PRECISION FREQUENCY AND PHASE SYNTHESIS TECHNIQUES IN INTEGRATED CIRCUITS FOR BIOSENSING, COMMUNICATION AND RADAR

Thesis by

Hua Wang

In Partial Fulfillment of the Requirements for the

degree of

Doctor of Philosophy

CALIFORNIA INSTITUTE OF TECHNOLOGY

Pasadena, California

2009

(Defended March 17, 2009)

To My Family

Acknowledgements

First and foremost, I would like to express my deepest gratitude to my advisor, Professor Ali Hajimiri for his support over the course of my Ph.D. study at Caltech. His high-level guidance and insightful technical advice have led me through this adventurous academic journey full of discovery. More importantly, his generous mentorship has nourished me to an extent far beyond the academic matters. His enthusiasm for knowledge and his persistence for perfection have been and will continue inspiring me to strive for excellence and success in my career.

I would also like to thank my other thesis committee members: Professor Axel Scherer, Dr. Sander Weinreb, Professor Azita Emami, and Professor Changhuei Yang for their assistance and helpful comments.

I am especially indebted to Dr. Yan Chen and Mr. David Wu for being my collaborators. Their active involvements and critical contributions are essential and irreplaceable for the success of my Ph.D. project.

I am grateful to all the members of Caltech High Speed Integrated Circuits and Caltech Millimeter-Wave IC groups. I am greatly indebted to Dr. Abbas Komijani, Dr. Arun Natarajan, Professor Ehsan Afshari, Dr. Xiang Guan, Professor James Buckwalter, and Professor Arjang Hassibi for helping me start my graduate study at Caltech. Particularly, I consider myself very lucky to have Professor Yujiu Wang as my colleague and moreover a great friend. I have benefitted from his helpful technical discussions and his consistent mental support. Our coffee breaks during the tape-out time are among the best moments in my Ph.D. memory. I also wish to thank my colleagues, Florian Bohn, Professor Sanggeun

Jeon, Shohei Kosai, Edward Keehr, Juhwan Yoo, Kaushik Sengupta, Jay Chen, Steven Bowers, Kaushik Dasgupta and Tomoyuki Arai for their support.

I would also like to thank Michelle Chen, Carol Sosnowski, Hamdi Mani, Naveed Near-Ansari, John Lilley, Gary Waters, Kent Potter, Jim Endrizzi, Laura Flower Kim, Tess Legaspi, Linda Dozsa, Tanya Owen, Trevor Roper and Lynn Hein, whose valuable work has created a friendly environment at Caltech for students.

This research has been supported by the Office of Naval Research and Caltech Innovation Initiative (CI^2) Research Grant.

I am also thankful to my supportive friends, Ali Vakili, Wen Li, Guangxi Wang, Tao Liu, Fangwei Shao, Ray Huang, Mike Liu, Jeffrey Chun-Hui Lin, Costis Sideris, Young Shik Shin, Udi Vermish, Jifeng Peng, Sawyer Fuller, Quoc (Brandon) Quach, Weiyi Li, Howard Ge and Frank Weigang Bao, who have always encouraged me.

Finally, I would like to give my greatest thanks to my family. The positive influences from my loving parents have taught me to keep my persistence and enthusiasm, which have proved to be vital for achieving my goals. The unreserved caring and understanding from my beloved fiancée, Ying Wang, has always been encouraging me to move forward. My family's unconditional love has been my greatest source of energy and the most invaluable gift I have ever received in my life.

Abstract

Today's CMOS technology provides circuit designers with a powerful implementation platform that supports innovation opportunities on both circuit-topology and system-architecture levels. Moreover, the versatility of CMOS implementation opens the door for a plethora of challenging and exciting interdisciplinary research.

This dissertation focuses on investigating novel techniques and applications for precision frequency and phase synthesis in CMOS. It consists of two parts: a CMOS compatible molecular-level biosensor and a multiple-beam/multi-band scalable CMOS phased array receiver system.

In the first part, a frequency-shift-based magnetic biosensing scheme is introduced to address the Point-of-Care (PoC) biomolecular diagnosis which requires high-sensitivity, ultra-portability and low-cost. Compared with existing biosensing schemes, the proposed scheme achieves a competitive sensitivity without using optical devices, external biasing fields or expensive post-processing steps. A discrete implementation first verifies the sensing mechanism and reveals several design insights. An integrated implementation based on standard 130nm CMOS process is then designed with differential sensing and temperature controlling schemes. Overall, with a differential uncertainty of 0.13ppm for relative frequency shift, the sensor achieves reliable detection of one single micron-size magnetic particle (D=4.5μm, 2.4μm and 1μm) as well as 1n-Molar real DNA samples labeled by magnetic nanoparticles (D=50nm).

