Portland State University

PDXScholar

Dissertations and Theses

Dissertations and Theses

9-2-2020

Analog Statistical Design for Manufacturability

Yu-Shan Wang

Portland State University

Follow this and additional works at: https://pdxscholar.library.pdx.edu/open_access_etds

Part of the [Electrical and Computer Engineering Commons](#)

Let us know how access to this document benefits you.

Recommended Citation

Wang, Yu-Shan, "Analog Statistical Design for Manufacturability" (2020). *Dissertations and Theses*. Paper 5605.

<https://doi.org/10.15760/etd.7477>

This Dissertation is brought to you for free and open access. It has been accepted for inclusion in Dissertations and Theses by an authorized administrator of PDXScholar. Please contact us if we can make this document more accessible: pdxscholar@pdx.edu.

Analog Statistical Design for Manufacturability

by

Yu-Shan Wang

A dissertation submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy
in
Electrical and Computer Engineering

Dissertation Committee:

John M. Acken, Chair

Branimir Pejcinovic

Xiaoyu Song

Tim Anderson

Portland State University
2020

Abstract

The manufacturing yield, overkill, and defect level limit the feasibility of analog circuits in SoCs. The conventional method of handling process and environmental variation is to assign a design margin such that the design meets specifications at several processes and environmental corners. However, checking only extreme corners limits performance in comparison to the more rigorous statistical approach of the computing manufacturing and quality figure of merit. The statistical approach requires transistor-level simulation of hundreds or even thousands of samples, not just a few corners, and thus is very time consuming.

This research offers a method for reducing the time required for the statistical approach by characterizing each of the many samples of building blocks once at the transistor level. The building blocks are scalable such that the statistics are preserved when a building block is adjusted to the requirement of a higher-level design. Many design scenarios can be rapidly explored by assembling and scaling the building block samples without SPICE simulation. This study employs a continuous-time low-pass filter design example to extract the requirements of the building block approach. The requirements include a method to assemble building blocks (biquad element for the example) into a filter design while preserving the statistics that would have been extracted by simulation of the entire filter at the transistor level. The assembly method for both linear and nonlinear response is proposed.

Table of Contents

List of Tables	iv
List of Figures	v
Chapter 1 Introduction	1
1.1 Analog Circuit Design	3
1.2 Synthesis of Analog Circuits	4
1.3 Outline of This Thesis	5
Chapter 2 Filter Design	8
2.1 Analog Filter	9
2.2 The Transfer Function	11
2.3 Filter Topology	14
2.4 Performance Metrics: Design Figure of Merit	20
2.5 Sensitivity	24
Chapter 3 Tolerance Design	32
3.1 Introduction	32
3.2 Tolerance Analysis	33
3.3 Design Centering	34
3.4 Reduction of Sensitivity of Tolerance	34
3.5 Manufacturing Figure of Merit (FOM) : Defect level and Yield loss .	35

Chapter 4	Analog Statistical Design in Linear and Nonlinear Response for Manufacturing	38
4.1	Introduction	38
4.2	Motivation and Contributions	41
4.3	Building Block Methodology for Manufacturing and Quality FoM Prediction	45
4.4	Biquad Library and Characterization	47
4.5	Estimation of Filter Frequency Response Using Biquad Library . . .	51
4.6	Estimation of Linearity (IIP ₃) using Biquad Library	53
4.7	Results: Comparison of Design Alternatives	57
4.8	Conclusions	65
Chapter 5	Copula-based Modeling of Analog Parametric Tests Separating Variation from Defects	68
5.1	Introduction	68
5.2	Methodology	69
5.3	Result	79
5.4	Recommendations	91
5.5	Conclusion	91
Chapter 6	Conclusion	93
	References	95

List of Tables

4.1	Filter use requirements and test specifications at 27°C.	45
4.2	Filter Design Specifications and Biquad Library Usage	58
4.3	IIP_3 Correlation between SPICE and synthesis	60
4.4	Filter Comparison from AC Manufacturing Results	63
4.5	IIP_3 Manufacturing Results	63
4.6	Run time comparison SPICE and Synthesis (Building Block Method)	65
5.1	Copula Properties	72
5.2	Statistics summary for Monte Carlo parameter, Device Model	80
5.3	Miscorrelating outliers movements from original model file to model file with $KT=0$ for PMOS	85
5.4	Monte Carlo Parameters RAND 1-4 for E1-6	87

List of Figures

1.1	Product life cycle	2
1.2	Circuit design cycle	4
1.3	Proposed circuit design cycle using library of building blocks - with no SPICE simulation for each design alternative	6
2.1	Circuit hierarchy. This research focused on synthesizing the perfor- mance of Gm cells and low-pass cascaded Gm-C biquad filters. . . .	10
2.2	Second-order low-pass filter phase response.	13
2.3	Second-order low-pass filter phase response with different Q and for the same ω_0	13
2.4	Gm C biquad	15
2.5	Cascaded biquads	16
2.6	Differential in and single ended out OTA schematic	16
2.7	Fully differential OTA schematic	16
2.8	Small signal model for a Gm cell including input and output para- sitic: (G_I, G_O)	18
2.9	Lowpass filter specification	21
2.10	IM3 falls inband of interest	24

2.11	Power spectrum at the output of a nonlinear system. Two tones at ω_1 and ω_2 are applied at input. The output contains tones at ω_1 and ω_2 with power of P_1 and P_2 . The output also contains 3rd order intermodulation terms ω_3 and ω_4 with power of P_L and P_u	25
2.12	Differential pair with active load	27
2.13	Filter nonlinearity	31
3.1	Design centering by designing the circuits at R'_T instead of R_T to achieve higher yield. R_A is the Region of Acceptance (acceptable circuit performance) and R_T is the Tolerance Region (expected manufacturing variation).	35
3.2	Tolerance tightening (Reduce R_T). Different design alternatives result in different R_T . Smaller R_T in design 2 represents less performance variability comparing to design 1 due to device variation. . .	35
3.3	Test process	37
4.1	Relationship of Figure-of-Merit to population category probabilities in the context of an analog filter. Without Test phase, a fraction (red) of the population of components would fail in Use condition. Test reduces the customer-perceived defect level (DL) by trimming and screening components. Reducing DL increases yield loss (YL) and overkill (OL).	42
4.2	(a) A 2N-order low pass cascaded biquad Gm-C filter using N biquads. (b) Each biquad is constructed with transconductor (Gm) and capacitors (C).	44

4.3	Building Block methodology: Biquad as building block replaces the compute intensive SPICE simulation of the filter in Test/Use and building block level design optimization by much less compute intensive sampling and scaling of the (K, Q, ω_0, IM_3) statistical model. SPICE simulations to establish the (K, Q, ω_0, IM_3) biquad model need be done only once.	48
4.4	Biquad library: building-block library for the Gm-C biquads comprising of small-signal and large signal characterization as well as process and mismatch data.	49
4.5	The database organization and storage of each biquad design characterization.	50
4.6	(a) Filter synthesis flow and (b) flow chart to estimate filter transfer function using analog library.	52
4.7	Generating multiple filter samples using bootstrap sampling for 4 th -order filter (a) using biquads from the same biquad design r_1 and (b) using biquads scaled from different biquad designs, r_1 and r_2 . Each line indicate a possible sampling combination.	53
4.8	Modeling of IM_3 voltage for cascaded stages using magnitude and phase of the generated and propagated IM_3 voltage terms from different stages.	55
4.9	Bilinear interpolation for (a) IM_3 magnitude and (b) IM_3 phase at arbitrary (Q_i, f_i) for a single process and mismatch indexed biquad sample.	55
4.10	Flow chart for (a) generating biquad IM_3 in synthesized filter and (b) calculating filter IIP_3	56

4.11	Ideal filter transfer functions (a) <i>4th</i> -order Chebyshev and (b) <i>8th</i> -order Butterworth. Additional plots are library biquad component transfer functions after scaling to filter design specifications. Use requirement upper and lower magnitude brick-walls are shown as dashed lines.	57
4.12	(a) <i>4th</i> -order Chebyshev and (b) <i>8th</i> -order Butterworth. Filter design alternatives IIP_3 values are compared by scatter-plot and by marginal histograms. Comparison of each simulation and estimation uses 250 filter samples. Each data point in the plots is one filter sample. SPICE IIP_3 evaluation is from transient analysis. IIP_3 synthesized are from interpolation. IIP_3 corner frequency is 140kHz for both filters	59
4.13	AC transfer functions of 1200 samples of two filter design alternatives (a) <i>4th</i> -order Chebyshev (b) <i>8th</i> -order Butterworth. Each filter is DC trimmed to 0dB at test. Black vertical bars locate test frequency and magnitude specifications. Passband test guard-band is 0.5dB @ 190kHz. Nominal trace denotes ideal AC response. Legend denotes results of test and use, see Table 4.1. For example, PF denotes pass at test and fail in use. Use specifications are the dashed lines. . . .	62

4.14	Bootstrap IIP_3 synthesis results for 250 (a) 4th-order Chebyshev filters and (b) 8th-order Butterworth filters. Use IIP_3 set the center-frequency test-tone =100kHz and at test guard-banded the center-frequency =140kHz. Frequency offset = $\mp$ 2.5% of the center-frequency. Vertical and horizontal lines mark the IIP_3 test and use limits. Each dot '.' in FP and PF quadrants is a test escape or test overkill, respectively. Counts in FP and PF quadrants assess manufacturing risks from the miscorrelation of test and use.	64
5.1	Synthesized sample from (a) Gaussian copula with $j=0.9$ and (b) Clayton copula with $j=9.74$	73
5.2	Bivariate Normal Joint Density $h(x,y)$ and histogram of x , $f(x)$, and y , $g(y)$	74
5.3	Joint density plots from Gaussian marginal with dependence structure of (a) Gaussian copula $h_1(x,y)$ and (b) Clayton copula h_2 . .	75
5.4	Joint density plots from Bimodal and Weibull marginal with dependence structure of (a) Gaussian copula $h_3(x,y)$ and (b) Clayton copula $h_4(x,y)$	75
5.5	OTA schematic	77
5.6	Flow chart for the proposed technique	79
5.7	Marginal distribution of (a) random (b) dv_{th0} (c) OTA gain at 27C, $f(x)$ (d) OTA gain at 100C, $g(y)$	80
5.8	Joint density $h(x,y)$ for OTA gain at 1VDD, 27C/80C	81
5.9	Empirical copula $c(u,v)$ for OTA gain at 1.0VDD, 27C/80C	82
5.10	(a) Symmetric analysis for empirical copula. Instances outside 98% density contour are highlighted in rectangular dot. (b) Miscorrelating outliers are highlighted in triangle dot.	83

5.11	Joint density $h(x,y)$ and marginal distribution $f(x)$ and $g(y)$ with mis-correlating outliers highlighted in orange	84
5.12	Empirical copula for OTA gain at 27C/100C and 1.0VDD with PMOS MP (a) original MC results (b)KT=0 (c) AT=0 (d)UTE=0	85
5.13	Three Factorial DoE- joint density	87
5.14	Three Factorial DoE- copula	88
5.15	Three Factorial DoE - copula with limits	88
5.16	Dependence structure for OTA gain at 1VDD and (a) 100C/27C (b) 0C/27C	90
5.17	Dependence structure for OTA between 27C and 80C at (a) 1.0VDD (b) 1.1VDD	90
5.18	Spearman's rho at various environment conditions	90

Chapter 1

Introduction

Advanced CMOS technology has enabled highly integrated, high-performance, and low power electronic devices such as computers and mobile phones. With the increased demand for more features in these electronic devices, the number of electronic circuits and their complexity have increased exponentially. Different integrated circuits (ICs) such as memory, microprocessors, peripherals are combined to build a system on a chip (SoC). To meet the time to market, the efficiency of the design of an SoC and verification of the required performance are crucial [1].

Fig. 1.1 shows a typical product cycle. Take an RF wireless transceiver used in a base station as an example; a customer requests a list of transceiver requirements such as the target standard, power envelope, form factor, and cost target. The system architect comes up with a high-level architecture with the corresponding circuit performance specifications that could meet the overall transceiver requirement. The architecture typically includes digital circuits(e.g., an encoder/decoder, signal processing units, and controllers), analog circuits (e.g., an amplifier, filters, and oscillators), and mixed-signal circuits (e.g., data converters and phase lock loops) [2,3]. The specifications for the circuit performance must be met over manufacturing process variations and different environmental conditions such as voltage and temperature variations. Next, design engineers design circuits based on the specifications and verify the design using circuit simulators [4,5]. The design verification includes sim-

Figure 1.1: Product life cycle

ulations at process, temperature, and voltage (PVT) variations. After the circuit's performance over PVT variations meets all the specifications and the final design choice is made, it is sent to a semiconductor manufacturing plant (also referred to as a fab or foundry) to be manufactured. The manufactured circuit will be tested based on the test limit [6] in the foundry and assembly facilities. If the circuit passes all tests, it will be shipped to the customer; otherwise, it will be discarded as a yield loss, which is the the number of circuits that will fail the design specifications out of the total production run [7,8].

1.1 Analog Circuit Design

Analog circuit design is a highly labor intensive and slow process. Even in advanced highly integrated systems, analog designs are done in much the same way as they were decades ago. One of the bottlenecks is their manual design and simulation, which are highly time consuming. Transistor SPICE simulations are often required to estimate circuit performance [9–11]. Monte Carlo simulations [12, 13] are used to evaluate the impact of manufacturing process variation on circuit performance. Furthermore, during the design phase, many different circuit alternatives for analog circuits are considered, such as filters and amplifiers [14]. Conventionally, each alternative design (e.g., different filter implementation) is simulated over PVT to verify the performance and make the final design choices. The complexity of circuit SPICE simulation increases with the size of the design (i.e., the transistor count), number of circuit design instances simulated (i.e., the number of Monte Carlo runs), and the number of combinations of environmental conditions (i.e., combinations of supply voltage and temperature). Often, it takes multiple iterations in the design cycle to finally meet the required specifications. Fig. 1.2 presents the design cycle for analog circuits. Based on the circuit specifications, the designer first comes up with several different circuit topologies based on factors such as his/her experience and rules of thumb and then evaluates each design alternative by simulating their performance metrics at different environmental conditions. The analog design alternatives with performance metrics that meet the specifications are further evaluated for a manufacturing figure of merit (FOM), such as yield analysis, for the final decision [15]. The yield analysis includes estimation of yield loss (YL), overkill (OL) and defect level (DL) and will be explained in Chapter 3. Early rejection of design alternatives is desirable when the analog circuitry is embedded in an SOC design. SOC design cycles are long, and reducing simulation of digital and analog circuits to the best performing

Figure 1.2: Circuit design cycle

combinations helps to reduce the design time and time to market. In addition to reducing the time to market, expensive simulation resources can also be spared if, early in a product life cycle, tools are available to identify design alternatives that are unlikely to meet the performance metric. Simulation resources are spared because such alternatives could be rejected at this earlier stage before expensive circuit simulations begin in earnest.

1.2 Synthesis of Analog Circuits

One way to evaluate design alternatives early and accelerate analog circuit design is to synthesize higher level circuits from a library of building blocks without SPICE simulations [16–18]. While synthesis for digital circuits with logic gates has been

used extensively [19, 20], synthesizing an analog circuit and predicting its manufacturing FoM is not as straightforward because of its complex analog operation ; that is, its continuous time and continuous signal level compared with the discrete time and discrete level in digital circuits [21–27]. The objective of this research is to introduce a method to estimate the performance metrics of different analog circuit design alternatives with minimal circuit SPICE simulations. As shown in Fig. 1.3, the estimation of circuit performance in this work was based on building a library of building blocks for higher-level analog circuit design alternatives. Each analog design alternative’s performance metrics at different environmental conditions was synthesized. The estimation of manufacturing FOM was obtained without circuit SPICE simulations of any of the analog circuit design alternatives. Circuit simulation was confined to just the library of building blocks, which was a much smaller circuit, and thus simulation was minimized when evaluating several design alternatives that met a given specification. This thesis provides a systematic method to build the library and model the performance metrics for analog circuit design alternatives from that library. This work provides a basis for a CAD tool for the early evaluation of analog design alternatives and for the early prediction of manufacturing risks in the product life cycle.

1.3 Outline of This Thesis

The remainder of this thesis is organized as follows. Chapter 2 provides a basic description of analog filter design, which is used to demonstrate the synthesis of analog circuits using a building block library. Chapter 3 introduces the concept of tolerance design, tolerance analysis, and the concept of tolerance adjustment and design centering for yield optimization. Chapter 4 demonstrates the synthesis method using the building block library for two filter performance metrics: ac transfer function and

Figure 1.3: Proposed circuit design cycle using library of building blocks - with no SPICE simulation for each design alternative

input-referred third-order intercept point (IIP3) 2.4.2 for a family of cascaded-biquad Gm-C filters (refer to section 2.3.1 and discusses the use of such a tool for predicting linear and nonlinear performance and selecting design alternatives. Chapter 5 describes the method of using rank statistics to model and predict the analog performance of an operational transconductance amplifier (OTA) and the yield under manufacturing process variation. Chapter 6 concludes the thesis.

