Deterministic Multi-Objective Variation-Aware Analog Circuit Sizing for Carbon Nanotube Technology

By
©Zahra Heshmatpour

A dissertation submitted to the School of Graduate Studies
in partial fulfillment of the requirements for the degree of

Master of Engineering

Faculty of Engineering & Applied Science
Memorial University of Newfoundland

Supervisory Committee

Prof. Lihong Zhang (Co-Supervisor)
Prof. Howard Heys (Co-Supervisor)

MAY 2022

St. John's, Newfoundland

Abstract

The modern CMOS technology, which is based on Moore's law semiconductor devices, continuously scales down to achieve higher speed and larger packing density. The need for downscaling of MOSFETs below 45nm technology node results in transistor malfunction due to short channel effects, mobility degradation and higher power consumption. Carbon nanotube field-effect transistors (CNFETs) are one of the promising candidates to substitute CMOS technology for next-generation integrated circuits thanks to their small size and excellent electrical properties, such as quasi-ballistic transport and high carrier mobility.

Process variation, such as CNT (Carbon nanotube) diameter, presence of metallic CNTs, misaligned CNTs, high metal-CNT contact resistance and CNT density variations, hinders wide adoption of CNFET technology. As a result, many techniques have been developed to overcome the fabrication variation for digital CNFET circuits while analog CNFET circuits generally lack a proper approach to avoid performance failure. This thesis starts with description of CNT parameters and then considers the most critical parameters that have the greatest impact on analog circuit performance.

Then we use a design centering approach for circuit sizing to obtain the optimal value of design parameters against carbon nanotube process variation to ensure to meet the performance specification and enhance the functional robustness. Subsequently, we present a modified method to solve the generalized boundary curve (GBC) optimization as a starting point to develop our fully deterministic multi-objective sizing flow. In the next step we take advantage of the normal boundary intersection (NBI) method in combination

with our modified GBC method to develop our multi-objective optimization analog CNFET sizing design methodology considering carbon nanotube parameter process variation.

The performance of our methodology is analysed by optimizing two operational amplifier (op-amp) circuits and two current conveyor (CCII) circuits, which are considered to be the common building blocks in the interface circuits to the analog world. The experimental results demonstrate that our proposed method can reach a better approximation to the Pareto front compared to the other state-of-the-art multi-objective methods.

Acknowledgments

I would like to thank my supervisors Prof. Lihong Zhang and Prof. Howard Heys for their mentorship, guidance, and vision toward studying emerging new technologies and their abundant support and assistance in my Master's study, research, writing papers and thesis. This work was supported by Memorial University of Newfoundland.

To loving memory of my grandmother

Table of Contents

List of Tables

List of Figs

List of Abbreviations

Word	Abbreviations
Carbon Nanotube	**CNT**
Carbon Nanotube Field Effect Transistor	**CNFET**
Characteristic Boundary Curve	**CBC**
Generalized Boundary Curve	**GBC**
Normal Boundary Intersection	**NBI**
Convex Hull of Individual Minima	**CHIM**
Interior Point OPTimizer	**IPOPT**
Sequential Quadratic Programing	**SQP**
Metal Oxide Semiconductor Field-Effect Transistor	**MOSFET**
Complementary Metal Oxide Semiconductor	**CMOS**
Multi walled CNT	**MWCNT**
Single walled CNT	**SWCNT**
Catalyzed Chemical Vapor Deposition	**CVD**

Chapter 1

Introduction to Carbon Nanotube Transistors

1.1. Introduction

Based on Moore's law, every two years, the dimensions of individual devices that can be incorporated in an integrated circuit decrease by a factor of two [1]. For many years, transistor downscaling has been the most essential and practical approach to achieve larger chip densities and lower power using CMOS technology. However, when transistor dimensions have been reduced to less than 22nm, CMOS dimensional scaling has been encountering challenges in terms of the manufacturing process and device performance.

Due to decreased gate control, increased leakage current, tunneling of electrons, considerable power consumption in CMOS circuits, and the fact that aggressive scaling affects the electron and hole mobility the transistor performance gains decrease in each generation [2]. Therefore, next-generation integrated circuits need a transistor with better semiconductor materials for the current channel to overcome the existing challenges.

Carbon nanotube field-effect transistors (CNFETs) have been introduced as a potential substitute to CMOS technology. Thanks to the cylindrical geometry of the channel, CNFETs have better gate control capability, leading to better scalability than other

alternative technologies. In addition, CNFET's high channel mobility [3], high electrostatics [4] and its ballistic transport of carriers enable the CNFET-based circuits to use lower supply voltages, making it serve as a good candidate for superior energy and performance efficiency.

The fundamental difference between a CNFET and a MOSFET is the channel under the gate. On the one hand, the CNFET channel already exists through arrays of semiconducting carbon nanotube; on the other hand, the MOSFET channel forms during the device operation [5]. The current in CNFETs highly depends on the variaty of parameters of carbon nanotubes (CNTs). Therefore, process variations and manufacturing defects have a massive impact on the performance of the circuits [6]. The process variation can include CNT diameter variations, presence of metallic CNTs, CNT density variations due to non-uniform spacing between CNTs during CNT growth as well as misalignment and mispositioning of the CNTs.

1.2. Overview of CNFET

CNFET consists of multiple CNTs forming a channel connecting the drain to the source, and the gate controls the current flow through the CNTs. CNTs are made of graphene sheets rolled up to make a cylinder with a diameter of a few nanometers; by doing so, the atomic electromigration is more convoluted, enabling carbon nanotubes to be highly stable structures that can manage high currents without any damage [7].

The resulting cylinder is called a single-walled carbon nanotube (SWCNT). If more than one SWCNT of varying diameters are folded concentrically, they form a multi-walled

CNT (MWCNT), as illustrated in Fig. 1. Depending on CNT atomic arrangement, SWCNT can be used as a semiconductor or conductor [8]. The angle of the nanotube's hexagonal carbon-atom lattice determines the chirality vector (n, m). A CNT is considered as a semiconductor unless the chirality vector has the following conditions: $n = m$ or $n - m = 3j$ (j is an integer), which determines that the CNT is metallic.

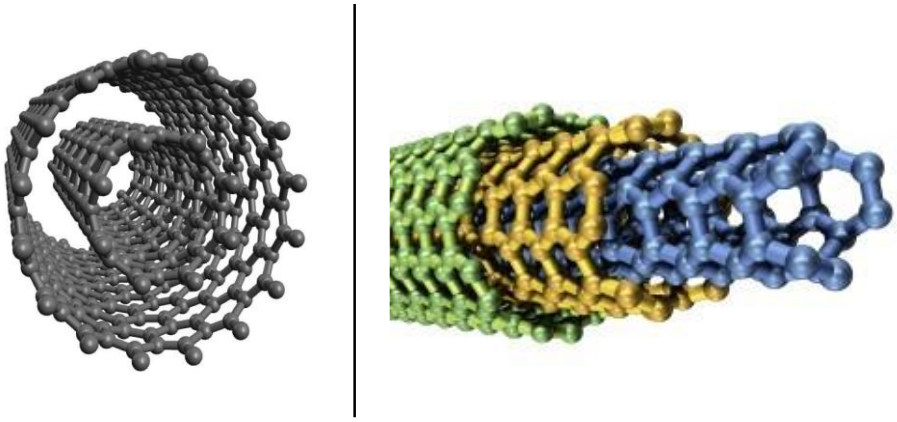

Fig. 1. MWCNT [9]

Ballistic transport is another important phenomenon in CNT. The average distance traveled by a carrier before encountering a scattering event is known as the mean free path. Ballistic transport in CNFET happens where electrons travel at high speeds through the CNT without scattering; in other words, it happens when the mean free path exceeds the carrier's traveling distance. Therefore, the electrons face no resistance, and no energy dissipation occurs through the channel. One carrier can be scattered by many diverse objects, such as ionized defects, impurities, and other carriers.

The nanotube forces carriers to only go back and forth, removing the potential of small-angle scattering events reversing the electron's path. Although optical phonons have

enough momentum to reverse carrier orientation, they are too energetic to exist at room temperature and low voltages. Furthermore, with variation-aware manufacturing of carbon nanotubes, static flaws such as impurities, faults, interfaces, and so on may be avoided [10].

In this thesis research, we have used a CNFET model from Stanford University for analyzing CNFET circuits. Table 1 illustrates all the available parameters in the Stanford VS-CNFET model and their description [11]. The intrinsic drain current and terminal charges in this library are calculated based on the virtual source (VS) concept [12].

Table 1. VS-CNFET Input Parameters [11]

Name	**Suggested Scope**	**Description**
type	-1 or 1	type of transistor. 1: nFET; -1: pFET
s	[2.5e-9:inf)	spacing between the CNTs (center-to-center) [m]
W	[s:inf)	transistor width [m]
Lg	[5e-9:100e-9]	physical gate length [m]
Lc	[1e-9:inf)	contact length [m]
Lext	(0:inf)	source/drain extension length [m] (or spacer length)
d	[1e-9:2e-9]	CNT diameter [m]
tox	[1e-9:10e-9]	gate oxide thickness [m]
kox	[4:25]	gate oxide dielectric constant
kcnt	1	CNT dielectric constant
ksub	[1:kox)	substrate dielectric constant
kspa	[1:16)	source/drain spacer dielectric constant
Hg	[0:inf)	gate height [m]
Efsd	[-0.1:0.5]	Fermi level to the band edge [eV] at the source/drain, related to the doping density. The larger the Efsd, the higher the doping density in the source/drain extensions.
Vfb	[-1:1]	flat band voltage [V] (for threshold voltage adjustment)
Rcmod	0 or 1 or 2	contact mode. 0: user-defined value, Rs0; 1: diameter-dependent transmission line model; 2: diameter-independent transmission line model (*R*c is calculated at *d* = 1.2 nm regardless of the input *d*)
Rs0	[0:inf)	User-defined series resistance (Ω)

The virtual source is a semi-empirical model that describes the current versus voltage characteristics of the transistor and is valid in all regions of operation. The VS also provides intrinsic charge descriptions that extend all the way to the ballistic regime. The VS model maintains the advantage of using only a limited number of input parameters, most of which have straightforward physical meanings and can be easily measured from device characterization [13].

Fig. 2, illustrates a CNFET alongside its parameters, where L_{gate} and W_{gate} are transistor length and width, respectively. The semiconducting intrinsic CNT, shown in Fig. 2, under the gate acts as the channel and heavily doped CNT regions outside the gate form the source/drain extension regions.

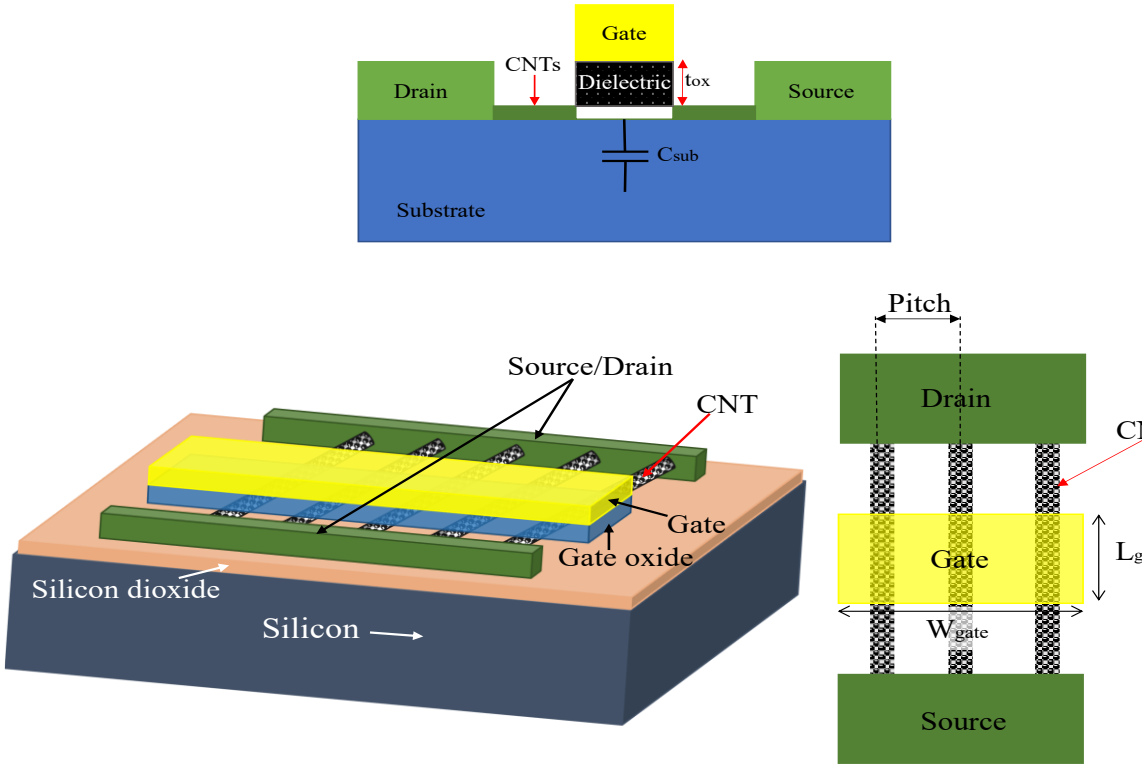

Fig. 2. CNFET structure

The CNFET width (W_{gate}) is determined by different parameters, like the number of tubes (N), CNT diameter (D_{CNT}), and inter CNT pitch (S). The relationship between W_{gate} and these parameters can be given by the following equation:

$$W_{gate} = (N - 1) \cdot S + D_{CNT}. \quad (1)$$

The CNT diameter is determined by its chirality. CNFET threshold voltage (V_{th}) and bandgap (E_g) are inversely proportional to the CNT diameter, which can be expressed by

$$D_{CNT} = \frac{a_{c-c} \cdot \sqrt{n^2 + m^2 + nm}}{\pi}, \quad (2)$$

$$V_{th} = \frac{V_\pi \cdot a_{c-c}}{\sqrt{3}. q \cdot D_{CNT}}, \quad (3)$$

$$E_g = \frac{2E_P \cdot a_{c-c}}{D_{CNT}}, \quad (4)$$

where a_{c-c} = 0.142 nm is the carbon–carbon distance, (n, m) is chirality vector, $V_\pi \approx$ 3.033 eV is the carbon $p - p$ bond energy, q is the electronic charge, and E_P = 3eV is the tight-binding parameter.

The variation of CNT parameters, which affects the current path in CNFET, should be considered during the circuit design to ensure the performance robustness of the circuit. In this work, we consider the variation of CNT diameter (D_{CNT}), number of CNT (implemented by assuming a constant value for inter CNT pitch and changing W_{gate}), and L_{gate} in our process-variation-aware sizing optimization.

1.3. Types of CNFET

There are three types of CNFETs [14]:

1.3.1. Partially Gated CNTFET

The partially gated CNTFET (PGCNFET) has a uniformly doped or intrinsic channel. The transistor operates in depletion mode because consistent doping exists across the channel. Based on the doping type, they exhibit n-type or p-type behaviour [15].

1.3.2. MOSFET-like CNFET

MOSFET-like CNFET (Fig. 3(a)) have three sections. The intrinsic zone lies below the gate, while the other two sections are highly n-type or p-type doped. When a gate potential is applied, they function in either a pure p- or n-type enhancement mode or a depletion mode, depending on the concept of barrier height modulation. This type of CNFET is promising because they are unipolar, lack of Schottky-barrier minimizes OFF leakage current, and the source-to-channel junction has a greater ON current in the operating state. Thus, we consider the MOSFET-like CNFET for our circuits in this thesis work [16].

1.3.3. Schottky-barrier CNFET

Schottky-barrier CNFET (SBCNFET shown in Fig. 3(b)) is made by contacting metal directly with carbon nanotubes, forming Schottky barriers at the nanotube-metal junction. They operate on the idea of direct tunneling via a Schottky barrier at the source-channel

junction, with the width of the barrier being regulated by the gate voltage. The existence of Schottky barriers in the ON state drastically restricts the transconductance of the nanotube transistors and decreases the current. SBCNFETs also have high ambipolar properties [17].

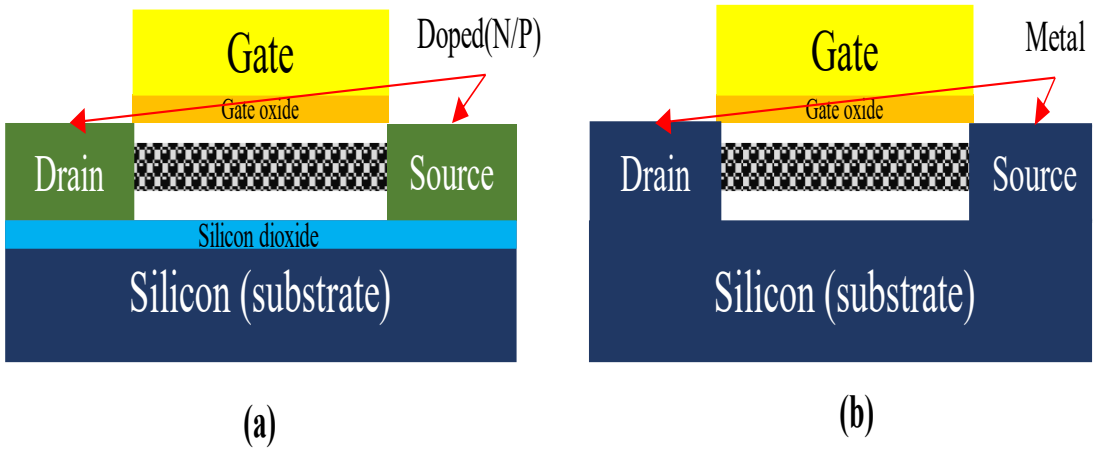

Fig. 3. CNFET types: (a) MOSFET like CNFET, (b) Schottky-barrier CNFET.

1.4. CNFET Structure

Due to thick SiO_2 layer and back gate design (Fig. 4), the initial CNFETs fabricated on silicon oxide had poor gate coupling. After the invention of CNFETs with top-gate geometry in 2002, a significant improvement happened to CNFET performance [18]. Later wrap-around gate CNFETs, also known as cylindrical gate-all-around CNFETs, were developed in 2008 [19]. This is an advancement over the top-gate structure for CNFETs.

The top-gate and cylindrical all-around gate are two types of gate positions available in the Stanford library.

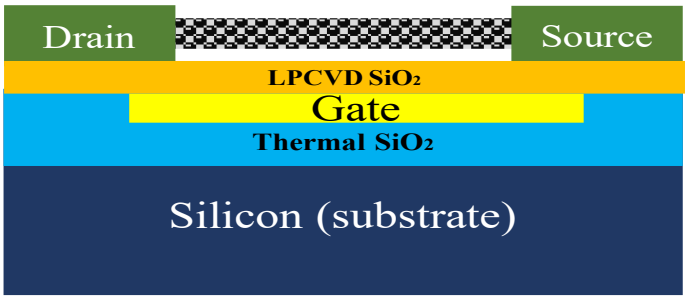

Fig. 4. Back gate CNFET [20]

1.4.1. Back Gate CNFET

The first techniques used for CNFET fabrication involve the way of depositing the CNTs on top of pre-patterned parallel lines of metal across a silicon dioxide substrate in a random style. The semiconductor CNTs that were between two metal strips made a basic field-effect transistor by considering one metal strip as drain contact and the other one as source contact. The silicon oxide substrate is considered as the gate oxide, with an added metal contact on the back to act as the gate. The back gate CNFET is illustrated in Fig.4.

This method has many flaws that result in the fabrication of non-optimized transistors. First of all, simply placing CNTs on top of the metal contact results in a minimal contact area; therefore, the metal contact has very little interaction with the CNT. Another drawback of this method is the back gate geometry. Its thickness makes low voltage switching challenging. Moreover, this production procedure results in poor contact between the gate dielectric and the CNT.

