter [83], [57], a $\Sigma\Delta$ analog-to-digital converter (ADC) [84], a power amplifier [85], and so on. The two-tone test has also been applied to a wider aspect, such as sensing the electro-chemical impedance of protein [86], and detecting the electro-thermal modulation of conductivity in passive antennas [87].

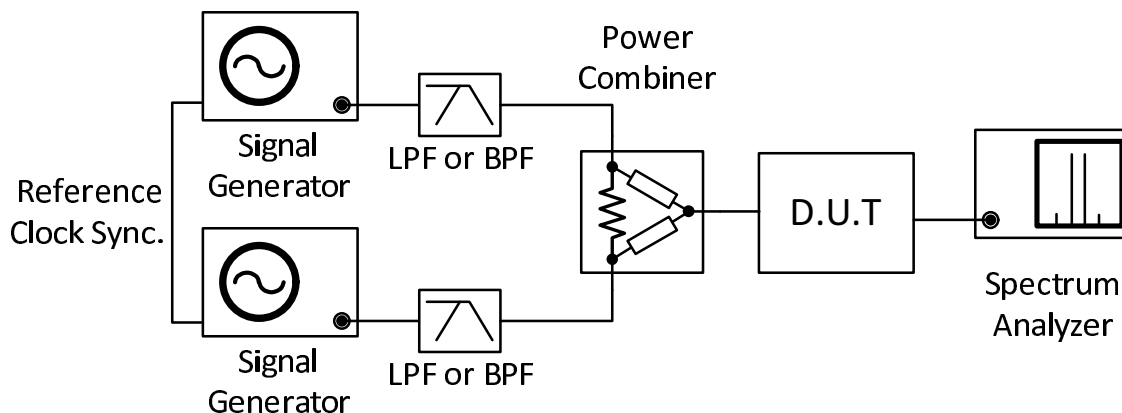


Figure 4.1: Traditional two-tone test configuration using testing equipment.

*Part of this chapter is reprinted from “On-chip two-tone synthesizer based on a mixing-FIR architecture” by C. Shi and E. Sanchez-Sinencio, IEEE Journal of Solid-State Circuits, vol. PP, pp. 1–12, copyright 2017 by IEEE.

Fig. 4.1 demonstrates the traditional two-tone test configuration. Two arbitrary waveform signal generators (AWG) are generating two single tones at different frequencies. Usually, the generated sinusoidal waveforms are not just two tones, but contain some other harmonic components. Thus, two passive low-pass filters (LPF) or band-pass filters (BPF) are required to suppress the residue harmonics. A power combiner is used to combine the two single tones and stimulate the DUT with the desired two-tone signal. The spectrum analyzer at the DUT's output will measure the IM3. However, this conventional test bench is bulky and costly. Later in this section, analysis will show that the IM3 tones in the stimulus are the main measurement error contributor. Therefore, we are proposing a compact on-chip two-tone synthesizer that focuses on suppressing the 3rd-order distortion.

Although the two-tone test is a critical measurement methodology, few papers have discussed potential for on-chip built-in implementation. [16] proposes a direct two-tone generation using two voltage-controlled oscillators (VCO) in a phase-locked loop (PLL). The two VCOs work at different frequencies, and their output waveforms are added together by a linear adder. However, [16] only reports the simulation results; thus, it is difficult to predict the circuit overhead and non-ideality of their proposed design. Digital-to-analog converter (DAC) can also be used for on-chip two-tone generation. [88] proposes a 10-bit current steering DAC with an improved dynamic element matching (DEM) technique. It achieves -62.16 dBc IM3 with 245 and 247 MHz two-tone signals at 500 MS/s. [89] further extends the range of IM3 < -61 dBc over 1.4 GHz in 40 nm technology. It should be noted that [88] and [89] achieve smaller footprints (0.034 mm² in 180 nm node and 0.016 mm² in 40 nm node, respectively) than that of the other state-of-the-art DACs. However, both of them did not include the extra digital area used for encoding the two-tone signal. In fact, a memory storage block for “replaying” the waveform look-up table or a digital

signal processor (DSP) will consume a considerable area overhead, such as the digital cores in [90] and [91]. They push the 3rd-order intermodulation below -80 dBc, but occupy 1.6 mm^2 [90] and 3.3 mm^2 [91] in the 65 nm node. In addition, [92] and [93] proposed 6-bit DACs with a design-for-testability (DFT) memory. Their compact design is appropriate to be embedded into a modern system-on-chip (SoC). However, the 5KB on-chip memory with 0.048 mm^2 is still larger than the DAC's 0.035 mm^2 core circuit, in 28 nm process. Therefore, these stand-alone high-linearity DACs should be considered oversized for the purpose of an on-chip linearity built-in self-test (BIST), which requires a trade-off between the circuit overhead and the third order distortion suppression.

Developing a compact two-tone signal synthesizer that can meet the emerging on-chip test demands of analog/RF circuits faces many design challenges. On the one hand, a built-in linearity optimization system has been proposed in [19] for an RF low noise amplifier (LNA). It integrates an envelope detector, an ADC, an IM3 calibration unit and an on-chip spectrum analyzer, which is introduced in [18]. Additionally, [18] further analyzes the measurement precision impacted by DAC bit numbers and FFT point numbers for an on-chip linearity BIST system. However, such a proposed system relies on external two-tone excitation signals, preventing its full integration. On the other hand, high-linearity single-tone sinusoidal synthesizers have been well researched and applied to the BIST architecture. [66] proposes a digital harmonic cancellation (HC) technique to generate a sinusoidal waveform from only square wave digital clocks. This achieves a 72 dB total harmonic distortion (THD) at 10 MHz output frequency. A high-linearity sine-wave synthesizer, based on a finite impulse response (FIR) filter architecture, is further proposed in [44] for high frequency operation. It demonstrates the sensitivity of timing mismatches for the harmonic cancellation effect and adopts an optimization loop to reduce the errors

in the clock distribution network. Similarly, the harmonic cancellation technique can be applied to the two-tone signal generation.

Due to the cumbersome approach of synthesizing and combining two high-frequency sine-wave tones with high linearity, we propose to generate the single tone at a relatively low frequency, move it to the desired high frequency band, and duplicate the single tone. In this section, an on-chip low-IM3 two-tone synthesizer is proposed, consisting of a cascade FIR architecture and a passive mixer. The cascade FIR architecture implements a harmonic cancellation technique to suppress a large number of odd-order frequency components. It benefits from operating at the low “baseband” frequency (ω_0), which reduces the impact of the delay mismatches in the clock distribution network. Thus, the accuracy of harmonic suppression can be improved. The passive mixer utilizes the nature of up-conversion, and mirrors the “baseband” to two side bands around the LO frequency ω_{LO} . Two tones with equal amplitudes can then be obtained. The mixer’s simple structure is able to minimize the linearity degradation introduced by the MOS transistor switches. The proposed design aims at synthesizing two tones with low IM3 from DC to 1 GHz.

This section is organized as follows. The impact on the linearity measurement accuracy by using weakly nonlinear stimulus is analyzed in Section 4.2, as well as the principles of the proposed two-stage cascade FIR architecture. The system architecture and detailed circuit implementation are introduced in Section 4.3. Section 4.4 analyzes the non-ideality effects in the proposed system, such as the timing and the current mismatches, the mixer’s linearity, the LO leakage, and the aliasing issue. Section 4.5 shows the measurement results, followed by conclusions in Section 4.6.

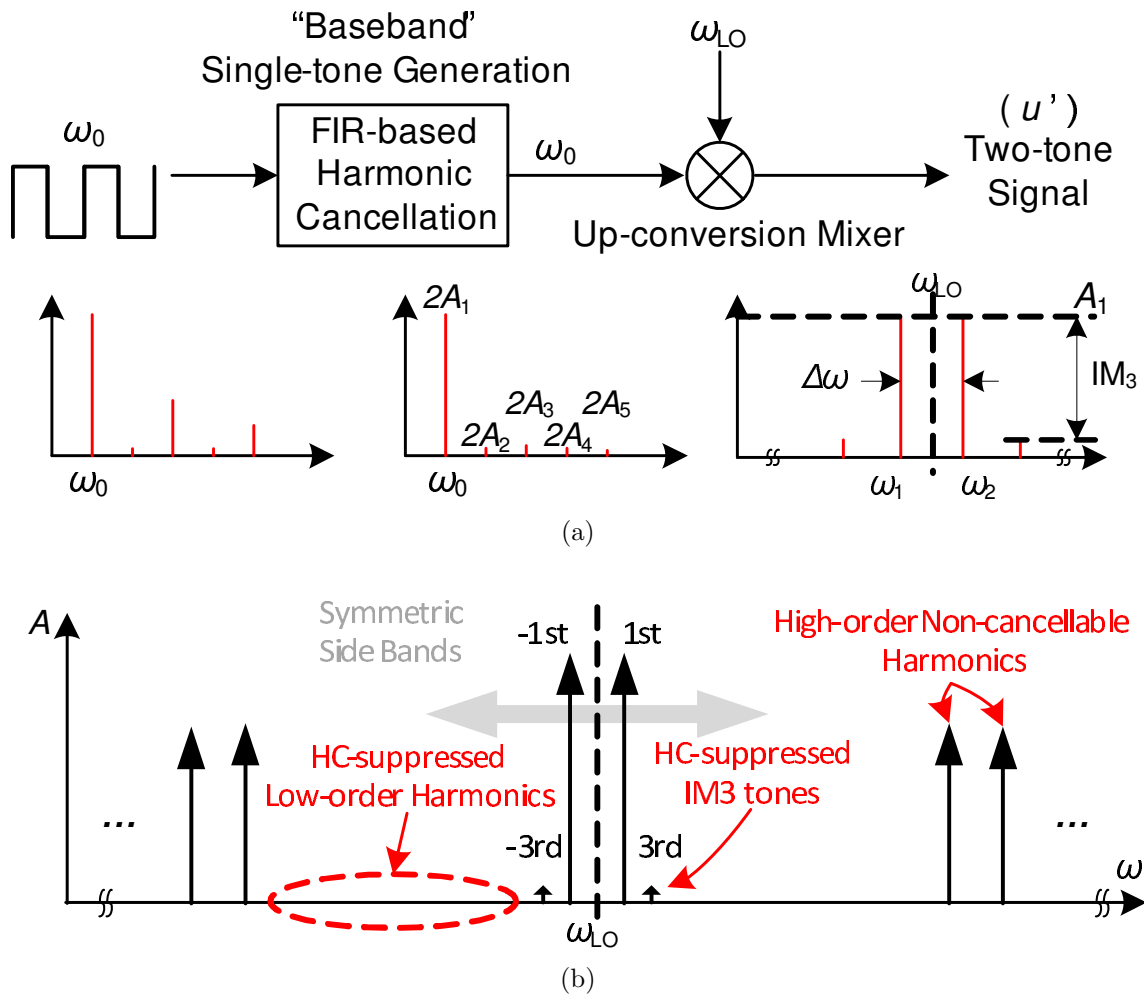


Figure 4.2: Two-tone generation architecture concept: (a) Mixing-FIR two-tone generation and (b) output two-tone signal spectrum.

4.2 Two-tone Generation

4.2.1 Two-tone Signal Generation Architecture

The systematic concept of the proposed mixing-FIR two-tone generator for on-chip linearity BIST is depicted in Fig. 4.2a, where ω_0 and ω_{LO} are external clock signals with different frequencies. It implements the fully differential circuit to cancel the even-order harmonics, and uses the FIR-based harmonic cancellation (HC) tech-

nique to suppress multiple odd-order harmonics. The “baseband” quasi-sinusoidal single-tone signal is generated at the frequency, $\omega_0 = \Delta\omega/2$, where $\Delta\omega = \omega_2 - \omega_1$ is the expected difference frequency between the two tones. A highly linear up-conversion mixer was also adopted to mirror the single-sideband to dual bands and move them around the desired frequency ω_{LO} . Note that we have $\omega_1 = \omega_{LO} - \omega_0$ and $\omega_2 = \omega_{LO} + \omega_0$. Ideally, the two-tone test method (Fig. 4.1) can generate mostly pure spectrum with only two tones. Correspondingly, as derived in [44], all odd-order harmonics can be suppressed if and only if the FIR architecture has an infinite number of FIR taps, which is not practical in the real circuit design. On the contrary, a limited number of FIR taps eliminates lower order harmonics but leaves higher order non-cancellable harmonics, and they are also up-converted by the mixer, as shown in Fig. 4.2b. The spectrum of the output two-tone signal in the proposed design is not theoretically pure. Its impact on IM3 measurement precision is analyzed in Section 4.2.2, and its aliasing issue is discussed in Section 4.4.4.

4.2.2 Linearity Test using Weakly Nonlinear Stimulus

The DUT is characterized as a nonlinear system (ignoring the DC term),

$$y = k_1u + k_2u^2 + k_3u^3 + \dots \quad (4.1)$$

where k_1 is the linear gain of the DUT, and k_2, k_3, \dots , indicate the DUT’s nonlinear coefficients.

If a single-tone cosine waveform is applied to the input,

$$u = A_1 \cos(\omega t) \quad (4.2)$$

where A_1 is the amplitude and ω is the angular frequency of the cosine signal. The

output becomes

$$y = \frac{k_2}{2}A_1^2 + \left(k_1 + \frac{3}{4}k_3A_1^2\right)A_1 \cos(\omega t) + \frac{k_2}{2}A_1^2 \cos(2\omega t) + \frac{k_3}{4}A_1^3 \cos(3\omega t) + \dots \quad (4.3)$$

The n -th order harmonic distortion (HD_n) is defined as the ratio of the n -th order harmonic magnitude to the fundamental magnitude, assuming $k_1 \gg \frac{3}{4}k_3A_1^2$. For example, the second order and third order harmonic distortions are expressed as

$$\text{HD}_2 \approx \frac{1}{2} \frac{k_2}{k_1} A_1 \quad (4.4)$$

$$\text{HD}_3 \approx \frac{1}{4} \frac{k_3}{k_1} A_1^2 \quad (4.5)$$

Also the total harmonic distortion (THD) is given by

$$\text{THD} = \sqrt{\text{HD}_2^2 + \text{HD}_3^2 + \dots} \quad (4.6)$$

Note that HD_n and THD are all correlated to the input signal's amplitude A_1 .

If the input of the DUT is a combination of two sinusoidal waveforms with the same amplitude of A_1 (two-tone test),

$$u = A_1 \cos(\omega_1 t) + A_1 \cos(\omega_2 t) \quad (4.7)$$

Apply (4.7) to (4.1), we have

$$y = \left(k_1 A_1 + \frac{9}{4} k_3 A_1^3\right) [\cos(\omega_1 t) + \cos(\omega_2 t)] + \frac{3}{4} k_3 A_1^3 [\cos((2\omega_2 - \omega_1)t) + \cos((2\omega_1 - \omega_2)t)] + \dots \quad (4.8)$$

The DUT's linearity can be evaluated by the normalized IM_3 [15], which is given by the ratio of the components at $2\omega_2 - \omega_1$ to the fundamental at ω_2 .

$$\text{IM}_3(\text{dBc}) = \frac{\frac{3}{4}k_3A_1^3}{k_1A_1 + \frac{9}{4}k_3A_1^3} \approx \frac{3}{4} \frac{k_3}{k_1} A_1^2 \quad (4.9)$$

when the low-distortion conditions, which include the mostly linear DUT criteria ($k_1 \gg k_2, k_3 \dots$) and the relatively low stimulus amplitude criteria ($k_1A_1 \gg k_3A_1^3$), are satisfied. We can find that

$$\text{IM}_3 = 3\text{HD}_3 \quad (4.10)$$

In this section, for short, we use IM_3 in dBc to present the normalized value. Moreover, the third order intercept point (IP_3) is defined at the point where $\text{IM}_3 = 0$ dBc, leading to

$$\text{IP}_3 = \frac{A_1}{\sqrt{\text{IM}_3}} = \sqrt{\frac{4}{3} \frac{k_1}{k_3}} \quad (4.11)$$

It should be noted that IP_3 does not rely on the input amplitude. Therefore, IP_3 is an absolute performance metric indicating the circuit linearity.

However, note that the analysis in [15] is based on ideal sinusoidal waveforms. In the proposed design, the harmonic cancellation is not perfect due to the mismatches and the PVT variations. The “baseband” single tone is still considered a weakly nonlinear waveform. Assuming an ideal up-conversion and ignoring the DC term, the two-tone output signals in Fig. 4.2a can be expressed as

$$u' = \cos(\omega_{LO}t) \cdot (2A_1 \cos(\omega_0t) + 2A_2 \cos(2\omega_0t) + \dots) \quad (4.12)$$

where $2A_1$ is used to have the same major tone power as that of (4.7). Harmonic coefficients A_i are annotated in Fig. 4.2a. Under the same low distortion conditions,

the actual measured IM3 can be approximated as the ratio of the IM3 components at $\omega_{LO} \pm 3\omega_0$ to the fundamental components at $\omega_{LO} \pm \omega_0$,

$$\begin{aligned} \text{IM}_3 &= \frac{\frac{3}{4}k_3A_1^3 + k_1A_3 + \frac{9}{2}k_3A_1^2A_3 + \dots}{k_1A_1 + \frac{9}{4}k_3A_1^3 + \dots} \approx \frac{3}{4} \frac{k_3}{k_1} A_1^2 + \varepsilon \\ \varepsilon &\approx \frac{A_3}{A_1} + \frac{9}{2} \frac{k_3}{k_1} A_1 A_3 + \frac{9}{4} \frac{k_3}{k_1} A_1 A_5 + \frac{9}{2} \frac{k_3}{k_1} \sum_{i=1}^{\infty} A_{2i+1} (A_{2i+3} + A_{2i+5}) \end{aligned} \quad (4.13)$$

where ε is the deviation from the conventional IM3. A_3/A_1 is the relative 3rd-order harmonic amplitude of the “baseband” signal at ω_0 . The derivation shows that all “baseband” harmonics will affect the IM3 measurement precision. However, the 3rd-order harmonic (A_3) has the most significant impact on the IM3 measurement accuracy if the proposed architecture is used.

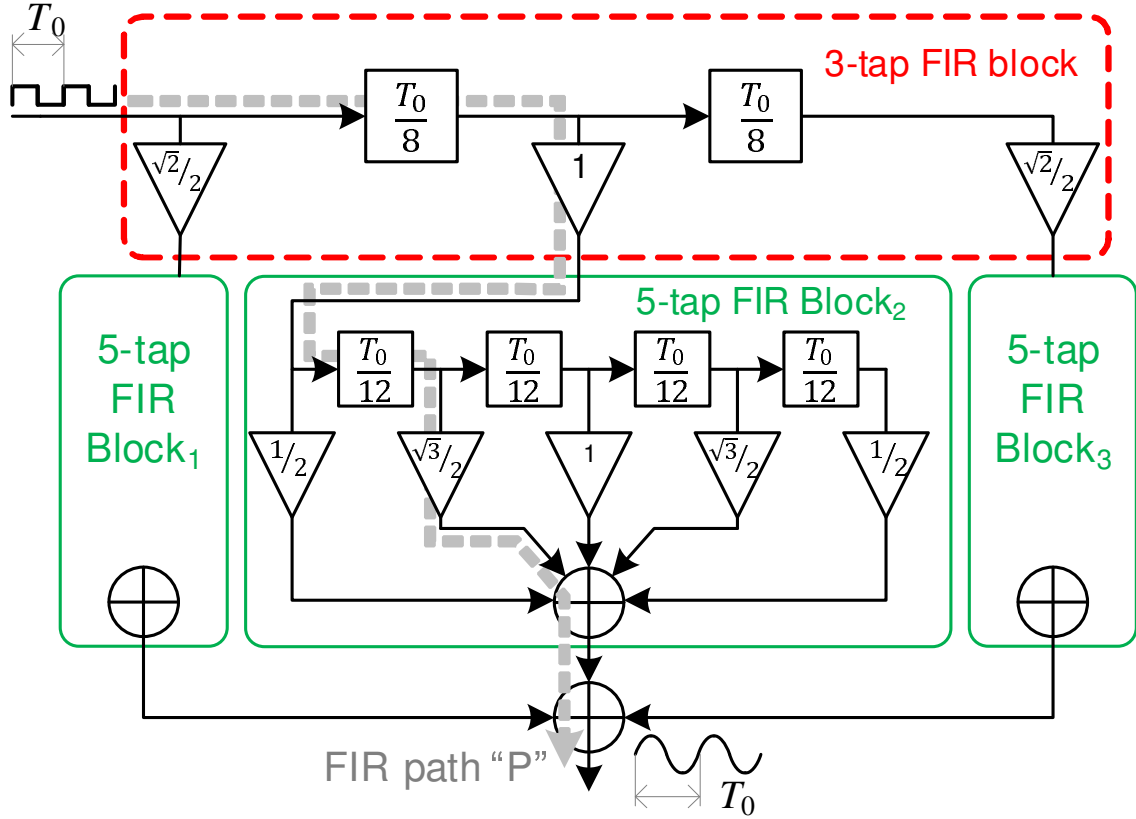
4.2.3 Cascade FIR-based Harmonic Cancellation

[44] has proposed a sinusoidal signal generator based on the finite impulse response (FIR) filter approach, whose implementation uses only multiple delayed square-wave clocks. The proposed M -tap FIR filter suppresses odd-order harmonics up to the $(2M + 1)$ -th order with designated tap coefficients,

$$c_i = \cos\left(\frac{(i+1)\pi}{M+1} - \frac{\pi}{2}\right), (i = 0, 1, \dots, M-1) \quad (4.14)$$

The 5-tap ($M = 5$) single-tone generation architecture implemented in [44] can achieve 55 dBc maximum spur free dynamic range (SFDR) without any tuning after fabrication.

In this section, we propose a two-stage cascade FIR filter architecture to further suppress the 3rd-order harmonic and push the residue odd-order harmonics to higher frequencies as shown in Fig. 4.3. The first stage is a 3-tap FIR block, and the



M	c_0	c_1	c_2	c_3	c_4
3	$\frac{\sqrt{2}}{2} \left(\frac{12}{17} \right)$	1	$\frac{\sqrt{2}}{2} \left(\frac{12}{17} \right)$		
5	$\frac{1}{2}$	$\frac{\sqrt{3}}{2} \left(\frac{13}{15} \right)$	1	$\frac{\sqrt{3}}{2} \left(\frac{13}{15} \right)$	$\frac{1}{2}$

Figure 4.3: Proposed two-stage cascade FIR harmonic cancellation and tap coefficients for the “baseband” single-tone generation.

second stage consists of three identical 5-tap blocks (annotated as Block₁, Block₂ and Block₃). T_0 is the period of the output quasi-sinusoidal waveform’s fundamental tone ($T_0 = 2\pi/\omega_0$), M is the number of FIR taps, and c_0, c_1, \dots are tap coefficients obtained from (4.14). For irrational c_i , we can approximate $\sqrt{3}/2 \approx 13/15$ and $\sqrt{2}/2 \approx 12/17$. The quantization errors are only +0.07% and -0.17%, respectively. Moreover, an FIR path is defined from the input to the output, which passes through multiple delays and through two exact tap coefficients as path “P” shown in Fig. 4.3. Section 4.3.1

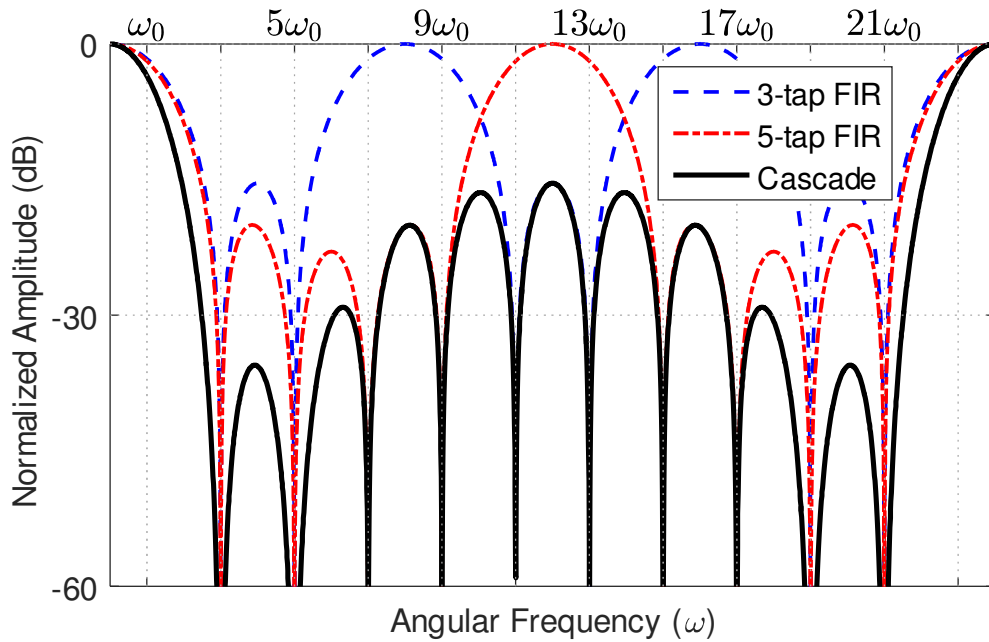


Figure 4.4: Frequency response of the cascade FIR architecture.

will show how to rearrange the FIR path for practical hardware implementation. Fig. 4.4 shows the frequency response of the proposed architecture. The first non-cancellable harmonic is pushed to the 23rd order, and all lower odd-order harmonics are eliminated. To achieve the same harmonic cancellation effect, Fig. 4.5 shows a single stage approach proposed in [44], which would require $M = 11$ (11 taps); thus, a high precision circuit becomes necessary to achieve a series of fractional coefficients obtained from (4.14) (0.259, 0.5, 0.707, 0.866, 0.966, and 1). Instead, the proposed cascade architecture separates the coefficients into two groups and each group has only two unique factors. The 3-tap FIR stage adopts coefficients 1 and $12/17$, while $1/2$ and $13/15$ can cover the 5-tap FIR stage (1 is not a unique coefficient as it can be represented by $2 \times 1/2$). This coefficient selection significantly reduces the hardware implementation complexity. In addition, the proposed two-

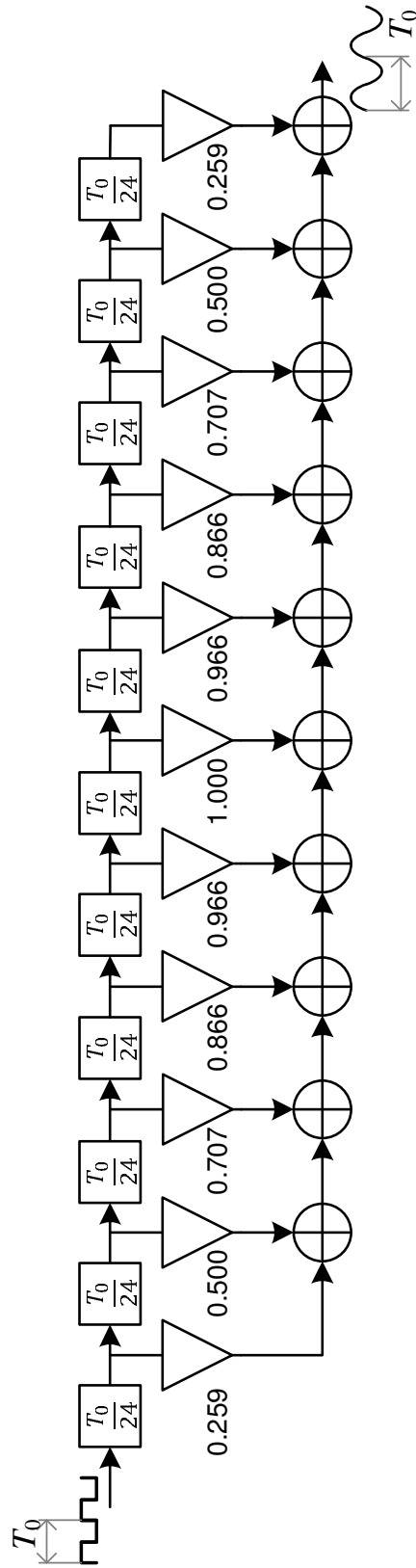


Figure 4.5: The architecture of the single-stage 11-tap FIR approach.

stage FIR architecture adopts one 3-tap FIR followed by three 5-tap FIR blocks, which are demonstrated in Fig. 4.6a. The input signals of all FIR blocks are shifted square waveforms. Alternatively, Fig. 4.6b shows another two-stage FIR scheme. Although this alternative scheme implements one 3-tap FIR plus only one 5-tap FIR block, the input of the 5-tap FIR becomes a quasi-sinusoidal analog signal. The FIR filter that accepts and processes analog signal is more complicated and more sensitive to PVT variations (i.e. Section 5 will demonstrate a switched capacitor FIR filter). Therefore, the simpler (from the aspect of hardware implementation) solution (Fig. 4.6a) is chosen in this design. Detailed implementation, which uses two different current biases and two different transistor sizes, is discussed later in Section 4.3.2.

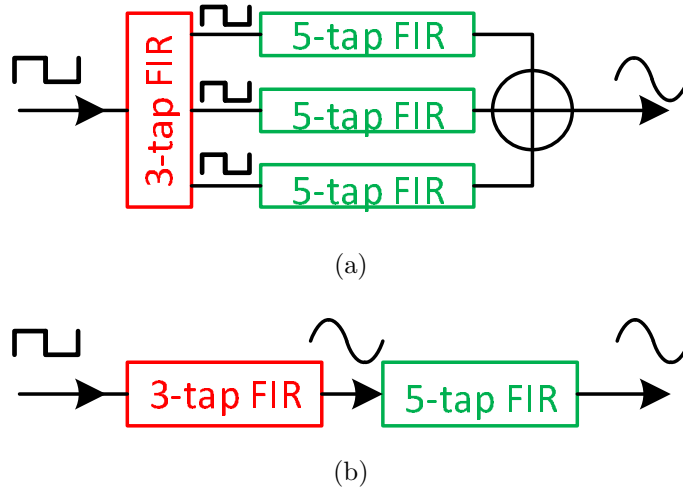


Figure 4.6: Comparison between two-stage FIR approaches: (a) Proposed two-stage FIR with three 5-tap FIRs and (b) a single 3-tap FIR followed by a single 5-tap FIR.

Moreover, the notching points of the FIR frequency responses will become finite due to the PVT variations as discussed in [44]. The cascade architecture can help

emphasize the cancellation of 3rd-order harmonic by stacking a 5-tap FIR over a 3-tap FIR, which both have a notching point at $3\omega_0$. Additionally, assuming the FIR architecture has an ideal response, as shown in Fig. 4.4, and its input is an ideal square wave, we can derive $A_1 = 1$, $A_{23} = 1/23$, $A_{25} = 1/25$, $A_{47} = 1/47$, \dots . The spur-free dynamic range (SFDR) shown in Fig. 4.2b is not better than 27.2 dB ($1/23$). However, applying all these components to (4.13), the measurement error (ε) is around 1.6% compared to the theoretical measured value in (4.9). It is negligible. Therefore, the design of this cascade FIR architecture should focus on improving the 3rd-order cancellation. Details will be discussed in Section 4.4.1.

4.3 Circuit Implementation

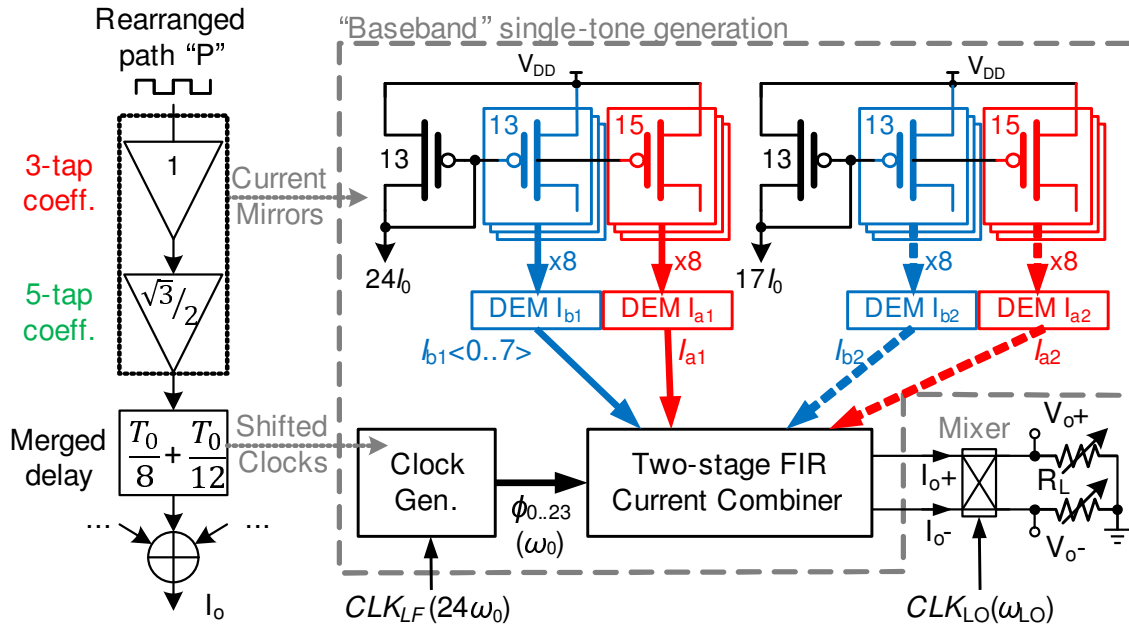


Figure 4.7: System architecture of the proposed two-tone synthesizer and the corresponding rearranged FIR path

4.3.1 System Architecture

To implement an FIR transfer function using an analog circuit, two major types of architecture can be considered—the switched capacitor (SC) filter or the current-steering FIR reconstruction filter. On the one hand, the current-steering architecture is able to drive the resistive loads, which has more flexibility than that of the passive SC filters. On the other hand, the current-steering architecture consists of mostly current mirrors and MOS switches. It is less complicated but more friendly to digital circuits and advanced technologies, compared to the active SC filters. Therefore, the current-steering architecture is more suitable for the proposed design, aiming at the BIST applications. In addition, Section 5 will give a more detailed discussion on the SC filters.

The conceptual two-tone generation architecture of Fig. 4.2a can be implemented by the system demonstrated in Fig. 4.7 after rearranging the FIR paths. To illustrate, path “P” in Fig. 4.3 and Fig. 4.7 show the proposed rearrangement, and Fig. 4.8 shows the complete architecture in Fig. 4.3 after the rearrangement. On the one hand, tap coefficients are put together and implemented by a current mirror array. In detail, the 3-tap coefficients are achieved by two different bias currents, while the current mirror ratios are used to produce the 5-tap coefficients. On the other hand, the FIR delays across the whole path are merged together. The total delay of path “P” is $\frac{T_0}{8} + \frac{T_0}{12} = \frac{5}{24}T_0$, which is implemented via a MOS switch driven by the clock ϕ_5 with $\frac{5}{24}T_0$ delay in the current combiner. Generally, 50% duty-cycled ϕ_k ($k = 0 \dots 23$) has a delay of $\frac{k}{24}T_0$. They are used to represent any FIR path delay in Fig. 4.3. Following the “baseband” single-tone generator, a differential quasi-sinusoidal current waveform I_O is then up-converted to the desired ω_{LO} band by an up-conversion mixer and outputs the voltage waveform V_O across R_L .

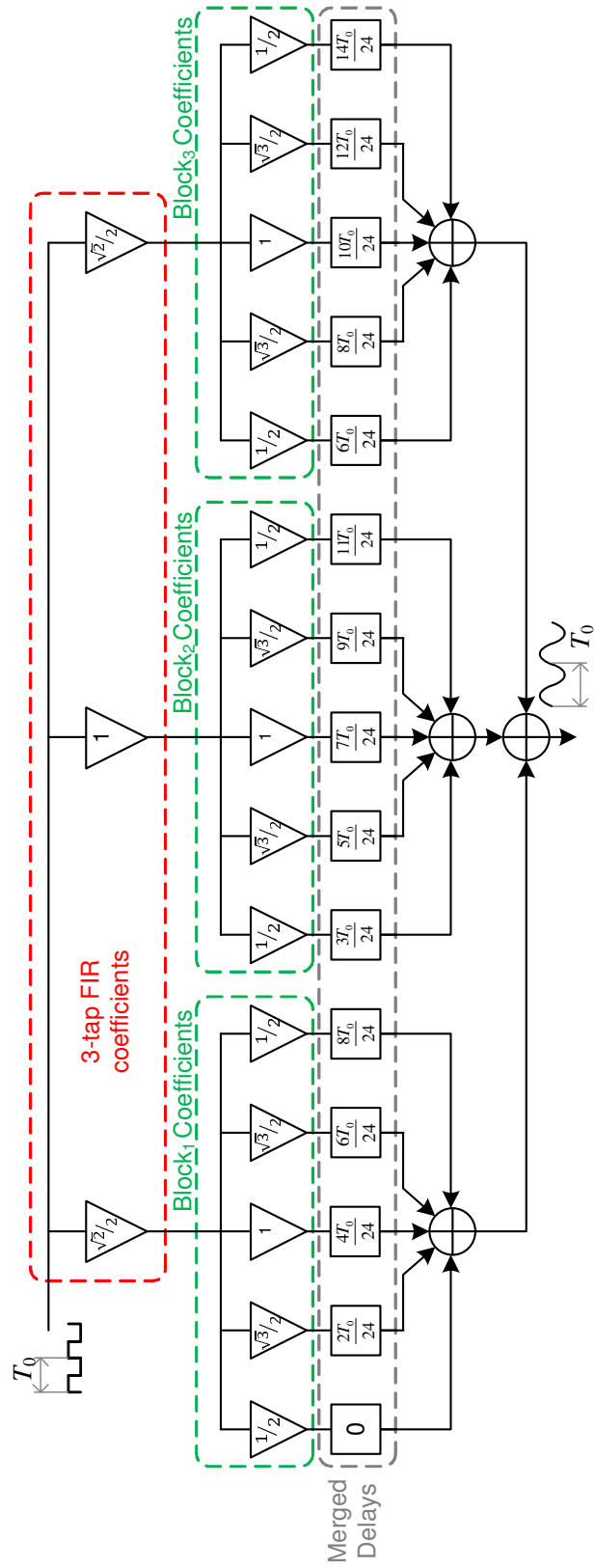


Figure 4.8: The proposed two-stage FIR architecture after the path rearrangement.

4.3.2 Current Mirror Array for FIR Tap Coefficients

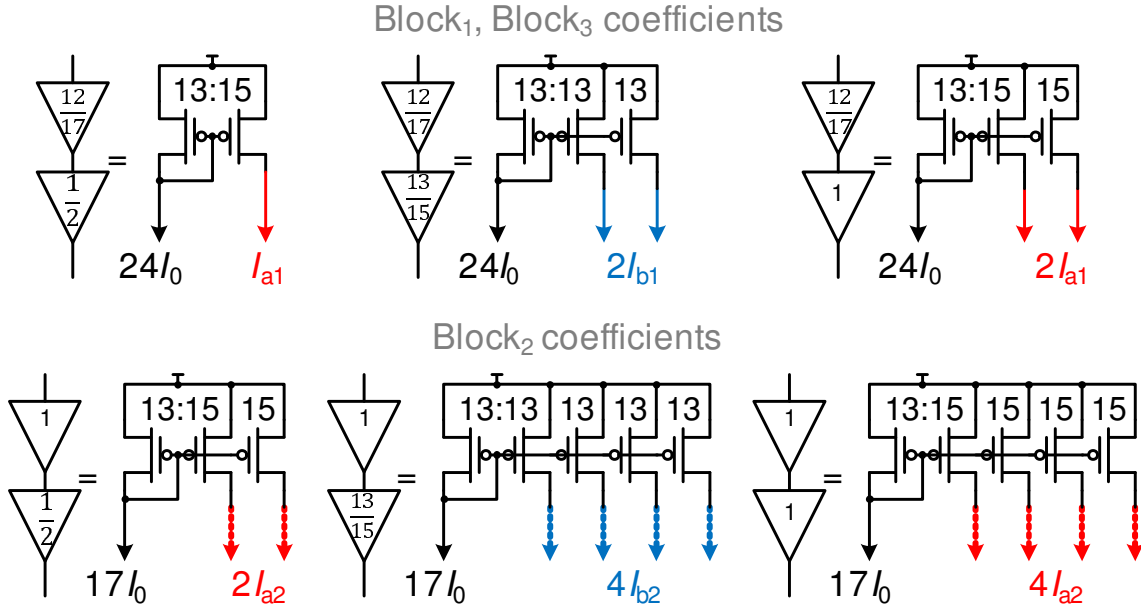


Figure 4.9: Current mirror implementation of the two-stage FIR coefficients.