Chapter 2

Filter Design

Electronic circuits are generalized into three categories: analog circuits, digital circuits and mixed-signal circuits [28], as shown in Fig. 2.1. Whereas digital circuits operate using digital (discrete) signals, analog circuits process continuous signals and are more sensitive to noise compared with digital circuits. Mixed-signal circuits have a mixture of analog and digital components. Examples of analog circuits are continuous-time filters, amplifiers, and oscillators. Each analog functionality can be realized in several ways. Take filters for example, types of filter include n^{th} order Butterworth and Chebyshev filters [29]. For each type of filter, different topologies exist for implementing the circuit such as cascaded-biquad, LC ladder, and multiple feedback. Each of these can be realized using trans-conductors along with capacitors (Gm-C) or operational amplifiers along with resistors and capacitors (Opamp-RC), among others. [30]. For Gm-C filters, there are different types of trans-conductor Gm cells for implementation. Furthermore, for each Gm cell topology, there are choices of different trans-conductance ratios among Gm cells (Gm ratio) in the filter and different design components such as capacitance to design for specification. Each combination of design choices (i.e., design freedoms) are design alternative decisions that must be made not only to meet the analog design requirement but also to maximize the yield for manufacturing. An early decision to choose design alternatives is desirable to reduce time-to-market, simulation resources, and design man power. The method

developed in this research can predict the yield and evaluate the feasibility of different filter design alternatives without SPICE simulations on the filters, thereby reducing the time and number of iterations for making design decision.

This research aims to develop a methodology to build a library of building blocks to enable analog filter design alternatives to be evaluated (see Fig. 1.3). The building block is a combination of Gm-C biquad with a different Gm ratio and capacitance ratio. The library can be used to synthesize each filter design alternative. The synthesis results evaluate the designs by predicting circuit performances and yield.

2.1 Analog Filter

Analog filters are often used in systems to remove the interference or reject unwanted signals outside of the target signal band. Based on the requirements such as attenuation, phase, and group delay, the type of a filter is chosen; then, a circuit topology is selected to implement it. The filter characteristic shapes the attenuation curve and is often one of the n^{th} order classical types such as a Butterworth, Bessel, Elliptic, or some form of Chebyshev. One simple method to realize a higher-order filter function is to factor both the numerator and denominator polynomials into a biquadratic or bilinear function. Higher-order filters are often built of cascaded first and second order blocks. A variety of circuit topologies exist for building second-order sections; for example, Sallen-Key, multiple feedback, and biquad filters. For some applications that use filters, the amplitude response is of greater interest than the phase response. However, in some applications, the phase response of the filter is critical. An example of this might be where a filter is an element of a process control loop. Here, the total phase shift is of concern because it may affect loop stability.

Often, when designers are selecting the topology for implementation, several rules of thumb exist for maximizing circuit performance and minimizing the performance

Figure 2.1: Circuit hierarchy. This research focused on synthesizing the performance of Gm cells and low-pass cascaded Gm-C biquad filters.

variation caused by different environmental conditions and manufacturing processes. These rules of thumb depend heavily on the designer's experience. For example, gain and quality factor allocation for each cascaded stage can be used to make the maximum and minimum signal strengths in the pass band in all sections as uniform as is practical. If this is achieved, the burden of attaining the overall gain is shared by all sections without making the signal in one particular section either too strong to cause op-amp saturation, and thus distortion of the signals, or too weak that it falls below the noise threshold. A quantitative comparison can also be conducted by an exhaustive study of the gains from the input to the outputs at all intermediate junctions for differently sequenced combinations. The sequence that provides the flattest variations of these intermediate transfer functions can be considered the most desirable choice [31].

The decision of which type of topology and implementation can sometimes be difficult unless tolerance analysis is performed using Monte Carlo simulation and yield is estimated [32–34]. In this chapter, the cascaded biquad Gm-C filter design used for demonstration is analyzed in detail. The filter AC response using Gm cells is modeled including parasitics. The sensitivity of quality factor (Q) to parasitics and that of linearity (IIP3) to device tolerance is analyzed. Chapter 4 uses the Gm-C filter design to demonstrate the method for tolerance analysis with minimum SPICE simulation and evaluate rules of thumb quantitatively.

2.2 The Transfer Function

A general n^{th} order filter transfer function is

$$H(s) = \frac{b_0 + b_1s + b_2s^2 + b_ms^m \dots}{s^n + a_0s^{(n-1)} + a_1s^{n-2} + \dots} \quad (2.1)$$

where coefficients a and b are determined by the type of filter characteristics. The

order of the filter n must satisfy $n \geq m$.

For a second-order low-pass filter, the transfer function is

$$H(s) = \frac{Kb}{s^2 + as + b} = \frac{K\omega_0^2}{s^2 + \frac{\omega_0}{Q}s + \omega_0^2} \quad (2.2)$$

where Q is the quality factor which determines the peaking in the amplitude response and the sharpness of the phase transition, and ω_0 is the normalized center frequency in radians per second. Higher Q will cause higher peaking in the amplitude response. For example, the second-order Butterworth has a Q of 0.707, producing a maximal flat response [35]. The two poles of the filter are

$$(p_1, p_2) = \frac{\omega_0}{2Q}(1 \pm \sqrt{4Q^2 - 1}) \quad (2.3)$$

For a second-order low-pass filter, the transfer function has a phase shift that can be approximated using

$$\phi(\omega) = \arctan\left(\frac{-\omega_0 \times Q \times \omega}{\omega_0^2 - \omega^2}\right) \quad (2.4)$$

Fig. 2.2 shows Eq. 2.4 from two decades below to two decades above the center frequency. When the signal frequency is the same as the center frequency ω_o , that is, the normalized frequency is 1, the signal exhibits a phase shift of -90. Fig. 2.3 shows the effect of a low-pass filter on the phase response as Q is varied. The phase can start to change well below the cutoff frequency for low values of Q .

Figure 2.2: Second-order low-pass filter phase response.

Figure 2.3: Second-order low-pass filter phase response with different Q and for the same ω_0 .

2.3 Filter Topology

2.3.1 Filter implementation

One topology for implementing a second-order filter is Gm-C biquad, as shown in Fig. 2.4. Assuming no current can enter the input of the transconductor, we obtain the following:

$$\begin{aligned}I_1 &= G_1 \times V_{IN} \\I_2 &= -G_2 \times V_{bp} \\I_4 &= G_4 \times V_{bp} \\I_3 &= -G_3 \times V_{OUT} \\I_1 + I_2 + I_3 &= s \times C_1 \times V_{bp} \\I_4 &= s \times C_2 \times V_{OUT}\end{aligned}$$

Solving the nodal equations, the low pass biquad transfer function using Gm-C implementation is

$$H(s) = \frac{V_{OUT}}{V_{IN}} = \frac{G_1 G_4}{C_1 C_2} \frac{1}{s^2 + \frac{G_2}{C_1} s + \frac{G_3 G_4}{C_1 C_2}} \quad (2.5)$$

Comparing 2.5 with 2.2, we may identify

$$K = \frac{G_1}{G_3} \quad (2.6)$$

$$\omega_o = \sqrt{\frac{G_3 G_4}{C_1 C_2}} \quad (2.7)$$

$$\frac{\omega_o}{Q} = \frac{G_2}{C_1} \quad (2.8)$$

and hence

Figure 2.4: Gm C biquad

$$Q = \frac{1}{G_2} \sqrt{\frac{C_1}{C_2} G_3 G_4} \quad (2.9)$$

From Eq. 2.6 to 2.9, we observe the following: (a) the biquad gain is the ratio of transconductance of G_1 and trans-conductance of G_3 . (b) Q depends on the ratio of C_1 and C_2 . (c) ω_0 depends on the product of C_1 and C_2 . An example of designing a biquad for a certain K , Q and ω_0 is as follows: (1) For a given G_3 , G_1 equals $K \times G_3$. (2) Assuming $G_2 = G_4 = G_3 \times \frac{1}{G_r}$, the quality factor can be re-written as $Q = \sqrt{G_r} \times \sqrt{C_1/C_2}$. Therefore, for a given G_r ratio, the capacitor ratio can be calculated: $C_1/C_2 = C_r = Q^2 \times G_r$. (3) With C_r , G_r , and G_2 , the absolute value for C_2 is derived: $C_2 = \sqrt{G_2^2 G_r / C_r \omega_0^2}$

A higher even n^{th} order filter may be constructed by cascading multiple biquads in series as shown in Fig. 2.5. The filter's transfer function is the product of the transfer functions of the individual biquads. The selection of design parameters such as frequency ω_o and quality factor Q of each of the filter's biquads is determined by the specification of the filter [36].

$$H_{filter}(s) = \prod_{i=1}^{n/2} \frac{K_i \omega_{o_i}^2}{s^2 + \frac{\omega_{o_i}}{Q_i} s + \omega_{o_i}^2} \quad (2.10)$$

Figure 2.5: Cascaded biquads

Figure 2.6: Differential in and single ended out OTA schematic

2.3.2 Gm cell implementation and its parasitic model

There are multiple ways to implement Gm cells in Fig. 2.4 [14]. Fig. 2.6 and 2.7 show two OTA designs used in this research. Both OTAs are designed in TSMC 45nm technology.

An ideal OTA design has infinite input and output impedance. Deviation from ideality of an OTA can be represented by the input and output parasitic transconductance models (G_I , G_O) and parasitic capacitances (C_I , C_O) as shown in Fig. 2.8.

Figure 2.7: Fully differential OTA schematic

The input and output admittance for the i^{th} Gm cell in Fig. 2.4 are

$$Y_{Ii} = sC_{Ii} + G_{Ii}, \quad Y_{Oi} = sC_{Oi} + G_{Oi} \quad (2.11)$$

Assuming the biquad is followed by another biquad using the same kind of OTAs, the KCL equations for the three nodes of the biquad are

$$-V_{in}Y_{I1} + I_{in} = 0 \quad (2.12)$$

$$V_{in}G_1 - V_{out}G_3 - V_{bp}G_2 - V_{bp}sC_1 - V_{bp}(Y_{O1} + Y_{O2} + Y_{O3} + Y_{I2} + Y_{I4}) = 0 \quad (2.13)$$

$$V_{bp}G_4 - I_{out} - V_{out}sC_2 - V_{out}(Y_{O4} + Y_{I3}) = 0 \quad (2.14)$$

$$I_{out} = V_{out}Y_{I1} \quad (2.15)$$

Applying Eqs. 2.11 to Eq. 2.12 to 2.15, we obtain

$$V_{in} - V_{out}G_3 - V_a(G_y + sC'_1) = 0 \quad (2.16)$$

$$V_aG_4 - V_{out}(G_x + sC'_2) = 0 \quad (2.17)$$

where

$$G_x = G_{I1} + G_{I3} + G_{O4} \quad G_y = G_2 + G_{O1} + G_{O2} + G_{O3} + G_{I2} + G_{I4} \quad (2.18)$$

$$C'_1 = C_1 + C_{O1} + C_{O2} + C_{O3} + C_{O2} + C_{O4} \quad C'_2 = C_2 + C_{I1} + C_{I3} + C_{O4} \quad (2.19)$$

Figure 2.8: Small signal model for a Gm cell including input and output parasitic: (G_I , G_O)

Following Schaumann et al. [37], a model of the transfer characteristic that takes account of the input and output parasitics of the OTA circuits in the biquad design is

$$H(s) = \frac{G_1 G_4}{C'_1 C'_2} \frac{1}{s^2 + \frac{C'_1 G_x + G_y C'_2}{C'_1 C'_2} s + \frac{G_3 G_4 + G_y G_x}{C'_1 C'_2}} \quad (2.20)$$

Comparing Eq. 2.20 with Eq. 2.2 gives

$$K = \frac{K'}{1 + uv} \quad Q = Q' \frac{\sqrt{1 + uv}}{v + uQ'^2} \equiv f_{u,v}(Q) \quad \omega_o = \omega'_o \sqrt{(1 + uv)} \quad (2.21)$$

where

$$K' = \frac{G_1}{G_3} \quad Q' = \frac{\sqrt{G_3 G_4}}{G_2} \sqrt{\frac{C'_1}{C'_2}} \quad \omega'_o = \sqrt{\frac{G_3 G_4}{C'_1 C'_2}} \quad (2.22)$$

and

$$u = \frac{G_2}{G_3 G_4} G_x \quad v = \frac{G'_2}{G_2} \quad (2.23)$$

Assuming parasitic capacitance is negligible compared with the filter capacitance (i.e., $C_{Ix} = C_{Ox} \ll C_1, C_2$), Eq. 2.22 is consistent with Eqs. 2.6-2.9. K , Q , and ω_o are biquad parameters with parasitic effect, whereas K' , Q' , and ω'_o are parasitic free biquad parameters.

Note that the “realized” quality factor Q' using a non ideal OTA can be expressed

in terms of the quality factor Q with parasitics as

$$Q' = \frac{2vQ}{\sqrt{1+uv} + \sqrt{1+uv(1-4Q^2)}} \equiv f_{u,v}^{-1}(Q) \quad (2.24)$$

It can be shown from Eq. 2.24 that Q' has a maximum possible value of

$$Q'_{max} = \frac{1}{2\sqrt{uv}} \quad (2.25)$$

From Eq. 2.25 and 2.23, the maximum quality factor that can be realized is limited by the parasitics G_O and G_I [38].

From Eq. 2.25

$$Q'_{max} = \frac{1}{2\sqrt{uv}} = \frac{1}{2\sqrt{\frac{G_2}{G_3G_4}G_x\frac{G'_2}{G_2}}} \quad (2.26)$$

This shows that Q'_{max} is a function of transconductance of Gm cells.

2.3.3 Simplified biquad parasitic model

Sometimes, for design simplicity and good layout matching to reduce performance variation caused by manufacturing, it is common to use the same Gm design for G_1 and G_3 and the same Gm design for G_2 and G_4 . For example, one filter implementation can be designed with gain $K = 1$ by assigning $G_1 = G_3 = G_{odd}$ and $G_2 = G_4 = G_{even}$. The nominal design values of OTA transconductances and capacitances in the design are given as multiples of g_m (Siemens) and C_a (Farads) such that $G_{odd} = scale_{odd} \times g_m$, $G_{even} = scale_{even} \times g_m$, $C_1 = scale_{C1} \times C_a$, and $C_2 = scale_{C2} \times C_a$. We define Gm ratio, Gr , as $\frac{G_{odd}}{G_{even}} = \frac{scale_{odd}}{scale_{even}}$ and capacitor ratio, Cr , as $\frac{C_1}{C_2} = \frac{scale_{C1}}{scale_{C2}}$. Now assume that parasitics of transconductance scale in the same way as nominal design values of transconductances and $G_i = 0$, from Eqs.

2.18 and 2.19, we get

$$G'_2 = scale_{even}g_m + (2scale_{odd} + scale_{even})G_o \quad (2.27)$$

$$G' = scale_{even}G_o \quad (2.28)$$

We now rewrite Eq. 2.23 as

$$u = \mu\gamma \quad v = 1 + \frac{1 + \mu + \eta}{\mu}\gamma \quad (2.29)$$

where

$$\mu = \frac{b}{a} \quad \eta = 1 \quad \gamma = \frac{G_o}{g_m} \quad (2.30)$$

γ is a measure of the effect of parasitics on the biquad transfer characteristic. For ideal Gm cell, γ is 0.

2.4 Performance Metrics: Design Figure of Merit

For filter design, performance parameters such as frequency response, linearity, noise, and power dissipation are commonly evaluated. In practice, there are trade-offs between these parameters, making the design a multi-dimensional optimization. In this section, small signal AC frequency response and large signal IIP3 are described for later demonstration.

2.4.1 Small signal AC frequency response

Before a filter can be designed, a set of filter specifications must be defined. For example, suppose that we would like to design a low-pass filter with a corner frequency ω_c ,

Figure 2.9: Lowpass filter specification

stopband frequency ω_s , dc gain, passband ripple, and stopband attenuation. Fig. 2.9 shows five of the specifications in small signal ac response for lowpass filter response.

- Corner frequency (ω_c): This is the frequency range in which we desire to let the signal through with minimal attenuation.
- Stopband frequency (ω_s): This is the lower boundary of the frequency range where the signal should be attenuated to a certain attenuation.
- Passband ripple (A_{pass}): The max magnitude variation in the passband, in decibels.
- Stopband attenuation (A_{stop}): The max magnitude level in the stopband, or in other words the min attenuation in the stopband, in decibels.
- DC gain ($|H(j0)|$): the filter gain at DC.

The passband is defined as the frequency range $1 \leq \omega \leq \omega_c$, the stopband as the frequency range $\omega \geq \omega_s$, and the transition band as the frequency range $\omega_c \leq \omega \leq \omega_s$.

2.4.2 Large signal linearity: IIP3

Analog circuits such as low noise amplifiers (LNAs), mixers, filters, power amplifiers (PAs) and other components can generate very large signal dynamics and results in non-linear behavior between input and output. Several parameters have been defined to characterize this non-ideal relationship between input and output: (1) 1dB compression points, (2) compression dynamic range, (3) spurious-free dynamic range, (4) desensitization dynamic range, and (5) intercept points. Since all the aforementioned terms indicate how good the linearity of a device is, relationships exist between them. This study focuses exclusively on the third-order intercept points, or IP3, because they reveal the most about how non-linearity negatively affects useful signals.

