

TRENDS IN ANALOG IC DESIGN: HIGHLY RECONFIGURABLE FILTERS,
PERFORMANCE OPTIMIZATION, AND IP PROTECTION

A Dissertation

by

ADRIANA CAROLINA SANABRIA BORBÓN

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, Samuel Palermo
Jiang Hu
Douglas Allaire
Head of Department, Miroslav M. Begovic

May 2021

Major Subject: Electrical Engineering

Copyright 2021 Adriana Carolina Sanabria Borbón

ABSTRACT

The continuous technology scaling and rapid growth of applications involving a vast and diverse network of interconnected devices increase analog integrated circuit (IC) design complexity. This work addresses three main trends of analog IC design: highly reconfigurable power-efficient analog circuits, automatic IC design for performance optimization, and analog IP protection against security threats.

The first part of this dissertation discusses the synthesis and design methodology of high-order and frequency-tunable low-pass active-R filter architectures for multi-standard wireless applications. Active-R filters use the inherent integrator-like behavior of amplifiers to realize their frequency response. The main advantages of this type of filter are high-frequency performance and a low integrated area since the only capacitor they require is the Miller capacitor used in internally compensated amplifiers. In this work, amplifiers with configurable unity-gain frequencies enable the continuous tuning of active-R filters. Three different filter architectures realize a fifth-order Butterworth prototype tunable in the 1–50 MHz frequency range.

The second part of this dissertation discusses the development of a computationally low-cost surrogate model for multi-objective optimization-based automated analog IC design. The surrogate has three main components: a set of Gaussian process regression models of the technology's parameters, a physics-based model of the MOSFET device, and a set of equations of the performance metrics of the circuit under design. The surrogate model is inserted into two different state-of-the-art optimization algorithms to prove its flexibility. The efficacy of our surrogate is demonstrated through simulation validation across process corners in three different CMOS technologies, using three representative circuit building-blocks that are commonly encountered in mainstream analog/RF ICs.

Finally, this dissertation presents an overview of analog IP security, including the threat models, protection techniques, and reported attacks. A novel Schmitt-trigger based key provisioning technique is proposed for increasing the security level of existing IP protection techniques. This

approach has a very small area overhead that remains constant and independent of the key size. Moreover, it consumes power only at power-up.

DEDICATION

To my parents and my husband for their love, encouragement, and support.

ACKNOWLEDGMENTS

I want to express my gratitude to everyone that supported me in the pursuit of this degree.

To my advisor Edgar Sanchez-Sinencio for his guidance and motivation. To my committee: professors Jiang Hu, Samuel Palermo, and Douglas Allaire for your valuable feedback. To my fellow graduate students Sergio, Congyin, Jiafan, Johan, Hatem, Joseph, Sanghoon, Fernando, Johan, Guillermo, and Zizhen our technical discussions and collaboration in different projects. To Shankari and professors Jiang hu and JV Rajendran for introducing me to the world of hardware security and improving my writing skills. To my former research advisors Jaime Vitola, Cesar Pedraza, and Esteban Tlelo-Cuautle for inspiring me to do research.

To my husband Suraj Prakash for his encouragement and companionship throughout this path. To my parents Pedro and Nancy, my brothers Pablo and Marco, my goddaughter Danna, and to all my family for believing on me. To my best friend Erika Garcia-Villatoro for her moral support.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supported by a dissertation committee consisting of Professors Edgar Sánchez-Sinenecio, Jiang Hu, and Samuel Palermo of the Department of Electrical and Computer Engineering and Professor Douglas Allaire of the Department of Mechanical engineering.

The circuit analyses and discussion depicted in Chapters 2 and 3 was enriched by the discussion with student members of the Analog and mixed signal group Suraj Prakash, Sergio Soto, Fernando Lavalle, Joseph Riad, and Hatem Osman at the Electrical and Computer Engineering Department. The results reported in Chapter 5 were generated in collaboration with Sergio Soto and Professor Douglas Allaire of the Department of Mechanical engineering. The research results reported in Chapters 6 and 7 were developed in collaboration with Professors Jiang Hu and JV Rajendran and my fellow students Nithyashankari G. Jayasankaran and Sandghoon Lee of the Department of Electrical and Computer Engineering.

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

Funding Sources

Graduate study was supported by a scholarship from Colciencias and a fellowship from Texas Instruments Inc. One research project had partial support from Qualcomm, Silicon, Labs and Texas Instruments. While the research project on hardware security was supported by the NSF grant 1815583.

TABLE OF CONTENTS

	Page
ABSTRACT	ii
DEDICATION	iv
ACKNOWLEDGMENTS	v
CONTRIBUTORS AND FUNDING SOURCES	vi
TABLE OF CONTENTS	vii
LIST OF FIGURES	xi
LIST OF TABLES.....	xvii
1. INTRODUCTION.....	1
2. AMPLIFIER MACROMODELING	4
2.1 General gain, pole, zero amplifier model	4
2.2 Transistor level OTA implementation and its modeling.....	6
2.2.1 Small-signal macromodel	7
2.2.2 Effect of large-signal non-idealities of the OTA	10
2.2.3 Frequency vs. power trade-off	11
3. CONTINUOUS-TIME ACTIVE FILTERS.....	12
3.1 Introduction.....	12
3.2 Active-RC filters	14
3.3 Gm -C filters	15
3.3.1 Effect of real operational transconductance amplifiers in Gm -C filters	17
3.3.2 Noise in Gm -C filters	17
3.4 Active-R filters	18
3.4.1 Effect of the amplifier's small-signal non-idealities in active-R filter imple- mentations	20
3.4.2 Active-R filter tuning.....	23
3.4.3 Linearity and noise trade-off in the second-order active-R filters	23
3.5 Gm -C vs Active-R filters	25
4. SYNTHESIS OF HIGH-ORDER CONTINUOUSLY TUNABLE LOW-PASS ACTIVE- R FILTERS FOR BASE-BAND COMMUNICATIONS.....	26

4.1	Introduction and motivation	26
4.2	High-order active-R filter implementation	27
4.2.1	Synthesis of a cascade of biquads (CoB) low-pass active-R filter	27
4.2.2	Synthesis of multiple-loop feedback (MLF) active-R low-pass filters	29
4.2.2.1	Follow the leader feedback (FLFB)	29
4.2.2.2	Inverse follow the leader feedback (IFLFB)	32
4.3	Circuit implementation	33
4.3.1	Amplifier with tunable ω_t	33
4.3.2	Common-mode feedback (CMFB) circuit	36
4.3.3	Resistor design	37
4.4	Experimental results	39
4.4.1	Experimental setup	39
4.4.2	Measurement results of the second-order active-R Q tuning	41
4.4.3	Measurement results of f_o tuning of the three fifth-order active-R architectures	41
4.4.4	Performance comparison	43
4.4.5	Limitations of this work	47
4.5	Conclusion	47
5.	GAUSSIAN PROCESS BASED SURROGATE FOR OPTIMIZATION AIDED AND PROCESS VARIATIONS AWARE ANALOG CIRCUIT DESIGN	48
5.1	Introduction	48
5.1.1	Motivation	48
5.1.2	Our Approach and Contributions	50
5.2	Multi-Objective Constrained Optimization for Automatic Circuit Design	51
5.2.1	Multi-Objective Optimization	51
5.2.1.1	Gradient-Based Optimization Algorithms	52
5.2.1.2	Evolutionary Optimization Algorithms	53
5.2.2	Analog Circuit Design as an Optimization Problem	53
5.3	Proposed Surrogate Model for Optimization-Based EDA Tools	54
5.3.1	General Optimization Architecture	54
5.3.2	Advanced Compact MOSFET (ACM) Model	55
5.3.3	Gaussian Process-Based Regression Models of the Process Characterization	56
5.3.3.1	Characterization of the Parameters of CMOS Transistors	57
5.3.3.2	Gaussian-Processes-Based Regression Models	58
5.3.4	Circuit Performance Equation-Based Model	60
5.3.5	Process Variations-Aware Automatic Design	61
5.4	Experimental Results	61
5.4.1	Error of the GPR-Based Surrogate Model	61
5.4.2	Experimental Setup	62
5.4.3	Active-RC Second Order Filter	64
5.4.3.1	Surrogate of the Filter's Performance Metrics	66
5.4.3.2	Results of Filter's Automatic Design	67
5.4.4	Capacitor-Less Low-Dropout (CL-LDO) Voltage Regulator	69

5.4.4.1	Surrogate of the LDO's Performance Metrics	71
5.4.4.2	Results of the LDO's Automatic Design	73
5.4.5	Current-Starved Voltage Controlled Oscillator (CSVCO).....	77
5.4.5.1	Surrogate of the CSVCO's Performance Metrics	78
5.4.5.2	Results of the CSVCO's Automatic Design	79
5.4.6	Summary	80
5.5	Conclusions.....	84
6.	ANALOG IP PROTECTION TECHNIQUES	86
6.1	Introduction.....	86
6.2	Supply chain attacks	87
6.3	Defenses techniques to protect analog IP	88
6.3.1	Locking the bias generation circuits	89
6.3.1.1	Current biasing	89
6.3.1.2	Voltage biasing	91
6.3.2	Locking the gain transistors of analog circuits.....	93
6.3.3	Locking analog and mixed-signal (AMS) circuits	93
6.3.4	Camouflaging of the analog circuit's layout	94
6.4	Key provisioning techniques.....	96
6.5	Breaking Analog IP protection techniques.....	97
6.5.1	Brute-force attacks	97
6.5.2	Reverse engineering	98
6.5.3	Removal/bypass attacks	98
6.5.4	SAT-based attack [1,2]	98
6.5.5	SMT-based attack [3].....	98
6.5.6	Optimization based attacks [4].....	99
6.6	Conclusion and future research directions	99
7.	SCHMITT-TRIGGER BASED KEY PROVISIONING TECHNIQUE	101
7.1	Introduction.....	101
7.1.1	Related works on key provisioning techniques	102
7.1.2	Limitations of existing key provisioning techniques.....	104
7.1.3	Contributions of this work	104
7.2	Background.....	106
7.2.1	Schmitt trigger (ST)	106
7.2.1.1	Internal-feedback Schmitt Trigger (ST)	107
7.2.1.2	External-feedback Schmitt Trigger (ST)	107
7.2.2	Output transition probabilities of the non-inverting Schmitt trigger (ST)	108
7.2.3	Window comparator	108
7.2.3.1	Inverter-based window comparator	109
7.2.3.2	OpAmp-based window comparator	109
7.3	Proposed approach	109
7.3.1	Threat model.....	109
7.3.2	Schmitt trigger (ST)-based key provisioning.....	110

7.3.3	Security metrics	113
7.3.3.1	Key size	113
7.3.3.2	Resiliency against brute force attacks	113
7.3.3.3	Resiliency against reverse engineering attacks.....	114
7.3.3.4	Resiliency against SAT/SMT attacks.....	114
7.3.3.5	Probability distribution of the common key (CK)	114
7.3.3.6	Uniqueness of the user key (UK)	115
7.4	Results and discussion	116
7.4.1	Experimental setup	116
7.4.2	Effective size of the common key (CK)	116
7.4.3	Uniqueness of the user key (UK)	118
7.4.4	Power and area overhead	119
7.4.5	Robustness against process and temperature variations	121
7.4.6	Test cases with analog locks	122
7.4.6.1	Bandpass fourth-order Gm-C filter	122
7.4.6.2	Low-dropout (LDO) voltage regulator	124
7.4.6.3	Cascode common-gate low-noise amplifier (CG-LNA)	124
7.4.6.4	Overhead of the key provisioning on the test cases.....	125
7.5	Conclusion.....	126
8.	CONCLUSION.....	127
	REFERENCES	129

LIST OF FIGURES

FIGURE	Page
1.1 Problem statement: challenges on analog integrated circuit design.....	1
2.1 Op-Amp based inverting amplifier.....	5
2.2 Frequency response of the inverting amplifier using different models of the amplifier $A_1(s)$ open-loop response.	5
2.3 Effect of a finite open-loop amplifier's DC gain on the inverting amplifier response. (a) Transfer function comparison (b) Percentage error across frequencies for different DC gains.	6
2.4 Effect of the amplifier's non-dominant pole location on the inverting amplifier response. (a) Transfer function comparison (b) Percentage error across frequencies for different ratios ω_{p2}/ω_t	6
2.5 Two-stage Miller compensated OTA. (a) Transistor level schematic. (b) Small-signal model.	7
2.6 Small-signal macromodel of the closed-loop first order Active-R.....	7
2.7 Sensitivity of the closed-loop response of the inverting amplifier with respect to the small-signal circuit parameters.	9
2.8 Loop gain of closed loop first order Active-R	9
2.9 Amplifier estimated power consumption vs second pole location f_{p2}/f_{u1} . (a) Total current in μm at $f_C = 1MHz$. (b) Gain in dB at $f_C = 1MHz$. (c) Power consumption in μm at $f_C = 50MHz$. (d) Gain in dB at $f_C = 50MHz$	11
3.1 Active integrators (a) Active-RC (b) Gm-C (c) Gm-OTA-C (d) Active-R.....	13
3.2 Active-RC (a) lossless and (b) lossy integrators.	14
3.3 Second-order Tow-Thomas active-RC filter.....	14
3.4 Gm-C biquad implementation.	16

3.5	Magnitude and phase error of the active-R biquad's transfer function estimated for different approximations of the amplifier's transfer function. The single-pole model in which the pole location is ω_{p1} , the DC gain A_0 , and $\omega_t = \omega_{p1} \cdot A_0$ leads to the error in (a) Magnitude (b) Phase. The two pole transfer function where $x = \omega_{p2}/\omega_t$ yields an error in (c) Magnitude (d) Phase.....	21
3.6	Simulated noise and linearity trade-off of the biquad active-R filter. Four different combinations of ω_{t1} , ω_{t2} and R_2 that implement the same f_o and Q are compared in terms of dynamic range. In the optimal design occurs when $\omega_{t1} > \omega_{t2}$. A larger $gm_1 = \omega_{t1}/C_1$ reduces the input-referred noise and maximize the loop gain to reduce distortion.....	24
4.1	Multi-standard receiver. This project proposes a frequency tunable low-pass base-band filter (red).	26
4.2	Implementation of a fifth order filter by cascading first-order and second-order stages. The order of the stages does not change the filter transfer function, but instead affects its noise performance.	27
4.3	Fifth-order follow the leader feedback (FLFB) active-R filter: (a) Flow diagram and (b) fully differential circuit schematic.	30
4.4	Modeled input-referred noise transfer function of each stage of the fifth-order FLFB filter for $f_o = 1MHz$	31
4.5	Fifth-order inverse follow the leader feedback (IFLFB) active-R filter: (a) Flow diagram and (b) fully differential circuit schematic.	32
4.6	Modeled input-referred noise transfer function of each stage of the fifth-order IFLFB filter with $f_o=1 MHz$	33
4.7	Two-stage fully-differential Miller amplifier with programmable unity gain frequency ω_t . While the configurable input pair provides coarse tuning, the external biasing current I_B allows continuous tuning. A transmission gate-based multiplexor implements each switch. The common-mode feedback circuit includes passive sensing and a single-stage error amplifier.	34
4.8	Simulation results of the f_t tuning range. The configurable input differential pair provides coarse-tuning (S1, S2, S3, S4, S5, S6), while the fine-tuning is done by changing the biasing current (I_B). With the proposed design we achieve a 90x range which is sufficient for the frequency tunable active-R filter implementation. ...	36
4.9	Simulated frequency response (magnitude and phase) of the common-mode feedback loop. The plots show the loop-gain for the low and high frequency configurations when the smallest differential pair is selected.....	37

4.10	Effect of the resistor values in the fully-differential first-order active R across f_o . The percentage of integrated noise due to the resistors is shown (a). The RMS error of the transfer function due to different resistor values (b).	38
4.11	Digitally programmable resistor implementing the resistor R_2 that enables Q and ω_o tuning of the second-order active-R filter.	39
4.12	Chip microphotograph of the three fifth-order low-pass active-R filter topologies: the cascade of biquads (CoB), the follow the leader feedback (FLFB), and the inverse follow the leader feedback (IFLFB). The three structures use the same amplifier but differ on the resistive feedback configuration. The designs are fabricated using the TSMC 0.18 μm CMOS process.	40
4.13	Measurement setup for the filter performance characterization. The PCB includes baluns, voltage buffers, potentiometers, and 50 Ω SMA connectors. The measurement equipment includes a dual-channel signal generator and a spectrum/network analyzer. The NI-DAQ board generates the control signals for the scan-chains.	40
4.14	Measured Q tuning of a biquad active-R filter using the digitally programmable resistor R_2 for different ratios ω_{t2}/ω_{t1} . Each TF corresponds to a different configuration of C[2:0].	41
4.15	Measured transfer function of the fifth-order Butterworth active-R filters: (a) CoB, (b) FLFB, and (c) IFLFB. The coarse-tuning steps depend on the digitally-programmable input differential pair. Any cutoff frequency within these steps is achieved by changing the amplifiers' biasing current.	42
4.16	Measurement of the two-tone test of the fifth-order CoB filter at $f_o = 1$ MHz. We applied tones at $f_1 = 300$ kHz and $f_2 = 330$ kHz, generating intermodulation tones at 270 kHz and 360 kHz.	44
4.17	Measured iB-IIP3 (red), iB-IIP2 (black), and OoB-IIP3 (green) of the 5th order CoB filter at $f_o = 1$ MHz. The iB tests are performed with tones at $f_1 = 300$ kHz and $f_2 = 330$ kHz. The OoB test is performed with tones at $f_1 = 5$ MHz and $f_2 = 9.7$ MHz.	44
4.18	Measured in-band IIP3 and FoM of the fifth-order active-R filters (a) CoB, (b) FLFB, and (c) IFLFB. The measurement data of the low (\diamond, \circ) and high (\blacklozenge, \bullet) I_B is shown for each coarse tuning step.	45
5.1	Proposed surrogate model inserted on a modular optimization framework for automatic IC design.	55
5.2	Schematics for device characterization of (a) Oxide Capacitance (C'_{ox}). (b) Threshold Voltage (V_{TH}). (c) Normalization current (I_S). (d) Saturation voltage (V_{DSAT}). early voltage (V_A).	58

5.3	Sample of characterization data of a given CMOS technology in typical corner (TT) (a) NMOS: V_{TH} . (b) PMOS: V_{TH} . (c) NMOS: I_S . (d) PMOS: I_S . (e) NMOS: V_{DSAT} . (f) PMOS: $ V_{DSAT} $. (g) NMOS: V_A . (h) PMOS: $ V_A $	59
5.4	Comparison of the percentage error of prediction of CMOS parameters using models based on curve-fitting and Gaussian process regression (GPR) for sizes (a) KL = 1, KWL = 4. (b) KL = 10, KWL = 50. These models were built from the characterization data.....	62
5.5	Active-RC second order filter under design (a) circuit topology. (b) Transistor level schematic of the second order internally compensated amplifier.	65
5.6	Pareto front of the optimization of filter ($FC = 100KHz$) using SQP and NSGA-II optimization algorithms in (a) TSMC 180 nm CMOS process. (b) IBM 130 nm CMOS process. (c) TSMC 65 nm CMOS process.....	68
5.7	Values of the optimization variables from solutions in the Pareto front for the filter optimization ($FC = 100 KHz$) obtained with: (a) SQP-180 nm, (b) SQP-130 nm, (c) SQP-65 nm, (d) NSGA-II-180 nm, (e) NSGA-II-130 nm, and (f) NSGA-II-65 nm.	69
5.8	Circuit schematic of the capacitor-less low dropout (LDO) with type-A single stage error amplifier internal frequency compensation provided by C_C	71
5.9	Small-signal macromodel of the LDO for the calculation of the power supply rejection (PSR).....	72
5.10	Small-signal macromodel of the LDO for the calculation of the phase margin.	72
5.11	Pareto front of the optimization of the LDO circuit using SQP and NSGA-II optimization algorithms and (a) TSMC 180 nm. (b) IBM 130 nm. (c) TSMC 65 nm CMOS process.....	74
5.12	Values of the optimization variables of the solutions in the Pareto front obtained with: (a) SQP-180 nm, (b) SQP-130 nm, (c) SQP-65 nm, (d) NSGA-II-180 nm, (e) NSGA-II-130 nm and (f) NSGA-II-65 nm.	76
5.13	Circuit schematic of a 5-stage current starved voltage controlled oscillator (VCO). The control voltage is generated with a biasing current and a diode connected transistor.....	78
5.14	Pareto front of the optimization of the CSVCO circuit using SQP and NSGA-II optimization algorithms and (a) TSMC 180 nm process. (b) IBM 130 nm process. (c) TSMC 65 nm CMOS process.	79

5.15	Values of the optimization variables of the solutions in the Pareto front obtained with: (a) SQP-180 nm, (b) SQP-130 nm, (c) SQP-65 nm, (d) NSGA-II-180 nm, (e) NSGA-II-130 nm and (f) NSGA-II-65 nm.	81
5.16	History of the objective function through the iterative optimization process using (a) SQP algorithm. (b) NSGA-II algorithm.....	82
6.1	Locking the current biasing of analog circuits. (a) Simple current mirror. (b) Configurable current mirror [5]. (c) Parameter biasing obfuscation [6]. Only the correct key provides the correct output current.....	90
6.2	Voltage biasing locking. (a) Obfuscated voltage divider. (b) Memristor-based voltage divider. (c) Analog neural network-based voltage biasing.....	91
6.3	Locking of analog and mixed-signal circuits (AMS). (a) MixLock [7] uses logic locking on the digital component of the AMS circuit. (b) AMSlock [8] locks the optimization core setting the tuning knobs of the analog part. (c) Shared dependencies [9] locks both the digital and the analog components of the AMS circuit.	93
6.4	Analog layout camouflaging techniques protect against reverse engineering-based attacks. (a) In multiple threshold voltage (VT) camouflaging [10], transistors with different threshold voltages (VT) are used to hide the circuit design. (b) Sizing camouflaging [11] uses fake contacts to hide the circuit connectivity and effective device sizing.	95
7.1	The key provisioning unit generates the common key using the user key which is unique to that chip instance.	102
7.2	EPIC protocol for remote activation of the locked chip using public key cryptography [12]. Master key (MK), random chip key (RCK), common key (CK), user key (UK), public (Pub), private (Pri).	103
7.3	In the combinational lock [5] the key provisioning generates the common key from the unique user key with the aid of a PUF [13].	103
7.4	(a) Voltage response of a non-inverting ST (top) transient response: while the output values are either V_{OL} or V_{OH} , the input voltage takes any value in between. (bottom) Voltage transfer characteristic. (b) Internal feedback or 6-T ST. The thresholds are defined by the sizing of the transistors M_3 and M_6 and the control voltages V_{C1} and V_{C2} . (c) The external feedback ST uses an amplifier with resistors implementing positive feedback. The programmable resistors can be controlled by a digital word $\{b_0, b_1, b_2, \dots, b_{T-1}\}$ or a control voltage V_C	106

7.5	(a) Voltage response of a window comparator (top) transient response: if $V_{TL} < V_{IN} < V_{TL}$ the output is V_{OH} , otherwise V_{OL} . (bottom) voltage transfer characteristic. The low and high thresholds are defined by the configurable resistors R_A , R_B , and R_C . (b) Inverter-based window comparator. (c) OpAmp-based window comparator.	108
7.6	A negative hysteresis ST and its voltage transfer characteristics. This is built from a conventional positive (non-inverting) ST and a window comparator.	111
7.7	Proposed approach. The UK consists of x segments that reuse the same circuitry. Each segment selects between a positive ST or negative ST, configures the threshold voltages, and also provides the input voltages to generate the required CK. DAC_1 is needed for digital UKs, but not for analog UKs. DAC_2 is needed for analog CKs, but not for digital CKs.	112
7.8	Effective key size of the CK generated by the proposed key provisioning compared to a true number (TRN). The effective CK's size is calculated when the hysteresis window (HW) is static or dynamic and using a positive ST (PST) alone or combined with a negative ST (NST).	117
7.9	Uniqueness of the user key. The hamming distance of all user keys is calculated for different common keys CK_1 , CK_2 , CK_3 , and CK_4	118
7.10	Worst case variation in threshold voltage due to process and temperature variations. The standard deviation σ and mean variance μ are given for each temperature value.	121
7.11	Estimated variation on the common key's entropy due to a $\pm\Delta$ voltage variation in V_{TL} and V_{TH} . The results of six experiments with different values of the number of bits describing the hysteresis window w , the input values m , and the number of inputs n are reported.	122
7.12	Second order Gm-C bandpass filter.....	123
7.13	Transfer function of 4th order Gm-C filter for the correct and incorrect keys. The specifications of the filter are met only for the correct key.	123
7.14	Applying the correct key to the locked LDO, gives the desired performance $PM > 45^\circ$ and $PSR > 70$ dB, whereas an incorrect key gives undesired performance.	124
7.15	$S_{11} \leq -30$ dB and $S_{21} > 25$ dB when correct key is applied to the locked LNA. Otherwise, for an incorrect key, the S-parameters do not satisfy the specifications.	125

LIST OF TABLES

TABLE	Page	
3.1	Summary of three basic active-R topologies: lossless integrator, first-order filter or lossy integrator, and second-order filter or biquad. The single-ended circuit schematic, the signal flow graph, and the transfer functions of each topology are presented. The amplifier's frequency response is represented by the ideal integrator model $A_x(s) = \omega_{tx}/s$	19
3.2	Error in the first-order active-R filter due to the frequency response of the amplifier where $GBW = A_0 \cdot \omega_{p1}$	21
4.1	Input impedance (Z_{in}) and output impedance (Z_o) of the single-ended first-order and second-order active-R filters.	28
4.2	Transistor sizing of the two-stage Miller amplifier in Fig. 4.7.	35
4.3	Continuous ω_o tuning range at each coarse tuning step of the CoB filter.	42
4.4	Comparison of performance metrics of the three proposed active-R topologies proposed in this work with the state-of-the-art in active CT filters.	46
5.1	Optimal options of the function <i>fitrpg</i> for the training of the Gaussian Process (GP) model from the characterization of the process parameters in typical corner.	60
5.2	Parameters of the optimization algorithms used for this experiment sequential quadratic programming (SQP) and NSGA-II.	63
5.3	Performance metrics of the Pareto front solutions obtained through SQP optimization of the 2nd-order active-RC low-pass filter optimization in a TSMC 180 nm process (Noise (V_N) reported at 10 kHz).	70
5.4	Circuit parameters and specification constraints for the optimization of the LDO in two different CMOS processes.	73
5.5	Performance metrics of the Pareto front solutions obtained through SQP optimization of the CL-LDO in a 130 nm process measured with simulations.	75
5.6	Performance metrics of the Pareto front solutions obtained through NSGA-II optimization of the CL-LDO in the IBM 130 nm process measured with simulations.	77
5.7	Summary of the test cases under optimization using the proposed surrogate model.	82

5.8	Summary of success rate from surrogate evaluation to simulation verification across corners.	83
6.1	Sources of information available to the attacker.	88
6.2	Attack success of the attacks reported in Section 6.5 on the defense techniques in Section 6.3. It denotes the successful (✓), unsuccessful (✗), potentially successful but not demonstrated (✓*) attacks. \approx denotes that the attack reduces the search space.....	97
7.1	Existing digital and analog locking techniques and their user key (UK) and common key (CK) types. The proposed key provisioning technique receives the UK and generates CK. It is compatible with analog and digital keys.	105
7.2	The entropy of the CK for all combinations of w , n , and m . w bits set the width of the HW . n is the number of m -bit input values applied in series.	117
7.3	Area overhead of the ST-based key provisioning implementation for keys in the digital or the analog domain.	119
7.4	Area overhead comparison with other techniques.	120
7.5	The area overhead incurred by the proposed key provisioning unit on different locked analog circuits.	125

1. INTRODUCTION

Analog and digital circuits are fundamentally different. The main differences are the integration scale and the automatic vs. customized circuit synthesis. The lower transistor count of analog circuits does not translate in lower design complexity. As shown in Fig. 1.1, analog integrated circuit (IC) design consists of finding the circuit's topology, device sizing, and biasing conditions that produce a specific response and meets several performance specifications. However, the technology scaling and recent application trends have increase the challenges of analog IC design. These challenges include: (i) conflicting trade-offs between specifications, (ii) transistor imperfections that increase the modeling complexity, (iii) lower supply voltages, (iv) sensitivity of the performance to process, voltage and temperature variations, (v) flexible and programmable response, (vi) low-power and low-area consumption requirement, and (vii) IP protection against different kinds of security attacks.

The main objectives are therefore:

- To study circuits sensitive to PVT. Investigate the implementation and fabrication of tunable Active-R low pass filters, as a power efficient filter implementation for baseband filters.

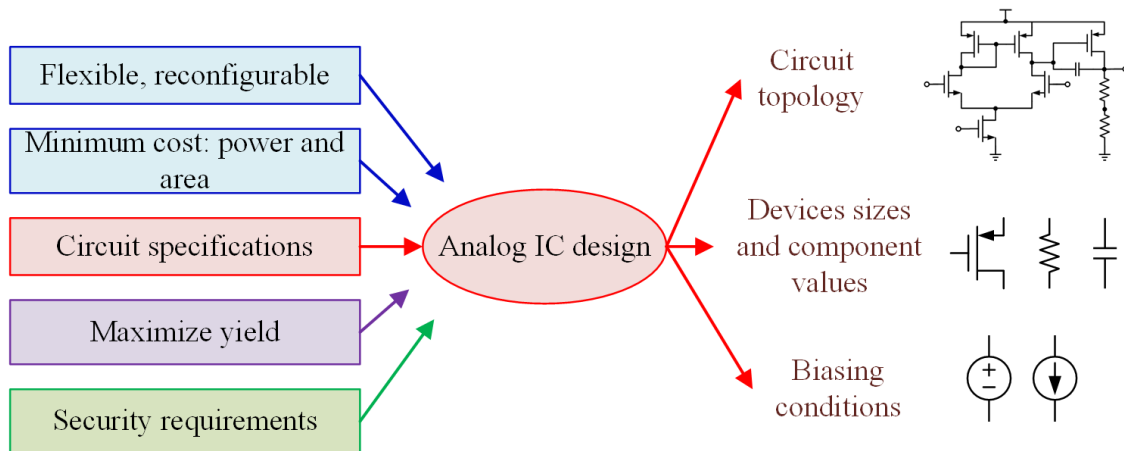


Figure 1.1: Problem statement: challenges on analog integrated circuit design.

- To research on optimization techniques applied to the automatic pre-silicon and post-silicon design of analog circuits. It includes to compare optimization algorithms, and their application to the design of analog circuits.
- To develop techniques for built-in-self testing (BIST) different performance metrics in frequency and time domain.
- To generate strategies for analog IP protection against overproduction. One option proposed is to apply the logic locking strategy to the digital core of the optimization platform. Another strategy consist on developing analog locks for IC security.

This dissertation proposes and discusses techniques to address these challenges. Figure 1.1 shows the scope of this research proposal.

The first part of this dissertation discusses the synthesis and circuit implementation of of highly tunable power-efficient high-order active-R low-pass filters for flexible radios used in base-band multi-standard wireless applications [14]. Active-R filters use the inherent integrator-like behavior of amplifiers to realize their frequency response. The main advantages of this type of filter are high-frequency performance and a low integrated area since the only capacitor they require is the Miller capacitor used in internally compensated amplifiers. In this work, amplifiers with configurable unity-gain frequencies enable the continuous tuning of active-R filters. Three different filter architectures realize a fifth-order Butterworth prototype tunable in the 1–50 MHz frequency range. These filters are designed, fabricated, and tested using the TSMC 0.18 μm process. The integrated area of each fifth-order topology is $\leq 0.33 \text{ mm}^2$, giving the smallest area per tuning range ratio. Also, the power consumption is in the range of 7.45 mW to 9.38 mW from a 1.8 V supply [14]. Compared with state-of-the-art active filters, the filters presented in this work have the largest tuning range without dynamic range degradation.

The second part of this work explored the application of optimization techniques to the automatic design of analog circuits. We describe pre-fabrication and post-fabrication optimization approaches capable of handling multiple conflicting performance metrics while reducing power and

area consumption. A computationally low-cost surrogate model for multi-objective optimization-based automated analog IC design was developed. This surrogate enables the performance evaluation of a given circuit. It is more accurate than the equation-based models and faster than circuit simulations. The surrogate has three main components: a set of Gaussian process regression models of the technology's parameters, a physics-based model of the MOSFET device, and a set of equations of the performance metrics of the circuit under design. This model was inserted into two different state-of-the-art optimization algorithms to prove it is compatible with gradient-based and population-based algorithms. The efficacy of our surrogate is demonstrated through simulation validation across process corners in three different CMOS technologies, using three representative circuit building-blocks that are commonly encountered in mainstream analog/RF ICs. The proposed surrogate is $69X$ to $470X$ faster at evaluation compared with circuit simulations [15].

The third part of this work describes the development of defense techniques for analog IP protection [16]. Specifically, we present a technique for analog performance locking enhanced with low-overhead key provisioning. Combinational locking has demonstrated being an effective technique to protect the performance of analog circuits using security keys. The locked circuit meets the specifications only under a specific configuration decided by the correct key, shared by all chip instances of the same design. However, increasing the security level is achieved by increasing the key size, which results in a larger area overhead. Moreover, the designer must ensure that each chip instance has its unique security key that cannot unlock other chips. The proposed Schmitt-trigger based key provisioning addresses these issues. The proposed key provisioning is compatible with existing analog locking techniques and has a constant area overhead regardless of key size [17]. This approach is tested with three analog/RF circuits to demonstrate its low overhead and effectiveness in security.

2. AMPLIFIER MACROMODELING

2.1 General gain, pole, zero amplifier model

Applications using amplifiers in feedback often assume an amplifier with a very high gain and infinite bandwidth [18]. Practical amplifiers have a large gain that remains constant over a bandwidth. Outside of this bandwidth, the gain decreases depending on the poles and zeros of the amplifier's transfer function (TF). The amplifier open-loop frequency response can be represented by its unity-gain frequency ω_t , DC gain A_{DC} , poles ω_{p_i} , and zeros ω_{z_j} :

- Ideal integrator

$$A(s) = \frac{\omega_t}{s} \quad (2.1)$$

- Single pole system

$$A(s) = \frac{A_{DC}}{(1 + s/\omega_{p1})} = \frac{\omega_t}{(s + \omega_{p1})} \quad (2.2)$$

- Two pole system

$$A(s) = \frac{A_{DC}}{(1 + s/\omega_{p1})(1 + s/\omega_{p2})} = \frac{\omega_t}{(s + \omega_{p1})(1 + s/\omega_{p2})} \quad (2.3)$$

- Two poles and one right half-plane (RHP) zero (Miller compensated amplifier)

$$A(s) = \frac{A_{DC}(s/\omega_{z1} - 1)}{(1 + s/\omega_{p1})(1 + s/\omega_{p2})} = \frac{\omega_t(s/\omega_{z1} - 1)}{(s + \omega_{p1})(1 + s/\omega_{p2})} \quad (2.4)$$

The amplifier is, most of the time, used in feedback configurations. For instance, the inverting amplifier configuration in Fig. 2.1 has a TF given by

$$T(s) = \frac{-a}{1 + \frac{1}{A(s)}(1 + a)} \quad (2.5)$$

The closed-loop inverting amplifier's TF is illustrated in Fig. 2.2 for the four different models of

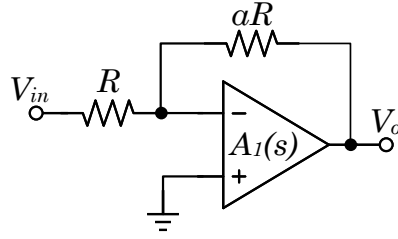


Figure 2.1: Op-Amp based inverting amplifier.

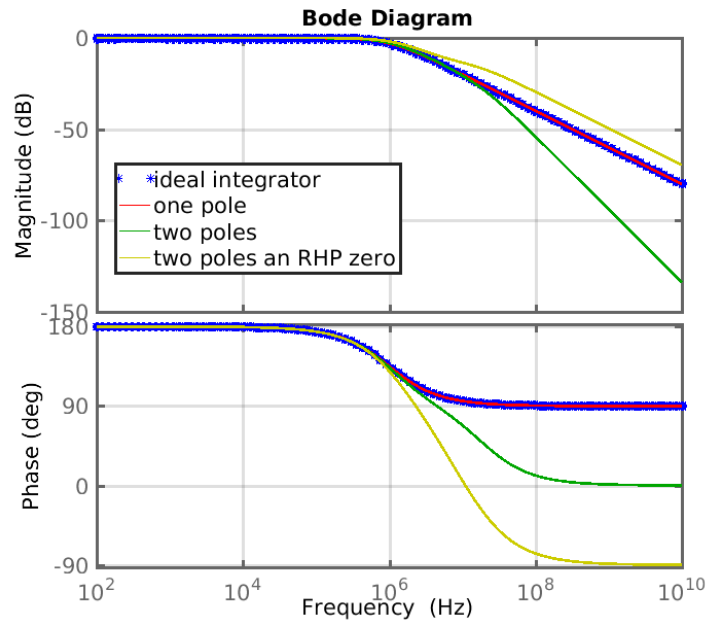


Figure 2.2: Frequency response of the inverting amplifier using different models of the amplifier $A_1(s)$ open-loop response.

the amplifier's open-loop gain.

The error between the ideal and the actual amplifier TFs is calculated for different amplifier models. Fig. 2.3 shows the effect of the finite DC gain, while Fig. 2.4 shows the error due to the non-dominant pole location. The ideal frequency response of the inverting amplifier is calculated with the ideal integrator macromodel. The actual frequency response of the filter $H(s)$ is calculated with different amplifier's models.

The discrepancy function D , also called correction factor, is the ratio between the actual TF and the ideal one ($D(s) = H(s)/H_{ideal}(s)$) [18, 19]. When there is no error i.e. $H(s) = H_{ideal}(s)$, the

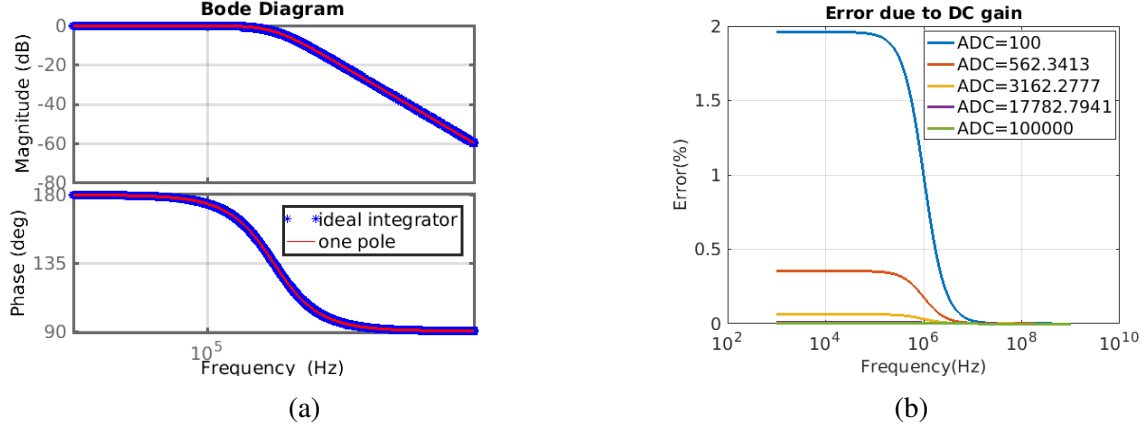


Figure 2.3: Effect of a finite open-loop amplifier's DC gain on the inverting amplifier response. (a) Transfer function comparison (b) Percentage error across frequencies for different DC gains.

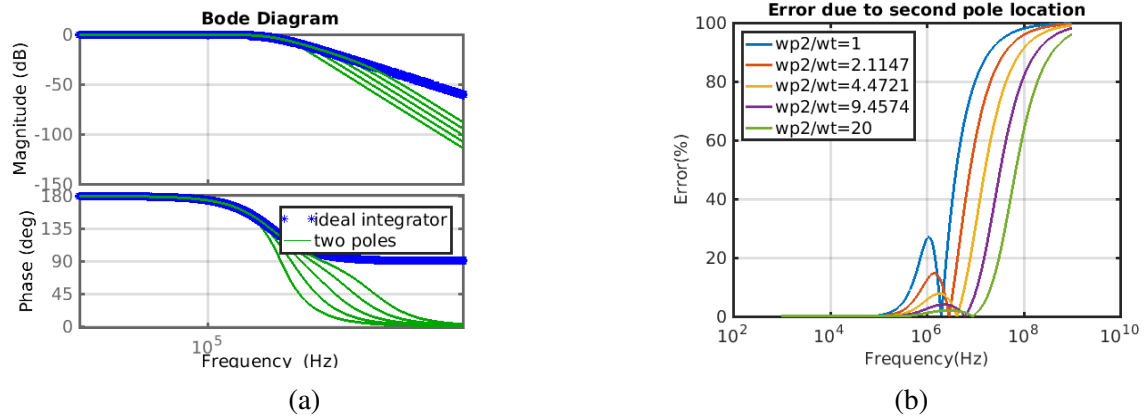


Figure 2.4: Effect of the amplifier's non-dominant pole location on the inverting amplifier response. (a) Transfer function comparison (b) Percentage error across frequencies for different ratios ω_{p2}/ω_t .

discrepancy's value is $D(j\omega) = 1\angle 0^\circ$. Moreover, the error function in terms of D is approximated by

$$\epsilon(s) = \frac{H(s) - H_{ideal}(s)}{H_{ideal}(s)} = \frac{H(s)}{H_{ideal}(s)} - 1 = D - 1. \quad (2.6)$$

2.2 Transistor level OTA implementation and its modeling

Consider the transistor level implementation of an operational transconductance amplifier (OTA) shown in Fig. 2.5(a). It is a two-stages OTA with internal frequency compensation provided by the

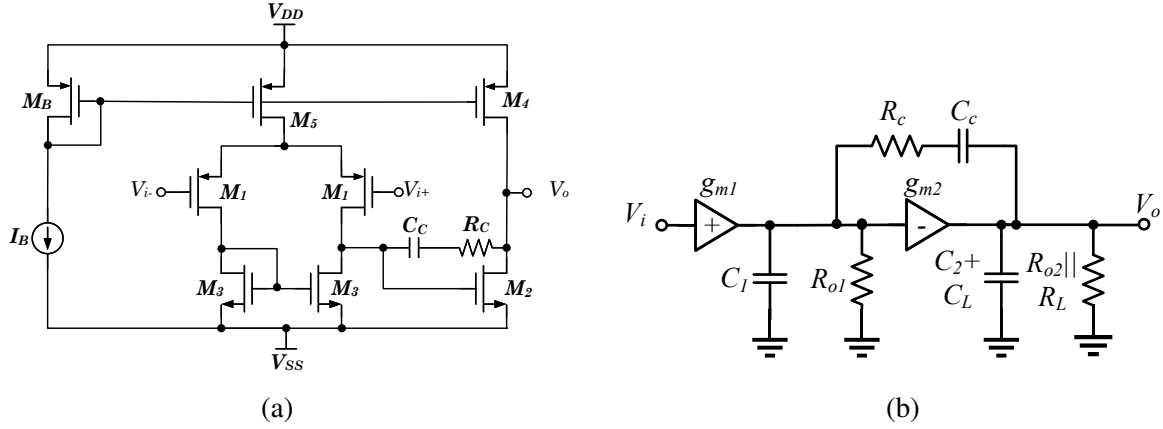


Figure 2.5: Two-stage Miller compensated OTA. (a) Transistor level schematic. (b) Small-signal model.

Miller capacitor C_C [20]. Moreover, R_C can cancel the RHP zero due to C_C or transform it into a left half-plane (LHP) zero, improving the loop stability.

2.2.1 Small-signal macromodel

Fig. 2.5(b) shows the small-signal macromodel of the two-stage Miller compensated shown in Fig. 2.5(a). Inserting the small-signal macromodel into the inverting amplifier of Fig. 2.1, leads to the circuit in Fig. 2.6. Its TF can be obtained in terms of the circuit parameters by solving the system extracted using modified nodal analysis (MNA).

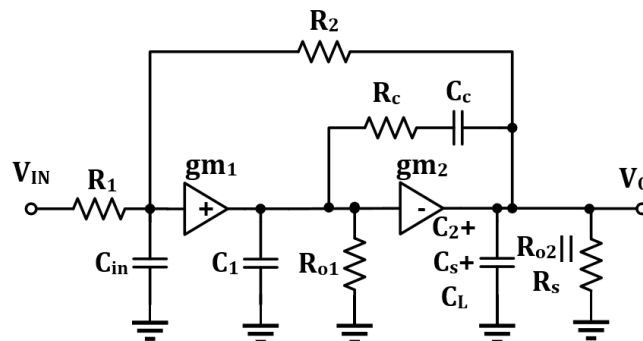


Figure 2.6: Small-signal macromodel of the closed-loop first order Active-R

$$\begin{bmatrix}
1/R_1 & -1/R_1 & 0 & 0 & 1 \\
-1/R_1 & 1/R_1 + 1/R_2 & 0 & -1/R_2 & 0 \\
s \cdot C_{in} & & & & \\
0 & -g_{m1} & 1/R_{o1} + s \cdot C_1 + 1/Z_c & -1/Z_c & 0 \\
0 & -1/R_2 & -1/Z_c + g_{m2} & s \cdot C_L + s \cdot C_2 + s \cdot C_s + \\
& & & 1/R_{o2} + 1/R_2 + 1/R_s + 1/Z_c & 0 \\
1 & 0 & 0 & 0 & 0
\end{bmatrix} \cdot \begin{bmatrix} V_{in} \\ V_1 \\ V_2 \\ V_o \\ I_{vin} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

The unity-gain frequency of this OTA is represented by the expression

$$\omega_t = \frac{g_{m1} \cdot R_{o1} \cdot g_{m2} \cdot R_{o2}}{R_{o1}(C_1 + C_C(1 + g_{m2} \cdot R_{o2})) + R_{o2}(C_{Ltot} + C_C)} \quad (2.7)$$

where g_{m1} and g_{m2} are the transconductances, while R_{o1} and R_{o2} are the output impedances of the first and second stages, respectively. The output impedances are defined as $R_{o1} = 1/(g_{ds1} + g_{ds3})$ and $R_{o2} = 1/(g_{ds2} + g_{ds4}) || R_s || R_L$. C_C is the Miller compensation capacitor, and C_{Ltot} is the total load capacitance modeling the equivalent parasitics of the pad, bondwire, and packaging. C_1 represents the total capacitance at the output node of the first stage. R_L represents the equivalent load resistance. Moreover, the sensitivity of closed loop gain w.r.t circuit parameters is shown in Fig. 2.7.