In the second part, a high-resolution compensation technique is proposed to address mismatch and offset issues encountered by practical phased array system. It employs a

dense Cartesian interpolation scheme with a scalable architecture and a wide operation bandwidth. As an implementation example, a 6-to-18GHz dual-band quad-beam phased array CMOS receiver is presented, which is capable of forming four spatially independent beams at two different frequencies across a tritave bandwidth. With the mismatch compensation, the array element has achieved a maximum RMS phase error of 0.5˚ with an RMS amplitude variation less than 1.5dB for the 360˚ interpolation over the full operation bandwidth. For a 4-element phased array receiver system based on the designed CMOS chip, the electrical array pattern is measured at 6GHz, 10.4GHz and 18GHz, with the worst case peak-to-null ratio of 21.5dB. In addition, a broadband inductorless design methodology based on Cherry-Hooper topology is proposed for chip area saving. As implementation examples, we will show a DC-19GHz 10dB gain broadband buffer amplifier, a DC-12GHz broadband phase rotator with 10-bit resolution and a beam-forming network in a 10.4GHz to 18GHz phased array receiver chip with dual-beam capability.

Contents

List of Figures

Chapter 1: Introduction

"Logic will get you from A to B. Imagination will take you everywhere."

--- Albert Einstein

Ever since its debut in Julius Edgar Lilienfeld's invention in 1925, CMOS technology has experienced enormous amounts of improvement in its modeling, fabrication and implementation techniques. Today, CMOS is unquestionably the dominant choice for commercial electronics with its "application spectrum" ranging from microprocessor, memory cell, image sensor, data convertor and highly integrated transceivers [1].

In the last decade, CMOS transistors' sizes have been continuously shrinking, providing faster transistors and higher integration levels. However, this process down-scaling is not the designers' panacea for a guaranteed circuit performance improvement. This is because firstly smaller transistors lead to various design challenges such as dynamic range limitations, mismatches and power handling capabilities. But more importantly, this passive dependence on process scaling would diminish designers' creativity and imagination to explore new opportunities.

We can gain some technology-direction insights by examining what CMOS has provided and will provide for circuit designers. First of all, CMOS process ensures an unparallelled integration level at a low price-tag together with a remarkable reliability. Moreover, it supports immense signal processing power for various digital assisting functionalities. Furthermore, CMOS circuits are capable of generating and detecting electromagnetic (EM) signals with high accuracy and sensitivity. In addition, CMOS

process provides high-quality metal layers which can be patterned to well-defined EM structures as direct interfaces between the physical world and the on-chip electronics.

Therefore, CMOS essentially provides us with a powerful platform that supports a plethora of innovation opportunities both on the circuit-topology and the system-architecture levels. More importantly, the versatility of CMOS implementation opens the door for various challenging and exciting interdisciplinary research, such as integrated biosensors/actuators [2]-[9] and high-efficiency on-chip antennas [10]-[12].

Guided by this philosophy, through my five-year Ph.D. study at Caltech I have devoted my research effort in finding novel circuit techniques and applications in precision frequency and phase synthesis based on CMOS technology. The research work consists of two parts: integrated molecular-level biosensor and multi-beam/multi-band phased array receivers.

In the first part, new frequency synthesis techniques with a long-term stability are investigated, which forms the basis of the proposed frequency-shift magnetic biosensor scheme. Compared with all other CMOS biosensors, this scheme achieves a competitive sensing performance (1nM DNA) without any external biasing fields or expensive post-processing steps. Moreover, this scheme achieves, to our knowledge, the best reported performance (one single magnetic particle D=1μm) among integrated CMOS magnetic sensors. This frequency-shift based scheme is therefore ideal for advanced point-of-care (PoC) medical applications, where high sensitivity, high portability and a low price are needed.

In the second part, sources of mismatch and offset in phased-array systems are investigated to minimize the degradation of the array performance. A Cartesian

interpolation based calibration scheme is proposed together with circuit techniques. The scheme is implemented in a dual-band quad-beam 6-to-18GHz phased array receiver in CMOS. By using this scheme, mismatch in practical array implementations are compensated for, resulting in a worst-case array peak-to-null ratio of 21.5dB for a four-element array.

The detailed organization is given in the following section.