An output $y(t)$ of a nonlinear system can be approximated by the Taylor's series expansion of input $x(t)$

$$y(t) = \alpha_0 + \alpha_1 x(t) + \alpha_2 x^2(t) + \alpha_3 x^3(t) + \dots \quad (2.31)$$

If two signals with different frequencies are applied to a nonlinear circuit, the output exhibits some components that are not harmonics of the input frequencies [39].

Assume the input of the circuit is

$$x(t) = A_1 \cos(\omega_1 t + \phi_1) + A_2 \cos(\omega_2 t + \phi_2)$$

After going through a nonlinear system, the output becomes

$$\begin{aligned} y(t) = & \alpha_1 (A_1 \cos(\omega_1 t + \phi_1) + A_2 \cos(\omega_2 t + \phi_2)) \\ & + \alpha_2 (A_1 \cos(\omega_1 t + \phi_1) + A_2 \cos(\omega_2 t + \phi_2))^2 \\ & + \alpha_3 (A_1 \cos(\omega_1 t + \phi_1) + A_2 \cos(\omega_2 t + \phi_2))^3 + \dots \end{aligned} \quad (2.32)$$

Expanding the right side and discarding dc terms and harmonics, the two input frequency tones (fundamental tones) at the output appear to be:

$$\omega = \omega_1, \omega_2 : (\alpha_1 A_1 + \frac{3}{4}\alpha_3 A_1^3 + \frac{3}{2}\alpha_3 A_1 A_2^2) \cos(\omega_1 t + \phi_1) + (\alpha_1 A_2 + \frac{3}{4}\alpha_3 A_2^3 + \frac{3}{2}\alpha_3 A_2 A_1^2) \cos(\omega_2 t + \phi_2)$$

The output also contains intermodulation products (IM_n), for example, third order intermodulation products (IM_3):

$$\omega_3 = 2\omega_1 \pm \omega_2 : \frac{3\alpha_3 A_1^2 A_2}{4} \cos(2\omega_1 t + 2\phi_1 + 2\omega_2 t + 2\phi_2) + \frac{3\alpha_3 A_1^2 A_2}{4} \cos(2\omega_1 t + 2\phi_1 - 2\omega_2 t - 2\phi_2)$$

$$\omega_4 = 2\omega_2 \pm \omega_1 : \frac{3\alpha_3 A_2^2 A_1}{4} \cos(2\omega_2 t + 2\phi_2 + 2\omega_1 t + 2\phi_1) + \frac{3\alpha_3 A_2^2 A_1}{4} \cos(2\omega_2 t + 2\phi_2 - 2\omega_1 t - 2\phi_1)$$

If ω_1 and ω_2 are close, the components of IM_3 at ω_3 and ω_4 also appear close to ω_1 and ω_2 . Fig. 2.10 shows how the undesired IM_3 of two strong interferences falls in the signal band of interest and corrupts the desired signal.

In low-pass filter applications, a primary concern is the filter mixing two or more input tones to create (and redistribute signal power to) inter-modulation product frequencies. To assess a filter's non-linear inter-modulation, a dual signal frequency signal at ω_1 and ω_2 to the circuit is commonly applied to evaluate the 3rd order, Input-referred Intercept Point (IIP3). This is also known as a 'two tone test' in which the amplitude of the dual signal frequency $A_1 = A_2 = A$ set to be sufficiently small such that higher-order nonlinear terms are negligible and the gain of the circuit is relatively constant and equal to α_1 .

Figure 2.10: IM3 falls inband of interest

In Eq. 2.33, IIP3 is computed (in dBV¹) using two measurements; input power and the power of the mixing product. The power difference $\Delta P = P_1 - P_L$ is shown in Fig. 2.11. In Fig. 2.11 and Eq. 2.33, input power $P_1 \neq P_2$ are averaged because tone powers differ in filter testing when the test-tones are near the passband edge. The higher the IIP3, the more linear the circuit.

$$IIP3 = \frac{\Delta P}{2} + \frac{P_1 + P_2}{2} = \frac{P_1 - P_L}{2} + \frac{P_1 + P_2}{2} \quad (2.33)$$

where P_L is the power of IM_3 in dBV at ω_3 .

2.5 Sensitivity

When designing an active filter, there are generally more elements than needed to satisfy the design equations. The remaining elements are the degrees of freedom determined by the circuit designer to simplify the design equations, to achieve preferred component values, or to ease the manufacturing process. The sensitivity of a design component to a design can be viewed as the slope of a design target specification (e.g., quality factor or transfer function) in multi-dimensional element space at the nominal point. When the circuit designer selects a different free design element, that is, when

¹Reference impedance used is 50 Ohms

Figure 2.11: Power spectrum at the output of a nonlinear system. Two tones at ω_1 and ω_2 are applied at input. The output contains tones at ω_1 and ω_2 with power of P_1 and P_2 . The output also contains 3rd order intermodulation terms ω_3 and ω_4 with power of P_L and P_u .

the nominal design point is changed, the sensitivities are also altered and must be re-evaluated. The designer may consequently attempt to select the free element such that a point of minimum sensitivity (slope) is reached, thereby reducing the slope of the performance measure for sensitivity optimization. For example, the designer may reduce the sensitivity of the quality factor in a filter to the OTA's parasitics through appropriate selection of the single free parameter - the Gm ratio of the filter. The sensitivity analysis not only allows the designer to select the best circuits from the many available in the active filter literature, but also provides insights on whether a chosen filter circuit can meet specifications over manufacturing process variation [37].

Sensitivity Definition

Given a component x , any performance criterion P , such as the quality factor or a pole or zero frequency, will generally depend on x ; that is, $P = P(x)$. If the performance measure P represents the transfer function $H(s)$ or its magnitude or phase, then P

is also a function of frequency, so that we may write $P = P(s, x)$. An intuitively appealing method for determining mathematically the deviation in P caused by an error $dx = x - x_0$ of the element x is provided by a Taylor expansion of $P(s, x)$ around the nominal value x_0 of the component:

$$P(s, x) = P(s, x_0) + \frac{\partial P(s, x)}{\partial x} \Big|_{x_0} dx + \frac{1}{2} \frac{\partial^2 P(s, x)}{\partial x^2} \Big|_{x_0} (dx)^2 + \dots \quad (2.34)$$

If we assume that $(dx/x_0) \ll 1$ and that the curvature of $P(s, x)$ in the vicinity of x_0 is not overly large, we can neglect the second and higher derivative terms in Eq. 2.34 to obtain

$$\Delta P(s, x_0) = P(s, x_0 + dx) - P(s, x_0) \simeq \frac{\partial P(s, x)}{\partial x} \Big|_{x_0} dx \quad (2.35)$$

In most situations, the designer is less interested in the absolute changes ΔP caused by absolute changes dx and more interested in the relative changes. We can therefore normalized Eq. 2.35 to yield

$$S_x^P = \frac{x_0}{P(s, x_0)} \frac{\partial P(s, x)}{\partial x} \Big|_{x_0} = \frac{\partial P/P}{\partial x/x} \Big|_{x_0} = \frac{d(\ln P)}{d(\ln x)} \Big|_{x_0} \quad (2.36)$$

s_x^P is the sensitivity to a small change in a single parameter and

$$\frac{\Delta P}{P} \simeq s_x^P \frac{dx}{x} \quad (2.37)$$

In Eq. 2.37, the number s_x^P , also referred to as the variability of a performance measure P , is s_x^P times as large as the relative change of the circuit parameter x on which P depends. The lesser the sensitivities to their elements are, the smaller are the circuit's performance deviations. In other words, circuits with less sensitivity have larger tolerances dx/x . With component manufacturing variation, such circuits

Figure 2.12: Differential pair with active load

are less likely to drift out of the acceptable range R_A of specification.(see section 3.2)

Sensitivity of Nonlinearity to Device Variation

A filter complete transfer function can be described with a linear filter transfer function followed by a nonlinear section caused by the device operation, as depicted in Fig. 2.13. Due to device variations, for example in threshold voltage V_T or mobility μ , the operation of the device may enter a nonlinear region, which results in α_3 change in Eq. 2.31. In addition, when V_T changes, the transconductance changes, and as a result, the transfer function H (a function of α_1) changes as well. As previously noted, when α_1 and α_3 change, IIP3 changes.

A Gm cell design such as a differential pair can be used to analyze the sensitivity of non-linearity to device variation.

Consider a differential pair in Fig. 2.12 as an example of a Gm cell:

$$V_{IN} = V_{IN+} - V_{IN-} = V_{GS1} - V_{GS2}$$

$$I_{D1} + I_{D2} = I_{SS}$$

where

$$I_{D1} = I_{DA} + i_{d1} = I_{DA} + i_d \quad I_{D2} = I_{DA} + i_{d2} = I_{DA} - i_d \quad (2.38)$$

$$I_{DA} = \frac{1}{2} I_{SS}$$

In the saturation region

$$I_D = \frac{k'}{2} \frac{W}{L} (V_{GS} - V_T)^2 \quad (2.39)$$

where $k' = \mu C_{ox}$

and

$$g_m = k' \frac{W}{L} (V_{GS} - V_T) = [2k' \frac{W}{L} I_D]^{0.5} \quad (2.40)$$

$$V_{GS1} = V_T + \left(\frac{I_{D1}}{\frac{k'}{2} \frac{W}{L}} \right)^{0.5}$$

$$V_{GS2} = V_T + \left(\frac{I_{D2}}{\frac{k'}{2} \frac{W}{L}} \right)^{0.5}$$

$$V_{IN} = V_{GS1} - V_{GS2} = \left(\frac{I_{D1}}{\alpha} \right)^{0.5} - \left(\frac{I_{D2}}{\alpha} \right)^{0.5} \quad (2.41)$$

where $\alpha = \frac{k'}{2} \frac{W}{L}$

Applying Eq. 2.38 to Eq. 2.41, we obtain

$$V_{IN}^2 = \left(\frac{I_{SS}}{\alpha} \right) \left\{ 1 - \left[\left(1 + \frac{i_d}{I_{DA}} \right) \left(1 - \frac{i_d}{I_{DA}} \right) \right]^{0.5} \right\}$$

which we can rewrite as

$$i_d = \left(\frac{\alpha I_{SS}}{2}\right)^{0.5} V_{IN} \left[1 - \frac{\alpha V_{IN}^2}{2I_{SS}}\right]^{0.5}$$

and again as

$$\frac{i_d}{I_{SS}} = \left(\frac{d}{2}\right) \left[1 - \left(\frac{d}{2}\right)^2\right]^{0.5} \quad (2.42)$$

where d is the normalized input voltages.

$$d = \frac{V_{IN}}{(V_{GS} - V_T)} = \left[\frac{2}{\left(\frac{I_{SS}}{\alpha}\right)}\right]^{0.5} V_{IN}$$

Expanding Eq. 2.42 in the power series, we obtain

$$\frac{i_d}{I_{SS}} = \left(\frac{d}{2}\right) \left[1 - \left(\frac{1}{2}\right)\left(\frac{d}{2}\right)^2 - \dots\right] \quad (2.43)$$

Now if a sinusoidal input voltage v_1 is applied to this Gm cell:

$$v_1 = V_{INA} \cos \omega_1 t$$

where V_{INA} is the zero-to-peak value of the input sinusoidal tone. We obtain

$$I_{D2} = I_{DA} - i_d = I_{SS} \left\{ \frac{1}{2} - \left(\frac{1}{2}\right) \frac{V_{INA}}{(V_{GS} - V_T)} \cos \omega_1 t + \left(\frac{1}{64}\right) \left[\frac{V_{INA}}{(V_{GS} - V_T)}\right]^3 \cos 3\omega_1 t - \dots \right\} \quad (2.44)$$

From Eq. 2.44

$$\alpha_1 = \frac{I_{SS}}{2} \frac{1}{(V_{GS} - V_T)} = g_m \quad (2.45)$$

$$\alpha_3 = \frac{I_{SS}}{64} \left[\frac{1}{(V_{GS} - V_T)} \right]^3 = \frac{1}{8I_{SS}^2} g_m^3 \quad (2.46)$$

Differentiating α_1 and α_3 in Eqs. 2.45 and 2.46 with respect to the device parameter V_T , we obtain

$$\frac{d\alpha_1}{dV_T} = \frac{I_{SS}}{2} \frac{1}{(V_{GS} - V_T)^2}$$

$$\frac{d\alpha_3}{dV_T} = \frac{3I_{SS}}{64} \frac{1}{(V_{GS} - V_T)^4}$$

or

$$\frac{d\alpha_1}{\alpha_1} = \frac{dV_T}{V_{GS} - V_T} \quad (2.47)$$

$$\frac{d\alpha_3}{\alpha_3} = 3 \frac{dV_T}{V_{GS} - V_T} \quad (2.48)$$

We obtain

$$S_{V_T}^{\alpha_3} = 3S_{V_T}^{\alpha_1} = \frac{3V_T}{V_{GS} - V_T} \propto g_m \quad (2.49)$$

The non-linearity term α_3 is three times more sensitive to V_T variation than α_1 . From Eq. 2.49 and 2.45, the sensitivity of the non-linearity term to device variation in Eq. 2.49 is proportional to g_m .

Sensitivity calculations can suggest potential deviations of the slope of the nominal filter function with respect to the circuit elements at the nominal design values. Even if a multi-parameter statistical sensitivity measure is used, the designer can arrive only at an estimate of the expected mean deviations caused by a set of components with known tolerances. However, in a practical design environment, filter requirements are

Figure 2.13: Filter nonlinearity

usually not prescribed precisely, but instead are only specified with certain tolerances or limits. For example, system requirements may call for a low pass filter to have no more than A_{pass} (in dB) attenuation in the passband and at least A_{stop} (in dB) attenuation in the stopband. In order to evaluate the manufacturing yield with a given fabrication process, and understand and reduce the impact of shifting to a more expensive and tighter tolerant fabrication process, statistical simulation, such as a Monte Carlo simulation, of the filter design is still required. The result of the Monte Carlo simulation can help to identify a better nominal design that still meets the imprecise passband and stopband requirements of $\leq A_{pass}$ and $\geq A_{stop}$, and at the same time results in a lower number of rejected samples through design centering (see section 3.3).

Chapter 3

Tolerance Design

3.1 Introduction

Electronic circuits are designed based on the specifications dictated by the customer. The circuit designer proposes a first-trial design based on rules of thumb deriving from his/her own experience or from an earlier circuit designed for a similar function and specification. After several simulations, possibly in the course of a number of iterations, the circuit is gradually modified until its performance eventually meets the specification. In mass production, all devices (e.g., transistors, capacitors, resistors) within the manufactured circuit have associated tolerances, that is, the device variation, such that the value of each device differs to some random degree from its nominal value. As the device values of each manufactured circuit are generally different from the ones designed, it is expected that each circuit's performance will differ from that of the simulated nominal circuit. Therefore, without some remedy, the manufacturing yield of a circuit is compromised, which is very costly both in time and in skilled manpower. Given that the unsatisfactory yield is caused by device tolerances (or variation) and that with zero tolerances (that is, with the nominal circuit) the specifications are not violated, a proposal to tighten the device tolerances seems reasonable. However, it is usually the case that the cost of a device is an inverse function of its tolerance. Therefore, while the yield may be increased through this approach, the circuit will cost more to manufacture. If we keep the device tolerances

fixed in order to hold the total device cost of the circuit constant, there are two possible ways to improve the yield: design centering and reduction of the sensitivity of tolerance. This chapter describes how to optimize the yield through design centering and reduction of sensitivity of tolerance by tolerance analysis [40].

3.2 Tolerance Analysis

In integrated circuit (IC) design, tolerance analysis is a method to understand the sources of variation in devices propagated across circuit designs and the performance of a design in achieving its design requirements within the process capabilities of manufacturing. In a two-dimensional circuit performance space, there exists a region, called the region of acceptability (R_A), where both circuit performances meet the requirement. In the same space, a region called the tolerance region (R_T) reflects the realized circuit performances due to the tolerances on the devices. Monte Carlo analysis is one of several different methods of tolerance analysis to estimate R_T . In a Monte Carlo simulation [12], each randomly varying parameter x_i in a device is replaced by a random number generator that produces pseudo random values r_i with the same statistical properties as those of x_i . For each parameter value generated, the performance measure P_i is evaluated and the process is repeated N times. Here, N must be large enough (usually $100 < N < 10,000$) to obtain results that are statistically significant (i.e., the statistical properties of measured P_n s may be estimated).

If the statistical device variation are modeled correctly by the samples r_i obtained from the random number generator, Monte Carlo simulation offers engineers a realistic picture of practical circuit performance . This means that the samples r_i must have the same statistical properties and distributions as the circuit parameters x_i they represent. A strong statistical model requires detailed knowledge of the fabrication or process environment and the availability of extensive manufacturing data from which

the statistical properties can be determined. Monte Carlo simulation is conventionally used to simulate the final design and provides a useful estimate of the manufacturing yield, that is, the percentage of circuits that will meet the design specifications when the device varies due to fabrication tolerances or drift during operation. However, as the circuit size and number of design alternatives grow, extensive computer time is required.

In this research, Monte Carlo SPICE simulation is limited due to the use of library characterization. Tolerance analysis for the final designs is performed by synthesizing the designs from the library without the need for time consuming Monte Carlo SPICE simulation.

3.3 Design Centering

Fig 3.1 illustrates the process of design centering, which is accomplished by adjusting the nominal values of the device parameters (e.g. transistor width/length, capacitance, resistances) while leaving their tolerances fixed (i.e., same process technology) such that R_T is more centrally located within R_A to increase the manufacturing yield. The primary goal of design centering is to identify a nominal circuit design that is centered in the R_A sufficiently far away from its borders such that the prescribed specifications are met.