Stability is an important concern in all feedback systems. One conventional approach to analyze stability is to check the loop gain. Using a macromodel of the amplifier, the first order active-R filter is represented by the schematic in Fig. 2.6.

The closed loop gain of this configuration (Fig. 2.6) is V_O/V_{IN} . Using the same amplifier macromodel, the loop gain V_R/V_T can be represented by Fig. 2.8.

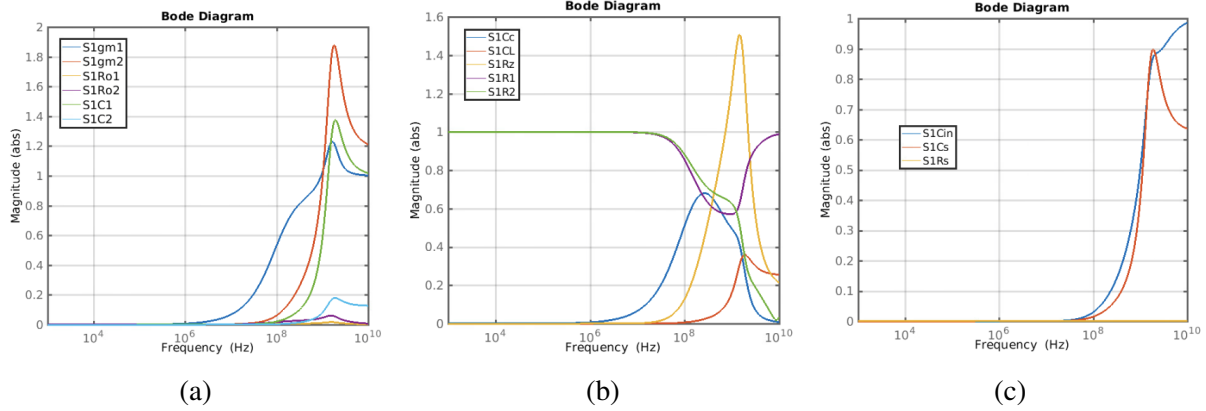


Figure 2.7: Sensitivity of the closed-loop response of the inverting amplifier with respect to the small-signal circuit parameters.

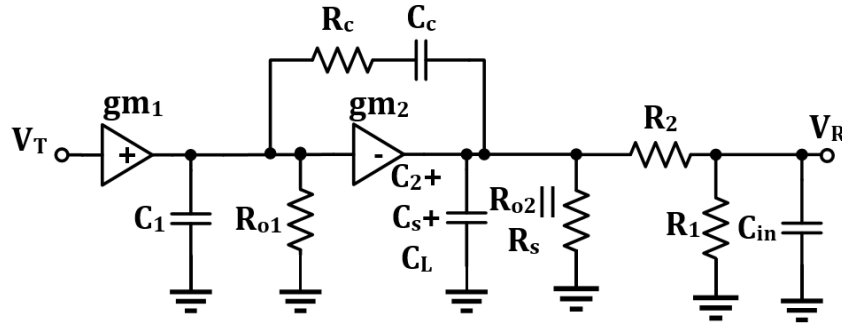


Figure 2.8: Loop gain of closed loop first order Active-R

$$\begin{bmatrix}
 1/R_{o1} + 1/R_{z+} & -1/R_z & 0 & 0 \\
 s \cdot C_1 & & & \\
 -1/R_z & 1/R_z + s \cdot C_c & -s \cdot C_c & 0 \\
 gm_2 & -s \cdot C_c & 1/R_{o2} + 1/R_2 + 1/R_s + & -1/R_2 \\
 & & s \cdot C_L + s \cdot C_c + s \cdot C_2 + s \cdot C_s & -1/R_2 \\
 0 & 0 & -1/R_2 & 1/R_1 + 1/R_2 + \\
 & & & s \cdot C_{in}
 \end{bmatrix} \cdot \begin{bmatrix} V_2 \\ V_z \\ V_o \\ V_R \end{bmatrix} = \begin{bmatrix} gm_1 \cdot V_T \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

The dominant pole is approximately given by

$$\omega_{p1} \cong \frac{-1}{g_{m2}R_{o1}R_{o2}C_C} \quad (2.8)$$

whereas, the non-dominant pole can be written as

$$\omega_{p2} \cong \frac{-g_{m2}}{C_1} \frac{1}{1 + \frac{C_{Ltot}}{C_C} + \frac{C_{Ltot}}{C_1}} \quad (2.9)$$

The third pole, $\omega_{p3} = 1/(R_C \cdot C_1)$ also contributes with attenuation and phase margin [21], and the zero is placed at:

$$\omega_z = \frac{1}{C_C(1/g_{m2} - R_C)} \quad (2.10)$$

2.2.2 Effect of large-signal non-idealities of the OTA

The maximum rate of change of the output of the first order active-R is limited by the amplifier bandwidth or its slew rate. Equation (2.11) represents this trade-off [18]. If this inequality is satisfied, the output has an exponential response since it is limited by the bandwidth. In other case, the output is slew rate limited and the output displays a linear behavior.

$$\left| \frac{R_2}{R_2 + R_1} V_m \cdot GBW \right| < SR \quad (2.11)$$

where V_m is peak amplitude of an input sine waveform and for a first order amplifier approximation where the unity gain frequency is equal to the gain bandwidth product $GBW = \omega_t$. This trade-off leads to an important design constraint since:

- An active-R filter cutoff frequency depends on the amplifier's bandwidth
- In a CMOS amplifier the unity gain frequency and the slew rate have common parameters, so their design is related.

Slew rate limitation is not desired since it causes signal distortion. In consequence, the amplifier should be designed such that the output is bandwidth limited instead of slew rate limited [18].

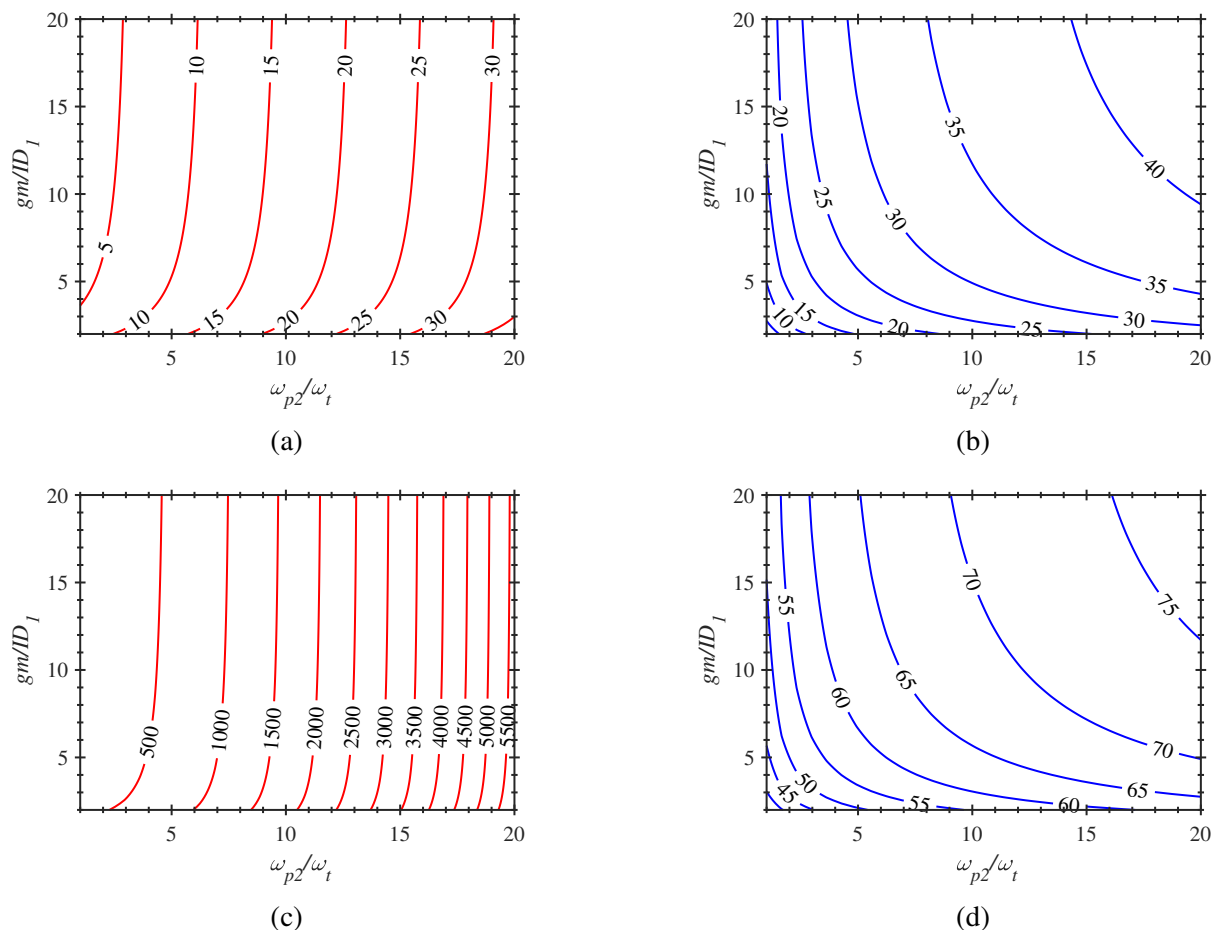


Figure 2.9: Amplifier estimated power consumption vs second pole location f_{p2}/f_{u1} . (a) Total current in μm at $f_C = 1MHz$. (b) Gain in dB at $f_C = 1MHz$. (c) Power consumption in μm at $f_C = 50MHz$. (d) Gain in dB at $f_C = 50MHz$.

2.2.3 Frequency vs. power trade-off

The design of the amplifier's second-stage includes several considerations. First, lowering the output impedance moves the non-dominant output pole to higher frequencies ensuring stability and reducing the error on the filter's transfer function. However, it also reduces the gain. The location of the non-dominant output pole of the amplifier creates a trade-off between gain, stability, and power consumption. This trade-off is illustrated in Fig. 2.9. As expected, lowering the impedance of internal nodes of the amplifiers to move the poles to higher frequencies increases the power consumption.

3. CONTINUOUS-TIME ACTIVE FILTERS¹

3.1 Introduction

Active elements such as amplifiers are used in filter implementation to solve several limitations of passive RLC filters. Active elements allowed to realize arbitrary sets of real-axis poles and zeros, minimize the loading effect, tune independently the filter parameters, and remove the need of inductors enabling on-chip integration [19]. We will limit our discussion to active low-pass filters (LPF) only, given that is the focus of this work. Other filter prototypes like band-pass, high-pass, and band-stop are not described on this dissertation but can be seen in [18, 19].

First- and second-order LPFs are built from integrators characterized by the transfer function $H(s) = \omega_u/s$, where ω_u is the integrator's unity-gain frequency. The general form of a first-order LPF is written in terms of its parameters: the DC gain H_0 and the cutoff (or -3dB) frequency ω_0 , as [18]

$$H(s) = \frac{H_0}{1 + s/\omega_0} \quad (3.1)$$

Meanwhile, the general form of of a second-order LPF is expressed as

$$H(s) = \frac{\pm H_0 \cdot \omega_0^2}{s^2 + (\omega_0/Q)s + \omega_0^2} \quad (3.2)$$

where Q denotes the filter's quality factor [18, 19].

Moreover, active filters are classified according to their components. Fig. 3.1 shows three different integrator implementations: active-RC, G_m -C, and G_m -OTA-C [22].

In this chapter we review traditional continuous-time active LPF topologies and summarize their conflicting performance metrics including noise, linearity, dynamic range, power consump-

¹Part of this chapter is reprinted with permission from "Efficient use of gain-bandwidth product in active filters: Gm-C and Active-R alternatives" by A. Sanabria-Borbón and E. Sánchez-Sinencio, in IEEE Proceedings of the 8th Latin American Symposium on Circuits & Systems (LASCAS), Feb. 2017; and permission from "Synthesis of High-Order Continuously Tunable Low-Pass Active-R Filters" by A. Sanabria-Borbón and E. Sánchez-Sinencio, in IEEE Transactions on Circuits and Systems I: Regular Papers, Jan. 2021.

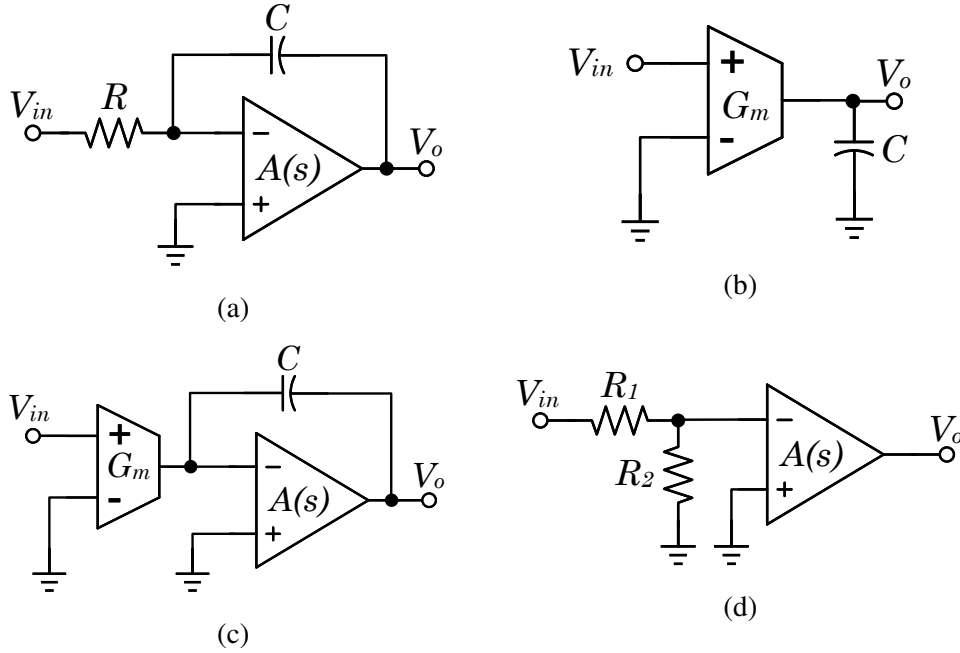


Figure 3.1: Active integrators (a) Active-RC (b) Gm-C (c) Gm-OTA-C (d) Active-R.

tion, and area.

Filter topologies combining the advantages of active-RC and G_m -C have been reported in [22, 23], and [24]. G_m -OTA-C filters operate at higher frequencies than active-RC filters, but they have the dynamic range limitations of G_m -C filters, resulting in lower power efficiency [22]. The G_m -assisted OTA-RC technique proposed in [22] combines the two filter techniques to achieve both high frequency and high linearity. The active- G_m -RC filter topology in [23] and [24] uses the amplifier's roll-off and a capacitor to implement a second-order TF. This approach offers power and area savings by using a single amplifier for a biquad implementation but increases the design complexity. In all these topologies, the use of capacitors limits the filter tunability and results in a large integrated area [19]. Thus, there is still a need for filter implementations with: (i) high power and area efficiency, (ii) a continuous wide frequency tuning range, and (iii) a sufficient dynamic range performance that does not degrade with frequency.

3.2 Active-RC filters

Active-RC filters are built with operational amplifiers, that have ideally infinite gain-bandwidth products (GBWs), connected in negative feedback. Hence, their passive elements (Rs and Cs) determine the filter transfer function (TF) [18]. Given a large gain-bandwidth amplifier, the resistors and capacitors connected on the amplifier terminals define the gain, poles, and zeros of the filter. Moreover, the negative feedback, which yields high linearity active-RC filters also limits their frequency performance [25]. Fig. 3.2 shows the active-RC lossless and lossy active-RC integrators [19].

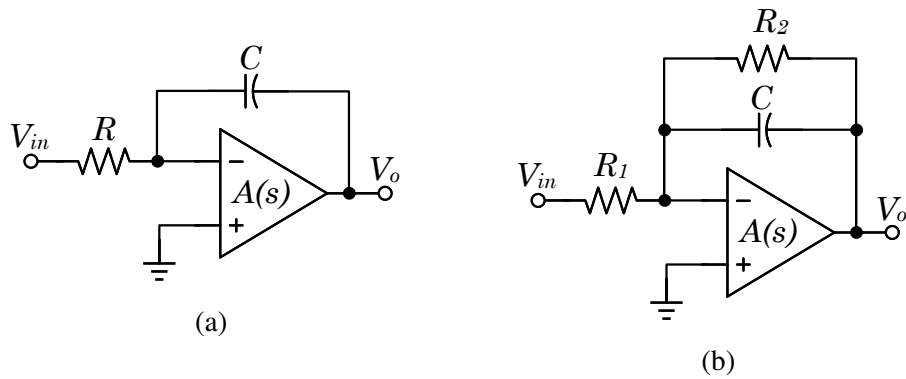


Figure 3.2: Active-RC (a) lossless and (b) lossy integrators.

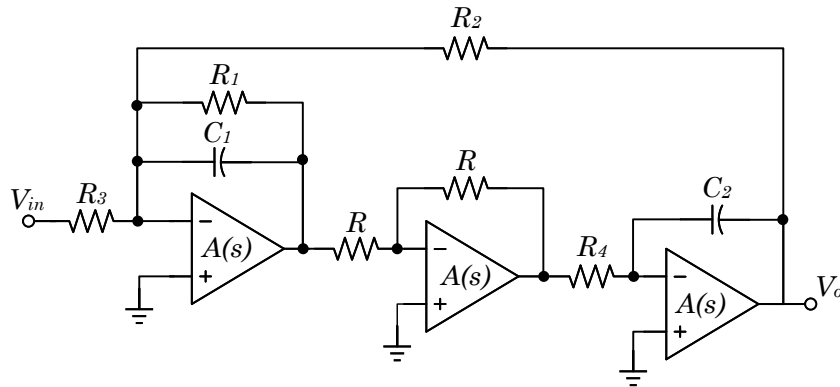


Figure 3.3: Second-order Tow-Thomas active-RC filter.

A second-order Active-RC LPF structure is shown in Fig. 3.3. Its TF is given by

$$H_{LP} = \frac{V_o}{V_{in}} = \frac{1/(R_3R_4C_1C_2)}{s^2 + s/(R_1C_1) + 1/(R_2R_4C_1C_2)} \quad (3.3)$$

and the filter characteristics are defined as

$$\omega_o = \sqrt{\frac{1}{R_2R_4C_1C_2}} \quad Q = \frac{R_1}{\sqrt{R_2R_4}} \sqrt{\frac{C_1}{C_2}} \quad H = \frac{R_2}{R_3} \quad (3.4)$$

Implementing an active-RC filter requires $\omega_t/\omega_o \geq 10$, where ω_t is the amplifier unity-gain frequency [26]. Moreover, [24, 26], and [27] show that implementing highly tunable active-RC filters requires large banks of passive components, yielding a large area overhead.

3.3 G_m -C filters

Also called OTA-C filters [28]. The frequency response of the G_m -C filters is controlled by their building elements transconductors (G_m), which act as voltage to current converters, and grounded capacitors. G_m -C filters can achieve higher frequencies than their active-RC counterparts at the expense of lower linearity due to their open-loop operations [29, 30]. The dependency of the filter TF on G_m enables continuous tuning, but it also calls for advanced linearization techniques [24, 31]. Moreover, G_m -C filters are sensitive to parasitic capacitances [22].

Ideally the transconductor has infinite input and output impedance. Tuning is achieved by adjusting the biasing current of the transconductor or capacitor banks. A single MOS transistor can be used as transconductor, however it has small linearity range since its output current does not follow a linear relationship with its input voltage. The differential pair doubles the linearity range compared with a single-transistor common-source amplifier [28]. Hence, $G_m - C$ filters are preferred for applications with small signal amplitudes like biomedical. Applications with large signal amplitudes require additional linearization techniques. Several OTA topologies with linearization techniques are reported in [32]. Increasing the OTA linearity is achieved by reducing the transconductor gain and eliminating the input offset voltage. Moreover, the use of fully differ-

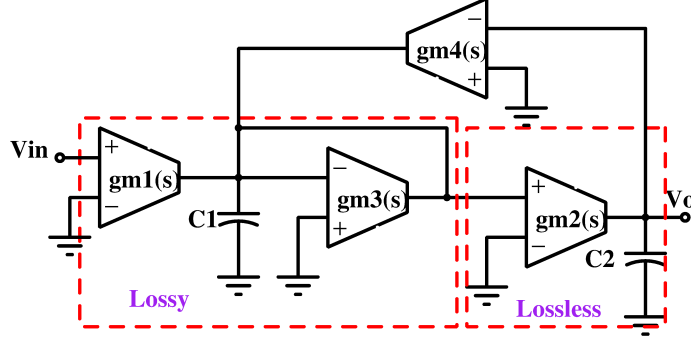


Figure 3.4: Gm-C biquad implementation.

entail or balanced filter architectures increases further the linearity by suppressing the even-order harmonics. It comes at the expense of a common-mode feedback loop.

The equivalent Gm-C implementation of this biquad is presented in Fig. 3.4. This topology is presented in the single ended version and using only one of the input terminals, in this way the implementation of the fully differential version is straightforward.

The transconductance gm_1 converts the input voltage into current to be integrated in the capacitor C_1 , gm_3 in unity feedback mimics a resistor that connected to gm_1 and C_1 make a lossy integrator. The lossless integrator is composed by gm_2 and C_2 , and finally gm_4 closes the feedback loop. The low-pass transfer function of this configuration is given by the Equation 3.5. Again, if the output is taken after one integrator only, a bandpass response is obtained.

$$H_{Gm-C}(s) = \frac{gm_1 gm_2}{C_1 C_2} \frac{1}{s^2 + s \frac{gm_3}{C_1} + \frac{gm_4 gm_2}{C_1 C_2}} \quad (3.5)$$

From the transfer function we can get the DC gain $H_0 = gm_1/gm_4$ the cut-off frequency and the quality factor as:

$$\omega_o = \sqrt{\frac{gm_4 gm_2}{C_1 C_2}} \quad Q = \frac{1}{gm_3} \sqrt{\frac{gm_4 gm_2 C_1}{C_2}} \quad (3.6)$$

3.3.1 Effect of real operational transconductance amplifiers in Gm -C filters

The transfer function of the Gm -C biquad topology (Eq. 3.5) assumes an ideal transconductance that has infinite bandwidth. However, a more realistic approximation [33] considers one dominant pole $\omega_p \ll \omega_o$, then the transfer function is given as $H(s) = N(s)/D(s)$ where $N(s)$ and $D(s)$ are presented in Equations 3.8 and 3.9.

$$gm = gm_o e^{-s/\omega_p} \cong gm_o (1 - s/\omega_p) \quad (3.7)$$

$$N(s) = gm_{o1} gm_{o2} \left(1 - \frac{s}{\omega_{p1}}\right) \left(1 - \frac{s}{\omega_{p2}}\right) \quad (3.8)$$

$$D(s) = C_2 s (C_1 s + gm_{o3} \left(1 - \frac{s}{\omega_{p3}}\right) + gm_{o2} gm_{o4} \left(1 - \frac{s}{\omega_{p2}}\right) \left(1 - \frac{s}{\omega_{p4}}\right)) \quad (3.9)$$

Assuming all poles are equal, i.e. ω_p , the ω_o and the bandwidth BW become:

$$\omega_{oa} = \sqrt{\frac{gm_{o2} gm_{o4} \omega_p^3}{(gm_{o2} gm_{o4} \omega_p + C_2 \omega_p^2 (-gm_{o3} + C_1 \omega_p))}} \quad (3.10)$$

$$BW_a = \frac{(-2gm_{o2} gm_{o4} \omega_p^2 + C_2 gm_{o3} \omega_p^3)}{(gm_{o2} gm_{o4} \omega_p - C_2 gm_{o3} \omega_p^2 + C_1 C_2 \omega_p^3)} \quad (3.11)$$

$$Q_a = \frac{Q}{1 - \frac{2\omega_{oa}}{\omega_{p1}} Q} \quad (3.12)$$

Equation 3.12 shows that the actual Q is enhanced as ω_p approaches ω_{oa} .

3.3.2 Noise in Gm -C filters

The Eq. 3.13 shows the expression for the input referred noise for the first order gm -C filter.

$$Vn_{in}^2 = \frac{4KT}{gm_1} \left(1 + \frac{gm_3}{gm_1}\right) \quad (3.13)$$

Additionally, the input referred noise of the biquad gm-C is presented in the Equation 3.14.

$$Vn_{in}^2 = \frac{4KT}{gm_1} \left[1 + \frac{gm_3}{gm_1} + \frac{gm_4}{gm_1} + gm_2 \left(\frac{1}{\omega C_2} \right)^2 \frac{gm_4^2}{gm_1} \right] \quad (3.14)$$

3.4 Active-R filters

The TF of an active-R filter depends on the amplifier's frequency response and resistor ratios [34–38]. Table 3.1 summarizes the circuit schematic, signal flow graph, TF, and filter parameters of three basic active-R filters [35, 39, 40]. This summary includes the lossless integrator, the lossy integrator or first-order filter, and the biquad or second-order filter [40]. The order of these active-R filters is equal to the number of amplifiers. Moreover, the filters' TF are evaluated with the amplifier's integrator model $A_x(s) = \omega_{tx}/s$.

This filter topology only contains amplifiers and resistors, avoiding the use of capacitors external to the amplifier. Hence, the filter characteristics ω_o , the quality factor Q , and the pass-band gain H_0 are set by ω_t and resistor ratios [34, 39–41]. Moreover, active-R filters are easily tunable and area inexpensive [42]. Compared with active-RC filters, active-R filters require a smaller ω_t/ω_o ratio. Therefore, active-R filters can achieve higher frequencies with the same power consumption or reduce the power required to implement the same ω_o . Compared with G_m -C filters, active-R filters can achieve a higher linearity for the same power and area consumption due to their closed-loop operation [43].

Although the idea of using the amplifier TF for filter realization was first proposed 40 years ago [39, 42], the fixed ω_t of the available commercial amplifiers limited their design and practical implementation [42]. However, with the advent of ASIC design and CMOS technologies, we can leverage the advantages of active-R topologies on implementing highly tunable baseband filters for multi-standard receivers.

The biquad active-R configuration resembles the active-RC Tow-Thomas biquad filter since: (i) it is built from a lossy and a lossless integrator, and (ii) it implements multiple TFs [18]. Assigning

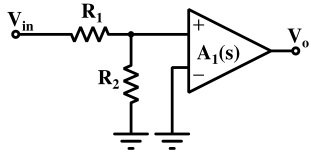
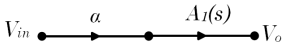
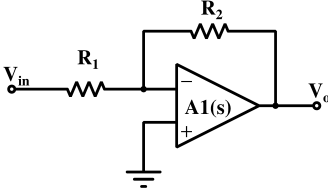
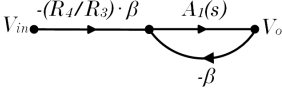
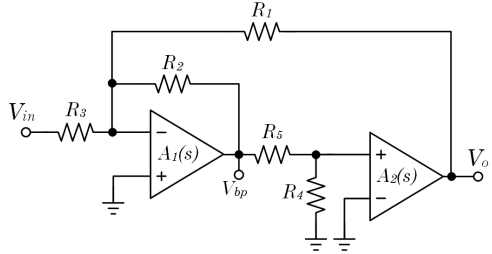
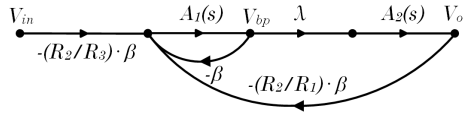
Lossless Integrator	Lossy Integrator	Second order Active-R
 	 	 
$H_L(s) = A_1(s) \cdot \alpha$ $H_L(s) = \frac{\omega_{t1} \cdot \alpha}{s}$ $\alpha = R_2 / (R_1 + R_2)$	$H_F(s) = \frac{-R_4 / R_3}{1 + 1 / (A_1(s) \cdot \beta)}$ $H_F(s) = \frac{-R_4 / R_3}{1 + s / (\omega_{t1} \cdot \beta)}$ $\omega_o = \omega_{t1} \cdot \beta$ $\beta = R_3 / (R_3 + R_4)$	$H_S(s) = \frac{-(R_2 / R_3) \cdot \lambda \cdot \beta \cdot \omega_{t1} \cdot \omega_{t2}}{s^2 + s \cdot \beta \omega_{t1} + \lambda \cdot \beta \cdot \omega_{t1} \cdot \omega_{t2} \cdot R_2 / R_1}$ $\lambda = R_4 / (R_4 + R_5), \quad \beta = (R_1 R_3) / (R_2 + R_1 R_3)$ $\omega_o = \sqrt{\frac{\lambda \cdot \beta \cdot \omega_{t1} \cdot \omega_{t2} \cdot R_2}{R_1}} \quad Q = \sqrt{\frac{\lambda \cdot \omega_{t2} \cdot R_2}{\beta \cdot \omega_{t1} \cdot R_1}}$ $H_0 = \frac{-R_1}{R_3}$

Table 3.1: Summary of three basic active-R topologies: lossless integrator, first-order filter or lossy integrator, and second-order filter or biquad. The single-ended circuit schematic, the signal flow graph, and the transfer functions of each topology are presented. The amplifier's frequency response is represented by the ideal integrator model $A_x(s) = \omega_{tx}/s$.

the same value to all resistors of the biquad simplifies its characteristics to $\omega_o = \sqrt{(\omega_{t1} \cdot \omega_{t2})/6}$, $Q = \sqrt{(3 \cdot \omega_{t2}) / (2 \cdot \omega_{t1})}$, and $H_0 = 1 \text{ V/V}$. Moreover, making $\omega_{t1} = \omega_{t2} = \omega_t$ yields $\omega_t / \omega_o < 3$. This example illustrates the main advantage of active-R filters, considering that an equivalent active-RC filter would require a much larger ω_t to ω_o ratio [18]. In this work, we leverage $\omega_{t1} \neq \omega_{t2}$ and a variable resistor R_2 to enable the biquad Q tuning. Other second-order active-R implementations are reported in [34, 39], and [35]. Moreover, a high-pass (HP) and a band-stop (BS) active-R filters are implemented by a different selection of the input and output nodes or combining the LP and BP responses with a summing block [38, 44].

3.4.1 Effect of the amplifier's small-signal non-idealities in active-R filter implementations

Modeling the amplifier as an integrator $A(s) = \omega_t/s$ is sufficient to estimate the TF of active-R filters. However, it does not account for the limitations of integrated CMOS amplifiers [42]. Such limitations include finite DC gain (A_0), a limited BW set by the dominant pole (ω_{p1}), non-dominant poles ($\omega_{p,n}$ for $n > 1$) and zeros ($\omega_{z,k}$ for $k \geq 1$), finite input impedance (Z_{in}), and non-zero output impedance (Z_o) [38]. The deviation between the real amplifier TF and its simplified integrator model results in an error on the filter characteristics [18]. Hence, the main challenge of an integrated monolithic implementation of active-R structures is to understand and minimize the error caused by the amplifier's limitations.

The error between the ideal and the actual filter's TFs is calculated for different amplifier models as shown in Table 3.2. The ideal frequency response of the first-order active-R filter was introduced in Table 3.1. The actual frequency response of the filter $H(s)$ is calculated with different amplifier's models. The discrepancy function D , also called correction factor, is the ratio between the actual TF and the ideal one ($D(s) = H(s)/H_{ideal}(s)$) [18, 19]. When there is no error i.e. $H(s) = H_{ideal}(s)$, the discrepancy's value is $D(j\omega) = 1\angle 0^\circ$. Moreover, the error function in terms of D is approximated by

$$\epsilon(s) = \frac{H(s) - H_{ideal}(s)}{H_{ideal}(s)} = \frac{H(s)}{H_{ideal}(s)} - 1 = D - 1. \quad (3.15)$$

The error function of the second-order filter is calculated using the same procedure. Fig. 3.5 shows the magnitude error ($\epsilon_M = |D(j\omega)| - 1$) and the phase error ($\epsilon_P = \angle D(j\omega)$) in the filter's TF for different amplifier's models. The error function of the second-order active-R filter is calculated using the same procedure. Fig. 3.5 shows the magnitude error and the phase error in the filter's TF for different amplifier's models. Considering a first dominant pole ω_{p1} , the amplifier frequency response becomes $A_i(s) = \omega_t/(s + \omega_{p1})$ [39]. Fig. 3.5(a) and 3.5(b) shows the error components, when the amplifier is represented by the single-pole model. The error is calculated for an amplifier

Amplifier model	$A(s)$	Error function $\epsilon(s)$
A constant A_0 with a single pole.	$\frac{GBW}{(s+\omega_{p1})}$	$\frac{\omega_{p1}}{s+GBW\cdot\beta+\omega_{p1}}$
A constant A_0 with a dominant and a non-dominant poles	$\frac{GBW}{(s+\omega_{p1})(1+s/\omega_{p2})}$	$\frac{s^2+s\cdot\omega_{p1}+\omega_{p1}\omega_{p2}}{s^2+s(\omega_{p1}+\omega_{p2})+(GBW\beta+\omega_{p1})\omega_{p2}}$
A constant A_0 with a dominant pole, a non-dominant pole, and a left-half-plane (LHP) zero.	$\frac{GBW(1+s/\omega_{z1})}{(s+\omega_{p1})(1+s/\omega_{p2})}$	$\frac{s^2(\omega_z-\omega_{p2})+s\cdot\omega_z\omega_{p1}+\omega_z\omega_{p1}\omega_{p2}}{s^2\omega_z+s((\omega_{p1}+\omega_{p2})\omega_z+GBW\omega_{p2}\beta)+(GBW\beta+\omega_{p1})\omega_{p2}\omega_z}$

Table 3.2: Error in the first-order active-R filter due to the frequency response of the amplifier where $GBW = A_0 \cdot \omega_{p1}$.

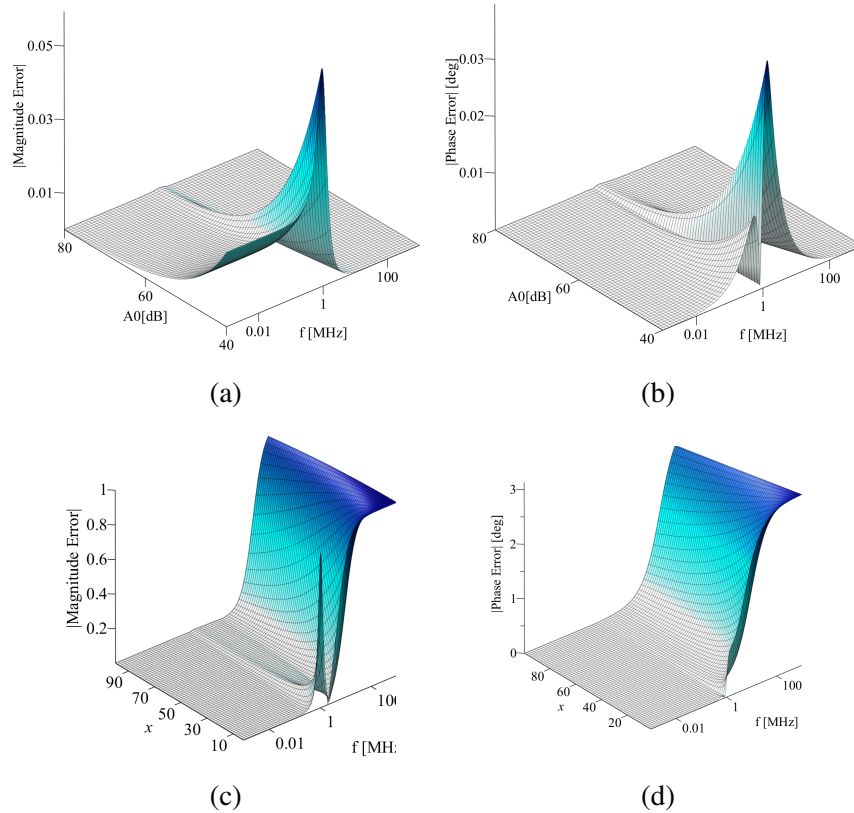


Figure 3.5: Magnitude and phase error of the active-R biquad's transfer function estimated for different approximations of the amplifier's transfer function. The single-pole model in which the pole location is ω_{p1} , the DC gain A_0 , and $\omega_t = \omega_{p1} \cdot A_0$ leads to the error in (a) Magnitude (b) Phase. The two pole transfer function where $x = \omega_{p2}/\omega_t$ yields an error in (c) Magnitude (d) Phase.

A_0 in the range from 40 dB to 80 dB, while keeping ω_t constant.

Using the single-pole amplifier's model $GBW/(s + \omega_{p1})$, where $GBW = A_0 \cdot \omega_{p1}$, yields the actual cutoff frequency ($\omega_{o,a}$) and the actual quality factor (Q_a) represented by Equations (3.16) and (3.17), respectively. Larger A_0 leads to a smaller error in the pass-band. Hence, the designer can determine the minimum amplifier's gain required to keep the error below a certain value. In another approach, [45] proposes the use of low power and low gain amplifiers for filter implementations. The filter TF error due to the low amplifier gain is compensated by the subsequent programmable gain amplifier (PGA) stage. Although this approach is valid for active-R filters, the amplifier's gain also needs to satisfy the linearity specifications.

$$\omega_{o,a} = \sqrt{\omega_{p1}^2 + \frac{(\omega_{t1}(\omega_{t2} + 2 \cdot \omega_{p1}))}{6}} \quad (3.16)$$

$$Q_a = \frac{\sqrt{9\omega_{p1}^2 + 1.5\omega_{t1} \cdot (\omega_{t2} + 2 \cdot \omega_{p1})}}{(\omega_{t1} + 6 \cdot \omega_{p1})} \quad (3.17)$$

The non-dominant pole ω_{p2} affects the active-R biquad TF at two frequency regions: around ω_o , and after ω_{p2} . At $\omega = \omega_o$, ω_{p2} causes additional peaking, also known as Q enhancement [18,43,46]. The two-pole amplifier model ($GBW/((s + \omega_{p1})(1 + s/\omega_{p2}))$) leads to $\omega_{o,a}$ and Q_a represented by (3.18) and (3.19), respectively. To derive these expressions, the non-dominant pole was approximated as $1/(1 + s/\omega_{p2}) \approx (1 - s/\omega_{p2})$ for $s/\omega_{p2} \ll 1$. Hence, (3.18) and (3.19) are valid only for $\omega \ll \omega_{p2}$. At $\omega \geq \omega_{p2}$, it causes additional attenuation and phase shift. In a filter, additional attenuation after ω_o is not necessarily a negative effect. However, the additional amplifier's phase shift can jeopardize the loop stability. In practical implementations, pushing non-dominant poles to higher frequencies increases the power consumption. Therefore, the designer should balance the trade-off between TF's error and power consumption. Moreover, inserting an LHP zero in the amplifier response can cancel the effect of the non-dominant pole [47].

$$\omega_{o,a} = \sqrt{\frac{\omega_{p2}^2(GBW^2 + 2 \cdot GBW \cdot \omega_{p1} + 6 \cdot \omega_{p1}^2)}{GBW^2 - 2 \cdot GBW \cdot \omega_{p2} + 6 \cdot \omega_{p2}^2}} \quad (3.18)$$

$$Q_a = \sqrt{\frac{(GBW^2 + 2 \cdot GBW \cdot \omega_{p1} + 6 \cdot \omega_{p1}^2)(GBW^2 - 2 \cdot GBW \cdot \omega_{p2} + 6 \cdot \omega_{p2}^2)}{4 \cdot (GBW^2 + (\omega_{p1} - \omega_{p2})GBW - 6 \cdot \omega_{p1} \cdot \omega_{p2})^2}} \quad (3.19)$$

In summary, reducing the filter error due to the amplifier non-idealities is done through: (i) increasing A_0 , (ii) increasing the ratio ω_{p2}/ω_t , and (iii) inserting an LHP zero that cancels or reduces the effect of ω_{p2} .

3.4.2 Active-R filter tuning

According to the first-order filter equations in Table 3.1, ω_o is defined by the product $\beta \cdot \omega_t$. However, as β also defines the filter DC gain, only ω_t can be used for frequency tuning. In the second-order filter all the circuit elements control both ω_c and Q . However, the 0 dB filter gain imposes the restriction $R_1 = R_3$. In this implementation we keep R_1 , R_3 , R_4 , and R_5 fixed and equal. R_2 , ω_{t1} , and ω_{t2} are programmable, enabling ω_o and Q tuning. First, Q tuning is performed by setting R_2 and the ratio ω_{t2}/ω_{t1} . Then, the values of ω_{t2} and ω_{t1} are both increased or decreased to achieve the ω_o spec, while keeping their ratio constant.

Q tuning is essential for two main reasons. First, it allows the realization of biquads with different Q s, required for high-order filter implementations. Also, a variable Q makes it easy to interchange the order of the filters. Second, it helps to compensate the Q -enhancement caused by the amplifier non-idealities and post-fabrication variations. It also allows relaxing the specifications of the amplifier TF.

3.4.3 Linearity and noise trade-off in the second-order active-R filters

As shown in [48], the noise of the second-order active-R filter shown in Table 3.1 can be analyzed by considering the total noise of each of its building blocks. Assume a single source represents the total noise power spectral density (PSD) of each lossy and lossless integrator. Each noise source includes the noise PSD of the amplifier and the resistors connected to its input terminals. It can be demonstrated that the output referred noise TF of the lossy integrator has a low-pass shape, while the noise TF of the lossless integrator has a band-pass shape [48]. The input-referred noise (IRN) is then dominated by the noise contribution of the lossy integrator (due to A_1 , R_1 , R_2 ,

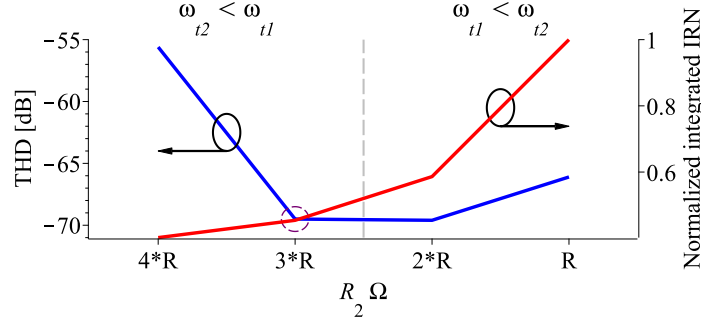


Figure 3.6: Simulated noise and linearity trade-off of the biquad active-R filter. Four different combinations of ω_{t1} , ω_{t2} and R_2 that implement the same f_o and Q are compared in terms of dynamic range. In the optimal design occurs when $\omega_{t1} > \omega_{t2}$. A larger $gm_1 = \omega_{t1}/C_1$ reduces the input-referred noise and maximize the loop gain to reduce distortion.

and R_3). In contrast, the lossy integrator's PSD (due to A_2 , R_4 , and R_5) is attenuated at DC, but it reaches its maximum around ω_o [48].

The linearity of the second-order active-R filter is affected differently by the linearity of each amplifier. A mechanism to improve the linearity on feedback-based filters consists of reducing the voltage swing at each amplifier's input [22]. In the biquad active-R filter, the lossy integrator's negative feedback attenuates the signal amplitude at the input of $A_1(s)$. However, the signal at the input of $A_2(s)$ is only attenuated by the voltage divider λ . Therefore, the biquad filter linearity improves by increasing the gain of A_1 and decreasing the gain of A_2 .

Consider the design of a biquad filter, which must meet certain ω_o and Q specifications. Since the circuit parameters R_2 , ω_{t1} , and ω_{t2} are programmable to allow tuning, there are several distinct designs that lead to the same ω_o and Q . However, these designs lead to different noise and linearity performance. This is illustrated in Fig. 3.6, where we assume $R_2 = (R, 2 * R, 3 * R, \text{ or } 4 * R)$. At each case, ω_{t1} and ω_{t2} are adjusted accordingly to meet the filter specifications.

Using the approximation $\omega_{ti} = gm_i/C_i$, an optimal design is achieved when $gm_1 > gm_2$, and gm_1 is sufficiently large. It reduces the IRN and increases the loop gain to attenuate the distortion of $A_1(s)$. Since the second amplifier has no local feedback, a larger gm_2 translates into a larger gain and higher distortion.

3.5 Gm-C vs Active-R filters

Conventional Active-RC filters require large GB/ω_o ratio to operate properly. Hence, they demand a higher bandwidth and power consumption. We explore two alternative implementation types which are more GB/ω_o efficient. Thus, a comparison between Active-R and Gm-C low pass filters is presented in [43]. For this purpose, a first order and a biquad low pass structure was implemented using both techniques; analysis and simulations allow to compare their performance in terms of power, linearity and noise. The effects of the amplifier non-ideal behavior are also addressed for both topologies, and we show how to leverage on those non-idealities.

In [43], the two biquads were designed for $f_o=10$ MHz, $Q=2$ and same power consumption; then, the performance is compared by simulation using IBM 130 nm CMOS technology. Under this conditions, both topologies have similar performance in terms of input referred noise, however the Active-R shows better linearity achieving a 21.3 dBm in-band IIP3, compared with a 3.3 dBm IIP3 of the Gm-C implementation. Therefore the dynamic range (DR) of the Active-R topology is 1.4 times larger.