1.4.2. Top-gate CNFET

For the construction of a top-gate CNFET solution-deposition of single-walled carbon nanotubes onto a silicon oxide substrate is the initial stage. An atomic force microscope or a scanning electron microscope is used to locate individual nanotubes. High-resolution electron beam lithography defines and shapes source and drain connections once an individual tube has been separated. By enhancing contact adhesion between contacts and CNT, a high temperature annealing stage lowers contact resistance. Evaporation or atomic layer deposition is then used to produce a thin top-gate dielectric on top of the nanotube. The operation is then completed by depositing the top-gate contact on the gate dielectric.

Because the gate contacts are electrically separated from one another, arrays of top gated CNFETs can be produced on the same wafer. Also, because the gate dielectric is thin, a more significant electric field for the nanotube may be created with a smaller gate voltage. We use the top-gate structure as shown in Fig. 5 to do our simulations.

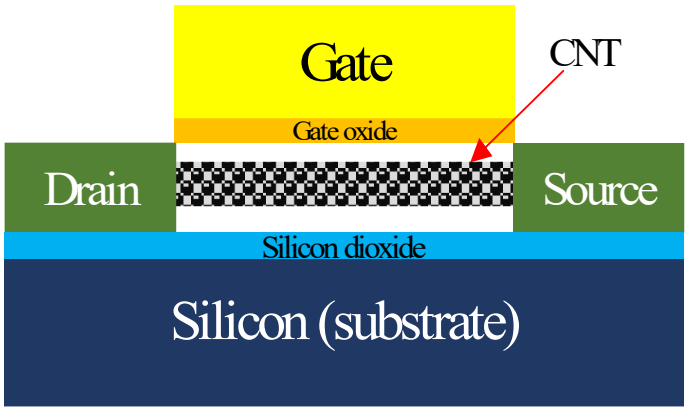

Fig. 5. Top gate CNFET

1.4.3. Cylindrical Gate-all around CNFET

Fig. 6 shows a cylindrical gate-all-around CNFET, which is an improved structure over the top-gate device design. Instead of only gating the section of the CNT closest to the metal gate contact, the whole circumference of the nanotube is gated in this device, which should improve the electrical performance of CNFETs by lowering leakage current and increasing the device on/off ratio.

The initial step in device construction is to wrap CNTs in a gate dielectric and gate contact using atomic layer deposition. After that, the wrapped nanotubes are solution-deposited on an insulating substrate, where the wrappings are partially etched away, revealing the nanotube ends. The contacts for the source, drain, and gate are then deposited on the CNT ends and the metallic outer gate wrapping, respectively [11].

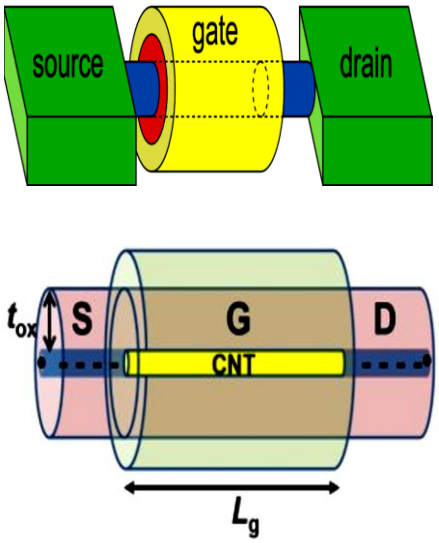

Fig. 6. Cylindrical gate-all around CNTFET [11, 21]

1.5. CNT Manufacturing Method

The first CNFETs were fabricated using carbon nanotubes on top of SiO_2 and a Si-substrate that worked as a back gate. Laser ablation synthesis was used to create the carbon nanotubes, randomly spread across SiO_2. This approach lost its usage because it relied on an unknown amount of carbon nanotubes bridged between the electrodes showing unsatisfactory results. As a result, carbon nanotubes were generated only in specified metallic catalyst islands using chemical vapor deposition of methane on the patterned substrate [22].

In the following sections, we explain several methods of making CNTs [23].

1.5.1. Arc Method

The carbon arc discharge technique is the most frequent and easiest approach to make CNTs due to its simplicity. However, because the process creates a complex combination of components, additional purification is required. In this method the CNTs are made by arc-vaporizing two carbon rods, which are put end to end in an enclosure filled with inert gas at low pressure. The discharge vaporizes one of the carbon electrode surfaces and leaves a small rod-shaped deposit on the other. The uniformity of the plasma arc and the temperature of the deposit formed on the carbon electrode are important factors in producing high yield CNTs.

1.5.2. Laser Methods

A dual-pulsed laser was used in 1996 to produce CNTs. The first laser vaporization pulse was followed by a second pulse to uniformly evaporate the object. The quantity of carbon deposited as soot is reduced when two subsequent laser pulses are used. The second laser pulse breaks up the bigger particles ablated by the previous one and feeds them into the nanotube structure as it grows. This process produces a sheet of ropes, and each rope is made up of a bundle of single-walled nanotubes oriented along a common axis. By adjusting the growth temperature, and other process variables, the average nanotube diameter and size distribution can be varied.

1.5.3. Catalyzed Chemical Vapor Deposition

A typical approach for producing different carbon compounds is catalyzed chemical vapor deposition (CVD) of hydrocarbons over a metal catalyst. Catalytic CVD of acetylene over cobalt and iron catalysts based on silica or zeolite may produce large volumes of CNTs. In the ethylene atmosphere, supported catalysts like iron, cobalt, and nickel, which include either a single metal or a combination of metals, appear to encourage the formation of single-walled nanotubes. Methane has been utilized as a carbon source in the past; it has been used to create nanotube chips, which contain isolated single-walled nanotubes at precise places. Fig. 7 illustrates the CVD process.

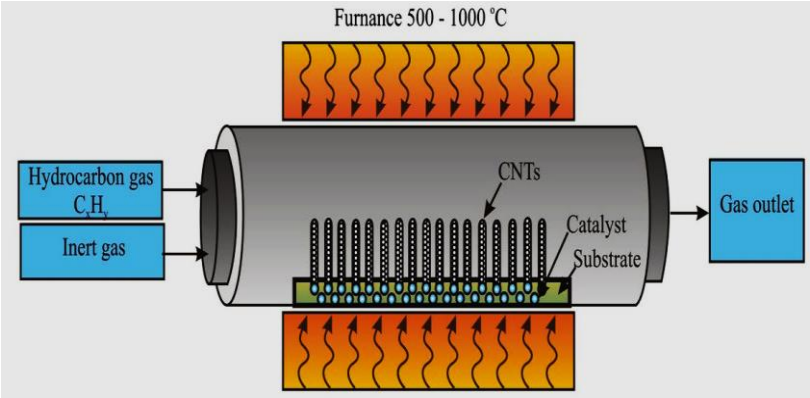

Fig. 7. CVD Process [24]

As shown above this process has a reaction chamber as well as a heating furnace. The CVD growth process is usually carried out in the reaction chamber, which is commonly made up of a quartz tube. The gases are pyrolyzed in a reaction chamber heated by a high-temperature electric tube furnace.

The temperature range of the reaction is frequently greater than 500°C. The catalyst is deposited on substrates first, then placed into the quartz-tube reactor. The temperature is then raised to a certain temperature (e.g., 700°C) as the reactor reaches 700°C, NH_3 gas is supplied to the chamber, followed by C_xH_x (e.g., C_2H_2) for the rest of the growth. The samples are retrieved from the reactor after they have grown and have been cooled to 250°C under vacuum [25].

1.6. Thesis Objective and Structure

The thesis objective is to ensure robust performance of CNFET based circuits under process variation, in this thesis we analyze CNT variations that can lead to device

malfunction. Our goal is to develop a method that helps us to design and optimize CNFET circuits for analog applications.

The core chapters in the thesis are Chapter 3 to Chapter 6. We especially target the design and optimization methodologies for the generic CNFETs circuits. The thesis structure is organized as follows:

• Chapter 2 discusses the previous works on CNFET variation and circuit optimization.

• Chapter 3 primarily investigates the effect of different parameters process variation on the performance of CNFET transistors and circuits.

• Chapter 4 presents a modified version of the GBC optimization.

• Chapter 5 investigates the variation-aware multi-objective optimization.

• Chapter 6 presents the optimization of CNFET based circuits and explores their performance evaluation.

• Chapter 7 presents a summary of the thesis and offers the promising topics for future research.

1.7. Summary

In this chapter, we review the fundamentals of CNFET and its parameters, as well as its different types. Moreover, all the CNFET fabrication methods discussed in this chapter need a purification process of CNTs to separate the CNTs from other particles, such as carbon nanoparticles, residual catalysts, and other unwanted materials. In the next chapter, we will review some of the previous work studying CNFET variations and optimization.

Chapter 2

Literature Review

2.1. Introduction

This chapter will discuss the previous study on CNFET circuits and optimization methods used to optimize different CNFET based circuits. Subsequently, we outline the contributions of this thesis.

2.2. CNFET Circuit Study

Here, we divide the published papers in this area into three different categories.

2.2.1. CNFET vs. CMOS

In this section we discuss the papers that compared the performance of CNFET circuits with that of CMOS circuits.

G. Jing *et al.* [26] improved prior CNFET models by creating a basic model for ballistic transistors that incorporates both the ballistic transport and quantum capacitance constraints. They used the model to study MOSFET-like carbon nanotube FETs (CNFETs) as well as MOSFETs near the scaling limit and compared the results. Based on their

analysis, the CNFET showed better performance for the on-current because of the high gate capacitance and improved channel transport. Also, CNFET exhibited higher channel velocity due to increased mobility and the band structure of CNFET.

F. Rahman *et al.* [27] presented a performance comparison and evaluation for a carbon nanotube-based operational amplifier (CNT op-amp), using MOSFET-like SWCNFETs. They compared the results of a simulation-based design with a focus on the performance analysis of a CNFET op-amp with reference to a silicon-based op-amp in the 32nm technology. In their comparative analysis, the CNFET op-amp showed a promising rise in operating bandwidth, gain-bandwidth product (GBP), and switching speed, and a significant reduction was observed in circuit power consumption. A two-fold increase in CMRR also demonstrated significantly improved noise performance. The findings showed that the CNFET is an excellent replacement for silicon-based transistors for low-power, high-speed analog applications.

2.2.2. CNFET Parameter Variation

In this section we study the papers that considered variation in CNFET parameter and its effect on the performance of CNFET circuits.

J. Zhang *et al.* [28] proposed an imperfection-immune approach called VLSI-compatible metallic-CNT removal (VMR) to overcome CNT process variations such as metallic CNTs, mispositioned CNTs, and variations in CNT density. VMR combines layout design and processing to address metallic CNT difficulties. They observed that in directional CNT development, there is a high degree of CNT correlation, and it can be

exploited to successfully overcome the problems posed by metallic CNTs and fluctuations in CNT density.

Vendra and Chrzanowska-Jeske [29] investigated the effect of CNT length variation on the functional yield, performance of CNFETs and change on gate delay. CNFETs that have aligned CNT growth have been found to increase the functional yield of CNFET-based circuits when the diameter and quantity of CNTs are varied (can be done by varying CNT pitch or gate width while keeping the other one constant). They examined circuits with correlated CNFETs because all the transistors in correlated sets use the same sets of tubes, which have the same likelihood of failure and drive current. CNTs can grow to be quite long, but not all of them do. Due to CNT diameter variation and early catalyst precipitation, CNTs will attain different lengths. Their results showed that changing tube length increases the failure risk of rows of linked transistors by 85% when compared to the correlated CNFETs with the same CNT.

Banerjee *et al.* [30] studied the impact of manufacturing defects on the on-current in digital CNFET circuits and determined the effectiveness of various CNT parameter variations on the performance of digital circuits. They presented an approach to quantify the impact of process variations on CNFET circuits and described the steps in fabrication that can lead to defects and changes in the device performance. They modeled some of the existing defects caused by imperfect CNTs manufacturing, such as parasitic CNFETs and the existence of pinholes in the gate dielectric, in order to analyze their effect on the circuit performance.

Amer *et al.* [31] proposed a self-healing analog method to overcome analog CNFETs variations by using non-volatile resistive RAM (RRAM). They split a CNFET into multiple smaller units called sub-CNFET and fabricated a RRAM cell under the drain or source contact of each sub-CNFET. This method leverages the programmability of RRAM to "self-heal" analog circuits in the presence of the metallic CNT, where the RRAM acts as a low-resistance or high-resistance to ensure the circuit performance. Although it sounds interesting, this method heavily depends on introducing RRAM devices and accurate control of the self-heal ability.

Hills *et al.* [32] developed techniques to overcome CNFET manufacturing challenges such as CNT variations and less robust processing. This technique allows to build large-scale CNFET circuits using standard VLSI manufacturing technology. They took advantage of RINSE (Removal of Incubated Nanotubes through Selective Exfoliation) and MIXED (Metal Interface engineering crossed with Electrostatic Doping) along with their method DREAM (Designing REsilience Against Metallic CNTs). They applied them to form a complete method for building CNFET based VLSI circuits. Using their method, they demonstrated some of the most complicated CNFET circuits such as static SRAM memory arrays, and a 16-bit RISC-V microprocessor.

Hills *et al.* [33] overcame the existing inconsistency in carbon nanotubes and the problem of inefficiently controlling intrinsic nanoscale deformities. They were able to demonstrate the build-up of a complex CNFET chip, which was previously entirely based on silicon. This 16-piece chip, which has over 14,000 complementary CNFETs, was developed by using common industry streams. It can perform the conventional 32-bit

instructions on 16-bit information and address. Based on their experiments, the authors proposed a new approach to carbon nanotube manufacturing, design strategies for overcoming nanoscale defects, and a viable path for electronic devices beyond silicon.

2.2.3. CNFET Circuit Optimization

In this section we discuss the papers that tried to optimize CNFET circuit in order to achieve a better performance.

Ansari and Tripathi [34] and Imran *et al.* [35] considered CNFET based CCII circuits and optimized a high-accuracy CCII+ circuit and an ultra-wideband CCII± circuit, respectively, both of which are based on CNFETs. They studied the effect of CNT diameter, inter CNT pitch, and the number of CNT variations. And they reported the corresponding variation of gain, bandwidth, input port impedance, and output port impedance. In addition, after they made a comparison between the CNFET-based circuit and its CMOS equivalent, they confirmed the superiority of CNFET performance. Although the authors strived to derive their optimized circuits based on their specified simulation settings, they did not form a generic optimization method for solving analog CNFET circuit design in practice.

Yasir and Alam [36] studied a different analog CNFET circuit optimization method by applying g_m/I_D (transconductance / drain current) technique. Due to the existing difference between CNFET and CMOS design parameters, they provided a step-by-step approach showing how to use this method on CNFETs circuits. By tackling short channel devices and moderate and weak inversion behavior of the transistors, this work can mitigate the drawbacks of using the square-law model of MOSFET, which loses its accuracy in

more advanced technologies with smaller dimensions [37]. However, since some equations are vague and the exact value of some parameters can not be readily achieved, the authors selected to use random values. Moreover, this work is only limited to the sizing optimization of one circuit, CCII, which is too narrow to be general at the methodology level.

2.3. Optimization Methods Applied to Analog Circuits

In this section, we review a few key methods used for optimization in analog circuits design. Deterministic (or gradient-based) methods and heuristic (or probabilistic) methods are the two main types of optimization algorithms. Deterministic approaches are algorithms based on mathematical methods that employ information from past and current phases to decide the direction of future optimization. A heuristic algorithm determines an optimal solution by iteratively trying to improve a candidate solution. The most prevalent type of evolutionary algorithm is genetic algorithms, which apply the principle of survival of the fittest and typically need a high number of function evaluations. Once a maximum number of iterations has been reached, the program will end. A heuristic algorithm can solve a problem in a faster and more efficient way than non huristic methods however in order to achieve higher computational speed heuristic algorithms sacrifice optimality, accuracy, and precision of the result.

2.3.1. Deterministic Optimization Methods

In this section we discuss previous works that used deterministic optimization methods to optimize their circuits.

Dong and Zhang [38] considered the design centering method to optimize their circuit while considering manufacturing defects and avoiding their effects on circuit performance. They also presented an efficient method to solve the generalized boundary curve (GBC) when dealing with the process-variation-aware sizing problem. This method was suggested for the layout migration purpose where, to make the circuit robust, the designers may already have a reasonably good idea about initial device size estimation and maximum step-size change during algorithmic search. However, this assumption may not always be appropriate for the general process-variation-aware sizing problems.

Stehr *et al.* [39] studied multi-objective optimization using non-heuristic methods such as weighted sum, objectives as constraints, and normal boundary intersection (NBI). The weighted sum method cannot generate an evenly spread Pareto points and clusters in the regions with a strong curvature. Moreover, it is not able to detect all points on nonconvex Pareto fronts. While the constraint method overcomes the limitation of nonconvex Pareto curves, it still has unevenly distributed points on the Pareto front. Therefore, they focused on using a simulation-based approach and applying NBI to analog circuit sizing. Although having high-accuracy individual minima is essential for the NBI method, the technique of finding individual minima is not mentioned in the paper. In addition, sequential quadratic programming (SQP) is used to solve the NBI optimization. But it is known that the SQP method cannot handle non-smooth functions effectively.

Schreiber and Kampe [40] presented an algorithm by applying the NBI method with a constraint rotation scheme to perform an efficient system-level optimization and calculate the performance space boundary points. Their proposed method was used to optimize a fully differential transconductance amplifier with folded Cascode. However, the circuit which is optimized by their method is prone to manufacturing variation. Their method lacks accuracy as they tried to change the equality constraint to inequality to use the goal attainment method [41] for solving NBI.

2.3.2. Heuristic Optimization Methods

In this section we study some of the common heuristic multi-objective optimization methods.

Coello and Lechuga [42] proposed a method called MOPSO to extend the particle swarm optimization (PSO) to solve the multi-objective optimization problems. They used the Pareto dominance idea to select a particle's movement direction. By saving previously discovered nondominated vectors in a global repository, the other particles could subsequently reuse the data to guide their individual flight. This work focused on the generation of non-dominated vectors and how to maintain diversity in their algorithm; however, the PSO is an unconstrained search method, and the proposed MOPSO lacks comprehensive consideration of constrained multi-objective optimization.

Deb *et al.* [43] suggested a non-dominated sorting-based multi-objective evolutionary algorithm called non-dominated sorting genetic algorithm II (NSGA II) to overcome the existing problems. They presented a fast non-dominated sorting approach with lower

computational complexity. They developed a fast non-dominated sorting approach with lower computational complexity. They considered a selection operator to form a mating pool by combining the parent and child populations that choose the finest N solutions according to spread and fitness. According to the experimental results, NSGA II outperformed two other famous methods Pareto archived evolution strategy (PAES) and strength Pareto evolutionary algorithm (SPEA).

2.4. Summary

In this chapter we have reviewed the previous studies on optimization method, CNFET parameter variation and optimizing CNFET circuits. Although several techniques have been proposed to overcome the process-variation problems in digital CNFET circuits, it is still an obstacle in manufacturing the analog CNFET circuits due to lack of such research.

Chapter 3

Variation Aware Parameter Selection and Flow of Methodology

3.1. Introduction

CNFET device manufacturing can cause variation in CNT alignment, CNT diameter [44], CNT density variations caused by non-uniform inter-CNT spacing [45], and type (metallic or semiconductor) of CNT. These changes can significantly impact CNFET performance and may cause CNFET failure, compromising its application. This chapter evaluates the impact of the main CNT parameter variations on the CNFET circuits.

3.2. CNFET Current-Voltage Characteristics

In this section we derive the drain current variation based on the gate-source and drain-source voltage variation for an N-CNFET and P-CNFET to better understand the CNFET transistors performance.

3.2.1. N-CNFET

An N-CNFET schematic is shown in Fig. 8. In order to examine its current and transconductance variation based on drain and gate voltage, we implemented the circuit in Fig. 8 in Hspice using Stanford University virtual source model [11].