The relationship between the FIR tap coefficients of Fig. 4.3 and the current branches of Fig. 4.7 is further explained in Fig. 4.9. Considering the 5-tap FIR coefficients $1/2 : 13/15 : 1 = 15 : 13 \times 2 : 15 \times 2$, one I_{a1} branch is used for the tap coefficient $1/2$, two I_{a1} branches for 1, and two I_{b1} branches for $13/15$. As a result, eight I_{a1} and eight I_{b1} can cover all the tap coefficients of Block₁ and Block₃ in Fig. 4.3. To implement Block₂, two I_{a2} branches are used for representing $1/2$, four I_{a2} for 1, and four I_{b2} for $13/15$. Note that the number of current branches used are doubled so as to reuse the 16-branch current mirror design and the dynamic element matching (DEM) blocks. Furthermore, we have 3-tap coefficients $12/17 : 1 = 24 : 17 \times 2$, this leads to the bias currents, $24I_0$ and $17I_0$, in Fig. 4.7, where I_0 is a unit current (i.e. we

have $I_{a1} = \frac{15}{13} \times 24I_0$). These current ratios work together with the clock connection pattern, which will be introduced below to achieve the function of the proposed cascade FIR architecture. Additionally, as shown in Fig. 4.7, four DEM current branch rotators shuffle each of the eight current channels and output $I_{a1}<0..7>$, $I_{b1}<0..7>$, $I_{a2}<0..7>$ and $I_{b2}<0..7>$ to the following current combiner, where the annotation $<i..j>$ is used to index the channel number from i to j .

4.3.3 Clock Divider and Current Combiner for FIR Tap Delays

The clock divider is implemented by a 24-bit cyclic shifted register. Its hardware implementation is discussed in Section 4.3.6.4. It is driven by CLK_{LF} at a frequency of $24\omega_0$, and cyclically shifts twelve 1s followed by twelve 0s. The output ports of 24 registers are used as $\phi_{0..23}$. Noting that $T_0 = 2\pi/\omega_0$, each ϕ_i and ϕ_{i+1} pair has a delay of $T_0/24$ (15° phase shift) between them.

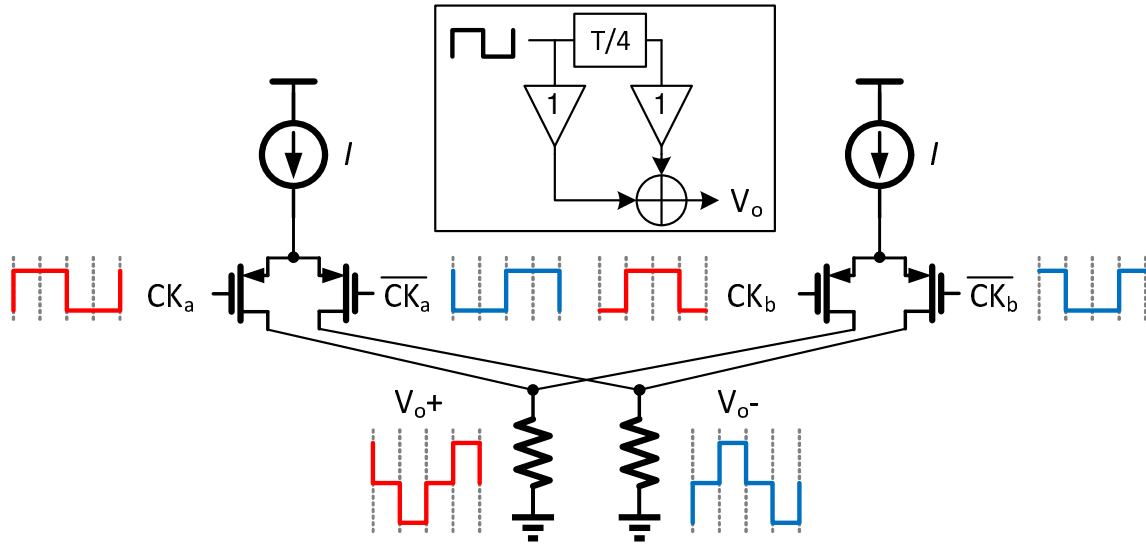


Figure 4.10: Current steering implementation with shifted clocks and the equivalent flow diagram. CK_b is CK_a delayed by $T/4$, where T is the clock period.

Different from the conceptual architecture in Fig. 4.3, which accepts the square waveform as the input of the FIR filter, the proposed current steering FIR implementation applies the shifted square clocks to the switching transistors just before the output node. Fig. 4.10 illustrates an example showing the equivalence between the current steering approach and the FIR flow diagram.

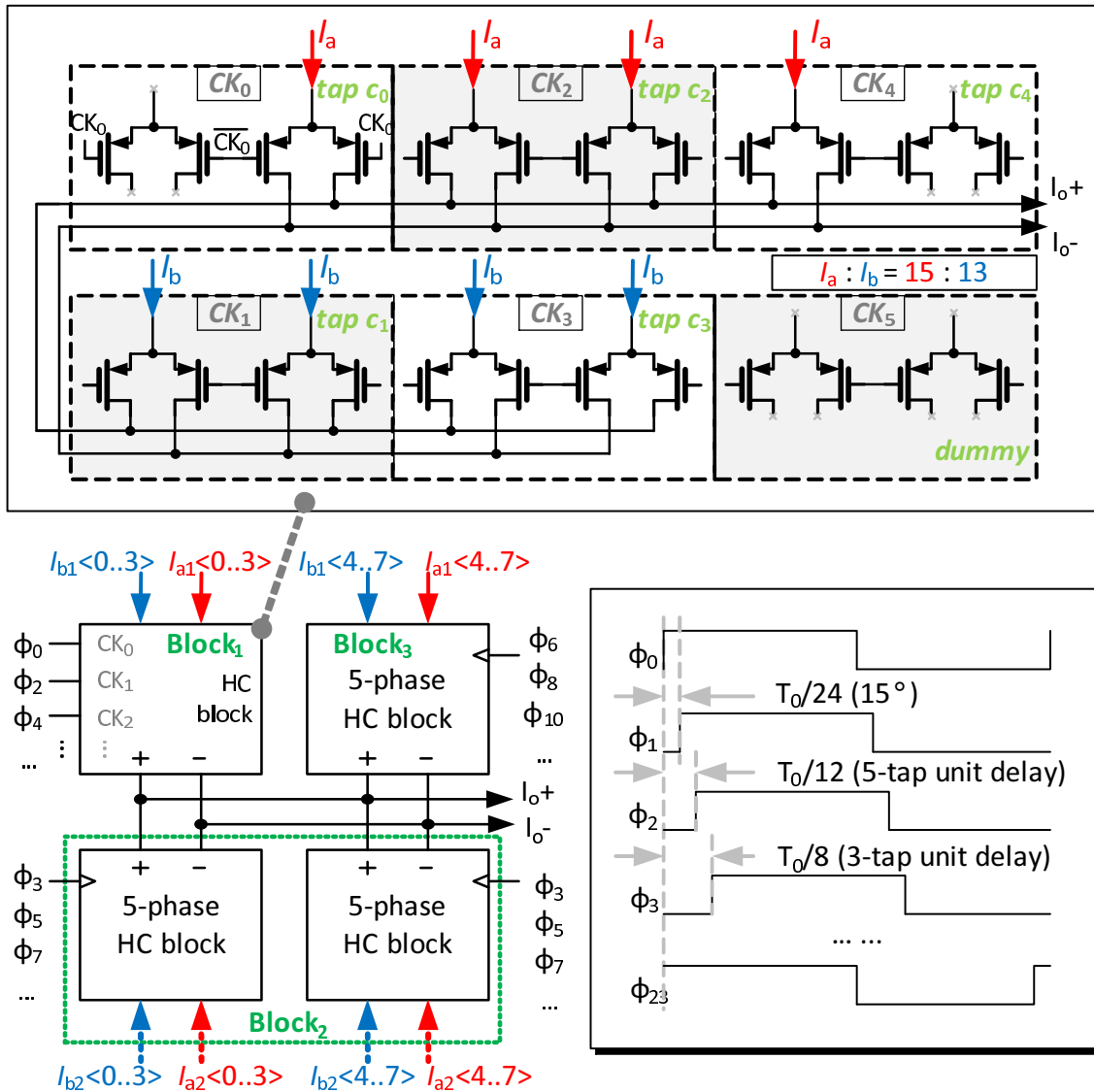


Figure 4.11: Current combiner topology and clock connections of the proposed two-stage FIR architecture.

In this design, the two-stage cascade FIR current combiner is implemented as an extra layer of PMOS switches, as demonstrated in Fig. 4.11. It consists of four 5-phase harmonic cancellation (HC) blocks. Each 5-phase HC switching block is switched by six clocks, $CK_{0..5}$, and the corresponding inverted phases, $\overline{CK_{0..5}}$. Fig. 4.11 shows that, CK_0 and CK_4 switch one I_a channel respectively. CK_2 switches two I_a channels. CK_1 and CK_3 switch two I_b channels separately. If $I_a : I_b = 15 : 13$ and the delay between CK_i and CK_{i+1} is $T_0/12$, this 5-phase HC block realizes a single 5-tap FIR block shown in Fig. 4.3. In order to balance the load of each clock signal, dummy switch pairs are added for CK_0 , CK_4 and CK_5 . Fig. 4.11 also illustrates the full current-combining topology. The 5-phase HC block, which is equivalent to Block₁ in Fig. 4.3, accepts $I_{a1} <0..3>$ and $I_{b1} <0..3>$ as equivalent tap coefficients. It also uses $\phi_0, \phi_2, \phi_4, \phi_6, \phi_8$, and ϕ_{10} for $CK_{0..5}$. Additionally, $\phi_{12}, \phi_{14}, \phi_{16}, \phi_{18}, \phi_{20}$ and ϕ_{22} are used as the inverted $\overline{CK_{0..5}}$. Similarly, Block₃ equivalent block are driven by $\phi_6, \phi_8, \dots, \phi_{22}, \phi_0, \phi_2$, and ϕ_4 . It takes the remaining $I_{a1} <4..7>$ and $I_{b1} <4..7>$. Block₂ is achieved by two identical 5-phase HC blocks, whose $CK_{0..5}$ and $\overline{CK_{0..5}}$ are $\phi_3, \phi_5, \dots, \phi_{23}$, and ϕ_1 . It takes all $I_{a2} <0..7>$ and $I_{b2} <0..7>$. This special pattern of clock connections guarantee a balanced load for each phase of $\phi_{0..23}$. The delay between ϕ_i and ϕ_{i+2} is $T_0/12$, and the delay between ϕ_i and ϕ_{i+3} is $T_0/8$, which are the unit delays in the 5-tap and 3-tap FIR blocks of Fig. 4.3, respectively. Finally, all the current branches are summed together to output the differential quasi-sinusoidal current waveform I_O . Table 4.1 summarizes all clock and current connection patterns for the proposed cascade FIR architecture, where the current flowing into I_{O+} is indicated by “+”, and “-” marks the current flowing into I_{O-} . Because the PMOS switches are adopted in this design, $\overline{\phi}$ is used to indicate the inverted clock phases.

Table 4.1: 24-phase Clock and Current Branch Connection Pattern (\emptyset represents the dummy connection)

$\bar{\phi}$	ϕ_0	ϕ_1	ϕ_2	ϕ_3	ϕ_4	ϕ_5
I_O	$+I_{a1}<0>$	\emptyset	$+I_{b1}<0>$	$+I_{a2}<0>$	$+I_{a1}<1>$	$+I_{b2}<0>$
	\emptyset	\emptyset	$+I_{b1}<1>$	\emptyset	$+I_{a1}<2>$	$+I_{b2}<1>$
	$-I_{b1}<6>$	\emptyset	$-I_{a1}<7>$	$+I_{a2}<4>$	\emptyset	$+I_{b2}<4>$
	$-I_{b1}<7>$	\emptyset	\emptyset	\emptyset	\emptyset	$+I_{b2}<5>$
$\bar{\phi}$	ϕ_6	ϕ_7	ϕ_8	ϕ_9	ϕ_{10}	ϕ_{11}
I_O	$+I_{b1}<2>$	$+I_{a2}<1>$	$+I_{a1}<3>$	$+I_{b2}<2>$	\emptyset	$+I_{a2}<3>$
	$+I_{b1}<3>$	$+I_{a2}<2>$	\emptyset	$+I_{b2}<3>$	\emptyset	\emptyset
	$+I_{a1}<4>$	$+I_{a2}<5>$	$+I_{b1}<4>$	$+I_{b2}<6>$	$+I_{a1}<5>$	$+I_{a2}<7>$
	\emptyset	$+I_{a2}<6>$	$+I_{b1}<5>$	$+I_{b2}<7>$	$+I_{a1}<6>$	\emptyset
$\bar{\phi}$	ϕ_{12}	ϕ_{13}	ϕ_{14}	ϕ_{15}	ϕ_{16}	ϕ_{17}
I_O	$-I_{a1}<0>$	\emptyset	$-I_{b1}<0>$	$-I_{a2}<0>$	$-I_{a1}<1>$	$-I_{b2}<0>$
	\emptyset	\emptyset	$-I_{b1}<1>$	\emptyset	$-I_{a1}<2>$	$-I_{b2}<1>$
	$+I_{b1}<6>$	\emptyset	$+I_{a1}<7>$	$-I_{a2}<4>$	\emptyset	$-I_{b2}<4>$
	$+I_{b1}<7>$	\emptyset	\emptyset	\emptyset	\emptyset	$-I_{b2}<5>$
$\bar{\phi}$	ϕ_{18}	ϕ_{19}	ϕ_{20}	ϕ_{21}	ϕ_{22}	ϕ_{23}
I_O	$-I_{b1}<2>$	$-I_{a2}<1>$	$-I_{a1}<3>$	$-I_{b2}<2>$	\emptyset	$-I_{a2}<3>$
	$-I_{b1}<3>$	$-I_{a2}<2>$	\emptyset	$-I_{b2}<3>$	\emptyset	\emptyset
	$-I_{a1}<4>$	$-I_{a2}<5>$	$-I_{b1}<4>$	$-I_{b2}<6>$	$-I_{a1}<5>$	$-I_{a2}<7>$
	\emptyset	$-I_{a2}<6>$	$-I_{b1}<5>$	$-I_{b2}<7>$	$-I_{a1}<6>$	\emptyset

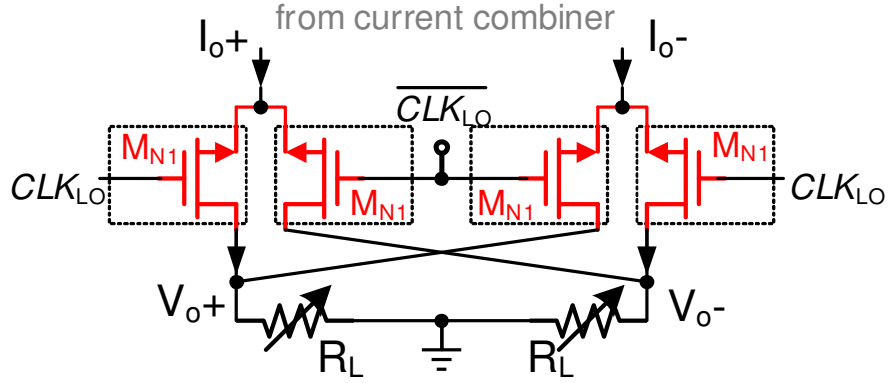


Figure 4.12: Schematic of a simple passive up-conversion mixer.

4.3.4 Up-Conversion Mixer

Fig. 4.7 adopts a up-conversion mixer to move the generated “baseband” single tone up to the desired high frequency band centered at ω_{LO} and converts the current waveform I_O to the output voltage waveform V_O . A simple MOS switch based mixer [94] (Fig. 4.12) can be implemented for this purpose. However, in this design, an improved architecture is used. In order to reduce the switches’ ON resistance, to extend the output voltage swing, and to minimize the impact on linearity, bootstrapped NMOS switches [95] are adopted as demonstrated in Fig. 4.13. The switching NMOS transistors, M_{N1} and M_{N2} , are implemented using twin-well devices and their source terminals are connected to the bulk. Two R_L resistor arrays are also implemented on-chip as the loads of the mixer. They are digitally controlled and can be manually adjusted (e.g. set to 200 Ohm in this design) for impedance matching and test purpose.

Fig. 4.14 shows a comparison between the simulated spectrum obtained at the output of a mixer without (Fig. 4.12) and with (Fig. 4.13) the bootstrapped switches. The load impedance is set to 400 Ohm ($R_L = 200 \Omega$). The mixer’s input current is

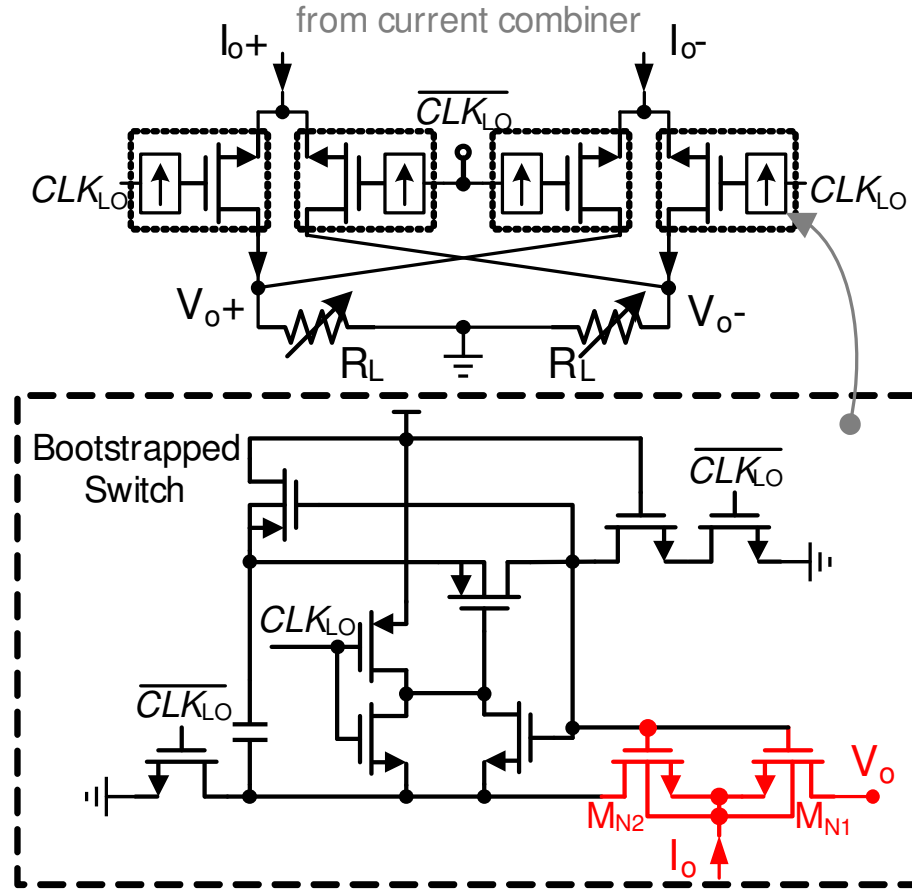


Figure 4.13: Passive up-conversion mixer with bootstrapped MOS switches.

set to induce a 400 mV peak-to-peak sine wave across the load. $f_0 = 1$ MHz and $f_{LO} = 400$ MHz are also set for the simulation. 22 dB reduction on the IM3 tone can be observed. To conclude, the bootstrapped switches significantly improve the linearity of the mixer.

The major drawback of the proposed mixer design is the LO leakage. Especially, the high gate switching voltage deteriorates the isolation between I_O and CLK_{LO} . However, analysis in Section 4.4 will show that the leakage tone still has limited impact on the two-tone test accuracy. In addition, several techniques can help to further extend the proposed mixer's working frequency range to cover more RF

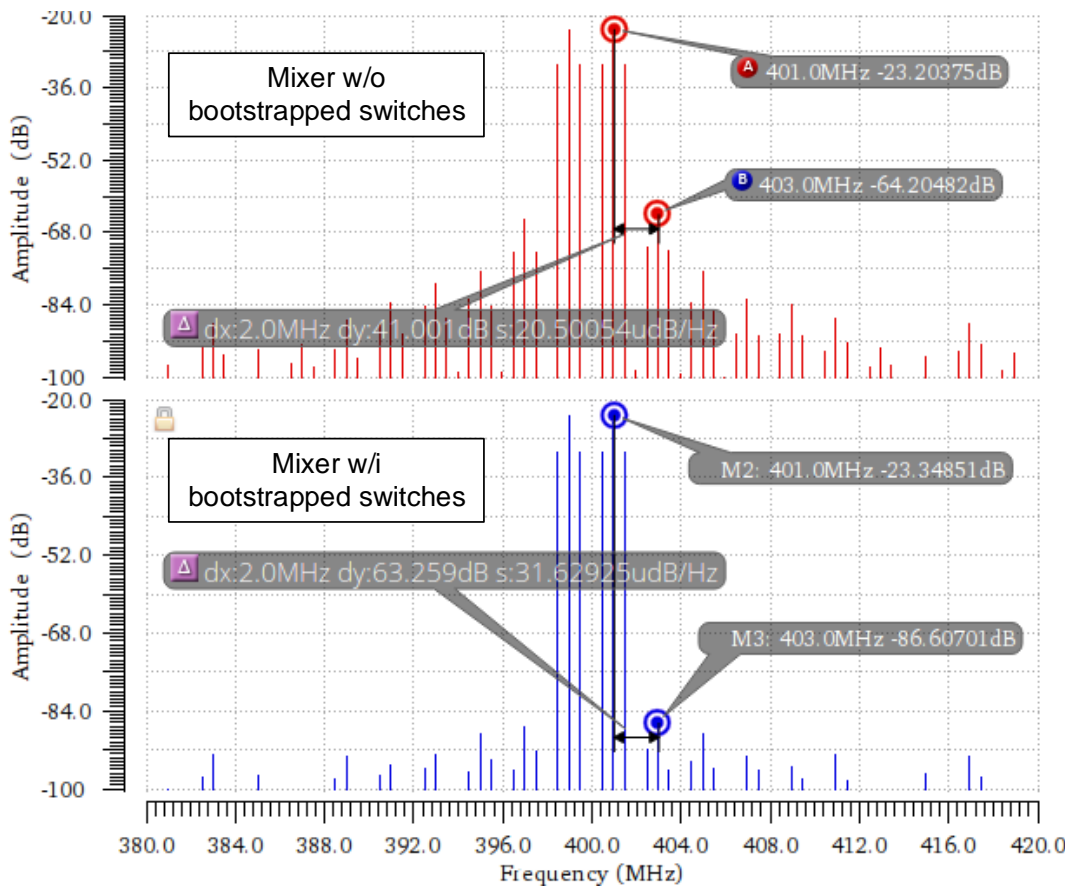


Figure 4.14: IM3 tone comparison between the mixers with and without the bootstrapped switches.

applications at the expense of larger area and power consumption. The local mixer array architecture suggested in [96] is able to improve the mixer’s linearity at high LO frequencies and it is compatible to the proposed design. I/Q modulation and LO cancellation techniques [97] have been well researched for communication systems, and they can be adopted to reduce the LO leakage.

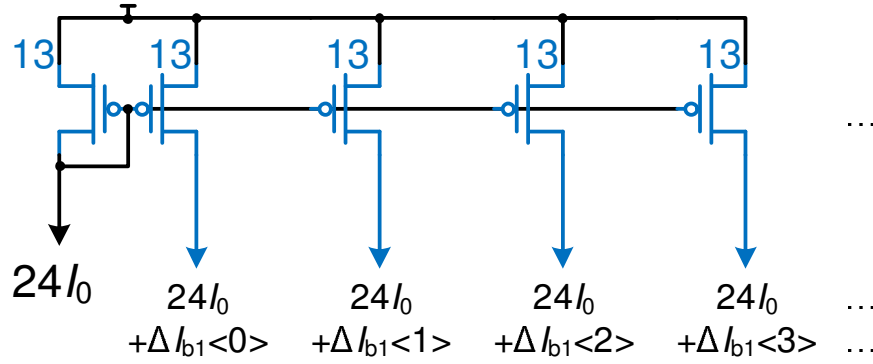


Figure 4.15: Mismatches existed in the current mirrors.

4.3.5 3-bit DEM Rotator

PVT variations will cause mismatches in a current mirror [98]. The variance of the current ratio is inversely proportional to the transistor area. Fig. 4.15 illustrates the mismatches existing in the current mirrors of the proposed design. Although the current mirrors are designed with 1 : 1(13 : 13) or 13 : 15 current factors, the mismatch induced error currents, $\Delta I_{\{a1,a2,b1,b2\}} <0..7>$, are randomly distributed among all current branches, even the transistors are physically adjacent to each other in the layout. As proposed in [98], larger transistor sizes can be adopted to reduce the variance.

Apart from increasing the transistor sizes, the dynamic element matching (DEM) technique can also be used to alleviate the mismatch issue. The working principle of the DEM technique has been introduced in [99]. In this design, each branch current of I_a or I_b is dynamically and arbitrarily picked from one of the corresponding eight mismatched current mirrors' outputs, leading to an average current level plus a white noise floor. It should be mentioned that [99] introduced two types of DEM—the mismatch scrambling DEM and the mismatch shaping DEM. The mismatch shaping DEM has a low noise level at low frequency range but a high noise level close to

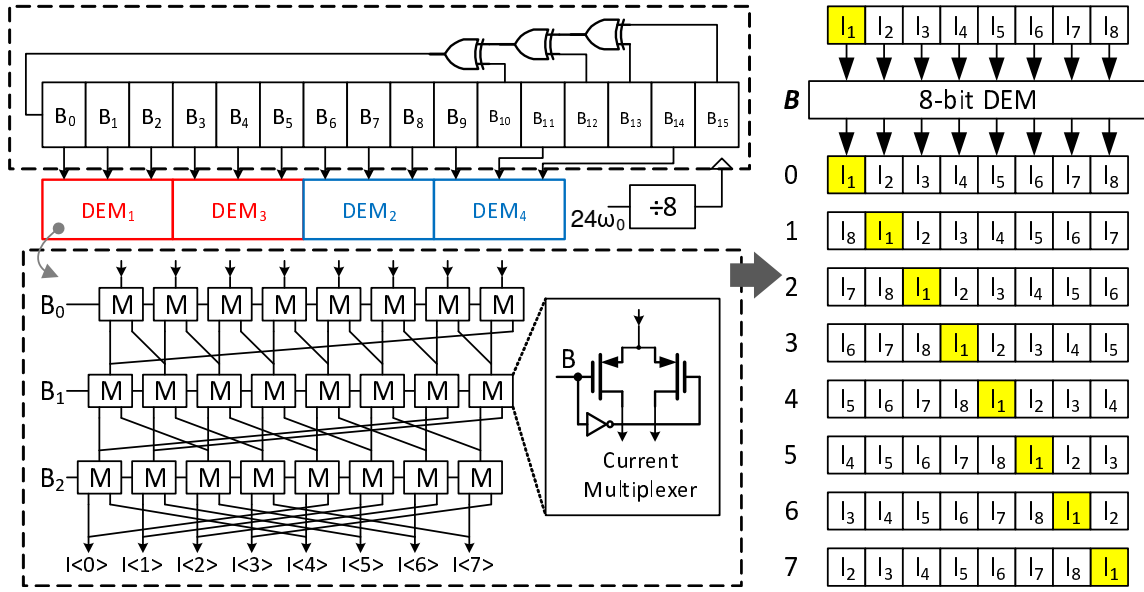


Figure 4.16: Current branch rotator with dynamic element matching and the corresponding current branch rotation patterns.

the switching frequency. However, a flat noise floor is usually acceptable for a signal generator. Therefore, the mismatch scrambling DEM is adopted in this design to generate a white noise floor by using a simple pseudo-random number generator. Additionally, [100] has first applied the DEM to improve the linearity of sine-wave synthesis, but it adopts a partial DEM configuration, which is more complicated than the proposed architecture.

Fig. 4.16 illustrates the structure of the DEM current branch rotator and the corresponding rotation patterns. Each current branch rotator is controlled by three control bits to shift the 8-channel current branches cyclically. A 16-bit linear feedback shift register (LFSR) generates the control bits, and each DEM rotator receives a different group of control bits, so as to avoid identical patterns. The chosen feedback taps, B_{10} , B_{12} , B_{13} and B_{15} , guarantee that the LFSR can cycle through the maximum number of 65535 states except for the all-zero state. All DEM current branch

rotators are driven by a clock of CLK_{LF} divided by 8, which is at a frequency of $3\omega_0$.

4.3.6 Design Procedure

4.3.6.1 System level

As the synthesizer is an excitation signal generator, the output swing is the major concern. The system level design procedure is

1. Set the target output swing to -4 dBm (approximately 400 mVpp).
2. Set 400Ω load resistor. Thus, $R_L = 200 \Omega$.
3. Consider an average mismatch deviation (σ) around 0.36%. Look up the corresponding transistor area ($W \cdot L$) in the design document and find that it is $8 \mu\text{m}^2$.
4. W/L of an unit current mirror PMOS starts from $8 \mu\text{m}/1 \mu\text{m}$. Build the current mirror array shown in Fig. 4.7. Using 13 units and 15 units separately to achieve the ratio of 13:15.
5. Create a PMOS switch pair with the minimum W/L.
6. Construct the current combiner shown in Fig. 4.11 and the current rotator (without the random number generator) shown in Fig. 4.16 by reusing the switch block created in step 5. Fix the input B=000 for all four rotators.
7. Connect the differential output of the “baseband” generator obtained from step 6 to R_L . Set $I_0 = 2.5 \mu\text{A}$ (a coarse approximation of producing 200 mV over $200 \Omega R_L$ via 16 branches).
8. Simulate the quasi-sinusoidal waveform generated by the “baseband” synthesizer and evaluate its HD_3 .
9. Increase the width of the PMOS in the switch block (step 5) and repeat the simulation in step 8 until the simulated linearity of the output signal meets the

target, $HD_3 < -80 \text{ dBc}$ at $f_0 = 1 \text{ MHz}$, for the “baseband” generator.

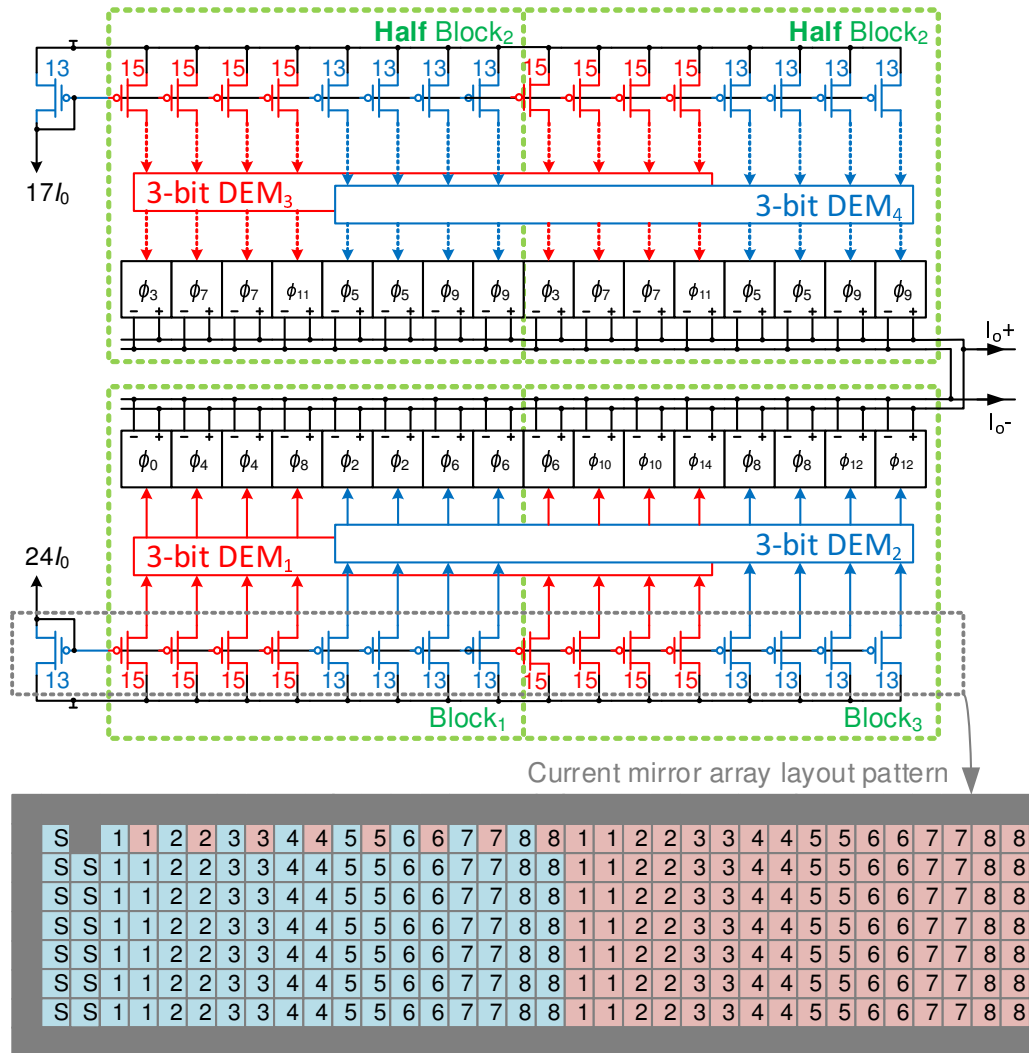
10. Size I_0 to obtain the target swing and repeat step 8 until the simulated HD_3 does not change significantly.
11. Insert the mixer design shown in Fig. 4.13. The bootstrapped switch design is from [50].
12. Add the LFSR random number generator shown in Fig. 4.16 and connect it to the four current branch rotators.
13. Simulate and evaluate the linearity of the whole system, sweeping the LO frequency from DC to 1 GHz. Iteratively increase the length of the current mirror PMOS, the width of switch PMOS and the width of the mixer switch (M_{N1} and M_{N2} in Fig. 4.13) if the linearity criteria cannot be satisfied.

Moreover, the LFSR selection for the DEM should be emphasized. There are four 3-bit current rotators in the proposed design. In order to avoid correlation between any two control bits and maximize the randomization effort, at least a 12-bit LFSR should be adopted for the DEM. Therefore, each control bit can be associated to only one position in the LFSR. It should be noted that the LFSR design pattern is fixed [101].

Additionally, some detailed schematic or layout design issues are discussed below.

4.3.6.2 *Current mirror array*

A detailed description of the current mirror array block can be found in Fig. 4.17 (the 3-bit DEM current branch rotator will be explained in Section 4.3.6.3, and the ϕ_i blocks are the expanded current combiner topology (Fig. 4.11), where ϕ_i is the clock assigned to the corresponding switch pair). The current mirror ratio 13 : 15 is achieved by different number of the unit transistors. A unit PMOS transistor has a size of $4 \mu\text{m}/2 \mu\text{m}$. PMOS with this size has an average mismatch deviation (σ)



- 1 PMOS in x13 current branch (I_b) group and branch #
- 1 PMOS in x15 current branch (I_a) group and branch #
- S Current mirror source group Dummy PMOS

Figure 4.17: Full display of the current mirror array arrangement and the layout patterns.

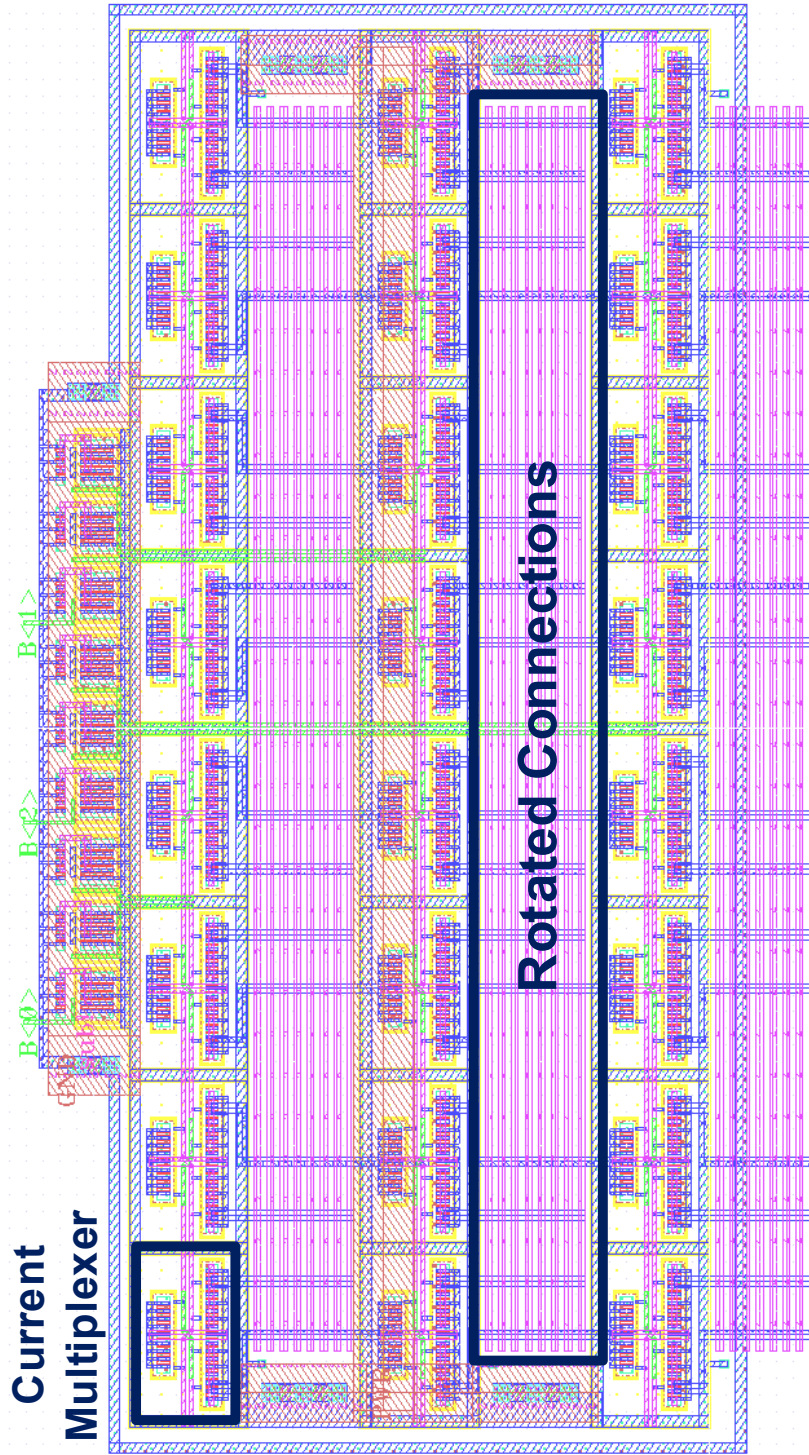


Figure 4.18: Layout of the current branch rotator.

around 0.36% according to the design document. Later in Section 4.4.1, we will analyze that this size value is conservative. Moreover, because $13 + 15 = 28 = 4 \times 7$, unit transistors of I_a and I_b branches are interleaved as demonstrated in Fig. 4.17. Total 237 units (13 source units + 28 x 8 branch units) construct the array with $24I_0$ source for $Block_1$ and $Block_3$, and this array is also duplicated for $Block_2$ with $17I_0$ source.