4. SYNTHESIS OF HIGH-ORDER CONTINUOUSLY TUNABLE LOW-PASS ACTIVE-R FILTERS FOR BASE-BAND COMMUNICATIONS¹

4.1 Introduction and motivation

The growing demand for multi-standard wireless applications motivated the development of flexible radios, also known as software-defined radios (SDRs) [24,30,49]. These architectures rely on tunable circuits to cover a wide range of frequency bandwidths without performance degradation [50]. Tunable filters are fundamental blocks in multi-standard wireless receivers, as shown in Fig. 4.1 [51]. Low-IF and zero-IF receivers require low-pass filters (LPFs) for channel selection and out-of-band (OoB) signal rejection [52,53]. Examples of standards with a channel bandwidth (BW) ranging from 1 to 50 MHz are Bluetooth ($BW = 1$ MHz), code division multiple access (CDMA, $BW = 1.23$ MHz), Home-RF ($BW = 1\text{--}3.5$ MHz), WCDMA ($BW = 5$ MHz), ZigBee ($BW = 5$ MHz), WiMAX IEEE 802.16 ($BW = 1.25\text{--}20$ MHz), IEEE 802.11 ($a/b/g/n$, $BW = 20\text{--}40$ MHz) [24,26,50,54,55]. Hence, there is a need for low power highly tunable and highly linear filters.

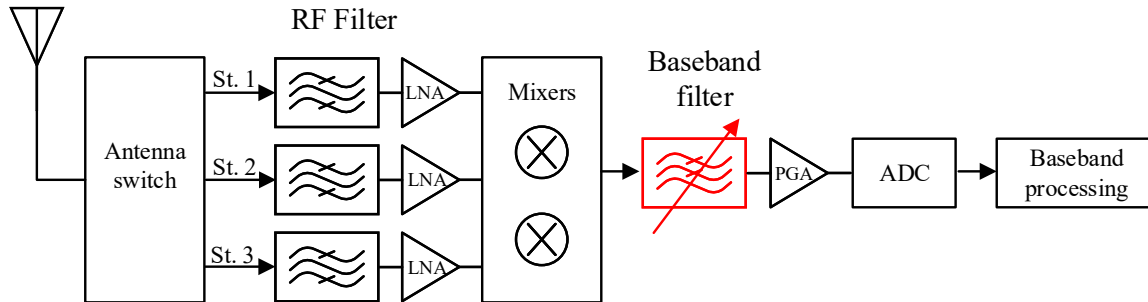


Figure 4.1: Multi-standard receiver. This project proposes a frequency tunable low-pass baseband filter (red).

¹Part of this chapter is reprinted with permission from “Synthesis of High-Order Continuously Tunable Low-Pass Active-R Filters” by A. Sanabria-Borbón and E. Sánchez-Sinencio, in IEEE Transactions on Circuits and Systems I: Regular Papers, Feb. 2021.

In this chapter, we discuss the circuit implementation and design considerations of high-order fully differential (FD) and frequency tunable active-R filters in CMOS technologies. Active-R filters use the internal poles of the amplifiers when implemented [35, 39, 42, 56].

The main contribution of this work is the implementation of active-R filters that are continuously tunable over a wide frequency range via programmable- ω_t amplifiers without a dynamic range degradation.

Three fifth-order Butterworth active-R LPF topologies were designed and fabricated using the TSMC 0.18 μm process. These filters operate at a ± 0.9 V supply and provide a continuously tunable ω_o in the range of 1 to 50 MHz. Compared with state-of-the-art filters, the performance of the proposed filters is competitive in terms of dynamic range and power efficiency, and FoM.

4.2 High-order active-R filter implementation

In filter design, the target attenuation in the stop-band defines the filter order. Thus, a particular communication standard with a higher attenuation or sharper roll-off spec calls for a higher filter order. This section discusses the synthesis of high-order and fully-differential active-R filters implemented with well-known architectures: CoB and MFL [19]. These filter architectures are built from the simple filter structures in Table 3.1.

4.2.1 Synthesis of a cascade of biquads (CoB) low-pass active-R filter

Any N th-order filter can be implemented as a cascade of simpler first-order and second-order stages [19]. The N th-order filter TFs are given based on the product of the individual stage TFs. Fig. 4.2 shows the block diagram of a fifth-order filter. Furthermore, lower filter orders (first,

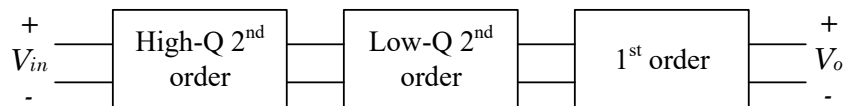


Figure 4.2: Implementation of a fifth order filter by cascading first-order and second-order stages. The order of the stages does not change the filter transfer function, but instead affects its noise performance.

1st	Z_{in}	$R_1 + \frac{R_2}{1+A(s)}$
	Z_o	$(R_1 + R_2) \parallel \frac{r_o}{1+A(s)\beta}, \beta = \frac{R_1}{R_1+R_2}$
2nd	Z_{in}	$R_3 + \frac{R_2}{1+A(s)} + \frac{R_1}{1+A(s)^2\lambda}$
	Z_o	$R_w \parallel \frac{r_o}{1+A(s)^2\lambda/R_w}, R_w = \left(R_1 + R_3 \parallel \frac{R_2}{1+A(s)} \right)$

Table 4.1: Input impedance (Z_{in}) and output impedance (Z_o) of the single-ended first-order and second-order active-R filters.

second, third, and fourth) are achieved by changing the input or output ports or by bypassing stages of the CoB architecture [26].

Since active-R filters inherently implement all-pole TFs, they are suited for the realization of Butterworth prototypes. Compared with other filter prototypes, Butterworth offers a maximally flat response in the passband and minimum group delay [19]. The prototype of a fifth-order Butterworth filter is written as [19]:

$$H(s) = \frac{1}{1+s} \cdot \frac{1}{1+0.68 \cdot s + s^2} \cdot \frac{1}{1+1.618 \cdot s + s^2} \quad (4.1)$$

This approximation assumes that each stage has an infinite Z_{in} and a zero Z_o [18]. In a practical implementation, the loading effect must be considered. The condition $|Z_{in,n+1}| \gg Z_{o,n}$ must be satisfied, when connecting each stage n to the following stage, $n+1$ for $n = 1, 2, \dots, N-1$. Table 4.1 summarizes the expressions of Z_{in} and Z_o of the single-ended first-order and second-order active-R filters. r_o represents the amplifier's output resistance. For CMOS amplifiers, we assume a very high input impedance.

Although the order of the stages connected in cascade does not change the filter TF, it can impact the total noise and linearity [19]. The IRN of the second-order filter is higher than that of the IRN of the first-order filter. This is because it has more elements contributing to the total noise. The filter stages described in this work have 0 dB gain in the pass-band, making the output referred noise equal to the input noise. It suggests that the order of the stages does not impact the

total IRN. However, the high- Q biquad has gain around f_o ($|H(j\omega_o)| > 1$). Hence, placing this stage first helps to attenuate the integrated IRN due to the following stages. The best configuration in terms of integrated IRN is 1) high- Q biquad, 2) low- Q biquad, and 3) first-order stage, as shown in Fig. 4.2.

4.2.2 Synthesis of multiple-loop feedback (MLF) active-R low-pass filters

MFL filters offer lower sensitivity to component variations than a CoB filter at the expense of higher design complexity [57, 58]. Here we discuss the realization of two MFL filter architectures i.e, follow the leader feedback (FLFB) and inverse follow the leader feedback (IFLFB).

4.2.2.1 Follow the leader feedback (FLFB)

FLFB is a particular case of multiple-loop feedback filters [23, 59]. The implementation of an N th-order LP FLFB consists of connecting N integrators $T = (-k)/(s + k)$ in cascade. Then, the output of each integrator is fed back to the input node through the feedback coefficients F_x , $x = 2, 3, \dots, N$ [57]. Fig. 4.3(a) shows the block diagram of a fifth-order FLFB LPF where k_{in} is a gain factor. Its TF is given by

$$H(s) = -\frac{k_{in}T^5}{1 + F_2T^2 + F_3T^3 + F_4T^4 + F_5T^5} \quad (4.2)$$

Another way to represent the prototype of a Butterworth filter is by its denominator's coefficients $B_N(s) = s^N + \sum_{i=0}^{N-1} a_i s^i$ [19]. Hence, the denominator of (4.1) is equivalent to $a_{i,i=0:4} = [1 \ 3.23 \ 5.23 \ 5.23 \ 3.23]$. The filter design consists of: (i) making the filter TF equal to the filter prototype, (ii) obtaining the system of equations by comparing the powers of s , (iii) solving the equation system to obtain the values of the variables F_x and k , and (iv) performing frequency

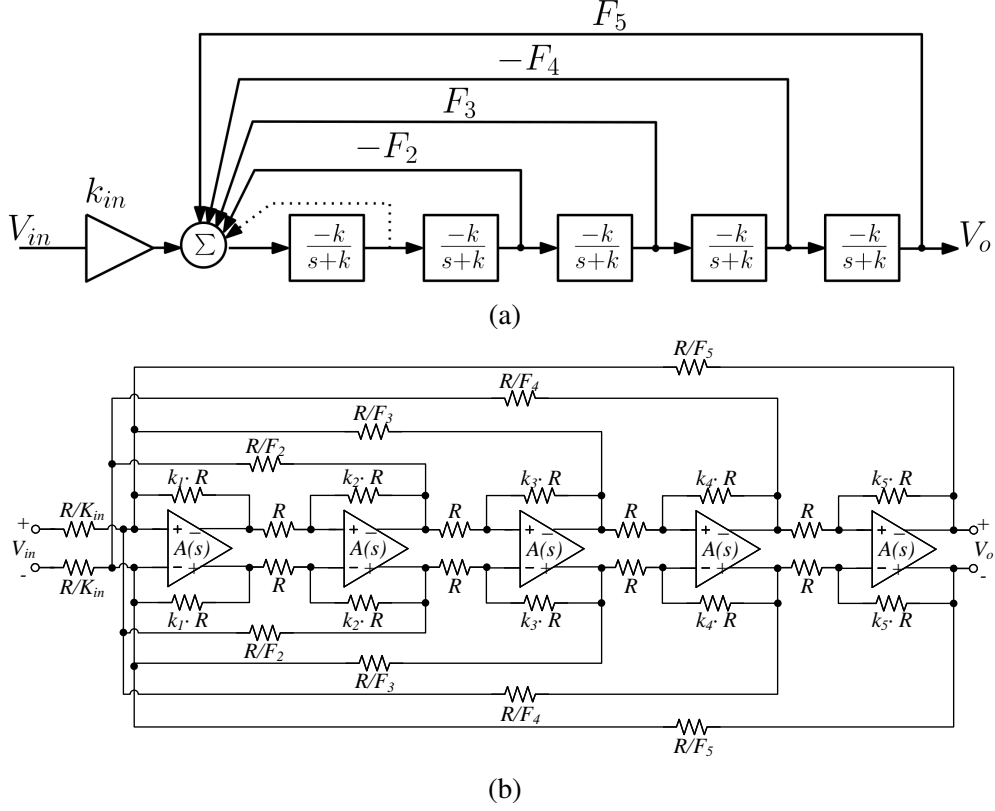


Figure 4.3: Fifth-order follow the leader feedback (FLFB) active-R filter: (a) Flow diagram and (b) fully differential circuit schematic.

transformation [59]. we can establish the set of equations

$$\left\{ \begin{array}{l} k^5(F_2 + F_3 + F_4 + F_5 + 1) = a_0 \\ k^4(3 \cdot F_2 + 2 \cdot F_3 + F_4 + 5) = a_1 \\ k^3(3 \cdot F_2 + F_3 + 10) = a_2 \\ k^2(F_2 + 10) = a_3 \\ 5 \cdot k = a_4 \end{array} \right. \quad (4.3)$$

$$T = \frac{-k}{s+k}. \quad (4.4)$$

The gain factor and the feedback coefficients can be further scaled to maximize the dynamic

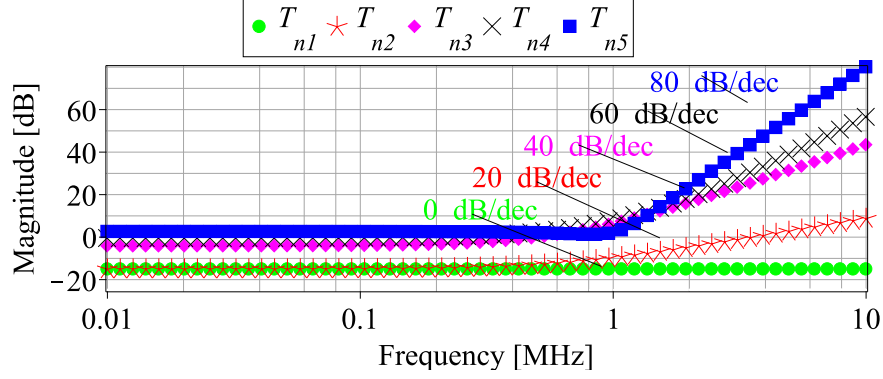


Figure 4.4: Modeled input-referred noise transfer function of each stage of the fifth-order FLFB filter for $f_o = 1MHz$.

range [59]. Hence, additional gain coefficients k_1, k_2, k_3, k_4 , and k_5 are optimized to control the signal amplitude at the internal nodes. Fig. 4.3(b) shows the circuit implementation of the fully-differential fifth-order FLFB active-R LPF. The first-order active-R filter (shown in Table 3.1) implements the integrator T . Note that all the loops have a negative feedback to guarantee stability. The feedback coefficients and the scaling factors are implemented by the resistor ratios. Hence, the value of R does affect the filter TF, but it impacts its noise performance. The input-referred PSD of the filter can be written as

$$e_{in}^2(\omega) = \sum_{i=1}^N e_i^2(\omega) |T_{ni}(j\omega)|^2 \quad (4.5)$$

where e_i^2 is the noise PSD of each stage T and $|T_{ni}(j\omega)|$ is the IRN transfer function of each stage i , shown in Fig. 4.4. e_i^2 includes the amplifier's noise (e_{Ai}^2) and the resistor's noise (e_R^2). Hence, we can write

$$e_i^2(\omega) = e_{Ai}^2(\omega) + e_{Ri}^2(\omega) \quad (4.6)$$

$$e_i^2(\omega) = e_{Ai}^2(\omega) + 4KT \left(\frac{1}{g_f + \sum g_{ik}} \right) \quad (4.7)$$

where g_f is the conductance associated with the feedback resistor, and $\sum g_{ik}$ represents the sum of all the other conductances connected to the input terminals of the amplifier at each stage i . Since

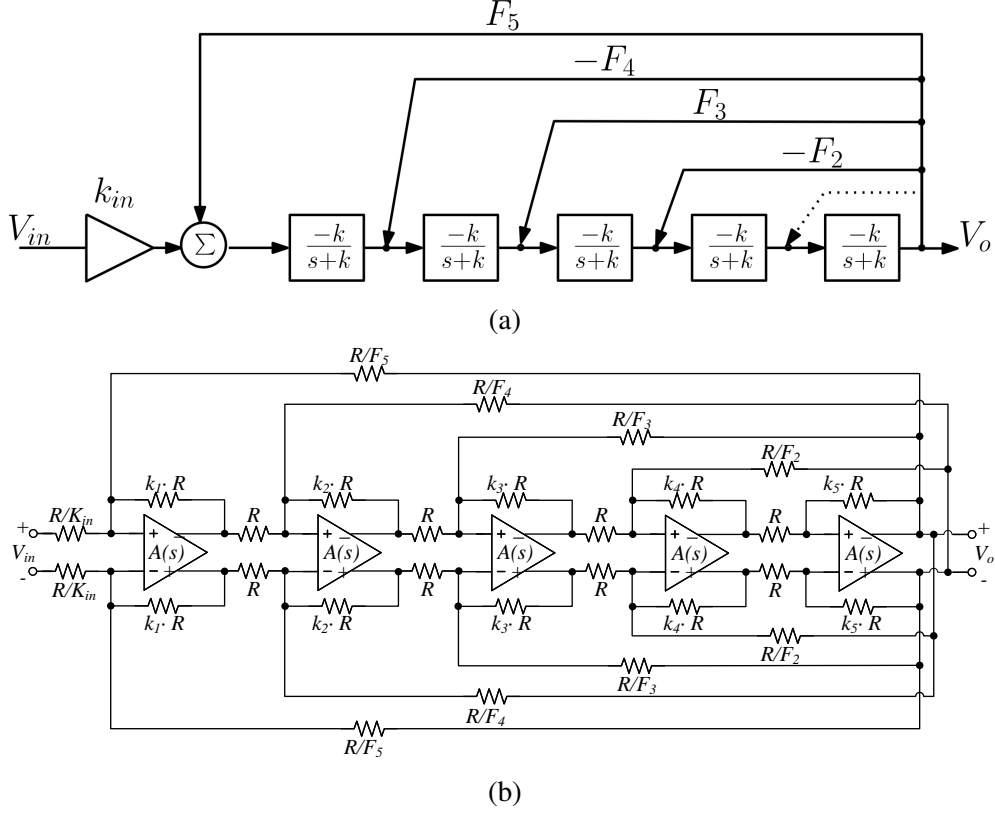


Figure 4.5: Fifth-order inverse follow the leader feedback (IFLFB) active-R filter: (a) Flow diagram and (b) fully differential circuit schematic.

all the feedback resistors are connected to the input terminal of the first integrator's amplifier, its noise contribution is more significant than that of the integrators that follow.

4.2.2.2 Inverse follow the leader feedback (IFLFB)

IFLFB is another MFL filter topology [60]. Similar to the FLFB, the IFLFB filter is based on the integrator T repeated in N times. However, as shown in the block diagram of Fig. 4.5(a), the feedback occurs from the filter's output to each stage T . The calculation of the feedback coefficients follows the same procedure as the FLFB filter. Also, the feedback coefficients are scaled to maximize dynamic range [36]. The fully-differential circuit implementation of the fifth-order IFLFB active-R LPF is shown in Fig. 4.5(b).

Compared with the FLFB filter topology, the IFLFB filter has a lower total IRN. The noise of the IFLFB filter can also be written as (4.5). Fig. 4.6 shows the IRN transfer functions $|T_{ni}(j\omega)|$.

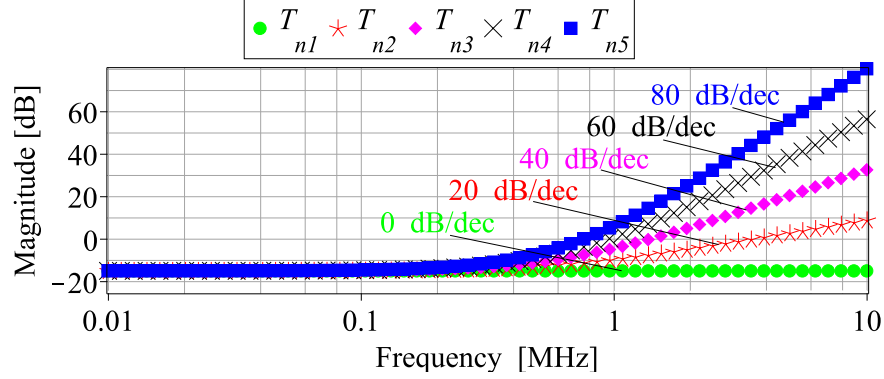


Figure 4.6: Modeled input-referred noise transfer function of each stage of the fifth-order IFLFB filter with $f_o=1$ MHz.

We observe that the iB-IRNs of all the stages have the same attenuation. Moreover, the 15 dB attenuation is larger than the attenuation of the stage's PSD on the FLFB configuration.

In summary, since the signal amplitude is optimized in both MFL topologies, they have a similar linearity. However, IFLFB has a better dynamic range than the FLFB topology due to its lower integrated noise.

4.3 Circuit implementation

In the previous section, we presented three topologies for implementing a fifth-order Butterworth active-R LPF. We discussed the system-level design of the fully-differential version of the CoB (Fig. 4.2), FLFB (Fig. 4.3(c)), and IFLFB (Fig. 4.3(d)). This section discusses the circuit implementation and design considerations of the CMOS amplifier and the resistor values building these filters. The three filter topologies use the same amplifier. Hence, we can compare their performance in terms of dynamic range since their area and power consumption are similar.

4.3.1 Amplifier with tunable ω_t

An amplifier with programmable ω_t enables the filter's ω_o tuning. The two-stage internally compensated OTA shown in Fig. 4.7 was selected for amplifier implementation. Its main characteristics are simplicity and robustness [47]. Its fully-differential structure increases the dynamic range by doubling the output swing, suppressing even-order harmonics, and canceling common-

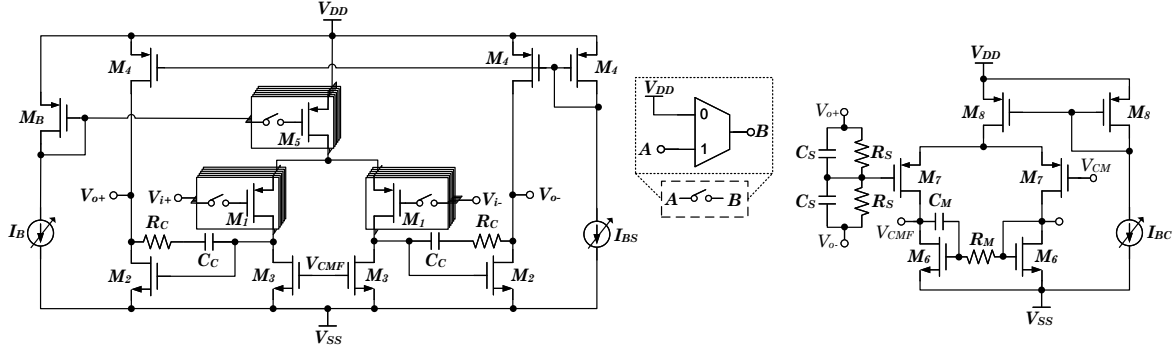


Figure 4.7: Two-stage fully-differential Miller amplifier with programmable unity gain frequency ω_t . While the configurable input pair provides coarse tuning, the external biasing current I_B allows continuous tuning. A transmission gate-based multiplexor implements each switch. The common-mode feedback circuit includes passive sensing and a single-stage error amplifier.

mode noise [61]. However, this OTA has limitations when driving low impedance loads. An output buffer can help drive such loads, but it comes at the expense of prohibitive large power and area overheads. The ω_t of this amplifier is given by [61]

$$\omega_t = \frac{g_{m1} \cdot R_{o1} \cdot g_{m2} \cdot R_{o2}}{R_{o1}(C_1 + C_C(1 + g_{m2} \cdot R_{o2})) + R_{o2}(C_{Ltot} + C_C)} \quad (4.8)$$

where g_{m1} and g_{m2} are the transconductances, while R_{o1} and R_{o2} are the output impedances of the first and second stages, respectively. The output impedances are defined as $R_{o1} = 1/(g_{ds1} + g_{ds3})$ and $R_{o2} = 1/(g_{ds2} + g_{ds4}) || R_s || R_L$. C_C is the Miller compensation capacitor, and C_{Ltot} is the total load capacitance modeling the equivalent parasitics of the pad, bondwire, and packaging. C_1 represents the total capacitance at the output node of the first stage. R_L represents the equivalent load resistance estimated using Table 4.1.

When the gain of the amplifier stages ($g_{m1} \cdot R_{o1}$, $g_{m2} \cdot R_{o2}$) are large, the amplifier's unity gain frequency can be approximated to $\omega_t \cong g_{m1}/C_C$. A configurable ω_t requires a programmable g_{m1} , or C_C , or both. Implementing C_C with a capacitor bank that has a sufficient range and resolution requires a huge area overhead. Instead, the dependency of g_{m1} on the transistor's aspect ratio $(W/L)_1$ and its current I_{D1} facilitate the coarse and fine tuning required to cover a wide tuning

	M_1	M_2	M_3	M_4	M_5	M_6	M_7	M_8
L[μm]	0.18	0.54	0.90	0.36	0.18	0.36	0.5	0.36
W[μm]	0.5	20	675	3.6	1.8	0.36	2	1
n	2	8	2	8	2	1	4	4
m	1:32	1	4	8	1:32	1	1	1

Table 4.2: Transistor sizing of the two-stage Miller amplifier in Fig. 4.7.

range. As shown in Fig. 4.7, M_1 and M_5 are implemented with transistor banks of weighed aspect ratios $(W/L)_x = 2^k \times (W/L)_{u,x}$ with $k = 0, 1, \dots, 5$. The tail transistor M_5 is scaled accordingly to keep $(gm/ID)_{M_1}$ and (V_{dsat,M_1}) constant. However, the variation of I_B that allows the fine-tuning of g_{m1} comes at the expense of variation on the gm/ID . The tuning range design is such that it allows only a $1.45X$ gm/ID variation, which does not have a significant impact on the amplifier's linearity [21].

Transmission gate (TG) based multiplexers implement the switches of the transistor's banks, as shown in Fig. 4.7. When the selector bit is '0', the transistor's gate terminal is connected to V_{DD} to ensure the OFF state and reduce leakage. Otherwise, when the selector bit is '1', the differential pair is turned ON with a small equivalent resistance in series with the transistor's gate terminal. C_C is implemented by a 1 pF mimcap distributed in 20 multipliers to improve the matching in a common centroid layout [26].

The design of the amplifier's second stage includes several considerations. In a fixed design, the tuning of g_{m1} causes the ratio of ω_t to ω_{p2} to change, degrading the phase margin (PM) of the amplifier and jeopardizing its stability. A configurable design, like the one in [27], allows us to keep the ratio ω_t/ω_{p2} constant at the expense of additional area overhead. This design keeps the sizing fixed and ensures a $PM > 80^\circ$ for all values of ω_t . It increases power consumption but requires a smaller integrated area. The amplifier TF has three poles and one zero [21, 61]. Their location is designed to reduce the error in the filter TF while ensuring stability. The sizing of all the MOSFET devices is summarized in Table 4.2.

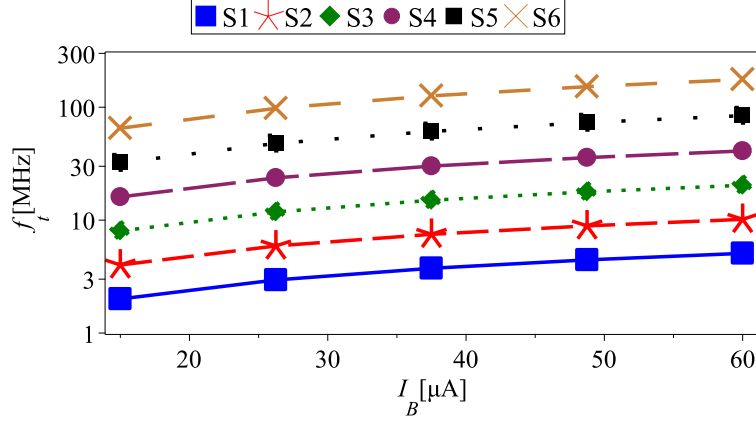


Figure 4.8: Simulation results of the f_t tuning range. The configurable input differential pair provides coarse-tuning (S1, S2, S3, S4, S5, S6), while the fine-tuning is done by changing the biasing current (I_B). With the proposed design we achieve a 90x range which is sufficient for the frequency tunable active-R filter implementation.

Moreover, $R_C = 1 \text{ k}\Omega$ transforms the RHP zero due to the Miller capacitor, into an LHP zero [47], which subsequently improves the PM and reduces the error in the filter TF [18,61].

The simulated 2–180 MHz f_t range is shown in Fig. 4.8. This range is sufficient to cover the design of the three filter topologies described in Section 4.2. Across this frequency range, the amplifier’s gain varies only between 63.21 dB and 65.36 dB, while the PM is larger than 81.95° for all the f_t values.

4.3.2 Common-mode feedback (CMFB) circuit

The CMFB circuit consists of the passive sensor ($R_s = 100 \text{ k}\Omega$ and $C_s = 60 \text{ fF}$) and the error amplifier, as shown in Fig. 4.7. The error amplifier is a single-stage OTA with a diode-connected load. Its output controls the active load of the core amplifier’s first stage (M_3) [62]. Adding $C_M = 1 \text{ pF}$ and $R_M = 50 \text{ k}\Omega$ to the error amplifier have two main effects on the CMFB loop [63]. First, it increases the loop-gain at low frequencies by 27.25 dB, improving the accuracy of the common-mode voltage. Second, it provides frequency compensation to ensure stability on a wide BW .

Three gain stages form the CMFB loop: the core amplifier’s two stages and the error amplifier. Despite tuning the amplifier’s ω_t modifies the CMFB loop’s BW , it always larger than f_o . It can

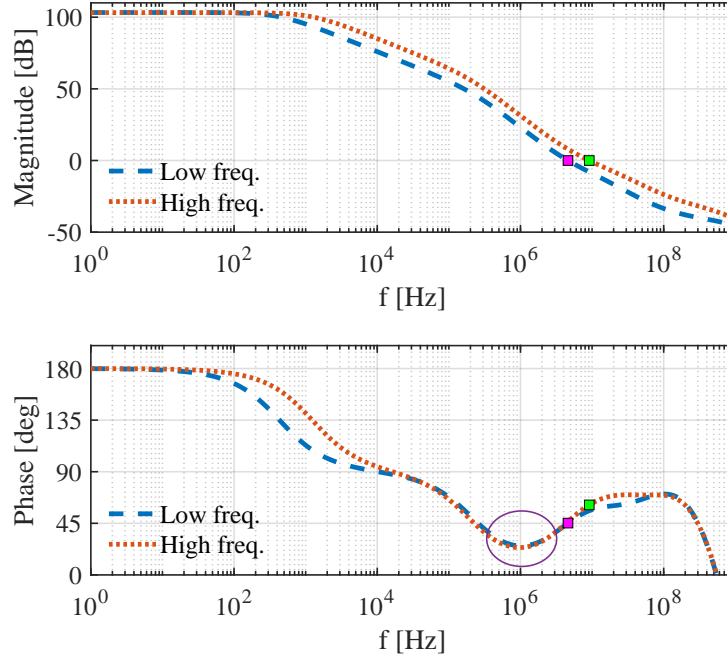


Figure 4.9: Simulated frequency response (magnitude and phase) of the common-mode feedback loop. The plots show the loop-gain for the low and high frequency configurations when the smallest differential pair is selected.

be demonstrated (see [51, 63]) that the loop has two low-frequency poles that can jeopardize the loop stability. However, two mid-frequency LHP zeros compensate the phase due to the poles and ensure stability, as shown in Fig. 4.9. Nevertheless, the worst-case PM is 45.6° for all the frequency range and considering process, voltage, and temperature (PVT) variations.

4.3.3 Resistor design

The active-R filter characteristics are defined by resistor ratios instead of their values. However, the resistor values affect the error in the filter TF and the total integrated noise. While small resistors have a low thermal noise contribution, they increase the error in the filter TF. A small resistor connected at the amplifier's output decreases the loop gain leading to a larger error in the pass band. It also reduces the frequency of the non-dominant poles jeopardizing stability. On the other hand, a large resistor does not degrade the loop gain, but it increases the total thermal noise. Also, large resistors interact with the amplifier's input capacitor, reducing non-dominant

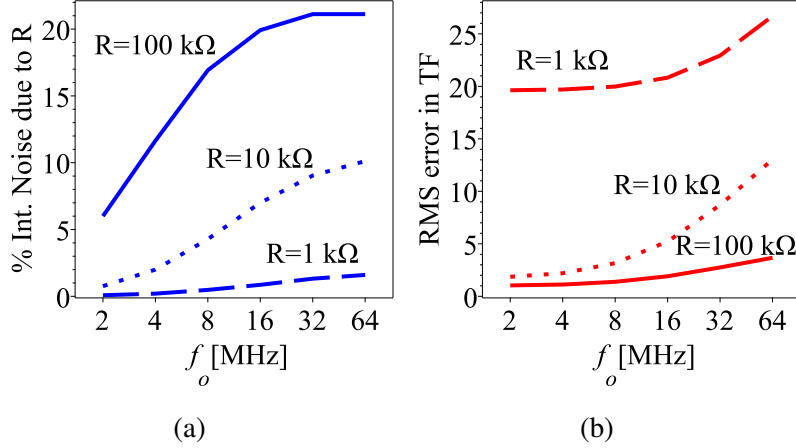


Figure 4.10: Effect of the resistor values in the fully-differential first-order active R across f_o . The percentage of integrated noise due to the resistors is shown (a). The RMS error of the transfer function due to different resistor values (b).

poles' frequency, which increases the phase shift. A simulation was performed in a first-order fully differential active-R filter with a configurable ω_o to observe the effect of three resistor sizes 1 k Ω , 10 k Ω , and 100 k Ω . Fig. 4.10(a) shows the percentage of the total in-band integrated noise that is due to the four resistors. When $R = 100\text{ k}\Omega$ resistor's noise contribution grows to 80% for the high BW condition, while it stays below 10% for $R = 1\text{ k}\Omega$. Fig. 4.10(b) shows the RMS error of the filter TF for the three resistor values. The TF error due to a 1 k Ω resistor is more than 20 times the error due to the 100 k Ω one. The $R = 10\text{ k}\Omega$ resistor shows the best trade-off between noise and error, as presented in Fig. 4.10.

Based on the noise-error trade-off illustrated in Fig. 4.10, the value of all CoB resistors is set to 10 k Ω , except for R_2 . The programmable resistor array shown in Fig. 4.11 implements R_2 of the second-order active-R filter. The switches are implemented with transmission gates for good linearity because they operate across a large voltage swing. Moreover, the terminal P_A is connected to the amplifier's input node, which experiences lower voltage swings than the amplifier's output node.

The reference resistor of the FLFB and IFLFB designs is $R = 10\text{ k}\Omega$. However, the actual resistor values are scaled with the feedback coefficients and the gain factors obtained from the sys-

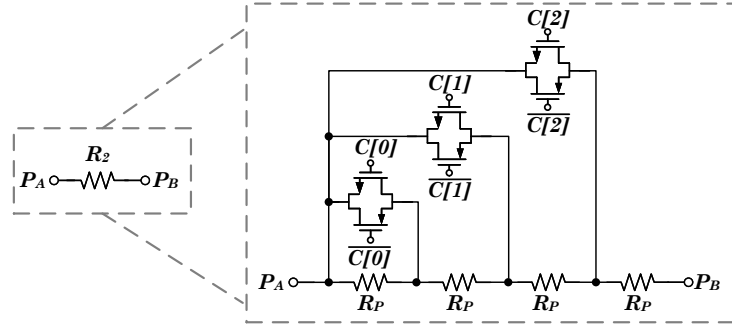


Figure 4.11: Digitally programmable resistor implementing the resistor R_2 that enables Q and ω_o tuning of the second-order active-R filter.

tem design, which consist of rational numbers. These resistor's rational values are approximated to the closest available size in the technology. Moreover, in the multi-loop topologies, a configurable ω_t enables ω_o tuning, whereas Q remains constant.

4.4 Experimental results

4.4.1 Experimental setup

The three active-R filter topologies: CoB, FLFB, and IFLFB, were fabricated using the TSMC 0.18 μm CMOS process. Each filter implements a 5th order low-pass Butterworth response with a continuously tunable cutoff frequency in the range 1–50 MHz.

Fig. 4.12 shows the chip microphotograph including the 44 pads. The CoB filter area is $413 \mu\text{m} \times 808 \mu\text{m}$, including the programmable resistors, and the scan-chain controlling the tuning knobs. Using a scan-chain for the digital control trades the number of pins with an area overhead. The areas of the FLFB and IFLFB filters are $408 \mu\text{m} \times 750 \mu\text{m}$ and $380 \mu\text{m} \times 750 \mu\text{m}$, respectively. The integrated area of the FLFB includes the scan-chain controlling the tuning knobs of the two multiple-feedback filters. The filters operate with a dual $\pm 0.9 \text{ V}$ supply.

Fig. 4.13 illustrates the measurement setup. At the input of the fully-differential filters, the Coilcraft PWB2010 balun provides a single-ended (SE) to differential (D) signal conversion. At each filter's output, a voltage buffer isolates the filter's output from the 50Ω impedance of the balun. The 1 GHz FastFET ADA4817-2 Op-Amp is connected in unity feedback to implement the

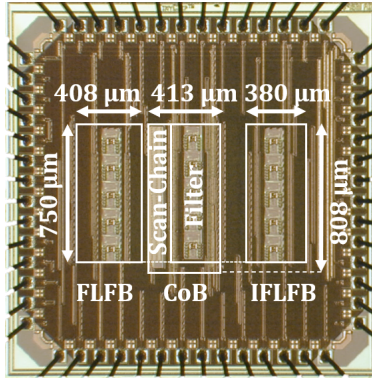


Figure 4.12: Chip microphotograph of the three fifth-order low-pass active-R filter topologies: the cascade of biquads (CoB), the follow the leader feedback (FLFB), and the inverse follow the leader feedback (IFLFB). The three structures use the same amplifier but differ on the resistive feedback configuration. The designs are fabricated using the TSMC 0.18 μm CMOS process.

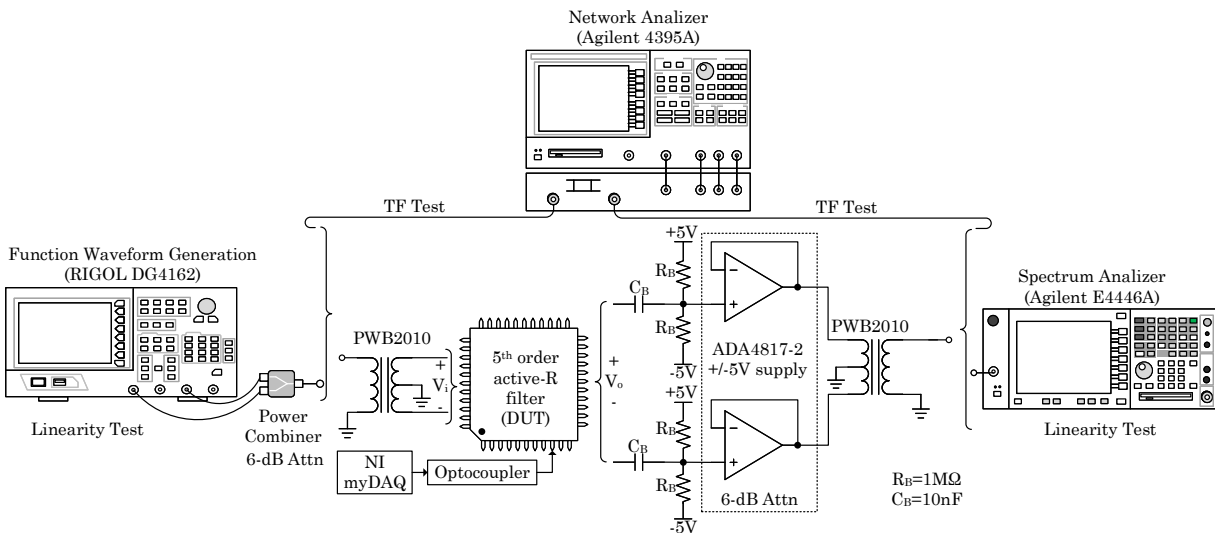


Figure 4.13: Measurement setup for the filter performance characterization. The PCB includes baluns, voltage buffers, potentiometers, and 50 Ω SMA connectors. The measurement equipment includes a dual-channel signal generator and a spectrum/network analyzer. The NI-DAQ board generates the control signals for the scan-chains.

buffer. A dual amplifier was used for better differential signal matching. A second balun combines the differential outputs into an SE one. Potentiometers control the bias current of the amplifiers used for the fine-tuning of the cutoff frequency. A NI-myDAQ board provides the control signals to the scan chains setting the coarse frequency tuning. Moreover, optocouplers act as an interface

to adjust the voltage levels between the NI-myDAQ and the fabricated IC. The filter TF (S_{21}) is measured using a network analyzer. This test enables the verification of the frequency tuning of the three filter topologies. The RIGOL DG4162 dual-channel signal generator, a power combiner, and the Agilent E4446A spectrum analyzer were used to characterize the filter linearity. The power combiner causes a 6 dB attenuation, while the buffer's load contributes an additional 6 dB attenuation.

4.4.2 Measurement results of the second-order active-R Q tuning

The integrated scan-chain sets the control bits of the programmable resistor R_2 (Fig. 4.11) of each active-R biquad filter. Fig. 4.14 shows the filter TF for each configuration code $C[2:0]$

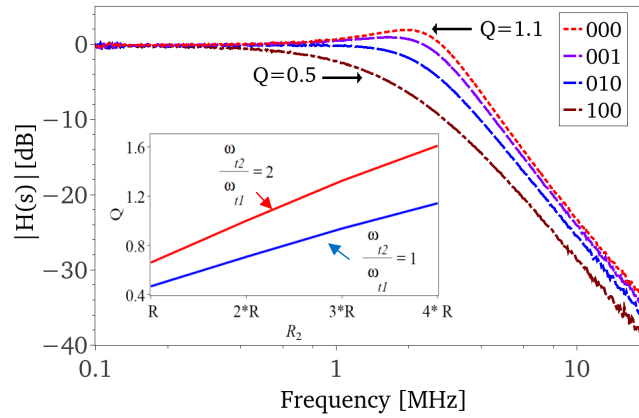


Figure 4.14: Measured Q tuning of a biquad active-R filter using the digitally programmable resistor R_2 for different ratios ω_{t2}/ω_{t1} . Each TF corresponds to a different configuration of $C[2:0]$.

controlling R_2 when $\omega_{t2} = \omega_{t1}$. Fig. 4.14 also summarizes the dependency of Q on the ratio ω_{t2}/ω_{t1} . The larger the ratio, the higher the Q is achieved.

4.4.3 Measurement results of f_o tuning of the three fifth-order active-R architectures

Section 4.3 shows how the coarse and fine tuning of the amplifier enables the 50X frequency range of each filter topology. Fig. 4.15(a) shows the measured TF of the fifth-order CoB active-R LPF across the six coarse frequency tuning steps. In this test, the amplifier's biasing is kept

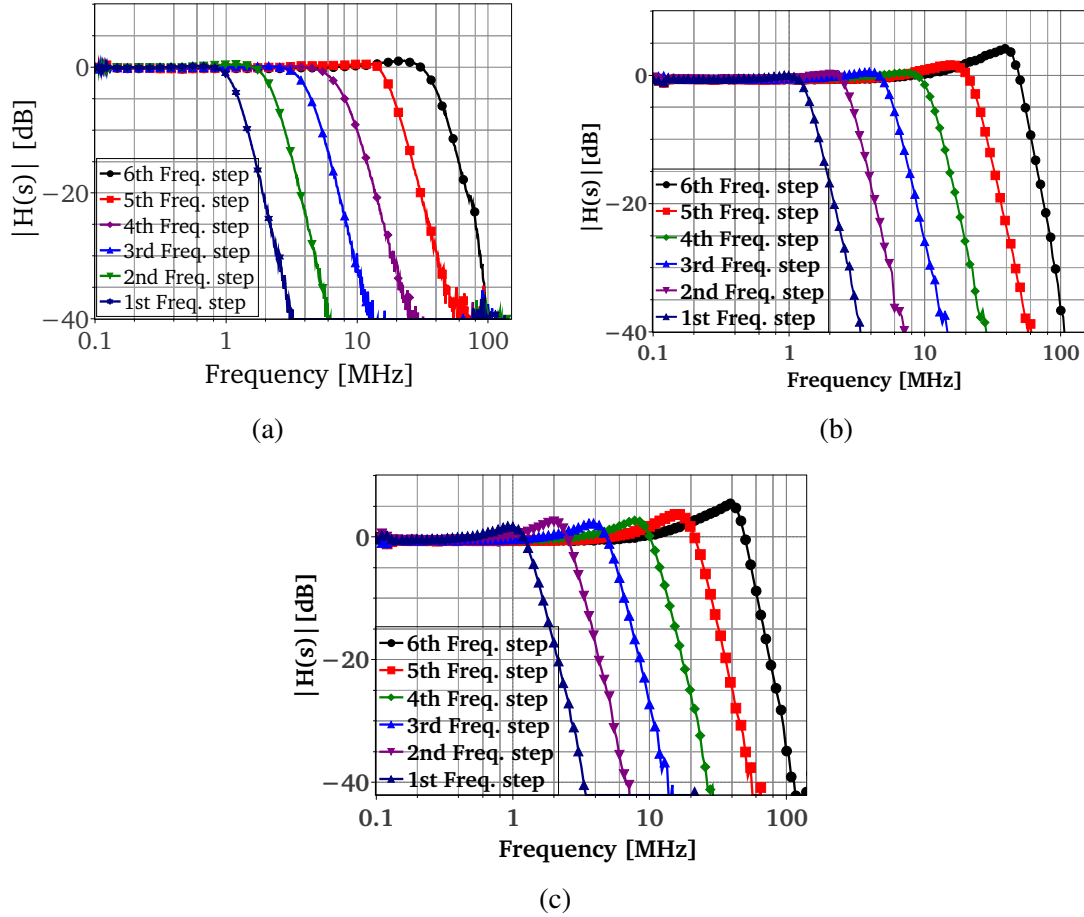


Figure 4.15: Measured transfer function of the fifth-order Butterworth active-R filters: (a) CoB, (b) FLFB, and (c) IFLFB. The coarse-tuning steps depend on the digitally-programmable input differential pair. Any cutoff frequency within these steps is achieved by changing the amplifiers' biasing current.

constant. The different TFs are achieved by selecting one differential pair at a time. Also, R_2 is configured accordingly. Table 4.3 summarizes the range of continuous frequency tuning, which is controlled by the amplifier's biasing.

Coarse tuning step	1	2	3	4	5	6
f_o [MHz]	1-2.5	2-5	4-10	8-20	16-40	32-75

Table 4.3: Continuous ω_o tuning range at each coarse tuning step of the CoB filter.