3.4 Reduction of Sensitivity of Tolerance

An additional method to the improve yield without changing the device tolerance (i.e., keeping the process technology the same) is to reduce the sensitivity of tolerance to the design. This is accomplished by identifying an alternative design that has less performance variability due to device tolerance, as demonstrated in Fig. 3.2. R_T represents the tolerance region for design #1. R'_T represent the tolerance region for

Figure 3.1: Design centering by designing the circuits at R'_T instead of R_T to achieve higher yield. R_A is the Region of Acceptance (acceptable circuit performance) and R_T is the Tolerance Region (expected manufacturing variation).

Figure 3.2: Tolerance tightening (Reduce R_T). Different design alternatives result in different R_T . Smaller R_T in design 2 represents less performance variability comparing to design 1 due to device variation.

design #2. To achieve a design that falls in R'_T rather than R_T , circuit designers can architect the circuit topology such that the performance variability over a satisfactory fraction of the manufactured samples is reduced.

3.5 Manufacturing Figure of Merit (FOM) : Defect level and Yield loss

Several figures of merit, such as yield loss and defect level, are evaluated before and after manufacturing circuits in mass volume in order to reduce costs. **Yield (Y)** is defined as the number of units passing the manufacturing test process divided by the number of units entering the process over a specified period of time. **Defect level (DL)** is the percentage of units shipped to customers that are defective. The defect

level can be estimated in terms of yield according to Eq. 3.1, where **test coverage (TC)** is the percentage of defects that can be found during the manufacturing test (with the test limit) versus the total number of possible defects [41,42].

$$DL = 1 - Y^{1-TC} \quad (3.1)$$

As depicted in Fig. 3.3, the manufacturing circuits are tested before shipping to the customer. The test limits in the **TEST** condition are determined by the limit that best represent the actual customer **USE** condition. The selection of test limits can be trivial, as the test process is usually confined to a constraint environment and for a short period of time to keep the cost down, while in the actual use condition, the circuits are operated in diverse scenarios for a longer period of time. When the test and use conditions are not perfectly correlated due to limited test resources such as test time and test equipment (i.e., test converge is not 1), there will be defective units that are not detected. The percentage of units that pass the test but fail in use (PF) is called test-escape (TE). Meanwhile, the percentage of units that pass in use but fail in the test (FP) is called over-kill (OL). In an ideal case, in which the test and use conditions are perfectly correlated, $TE=0$ and $OL=0$.

Figure 3.3: Test process

Chapter 4

Analog Statistical Design in Linear and Nonlinear Response for Manufacturing

4.1 Introduction

Mixed-signal Application Specific IC (ASIC) and System-on-Chip (SoC) integrate analog components such as PLLs, IOs, filters, analog-to-digital and digital-to-analog converters, thermal sensors, etc. within larger digital systems with component counts and areas considerably less than the digital system. Several analog design solutions exist that meet the required specifications. However, each design alternative, that is, different topology, transistor implementation, responds differently to manufacturing variations and may require different testing strategies [43] resulting in different sensitivities to the final SoC product manufacturing and quality Figures-of-Merit (FoM), viz. yield, test-escape, overkill and defect level.

Although analog design methods are well established, a persistent problem is characterization in the presence of process variation. It is difficult to include manufacturing and quality figures-of-merit affected by process variation in the design methodology since statistical data for all performance metrics are needed. In principle, the statistics can be obtained from Monte Carlo SPICE simulation of many samples of the circuit in transistor level. Conventional Monte Carlo approach is compute-intensive and limits the circuit complexity that can be explored, to find an optimal design alternative for given Test and Use scenarios. Efficient Monte Carlo

sampling methods such as quasi-Monte Carlo [32], latin hypercube sampling [33], Bayesian inference [34], etc. have been proposed. However, these approaches still require transistor simulation on each analog design alternative to estimate their corresponding yields. Fast estimation of performance metrics of analog systems using analytical performance models are proposed in [44, 45] which can be used to quantify manufacturing and quality FoMs. These approaches require finding the sensitivity of various parameters of interest of the analog systems to different process and environmental conditions which is non trivial for complex systems in nanoscale technologies, since bias dependent short-channel effects (such as V_T roll-off, velocity saturation and drain-induced barrier lowering or DIBL) [46] and layout dependent effects (such as well-proximity and shallow trench isolation or STI stress) [47, 48] increase the dimensionality of analytical models. Not having a highly efficient methodology to obtain manufacturing FoMs for analog design alternatives limits the ability to analyze trade-offs that can be made for different designs, component counts, areas, manufacturing yield, and quality early in the product life cycle. When a final design does not meet manufacturing specifications, difficult choices must be made to either tolerate low SoC yield or integrate a different analog design into the SoC and invest more expensive design time.

An alternate approach to analog system design is to synthesize analog circuits, including topology selection, transistor sizing, and hardening synthesized circuits for process variation and layout generation. Reasoning-based topology synthesis method to design a low power opamp is shown in [49]. Higher level circuits, such as filters, are synthesized using the concept of building blocks (op-amps, resistors and capacitors) in [50] using tabu search heuristic that sizes transistors for optimal area and performance. Synthesis of more complex analog circuits such as ADC, PLL and filters by leveraging digital design tools for accelerating circuit layout have been proposed

in [51–54]. These approaches still require SPICE simulations to obtain quantitative statistical measures for the performance of analog circuits and estimate the manufacturing and quality FoMs. Parametric variations aware circuit synthesis is described in [55–57] where the synthesis algorithm generates transistor sizing taking into account process and supply variations to generate OTAs. Variation aware VCO design synthesis using DoE assisted Monte Carlo simulations is shown in [58]. However, the yield aware analog circuit synthesis is still limited to small building blocks such as op-amps and VCOs.

This chapter proposes an analog system design methodology that enables evaluation of manufacturing and quality FoMs in order to evaluate different design alternatives in early product life cycle. The analog systems are built using building blocks [50]. Instead of using analytical models [44, 45], the linear and non-linear responses of the building blocks are pre-characterized for process and environmental variations using SPICE simulations. Using the pre-characterized library of building blocks, circuit performance, parametric manufacturing yield and defect level of different topologies are evaluated without additional SPICE simulations for 200kHz anti-aliasing or reconstruction filter that can be used for GSM/DECT receiver [59].

This chapter is organized as follows: The motivation and main contribution of this work are described in Section-4.2. The proposed building block approach for design for manufacturing is described with filter as example in Section-4.3. The characterization of the building block library is shown in Section-4.4. Sections-4.5 and 4.6 describe the methodology used to estimate linear and non-linear performance of filters. Section-4.7 shows the results of evaluation for different filter design alternatives to maximize yield and minimize defect level.

4.2 Motivation and Contributions

Fig. 4.1 shows the flow of circuit components from the silicon fabrication factory (Fab) through the Test factory and finally to the customer (Use). Because of process variation in the Fab, a fraction of devices produced by the Fab, shown in red in the figure, will not satisfy the performance requirements in the published datasheet, called the Use Specification. The Use Specification gives the performance limits assumed by the customer's system designer when designing a system using the components. The Use Specification is also the component producer's definition of a *good* component. For high-performance components, the fraction of components produced by the Fab not meeting the Use Specification is usually high due to the circuit is operating at the technology limit and must be controlled by screening the population of components produced by the Fab using the manufacturing Test operation in the producer's factory.

Producer and OEM, that is, the party using the design component, manufacturing cost models require quantitative statistical Figure-of-Merit (FoMs). However the usual method of simulating analog circuits at process and environmental corners does not quantify parametric variation at early design phase reflected in manufacturing FoMs such as yield loss at the producer and shipped defect levels (quality) received by the OEM.

Once Test and Use specifications are set, manufacturing and quality FoM, yield loss (YL), overkill (OL) and defect level (DL), are computed as follows:

$$\begin{aligned}
 YL &= \frac{FF + FP}{FF + FP + PF + PP} \\
 OL &= \frac{FP}{FF + FP + PF + PP} \\
 DL &= \frac{PF}{PF + PP}
 \end{aligned} \tag{4.1}$$

Figure 4.1: Relationship of Figure-of-Merit to population category probabilities in the context of an analog filter. Without Test phase, a fraction (red) of the population of components would fail in Use condition. Test reduces the customer-perceived defect level (DL) by trimming and screening components. Reducing DL increases yield loss (YL) and overkill (OL).

where FF is the number of samples that fail in both test and use conditions, FP is the number of samples that fail in test condition, but passes in use condition, PF is the number of samples passing in test condition, but failing at use condition and PP is the number of samples that pass in both test and use conditions. YL is the yield loss at Test and does not depend on the Use condition. OL is the part of yield loss attributed to the tester and test method, called overkill. DL is the end-user perceived fraction not meeting the Use specifications, and hence it is the quality figure of merit [60].

The proposed methodology implements analog systems by employing scalable building blocks of circuits which preserve correct process statistics when used in a larger system. The library of building blocks is built with transistor level simulation of many process samples (Section-4.4) avoiding the need for complex analytical multidimensional modeling, especially for non-linear responses in highly scaled processes. The methodology

1. Minimizes SPICE simulation to predict linear and non-linear response of design alternatives by reusing small and large-signal SPICE simulations from an analog building block library.
2. Estimates - early in the design cycle - manufacturing variation effects on a filter design's linear and non-linear response and verifies test specification setting for improved correlation between test and use.
3. Provides a quantitative comparison of manufacturing and quality FoMs for filter design alternatives and without transistor level SPICE simulation.

The methodology is demonstrated by evaluating two different design alternatives for a continuous-time low pass filter with a pass-band gain of 0dB, bandwidth of 200kHz and stop-band rejection of >30dB at 435kHz in a 45nm CMOS technology to maximize the manufacturing and quality FoMs. Cascaded biquads (Fig. 4.2) are used

Biquad Transfer Function

Figure 4.2: (a) A 2N-order low pass cascaded biquad Gm-C filter using N biquads. (b) Each biquad is constructed with transconductor (Gm) and capacitors (C).

to realize the low-pass filter transfer function, $H_F(s)$, in (4.2). A biquad allows for a scalable design since different transfer functions can be obtained by minor changes to the component values without changing the circuit topology [36]. The biquads can be implemented either using Opamp-RC or Gm-C topologies depending on the power, noise and linearity requirements of the filter and for this demonstration, a Gm-C topology was chosen since it can be implemented with a lower power to achieve a given bandwidth as well as being more compact compared to OpampRC filters [37]. The gain (K), pole frequency (ω_0) and quality factor (Q) of the individual biquads, $H_i(s)$, are chosen to achieve the desired frequency response.

$$H_F(s) = \prod_{i=1}^N H_i(s) \quad (4.2)$$

$$H_i(s) = \frac{K_i \omega_{0i}^2}{s^2 + s(\omega_{0i}/Q_i) + \omega_{0i}^2} \quad (4.3)$$

4.3 Building Block Methodology for Manufacturing and Quality FoM Prediction

As described in the previous section, a cascaded biquad filter design is used to demonstrate the design methodology for manufacturing. Filter *use requirement* such as DC gain, cut-off frequency, in-band-ripple, stop-band attenuation and linearity are listed in Table 4.1. These design specifications can be achieved with continuous time filter architectures such as cascaded biquads or ladder filters. Either of these architectures can be realized using active circuit topologies such as OPAMP-RC or Gm-C. Furthermore, the filter specification can be met by a plethora of filter order and types (e.g. Butterworth, Chebyshev, elliptical, etc.). Since the specifications can be met using different filter types, order, topology or architecture, it is important to evaluate which design alternative will perform best considering all aspects of the product. Table 4.1 also lists the test specifications for the filter. This is required because during the test phase, it is impractical (and often not needed) to characterize the filter AC and non-linear response across entire use condition due to the large test times involved. In order to reduce the test time and cost, the filter is often only tested at few frequencies. The manufacturing FoMs are evaluated at these test specifications.

Table 4.1: Filter use requirements and test specifications at 27°C.

Passband		Use		Test
Cut-off	f_p	200kHz		
Ripple	A_r	± 3.5 dB		± 3 dB @ 140kHz
Gain	K_p	0dB		Trim
Stopband				
Cut-off	f_s	435kHz		
Gain	A_s	< -30 dB		< -12 dB @ 300kHz
Linearity				
IIP_3	IIP_3	> -16 dBV @ 100kHz		> -16 dBV @ 140kHz

Fig. 4.3 shows the building block methodology for cascaded filter design where the biquad is used as the building block. To realize a given transfer function, a cascade of these biquads with slight variation in the component values can be used [36]. The basic building block (in this case the biquad) can be designed using traditional analog design flow or using synthesis methodologies described in [49, 58, 61]. Each biquad circuit is Monte Carlo simulated for many samples to capture the effect of process variation. Optionally, advanced simulation methodologies [32–34] can also be used.

The specifications in Table 4.1 can be broadly classified into small-signal (or linear) characteristics or large-signal (or non-linear) characteristics. Transfer-function (DC gain, cut-off frequency, inband ripple, stop band rejection, etc.) are dictated by the small signal responses and can be evaluated using transfer functions of cascaded blocks as shown in (4.2). Characteristics such as gain compression, harmonic distortion and intermodulation can be determined by the non-linear response of the system as described in (4.4)

$$v_{out} = \sum_{i=0}^n \alpha_i v_{in}^i \quad (4.4)$$

where v_{out} is the output voltage, v_{in} is the input and α_i are the coefficient of the Taylor series describing the large signal behavior of the system. The third order intermodulation (IM_3) is a useful metric to quantify non-linearity [39] when the third order distortion (or α_3) is the dominant source as is the case with most circuits.

Transfer function parameters, K , Q and ω_0 and large signal parameter IM_3 are extracted for each sample and stored in the Biquad library: for parameters ($K_{p,m}$, $Q_{p,m}$ and $\omega_{0p,m}$), each Monte Carlo sample are extracted from the simulated AC transfer function by least square fitting to the ideal transfer function (4.3). The transient simulation is performed to estimate parameter IM_3 for each Monte Carlo sample. Filter design has two parts: First the *design center* is established, and second, variation around the center is characterized. Mean values of biquad K , Q

and ω_0 parameters from the Biquad Library are used to establish the design center. Then, to characterize variation, K , Q and ω_0 parameters of a sample of filter instances is created by *scaling* from biquad K , Q , ω_0 , and IM_3 parameters sampled from the Biquad Library. Scaling transforms samples of Biquad Library parameters to samples of biquads that match filter requirements while preserving the proper statistics of the parameters. Bootstrap sampling or generation from the parametric statistical model is used to create K , Q , ω_0 and IM_3 samples from the Biquad Library. Bootstrap sampling method is a resampling method by independently sampling with replacement from an existing relatively small sample data and performing inference among these resampled data. This method enables us to construct reliable confidence intervals and conduct valid hypotheses test [62].

Filter design requires exploration of many hypothetical design, test, and use scenarios as shown in Fig. 4.3. Construction of the filter transfer function and nonlinear response from samples of biquad K , Q , ω_0 and IM_3 parameters, rather than circuit simulations requires much less computation. In fact, the what-if design/test/use scenarios for the filter design examples given in this demonstration were coded in Python and were executed in real time on a PC.

4.4 Biquad Library and Characterization

The generation and characterization of the analog library is key to the entire design flow. For cascaded biquad filter, the building-block is a biquad shown in Fig. 4.2. The frequency response of the biquad can be represented using a closed form equation as a function of Gms and the capacitors. The library contains practical values for the ratio of Gm_1 to Gm_4 . This is analogous to digital standard-cell library which typically

Figure 4.3: Building Block methodology: Biquad as building block replaces the compute intensive SPICE simulation of the filter in Test/Use and building block level design optimization by much less compute intensive sampling and scaling of the (K, Q, ω_0, IM_3) statistical model. SPICE simulations to establish the (K, Q, ω_0, IM_3) biquad model need be done only once.

contains a subset of all possible sizing for digital gates. The analog library used in this demonstration (Fig. 4.4) consists of three biquad designs (r_k where $k \in [1, 2, 3]$) which has unique values of biquad Gm ratios (Gm_{1-4}) for demonstration. Each biquad is characterized for four different Qs (Q_n , where $n \in [1, 2, 3, 4]$). The Q of the biquad is tuned by the capacitor ratios of the biquad. The biquad is characterized for its frequency response using small-signal simulations. Non-linear response at different frequencies ($f_{norm,j}$, where $j \in [1, 2, \dots, 8]$) normalized to the pole frequency (ω_0) are characterized using two test tones separated by 5% around f_{norm} for large-signal simulations [39]. As we will see in Section-4.6, both magnitude and phase of the IM_3 are required to estimate the overall filter IM_3 from the building blocks. The large

Figure 4.4: Biquad library: building-block library for the Gm-C biquads comprising of small-signal and large signal characterization as well as process and mismatch data.

signal simulation is performed to estimate the IM_3 of each biquad and stored in the library. In order to capture process variation (i.e., fast, slow, normal corners) and device mismatch due to manufacturing variability, Monte Carlo SPICE simulations are run on each biquad $B_{r,Q}$. The effect of process ($p \in [1, 2, \dots, 250]$) and mismatch ($m \in [1, 2, \dots, 5]$) on $K_{p,m}$, $Q_{p,m}$, $\omega_{0p,m}$ and IM_3 are stored and each biquad sample in this library can be indexed by its four vectors, viz. Gm ratio, Q , process and mismatch identifiers ($B_{r,Q,p,m}$). Note that the biquad filter transfer function parameters ($K_{p,m}$, $Q_{p,m}$ and $\omega_{0p,m}$) for each Monte Carlo sample are extracted from the simulated AC transfer function by least square fitting to the ideal transfer function (4.3). Performing Monte Carlo SPICE simulations on a small building-block is very efficient and eliminates the need for doing these simulations for higher level circuits. The fully characterized building-block now contains all the information necessary for the subsequent steps and is shown in Fig. 4.4.