4.3.6.3 Current branch rotator

The layout of the current branch rotator is shown in Fig. 4.18. The 3x8 current multiplexers, which are also demonstrated in Fig. 4.16, consists of PMOS switches with a size of $6 \mu\text{m}/0.12 \mu\text{m}$. The rotated connections follow the pattern shown in Fig. 4.16.

4.3.6.4 24-phase clock generator

As described in Section 4.3.3, each ϕ_i and ϕ_{i+1} pair has a fixed delay of $T_0/24$ (15° phase shift) between them, as shown in the timing diagram (Fig. 4.19). If each clock is treated as a 0/1 series, we can find a cyclic pattern, which is annotated in Fig. 4.19. This pattern is achievable by using a cyclic shift register.

The proposed 24-phase clock generator is illustrated in Fig. 4.20. The main part of the generator is a 24-bit cyclic shift register, consisting of 12 settable D flip-flops followed by 12 resettable D flip-flops. The active-low $\overline{\text{RST}}$ signal will set/reset the whole register to 12 ones followed by 12 zeros, and this pattern will be shifted cyclically at the falling edge of the CLK signal, which has a frequency of $24f_0$. The outputs (Q) of 24 bits naturally form the 24-phase clocks, $\phi_0 \cdots \phi_{23}$. Particularly, retiming technique [102] is applied to reduce the phase error across different clock phases. D flip-flop retimers are implemented to sample each bit in the shift register at the rising edge of CLK.

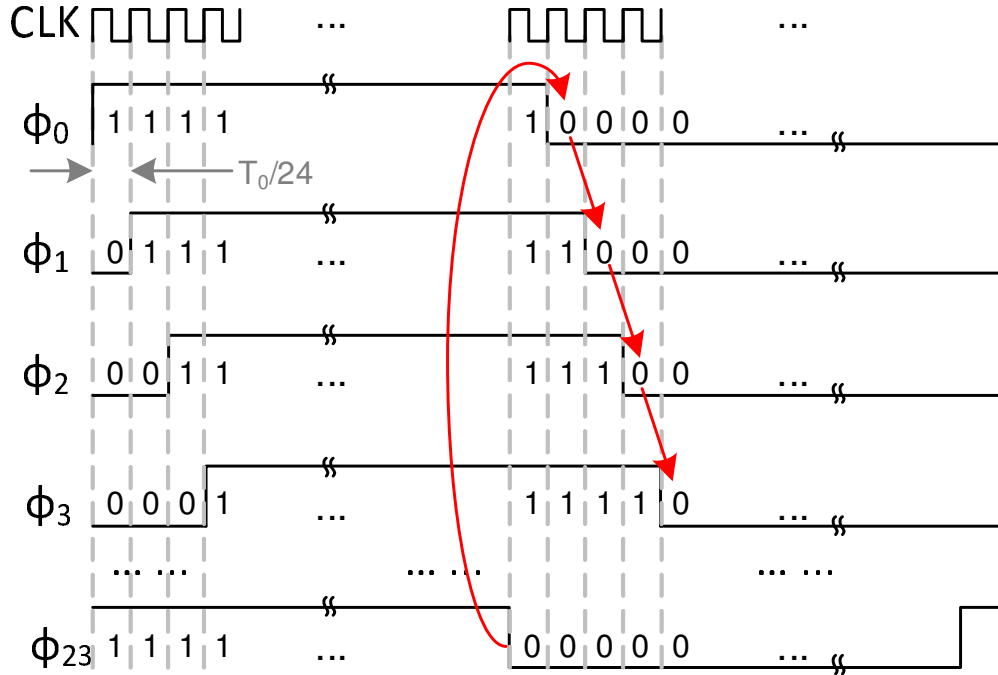


Figure 4.19: The timing of the 24-phase clock output $\phi_{0...23}$.

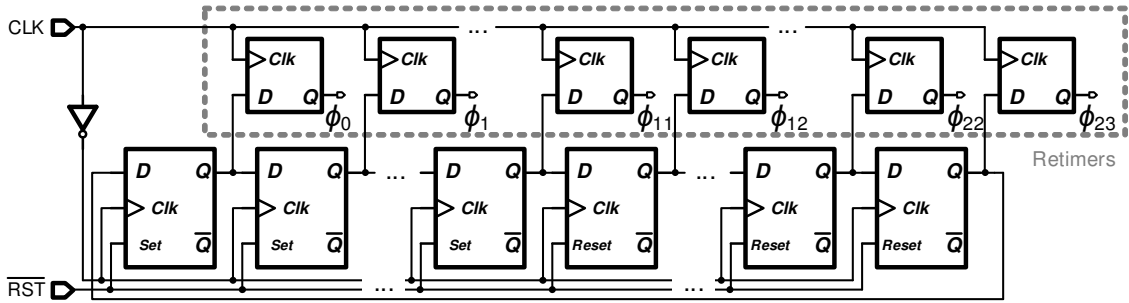


Figure 4.20: Schematic of the 24-phase clock generator.

4.4 Non-ideality Analysis

4.4.1 Baseband Current Mismatch and Phase Error

For the “baseband” single tone signal generation, there are two major error sources—the amplitude error induced by current mirror mismatches and the clock

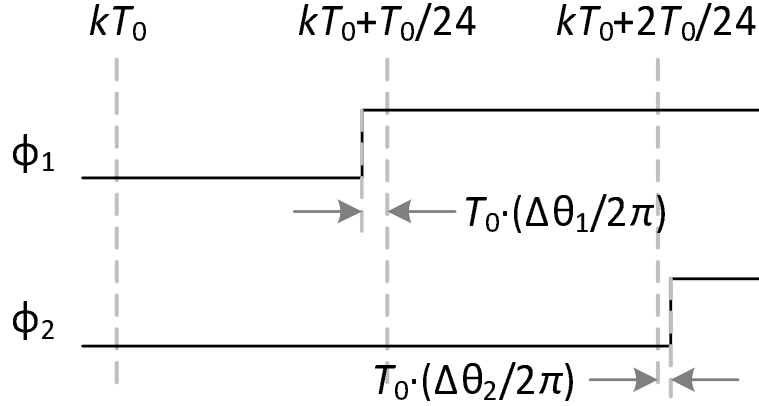


Figure 4.21: Definition of the phase error $\Delta\theta_i$.

phase error induced by delays in the clock divider and distribution network. To analyze the impact factors on the linearity of the generated “baseband” single tone signal, the current errors are annotated as $\Delta I_{\{a1,a2,b1,b2\}} \langle 0..7 \rangle$. One example has been shown in Fig. 4.15. I_i is used to represent all current branches switched by clock ϕ_i . The detailed combination pattern can be found in Table 4.1. For example, we have $I_0 = +I_{a1} \langle 0 \rangle - I_{b1} \langle 6 \rangle - I_{b1} \langle 7 \rangle$. The corresponding error term is denoted as ΔI_i . Moreover, θ_i is the phase of the clock ϕ_i . $\phi_{0..23}$ are equally distributed clocks, and thus $\theta_i = \pi/12 \cdot i$, $i = 0, 1, \dots, 23$. $\Delta\theta_i$ is the annotation of the i -th phase error, which is defined in Fig. 4.21. Consider the Fourier series of the “baseband” output current I_O ,

$$I_O(t) = \sum_{k=1}^{+\infty} X_k \cos(k\omega_0 t) + Y_k \sin(k\omega_0 t) \quad (4.15)$$

Noting that $I_i = -I_{i+12}$ for $i = 0, 1, \dots, 11$, X_k and Y_k in (4.15) can be expressed as

$$X_k = -\frac{2}{k\pi} \sum_{i=0}^{11} (I_i + \Delta I_i) [M_{k,i} \sin(k\theta_i) + N_{k,i} \cos(k\theta_i)] \quad (4.16)$$

$$Y_k = \frac{2}{k\pi} \sum_{i=0}^{11} (I_i + \Delta I_i) [M_{k,i} \cos(k\theta_i) - N_{k,i} \sin(k\theta_i)] \quad (4.17)$$

$$M_{k,i} = \cos(k \cdot \Delta\theta_i) - (-1)^k \cos(k \cdot \Delta\theta_{i+12}), N_{k,i} = \sin(k \cdot \Delta\theta_i) - (-1)^k \sin(k \cdot \Delta\theta_{i+12}) \quad (4.18)$$

where $A_k = \sqrt{X_k^2 + Y_k^2}$, $k = 1, 2, 3, \dots$ are the magnitude components of $I_O(t)$, also known as the “baseband” components described in Fig. 4.2a and (4.12).

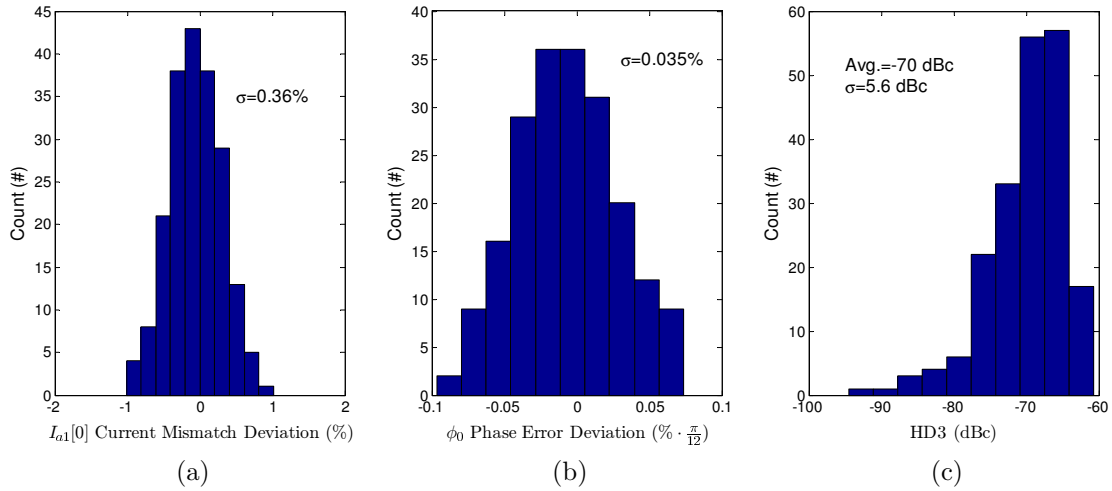


Figure 4.22: Monte-Carlo simulation results: (a) $I_{a1} <0>$ branch current deviation, (b) phase error deviation of ϕ_0 , and (c) distribution of the HD3 calculated from the current and the clock phase mismatches.

Two-hundred runs of Monte-Carlo simulation are carried out to obtain the mismatch distribution in the current mirror and the clock distribution network; thus, I_i , ΔI_i , and $\Delta\theta_i$ can be obtained. It should be mentioned that DEM is not applied in this analysis. Fig. 4.22 shows the numeric simulation results with the fundamental “baseband” frequency $f_0 = 1$ MHz. Fig. 4.22a demonstrates the simulated percentage of the deviation from the expected current ($\Delta I_i / I_i \cdot 100\%$) in branch $I_{a1} <0>$. More-

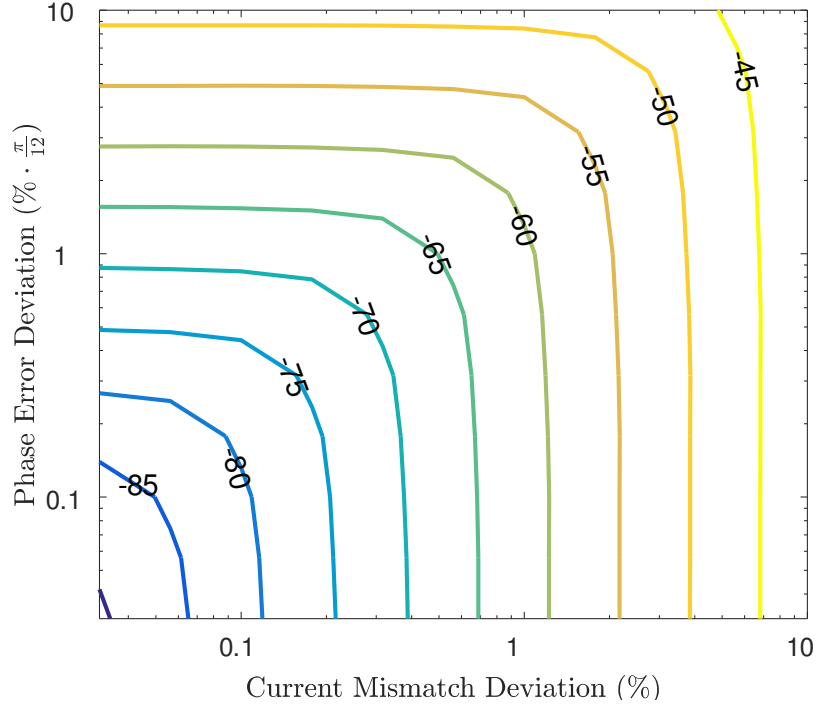


Figure 4.23: Contour of the simulated average HD3 by sweeping the current mismatch and the clock phase error.

over, the percentage of the deviation from the expected clock phase in ϕ_0 is also illustrated in Fig. 4.22b. By applying the statistical data I_i , ΔI_i , and $\Delta\theta_i$ to (4.15), we can obtain the distribution of the I_O waveform third order harmonic distortion (HD3) in Fig. 4.22c. It shows an average of -70 dBc HD3. Further numerical analysis uses the Monte-Carlo simulation data as a baseline and sweeps the clock phase error deviation and the current mismatch deviation. The newly generated data is fed into (4.15), leading to the average HD3 performance, which is shown in Fig. 4.23. It can be found that, to achieve -70 dBc HD3, about 0.3% errors for both the current mismatch and clock phase error are required. 0.1% errors make the HD3 below -80 dBc. Moreover, if an error factor is much bigger than the other, it will become the dominant error source. For instance, 1% current mismatch and 1% phase error result

in about -60.5 dBc HD3. If the phase error is reduced to 0.03% , we still have -61.4 dBc HD3. Thus, to increase the “baseband” linearity, both of the current mismatch and the clock phase error should be improved. On the one hand, because the delays in the clock distribution network are almost fixed after fabrication, lower ω_0 , which means longer clock period, can help decrease the percentage of delay errors. On the other hand, the DEM technique is introduced to alleviate the current magnitude mismatch when it becomes the dominant error source. To sum up, the mismatched currents and clock phases are both important factors when dealing with the linearity of the “baseband” single tone generation.

4.4.2 3-bit DEM Rotator

The DEM technique can randomize the branch currents and thus relaxes the matching requirement of the current mirrors. However, it also raises the noise floor of the “baseband” sinusoidal output, deserving further analysis. A behavior model of the “baseband” synthesizer, which adopts the functionality of the proposed DEM approach in Fig. 4.16, is implemented with the aforementioned current mismatch data ($\sigma = 0.36\%$). Ideal clock phases are used so as to reveal the intrinsic noise floor induced by the DEM. Fig. 4.24a and Fig. 4.24b demonstrate the normalized power spectrum density (PSD) before and after turning on the DEM for one group of current mismatch data. The DEM reduces the HD3 from a -68 dBc harmonic tone to a noise floor of -102 dBc. The average noise floor obtained from two-hundred Monte-Carlo runs is -105 dBc. Moreover, the distribution of ΔI_i will also affect the noise level. Fig. 4.25 shows the average noise floor versus different current mismatch deviation. The proposed DEM provides an average 35 dB suppression to the 3rd order harmonic. To conclude, the DEM technique significantly reduces the HD3 induced by the current mismatch and thus makes the “baseband” synthesizer more

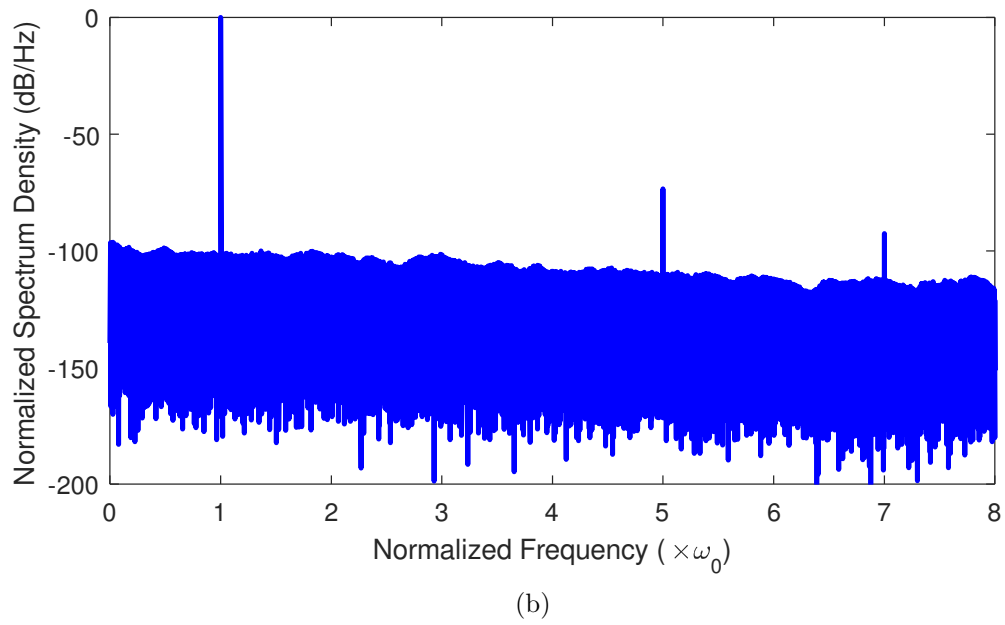
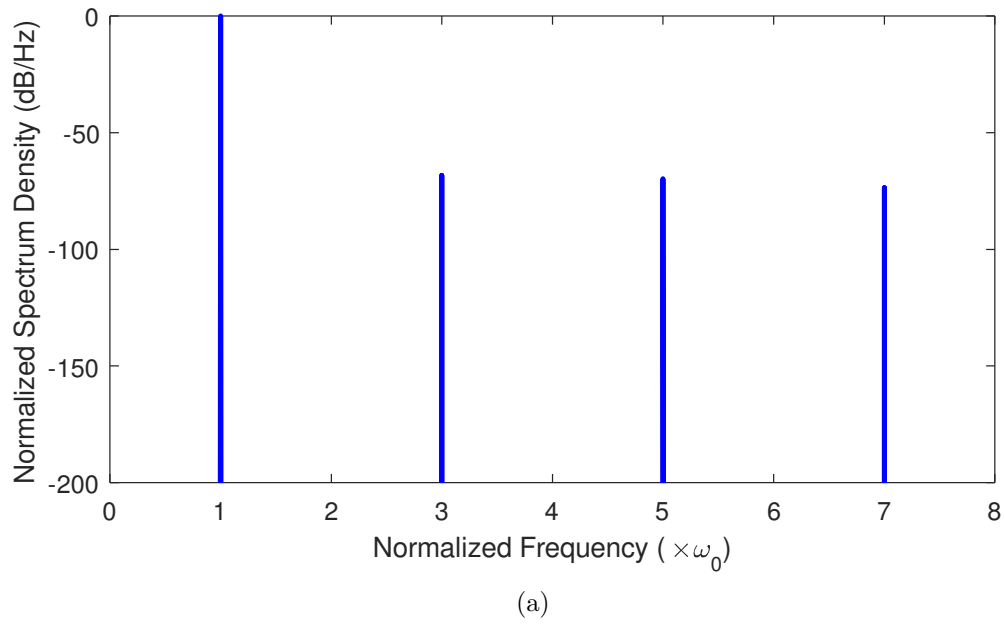


Figure 4.24: Simulated power spectrum density with current mismatches plus (a) DEM OFF or (b) DEM ON.

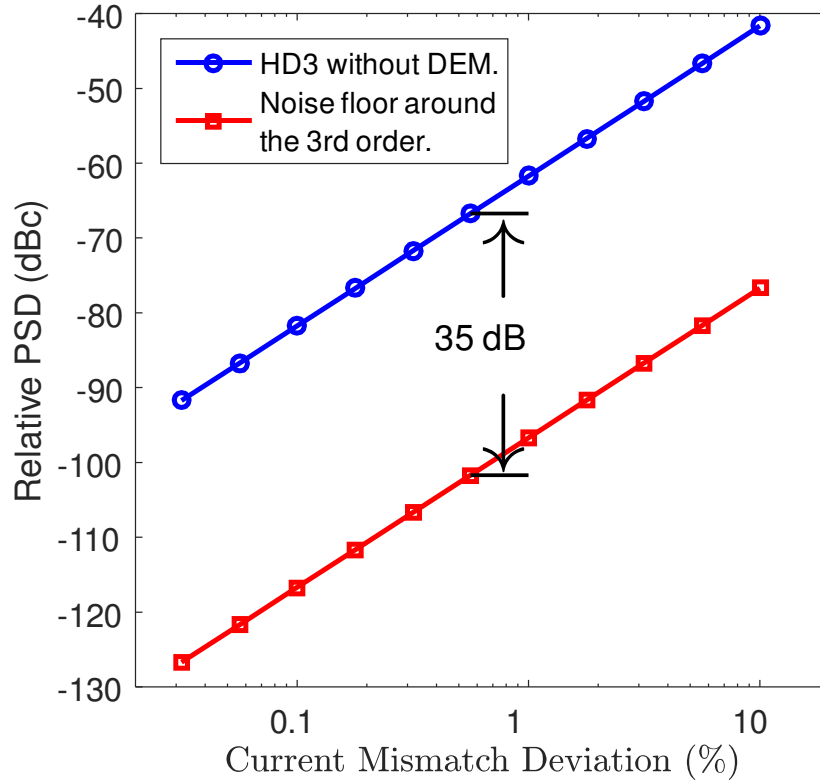


Figure 4.25: Simulated HD3 suppression by using the proposed DEM technique.

sensitive to the clock phase error. The chosen $\sigma = 0.36\%$ current mismatch deviation in this design is considered conservative. In other words, the current mirror array could be designed smaller without much linearity degradation penalty.

4.4.3 Nonlinear Up-conversion

The linearity of the passive CMOS mixer has been analyzed in [103], and its conclusion can be applied to the design of the passive mixer in the proposed architecture. Considering the “baseband” single tone generator is an equivalent current-steering DAC, the input impedance of the mixer is large due to the current mirror structure. This impedance is optimal for the mixer’s linearity. Particularly, a low-pass filter

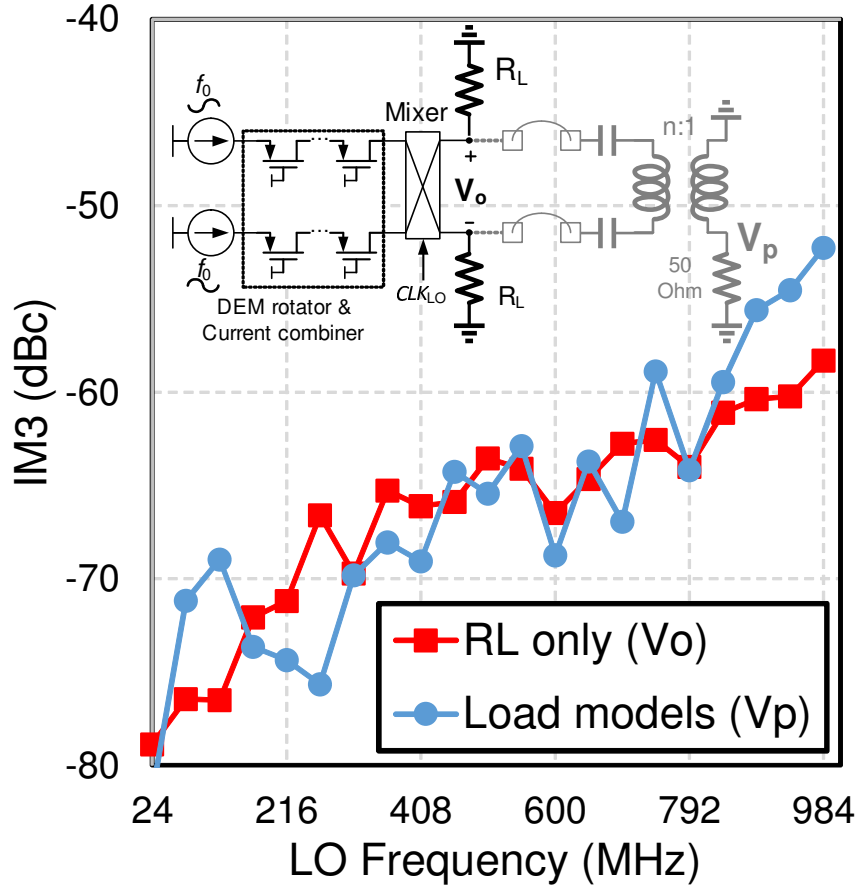


Figure 4.26: Simulated IM3 of the passive mixer ($f_0 = 1\text{MHz}$).

cannot be implemented to purify the “baseband” signal, such as the one proposed in [44]. Any extra capacitance introduced by an LPF may reduce the mixer’s input impedance at high frequencies, and, thus, degrade its linearity performance. Moreover, [103] also suggests using low load impedance to improve the linearity. $R_L = 200 \Omega$ load resistance (see Fig. 4.7) is adopted in this design. It is reasonable for an on-chip BIST application, which allows the proposed design to drive most DUTs on chip and significantly reduces the area and power overhead, compared to a 50Ω load. LO frequency is also a factor that affects the mixer’s linearity. To evaluate the mixer’s frequency-dependent linearity degradation, a schematic with only resistive

load is simulated, compared to another simulation taking the bondwire and the balun models into consideration. The simulated IM3 obtained on V_O and V_P at $f_0 = 1$ MHz are depicted in Fig. 4.26, as well as the detailed schematic. It can be found that the external components used for test purpose introduce more fluctuation to the output IM3 due to the changing impedance matching conditions. However, in the on-chip BIST application scheme, this fluctuation should not be a concern. From Fig. 4.26, we can find the simulated IM3 keeps increasing when the LO frequency rises. To the contrary, if given a fixed ω_0 , the “baseband” HD3 is almost fixed because the clock phase errors and the current mismatches will not change too much. As the output linearity is determined by both the “baseband” generator and the mixer, at high LO frequency, the output IM3 will increase as the mixer dominates the linearity degradation. While at low LO frequency the output IM3 will keep almost the same, because the “baseband” becomes the major source of IM3 tones.

4.4.4 Aliasing of the Residue Harmonics

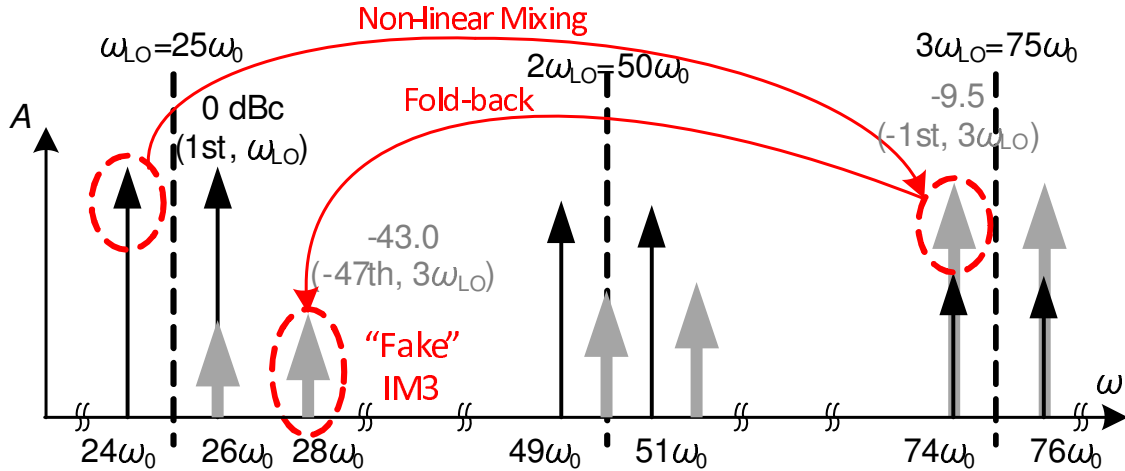


Figure 4.27: “Fake” IM3 induced by fold-back harmonics.

In Section 4.2.3, we showed that the first non-cancellable harmonic is pushed to the 23rd order. As a result, multiple residue harmonics are located at $(24k \pm 1) \cdot \omega_0$, where $k = 1, 2, \dots$. The equivalent input of the cascade FIR architecture is a square wave, the closest two significant harmonics, the 23rd and the 25th order, still have high amplitudes $(\frac{1}{24k \pm 1})$, -27.2 dBc and -28.0 dBc, respectively. The up-conversion mixer adopts the passive switching structure, which has a strong nonlinear mixing behavior, leading to aliasing issues. By replacing $\cos(\omega_{LO}t)$ in (4.12) with the Fourier series of a square wave, $\frac{4}{\pi} \sum_{l=0}^{\infty} \frac{1}{2l+1} \cos[(2l+1)\omega_{LO}t]$, the aliasing will spread the residue harmonics across the whole spectrum, locating at $(2l+1)\omega_{LO} \pm (24k \pm 1)\omega_0$, where $l = 1, 2, \dots$. Without a carefully chosen ω_{LO} and ω_0 , some residue harmonic may fold back to the IM3 frequency, leading to a high “fake” IM3. As shown in Fig. 4.27, assuming we have $\omega_{LO} = 25\omega_0$, the two major tones are at $\omega_1 = 24\omega_0$ and $\omega_2 = 26\omega_0$, and the IM3 tones emerge at $2\omega_1 - \omega_2 = 22\omega_0$ and $2\omega_2 - \omega_1 = 28\omega_0$. If we have $l = 1$ and $k = 2$, the nonlinear mixing first duplicates the “baseband”, centering at $3\omega_{LO}$ and attenuates the whole band by 9.5 dB. The -47 th order residue harmonic (where the minus indicates the left side band of $3\omega_{LO}$) of the duplicated band will fold back to $3\omega_{LO} - 47\omega_0 = 28\omega_0$. It has an amplitude of -43.0 dBc, which is much higher than the originally suppressed IM3 tone. To prevent having high “fake” IM3 tones, one solution is to set ω_{LO} an integer multiple of ω_0 , but avoid $(12m \pm 1)\omega_0$ or $(12m \pm 2)\omega_0$, where m is an integer number. Fractional multiple of ω_0 can also be used. However, this will make it more complicated to judge whether the “fake” IM3 tones exists. It will also prevent ω_{LO} from being divided to obtain ω_0 ; and thus an extra clock source is needed.

4.4.5 LO Leakage and Imbalance

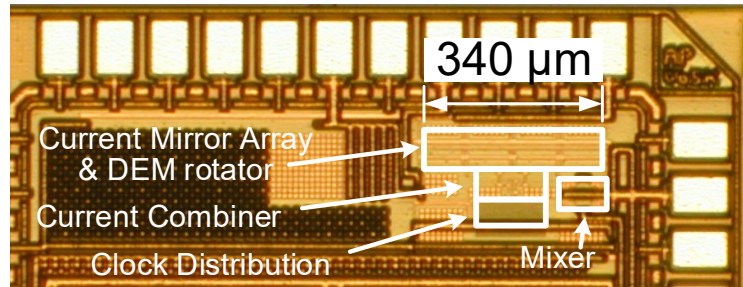
The analysis above covers different error sources that cause the degradation of IM3. Apart from the two fundamental tones and the corresponding two IM3 tones that will appear in the output spectrum, there are some other tones emerging around ω_{LO} . On the one hand, the usage of a passive mixer raises a concern that the output waveform V_O contains the LO leakage signal, which appears at ω_{LO} . On the other hand, it is difficult to make a fully symmetric layout design or achieve a 50% duty cycle LO clock at high frequency; thus, imbalance is introduced to the output waveform. To take these non-ideal factors into consideration, we can rewrite (4.12) as

$$u'' = u' + B \cos(\omega_{LO}t) + C \cos[(\omega_{LO} \pm 2\omega_0)t] \quad (4.19)$$

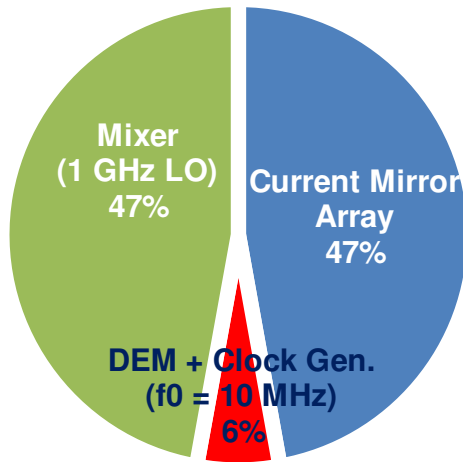
where B is the amplitude of the LO leakage signal, and C is used to approximate the imbalance. Therefore, by counting A_1 and A_3 , extra terms are added into the error in (4.13),

$$\varepsilon' \approx \varepsilon + \frac{3}{4} \frac{k_3}{k_1} \left[\frac{A_3}{A_1} (3B^2 + 6C^2) + 3C^2 + 6BC \right] \quad (4.20)$$

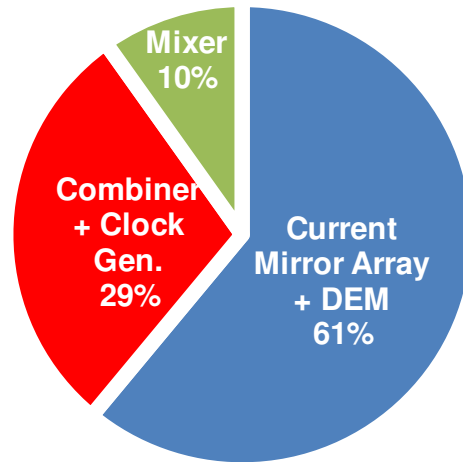
It can be concluded that, even if the LO leakage amplitude (B) is high, the small coefficient A_3/A_1 will markedly lower its impact in the two-tone test. More emphasis should be put on the symmetry of C . Nevertheless, even $C > A_3$, its effect can be neglected, providing $C \ll A_1$. This observation makes the design constraint of the mixer much more relaxed—relative high LO leakage and some asymmetry in the output waveform are acceptable.



(a)



(b)



(c)

Figure 4.28: (a) Die photograph, (b) power distribution, and (c) block area distribution of the proposed two-tone generator.

4.5 Experiment Results

The proposed two-tone signal generator is fabricated in 130 nm standard CMOS technology. The chip die micrograph is shown in Fig. 4.28, and the percentage of each block's area and power is also compared. The total silicon area is 0.056 mm^2 . From a 1.5 V power supply, the current mirror array draws around 2 mA, while the mixer consumes a maximum 2 mA at the LO frequency of 1 GHz. The other circuits are working under a 1.2 V supply.

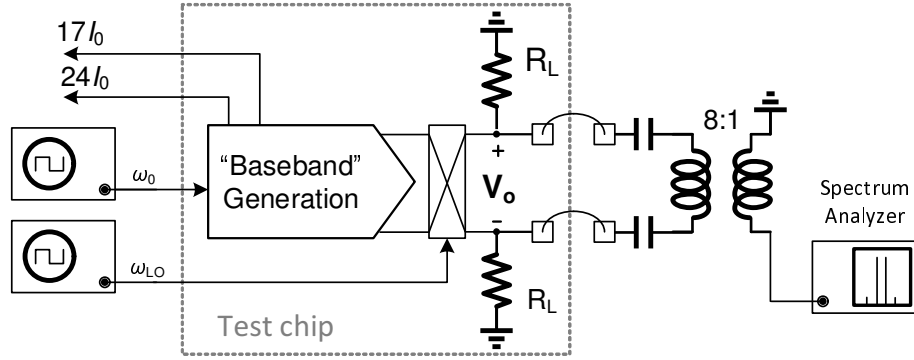
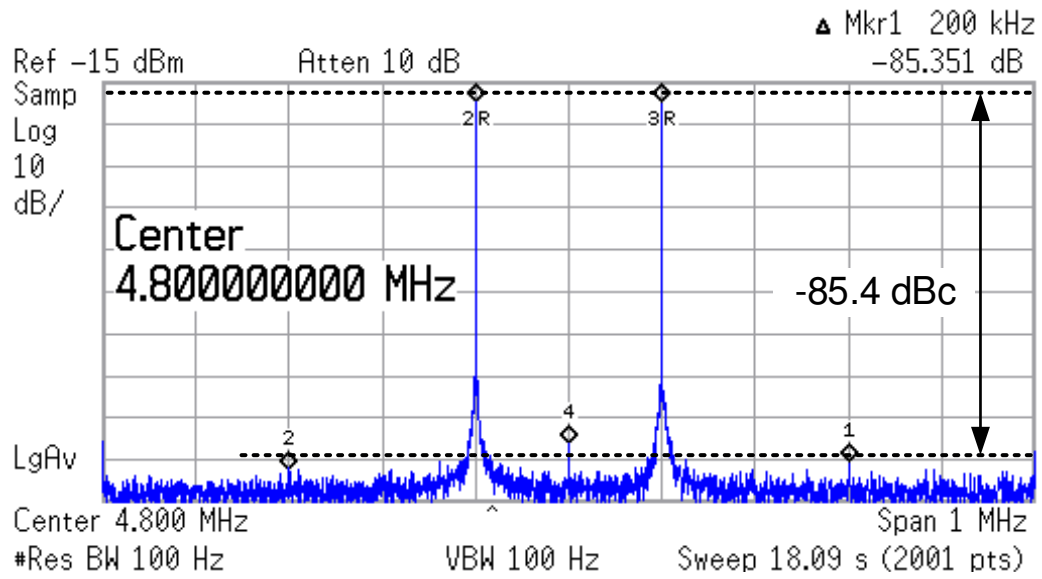


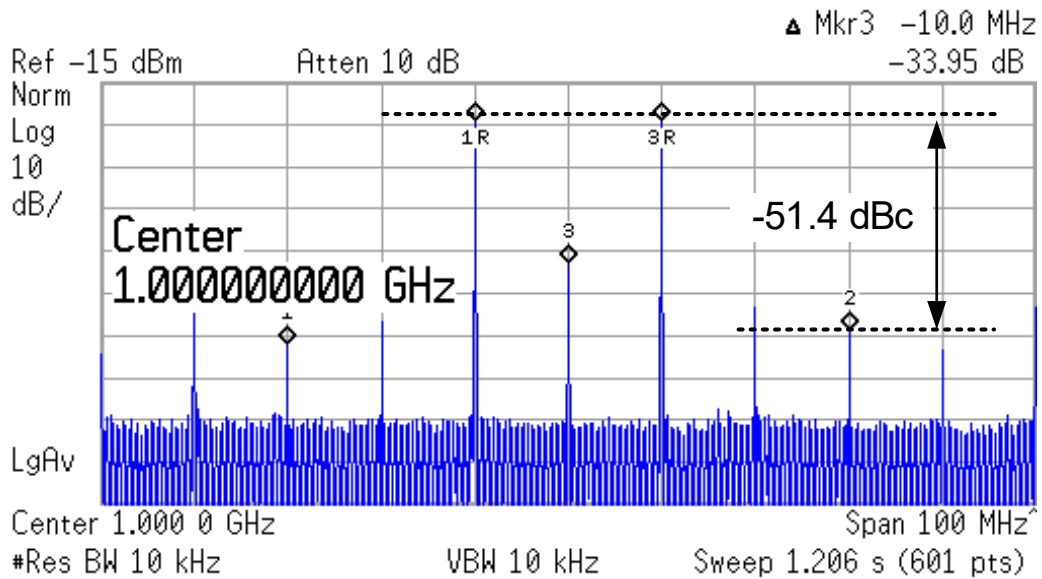
Figure 4.29: Test bench configuration for the proposed two-tone generator.

Moreover, the test bench is set up as illustrated in Fig. 4.29. The differential output impedance is set to 400Ω (200Ω for each R_L). Current biases $17I_0$ and $24I_0$ is provided from external sources. Two clock generators feed the test chip with clocks at f_0 and f_{LO} , respectively. A balun with an 8:1 impedance ratio is used to bridge the synthesizer and the spectrum analyzer.