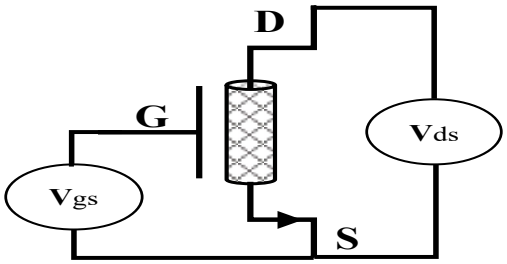

Fig. 8. N-CNFET

First, we examine the variation of I_d and g_m values based on V_{gs} variation and present the results in Fig. 9 and Fig. 10, where each line has a different value of V_{ds}.

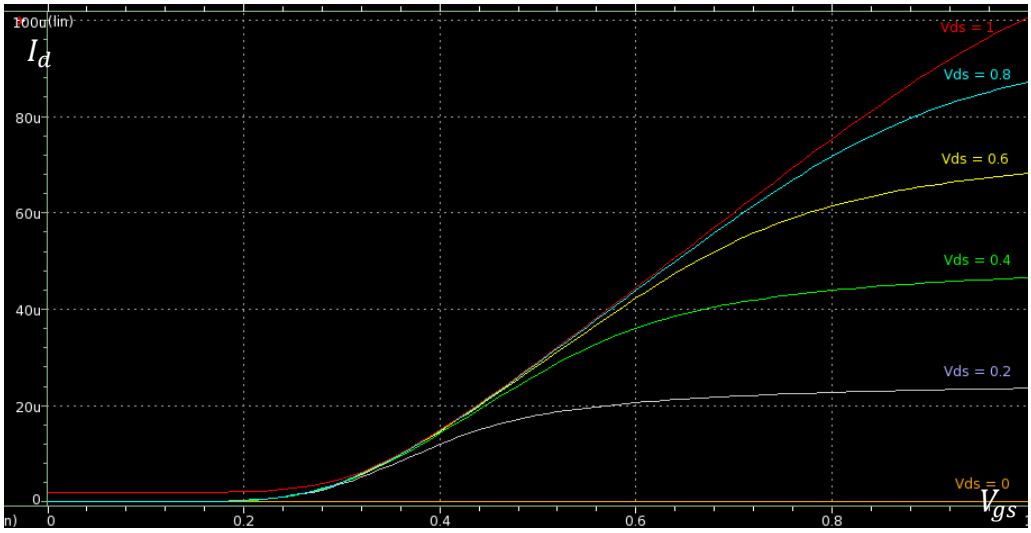

Fig. 9. $I_d - V_{gs}$ graph.

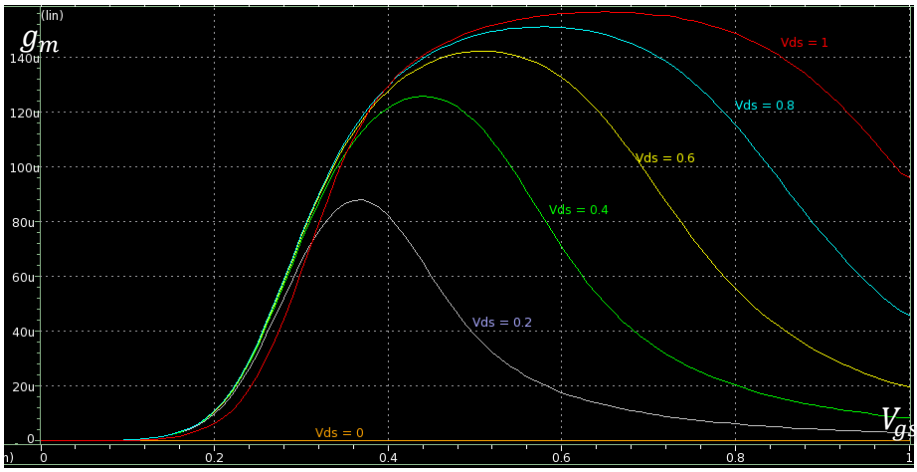

Fig. 10. $g_m - V_{gs}$ graph.

As well, the variation of I_d and g_{ds} values based on V_{ds} variation is considered and presented in Fig. 11 and Fig. 12, where each line has a different value of V_{gs}.

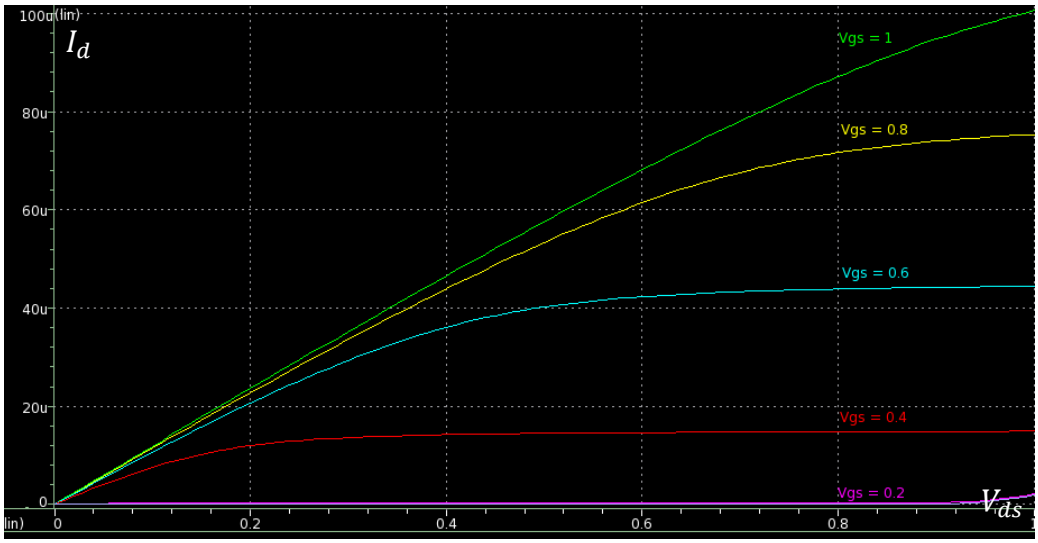

Fig. 11. $I_d - V_{ds}$ graph.

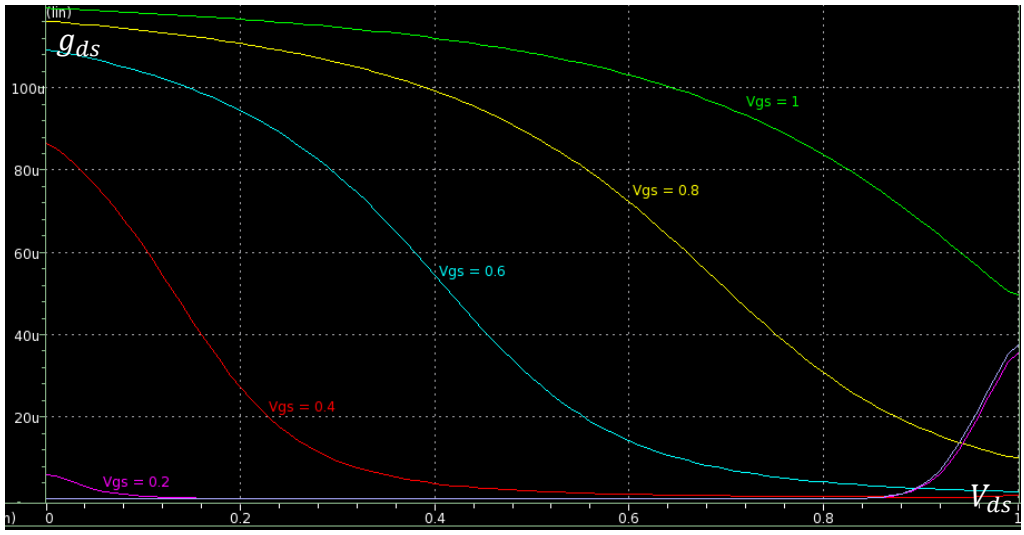

Fig. 12. $g_{ds} - V_{ds}$ graph

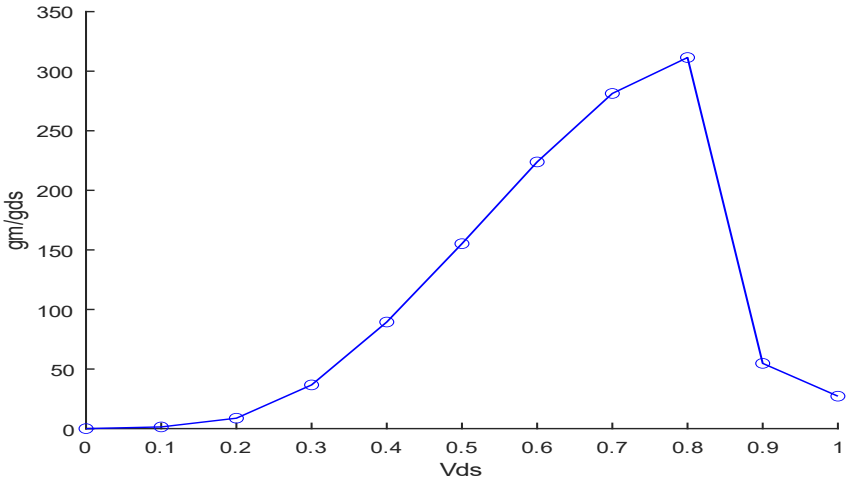

Fig. 13. $\frac{g_m}{g_{ds}}$ vs. V_{ds} graph

Based on the $I_d - V_{ds}$ curve, we can see that transistor is either in linear region or saturation region and $I_d - V_{gs}$ helps in identifying transistor threshold voltage. The $\frac{g_m}{g_{ds}}$ (i.e., transistor intrinsic gain) curve shows its strong dependence on V_{ds}. Compared to the regular CMOS transistors, CNFETs feature much larger (over 10 times) intrinsic gain.

3.2.2. P-CNFET

A P-CNFET schematic is shown in Fig. 14. In order to examine its current and transconductance variation based on drain and gate voltage, we implemented the circuit in Fig. 15 and Fig. 16 in Hspice using the Stanford University virtual source model. [11]

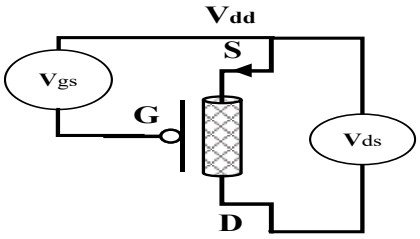

Fig. 14. P-CNFET

The variation of I_d and g_m values based on V_{gs} is examined and the results are presented below, where each line has a different value of V_{ds} ($V_{dd} = V_{ds}$).

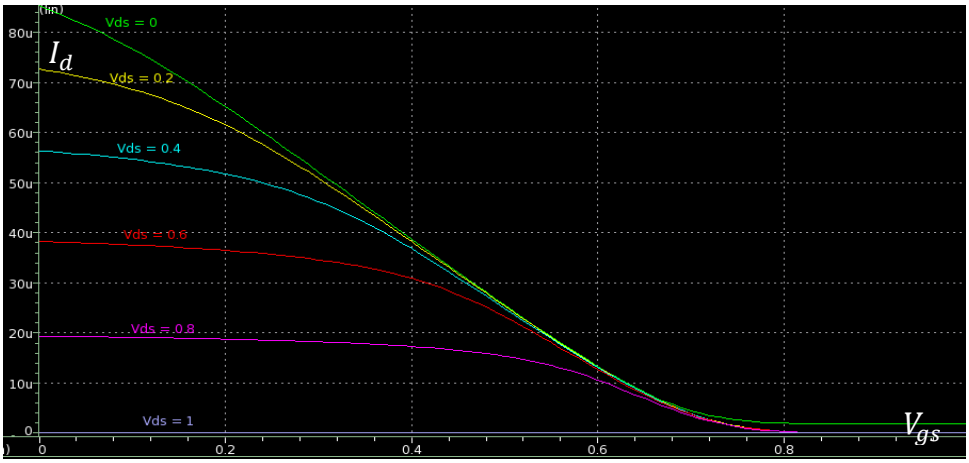

Fig. 15. $I_d - V_{gs}$ graph.

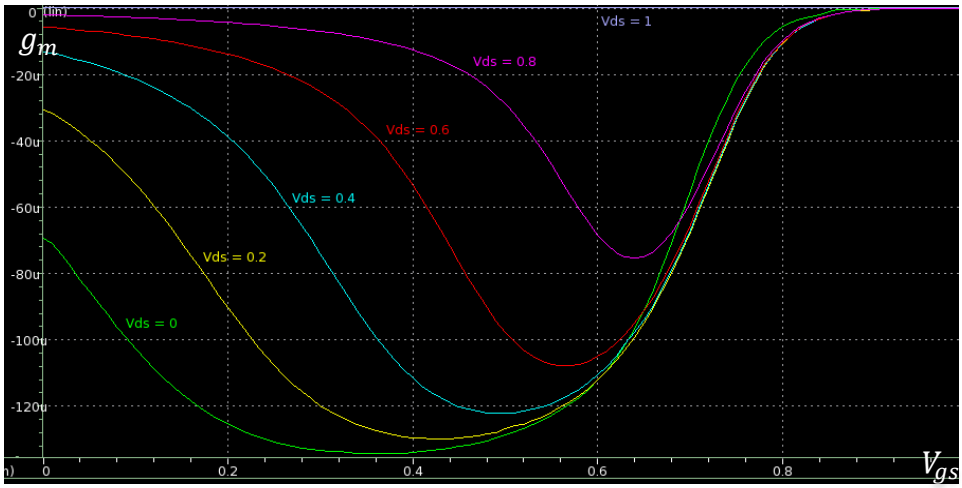

Fig. 16. $g_m - V_{gs}$ graph.

Also, the variation of I_d and g_{ds} values based on V_{ds} variation is considered and presented in Fig. 17 and Fig. 18 (each line has a different value of V_{gs}).

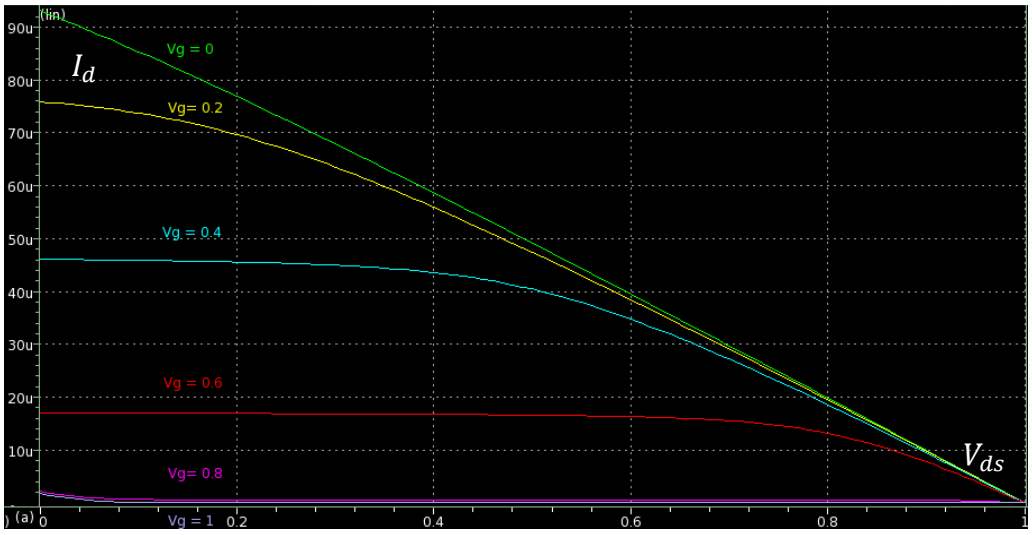

Fig. 17. $I_d - V_{ds}$ graph,

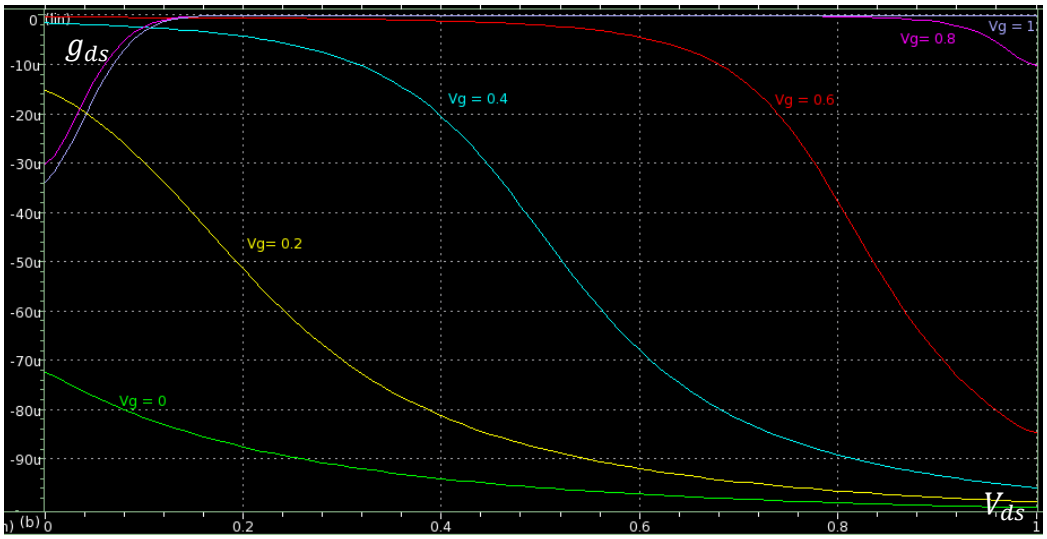

Fig. 18. $g_{ds} - V_{ds}$ graph

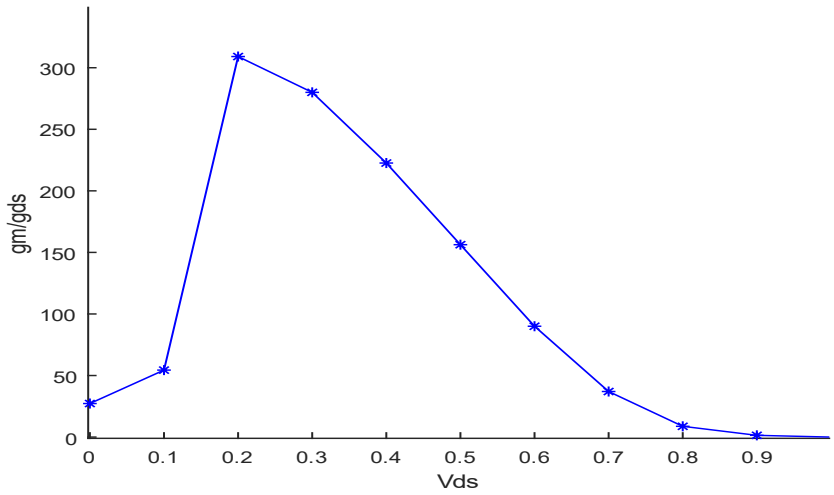

Fig. 19. $\frac{g_m}{g_{ds}}$ vs. V_{ds} graph

Based on the $I_d - V_{ds}$ curve we can see that transistor is either in linear region or saturation region and $I_d - V_{gs}$ helps in identifying transistor threshold voltage. $\frac{g_m}{g_{ds}}$ variation curve based on V_{gs} shows the much higher transistor intrinsic gain offered by the CNFETs over regular CMOS transistors.

3.3. Variation of CNFET Parameters

In the Stanford VS-CNFET compact model, the drain current depends on the CNT diameter, gate length, gate width, gate height, and oxide thickness. Also, the carrier injection mobility and velocity are functions of the CNT diameter and the gate length. We have studied the impact of the CNT diameter, gate length, and gate width on the drain current (as they are the parameters that have greater effect on the current and are controllable by the designer). Two amplifiers using N-CNFET, P-CNFET, and a differential pair as shown in Fig. 20 and Fig. 21 are considered for this analysis.

3.3.1. Parameter Variation of CNFET-based Circuits

An N-CNFET, and P-CNFET based amplifier are shown in Fig. 20.