1.1 Organization

A brief background review of current CMOS biosensor technology is given in Chapter 2. The focus is on sensors with applications in biomolecular detection. Three major types of sensors are covered: CMOS fluorescence biosensors, CMOS electrochemical biosensors and CMOS magnetic biosensors. Each sensor scheme is presented with a reported implementation example followed by a discussion on the fundamental advantages and limitations.

Chapter 3 introduces our proposed sensing mechanism, i.e., frequency-based magnetic biosensing. Based on theoretical modeling, it will be demonstrated that the sensor transducer gain is determined by the sensing inductor design and the sensor noise floor is limited by the $1/f^3$ phase noise of the sensor oscillator. Furthermore, a sensor design scaling law is derived which guides subsequent sensor implementation.

Two sensor implementation examples are presented in Chapter 4. The first one is a discrete type thin-film design, whose measurement results confirm the validity of the proposed sensing scheme and sensor modeling. The second example is a CMOS implementation with eight parallel sensing cells. By applying differential sensing and a

temperature regulation scheme, the sensor achieves 0.13ppm as its ultra-low frequency shift measurement noise floor and the aforementioned detection capability.

To further improve the sensitivity by lowering the noise-floor, Chapter 5 proposes a novel low noise technique based on correlated double frequency counting (CDFC) with negligible power and design overhead. Theoretical derivations as well as practical circuit implementations are presented. In addition, as a modification of the basic CDFC scheme, an Interleaving-N CDFC technique is proposed, together with implementation considerations. This modified scheme can further suppress the noise floor and improve the sensitivity.

In Chapter 6, the phased array concept and performance degradation due to mismatch and offsets are first discussed. A Cartesian calibration scheme is proposed and implemented. A 6-to-18 GHz dual-band quad-beam CMOS phased array receiver system is then presented. The functionality of the compensation scheme is verified with measurement results of a 4-element electrical array system.

Chapter 7 introduces a new design methodology for Cherry-Hooper amplifiers to achieve inductorless bandwidth extension for chip area saving. Conventional broadband techniques and the Cherry-Hooper topology are first presented. Then the design methodology for the Cherry-Hooper amplifiers is proposed based on circuit insights and analysis for linear transfer function, weak nonlinearity and noise performance. Finally, three implementation examples are given to confirm the validity of the methodology.

Finally, Chapter 8 summarizes this dissertation and proposes potential research directions for future work.

Chapter 2: CMOS Biosensor

The purpose of this chapter is to present the current landscape of the field for CMOS biosensors in a concise summary. Nowadays, tremendous research and development efforts have been devoted in this exciting and booming area, particularly with an emphasis on biomolecular-level sensing. In this chapter, three reported molecular-level sensing modalities will be discussed as the examples for the state-of-the-art technology.

This chapter is organized as follows. In Section 2.1, the concept of biosensor and biosensing will be introduced first. CMOS technology, as a low-cost, high-yield and high-integration process will then be presented and shown as a promising platform for biosensor implementations. A brief survey on currently developed CMOS biosensor will be given in Section 2.2. As the first two examples, CMOS electrochemical biosensor in Section 2.2.1 and CMOS fluorescent biosensor in Section 2.2.2 will be shown. However, these two types of biosensor typically experience intrinsic noise floor from affinity binding process between the analytes and the probe molecules. To overcome this issue, CMOS magnetic biosensor with magnetic micro or nano particles as sensing labels has been proposed. An implementation example of this biosensor type will be demonstrated in Section 2.2.3. The limitation of the reported magnetic biosensor will also be discussed. Finally, a conclusion will be given in Section 2.4 for summary.

2.1 Introduction

A biosensor is a device for analyte detection that normally contains a biological component and a physicochemical detector component [13]. The sensor first requires

certain biological element or environment to interact with the target analytes. Then the chemical signals resulting from this interaction are transformed to other types of signals, e.g., electrical signal, to facilitate the measurement. This signal type transformation is the transducer mechanism of the sensor based on which the sensor type is normally classified. In addition, the sensor system can include further data processing techniques, such as filtering and correlation, to condition the measurement results in a desired way.

Therefore, biosensor/biosensing is inherently an interdisciplinary research field which presents challenges and requires joint explorations in various areas including biology, electromagnetics, stochastic modeling and signal processing, etc. Advance in this field is believed to have a huge impact on our daily life and have the potential to completely revolutionize the landscape of future medical service.