Measurement results are shown in Fig. 4.30. Fig. 4.30a illustrates that an IM3 of -85.4 dBc is achieved at a relative low LO frequency ($f_{LO}=4.8$ MHz), and $CLK_{LF}(24f_0)$ is set to 2.4 MHz. The measured IM3 rises to -51.4 dBc when f_{LO} is increased to 1 GHz, as demonstrated in Fig. 4.30b. Note that a high LO leakage tone appears at f_{LO} , and the even-order distortion at $f_{LO} \pm 2f_0$ also emerge. For (4.20), we have $B = -33.9$ dBc, $C = -44$ dBc, and $A_3/A_1 = -51.4$ dBc, yielding $\epsilon' \approx \epsilon + \frac{3}{4} \frac{k_3}{k_1} A_1^2 \times 0.09\%$. Therefore, even in the worst case, the LO leakage and $f_{LO} \pm 2f_0$ distortion has limited impact on the IM3 measurement. Fig. 4.30c shows the change of measured IM3 in the f_{LO} range from 48 to 1008 MHz with a fixed $f_0=1$ MHz. Better than -68 dBc IM3 can be achieved when $f_{LO} < 480$ MHz. The results for f_{LO} range from 2.4 to 153.6 MHz are further shown in Fig. 4.30d, in which the IM3 values with different f_0 are also compared. The measured IM3 is <

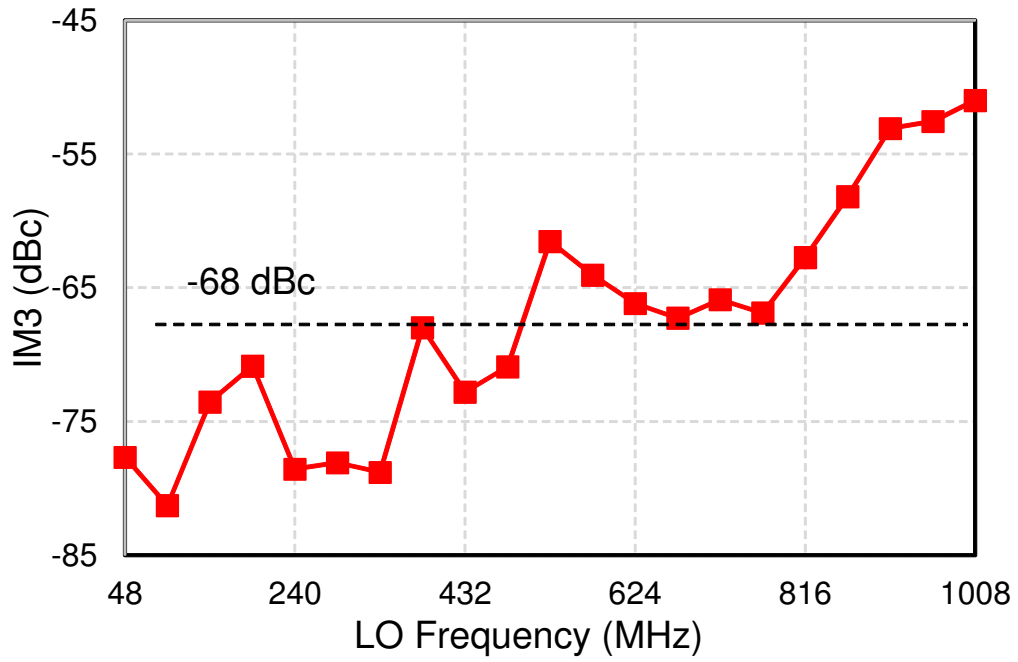


(a)

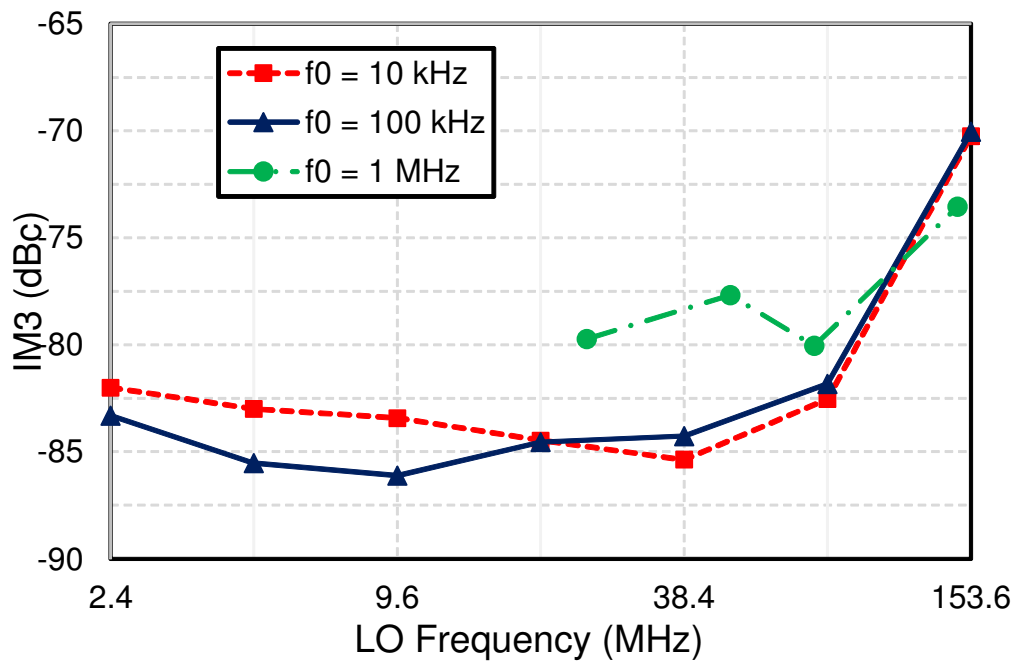


(b)

Figure 4.30: Measurement results: (a) IM3 ($f_{LO}=4.8$ MHz and $f_0=100$ kHz), (b) IM3 ($f_{LO}=480$ MHz and $f_0=1$ MHz), (c) IM3, where f_{LO} is swept from 24 MHz to 1008 MHz ($f_0=1$ MHz), and (d) IM3, where f_{LO} is swept from 2.4 MHz to 153.6 MHz (10 kHz, 100 kHz, and 1 MHz f_0).



(c)



(d)

Figure 4.30: Continued.

-75 dBc within 76.8 MHz f_{LO} , and $IM3 < -80$ dBc can further be obtained with $f_0 \leq 100$ kHz. These results are close to the simulation results in Fig. 4.22 and Fig. 4.26. Moreover, it can be concluded that the “baseband” single tone generator has around -85 dBc $IM3$ limit, which dominates the linearity for $f_{LO} < 100$ MHz. Above 100 MHz, the passive mixer becomes the major contributor of the linearity degradation. Additionally, the balance between the two-tone amplitudes exhibits < 0.1 dB difference across the whole frequency range.

Section 4.4.1 concludes that the DEM technique can provide a further improvement of linearity when the current mismatch becomes the dominant error source. Fig. 4.31b shows that the $IM3$ improves almost 14 dB when the DEM is ON, compared to the results in Fig. 4.31a with the DEM turned off. Around the $IM3$ tones, the measured noise floor is about -108 dBc without DEM and it raises 1 dB after the DEM is turned on. We can conclude that the proposed DEM technique achieves low induced noise, which is consistent with Fig. 4.24b. Fig. 4.32 further demonstrates the measured $IM3$ improvement (for $f_0 = 100$ kHz and $f_0 = 1$ MHz) after turning on the DEM. As expected, the linearity improvement is significant when the current mismatch becomes the dominant error source. While at high LO frequency, the mixer distortion becomes dominant; and thus the DEM is not as effective as working at low LO frequency.

Table 4.2 compares the performance of the proposed synthesizer to the other state-of-the-art works. The reported $IM3$ upper limits and the corresponding frequency ranges are especially significant. Single-tone sinusoidal synthesizers are also used as references. Furthermore, to evaluate whether the signal generator is suitable

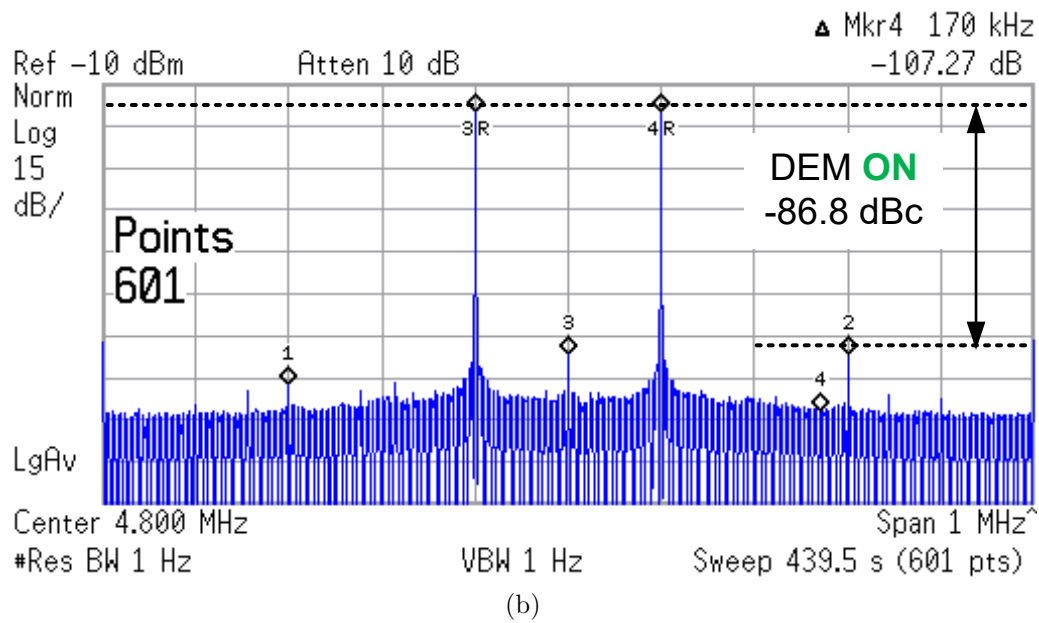
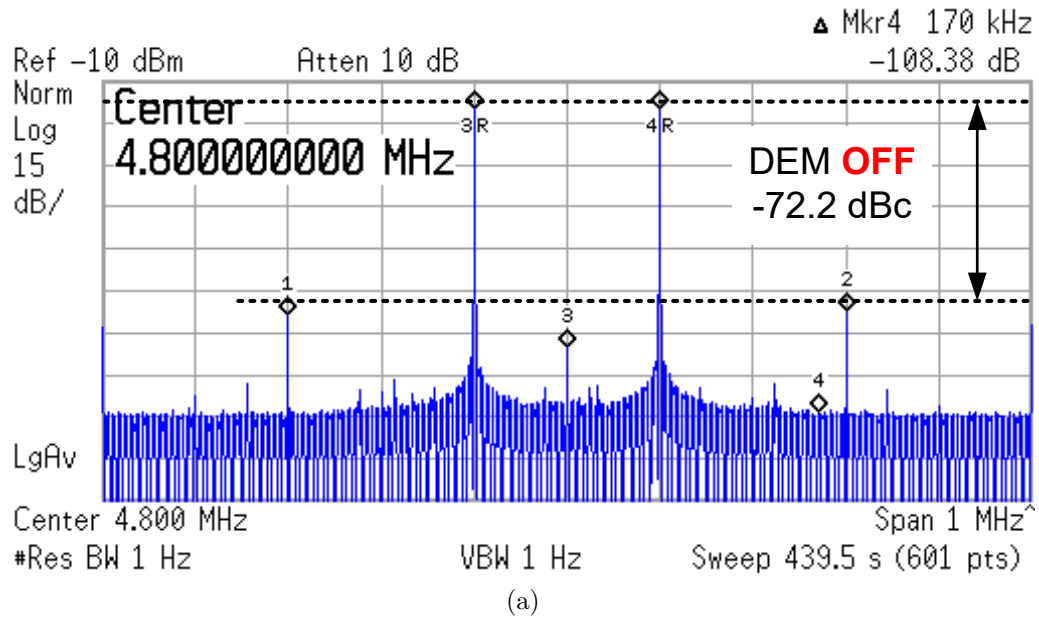


Figure 4.31: Measured IM3 of the two-tone signals ($f_{LO} = 4.8$ MHz, $f_0 = 100$ kHz) (a) without DEM or (b) with DEM.

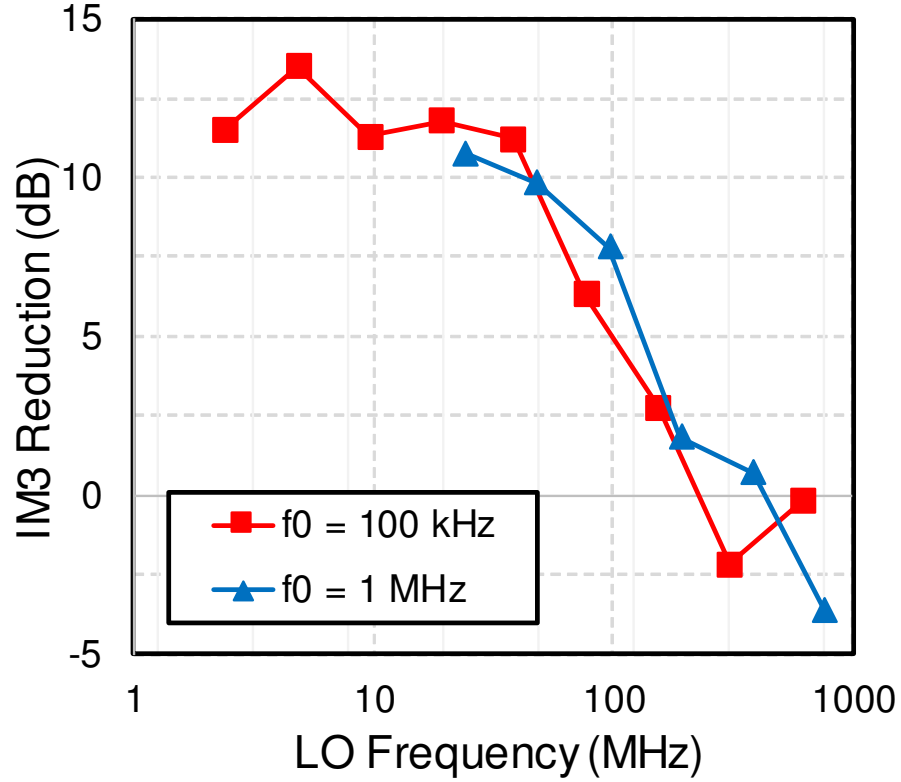


Figure 4.32: Measured IM3 improvement by turning on DEM.

for on-chip BIST application, the Figure of Merit (FoM) in [44] can be modified to

$$\text{FoM}_{\text{BIST}} = \frac{f_{LO}(\text{MHz}) \times 2^{\frac{-\text{IM3}(\text{dB})}{6}}}{P_{\text{total}}(\text{mW}) \times \frac{A(\text{mm}^2)}{L_{ch}^2(\mu\text{m}^2)}} \quad (4.21)$$

where f_{LO} and IM3 are the characteristic LO frequency and the corresponding reported IM3. L_{ch} is the process node in μm . To sum up, the proposed calibration-free synthesizer architecture shows comparable linearity performance to the other state-of-the-art works but has smaller hardware overhead. It is suitable for on-chip linearity BIST applications.

Table 4.2: Comparison of Two-/Single-tone Generation Performance

Ref.	This work	[16] ¹	[88]	[89]	[91]	[90]	[104]	[66] ³	[44] ³
Max Freq.	1 GHz	N.A.	250 MHz	800 MHz / 1.4 GHz	1 GHz	5.26 GHz	1.062 GHz	11 MHz	850 MHz
Area (mm ²)	0.056	N.A.	0.034	0.016	3.3 ²	1.6 ²	0.23	0.186	0.08
Process (nm)	130	65	180	40	65	65	250	130	180
Supply (V)	1.2/1.5	N.A.	1.8	1.2	1.0/2.5	1.2/3.3	2.5	1.2	1.8
IM3/HD3 (dBc)	-66 @ 720 MHz			-61.5 @ 1.36 GHz ⁴	-80 @ 950 MHz	-62 @ 4.1 GHz	-64.7 @ 1.0 GHz	-72 @ 10 MHz	-70 @ 750 MHz
	-68 @ 480 MHz	-75	-62.16 @ 246 MHz	-70 @ 800 MHz	-85 @ 300 MHz	-82 @ 1.9 GHz			
	-70 @ 336 MHz			-75 @ 20 MHz					
	-80 @ 20 MHz								
I_{load} (mA)	2	N.A.	10	16	16	20	5.1	N.A.	N.A.
R_L (Ω)	400	50	50	50	50	50	100	N.A.	N.A.
P_{total} (mW)	6	N.A.	24	40	681	380	122	3.34	57
Type	Mixing-FIR	Dual-VCO	DAC	DAC	DAC	Mixing-DAC	Mixing-DAC	HC	HC(FIR)
FoM _{BIST} ($\times 10^4$)	8.70	N.A.	1.30	0.41	0.018	0.12	0.38	N.A.	N.A.

¹ Simulation result; ² Area of digital circuits are included; ³ Single-tone generation HD3 for reference;

⁴ IM3 is measured @ 2.8 GS/s, the others are measured @ 1.6 GS/s.

4.6 Conclusions

A sine-wave synthesizer is proposed to generate two low-distortion sinusoidal signals for on-chip linearity BIST. It is driven by only square-wave digital clocks and can cover a wide frequency range from DC to 1 GHz. A cascade FIR architecture with three 5-taps and one 3-tap current-steering FIR blocks are used in the “baseband” single tone generator to suppress the 3rd-order component, which has the most significant impact on the precision of the IM3 two-tone test. Due to the low frequency “baseband” clock, phase error in the clock distribution network can be minimized. The dynamic element matching technique is further implemented to improve matching among current branches in the current mirror array, which implements the FIR tap coefficients. Thus, strong cancellation of the 3rd-order harmonic can be achieved without calibration for the single tone. The generated single tone is then up-converted by a passive mixer using a high speed clock at the LO frequency. The up-converting produces two sinusoidal tones and the mirroring mechanism guarantees balance between the amplitudes of these two tones. Moreover, the compact synthesizer design, which has only logic gates, MOS switches, and current mirror array, is scalable for future advanced IC processes. This approach is one step towards the built-in self-test and in-situ optimization for integrated circuit design.

5. CT+DT HYBRID BASEBAND CHAIN USING HARMONIC CANCELLATION FOR ON-CHIP LINEARITY TEST

5.1 Background

Charge-domain filter (CDF), or discrete-time filter (DTF), is a competitive base-band architecture of modern multi-mode/multi-band communication systems, which is attracting a lot of research attention. It consists of several switched capacitor samplers controlled by multi-phase clocks, achieving finite-impulse-responses (FIRs) and/or infinite-impulse-responses (IIRs), which are frequency responses. Compared to the conventional analog continuous-time (CT) filters, DTF's bandwidth can be reconfigured by simply changing the sampling clock frequency, which makes it flexible for software-defined radio (SDR) applications [105, 106]. In addition, high-order FIR response shows a rapid out-of-band roll-off rate. This characteristic is good for adjacent channel rejection (ACR) in many communication systems [107].

The z-domain transfer function of an FIR filter $\alpha_0 + \alpha_1 z^{-1} + \dots + \alpha_n z^{-n}$ can be constructed by a combination of delays z^{-i} and tap coefficients α_i . Correspondingly, in a DTF, the MOS switches controlled by delayed clocks represent the delays, while the capacitance values determine the tap coefficients, as shown in Fig. 5.1. A four-tap DTF with uniform tap coefficients is proposed in [108]. That work also indicated that, for a decimation ratio of four, six clock phases (or unit samplers) would have been sufficient, including four sampling and one decimation phase. Various capacitance values can also be used for different clock phases. [109] demonstrated such a down-conversion DTF that implements a band-pass FIR transfer function, using six capacitance values to achieve 11 different filter tap coefficients.

The decimation technique was implemented in down-conversion DTFs [106, 108,

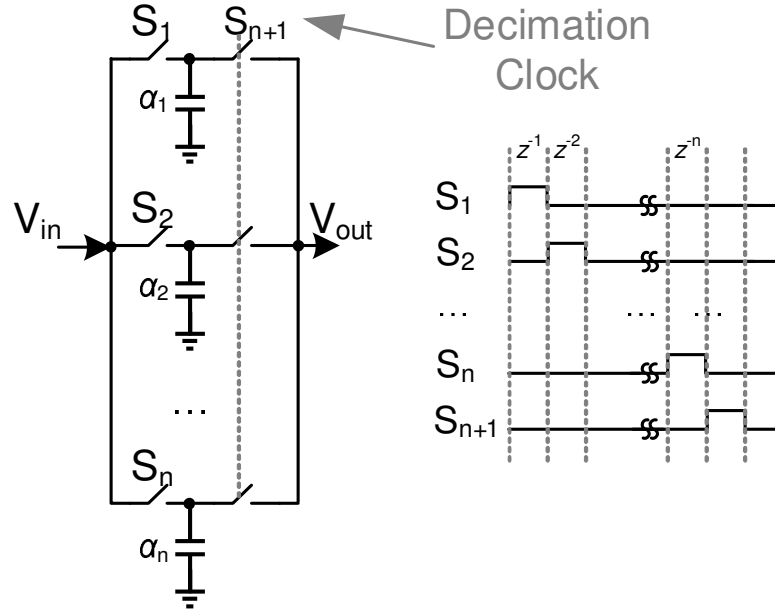


Figure 5.1: DTF architecture with the decimation technique.

109, 110, 111, 112], in which the output sampling frequency of the DTF was reduced, which is also illustrated in Fig. 5.1. These DTF designs play a full or partial role of a down-conversion mixer thereby removing the speed constraint of the baseband analog-to-digital converter (ADC). However, the reduced output sampling frequency also prevents the DTF from being cascaded to construct high-order DTFs without losing bandwidth. To solve this issue, [113] proposed a hybrid DTF architecture, which adopts a non-decimation DTF as the first stage. A four-stage cascade non-decimation DTF was further proposed in [107] with a bandwidth calibration scheme. Moreover, [114] proposed a hybrid baseband chain for long-term evolution (LTE) applications, which combines a continuous-time low-pass filter (LPF) with a DTF. On the one hand, the CT LPF was adopted to compensate the passband distortion of the DTF, achieving a flat in-band response. On the other hand, the DTF is implemented by a two-stage 3-tap moving average (MA-3²) architecture. This hybrid

cascade architecture enhances the filter’s stop-band attenuation.

Furthermore, few publications have discussed the on-chip built-in linearity measurement techniques for receiver chains. [19] proposed a built-in linearity optimization system, which integrates an envelope detector, an analog-to-digital converter (ADC), a third-order intermodulation (IM3) calibration unit and an on-chip spectrum analyzer, which are introduced in [18]. [18] further analyzed the precision requirements for both the DAC and the FFT algorithm. However, the IM3 tones are usually much lower than the major test tones in the receiver linearity test. To detect the weak IM3 tones, a high resolution ADC and long FFT sampling sequence are needed, resulting in a significant hardware overhead introduced by the built-in self-test circuits.

In this section, a dual-mode discrete-time filter with a harmonic cancellation technique is proposed for a CT+DT hybrid receiver baseband chain; plus a linearity measurement and design method is constructed based on the behavior of this specific DTF. The proposed DTF works as the last filtering stage of the CT+DT baseband chain in the normal operation mode; while in the tone suppression mode, the bandwidth of the DTF is changed to suppress the two test tones and thus expose the IM3 tones. A power detector is used to measure the power of the remaining tones at the output, achieving the linearity measurement by comparing the output power obtained from the two different modes. This section is organized as follows. The system architecture and the principle of linearity measurement is proposed in Section 5.2. The detailed DTF implementation is introduced in Section 5.3. Section 5.5 further discusses the factors that impact the measurement precision of the proposed method. Section 5.6 shows the experimental procedures and results, followed by conclusions in Section 5.7.

5.2 Linearity Measurement of a Hybrid Chain

5.2.1 Hybrid Chain System Architecture

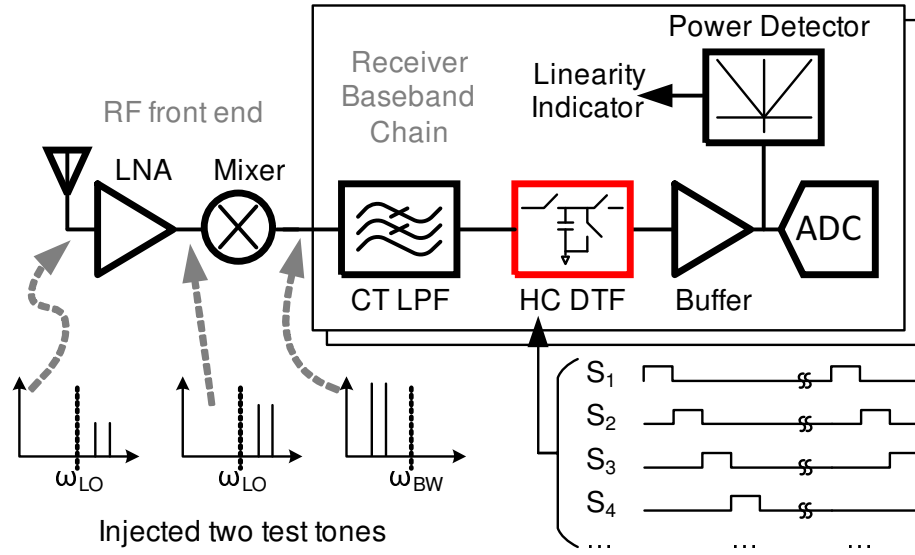


Figure 5.2: Receiver architecture using CT+DT hybrid baseband chain.

Fig. 5.2 illustrates the receiver architecture using a CT+DT hybrid baseband chain. It consists of an RF front end and two baseband chains for quadrature demodulation. In the baseband chain design, a CT+DT hybrid architecture was adopted, including a CT LPF and a DTF implementing the harmonic cancellation (HC) technique. The proposed HC DTF has two modes—the normal operation mode and the tone suppression mode. In the normal operation mode, the DTF works together with the CT LPF and provides a high order filter response for the baseband. To test the linearity of the proposed receiver architecture, the two test tones should be first injected into the chain, as indicated in Fig. 5.2. The HC DTF can be switched between the normal operation mode and the tone suppression mode with designated

sampling frequencies (ω_S) to pass or suppress the two test tones, so as to expose the test tones or the IM3 tones to the following circuit. A power detector is attached to the end of the proposed baseband chain to measure either tones' power. In this way, a composite linearity, which includes all preceding blocks starting from the injection point, can be observed. It should be emphasized that, the proposed linearity test mechanism reuses the building blocks of the receiver, and the overhead can be considered as zero. The power detector is usually adopted in receivers as the received signal strength indicator (RSSI), while an analog-to-digital converter (ADC) is not used for the proposed linearity test. In addition, to demonstrate the proposed linearity test concept, a programmable 2nd-order Tow-Thomas biquad using Miller-compensated two-stage amplifiers is implemented as the CT LPF in this design. It has a variable bandwidth of ω_0 and a fixed quality factor of $Q = 1$. According to the analysis in [114], $\omega_S = \frac{16}{3}\omega_0$ should be set for the HC DTF's normal operation mode to obtain the best in-band flatness of the hybrid baseband. The filters' frequency response and details about different modes will be further discussed in Section 5.2.3.

5.2.2 Harmonic Cancellation for Two-tone Suppression

The harmonic cancellation technique based on the finite impulse response (FIR) filter approach has been proposed in [44]. The proposed M -tap FIR filter suppresses odd order harmonics up to the $(2M + 1)$ th order with designated tap coefficients

$$c_i = \cos\left(\frac{(i+1)\pi}{M+1} - \frac{\pi}{2}\right), (i = 0, 1, \dots, M-1) \quad (5.1)$$

and a fixed delay of $T_S/(2M + 2)$ between two adjacent taps. T_S is the period of the sampling clock, and we have $\omega_S = 2\pi f_S = 2\pi \cdot T_S^{-1}$. According to the analysis in Section 3.2.3 and Table 3.1, a 3-tap ($M = 3$) FIR filter has two notching frequencies, $3f_S/8$ and $5f_S/8$. Assume this FIR filter is connected to a nonlinear circuit that

is excited by two tones at $3f_S/8$ and $5f_S/8$. Then at the output, the two major tones will be eliminated, but the IM3 tones at $f_S/8$ and $7f_S/8$ remain the same. Therefore, the 3-tap harmonic cancellation FIR filter can be used to conduct an IM3 measurement. The z-domain transfer function of such an FIR filter is defined as

$$H(z) = 1 + \sqrt{2}z^{-1} + z^{-2} \quad (5.2)$$

In the real circuit design, $^{17}/_{12}$ can be used to approximate the irrational coefficients $\sqrt{2}$ in (5.2). It leads to only 0.17% quantization error. Furthermore, the detailed measurement procedure is explained in Section 5.2.3 and the circuit implementation is introduced in Section 5.3.

5.2.3 Linearity Measurement Method

It should be emphasized that, in the analysis below, ω_{LO} is the local oscillator frequency and ω_0 represents the bandwidth of the continuous-time filter. Moreover, the simulation results show that the hybrid baseband chain achieves the best in-band flatness when the sampling frequency of the DTF is set to $\omega_S = \frac{16}{3}\omega_0$. This frequency is named as the normal operation mode of the DTF.

Fig. 5.3a and Fig. 5.3b show the proposed procedures of the in-band and the out-of-band IP3 measurement for the proposed hybrid receiver architecture, respectively. During the in-band IP3 test (Fig. 5.3a), the two tones located at $\omega_{LO} + 0.6\omega_0$ and $\omega_{LO} + \omega_0$ are used as the excitation test signals. Particularly, different from the traditional in-band IP3 test method, which puts the two test tones at symmetrical frequencies with respect to ω_{LO} in the double side band (DSB), the proposed method requires both of the two test tones to remain in the same side band. After the RF down-conversion, the two major tones are at $0.6\omega_0$ and ω_0 in the single side band (SSB). IM3 tones at $0.2\omega_0$ and $1.4\omega_0$ also emerge. Then, the spectrum that has the

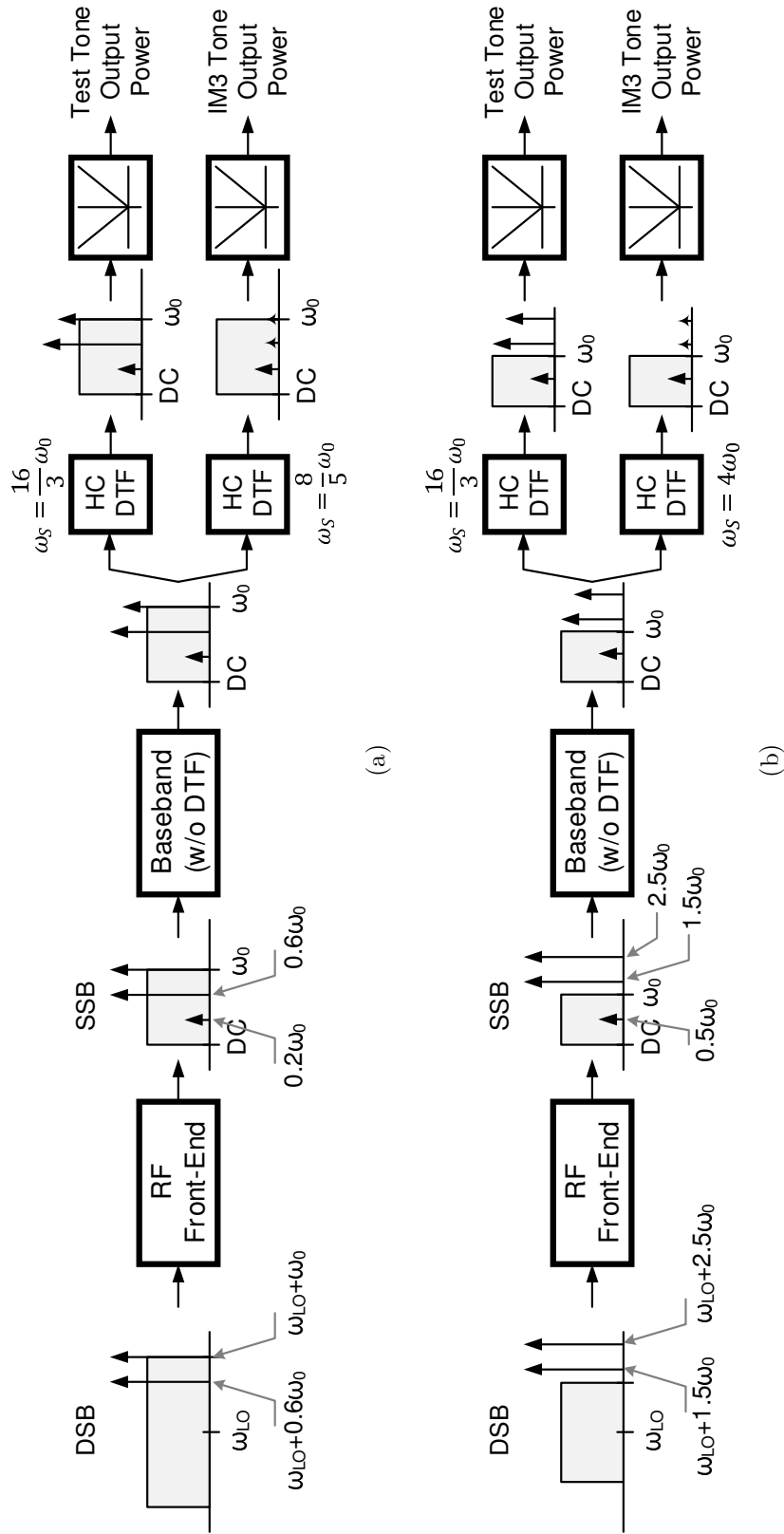
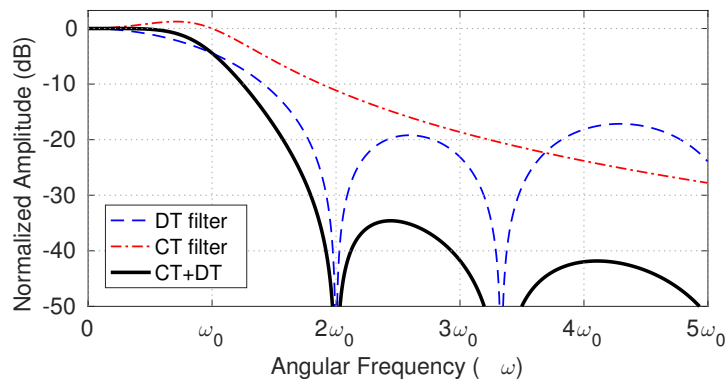


Figure 5.3: Proposed IP3 measurement procedure: (a) In-band IP3 and (b) out-of-band IP3.

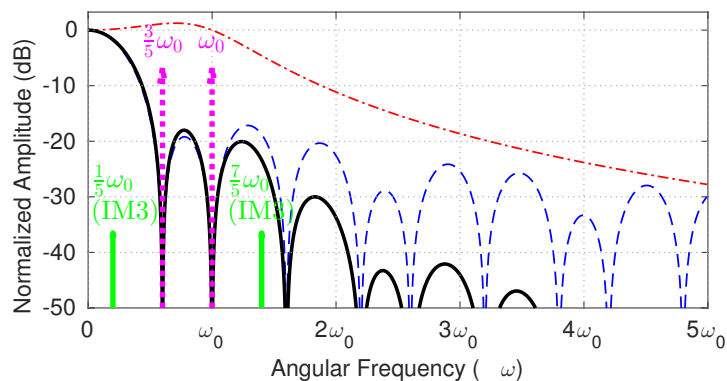
two test tones and the two induced IM3 tones passes through the HC DTF which is configured in the normal operation mode ($\omega_S = \frac{16}{3}\omega_0$). Because the power of the two test tones is much higher than that of the two generated IM3 tones, the power detector usually picks the strongest test tone power. Moreover, the HC DTF can also be configured to the in-band IP3 tone suppression mode, which sets $\omega_S = \frac{8}{5}\omega_0$. In this case, the HC DTF notches $0.6\omega_0$ and ω_0 . The two major tones are suppressed, and thus, the IM3 tone power is exposed to the power detector. Particularly, the IM3 tone at $0.2\omega_0$ falls inside the CT filter bandwidth ω_0 . It can be detected without significant loss. The out-of-band IP3 test follows a similar procedure, except that the test tones are moved to $\omega_{LO} + 1.5\omega_0$ and $\omega_{LO} + 2.5\omega_0$. Correspondingly, the tone suppression mode for out-of-band IP3 configures the HC DTF with $\omega_S = 4\omega_0$, so as to eliminate the two test tones (Fig. 5.3b). In this case, the IM3 tone at $0.5\omega_0$ falls inside the bandwidth ω_0 . To conclude, with the designated two tone excitations, the normal operation mode feeds the test tone power to the power detector, while the tone suppression mode outputs the IM3 tone power. Fig. 5.4a, Fig. 5.4b, and Fig. 5.4c further illustrate the composite frequency responses of the baseband chain in the normal operation, the in-band IP3 tone suppression, and the out-of-band IP3 tone suppression modes. The positions of the two test tones are also annotated in these figures. It should be noted that the DTF not only carries the harmonic cancellation response but is also attenuated by a sinc-like response due to the zero-order-hold mechanism, which is discussed in [114]. The two suppressed test tones and the residual IM3 tones are also annotated in Fig. 5.4b and Fig. 5.4c.

5.2.4 Circuit Implementation

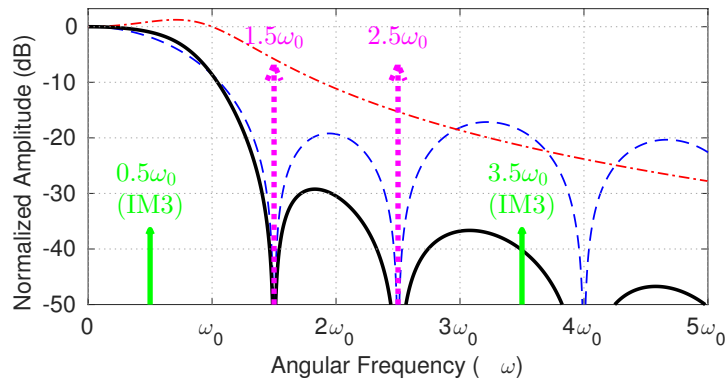
To verify the conceptual linearity test method, a single receiver baseband path is implemented with a 2nd-order programmable active-RC biquad and a HC-3² DTF.



(a)



(b)



(c)

Figure 5.4: Frequency responses with different sampling clock settings: (a) Normal operation mode ($\omega_S = \frac{16}{3}\omega_0$), (b) test tone suppression for in-band IP3 test ($\omega_S = \frac{8}{5}\omega_0$), and (c) test tone suppression for out-of-band IP3 test ($\omega_S = 4\omega_0$).

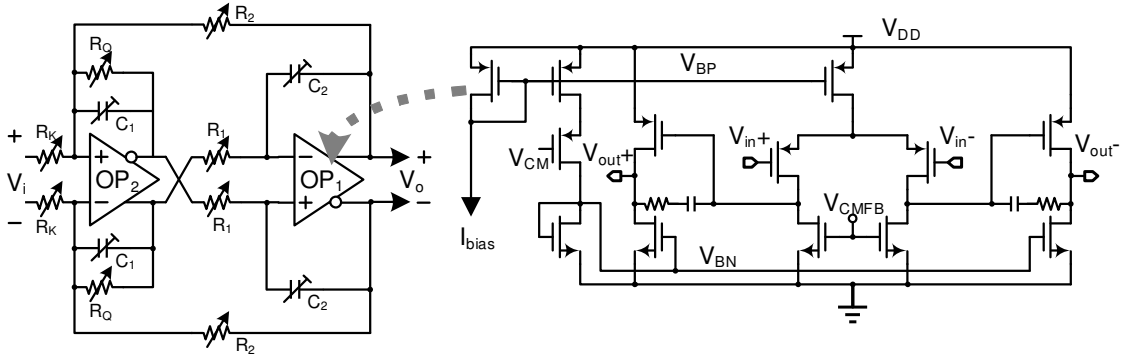


Figure 5.5: Schematic of the Tow-Thomas LPF biquad and the Miller-compensated amplifier.