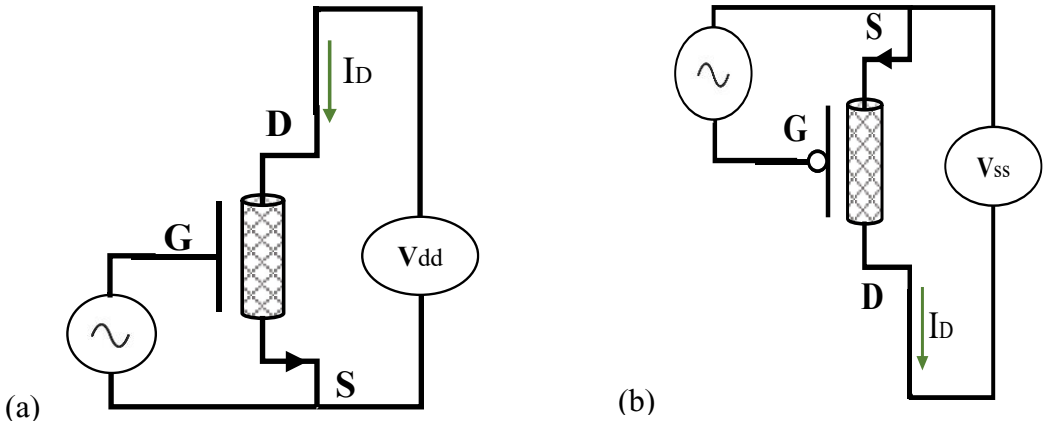

Fig. 20. Circuits: (a) N-CNFET amplifier, (b) P-CNFET amplifier.

The nominal parameter values of the transistor are given in Table 2.

Table 2 Nominal parameter values

Diameter (D)	Gate Length (L_g)	Gate Width (W_g)	Pitch (S)
1.49 nm	16 nm	30 nm	4 nm

3.3.2. Monte Carlo Simulation

By using transmission electron microscopy (TEM) and optical response analysis, the CNTs diameter was examined to follow a Gaussian distribution [30]. Also, gate width and length have been shown to have a Gaussian distribution variation as shown in Table 3. The VS-CNFET Stanford university model does not take into account variability, and only permits transistors with semiconducting CNTs of the same diameter, and a single-spacing value between all CNT pairs in a CNFET. To demonstrate fluctuations in a single process parameter impact on circuit performance, we employed HSPICE Monte Carlo analysis as a method to evaluate variability.

Table 3. Probability distributions of CNT specific variations

CNT specific variations	**Distribution (10%variation)**
Diameter variations	Gaussian: $\mu = 1.49\text{nm}, \sigma = 0.14\text{nm}$
Gate Length variations	Gaussian: $\mu = 16\text{nm}, \sigma = 1.6\text{nm}$
Gate Width variations	Gaussian: $\mu = 30\text{nm}, \sigma = 3\text{nm}$

For each simulation set with N=1000 Monte Carlo trials, we assume a single parameter to change in a Gaussian distribution with μ as the mean and σ as the standard deviation while other parameters are fixed. To indicate the dispersal of performance, we use the

coefficient of variation (CV= (σ/μ)) for the performance outcomes of the Monte Carlo simulation. A greater value of CV shows a stronger influence of parameter variation. We considered the circuits in Fig. 20 with the nominal parameters shown in Table 2 and observed its gain variation based on CNFET parameter variation. The results of the Monte Carlo simulations are shown in Table 4.

Table 4. Monte Carlo result of N-CNFET circuit and P-CNFET circuit

Parameter variation	Diameter	Gate Length	Gate Width
Variation of I_d (μA)	[274.11 – 444.77]	[396.64 – 404.61]	[352.24 – 447.77]
CV= (σ/μ)	0.0276	0.0014	0.0053

According to Table 4 the variations in CNT diameter have the maximum impact, while the variations in the gate width and gate length have less impact on the circuit gain. The change due to W_g variation is because the I_{ON} is affected by variations in W_g only if the variation is large enough to affect CNT count in the device. (Here the gate width changes from 27 to 33 causing the change in the existing number of CNTs.)

3.4. Case Study

In this section we study the effect of CNFET parameter variation on differential pair circuit shown in Fig. 21. The nominal values of all the transistors are the same as Table 2.

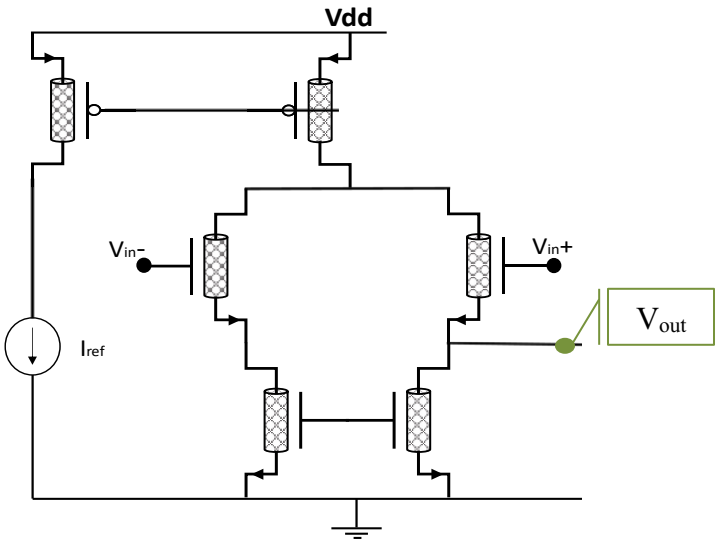

Fig. 21 . Differential pair circuit

3.4.1. CNT Diameter

First, we study the effect of diameter variation on the gain and -3dB bandwidth of the circuit. The results are shown in Fig. 22.

As shown in Fig. 22, the bandwidth increases with the increase of the nanotube diameter. This is because as the CNT diameter is increased, the transconductance goes up. Also, the DC gain decreases with the diameter increase because the reduction in the output resistance with diameter is more than the rise in its transconductance. This variation occurred because CNFET electrical behavior and circuit performance directly depend on the CNT diameter.

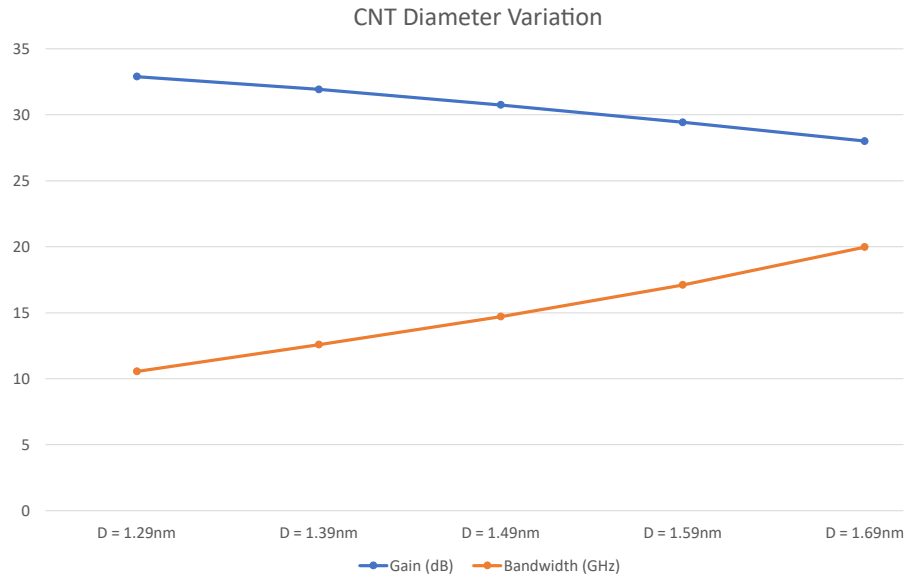

Fig. 22 . Variation of gain and bandwidth based on diameter variations

3.4.2. Gate Length

Now, we study the effect of gate length variation on the gain and bandwidth of the circuit. As the gate length increases, by considering a suitable gate voltage, ballistic transport in CNFET and the fact that we do not consider parasitic effects in our simulations, the current would not decrease. So, we can see that increasing the gate length increases the gain to a small degree and decreases the bandwidth. Gain and bandwidth variations based on gate length are shown in Fig. 23.

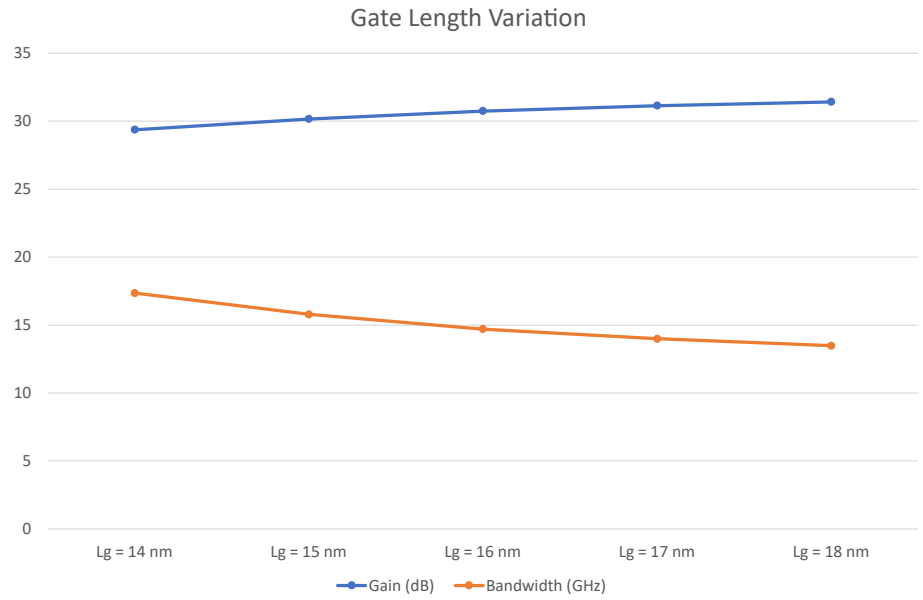

Fig. 23. Variation of gain and bandwidth based on gate length variation

3.4.3. Gate Width

In this section, we study the effect of gate width variation on the gain and bandwidth of the circuit. The results are shown in Fig. 24.

In an ideal CNFET the CNTs are uniformly spaced and the number of CNTs in a CNFET is gate width (W_g) divided by pitch (S). Therefore, as W_g and S change, the number of CNT changes in steps, so the CNFET current changes in steps as well. Therefore, slight variations in W_g and S do not have a considerable impact on gain unless the nominal values lie at or near a step edge. However, increasing W_g will decrease the bandwidth to a slight degree.

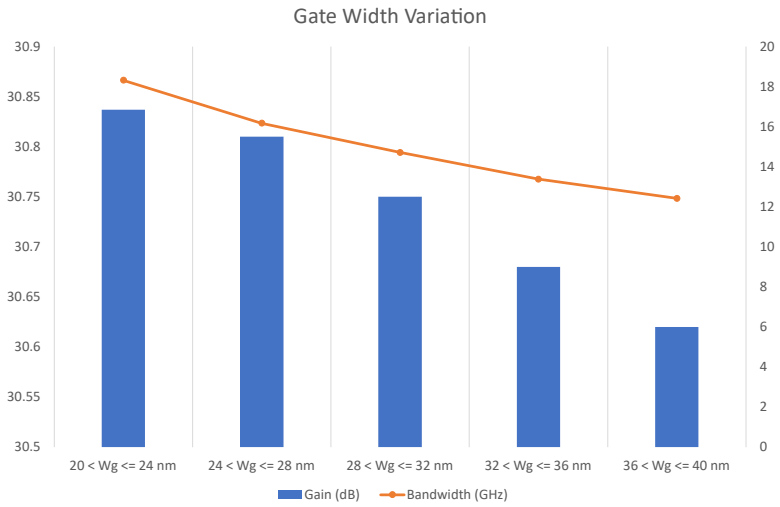

Fig. 24. Variation of gain and bandwidth based on gate width variation

3.4.4. Monte Carlo Simulation

Here we consider the circuit in Fig. 21 and observed the effectiveness of each CNFET parameter variation on its gain.

Table 5. Monte Carlo result of differential pair circuit

Parameter variation	Diameter	Gate Length	Gate Width
Variation of gain	[28.21 – 31.13]	[28.64 – 30.23]	[29.46 – 29.89]
CV= (σ/μ)	0.01589	0.00657	0.000071

It is observed that the variations in CNT diameter have the maximum impact, while the variations in the gate width and gate length have less impact.

3.5. Flow of Methodology

Our optimization flow is illustrated in Fig. 25. We first pass the CNT process variation information, circuit specifications and design parameters to our proposed sizing engine. The sizing engine includes a single-objective optimization process using our modified GBC algorithm as the first optimization phase. By using this design centering method, we can ensure the design robustness to get the optimum points for each objective and desired functionality. This is followed by a multi-objective optimization process as the second optimization phase using NBI to get the Pareto front and achieve fully deterministic multi-objective optimization methodology. For solving NBI, we take advantage of Multi JuMP library based on Julia language [46].

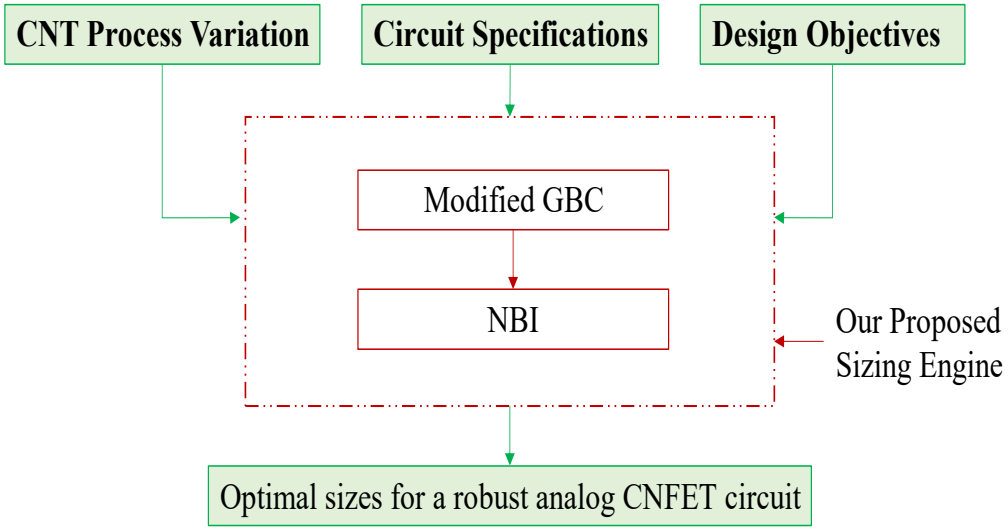

Fig. 25. Our proposed modified GBC+NBI method

3.6. Summary

CNFETs are prone to CNFET manufacturing imperfection that causes variations in CNT parameters affecting their performance. In this thesis we consider the following variations of CNT: 1) CNT diameter variations, 2) CNFET gate length variations, and 3) CNFET gate width variations.

In this chapter, we presented a detailed variability analysis of CNT specific variations affecting CNFET devices and the exemplary circuit of one single-stage amplifier, and then described the flow of our proposed optimization methodology. In the next chapter, we will discuss the first optimization phase of our proposed methodology, a modification of the GBC method [38].

Chapter 4

A Modified Generalized Boundary Curve Method

4.1. Introduction

In our work, we use design centering to achieve a robust circuit design and the best yield functionality while having a safety margin in our parameter distribution. The worst-case distance is considered as a robustness variable in design centering, and the goal is to maximize it while the linearization limit does not exceed the size variation.

In this chapter, the GBC method presented in [38] is modified to solve the sizing optimization problem without having to ensure any prior knowledge of the initial estimate and step-size change. Thus, we can get the optimized circuit sizes that can tolerate the process variation of CNT parameters in the analog CNFET circuits.

4.2. Bézier Curves

Bézier curves are parametric curves, which use Bernstein polynomials as their foundation. They are often used to produce smooth curves in various engineering applications [47]. For instance, they are used in computer aided design (CAD), vector-based graphic (SVG), computer aided manufacturing (CAM), and advanced animations.

We use n control points $(P_0, \ldots, P_{n-1})$ to define the Bézier curve :

$$P(t) = \sum_{i=0}^{n-1} (\mathrm{B}_{\mathrm{i}}^{\mathrm{n}}(t) \cdot P_i), \qquad t \in [0,1] \tag{1}$$

where $\mathrm{B}_{\mathrm{i}}^{\mathrm{n}}(t)$ is the Bernstein polynomial:

$$\mathrm{B}_{\mathrm{i}}^{\mathrm{n}}(t) = \binom{n}{i} t^i (1-t)^{n-i}, \quad \binom{n}{i} = \frac{n!}{i!\,(n-i)!} \tag{2}$$

The main characteristics of the Bézier curve are as follows: for any number of the control points, the curve starts at the first control point and ends at the last control point, and all the points are not located on curve but the curve is located inside the convex hull of the control points. The curve order is equal to the number of the control points minus one. For two points, we will have a straight line, while for three points, we will have a quadratic curve. By choosing four points, the resulting curve is cubic as illustrated in Fig. 26 [48]. In addition, all the derivatives can be determined analytically.

Bézier curve was initially used to effectively calculate the curvature and curvature radius of the characteristic boundary curve (CBC) [49]. They tried to find a point with a good ratio between correction effect and correction effort, which is located near the bend of the curve. The smaller the radius gets, the higher the curvature becomes. Hence, they considered the CBC solution as a point with the smallest curvature radius.

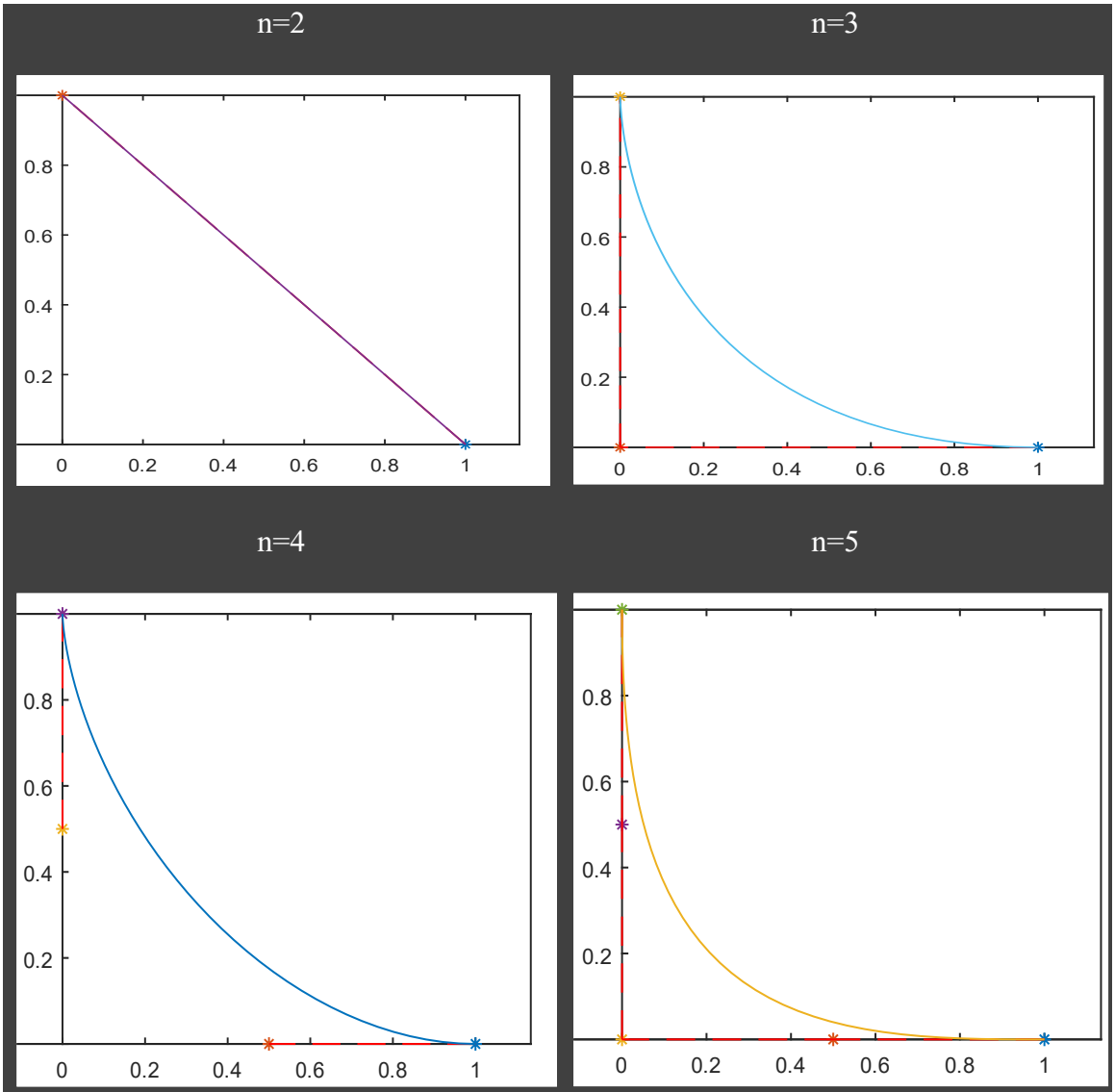

Fig. 26. Various Bézier curves

4.3. Single-Objective Optimization

In the single-objective optimization, an objective function is minimized, while satisfying a set of inequality and equality constraints. The mathematical single-objective optimization formulation is defined as follows:

$$\underset{x}{Min}:\ F(x), \qquad s.t. \quad h(x) \geq 0, g(x) = 0 \tag{3}$$

where F is the objective function, while h and g are vectors of inequality and equality constraints respectively, and x is the design vector with n design variables.