Fig. 4.15(b) and 4.15(c) show the measured TFs of the FLFB and IFLBF filters, respectively. Each figure shows the TFs achieved by coarse tuning. The amplifiers' external biasing enables continuous frequency tuning within these steps. The two filter topologies display a tuning range of at least 50X. The multiple-feedback active-R structures display a Q enhancement. It increases at a high ω_o reaching a maximum gain of 5 dB. Although, the frequency tuning of these topologies is provided by the programmable amplifier, the Q calibration was not enabled in this design. Hence, the Q enhancement of this topology cannot be calibrated post-fabrication. This issue can be resolved by allowing independent biasing of the amplifiers and programmable resistors for the feedback realization. This solution enables post-fabrication calibration at the expense of extra area overhead.

4.4.4 Performance comparison

The performance of the three fabricated filters was characterized and compared. A two-tone ($f_{IM,L}, f_{IM,H}$) test characterized the in-band (iB) and out-of-band (OoB) linearity of the filters. The test tones were adjusted keeping the ratio $f_o/f_{IM,L} = 3.33$ constant for all f_o values. Note that both the fundamental tones and their intermodulation products had a 12 dB attenuation. However, this did not affect the IIP3 measurement. The measured IIP3 included not only the filter distortion but also the distortion of the buffer in the measurement setup. The distortion of the filter was extracted following the procedure in [45].

Fig. 4.16 and Fig. 4.17 show the linearity measurements of the CoB topology at $f_o = 1$ MHz. The estimated linearity measurements are 27.5 dBm iB-IIP3, 56 dBm iB-IIP2, and 16 dBm OoB-IIP3. The OoB-IIP3 varies from 15.77 dBm to 26.78 dBm across the f_o range. The OoB linearity is lower than the iB linearity due to the low loop-gain of the active-R topology outside its pass-band. The iB-IIP3 of the three configurations are reported in Fig. 4.18. Furthermore, the OoB-IIP3 varies from 15.75 dBm to 26.98 dBm across the f_o range. The iB-IIP2 is larger than 55.21 dBm for all values of f_o .

The comprehensive figure-of-merit (FoM) in (4.9) allows us to compare the performance of active filters [23]. This FoM includes the cutoff frequency (f_o), the order of the filter (N), the

total power consumption (PW), and the dynamic range ($IMFDR_3$).

$$F_{oM} = 10 \cdot \log_{10} \frac{IMFDR_3 \cdot f_o \cdot N}{PW} \cdot \frac{f_{IM3,LOW}}{f_{poles}} \quad (4.9)$$

where, $IMFDR_3$ is calculated with the iB-IRN ($V_{N,in}$) and the iB-IIP3 as

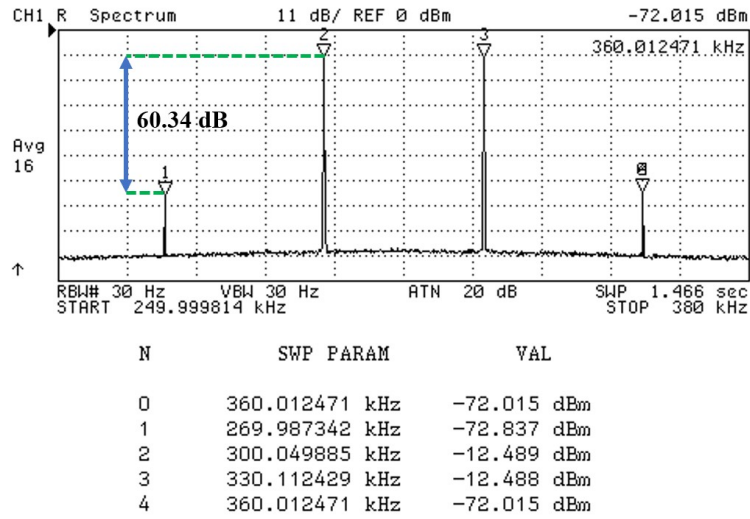


Figure 4.16: Measurement of the two-tone test of the fifth-order CoB filter at $f_o = 1$ MHz. We applied tones at $f_1 = 300$ kHz and $f_2 = 330$ kHz, generating intermodulation tones at 270 kHz and 360 kHz.

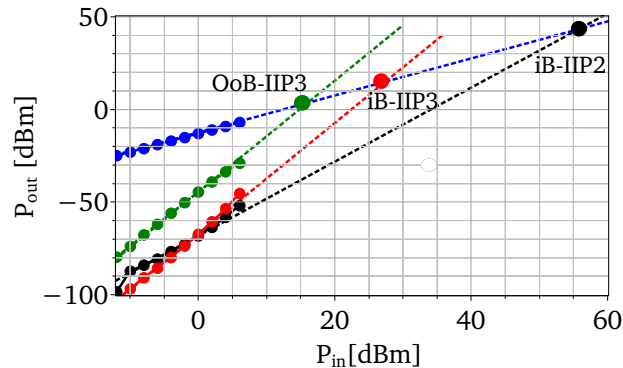
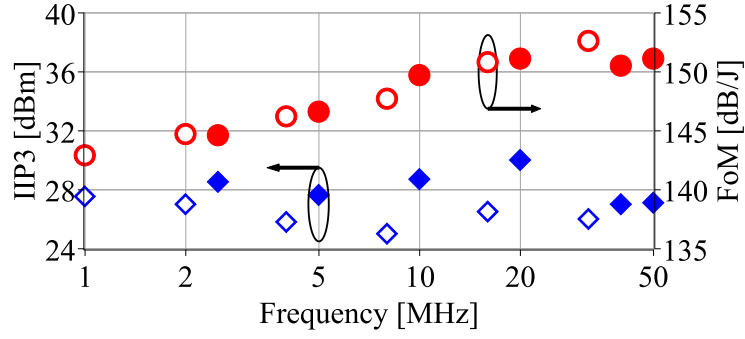
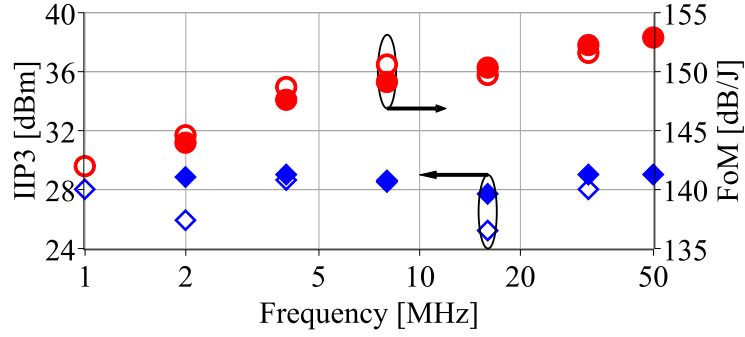


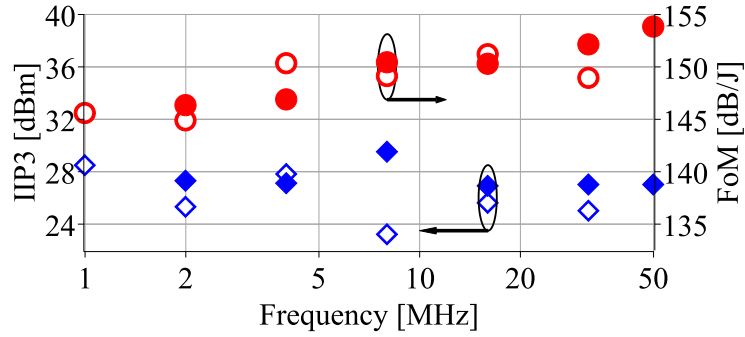
Figure 4.17: Measured iB-IIP3 (red), iB-IIP2 (black), and OoB-IIP3 (green) of the 5th order CoB filter at $f_o = 1$ MHz. The iB tests are performed with tones at $f_1 = 300$ kHz and $f_2 = 330$ kHz. The OoB test is performed with tones at $f_1 = 5$ MHz and $f_2 = 9.7$ MHz.



(a)



(b)



(c)

Figure 4.18: Measured in-band IIP3 and FoM of the fifth-order active-R filters (a) CoB, (b) FLFB, and (c) IFLFB. The measurement data of the low (\diamond, \circ) and high (\blacklozenge, \bullet) I_B is shown for each coarse tuning step.

$$IMFDR_3 = \left(\frac{IIP3}{V_{N,in}} \right)^{4/3}. \quad (4.10)$$

Fig. 4.18 shows the measured iB-IIP3 and the FoM for each filter topology across their f_o tuning range. At each coarse frequency step, we measured the integrated IRN and linearity at both

Parameter	De Matteis JSSC'09 [64]	Amir JSSC'09 [26]	Lo JSSC'09 [53]	Acosta TCAS-I'09 [29]	Thyagarajan JSSC'11 [22]	Ye TCAS-I'13 [45]	Vigramham JSSC'14 [25]	Wang TCAS-II'15 [50]	De Matteis JSSC'17 [23]	Lavalle TVLSI'20 [27]	This work CoB	This work FLFB	This work IFLFB	
Topology	Active-gm-RC	Active-RC	Gm -C	Gm -C	Active-RC	Active-RC	Active-RC	Active-RC	Active-RC	Active-gm-RC	Active-RC	Active-R	Active-R	Active-R
Order-type.	4-But.	5-Cheb.	3-But.	3-Cheb.	5-Cheb.	6-But.	4-But.	4-But.	4-But.	4-But.	5-But.	5-But.	5-But.	
Technology [nm]	130	130	180	500	180	180	65	65	180	130	180	180	180	
Supply [V]	0.55	1	1.2	5	1.8	1.8	0.6	1.2	1.8	0.6	1.8	1.8	1.8	
Power [mW]	3.5	3/7.5	4.1/11.1	30	5.6	0.4	26.2	1.72/9.6	12.6	5.56/23.77	7.45/9.38	7.45/9.38	7.45/9.38	
Power/pole [mW]	0.87	0.6/1.5	1.36/3.7	60	1.12	0.06	6.55	2.4	3.15	1.39/5.94	1.49/1.87	1.49/1.87	1.49/1.87	
f_0 [MHz]	11.3	1/20	0.5/20	6/12	20	10	70	0.2/20	22.5	20/160	1/50	1/50	1/50	
Continuous tuning	No	Yes	Yes	Yes	No	No	No	No	No	No	Yes	Yes	Yes	
IRN [nV/$\sqrt{\text{Hz}}$]	32.7	85/52	425/12	64	219	63.7	43.6	87	18.3	54/41	172/99	214/81	121/60	
Int. IRN [μV_{RMS}]	110	85/232	343/54	156	980	201	365	38.9	87	243	172/630	214/576	121/424	
IIP3 [dBm]	10	31/26	22.3/19	33	43.3	14.4	18	19	21.5	9.56/21	27.5/27.1	28/29	28.5/27	
DR [dB]	60	71.4	–	79.3	60.3	58	65	–	–	64.2/51	71.2-61.5	70.1-62.3	70.3-64.6	
Area [mm²]	0.43	1.53	0.23	1.5	0.2	0.21	0.38	0.8	0.35	0.24	0.33	0.30	0.28	
FoM [dB/J]	144.4	143.2	132.8/152.9 [†]	145.5/148.5	152	126	134	144.8	150.8	149.8	142.9/151.1	141.9/152.8	145.6/153.3	

[†] Assuming an IIP3 test tone at $f_{IM3,LOW} = 0.3 * f_o$

Table 4.4: Comparison of performance metrics of the three proposed active-R topologies proposed in this work with the state-of-the-art in active CT filters.

the low-frequency and the high-frequency extremes. The three filter topologies exhibit similar linearity with a 23.3 dBm to 30 dBm IIP3 range across the whole tuning range. Moreover, in the three cases, the FoM increases proportionally to f_o . This is because increasing f_o only requires a small power increment. However, the estimated FoM of the IFLFB filter is slightly better than the other two topologies due to its lower integrated IRN.

Table 4.4 summarizes the performance metrics of the three active-R topologies presented in this work. These results are compared with previously reported filters on similar frequency ranges and characteristics. The active-R filters presented in this work have the largest continuous tuning range and competitive linearity and noise. The large IRN of active-R topologies is compensated with a high iB linearity and power efficiency. The results feature a superior FoM at a high f_o compared with the state-of-the-art filters.

4.4.5 Limitations of this work

The close relationship between the TF of the active-R filter and its amplifier, which allows frequency tuning, makes them sensitive to process, voltage, and temperature (PVT) variations. Such sensitivity can be reduced by employing a robust biasing circuit [47]. Additionally, analog circuits can include an automatic calibration techniques to compensate for the variations. Reported calibration techniques are based in configurations such as: master-slave [19], phase-locked-loop (PLL) [40], band-pass-based oscillators [65], charge comparator-based frequency controllers [66], replica-circuit-based tuning [52, 67], and on-chip optimization [68]. All these techniques are compatible with the filters presented in this work.

4.5 Conclusion

In this work, we revisited the concept of active-R filters and discussed their implementation and design considerations using CMOS technologies. Three fifth-order fully differential and continuously tunable active-R LPF topologies were designed, fabricated, and tested in the TSMC 0.18 μm . Amplifiers with programmable ω_t enable continuous frequency tuning over the range from 1 to 50 MHz without dynamic range degradation. The measured FoM of the filters reported in this work is competitive with the state-of-the-art filters. This work demonstrates that active-R topologies are viable for multi-standard receivers since they can achieve higher frequencies than active-RC filters and better dynamic range than Gm -C filters. Also, active-R filters can achieve a wide tuning range with a small integrated area (0.33 mm^2).

5. GAUSSIAN PROCESS BASED SURROGATE FOR OPTIMIZATION AIDED AND PROCESS VARIATIONS AWARE ANALOG CIRCUIT DESIGN ¹

5.1 Introduction

5.1.1 Motivation

While highly automated Computer-aided design (CAD) tools are now commonly used for optimization, synthesis, placement, and routing of digital circuits, despite significant efforts, design automation has not been yet standardized for analog design [69]. The creation of automatic design tools that can address all the challenges of analog/RF integrated circuit (IC) design is not a trivial task. Conflicting design specifications, large non-continuous non-convex search spaces, increasingly complex device models in modern technologies, stringent power and area requirements, as well as shorter time-to-market cycles, are some of the main challenges of analog/RF design [20].

The traditional analog/RF design process starts by selecting the topology and translating the specifications from the architecture level to the circuit level [20, 70]. Subsequently, the sizes of the devices and the biasing conditions are selected such that the performance of the circuit meets the specifications. Commonly, the sizing and biasing steps are iterative processes that use simplified approximations followed by verification with circuit simulations. Therefore, achieving a first-time design that meets the specifications relies heavily on the experience of the designer and computationally expensive simulations [71]. Due to the stringent time-to-market demands, this procedure commonly lacks exploration of the available design space, and it does not guarantee that the solution found is near-optimal. In fact, circuits are often over-designed to ensure accuracy and robustness, which can lead to an excessive cost on silicon area or power consumption.

To reduce design effort while achieving high-performance circuits, extensive research has been done on optimization-based electronic design automation (EDA) tools for automatic analog/RF

¹Reprinted with permission from "Gaussian-Process-Based Surrogate for Optimization-Aided and Process-Variations-Aware Analog Circuit Design" by Sanabria-Borbón, A.C.; Soto-Aguilar, S.; Estrada-López, J.J.; Allaire, D.; Sánchez-Sinencio, E., 2020, *Electronics*, 9, 685.

IC design [69, 70, 72, 73]. A circuit sizing and biasing task can be formulated as a constrained multi-objective optimization problem by translating the circuit parameters, performance metrics, and specifications into design variables, objective functions, and constraints. The main advantage of using an optimization algorithm for circuit sizing is that it considers all the design variables and design specifications simultaneously while exploring the available solution space.

In general, optimization techniques for EDA tools can be classified depending on whether the optimization approach uses equation-based models (low-fidelity, low-cost) or circuit simulations (high-fidelity, high-cost) [74] for the evaluation of the fitness function [75]. These two approaches display a trade-off between the accuracy of the model and the complexity of its evaluation. Optimization algorithms require a large number of evaluations proportional either to the number of iterations, the number of runs or the size of the population, depending on the type of algorithm. Therefore, inserting the circuit simulator (high-fidelity model) in the optimization loop could become impractical for large-scale problems due to the required computational cost. On the contrary, it has been shown that equation-based surrogate models can provide a sufficient functional level description with less computational complexity, at the cost of a reduction in accuracy [70]. Another advantage of these models is that after being created, they can be stored and reused [76].

Several techniques for generating equation-based models are summarized in Reference [73]. The equations can be generated manually by the experienced designer or automatically by symbolic tools [77]. As an example, Binkley et al. [78] present an all-equation optimization approach that uses the inversion level concept for the sizing of individual CMOS transistors.

Various approaches have proposed to leverage the low complexity of analytic equation-based models as with the accuracy of the circuit simulations. For instance, ASTRX/OBLX [79] uses asymptotic waveform evaluation to speed-up small-signal analysis in the simulator. Also, in Reference [77] the search space of the simulator-based optimization is reduced by using an optimization watchdog feature. Another approach combines simulation-based genetic optimization with multivariate regression techniques for an efficient space boundary exploration [80]. In Reference [76] Gaussian process-based regression models were used as surrogate models in Bayesian optimiza-

tion, although effective, the surrogate is completely-application specific, requiring new models to be generated for each circuit topology and technology. Moreover, in [81] gm/I_D method [21] and circuit equations are used for a smart reduction of the search space before optimization using the simulator. Still, all these approaches require a large number of computationally expensive circuit simulations within the optimization loops. A more complete surrogate model composed of circuit equations, a physics-based transistor model, and mathematical expressions of the process parameters extracted from curve fitting was reported in [82]. Although the surrogate proved to be an effective tool for circuit design, its accuracy was limited by the selection of the fitting method and it does not address process variations.

5.1.2 Our Approach and Contributions

In this work, we propose a low-cost yet high-accuracy surrogate model embedded in an optimization-based framework for automatic sizing and biasing of analog circuits. The main purpose of this surrogate implementation is to combine the low complexity of equation-based models of the performance of the circuits and the MOS transistor parameters with a highly accurate technology characterization. To achieve this goal, the proposed surrogate has three main components: Gaussian process regression (GPR) models of technology parameters across corners, a physics-based model of the MOSFET device, and a set of equations of the performance metrics of the circuit under design. Our proposed high-accuracy surrogate is modular and flexible to different circuit topologies and fabrication technologies. Each of the building blocks can be replaced according to the needs of the design. For example, moving from one technology process to another only requires characterizing the new technology and creating the GPR models of the device's parameters. The surrogate uses the updated GPR models for the evaluation of the advanced compact model (ACM) equations and the topology-specific metrics.

The surrogate allows evaluating the circuit performance metrics for a set of design variables (inputs). Therefore, it can be embedded in an optimization algorithm where the optimization variables correspond to the circuit design variables, and the objectives and constraints represent circuit performance metrics. Thus, the proposed approach is not limited to a particular optimization algo-

rithm, allowing our framework to utilize the most well-suited algorithm for the particular circuit being optimized. To prove the flexibility of our approach, the proposed surrogate is embedded in two state-of-the-art optimization algorithms: a gradient-based method and a heuristics-based algorithm. Also, the generality of the surrogate model to different technologies and circuit topologies is demonstrated for several test cases. As proof of concept, our technique is tested using three CMOS fabrication processes (TSMC 180 nm, IBM 130 nm, TSMC 65 nm), and three circuit examples: a second-order Butterworth active-RC filter, a capacitor-less low dropout (CL-LDO) voltage regulator, and a current-starved voltage controlled oscillator (CSVCO).

The main contributions of this work are:

1. A high-accuracy surrogate model for circuit optimization with low-computational effort compared with circuit simulations.
2. The use of Gaussian processes regression models for high accuracy prediction of device parameters across corners based on the technology characterization.
3. A flexible optimization framework easily configurable for different fabrication processes, circuit topologies, and optimization algorithms.

5.2 Multi-Objective Constrained Optimization for Automatic Circuit Design

In this section, we provide a background on the multi-objective optimization techniques that are used for the optimization framework demonstration of our proposed surrogate model.

5.2.1 Multi-Objective Optimization

Multi-objective optimization algorithms perform optimization of multiple specifications simultaneously [76]. Therefore, instead of a single solution, multi-objective algorithms provide the trade-off of the multiple objectives represented in the Pareto front [83]. The Pareto front is a set of solution points that can not be improved in one objective without getting degraded in another [80, 84]. The general formulation of a constrained optimization problem is as follows:

$$\text{minimize } f_1(\mathbf{x}), f_2(\mathbf{x}), f_n(\mathbf{x})$$

$$\text{s.t. } \mathbf{x}_{lb} \leq \mathbf{x} \leq \mathbf{x}_{ub}, \mathbf{e}(\mathbf{x}, \mathbf{p}) = 0, \text{ and } \mathbf{g}(\mathbf{x}, \mathbf{p}) < 0.$$

Therefore, the goal is to minimize the objective functions f_i where $i = 1, 2, \dots, n$, while being subject to the minimum \mathbf{x}_{lb} and maximum \mathbf{x}_{ub} boundaries of the design variables \mathbf{x} , equality constraints \mathbf{e} , and inequality constraints \mathbf{g} .

Optimization algorithms can be classified in two main categories according to the operations performed to find solutions: deterministic or gradient-based [85, 86], and stochastic or heuristics-based [79, 84, 87–89]. If a problem has an objective function that allows the calculation of gradients and a search space with a global minimum, using deterministic algorithms is usually the fastest way to find a solution. However, depending on the complexity of the solution space, gradient-based approaches can get stuck on local optimal solutions, resulting in poor exploration of the search space. When gradients are not available or when the solution space is non-convex or non-continuous, heuristic algorithms may be preferable.

5.2.1.1 Gradient-Based Optimization Algorithms

There are several deterministic algorithms for nonlinear constrained optimization [90]. Such algorithms are iterative and are designed to provide, at least, local optimum solutions. These algorithms use different approaches to include constraints like interior point methods, penalty functions and augmented Lagrangian methods [90].

Sequential quadratic programming (SQP) is a popular gradient-based optimization method for nonlinear constrained optimization. SQP closely mimics Newton's method as it generates search directions at each iteration by solving quadratic sub-problems [90]. SQP is particularly effective at handling problems with significant non-linearity in their constraints [90], which makes this algorithm well suited for our application, as many analog circuit specifications exhibit non-linear relationships with circuit design parameters. In fact, SQP has been successfully employed previously for circuit automatic design [82, 84]. The minimization function in the gradient-based optimization is built as the weighted sum of the optimization functions. It also includes the linearization function of the constraints using Lagrangian multipliers as described in Reference [90].

5.2.1.2 Evolutionary Optimization Algorithms

Evolutionary algorithms (EAs), also called genetic algorithms (GAs), are population-based iterative optimization processes [89]. In EAs each possible solution is encoded in a chromosome. Then, the execution begins with a randomly generated initial population of N Chromosomes. At each iteration or generation, the operators of crossover and mutation are used to evolve, that is, to generate a new population from the previous one based on fitness. The crossover operator combines the parent's population to generate offspring while the mutation operator introduces random modifications to certain individuals to increase the design space exploration [72]. Although there are several kinds of EAs, non-dominated genetic algorithms are popular for procuring diversity on the Pareto front [89].

Non-dominated sorting genetic algorithm II (NSGA-II) is an example of EA, proposed as an improved version of NSGA [91]. NSGA-II is computationally fast and it uses elitism and crowding distance calculations to maintain diversity in the non-dominated Pareto front [91]. Also, this algorithm uses a real coded GA as search engine, simulated binary crossover (SBX) and polynomial mutation [89,91]. These operators determine how the generated children will be different from their parents, and therefore, they define the space exploration. NSGA-II has been successfully used in the past for the optimization of circuits [72, 89, 92, 93].

5.2.2 Analog Circuit Design as an Optimization Problem

The formulation of the design of analog circuits as an optimization problem starts with the definition of the netlist of the circuit topology, and the characterization of the fabrication process available for this task [94]. Then, the optimization variables are defined from the design or tuning parameters, and the specifications or figures of merit (FoMs) are assigned to optimization objectives and constraints. Depending on the topology of the circuit and the target application, some specifications are better defined as optimization objectives and others as constraints. For example, a particular application may require an amplifier with a certain gain and bandwidth at the minimum power consumption possible. Thus, for the scope of this work, the design of an analog circuit is

defined as a constrained optimization problem; where the search space of all viable solutions is large and of unknown shape, making a brute force approach of trying all possible combinations non-practical.

5.3 Proposed Surrogate Model for Optimization-Based EDA Tools

A surrogate model or metamodel is a ‘model of a model’ used in EDA tools to replace the computationally expensive circuit simulation models and speed up optimization tasks [94]. The main characteristics of a surrogate model are accuracy, efficiency, robustness, simplicity, and transparency. In this work, we propose the creation of a low-cost surrogate model to be embedded in a modular and flexible optimization framework for the automatic design of analog/RF circuits.

5.3.1 General Optimization Architecture

The structure of a general and modular optimization framework for automatic IC design is shown in Fig. 5.1. Optimization is an iterative process that starts by generating the initial values of the variables at random. At each iteration, the optimization algorithm uses the proposed surrogate model for the evaluation of the objective function and the constraints. When the stop criteria are met, the optimization ends providing a set of solutions. Examples of stop criteria are the maximum number of evaluations, the maximum number of iterations, and the minimum tolerance on the objective function. Finally, the solutions are verified with circuit simulations under process corners, and only the ones that meet the specifications are selected to build the Pareto front.

The proposed surrogate model has three main components described from bottom to top:

1. The Gaussian process regression models of parameters of the process technology trained from characterization data.
2. The physics-based model of the parameters of the MOS transistor.
3. The circuit equations of the performance metrics of the circuit topology.

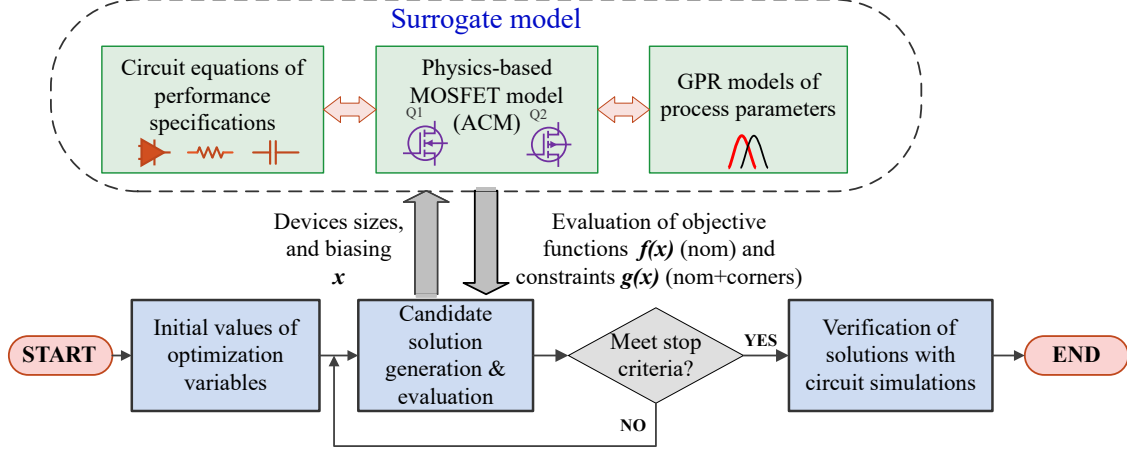


Figure 5.1: Proposed surrogate model inserted on a modular optimization framework for automatic IC design.

5.3.2 Advanced Compact MOSFET (ACM) Model

ACM is a model based on the device physics of the MOS transistors similar to other charge-based models like EKV and BSIM5 [28, 86, 95]. The ACM model provides higher accuracy compared to the square-law model traditionally used for circuit design. This model is built with a small set of equations valid for all the regions of operation of the MOSFET transistor, providing continuous modeling of the I/V characteristics of the device from deep saturation (strong inversion) to sub-threshold operation (moderate to weak inversion). A particular subset of the ACM equations are of interest for the formulation of our surrogate: the source trans-conductance g_m (Equation (5.1)), the inversion level i_f (Equation (5.2)), and the normalization current I_S (Equation (5.3)).

$$g_m = \frac{2I_D}{\phi_t n (1 + \sqrt{1 + i_f})} \quad (5.1)$$

$$i_f = \frac{I_D}{I_S} \quad (5.2)$$

$$I_S = \frac{W}{L} \mu n C'_{ox} \frac{\phi_t^2}{2}, \quad (5.3)$$

where I_D is the drain current, ϕ_t is the thermal voltage, (W/L) is the device aspect ratio, n is the

slope factor, μ is the mobility, and C'_{ox} is the oxide capacitance per unit area. The inversion level of the transistor is of particular importance as it provides information about the operating point of the transistor without computing the potential at the gate of the device.

Although the base ACM model does not include short-channel effects or the dependency of the mobility on the transversal field, the model can be expanded to account for these effects [28, 95]. However, many of the additional parameters needed to properly model these higher-order effects can be challenging to extract from simulation accurately. Instead, in our proposed model, we rely on non-parametric regression to account for higher-order effects and process variations.

5.3.3 Gaussian Process-Based Regression Models of the Process Characterization

The scaling trend on CMOS devices has resulted in not only smaller channel lengths and thinner gate-oxides, but also in the introduction of increasingly complex device structures and doping profiles [20, 28]. To accurately predict the behavior of modern MOSFETs, device models have become more elaborated, introducing hundreds of device parameters, and making them impractical to use outside of the simulators available in specialized CAD software tools. To build our surrogate model, the behavior of the key parameters of the device must be accurately formulated as functions of design variables that are at the control of the designer. For our surrogate model, the parameters required are the oxide capacitance per unit of area (C'_{ox}), the normalization current (I_S), the threshold voltage (V_{TH}), the saturation voltage (V_{DSAT}) and the early voltage (V_A), the latter of which is used to estimate the output conductance of the transistor ($g_{ds} = I_D/V_A$).

Once the parameters have been characterized against design variables, a prediction model must be created that can accurately approximate the data and can be used in conjunction with the remaining parts of our surrogate. Since the behavior of the parameters in question is complex and irregular, particularly with devices that exhibit significant short-channel and narrow-channel effects, using traditional equation-based curve-fitting techniques can become challenging and impractical. Therefore, it is attractive to consider non-parametric methods, that when properly configured, can approximate such complex irregular behaviors with significant ease and low error. In this work, Gaussian process regression is used to generate prediction models for the parameters in our surro-

gate.

5.3.3.1 Characterization of the Parameters of CMOS Transistors

In theory, the normalization current I_S (Equation (5.3)) depends only on the aspect ratio W/L of the transistor. However, in modern-day technologies, I_S also varies with the dimensions of the device, as they can have a direct impact on the device's mobility (μ). This makes the inversion level of a transistor a function of both W and L . Moreover, both the saturation voltage V_{DSAT} and the early voltage V_A exhibit a dependency on the inversion level of the transistor [95–97]. Finally, due to a combination of both short-channel and narrow-channel effects in modern CMOS processes, the threshold voltage (V_{TH}) is also a function of both W and L [28].

The transistor is configured in a particular setup to characterize each parameter of CMOS transistors. Fig. 5.2(a) shows the extraction setup for the oxide capacitance, where an AC signal is injected at a specific frequency while sweeping the DC bias at the gate, the capacitance is then computed from the impedance seen at the gate when the transistor is in the accumulation region. Fig. 5.2(b) shows the schematic for the characterization of V_{TH} , where W and L are varied. Also, the transistor is configured as shown in Fig. 5.2(c) to extract the normalization current for each value of W and L . The voltage at the source terminal controls the forward inversion level while the diode-connection configuration avoids the effects of the reverse inversion [95]. The normalization current is then extracted from the gm/I_D curve of the transistor based on Equation (5.1). Finally, the information of the normalization current is used to bias the transistor in specific forward inversion levels (Fig. 5.2(d) for which the early voltage and the saturation voltage are extracted. Fig. 5.3 shows some samples of the data extracted from the characterization of the TSMC 180 nm process. The sizes of the transistor are normalized with respect to the minimum dimensions of the technology, such that $L = KL * L_{min}$ and $(W/L) = KWL * (W_{min}/L_{min})$. Moreover, the characterization data is extracted for all process corners of interest and for each fabrication process to build the surrogate.

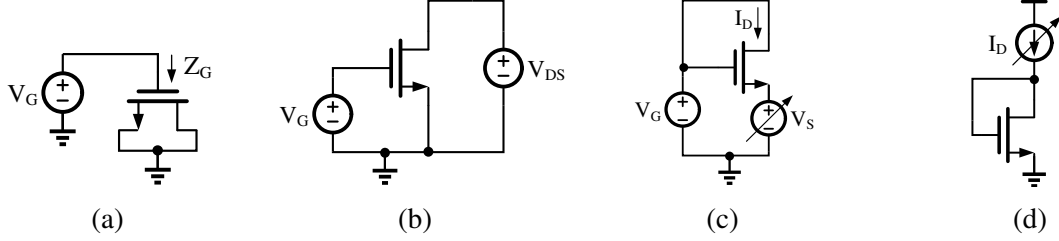


Figure 5.2: Schematics for device characterization of **(a)** Oxide Capacitance (C'_{ox}). **(b)** Threshold Voltage (V_{TH}). **(c)** Normalization current (I_S). **(d)** Saturation voltage (V_{DSAT}). early voltage (V_A).

5.3.3.2 Gaussian-Processes-Based Regression Models

Large data-sets of the device's parameters are acquired through the characterization process. The next step to complete our surrogate is to generate regression models capable of accurately estimating the required device's parameters over all the possible design values. One alternative of doing this is through polynomial regression, as done previously in Reference [82]. However, the parametric nature of this approach limits the precision achievable in modeling the complex combination of short-channel and narrow-channel effects on modern CMOS devices. Non-parametric methods, on the other hand, can achieve higher precision by allowing the number of parameters to increase as dictated by the sample size.

Gaussian Process Regression (GPR) is a type of non-parametric method used in classification and regression models for Bayesian optimization or machine learning [76]. GPR models have gained popularity due to its probabilistic nature, which can model fairly complicated functional forms with high accuracy [76, 89, 98]. Instead of specifying a particular function for regression, a Gaussian Process defines a probability distribution over a function-space defined by the data-set. When a GPR model is evaluated, an inference takes place, providing a function $y = f(x)$ within the function-space. The mean function $m(x)$ and covariance function $k(x_1, x_2)$ are what define a GPR model and determine the function-space generated from the data-set. While the mean function is commonly assumed constant and in many cases set to zero, the covariance function establishes the expected similarity or nearness between data points. By carefully choosing the

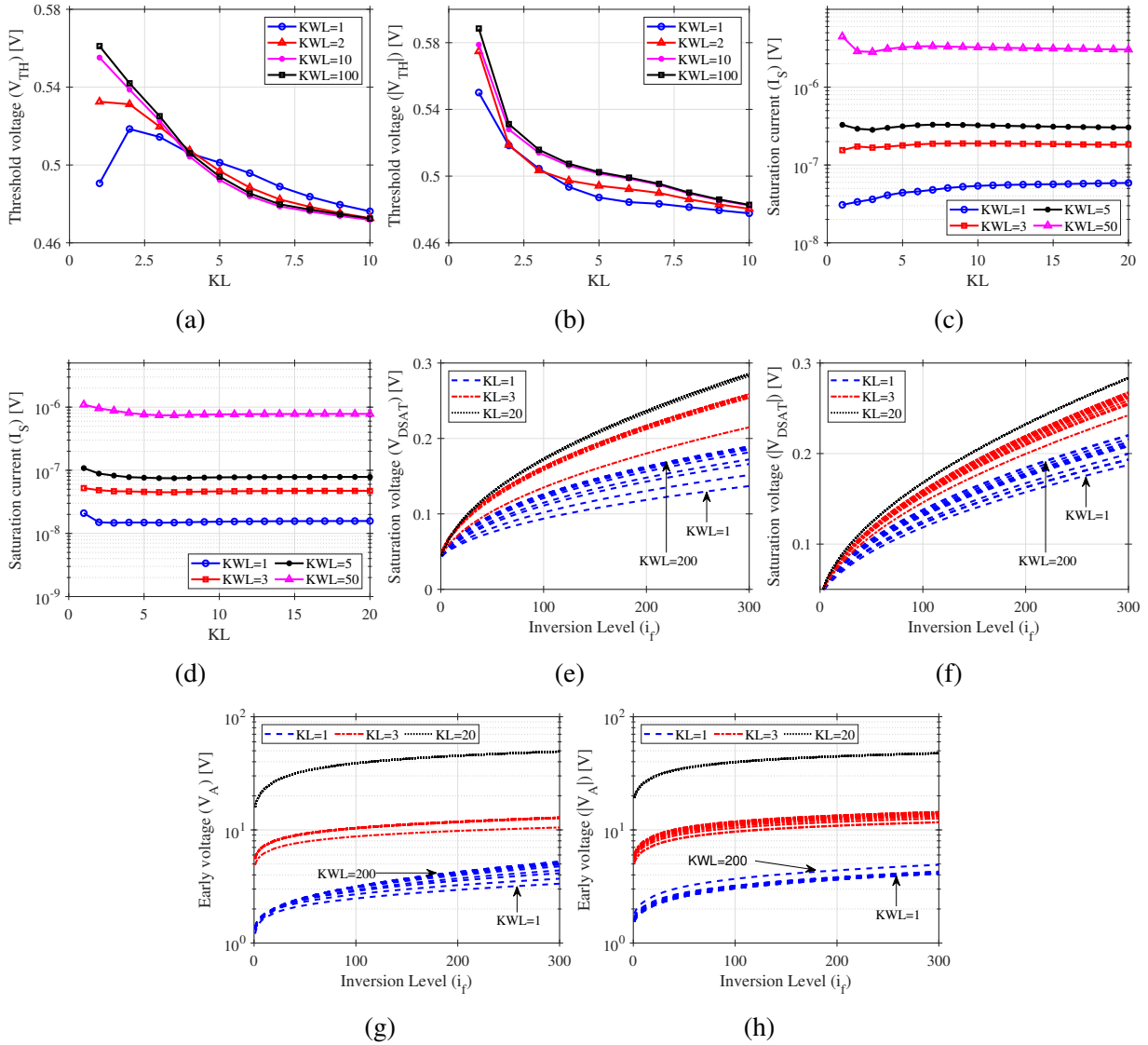


Figure 5.3: Sample of characterization data of a given CMOS technology in typical corner (TT) (a) NMOS: V_{TH} . (b) PMOS: V_{TH} . (c) NMOS: I_S . (d) PMOS: I_S . (e) NMOS: V_{DSAT} . (f) PMOS: $|V_{DSAT}|$. (g) NMOS: V_A . (h) PMOS: $|V_A|$.

covariance function one can embed the model with prior knowledge about the objective function to improve the predicting accuracy of the model [99]. A regression GP model is built from a training set $D(X, \mathbf{y})$ where $X = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N\}$ denotes the input vectors with N observations, and \mathbf{y} represents the corresponding outputs [99, 100].

The mathematical software MATLAB[®] has in-built functions for training GPR models, make

	I_S NMOS	I_S PMOS	V_{TH} NMOS	V_{TH} PMOS	V_{DSAT} NMOS	V_{DSAT} PMOS	V_A NMOS	V_A PMOS
KernelFunc	Exp.	ArdExp.	ArdExp.	ArdExp.	Exp.	Exp.	Exp.	ArdExp.
BasisFunc	Linear	None	Constant	None	Constant	Constant	None	None
FitMethod	Fic	Fic	Sr	Sr	Sd	Sd	Fic	Sr
ActiveSetMethod	Sgma	Sgma	Random	Random	Entropy	Random	Sgma	Random
PredictMethod	Exact	Exact	Exact	Exact	Exact	Exact	Exact	Exact
ResubLoss	1.37×10^{-5}	3.5×10^{-7}	6.26×10^{-8}	11.52×10^{-8}	5.18×10^{-7}	2.35×10^{-7}	6.29×10^{-6}	2.47×10^{-5}

Table 5.1: Optimal options of the function *fitrpg* for the training of the Gaussian Process (GP) model from the characterization of the process parameters in typical corner.

output predictions, calculate the regression loss, and even perform hyper-parameter optimization [101]. The function for the training of GP regression models *fitrpg* has a large number of parameters with several configuration options [101]. Therefore, in this work, a script selects the options of the *fitrpg* function that minimizes the prediction error. Such parameters are the kernel or the covariance function, the fitting method, the prediction method, and the active set selection method. For instance, Table 5.1 summarizes the best configuration for the training of GP models for the TSMC 180 nm technology. The GPR models of C'_{ox} , V_{TH} , V_A , I_S , and V_{DSAT} are created using the optimal parameters. The prediction function receives the GPR model and a set of design variables' values and returns the corresponding value of the parameter. The main advantage of using a regression model is that it can predict the parameter values not only for the characterization data but also for points in between.

5.3.4 Circuit Performance Equation-Based Model

So far, the components of the surrogate model represent the fabrication process and the MOS-FET device but are independent of the circuit topology. The third component corresponds to the equations- based model describing the performance metrics of an analog IC circuit. Such equations are derived using macromodel device representation and conventional circuit analysis techniques [20]. For the scope of this work, the equations were extracted manually. However, tools for automatic model generation using symbolic analysis [102], graph-based analysis [69] or signal

path analysis could be used as an alternative if needed.

In Section 5.4, we compile three analog circuits with equations of their performance metrics that will be used to demonstrate the usefulness of the surrogate model. The circuit equations use the GPR models to compute the small-signal parameters needed for each particular transistor, provided the device dimensions and current bias are known. C'_{ox} , V_{TH} and I_S are evaluated directly from device dimensions. I_S is used to compute the inversion level of the transistor (Equation (5.2)), which is necessary to calculate the device's transconductance (Equation (5.1)) and evaluate the GPR models for V_{DSAT} and V_A .

5.3.5 Process Variations-Aware Automatic Design

Although smaller technologies allow high device integration, they tend to suffer from higher process variations in the fabrication process that are particularly harmful to the performance of analog integrated circuits. Several techniques for process-aware design rely on either worst-case analysis or Monte Carlo analysis. The worst-case analysis use corner models to estimate performance variations. Despite being a simple and fast approach, it might be pessimistic and can lead to over-design. On the other hand, Monte-Carlo analysis is a statistical method that can approximate variations with low error. However, it requires hundreds of samples, making it computationally expensive for large designs with time-consuming simulations.

Since the main goal of this surrogate model is to reduce the computational complexity of models, we use worst-case analysis to ensure robust solutions. The GPR models created from the characterization of the technology includes information of the nominal and corner models. This allows the surrogate to estimate the robustness of the solutions to process variations.

5.4 Experimental Results

5.4.1 Error of the GPR-Based Surrogate Model

In this subsection, we evaluate the accuracy of the GPR model used in this work. As mentioned in Section 5.3 the process-dependent parameters of CMOS devices are characterized in terms of design variables and stored. Fig. 5.4 shows the comparison of the percentage error of the param-

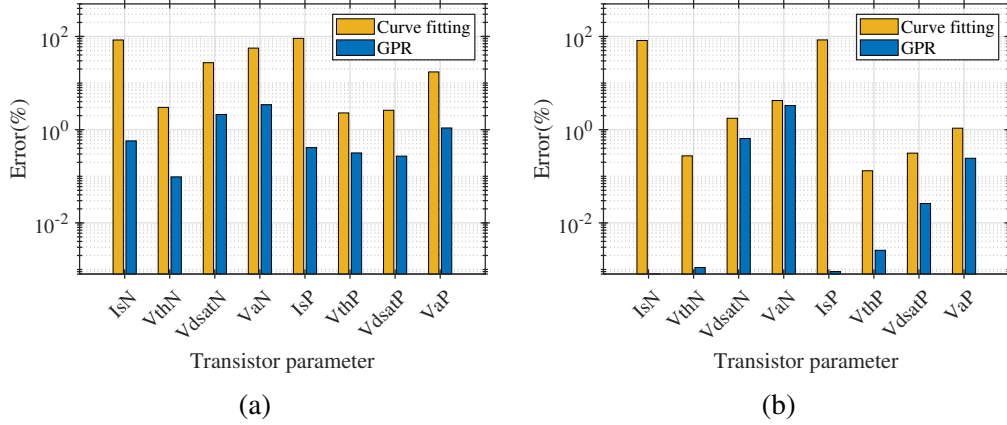


Figure 5.4: Comparison of the percentage error of prediction of CMOS parameters using models based on curve-fitting and Gaussian process regression (GPR) for sizes (a) $KL = 1$, $KWL = 4$. (b) $KL = 10$, $KWL = 50$. These models were built from the characterization data.

eter’s prediction using the curve-fitting-based models from Reference [82], and the GPR models developed in this work, with respect of circuit simulations (high-fidelity model). The accuracy of the models is evaluated for a small (Fig. 5.4(a)) and large (Fig. 5.4(b)) device size of the NMOS and PMOS devices on the TSMC 180 nm process. Note that the Y-axis has a logarithmic scale to visualize the error of the GPR that in all cases is lower than using a curve-fitting approach for prediction. Because the curve fitting model highly depends on the type of function and the number of coefficients used for the fitting, its accuracy is limited. As shown in Fig. 5.4, the prediction error of the GPR models is smaller than 4% for all the parameters. Note that the prediction error of the NMOS’ normalization current (I_{sN}) in Fig. 5.4(b) is lower than 0.0001%, and therefore, it is not visible on the plot.