The schematic of the active-RC biquad and the detailed amplifier design are demonstrated in Fig. 5.5, where V_{CM} is the common mode reference voltage and V_{CMFB} is the voltage from the common mode feedback (CMFB) loop. The HC-3² DTF improves the MA-3² architecture in [114]. Compared to the MA-3² DTF, HC-3² DTF uses more clock phases but fewer capacitors, which is more area efficient. The implementation of the HC-3² DTF is introduced later in Section 5.3.

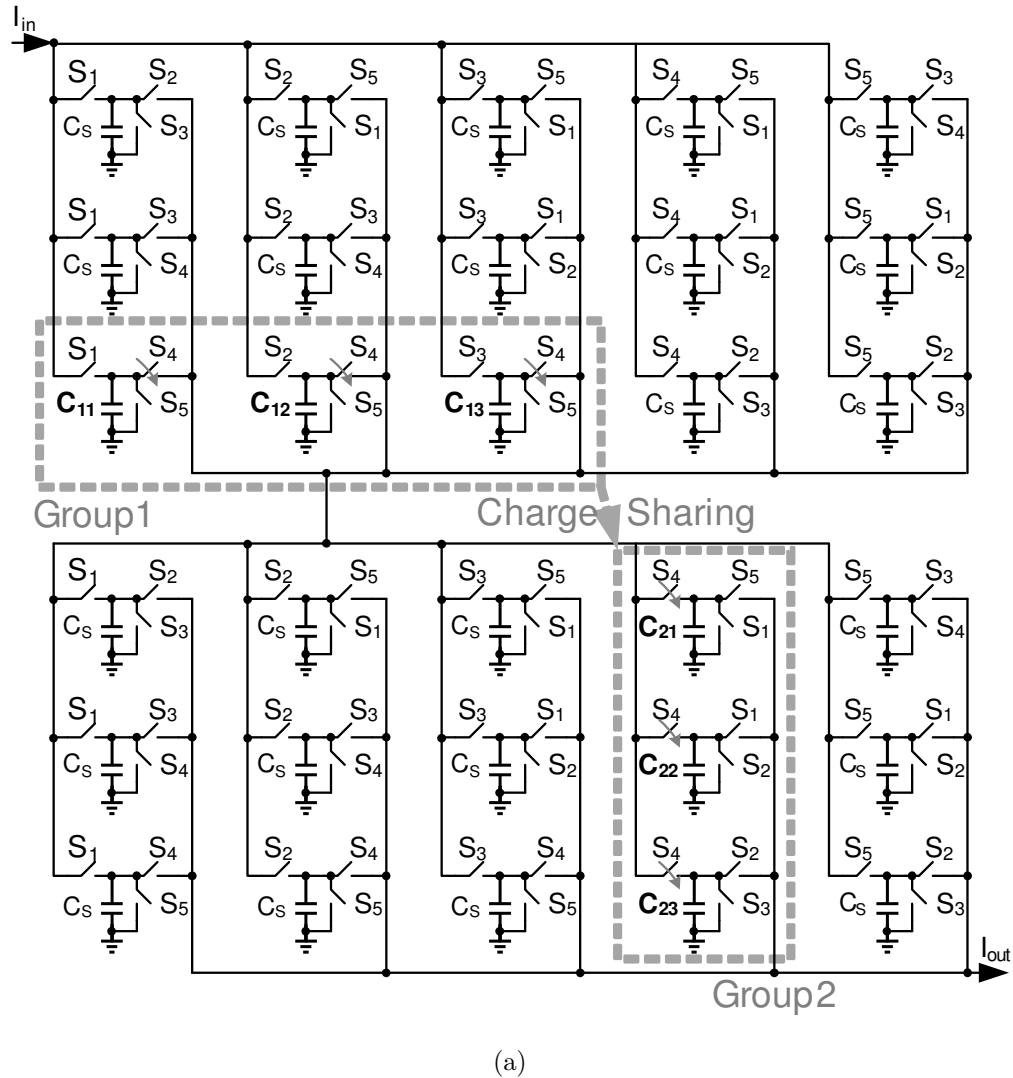
5.3 Evolutionary Charge-domain Filter Design

5.3.1 5-phase MA-3² Filter

[114] proposed a two-stage 3-tap switched capacitor (SC) moving average (MA-3²) filter. It has two identical stages, and the final z -domain transfer function can be derived as

$$H_{MA-3^2}(z) = \frac{1}{2} \times \frac{1}{3} (1 + z^{-1} + z^{-2}) \times \frac{1}{3} (1 + z^{-1} + z^{-2}) \quad (5.3)$$

where the $1/2$ attenuation is because of the charge transfer between two stages. In the real circuit implementation, multi-phase clocks are used to implement the unit delay



(a)

SW	S ₁	S ₂	S ₃	S ₄	S ₅
C ₁₁	I _{in}			Share	Reset
C ₁₂		I _{in}		Share	Reset
C ₁₃			I _{in}	Share	Reset
C ₂₁	Reset			Share	I _{out}
C ₂₂	I _{out}	Reset		Share	
C ₂₃		I _{out}	Reset	Share	

(b)

Figure 5.6: Minimum two-stage MA-3² filter with (a) filter architecture and (b) timing diagram for the five-phase MA-3² filter.

(z^{-1}). six-phase clocks are used in [114]. However, as discussed in [108], to achieve the three delays in (5.3), three sampling phases and one decimation (charge sharing) phase are required. In addition, for the type of current input DTF, an extra reset phase is needed. Therefore, five-phase clock is the minimum requirement needed to achieve the frequency response of (5.3). The minimum MA-3² filter is demonstrated in Fig. 5.6a. It consists of multiple SC cells, in which a capacitor is connected to three MOS switches to conduct different operations. Fig. 5.6b demonstrates the full five-phase clock cycles containing S_1, \dots, S_5 and the corresponding operations. S_i indicates that the switch is only activated (turned on) during the i -th clock period as shown in Fig. 5.2.

As shown in Fig. 5.6a, dashed boxes of Group 1 and Group 2 demonstrates one path of the charge transfer. The Group 1 array in the first stage contains C_{11} , C_{12} and C_{13} . C_{11} is charged (through I_{in}) during the clock phase S_1 . Then, C_{12} is charged during S_2 , and C_{13} is charged by S_3 . S_{i+1} is S_i delayed by a unit clock period, which is equivalent to z^{-1} . S_4 sums the charges of Group 1 and shares them to Group 2. The sum-and-share achieves the transfer function

$$H(z) = \frac{C_{13} + C_{12}z^{-1} + C_{11}z^{-2}}{C_{13} + C_{12} + C_{11}} \quad (5.4)$$

For the MA-3² filter, we have $C_{11} : C_{12} : C_{13} = 1 : 1 : 1$. Moreover, the charge sharing operation only transfers half of the charges from Group 1 to Group 2; thus, the proposed MA-3² filter has a 6 dB attenuation. After the S_4 sum-and-share operation, C_{11} , C_{12} and C_{13} in Group 1 can be reset in any phase before their next charging phases, such as S_5 shown in Fig. 5.6b. At the same time, the charges on C_{21} , C_{22} and C_{23} in Group 2 will be outputted (discharged through I_{out}) one-by-one to the next stage, followed by the reset operations. It should be noted that, except

for the sum-and-share S_4 phase, the Group 1 capacitors and Group 2 capacitors are isolated from each other. To sum up, a five-phase operation is the minimum setting for (5.3). Three charging/discharging phases are needed to achieve the three delays. The sum-and-share phase is adopted to transfer charges from the first stage to the second, and the reset phase is used to clear the capacitor before the next charging or sharing. In addition, the non-decimation operation principle requires the DTF to sample the input at all five clock phases, leading to five duplicated paths of Group 1 and Group 2. The total hardware cost of the minimum two-stage five-phase MA-3² filter is 30 capacitors plus 90 switches.

5.3.2 8-phase MA-3² Filter

On the one hand, as mentioned previously, the five-phase clock is the minimum configuration for the transfer function (5.3). In other words, more phases can be used for the moving average (MA) filter, although the increased phase number leads to more SC cells, which consume more area. A six-phase design has been demonstrated in [114]. On the other hand, Section 5.2.2 indicated that a 3-tap FIR HC architecture requires a fixed delay of $T_S/8$ between adjacent taps to achieve the notching frequencies, $3f_S/8$ and $5f_S/8$, precisely. Therefore, we take the first step and extend the MA-3² filter, which gives us eight phases, achieving the designated $T_S/8$ delay for every SC cell. Fig. 5.7 demonstrates the extended eight-phase MA-3² filter with clock phases S_1, \dots, S_8 . Similar to the five-phase configuration, Group 1 and Group 2 SC arrays are annotated to indicate a single path of the charge transfer. The corresponding timing diagram is also illustrated in Fig. 5.9a. S_1, S_2 and S_3 are assigned to charge C_{11}, C_{12} and C_{13} . S_5, S_6 and S_7 are used to discharge C_{21}, C_{22} and C_{23} . S_4 remains the sum-and-share operation for all six capacitors. Moreover, the increased number of phases makes it more flexible to allocate the reset phases, which should

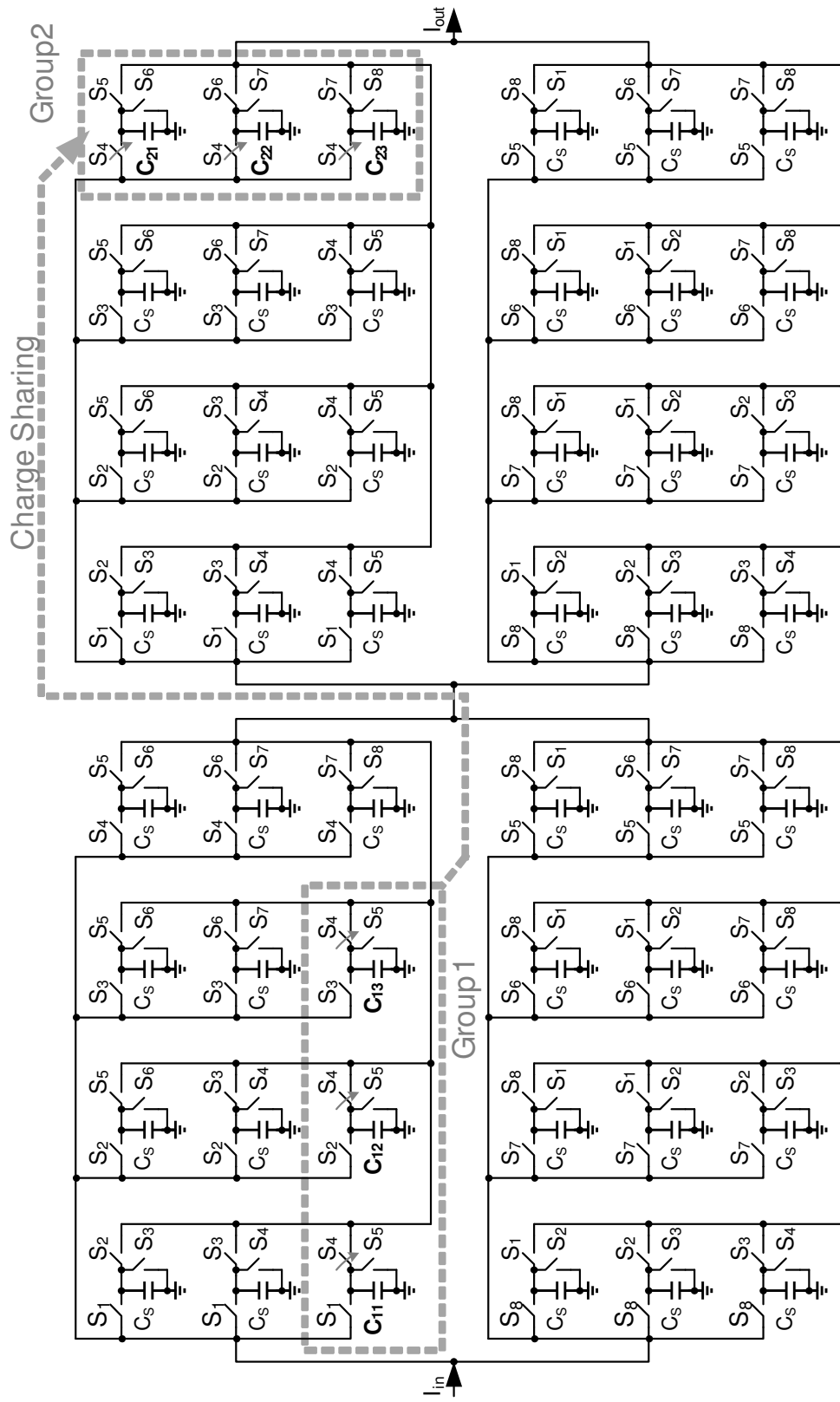


Figure 5.7: 8-phase MA-3² filter before the compaction.

be between S_4 and the corresponding charge/discharge phases. Fig. 5.9a shows an arrangement of the reset phases, although it is not the only scheme. This eight-phase MA-3² filter consists of 48 capacitors and 144 switches.

5.3.3 Stage Compaction and HC-3² Filter

Although the eight-phase MA-3² filter (Fig. 5.7) takes up more area than the five-phase filter (Fig. 5.6a), it is one important step towards a more compact design. First of all, moving all capacitors' reset phases to S_8 will not change the behavior of this filter. After this change, S_1 , S_2 and S_3 are assigned to solely charging operations (I_{in}), which has no impact on Group 2 capacitors. Meanwhile, S_5 , S_6 , and S_7 are associated with only discharging operations (I_{out}), during which Group 1 is kept intact. To the contrary, S_4 and S_8 now involves all capacitors, regardless of the group. This means that Group 1 capacitors can be reused for S_5 , S_6 , and S_7 , and the second stage can be removed. As a result, the eight-phase DTF architecture after compaction is demonstrated in Fig. 5.8, and the modified timing diagram is shown in Fig. 5.9b. Furthermore, harmonic cancellation technique is adopted to construct the proposed HC-3² DTF. $C_{11} : C_{12} : C_{13} = 1 : 17/12 : 1$ is chosen to achieve the specific tone suppression frequency responses illustrated in Fig. 5.4. This ratio is maintained for all eight duplicated groups, which are indicated by two different capacitor sizes in Fig. 5.8. In the circuit design, the smaller capacitor is 302 fF with a 12 μm x 12 μm area, while the larger one is 427 fF with a 17 μm x 12 μm area. To sum up, the proposed compact eight-phase HC-3² DTF consumes 24 capacitors and 96 MOS switches, which is even smaller than the “minimum” five-phase MA-3² solution. Different types of DTF and the corresponding hardware costs are summarized and compared in Table 5.1. It should be mentioned that the proposed HC-3² filter has no attenuation, compared to the 6 dB attenuation in MA-3² filter due to the charge

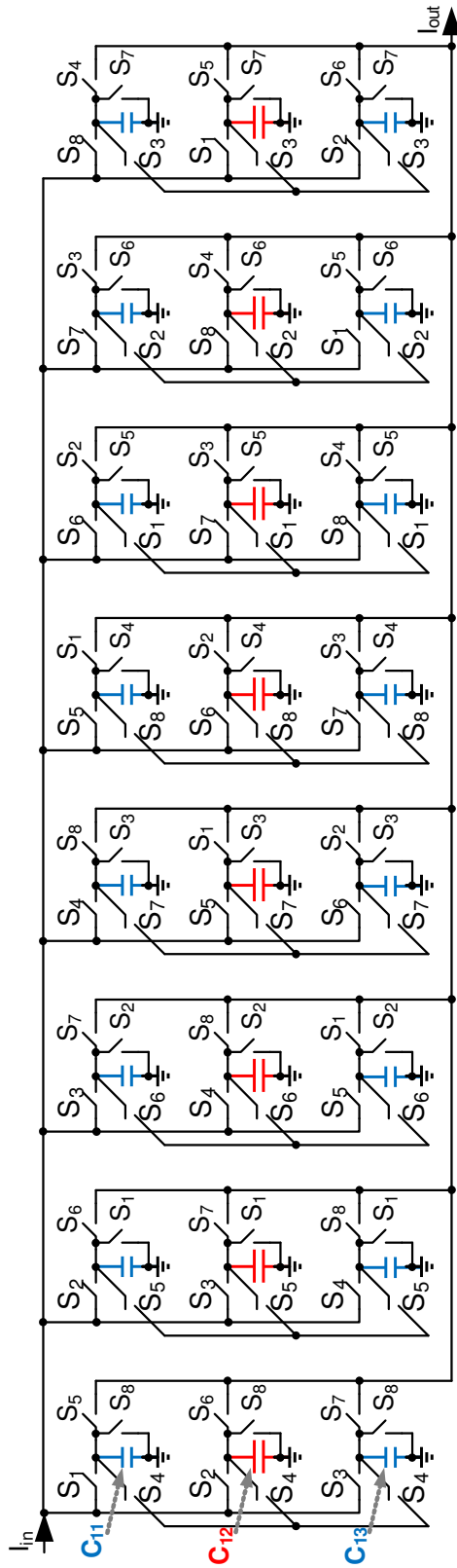


Figure 5.8: Compacted 8-phase HC-3² filter with weighted capacitors ($C_{11} : C_{12} : C_{13} = 1 : 17/12 : 1$).

SW	S ₁	S ₂	S ₃	S ₄	S ₅	S ₆	S ₇	S ₈
C ₁₁	I _{in}			Share	Reset			
C ₁₂		I _{in}		Share	Reset			
C ₁₃			I _{in}	Share	Reset			
C ₂₁				Share	I _{out}	Reset		
C ₂₂				Share		I _{out}	Reset	
C ₂₃				Share			I _{out}	Reset

(a)

SW	S ₁	S ₂	S ₃	S ₄	S ₅	S ₆	S ₇	S ₈
C ₁₁	I _{in}			Share	I _{out}			Reset
C ₁₂		I _{in}		Share		I _{out}		Reset
C ₁₃			I _{in}	Share			I _{out}	Reset
C₂₁								
C₂₂								
C₂₃								

(b)

Figure 5.9: Timing diagram for DT filters: (a) Eight-phase MA-3² filter and (b) compacted 8-phase HC-3² filter

being shared to the capacitors in the second stage.

Table 5.1: Hardware cost comparison of the DTF architectures

Architecture	No. of Capacitors	No. of Switches
Five-phase MA-3 ²	30	90
Six-phase MA-3 ² [114]	36	108
Eight-phase MA-3 ²	48	144
Eight-phase compact HC-3 ²	24	96

5.4 Design Procedure

5.4.1 Operational Amplifier

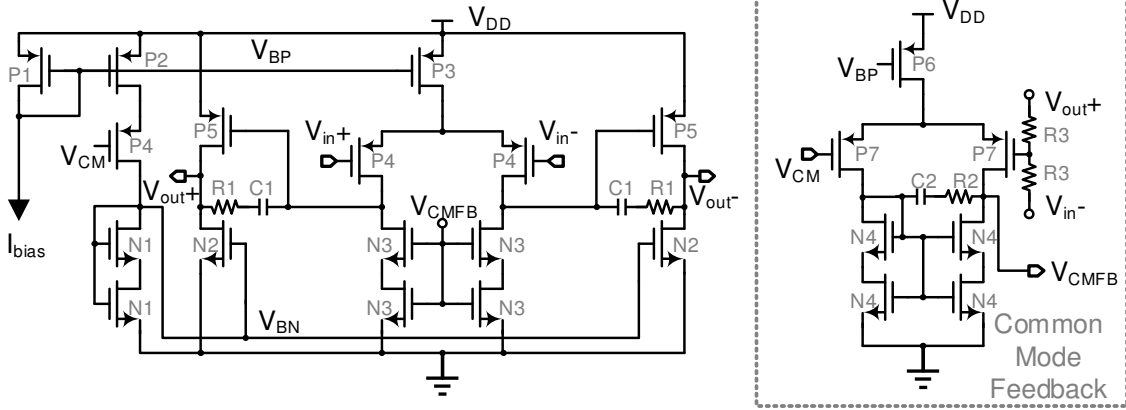


Figure 5.10: Detailed schematic of the the Miller-compensated amplifier.

To demonstrate the proposed hybrid baseband chain, an active-RC CT LPF biquad is implemented. For simplicity, each amplifier in the biquad adopts the differential miller-compensated two-stage architecture, which is illustrated in Fig. 5.10. It should be mentioned that, the bandwidth of the programmable LPF is designed up to 30 MHz. At least 10 times, 300 MHz, is required for the amplifiers' gain-bandwidth product (GBW). We design the amplifier with some margin, and thus extend the GBW to 500 MHz. In addition, the supply voltage is 1.2 V, and we assign each amplifier with 1 mA current budget.

Firstly, judiciously assign currents to each branch, considering $I_{bias} = 12 \mu\text{A}$. Set the first-stage-to-second-stage current ratio to 1:16, and the common mode feedback

(CMFB) amplifier consumes half current of the first stage. Therefore, we have

$$\left(\frac{W}{L}\right)_{P1} : \left(\frac{W}{L}\right)_{P2} : \left(\frac{W}{L}\right)_{P6} : \left(\frac{W}{L}\right)_{P3} = 1 : 2 : 2 : 4 \quad (5.5)$$

To reduce the area occupied by the NMOSFETs, self-cascode structure [115] is adopted to achieve high current factor ($I_{N1} : I_{N2} = 1 : 16$),

$$\left(\frac{W}{L}\right)_{N1} : \left(\frac{W}{L}\right)_{N2} : \left(\frac{W}{L}\right)_{N3} = 1 : 8 : 1 \quad (5.6)$$

As a result, the total current drawn by an amplifier is $876 \mu\text{A}$, which is below the current budget. The current for the first stage is $48 \mu\text{A}$, and the second stage draws $768 \mu\text{A}$, respectively. At the very beginning, the CMFB feedback is disconnected. Instead, V_{CMFB} is connected to V_{BN} , and N3 is set to the same size as that of N1. $L = 0.6 \mu\text{m}$ is set for all current mirror transistors, including P1, P2, P3, P6, N1, N2, and N3, to obtain accurate current copy. For the IBM 130 nm process, the mobility of an NMOSFET is $440 \text{ cm}^2/\text{V} \cdot \text{s}$, and the mobility of a PMOSFET is $94 \text{ cm}^2/\text{V} \cdot \text{s}$, which is approximately 1/4 of that of the NMOSFET. Therefore, $W_{P1} = 1.2 \mu\text{m}$ and $W_{N1} = 1.2 \mu\text{m}$ are set as the start points. The other current mirror transistor sizes can be obtained proportionally.

Moreover, to design the biquad using the ideal transfer function, the amplifier should have a high DC gain. Assume the total gain of the Miller-compensated amplifier is 50 dB. Distribute it between the two stages,

$$\begin{aligned} A_{v1} &= g_{m,P4} \cdot (r_{ds3} || r_{ds,P4}) = 30 \text{ dB} \\ A_{v2} &= g_{m,P5} \cdot (r_{ds,N2} || r_{ds,P5}) = 20 \text{ dB} \end{aligned} \quad (5.7)$$

where r_{ds3} is the equivalent output impedance of the cascode N3 transistors. To

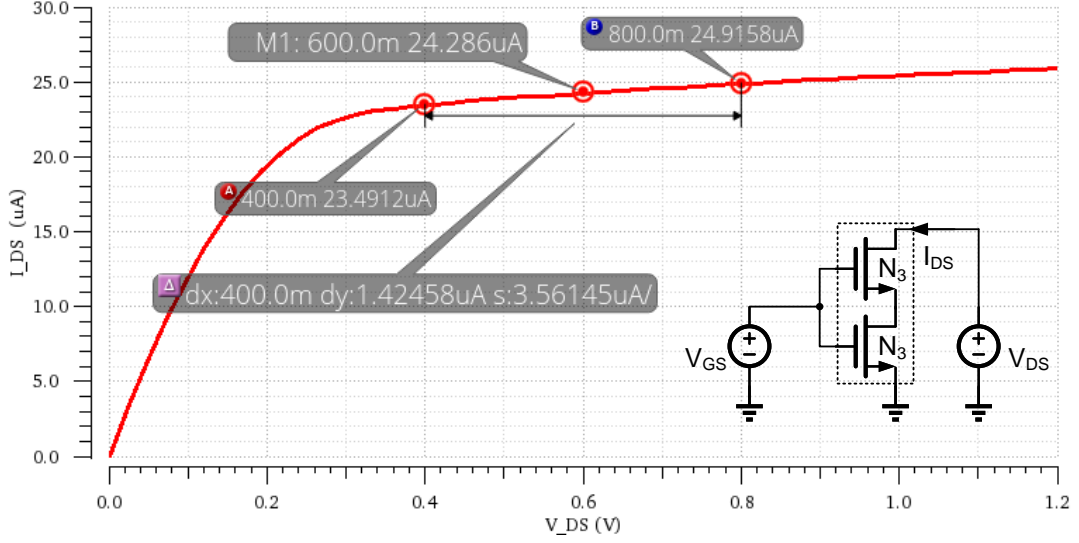


Figure 5.11: I_{DS} vs V_{DS} simulation result for N3 cascode transistors.

obtain the corresponding r_{ds} , the curve of I_{DS} versus V_{DS} is simulated. In this design, the current passing through N3 is $24 \mu\text{A}$. V_{GS} is first set to match the branch current at about $V_{DS} = 600 \text{ mV}$. The test bench schematic and the simulation result are shown in Fig. 5.11. Thus, we have $r_{ds3} \approx 280 \text{ k}\Omega$. Similarly, the current passing through N2 is $384 \mu\text{A}$, and the simulation result gives $r_{ds,N2} \approx 9.65 \mu\Omega$. Finally, we have $g_{m,P4} \approx 1.13 \times 10^{-4} \Omega^{-1}$ and $g_{m,P5} \approx 1.04 \times 10^{-3} \Omega^{-1}$. The transistor sizes can be calculated,

$$g_m = \sqrt{2\mu C_{ox} \left(\frac{W}{L}\right) I_{DS}} \quad (5.8)$$

where $\mu = 94 \text{ cm}^2/\text{V} \cdot \text{s}$ and $C_{ox} = \epsilon_{SiO_2}/t_{ox} = 1.03 \times 10^{-2} \text{ F}/\text{m}^2$, and I_{DS} is the bias current. Assuming $r_{ds3} = r_{ds,P4}$ and $r_{ds,N2} = r_{ds,P5}$ for (5.7), $(W/L)_{P4} \approx 6$ and $(W/L)_{P5} \approx 30$ can be used as the start point. $L = 0.36 \mu\text{m}$ is set for P4 and P5. In addition, R3 resistors are added to measure the output common mode voltage. R3 should be set 10 times of $r_{ds,N2} || r_{ds,P5}$, so as to avoid the loading effect. R3 is set to

48 $k\Omega$ here.

As mentioned previously, 500 MHz GBW is the design target. the capacitance value of C_1 can be derived from

$$GBW \cdot 2\pi = \frac{g_{m,P4}}{C_1} = 500MHz. \quad (5.9)$$

$C_1 = 36$ fF is relatively small and thus we'll try increasing $g_{m,P4}$ in the following iterative design steps.

The initial design of the CMFB amplifier simply copies half of the first stage in the main amplifier. In other words, the sizes of N4, P6 and P7 are half of the sizes of N3, P3 and P4, respectively. Later, R1 is added to improve the phase margin of the main amplifier. V_{CMFB} is connected as shown in Fig. 5.10. C2 and R2 are also added to induce a zero in the CMFB (a.k.a. DC offset cancellation (DCOC)) loop, making the loop more stable.

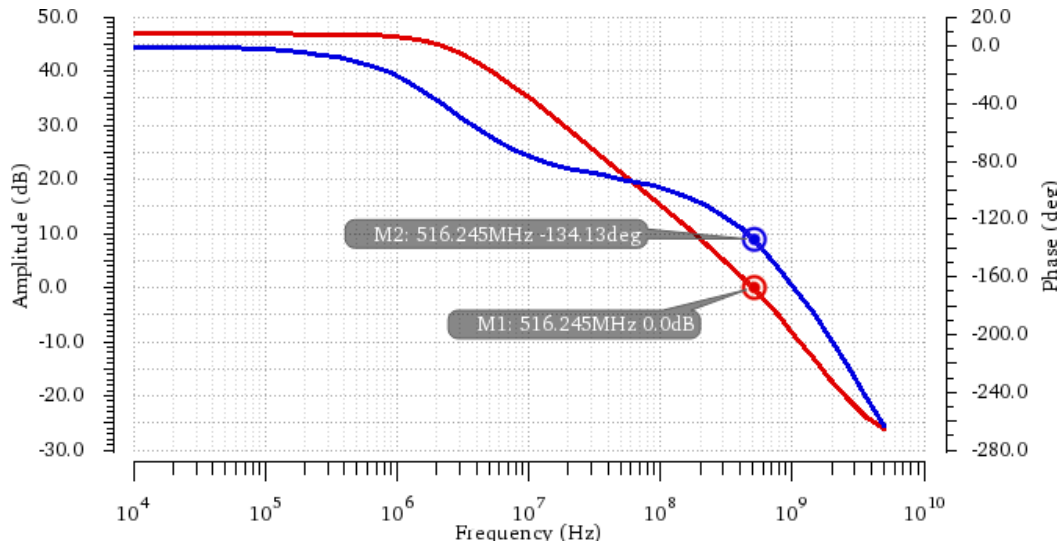


Figure 5.12: Simulated transfer function of the miller-compensated amplifier.

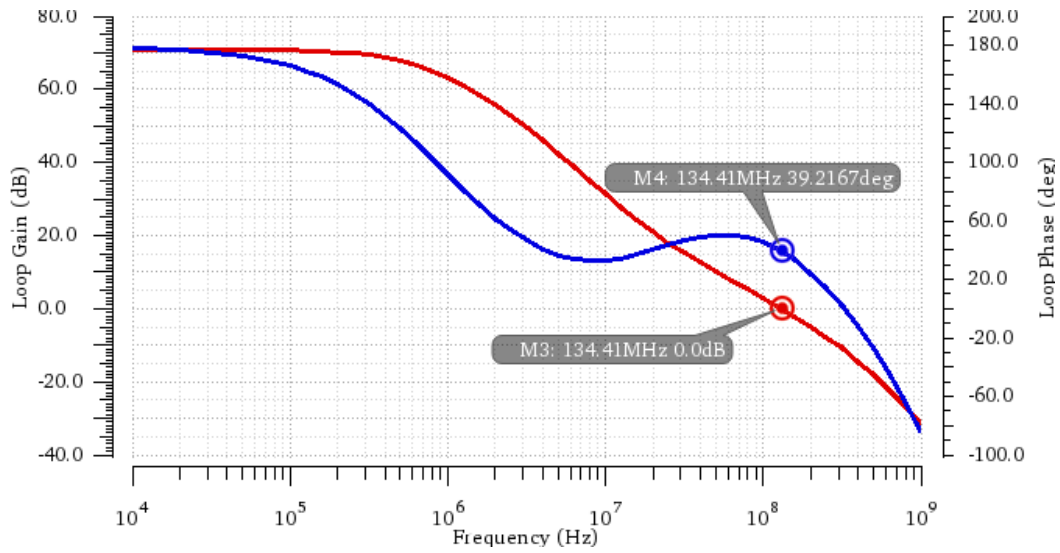


Figure 5.13: Simulated transfer function of the amplifier’s DCOC loop.

After several iterative design steps, the final simulated transfer function of the amplifier is shown in Fig. 5.12. The DC gain is 47 dB, the gain-bandwidth product (GBW) is around 516 MHz, and the phase margin is 46°. The stability of the DCOC loop is also simulated and its transfer function is illustrated in Fig. 5.13. The DC gain of the DCOC loop is 70 dB, having a bandwidth of 134 MHz and a loop phase of 39°. Simulation results are summarized in Table 5.2. Table 5.3 summarizes the final design parameters for all components.

Table 5.2: Summary of the Miller-compensated amplifier parameters

Parameter	Value
DC gain	47 dB
Gain-bandwidth	516 MHz
Phase margin	46 deg
DCOC loop gain	70 dB
Loop gain-bandwidth	134 MHz
Loop phase	39 deg

Table 5.3: Design parameter of the Miller-compensated amplifier

Comp.	Parameter (W/L)	Value	Comp.	Parameter (W/L)	Value
P1	1.6 $\mu\text{m}/0.6 \mu\text{m}$		N2	12.8 $\mu\text{m}/0.5 \mu\text{m}$	
P2	3.2 $\mu\text{m}/0.6 \mu\text{m}$		N3	3.2 $\mu\text{m}/0.5 \mu\text{m}$	
P3	6.4 $\mu\text{m}/0.6 \mu\text{m}$		N4	1.6 $\mu\text{m}/0.5 \mu\text{m}$	
P4	8 $\mu\text{m}/0.5 \mu\text{m}$		R1	0.4 $\mu\text{m}/0.8 \mu\text{m}$	1.03 k Ω
P5	12.8 $\mu\text{m}/0.36 \mu\text{m}$		C1	8.5 $\mu\text{m}/5.24 \mu\text{m}$	95.6 fF
P6	3.2 $\mu\text{m}/0.6 \mu\text{m}$		R2	0.2 $\mu\text{m}/4.5 \mu\text{m}$	9.69 k Ω
P7	2 $\mu\text{m}/0.5 \mu\text{m}$		C2	11.5 $\mu\text{m}/20 \mu\text{m}$	481.4 fF
N1	1.6 $\mu\text{m}/0.5 \mu\text{m}$		R3	0.4 $\mu\text{m}/9 \mu\text{m}$	34.3 k Ω

5.4.2 Programmable CT Filter

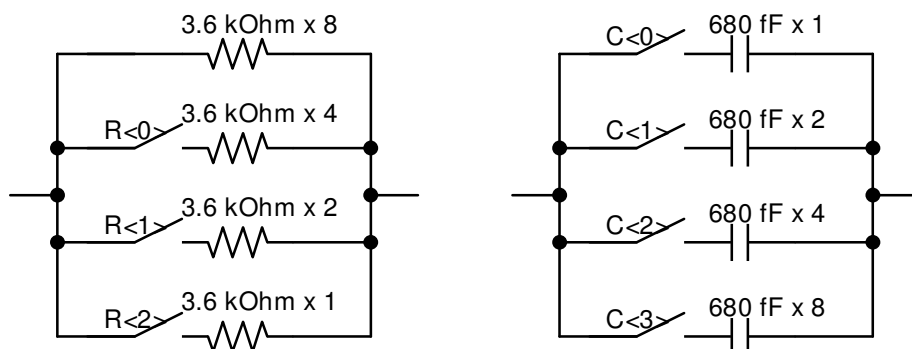


Figure 5.14: Configuration of the capacitor and resistor arrays.

The LPF biquad architecture has been demonstrated in Fig. 5.5. Its transfer function is

$$H_{CTF}(s) = \frac{V_{out}}{V_{in}} = \frac{G_{LPF}}{s^2 + \frac{\omega_0}{Q}s + \omega_0^2} \quad (5.10)$$

where the DC gain, the bandwidth, and the quality factor are defined as

$$G_{LPF} = -\frac{R_Q}{R_K}, \quad \omega_0 = (R_1 R_2 C_1 C_2)^{-1}, \quad Q = \sqrt{\frac{R_Q^2 C_1}{R_1 R_2 C_2}} \quad (5.11)$$

In this design, $Q = 1$, we have $R_K = R_Q = R_1 = R_2 = R$ and $C_1 = C_2 = C$. Therefore, all resistor arrays reuse the same design. Each array R has three control bits. The unit resistor is selected equal to the output impedance of the Miller-compensated amplifier, $3.6 \text{ k}\Omega$, and thus it will not induce a significant loading effect even the biquad is set to the maximum bandwidth. The four capacitor arrays are also identical, and each has four control bits. The configuration of the resistor and the capacitor arrays are shown in Fig. 5.14. The programmable RC time-constant can cover the bandwidth from 500 KHz to 30 MHz. Finally, the layout of the 2nd order Tow-Thomas LPF biquad is demonstrated in Fig. 5.15.

5.4.3 HC-3² DT Filter

The proposed HC-3² DT Filter consists of a 8-phase clock generator and a switch-capacitor (SC) array. The 8-phase clock generator is similar to the 24-phase version introduced in Fig. 4.20.

One critical design issue of the proposed DTF is how to select the capacitance value of the SC array at the beginning. We can find that the total sampling capacitance of the HC-3² DTF during one clock phase ($C_{11} + C_{12} + C_{13}$ in Fig. 5.8) is equal to the filter's load, C_L . Consider the CT LPF has a maximum output swing (ΔA) of 0 dBm (632 mVpp) and a maximum bandwidth (f_{3dB}) of 30 MHz. The maximum derivative of the CT LPF's output sinusoidal waveform is $2\pi f_{3dB} \cdot \Delta A$. Assuming all the current of the amplifier's second stage, $I_{ds,P5} = 384 \mu\text{A}$, is used to charge C_L , to

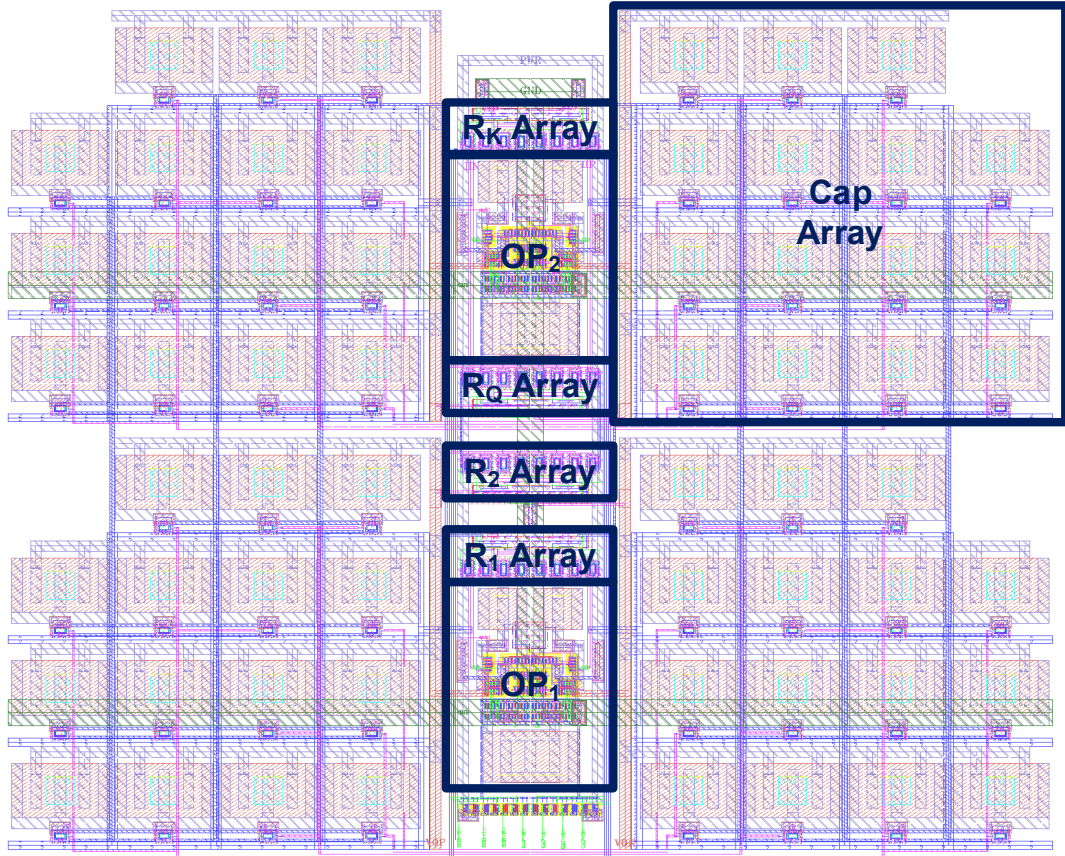


Figure 5.15: Layout of the CT LPF biquad.

avoid a significant in-band distortion, we can derive the slew rate constraint,

$$SR = \frac{I_{ds,P5}}{C_L} \geq 2\pi f_{3dB} \cdot \Delta A \quad (5.12)$$

As a result, $C_L \leq 3.2$ pF is the upper limit for the capacitance value.