4.3.1. Design Centering

A major difficulty that the designers face is to develop systems or models that perform reliably in the face of uncertainty and environmental variations. This issue is addressed in the design centering, which is described as finding a design that meets specified requirements with a significant probability of continuing to do so even if parameters or specifications alter randomly. Design optimization seeks to identify the design that best satisfies the specifications, whereas design centering seeks to find the design that most robustly meets the specifications. We use the design centering concept to determine the nominal values of the CNT parameters in a way that the circuit not only tries to fulfill its specifications but also is robust against manufacturing variations. The design centering solution refers to a restricted part of the Pareto front of objective functions, which is determined by the circuit specifications and robustness.

4.3.2. Deterministic Circuit Sizing Using GBC

Deterministic optimization is a subset of numerical optimization that converges to the global optimum in finite time and focuses on finding global solutions to optimization problems. It assures that the reported answer is the global solution within a certain tolerance. In order to optimize the circuit performance and achieve a robust circuit design under

variation, we first consider initial device sizes S_n (i.e., d_{cnt}, W_g, L_g) and a performance vector $f = [f_1, f_2, \ldots, f_n]$ of the output quantities of circuit simulation (i.e., performance features) such as gain, phase margin, slew rate, and power supply. Variable f_i refers to the value of the respective performance of the circuit. Then we linearize each performance regarding S_n and find the worst-case process corner to calculate the minimum deviation of S_n which transfers the performance from worst-case to our specified corner also known as parameter distance:

$$\delta_i(x) = \delta_i(s_n) + \frac{g_{s,i}^t}{\|g_{s,i}\|}, \tag{4}$$

where δ_i is the worst-case parameter distance, s_n shows the initial device sizes, and $g_{s,i}$ represents the performance gradient.

Now to have a robust sizing in the worst-case process corner, we need to maximize the worst-case parameter distance ($\delta_i(x)$) while the size change $x = s_{n+1} - s_n$ is in the linear approximation limit. Maximizing $\delta_i(x)$ matches minimizing the following function using GBC method:

$$min \ (\sum_i exp\left(-\alpha \cdot \delta_i(x)\right))^2 + \lambda \cdot \|x\|^2, \quad \lambda \geq 0 \tag{5}$$

where λ determines the weight of x and α is a scaling factor.

This method is based on circuit performance gradients and searches the solution space in a specified route accordingly and arrives at a unique set of circuit sizes using linearized approximations.

4.4. Modified GBC

In the GBC method, the suggested cost function simultaneously considers both the parameter correction norm and the linearization error reduction [50]. It uses the nonlinear cost function, which is calculated by linearized objectives. Therefore, in this method, the error of linearization for the nonlinear cost function is minimum. The system is given by:

$$\delta_i(x) = \delta_i(s_n) + \frac{g_{s,i}^t}{\|g_{s,i}\|}, \tag{6}$$

$$\bar{\varphi}(x) = \sum_{i=0}^{n_\beta} exp\left(-\alpha \cdot \bar{\delta}_i(x)\right), \tag{7}$$

$$\bar{\Phi}(x) = \bar{\varphi}^2(x) + \lambda \cdot \|x\|^2 \qquad \lambda \geq 0, \tag{8}$$

where δ_i is the worst-case parameter distance, $\bar{\varphi}(x)$ is the linearized cost function, $\bar{\Phi}(x)$ is modified version of $\bar{\varphi}(x)$ (to balance the error reduction and the norm of the parameter correction), s_n shows the initial device sizes, $g_{s,i}$ represents the performance gradient, $\bar{\varphi}(x)$ is the approximated objective function, α is a positive constant for scaling purpose, $\bar{\Phi}(x)$ is the GBC function and λ controls the weight of the norm of x. Then we need to find the step size that minimizes (8) so that the optimization functions are:

$$x_c(\lambda) = \min \bar{\Phi}(x), \tag{9}$$

$$\bar{\varphi}_c(\lambda) = \bar{\varphi}\left(x_c(\lambda)\right). \tag{10}$$

To solve the optimization problem above, both $x_c(\lambda)$ and $\bar{\varphi}_c(\lambda)$ are transformed to have the solution in the [0, 1] interval:

$$a_c(\lambda) = \frac{\|x_c(\lambda)\| - \|x_c(\lambda \to \infty)\|}{\|x_c(0)\| - \|x_c(\lambda \to \infty)\|}. \tag{11}$$

$$r_c(\lambda) = \frac{\bar{\varphi}_c(\lambda) - \bar{\varphi}_c(\lambda \to \infty)}{\bar{\varphi}_c(0) - \bar{\varphi}_c(\lambda \to \infty)}. \tag{12}$$

The GBC curve is shown in Fig. 27, and λ varies from 0 to ∞; the optimum λ is in the shaded area. The slope of the GBC is given as follows:

$$m_c = -\lambda \cdot \frac{\|x_c(\lambda)\|}{\bar{\varphi}_c(\lambda)} \cdot \frac{\|x_c(0)\| - \|x_c(\lambda \to \infty)\|}{\bar{\varphi}_c(0) - \bar{\varphi}_c(\lambda \to \infty)}. \tag{13}$$

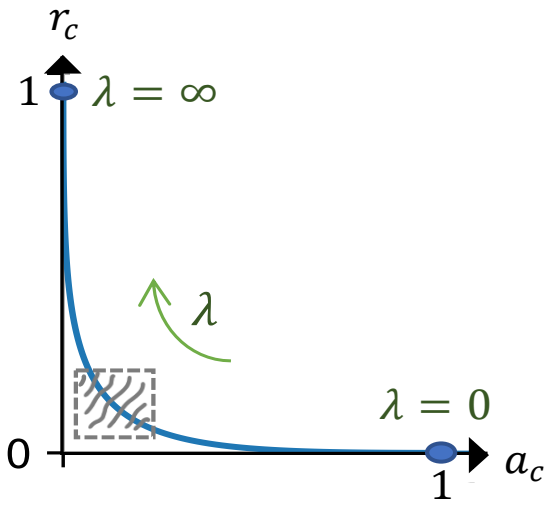

Fig. 27. GBC curve

We now discuss a detailed method of estimating size change, which will be the core part of our modified GBC method. Since the work in [38] aims to solve the GBC for layout migration problems, to solve (8) they consider a x_{max} as the right boundary and define x_{est} as x_{max}/m (m is an integer). However, in other application scenarios it may not be possible to suggest any value for those variables. To overcome this challenge, we must develop a general scheme to derive x_{est} and estimate the range of x.

In this regard, we propose to use a Bézier curve to estimate the curve as shown in Fig. 20. The Bézier curve equation can be defined as follows:

$$B(t) = \sum_{i=0}^{n-1} \binom{n}{i} (1-t)^{n-i} \cdot t^i \cdot P_i\,, \qquad 0 \le t \le 1 \tag{14}$$

where the Bézier curve is a function of t, n is the number of the control points we have, and P_i is the coordinates of the control points.

To have a better estimation, we opt to use 5 points to build the curve: [P_0 = [1, 0], P_1 = [0.25, 0], P_2 = [0, 0], P_3 = [0, 0.25], P_4 = [0, 1]]. The equation in this case will become:

$$\begin{aligned} B(t) &= (1-t)^4 \cdot P_0 + 4 \cdot t^1 \cdot (1-t)^3 \cdot P_1 \\ &+6 \cdot t^2 \cdot (1-t)^2 \cdot P_2 + 4 \cdot t^3 \cdot (1-t)^1 \cdot P_3 + t^4 \cdot P_4 \end{aligned} \tag{15}$$

As mentioned before, the best λ is in the shaded area. So for its estimation, we can consider a situation that (15) has the same a_c and r_c $\big((a_c, r_c) = (t, t)\big)$, then we have $x_c(\lambda \to \infty) = \min \bar{\varphi}^2(0)$ and $x_c(0) = \min \bar{\varphi}^2(x)$. Based on the results of $x_c(0)$ and $x_c(\lambda \to \infty)$, we can get the value of $\bar{\varphi}_c(0)$ and $\bar{\varphi}_c(\lambda \to \infty)$. Now based on (11) and (12), we can easily get the value for $\|x_c(\lambda)\|$ and $\bar{\varphi}_c(\lambda)$.

Then at point (t, t), we can calculate the slope of our Bézier curve based on (15). Since we use this Bézier curve to approximate the GBC curve, the slope value can also be expressed by using (13) and we have already calculated every parameter in (13) except λ. By solving this slope equation enclosing (13), we can then get an estimation of λ. In this way we can get our estimated x_{est} from

$$\Sigma\, exp^2\big(-\alpha \cdot \delta_i(x_{est})\big) = \lambda_{est} \cdot \|x_{est}\|^2 \tag{16}$$

The process of the modification is illustrated in Algorithm 1.

Algorithm 1: Modified section of GBC	
	Input: $x_c(\lambda \to \infty), \bar{\varphi}_c(\lambda \to \infty), x_c(0), \bar{\varphi}_c(0)$
	Output: *value of* x_{est}*// optimal value*
1	$B(t) = \sum_{i=0}^{n} \binom{n}{i} (1-t)^{n-i} \cdot t^i \cdot P_i$ *// form a Bézier curve for GBC*
2	*Get the point that* $(a_c, r_c) = (t, t)$
3	*Calculate* $\|x_c(\lambda)\|$ *and* $\bar{\varphi}_c(\lambda)$ *for (t, t) point*
4	*Calculate Bézier curve slope at (t, t) to get* m_k
5	*Solve Eq. (13) to get* λ_k*:* $m_k = m_c$
6	$\Sigma\, exp^2\left(-\alpha \cdot \delta_k(x_{est})\right) = \lambda_k \cdot \|x_{est}\|^2$ *// to get* x_{est}
7	$x_{init\_k} = x_{est}$
8	$\boldsymbol{min}\ \Sigma\, exp^2\left(-\alpha \cdot \delta_k(x_k)\right) + \lambda_k \cdot \|x_k\|^2$ *// to get* x_k
9	$x_{max} = n \cdot x_{est}$ //to get x_{max}
10	$right = x_{max}$, $left = 0$

The rest of GBC solving follows the method in [38] to find the optimal λ value, which is in the shaded area of Fig. 27. The expression in (5) is composed of an exponential term, which is the dominant section of our minimization task when $\lambda = 0$, and an x term which determines the minimum when $\lambda = \infty$. These two dominating conditions are identified on the boundary curve as points (1, 0) and (0, 1). So, solving the GBC is equal to finding an optimal λ value so that neither of the x term nor the exponential term dominates the cost function. The λ_{opt} determines a specific x_{opt}, which will minimize (5).

In order to make sure none of the x term and the exponential term dominates each other, we need to tune our estimation of x and the value of λ so that if x is greater than x_k, the value of λ would decrease, while if x is smaller than x_k, the value of λ would increase. This process will continue until the gap between x and x_k is smaller than x_{th}. While updating x, another important variable that we consider is the distance of x and λ from the origin ($dist_k$), the closer they are to the origin, the more accurate the result becomes.

Therefore, if $dist_{k+1}$ is smaller than $dist_k$, we will update the x value. The GBC function is shown in Algorithm 2.

Algorithm 2: GBC function

Input: $x_k,\ \lambda_k, right, left$

Output: *value of* x_{opt} *// optimal value*

function *GBC*$(x_k,\ \lambda_k)$

  $A = (x_k, \lambda_k);\ B = (0,0);\ dist_k = norm(A, B);$ *// $dist_k$ is the distance from origin.*

  while *(True)*

    if $(|x_k - x_{init\_k}| < ||x_{th}||)$

      break*;*

    elseif $(x_k < x_{init\_k})$

      $left = x_{init\_k};$

    else

      $right = x_{init\_k};$

    end;

    $x_{init\_k+1} = (left + right)/2;$

    min $\Sigma exp^2\left(-\alpha \cdot \delta_{k+1}\left(x_{init\_k+1}\right)\right) = \lambda_{k+1} \cdot \left\|x_{init\_k+1}\right\|^2$ *//to get λ_{k+1}*

    min $\Sigma exp^2\left(-\alpha \cdot \delta_{k+1}(x_{k+1})\right) + \lambda_{k+1} \cdot \|x_{k+1}\|^2$ *// to get x_{k+1}*

    $A = (x_{k+1}, \lambda_{k+1});\ B = (0,0);\ dist_{k+1} = norm(A, B);$

    if *($dist_{k+1} < dist_k$)*

      $x_{opt} = x_{k+1};\ \lambda_{opt} = \lambda_{k+1};\ dist_{opt} = dist_{k+1};$

    end;

  end;

  return $x_{opt};$

end function;

Here we compare some random x_{est} results with our modified results.

Table 6. λ_{opt} value based on different x_{est}

	x_{est_MGBC} (Our modified GBC x_{est})	$x_{est} \gg x_{est\_MGBC}$	$x_{est} \ll x_{est\_MGBC}$	x_{est} used as initial value in [38]
λ_{opt}	0.0031309	0.0001803	0.0092750	0.0031310

As discussed earlier choosing λ_{opt} is so important for making sure that neither the x term nor the exponential term dominates the cost function. So, based on Table 6 we can see that choosing wrong estimation can lead to wrong answer, however, it can be solved by increasing the iterations which increase optimization timing and accuracy is not guaranteed.

4.5. Summary

In this chapter, we proposed a new modification to the GBC method. In the next chapter we extend our single-objective method to multi-objective optimization by applying the GBC results as individual optima to multi-objective variation-aware optimization with the deterministic normal boundary intersection (NBI) method to achieve an estimation of the Pareto front.

Chapter 5

Multi-Objective Variation-Aware Analog CNFET Sizing Methodology

5.1. Introduction

A wide range of challenges in circuit design optimization entail targeting various performance criteria; yet it is highly unlikely that these opposing objectives would be met by the same design. As a result, while deciding on the final design of a circuit, the designer must make some compromises between the competing objectives. Multi-objective optimization is used to optimize more than one objective function and the answer is a set of solutions that define the best trade-off between competing objectives. In this chapter we will first explain the multi-objective optimization. And then we will discuss the NBI method and how we use it to develop our fully deterministic multi-objective method that considers process variation to give a robust solution.

5.2. Multi-objective Optimization

The mathematical formulation of the multi-objective optimization can be expressed as follows:

$$\underset{x \in C}{Min:}\ F(x) = \begin{bmatrix} f_1(x) \\ . \\ . \\ . \\ f_n(x) \end{bmatrix}, \quad s.t.\ \ h(x) \geq 0, g(x) = 0 \tag{1}$$

where $F(x)$ is a vector with n objective functions labelled as $f_1(x), \dots, f_n(x)$, x is a vector with n optimization parameters, and $C \in \mathbb{R}^n$ is the solution space of F.

The multi-objective optimization is used to get the points that are not dominated by any other points within the F domain [51]. A vector $a = [a_1, a_2, \dots, a_n]$ dominates a vector $b = [b_1, b_2, \dots, b_n]$ if:

$$\forall\ (a_i \leq b_i)\ \wedge\ \exists\ (a_i < b_i), \qquad i \in \{1,2, \dots, n\} \tag{2}$$

The set of non-dominated points are the Pareto optima (F^*). Since no single x would generally minimize every f_i simultaneously, a concept of optimality is defined as Pareto optimality. Fig. 28 shows the feasible region for two objectives as an example and indicates that point $F^{\#}$ is dominated by all the points located on the arc between F_1^* and F_2^*. Here ∂F is the domain of F and $\partial F^{\prec}$ is the feasible boundary between F_1^* and F_2^* called Pareto optimal front where all the points on the arc are non-dominated.

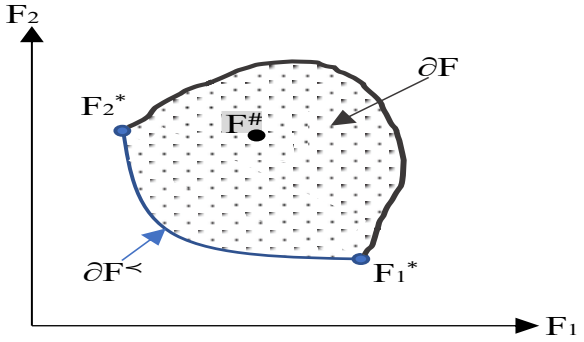

Fig. 28. Feasible region and Pareto front.

It is not always possible to optimize all the circuit performances simultaneously with a unique set of parameter sizes. Thus, there could be a trade-off between performance improvements such that boosting one performance might affect another. In this situation, we take advantage of the Pareto optimality concept in our optimization [52]. This Pareto optimum boundary is crucial to the designer because it shows different optimization capability or potential of a circuit while considering the trade-offs involved in the multi-objective optimization. The main goal of the multi-objective optimization is to get the Pareto front as accurately as possible.

There are different stochastic and deterministic approaches to solve the multi-objective optimization problem [53]. In the stochastic methods, the global optimal is not guaranteed; in addition, we need extra data for learning and testing. On the other hand, in the deterministic optimization, the multi-objective optimization problem is transformed into a single-objective equation solving process that will proceed repeatedly, and thus achieving the global minimum can be guaranteed for each specific run.

In this work, we are motivated to develop a genuine deterministic multi-objective sizing method for analog CNFET circuits. We first pass the CNFET library, CNT process variation information, design parameters, and circuit specifications to our proposed sizing engine. This sizing engine is made up of two blocks: the GBC block generates multiple single-objective optimizations as initial optima while the NBI block tackles the comprehensive multi-objective optimization challenges all together. The optimal CNFET sizes will be output as the final results, which will contribute to the Pareto front in terms of variation-aware optimum analog CNFET circuit performances.

5.3. NBI Method

The NBI method has been applied in different disciplines for multi-objective optimization [54]. The most critical topics in the NBI method are provided as follows.

5.3.1. Individual Optimum f_i^*

The optimum f_i^* points are the global minima for each objective, which are generated from our modified GBC method in this work. They play essential roles in the NBI method because they form the boundary points; the non-dominated points of the Pareto front are located between those individual minima.

5.3.2. Convex Hull of Individual Minima

Assume x_i^* as the respective global minimizers of $f_n(x), i = 0, 1, \dots, n-1$ and $x \in C$. Let ϕ be an $n \times n$ matrix known as pay-off matrix with each column containing $f_i^* - f^*$ value. Then the set of points in R^{n} that are convex combinations of multiple $f_i^* - f^*$ ($\{\phi \cdot \beta : \beta \in R^{\mathrm{n}},\ \beta_i \geq 0,\ \Sigma\beta_i = 1\}$) is referred to as the Convex Hull of Individual Minima (CHIM). To achieve the Pareto front points, NBI will start the search from points on CHIM.