The analog building block characterization must identify and retain non-ideal effects as well as include the effects of component variation from processing and the effects of component mismatch. The characterization could also include the variation

1	K_1	Q_1	$\omega_{0,1}$	$ IM_{3,1} $	$\angle IM_{3,1}$
2	K_2	Q_2	$\omega_{0,2}$	$ IM_{3,2} $	$\angle IM_{3,2}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
pm	K_{pm}	Q_{pm}	$\omega_{0,pm}$	$ IM_{3,pm} $	$\angle IM_{3,pm}$

Figure 4.5: The database organization and storage of each biquad design characterization.

of biquad capacitor (C1 and C2) as well as parasitic interconnect components in the SPICE simulations. Interconnect parasitics can be estimated [63] and be absorbed into the filter design capacitor scaling step. Integrated Gm-C biquad circuits are dominated by two non-ideal Gm effects; 1) frequency dependent Gm due to internal parasitic poles and; 2) finite input (G_i) and output (G_o) admittances [37]. Parasitic poles alter an Gm-C biquad transfer function and a finite Gm output admittance limits an Gm-C biquad operating Q. In this work, the effects of a parasitic pole changed biquad parameterization by less than 1 part in 10^5 and were ignored in filter design. The non-ideal Gm output admittance was significant, $G_o \approx 0.05G_m$ and was included during filter design.

The database organization and storage of its design characterization can be visualized as a table shown in (4.5). Each row retains the characterization data for one Monte-Carlo sample of a specific library biquad design $\mathbf{r}_k$ of quality factor Q_n and normalized frequency $\mathbf{f}_{norm,j}$ used for IM_3 simulations.

The biquad sample index unique to a specific biquad design and specific Monte Carlo sample ranges from 1 to $p \times m$ (1250 in this demonstration). The data in each row is one example of the non-ideal effects and transistor variation on a biquad design. Simulation data for each biquad design are randomly selected during bootstrap generation of a sample of a filter and will be described in Section-4.5.

4.5 Estimation of Filter Frequency Response Using Biquad Library

The quality factor, Q_i , and the pole-frequency, ω_{0i} for the i^{th} biquad in the filter that needs to be generated could be different from that in the analog library. Based on the biquad transfer function in Eq. 4.3, the biquads in the library can be mapped to the required Q and ω_0 for i^{th} stage biquad by scaling the capacitors C_1 and C_2 from a biquad design r using the following equations

$$\begin{aligned} S_{1i}(Q_i, \omega_{0i} | Q_r, \omega_{0r}) &= \frac{\omega_{0r}}{\omega_{0i}} \frac{Q_i}{Q_r} & C_{1i} &= S_1 \cdot C_{1r} \\ S_{2i}(Q_i, \omega_{0i} | Q_r, \omega_{0r}) &= \frac{\omega_{0r}}{\omega_{0i}} \frac{Q_r}{Q_i} & C_{2i} &= S_2 \cdot C_{2r} \end{aligned} \tag{4.5}$$

The flow chart explaining the filter generation is shown in Fig. 4.6(a).

Multiple filter samples can be generated by bootstrapping. Fig. 4.7 shows an example to generate multiple filter samples for a 4th-order filter. In order to generate a filter sample, biquads from the same process index, p , are selected from the library. If biquad design r_k is used more than once, the filter sample is generated by selecting biquads from the same process p , but with different mismatch indexes m . For example, if all biquads used in the filter are from the same biquad design, there are $P(M, N)$ filter samples can be generated for each process p where M is the total number of mismatch simulations for each process index in the biquad library and $2N$ is the order of the synthesized filter. Repeating this for each process index, a total of $P \times P(M, N)$ filter samples can be created. If the filter uses different biquad designs for each stage, then it is possible to generate even higher number of samples. All permutations of biquads from r_i yields $P \times P(M, 1)^N$ filter samples. The small-signal filter performance (filter transfer function) for the generated filter samples is calculated using (4.2).

Figure 4.6: (a) Filter synthesis flow and (b) flow chart to estimate filter transfer function using analog library.

Figure 4.7: Generating multiple filter samples using bootstrap sampling for 4th-order filter (a) using biquads from the same biquad design r_1 and (b) using biquads scaled from different biquad designs, r_1 and r_2 . Each line indicate a possible sampling combination.

4.6 Estimation of Linearity (IIP₃) using Biquad Library

There are well known equations to calculate linearity of cascaded systems, taking into account the gain and non-linearity of individual stages [39]. For example, the magnitude of third-order intermodulation distortion product (IM₃), A_{IP3} , of a cascaded system with three or more stages is given by

$$\frac{1}{A_{IP3}^2} = \frac{1}{A_{IP3,1}^2} + \frac{G_1^2}{A_{IP3,2}^2} + \frac{G_1^2 G_2^2}{A_{IP3,3}^2} + \dots \quad (4.6)$$

where $A_{IP3,1-3}$ are input IP3 of stages 1–3, and G_1 and G_2 are the voltage gains for stage-1 and stage-2 respectively. However (4.6) assumes that the intermodulation products from different stages add in power and does not take into account the phase of the IM₃ products from different stages. As a result, the calculation does not

yield accurate results. Fig. 4.8 shows the mechanism by which the intermodulation products generated in different stages interact with each other. Two fundamental frequency tones at f_1 and f_2 , applied to the input goes through $\mathbf{H}_1(\mathbf{f})$ and $\mathbf{H}_2(\mathbf{f})$ to the output of the second stage. IM_3 products are generated by the first stage at frequencies f_3 and f_4 . Similarly, IM_3 products are generated in stage-2 from the fundamental tones at its input. The total IM_3 voltage at the output of the second stage is given by the vector sum of IM_3 voltage propagated from stage-1 and the IM_3 voltage generated in stage-2. At the output of stage-i, the IM_3 voltage can be derived

$$\begin{aligned}
|IM_3|_i = & ||IM_3|_{i-1} \times G_i(f_3) \times \angle IM_{3,i-1} \\
& + |IM_3|_i \times G_{i-1}^2(f_1) \times G_{i-1}(f_2) \times \angle IM_{3,i} \\
& + 2 \times \theta_{i-1}(f_1) - \theta_{i-1}(f_2)|
\end{aligned} \tag{4.7}$$

where G_i is the voltage gain, $\theta_i(f_t)$ is the linear phase shift at (f_t) , $|IM_3|_i$ is the magnitude and $\angle IM_{3,i}$ is the phase of IM_3 voltage at stage-i. The power of IM_3 in dBV (i.e., P_L in Eq. 2.33) at stage-i is $20\log(|IM_3|_i)$. As shown in Fig. 4.8, the final non-linearity at the output of two stages can be expressed in terms of the non-linearity of individual stages and the appropriate frequency response of the different stages. By storing this information in the biquad library and extending this concept to N-stages, it is possible to accurately estimate the IIP_3 of a $2N^{\text{th}}$ biquad filter.

Filter synthesis described in Section-4.5 scales the capacitors using (4.5) of the biquad library elements to realize the required filter. Since the element stored in the biquad library can be characterized in a different Q and ω_0 from the ones used in filter to synthesize, the final biquads in the synthesized filter will have different IM_3 characteristics compared to the elements in the biquad library since the cap scaling changes the Q and ω_0 . Additionally, the frequencies at which the IM_3 in biquad library are characterized could be different from where the test-tones are applied for the filter use condition. Both these challenges can be addressed by performing

Figure 4.8: Modeling of IM_3 voltage for cascaded stages using magnitude and phase of the generated and propagated IM_3 voltage terms from different stages.

a bilinear interpolation of the IM_3 voltage magnitude and phase ($|IM_3|$ and $\angle IM_3$) from a grid of Q and normalized test-tone frequency, f_{norm} , of the selected filter biquad sample as shown in Fig. 4.9. Since the IM_3 voltage magnitude and phase varies with process and mismatch, it is important to note that the bilinear interpolation has to be performed on each process and mismatch indexed biquad.

Multiple filter samples around the design average can be generated by bootstrap-

Figure 4.9: Bilinear interpolation for (a) IM_3 magnitude and (b) IM_3 phase at arbitrary (Q_i, f_i) for a single process and mismatch indexed biquad sample.

Figure 4.10: Flow chart for (a) generating biquad IM_3 in synthesized filter and (b) calculating filter IIP_3 .

ping as described in Section-4.5 and the synthesis of multiple filter IIP_3 response is shown in Fig. 4.10. This section outlined the use of bootstrap sampling of the biquad standard-cells in a filter design. Scaling and interpolation of biquad characterization data reflects the cell's use in specific filter designs. In the next section, efficient bootstrap filter sampling is used to permit sufficient sample sizes to obtain statistically meaningful manufacturing FoMs such as AC and IIP_3 yield loss or customer defect level.

Figure 4.11: Ideal filter transfer functions (a) 4^{th} -order Chebyshev and (b) 8^{th} -order Butterworth. Additional plots are library biquad component transfer functions after scaling to filter design specifications. Use requirement upper and lower magnitude brick-walls are shown as dashed lines.

4.7 Results: Comparison of Design Alternatives

Two filter design alternatives are used for demonstration. The two design alternatives—a 4^{th} -order 1dB ripple Chebyshev filter and an 8^{th} -order Butterworth filter—meet the same requirements in use and are tested to the same specification. Test and use specifications are summarized in Table-4.1. The use requirement low-pass are shown with the dashed lines in Fig. 4.11(a-b). The ideal filter transfer functions are shown as the bold trace. The 4^{th} -order Chebyshev and 8^{th} -order Butterworth are designed based on a common library of pre-characterized biquad building block. The component biquad ideal transfer functions are included in Fig. 4.11.

Each filter’s biquad library usage and biquad parameters are summarized in Table-4.2. In general, many library biquad building block could be selected to complete a given filter’s design. In this demonstration, specific biquad choices (and biquad ordering) were made to highlight important features of design scaling, performance estimation and the computation of manufacturing statistical Figures-of-Merit (FoM). The biquad ordering determines the AC signal flow from input to output. The filter’s

input is accepted by biquad #1 and continues, in sequence, to biquad #N which delivers the filter's output. The maximum value of N is set by the filter order. The filter biquad ordering is from minimum to maximum Q.

Table 4.2: Filter Design Specifications and Biquad Library Usage
Filter Design Alternative

Filter Biquad (#c)	4th-Cheb		8th-Butter			
	#1	#2	#1	#2	#3	#4
$K_{\#c}[\text{V/V}]$	0.944	0.944	1.000	1.000	1.000	1.000
$Q_{\#c}$	0.785	3.559	0.510	0.601	0.900	2.563
$f_{p\#c} [\text{kHz}]$	106	199	218	218	218	218
$S_1 \cdot C_1 [\text{pF}]$	237	1057	78	87	134	550
$S_2 \cdot C_2 [\text{pF}]$	287	72	219	185	120	58
Library	#1	#2	#1	#2	#3	#4
Gm	1:1:1:1	4:1:4:1	1:1:1:1	1:1:1:1	1:1:1:1	2:1:2:1
$C_1 [\text{pF}]$	140	540	140	135	135	270
$C_2 [\text{pF}]$	67	270	270	270	135	135
S_1	1.691	1.953	0.541	0.642	0.989	2.032
S_2	4.247	0.266	0.810	0.683	0.887	0.427

OTA transconductance base value, $g_m=170\mu\text{S}$

Section-4.5 explains the flexibility to set biquad Q , and ω_0 is achieved by scaling two biquad design capacitors, C_1 and C_2 . In Table-4.2 each component biquad $S_{1\#c}$ and $S_{2\#c}$ scaling parameter values (and each biquad's capacitor values) are given.

The filter generation methodology described in Section-4.5 is used to create 250 filter samples each for the 4th-order Chebyshev filter and 8th-order Butterworth filter and estimate each sample's AC and IM_3 (and related IIP_3) response. Sample size is an important consideration when deciding which of design alternative meets manufacturing yield and quality requirements. Roughly speaking larger sample sizes increase decision confidence. For a given sample size the design alternative with fewer test fails is the design alternative with higher manufacturing yield. Similarly, customer quality level increases with the design alternative that has fewer test escapes. For a sample size of 250, a greater than 90% confidence level for yield and quality FoMs is

possible when test fails or test escapes counts are less than ten.

To select a filter design without direct SPICE simulation requires validation of the alternative bootstrap estimates to the traditional SPICE simulation. Building block characterization data for AC and inter-modulation were used to compute bootstrap estimates which were compared to results obtained from SPICE simulation. Sets of sample filters were assembled by scaling a bootstrap selected samples of biquad standard-cells. To permit direct comparison, the generation of each filter sample netlist was carefully controlled and SPICE netlists were manually confirmed to be identical. AC validation was established by matching the bootstrap transfer function to corresponding SPICE simulated transfer function.

Figure 4.12: (a) *4th*-order Chebyshev and (b) *8th*-order Butterworth. Filter design alternatives IIP_3 values are compared by scatter-plot and by marginal histograms. Comparison of each simulation and estimation uses 250 filter samples. Each data point in the plots is one filter sample. SPICE IIP_3 evaluation is from transient analysis. IIP_3 synthesized are from interpolation. IIP_3 corner frequency is 140kHz for both filters

Bootstrap estimation of filter inter-modulation based on pre-characterized biquad introduces two potential sources of miscorrelation to direct SPICE filter simulation.

1) Pre-characterization set equal test-tone power levels whereas within the filter the power levels may be different. The power difference is greatest when the test-tone

center-frequency is near the filter’s low-pass pass-band edge. 2) Pre-characterization IM_3 data is collected for a limited number of specific test-tone frequencies and offsets and biquad Q and cut-off frequencies.

Table 4.3: IIP_3 Correlation between SPICE and synthesis

		Filter Design Alternative			
		4th-Cheb		8th-Butter	
Center Frequency [kHz]		100	140	100	140
IIP3 Median [dB]	SPICE	-13.5	-13.3	-14.2	14.1
	Synthesis	-13.5	-13.4	-13.8	13.3
Std. Dev. IIP3 [dB]	SPICE	1.1	1.7	1.1	0.9
	Synthesis	1.1	1.9	1.2	0.9
Synthesis vs. SPICE Correlation		0.98	0.97	0.80	0.93

Validation of filter bootstrap estimates of IM_3 used the correlation of IIP_3 between SPICE simulated and synthesized at a center-frequency of 140kHz. Fig.4.12 (a) and (b) plots the bootstrap synthesized IIP_3 estimate versus SPICE for 250 samples each of the 4th-order Chebyshev filter and the 8th-order Butterworth. The plots emphasize the IIP_3 correlation (i.e., no significant outliers) and the four marginal histograms long-tailed structures. Table 4.3 summarizes statistics and correlation of the two different method’s for computing IIP_3 for each filter design. The agreement of AC transfer functions and IIP_3 marginal distribution shapes and Pearson’s correlations ≥ 0.8 allows the computationally less demanding bootstrap estimates to replace exhaustive SPICE simulation as a means to estimate statistical FoMs for different filter designs.

The combination of test limits and use limits divides the manufactured filters into four bins (see Table 4.4 and Table 4.5). Test frequency selection and guard-band test limit settings are a trade-off between test-time and test-cost. Manufacturing yield loss (YL) from true fails and overkill (OL) and user observed defect level (DL) are FoM statistics that quantify the trade-off between test and use. For the test and use settings, the four classification bin counts are used to compute YL (and the yield loss

overkill component) and DL FoMs. (Eq. 1).

Bootstrap estimation and final design selection considered first AC response assessment of manufacturing yield and quality. Fig.4.13 (a) and (b) plots transfer function magnitudes for a bootstrap generated synthesized *4th*-order Chebyshev filter and a *8th*-order Butterworth filter, respectively. The traces are coded by line style and color for PASS and FAIL at test (manufacturing yield FoM) and at use (quality FoM). For reference, the blue overlay trace is each filter's ideal (i.e., ideal OTAs and no manufacturing variation) transfer function. All green and red traces are filter transfer functions in the presence of non-ideal OTAs and process variation and transistor mismatch. Green traces denote filters which PASS all AC use specifications. Red traces denote test escape filters which FAIL one or more AC use specifications. Solid line traces denote filters which PASS test (PP or PF) and dashed line traces denote filters which FAIL one or more AC test specifications (FP or FF).

For the AC magnitude test, bootstrap estimated the *4th*-order Chebyshev filter yield loss is 1.5X greater than the *8th*-order Butterworth filter yield loss, Table 4.4. For the sample size, the filter defect levels are statistically insignificant. The computed bootstrap estimates of filter yield loss etc. are the result of non-ideal OTAs and manufacturing variation and not from random point defects. For example, the *4th*-order Chebyshev yield loss is largely explained by the one high-Q biquad stage increased sensitivity to device variation.