Furthermore, the schematic of the switched capacitor array is illustrated in Fig. 5.8 and the corresponding layout is demonstrated in Fig. 5.16. In the layout, each row represents one column in Fig. 5.8. To implement the differential circuitry, the ca-

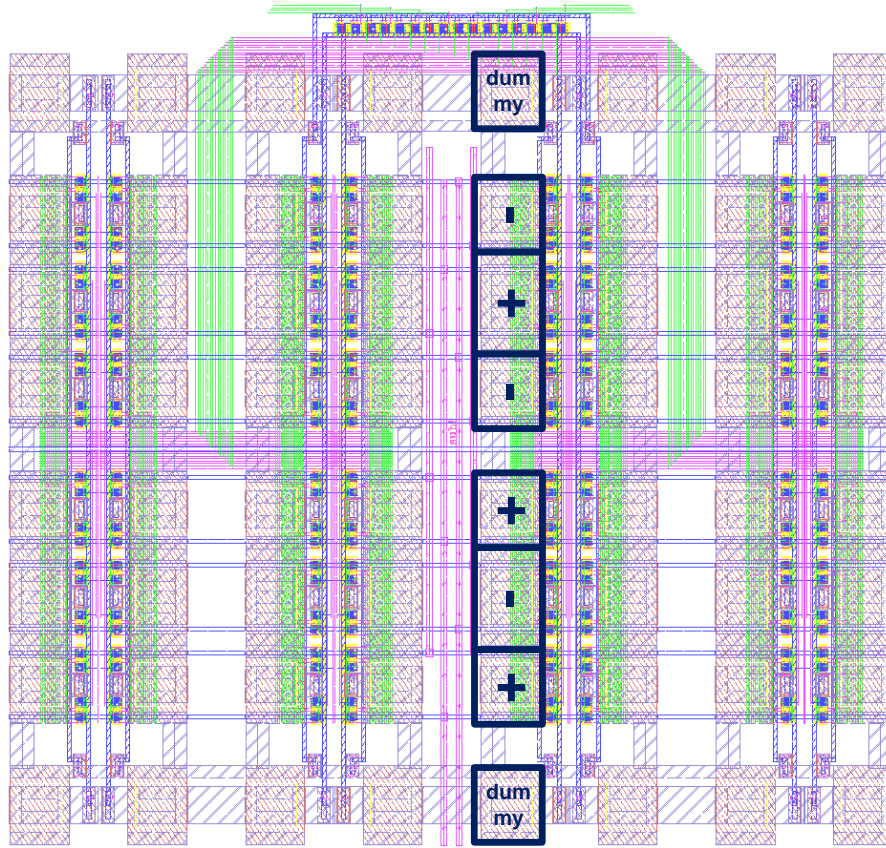


Figure 5.16: Layout of the HC-3² DT Filter.

capacitor number is doubled. Common-centroid layout pattern is used as indicated by “+” and “-”, representing the positive and the negative half respectively.

It should be mentioned that, in this design, the Gm stage between the CT LPF and the HC-3² DTF is not implemented for the sake of simplicity and low area overhead. As shown in Fig. 5.17a, in this design, the HC-3² DTF directly samples and holds the output of the CT LPF, while a PGA is inserted into the chain of [114] (Fig. 5.17b), playing the role of a 1st order integration. This Gm stage is an

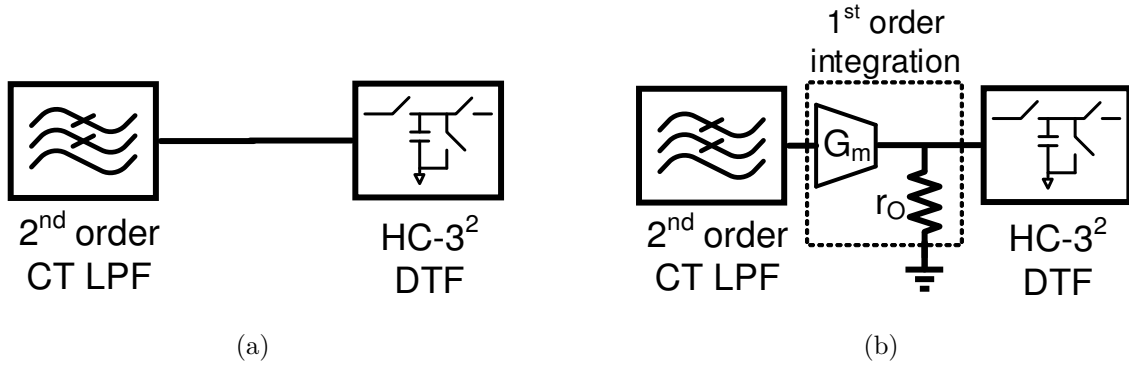


Figure 5.17: Hybrid baseband chain architecture (a) without the integration stage or (b) with the integration stage.

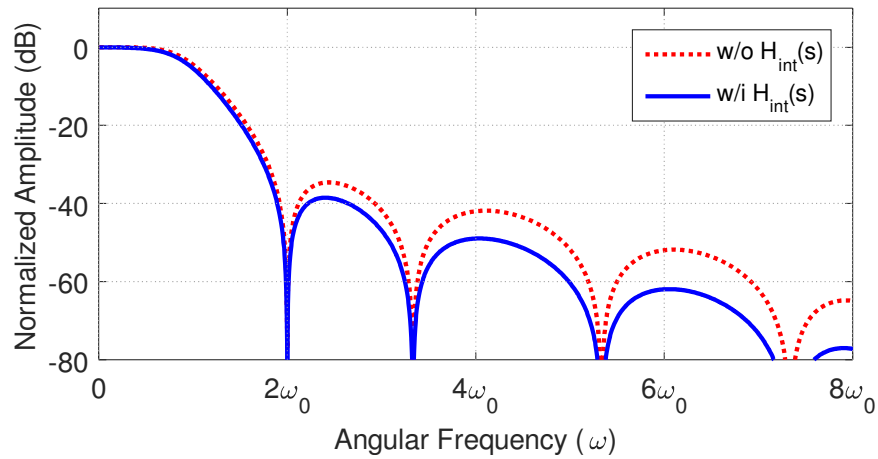


Figure 5.18: Comparison of the frequency responses between the baseband chains without and with the integration stage.

additional LPF transfer function applied to the baseband chain,

$$H_{int}(s) = \frac{g_m r_O}{s r_O C_L + 1} \quad (5.13)$$

where r_O is the output impedance of the G_m stage. This integration stage can

further suppress the high-frequency interference. If $g_m r_O = 1$ and $(r_O C_L)^{-1} = 2\omega_0$, the baseband chain frequency responses without (the same as Fig. 5.4a) and with the integration stage are compared in Fig. 5.18. However, this extra Gm stage suffers from the degraded sinc-type function due to the limited r_O [116].

5.5 Analysis of Measurement Precision

5.5.1 Notching Degradation

5.5.1.1 Capacitance Mismatch

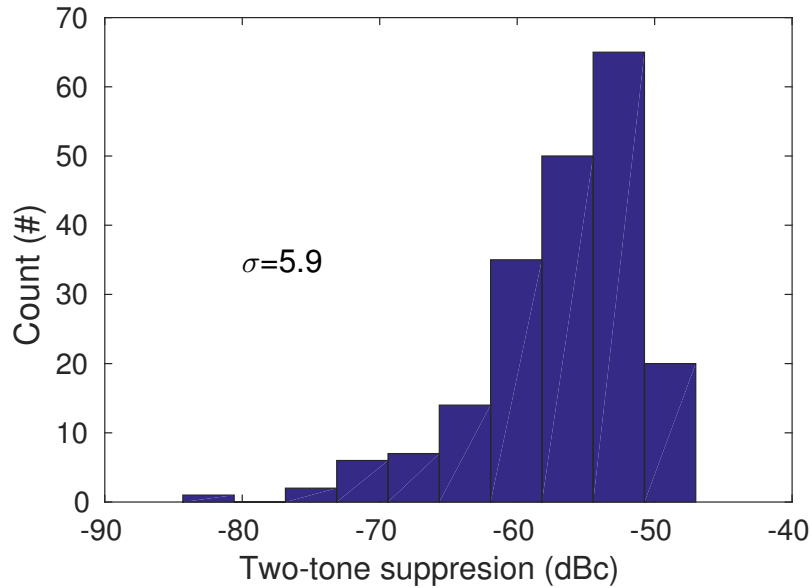


Figure 5.19: Monte-Carlo simulated small-signal two-tone suppression at $3f_0$ and $5f_0$ ($f_s = 8f_0$).

Even after the compaction procedure proposed in Section 5.3.3, there still remains 24 capacitors in the proposed HC-3² DTF architecture. Capacitance matching is sensitive to the PVT variations. Mismatches will lead to the finite attenuation at

the notching frequency points as discussed in [44]. To evaluate the impact of the PVT variations, two-hundred runs of Monte-Carlo simulation are carried out to obtain the frequency responses of the proposed HC-3² DTF. Attenuation at the two notching frequency points, $3f_0$ and $5f_0$, is examined, where the sampling frequency $f_S = 8f_0 = 8$ MHz. Small-signal simulation results are depicted in Fig. 5.19. An average of 56.6 dBc attenuation can be observed from the higher residual tone, either at $3f_0$ or $5f_0$.

5.5.1.2 DTF Sampling Frequency

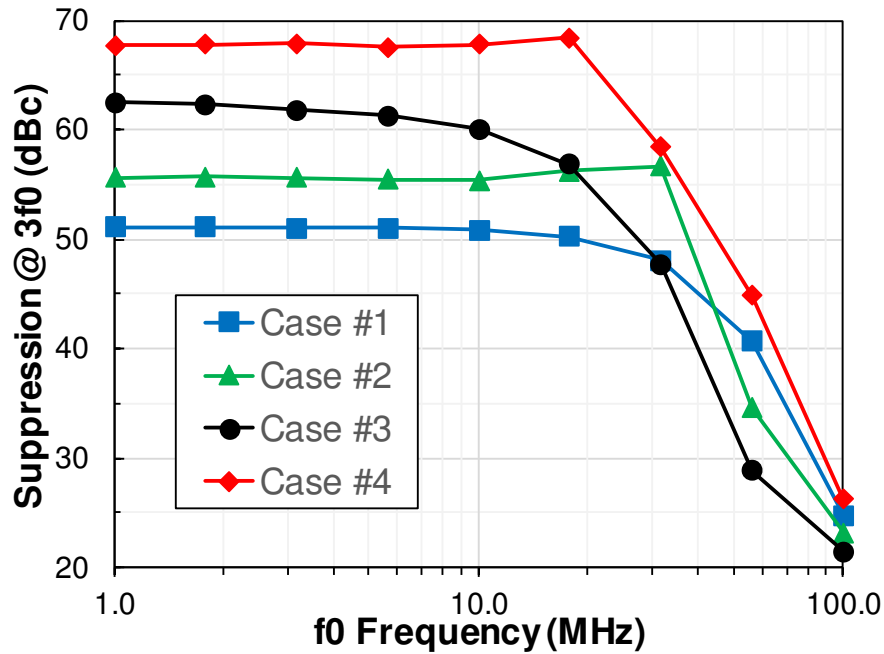


Figure 5.20: Simulated small-signal tone suppression at $3f_0$ versus the sampling frequency, $f_S = 8f_0$ is swept from 8 MHz to 800 MHz. Four cases were arbitrarily picked from the Monte-Carlo simulation results.

The sampling frequency will also affect the notching characteristic of the proposed HC-3² DTF. Four matching conditions of the capacitors are arbitrarily picked from

the aforementioned Monte-Carlo simulations, and the sampling frequency is swept from 8 MHz to 800 MHz. The small-signal simulation results are shown in Fig. 5.20. We can find that the filter's attenuation at $3f_0$ starts degrading when $f_0 = 10$ MHz ($f_S = 80$ MHz). The higher the sampling frequency f_S is the lower the suppression is.

5.5.2 Measurement Precision

To obtain the IP3 point, the conventional IP3 measurement method adopts the extrapolation, which is demonstrated in Fig. 5.21a. Two test tones are injected into the circuit-under-test (CUT), and the input power (P_{IN}) is swept. The CUT's output is connected to a spectrum analyzer to read out the output power (P_{OUT}) of the test tones and the IM3 tones. At low P_{IN} , the power increment of the IM3 tones is roughly three times of that of the test tones (with a slope of k). Draw extended lines across the measured points at slopes of k and $3k$ separately. These two lines will intercept at the 3rd order interception point (IP3).

A similar procedure can be applied to the power detector (PD) based measurement method proposed in Section 5.2.3. However, as discussed in Section 5.5.1, the notching of the proposed HC-3² DTF is not ideal; thus, the suppression on the two test tones are limited. The impact of the limited suppression is demonstrated in Fig. 5.21b for the proposed linearity test methodology. When P_{IN} is low, the PD output is dominated by the residual test tone power but not the IM3 tone power. Therefore, the interception method should be modified accordingly. We cannot use measurement data points to extend the $3k$ -slope line. Instead, draw a straight line tangent to the PD output curve at the point X , where the PD curve's slope is closest to $3k$.

On the one hand, we can find in Fig. 5.21a that when P_{IN} increases, the IM3

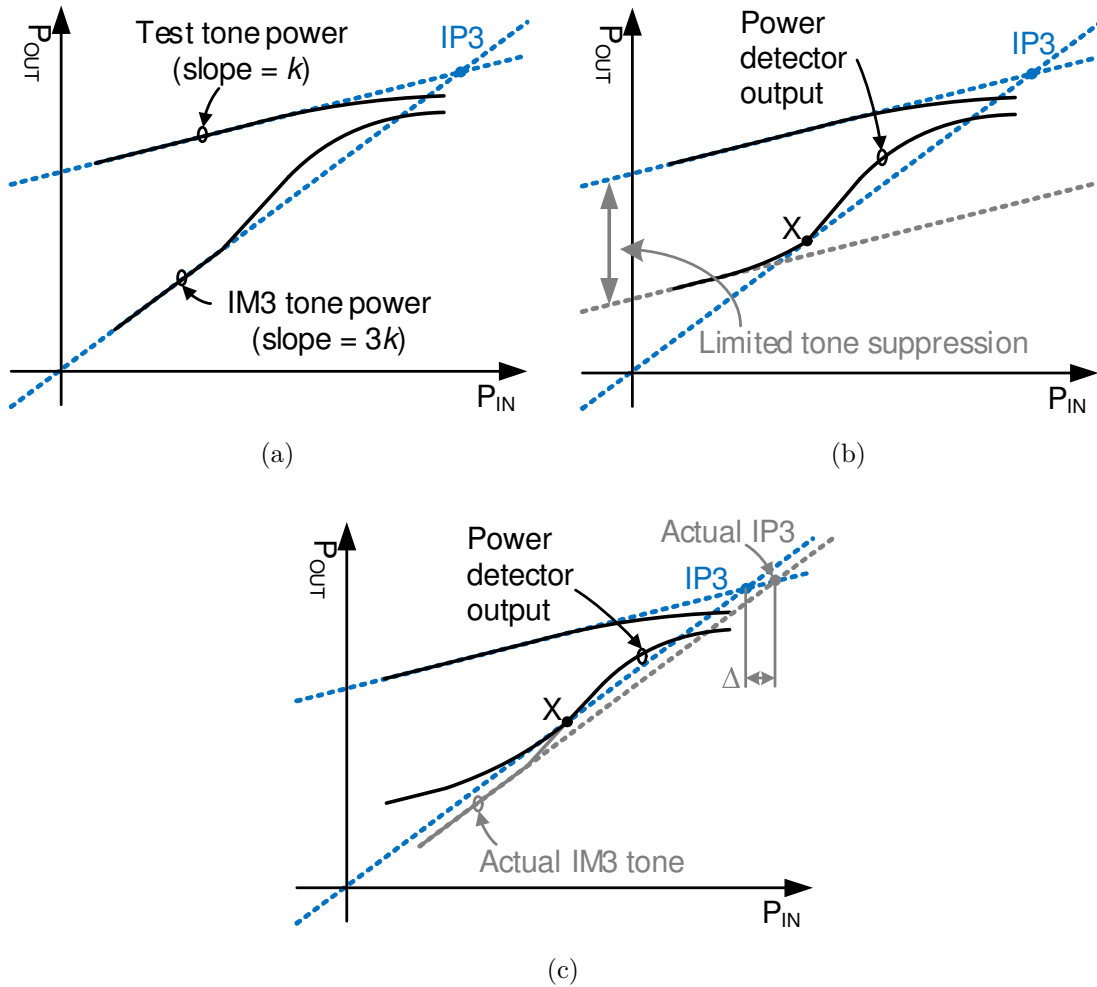


Figure 5.21: IP3 measurement methodology: (a) Conventional measurement method, (b) PD-based measurement method, and (c) PD-based measurement induced error.

tone power doesn't follow the $3k$ slope any more. The distortion pattern may be different, depending on the different CUT architectures. On the other hand, based on the power detector output, the IM3 tone power may overtake the residual test tone power after the former has started distorting, as illustrated in Fig. 5.21c. This induced an error Δ in the interception method. It should be noticed that, the higher the actual IP3 value is, the larger the Δ is. Although it cannot guarantee an absolute

IP3 value as accurate as that obtained by a spectrum analyzer, the proposed PD-based IP3 measurement method can still be used as an indicator of the linearity performance. Higher measured value means better linearity and vice versa. Later in Section 5.6, the experimental results will show its application in the linearity optimization.

5.6 Experiment Results

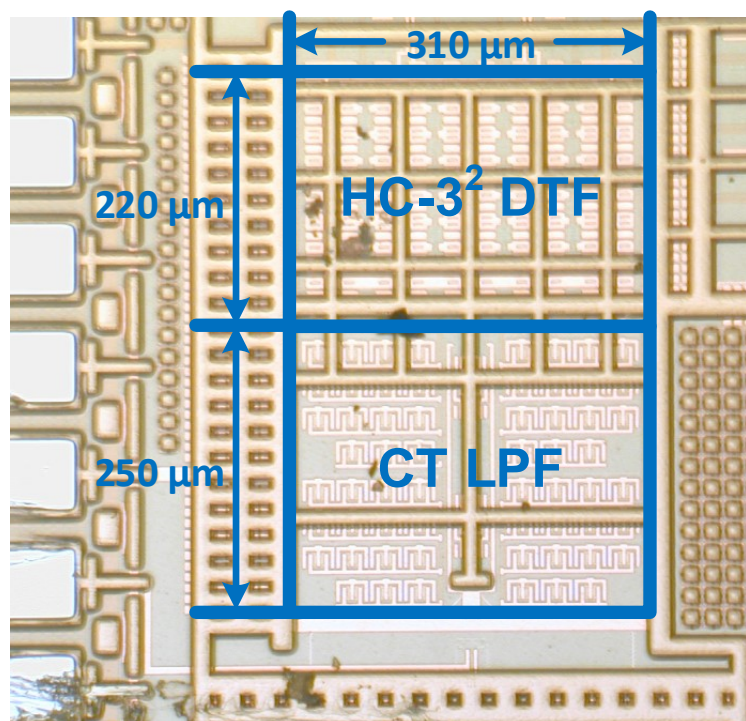


Figure 5.22: Die photograph of the proposed baseband chain.

To verify the concept, one path of the proposed CT+DT baseband chain was fabricated in 130 nm standard CMOS technology. The chip die micrograph is shown

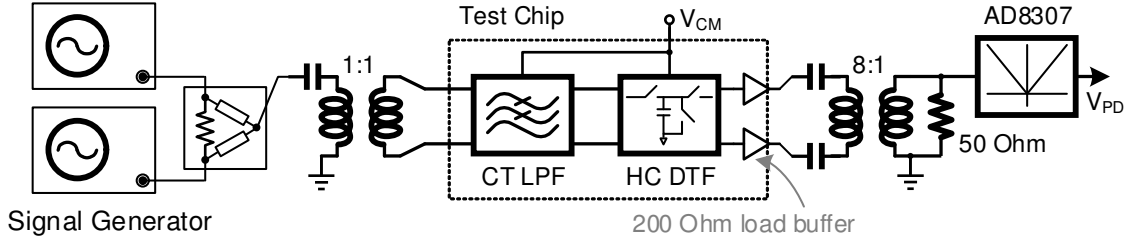


Figure 5.23: Configuration of the chip test bench.

in Fig. 5.22. The total silicon area is 0.146 mm^2 , in which the HC- 3^2 DTF occupies $310 \mu\text{m} \times 220 \mu\text{m}$, and the programmable CT LPF occupies $310 \mu\text{m} \times 250 \mu\text{m}$. The chip test bench configuration is demonstrated in Fig. 5.23. The common mode voltage V_{CM} is connected outside the chip so as to achieve the tunability of the baseband linearity performance. A logarithmic amplifier AD8307, which maps the input power (in dBm scale) to the output voltage V_{PD} (in voltage scale), is the power detector.

The baseband chain is configured to a bandwidth (f_0) of 6 MHz. Test signals with two tones, located at 3.6 MHz and 6 MHz, are injected into the CT LPF. The DT filter is first set in the normal operation mode, applying the sampling clock at $16f_0/3 = 32 \text{ MHz}$. The output spectrum of the baseband chain is shown in Fig. 5.24a. Then the in-band IIP3 measurement puts the DT filter in tone-suppression mode. The sampling clock is reduced to $8f_0/5 = 9.6 \text{ MHz}$, and the corresponding measured spectrum is demonstrated in Fig. 5.24b. The suppression on the two test tones is about 40 dB.

Detailed IIP3 measurement is conducted by sweeping the input test tone power from -15 to 2 dBm . Experimental results at $V_{CM} = 400 \text{ mV}$ are shown in Fig. 5.25. Fig. 5.25a demonstrates the power values measured by the spectrum analyzer and the measured in-band input IP3 (IIP3) is $+15.5 \text{ dBm}$, as indicated in the plot. Cor-

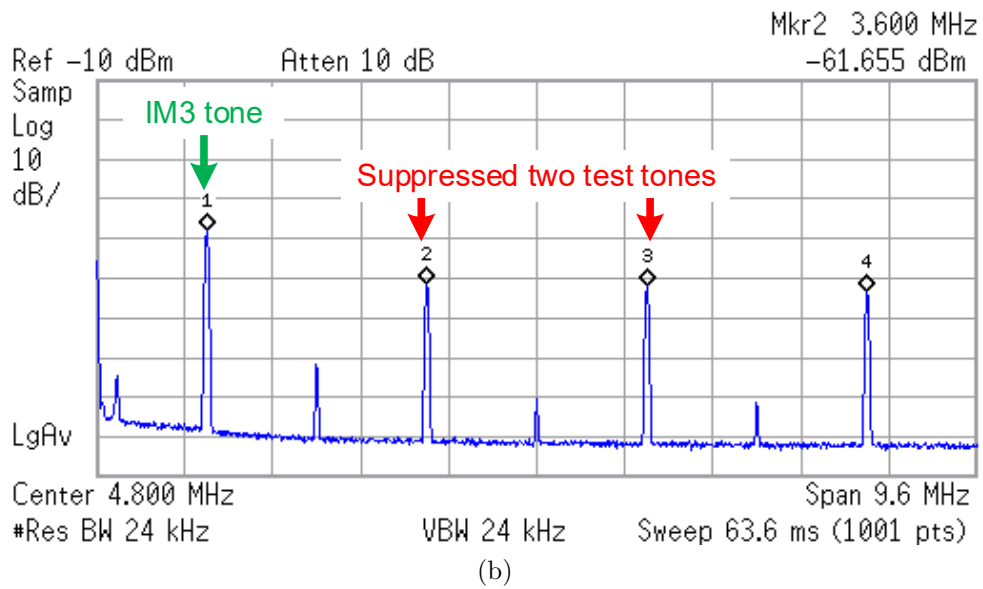
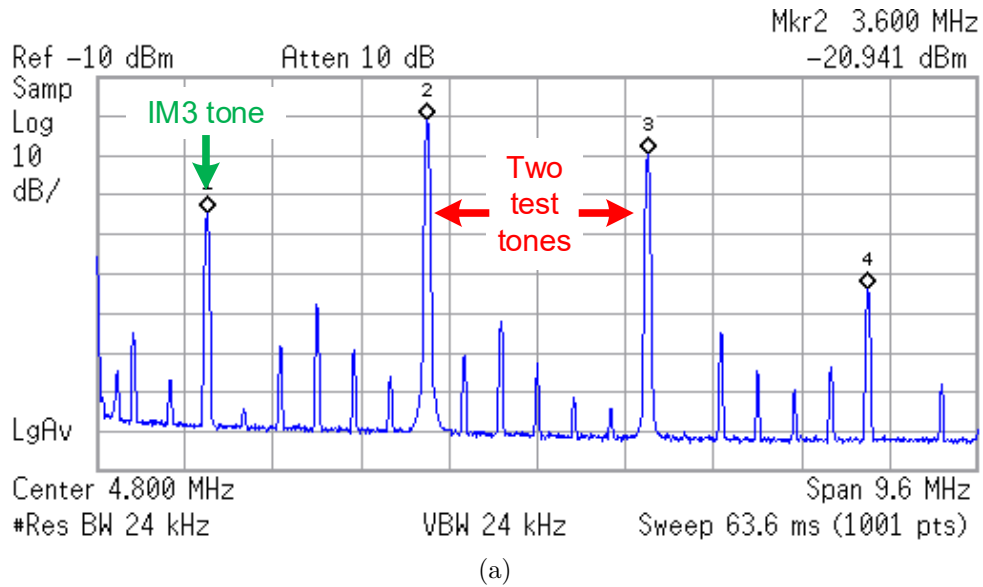
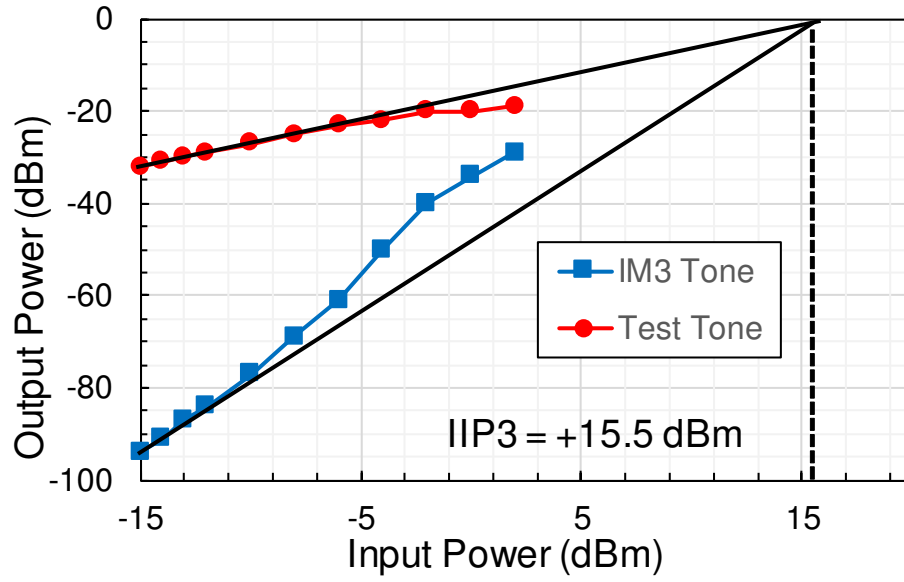
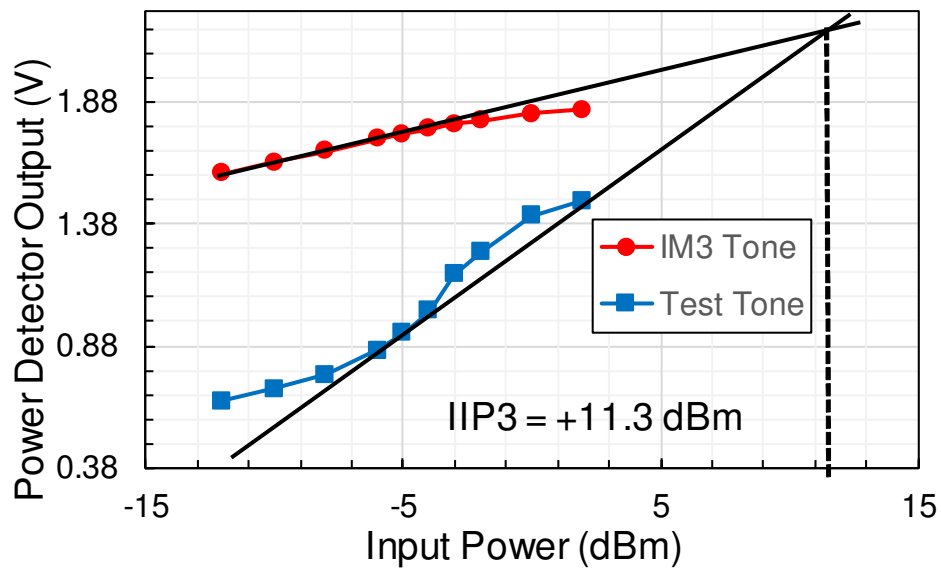


Figure 5.24: Measured output spectrum of the proposed baseband: (a) Normal operation mode and (b) tone-suppression mode for in-band IIP3 measurement.

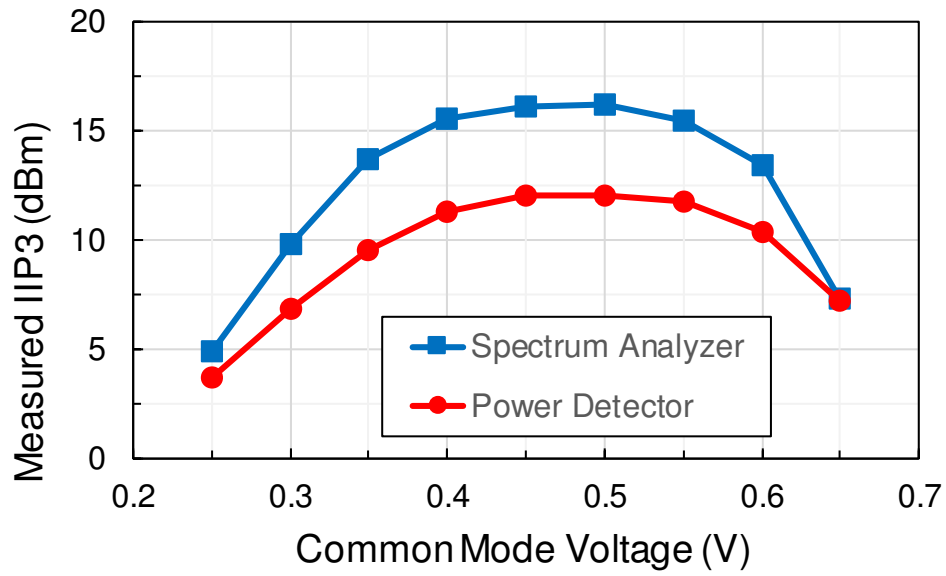


(a)

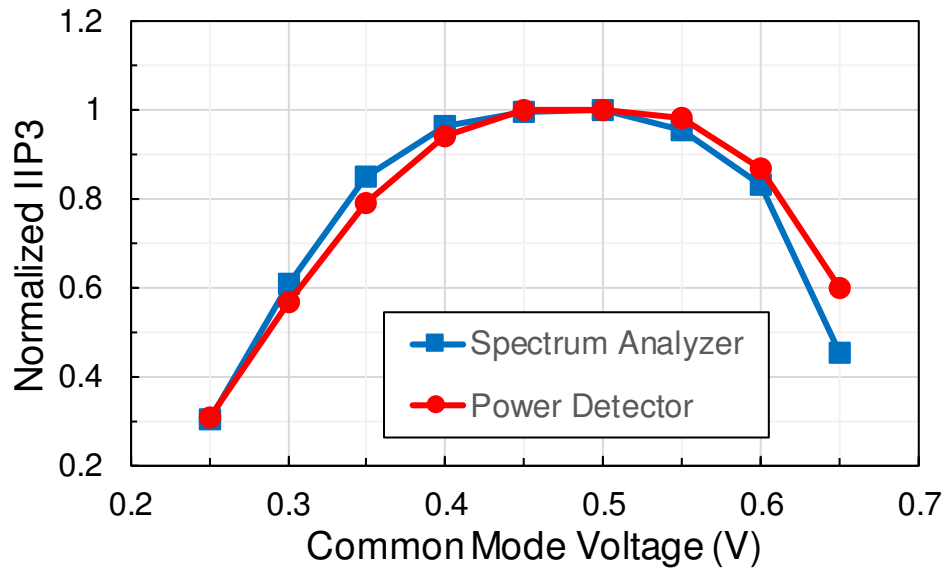


(b)

Figure 5.25: In-band IIP3 measurement results: (a) IIP3 measured by a commercial spectrum analyzer and (b) IIP3 measured by the proposed PD-based method.



(a)



(b)

Figure 5.26: In-band IIP3 measurement results: (a) IIP3 measured by a commercial spectrum analyzer and by the power detector and (b) normalized IIP3 measurement results.

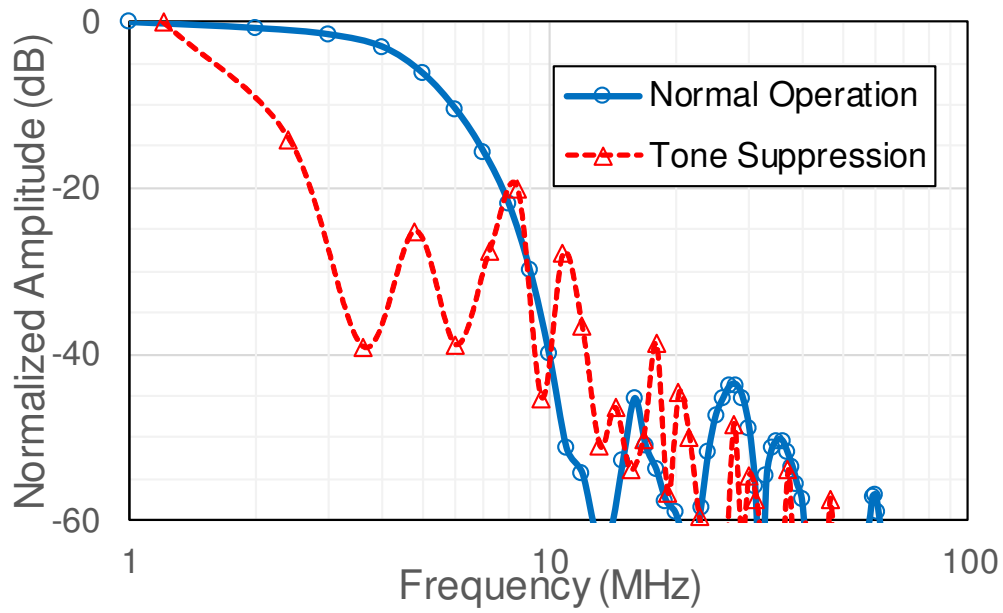
respondingly, the proposed PD-based measurement gives an approximate IIP3 value of +11.3 dBm as shown in Fig. 5.25b. The measured curves and the errors are consistent with the description in Section 5.5.2. The IIP3 measurement using a spectrum analyzer and the PD-based measurement are compared in Fig. 5.26a. Mismatch occurs between the two groups of measured values. However, Fig. 5.26b shows that if values in each group are normalized to the maximum value, the normalized IIP3 are well matched between the two groups. This means that the proposed PD-based linearity measurement can still be used as a linearity indicator and adopted to optimize the circuit's performance. Fig. 5.26b shows that

the proposed baseband achieves the best linearity when V_{CM} ranges from 450 to 500 mV.

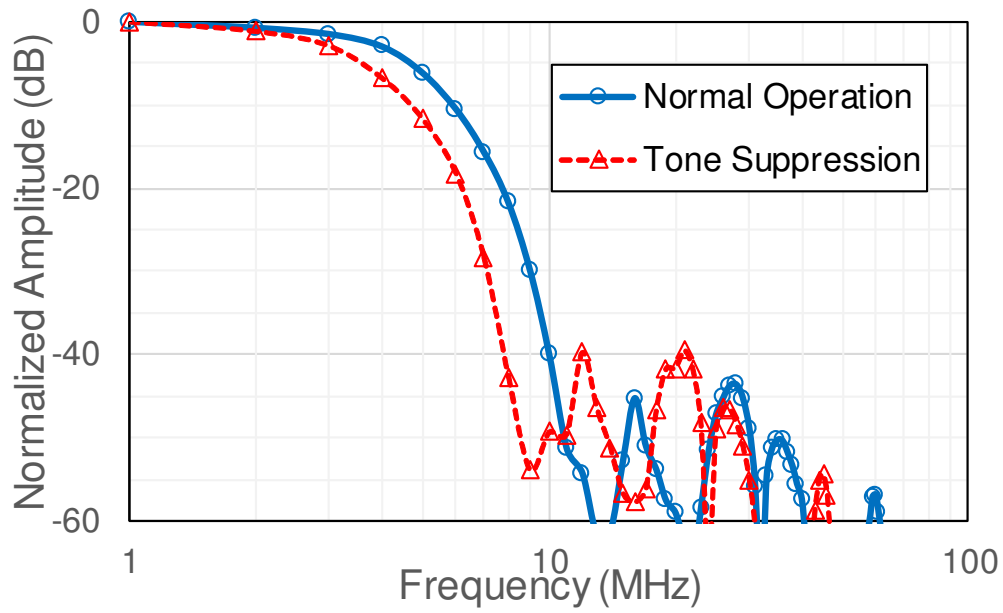
The complete frequency response of the proposed hybrid baseband is illustrated in Fig. 5.27. The normal operation mode and the tone suppression mode are compared. The response for in-band IP3 measurement setting is shown in Fig. 5.27a, while the response for out-of-band IP3 measurement setting is shown in Fig. 5.27b.

5.7 Conclusion

A continuous-time (CT)+discrete-time (DT) baseband chain is proposed with the testability of chain linearity performance. The proposed DT filter implements a 3-tap harmonic cancellation architecture tuned by a variable sampling frequency. The compacted architecture adopts the minimum number of capacitors for the same filter behavior among possible solutions. Different sampling frequencies switch the DT filter between the normal operation mode or the tone suppression mode. In the tone suppression mode, the DT filter notches the two test tone signals as they pass through the IM3 tones. A power detector is adopted to pick either the test tone power in the normal operation mode or the residual IM3 tone power in the



(a)



(b)

Figure 5.27: Measured frequency response of the proposed baseband: (a) Normal operation and tone suppression mode for in-band IP3 measurement and (b) normal operation and tone suppression mode for out-of-band IP3 measurement.

tone suppression mode. A comparison between the two types of measured power indicates chain linearity, which can be used to optimize chain performance. Moreover, the proposed design fully utilizes all existing components in a hybrid receiver chain and thus has zero overhead for the additional testability. This approach represents one step in the right direction for built-in testing/optimization of integrated circuit design.

6. ANALOG LTI SYSTEM AC/DC BIST BASED ON A TIME-TO-DIGITAL CONVERTER

6.1 Background

Testability is a very important concern in modern IC design. As the complexity of integrated digital systems have increased tremendously in last decades, digital on-chip built-in self-test (BIST) techniques have been well researched to reduce the testing cost and accelerate product time-to-market [2]. However, for analog BIST, current solutions are not mature compared to their digital counterparts. On one hand, analog excitation and output response are not 0/1 series but continuously changing waveforms. This imposes a complex design of the test pattern generator (TPG) and output response analyzer (ORA). On the other hand, multiple functional tests are preferred to characterize the full capability of an analog circuit. Therefore, to build an analog BIST system is to make a trade-off between hardware overhead and functional testability.