Today, biological and medical research has reached the microscopic level which involves characterizing the interactions between biomolecules and their functionalities in the cells. From information point of view, biomolecules are used ubiquitously in biological entities as powerful and reliable machinery to store, transmit and process biological information [14]. This includes well-known nucleotides such as DNAs and RNAs, and proteins such as enzymes and antibodies. Then, to detect those biomolecules both qualitatively and quantitatively in the given samples is the pre-requisite for molecular-level biological/medical research. This molecular sensing application presents challenges on the sensing technologies and demands advanced features, such as high sensitivity, small footprint, high parallel processing capabilities and low cost.

On the other hand, CMOS technology presents itself as a promising and powerful tool for biosensor implementation. Equipped with millions of transistors reliably integrated

onchip, CMOS provides immense signal processing power with well predicted noise/distortion performance, resulting in high-sensitivity performance. Also, with the advance of lithography, CMOS has achieved its critical fabrication features of several tens of nanometers. This resolution is sufficient to meet any form-factor requirements for current biosensing schemes. Moreover, as a low cost and standard process, CMOS supports well-scaled design both at the chip level and the module level. This ensures implementation of parallel detections for multiple analytes simultaneously. Furthermore, capable of generating and detecting electromagnetic (EM) signals with high accuracy and sensitivity, CMOS circuits are potentially viable to be directly used as sensors and actuators for innovative design methodology. In addition, CMOS integrated circuits can augment traditional BioMEMS as the signal generation/processing back-bone to achieve overall low system form-factor for implantable and ultra-portable applications.

Therefore, incorporating CMOS technology in biosensor device and system design would potentially improve the performance of existing sensing schemes and open the door for many novel sensing modalities.

2.2 A Brief Survey on CMOS Biosensor Technology

As mentioned in Section 2.1, applying CMOS for biosensing systems will have a dramatic impact in biosensor design, particularly in biomolecular sensing. In this section, we will introduce several CMOS biosensor designs as examples to demonstrate the huge potentials for this emerging and exciting research field.

2.2.1 CMOS Fluorescent Biosensor

Currently the dominant biosensor and microarray detection technology is based on fluorescence spectroscopy (wavelength typically between 400nm and 800nm) with fluorescent labels as the reporters for target analyte molecules [15]. Although tremendous research efforts have been devoted in developing other biosensing modalities, the fluorescence-based detection remains as the most sensitive and robust method, particularly for DNA detection applications. The performance advantages of this detection method over other methods originate from the uniqueness of the fluorescence phenomenon, which makes the generated signals very specific and less susceptible to biological interference. Furthermore, fluorescent groups generally have a much smaller mass and volume compared with target molecules, which minimizes their artifact effects on the biochemical properties of the target molecules and ensures the feasibility for dynamic measurement on biochemical reactions of the target molecules. In addition, the target molecules under investigation can be genetically modified with fluorescent tags attached, which potentially simplifies the experiment procedures as well as providing a mean of tracing the biological pathway of the target molecules. However, because the excitation and emission lights are normally close in the spectrum, to ensure that the photo detectors respond only to the emission light, high quality optical filtering systems are required for the fluorescence-based sensors, e.g., microarray reader or fluorescent microscopy. This partially explains the high cost and complexity of this type of system.

The general working mechanism for fluorescence-based sensing is as follows. The target molecules are first introduced to the sensor surface and immobilized by the pre-deposited probe molecules. Then the fluorescent tags are introduced, which can specifically

bind to the target molecules. After this tagging step, the excitation light, also called absorption light, is introduced. Due to the fluorescence behavior of the tags, an emission light will be sent out at a lower energy level, i.e. a longer wavelength. This emission light is then detected by the optical sensor. The information regarding the target molecules can be concluded from the emission light intensity both qualitatively and quantitatively. This is shown in Figure 2.1. Figure 2.2 shows the fluorescence spectrums of commonly used fluorescent dyes, Alexa Fluor 610 and Cy3 [16].