Estimation of IIP_3 yield loss and defect level FoMs for the bootstrapped *4th*-order Chebyshev filter and the *8th*-order Butterworth show design selection may result in different trade-offs than the AC results. In Fig.4.14, filter IIP_3 estimates are plotted with test and use bin limits as lines. Each sample filter is a single dot. Fig.4.14 adopts the AC binning color scheme. The green shaded area surrounds the filters that PASS IIP_3 in use. The red area surrounds the filters that FAIL IIP_3

Figure 4.13: AC transfer functions of 1200 samples of two filter design alternatives (a) *4th*-order Chebyshev (b) *8th*-order Butterworth. Each filter is DC trimmed to 0dB at test. Black vertical bars locate test frequency and magnitude specifications. Passband test guard-band is 0.5dB @ 190kHz. Nominal trace denotes ideal AC response. Legend denotes results of test and use, see Table 4.1. For example, PF denotes pass at test and fail in use. Use specifications are the dashed lines.

in test. The x-axis is a filter test IIP_3 estimate at guard-banded center-frequency =140kHz. The test limit is plotted as a vertical (red) dashed line. The y-axis is the use IIP_3 estimate at center-frequency =100kHz. The use limit is plotted as a horizontal solid (red) line. The limits divide sample filter responses into four distinct groups shown in the plots four corners.

The IIP_3 comparison for the test and use limits is summarized in Table 4.5. Yield loss for *4th*-order Chebyshev is 4X the Butterworth filter. Based on IIP_3 yield loss the preferred design alternative is the *8th*-order Butterworth. While the overall IIP_3 yield loss is consistent with the AC, the source of the loss is different. For both design alternatives the yield loss is the result of test overkill, that is the component of yield loss from the miscorrelation between test and use. Of particular concern is when test overkill is a large multiple of filters FAIL the IIP_3 limit at the guard-banded

Table 4.4: Filter Comparison from AC Manufacturing Results

		Filter Design Alternative	
	Bin	4th-Cheb	8th-Butter
Test/Use	F/F	23	13
	F/P	4	2
	P/F	4	3
	P/P	219	232
Yield Loss (YL)		10.8%	6.0%
Overkill (OL)		1.6%	1.0%
Defect Level (DL)		1.8%	1.3%
AC test frequencies 140kHz and 300kHz			

test center-frequency but meet IIP_3 at the use center-frequency. For example, in Table 4.5 the 4th-order Chebyshev has a 'kill ratio' of FP/FF=34/6 meaning 5 or 6 filters are removed as FAILs (and not shipped to the customer) for each true FAIL at use. Test escapes are assessed by customers as a defect level. The bootstrap estimates suggest in the customer view (that is for an SOC the performance in use) the 8th-order Butterworth alternative has nearly 4X as many faulty units delivered downstream. The trade-off is balancing the choice of the guard-band center-frequency, different test limits or some combination to reduce the kill-ratio by the risk of increasing the number of test escapes (PF =2 and =9 for the two filters, respectively).

Table 4.5: IIP_3 Manufacturing Results

		Filter Design Alternative	
	Bin	4th-Cheb	8th-Butter
Test/Use	F/F	6	3
	F/P	34	8
	P/F	2	9
	P/P	208	230
Yield Loss (YL)		16%	4.4%
Overkill (OK)		13.6%	3.2%
Defect Level (DL)		1.0%	3.8%
Test center frequency = 140kHz			
Use center frequency = 100kHz

Figure 4.14: Bootstrap IIP_3 synthesis results for 250 (a) 4th-order Chebyshev filters and (b) 8th-order Butterworth filters. Use IIP_3 set the center-frequency test-tone =100kHz and at test guard-banded the center-frequency =140kHz. Frequency offset $=\pm 2.5\%$ of the center-frequency. Vertical and horizontal lines mark the IIP_3 test and use limits. Each dot '.' in FP and PF quadrants is a test escape or test overkill, respectively. Counts in FP and PF quadrants assess manufacturing risks from the miscorrelation of test and use.

The filter design using the biquad building blocks offers other benefits for design comparison. In the design description Table 4.2 the 4th-order Chebyshev filter input biquad section was of type 1:1:1:1 and output section of type 4:1:4:1. The wider transistors in the latter biquad means the biquad requires about 4X more power. The 8th-order Butterworth filter used three biquad sections of type 1:1:1:1 and one of type 2:1:2:1. By using more of the lower power 1:1:1:1 biquads, the 8th-order Butterworth filter power consumption is SPICE estimated to be 137% of the 4th-order Chebyshev and by comparing their biquad standard-cell counts 122%. Using capacitor area as the dominant layout component, the 8th-order Butterworth filter area is $\sim 14\%$ smaller than the 4th-order Chebyshev filter.

The computation time for SPICE simulation and bootstrapped estimate of the

Table 4.6: Run time comparison SPICE and Synthesis (Building Block Method)

250 Sample Statistics	Filter Design Alternative		Total
SPICE Run Time	4th-Cheb	8th-Butter	
AC	26s	77s	113s
IIP_3	3.2 hrs	6.3 hrs	9.5 hrs
Synthesis Run Time	4th-Cheb	8th-Butter	Total
AC	2.6s	5.2s	7.8s
IIP_3	166s	332s	498s

FoMs is summarized in Table 4.6. To obtain the manufacturing statistics requires sufficient sample size and at the modest sizes used in this study, the difference in computational time is substantial. Using SPICE to compute 250 samples is 9.5hrs. Using the building block method takes roughly 10mins.

Summarizing the entire FoM analysis for the two design alternatives leads one to select the *8th*-order Butterworth filter as the final design. This conclusion should not be confused with a recommendation that maximally-flat filters are in general the preferred design. Rather this conclusion demonstrates the potential of a coordinated synthesized, biquad building block approach which includes sufficient characterization of the biquads to compare 5, 10 or 100 different filter design alternatives that meet the original use requirements and select the design which has the best chance of simultaneously meeting manufacturing, test and use requirements.

4.8 Conclusions

This chapter demonstrated a computationally efficient methodology to give analog filter designers performance of different filter designs at test and at use condition and provide statistically significant manufacturing Figures-of-Merit. The foundation of the efficiency is the reuse of the linear and non-linear response data from pre-characterized biquad building block library to manufacturing variation. Bootstrapped

filter generation and reuse of building block SPICE simulations provides important manufacturing FoM statistics early in the design phase without SPICE simulation of each design alternative. Trends in multiple manufacturing FoMs provides as complete a picture as possible to all the stakeholders; system engineers, circuit engineers, manufacturing engineers and test engineers and reduces the risk of selecting an alternative which is more sensitive to advanced technology manufacturing variation. Yield loss and defect level FoMs allow design and test engineers to adjust limits to improve manufacturing test and product use correlation and gain the feel for a design alternative's sensitivity to manufacturing variation. Analyzing the trade-offs between different filter designs, specification settings, manufacturing yield and quality are essential when small analog filters are integrated into (a more expensive to produce) SoC or ASIC. The estimation of every possible filter design's yield and filter defect level requires accounting for non-idealities and for process variation and mismatch variation effects. Replacing intensive Monte Carlo SPICE simulations of each design alternative with bootstrap synthesis using the pre-characterized biquads from the library efficiently and quickly estimates the power, area and manufacturing and quality FoMs of any filter design alternative.

The biquad building block design is set by a small number of independent design degrees-of-freedom which can be varied systematically for pre-characterization and Monte-Carlo samples of each set stored in a library. The low-cost of data storage allows linear and non-linear response of each Monte-Carlo sample of the building block to be stored and simple models developed to estimate both linear and non-linear responses of a random sample of each filter design alternative and preserve the inherent correlation between different responses without the use of ad-hoc and difficult to parametrize statistical models.

This chapter describes design selection for manufacturability of analog circuits,

specifically continuous-time filters. The methodology adapts to analog design the common practice of a standard-cell library in digital circuit design. The results demonstrate linear and non-linear response predictions for different design alternatives can select a preferred final design early in the design flow and reduce the risk of expensive design iterations. Future work is to establish the building block library and model to synthesize manufacturing FoM for other type of analog circuits. There are three rules suggested when exploring potential building blocks for any analog circuit under test (CUT): (i) repetitive common circuits found in CUT (ii) circuits that have similar architecture as the CUT (iii) combination of sub-circuits which can represent the CUT performance. The selection of building blocks is designed based on characterizing the building block's responses such as gain, bandwidth, linearity, noise and possibly others at different environmental conditions. The building block is identified that can best describe the CUT behavior. Potential analog circuits that could benefit from the methodology proposed are cascaded transimpedance amplifier, ring oscillators, ADCs and DACs, etc.

Chapter 5

Copula-based Modeling of Analog Parametric Tests Separating Variation from Defects

5.1 Introduction

This chapter applies the copula to describe the dependency of analog circuit design performance to process variation and temperature changes. The copula is shown to provide unique insight into the dependence. The copula reveals that process variation and not defects can produce outlier behavior which would normally be interpreted as the result of defects. Integrated circuits passing at test and failing in use are a concern in both analog and digital domains. Consistent test escapes (i.e., chips failing in use the same way) require remedial action by the consumer and the producer. The “No Trouble Found” integrated circuit failing in systems and passing retest at ATE can account for 30-40% of returns [64] [65]. The resolution of NTF often centers on mis-correlation between test and use. In this chapter, Monte Carlo circuit simulations are used to characterize performance at various temperatures and supply voltages. The process variation induced changes in circuit performance correlation are studied by combining the Monte Carlo with a 3-Factorial Design of Experiments (DoE) [66,67]. The correlation of circuit performance is modeled using copulas. Copulas in test have been reported before [42] and [68]. In [4] the copula is used to describe the correlation between built-in-self test results and circuit performance to estimate the figures-of-merit such as test escapes and yield loss. In this chapter the copula re-

veals the temperature performance dependency structure is not well described by a bivariate (two-variable) Gaussian distribution. The DoE shows the manufactured unit “outliers” are not the result of process defects and instead are a significant fraction of normal process variation. To establish circuit functionality, analog circuits are tested under different environmental conditions such as temperature and supply voltage. Test time increases with each new environmental condition. At some point in the volume ramp testing at all environmental conditions is simply not feasible to meet test throughput requirements or remain within cost constraints. One approach to reduce test time is to drop the tests that are guaranteed to pass. To accomplish this, one can find the correlation between two parametric performance outputs at their respective conditions. When performance between two conditions is correlated, results from the first condition test can replace the second condition test by predicting the response through a correlation model [69]. In section 2, is a review of the copula dependence. A common analog building block (OTA) demonstrates the copula modeling of temperature and voltage correlation in section 3. In section 4, the recommendations and observations from this study are discussed. Section 5 concludes this chapter.

5.2 Methodology

5.2.1 Concept of Dependence and Copulas

The association between two random variables x and y whose cumulative distribution are $F(x)$ and $G(y)$ is completely described through the cumulative distribution function, $H(x, y)$,

$$H(x, y) = P[X \leq x, Y \leq y] \quad (5.1)$$

If random variables x, y are independent the cumulative distribution $H(x, y)$, satisfies

$$H(x, y) = F(x)G(y) \quad (5.2)$$

or when using the joint density PDF independence is written

$$h(x, y) = f(x)g(y) \quad (5.3)$$

where $G(x)$, $f(x)$ and $G(y)$, $g(y)$ are the marginal cumulative and probability density functions of x and y , respectively. The independence condition suggests a general expression for dependencies in $h(x, y)$

$$h(x, y) = f(x)g(y)(h(x, y))/(f(x)g(y)) \quad (5.4)$$

The *dependence functions* $(h(x, y))/(f(x)g(y))$ of the joint density of x and y is a simple way to display the dependence of the random variable pair. For example, from Eq. 5.4 if x and y are independent, then the dependence function $(h(x, y))/(f(x)g(y))$, is 1.

The random variables u and v are the cumulative density of $F(x)$ and $G(y)$, respectively.

$$u = F(x) \text{ where } u = \text{uniform}[0, 1]; v = G(y) \text{ where } v = \text{uniform}[0, 1] \quad (5.5)$$

In other words, the cumulative density assigns the (x, y) pair in joint distribution parametric space to a normalized ranked pair (u, v) in an always uniform distributed in rank space.

According to Sklar's Theorem [70], if $H(x, y)$ is the cumulative distribution and $F(x)$ and $G(y)$ are the marginal, then there exists a copula C such that

$$C(u, v) = H(F^{-1}(x), G^{-1}(y)) \quad (5.6)$$

If $F(x)$ and $G(y)$ are continuous, then the copula is unique. When $H(x,y)$, $F(x)$, $G(y)$ and $C(u,v)$ are differentiable

$$c(u, v) = \partial^2 C(u, v) / \partial u \partial v = (\partial^2 H(x, y)) / (\partial F(x) \partial G(y)) \quad (5.7)$$

where $c(u,v)$ is called the copula density, similar to the PDF of the cumulative distribution. The relationship between $h(x,y)$ and the copula density $c(u,v)$ can be shown to be

$$h(x, y) = f(x)g(y)c(u, v) \quad (5.8)$$

Substituting dependence function in Eq. 5.4 with Eq. 5.8

$$(h(x, y)) / (f(x)g(y)) = c(u, v) \quad (5.9)$$

shows the abstract idea of a general dependence function in (4) to be the copula density.

The importance of the decomposition is best seen in 5.8. Typically, the model for correlation between two random variables is obtained through the joint density $h(x,y)$. When correlation is approached through $h(x,y)$, the random variables' marginal distributions and correlation are intertwined and difficult to separate. Equation (8) shows when $h(x,y)$ is expanded using the dependence structure of a copula density, the marginal distributions are explicit and can be separately studied. From a modeling perspective, the separation of marginal distributions from a joint density's dependence structure provides additional model choices for the joint density.

Copula Type	Gaussian Copula	Clayton Copula
$C(u, v)$	$H(F^{-1}(u), G^{-1}(v); \theta)$	$(u^{-\theta} + v^{-\theta} - 1)^{-1/\theta}$
$c(u, v)$	$(1 - \theta^2)^{-1/2}$ $\times \exp\left\{\frac{-1}{2}(1 - \theta^2)^{-1}(x^2 + y^2 - 2\theta xy)\right\}$ $\times \exp\left\{\frac{1}{2}(x^2 + y^2)\right\}$ <i>where $x = F^{-1}(u), y = G^{-1}(v)$</i>	$(1 + \theta) \times (uv)^{-\theta-1}$ $\times (u^{-\theta} + v^{-\theta} - 1)^{-2-1/\theta}$
$\theta - \text{domain}$	$-1 \leq \theta < 1$	$\theta \in (0, \infty)$
Kendall's τ	$\frac{2}{\pi} \arcsin(\theta)$	$\frac{\theta}{\theta + 2}$
Spearman's ρ	$\frac{6}{\pi} \arcsin(\frac{\theta}{2})$	*

Table 5.1: Copula Properties

In this chapter, other alternatives are explored and evaluated for dependence structures when $c(u, v) \neq 1$. A wide range of copulas are reported in the literature such as the Gaussian, the Clayton, Frank, Gumbel and so on [71]. To demonstrate the opportunities for modeling a joint density, Table 4.1 summarizes the mathematical properties of Gaussian copula and Clayton copula. Similar to the joint density's cumulative and PDF, $H(x, y)$ and $h(x, y)$, a copula can be mathematically described by its $C(u, v)$ or $c(u, v)$. A copula typically has one or two dependence parameter(s). The copulas in Table 4.1 each have a single dependence parameter denoted by θ . Again drawing on relationship between correlation coefficient and the joint density, copulas have a connection to the rank based population correlation statistics of Kendall's tau (τ) and of Spearman's rho (ρ). Tau and rho are measures correlation and can be used interchangeably [72]. Table 4.1 summarizes the relationship between copula's dependence parameter and either Kendall's 'tau' or Spearman's 'rho'. Either data sample statistic is a convenient way to determine the copula's parameterization.

The advantage of separating marginals from dependency is shown in Fig. 5.1. Fig. 5.1 displays scatter-plots for synthesized samples of the copula densities for the

Figure 5.1: Synthesized sample from (a) Gaussian copula with $\rho=0.9$ and (b) Clayton copula with $\rho=9.74$

Gaussian copula and Clayton copula, respectively. For each the x-axis is a uniform random variable and similarly for the y-axis. Recall the cumulative distribution (5) for any random variable, u and v is a uniform random variable.

The large number of possible copula functions means selection can be difficult for any particular situation. To select a particular copula function is somewhat of an art and relies on matching observed properties and computing goodness-of-fit statistics. One obvious way to match a copula function to data is symmetry. Symmetry above and below the 45° line (i.e., equal points above and below the 45 line) is a property the Gaussian and the Clayton copulas share. Symmetry is not general and the symmetry of the Gaussian copula will be seen to provide a poor fit to the observed OTA temperature response. Another copula density property which has proved to be helpful is the Kendall's tau correlation in the tails [3], that is, the correlation as the data approaches the limits of $(0,0)$ or $(1,1)$. Left and right tail dependences are denoted α_0 and α_1 , respectively. Mathematically, the Gaussian copula's tail dependence is identical, $\alpha_0=\alpha_1=0$ and is suggested by the scatter-plot in Fig. 5.1(a). In contrast, the Clayton copula in Fig. 5.1(b) exhibits very different tail dependence.