5.4.2 Experimental Setup

The surrogate model is built for three technology nodes and three different circuit topologies. For each process, the model includes data of the nominal corner (TT) and the worst-case corners (SS, FF); all of them are considered simultaneously in the optimization. Then, the model is inserted on the iterative loop of state-of-the-art optimization algorithms, SQP and NSGA-II, to demonstrate the compatibility of our approach with different algorithms. The configuration of the main param-

Parameter of the Algorithm	SQP	NSGA-II
Algorithm implementation	fmincon: sqp [101]	NSGA2 toolbox [91]
Multi-start/Runs	10	10
Stop criteria	Function tolerance = 1×10^{-17}	Max. generations = 500
Other parameters	Max. Fun. evaluations = 8×10^5	Population size = 30
	Max. iterations = 2×10^6	Dist. index for crossover = 20
	Constraint tolerance = 1×10^{-6}	Dist. index for mutation = 20

Table 5.2: Parameters of the optimization algorithms used for this experiment sequential quadratic programming (SQP) and NSGA-II.

eters of the algorithms used for the experiments is summarized in Table 5.2. The stop criterion is probably the most relevant parameter. The execution of the SQP algorithm ends when the solver takes a step that is smaller than the function tolerance, and the constraints are satisfied within the constraint tolerance [101]. If the execution does not achieve the minimum function tolerance, the algorithm stops when it reaches the maximum number of iterations or the maximum number of function evaluations. On the other hand, the NSGA-II algorithm stops when it reaches the maximum number of generations. Additionally, each optimization algorithm runs 10 times for the same optimization problem, keeping only the best results. Another difference in the operation of the algorithms is the construction of the Pareto front. While NSGA-II builds the Pareto using in-built functions for ranking and crowding distance calculation, that is not the case of the SQP algorithm. The gradient-based optimization with nonlinear constraints minimizes the weighted sum of all the objective functions. Therefore, each point of the Pareto front is obtained by modifying the weights of the objectives. Finally, NSGA-II requires distribution indexes for the operators of crossover and mutation [91].

The surrogate is self-contained and easily adaptable to other optimization techniques like differential evolution [77, 81, 89], simulated annealing [79], particle swarm optimization [71], Bayesian optimization [76], and even learning-based techniques like neural networks or support vector machines [75]. The optimization framework is implemented in MATLAB[®] using in-built functions

combined with optimization toolboxes [91, 101]. Finally, the solutions are evaluated by circuit simulation (high-fidelity model) using the Cadence[®] Spectre[®] Simulator, and any solutions not found in compliance with the specifications, on all the corners, are discarded. All experiments were executed in a Linux workstation with an Intel Xeon CPU with frequency 2.3 GHz and 131 GB of RAM memory.

5.4.3 Active-RC Second Order Filter

Filters are key building blocks in signal processing, allowing the suppression or selection of specific frequency bands. The requirements of the filter type, order, bandwidth, and selectivity may change depending on the application. In this test case, the proposed framework provides designs of the Tow-Thomas 2nd-order Butterworth active-RC low-pass filter (LPF) as shown in Fig. 5.5(a). This filter topology is widely used because of its ease of tunability and relatively low sensitivity to component variation. The transfer function of the filter in standard second-order form is shown on Equation (5.4), where Q represents the selectivity (5.5) and ω_0 indicates the center frequency (5.6). These expressions do not take into account the effect of the frequency response of the amplifiers, which can lead to Q-enhancement and center-frequency errors [18]. Instead, these effects are taken into consideration on the bandwidth requirements of the amplifiers.

$$H_{LP}(s) = \frac{-\omega_0^2}{s^2 + s(\omega_0/Q) + \omega_0^2} \quad (5.4)$$

$$Q = R_Q/R \quad (5.5)$$

$$\omega_0 = 1/RC. \quad (5.6)$$

The two-stage internally-compensated topology shown in Fig. 5.5(b) is used on the implementation of the amplifiers in the filter (A_1 , A_2 , and A_3). This amplifier consists of two gain stages: 1st-stage a PMOS differential pair with NMOS active-load, and 2nd-stage an NMOS common-source amplifier. The miller-capacitor (C_C) performs stability compensation via pole-splitting.

From the transfer function of the filter, it can be straightforward to obtain a set of capacitor and

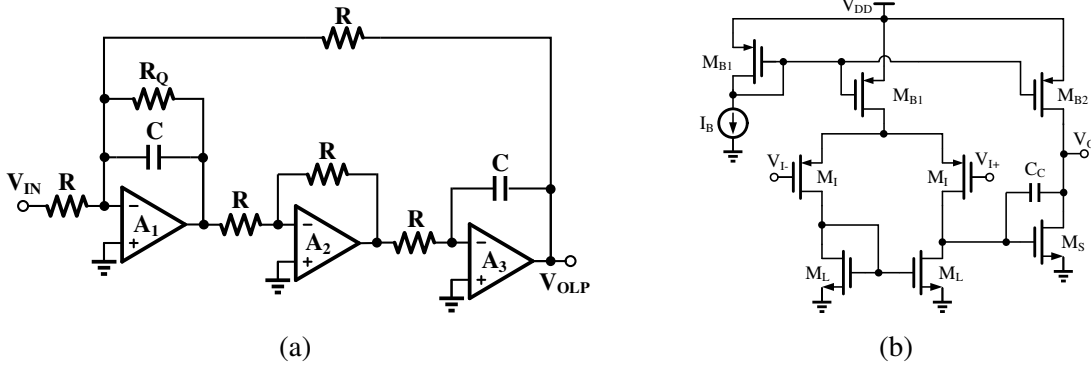


Figure 5.5: Active-RC second order filter under design (a) circuit topology. (b) Transistor level schematic of the second order internally compensated amplifier.

resistor values to meet certain specifications of ω_o and Q . However, on an integrated solution, the selection of these values can have a significant impact on the power and area cost of the circuit. Depending on the technology, the density of the available passives will change, and the sizing of the resistors and capacitors to implement a specific time-constant will have direct consequences on the loading and stability requirements of the amplifiers as well as the noise performance of the filter. To successfully design a power and area efficient filter, one must optimize the resistor and capacitors in conjunction with the sizing of transistors in the amplifiers to guarantee a certain desired performance. Within this context, the optimization problem is defined as follows:

$$\begin{aligned} & \text{Minimize } (\alpha_a \cdot \text{Area}(\mathbf{x}, \mathbf{p}) + \alpha_p \cdot \text{Power}(\mathbf{x}, \mathbf{p})) \\ & \mathbf{x} = [KI_B, KL, KWL_{B1}, KWL_I, KWL_L, KWL_S, KWL_{B2}, KC_C, KC] \\ & \text{subject to } \mathbf{x}_{lb} \leq \mathbf{x} \leq \mathbf{x}_{ub}, \text{ and } \mathbf{g}(\mathbf{x}, \mathbf{p}) \leq 0 \\ & \mathbf{g}(\mathbf{x}, \mathbf{p}) = [A0, UGF, PM, OS, ICMR, SR, VN], \end{aligned}$$

where \mathbf{x} is the set of normalized optimization variables, which includes: biasing conditions ($I_B = KI_B \times 1\mu A$), transistor's length ($L = KL \times L_{min}$), aspect ratios ($W/L_i = KWL_i$) and capacitor values ($C_C = KC_C \times 100fF$, $C = KC \times 1pF$), with L_{min} as the minimum channel length of the technology. The set of constant parameters is $\mathbf{p} = [Q, FC]$, which is meant to represent system-level requirements outside of the control of the design, where $FC = 2\pi\omega_0$ for $Q = 0.707$ (Butterworth filter). The resistor values of the filter are not considered design variables, as they

are automatically derived by specifying the capacitor (C) value and the filter specifications. In addition, the non-equality constraints \mathbf{g} considered are the amplifier's DC gain (A_0), unity-gain frequency (UGF), phase margin (PM), input common-mode range ($ICMR$), output swing (OS), slew rate (SR) and input-referred spot noise (VIN). Although linearity is not explicitly considered as a constraint in the surrogate, it can be adjusted as needed by changing the gain specification of the amplifiers, taking advantage of the linearization properties of negative feedback [103]. For other topologies that can not rely on negative feedback to improve linearity, a specific linearity specification could be included in the surrogate.

5.4.3.1 Surrogate of the Filter's Performance Metrics

To complete the surrogate model for the circuit, a system of equations was formulated using known approximations of the performance metrics of the amplifier. The amplifier's DC gain is $A_0 = g_{m_I} g_{m_S} R_1 R_2$, where $R_1 = 1/(g_{dS_L} + g_{dS_I})$ and $R_2 = 1/(g_{dS_S} + g_{dS_{B2}} + 1/R_L)$ are the output impedances associated with each stage of the amplifier, with R_L as the load seen by the amplifier due to the resistors in the filter. The unity gain frequency is shown on Equation (5.7), where ω_{p1} and ω_{p2} are the poles at the output of each gain stage in the amplifier, given by Equations (5.8) and (5.9) respectively. The phase margin is given by $PM \approx 180^\circ - \tan^{-1}(UGF/\omega_{p1}) - \tan^{-1}(UGF/\omega_{p2}) - \tan^{-1}(UGF/\omega_z)$, with $\omega_z = g_{m_s}/C_C$ as the feed-forward zero created by the miller capacitor. The input common-mode range is computed as: $ICMR = V_{DD} - |V_{DSAT_B}| - |V_{DSAT_I}| - V_{TH_L} - V_{DSAT_L}$, and the output swing as: $OS = V_{DD} - |V_{DSAT_{B2}}| - V_{DSAT_{S2}}$, where V_{DD} represents the nominal supply-voltage for the given technology. The slew-rate is approximated to be the minimum of either $SR_1 = I_B/C_C$ or $SR_2 = I_{B2}/(C_C + C)$, where $I_{B2} = (KWL_{B2}/KWL_{B1}) \times I_B$ is the current bias of the amplifier's second stage. Noise equations to compute input-referred spot noise from both white and flicker noise sources are also included [104].

$$UGF \approx \left(\frac{1}{8\pi} \right) \sqrt{\sqrt{4A_0^2 \omega_{p1}^2 \omega_{p2}^2 - 2\omega_{p1}^2 \omega_{p2}^2 + \omega_{p1}^4 + \omega_{p2}^4} - \omega_{p1}^2 - \omega_{p2}^2} \quad (5.7)$$

$$\omega_{p1} = \frac{1}{R_1 C_C + R_2 (C_C + C_L) + g_{m_s} R_1 R_2 C_C} \quad (5.8)$$

$$\omega_{p2} = \frac{R_1 C_C + R_2 (C_C + C_L) + g_{m_s} R_1 R_2 C_C}{R_1 R_2 C_C C_L} \quad (5.9)$$

Finally, for the objective functions: power consumption is given by $PWR = 3 \times (2I_B + I_{B2})(V_{DD})$ and the area cost (A) of the design is calculated as shown by Equation (5.10). Where C_{PA} is the capacitance per unit of area, R_S is the sheet resistance and W_R is the resistor width, all of which are process-dependent parameters.

$$A = 3 \times (L)^2 ((2KWL_{B1} + KWL_{B2}/KWL_{B1})KWL_{B1} + 2KWL_I + 2KWL_L + KWL_S) + 2 \times ((C_C + C)/C_{PA}) + (5 + Q) (R/R_S) (W_R)^2. \quad (5.10)$$

The surrogate also includes equations to make sure that all transistors are properly biased outside the triode region ($V_{DSAT} < V_{DS}$) based on the common-mode signal level at the input (V_{CM}) and above weak inversion ($i_f > 1$) [28,97].

5.4.3.2 Results of Filter's Automatic Design

As described in Section 5.3, once the optimizer generates a set of solutions, they are verified through circuit simulation in the high-fidelity model to obtain the final performance metrics. Fig. 5.6 shows the design trade-off between minimization objectives (power, area) for the three technology processes (TSMC 180 nm, IBM 130 nm, TSMC 65 nm) and both optimization algorithms (SQP, NSGA-II) with a specification of $FC = 100$ KHz. From the Pareto we can take notice that the solutions generated in the 180 nm process can provide lower area costs than in the 130 nm, this is due to the fact that the sheet resistance (R_S) of the selected resistor in the 130 nm process was smaller, resulting in a larger area consumption in comparison with the solutions in the 180 nm process. Similarly, the higher density of passives on the 65 nm process allows for much lower area metrics compared to the other technologies. However, the power consumption is also higher, likely due to the reduced intrinsic gain of the smaller process. Also, worth mentioning is

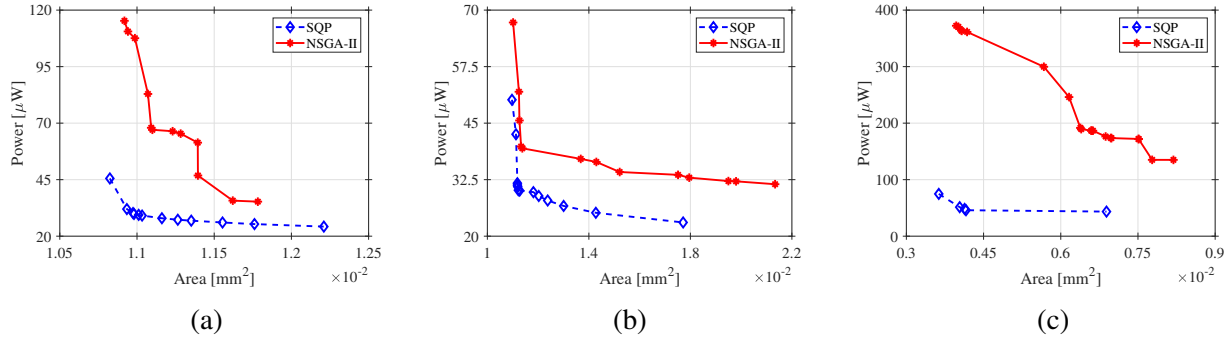


Figure 5.6: Pareto front of the optimization of filter ($FC = 100KHz$) using SQP and NSGA-II optimization algorithms in (a) TSMC 180 nm CMOS process. (b) IBM 130 nm CMOS process. (c) TSMC 65 nm CMOS process.

that even though the NSGA-II optimizer generated a larger set of solutions, only a small subset of them were competitive enough to be included in the Pareto comparison with SQP.

Fig. 5.7 shows a set of box plots comparing the design variables (x) in the Pareto's solutions, allowing us to gain insight in which design variables are more tightly restricted for an optimal design, an overlay box with each variable allowed range is included for reference. For example, in all cases the biasing current (I_B) is kept low, even going so far as to reach the lower boundary set for the variable in some cases, which relates to the direct impact its value has on the power consumption objective.

Table 5.3 shows a summary of constraints for one set of Pareto solutions (SQP optimization, 180 nm technology) evaluated through process variations for different desired filter cut-off frequencies (FC), where the UGF specification was set to be at least 10 times larger than the FC parameter to minimize the effect of Q-enhancement and center-frequency errors. The comparison on the table reveals how different constraints become active or inactive as the bandwidth requirements change, where an active constrain is one that is satisfied by a narrower margin, indicating it might be a bottleneck for the design. For example, as the frequency requirements increase, the A_0 specification is more narrowly satisfied while the SR constrain is relaxed, which is a direct consequence of the increased biasing required to reach a higher UGF specification. The UGF requirement is also more tightly met at higher FC specifications, which is to be expected since it

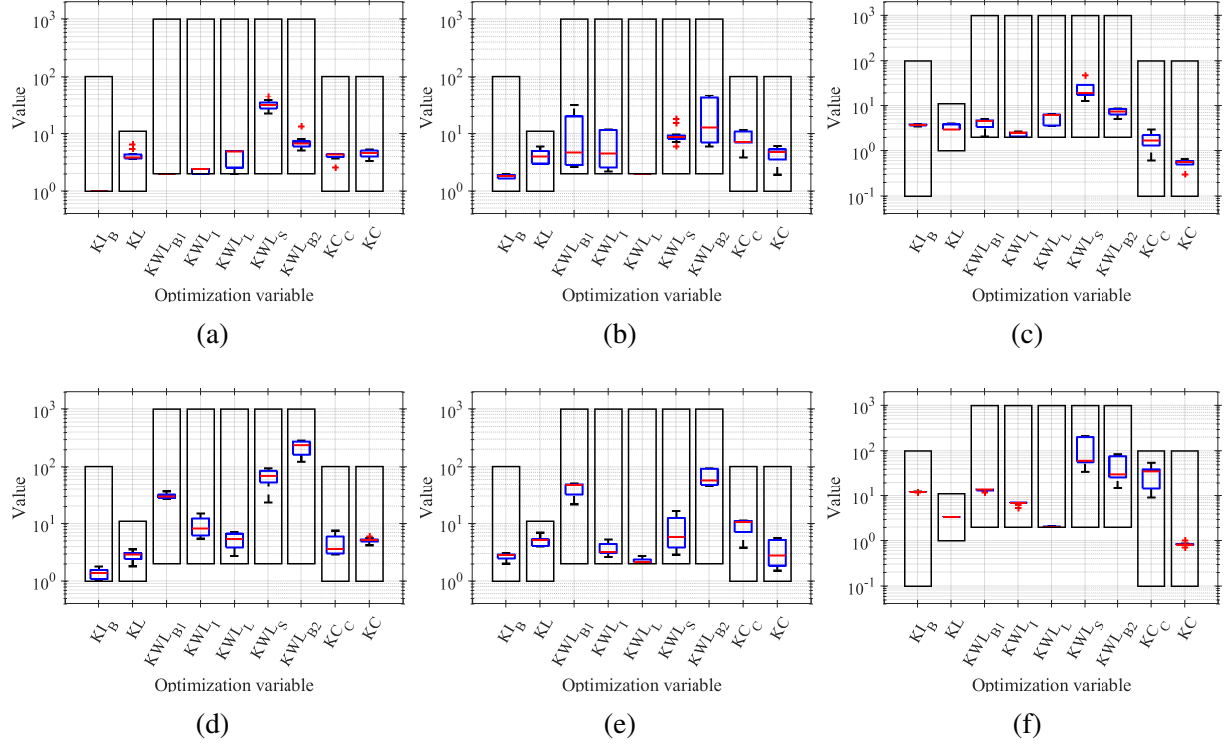


Figure 5.7: Values of the optimization variables from solutions in the Pareto front for the filter optimization ($FC = 100$ KHz) obtained with: (a) SQP-180 nm, (b) SQP-130 nm, (c) SQP-65 nm, (d) NSGA-II-180 nm, (e) NSGA-II-130 nm, and (f) NSGA-II-65 nm.

directly conflicts with the minimization of power consumption in the optimizer. All solutions are viable and comply with the desired specifications across corners. In all cases, the cut-off frequency specification is satisfied within a worst-case 8% error, which is consistent with the UGF requirement set for the optimizer. Thus, for the case of the 2nd-order Butterworth LPF, the proposed surrogate model design framework was demonstrated to generate viable solutions across process variation, technologies, frequency requirements, and optimization algorithms, all while providing insightful trade-offs in the constraints and objectives of the design.

5.4.4 Capacitor-Less Low-Dropout (CL-LDO) Voltage Regulator

In the second case example, we present the optimization of a capacitor-less low-dropout (CL-LDO) voltage regulator with a PMOS pass transistor and a single-stage error amplifier as shown in Fig. 5.8 [105]. This circuit provides a clean and stable output voltage V_O set by the reference

Parameter	Spec	FC = 100KHz			FC = 1MHz			FC = 10MHz		
		Min	Mean	Max	Min	Mean	Max	Min	Mean	Max
A_0 [dB]	>40	55.54	57.59	61.35	52.21	55.69	58.58	42.47	45.65	49.24
UGF/FC	>10	19.28	20.98	31.87	13.59	15.03	16.26	11.71	12.27	13.34
PM [°]	>40	42.17	45.49	48.16	42.31	45.75	49.21	44.62	46.91	48.68
$ICMR$ [V]	>0.6	1.14	1.18	1.23	1.11	1.20	1.30	0.76	0.85	0.95
OS [V]	>1	1.29	1.36	1.65	1.25	1.32	1.37	1.39	1.43	1.47
SR [V/ μ s]	>0.4	2.13	2.38	3.63	10.40	12.24	15.06	85.84	91.49	103.63
		0.56	0.66	1.15	4.73	5.53	6.41	56.02	57.15	60.07
V_N [μ V/ $\sqrt{\text{Hz}}$]	<2	0.89	1.04	1.22	1.22	1.34	1.41	1.05	1.11	1.18

Table 5.3: Performance metrics of the Pareto front solutions obtained through SQP optimization of the 2nd-order active-RC low-pass filter optimization in a TSMC 180 nm process (Noise (V_N) reported at 10 kHz).

voltage V_{REF} and the voltage divider implemented by R_{F1} and R_{F2} . The dropout voltage $V_{DO} = V_{IN} - V_O$ is the minimum difference between input and output voltage to maintain regulation. The pass transistor M_4 is sized such that its output impedance $r_{ds} = 1/g_{ds}$ complies with the load current and the voltage across V_{DO} . This topology uses a type-A error amplifier due to its inherent good power supply rejection (PSR) [105]. Most importantly, this LDO topology does not require an external large capacitor at the output for stability purposes. Instead, the circuit is internally compensated by an integrated compensation capacitor C_C that adds to the already large parasitic capacitors of the pass transistor. The CL-LDO is a particularly good fit for optimization-aided design, as the circuit requirements must be satisfied across the load range, particularly balancing the trade-off between good PSR and sufficient stability.

For the purpose of this example, the LDO shown will be designed to meet a set of specifications of power supply rejection (PSR) at different frequencies, phase margin, input common-mode range and output swing of the error amplifier. With the objective of minimizing the quiescent power $P_{quiescent}[W]$ and the power supply rejection at DC ($PSR@DC[V/V]$). The total active area is

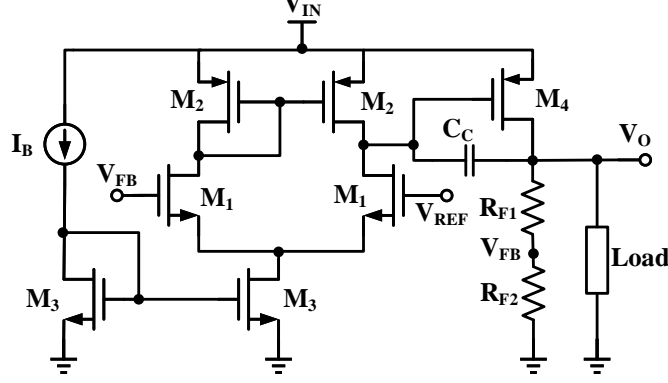


Figure 5.8: Circuit schematic of the capacitor-less low dropout (LDO) with type-A single stage error amplifier internal frequency compensation provided by C_C .

not considered for optimization because it is dominated by the area of the pass transistor, which is defined by the range of the load current. The optimization problem is defined as follows:

$$\text{minimize } \mathbf{P}_{\text{quiescent}}(\mathbf{x}, \mathbf{p}), \text{PSR@DC}(\mathbf{x}, \mathbf{p})$$

$$\text{subject to } \mathbf{x}_{\text{lb}} \leq \mathbf{x} \leq \mathbf{x}_{\text{ub}}, \text{ and } \mathbf{g}(\mathbf{x}, \mathbf{p}) \leq 0$$

$$\mathbf{x} = [KI_B, KL_A, KWL_3, KWL_1, KWL_2, KWL_4, V_{REF}, KR_B, KC_C, KL_4],$$

where the set of normalized design variables \mathbf{x} , includes the biasing conditions of the amplifier ($I_B = KI_B \times 1\mu A$), transistors lengths ($L_i = KL_i \times L_{min}$) and aspect ratios ($W/L = KWL_i$), reference voltage (V_{REF}), the output sampling resistor (KR_B) and compensation capacitor ($KC_C = KC_C \times 1pF$). Note that $KR_B \times 1k\Omega = R_{F1} + R_{F2}$.

5.4.4.1 Surrogate of the LDO's Performance Metrics

Small signal analysis techniques are used to build the component of the surrogate related to the performance metrics of the circuit under design.

Fig. 5.9 shows the small-signal macromodel for calculating the PSR of the LDO. The transfer function of the macromodel ($V_O(s)/V_{IN}(s)$) was obtained by using circuit analysis techniques like modified nodal analysis (MNA). Similarly, the macromodel for obtaining the expression of the loop gain and the phase margin is shown in Fig. 5.10.

In order to simplify the analysis let us define the auxiliary variables: $R_1 = 1/(r_{ds1} || r_{ds2})$,

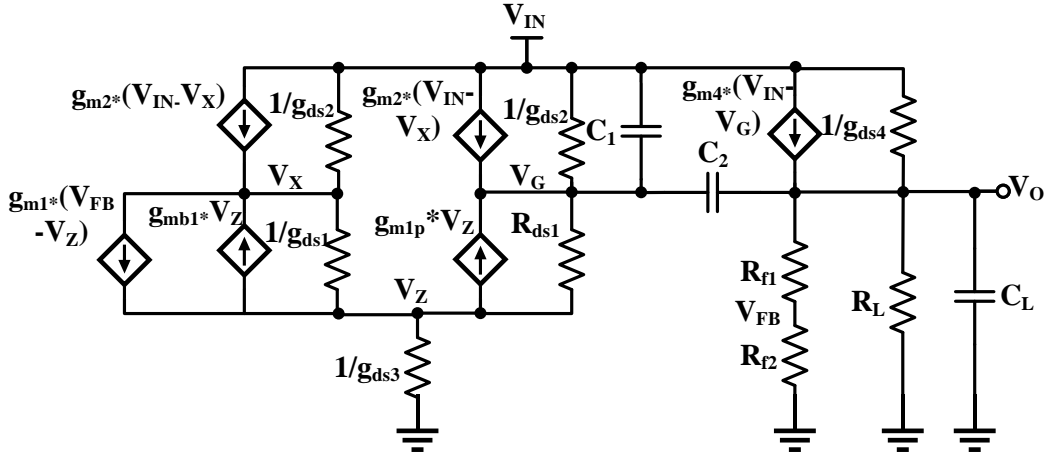


Figure 5.9: Small-signal macromodel of the LDO for the calculation of the power supply rejection (PSR).

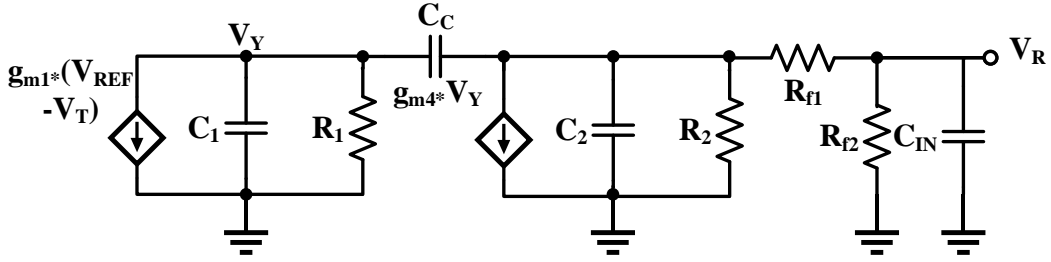


Figure 5.10: Small-signal macromodel of the LDO for the calculation of the phase margin.

$R_2 = r_{ds4} || V_O/I_L$, $C_1 = C_{gs4} + C_{gb4} + C_{EA}$, $C_2 = C_{gd4} + C_C$ and $C_{IN} = C_{gs1}$. The symbolic expressions of three poles (5.11)–(5.13) and the zero $f_z = g_{mP}/(2\pi C_2)$ were extracted from the loop transfer function.

The phase margin is evaluated as $PM = 180^\circ - \tan^{-1}(UGF/f_{p1}) - \tan^{-1}(UGF/f_{p2}) - \tan^{-1}(UGF/f_{p3}) - \tan^{-1}(UGF/f_z)$, where UGF is the unity gain frequency. While these expressions are impractical for hand calculations, they are compatible with the proposed surrogate model enabling high accuracy. Finally, biasing conditions are included in the constraints to ensure that the transistors do not operate in the linear region $V_{DSAT} < V_{DS}$. Also, the minimum inversion level is controlled such that the transistors will operate in moderate to strong inversion

levels $i_f > 1$. Thus, the equation $V_{DSAT_4} < (V_{IN} - V_O)$ is included in the set of constraints. Biasing equations of the error amplifier are also included: $(V_{TH_1} + V_{DSAT_1} + V_{DSAT_3}) < V_{REF}$ and $(V_{REF} < (V_{IN} + V_{TH_1} - |V_{TH_2}| - |V_{DSAT_2}|)$.

$$f_{p1} \approx \frac{R_2 + R_{f1} + R_{f2}}{(2*\pi)(C_L R_2 R_{f1} + C_{IN} R_2 R_{f2} + C_L R_2 R_{f2} + C_{IN} R_{f1} R_{f2} + C_2 R_2 (R_{f1} + R_{f2}) + C_1 R_1 (R_2 + R_{f1} + R_{f2}) + C_2 R_1 (R_2 + R_{f1} + R_{f2} + g_{mP} R_2 (R_{f1} + R_{f2})))} \quad (5.11)$$

$$f_{p2} \approx \frac{C_L R_2 R_{f1} + C_{IN} R_2 R_{f2} + C_L R_2 R_{f2} + C_{IN} R_{f1} R_{f2} + C_2 R_2 (R_{f1} + R_{f2}) + C_1 R_1 (R_2 + R_{f1} + R_{f2}) + C_2 R_1 (R_2 + R_{f1} + R_{f2} + g_{mP} R_2 (R_{f1} + R_{f2}))}{(2*\pi)(C_{IN} C_L R_2 R_{f1} R_{f2} + C_2 C_{IN} (R_2 R_{f1} + R_1 (R_2 + R_{f1} + g_{mP} R_2 R_{f1})) R_{f2} + C_2 C_L R_1 R_2 (R_{f1} + R_{f2}) + C_1 R_1 (C_{IN} (R_2 + R_{f1}) R_{f2} + C_2 R_2 (R_{f1} + R_{f2}) + C_L R_2 (R_{f1} + R_{f2})))} \quad (5.12)$$

$$f_{p3} \approx \frac{C_{IN} C_L R_2 R_{f1} R_{f2} + C_2 C_{IN} (R_2 R_{f1} + R_1 (R_2 + R_{f1} + g_{mP} R_2 R_{f1})) R_{f2} + C_2 C_L R_1 R_2 (R_{f1} + R_{f2}) + C_1 R_1 (C_{IN} (R_2 + R_{f1}) R_{f2} + C_2 R_2 (R_{f1} + R_{f2}) + C_L R_2 (R_{f1} + R_{f2}))}{(2*\pi) C_{IN} (C_2 C_L + C_1 (C_2 + C_L)) R_1 R_2 R_{f1} R_{f2}} \quad (5.13)$$

5.4.4.2 Results of the LDO's Automatic Design

As with the previous test case, the CL-LDO circuit was designed using three different fabrication processes (TSMC 180 nm, IBM 130 nm, and TSMC 65 nm), and two optimization algorithms: SQP and NSGA-II, for a total of 6 experiments. The circuit parameters and the specification constraints are summarized on Table 5.4.

The resulting set of solutions build the Pareto fronts in Fig. 5.11. The trade-off between power consumption and PSR is evident in the Pareto fronts. The PSR of this LDO topology is related to the gain of the error amplifier. Since higher gain requires higher power consumption, the two objectives can not be improved simultaneously. Comparing the Pareto fronts obtained using both optimization algorithms we do not observe a clear dominance of one algorithm on finding the best solutions for all designs. Regardless, the surrogate model fits well both optimization algorithms and enables them to find different sets of valid solutions. The large set of solutions found by the

Process	V_{IN}	V_O	C_L	$I_{L,min}$	$I_{L,max}$	PSR@1 kHz	PSR@10 kHz	PSR@100 kHz	PM
TSMC 180 nm	1.8 V	1.6 V	100 pF	533.3 μ A	5.3 mA	<-50 dB	<-45 dB	<-25 dB	>45°
IBM 130 nm	1.2 V	1 V	100 pF	333.3 μ A	3.3 mA	<-40 dB	<-40 dB	<-25 dB	>45°
TSMC 65 nm	1.2 V	1 V	100 pF	333.3 μ A	3.3 mA	<-40 dB	<-40 dB	<-25 dB	>45°

Table 5.4: Circuit parameters and specification constraints for the optimization of the LDO in two different CMOS processes.

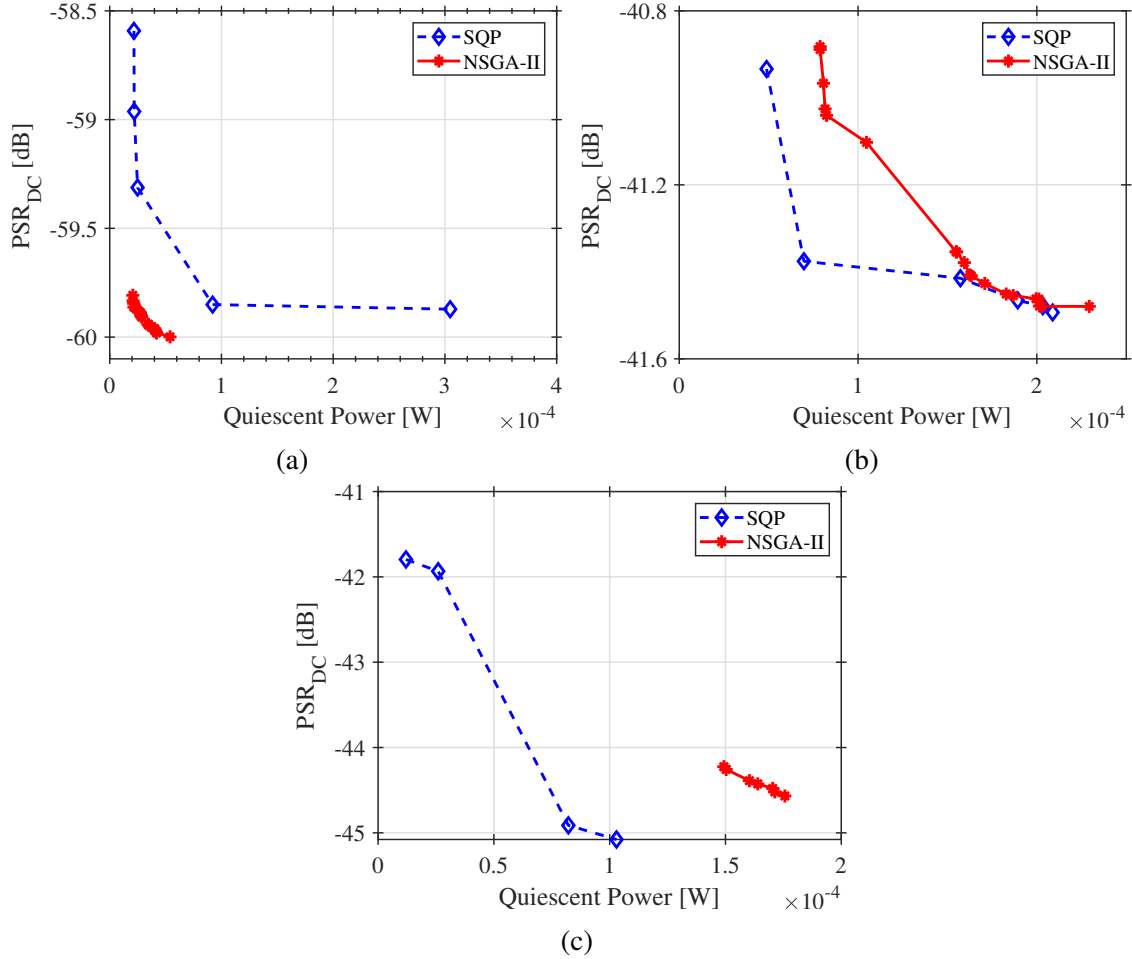


Figure 5.11: Pareto front of the optimization of the LDO circuit using SQP and NSGA-II optimization algorithms and (a) TSMC 180 nm. (b) IBM 130 nm. (c) TSMC 65 nm CMOS process.

algorithms and verified with circuit simulations prove the high-accuracy of the surrogate model including process variations.

The variable with the smallest distribution is V_{REF} that consistently tends to the same value across all solutions. We also observe a clear tendency for the algorithms to maximize KL_A (since its maximum value is 10) to increase the gain of the error amplifier and therefore increase the IPSR_I. The aspect ratio of the pass transistor KWL_A in most of the cases tends to the maximum allowed value. The other variables have a larger distribution of values over their ranges and their tendency is similar for both algorithms. Moreover, the range of the values obtained using NSGA-II is narrower than the one obtained with SQP.

Box plots of the values of each optimization variable were created to determine the diversity of the solutions in the Pareto front as shown in Fig. 5.12. For comparison purposes, Fig. 5.12 also shows the whole range of each design variable.

To verify the constraints, we extracted the min, mean and max values of the specifications estimated through circuit simulations. The results of the optimization case SQP/130 nm are reported in Table 5.5, while the ones of the case NSGA-II/130 nm are presented in Table 5.6. Within these constraints evaluated for the nominal corner only, the PSR@10kHz and the PSR@100kHz appears to be active constraints, while other constraints like phase margin seem more relaxed (non-active). The fact that the designs display over-design of the phase margin specification has several possible explanations. First, that despite it is not an active constraint in the nominal corner TT, it becomes one once it is evaluated for the case of the worst-corners. Second, that there is some discrepancy in the surrogate model, particularly on the estimation of the capacitances and resistances associated with the dominant and output nodes. Third, that given the range or search space of each variable, there is no such solution that could reduce the phase margin while still satisfying the PSR constraints.

Metric	Spec.	Low Load			High Load		
		Min	Mean	Max	Min	Mean	Max
PSR@1kHz [dB]	>40	40.26	48.81	54.40	48.62	54.08	63.18
PSR@10kHz [dB]	>40	40.04	45.58	50.71	44.77	48.67	54.41
PSR@100kHz [dB]	>25	27.87	30.99	34.51	27.02	30.74	35.29
Phase Margin [°]	>45	49.33	59.91	68.62	73.40	81.07	85.64

Table 5.5: Performance metrics of the Pareto front solutions obtained through SQP optimization of the CL-LDO in a 130 nm process measured with simulations.

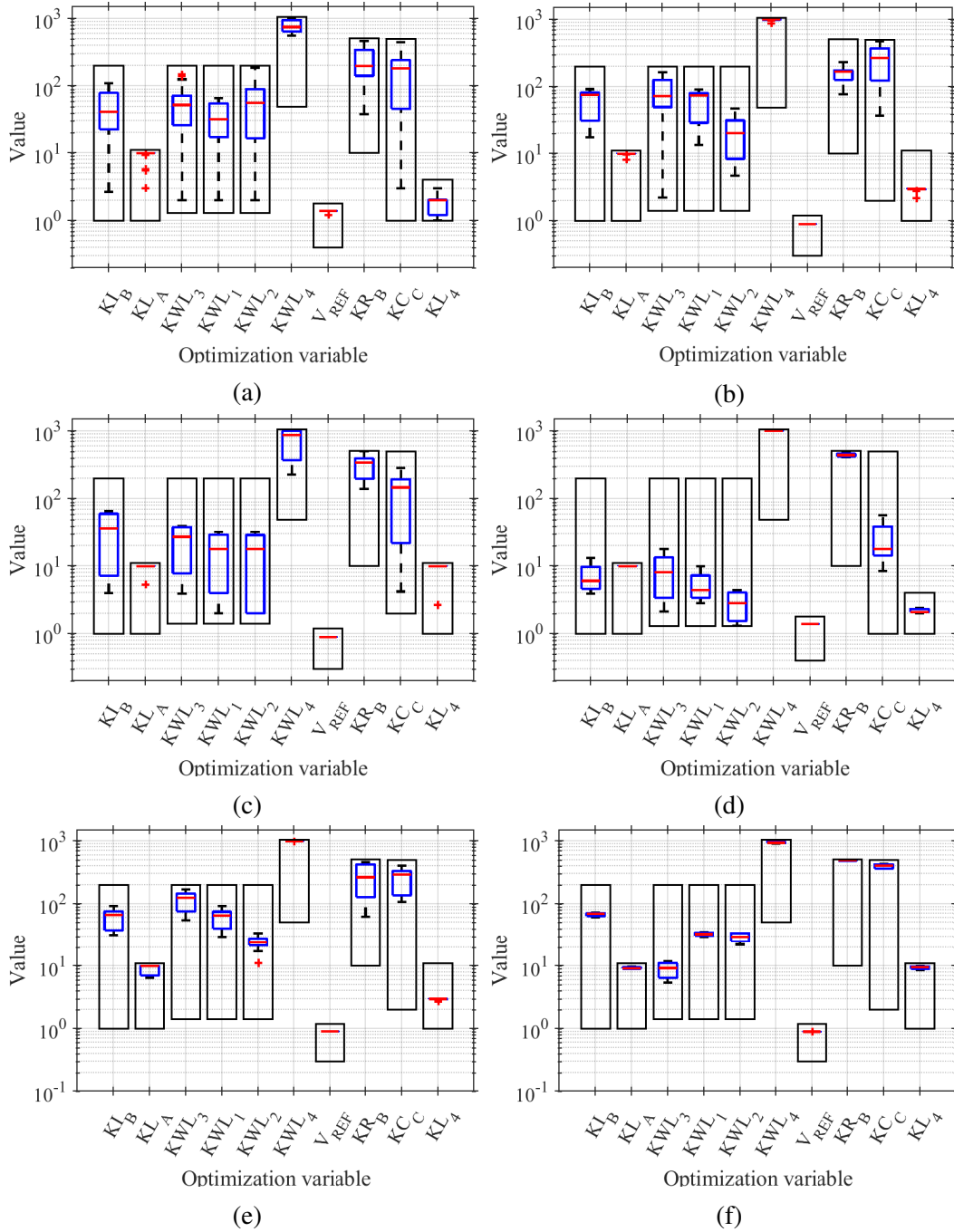


Figure 5.12: Values of the optimization variables of the solutions in the Pareto front obtained with: (a) SQP-180 nm, (b) SQP-130 nm, (c) SQP-65 nm, (d) NSGA-II-180 nm, (e) NSGA-II-130 nm and (f) NSGA-II-65 nm.

Metric	Spec.	Low Load			High Load		
		Min	Mean	Max	Min	Mean	Max
PSR@1kHz [dB]	>40	40.37	48.97	59.72	45.78	56.67	74.05
PSR@10kHz [dB]	>40	40.19	47.12	54.45	45.31	50.60	56.27
PSR@100kHz [dB]	>25	31.54	33.75	36.49	31.71	33.93	37.06
Phase margin [°]	>45	50.39	53.32	55.99	70.85	73.78	77.42

Table 5.6: Performance metrics of the Pareto front solutions obtained through NSGA-II optimization of the CL-LDO in the IBM 130 nm process measured with simulations.

5.4.5 Current-Starved Voltage Controlled Oscillator (CSVCO)

Oscillators are a fundamental block in transceivers and Phase-Locked-Loops (PLLs) [51]. One of the most conventional implementations of an oscillator is the ring oscillator, which consists of an odd number of inverter cells arranged in a feedback loop. Current limiting transistors are added to control the frequency of oscillation, to the top and bottom of the inverter cell, commonly known as current starving.

In this test case we optimize the performance of the five-stage CSVCO shown in Fig. 5.13 [87]. As shown in the schematic, the control voltage is generated by a current source I_B and a diode-connected transistor.

The optimization problem is defined as follows:

$$\begin{aligned}
& \text{minimize } \{\mathbf{P}(\mathbf{x}), \mathcal{L}(\Delta f, x)\} \\
& \text{subject to } \mathbf{x}_{lb} \leq \mathbf{x} \leq \mathbf{x}_{ub}, \mathbf{e}(\mathbf{x}, \mathbf{p}) = 0, \text{ and } \mathbf{g}(\mathbf{x}, \mathbf{p}) \leq 0 \\
& \mathbf{x} = [KI_B, KL_I, KWL_N, KWL_{NI}, KWL_P, KWL_{PI}, KLC],
\end{aligned}$$

where the set of normalized design variables \mathbf{x} , includes the biasing conditions for the current limiting transistors ($I_B = KI_B \times 1\mu A$), transistors lengths ($L_i = KL_i \times L_{min}$) and aspect ratios ($W/L = KWL_i$) for the transistor in the inverter and the biasing circuit.

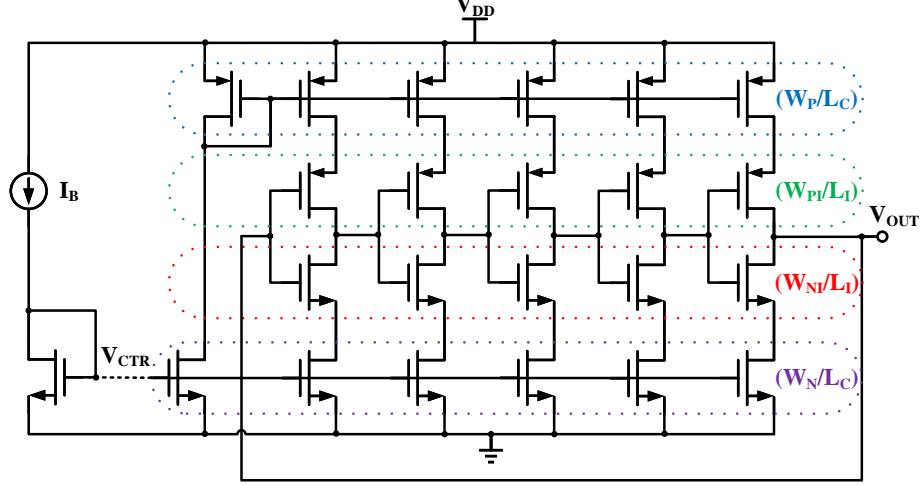


Figure 5.13: Circuit schematic of a 5-stage current starved voltage controlled oscillator (VCO). The control voltage is generated with a biasing current and a diode connected transistor.

5.4.5.1 Surrogate of the CSVCO's Performance Metrics

Equations (5.14) and (5.15) describe some of the most important performance metrics of the CSVCO such as oscillation frequency (f_{osc}), total power (P), and phase noise (\mathcal{L}) [87, 106, 107].

$$f_{osc} = \frac{I_D}{N \cdot V_{DD} \cdot C_{tot}} \quad C_{tot} \approx \frac{5}{2}(C_{ox} * W_{NI} * L + C_{ox} * W_{PI} * L) \quad (5.14)$$

$$P = P_{short-circuit} + P_{average} \quad \mathcal{L}(\Delta f) = \frac{8kTV_{DD}f_{osc}^2}{3\eta PV_{char}\Delta f^2}, \quad (5.15)$$

where, I_D is the inverter's current, V_{DD} is the supply voltage, C_{tot} is the total capacitance at the output of each inverter stage. W_{NI}/L and W_{PI}/L represent the aspect ratios of the NMOS and PMOS transistors, respectively.