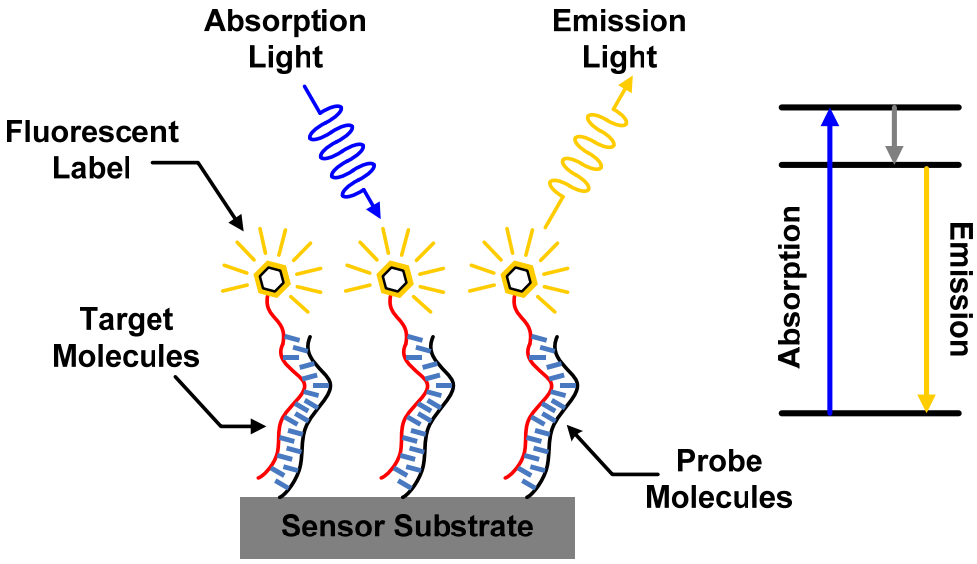


Figure 2.1: Mechanism for fluorescence emission

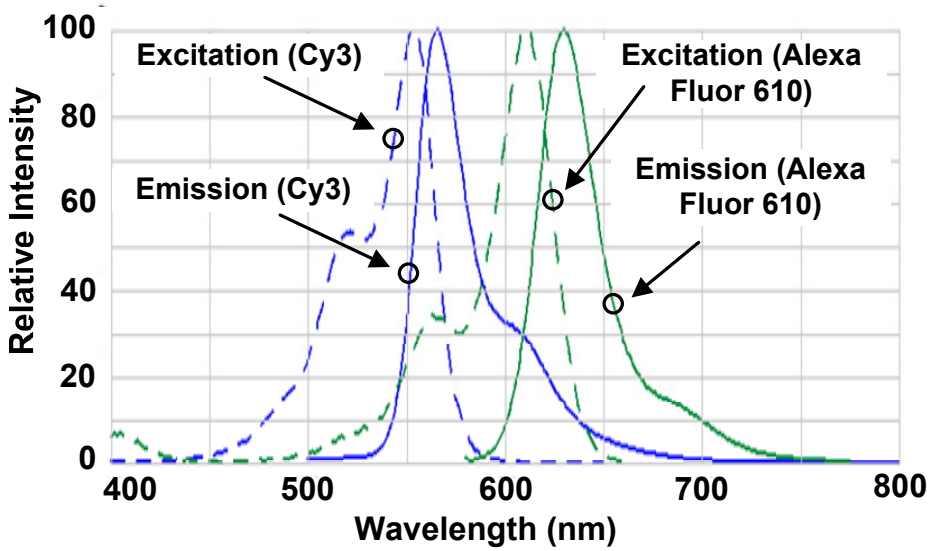


Figure 2.2: Fluorescence spectrum for Cy3 and Alexa Fluor 610

As an integrated design example, one fluorescence-based CMOS biosensor array system will be presented in the following part of this section [2]. The system includes a transducer array, molecular capture probes, readout circuitry, and in-pixel ADCs. Since the excitation light and emission light are close in the spectrum, to ensure a large sensor dynamics range without saturation, a multilayer thin-film dielectric long-pass optical filter with 20 layers of ZnS (n=2.30) and Na3AlF6 (n=1.35) is fabricated based on a fiber-optical faceplate (FOF) to filter out the excitation lights before they reach the CMOS sensor. The photodetector is designed to take advantage of the n-well/p-sub diodes in CMOS technology. Each diode is $50\times50\mu m^2$ and the pitch between each pixel is 250μm to be compatible with microarray probe spotting instruments, reduce the optical crosstalk between adjacent photodetectors, and provide sufficient silicon area for the in-pixel circuit design of amplification and signal processing. The lateral view of the sensor structure is shown in Figure 2.3.