One interesting approach is the on-chip spectrum analyzer, a fully integrated test system which imitates an external spectrum analyzer. In this approach, a phase-locked loop (PLL) with frequency sweep ability drives a signal generator (SG) to output sinusoidal waveform and excite the circuit-under-test (CUT). The CUT's output is down-converted, filtered, amplified, and fed into an envelope detector or an analog-to-digital convertor (ADC) [4][51]. Moreover, direct digital synthesizer (DDS) is another way of generating the excitation signal. Instead of PLL and SG, a memory block and digital-to-analog convertor (DAC) are required to store and replay the waveform. [117] demonstrates the ability of measuring the CUT's linearity (IP3) by applying multi-tone DDS. On-chip spectrum analyzer is a very powerful BIST

tool designed to do most measurement tasks that an external instrument can do. Nevertheless, long test time and huge hardware overhead remain open problems. In [64], an on-chip spectrum analyzer occupies a significant 7.75% of the whole chip area. Furthermore, complicated BIST design raises a paradox—how to verify the analog block itself in the BIST. On the contrary, oscillation-based BIST (OBIST) is a light-weight solution with very low overhead [118]. In an OBIST circuit, a proper feedback network is connected to the CUT to make the CUT oscillate. By measuring the amplitude and oscillation frequency, fault in the CUT is detected. However, the major problem of OBIST is its specificity. Different CUTs require different feedback networks, which increases the design complexity for a SoC containing multiple analog blocks.

A time domain test [41, 119, 120, 121] is an alternative solution for compact on-chip analog BIST, as it's possible to remove the bulky frequency-sweep and ADC components. To detect faults in analog circuits, [41] suggested generating an impulse response. However, an ideal impulse is difficult to generate in practice. Instead, [119] utilized a square waveform, and [120] established principles for a ramp response test. Compared to a sine wave, harmonics of a step or ramp will spread across the whole spectrum. Applying such waveform is equivalent to running a frequency sweep without FS, SG or DDS. Besides, recent research focuses on time-to-digital converter (TDC), which is fully compatible to digital circuits, to do the time-domain measurement [121]. TDC's time resolution increases with the scaling down of the IC process. However, in [121], the author proposed a TDC-based BIST technique only for on-chip DC voltage measurement.

In this section, we take a more general approach to characterize on-chip linear-time-invariant (LTI) analog blocks for both DC gain and pole/zero positions via delay measurement. System architecture is very compact with respect to the CUT, con-

taining only a DAC for ramp generation and a TDC for delay measurement. Ramp signal is used to excite the CUT. Then, the measured input-to-output delay reveals pole/zero positions, and the ramp rise time can predict the CUT's voltage gain. Although a step excitation is much easier to obtain, ramp signal is preferred from a systematic view, because the step response involves an exponential term, which requires not only basic algebra but also an exponential equation solver, increasing the complexity of data processing. In addition, ramp excitation can use only simple inverters for voltage comparison in the measurement block, and an error cancellation technique is further developed to eliminate errors with variable ramp slopes.

6.2 Proposed Analog BIST Approach Using Only Digital I/O

6.2.1 System Architecture

The full conceptual architecture of the proposed analog BIST is demonstrated in Fig. 6.1(a). A differential ramp is generated by a clock-driven ramp generator and is fed to the input m_{IN} of the CUT. m_{OUT} is the output ramp of the CUT. The following Trigger/Multiplexer (T/M) block is an interface between the CUT and the TDC. It adopts two comparators with different threshold voltages V_{T1} and V_{T2} , and converts analog ramps to time delay. This time delay signal is then sent to the TDC block. A ring-oscillator TDC is implemented to digitalize the delay time between its inputs, START and STOP, and the result is read by an external digital automatic test equipment (ATE). Data is processed by the ATE and interpreted to the CUT performance. The whole BIST system is fully digital except for the output buffers of the ramp generator, which are used to drive the CUT.

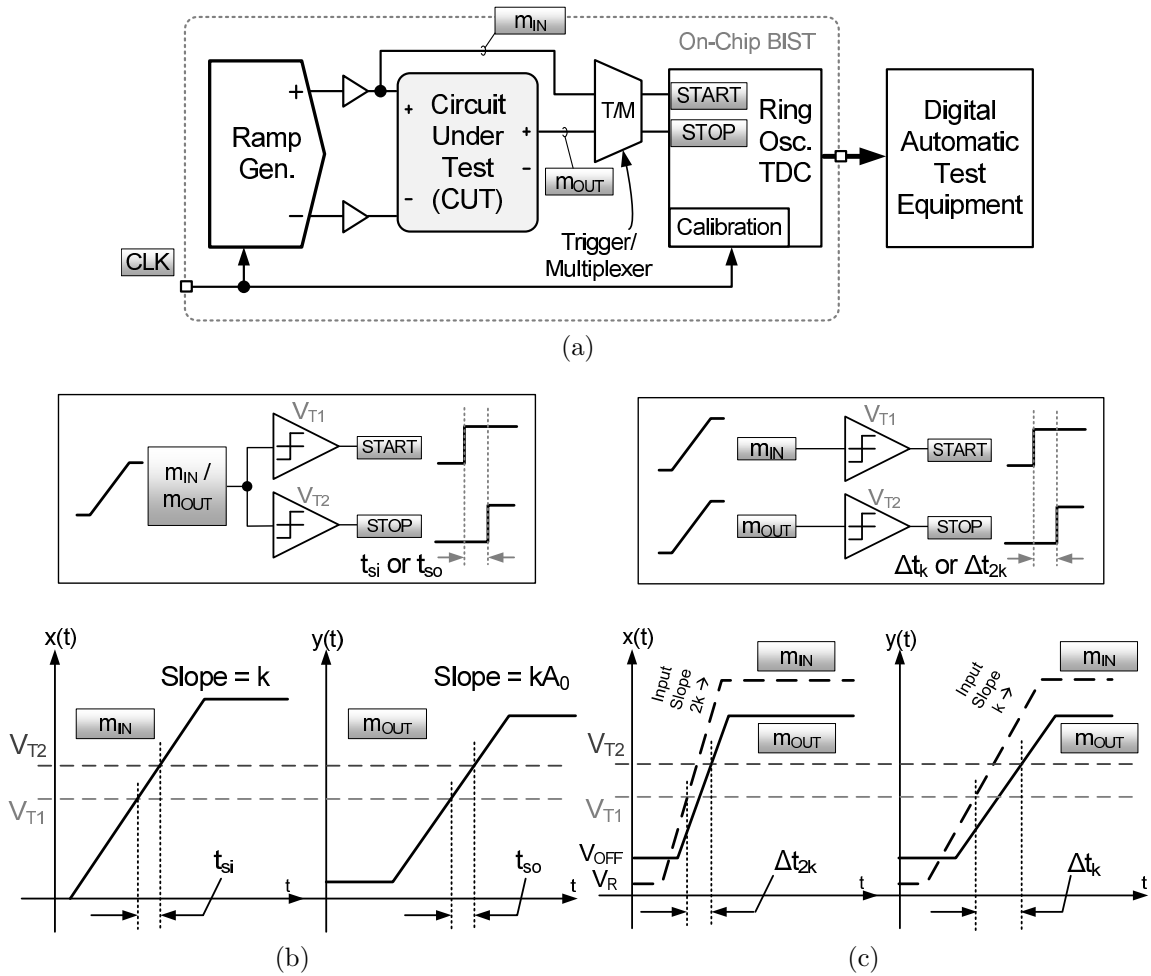


Figure 6.1: Proposed BIST approach: (a) BIST system architecture, (b) DC gain measurement, (c) t_d measurement with error cancellation.

6.2.2 Time-domain measurement principle

An LTI system can be described by a s-domain transfer function, which has a general form of

$$H(s) = A_0 \frac{\left(1 + \frac{s}{z_0}\right) \left(1 + \frac{s}{z_1}\right) \cdots \left(1 + \frac{s}{z_m}\right)}{\left(1 + \frac{s}{p_0}\right) \left(1 + \frac{s}{p_1}\right) \cdots \left(1 + \frac{s}{p_n}\right)} \quad (6.1)$$

Measurement of an LTI system yields information on its gain A_0 , poles p_i and zeroes z_j . For (6.1), if input is a ramp with slope k , $x(t) = kt$, then the output ramp

response is $r(t) = k\mathcal{L}^{-1}\left[\frac{1}{s^2}H(s)\right]$, where $\mathcal{L}^{-1}()$ is the inverse Laplace transform. For instance, a simple LTI system $A_0\frac{(1+s/z_0)}{(1+s/p_0)}$, which has one real pole p_0 , one real zero z_0 and DC gain of A_0 , generates the ramp response ($t \geq 0$)

$$r(t) = kA_0 \left[t - \left(\frac{1}{p_0} - \frac{1}{z_0} \right) + \left(\frac{1}{p_0} - \frac{1}{z_0} \right) e^{-p_0 t} \right] \quad (6.2)$$

When $t \rightarrow \infty$, the exponential term could be ignored. In other words, after some time, the output becomes

$$r(t) \approx kA_0 (t - t_d) \quad (6.3)$$

where kA_0 is the slope of the output ramp, and t_d is a fixed delay $t_d = p_0^{-1} - z_0^{-1}$. Further analysis shows that (6.3) is also true for a higher-order LTI system (as defined by (6.1)) except for different t_d . Physically, t_d is the phase delay of the LTI system at DC, and it can also be expressed as the equivalent time constant, which is the sum of all time constants in the LTI system,

$$t_d = \sum_{i=0}^n \frac{1}{p_i} - \sum_{j=0}^m \frac{1}{z_j} + \sum_{k=0}^r \frac{1}{\omega_k Q_k} \quad (6.4)$$

where p_i and z_j are real poles and real zeroes. For each complex pole pairs $1/(s^2 + \frac{\omega_l}{Q_l}s + \omega_l^2)$, resonance frequency ω_k and quality factor Q_k are used to represent the time constant. Particularly, if a circuit contains only one dominant pole, and its other poles and zeroes are located very far away from the dominant one, we can approximate its 3dB bandwidth as $\omega_{3dB} \approx t_d^{-1}$. This dominant pole scenario is very common for operational amplifiers. Furthermore, in a 2nd order low-pass system, the delay measurement is described by the last term of (6.4), $\omega_0 Q_0 \approx t_d^{-1}$. For higher order systems, although it is difficult to obtain the exact position of each pole or zero from one-time measurement, t_d is still a good signature, which lets us know whether the

circuit is working properly by comparing measured t_d to its simulated value in the design phase.

We can use (6.3) to get the DC gain A_0 and the equivalent time constant t_d of an LTI system.

6.2.2.1 Measure A_0

As introduced in the system architecture, two comparators with different threshold voltages V_{T1} and V_{T2} are purposely implemented. Fig. 6.1(b) shows the use of these two comparators. By connecting the CUT input ramp m_{IN} and output ramp m_{OUT} to the two comparators simultaneously, the delays between the trigger time determines the ramp slope, $t_{si} = \Delta V_{LAT}/k$ and $t_{so} = \Delta V_{LAT}/kA_0$, where $\Delta V_{LAT} = V_{T2} - V_{T1}$. Hence, the DC gain is obtained

$$A_0 = \frac{t_{si}}{t_{so}} \quad (6.5)$$

6.2.2.2 Measure t_d with offset error cancellation

t_d measurement is demonstrated in Fig. 6.1(c). The input and output ramp are connected to the two comparators separately. On the one hand, due to the DAC reference voltages, and the output swing limit and the offset voltage of the output buffers, the input ramp may start from an arbitrary voltage level V_R , which can be expressed as $x(t) = kt + V_R$. The input ramp triggers the V_{T1} comparator at $t_{in} = \frac{1}{k}(V_{T1} - V_R)$. On the other hand, the CUT may also contribute additional offset voltage, leading to another start level V_{OFF} for the output ramp $r(t) = kA_0(t - t_d) + V_{OFF}$, which triggers the V_{T2} comparator at $t_{out} = \frac{1}{kA_0}(V_{T2} - V_{OFF}) + t_d$. We can rewrite (6.3) to consider all offset voltages,

$$\Delta t = t_{out} - t_{in} = \frac{1}{k}e_v + t_d \quad (6.6)$$

$$e_v = \frac{V_{T2} - V_{OFF}}{A_0} - V_{T1} + V_R$$

An error term e_v is emerging, which is related to CUT's DC gain, offset voltage, and comparator threshold voltages, plus the objective delay t_d . To eliminate e_v and get the true equivalent time constant, two input ramp slopes should be used. Applying the ramp signals of two different slopes, e.g., k and $2k$, gives us $\Delta t_k = \frac{1}{k}e_v + t_d$ and $\Delta t_{2k} = \frac{1}{2k}e_v + t_d$. No matter how large is the error term e_v , t_d can be obtained

$$t_d = 2\Delta t_{2k} - \Delta t_k \quad (6.7)$$

In short, accurate DC and AC measurements can be achieved by applying a ramp excitation and obtained by a delay measurement. The input and output ramp slopes can be converted to time delays, and their ratio reveals a DC gain, A_0 (Fig. 6.1(b)). The input-to-output ramp delay determines t_d , which indicates pole/zero positions (Fig. 6.1(c)). Both measurements adopt the same comparator set with different threshold V_{T1} and V_{T2} . Moreover, both measurements are insensitive to PVT variation. DC gain measurement relies only on the ratio of slopes rather than individual values, while t_d measurement uses two different slopes to eliminate offset error during data processing. Although measurement configuration is straightforward as indicated by (6.5) and (6.7), a certain procedure must be followed. LTI systems which need to be measured should have non-zero/non-infinite DC gain. That is to say, pure band-pass or high-pass circuits, which have zero DC gain, cannot be tested by the proposed method.

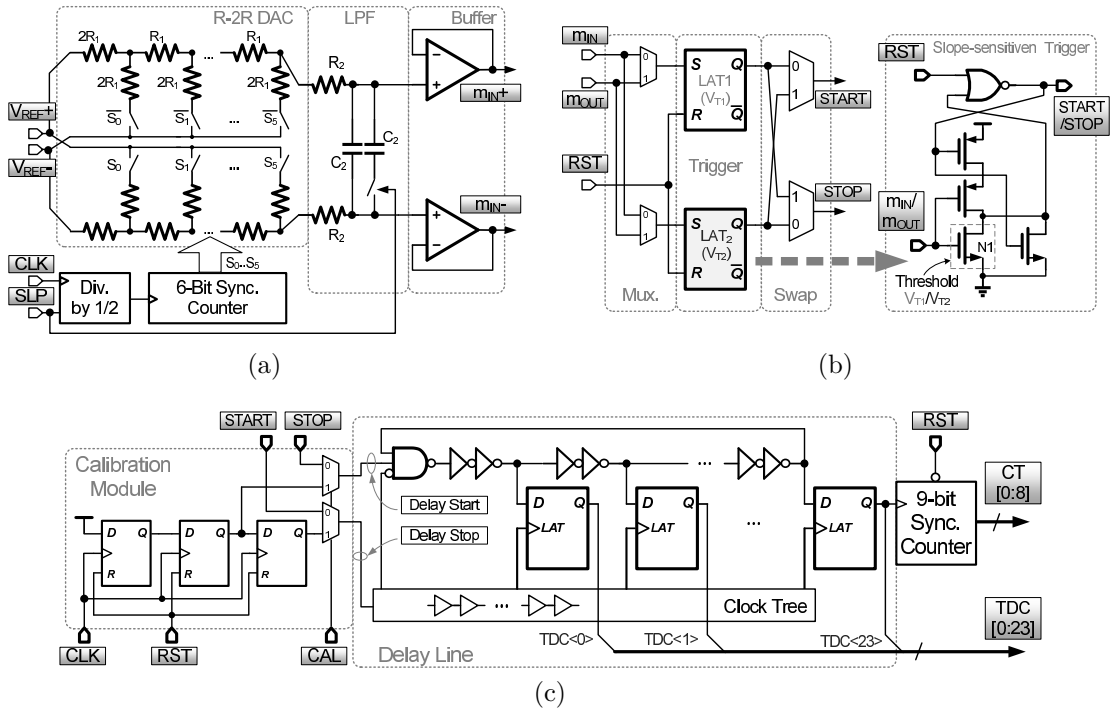


Figure 6.2: Schematic of the system blocks: (a) Ramp generator, (b) trigger and multiplexer, and (c) time-to-digital convertor.

6.3 Block Circuits Design

6.3.1 Ramp Generator

The schematic of the ramp generator is demonstrated in Fig. 6.2(a). The CLK signal drives the 6-bit counter to count up, controlling the DAC to generate a differential stair ramp. Then the stair waveform passes through a low-pass filter (LPF), resulting in an approximate quasi-ramp signal. This quasi-ramp is slightly distorted and introduces a linearity issue for the measurement, which will be discussed later. To isolate the CUT load from the LPF, two-stage single-ended amplifiers are implemented to buffer the quasi-ramp output. As discussed above, the offset error cancellation technique for t_d measurement requires accurate slope control. In Fig.

6.2(a), signal SLP is added to change the clock divider's ratio between 1 and 2, and, thus, outputs CLK/2 for slope $2k$ or CLK for k . The LPF is also controlled by SLP. LPF bandwidth will be reduced to half for slope $2k$. Moreover, R-2R DAC is adopted because of its constant output impedance R_1 , and thus the LPF has a fixed bandwidth of $\omega_{3dB} = 1/[(R_1 + R_2)C_2]$. It will not change for any DAC input bits. Besides, the mechanism of clock-driven ramp generation guarantees that the slope error of the quasi-ramp depends on clock jitter only.

In one test cycle, the output quasi-ramp waveform changes from V_{REF-} to V_{REF+} in a period of T_{Ramp} . Then, the approximate slope of the quasi-ramp signal becomes

$$k \approx \frac{V_{SW}}{T_{Ramp}} = \frac{f_{CLK}V_{SW}}{N-1} \quad (6.8)$$

where f_{CLK} is the clock frequency, V_{SW} is the ramp swing, and the counter output has N steps.

6.3.2 Trigger/Multiplexer

The Trigger/Multiplexer block is designed for converting analog ramp signal to time delay, which is presented by the rising edges of START and STOP. The schematic of a T/M block is shown in Fig. 6.2(b). In the first multiplexer stage, the rising ramp m_{IN} and m_{OUT} are routed to LAT1 or LAT2, depending on the measurement mode. For t_d measurement, the output delay will be generated between m_{IN} and m_{OUT} , which is described by (6.6). If the input ramp slope (t_{si} in (6.5)) will be measured, both LAT1 and LAT2 will accept m_{IN} . Similarly, connecting m_{OUT} to both latches will give the slope of the CUT's output (t_{so} in (6.5)). In the following stage of slope-sensitive triggers, two RS-latches play the role of voltage comparator. They are NOR type RS latches and share the same sizes of transistors except for N1. Thus the two latches have separate thresholds V_{T1} and V_{T2} . During

testing, the operator must first set RST=1 and tie down the ramp signal to reset the latches. The next step is to set RST=0 and let the ramp signal climb up. When the ramp arrives at the specific threshold voltage, it will generate a rising edge on the output START/STOP. RS-latch is chosen because after it is triggered, its output will not change even if the ramp voltage drops back below its threshold, avoiding noise-induced output bit flipping. Furthermore, in the last stage, the START and STOP signal can be swapped, which allows the TDC to measure negative delays.

6.3.3 Time-to-digital Converter

Fig. 6.2(c) illustrates the TDC block, which can generate a composite digital word W_{TDC} to represent the delay t_{TDC} measured by the TDC. Considering that the TDC has the finest time resolution δ_{TDC} , we can derive

$$W_{TDC} = \left\lfloor \frac{t_{TDC}}{\delta_{TDC}} \right\rfloor \quad (6.9)$$

The fine measurement delay line is constructed by a 24-stage ring oscillator. Each stage has a delay of δ_{TDC} . D-latch array is implemented to latch the line status when receiving the delay stop signal, which is propagated by a clock tree. The last bit of the delay line is also fed into a 9-bit synchronous counter for coarse measurement. Finally, TDC<0:23> is the thermometer code of the delay line, while CT<0:8> is the code of the counter. W_{TDC} can be obtained by combining these two codes, and, thus, it allows the TDC to operate in a wide time range. In addition, a 3-stage resettable D-flipflop chain is adopted to calibrate δ_{TDC} . During the calibration procedure (CAL=1), all D-flipflops are reset, and then CLK will drive logic 1 propagating along the chain. TDC is activated at the second clock edge and disabled at the third. One clock cycle is measured and its output W_{TDC} indicates $\delta_{TDC} \approx (W_{TDC} \cdot f_{CLK})^{-1}$.

6.4 Measurement Analysis

6.4.1 Quantization Error

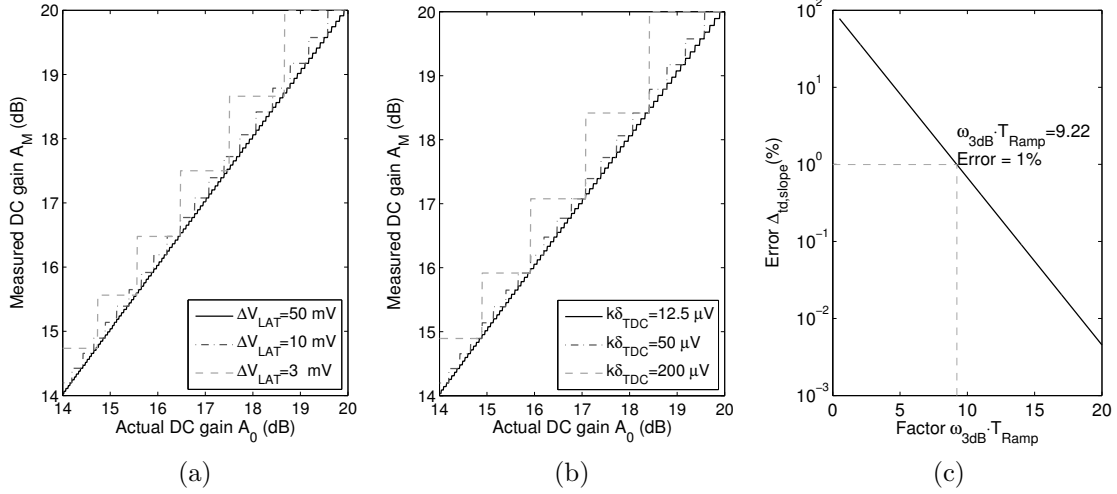


Figure 6.3: Measurement analysis: (a) Measure DC gain, $k\delta_{TDC} = 50$ μ V, sweep ΔV_{LAT} , (b) measure DC gain, $\Delta V_{LAT} = 10$ mV, sweep $k\delta_{TDC}$, (c) ramp-slope-induced error in a 1-pole system.

Quantization procedure will affect the measurement accuracy. For DC gain measurement, t_{si} and t_{so} in (6.5) are quantized by (6.9). Thus, the DC gain measured by TDC can be written as

$$A_M = \frac{W_{TDC,ti}}{W_{TDC,to}} = \frac{\left\lfloor \frac{\Delta V_{LAT}}{k\delta_{TDC}} \right\rfloor}{\left\lfloor \frac{\Delta V_{LAT}}{kA_0\delta_{TDC}} \right\rfloor} \quad (6.10)$$

Fig. 6.3(a) shows the measured A_M for different ΔV_{LAT} , while Fig. 6.3(b) illustrates the impact of $k\delta_{TDC}$. Increasing ΔV_{LAT} or decreasing $k\delta_{TDC}$ (to achieve a lower input ramp slope or higher TDC resolution) can help improve the measurement

precision. For instance, assume $\Delta V_{LAT} = 10$ mV and $k\delta_{TDC} = 15$ μ V, the error of the measured DC gain is less than 0.6% in dB (for $A_M < 20$ dB). Quantization error can also limit the range of t_d measurement. As t_d is directly quantized by TDC, its resolution is indeed TDC's resolution δ_{TDC} . Assume a TDC with $\delta_{TDC} = 40$ ps and let $t_d = \delta_{TDC}$. We can find a maximum of 50% measurement error for a 1-pole system with bandwidth around $1/2\pi\delta_{TDC} \approx 3.98$ GHz.

6.4.2 Exponential-term-induced Error

This section presents an analog test challenge where the test had to be finished in a given time. Hence, a limited T_{Ramp} was selected. As a result, the excitation ramp became too fast for poles/zeros located at low frequencies because the exponential terms in (6.3) cannot be neglected. Consider a normalized 1-pole system ($z_0 = \infty$ and $A_0 = 1$ for (6.2)), the actual delay measured is $\frac{1}{p_0} - \frac{1}{p_0}e^{-p_0t}$, where the pole is equal to the 3dB bandwidth, $p_0 = \omega_{3dB}$. If we assume $V_{T1} = V_{T2} = \frac{1}{2}V_{SW}$, the additional error delay induced by the exponential term is expressed as

$$\Delta_{t_d, slope} = \left| e^{-\frac{1}{2}\omega_{3dB}T_{Ramp}} \right| \quad (6.11)$$

Fig. 6.3(c) shows that, in a 1-pole system with bandwidth ω_{3dB} , factor $\omega_{3dB}T_{Ramp} = 9.22$ can produce 1% error delay. This criteria will become more stricter in real circuit because of following reasons. The comparators in the T/M block are purposely designed with different threshold voltages, causing different exponential terms for t_{out} and t_{in} in (6.6). A more complicated LTI system has more poles and zeroes, and thus add more exponential terms into (6.11). Positive DC gain will also amplify the exponential terms. Hence, $T_{ramp} > 9.22/\omega_{3dB}$ is required for the lowest bandwidth ω_{3dB} to be measured.

6.4.3 Linearity Analysis

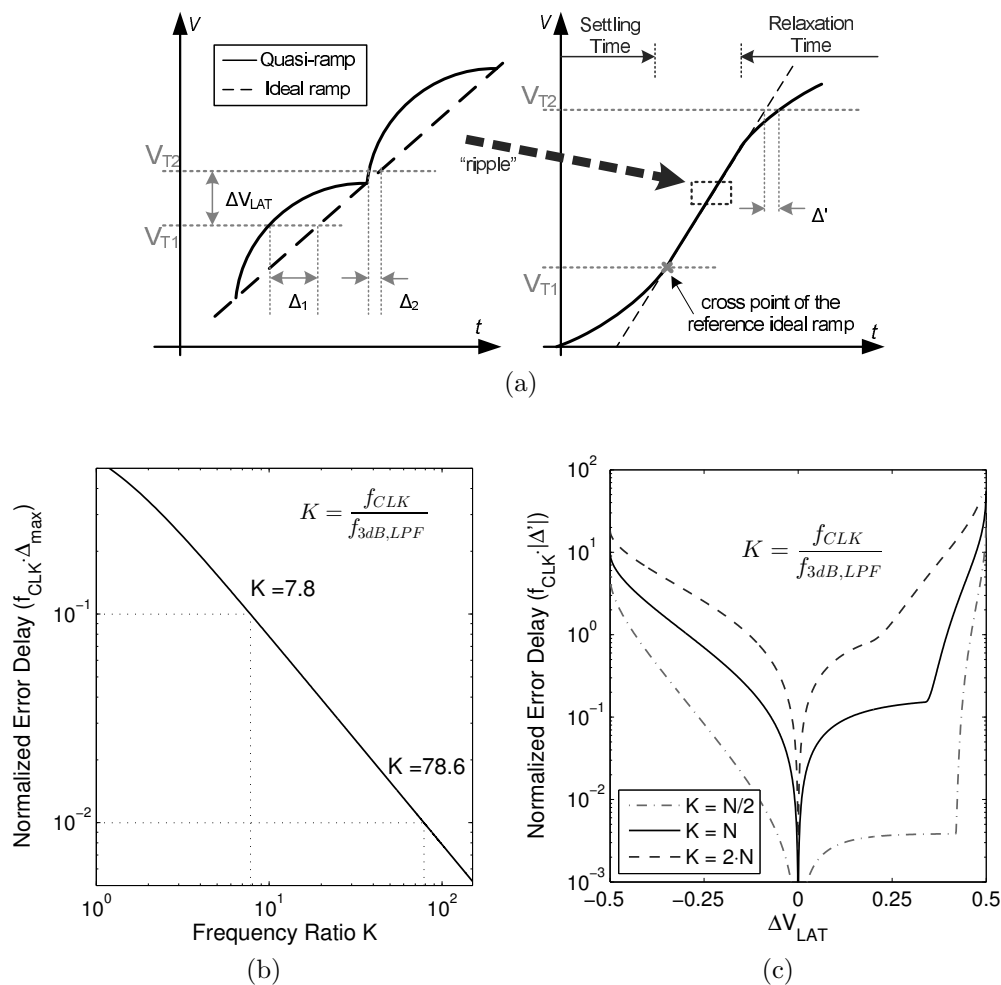


Figure 6.4: Measurement analysis: (a) Nonlinearity-induced error sources, (b) periodic-ripple-induced error delay, (c) settling/relaxation-time-induced error delay.

In the ramp generator, the output is, in fact, a non-linear quasi-ramp as its shape is slightly different from an ideal ramp. This nonlinearity has two aspects, periodic nonlinearity and settling/relaxation nonlinearity. As illustrated in Fig. 6.4(a), the quasi-ramp exhibits a settling/relaxation time near the start/end ramp corner due to

the LPF's characteristic. And the "ripple" on the ramp introduces additional periodic nonlinearity. This "ripple" is constructed by the original staircase waveform's high order harmonics, which are not fully suppressed by the LPF. For simplicity, we assume that the dominant pole frequency of CUT is much higher than the LPF bandwidth and $A_0 = 1$, $V_{OFF} = V_R = 0$. As a result, the output waveform will exactly follow the input quasi-ramp, with some time shift. Thus the analysis will focus on the error occurring on the quasi-ramp itself.

In the detailed "ripple" plot of Fig. 6.4(a), the distorted waveform crosses threshold voltages at different time points compared to an ideal ramp, resulting in an extra error delay Δ_1 for V_{T1} and Δ_2 for V_{T2} . Fig. 6.4(b) shows the plot of the normalized error delay $f_{CLK} \cdot \Delta_{max}$ vs. frequency ratio K for $N = 64$. Δ_{max} is the maximum possible periodic-ripple-induced error delay $max |\Delta_2 - \Delta_1|$; f_{CLK} is the ramp generator's clock frequency, and $f_{3dB,LPF}$ is the LPF's bandwidth frequency. Frequency ratio K is defined as $f_{CLK}/f_{3dB,LPF}$. When $K \approx 78.6$, Δ_{max} is 1% of the clock period. Therefore, to reduce the error delay, increasing the clock frequency or reducing the LPF bandwidth are efficient.

Another source of the nonlinearity comes from the settling/relaxation procedure. As demonstrated in Fig. 6.4(a), if we draw a line with slope k , making it cross the intersection point of V_{T1} and the quasi-ramp, the time difference between the extended line and the output ramp at threshold V_{T2} is Δ' . Δ' will become larger if either V_{T1} or V_{T2} is close to the lower or upper bound of the ramp swing where the quasi-ramp has not fully settled or has started relaxing. Decreasing ΔV_{LAT} and placing V_{T1} and V_{T2} close to the middle of the ramp swing could help to reduce Δ' . On the other hand, the LPF bandwidth cannot be set too low as this will prolong the settling/relaxation time. Assuming V_{T1} is 0.5 V, $N = 64$ with sweeping ΔV_{LAT} and frequency ratio K , normalized error delays are obtained in Fig. 6.4(c).

To sum up, the total nonlinearity-induced error delay can be expressed as

$$\delta_{distort} = \Delta_2 - \Delta_1 + \Delta' \quad (6.12)$$

This error is added each time when the delay is measured. In order to reduce the impact of quasi-ramp nonlinearity. A trade-off should be made on LPF's bandwidth and ΔV_{LAT} . Values of V_{T1} and V_{T2} should be carefully tuned for $V_{T2} > V_{T1} > 0.5V_{SW}$. And K should be chosen between $0.5N$ and N to optimize $\delta_{distort}$.

6.5 Experimental Results

The proposed circuit is fabricated in 65 nm standard CMOS technology. As shown in Fig. 6.5(a), the ramp generator only occupies $125 \mu\text{m} \times 115 \mu\text{m}$, and the TDC occupies about $150 \mu\text{m} \times 80 \mu\text{m}$, compared to a $340 \mu\text{m} \times 300 \mu\text{m}$ 2nd order reconfigurable active biquad with programmable RC bank.

The biquad is first reconfigured as a 1st order low pass active filter. t_d measurement is performed with 48 MHz input clock frequency. 36 different RC combinations are measured, and the delay results are interpreted according to the filter's bandwidth frequency f_{3dB} . As shown in Fig. 6.5(b), f_{3dB} points measured by the proposed technique are close to the frequencies measured by the oscilloscope. Measurement precision is higher for low 3dB frequency, achieving 7.8% average error. We can also observe the error on the lowest two frequency points in Fig. 6.5(b) due to the fast ramp (as defined by (6.11)).

DC gain measurement is also tested. The biquad is configured as a programmable gain amplifier (PGA). The DC gain results measured by an oscilloscope and by the proposed technique are compared in Fig. 6.5(c). The maximum measured gain error is 0.42 dB.

Furthermore, the programmable biquad is turned to a 2nd order low pass active-

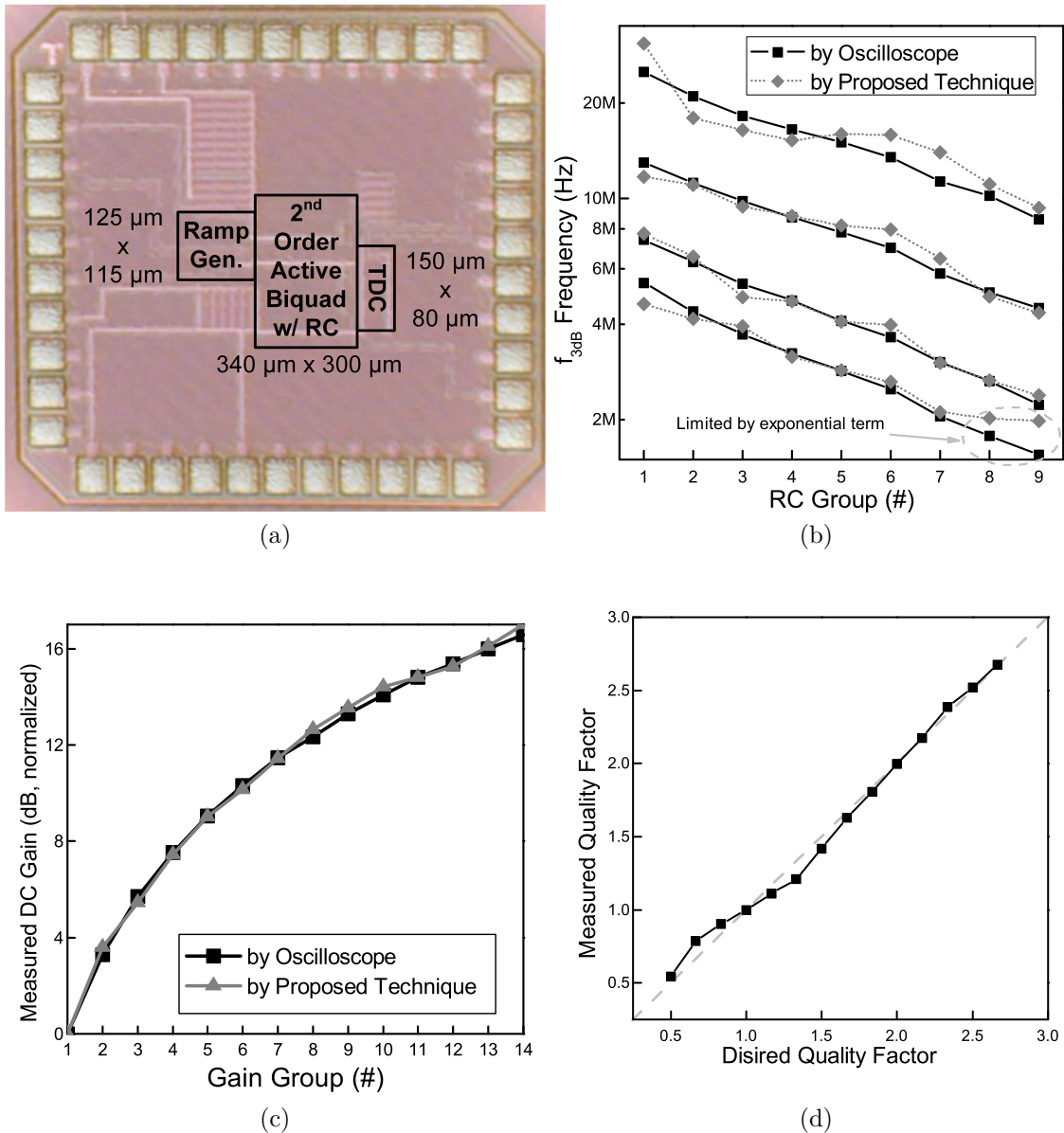


Figure 6.5: Experimental result: (a) Photo diagram of the chip die, (b) f_{3dB} measurement results of a 1st order low-pass active filter, (c) DC gain measurement results of an inverting amplifier, and (d) quality factor measurement results of a 2nd order low-pass filter.

RC filter for obtaining the quality factor (Q) based on the t_d measurement. The active-RC filter is set to a desired Q value by changing internal resistor ratio. Firstly,

we set desired Q to 1 and do a t_d measurement to obtain the filter's bandwidth ω_0 . Then, we set different resistor ratios and measure new t_d . Q can be obtained, according to $\omega_0 Q = \frac{1}{t_d}$ in (6.4). The measurement results are summarized in Fig. 6.5(d), and using the proposed technique achieves a maximum difference $\Delta Q = 0.12$ between the desired and measured quality factor values.

6.6 Conclusions

A low-overhead analog BIST technique with fully digital I/O is proposed in this section. Ramp response is deployed instead of traditional frequency-sweep approach. AC performance was obtained from input-to-output delay, and DC gain was collected by measuring the slope-related delays. Generic measurement methods enabled AD/DC testing for different on-chip analog LTI blocks with a single tester architecture, which consists of a clock-driven ramp generator and a TDC. Moreover, with an error cancellation technique in data processing, the BIST method is insensitive to the PVT variation by applying basic algebra. In addition, the TDC block can further adopt digital BIST, as it is constructed by logic gates only. The ramp generator's output can also be measured during the DC gain test, which indicates the correctness of the analog parts in the tester itself. In other words, the proposed approach is fully self-testable. The proposed BIST approach is an multi-functional, robust and economic solution for analog on-chip testing in future mixed-signal IC systems.

7. CONCLUSIONS AND FUTURE WORKS

7.1 Conclusions

This dissertation proposes concepts to achieve a fully integrated in-situ design validation and optimization hardware for analog circuits.

Section 2 implements a digital multi-dimensional optimization engine to adaptively adjust analog circuit components in order to achieve a desired performance. The proposed self-validation circuit synthesizes a sinusoidal waveform to stimulate the target analog circuit, and its response is quantized by a 10-bit low power SAR ADC. Measurement data is processed digitally to evaluate the error between the measured and the desired performance. The optimization engine further changes multiple design variables simultaneously to find a solution with minimum error. An instability detection feature is adopted to make the analog circuit insensitive to PVT variations and work as expected even under extreme conditions. To verify this method, a 2nd/4th active-RC BPF with variable-GBW amplifiers and programmable RC constants as well as a 2nd order Gm-C band-pass biquad with programmable transconductance are applied as study cases. This method is ideal for use in our study cases because of its ability to reduce power consumption by a factor of 4 from the conventional design specs, while meeting most of the original specifications. This self-contained system, which integrates all analog and digital blocks on-chip, was fabricated using 180 nm standard CMOS technology, occupying a 0.4 mm² silicon area.