Therefore, the optimization will provide the circuit sizing such that it meets the constraint of oscillation frequency while minimizing power and phase noise. Note that Δf is the offset frequency from the carrier where the phase noise is sampled, and it is set to $\Delta f = 1$ MHz for this experiment. Similar to the previous test cases, several constraints are also included to ensure that the transistors operate within moderate to strong inversion. However, this model does not account for errors in the current copy done by the current mirrors.

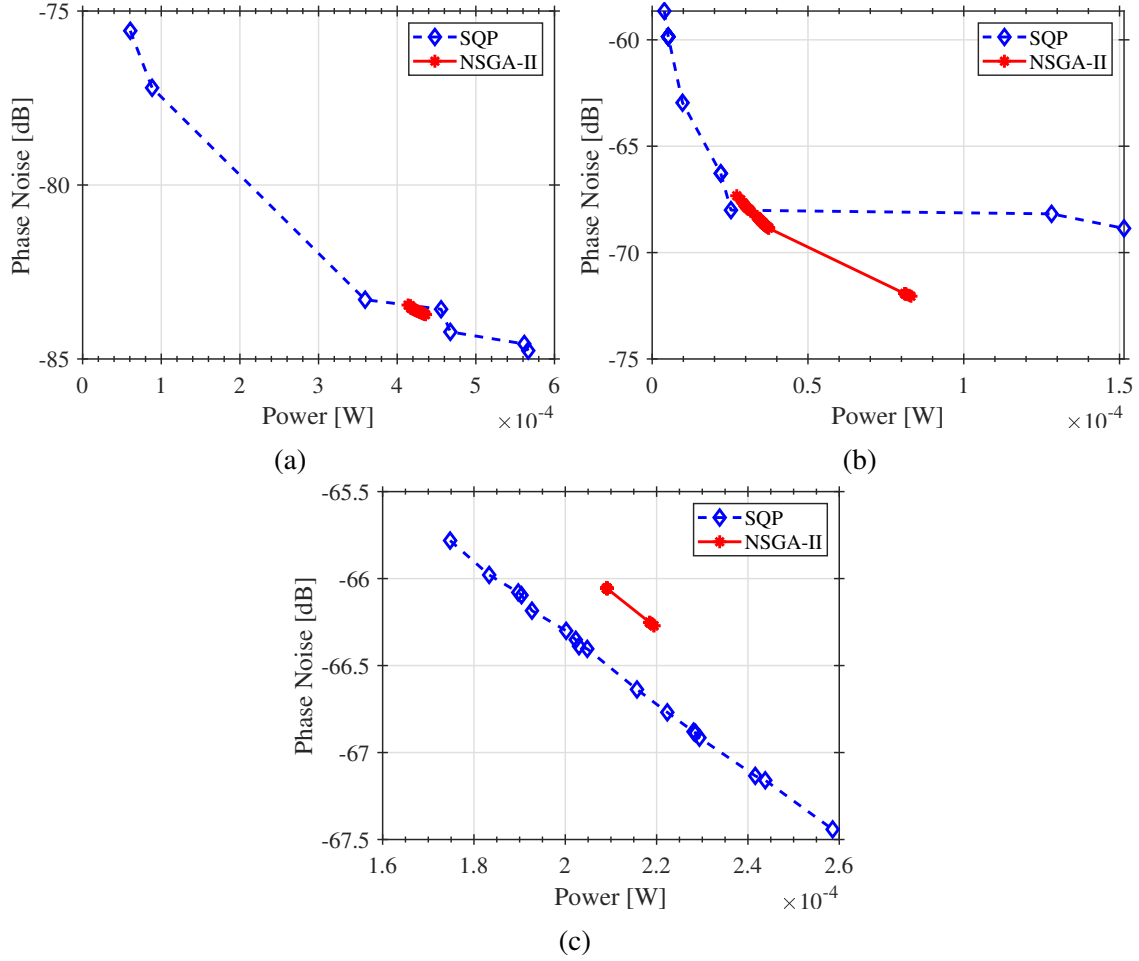


Figure 5.14: Pareto front of the optimization of the CSVCO circuit using SQP and NSGA-II optimization algorithms and **(a)** TSMC 180 nm process. **(b)** IBM 130 nm process. **(c)** TSMC 65 nm CMOS process.

5.4.5.2 Results of the CSVCO's Automatic Design

Fig. 5.14 shows the trade-off between phase noise and power consumption when sizing the CSVCO for a frequency of oscillation $f_{osc} = 1$ GHz for the TSMC 180 nm and IBM 130 nm processes, or $f_{osc} = 10$ GHz for the TSMC 65 nm process. From the Pareto fronts, we observe that the range of solutions found by the NSGA-II algorithm is narrower to the one found by the SQP.

That is also reflected in Fig. 5.15, where we observe a larger distribution of the variables of the solutions found using the SQP algorithm than the NSGA-II. Still, in both cases, the surrogate is accurate enough to allow both algorithms to find valid solutions across coroners verified through

simulation.

The constraints were also verified with circuit simulation to ensure that all solutions included in the Pareto provide a sustained oscillation at $f_{osc} = 1$ GHz (TSMC 180, IBM 130) or $f_{osc} = 1$ GHz (TSMC 65), and that under corners, the error of this frequency is no larger than 1%.

5.4.6 Summary

The main contribution of this work is a surrogate model that is computationally inexpensive, allowing fast evaluation of the objective functions and constraints for the optimization-aided automatic circuit design.

An example of the evolution of the objective function through the optimization processes of both algorithms is illustrated in Fig. 5.16. As expected, the objective function, considered here as a linear combination of the minimization objectives, reduces with the iterations (generations) until reaching the stop criteria. The SQP algorithm stops when the calculated step is smaller than the function tolerance, while the NSGA-II stops after reaching the maximum number of generations.

The test cases presented in this work are summarized in Table 5.7. The time required for the evaluation of objectives and constraints is compared when using the proposed surrogate and when using circuit simulations (using Ocean Cadence). Using the circuit simulator embedded in the optimization algorithm requires to write the parameters in the ocean file, source the software, run the simulation, and read the results. For example, in the case of the LDO, these operations take 0.6 s, 24.82 s, 1.18 s, and 5.5 ms, respectively. On the other hand, using the GP models require to load the models (41.58 s) and evaluate the model (17.84 ms). However, launching the simulation software and loading the GPR models happens only once in every execution of the optimization algorithm although these times seem large, they are negligible compared with the whole optimization run. Instead, the real target is to minimize the time required to evaluate a single solution since that will be repeated thousands of times in an execution. The number of evaluations in a single run depends on the number of iterations or generations, the size of the population, the stop criteria of the algorithm, and the number of multi-runs. Table 5.7 shows the comparison of the evaluation of a single candidate solution for all corners when using our surrogate vs the circuit

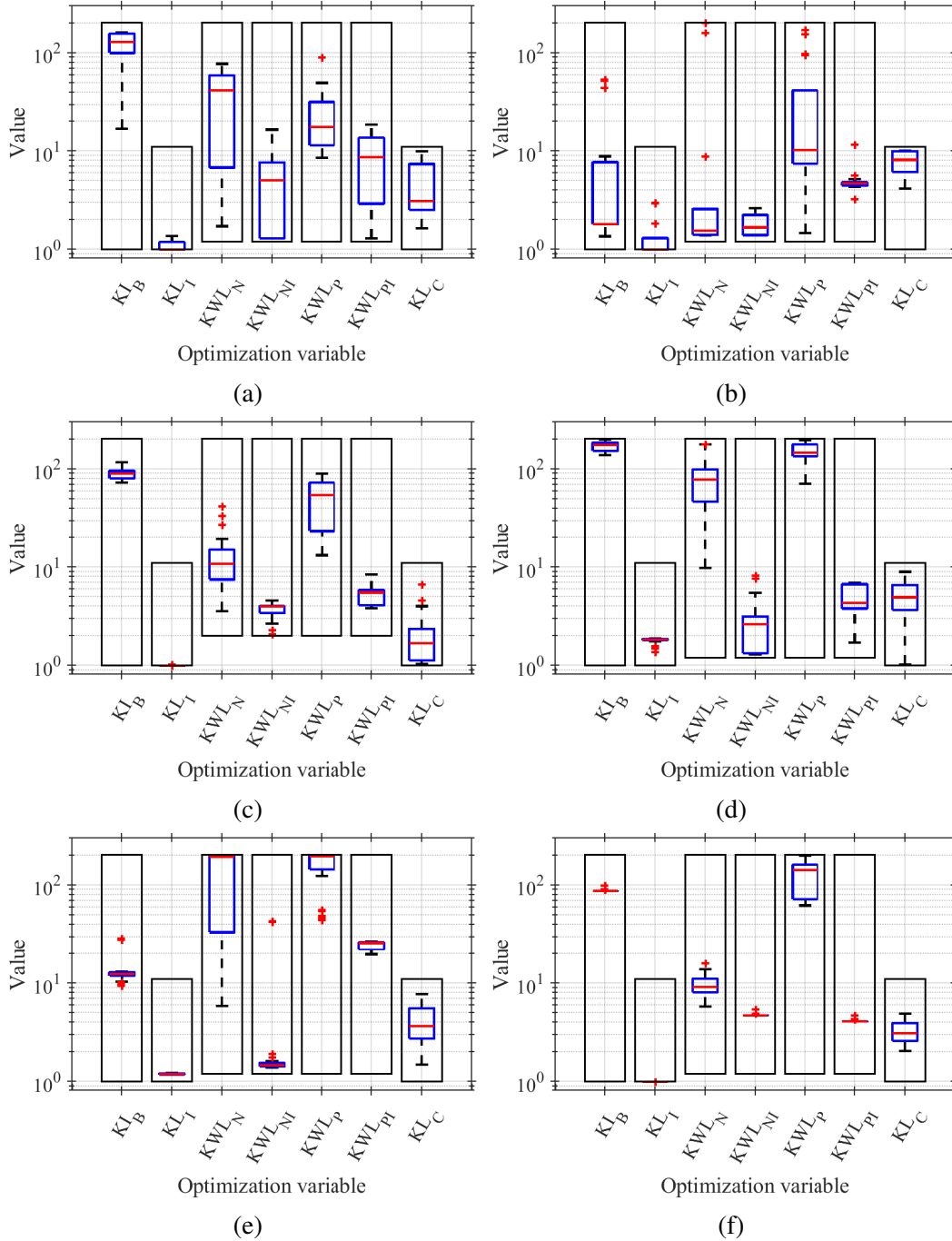


Figure 5.15: Values of the optimization variables of the solutions in the Pareto front obtained with: (a) SQP-180 nm, (b) SQP-130 nm, (c) SQP-65 nm, (d) NSGA-II-180 nm, (e) NSGA-II-130 nm and (f) NSGA-II-65 nm.

simulator.

Note that the simulation time is heavily dependent on the type of analysis require to quantify

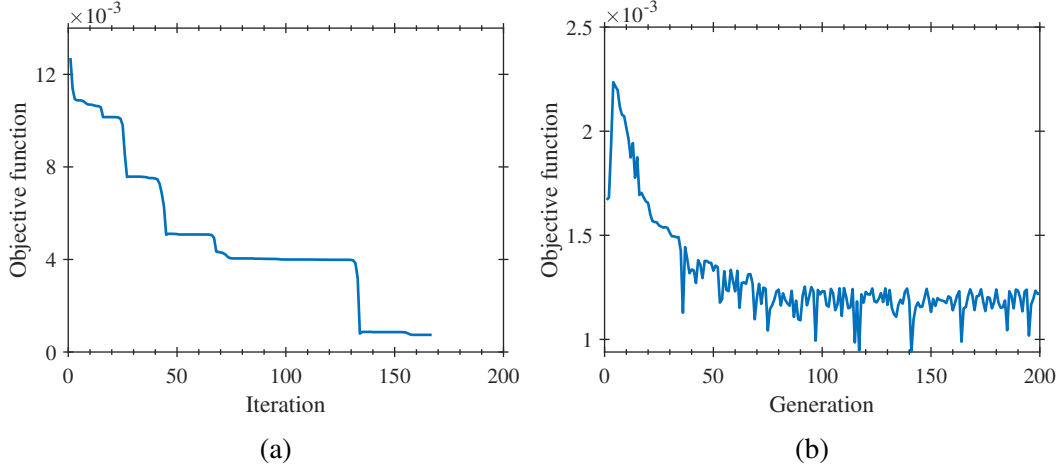


Figure 5.16: History of the objective function through the iterative optimization process using (a) SQP algorithm. (b) NSGA-II algorithm.

Circuit	No. of Design Variables	No. of Constraints	Evaluation Time Surrogate [s]	Evaluation Time Simulation [s]	Evaluation Time Improvement
Filter	9	21	0.123	57.708	470X
LDO	10	24	0.051	3.54	69X
VCO	7	15	0.063	18	285X

Table 5.7: Summary of the test cases under optimization using the proposed surrogate model.

certain metrics. For instance, measuring the slew rate requires a lengthy transient simulation in comparison to the phase margin that can be estimated with a faster AC simulation.

All results presented in the three test cases have been validated through simulation to verify all constraints are satisfied across process corners. Although the surrogate model uses highly accurate equations, avoiding simplification as much as possible, a margin of error in prediction is expected with respect to the high-fidelity model. One method to quantify the effectiveness of the surrogate for optimization-aided design is the success rate, which refers to the percentage of solutions generated through optimization using the surrogate that successfully comply with all specifications after being validated by circuit simulation. Table 5.8 summarizes the success rate of all the test cases across the three technology processes considered. Evaluating the solutions that fail to satisfy cir-

Circuit	TSMC180		IBM130		TSMC65	
	SQP	NSGA-II	SQP	NSGA-II	SQP	NSGA-II
Filter	90.48%	68.57%	100.00%	100.00%	57.14%	100.00%
LDO	76.67%	70.00%	70.00%	83.33%	56.67%	75.45%
VCO	60.00%	66.67%	70.00%	93.33%	86.67%	60.00%

Table 5.8: Summary of success rate from surrogate evaluation to simulation verification across corners.

cuit simulations can help identify opportunities to enhance the surrogate’s precision. For example, in the case of the filter design in the TSMC 65 nm process, the specification of the Op-Amp’s DC gain is the most significant active-constraint. Due to the lower intrinsic gain of the smaller process, the constraint is usually narrowly met. Then, even small errors in precision on the surrogate can cause the solution to fall under the constraint limit. Most of the solutions deemed invalid in the SQP-optimized case fall under a 1 dB error in the DC gain specification, and if accounted for, result in an overall success rate improvement from 57.14% to 83.33%. In the LDO optimization, some of the solutions fail to meet all specifications for both low-load and high-load conditions. This error is mainly caused by the error in the estimation of the pass transistor parameter’s, the largest device; this error reduces by increasing the resolution in the characterization of the early voltage. The error in the VCO optimization is mostly caused by the mismatch on the current copy of the current mirrors, even a small error in the current copy causes a deviation in the oscillation frequency larger than 1% when verified across corners.

Introducing additional details to the circuit model, such as current-mirror mismatch modeling could help improve the surrogate model in this regard. However, the additional effort required to increase the complexity of the model may not be justified if the valid solutions generated by the surrogate prove to be sufficient for the target application.

5.5 Conclusions

In this work, we presented a low-computational cost and accurate surrogate model for automatic IC sizing. This surrogate has three main elements that make it modular and reusable for the design of analog circuits across topologies and CMOS processes. In particular, Gaussian process regression is used to generate high accuracy prediction models of key device parameters based on characterization data of devices through process variations, expanding the capabilities of the transistor model to account for short-channel and narrow-channel effects. The process-aware nature of our surrogate reduces the iterative circuit verification to reach a viable circuit solution, despite the required initial effort to create it. Moreover, the created model can be easily re-used, such that a circuit can be re-designed for new applications by only updating the objectives and constraints. The low-cost surrogate allows for faster evaluation, which can enable the optimization of larger problems, demonstrating an improvement in evaluation time from $69X$ to $470X$ compared to the high-fidelity model.

The topology-dependent component of the surrogate requires the analog designer to obtain the expressions of the performance metrics. Therefore, to add a new topology, some initial effort is required. However, once the model is created, the designer can store it in a database, and reuse it for optimization across technologies and various performance metrics. Having a database of the models of different topologies could also allow a comparison of their performances to aid the designer in the topology selection. For instance, the optimization of RF circuits like power amplifiers requires extracting the design equations, performance specifications, and optimization objectives from works like [108] to build the architecture-dependent module of the surrogate model.

The proposed surrogate is integrated into a multi-objective constrained optimization framework with interchangeable state-of-the-art optimization algorithms. Then, the usage of our surrogate for automatic analog circuit design was tested on three different circuit topologies using three fabrication processes. The ability of our proposed surrogate to evaluate the circuit performance from the design variables was demonstrated by the generation of viable solutions across process corners independent of the optimization algorithm. Additionally, the use of our surrogate in conjunction

with multi-objective optimization allows the designer to improve the exploration of the space of solutions and to gain insight into design trade-offs through the Pareto front.

Future work should focus on the characterization and modeling of the post-layout parasitic components and include them in the surrogate to enable efficient parasitic-aware automatic analog design.

6. ANALOG IP PROTECTION TECHNIQUES¹

6.1 Introduction

The increased cost in manufacturing integrated circuits (IC) has led to the globalization of the IC supply chain. The distributed supply chain results in security threats against the intellectual property (IP) of ICs [109]. Researchers develop different design-for-trust (DfTr) techniques to secure digital circuits against these threats. These techniques include watermarking, IC metering, logic locking, camouflaging, and split manufacturing [2, 12, 110–112]. Similar to digital circuits, analog circuits are also susceptible to supply chain attacks. According to [113], analog ICs rank one in the top five counterfeited circuits.

The following details show the reason for this ranking and the necessity for strong analog IP protection techniques:

1. Despite the dominance of digital circuits, analog/RF circuits are still crucial to perform signal acquisition, filtering, amplification, and transmission [20].
2. The low transistor count, large transistor footprint, and reuse of some circuit topologies make it easier to obtain the analog circuit's netlist using reverse engineering techniques [114].
3. The design of analog/RF IPs is cumbersome due to the customization, expertise, and manual effort that goes into the schematic design and layout process. Hence, the IP of integrated circuits has great value [115].
4. The entry and exit points of major applications, like wireless communications and IoT, are in the analog domain. Hence, these points must be secured to avoid easy accessibility to the attacker [116–118].

¹Reprinted with permission from "Analog/RF IP Protection: Attack Models, Defense Techniques, and Challenges" by A. Sanabria-Borbón, N. G. Jayasankaran, J. Hu, J. Rajendran and E. Sánchez-Sinencio, in *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 68, no. 1, pp. 36-41, Jan. 2021,

Therefore, several techniques have been published to protect the IP of analog and mixed-signal (AMS) circuits. These techniques are either based on locking the analog circuit with a key input [5, 6, 8, 119–121] or camouflaging the analog circuits [10, 11]. Attack techniques to evaluate the resilience of defense schemes are proposed in [3, 4].

We describe the different supply chain attacks in Section 6.2. In Section 6.3, the various defense techniques to protect analog-only and AMS circuits are discussed. Various forms of key provisioning are outlined in Section 6.4. We then explain the resilience offered by these defenses against the different attacks available in the literature in Section 6.5. The future research directions and the conclusion are given in Section 6.6.

6.2 Supply chain attacks

Depending on where the attacker is located within the IC supply chain and the resources he/she has access to (please refer to Table 6.1), there are different types of supply chain attacks [109] that are described as follows:

- **IP piracy.** In this attack, the attacker steals and claims ownership of an IP. He/She can be in the foundry having access to the layout, process design kit details, and sufficient resources to modify it based on his/her requirements and claim that to be his/her product.
- **Overproduction.** This attack is a subset of IP piracy. Here, the attacker has minimal resources, sufficient to overproduce the chip but not sufficient to change the existing mask and pirate the design.
- **Counterfeiting.** The attacker takes a used IC, refurbishes it, and markets it as a new IC [122].
- **Reverse engineering (RE).** An attacker as untrusted end-user purchases a chip, de-packages it, and takes high resolutions pictures of each layer of the chip using a scanning electron microscope. The pictures are then processed using an image processing software, which helps in annotating the netlist [114, 123].

- **Hardware Trojans.** They are malicious circuits inserted in the design, whose activation in the design can cause either denial of service, performance degradation, or stealing secret information such as crypto keys [124–127].

The following section discusses the different analog-only and AMS locking techniques.

6.3 Defenses techniques to protect analog IP

The DfTr techniques developed to secure digital circuits cannot secure analog circuits for the following reasons.

1. the analog circuits are designed using transistor-level schematics [5, 119], whereas the digital circuits are designed either in RTL-level [130] or gate-level [131].
2. the digital circuits give incorrect responses, even if there is a one-bit difference in key input. In contrast, the analog circuits can provide close to the desired response for certain incorrect keys.

Source	Information acquired
Circuit netlist extracted from the GDS files or reverse engineering [114]	<ol style="list-style-type: none"> 1. Circuit topology, architecture, pin connectivity. 2. Transistor sizes (W, L) and element values (R, C, L). 3. Programmable circuit elements. Connectivity between security key and circuit devices.
Process design kit (PDK) [128]	<ol style="list-style-type: none"> 1. Available devices: transistors (CMOS, BJT), resistors, capacitors, inductors, diodes. 2. Process parameters and device models. 3. Parasitic modeling.
Circuit's datasheet [129]	<ol style="list-style-type: none"> 1. Circuit functionality and applications. 2. Specifications and operating conditions. 3. Chip pin-count and layout recommendations.

Table 6.1: Sources of information available to the attacker.

Hence, there are new techniques developed to secure analog-only and AMS circuits [5, 6, 8, 9, 119–121]. Analog locking consists of hiding some design parameters to prevent its unauthorized use. Here we classify the analog locking techniques according to the obfuscated parameter. We will discuss techniques that obfuscate the circuit current or voltage biasing. These are perhaps the most general techniques, given that all analog topologies require biasing of some sort. However, the obfuscated biasing block is easy to identify and remove or bypass. If only the biasing is unknown, an experienced analog designer might estimate the correct biasing and remove the lock altogether. Then we will describe some techniques that obfuscate the equivalent sizing of the gain devices of the protected analog lock.

Other techniques exploit the digital components of AMS circuits. These architectures have both analog and digital circuits integrated within the same die and working collaboratively. An example of an AMS circuit is a signal acquisition system with an analog front-end followed by digital signal processing. Another example is the digital-aided calibration that compensates for the performance degradation of analog circuits due to variations. In this context, several defense techniques leverage the advantages of logic locking to secure AMS circuits. Finally, we will look into techniques that perform camouflaging at the layout level to protect against reverse engineering attacks.

6.3.1 Locking the bias generation circuits

The proper biasing of analog circuits is fundamental for their correct operation. Therefore, the defenders lock the bias circuit that generates the current or voltage biasing. Hence, the precise value of the bias current or voltage is generated only for the correct key, and the analog circuit performs as expected. Otherwise, incorrect bias current or voltage is generated, and hence, the analog circuit's performance deviates from the expected response.

6.3.1.1 Current biasing

The simple current mirror shown in Fig. 6.1(a) is a common biasing circuit of analog circuits. It produces an output current (I_{OUT}) that is a weighted copy of the input current (I_{REF}).

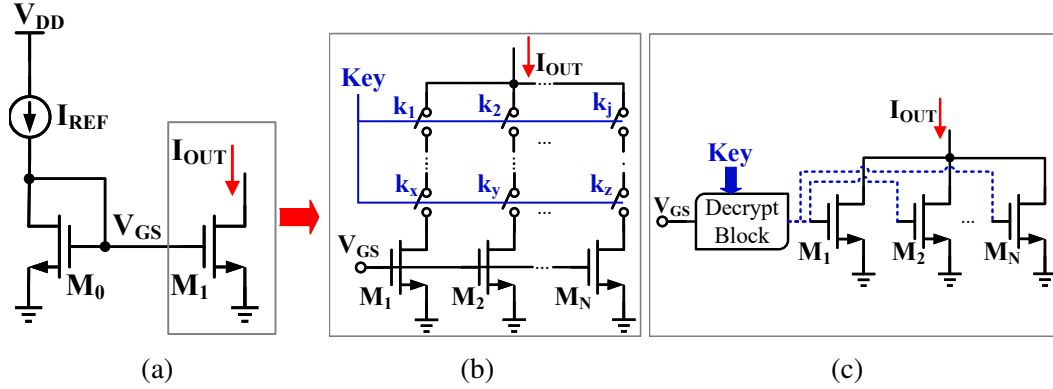


Figure 6.1: Locking the current biasing of analog circuits. (a) Simple current mirror. (b) Configurable current mirror [5]. (c) Parameter biasing obfuscation [6]. Only the correct key provides the correct output current.

- **Configurable current mirror (CCM) [5]**

In this technique, a simple current mirror is replaced with a configurable current mirror (CCM), as shown in Fig. 6.1(b). A transistor array, form by scaled mirroring transistors and analog switches, substitutes the current mirror's single output transistor. A satisfiability modulo theories (SMT) solver designs the transistor array's sizing to guarantee that the circuit only meets the specifications for the correct key. Moreover, using both NMOS and PMOS current mirrors results in a non-monotone relation between a key change and the corresponding output current variation.

This technique ensures that each chip has a unique key. An on-chip physically unclonable function (PUF) [132] generates a unique chip identifier, which is XORed with the inserted key to produce the CCM's control bits. The main advantage of this approach is that the current mirror is a ubiquitous biasing block. This lock can protect a wide range of circuits that can use the current mirror as their biasing circuit. However, the transistor array size is limited by the number of switches that can be piled up and effectively controlled, given the circuit headroom and the technology supply voltage.

In parameter biasing obfuscation [6], the current bias is obfuscated by replacing the single output transistor with a parallel array of weighted devices, as shown in Fig. 6.1(c).

- **Range controlled floating gate transistors [133]**

This work assumes that analog floating gate devices are inserted and programmed to compensate for the effect of variations. This locking technique only allows the calibration when the correct key is applied. In this approach, floating gate transistors are used to create (compensate) an amplifier’s offset voltage by introducing (reducing) the current imbalance of the differential branches. It consists of a two-step process unlocking/calibration process.

6.3.1.2 Voltage biasing

- **Parameter biasing obfuscation [6]**

In this approach, the effective device sizes of a transistor-based voltage divider are obfuscated. A bias voltage is generated using a transistor-based voltage divider, as shown in Fig. 6.2(a). The effective sizes of the PMOS and NMOS devices are obfuscated by replacing them with a parallel array of devices. The key bits control the V_{GS} voltage of all the transistors in the array. In a practical implementation, the V_{GS} voltage is generated using diode-connected transistors, resulting in a simplified version of the CCM-based lock.

- **Memristor-based protection [119]**

This technique secures a sense amplifier used for memory readout using memristors. Any mismatch in the input differential pair of the sense amplifier due to process variations generates an

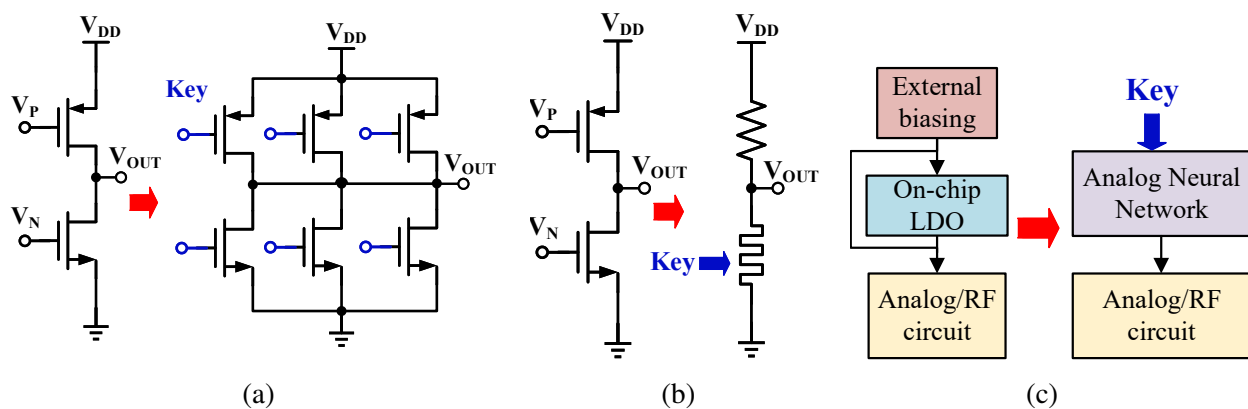


Figure 6.2: Voltage biasing locking. (a) Obfuscated voltage divider. (b) Memristor-based voltage divider. (c) Analog neural network-based voltage biasing.

offset voltage that can alter the circuit's output. This approach inserts an intentional mismatch in the input differential pair, making the readout circuit's output unreliable. The input key programs a memristor crossbar implementing a voltage divider as shown in Fig. 6.2(b). The output voltage biases the differential pair transistors' body terminal. Only the correct key provides the body biasing that compensates the offset voltage enabling the correct reading of the memory data. Otherwise, the sense amplifier corrupts the data read from the memory.

This approach has several advantages. The memristor can be precisely tuned within a wide range, and it breaks when a wrong key is applied, significantly increasing the effort of a brute force attack. However, this defense technique has several implementation challenges. First, the design of the sense amplifier's differential pair. The proposed approach uses an NMOS differential pair, which requires a triple-well process to enable individual transistor body biasing. Using a PMOS differential pair instead is compatible with more available and less expensive fabrication processes. Moreover, the tuning range of the body voltage is limited by the forward biasing voltage of the parasitic diodes. Exceeding this voltage causes latch-up that can damage the circuit. Also, memristors are not compatible with standard CMOS processes, which prohibits their on-die integration.

- **Neural network-based voltage biasing [121]**

In this approach, the voltage biasing of a circuit is obfuscated using an analog neural-network (ANN), as shown in Fig. 6.2(c). After the chip fabrication, the ANN is trained, and the resulting weights are stored in floating-gate (FGT) structures. It enables assigning unique keys to each chip. The key is a set of analog values applied at the ANN inputs by: i) generating the analog voltages off-chip and applying them through external pins, ii) using an on-chip combination of digital memory and digital to analog converter (DAC), or iii) using an on-chip analog storage element such as an FGT. The on-chip ANN uses a $n \times m$ array of synapses and neurons. While the synapses perform four-quadrant multiplications, the neurons implement the activation function. This approach has an ultra-low power consumption (sub- μW), but it has a very large overhead due to the ANN and additional circuitry for its training and operation.

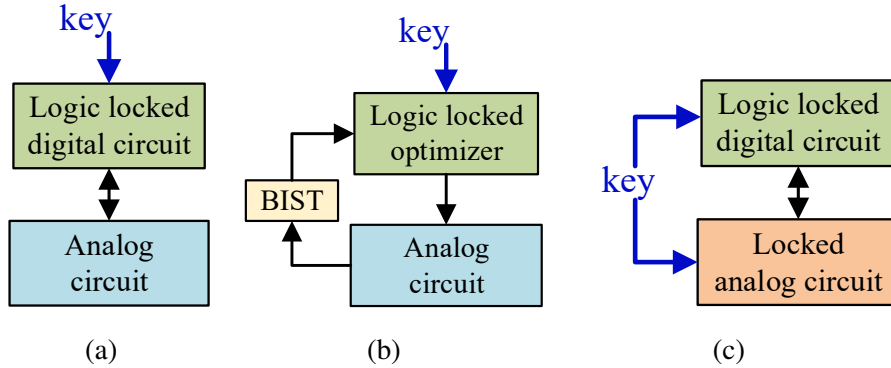


Figure 6.3: Locking of analog and mixed-signal circuits (AMS). (a) MixLock [7] uses logic locking on the digital component of the AMS circuit. (b) AMSlock [8] locks the optimization core setting the tuning knobs of the analog part. (c) Shared dependencies [9] locks both the digital and the analog components of the AMS circuit.

As per our knowledge, this is the first locking technique that uses an analog key instead of a digital one. This work suggests that the continuous nature of analog voltages make analog keys significantly more difficult to guess in comparison to a digital one. However, the number of possible keys is not unlimited. There is a maximum available voltage range often given by the supply voltages, which shrinks in newer processes. Also, there is a minimum voltage resolution, below which the lock can not distinguish a correct versus an incorrect key.

6.3.2 Locking the gain transistors of analog circuits

- **Transistor sizing obfuscation [134,135]**

In this approach, the sizing of the transistors of the analog circuits providing gain is obfuscated. Each device is replaced by a parallel transistor array [134] or a transistor mesh structure [135]. The array or mesh is built with transistors of different sizes and switches controlled by the security key. Obfuscating the effective size of the gain transistors results in hiding their transconductance and output impedance. By extension, the circuit’s performance metrics are protected.

6.3.3 Locking analog and mixed-signal (AMS) circuits

- **MixLock [7]**

This technique consists of logic locking the digital components of AMS circuits, as shown in

Fig. 6.3(a). The stripped Functionality Logic Locking (SFL) technique protects the functionality of the digital circuit blocks. This work inherits the low-overhead and fully design automation from hardware security in the digital domain. It is non-intrusive to the analog circuits preventing performance degradation, but it also increases the risk of removal attacks. This defense technique is also widely applicable except for purely analog circuits.

- **AMSlock [8]**

As shown in Fig. 6.3(b), this technique uses logic locking to secure the optimization core of a built-in self-test (BIST) on-chip PVT compensation approach in [68, 136]. The optimization core calculates the error between the circuit specifications and its measured performance and generates the corresponding tuning knobs that minimize that error. This work locks the optimization engine using the SFL technique such that it only reduces the error when the correct key is applied. Otherwise, the optimization output sets the analog tuning knobs to a value that does not reduce the error between the measured and desired responses. A challenge of this approach is to design the tuning range of the tuning knobs. The resolution should be small enough to compensate for performance errors due to variations and mismatch. Furthermore, the variation in the tuning knobs should be able to cause a significant performance degradation to distinguish the correct key from all the incorrect ones.

- **Shared dependencies [9]**

This approach proposed the locking of both the analog and digital blocks of analog and mixed-signal (AMS) architectures, as illustrated in Fig. 6.3(c). Additional functional and behavioral dependencies between the digital and analog components are explored for increasing the attack effort by a factor of three. While parameter-biasing obfuscation lock the analog circuits, random logic locking (RLL) [12] or SFL [2] protect the digital components.

6.3.4 Camouflaging of the analog circuit's layout

Until now, we assume the attacker can obtain the correct netlist using RE techniques. Camouflaging techniques increase the effort of RE-based attacks by hiding information at the layout

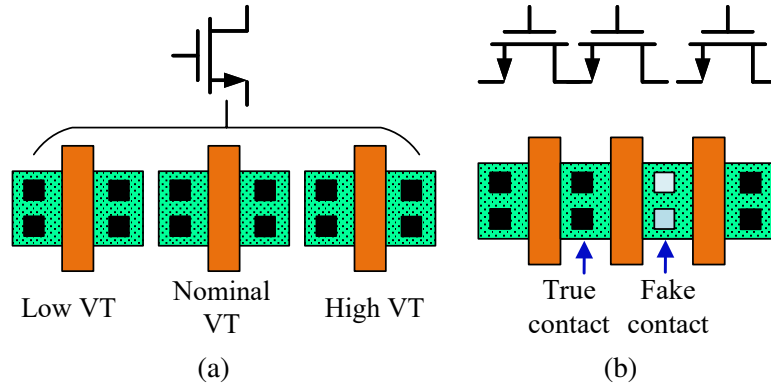


Figure 6.4: Analog layout camouflaging techniques protect against reverse engineering-based attacks. (a) In multiple threshold voltage (VT) camouflaging [10], transistors with different threshold voltages (VT) are used to hide the circuit design. (b) Sizing camouflaging [11] uses fake contacts to hide the circuit connectivity and effective device sizing.

level.

- **Multiple threshold voltage (VT) camouflaging [10]**

This technique uses the process design kit (PDK) options to provide resiliency against RE-based attacks. Several CMOS fabrication processes offer transistors with different threshold voltages (VT). For instance, a process can include nominal-VT, low-VT, and high-VT devices. As shown in Fig. 6.4(a), this approach assumes that the RE process reveals no information about the device doping, and the three device options look exactly alike. Therefore, the attacker cannot identify which kind of transistor is being used. This defense technique uses this advantage to replace some of the original unprotected design's transistors (most probably using only nominal-VT devices) with a choice of LVT or HVT.

This approach relies on the considerable performance variation from using transistors with different VTs. However, it also becomes a challenge for the IP owner to protect his/her design. The camouflaged design must achieve the same performance metrics as the original one, which entitles a considerable redesign effort. The small number of VT options, the low transistor count of analog circuits, and the matched devices result in a small number of combinations the attacker needs to consider before finding the correct one.

- **Sizing camouflaging [11]**

This approach uses obfuscated layout components that contain several fake contacts to hide the connectivity and the effective size of the transistors, resistors, and capacitors in an analog circuit. Fig. 6.4(b) illustrates how this technique protects against RE-based attacks [131]. If the attacker can not single out the fake contacts, the netlist extracted through RE techniques has the wrong device sizes and its performance does not meet the specifications. According to [11], the fake contacts used in this approach are compatible with CMOS technologies, do not require any change in the fabrication process, and are very difficult to differentiate from true contacts using existing RE techniques. Therefore, it is enough to obfuscate only a few components resulting in low area overhead. A challenge of this approach is to ensure that the parasitic components associated with the fake contacts do not perturb the analog circuit's performance. Hence, the obfuscated component must undergo a thorough post-layout verification.

6.4 Key provisioning techniques

Most of the protection techniques summarized in this section rely on a security key known only to the circuit IP owner. Key management refers to the process of inserting the key in the analog circuit to make it functional without directly revealing the secret key to the user. There are different techniques to provide the security key to the analog lock. The most common one is storing the key in a tamper-proof memory delivered along with the protected chip. However, if the attacker can perform probing at the connections between the memory and the chip, the secret key might be exposed. Techniques that use a non-volatile on-chip storage like fuses, EEPROM memory, or FGTs [12, 118, 121, 137] share the same vulnerability.

If that secret key is the same for all the protected circuit instances, the attacker can use that key to unlock all chips. A key provisioning unit addresses this issue by assigning a unique key to each chip instance. Ending piracy of integrated circuits [12] was the first work to create a protocol for remote chip authentication. It uses public-key cryptography to enable the chip's activation and testing, even at an untrusted entity without revealing the secret key. This approach has a large area overhead due to the need for on-chip PUF or TRNG and RSA modules. The key provisioning

Defense	Threat model	Resiliency against attack					
		Brute-force	RE-based	Removal/bypass	SAT-based	SMT-based	GA-based
Combinational lock CCM [5]	IP piracy	✗	NA	✗	✗	✓	✓
Parameter-biasing obfuscation [6]	IP piracy	✗	NA	✗	✗	✓	✓
Range controlled FGT [133]	IP piracy	✗	NA	✗	✗	✓*	✓*
Analog neural network [121]	Illegitimate access	✗	NA	✗	✗	✓	✓*
Memristor-based protection [119]	IP piracy	✗	NA	✗	✗	✓	✓*
Transistor sizing obfuscation [134, 135]	IP piracy	✗	NA	✗	✗	✓	✓
MixLock [7]	IP piracy	✗	NA	✗	✓	✗	✗
AMSlock [8]	Overproduction	✗	NA	✗	✓	✗	✗
Shared dependencies [9]	IP piracy	✗	NA	✗	✓	✗	✓*
Multiple-VT camouflaging [10]	Reverse engineering	✗	✗	✗	✗	≈	✗
Sizing camouflaging [11]	Reverse engineering	✗	✗	✗	✗	✗	✗

Table 6.2: Attack success of the attacks reported in Section 6.5 on the defense techniques in Section 6.3. It denotes the successful (✓), unsuccessful (✗), potentially successful but not demonstrated (✓*) attacks. ≈ denotes that the attack reduces the search space.

technique in [5] uses an on-chip PUF to generate a unique chip identifier. It is XORed with a unique user key to generate the common key that is the same for all the instances of the same protected chip. A recent Schmitt trigger-based key provisioning [17] takes in the unique user key and generates the common key. It has a minimal area overhead, is compatible with both analog and digital keys, and has proven to increase analog locks’ security.

6.5 Breaking Analog IP protection techniques

Here we describe potential attacks to the protection techniques presented in Section 6.3. Table 6.2 summarizes the attack success on these techniques.

6.5.1 Brute-force attacks

The attacker has access to an unlocked chip and also knows its expected performance according to its datasheet. He/she performs circuit simulations for all possible keys one at a time and observing the circuit’s response. Following this procedure, the attacker can eventually find out a single key that makes the circuit perform as per specification. The defender tries to make this attack infeasible by minimizing the probability of finding the correct key, which is often associated with increasing the key size. An 80-bit key is considered secure, given the existing computational

capabilities.

6.5.2 Reverse engineering

Reverse engineering refers to a set of techniques that derive IC/IP proprietary information such as schematic netlist and layout. Although several companies use RE to discover if other vendors are using their licensed IPs, attackers use it to steal the IP information and obtain profit. All of the IP protection techniques in Section 6.3 are susceptible to reverse engineering, except probably for the ones based on layout camouflaging.

6.5.3 Removal/bypass attacks

The attacker identifies the obfuscated component controlled with the security key. The attacker can remove it and replace it with the equivalent component. This attack requires sufficient analog IC design expertise to: i) identify the obfuscated component, ii) re-design the equivalent replacement that makes the circuit to meet the specifications, and iii) perform changes at schematic and layout level.

6.5.4 SAT-based attack [1,2]

It is a Boolean satisfiability-based attack on combinational logic locked circuits. It can infer the correct key after performing a few input/output observations from the oracle. This attack targets the security of AMS circuits that rely on locking only the digital component.

6.5.5 SMT-based attack [3]

SMT formulations can handle continuous variables. Hence, they have been used to effectively break most of the existing analog locking techniques. The attacker obtains the circuit's netlist either from an untrusted foundry or a reverse engineering process, the PDK information, and the circuit specifications from the datasheet. He/she then identifies the obfuscated component by analyzing the connectivity between the key input and the circuit. Next, analog expertise is required to formulate a set of equations. Some equations represent the relation between the obfuscated component and the circuit parameters, while others show the circuit performance metrics as a function

of the circuit parameters. An SMT-based solver uses the equation set to find the correct key. If the solver returns more than one key, a functional chip is used as an oracle to set apart the correct key.

6.5.6 Optimization based attacks [4]

This attack approach uses a genetic algorithm (GA) to find the obfuscated components or the secret key of analog locks. Initially, it requires the circuit netlist and an unlocked chip to use it as an oracle. The objective function then captures the error between the simulated circuit's performance after applying a particular key and the desired one measured in the oracle. A GA is well suited to perform this attack since the random initialization of the population and the mutation operator allow a thorough exploration of the search space. Also, it can handle multi-objective optimization. Compared with the SMT-based attack, this attack does not require prior knowledge of the circuit and returns a single key.

6.6 Conclusion and future research directions

This tutorial summarized the supply chain attacks against analog IP and the existing protection techniques against those threats. However, most of the analog locking based techniques can be broken using SMT-based or optimization-based attacks. The research scope of analog IP protection techniques include: (i) advanced defense techniques that are resilient to SMT- and optimization-based attacks, (ii) key provisioning units that enhance the security of the existing techniques with low overhead, (iii) standardization of security metrics that allow benchmarking and comparison of the techniques, and (iv) techniques that protect the IP ownership based on watermarking or fingerprinting.

Analog IC protection will soon become essential as regulation of electronic hardware security is imminent. As the number of attacks on electronic devices increases, the industry needs to implement security properties to protect their products' functionality, the user's data, and their design's IP. In this context, the current cost-driven analog design paradigm will have to embrace security as a critical design factor. Moreover, security features must have low power and area overheads while being robust to variations. Hence, a new design paradigm should consider security from the start

of the design process and not as an addendum in the latest stages.

The defense techniques to protect analog IP have focused on performance locking, camouflaging, and trojan detection [138, 139]. We anticipate two crucial requirements to make these techniques practical in real applications. First, the standardization of quantifiable security metrics. Second, the validation of the proposed approaches with silicon fabrication to quantify their robustness to PVT variations and overhead.

There is also scope for research in chip authentication and ownership protection. Analog techniques based on watermarking [140], fingerprinting, and IC metering should be explored.

7. SCHMITT-TRIGGER BASED KEY PROVISIONING TECHNIQUE¹

7.1 Introduction

The increased cost in fabricating integrated circuits (ICs) has led many semiconductor companies to go fabless. These companies face challenging security threats due to the outsourcing of IC fabrication. Security threats include intellectual property (IP) piracy, overproduction, reverse engineering, counterfeiting, and hardware Trojans [109]. Several design-for-trust (DfTr) techniques such as logic locking, camouflaging, and split manufacturing are proposed to secure digital circuits [2, 12, 110, 111] and analog circuits [5, 6, 8, 9, 119–121] against these threats. Logic locking is the most preferred DfTr technique as it protects the circuit from an untrusted foundry and an untrusted end-user, whereas other techniques protect the circuit from only one of them.

In digital logic locking [2, 12], the circuit is encrypted by inserting key-gates, additional gates connected to the key inputs. When the correct key is applied, the design functions as intended. Otherwise, applying an incorrect key produces an incorrect output.

Similarly, in analog locking [5, 121, 135], the key inputs control design parameters like the biasing (voltage or current) or the effective sizes of the transistors (channel length L and width W). Since these parameters have a direct impact on the circuit's response, its performance metrics are locked. Only the correct key configures these parameters such that the circuit performs as per specifications. Otherwise, the error between the measured circuit's response and the specified one is larger than the acceptable tolerance. All instances of the protected circuit share the same key, *a.k.a.*, the common key (CK). This key is the designer's secret and is available only to the authorized user.