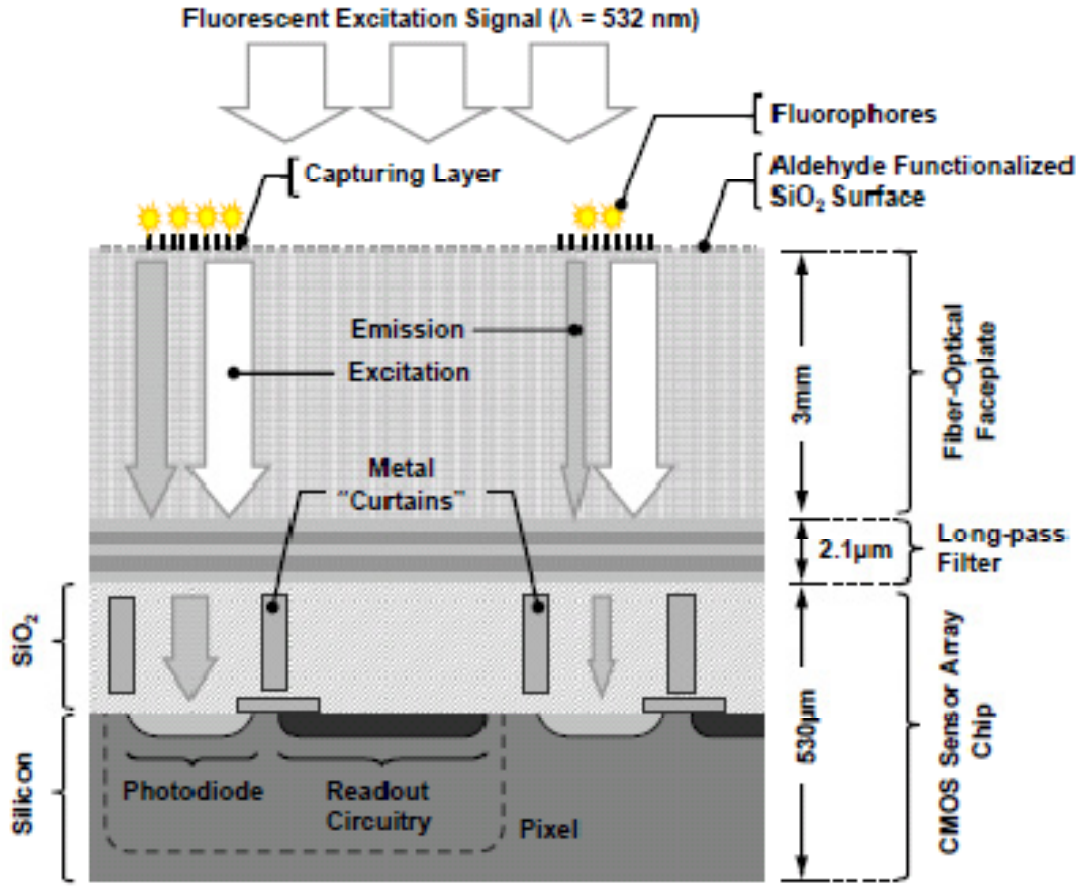


Figure 2.3: Cross section of the CMOS fluorescent-based biosensor microarray

The current response from the photodiode detector is integrated through a capacitive transimpedance amplifier (TIA) and further amplified by a chopper-stabilized amplifier. Real DNA hybridization kinetics are measured with respect to time for different DNA concentrations of 0.5nmol, 0.25nmol and 0.125nmol, and shown in Figure 2.4.