The excitation sinusoidal signal synthesizer of Section 2 is elaborated in Section 3. The proposed synthesizer utilizes 50% duty cycle and differential-mode circuitry to eliminate the even order harmonics, and it also implements a 5-phase 3-amplitude

harmonic cancellation technique to suppress the 3rd, 5th, 7th, and 9th order harmonics. The compact system architecture consists of a 12-phase ring oscillator, a weighted resistor summing network, and an RC output filter. Phase shifters are adopted in the ring oscillator to enable the control of an external harmonic cancellation optimization algorithm. The proposed application of the optimization algorithm compensates the errors in the circuit and further improves the linearity of the output waveforms. This synthesizer is fabricated in 180 nm standard CMOS technology, occupies a 0.08 mm² silicon area and achieves the spur-free dynamic range (SFDR) of 59 to 70 dBc from 150 to 850 MHz after the optimization procedure. It can operate from a 1 to 1.8 V supply voltage and achieve a power consumption from 9.11 to 57 mW.

To extend the application of the sine-wave synthesizer proposed in Section 3 to two-tone linearity test, a low-distortion current-steering two-tone sinusoidal signal synthesizer based on a mixing-FIR architecture is proposed in the Section 4. The proposed robust synthesizer adopts only digital blocks. It implements a two-stage cascade FIR harmonic cancellation technique that generates a single tone quasi-sinusoidal waveform and suppress the harmonics up to the 23rd order. The single tone signal is further up-converted to the desired LO frequency band, thus producing the desired two-tone sinusoidal signals. The proposed synthesizer is fabricated in 130 nm standard CMOS technology, occupying a 0.056 mm² silicon area. Measurement shows better than -68 dBc IM3 below 480 MHz LO frequency without calibration. For the LO frequency < 76.8 MHz and the two tone difference < 2 MHz, an IM3 less than -75 dBc can be achieved. The imbalance between the two tone amplitudes is measured < 0.1 dB across the whole frequency range.

An on-chip RF receiver linearity built-in test methodology for hybrid baseband chain is proposed in Section 5. It's one step towards the fully on-chip optimization

of analog circuits. The proposed baseband chain consists of a continuous-time (CT) lowpass filter and a discrete-time (DT) finite impulse response (FIR) filter with a compacted two-stage 3-tap harmonic cancellation (HC-3²) architecture to achieve notching at specific programmable frequency points. To measure the chain linearity before the DT filter, two-tone test signals are injected and the residue power after the DT filter is measured by a power detector. By changing the sampling frequency, the DT filter can switch between the normal operation mode, which allows the two test tones to pass through, or the tone suppression mode, which notches the two test tones and exposes the IM3 tone power. A comparison between the power measured in the proposed two modes reveals the chain linearity performance. One single path of the proposed receiver baseband chain is fabricated in 130 nm standard CMOS technology, occupying a 0.146 mm² silicon area. Measurement results shows a suppression of 40 dB on the test tone power, and the proposed on-chip linearity test can be conducted with the proposed technique.

Furthermore, Section 6 proposed a low cost built-in analog tester with fully digital input/output for linear-time-invariant (LTI) analog blocks. A single tester architecture is implemented, which consists of a DAC-based ramp generator and a ring-oscillator time-to-digital converter. This implementation allows full characterization of AC response and DC gain for passive or active LTI blocks. This measurement procedure is insensitive to PVT variation due to its ability to apply an error cancellation technique during data processing. The excitation and measurement circuits are fabricated in 65 nm standard CMOS process and occupies 0.026 mm² area.

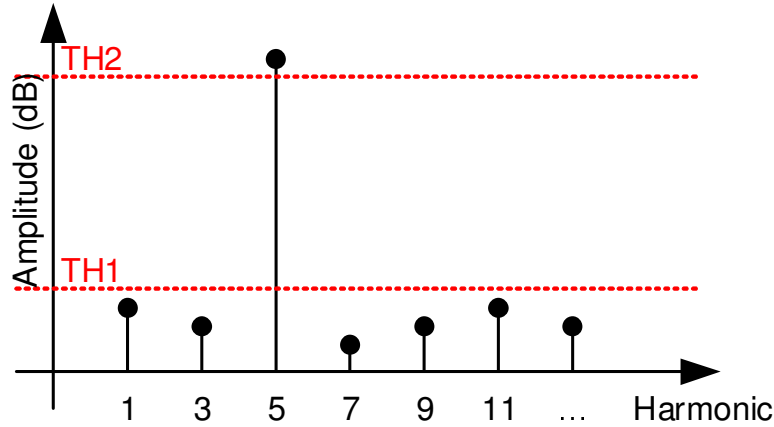


Figure 7.1: Expected spectrum of enforcing an arbitrarily selected harmonic.

7.2 Future Works

7.2.1 Strengthen or cancel arbitrary harmonic(s)

As introduced in Section 3.2.2.2, [76] demonstrated a digital harmonic synthesis block (DHSB) that added multiple shifted square-wave clocks to enforce a selected harmonic via the manipulation of the phase shift (i -th clock is shifted by θ_i). The objective is to find a group of phase shifts $\vec{\theta} = \{\theta_1, \theta_2, \dots, \theta_M\}$, where M is the total number of shifted clocks, to achieve the best cancellation. One expected example is illustrated in Fig. 7.1. The amplitude of the selected fifth order harmonic is higher than a threshold $TH2$, while the amplitudes of the other harmonics are lower than $TH1$. There's no easy way to solve such a problem. Even in [76], the unwanted harmonics near the selected harmonic are still high. However, [76] only manipulated $\vec{\theta}$, we can further use different amplitude sizing factor $\vec{\alpha} = \{\alpha_1, \alpha_2, \dots, \alpha_M\}$ to elaborate the original design, although this makes the problem more complicated.

An optimization method can be used to solve this problem. At the very beginning, a cost function should be defined to describe the expected outcome shown in Fig. 7.1.

One possible definition is

$$F_{cost,k}(\vec{\alpha}, \vec{\theta}) = \begin{cases} D_k - TH1 & D_k > TH1 \text{ and } k \text{ is not selected,} \\ TH2 - D_k & D_k < TH2 \text{ and } k \text{ is selected,} \\ 0 & \text{otherwise} \end{cases} \quad (7.1)$$

$$F_{cost}(\vec{\alpha}, \vec{\theta}) = \sum_{k=1}^M F_{cost,k}(\vec{\alpha}, \vec{\theta}) \quad (7.2)$$

where D_k is the amplitude of the k -th order harmonic obtained in (3.6). Obviously, when the expected spectrum is achieved, we have $F_{cost}(\vec{\alpha}, \vec{\theta}) = 0$. Moreover, only limited number of factors can be used for α_i and θ_i in order to allow the hardware implementation. α_i is limited to a set (Λ) of integer numbers (e.g. $\Lambda = \{2, 3, 4\}$). And θ_i is only allowed among slices of 2π with a step of $\frac{2\pi}{N}$, where N is an integer number. These slices construct a set $\Theta = \{\frac{1}{N} \cdot 2\pi, \frac{2}{N} \cdot 2\pi, \dots, \frac{N}{N} \cdot 2\pi, \}$. It should be noted that, M and N are also parameters to be optimized. The optimization problem is defined as

$$\begin{aligned} \text{minimize} \quad & F_{cost}(\vec{\alpha}, \vec{\theta}) \\ \text{s.t.} \quad & \alpha_i \in \Lambda \\ & \theta_i \in \Theta \\ & M_{min} < M \leq M_{max} \\ & 0 < N \leq N_{max} \end{aligned} \quad (7.3)$$

where M_{min} and M_{max} are limited by the architecture (i.e. current mirror, resistor array, etc.) and the corresponding matching tolerance (i.e. for a resistor array, achieving 1:2 is easy, but 10001:10002 is too accurate to be implemented), and N_{max}

is also restricted by the architecture (i.e. divider, delay line, etc.) and the timing resolution.

Solving the optimization problem defined in (7.3) and building a hardware implementation is still a challenge for the future work.

7.2.2 Improve the notching of HC-3² DTF

To improve the precision of the IP3 measurement proposed in Section 5, the notching depth at the test tone frequencies should be improved. As analyzed in Section 3.4, the effect of the harmonic cancellation is limited by the amplitude mismatches, which is the mismatches of the capacitance values in the proposed HC-3² DTF, and the clock phase mismatches. On the one hand, capacitance value mismatches can be reduced via larger capacitors and refined layout patterns. Increasing the area of every unit capacitor can help reduce the mismatches. However, larger area will result in longer wires for clock signals, inducing more clock phase errors. Trade-off should be made between the capacitor sizes and the clock wire lengths. The layout of the capacitor array can also be improved. Instead of drawing the common-centroid layout for one phase group, which is illustrated in Fig. 5.16, the cancellation will benefit from further splitting the unit capacitor and accommodating them in a fully common-centroid layout pattern for all phase groups. On the other hand, to reduce the clock phase mismatches, adjustable delays (such as the variable capacitor array shown in Fig. 3.13) can be applied to all clock signal paths. Set the HC-3² DTF in the suppression mode and excite the DTF using two test tones with medium or low power (to avoid strong IM3 tone generation). In this scheme, the power detector can be used to evaluate the test tone suppression, in other words, the depth of the notching. Define the measured output power as the cost function $F(\boldsymbol{v})$. Then, similar to the optimization problem introduced in Section 3, we can adjust

the delays to obtain the minimum $F(\mathbf{v})$, which means the DTF achieves the deepest suppression at the notching frequencies.

7.2.3 Time-domain measurement of LTI system with multiple poles and zeros

One drawback of the LTI system measurement methodology proposed in Section 6 is that multiple poles and zeros cannot be distinguished in one measurement, because the measured delay t_d is the sum of all time constants in the LTI system, which is defined by (6.4). A possible work around method is to carry out multiple measurements with the movement of poles and zeros. For instance, an LTI system has two poles, p_1 and p_2 , and it can also move p_2 to $2p_2$. In this case, we can do two measurements with p_2 and $2p_2$ respectively. The two time delays obtained are

$$t_{d1} = \frac{1}{p_1} + \frac{1}{p_2} \quad (7.4)$$

$$t_{d2} = \frac{1}{p_1} + \frac{1}{2p_2} \quad (7.5)$$

Solve the equation, we can get

$$p_1 = \frac{1}{2t_{d2} - t_{d1}} \quad (7.6)$$

$$p_2 = \frac{1}{2(t_{d1} - t_{d2})} \quad (7.7)$$

More measurements can be made to obtain more pole and zero locations, following the same rule.

REFERENCES

- [1] International Roadmap Committee, *Test and Test Equipment, International technology roadmap for semiconductors 2.0*. [Online]. Available: https://www.semiconductors.org/clientuploads/Research_Technology/ITRS/2015/ITWGs/0_2015%20ITRS%202.0%20Test%20.pdf.
- [2] L.-T. Wang, C. E. Stroud, and N. A. Touba, *System-on-chip test architectures: nanometer design for testability*. Morgan Kaufmann, 2010.
- [3] M. Mendez-Rivera, J. Silva-Martinez, and E. Sanchez-Sinencio, "On-chip spectrum analyzer for built-in testing analog ICs," in *IEEE International Symposium on Circuits and Systems (ISCAS)*, vol. 5, pp. V-61-V-64 vol.5, IEEE, 2002.
- [4] M. G. Méndez-Rivera, A. Valdes-Garcia, J. Silva-Martinez, and E. Sánchez-Sinencio, "An on-chip spectrum analyzer for analog built-in testing," *Journal of Electronic Testing*, vol. 21, no. 3, pp. 205–219, 2005.
- [5] A. P. Jose, K. A. Jenkins, and S. K. Reynolds, "On-chip spectrum analyzer for analog built-in self test," in *IEEE VLSI Test Symposium (VTS)*, pp. 131–136, May 2005.
- [6] P. Shoghi, T. P. Weldon, and C. J. Barnwell, "Experimental results for a successive detection log video amplifier in a single-chip frequency synthesized radio frequency spectrum analyzer," in *IEEE Southeastcon 2009*, pp. 379–382, IEEE, March 2009.
- [7] K. Nose and M. Mizuno, "A 0.016 mm², 2.4 GHz RF signal quality measurement macro for RF test and diagnosis," *IEEE Journal of Solid-State Circuits*,

- vol. 43, pp. 1038–1046, apr 2008.
- [8] D. Vázquez, G. Huertas, G. Leger, E. Peralías, A. Rueda, and J. Huertas, “On-chip evaluation of oscillation-based-test output signals for switched-capacitor circuits,” *Analog Integrated Circuits and Signal Processing*, vol. 33, no. 2, pp. 201–211, 2002.
- [9] D. Vázquez, G. Huertas, A. Luque, M. Barragán, G. Leger, A. Rueda, and J. Huertas, “Sine-wave signal characterization using square-wave and $\Sigma\Delta$ -modulation: Application to mixed-signal BIST,” *Journal of Electronic Testing*, vol. 21, no. 3, pp. 221–232, 2005.
- [10] E. Alon, V. Abramzon, B. Nezamfar, and M. Horowitz, “On-die power supply noise measurement techniques,” *IEEE Transactions on Advanced Packaging*, vol. 32, pp. 248–259, May 2009.
- [11] K. Sankaragomathi, W. Smith, B. Otis, and V. Sathe, “A deterministic-dither-based, all-digital system for on-chip power supply noise measurement,” in *IEEE/ACM International Symposium on Low Power Electronics and Design (ISLPED)*, pp. 283–286, Aug 2014.
- [12] S. Pant and D. Blaauw, “Circuit techniques for suppression and measurement of on-chip inductive supply noise,” in *European Solid-State Circuits Conference (ESSCIRC)*, pp. 134–137, Sept 2008.
- [13] C. Iorga, Y.-C. Lu, and R. Dutton, “A built-in technique for measuring substrate and power-supply digital switching noise using PMOS-based differential sensors and a waveform sampler in system-on-chip applications,” *IEEE Transactions on Instrumentation and Measurement*, vol. 56, pp. 2330–2337, Dec 2007.

- [14] E. Alon, V. Stojanovic, and M. Horowitz, “Circuits and techniques for high-resolution measurement of on-chip power supply noise,” *IEEE Journal of Solid-State Circuits*, vol. 40, pp. 820–828, April 2005.
- [15] J. M. W. Rogers and C. Plett, *Radio Frequency Integrated Circuit Design*. Norwood, MA, USA: Artech House, Inc., 1st ed., 2003.
- [16] S. Ahmad, K. Azizi, I. Zadeh, and J. Dabrowski, “Two-tone PLL for on-chip IP3 test,” in *IEEE International Symposium on Circuits and Systems (ISCAS)*, pp. 3549–3552, May 2010.
- [17] M. Barragan, R. Fiorelli, D. Vazquez, A. Rueda, and J. Huertas, “On-chip characterisation of RF systems based on envelope response analysis,” *IEE Electronics Letters*, vol. 46, pp. 36–38, January 2010.
- [18] H. Chauhan, Y. Choi, M. Onabajo, I.-S. Jung, and Y.-B. Kim, “Accurate and efficient on-chip spectral analysis for built-in testing and calibration approaches,” *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 22, pp. 497–506, March 2014.
- [19] Y. Choi, C. h. Chang, I. S. Jung, M. Onabajo, and Y. B. Kim, “A built-in calibration system with a reduced FFT engine for linearity optimization of low power LNA,” in *IEEE Int. Symp. Defect and Fault Tolerance in VLSI and Nanotechnology Systems (DFT)*, pp. 222–227, Oct 2014.
- [20] H. Graeb, S. Zizala, J. Eckmueller, and K. Antreich, “The sizing rules method for analog integrated circuit design,” in *IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*, pp. 343–349, Nov 2001.
- [21] W. Sansen, “1.3 analog CMOS from 5 micrometer to 5 nanometer,” in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, pp. 1–6, Feb 2015.

- [22] H. Iwai, “Si roadmap for 22 nm and beyond (invited paper),” *Microelectronic Engineering*, vol. 86, no. 7-9, pp. 1520–1528, 2009.
- [23] E. Sanchez-Sinencio and M. Majewski, “A nonlinear macromodel of operational amplifiers in the frequency domain,” *IEEE Transactions on Circuits and Systems*, vol. 26, pp. 395–402, Jun 1979.
- [24] H. Liu, A. Singhee, R. Rutenbar, and L. Carley, “Remembrance of circuits past: macromodeling by data mining in large analog design spaces,” in *Design Automation Conference (DAC)*, pp. 437–442, 2002.
- [25] R. Rutenbar, G. Gielen, and J. Roychowdhury, “Hierarchical modeling, optimization, and synthesis for system-level analog and RF designs,” *Proceedings of the IEEE*, vol. 95, pp. 640–669, March 2007.
- [26] H. Y. Koh, C. Sequin, and P. Gray, “OPASYN: a compiler for CMOS operational amplifiers,” *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 9, pp. 113–125, Feb 1990.
- [27] G. Peiwang and M. Xiaoqing, “A sensitivity-based heuristic search for constrained optimization in complex systems,” *Journal of Systems Engineering and Electronics*, vol. 10, pp. 75–80, March 1999.
- [28] G. Gielen, H. Walscharts, and W. Sansen, “Analog circuit design optimization based on symbolic simulation and simulated annealing,” *IEEE Journal of Solid-State Circuits*, vol. 25, pp. 707–713, Jun 1990.
- [29] J. Trejos, W. Castillo, J. González, and M. Villalobos, *Data Analysis, Classification, and Related Methods*, ch. Application of Simulated Annealing in some Multidimensional Scaling Problems, pp. 297–302. Berlin, Heidelberg: Springer Berlin Heidelberg, 2000.

- [30] I. Guerra-Gomez, E. Tlelo-Cuautle, C. Reyes-Garcia, G. Reyes-Salgado, and L. de la Fraga, "Non-sorting genetic algorithm in the optimization of unity-gain cells," in *International Conference on Electrical Engineering, Computing Science and Automatic Control (CCE)*, pp. 1–6, Jan 2009.
- [31] J. Zhang, H.-H. Chung, A.-L. Lo, and T. Huang, "Extended ant colony optimization algorithm for power electronic circuit design," *IEEE Transactions on Power Electronics*, vol. 24, pp. 147–162, Jan 2009.
- [32] R. Thakker, M. Baghini, and M. Patil, "Low-power low-voltage analog circuit design using hierarchical particle swarm optimization," in *International Conference on VLSI Design*, pp. 427–432, Jan 2009.
- [33] F. De Bernardinis, M. Jordan, and A. SangiovanniVincentelli, "Support vector machines for analog circuit performance representation," in *Design Automation Conferences (DAC)*, pp. 964–969, June 2003.
- [34] N. Kahraman and T. Yildirim, "Technology independent circuit sizing for fundamental analog circuits using artificial neural networks," in *Ph.D. Research in Microelectronics and Electronics (PRIME)*, pp. 1–4, June 2008.
- [35] X. Li, B. Taylor, Y. Chien, and L. Pileggi, "Adaptive post-silicon tuning for analog circuits: concept, analysis and optimization," in *IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*, pp. 450–457, Nov 2007.
- [36] S. Ray and B.-S. Song, "A 13-b linear, 40-MS/s pipelined ADC with self-configured capacitor matching," *IEEE Journal of Solid-State Circuits*, vol. 42, pp. 463–474, March 2007.

- [37] D. Han, B. S. Kim, and A. Chatterjee, “DSP-driven self-tuning of RF circuits for process-induced performance variability,” *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 18, pp. 305–314, Feb 2010.
- [38] D. Maliuk and Y. Makris, “On-chip intelligence: A pathway to self-testable, tunable, and trusted analog/RF ICs,” in *IEEE International Midwest Symposium on Circuits and Systems (MWSCAS)*, pp. 1077–1080, Aug 2014.
- [39] Keysight Technologies, *33600A Series Trueform Waveform Generators Data Sheet*, 2014. [Online]. Available: <http://literature.cdn.keysight.com/litweb/pdf/5991-3272EN.pdf>.
- [40] C. Shi, H. Yang, H. Xiao, J. Liu, and H. Liao, “A dual loop dual VCO CMOS PLL using a novel coarse tuning technique for DTV,” in *International Conference on Solid-State and Integrated-Circuit Technology (ICSICT)*, pp. 1597–1600, Oct 2008.
- [41] W. San-Um and T. Masayoshi, “Impulse signal generation and measurement technique for cost-effective built-in self test in analog mixed-signal systems,” in *IEEE International Midwest Symposium on Circuits and Systems (MWSCAS)*, pp. 1195–1198, Aug 2009.
- [42] B. Xia, S. Yan, and E. Sanchez-Sinencio, “An RC time constant auto-tuning structure for high linearity continuous-time $\Sigma\Delta$ modulators and active filters,” *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 51, pp. 2179–2188, Nov 2004.
- [43] M. Soda, Y. Bando, S. Takaya, T. Ohkawa, T. Takaramoto, T. Yamada, S. Kumashiro, T. Mogami, and M. Nagata, “On-chip sine-wave noise generator for analog IP noise tolerance measurements,” in *IEEE Asian Solid State Circuits Conference (A-SSCC)*, pp. 1–4, Nov 2010.

- [44] C. Shi and E. Sanchez-Sinencio, "150-850 MHz high-linearity sine-wave synthesizer architecture based on FIR filter approach and SFDR optimization," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 62, pp. 2227–2237, Sept 2015.
- [45] Keysight Technologies, *NFA X-Series Noise Figure Analyzer, Multi-touch N8973B, N8974B, N8975B, N8976B Data sheet*, 2016. [Online]. Available: <http://literature.cdn.keysight.com/litweb/pdf/5992-1270EN.pdf?id=2702444>.
- [46] S. Wang and M. Tehranipoor, "Light-weight on-chip structure for measuring timing uncertainty induced by noise in integrated circuits," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 22, pp. 1030–1041, May 2014.
- [47] M. Onabajo, J. Altet, E. Aldrete-Vidrio, D. Mateo, and J. Silva-Martinez, "Electrothermal design procedure to observe RF circuit power and linearity characteristics with a homodyne differential temperature sensor," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 58, pp. 458–469, March 2011.
- [48] Y. Chen, R. C. Jaeger, and J. C. Suhling, "Delta-Sigma based CMOS stress sensor with RF output," in *IEEE Asian Solid-State Circuits Conference (A-SSCC)*, pp. 243–246, Nov 2006.
- [49] S. S. Haykin and M. Moher, *Communication Systems, 4E*, pp. 95–98. Hoboken, N.J. : John Wiley & Sons, 2001.
- [50] J. Jin, Y. Gao, and E. Sanchez-Sinencio, "An energy-efficient time-domain asynchronous 2b/step SAR ADC with a hybrid R-2R/C-3C DAC structure," *IEEE Journal of Solid-State Circuits*, vol. 49, pp. 1383–1396, June 2014.

- [51] A. Valdes-Garcia, F. A. L. Hussien, J. Silva-Martinez, and E. Sanchez-Sinencio, "An integrated frequency response characterization system with a digital interface for analog testing," *IEEE Journal of Solid-State Circuits*, vol. 41, pp. 2301–2313, Oct 2006.
- [52] J. E. Volder, "The CORDIC trigonometric computing technique," *IRE Transactions on Electronic Computers*, vol. EC-8, pp. 330–334, Sept 1959.
- [53] J. Wang, C. Shi, E. Sanchez-Sinencio, and J. Hu, "Built-in self optimization for variation resilience of analog filters," in *IEEE Computer Society Annual Symposium on VLSI (ISVLSI)*, pp. 656–661, July 2015.
- [54] B. Wu and Y. Chiu, "A 40 nm CMOS derivative-free IF active-RC BPF with programmable bandwidth and center frequency achieving over 30 dbm IIP3," *IEEE Journal of Solid-State Circuits*, vol. 50, pp. 1772–1784, Aug 2015.
- [55] P. Kallam, E. Sanchez-Sinencio, and A. Karsilayan, "An enhanced adaptive Q-tuning scheme for a 100-MHz fully symmetric OTA-based bandpass filter," *IEEE Journal of Solid-State Circuits*, vol. 38, pp. 585–593, Apr 2003.
- [56] L. Thomas, "The biquad: Part i-some practical design considerations," *IEEE Transactions on Circuit Theory*, vol. 18, pp. 350–357, May 1971.
- [57] L. Ye, C. Shi, H. Liao, R. Huang, and Y. Wang, "Highly power-efficient active-RC filters with wide bandwidth-range using low-gain push-pull Opamps," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 60, pp. 95–107, Jan 2013.
- [58] P. Allen and D. Holberg, *CMOS Analog Circuit Design*. The Oxford Series in Electrical and Computer Engineering, OXFORD UNIV PR, 2011.

- [59] E. Sanchez-Sinencio and J. Silva-Martinez, "CMOS transconductance amplifiers, architectures and active filters: a tutorial," *IEE Proceedings - Circuits, Devices and Systems*, vol. 147, pp. 3–12, Feb 2000.
- [60] C. Azeredo-Leme, "Clock jitter effects on sampling: A tutorial," *IEEE Circuits and Systems Magazine*, vol. 11, pp. 26–37, Aug. 2011.
- [61] D.-U. Lee, A. Gaffar, R. Cheung, O. Mencer, W. Luk, and G. Constantinides, "Accuracy-guaranteed bit-width optimization," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 25, pp. 1990–2000, Oct 2006.
- [62] C. L. Wei, Y. W. Wang, and B. D. Liu, "Wide-range filter-based sinusoidal wave synthesizer for electrochemical impedance spectroscopy measurements," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 8, pp. 442–450, June 2014.
- [63] J. Wibbenmeyer and C. I. H. Chen, "Built-in self-test for low-voltage high-speed analog-to-digital converters," *IEEE Transactions on Instrumentation and Measurement*, vol. 56, pp. 2748–2756, Dec 2007.
- [64] G. Banerjee, M. Behera, M. A. Zeidan, R. Chen, and K. Barnett, "Analog/RF built-in-self-test subsystem for a mobile broadcast video receiver in 65-nm CMOS," *IEEE Journal of Solid-State Circuits*, vol. 46, pp. 1998–2008, Sept 2011.
- [65] S. H. W. Chiang, H. Sun, and B. Razavi, "A 10-bit 800-MHz 19-mW CMOS ADC," *IEEE Journal of Solid-State Circuits*, vol. 49, pp. 935–949, April 2014.
- [66] M. M. Elsayed and E. Sanchez-Sinencio, "A low THD, low power, high output-swing time-mode-based tunable oscillator via digital harmonic-cancellation

- technique,” *IEEE Journal of Solid-State Circuits*, vol. 45, pp. 1061–1071, May 2010.
- [67] F. Bahmani and E. Sanchez-Sinencio, “Low THD bandpass-based oscillator using multilevel hard limiter,” *IET Circuits, Devices and Systems*, vol. 1, pp. 151–160, April 2007.
- [68] S. W. Park, J. L. Ausin, F. Bahmani, and E. Sanchez-Sinencio, “Nonlinear shaping SC oscillator with enhanced linearity,” *IEEE Journal of Solid-State Circuits*, vol. 42, pp. 2421–2431, Nov 2007.
- [69] B. K. Vasan, S. K. Sudani, D. J. Chen, and R. L. Geiger, “Low-distortion sine wave generation using a novel harmonic cancellation technique,” *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 60, pp. 1122–1134, May 2013.
- [70] P. Aluthwala, N. Weste, A. Adams, T. Lehmann, and S. Parameswaran, “A simple digital architecture for a harmonic-cancelling sine-wave synthesizer,” in *IEEE International Symposium on Circuits and Systems (ISCAS)*, pp. 2113–2116, June 2014.
- [71] M. J. Barragan, G. Leger, D. Vazquez, and A. Rueda, “On-chip sinusoidal signal generation with harmonic cancelation for analog and mixed-signal BIST applications,” *Analog Integrated Circuits and Signal Processing*, vol. 82, no. 1, pp. 67–79, 2015.
- [72] T. Yoo, H. C. Yeoh, Y. H. Jung, S. J. Cho, Y. S. Kim, S. M. Kang, and K. H. Baek, “A 2 GHz 130 mW direct-digital frequency synthesizer with a nonlinear DAC in 55 nm CMOS,” *IEEE Journal of Solid-State Circuits*, vol. 49, pp. 2976–2989, Dec 2014.

- [73] C. Y. Yang, J. H. Weng, and H. Y. Chang, "A 5-GHz direct digital frequency synthesizer using an analog-sine-mapping technique in 0.35- μm SiGe BiCMOS," *IEEE Journal of Solid-State Circuits*, vol. 46, pp. 2064–2072, Sept 2011.
- [74] H. C. Yeoh, J. H. Jung, Y. H. Jung, and K. H. Baek, "A 1.3-GHz 350-mW hybrid direct digital frequency synthesizer in 90-nm CMOS," *IEEE Journal of Solid-State Circuits*, vol. 45, pp. 1845–1855, Sept 2010.
- [75] Wikipedia, *Fourier Series*. [Online]. Available: https://en.wikipedia.org/wiki/Fourier_series.
- [76] M. M. Abdul-Latif, M. M. Elsayed, and E. Sanchez-Sinencio, "A wideband millimeter-wave frequency synthesis architecture using multi-order harmonic-synthesis and variable N-push frequency multiplication," *IEEE Journal of Solid-State Circuits*, vol. 46, pp. 1265–1283, June 2011.
- [77] T. Saramaki, "Finite impulse response filter design," *Handbook for digital signal processing*, pp. 155–277, 1993.
- [78] B. Nauta, "A cmos transconductance-C filter technique for very high frequencies," *IEEE Journal of Solid-State Circuits*, vol. 27, pp. 142–153, Feb 1992.
- [79] M. R. Jan, C. Anantha, and N. Borivoje, *Digital Integrated Circuits—A Design Perspective*. Pearson Education, 2003.
- [80] G. Yu, "Min-max optimization of several classical discrete optimization problems," *Journal of Optimization Theory and Applications*, vol. 98, no. 1, pp. 221–242, 1998.
- [81] A. A. Abidi, "Phase noise and jitter in CMOS ring oscillators," *IEEE Journal of Solid-State Circuits*, vol. 41, pp. 1803–1816, Aug 2006.

- [82] Cadence Forum, *How to generate a clock signal with random noise in Cadence Spectre?* [Online]. Available: https://community.cadence.com/cadence_technology_forums/f/38/t/33222.
- [83] M. Mobarak, M. Onabajo, J. Silva-Martinez, and E. Sanchez-Sinencio, "Attenuation-predistortion linearization of CMOS OTAs with digital correction of process variations in OTA-C filter applications," *IEEE Journal of Solid-State Circuits*, vol. 45, pp. 351–367, Feb 2010.
- [84] S. C. Lee and Y. Chiu, "A 15-MHz bandwidth 1-0 MASH $\Sigma\Delta$ ADC with nonlinear memory error calibration achieving 85-dBc SFDR," *IEEE Journal of Solid-State Circuits*, vol. 49, pp. 695–707, March 2014.
- [85] D. Zhao and P. Reynaert, "A 60-ghz dual-mode class AB power amplifier in 40-nm CMOS," *IEEE Journal of Solid-State Circuits*, vol. 48, pp. 2323–2337, Oct 2013.
- [86] J. S. Daniels, E. P. Anderson, T. H. Lee, and N. Pourmand, "Simultaneous measurement of nonlinearity and electrochemical impedance for protein sensing using two-tone excitation," in *IEEE Annual International Conference of the Engineering in Medicine and Biology Society*, pp. 5753–5756, Aug 2008.
- [87] J. R. Wilkerson, I. M. Kilgore, K. G. Gard, and M. B. Steer, "Passive intermodulation distortion in antennas," *IEEE Transactions on Antennas and Propagation*, vol. 63, pp. 474–482, Feb 2015.
- [88] W. T. Lin and T. H. Kuo, "A compact dynamic-performance-improved current-steering DAC with random rotation-based binary-weighted selection," *IEEE Journal of Solid-State Circuits*, vol. 47, pp. 444–453, Feb 2012.

- [89] W. T. Lin, H. Y. Huang, and T. H. Kuo, "A 12-bit 40 nm DAC achieving SFDR > 70 dB at 1.6 GS/s and IMD < 61 db at 2.8 GS/s with DEMDRZ technique," *IEEE Journal of Solid-State Circuits*, vol. 49, pp. 708–717, March 2014.
- [90] E. Bechthum, G. I. Radulov, J. Briaire, G. J. G. M. Geelen, and A. H. M. van Roermund, "A wideband RF mixing-DAC achieving IMD < -82 dbc up to 1.9 GHz," *IEEE Journal of Solid-State Circuits*, vol. 51, pp. 1374–1384, June 2016.
- [91] S. Su and M. S. W. Chen, "27.1 a 12b 2GS/s dual-rate hybrid DAC with pulsed timing-error pre-distortion and in-band noise cancellation achieving >74 dBc SFDR up to 1GHz in 65nm CMOS," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, pp. 456–457, Jan 2016.
- [92] G. I. Radulov, P. J. Quinn, and A. H. M. van Roermund, "A 28-nm CMOS 1 V 3.5 GS/s 6-bit DAC with signal-independent delta-I noise DfT scheme," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 23, pp. 44–53, Jan 2015.
- [93] G. I. Radulov, P. J. Quinn, and A. H. M. van Roermund, "A 28-nm CMOS 7-GS/s 6-bit DAC with DfT clock and memory reaching SFDR >50 dB up to 1 GHz," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 23, pp. 1941–1945, Sept 2015.
- [94] B. W. Cook, A. Berny, A. Molnar, S. Lanzisera, and K. S. J. Pister, "Low-power 2.4-GHz transceiver with passive RX front-end and 400-mV supply," *IEEE Journal of Solid-State Circuits*, vol. 41, pp. 2757–2766, Dec 2006.
- [95] A. M. Abo and P. R. Gray, "A 1.5-V, 10-bit, 14.3-MS/s CMOS pipeline analog-to-digital converter," *IEEE Journal of Solid-State Circuits*, vol. 34, pp. 599–

606, May 1999.

- [96] E. Bechthum, G. Radulov, J. Briaire, G. Geelen, and A. van Roermund, “Classification for synthesis of high spectral purity current-steering mixing-DAC architectures,” *Analog Integrated Circuits and Signal Processing*, vol. 85, no. 3, pp. 497–504, 2015.
- [97] A. Eroglu, “Non-invasive quadrature modulator balancing method to optimize image band rejection,” *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 61, pp. 600–612, Feb 2014.
- [98] M. J. M. Pelgrom, A. C. J. Duinmaijer, and A. P. G. Welbers, “Matching properties of MOS transistors,” *IEEE Journal of Solid-State Circuits*, vol. 24, pp. 1433–1439, Oct 1989.
- [99] I. Galton, “Why dynamic-element-matching DACs work,” *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 57, pp. 69–74, Feb 2010.
- [100] P. D. Aluthwala, N. Weste, A. Adams, T. Lehmann, and S. Parameswaran, “Partial dynamic element matching technique for digital-to-analog converters used for digital harmonic-cancelling sine-wave synthesis,” *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. PP, no. 99, pp. 1–14, 2016.
- [101] Wikipedia, *Linear-feedback shift register*. [Online]. Available: https://en.wikipedia.org/wiki/Linear-feedback_shift_register.
- [102] B. Razavi, “Challenges in the design high-speed clock and data recovery circuits,” *IEEE Communications Magazine*, vol. 40, pp. 94–101, Aug 2002.
- [103] H. Khatri, P. S. Gudem, and L. E. Larson, “Distortion in current commutating passive CMOS downconversion mixers,” *IEEE Transactions on Microwave Theory and Techniques*, vol. 57, pp. 2671–2681, Nov 2009.

- [104] S. M. Taleie, T. Copani, B. Bakkaloglu, and S. Kiaei, “A linear $\Sigma - \Delta$ digital IF to RF DAC transmitter with embedded mixer,” *IEEE Transactions on Microwave Theory and Techniques*, vol. 56, pp. 1059–1068, May 2008.
- [105] B. Malki, B. Verbruggen, E. Martens, P. Wambacq, and J. Craninckx, “A 150 kHz-80 MHz BW discrete-time analog baseband for software-defined-radio receivers using a 5th-order IIR LPF, active FIR and a 10 bit 300 MS/s ADC in 28 nm cmos,” *IEEE Journal of Solid-State Circuits*, vol. 51, pp. 1593–1606, July 2016.
- [106] R. Bagheri, A. Mirzaei, S. Chehrazi, M. E. Heidari, M. Lee, M. Mikhemar, W. Tang, and A. A. Abidi, “An 800-MHz-6-GHz software-defined wireless receiver in 90-nm CMOS,” *IEEE Journal of Solid-State Circuits*, vol. 41, pp. 2860–2876, Dec 2006.
- [107] M. F. Huang, M. C. Kuo, T. Y. Yang, and X. L. Huang, “A 58.9-dB ACR, 85.5-dB SBA, 5–26-MHz configurable-bandwidth, charge-domain filter in 65-nm CMOS,” *IEEE Journal of Solid-State Circuits*, vol. 48, pp. 2827–2838, Nov 2013.
- [108] S. Lindfors, A. Parssinen, and K. A. I. Halonen, “A 3-V 230-MHz CMOS decimation subsampler,” *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, vol. 50, pp. 105–117, Mar 2003.
- [109] D. Jakonis, K. Folkesson, J. Dbrowski, P. Eriksson, and C. Svensson, “A 2.4-GHz RF sampling receiver front-end in 0.18- μm CMOS,” *IEEE Journal of Solid-State Circuits*, vol. 40, pp. 1265–1277, June 2005.
- [110] P. K. Prakasam, M. Kulkarni, X. Chen, Z. Yu, S. Hoyos, J. Silva-Martinez, and E. Sanchez-Sinencio, “Applications of multipath transform-domain charge-

- sampling wide-band receivers,” *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 55, pp. 309–313, April 2008.
- [111] M. F. Huang and T. L. Chiu, “A quadrature charge-domain filter with frequency downconversion for rf receivers,” *IEEE Transactions on Microwave Theory and Techniques*, vol. 58, pp. 1323–1332, May 2010.
- [112] Y. Zhou, N. M. Filiol, and F. Yuan, “A quadrature charge-domain sampling mixer with embedded FIR, IIR, and N-Path filters,” *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 62, pp. 1431–1440, May 2015.
- [113] C. Park, J. Yoon, and B. Kim, “Non-decimation FIR filter for digital RF sampling receiver with wideband operation capability,” in *IEEE Radio Frequency Integrated Circuits Symposium (RFIC)*, pp. 487–490, June 2009.
- [114] S. H. Shin, S. J. Kweon, S. H. Jo, Y. C. Choi, S. Lee, and H. J. Yoo, “A 0.7-MHz 10-MHz CT+DT hybrid baseband chain with improved passband flatness for LTE application,” *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 62, pp. 244–253, Jan 2015.
- [115] R. Fiorelli, A. Arnaud, and C. Galup-Montoro, “Series-parallel association of transistors for the reduction of random offset in non-unity gain current mirrors,” in *IEEE International Symposium on Circuits and Systems (ISCAS)*, vol. 1, pp. I-881–4 Vol.1, May 2004.
- [116] A. Mirzaei, S. Chehrazi, R. Bagheri, and A. A. Abidi, “Analysis of first-order anti-aliasing integration sampler,” *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 55, pp. 2994–3005, Nov 2008.
- [117] F. F. Dai, C. Stroud, and D. Yang, “Automatic linearity and frequency response tests with built-in pattern generator and analyzer,” *IEEE Transactions on Very*

- Large Scale Integration (VLSI) Systems*, vol. 14, pp. 561–572, June 2006.
- [118] G. Huertas, D. Vazquez, E. J. Peralias, A. Rueda, and J. L. Huertas, “Testing mixed-signal cores: a practical oscillation-based test in an analog macrocell,” *IEEE Design Test of Computers*, vol. 19, pp. 73–82, Nov 2002.
- [119] Z. Czaja, “Self-testing of analog parts terminated by ADCs based on multiple sampling of time responses,” *IEEE Transactions on Instrumentation and Measurement*, vol. 62, pp. 3160–3167, Dec 2013.
- [120] A. Balivada, J. Chen, and J. Abraham, “Analog testing with time response parameters,” *IEEE Design Test of Computers*, vol. 13, pp. 18–25, Summer 1996.
- [121] R. Vasudevamurthy, P. K. Das, and B. Amrutur, “Time-based all-digital technique for analog built-in self-test,” *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 22, pp. 334–342, Feb 2014.