The CK is either stored in a tamper-proof memory, as in [2], or generated by a key provisioning unit [5, 12]. In [2], if the attacker finds the CK, then all the instances of the same design can be unlocked using this key [1, 3]. A key provisioning technique helps in addressing this issue. As

¹Reprinted with permission from "Schmitt Trigger-Based Key Provisioning for Locking Analog/RF Integrated Circuits" by A. Sanabria-Borbon, N. G. Jayasankaran, S. Lee, E. Sanchez-Sinencio, J. Hu, J. Rajendran, in Proceedings of the international test conference (ITC) 2020.

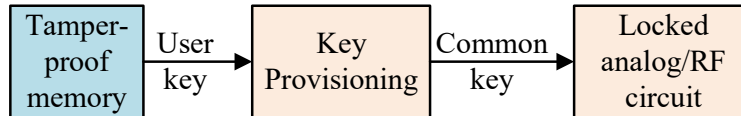


Figure 7.1: The key provisioning unit generates the common key using the user key which is unique to that chip instance.

shown in Fig. 7.1, this block takes in the user key (UK), and generates the CK, which is equal for all the instances of the locked analog circuit. A key provisioning unit ensures that each chip instance can be unlocked only by the UK, which is unique to that instance.

7.1.1 Related works on key provisioning techniques

Ending piracy of integrated circuits (EPIC) was the first work proposed on key provisioning [12]. It uses a physically unclonable function (PUF) or a true random number generator (TRNG), and RSA encryption to remotely activate a locked chip. The protocol for remote activation is as follows:

Step 1: As illustrated in Fig. 7.2, the designer locks the circuit with a CK and embeds his/her public master key (MK-Pub) in the circuit. Only the designer knows the secret CK. A PUF/TRNG and an RSA module are also inserted in the chip. The locked design is sent to the untrusted foundry, where the chip is manufactured and tested. The testing process does not require to load the key into the chip [141].

Step 2: On the first power-up, the manufactured chip generates the public and the private random chip keys RCK-Pub and RCK-Pri, respectively, using the PUF/TRNG. The foundry sends the RCK-Pub to the designer.

Step 3: The designer encrypts the CK with RCK-Pub. This can be decrypted only with the RCK-Pri generated inside the locked chip by a PUF/TRNG. For authentication, the encrypted CK is signed using the MK-Pri to generate the UK.

Step 4: The UK is sent to the foundry to activate the locked chip. The RSA module inside the locked chip authenticates the UK with MK-Pub and then decrypts it using RCK-Pri to obtain CK, thereby activating the locked chip.

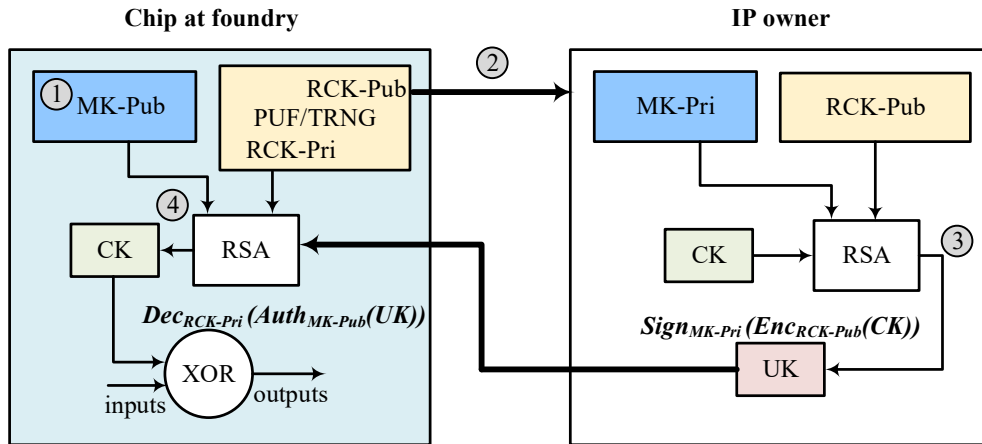


Figure 7.2: EPIC protocol for remote activation of the locked chip using public key cryptography [12]. Master key (MK), random chip key (RCK), common key (CK), user key (UK), public (Pub), private (Pri).

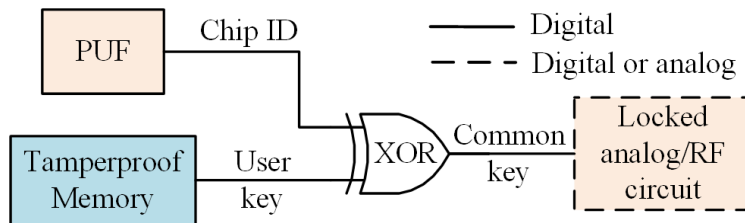


Figure 7.3: In the combinational lock [5] the key provisioning generates the common key from the unique user key with the aid of a PUF [13].

Fig. 7.3 shows another key provisioning technique [5]. In this work, a PUF produces an individual chip ID for each chip instance [13]. This chip ID is XORed with the UK to provide the CK. The UK is unique for each chip instance.

In general, a key provisioning technique should have the following properties:

1. each chip instance should have a unique UK,
2. given the UK, the attacker should not be able to recover the CK, i.e., the output of the provisioning unit should be unintelligible to the attacker, given the UK, and
3. low power and area overheads.

7.1.2 Limitations of existing key provisioning techniques

The EPIC work [12], which uses a PUF and an RSA module, remotely activates the locked chip. Hence, there is no requirement for a tamperproof memory as the UK is public. However, using the RSA module for one-time CK generation cannot be justified for the area overhead it incurs [12]. Moreover, a standalone PUF has the following limitations:

- It is not suitable for generating the CKs as their output is not deterministic and stable with process, voltage, and temperature (PVT) variations [142]. There have been many attacks on the PUFs based on statistical and machine learning techniques [143, 144].
- Its size increases as the size of the chip ID required increases. For example, in [145], each six-transistor Schmitt trigger (ST), generates a one-bit response. Hence, for an n -bit chip ID, the outputs of n PUFs should be concatenated.
- It is not compatible with analog UKs.

Although many logic locking approaches have been proposed across digital and analog domains [2, 5, 12, 119], there has been very little systematic research on key provisioning techniques. Therefore, to address the limitations in existing techniques, we propose a generalized ST-based key provisioning unit with low-area overhead that is compatible with any digital or analog locking approaches, like those listed in Table 7.1. In the proposed technique, the CK and the UK can take either digital or analog values.

7.1.3 Contributions of this work

We propose an ST-based key provisioning technique. This circuit takes in the UK, which is unique for each chip. It generates the CK required to unlock the analog and digital circuits locked using various logic locking techniques [5, 6, 8, 119, 120]. The contributions of this work are:

1. We propose a key provisioning technique based on the ST circuit. The conventional ST operation is enhanced with dynamic hysteresis and inversion of the thresholds to create a CK with the desired security properties.

Locking technique	Digital		Analog	
	UK	CK	UK	CK
EPIC [12]	✓	✓	✗	✗
Stripped functionality logic locking [2]	✓	✓	✗	✗
Combinational lock [5]	✓	✓	✗	✗
Memristor-based protection [119]	✓	✓	✗	✗
Parameter-biasing obfuscation [6]	✓	✓	✗	✗
AMS lock [8]	✓	✓	✗	✗
Mixlock [7, 120]	✓	✓	✗	✗
Analog performance locking [121]	✗	✗	✓	✓

Table 7.1: Existing digital and analog locking techniques and their user key (UK) and common key (CK) types. The proposed key provisioning technique receives the UK and generates CK. It is compatible with analog and digital keys.

2. The ST-based technique generates a unique UK for each chip. We use the Hamming distance as a metric of the uniqueness of the UK.
3. The UK contains most of the ST’s configuration. The remaining information is stored in on-chip fuses, written before the chip’s distribution. It increases the effort of reverse engineering attacks.
4. The proposed technique for key provisioning has a smaller area overhead compared to the existing approaches [5, 12]. In our technique, the UK is divided into segments applied serially to reuse the same circuitry. Hence, the area remains constant and independent of the key size.
5. The output, or CK, is deterministic for every input and robust to PVT variations.
6. The efficacy of this key provisioning technique is demonstrated on different locked analog circuits: a Gm-C bandpass filter (BPF), a common-gate low-noise amplifier (CG-LNA), and a low-dropout voltage regulator (LDO).
7. We present a new metric to evaluate key provisioning, namely entropy. This metric is used to

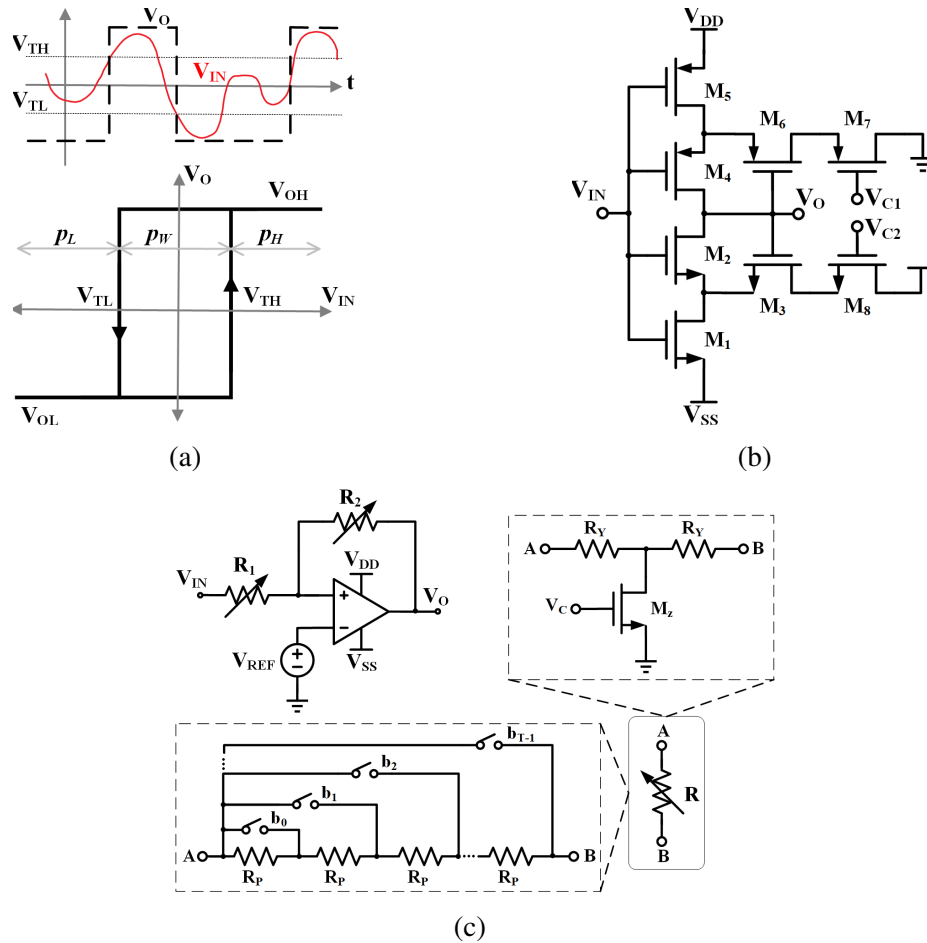


Figure 7.4: (a) Voltage response of a non-inverting ST (top) transient response: while the output values are either V_{OL} or V_{OH} , the input voltage takes any value in between. (bottom) Voltage transfer characteristic. (b) Internal feedback or 6-T ST. The thresholds are defined by the sizing of the transistors M_3 and M_6 and the control voltages V_{C1} and V_{C2} . (c) The external feedback ST uses an amplifier with resistors implementing positive feedback. The programmable resistors can be controlled by a digital word $\{b_0, b_1, b_2, \dots, b_{T-1}\}$ or a control voltage V_C .

estimate the effective size of the key generated by the proposed key provisioning.

7.2 Background

7.2.1 Schmitt trigger (ST)

An ST is a comparator with hysteresis that uses positive feedback to amplify the difference between the input voltage (V_{IN}) and the threshold voltages (V_{TL} and V_{TH}). This difference produces an output voltage (V_O) that takes either low (V_{OL}) or high (V_{OH}) voltage values. Hysteresis refers

to the dependency of the current output on the previous output [18]. Fig. 7.4(a) shows an example of the input and output waveforms of a non-inverting ST in the transient domain, on top, and its voltage transfer characteristic (VTC), on the bottom. The hysteresis window (HW) is the region in which the current output depends on the previous output. The width of this window is given by $HW = V_{TH} - V_{TL}$. Therefore, varying the threshold voltages varies the width of the HW . This work considers the following ST topologies: (i) internal feedback STs based on inverters, and (ii) external feedback STs based on operational amplifiers [18, 146, 147]. Note that although this work discusses only the non-inverting ST configuration, an inverting ST configuration can also be used.

7.2.1.1 Internal-feedback Schmitt Trigger (ST)

The CMOS 6T-ST circuit is based on six transistors ($M_1 - M_6$) and internal feedback, as shown in Fig. 7.4(b). The transistor sizes and the technology parameters define the threshold voltages of the ST. An additional transistor pair (M_7, M_8) with the corresponding control voltages (V_{C1} and V_{C2}) allows changing the width of the HW [146, 148].

7.2.1.2 External-feedback Schmitt Trigger (ST)

A non-inverting ST can be implemented with a high gain amplifier and external positive feedback realized by the programmable resistors R_1 and R_2 , as shown in Fig. 7.4(c) [18]. The ST's thresholds voltages can be written as

$$V_{TL,TH} = \frac{V_{REF}(R_1 + R_2) - (R_1 \times V_{OH,OL})}{R_2} \quad (7.1)$$

where V_{REF} is a reference voltage applied to the inverting input terminal of the amplifier. The amplifier's output swing defines the values of V_{OL} and V_{OH} [18].

The implementation of the programmable resistors varies depending on whether the controlling input is digital or analog, as shown in Fig. 7.4(c). In the former case, an array of T resistors are connected via switches. These switches are controlled by the digital input $\{b_0, b_1, b_2, \dots, b_{T-1}\}$, which in turn determines the equivalent resistance. Similarly, a T-network formed by R_Y and M_Z implements an analog programmable resistor, as illustrated in Fig. 7.4(c). The effective resistance

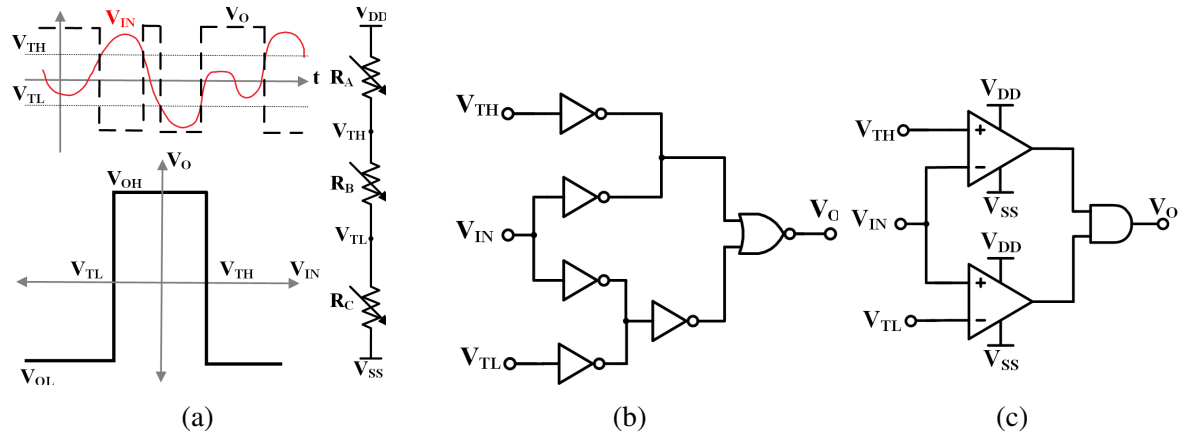


Figure 7.5: (a) Voltage response of a window comparator (top) transient response: if $V_{TL} < V_{IN} < V_{TH}$ the output is V_{OH} , otherwise V_{OL} . (bottom) voltage transfer characteristic. The low and high thresholds are defined by the configurable resistors R_A , R_B , and R_C . (b) Inverter-based window comparator. (c) OpAmp-based window comparator.

of the T-network is a function of R_Y and the on-resistance of M_z , controlled by the voltage V_C [18].

7.2.2 Output transition probabilities of the non-inverting Schmitt trigger (ST)

The comparison of the input voltage with the threshold voltages that leads to the output being low or high defines the output transition probabilities. As shown in Fig. 7.4(a), p_L is the probability of $V_{IN} < V_{TL}$ thus, $V_{OUT} = V_{OL}$ and p_H is the probability of $V_{IN} > V_{TH}$ thus, $V_{OUT} = V_{OH}$. p_W is the probability of $V_{TL} < V_{IN} < V_{TH}$, where the output voltage retains the previous value. The V_{TL} and V_{TH} are configurable via the resistor settings, as illustrated in the previous section. Hence, it is possible to change the output transition probabilities by modifying the threshold voltages, i.e., the width of the HW . This work leverages the varying output transition probabilities for increasing the security of the generated CK.

7.2.3 Window comparator

Similar to the ST, in the window comparator, the output voltage is determined by the comparison of V_{IN} with V_{TL} and V_{TH} . However, as shown in Fig. 7.5(a), the output voltage equals V_{OH} if the input voltage lies between the thresholds, i.e., $V_{TL} < V_{IN} < V_{TH}$. Otherwise, the output voltage equals V_{OL} . As shown in Fig. 7.5(a), a voltage divider formed by the programmable resistors

R_A , R_B , and R_C generates the required threshold voltages V_{TL} and V_{TH} . Equations (7.2) and (7.3) give the relationship between the threshold voltages and the resistors.

$$V_{TL} = (V_{OH} - V_{OL}) \times \frac{R_C}{R_A + R_B + R_C} \quad (7.2)$$

$$V_{TH} = (V_{OH} - V_{OL}) \times \frac{R_C + R_B}{R_A + R_B + R_C} \quad (7.3)$$

Similar to the ST circuit, we discuss two possible implementations of the window comparator. While Fig. 7.5(b) shows a window comparator built from logic gates, Fig. 7.5(c) shows an implementation based on amplifiers.

7.2.3.1 Inverter-based window comparator

A window comparator compatible with the internal feedback ST is shown in Fig. 7.5(b). This comparator is based on digital gates and has a transistor count of 14 [149]. Since the voltage divider in Fig. 7.5(a) sets the threshold voltages, the technology's standard gates can be used on this implementation.

7.2.3.2 OpAmp-based window comparator

The amplifier-based window comparator uses two high gain amplifiers as level detectors whose outputs are sent to the AND gate to produce the final output V_O , as shown in Fig. 7.4(c) [18]. Thus, only when the outputs of the two amplifiers are high, the output of the AND gate is high as well.

7.3 Proposed approach

7.3.1 Threat model

Our threat model is identical to the one considered by analog IP protection techniques [5, 6, 8, 119]. The attacker can be in the foundry or can be an end-user. The attacker in an untrusted foundry has access to minimal resources sufficient to overproduce the chip and sell the excess chips in the black market. However, he/she cannot modify the existing layout or perform internal probing. The attacker can gain access to:

1. The layout or design masks from the untrusted foundry.
2. The process design kit (PDK) details from the foundry.
3. The netlist of the circuit extracted through reverse engineering techniques.
4. A functional chip, which has the key loaded, *a.k.a.*, oracle.
5. The circuit specifications of the chip from the datasheet.

7.3.2 Schmitt trigger (ST)-based key provisioning

For a given analog input voltage and a *HW* configuration, the ST generates the corresponding binary output. As described in Section 7.2.1, the output bit is a function of the input voltage, the lower and upper threshold voltages, and the previous output. When a sequence of n input voltages is applied to the ST, it delivers a series of n 1-bit outputs. These outputs are concatenated to form an n -bit output.

The proposed approach uses this operation for key provisioning. While the UK defines the thresholds and the input values, the generated digital output corresponds to the CK. The CK controls the locked circuit. To increase the attack effort, we use variable *HW* settings, and to achieve uniform distribution of the CK, we use positive and negative STs.

Hysteresis window (*HW*) settings. A fixed configuration of the *HW* leads to a weak defense approach. It is because the values of the thresholds can be obtained by applying increasing and decreasing input voltage sweeps and observing the corresponding output transitions. Therefore, we propose to have a dynamic hysteresis configuration. Depending on the chosen ST topology, the width of the *HW* can be changed by varying the input control voltages V_{C1} and V_{C2} , as illustrated in Fig. 7.4(b), or by tuning the resistors R_1 and R_2 for the topology shown in Fig. 7.4(c). Some particular settings of the *HW* configuration are permanently written before to ensure the uniqueness of the UK for each chip instance. Those settings are stored on-chip fuses written by the holder of the IP rights.

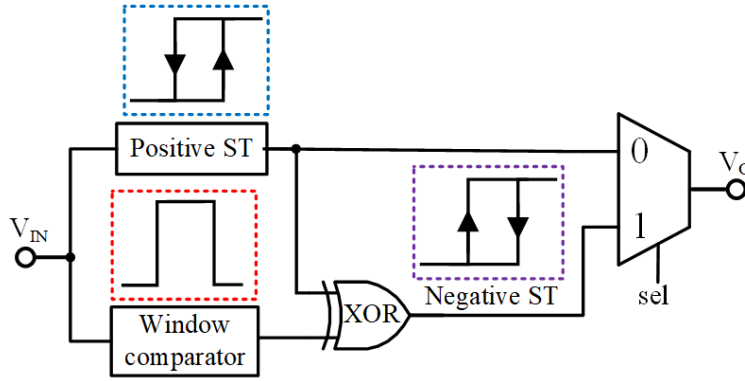


Figure 7.6: A negative hysteresis ST and its voltage transfer characteristics. This is built from a conventional positive (non-inverting) ST and a window comparator.

Positive and negative STs. The VTC of the non-inverting ST, *a.k.a.*, positive ST, is illustrated in Fig. 7.4(a). Its transition probabilities prevent the ST's output from having a uniform distribution. To compensate for that, we introduce the negative hysteresis ST. A negative ST has the values of lower and higher threshold voltages interchanged. Hence, when the input voltage is within the HW , the current output is equal to the previous output inverted. In consequence, the transition probabilities of the negative ST are complementary to the positive ST.

The output response of the negative ST is achieved by XORing the responses of the positive ST and the window comparator, configured for the same threshold voltages, as illustrated in Fig. 7.6. A multiplexor (MUX) selects between the response of the positive or negative ST. A one-bit *sel* controls the select line of this MUX. It is necessary to set the *sel* to 0 and 1 with equal probability to ensure that the output response has a distribution closer to a uniform one.

Fig. 7.7 illustrates the proposed ST-based key-provisioning. The UK and the CK can take either digital or analog values. The UK is divided into x segments that are applied in series to reuse the same circuitry. Each segment consists of three parts: (i) w bits (or analog voltage values) to configure the width of the HW , (ii) n input values consisting of a m -bit (or an analog voltage) each, and (iii) a one-bit *sel* (or a single voltage) that selects between the positive and the negative STs, as shown in Fig. 7.7.

The operation of the proposed key provisioning is as follows. In each segment, the w bits (or

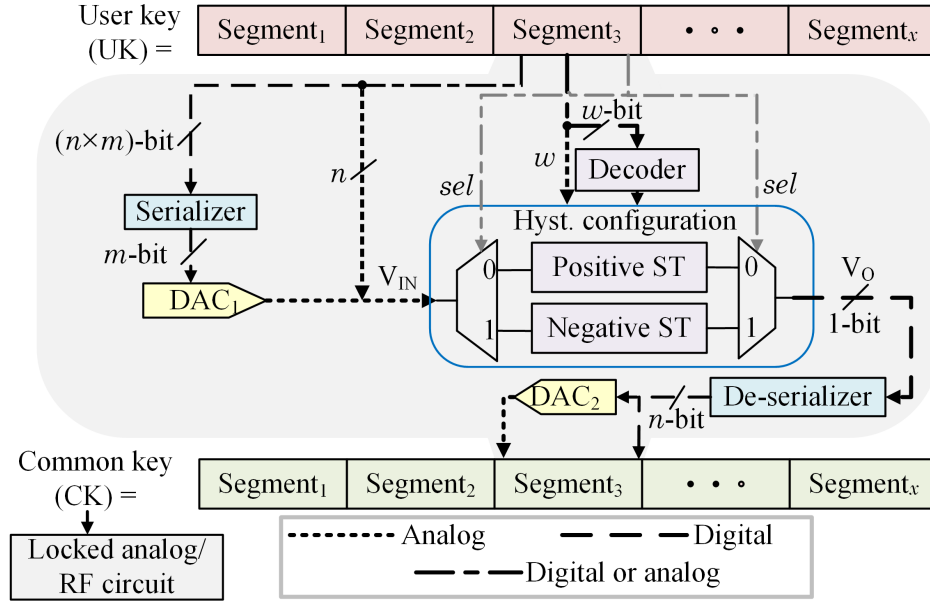


Figure 7.7: Proposed approach. The UK consists of x segments that reuse the same circuitry. Each segment selects between a positive ST or negative ST, configures the threshold voltages, and also provides the input voltages to generate the required CK. DAC_1 is needed for digital UKs, but not for analog UKs. DAC_2 is needed for analog CKs, but not for digital CKs.

the analog voltage values) configure the programmable resistors that set V_{TH} and V_{TL} , defining the width of the HW . For a digital UK, a decoder is required to generate the control bits of the digitally programmable resistors. Otherwise, an analog UK includes the control voltages for the T-network resistors implementation, as shown in Fig. 7.4(c). Also, when the UK is digital, a serializer receives $n \times m$ bits and delivers at the output m -bits at the time. Each m -bits are fed to the digital-to-analog converter (DAC_1) to generate an analog input voltage V_{IN} . Then, V_{IN} is applied to the ST to produce a single-bit output V_O . This process is repeated n times per segment, keeping w fixed, and generating a n -bit segment of the CK. The delays of the serializer, the DAC, the STs and the MUX define the total time required to generate a single bit of the CK.

Finally, this process is repeated for each segment to produce the CK of size $x \times n$ bits. Thus, the size of the search space of the CK is $2^{x \times n}$. The probability of each output outcome is equal to the product of the transition probabilities of all the input values applied serially. Consequently, by dedicating a key-bit per segment sel to select between positive or the negative ST randomly,

the distribution of CK becomes closer to a uniform distribution. Although the output key is in the digital domain, the key provisioning can also generate analog CKs by including the DAC_2 shown in Fig. 7.7.

7.3.3 Security metrics

This section discusses the security metrics when both the UK and the CK are digital.

7.3.3.1 Key size

The UK size is dependent on x , w , n , and m , which are the number of segments of the UK, the number of bits configuring the HW , the number of inputs per segment, and the number of bits representing each input, respectively. As illustrated in Fig. 7.7, the size of the UK is given by $x \times (w + (n \times m) + 1)$. Hence, the possible values of the UK are $2^{x \times (w + (n \times m) + 1)}$. As explained in Section 7.3.2, the CK is the concatenation of the output responses of ST from all the segments. Hence, the size of the CK is $x \times n$. The probability of attaining a specific CK is equivalent to the product of the output transition probabilities of the individual segments explained in Section 7.2.2.

Unlike other key provisioning techniques where the circuit size increases as the UK's size increases, our proposed approach does not incur any extra area overhead as the UK's size increases. Any desired size of the UK can be achieved by increasing x or n per x . However, increasing the UK size impacts the chip activation time, i.e., the time taken to generate the CK once the chip is turned on. This time delay is considered non-critical as it is a one-time delay at power-up.

7.3.3.2 Resiliency against brute force attacks

In a brute force attack, the attacker obtains an unlocked chip and use it as an oracle to find the correct UK. He/She explores the whole search space by trial and error or using advanced techniques like optimization algorithms. The robustness of defense approaches to brute force attacks is related to the effort required to find the correct key. Hence, a defense technique is provable secure if the key size is such that the effort required to break it is impractical. A UK size of 80 bits is considered unbreakable through a brute force attack [121]. Therefore, the UK can be designed such that $x(w + (n \times m) + 1) \geq 80$.

7.3.3.3 Resiliency against reverse engineering attacks

As specified in our threat model described in Section 7.3.1, the attacker can obtain the reverse-engineered netlist of the chip using services such as Chipworks [150]. The extracted netlist includes the key provisioning unit and the locked analog circuit. Even if the attacker can predict the CK from the locked analog circuit using mathematical formulations [3], he/she cannot predict the UK due to:

- The CK ports are not controllable and observable, i.e., the attacker cannot perform internal probing on the CK ports. The only way to control the CK is via the UK.
- The attacker can neither remove nor bypass the key provisioning unit to obtain direct access to the CK ports. As this work assumes resilience only against overproduction, the attacker has the resource only to overproduce the netlist but cannot perform any modifications to the existing netlist.
- Some of the bits configuring the *HW* are set permanently using on-chip fuses. This information cannot be obtained through reverse engineering techniques.

7.3.3.4 Resiliency against SAT/SMT attacks

The satisfiability-based (SAT) attack is based on Boolean logic. Although the output of the ST is Boolean, its input is a continuous analog voltage. Therefore, a SAT attack cannot be formulated on the proposed key-provisioning circuit. However, satisfiability modulo theories (SMT) can handle non-Boolean variables. In [3], the SMT-based attack was demonstrated successful on breaking most of the existing analog defense techniques. Although, equations of the working of the ST can be easily formulated, the *HW* configuration bits stored in on-chip fuses are not available to the attacker preventing him/her from formulating the SMT constraints.

7.3.3.5 Probability distribution of the common key (CK)

The effort of finding the correct UK increases as the distribution of the CK approaches a uniform distribution. We quantify how close is the distribution of the CK, provided by the proposed

approach, to a uniform one. Also, we analyze the effect of having a non-uniform distribution in the security level. There are several metrics available to measure the randomness of the generated CK. We use the entropy E as a metric of key unpredictability [118]. The entropy is calculated using Equation (7.4), where P_i is the probability of each outcome i or value of the CK.

$$E(CK) = \sum_i P_i \cdot \log_2 \left(\frac{1}{P_i} \right) \quad (7.4)$$

If the probability of a single-bit taking the value ‘0’ or ‘1’ is equal, $P(0)=P(1)=0.5$, the entropy equals the number of bits, one in this case. Otherwise, if $P(0) \neq P(1)$, the bit does not have a random distribution, and the entropy is smaller than one-bit. Hence, the entropy also determines the effective key size.

7.3.3.6 Uniqueness of the user key (UK)

In the case an attacker manages to find the correct UK of a chip, he/she should not be able to unlock other chip instances using this UK. Hence, each chip should have a unique UK value. This security property imposes two restrictions:

- The correct UKs of two different chips producing the same CK should be statistically different.
- The correct UK of one chip should not activate another chip.

To address the first restriction, we use the Hamming distance metric. The Hamming distance between two binary numbers is defined as the number of bit positions at which their values differ. The key provisioning unit should be designed such that several UKs that produce the same CK have a Hamming distance equivalent to 50% of the UK’s size.

The second restriction is met by hardcoding certain bits of the HW configuration using on-chip fuses. These fuses are written by the IP owner in a trusted facility after fabrication. These fuse settings ensure a unique configuration of the threshold voltages for each chip. Thus, two key provisioning units having different HW configurations generate different CKs for the same UK. Moreover, to generate the same CK with different HW configurations, the UKs must be different.

7.4 Results and discussion

7.4.1 Experimental setup

The transistor-level circuit simulations of our proposed ST-based key provisioning technique are performed using the Spectre[®] simulation platform. This technique is implemented using the IBM 180 nm CMOS process with a 1.8 V supply. Each DAC is built using the R-2R DAC topology [151]. Also, integrated polysilicon resistors realize the programmable resistors. The serializer/de-serializer is coded using Verilog HDL and is synthesized using the chosen CMOS process. The security metrics, such as the uniqueness of the UK and the effective CK's size, are determined from the behavioral model of the key provisioning unit implemented using MatLab[®].

7.4.2 Effective size of the common key (CK)

The following experiment calculates the entropy of the CK generated by the proposed technique in response to a single UK's segment ($x = 1$). This experiment is repeated for different combinations of w , m , and n . The resolution of the threshold voltages and the input voltage depends on w and m , respectively. We evaluate the outputs of all the possible combinations of HW settings, input voltage values, and the type of the ST (positive or negative hysteresis). Then, we calculate P_i , the probability of each output value i .

The entropy of the CK is calculated using Equation (7.4). Table 7.2 lists the entropy for different combinations of w , m , and n . If the CK generated by the key provisioning unit has a uniform distribution, the calculated entropy equals n , which is the number of inputs applied sequentially. Hence, the closer the value of entropy approaches n , the closer is the distribution of CK to the uniform distribution. The entropy thus quantifies the effective number of information bits of the CK.

From the results in Table 7.2, the effective key size (entropy) is smaller than the actual key size (n). However, the degradation in the effective key size is less than one bit. This information is useful for designing the ST-based key provisioning unit. For example, consider $w = m = 4$, and the desired number of bits of CK is 4. Selecting $n = 4$ translates to only 3.87 bits of CK that is insufficient. Therefore, $n = 5$ is chosen to achieve the desired level of security.

w	2			3			4		
$m \backslash n$	3	4	5	3	4	5	3	4	5
2	2.97	3.93	4.87	2.97	3.93	4.87	2.97	3.93	4.87
3	2.96	3.89	4.81	2.95	3.87	4.77	2.95	3.89	4.80
4	2.93	3.84	4.72	2.95	3.88	4.78	2.95	3.87	4.77

Table 7.2: The entropy of the CK for all combinations of w , n , and m . w bits set the width of the HW . n is the number of m -bit input values applied in series.

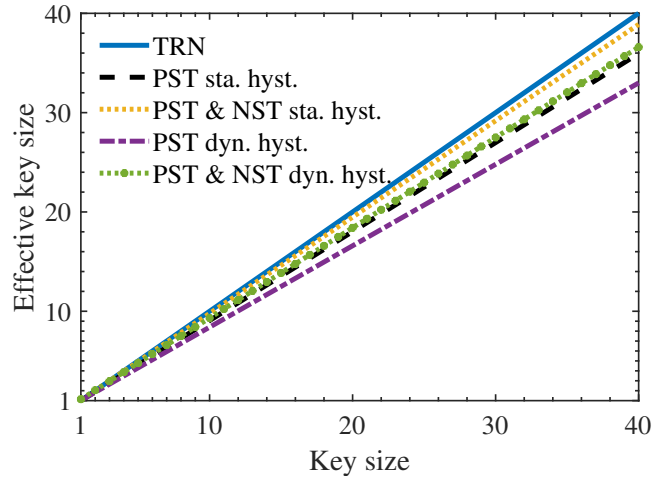


Figure 7.8: Effective key size of the CK generated by the proposed key provisioning compared to a true number (TRN). The effective CK's size is calculated when the hysteresis window (HW) is static or dynamic and using a positive ST (PST) alone or combined with a negative ST (NST).

We extrapolated the results of this experiment to estimate the effective key size of larger CKs. Fig. 7.8 shows the discrepancy between the effective size of a key generated with a TRNG and the proposed technique under different configurations. We estimated the effective key size for static or dynamic HW , and only positive ST (PST) or positive and negative ST (PST & NST). In all cases, the HW is centered at the middle of the supply voltage. The CK's entropy is highly dependent on the threshold values. In a static configuration, the smaller the HW , the larger the entropy. The effective key size of the dynamic hysteresis configuration considers all possible hysteresis widths. For instance, Fig. 7.8 shows the effective key calculated for a static configuration with a 0.4V HW

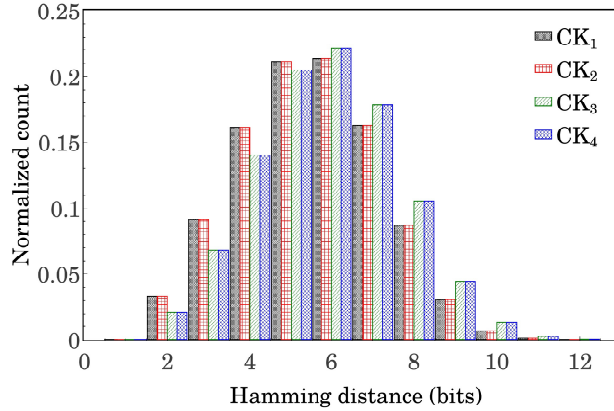


Figure 7.9: Uniqueness of the user key. The hamming distance of all user keys is calculated for different common keys CK_1 , CK_2 , CK_3 , and CK_4 .

versus a dynamic window with $w = m = 4$. Although a static configuration can have higher entropy than a dynamic one, it is a weak approach since it reduces the attacker’s effort to find the correct key. Moreover, having both PST and NST increases the effective key size compared with using a PST alone, at the expense of extra area. This experiment demonstrates that a dynamic HW with PST & NST is a good design for increasing the security level.

7.4.3 Uniqueness of the user key (UK)

We calculate the Hamming distances between each UK and the other UKs that generate the same CK to determine their uniqueness. For instance, in a ST-based key provisioning unit with $x = 1$, $w = 2$, $m = 2$, and $n = 5$, the size of UK and CK equals 13 bits and 5 bits, respectively. Due to the probability distribution of the CK, there are $\approx 2^{13}/2^5$ distinct UKs that produce the same CK. We calculate the Hamming distance between all the possible values of UKs generating the same CK. The experiment is repeated for all the values of CK. Fig. 7.9 shows a histogram of the Hamming distances between the UKs for four different CKs. This plot follows a Gaussian response, where the mean Hamming distance between UKs is equal to 6 bits, i.e., approximately equal to 50% of the UK’s size (13 bits).

User key	Common key	Internal feedback ST [mm ²]	External feedback ST [mm ²]
Digital	Digital	0.010	0.014
Digital	Analog	0.013	0.017
Analog	Digital	0.001	0.003
Analog	Analog	0.004	0.005

Table 7.3: Area overhead of the ST-based key provisioning implementation for keys in the digital or the analog domain.

7.4.4 Power and area overhead

The ST-based key provisioning unit consumes power only during a short period at power-up. During this time, the circuit acquires the UK and generates the corresponding CK. Then, the CK is stored in a shift register, applied to the locked circuit, and the key-provisioning unit is powered down. Therefore, **there is no power consumption during runtime** [8].

However, the key provisioning circuit is integrated on-chip and incurs an area overhead. The area overhead is calculated for different variations of the proposed approach, depending on the ST's circuit topology and whether it uses analog or digital keys. The positive and negative STs can be implemented either with internal or external feedback (Fig. 7.4 and Fig. 7.5), yielding different area overheads.

As the proposed approach is compatible with both digital and analog keys, the area overhead of each configuration accounts for the implementation of all required circuit blocks, according to Fig. 7.7. Hence, the integrated area in Table 7.3 includes the positive ST, the window comparator, the MUX for selection, the programmable resistors, the decoder (5-bit decoder for digital UKs only), the serializer, DAC_1 (5-bit DAC for digital UKs only), the de-serializer, and DAC_2 (5-bit DAC for analog CKs only). All this circuitry is required to process one key segment at the time. Hence, the circuit is reused for the x segments that form the UK. Moreover, increasing the UK size can be done by increasing x without any change in the circuit implementation.

We also compare the area overhead of our approach with other key-provisioning techniques

UK size [bits]	EPIC [12] [mm ²]	PUF ID XOR UK [5][mm ²]	Analog NN [121] [mm ²]	This work [mm ²]
80	0.282	0.017	>100	0.014
128	0.282	0.027	>100	0.014
256	0.282	0.054	>100	0.014

Table 7.4: Area overhead comparison with other techniques.

when both the UK and the CK are digital. Table 7.4 summarizes the comparison. The area of the previous works is not reported for these key sizes but estimated from their reported results [5, 8, 12, 121]. In [12], generating a 64-bit key incurs in an area overhead given by a TRNG and the RSA core implementation [12]. While the integrated area of the TRNG is 0.036mm² in a 130 nm process, the RSA requires around 10,000 two-input gates. These numbers were scaled to the 180 nm process for comparison. In [5], the circuit overhead is given by the PUF and digital circuitry. Its area was estimated from the reported results of three different circuits implemented in the 180 nm process, with different key sizes.

In [121], the area of the neural-network-based key provisioning is not reported. However, it can be estimated from the picture of the experimental setup. AMSlock [8] is not included in the comparison because its operation differs from a key provisioning technique.

From the comparison in Table 7.4, we observe that the proposed approach has the best area efficiency than all the other techniques for all the key sizes. In contrast with the PUF-based key generation, in the proposed approach, the area efficiency increases with an increase in the key size.

Another aspect of the overhead is the execution time. The time required for the generation of each CK's segment includes the configuration time t_1 and the evaluation time t_2 . The HW is set during t_1 . During t_2 , the ST receives a sequence of n inputs and generates the corresponding outputs. Hence, the total time t_t is a product of the time per segment and the number of segments $t_t = x \times (t_1 + t_2)$. On average, it takes $t_t=1.8\mu s$ to produce an 80-bit CK.

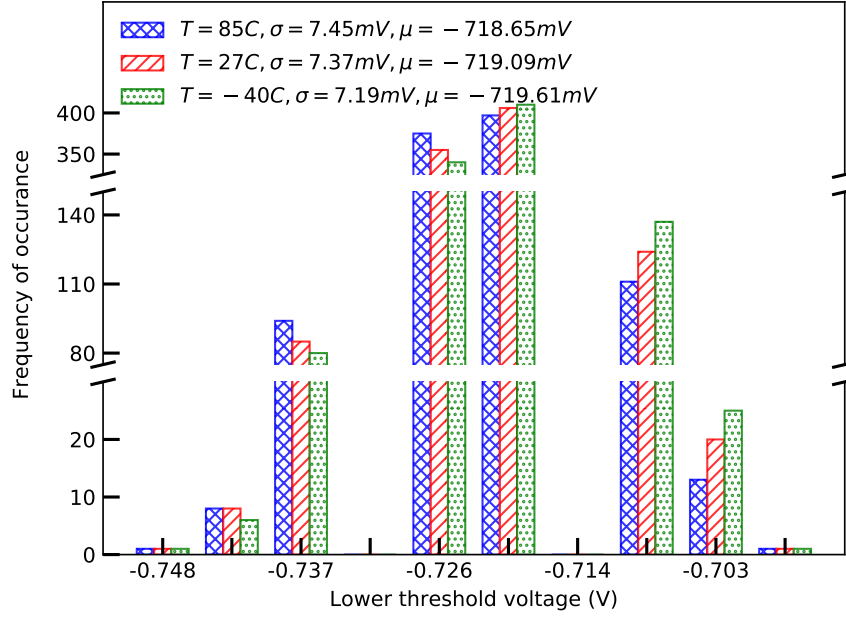


Figure 7.10: Worst case variation in threshold voltage due to process and temperature variations. The standard deviation σ and mean variance μ are given for each temperature value.

7.4.5 Robustness against process and temperature variations

Since the HW is represented by w bits, the threshold values are not continuous but discrete. The resolution step is given by $(V_{DD} - V_{SS})/(2^w)$. On top of that value, process and temperature variations can modify the threshold values. The proposed approach is considered robust to variations if the deviation caused by them is small compared with the resolution step of the thresholds.

A 1000-sample Monte Carlo simulation was performed to estimate the variation in the threshold voltages due to process and temperature variations. As shown in Fig. 7.10, V_{TH} variations are smaller than 3×7.455 mV based on a 1000-sample Monte-Carlo simulation results from -40° C to 85° C. As the output voltage saturates to either higher or lower supply voltage, the output of the ST is insensitive to voltage variations. Hence, it does not incur any performance degradation.

We also evaluated the impact of variations on the distribution of the CK. Fig. 7.11. The entropy was calculated for various combinations of w , m , and n similar to the results reported on Table 7.2. However, in this experiment, both thresholds have an additional $\pm\Delta$ error. The results demonstrate

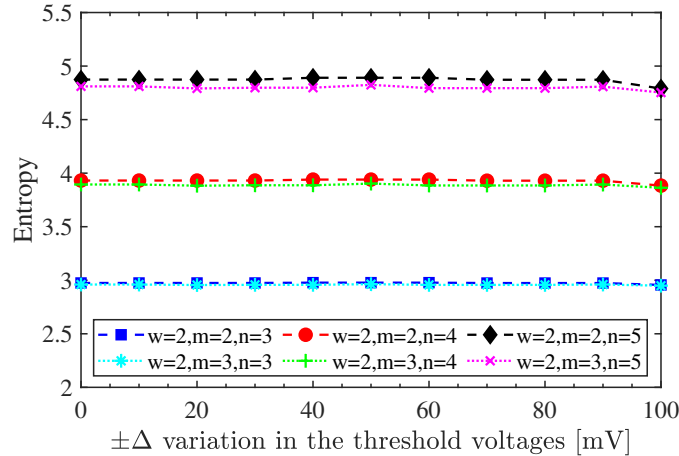


Figure 7.11: Estimated variation on the common key's entropy due to a $\pm\Delta$ voltage variation in V_{TL} and V_{TH} . The results of six experiments with different values of the number of bits describing the hysteresis window w , the input values m , and the number of inputs n are reported.

a worst-case degradation of 0.1-bit for as much as a 100 mV error in the threshold voltages. It is around $5X$ the $\pm 3\sigma$ variation estimated due to process and temperature changes.

7.4.6 Test cases with analog locks

The proposed key provisioning approach is demonstrated on three different locked analog/RF IC designs. They represent three broad areas of application. An active filter, used in signal processing [19]; a low-dropout (LDO) voltage regulator used in power management [152]; and a low noise amplifier, a fundamental block in RF receivers [51]. All the circuits were implemented using the IBM 180 nm process and powered by a 1.8 V supply. We simulated each circuit's performance when applied: i) the correct UK, and ii) several incorrect UKs.