5.3.3. Quasi-normal Vector to Convex Hull:

The search along a group of normal vectors intersecting the boundary is valid even if we do not have the exact normal direction toward the CHIM. We can use quasi-normal direction $\hat{n}$ pointing toward the origin with negative components. We consider the quasi-

normal component to be an equally weighted linear combination of the columns of $\phi \cdot e$ where $e = [-1, -1, \dots, -1]^t$ (-1 is used to ensure that it points toward the origin).

The idea behind this method is motivated by the fact that the intersection point between the boundary ∂F and the normal pointing towards the origin starting from any point in the CHIM is a point on the portion of ∂F containing the efficient points. This point is also a Pareto optimal point. The NBI formulation is

$$\begin{bmatrix} x^* \\ t^* \end{bmatrix} = \underset{[x\ t]}{argmax}\ t, \quad s.t. \ \ \phi \cdot \beta + t \cdot \hat{n} = F - F^* \tag{3}$$

$$\beta_i \geq 0, \quad \Sigma \beta_i = 1 .$$

In this formulation, F is the vector of objective functions, F^* is a vector of optimum values for each objective, t is the distance of Pareto points from CHIM, $\phi = [F_i^* - F^*]$, β is vectors of barycentric coordinates, and $\phi \cdot \beta$ represents a point in the CHIM.

The NBI method solves such a problem by dividing the multi-objective optimization into several single-objective optimizations while adding an equality constraint for each objective (3). The cost parameter t affects all the performance constraints. The vector constraint $\phi \cdot \beta + t \cdot \hat{n} = F - F^*$ ensures that point x is actually mapped by F to a point on the normal while the remaining constraints ensure feasibility of x with respect to the original problem.

To solve (3), we can achieve β by considering a variable as step size δ and the number of objectives so that $\beta = [\beta_1, \dots, \beta_n]$ can be defined. Fig. 29 illustrates the tree of values for n = 2 and δ = 0.2. Then we divide (3) to sub problems called NBI_β. Since the specific minima of the functions are already available from the first optimization phase conducted

by the modified GBC method, we can start at x_i^* and solve a nearby subproblem, then at another near the one we just solved, and so on. Then by solving the equation for different β, the Pareto front can be achieved.

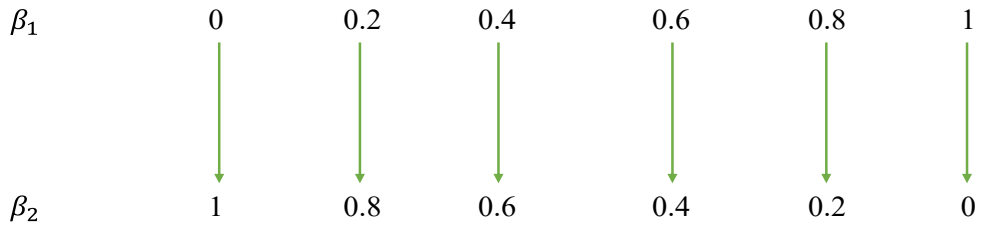

Fig. 29. Generating β for n = 2 and δ = 0.2

Since we aim to maximize t in the NBI subproblem, this means that our maximization subproblem tries to find a feasible point x as far from a chosen point $\phi \cdot \beta$ as possible. The NBI process of optimization is illustrated in Fig. 30. F_1^* and F_2^* are two global minima generated by our modified GBC method for individual single-objective problems. Based on the CHIM, the NBI method is to identify the Pareto front printed in blue as shown in Fig. 30.

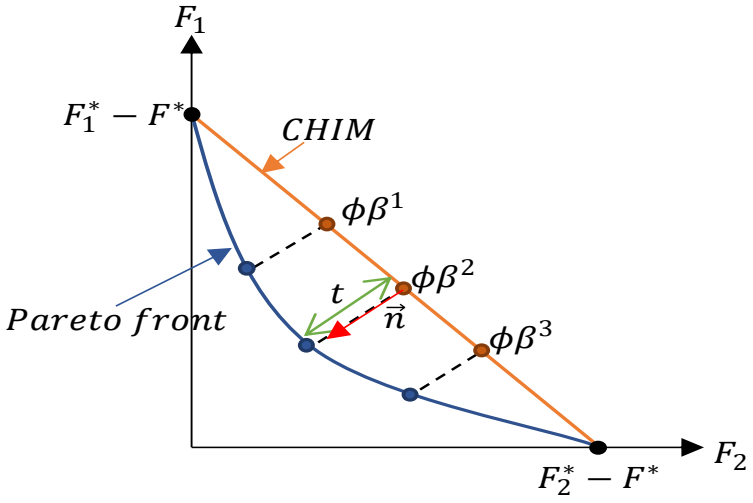

Fig. 30. NBI method.

As an example, we can consider minimizing:

$$f_1 = x_1^2 + x_2^2 + x_3^2$$

$$f_2 = 3x_1 + 2x_2 + x_3$$

$$s.t. \quad 2x_1 + x_2 + 3x_3 = 7$$

$$x_1^2 + x_2^2 + x_3^2 \leq 9$$

We consider f_1^* and f_2^* as individual minima. So, our subproblems can be:

1. $\begin{bmatrix} 0 & f_1^* - f_2^* \\ f_2^* - f_1^* & 0 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix} + t \cdot \begin{bmatrix} f_2^* - f_1^* \\ f_1^* - f_2^* \end{bmatrix} = \begin{bmatrix} f_1 - f_1^* \\ f_2 - f_2^* \end{bmatrix}$

2. $\begin{bmatrix} 0 & f_1^* - f_2^* \\ f_2^* - f_1^* & 0 \end{bmatrix} \cdot \begin{bmatrix} 0.2 \\ 0.8 \end{bmatrix} + t \cdot \begin{bmatrix} f_2^* - f_1^* \\ f_1^* - f_2^* \end{bmatrix} = \begin{bmatrix} f_1 - f_1^* \\ f_2 - f_2^* \end{bmatrix}$

3. $\begin{bmatrix} 0 & f_1^* - f_2^* \\ f_2^* - f_1^* & 0 \end{bmatrix} \cdot \begin{bmatrix} 0.4 \\ 0.6 \end{bmatrix} + t \cdot \begin{bmatrix} f_2^* - f_1^* \\ f_1^* - f_2^* \end{bmatrix} = \begin{bmatrix} f_1 - f_1^* \\ f_2 - f_2^* \end{bmatrix}$

⋮ ⋮

Now in order to develop our variation-aware multi-objective optimization we will consider simultaneously solving the GBC functions for each objective and then applying NBI method to develop our multi-objective optimization method. Algorithm 3 explains how our proposed NBI method works along with the input from the GBC method. First, the optima points and their corresponding x (generated by the modified GBC method) will be passed to function *NBI*(). Subsequently, an iterative process starts to form individual NBI_β subproblems and continues to solve (3) until it reaches the entire Pareto front.

Algorithm 3: NBI

Input: $F(x)$, x_n^* // $f_n(x_n^*) = f_n^*$
// at x_n^* , f_n *is the global optimum of each objective*
Output: *Pareto front*

```
function NBI(F(x), x_n^*)
    for i 1:n
        for j 1:n
            φ_ij = f_i(x_j^*) − f_i(x_i^*)
        End for;
    End for;
    z = 1 // Generate z weights, {1, ..., N_β}
    while (z < N_β)
        Form NBI_β subproblem and solve (3);
        z = z + 1;
    End while;
    return Pareto front;
end function;
```

5.4. Summary

This chapter discussed the multi-objective optimization in order to optimize the objectives while considering the trade-offs among them. We elaborated on our NBI method and explained and how we solved it using Julia. In the next chapter we will present the experimental results based on our method for different circuits and then compare the results with other state-of-the-art optimization methods.

Chapter 6

Experimental Results of Our Optimization Method for Analog CNFET Circuits

6.1. Introduction

This section presents the test results for the GBC+NBI method for different analog circuits and then compares the results with other multi-objective methods using well-known techniques to evaluate our proposed method. In order to solve NBI, we use MultiJuMP [46], which takes advantage of Ipopt. Ipopt is an open-source program that provides an interior-point line-search using the filter approach for large-scale nonlinear optimization. Depending on the applications, the computational effort spent during optimization using Ipopt is often focused on solving linear systems or computing problem functions and derivatives [55]. For the simulations in the thesis, we used Python, Julia, Perl and MATLAB.

6.2. Multi-objective Metrics

To compare our multi-objective GBC+NBI method with other state-of-the-art multi-objective optimizers we consider two indicators:

- Hypervolume
- Spacing Metric

6.2.1. Hypervolume

The hypervolume indicator [56] measures the size of the objective space that is dominated by a group of solutions such as the area in Fig. 31. A bounded space needs to be made by a user-defined reference point and the Pareto front to define the hypervolume indicator. A higher hypervolume indicates its superiority of the method to be estimated in terms of Pareto front.

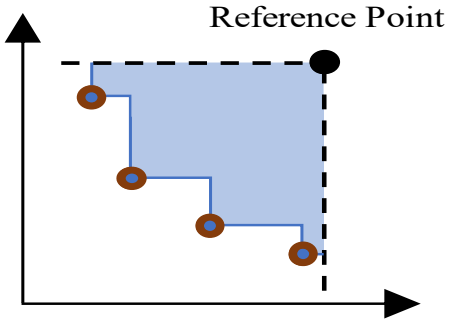

Fig. 31. Hypervolume

6.2.2. Spacing Metric

The spacing metric [57] is used to measure the uniformity of the solutions found as Pareto optimal front. The S metric is defined as

$$S^2 = \frac{\sum_{i=1}^{N}(x_i - \bar{x})^2}{N-1}, \tag{18}$$

where N is the number of nondominated solutions in the data set, x_i is the sum of the differences in the objective function values between solution i and its two nearest neighbors for each objective, and $\bar{x}$ is the mean of all the x_i that have been calculated.

The spacing metric approaches zero when the Pareto optimal solutions are nearly uniformly spaced.

6.3. Simulation Circuits

In [11], a detailed version of the CNFET model has been introduced to enable researchers to study the process variation, design trade-offs, tunneling currents, and parasitic effect of CNFETs at smaller dimensions. We use this SPICE-based virtual source model provided by Stanford University to do the circuit simulations in this work.

6.3.1. Second Generation Current Conveyor (CCII)

An op-amp has low slew rate at its output, and it suffers from contrast relation between gain and bandwidth. Showing unreliable frequency response, it is not very applicable for high frequency applications. On the other hand, the current-mode (CM) devices offer significant superiority over voltage-mode ones, such as low voltage, low power, wide bandwidth, and other better performance characteristics making them a great option for analog signal processing.

Sedra and Smith created the first-generation current conveyor in 1968. They later developed a new form of current conveyor called as second-generation current conveyor in 1970 [58, 59]. The current conveyor (CC) is a CM device, whose application in current

processing and voltage processing circuits such as integrators, impedance convertors and filters has led to introducing new variants of CC [60]. One of the most important CC is the second-generation CC (CCII) and Fig. 32 shows its block diagram, when $I_Z = +I_X$ the CCII is CCII+ called and when $I_Z = -I_X$ it is called CCII−.We consider two different CCII+ circuits in our experimental section.

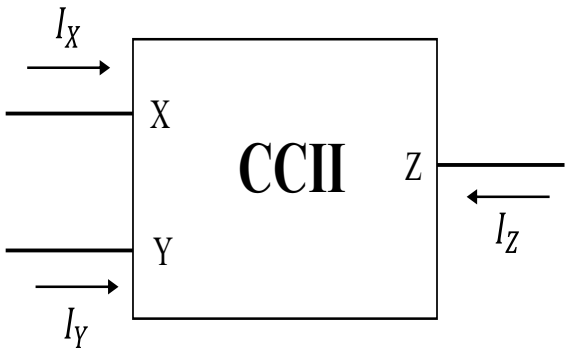

Fig. 32. CCII block diagram.

The CCII+ characteristics are expressed as follows:

$$I_Y = 0,\ V_X = V_Y,\ I_Z = +I_X \tag{19}$$

where V_Y and V_X are the voltages at port Y and X respectively, and I_Y, I_X and I_Z are the currents of ports Y, X and Z, respectively.

We can consider the CCII circuit as an ideal transistor where port Y is the gate with $I_Y = 0$ and ports X and Z acting as the source and drain; it is crucial to have the impedance at port Y and Z as high as possible and the port X impedance as low as possible.

6.3.1.1. High Accuracy CCII Circuit

We consider optimizing a high accuracy CCII+ circuits using our suggested method. As illustrated in Fig. 33, the netlist of the circuit was passed to our modified GBC method first to get the individual variation-aware optimum points in Table 7 (Here we consider unity gain bandwidth). Then we report the Pareto front of our proposed method GBC+NBI in comparison with WS GBC (weighted sum GBC), SQP (solving GBC+NBI while using SQP to solve NBI instead of the Julia solver), NSGA II, and multi-objective particle swarm (MOPSO), as shown in Fig. 34 and Table 8.

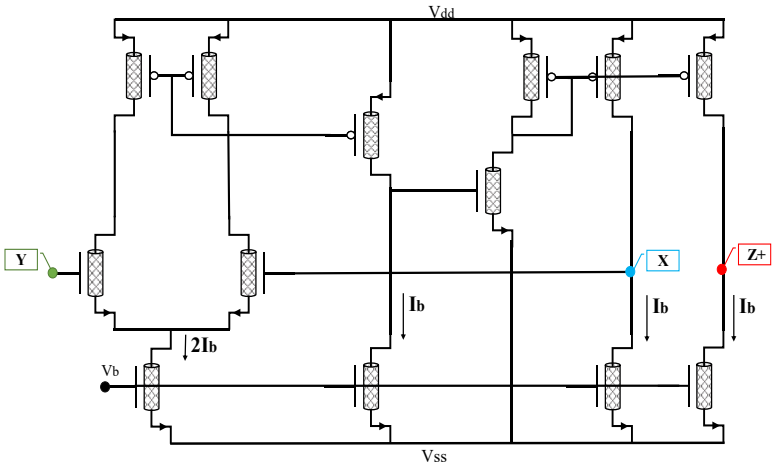

Fig. 33. High accuracy CCII+

Table 7. Single-objective optimization performance for the high accuracy CCII (from the modified GBC method)

Circuit performance being optimized	Port Z impedance (MΩ)	Bandwidth (GHz)
Bandwidth	7.658	90.57
Port Z impedance	0.627	356.40

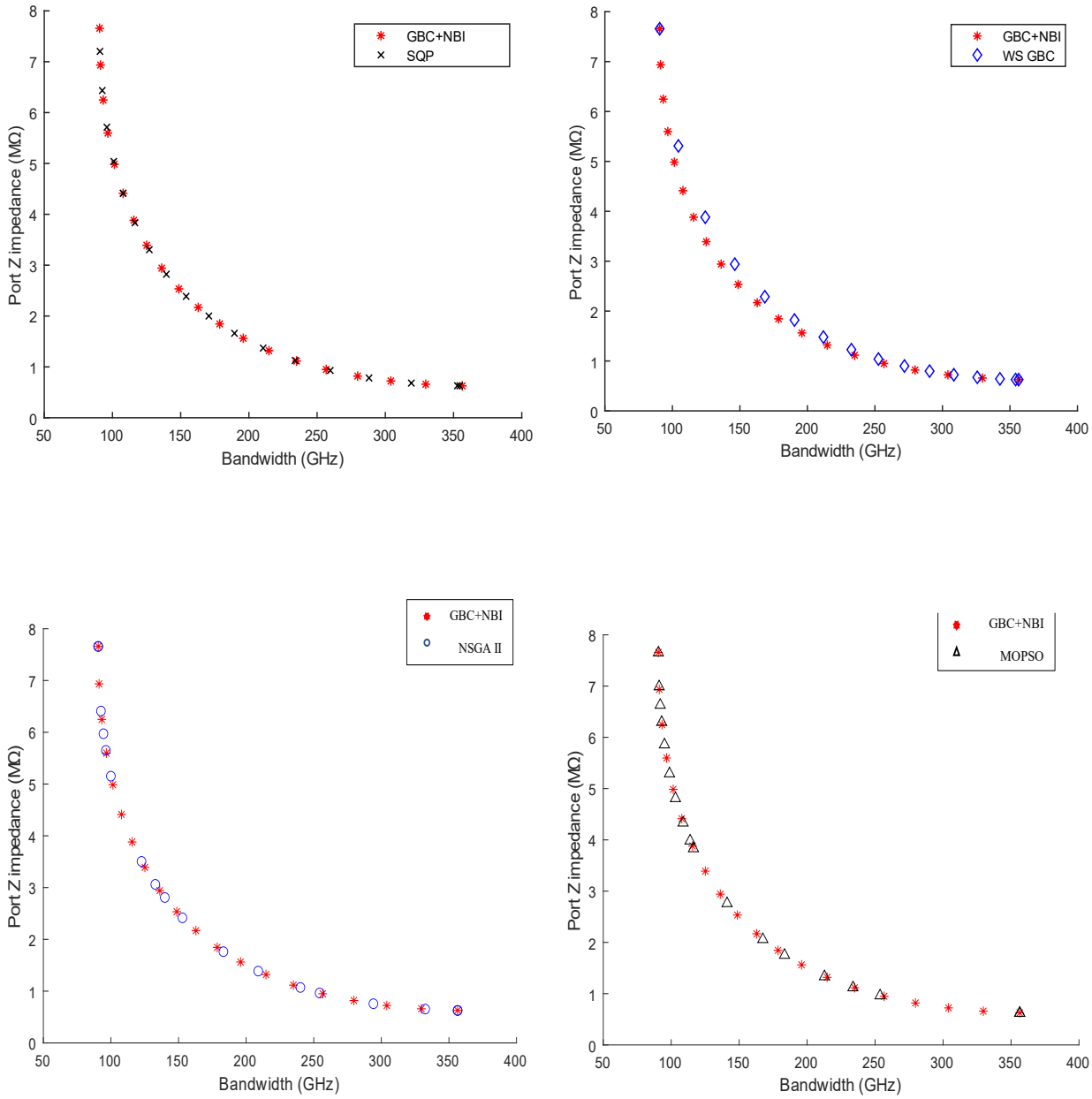

Fig. 34. High accuracy CCII Pareto front Comparison

Fig. 34 shows the Pareto front achieved by different multi-objective optimization methods We considered Port Z impedance and bandwidth as objective and the distribution and the range of them are shown in the above figure. We can observe that our method has a more uniform estimation of Pareto front.

Table 8. Performance comparison of multi-objective optimization for the high accuracy CCII

Method	Hypervolume	Spacing metric	Time(s)
NSGAII	1483	0.706	219
MOPSO	1436	0.989	223
SQP	1517	0.510	217
Weighted sum GBC	1446	0.844	205
GBC+NBI	1532	0.344	204

As we can see from the table our method GBC+NBI has higher hypervolume in comparison to others and lower spacing metric which can lead to the conclusion that our method is able to deliver better performance results compared to other state-of-the-art multi-objective optimization methods.

6.3.1.2. Wide Bandwidth CCII

The input of the wide bandwidth CCII in Fig. 35 (transistors 1-4) is a mixed trans-linear loop. The two current mirrors formed by transistors 5, 6, 7 and 10, 11 allow the

mixed loop to be dc biased by the current. The current mirrors built by transistors 12, 13 and 8, 9 are cross coupling to generate current at Z. Ideally, the input resistance at port Y should be infinite, since it draws no current at port Y it has a high input impedance and at port X it should have a low impedance. Similarly, we report the variation-aware single-objective optimal values (through our modified GBC) for port Z impedance and unity gain bandwidth of the circuit in Table 9. Similar comparison of the Pareto fronts with different multi-objective optimization methods are reported in Fig. 36 and Table 10.