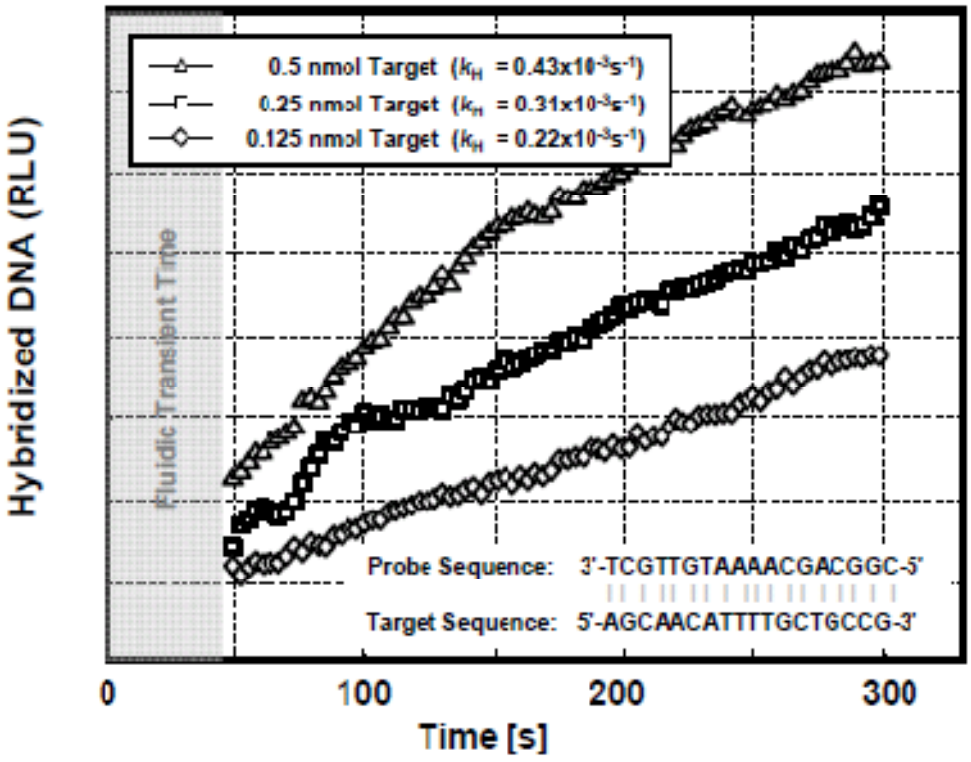


Figure 2.4: Measured real-time DNA hybridization kinetics for different concentrations

Among all the reported CMOS biosensor modalities, this design demonstrates the best measured sensitivity so far. This is due to the strong signaling of fluoresce labels. However, as we mentioned at the beginning of this section, the practical implementation requires complicated and expensive post-processing steps to fabricate the optical filter and fiber-optical faceplate structures. Moreover, as the excitation source, a high quality external laser generator is needed which inevitably limits the overall system integration level and form-factor.

2.2.2 CMOS Electrochemical Biosensor

Since most biomolecules carry extra electrical charges in aqueous solutions (or can be specifically attached with redox-active molecular-tags), it is therefore feasible to implement

simple detection schemes or even label-free detection by utilizing the electrochemical responses when the biomolecules are under pre-defined electrical excitations. Also, this type of sensor could achieve a very low-cost system implementation by fundamentally eliminating the expensive optical devices for excitation and detection, such as those used by fluorescent detections presented in Section 2.2.1.

In its generic form, this type of electrochemical biosensor is typically composed of one working electrode and one reference electrode with shared or separated ground electrodes. The electrical excitations are either DC signals or low frequency continuous waves (typically in the kHz range). In terms of the fabrication, additional post processing steps may be required to open the passivation layer and form the electrodes with desired metal, such as Au or Pt to facilitate deposition of probe molecules. The detailed sensor circuit design of the electrochemical sensors is determined by the operation technique, which includes impedance spectroscopy (IS), amperometric analysis, redox cycling and cyclic-voltammetry, etc.

As an example, a programmable electrochemical biosensor array based on cyclic-voltammetry operation will be presented in this section [5]. The sensor system is implemented in a 0.6μm 5V 3M2P standard CMOS process with additional lift-off processing steps to sputtering the Ti:W (20nm) and Pt(200nm) metal layers for forming electrodes. Then a 1.6μm -thick passivation layer stack (SiO_2 and Si_3N_4) is deposited for corrosion protection and opened through a reactive-ion etching (RIE) step. Every sensor cell contains one working electrode, one reference electrode and one counter electrode. Finally, a polymer layer (ploypyrrole) is required to be deposited onto the electrode surface for a desired electrical-eletrolyte interface.

In its cyclic voltammetry (CV) operation, the electrical potential on the working electrode is periodically varied with respect to the reference electrode as the excitation signal, while the respective current through the working electrode is measured as the response. This results in specific current-voltage (I–V) curves for different electrical-electrolyte surface. For example, for DNA sensing, since the DNA molecules have overall negative electrical charges due to the phosphate backbone, they will repel chloride ions from the electrode surface. Therefore, if the complementary DNA is hybridized onto the electrode, the increase in the negatively charged phosphate backbone would directly change the kinetics of the chloride ions, and thus alter the shape of the resulting I-V curves. This is shown in Figure 2.5, as follows.