Figure 5.2: Bivariate Normal Joint Density $h(x,y)$ and histogram of x , $f(x)$, and y , $g(y)$.

The dependence is strong in the left tail $(0, 0)$ but weak, $d_0=1$ and in the right tail $(1, 1) = 1=0$ [8].

A bivariate normal joint probability density distribution can be constructed by first selecting parameters for each Gaussian marginal distribution $f(x)$ and $g(y)$, the Gaussian dependence structure in Table 4.1 (Appendix) and synthesizing points based on simple algorithms. Fig. 5.2 shows an example bivariate Normal joint probability density distribution with marginal means $m=2$ and standard deviations of $sv=0.5$ and $f(x)$, $g(y)$ correlation coefficient, $= 0.9$.

Fig. 5.3 demonstrates the flexibility of the dependence structures (copulas), marginal distributions in creating joint densities. Combinations of different marginal distributions and copulas generate very different looking joint densities, $h(x,y)$. Fig. 5.3 displays the results for a Gaussian marginal distributions for $f(x)$ and $g(y)$ with

Figure 5.3: Joint density plots from Gaussian marginal with dependence structure of (a) Gaussian copula $h_1(x,y)$ and (b) Clayton copula h_2

Figure 5.4: Joint density plots from Bimodal and Weibull marginal with dependence structure of (a) Gaussian copula $h_3(x,y)$ and (b) Clayton copula $h_4(x,y)$.

different dependence structures $c(u,v)$. Fig. 5.3 (a) combines Gaussian marginals with a Gaussian copula and Fig. 5.3(b) combines the same Gaussian marginals with the Clayton copula to generate sample joint densities $h_1(x,y)$ and $h_2(x,y)$, respectively. In Fig. 5.4, a bimodal marginal distribution substitutes the Gaussian marginal for $f(x)$ and a Weibull marginal distribution substitutes the Gaussian for $g(y)$. Fig. 5.4 (a) retains the Gaussian copula and Fig. 5.4 (b) retains the Clayton copula to generate joint density plots $h_3(x,y)$ and $h_4(x,y)$, respectively.

Fig. 5.3(a) and Fig. 5.4(a) shows the symmetric shape of dependence structure (copula) and changing the marginal distribution, $f(x)$ and $g(y)$, can generate signif-

icantly different shapes of joint density. Alternatively, combining different copulas with fixed marginal distributions present very different joint densities despite the fact the underlying dependency structure is simple and fixed. In practice, when densities similar to $h3(x,y)$ and $h4(x,y)$ are observed, choices such as sums of Gaussians or other complex parameterization methods of the joint density are thought to be the only modeling alternative. Fig. 5.3 and Fig. 5.4 shows the use of copulas to model dependency separates marginal distributions from dependence structures and provides a complete and scale-free description of dependence. The copula based approach provides freedom to model the joint density with forms other than Gaussian. The appropriate copula for a particular application is the one which best captures dependence features of the data.

5.2.2 Using copulas to model the dependence structure and develop a statistical model for correlation in circuit performance over temperatures

In this section, the copula based approach is used to model the dependence structure of circuit performance at two temperatures. To demonstrate this technique, a common analog building block, the OTA is used to demonstrate a statistical model of temperature correlation for OTA AC gain performance. The OTA schematic, designed in 45nm technology, is shown in Fig. 5.5. The process technology device file is one of several the ICDT Laboratory has obtained. The OTA design has a nominal gain of 250V/V and bandwidth of 500kHz at the operating conditions ($V_{DD}=1.1V$ and 27 C).

Figure 5.5: OTA schematic

5.2.3 Methodology: Identifying Dependency

Process variation is introduced using the Monte Carlo models for the NMOS and PMOS. The process variation model file uses a version of the compact modeling technique to describe the correlated variation of the device model parameters [73]. For the 45nm technology Monte Carlo, four independent random variables, rand1-4, are used to describe the process variation and the correlation between twelve device model parameters for NMOS and eleven device model parameters for PMOS. For this study, Monte Carlo techniques were used to generate 1,000 design instances by independently sampling the recommended distribution for rand1-4, Gaussian with fixed mean $m=0$ and standard deviation $sv=1/3$. Key circuit performance parameters (such as gain, bandwidth, etc.) were computed for the 1,000 OTAs at five temperatures (0°C, 27°C, 65°C, 80°C and 100°C) and three VDD supply voltages (1.0V, 1.1V and 1.2V).

The 1,000 instances simulation results provide the data set for studying the OTAs dependency structure. Conventional statistics characterize OTA performance such as the marginal distributions mean $(\bar{x}, \bar{y})$, the standard deviation (s_x, s_y) , the goodness-of-fit to model marginal distributions. Correlation coefficient Kendall's tau) for data is evaluated to complete the list of conventional statistics. In addition, an empirical (observed) copula density for the response to temperature variation is computed based

on the 1,000 instances. The empirical copula is generated by separately sorting the values in each joint distribution pair of (x,y) and assigning to each ranked value the corresponding normalized value. The normalized ranked values are combined into a (u,v) pair of the copula density. Correlation coefficient Kendall's tau for data is evaluated to complete the connection to the copula.

An analytical derivation of dependency structure is difficult, if not impossible, for even a simple circuit such as an OTA. A more productive and practical approach is to estimate the dependence empirically by sampling. The Monte Carlo simulations could be replaced by measured response to accomplish the same dependency modeling goal. In addition to the Monte Carlo simulation sample, additional simulations were completed for a three-factorial (3!) design-of-experiment (DoE). Modeling process variation with compact model method makes the DoE convenient because to capture the correlations between MOS parameters, only the independent random variables, rand1-4, need to be set to $\pm 3\text{sv}$ and 0. The four random variables require 81 configurations ranging from (1,1,1,1) to (0,0,0,0) to (-1,-1,-1, -1).

This modeling methodology is summarized in the flowchart shown in Fig. 5.6.

The first and simplest step is establishing the statistical model for the marginal distributions, $f(x)$ and $g(y)$. The marginal distributions can be selected without regard to the dependency. Goodness-of-fit tests can be used to parameterize the marginals and judge the final fitting.

A single comparison demonstrates the concept of fitting dependency to a copula. That is, the Monte Carlo empirical copula is compared to the Gaussian copula. Particular attention is played to two elements of the dependency. First, density contours of the empirical copula are aligned with density contours of a Gaussian copula. Recall the Gaussian copula density is parameterized by a single parameter either Kendall's tau or Spearman's rho. Second, the symmetry of the empirical copula and the Gaus-

Figure 5.6: Flow chart for the proposed technique

sian model are compared. For example, if points contained within a contour of the empirical and Gaussian copula are consistent, then the copula model of dependency is consistent with the observed, empirical copula.

5.3 Result

5.3.1 Non-Gaussian Dependence

Fig. 5.7 shows histograms of the marginal distribution of 1,000 Monte Carlo instances for one the Monte Carlo process variation parameters rand1, one of the device model parameters NMOS dvth0 (models the threshold voltage variation at zero substrate bias) and two OTA performance responses, the OTA AC gain at 27C, 1.0VDD-- and OTA AC gain at 80C, 1.0VDD. Table 4.2 summarize the distribution statistics where OTA AC gain at 27C is $f(x)$ and OTA AC gain at 80C is $g(y)$. By construction, each of the four Monte Carlo parameters (rand1-4) will pass Shapiro-Wilks W-test-

Figure 5.7: Marginal distribution of (a) random (b) $dvth0$ (c) OTA gain at 27C, $f(x)$ (d) OTA gain at 100C, $g(y)$

parameter and OTA gain

Table 5.2: Statistics summary for Monte Carlo parameter, Device Model

	Mean	Standard deviation	Shapiro-Wilk W Normality Test
(a) Monte Carlo parameter <i>rand1</i>	0.01	0.33	W=0.999 Prob <W=0.7794
(b) Device parameter <i>dvth0</i>	-6.5e-12	1.87e-10	W=0.999 Prob <W=0.469
(c) OTA gain at 27C, $f(x)$	212.60	7.21	W=0.999 Prob <W=0.8986
(d) OTA gain at 80C, $g(y)$	170.39	6.51	W=0.998 Prob <W=0.3011

for-normality. Similarly, the 23 device model parameters are computed as linear combinations of $rand1-4$ and by construction will pass the W test-for-normality. Finally, the marginal distribution of the OTA AC gain 27C and 80C each pass the W test-for-normality. Although AC analysis linearizes the transistors at the DC operating point, a response such as AC gain is not guaranteed to be normal because the MOS circuitry is by its nature a non-linear system.

Fig. 5.8 shows the marginal joint density $h(x,y)$ of OTA gain at 27C/1.0VDD and 80C/1.0VDD as well as marginal distribution $f(x)$ and $g(y)$ in parametric space. The empirical copula $c(u,v)$ for dependence structure is constructed in rank space and shown in Fig. 5.9. The uniform distributions of the u and v marginals are plotted to emphasize the uniformity of the copula margins.

Figure 5.8: Joint density $h(x,y)$ for OTA gain at 1VDD, 27C/80C

When marginal distributions pass such normality tests the natural assumption is to model their joint density with the bivariate Gaussian. Hints of the trouble to come can be seen in Fig. 5.8. The upper edge of the scatter-plot there is much more distinct (sharp line) compared to the lower edge. The similarity (a requirement by symmetry) is seen in the scatter along the upper and lower edges of the Gaussian scatter-plots in Fig. 5.2. The empirical copula scatter-plot in Fig. 5.9 compared to Fig. 5.1 suggests the difference more clearly with many points scattered below the diagonal and almost none in the upper left hand corner.

Fig. 5.10 is the same empirical copula scatter-plot with a 98% density contour added. The added contour assumes the dependency is a Gaussian copula with a correlation coefficient equal to the observed value. Notice how the 2% outside of 98% density contour are miscorrelating. The instances reflect an outlier behavior and all lie below the density contour and none lie above the contour. The 2% of points

Figure 5.9: Empirical copula $c(u,v)$ for OTA gain at 1.0VDD, 27C/80C

are not evenly distributed above and below the contour as highlighted by the solid triangles in Fig. 5.10(b).

The empirical copula is not well described by a Gaussian copula because there is a larger than expected number of instances below the diagonal and outside the 98% contour.

The six instances below the 98% contour are identified as miscorrelating outliers because their rank of OTA AC gain at 80C miscorrelates (i.e. is too small) compared to the expected OTA AC gain ranks at 27C. Although 2% is a small fraction, the result of this miscorrelation is obtained from modeled process variation and not the result of explicit introduction and modeling of a defect. The miscorrelating outliers are caused by intrinsic process variation and are important to spot to avoid being falsely treated as defect.

The six miscorrelating outliers identified from copula $c(u,v)$ are less easily observed

Figure 5.10: (a) Symmetric analysis for empirical copula. Instances outside 98% density contour are highlighted in rectangular dot. (b) Miscorrelating outliers are highlighted in triangle dot.

in joint density $h(x,y)$ as shown in Fig. 5.11. Furthermore, the miscorrelating outliers cannot be detected from marginal distributions $f(x)$ or $g(y)$ since miscorrelating outliers fall in the main population of marginal distribution.

A non-Gaussian dependence structure means traditional bivariate Gaussian modeling of the density could lead to potentially misleading interpretations of fails, poor choices in test set point limit setting or worse expensive searches for defects in manufactured devices which would eventually have to be added to the “No Trouble Found” category.

5.3.2 Temperature and Non-Gaussian Dependence

Temperature dependent model parameters are evaluated to isolate a root cause for the six temperature miscorrelating outliers. A series of simulations turned off the temperature dependent device model parameters for transistor threshold voltage (KT), transistor mobility (UTE) and transistor saturation voltage (AT). Turn off means each parameter was set to zero and each of the 1,000 instances were resimulated at 27C and 1.0VDD. The model parameter PMOS KT was revealed as the main cause for the miscorrelating outliers. That is, for $KT_p=0$ the original miscorrelating outliers

Figure 5.11: Joint density $h(x,y)$ and marginal distribution $f(x)$ and $g(y)$ with miscorrelating outliers highlighted in orange

merged into main scatter of the AC gain copula density. Fig. 5.12 shows temperature dependence miscorrelation is eliminated when $KT=0$ while $UTE=0$ and $AT=0$ increase the overall correlation between two temperatures but the outliers remain miscorrelated. The effect of $KT=0$ and $KT=0$ on the ranks of the six miscorrelating outliers are summarized in Table 5.3.

5.3.3 Non-Gaussian Dependence and Test Metrics

A $3!$ DoE was used to estimate the population fraction of the miscorrelating outliers and as result estimate the DPM fraction in production. The DoE used the $[-3\sigma, 0, 3\sigma]$ level for each of the four Monte Carlo parameters, rand1-4 modeling process variation. Eighty-one combinations were simulated at 27C/1.0VDD and 80C/1.0VDD. The eighty-one $3!$ DoE results and the values from the original 1,000 Monte Carlo

Figure 5.12: Empirical copula for OTA gain at 27C/100C and 1.0VDD with PMOS MP (a) original MC results (b)KT=0 (c) AT=0 (d)UTE=0

Table 5.3: Miscorrelating outliers movements from original model file to model file with KT=0 for PMOS

Instance #	Normalized rank (u,v) (original model file)	Normalized rank (u,v) (KT=0 for pmos)
29	(0.757,0.205)	(0.766,0.761)
114	(0.691,0.219)	(0.704,0.760)
121	(0.600,0.110)	(0.615,0.673)
525	(0.847,0.410)	(0.856,0.851)
631	(0.290,0.043)	(0.312,0.434)
807	(0.535,0.081)	(0.552,0.618)

instances are combined and plotted in Fig. 5.13 and Fig. 5.14. In Fig. 5.14, note there are a few DoE combinations above the main distribution however, the DoE combinations in the upper left corner do not display the same miscorrelation as the combinations with OTA AC gain at 80C between 0 and 0.1.

The DoE effectively finds the combinations of rand1-4 that cause the miscorrelating outliers. The compact modeling eases these calculations by reducing combinatorial count (i.e., 81) to a practical level compared to the impossible number of required combinations if each of the 12 device parameters were independently varied (3121500K). The results shown in Fig. 5.14 allow a simple estimate of the expected number of miscorrelating outliers in production. The ranks of the DoE within the Monte Carlo were determined by combining the DoE and Monte Carlo instances. The DoE settings of the rand1-4 are listed in Table 4.4 for the six DoE extreme miscorrelating outliers. A setting at $\pm 3\text{sv}$ represents a small fraction of parts ($p=1.35\text{E}(-3)$) in the tail. The DoE setting of 0sv represents almost the entire distribution ($p\approx 1$). Examining the rand1-4 settings for combinations E2 and E5 estimates the fraction of die to be ≈ 4 DPPM. Combinations E1, E3, E4, E6 combined are a vanishingly small fraction, E1 DPPM. While the numbers are small for this case the intent of the study is to demonstrate the need to understand the miscorrelation to avoid more significant challenges to yield, NTF and the like in products.

For a given set of test limits at the two conditions the copula can be used to compare test metrics such as over-kill and test-escape tradeoffs. For example, suppose 27C is the normal use condition and 80C is the preferred test condition. Assuming the use limit and test limit specifications as shown in Fig. 5.15, the six miscorrelating outliers, E1-E6, are fails at the test conditions and are passes at the use conditions. The miscorrelation at these test limits results in an over-kill and lower yield.

Figure 5.13: Three Factorial DoE- joint density

Table 5.4: Monte Carlo Parameters RAND 1-4 for E1-6

	rand1	rand2	rand3	rand4
E1	3σ	0	-3σ	-3σ
E2	3σ	0	-3σ	0
E3	3σ	0	-3σ	3σ
E4	0	-3σ	-3σ	-3σ
E5	0	-3σ	-3σ	0
E6	0	-3σ	-3σ	3σ

Figure 5.14: Three Factorial DoE- copula

Figure 5.15: Three Factorial DoE - copula with limits

5.3.4 Non-Gaussian Dependence Parameterization

Using results from Monte Carlo or measured devices the copula based model can anticipate miscorrelation effects as at different environmental condition settings. The Monte Carlo simulations at (0C, 27C, 65C, 80C and 100C) reveal temperature dependent correlation as well as a supply voltage dependent correlation after simulating the instances at (1.0VDD, 1.1VDD and 1.2VDD).

Fig. 5.16 is a scatter-plot of the 1,000 Monte Carlo at 27C and 100C and a second scatter plot comparing 27C and 0C for 1.0VDD. Note the miscorrelating outliers always appear as lower ranked instances for the higher of the two temperatures. The correlation increases for the smaller temperature differences. In Fig. 5.17 two voltages are plotted, 1.0VDD and 1.1VDD. Fig. 5.17 shows dependence structure at 27C and 80C is also dependent on voltage. Fig. 5.16 and Fig. 5.17 show the dependence is a combined function of temperature and supply voltage.

Recall the Gaussian copula was used to identify the miscorrelation outliers. The combined temperature and voltage dependence and the effects of miscorrelation can be modeled easily because the Gaussian copula is fully parameterized by Spearman's rho (r). r at different environmental conditions are fitted with (10) and shown in Fig. 5.18. A simple fit suggests the voltage and temperature dependence is a simple product of voltage and temperature

$$r = A(V - V_{ref})(T - T_{ref}) + 1 \quad (5.10)$$

where the fitting parameters for the Monte Carlo data are $A = -2 \times 10^{-3}$, $T_{ref} = 27$ and $V_{ref} = 1V$. A full model of the voltage and temperature dependence requires research however, the benefits of constructing separate models for the marginal distributions and for the dependency structure alone have been demonstrated in the literature [3].