Table 9. Single-objective optimization performance for the wideband CCII (from the modified GBC method)

Circuit performance being optimized	**Port Z impedance (MΩ)**	**Bandwidth (GHz)**
Bandwidth	1.442	408.9
Port Z impedance	3.054	211.1

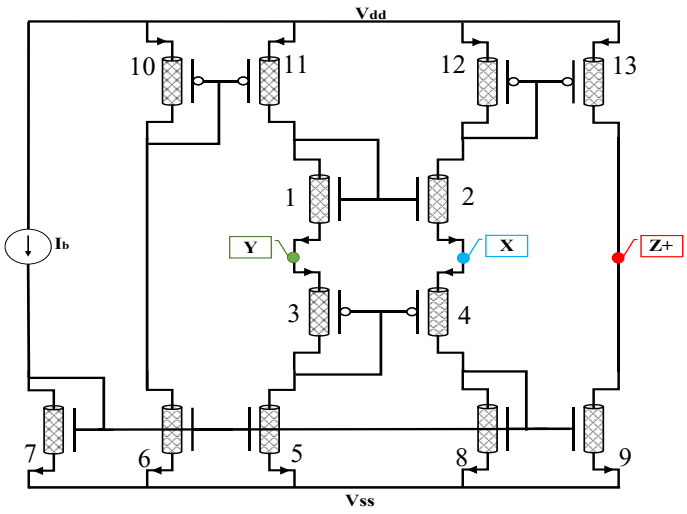

Fig. 35. Wide bandwidth CCII

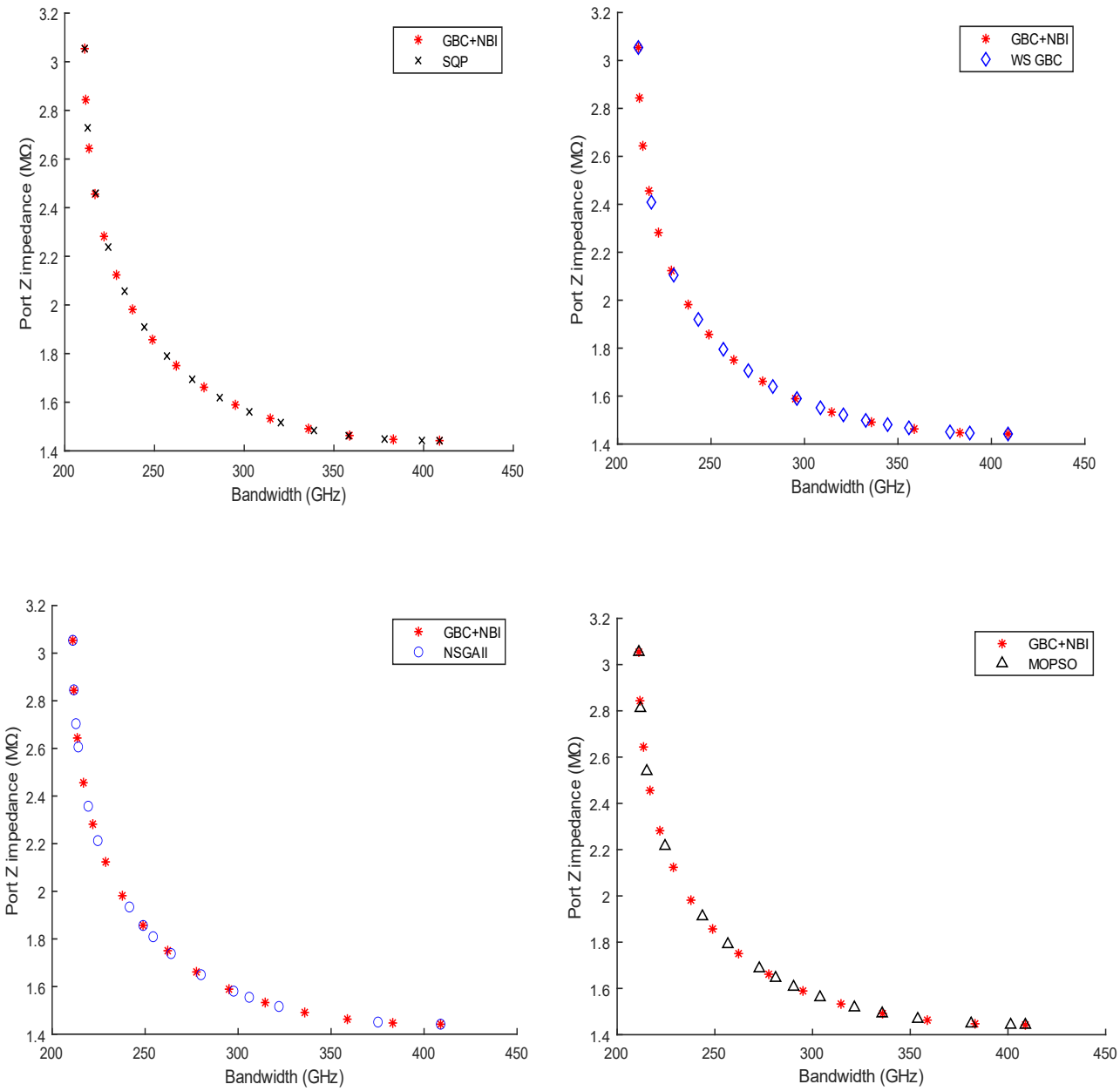

Fig. 36. Wide bandwidth CCII Pareto front comparison

Fig. 36 illustrates the Pareto front achieved by different multi-objective optimization methods. We considered Port Z impedance and bandwidth as objective and the distribution and the range of them are shown in the above figure. We can conclude that our method is superior due to more uniform estimation of Pareto front.

Table 10. Performance comparison of multi-objective optimization for the wide bandwidth CCII

Method	Hypervolume	Spacing metric	Time(s)
NSGAII	97640	0.72	173
MOPSO	97438	0.68	176
NBI with SQP	97751	0.62	171
Weighted sum GBC	97657	0.76	167
GBC+NBI	98070	0.53	165

As shown in the table, our technique GBC+NBI has a greater hypervolume and a smaller spacing metric than the others, implying that our method produces more superior performance than other state-of-the-art multi-objective optimization methods.

6.3.2. Folded Cascode Op-amp

In implementation of the operational folded cascode amplifier a differential pair is used as the input stage to the amplifier, functioning as the common source component of the cascade. The input transistors' drains are connected to two common gate transistors of opposing polarity. After that, the circuit is completed by connecting the common gate

transistors to an active current source load. In low voltage supply circuits, these topologies allow the input common-mode level to be near to the power supply voltage while also offering a strong output swing, a broad input common-mode range, and ideally steering.

We now report the variation-aware single-objective optimal values (through GBC) for the gain and -3dB bandwidth of the folded cascode op-amp as shown in Fig. 37 and Table 11. Similar comparison of the Pareto fronts with different multi-objective optimization methods is reported in Fig. 38 and Table 12.

Table 11. Single-objective optimization performance for the folded Cascode Op-amp (from the modified GBC method)

Circuit performance being optimized	**Gain (dB)**	**Bandwidth (GHz)**
Gain	46.458	2.372
Bandwidth	40.601	4.041

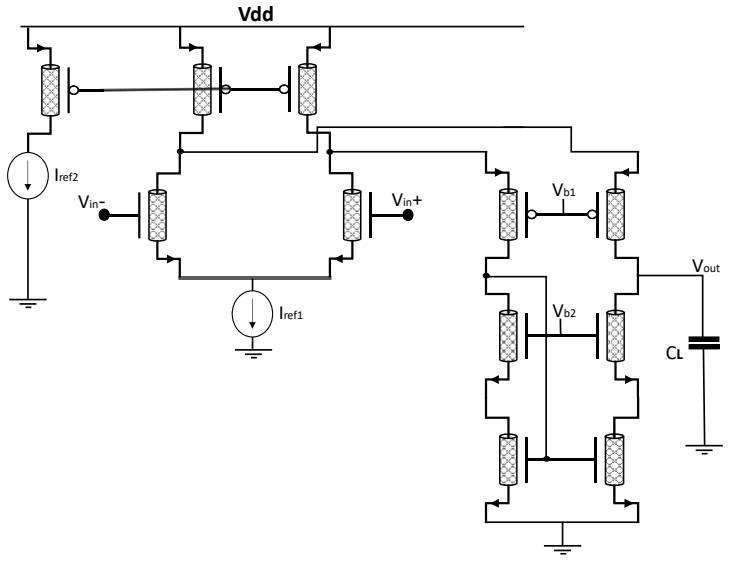

Fig. 37. Folded Cascode Op-amp

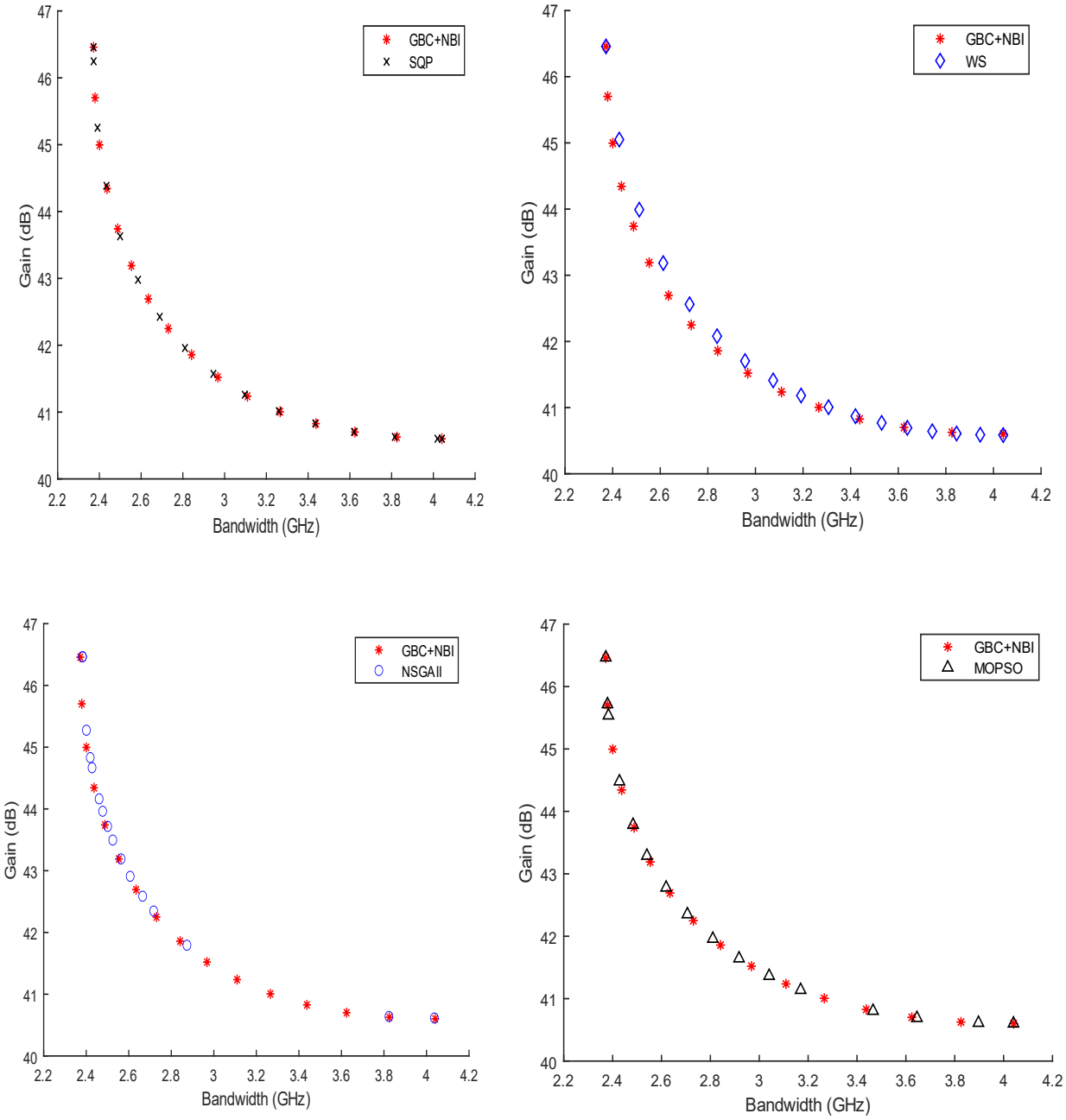

Fig. 38. Folded cascode op-amp comparison

Fig. 38 shows the result of different multi-objective optimization methods. We considered gain and bandwidth as objective and the distribution and the range of them are shown in the above figure. We can observe that our method has a more uniform estimation of Pareto front.

Table 12. Performance comparison of multi-objective optimization for the folded Cascode Op-amp

Method	Hypervolume	Spacing metric	Time(s)
NSGAII	7.805	0.5982	211
MOPSO	9.203	0.3788	214
NBI with SQP	9.197	0.5391	208
Weighted sum GBC	8.930	0.5157	197
GBC+NBI	9.231	0.3678	195

As indicated in the table above, our methodology GBC+NBI has a better hypervolume and a less spacing metric than the others, meaning that our method produces better outcomes than the other state-of-the-art multi-objective optimization methods.

6.3.3. Two-stage Amplifier

Two-stage amplifiers have been the most frequently used multistage amplifier due to its high gain and output swing. The input stage of a two-stage amplifier is a differential pair, and the second stage is a high gain stage driven by the first stage output.

We now report the variation-aware single-objective optimal values (through GBC) for the gain and bandwidth of the circuit in Fig. 39 and Table 13 (Here we consider -3dB bandwidth). Similar comparison of the Pareto fronts with different multi-objective optimization methods are reported in Fig. 40 and Table 14.

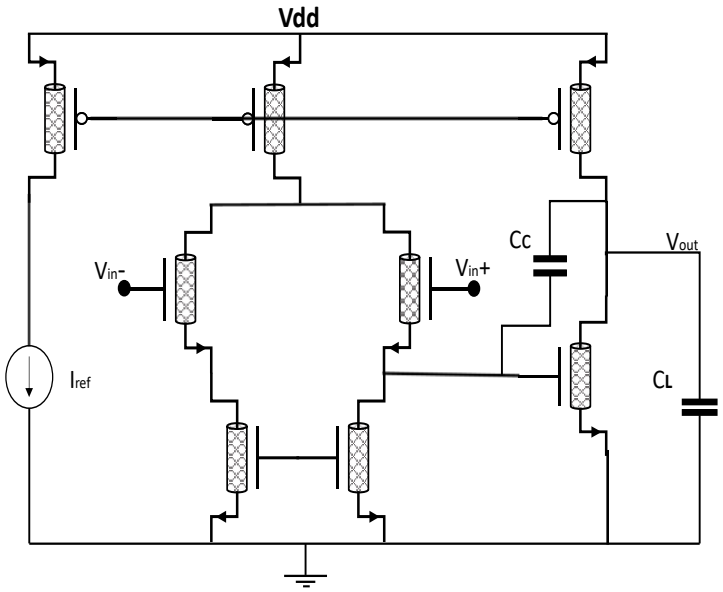

Fig. 39. Two-stage amplifier

Table 13. Single-objective optimization performance for two-stage amplifier (from the modified GBC method)

Circuit performance being optimized	**Gain (dB)**	**Bandwidth (MHz)**
Gain	73.6435	2.61
Bandwidth	38.1236	91.88

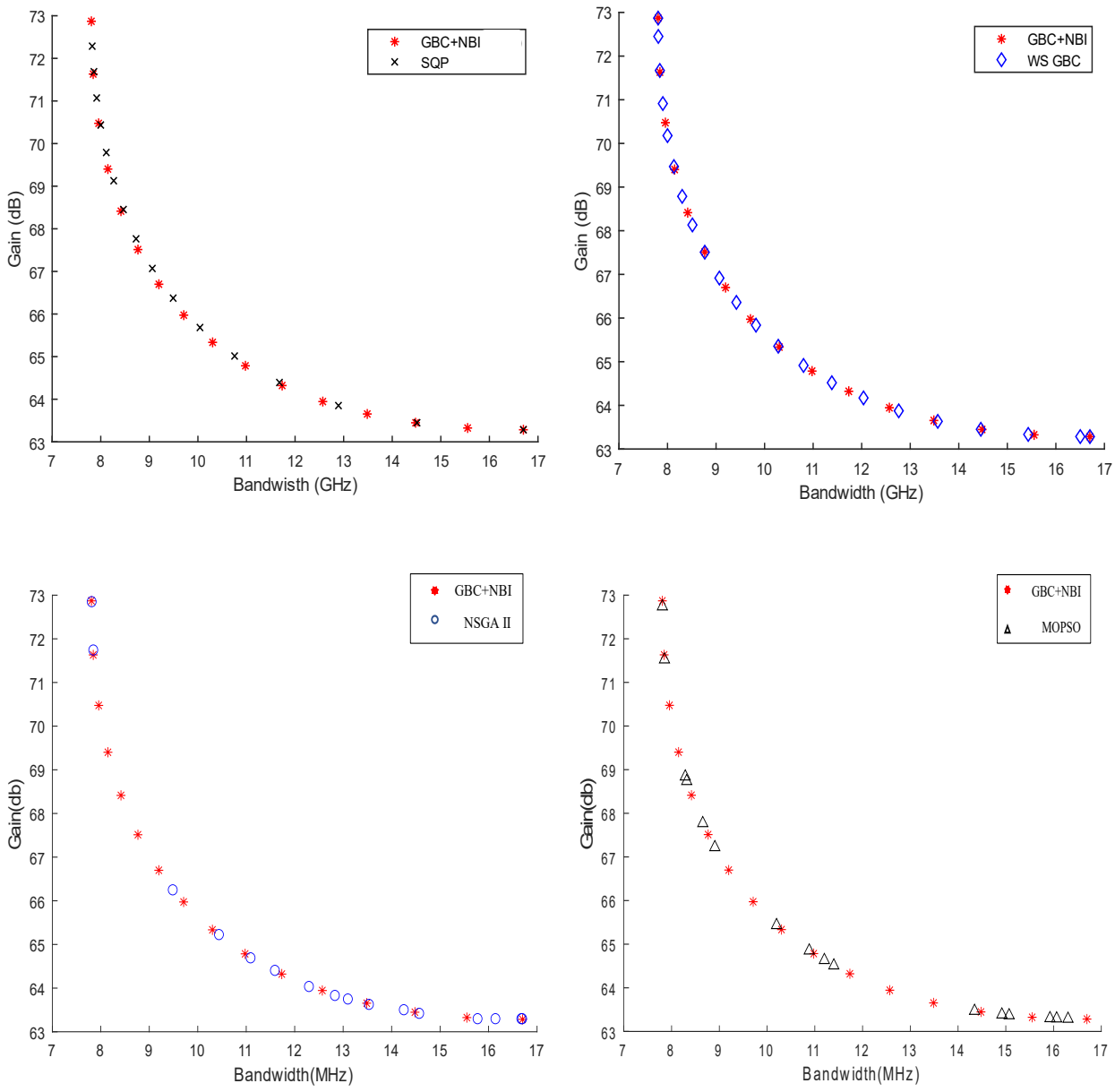

Fig. 40. Two-stage op-amp Pareto front comparison

Fig. 40 displays the two-stage op-amp Pareto front achieved by multiple multi-objective optimization methods and shows a comparison between them. Based on our observation we can conclude that our method has a more uniform estimation of Pareto front.

Table 14. Performance comparison of multi-objective optimization for two-stage Op-amp

Method	Hypervolume	Spacing metric	Time(s)
NSGAII	68.24	0.15	206
MOPSO	70.17	0. 13	211
NBI with SQP	72.38	0. 09	195
Weighted sum GBC	71.63	0.08	184
GBC+NBI	73.46	0.06	180

The results above exhibit that our proposed GBC+NBI method is able to generate better outcomes, as exemplified by higher hypervolumes and lower spacing metrics for those circuits. Compared to [39] using sequential quadratic programming (SQP) to solve the optimization, in this work we have used Julia MultiJuMP [46] for solving (17). MultiJuMP uses an interior-point optimizer (Ipopt) to solve the multi-objective optimization, which has advantages over SQP, such as low memory usage and the ability to efficiently solve large problems.