7.4.6.1 Bandpass fourth-order Gm-C filter

A 4^{th} order Gm-C filter is implemented as a cascade of two 2^{nd} order filters. The circuit schematic of a 2^{th} order Gm-C filter is presented in Fig. 7.12. Its transfer function is

$$H_{BPF} = \frac{V_{out}}{V_{in}} = \frac{g_{m1}C_1s}{s^2C_1C_2 + sg_{m3}C_2 + g_{m2}g_{m4}}. \quad (7.5)$$

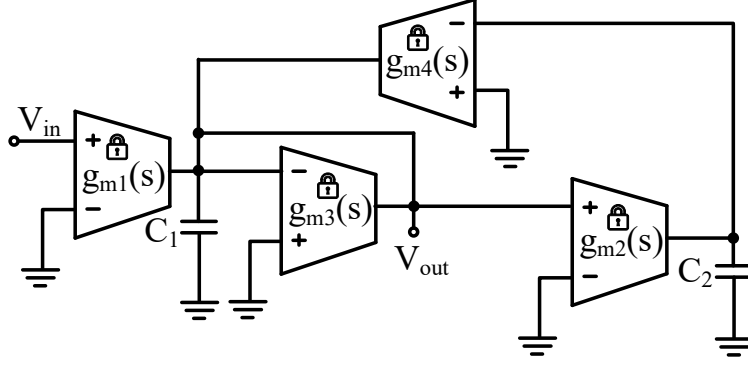


Figure 7.12: Second order Gm-C bandpass filter.

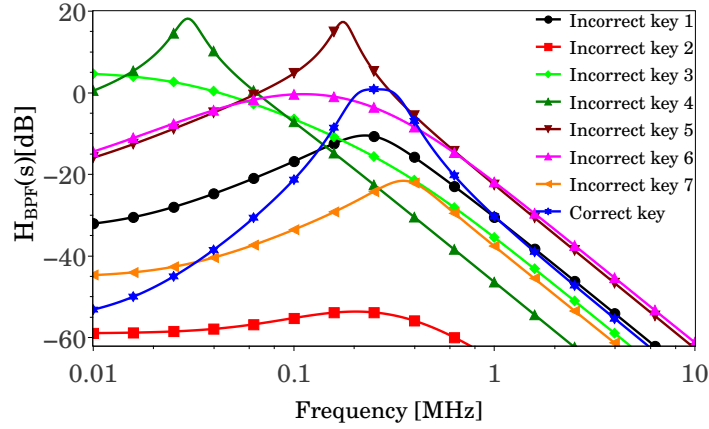


Figure 7.13: Transfer function of 4th order Gm-C filter for the correct and incorrect keys. The specifications of the filter are met only for the correct key.

The performance metrics of the filter are center frequency $\omega_o = \sqrt{\frac{g_{m2}g_{m4}}{C_1C_2}}$, quality factor $Q = \frac{1}{g_{m3}} \sqrt{\frac{g_{m2}g_{m4}C_1}{C_2}}$, gain of the passband $H(j\omega_o) = \frac{g_{m1}C_1}{g_{m3}C_2}$, and bandwidth $BW = (\omega_o/Q)$ [19]. The bias current to each OTA is provided by a non-monotonic, non-concave configurable current mirror (CCM) based lock [5]. Each CCM-lock is controlled by 12 bits. Hence, the total size of the CK is 96 bits. The ST-based key provisioning block is designed with $x = 12$, $n = 8$, $w = 5$ and $m = 5$ to produce the CK of size 96 bits. Also, the size of the UK is 552 bits. As illustrated in Fig. 7.13, the correct key sets the performance metrics of the filter equal to the desired values, i.e., $f_o = \omega_o/(2\pi) = 268$ kHz, $BW = 154$ kHz, and $H(j\omega_o) = 0$ dB. For the incorrect keys, as shown in the figure, the circuit specification is not met.

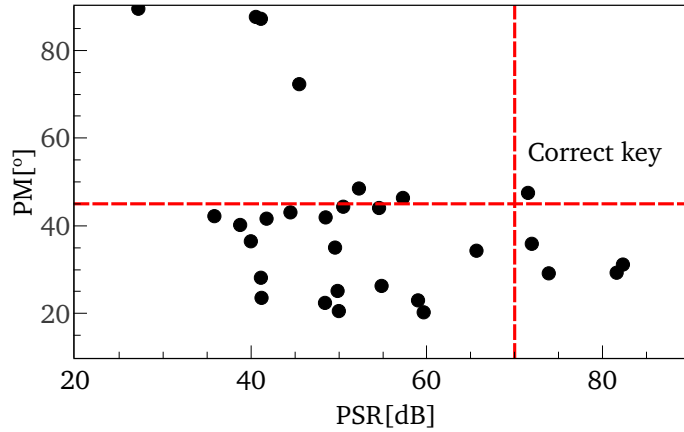


Figure 7.14: Applying the correct key to the locked LDO, gives the desired performance $PM > 45^\circ$ and $PSR > 70$ dB, whereas an incorrect key gives undesired performance.

7.4.6.2 Low-dropout (LDO) voltage regulator

A capacitor-less LDO voltage regulator with a single-stage error amplifier [152] is locked with an 18-bit CK. The key controls the biasing of the error amplifier by CCM-based lock and a configurable capacitor bank. The LDO is designed for an input voltage $V_{in}=1.8$ V and an output voltage $V_{out}=1.6$ V, resulting in a dropout voltage of 200 mV. This LDO is designed to provide a stable output voltage under the load conditions $I_L=(100 \mu\text{A}, 20 \text{ mA})$, with a load capacitor $C_L=1$ nF.

The 18-bit CK is produced by the ST-based key provisioning with the following configuration, $w = 5$, $m = 5$, $n = 3$, and $x = 6$. This lock secures two fundamental performance metrics of LDO: the phase margin $PM > 45^\circ$ and the power supply rejection $PSR(@1\text{KHz}) > 70$ dB. As illustrated in Fig. 7.14 the LDO's performance metrics meet the specifications across the given current load range for the correct key. However, for an incorrect key, the measured PM and PSR do not meet the design specifications.

7.4.6.3 Cascode common-gate low-noise amplifier (CG-LNA)

The cascode CG-LNA is a popular LNA topology [51]. In this circuit, a 24-bit CK controls the CCM providing the bias current of the gain transistor and the configurable tank load. The ST-based key provisioning block is configured with $w = 5$, $m = 5$, $n = 6$, and $x = 4$. Thus, the size of the UK is 184 bits. The circuit specifications of the secured cascode CG-LNA are input matching

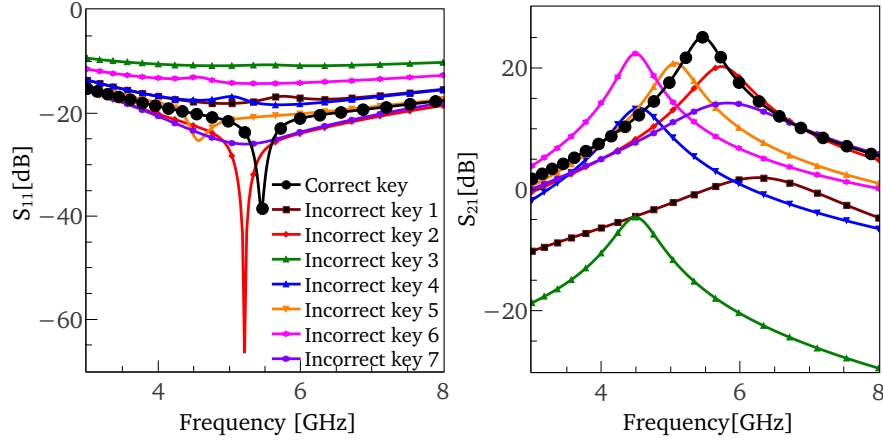


Figure 7.15: $S_{11} \leq -30$ dB and $S_{21} > 25$ dB when correct key is applied to the locked LNA. Otherwise, for an incorrect key, the S-parameters do not satisfy the specifications.

$S_{11} < -30$ dB and gain $S_{21} > 25$ dB. Fig. 7.15 shows the impact of the correct and incorrect UKs on the performance of the CG-LNA. Only the correct key gives the desired performance of $S_{11} = -39$ dB and $S_{21} = 26$ dB.

7.4.6.4 Overhead of the key provisioning on the test cases

The area overhead and the energy consumption of the proposed ST key are reported in Table 7.5. The area overhead is compared with the original area of the locked analog circuit, and it includes the elements of Fig. 7.7(d): serializer, DAC, ST, window comparator, MUX, decoder, programmable resistors, and de-serializer.

From Table 7.5, we observe that for analog/RF circuits with an area larger than 0.5 mm^2 , the

Circuit under test	UK size [bits]	CK size [bits]	Original Area [mm^2]	Area overhead		Energy [nJ]
				Area [mm^2]	Pct. %	
GM-C BPF	552	96	0.692	0.014	2.02	13.35
LDO	126	18	0.163	0.014	8.58	2.53
LNA	144	24	0.724	0.014	1.93	3.35

Table 7.5: The area overhead incurred by the proposed key provisioning unit on different locked analog circuits.

overhead is smaller than 3%. As expected, the area overhead is more considerable for smaller circuits like the LDO. However, the LDO is often integrated to provide a stable voltage to other circuits in a larger architecture. Hence, a locked LDO enables the IP protection of different circuit blocks by controlling their supply voltages and reducing the area overhead of our approach.

7.5 Conclusion

An ST-based key provisioning circuit has been designed and characterized for the security metrics considered. Our approach leverages a highly configurable circuit based on hysteresis comparators for a high resiliency to overproduction attacks. Increasing the sizes of both the CK and the UK is done by reusing the integrated circuitry. Hence, compared to the previous key provisioning techniques, the proposed techniques incur lesser area overhead. It takes approximately $\frac{1}{20.1}$ times the size of [12], half the size of [5], and $\frac{1}{7142}$ times the size of [121]. The proposed key provisioning only consumes power at the power-up time. Therefore, power overhead is not a concern for our approach. The chip activation time increases with the key size. The proposed method takes on average $1.8\mu s$ to acquire the UK and generate an 80-bit CK. This delay occurs during the power-up.

The IP rights holder designs the security metrics of this approach through the circuit settings. Experimental results demonstrate the efficacy of this approach on securing the performance of analog/RF circuits for both digital and analog keys. PVT variations do not affect the entropy of the generated key. Additional settings for the tuning of the center of the HW can be studied to increase the entropy of the CK further. Moreover, enabling a dynamic segment length could increase the resilience to brute force attacks.

8. CONCLUSION

In this work, we elaborated on several trends in analog circuit design. First, we discussed the application of tunable circuits that adapt their performance for a set of specifications across multiple standards. Chapters 2 and 3 summarize some basic concepts of continuous-time filters. Our contribution is presented in Chapter 4. We discussed the design considerations and silicon integration of a fifth-order low-pass active-R filter continuously-tunable in the frequency range from 1 MHz to 50 MHz. We revisited the active-R filter architecture and demonstrated its applications on baseband filtering for a multi-standard receiver. We presented the system-level and transistor-level design considerations of a frequency tunable CMOS amplifier used in the filter implementation. Compared with state-of-the-art active-RC and G_m -C filters, the proposed filter provides a wide tuning range, competitive dynamic range, and low area consumption.

Another relevant trend on analog IC design consists of using optimization algorithms for automatic sizing of active and passive devices. In Chapter 5 we described a novel surrogate model that combines circuit equations and characterization data to provide an accurate yet fast evaluation of the circuit performance. The surrogate model is inserted into either a gradient-based or a heuristics-based optimization algorithm for the automatic sizing of various analog circuits. Such an algorithm receives the circuit topology, the performance specifications, and the metric to minimize. After execution, it delivers the circuit design that complies with the specifications while minimizing the cost function.

Finally, this dissertation presents a contribution to a new area of IC design known as hardware security. It focuses on the protection of the intellectual property of analog circuits. Chapter 6 summarizes the state-of-the-art on threat models, defense, and attack mechanisms on analog hardware security. In Chapter 7 we present a novel key-provisioning technique that enables protecting chip instances with unique keys and extending the security level without a significant area and energy overheads. The Schmitt-trigger-based circuit provides a common-key equal for all the chip instances, from a user-key unique for each chip. We demonstrated this technique in the locking of

three different analog circuits. Furthermore, several security metrics were defined and characterized for the benchmark of key-provisioning techniques.

REFERENCES

- [1] P. Subramanyan, S. Ray, and S. Malik, “Evaluating the Security of Logic Encryption Algorithms,” *IEEE International Symposium on Hardware Oriented Security and Trust*, pp. 137–143, 2015.
- [2] M. Yasin, A. Sengupta, M. T. Nabeel, M. Ashraf, J. Rajendran, and O. Sinanoglu, “Provably-Secure Logic Locking: From Theory To Practice,” *ACM SIGSAC Conference on Computer & Communications Security*, pp. 1601–1618, 2017.
- [3] N. G. Jayasankaran, A. Sanabria-Borbón, A. Abuellil, E. Sánchez-Sinencio, J. Hu, and J. Rajendran, “Breaking analog locking techniques,” *IEEE Transactions on Very Large Scale Integration Systems*, pp. 1–14, 2020.
- [4] R. Y. Acharya, S. Chowdhury, F. Ganji, and D. Forte, “Attack of the Genes: Finding Keys and Parameters of Locked Analog ICs Using Genetic Algorithm,” *arXiv.org*, 2020.
- [5] J. Wang, C. Shi, A. Sanabria-Borbon, E. Sánchez-Sinencio, and J. Hu, “Thwarting Analog IC Piracy via Combinational Locking,” *IEEE International Test Conference*, pp. 1–10, Oct. 2017.
- [6] V. V. Rao and I. Savidis, “Parameter Biasing Obfuscation for Analog IP Protection,” *IEEE Latin American Test Symposium*, pp. 1–6, 2017.
- [7] J. Leonhard, M.-M. Louërat, H. Aboushady, O. Sinanoglu, and H.-G. Stratigopoulos, “Mixed-Signal Hardware Security Using MixLock: Demonstration in an Audio Application,” *IEEE International Conference on Synthesis, Modeling, Analysis and Simulation Methods and Applications to Circuit Design*, 2019.
- [8] N. G. Jayasankaran, A. S. Borbon, E. Sanchez-Sinencio, J. Hu, and J. Rajendran, “Towards Provably-Secure Analog and Mixed-Signal Locking Against Overproduction,” *IEEE Transactions on Emerging Topics in Computing*, pp. 7:1–7:8, 2020.

- [9] K. Juretus, V. Venugopal Rao, and I. Savidis, “Securing Analog Mixed-Signal Integrated Circuits Through Shared Dependencies,” *ACM Great Lakes Symposium on VLSI*, pp. 483–488, 2019.
- [10] A. Ash-Saki and S. Ghosh, “How Multi-Threshold Designs Can Protect Analog IPs,” *IEEE International Conference on Computer Design*, pp. 464–471, 2018.
- [11] J. Leonhard, A. Sayed, M. Louërat, H. Aboushady, and H. Stratigopoulos, “Analog and Mixed-Signal IC Security Via Sizing Camouflaging,” *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, pp. 1–1, 2020.
- [12] J. A. Roy, F. Koushanfar, and I. L. Markov, “Ending Piracy of Integrated Circuits,” *IEEE Computer*, vol. 43, no. 10, pp. 30–38, 2010.
- [13] S. Lin, Y. Cao, X. Zhao, X. Wang, and X. Pan, “A Compact Ultra-low Power Physical Unclonable Function Based on Time-Domain Current Difference Measurement,” *IEEE International Symposium on Circuits and Systems*, pp. 277–280, 2016.
- [14] A. C. Sanabria-Borbón and E. Sánchez-Sinencio, “Synthesis of high-order continuously tunable low-pass active-r filters,” *IEEE Transactions on Circuits and Systems I: Regular Papers*, pp. 1–14, 2021.
- [15] A. C. Sanabria-Borbón, S. Soto-Aguilar, J. J. Estrada-López, D. Allaire, and E. Sánchez-Sinencio, “Gaussian-process-based surrogate for optimization-aided and process-variations-aware analog circuit design,” *Electronics*, vol. 9, no. 4, 2020.
- [16] A. Sanabria-Borbón, N. G. Jayasankaran, J. Hu, J. Rajendran, and E. Sánchez-Sinencio, “Analog/rf ip protection: Attack models, defense techniques, and challenges,” *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 68, no. 1, pp. 36–41, 2021.
- [17] A. Sanabria-Borbon, N. G. Jayasankaran, S. Lee, E. Sanchez-Sinencio, J. Hu, and J. Rajendran, “Schmitt Trigger-Based Key Provisioning for Locking Analog/RF Integrated Circuits,” *IEEE International Test Conference*, pp. 1–10, 2020.

- [18] S. Franco, *Design with Operational Amplifiers and Analog Integrated Circuits*. New York, NY, USA: McGraw-Hill, Inc., 4 ed., 2015.
- [19] R. Schaumann, H. Xiao, and V. V. Mac, *Design of Analog Filters*. New York, NY, USA: Oxford University Press, Inc., 2nd ed., 2009.
- [20] B. Razavi, *Design of Analog CMOS Integrated Circuits*. Irwin Electronics & Computer E, McGraw-Hill Education, 2016.
- [21] P. G. A. Jespers and B. Murmann, *Systematic Design of Analog CMOS Circuits: Using Pre-Computed Lookup Tables*. Cambridge University Press, 2017.
- [22] S. V. Thyagarajan, S. Pavan, and P. Sankar, "Active-RC Filters Using the Gm-Assisted OTA-RC Technique," *IEEE Journal of Solid-State Circuits*, vol. 46, pp. 1522–1533, Jul. 2011.
- [23] M. De Matteis, A. Pipino, F. Resta, A. Pezzotta, S. D'Amico, and A. Baschirotto, "A 63-dB DR 22.5-MHz 21.5-dBm IIP3 Fourth-Order FLFB Analog Filter," *IEEE Journal of Solid-State Circuits*, vol. 52, pp. 1977–1986, Jul. 2017.
- [24] V. Giannini, J. Craninckx, S. D'Amico, and A. Baschirotto, "Flexible baseband analog circuits for software-defined radio front-ends," *IEEE Journal of Solid-State Circuits*, vol. 42, pp. 1501–1512, Jul. 2007.
- [25] B. Vigham, J. Kuppambatti, and P. R. Kinget, "Switched-Mode Operational Amplifiers and Their Application to Continuous-Time Filters in Nanoscale CMOS," *IEEE Journal of Solid-State Circuits*, vol. 49, pp. 2758–2772, Dec. 2014.
- [26] H. Amir-Aslanzadeh, E. J. Pankratz, and E. Sanchez-Sinencio, "A 1-V +31 dBm IIP3, Reconfigurable, Continuously Tunable, Power-Adjustable Active-RC LPF," *IEEE Journal of Solid-State Circuits*, vol. 44, pp. 495–508, Feb. 2009.
- [27] F. Lavallo-Aviles and E. Sánchez-Sinencio, "A 0.6-V Power-Efficient Active-RC Analog Low-Pass Filter With Cutoff Frequency Selection," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, pp. 1–13, May 2020.

- [28] M. C. Schneider and C. Galup-Montoro, *CMOS Analog Design Using All-Region MOSFET Modeling*. Cambridge University Press, 2010.
- [29] L. Acosta, M. Jimenez, R. G. Carvajal, A. J. Lopez-Martin, and J. Ramirez-Angulo, "Highly Linear Tunable CMOS G_m - C Low-Pass Filter," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 56, pp. 2145–2158, Oct. 2009.
- [30] A. Tasic, W. A. Serdijin, L. E. Larson, and G. Setti, *Circuits and Systems for Future Generations of Wireless Communications*. 2009.
- [31] W. Sansen, *Continuous-time filters*, pp. 567–602. Boston, MA: Springer US, 2006.
- [32] E. Sanchez-Sinencio and J. Silva-Martinez, "CMOS transconductance amplifiers, architectures and active filters: a tutorial," *IEE Proceedings - Circuits, Devices and Systems*, vol. 147, no. 1, pp. 3–12, 2000.
- [33] J. Ramirez-Angulo and E. Sanchez-Sinencio, "Active compensation of operational transconductance amplifier filters using partial positive feedback," *IEEE Journal of Solid-State Circuits*, vol. 25, pp. 1024–1028, Aug. 1990.
- [34] S. K. Sanyal, U. C. Sarker, and R. Nandi, "A novel microprocessor-controlled active-R multifunction network: design of programmable filter, oscillator, and FSK/PSK wave generator," *IEEE Transactions on Circuits and Systems*, vol. 37, pp. 1085–1091, Sep. 1990.
- [35] A. Mitra and V. Aatre, "Low sensitivity high-frequency active R filters," *IEEE Transactions on Circuits and Systems*, vol. 23, pp. 670–676, Nov. 1976.
- [36] R. Schaumann, R. E. Chalstrom, and K. R. Laker, "Optimisation of sensitivity and dynamic range of i.f.l.f. active filters," *Electronics Letters*, vol. 13, pp. 367–368, Jun. 1977.
- [37] K. R. Rao and S. Srinivasan, "A bandpass filter using the operational amplifier pole," *IEEE Journal of Solid-State Circuits*, vol. 8, no. 3, pp. 245–246, 1973.

- [38] M. Soderstrand, "Active R ladders: High frequency, high order, low sensitivity active R filters without external capacitors," *IEEE Transactions on Circuits and Systems*, vol. 25, pp. 1032–1038, Dec. 1978.
- [39] A. Nedungadi and S. Venkateswaran, "Generalized second-order active R filters," *International Journal of Circuit Theory and Applications*, vol. 10, no. 4, pp. 311–322, 1982.
- [40] J. R. Brand and R. Schaumann, "Active R filters: review of theory and practice," *IEEE Journal on Electronic Circuits and Systems*, vol. 2, pp. 89–, Jul. 1978.
- [41] C. F. Ho and P. L. Chiu, "Realisation of active-R filters using amplifier pole," *Proceedings of the Institution of Electrical Engineers*, vol. 123, pp. 406–410, May 1976.
- [42] V. Kapustian, B. Bhattacharyya, and M. Swamy, "Frequency limitations of active-R filters using operational amplifiers," *Journal of the Franklin Institute*, vol. 308, no. 2, pp. 141 – 151, 1979.
- [43] A. C. Sanabria-Borbon and E. Sanchez-Sinencio, "Efficient use of gain-bandwidth product in active filters: Gm-C and Active-R alternatives," pp. 1–4, Feb. 2017.
- [44] M. T. Ahmed, M. A. Sibdiqi, and M. T. Javed, "A general synthesis technique for active-R networks," *International Journal of Electronics*, vol. 54, pp. 417–425, Jul. 1983.
- [45] 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.
- [46] P. Bhushan Mital and U. Kumar, "Effects of Non-Idealities of Op-Amps on Active Filters: an Analytical Study," *Active and Passive Electronic Components*, vol. 17, pp. 179–201, Jun. 1994.
- [47] J. Mahattanakul and J. Chutichatuporn, "Design procedure for two-stage CMOS Opamp with flexible noise-power balancing scheme," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 52, pp. 1508–1514, Aug. 2005.

- [48] V. Kapustian, B. Bhattacharyya, and M. Swamy, "Noise performance of active-R filters," *Journal of the Franklin Institute*, vol. 308, no. 2, pp. 153 – 162, 1979.
- [49] B. Brannon, "Where Zero-IF Wins : 50 % Smaller PCB Footprint at 1/3 the Cost," *Analog Dialogue*, no. Sep., pp. 1–7, 2016.
- [50] Y. Wang, L. Ye, H. Liao, R. Huang, and Y. Wang, "Highly Reconfigurable Analog Baseband for Multistandard Wireless Receivers in 65-nm CMOS," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 62, pp. 296–300, Mar. 2015.
- [51] B. Razavi, *RF Microelectronics*. Upper Saddle River, NJ, USA: Prentice Hall Press, 2nd ed., 2011.
- [52] 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.
- [53] T. Y. Lo, C. C. Hung, and M. Ismail, "A wide tuning range G_m -C filter for multi-mode CMOS direct-conversion wireless receivers," *IEEE Journal of Solid-State Circuits*, vol. 44, pp. 2515–2524, Sep. 2009.
- [54] T. H. Lee, *The Design of CMOS Radio-Frequency Integrated Circuits*. Cambridge: Cambridge University Press, 2 ed., Dec. 2003.
- [55] A. Kumar, *A wide dynamic range high-Q high-frequency bandpass filter with an automatic quality factor tuning scheme*. PhD thesis, Georgia Institute of Technology, 2009.
- [56] F. Capparelli and A. Liberatore, "Active bandpass network with only resistors as passive elements," *Electronics Letters*, vol. 8, pp. 43–44, Jan. 1972.
- [57] K. Laker and M. Ghauri, "Synthesis of a low-sensitivity multiloop feedback active RC filter," *IEEE Transactions on Circuits and Systems*, vol. 21, pp. 252–259, Mar. 1974.
- [58] D. J. Perry, "New multiple feedback active RC network," *Electronics Letters*, vol. 11, pp. 364–365, Aug. 1975.

- [59] A. Sedra and P. Brackett, *Filter Theory and Design: Active and Passive*. Matrix series in circuits and systems, Matrix Publishers, 1978.
- [60] P. Mohan, *VLSI Analog Filters Active RC, OTA-C, and SC*. Upper Saddle River, NJ, USA: Springer, 2013.
- [61] P. R. Gray, P. Hurst, S. H. Lewis, and R. G. Meyer, *Analysis and Design of Analog Integrated Circuits*. Wiley Publishing, 5th ed., 2009.
- [62] J. F. Duque-Carrillo, "Control of the common-mode component in CMOS continuous-time fully differential signal processing," *Analog Integrated Circuits and Signal Processing*, vol. 4, pp. 131–140, Sep. 1993.
- [63] M. Gambhir, V. Dhanasekaran, J. Silva-Martinez, and E. Sánchez-Sinencio, "Low-power architecture and circuit techniques for high-boost wide-band Gm-C filters," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 54, pp. 458–468, Mar. 2007.
- [64] M. De Matteis, S. D'Amico, and A. Baschirotto, "A 0.55 V 60 dB-DR Fourth-Order Analog Baseband Filter," *IEEE Journal of Solid-State Circuits*, vol. 44, pp. 2525–2534, Sep. 2009.
- [65] S. Pavan and Y. P. Tsvividis, "An analytical solution for a class of oscillators, and its application to filter tuning," *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, vol. 45, no. 5, pp. 547–556, 1998.
- [66] J. Silva-Martinez, M. S. J. Steyaert, and W. Sansen, "A 10.7-MHz 68-dB SNR CMOS continuous-time filter with on-chip automatic tuning," *IEEE Journal of Solid-State Circuits*, vol. 27, no. 12, pp. 1843–1853, 1992.
- [67] S. Kousai, M. Hamada, R. Ito, and T. Itakura, "A 19.7 MHz , Fifth-Order Active- RC Chebyshev LPF for Draft IEEE 802.11n With Automatic Quality-Factor Tuning Scheme," vol. 42, pp. 2326–2337, Nov. 2007.
- [68] J. Wang, C. Shi, E. Sanchez-Sinencio, and J. Hu, "Built-In Self Optimization for Variation Resilience of Analog Filters," *IEEE Computer Society Annual Symposium on VLSI*, pp. 656–661, 2015.

- [69] L. Scheffer, L. Lavagno, and G. Martin, *EDA for IC implementation, circuit design, and process technology*. edited by Louis Scheffer, Luciano Lavagno, Grant Martin. Electronic design automation for integrated circuits handbook, CRC Taylor & Francis, 2006.
- [70] G. G. E. Gielen and R. A. Rutenbar, “Computer-aided design of analog and mixed-signal integrated circuits,” *Proceedings of the IEEE*, vol. 88, no. 12, pp. 1825–1854, 2000.
- [71] M. Fakhfakh, Y. Cooren, A. Sallem, M. Loulou, and P. Siarry, “Analog circuit design optimization through the particle swarm optimization technique,” *Analog Integrated Circuits and Signal Processing*, vol. 63, pp. 71–82, Apr 2010.
- [72] M. Fakhfakh, E. Tlelo-Cuautle, and P. Siarry, *Computational intelligence in analog and mixed-signal (AMS) and radio-frequency (RF) circuit design*. Springer, Cham, 2015.
- [73] R. A. Rutenbar, G. G. E. 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.
- [74] M. Goswami and S. Kundu, “Constrained low-power cmos analog circuit design via all-inversion region mos model,” *2014 IEEE Fourth International Conference on Consumer Electronics Berlin (ICCE-Berlin)*, pp. 277–278, Sept 2014.
- [75] F. Rocha, R. Martins, N. Lourenço, and N. Horta, *Electronic Design Automation of Analog ICs combining Gradient Models with Multi-Objective Evolutionary Algorithms*. Springer-Briefs in Applied Sciences and Technology, Cham: Springer International Publishing, 2014.
- [76] W. Lyu, P. Xue, F. Yang, C. Yan, Z. Hong, X. Zeng, and D. Zhou, “An efficient Bayesian optimization approach for automated optimization of analog circuits,” *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 65, no. 6, pp. 1954–1967, 2018.
- [77] M. Kubar and J. Jakovenko, “A Powerful Optimization Tool for Analog Integrated Circuits Design,” *RADIOENGINEERING*, vol. 22, pp. 921–931, September 2013.
- [78] D. M. Binkley, C. E. Hopper, S. D. Tucker, B. C. Moss, J. M. Rochelle, and D. P. Foty, “A cad methodology for optimizing transistor current and sizing in analog cmos design,”

- IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 22, pp. 225–237, Feb 2003.
- [79] E. S. Ochotta, R. A. Rutenbar, and L. R. Carley, “Synthesis of high-performance analog circuits in astrx/oblx,” *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 15, pp. 273–294, March 1996.
- [80] B. De Smedt and G. G. E. Gielen, “Watson: design space boundary exploration and model generation for analog and rfc design,” *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 22, pp. 213–224, Feb 2003.
- [81] A. Sanabria Borbon, E. Tlelo-Cuautle, and L. de la Fraga, *Optimal Sizing of Amplifiers by Evolutionary Algorithms with Integer Encoding and g_m/I_D Design Method*, vol. 731, pp. 263–279. 09 2018.
- [82] S. Soto-Aguilar, A. Sanabria-Borbón, and E. Sánchez-Sinencio, “Surrogate-based Optimization-aided Design for Low Power Analog Circuits,” *Midwest Symposium on Circuits and Systems*, vol. 2018-August, pp. 566–569, 2019.
- [83] H. Graeb and U. Schlichtmann, “Pareto optimization of analog circuits considering variability,” no. 01, pp. 283–299, 2009.
- [84] D. Mueller-Gritschneider and H. Graeb, “Computation of yield-optimized pareto fronts for analog integrated circuit specifications,” in *Proceedings -Design, Automation and Test in Europe, DATE*, pp. 1088–1093, EDAA, 2010.
- [85] N. Damera-Venkata and B. L. Evans, “An automated framework for multicriteria optimization of analog filter designs,” *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, vol. 46, pp. 981–990, Aug 1999.
- [86] M. D. M. Hershenson, S. P. Boyd, and T. H. Lee, “Optimal design of a cmos op-amp via geometric programming,” *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 20, pp. 1–21, Jan 2001.

- [87] P. K. Rout, D. P. Acharya, and G. Panda, "A multiobjective optimization based fast and robust design methodology for low power and low phase noise current starved vco," *IEEE Transactions on Semiconductor Manufacturing*, vol. 27, pp. 43–50, Feb 2014.
- [88] P. Mandal and V. Visvanathan, "Cmos op-amp sizing using a geometric programming formulation," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 20, pp. 22–38, Jan 2001.
- [89] B. Liu, G. Gielen, and F. V. Fernández, *Automated Design of Analog and High-frequency Circuits*, vol. 501 of *Studies in Computational Intelligence*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2014.
- [90] J. Nocedal and S. J. Wright, *Numerical optimization. 2nd ed. Jorge Nocedal, Stephen J. Wright*. Springer series in operations research and financial engineering, New York : Springer, 2006.
- [91] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: Nsga-ii," *IEEE Transactions on Evolutionary Computation*, vol. 6, pp. 182–197, April 2002.
- [92] E. Tlelo-Cuautle and A. C. Sanabria-Borbon, "Optimising operational amplifiers by evolutionary algorithms and g_m/I_d method," *International Journal of Electronics*, vol. 103, pp. 1665–1684, Oct 2016.
- [93] N. Lourenço, "GENOM-POF : Multi-Objective Evolutionary Synthesis of Analog ICs with Corners Validation," pp. 1119–1126, 2012.
- [94] S. P. Mohanty, *Nanoelectronic Mixed-Signal System Design*. McGraw-Hill Education, 2015.
- [95] A. I. A. Cunha, M. C. Schneider, and C. Galup-Montoro, "An MOS transistor model for analog circuit design," *IEEE Journal of Solid-State Circuits*, vol. 33, pp. 1510–1519, Oct 1998.

- [96] R. L. Radin, G. L. Moreira, C. Galup-Montoro, and M. C. Schneider, "A simple modeling of the early voltage of mosfets in weak and moderate inversion," in *2008 IEEE International Symposium on Circuits and Systems*, pp. 1720–1723, May 2008.
- [97] D. M. Binkley, *Tradeoffs and Optimization in Analog CMOS Design*. John Wiley and Sons Ltd., 2008.
- [98] B. Liu, Y. He, P. Reynaert, and G. Gielen, "Global optimization of integrated transformers for high frequency microwave circuits using a gaussian process based surrogate model," in *2011 Design, Automation Test in Europe*, pp. 1–6, March 2011.
- [99] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*. MIT Press, 2006.
- [100] S. Zhang, W. Lyu, F. Yang, C. Yan, D. Zhou, X. Zeng, and X. Hu, "An efficient multi-fidelity bayesian optimization approach for analog circuit synthesis," in *Proceedings of the 56th Annual Design Automation Conference 2019, DAC '19*, (New York, NY, USA), Association for Computing Machinery, 2019.
- [101] Matworks, "Matlab documentation r2019b," 2019.
- [102] G. G. Gielen and W. M. Sansen, *Symbolic Analysis for Automated Design of Analog Integrated Circuits*. USA: Kluwer Academic Publishers, 1991.
- [103] W. Vereecken and M. Steyaert, *Ultra-Wideband Pulse-based Radio*. Germany: Springer Netherlands, 1 ed., 2009.
- [104] L. T. Bruton, F. N. Trofimenkoff, and D. H. Treleaven, "Noise performance of low-sensitivity active filters," *IEEE Journal of Solid-State Circuits*, vol. 8, pp. 85–91, Feb 1973.
- [105] J. Torres, M. El-Nozahi, A. Amer, S. Gopalraju, R. Abdullah, K. Entesari, and E. Sanchez-Sinencio, "Low Drop-Out Voltage Regulators: Capacitor-less Architecture Comparison," *IEEE Circuits and Systems Magazine*, vol. 14, no. 2, pp. 6–26, 2014.

- [106] A. Hajimiri, S. Limotyrakis, and T. H. Lee, “Jitter and phase noise in ring oscillators,” *IEEE Journal of Solid-State Circuits*, vol. 34, pp. 790–804, June 1999.
- [107] D. Ghai, S. P. Mohanty, and E. Kougianos, “Design of parasitic and process-variation aware nano-cmos rf circuits: A vco case study,” *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 17, pp. 1339–1342, Sep. 2009.
- [108] S. Prakash, H. Martínez-García, M. H. Naderi, H. Lee, and J. Silva-Martinez, “An agile supply modulator with improved transient performance for power efficient linear amplifier employing envelope tracking techniques,” *IEEE Transactions on Power Electronics*, vol. 35, no. 4, pp. 4178–4191, 2020.
- [109] M. Rostami, F. Koushanfar, and R. Karri, “A Primer on Hardware Security: Models, Methods, and Metrics,” *Proceedings of the IEEE*, vol. 102, no. 8, pp. 1283–1295, 2014.
- [110] J. Rajendran, Y. Pino, O. Sinanoglu, and R. Karri, “Security Analysis of Logic Obfuscation,” *IEEE/ACM Design Automation Conference*, pp. 83–89, 2012.
- [111] Y. Xie and A. Srivastava, “Mitigating SAT Attack on Logic Locking,” *Cryptographic Hardware and Embedded Systems*, pp. 127–146, 2016.
- [112] X. Xu, B. Shakya, M. M. Tehranipoor, and D. Forte, “Novel Bypass Attack and BDD-based Tradeoff Analysis Against All Known Logic Locking Attacks,” *Cryptographic Hardware and Embedded Systems*, pp. 189–210, 2017.
- [113] IHS Technology Press Release, “Top 5 most counterfeited parts represent a \$169 billion potential challenge for global semiconductor industry,” 2012.
- [114] R. Torrance and D. James, “The state-of-the-art in semiconductor reverse engineering,” *IEEE/ACM Design Automation Conference*, pp. 333–338, 2011.
- [115] R. A. Rutenbar, “Design Automation for Analog: The Next Generation of Tool Challenges,” *IEEE/ACM International Conference on Computer-Aided Design*, p. 458–460, 2006.

- [116] O. Alrawi, C. Lever, M. Antonakakis, and F. Monrose, “Sok: Security evaluation of home-based iot deployments,” in *2019 IEEE Symposium on Security and Privacy (SP)*, pp. 1362–1380, 2019.
- [117] G. Zhang, C. Yan, X. Ji, T. Zhang, T. Zhang, and W. Xu, “Dolphinattack: Inaudible voice commands,” *CCS ’17*, (New York, NY, USA), p. 103–117, Association for Computing Machinery, 2017.
- [118] M. Alioto, “Trends in Hardware Security: From Basics to ASICs,” *IEEE Solid-State Circuits Magazine*, vol. 11, no. 3, pp. 56–74, 2019.
- [119] D. H. K. Hoe, J. Rajendran, and R. Karri, “Towards Secure Analog Designs: A Secure Sense Amplifier Using Memristors,” *IEEE Computer Society Annual Symposium on VLSI*, pp. 516–521, 2014.
- [120] J. Leonhard, M. Yasin, S. Turk, M. T. Nabeel, M.-M. Louërat, R. Chotin-Avot, H. Aboushad, O. Sinanoglu, and H.-G. Stratigopoulos, “MixLock: Securing Mixed-Signal Circuits via Logic Locking,” *IEEE/ACM Design Automation and Test in Europe*, 2019.
- [121] G. Volanis, Y. Lu, S. G. R. Nimmalapudi, A. Antonopoulos, A. Marshall, and Y. Makris, “Analog Performance Locking through Neural Network-Based Biasing,” *IEEE VLSI Test Symposium*, pp. 1–6, 2019.
- [122] U. Guin, K. Huang, D. DiMase, J. Carulli, M. Tehranipoor, and Y. Makris, “Counterfeit Integrated Circuits: A Rising Threat in the Global Semiconductor Supply Chain,” *Proceedings of the IEEE*, vol. 102, no. 8, pp. 1207–1228, 2014.
- [123] K. Shamsi, M. Li, K. Plaks, S. Fazzari, D. Z. Pan, and Y. Jin, “Ip protection and supply chain security through logic obfuscation: A systematic overview,” *ACM Trans. Des. Autom. Electron. Syst.*, vol. 24, Sept. 2019.
- [124] S. Bhunia, M. Hsiao, M. Banga, and S. Narasimhan, “Hardware Trojan Attacks: Threat Analysis and Countermeasures,” *Proceedings of the IEEE*, vol. 102, no. 8, pp. 1229–1247, 2014.

- [125] M. Tehranipoor and F. Koushanfar, “A Survey of Hardware Trojan Taxonomy and Detection,” *IEEE Design and Test of Computers*, vol. 27, no. 1, pp. 10–25, 2010.
- [126] Q. Wang, R. L. Geiger, and D. Chen, “Hardware trojans embedded in the dynamic operation of analog and mixed-signal circuits,” pp. 155–158, 2015.
- [127] Y. Wang, Q. Wang, D. Chen, and R. L. Geiger, “Hardware trojan state detection for analog circuits and systems,” *NAECON 2014 - IEEE National Aerospace and Electronics Conference*, pp. 364–367, 2014.
- [128] C. S. Chang, C. P. Chao, J. G. J. Chern, and J. Y. C. Sun, “Advanced CMOS Technology Portfolio for RF IC Applications,” *IEEE Transactions on Electron Devices*, vol. 52, no. 7, pp. 1324–1334, 2005.
- [129] Texas Instruments, “Universal Active Filter.” <https://www.ti.com/lit/ds/symlink/uaf42.pdf>, 2010. Last accessed on 04/25/2020.
- [130] Y. M. Alkabani and F. Koushanfar, “Active Hardware Metering for Intellectual Property Protection and Security,” *USENIX Security Symposium*, 2007.
- [131] J. Rajendran, M. Sam, O. Sinanoglu, and R. Karri, “Security Analysis of Integrated Circuit Camouflaging,” *ACM SIGSAC Conference on Computer & Communications Security*, pp. 709–720, 2013.
- [132] C. Herder, M. Yu, F. Koushanfar, and S. Devadas, “Physical Unclonable Functions and Applications: A Tutorial,” *Proceedings of the IEEE*, vol. 102, no. 8, pp. 1126–1141, 2014.
- [133] S. G. Rao Nimmalapudi, G. Volanis, Y. Lu, A. Antonopoulos, A. Marshall, and Y. Makris, “Range-Controlled Floating-Gate Transistors: A Unified Solution for Unlocking and Calibrating Analog ICs,” *IEEE/ACM Design Automation and Test in Europe*, pp. 286–289, 2020.
- [134] V. V. Rao and I. Savidis, “Transistor sizing for parameter obfuscation of analog circuits using satisfiability modulo theory,” *2018 IEEE Asia Pacific Conference on Circuits and Systems (APCCAS)*, pp. 102–106, 2018.

- [135] V. Rao and I. Savidis, “Mesh based obfuscation of analog circuit properties,” *2019 IEEE International Symposium on Circuits and Systems (ISCAS)*, pp. 1–5, 2019.
- [136] S. Lee, C. Shi, J. Wang, A. Sanabria, H. Osman, J. Hu, and E. Sánchez-Sinencio, “A Built-In Self-Test and In Situ Analog Circuit Optimization Platform,” *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. PP, pp. 1–14, March 2018.
- [137] G. E. Suh and S. Devadas, “Physical Unclonable Functions for Device Authentication and Secret Key Generation,” in *IEEE/ACM Design Automation Conference*, pp. 9–14, 2007.
- [138] A. Antonopoulos, C. Kapatsori, and Y. Makris, “Security and trust in the analog/mixed-signal/rf domain: A survey and a perspective,” in *2017 22nd IEEE European Test Symposium (ETS)*, pp. 1–10, 2017.
- [139] K. Subramani, G. Volanis, M. Bidmeshki, A. Antonopoulos, and Y. Makris, “Trusted and secure design of analog/rf ics: Recent developments,” in *2019 IEEE 25th International Symposium on On-Line Testing and Robust System Design (IOLTS)*, pp. 125–128, 2019.
- [140] A. Kahng, J. Lach, W. Mangione-Smith, S. Mantik, I. Markov, M. Potkonjak, P. Tucker, H. Wang, and G. Wolfe, “Watermarking techniques for intellectual property protection,” *IEEE/ACM Design Automation Conference*, pp. 776–781, 1998.
- [141] M. Yasin, S. M. Saeed, J. Rajendran, and O. Sinanoglu, “Activation of Logic Encrypted Chips: Pre-test or Post-Test?,” in *IEEE/ACM Design, Automation Test in Europe*, pp. 139–144, 2016.
- [142] Z. Paral and S. Devadas, “Reliable and Efficient PUF-Based Key Generation Using Pattern Matching,” *IEEE International Symposium on Hardware Oriented Security and Trust*, no. 978, pp. 128–133, 2011.
- [143] M. Khalafalla and C. Gebotys, “PUFs Deep Attacks: Enhanced Modeling Attacks Using Deep Learning Techniques to Break the Security of Double Arbiter PUFs,” *Design, Automation Test in Europe Conference Exhibition*, pp. 204–209, 2019.

- [144] J. Delvaux, "Machine-Learning Attacks on PolyPUFs, OB-PUFs, RPUFs, LHS-PUFs, and PUF-FSMs," *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 8, pp. 2043–2058, 2019.
- [145] C. W. Lin and S. Ghosh, "A Family of Schmitt-Trigger-Based Arbiter-PUFs and Selective Challenge-Pruning for Robustness and Quality," *IEEE International Symposium on Hardware Oriented Security and Trust*, pp. 32–37, 2015.
- [146] A. Pfister, "Novel CMOS Schmitt Trigger with Controllable Hysteresis," *Electronics Letters*, vol. 28, no. 7, pp. 639–641, 1992.
- [147] P. E. Allen and D. R. Holberg, *CMOS Analog Circuit Design*. Oxford University Press, 2002.
- [148] B. L. Dokic, "CMOS Schmitt Triggers.," *IEE Proceedings, Part G: Electronic Circuits and Systems*, vol. 131, no. 5, pp. 197–202, 1984.
- [149] C.-K. Pham, "CMOS Schmitt Trigger Circuit with Controllable Hysteresis Using Logical Threshold Voltage Control Circuit," *IEEE/ACIS International Conference on Computer and Information Science*, pp. 48–53, 2007.
- [150] Chipworks, "Reverse Engineering Software." <http://www.chipworks.com/en/technical-competitive-analysis/resources/reerse-engineering-software>, 2016.
- [151] B. Greenley, R. Veith, D. Y. Chang, and U. K. Moon, "A Low-Voltage 10-Bit CMOS DAC in 0.01 – mm² Die Area," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 52, no. 5, pp. 246–250, 2005.
- [152] S. Hong and G. Cho, "High-Gain Wide-Bandwidth Capacitor-Less Low-Dropout Regulator (LDO) for Mobile Applications Utilizing Frequency Response of Multiple Feedback Loops," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 63, no. 1, pp. 46–57, 2016.