Figure 5.16: Dependence structure for OTA gain at 1VDD and (a) 100C/27C (b) 0C/27C

Figure 5.17: Dependence structure for OTA between 27C and 80C at (a) 1.0VDD (b) 1.1VDD

Figure 5.18: Spearman's rho at various environment conditions

5.4 Recommendations

The dependence of between two circuit test responses has been outlined. The approach presented separates the marginal distribution and the joint density. The separation provides three advantages.

First, recognizing the copula as the entire description of dependency means a much wider array of dependencies can be deployed when modeling. This wider array opens the possibility of simplifying the description of joint densities by avoiding complex, parametric settings commonly found in techniques such as Gaussian kernel expansions.

Second, the dependency can be obtained through the use of Monte Carlo or production unit sampling and does not require any significant analytic predictions of the dependency.

Third, once the dependency is identified a wide array of test related issues can be addressed such as test set point limit setting, and estimating fractions of specific response categories. The separation of dependency from marginal response provides an opportunity to parameterize the margins with simple uni-variate goodness-of-fit tests and similarly for selecting the dependency copula for a specific system [74].

5.5 Conclusion

A common problem in test is the miscorrelation between test and use. Often miscorrelation is assumed to be the result of defects. Using a common analog building block the OTA, this study demonstrates miscorrelation is not always defect behavior and instead can be the result of normal process variation [75].

Using a CMOS OTA design as a case study, the dependence structure in analog parametric test over environmental conditions is characterized. Estimation of test matrices such as overkill due to falsely treating intrinsic process variation as defects

is demonstrated. This study conveys the benefit of using copula for modeling. Understanding the dependence structure can improve test quality by avoiding over-kill.

The use of a statistical model is proposed to describe the correlation of analog parametric test and others continuous valued test responses. The copula as a formulation for joint density reveals a dependence structure other than Gaussian for a simple analog building block. The miscorrelation of outliers was shown to be observed more efficiently by a copula than by examining the marginal joint density. Such a correlation model allows for modeling miscorrelation effects without having to include complicated methods of defect insertion.

Chapter 6

Conclusion

This thesis built a model based on the parametric statistics of building blocks (GmC biquad) to predict the yield for different filter design alternatives. The proposed methodology selects a base circuit from the numerous available cascaded-biquad GmC filter configurations. The tool considers that components will drift due to environmental effects, such as temperature and voltage variations, as well as the process variation. The methodology also models the active devices, such as those resulting from non-ideal OTA characteristics (i.e. parasitic transconductance and parasitic capacitance), that cause practical components to deviate from their ideal behavior. The methodology of establishing a building block library is summarized in the following steps: (1) Identify the type of circuits applicable to the method and identify the building block of analog circuit design alternatives (2) Identify the standard cell and specify the items and boundaries required for each standard cell to be characterized in the library. (3) Collect standard cells' responses within the characterization boundary (4) Model the standard cells' responses. (5) Model the dependency of the standard cells' responses on the circuit designs. (6) Compute the necessary statistics to test the hypothesis. This research demonstrated the implementation of each step of building a library for higher-order cascaded biquad Gm-C filters as an example.

The thesis also explored the possibility of predicting manufacturing FOM of analog circuits through rank statistics. A copula-based model for analog parametric

performance response was first built from a small set of training data. Monte Carlo simulations were used to synthesize the analog response to manufacturing process variation models. The copula was employed to characterize the temperature dependence relationship of an example circuit. The method was demonstrated using a 45nm-design for a CMOS operational trans-conductance amplifier (OTA). The dependence structure in the performance of the OTA between two temperatures is a function of temperature and supply voltage. The study demonstrated that temperature miscorrelation can be the result of process variation and does not require the introduction of defects.

References

- [1] B. Lojek, *History of semiconductor engineering*. Springer, 2007.
- [2] A. S. Sedra, D. E. A. S. Sedra, K. C. Smith, and K. C. Smith, *Microelectronic circuits*. New York: Oxford University Press, 1998.
- [3] P. R. Gray, P. Hurst, R. G. Meyer, and S. Lewis, *Analysis and design of analog integrated circuits*. Wiley, 2001.
- [4] “Advanced design system.” <https://www.keysight.com/en/pc-1297113/advanced-design-system-ads?&cc=US&lc=eng>. ADS.
- [5] “Cadence virtuoso.” https://www.cadence.com/en_US/home/tools/custom-ic-analog-rf-design/circuit-design/virtuoso-schematic-editor.html. Cadence.
- [6] P. E. Allen and D. R. Holberg, *CMOS analog circuit design*. Elsevier, 2011.
- [7] L. S. Milor, “A tutorial introduction to research on analog and mixed-signal circuit testing,” *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, vol. 45, no. 10, pp. 1389–1407, 1998.
- [8] B. Vinnakota, *Analog and mixed-signal test*. Prentice Hall PTR, 1998.
- [9] P. Antognetti and G. Massobrio, *Semiconductor device modeling with SPICE*. McGraw-Hill, Inc., 1990.

- [10] M. V. Dunga, X. Xi, J. He, W. Liu, K. M. Cao, X. Jin, J. J. Ou, M. Chan, A. M. Niknejad, and C. Hu, “Bsim4. 6.0 mosfet model,” *University of California, Berkeley*, 2006.
- [11] “Bsim4 model.” <https://bsim.berkeley.edu/models/bsim4/>. bsim4model.
- [12] B. A. Berg, *Markov chain Monte Carlo simulations and their statistical analysis: with web-based Fortran code*. World Scientific Publishing Company, 2004.
- [13] R. E. Caflisch, “Monte carlo and quasi-monte carlo methods,” *Acta numerica*, vol. 7, pp. 1–49, 1998.
- [14] D. A. Johns and K. Martin, *Analog integrated circuit design*. John Wiley & Sons, 2008.
- [15] L. Lavagno, I. L. Markov, G. Martin, and L. K. Scheffer, *Electronic Design Automation for IC System Design, Verification, and Testing*. CRC Press, 2017.
- [16] N. Horta, “Analog and mixed-signal ic design automation: Synthesis and optimization overview,” in *Proc. 5th Conference on Telecommunications, Tomar, Portugal*, 2005.
- [17] N. Lourenço, *LAYGEN: Automatic layout generation of analog ICs, from a system to device level using both hierarchical template descriptions and intelligent computing techniques*. PhD thesis, Master thesis, Dept. Electrical and Computer Engineering, Instituto Superior, 2007.
- [18] T. Eeckelaert, R. Schoofs, G. Gielen, M. Steyaert, and W. Sansen, “An efficient methodology for hierarchical synthesis of mixed-signal systems with fully integrated building block topology selection,” in *Proceedings of the conference on Design, automation and test in Europe*, pp. 81–86, EDA Consortium, 2007.

- [19] M. Adamski and A. A. Barkalov, *Architectural and sequential synthesis of digital devices*. University of Zielona Góra Press, 2006.
- [20] R. Rudell, “Tutorial: Design of a logic synthesis system,” in *33rd Design Automation Conference Proceedings, 1996*, pp. 191–196, IEEE, 1996.
- [21] R. Harjani, R. A. Rutenbar, and L. R. Carley, “Oasys: A framework for analog circuit synthesis,” *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 8, no. 12, pp. 1247–1266, 1989.
- [22] Y. Gu and C. Wang, “Transgraph: Hierarchical exploration of transition relationships in time-varying volumetric data,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 17, no. 12, pp. 2015–2024, 2011.
- [23] T. R. Dastidar, P. Chakrabarti, and P. Ray, “A synthesis system for analog circuits based on evolutionary search and topological reuse,” *IEEE Transactions on evolutionary computation*, vol. 9, no. 2, pp. 211–224, 2005.
- [24] Dian Zhou and Guanming Huang, “Design automation of analog circuit considering the process variations,” in *2013 IEEE 10th International Conference on ASIC*, pp. 1–4, Oct 2013.
- [25] A. A. I. Ahmed and L. Zhang, “Fast parasitic-aware synthesis methodology for high-performance analog circuits,” in *2012 IEEE International Symposium on Circuits and Systems (ISCAS)*, pp. 2155–2158, May 2012.
- [26] W. Zeng, H. Zhu, X. Zeng, D. Zhou, R. Liu, and X. Li, “C-yes: An efficient parametric yield estimation approach for analog and mixed-signal circuits based on multicorner-multiperformance correlations,” *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 36, pp. 899–912, June 2017.

- [27] Z. Rojec, A. Burmen, and I. Fajfar, “Analog circuit topology synthesis by means of evolutionary computation,” *Engineering Applications of Artificial Intelligence*, vol. 80, pp. 48 – 65, 2019.
- [28] B. Razavi, *Design of Analog CMOS Integrated Circuits*. McGraw-Hill electrical engineering series, McGraw-Hill, 2001.
- [29] K. L. Su, *Analog filters*. Springer Science & Business Media, 2012.
- [30] R. Schaumann, H. Xiao, and V. V. Mac, *Design of Analog Filters 2nd Edition*. Oxford University Press, Inc., 2009.
- [31] R. Schaumman, M. Ghausi, and K. Laker, “Design of analog filters: passive, active rc and switched capacitors,” *London, UK: Prentice Hall, section*, vol. 7, no. 1, p. 429, 1990.
- [32] A. Singhee and R. A. Rutenbar, “Why quasi-monte carlo is better than monte carlo or latin hypercube sampling for statistical circuit analysis,” *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 2010.
- [33] J. Jaffari and M. Anis, “On efficient lhs-based yield analysis of analog circuits,” *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 2011.
- [34] C. Fang, F. Yang, X. Zeng, and X. Li, “Bmf-bd: Bayesian model fusion on bernoulli distribution for efficient yield estimation of integrated circuits,” 2014.
- [35] F. Taylor and A. Williams, *Electronic Filter Design Handbook*. McGraw-Hill, USA, 1995.
- [36] J. E. Kardontchik, *Introduction to the Design of Transconductor-Capacitor Filters*. Kluwer Academic Publishers Boston:, 1992.

- [37] R. Schaumann, M. Ghausi, and K. Laker, *Design of Analog Filters: Passive, Active RC, and Switched Capacitor*. Prentice-Hall series in electrical and computer engineering, Prentice-Hall, 1990.
- [38] G. Shirley, “Analog design for manufacturing,” 2016.
- [39] R. Behzad, *RF Microelectronics*, vol. 2. Prentice Hall New Jersey, 1998.
- [40] R. Spence and R. S. Soin, *Tolerance Design of Electronic Circuits*. Addison-Wesley, 1988.
- [41] K. M. Butler, J. M. Carulli, and J. Saxena, “Modeling test escape rate as a function of multiple coverages,” in *2008 IEEE International Test Conference*, pp. 1–9, Oct 2008.
- [42] C. G. Shirley and W. R. Daasch, “Copula models of correlation: A dram case study,” *IEEE Transactions on Computers*, vol. 63, no. 10, pp. 2389–2401, 2013.
- [43] B. Gilbert, *Design for Manufacture*. Springer US, 1st ed., 2002.
- [44] S. Sun, X. Li, and C. Gu, “Structure-aware high-dimensional performance modeling for analog and mixed-signal circuits,” 2013.
- [45] X. Li, “Finding deterministic solution from underdetermined equation: large-scale performance variability modeling of analog/rf circuits,” *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 2010.
- [46] Y. Tsividis and C. McAndrew, *Operation and Modeling of the MOS Transistor*. Oxford Univ. Press, 2011.
- [47] J. V. Faricelli, “Layout-dependent proximity effects in deep nanoscale cmos,” 2010.

- [48] P. G. Drennan, M. Kniffin, and D. Locascio, “Implications of proximity effects for analog design.,” 2006.
- [49] F. Jiao and A. Doboli, “A low-voltage, low-power amplifier created by reasoning-based, systematic topology synthesis,” *IEEE International Symposium on Circuits and Systems*, 2015.
- [50] A. Doboli and R. Vemuri, “Exploration-based high-level synthesis of linear analog systems operating at low/medium frequencies,” *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 2003.
- [51] S. Weaver, B. Hershberg, and U. K. Moon, “Digitally synthesized stochastic flash adc using only standard digital cells,” *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 61, pp. 84–91, Jan 2014.
- [52] W. Deng, D. Yang, T. Ueno, T. Siriburanon, S. Kondo, K. Okada, and A. Matsuzawa, “A fully synthesizable all-digital pll with interpolative phase coupled oscillator, current-output dac, and fine-resolution digital varactor using gated edge injection technique,” *IEEE Journal of Solid-State Circuits*, vol. 50, pp. 68–80, Jan 2015.
- [53] S. M. Newton and P. R. Kinget, “A 4th-order analog continuous-time filter designed using standard cells and automatic digital logic design tools,” *IEEE International Symposium on Circuits and Systems*, 2016.
- [54] S. Li, B. Xu, D. Z. Pan, and N. Sun, “A 60-fj/step 11-enob vco-based ctdsm synthesized from digital standard cell library,” in *2019 IEEE Custom Integrated Circuits Conference (CICC)*, pp. 1–4, 2019.
- [55] T. Mukherjee, L. R. Carley, and R. A. Rutenbar, “Efficient handling of operating range and manufacturing line variations in analog cell synthesis,” *IEEE Trans-*

- actions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 19, pp. 825–839, Aug 2000.
- [56] T. Mukherjee, L. R. Carley, and R. A. Rutenbar, “Synthesis of manufacturable analog circuits,” pp. 586–593, Nov 1994.
 - [57] R. A. Rutenbar, G. G. E. Gielen, and B. A. Antao, “Efficient analog circuit synthesis with simultaneous yield and robustness optimization,” pp. 699–702, 2002.
 - [58] E. K. Saragu P. Mohanty, “Incorporating manufacturing process variation awareness in fast design optimization of nanoscale CMOS VCOs,” *IEEE Transactions on Semiconductor Manufacturing*, 2014.
 - [59] J. C. Rudell, J. A. Weldon, J.-J. Ou, L. Lin, and P. Gray, “An integrated gsm/dect receiver: Design specifications,” *UCB Electronics Research Laboratory Memorandum*, 1998.
 - [60] C. G. Shirley and W. R. Daasch, “Copula models of correlation: A dram case study,” *IEEE Transactions on Computers*, vol. 63, pp. 2389–2401, Oct 2014.
 - [61] Z. Zhao and L. Zhang, “An automated topology synthesis framework for analog integrated circuits,” *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, pp. 1–1, 2020.
 - [62] B. Efron and R. Tibshirani, *An Introduction to the Bootstrap*. Chapman & Hall/CRC Monographs on Statistics & Applied Probability, Taylor & Francis, 1994.
 - [63] A. Agarwal, H. Sampath, V. Yelamanchili, and R. Vemuri, “Accurate estimation of parasitic capacitances in analog circuits,” in *Proceedings of the conference on*

- Design, automation and test in Europe-Volume 2*, p. 21364, IEEE Computer Society, 2004.
- [64] S. Davidson, “Understanding ntf components from the field,” in *IEEE International Conference on Test, 2005.*, pp. 10–pp, IEEE, 2005.
 - [65] Z. Conroy, G. Richmond, X. Gu, and B. Eklow, “A practical perspective on reducing asic ntfs,” in *IEEE International Conference on Test, 2005.*, pp. 7–pp, IEEE, 2005.
 - [66] O. Kempthorne, “The design and analysis of experiments.,” 1952.
 - [67] D. C. Montgomery, *Design and Analysis of Experiments*. Wiley, 2013.
 - [68] A. Bounceur, S. Mir, and H.-G. Stratigopoulos, “Estimation of analog parametric test metrics using copulas,” *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 30, no. 9, pp. 1400–1410, 2011.
 - [69] L. Fang, M. Lemnawar, and Y. Xing, “Cost effective outliers screening with moving limits and correlation testing for analogue ics,” in *2006 IEEE International Test Conference*, pp. 1–10, IEEE, 2006.
 - [70] M. Sklar, “Fonctions de repartition an dimensions et leurs marges,” *Publ. inst. statist. univ. Paris*, vol. 8, pp. 229–231, 1959.
 - [71] R. B. Nelsen, *An introduction to copulas*. Springer Science & Business Media, 2007.
 - [72] P. K. Trivedi, D. M. Zimmer, *et al.*, “Copula modeling: an introduction for practitioners,” *Foundations and Trends® in Econometrics*, vol. 1, no. 1, pp. 1–111, 2007.

- [73] M. B. Yelten, P. D. Franzon, and M. B. Steer, “Process mismatch analysis based on reduced-order models,” in *Thirteenth International Symposium on Quality Electronic Design (ISQED)*, pp. 648–655, IEEE, 2012.
- [74] E. W. Frees and E. A. Valdez, “Understanding relationships using copulas,” *North American actuarial journal*, vol. 2, no. 1, pp. 1–25, 1998.
- [75] W. M. Lindermeir, T. J. Vogels, and H. E. Graeb, “Analog test design with i/sub dd/ measurements for the detection of parametric and catastrophic faults,” in *Proceedings Design, Automation and Test in Europe*, pp. 822–827, Feb 1998.