To further evaluate our method, we considered to optimize four objectives (gain, bandwidth, gain margin and phase margin) for the two-stage circuit (Fig. 39). The results and their comparisons with other methods are reported in Table 15 and Table 16.

Table 15. Single-objective optimization performance for two-stage amplifier (from the modified GBC method)

Circuit performance being optimized	**Gain (dB)**	**Bandwidth (MHz)**	**Gain margin**	**Phase margin**
Initial values	69.8988	8.06	33.8555	219.5028
Gain	73.6435	2.61	33.7365	216.9265
Bandwidth	38.1236	91.88	33.9461	229.8128
Gain margin	64.6358	6.02	34.0086	220.9738
Phase margin	38.5927	26.27	34.4398	241.2166

Table 16. Performance comparison of four objectives optimization for two-stage Op-amp

Method	Hypervolume	Spacing metric	Time(s)
NSGAII	113368	10.2639	329
MOPSO	105386	10.6535	331
NBI with SQP	97717	15.6997	325
Weighted sum GBC	100039	15.9868	312
GBC+NBI	113973	9.7555	310

6.4. Summary

In this chapter we presented different analog CNFET circuits and applied our multi-objective GBC+NBI method to optimize their performance while considering process variation. Then in terms of the performance metrics, we compared our proposed methodology with the other state-of-the-art multi-objective methods, which indicated the superiority of our work.

Chapter 7

Conclusion and Future Work

7.1. Thesis Summary

This thesis presents an overview of variability issues of analog CNFET circuits due to CNT process variation. The imperfect CNFET fabrication process can affect circuit performance drastically. And then we discuss our proposed solutions to design robustness against process variation in the context of multi-objective optimization.

In the first three chapters, CNFET parameters, models, and manufacturing methods have been discussed. The relationship between CNT parameters and their effects on circuit performance is also covered. Then the previous research on CNT manufacturing and its optimization have been surveyed and discussed. We have studied the sensitivity of CNT parameter variation thoroughly to decide the most important parameters that impact on the circuit performance and use them in our optimization to ensure circuit robustness.

Chapter 4 describes design centering concept using GBC. We present the basics and then discuss our modified version of the GBC method to achieve better accuracy and decrease the optimization timing.

In Chapter 5 we present multi-objective optimization and review the NBI method, which is a deterministic approach to do the optimization. The relationship between GBC and NBI methods is discussed. In particular, we describe how we use the optimized function from the GBC method and its optimum value for each objective to develop our variation-aware multi-objective optimizer.

At the end in Chapter 6 our fully deterministic multi-objective method is evaluated using a high accuracy CCII+ circuit, a wide bandwidth CCII, a two-stage amplifier, and a folded Cascode circuit. Based on our testing and experimental results, the design centering modified GBC+NBI method leads to better performance much closer to the Pareto front than the other start-of-the-art multi-objective optimization methods.

7.2. Thesis Contributions

The key contributions made in this thesis are to design and optimize CNFET circuits for analog applications. They are listed as follows:

- We systematically studied the robustness strategy against process variation of CNT parameters in analog CNFET circuit sizing design.
- We propose a new modified GBC method without needing prior knowledge of the range of parameter variation.

- We develop a fully deterministic multi-objective optimization method for reaching robust CNFET analog circuit design.
- We verify our techniques using experimental simulation.

7.3. Future work

As the CNFET transistors are becoming more accessible, by having their layout library we can develop a layout retargeting tool to support analog CNFET circuit design in a more efficient way. The future work including layout optimization may improve CNFET contact resistance, CNT density and CNT alignment to further increase performance of analog circuit circuits.

The transistors with greater reliability, lower manufacturing cost, and improved device characteristics will contribute to the most desirable development in CNT technology. For example, extrinsic factors such as the Schottky barrier between the CNT and metal contacts, numerous CNTs at a single gate, channel fringe capacitances, parasitic source/drain resistance, and series resistance owing to scattering effects might be added to the inner CNT transistors.

Further study of diameter variation by applying a method to consider different diameter for each CNT in one transistor and considering the probability of the existence of m-CNT in analog circuit to find the optimal number of CNTs that are needed per transistor to make sure the circuit continues its functionality.

Another important topic to consider in future work is parasitic resistance and capacitance in CNFET. The parasitic resistance is the series resistance consisting of two

parts: the metal-CNT contact resistance that is considered by the transmission line model [61] and the source/drain extension resistance which is based on the 1D quantum transport theory [62]. Also, the parasitic capacitance is composed of two parallel capacitances: the gate-to-contact capacitance and the gate-to-CNT fringe capacitance (their analytical model is adopted from [63]).

References

[1] G. E. Moore, "Cramming more components onto integrated circuits," in *IEEE Solid-State Circuits Society Newsletter*, vol. 11, no. 3, pp. 33-35, 1965.

[2] L. Zhang and U. Kleine, "A novel analog layout synthesis tool," in *IEEE International Symposium on Circuits and Systems (ISCAS)*, pp. V-V, 2004.

[3] P. Avouris, T. Hertel, R. Martel, T. H. R. H. S. Schmidt, H. R. Shea, & R. E. Walkup, "Carbon nanotubes: Nano mechanics, manipulation, and electronic devices," in *Applied Surface Science*, Vol. 141, Nov. 3–4, pp. 201-209, 1999.

[4] Zhou, Xinjian, Ji-Yong Park, Shaoming Huang, Jie Liu, and Paul L.McEuen., "Band structure, phonon scattering, and the performance limit of single-walled carbon nanotube transistors.," in *Physical Review Letters*, 95. 146805., 2005.

[5] L. Zhang, "VLSI circuit layout," in *Wiley Encyclopedia of Computer Science and Engineering*, vol. 5 pp. 3034-3044, 2009.

[6] X. Dong and L. Zhang,, "Process-variation-aware rule-based optical proximity correction for analog layout migration," in *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 36, no. 8, pp. 1395-1405, 2017.

[7] J. Appenzeller, "Carbon Nanotubes for High-Performance Electronics—Progress and Prospect," in *in Proceedings of the IEEE*, 2008.

[8] Srivastava, Ashok, et al., "Current Transport Modeling of Carbon Nanotube Field Effect Transistors," in *Physica Status Solidi. A, Applications and Materials Science 206. 1569 - 1578.*, 2009.

[9] [Online]. Available: https://www.cheaptubes.com/product-category/multi-walled-carbon-nanotubes/.

[10] D. Srivastava, M. Menon, K. Cho., "Computational Nanotechnology with Carbon Nanotubes and Fullerenes," in *IEEE Educational Activities Department*, vol. 3, no. 4, pp. 42-55, 2001.

[11] C. Lee, E. Pop and H. -. P. Wong,, "Compact modeling and design optimization of carbon nanotube field-effect transistors for the sub-10-nm technology nodes," in *73rd Annual Device Research Conference (DRC)*, pp. 275-276, 2015.

[12] A. Khakifirooz, O. M. Nayfeh, and D. Antoniadis,, "A Simple Semi-empirical Short-Channel MOSFET Current–Voltage Model Continuous Across All Regions of Operation and Employing Only Physical Parameters," *IEEE Transactions on Electron Devices,* vol. 56, no. 8, pp. 1674-1680, 2009.

[13] S. Rakheja, D. Antoniadis,, "nanoHUB," MVS (Version 1.1.1) Nanotransistor Model (Silicon), DEC 2015. [Online]. Available: https://nanohub.org/publications/15/4.

[14] A. G. Perri, Dispositivi Elettronici Avanzati, 2011.

[15] R. Marani and A. G. Perri, "The next generation of FETs: CNTFETs.," *International Journal of Advances in Engineering & Technology (IJAET),* Vols. Vol. 8, no. 5, 2015.

[16] P. L. McEuen, M. S. Fuhrer, and H. Park,, "Single walled carbon nanotube electronics," in *IEEE Transactions on Nanotechnology*, vol. 1, no. 1, pp. 78-85, 2002.

[17] S. Heinze, J. Tersoff, R. Martel, V. Derycke, J. Appenzeller, Ph. Avouris,, "Carbon Nanotubes as Schottky Barrier Transistors," in *Physical review letters 89. 106801.*, 2002.

[18] S. J. Wind, J. Appenzeller, R. Martel, V. Derycke, and P. Avouris, "Fabrication and electrical characterization of top gate single-wall CNFETs," *Journal of Vacuum Science and Technology B,* vol. 20, pp. 2798-2801, 2002.

[19] Z. Chen, D. Farmer, S. Xu, R. Gordon, P. Avouris and J. Appenzeller,, "Externally Assembled Gate-All-Around Carbon Nanotube Field-Effect Transistor," *IEEE Electron Device Letters,* vol. 29, no. 2, pp. 183-185, 2008.

[20] M. Zhang, P. C. H. Chan, Y. Chai, Q. Liang and Z. K. Tang, "Novel Local Silicon-Gate Carbon Nanotube Transistors Combining Silicon-on-Insulator Technology for Integration," in *IEEE Transactions on Nanotechnolog*, vol. 8, no. 2, pp. 260-268, 2009.

[21] S. Dandeliyaa and A. Srivastava, "Special topics in nanoelectronics," ABV-IIITM-Gwalior (MP) India, https://tiiciiitm.com/profanurag/STNE-2.1.pdf.

[22] J. Kong, H. T. Soh, A. M. Cassell, C. F. Quate, and H. Dai,, "Synthesis of individual single-walled carbon nanotubes on patterned silicon wafers," *Nature,* vol. 395, pp. 878- 881, 1998.

[23] [Online]. Available: https://www.cheaptubes.com/carbon-nanotubes-history-and-production-methods-2/.

[24] O. Zaytseva, G. Neumann,, "Carbon nanomaterials: production, impact on plant development, agricultural and environmental applications," in *Chem. Biol. Technol. Agric. 3*, 2016.

[25] Y. L. Y. W. Z. Ren, Chemical Vapor Deposition of Carbon Nanotubes, Springer, Berlin, Heidelberg: Aligned Carbon Nanotubes. NanoScience and Technology, 2012.

[26] S. D. M. L. M. B. P. M. A. J. H. D. H. K. a. P. M. J. Guo, "Assessment of silicon MOS and carbon nanotube FET performance limits using a general theory of ballistic transistors," in *Digest. International Electron Devices Meeting*, pp. 711-714, 2002.

[27] F. Rahman, A. M. Zaidi, N. Anam and A. Akter,, "Performance evaluation of a 32-nm CNT-OPAMP: Design, characteristic optimization and comparison with CMOS technology," in *14th International Conference on Computer and Information Technology (ICCIT)*, pp. 583-588,, 2011.

[28] J. Zhang, N. Patil, A. Lin, H. . -S. Philip Wong and S. Mitra,, "Carbon nanotube circuits: Living with imperfections and variations,," in *Automation & Test in Europe Conference & Exhibition (DATE)*, pp. 1159-1164, 2010.

[29] S. K. Vendra and M. Chrzanowska-Jeske,, "Carbon Nanotube Length Variation in Correlated CNFETs," in *IEEE 20th International Conference on Nanotechnology (IEEE-NANO)*, pp. 57-61, 2020.

[30] S. Banerjee, A. Chaudhuri, and K. Chakrabarty,, "Analysis of the Impact of Process Variations and Manufacturing Defects on the Performance of Carbon-Nanotube FETs," *IEEE Transactions on Very Large-Scale Integration (VLSI) Systems,* vol. 28, no. 6, pp. 1513-1526, 2020.

[31] A. G. Amer, R. Ho, G. Hills, A. P. Chandrakasan and M. M. Shulaker, "SHARC: Self-Healing Analog with RRAM and CNFETs," in *IEEE International Solid- State Circuits Conference - (ISSCC)*, pp. 470-472, 2019.

[32] G. Hills, C. Lau, T. Srimani, M. D. Bishop, P. Kanhaiya, R. Ho, A. Amer, and M. M. Shulaker., "Advances in Carbon Nanotube Technologies: From Transistors to a RISC-V Microprocessor," in *Proceedings of the 2020 International Symposium on Physical Design (ISPD '20)*, pp. 33–38, 2020.

[33] G. Hills, C. Lau, A. Wright, S. Fuller, M. D. Bishop, T. Srimani, P. Kanhaiya, R. Ho, A. Amer, Y. Stein, D. Murphy, Arvind, A. Chandrakasan, and M. M. Shulaker,, "Modern microprocessor built from complementary carbon nanotube transistors," *Nature,* vol. 572, no. 7771, p. 595–602, 2019.

[34] M. S. Ansari and S. K. Tripathi,, "Carbon Nanotubes as FET Channel: Analog Design Optimization considering CNT Parameter Variability," in *IOP Conf. Ser.: Mater. Sci. Eng. 225 012167* , 2017.

[35] A. Imran, M. Hasan, A. Islam and S. A. Abbasi,, "Optimized Design of a 32-nm CNFET-Based Low-Power Ultrawideband CCII," in *IEEE Transactions on Nanotechnology*, vol. 11, no. 6, pp. 1100-1109, 2012.

[36] M. Yasir and N. Alam., "Design of CNTFET-Based CCII Using gm/ID Technique for Low-Voltage and Low-Power Applications," in *Journal of Circuits, Systems and Computers*, vol. 29, no. 09, 2020.

[37] T. Liao and L. Zhang,, "Parasitic-aware GP-based many-objective sizing methodology for analog and RF integrated circuits,,," in *22nd Asia and South Pacific Design Automation Conference (ASP-DAC)*, pp. 475-480, 2017.

[38] X. Dong and L. Zhang,, "PV-Aware Analog Sizing for Robust Analog Layout Retargeting with Optical Proximity Correction," *ACM Transactions on Design Automation of Electronic Systems (TODAES),* vol. 23, no. 6, p. 1–19, 2018.

[39] G. Stehr, H. Graeb and K. Antreich,, "Performance trade-off analysis of analog circuits by normal-boundary intersection,," in *Proceedings of the 40th annual Design Automation Conference, Association for Computing Machinery*, pp. 958–963, 2003.

[40] D. Schreiber and J. Kampe,, "OTA Performance Space Modeling for System-Level Optimization of SC-circuits,," in *IEEE International Conference on Electronics, Circuits and Systems (ICECS)*, pp. 1-4, 2020.

[41] D. Mueller-Gritschneder, H. E. Graeb, and U. Schlichtmann, , "A successive approach to compute the bounded pareto front of practical multiobjective optimization problems,," *SIAM Journal on Optimization,,* vol. 20, no. 2, p. 915–934, 2009.

[42] C. A. Coello Coello and M. S. Lechuga, "MOPSO: a proposal for multiple objective particle swarm optimization," in *Proceedings of the 2002 Congress on Evolutionary Computation. CEC'02 (Cat. No.02TH8600)*, vol.2 pp. 1051-1056 , 2002.

[43] K. Deb, A. Pratap, S. Agarwal and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II,"," in *IEEE Transactions on Evolutionary Computation*, vol. 6, no. 2, pp. 182-197, 2002.

[44] N. Patil , A. Lin , ER. Myers , K. Ryu , A. Badmaev , C. Zhou, et al., "Wafer-scale growth and transfer of aligned single-walled carbon nanotubes.," in *IEEE Transactions on Nanotechnology*, vol. 8, no. 4, pp. 498-504, 2009.

[45] C. G. Almudéver and A. Rubio, "Variability and reliability analysis of CNFET in the presence of carbon nanotube density fluctuations.," in *Proceedings of the 19th International Conference Mixed Design of Integrated Circuits and Systems (MIXDES)*, pp. 124-129, 2012.

[46] A. N. Riseth, M. Besançon, and A. Naiem,, "anriseth/MultiJuMP.jl: Utopia and Nadir points," Zenodo, 15 Feb 2019. [Online]. Available: Available: https://doi.org/10.5281/zenodo.2565839.

[47] M. E. Mortenson, "Mathematics for Computer Graphics Applications: An Introduction to the Mathematics and Geometry of CAD/Cam, Geometric Modeling, Scientific Visualizati," in *Industrial Press Inc.*, 1999.

[48] [Online]. Available: https://pomax.github.io/bezierinfo/.

[49] R. Schwencker, J. Eckmueller, H. Graeb and K. Antreich, "Automating the sizing of analog cmos-circuits by consideration of structural constraints.," in *Design, Automation and Test in Europe (DATE)*, pp. 323-327, 1999.

[50] R. Schwencker, F. Schenkel, H. Graeb, and K. Antreich., "The generalized boundary curve-a common method for automatic nominal design and design centering of analog circuits," in *In Proceedings of the Design, Automation and Test in Europe Conference and Exhibition (DATE)*, pp. 42-47, 2000.

[51] Z. Liu and L. Zhang,, "Performance-constrained template-driven retargeting for analog and RF layouts," in *Proceedings of the 20th Symposium on Great Lakes Symposium on VLSI*, pp. 429–434, 2010.

[52] T. Liao and L. Zhang,, "Efficient parasitic-aware hybrid sizing methodology for analog and RF integrated circuits,," *Integration, the VLSI Journal,* vol. 62, pp. 301-313, 2018.

[53] T. Liao and L. Zhang,, "Layout-dependent effects aware gm/ID-based many-objective sizing optimization for analog integrated circuits," in *IEEE International Symposium on Circuits and Systems (ISCAS),*, pp. 1-5, 2018.

[54] S. Motta, R., Afonso, S.M.B. & Lyra, P.R.M, "A modified NBI and NC method for the solution of N-multi-objective optimization problems," in *Structural and Multidisciplinary Optimization 46*, 239–259, 2012.

[55] A. Waechter, "Getting Started with IPOPT in 90 minutes: Short tutorial," in *In Combinatorial Scientific Computing*, 2009.

[56] N. Beume, B. Naujoks, and M. Emmerich, "SMS-EMOA: Multiobjective selection based on dominated hypervolume," *European Journal of Operational Research (EJOR),* vol. 181, no. 3, p. 1653–1669, 2007.

[57] E. Masazade, R. Rajagopalan, P. Varshney, C. Mohan, G. Sendur, and M. Keskinoz,, "A Multiobjective Optimization Approach to Obtain Decision

Thresholds for Distributed Detection in Wireless Sensor Networks," in *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 40, no. 2, pp. 444-457, 2010.

[58] K . C. Smith and A. S. Sedra,, "The current conveyor A new circuit building block,," in *Proceedings of the IEEE*, vol. 56, no. 8, pp. 1368-1369, 1968.

[59] A . S. Sedra and K. C. Smith,, "A second generation current conveyor and its applications,," in *IEEE Transactions on Circuit Theory*, vol. 17, no. 1, pp. 132-134, 1970.

[60] S. Minaei, O. K. Sayein, and H. Kuntman,, "A new CMOS electronically tunable current conveyor and its application to current mode filter,," in *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 53, no. 7, pp. 1448-1457, 2006.

[61] Paul M. Solomon, "Contact Resistance to a One-Dimensional Quasi-Ballistic Nanotube/Wire," *IEEE Electron Device Lett,* vol. 32, no. 3, pp. 246-248, 2011.

[62] S. Datta, Quantum Transport: Atom to Transistor, Cambridge: U.K.: Cambridge Univ. Press, 2006.

[63] J. Deng and H.-S. P. Wong, , "Modeling and Analysis of Planar-Gate Electrostatic Capacitance of 1-D FET With Multiple Cylindrical Conducting Channels,," *IEEE Trans. Electron Devices,,* vol. 54, no. 9, pp. 2377-2385, 2007.

Appendix: Published Papers

1) Z. Heshmatpour, L. Zhang, H. Heys "Variation-Aware Analog Circuit Sizing in Carbon Nanotube," in Proc. IEEE International Symposium on Circuits and Systems (ISCAS), January 2022, (Accepted).

2) Z. Heshmatpour, L. Zhang, H. Heys " Multi-Objective Variation-Aware Sizing for Analog CNFET Circuits," in Proc. IEEE International Symposium on Quality Electronic Design (ISQED), December 2021, (Accepted).

3) Z. Heshmatpour, L. Zhang, H. Heys " Robust CNFET Circuit Sizing Optimization," in Proc. IEEE International Symposium on VLSI Design, Automation and Test Design (VLSI_DAT), December 2021, (Accepted).