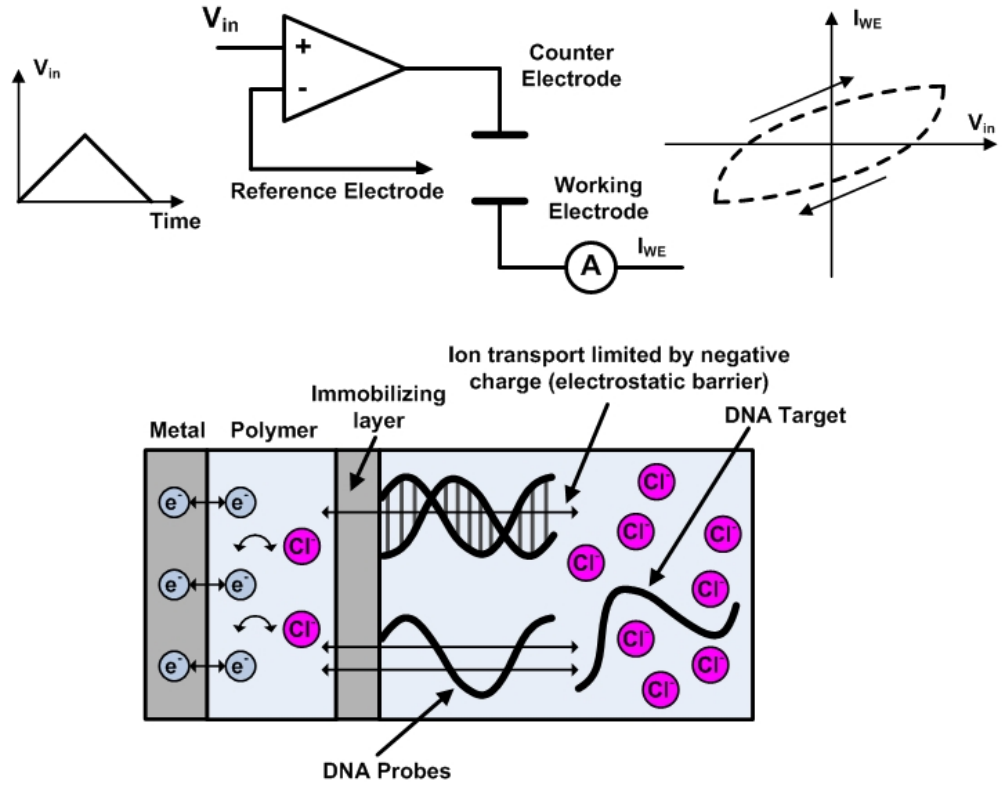


Figure 2.5: Illustration of a cyclic voltammetry experiment and description of the label-free electrochemical DNA hybridization detection principle

Two types of experiments are performed. One measures the CV response with and without complementary 30-mer target DNAs. At 100nM concentration, the average change for the I–V curve enclosed area is -38%. In the other experiment, the HIV-1 DNA is tested, which results in a 21% area increase. Both results are summarized in Figure 2.6.

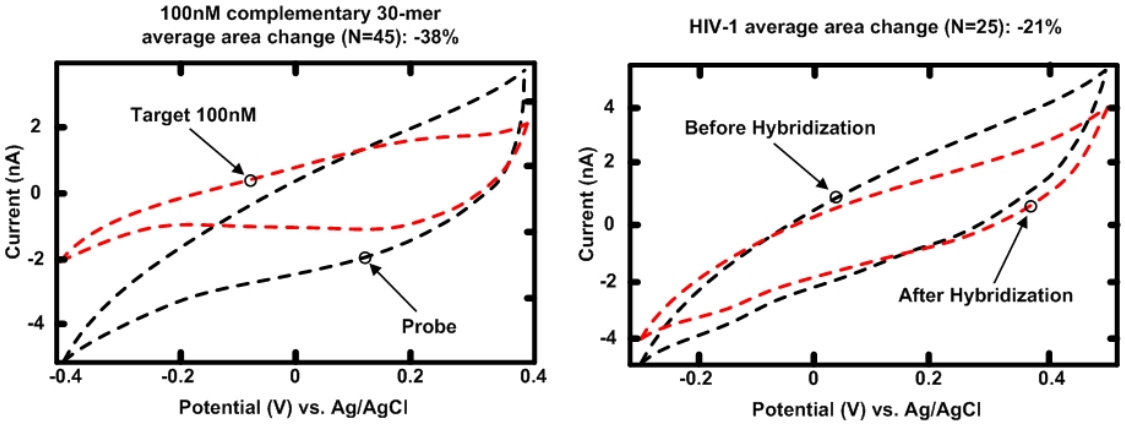


Figure 2.6: Measurement results for the CV electrochemical biosensor

Although compact and simple in implementation and potentially label-free in operation, electrochemical biosensors are highly sensitive to the offset and background noise. Moreover, since operation of all the sensors in this category relies on the electrode-electrolyte chemical interface, which is practically dominated by the biological short-noise through hybridization effect. This leads to the fundamental limitation on the sensor noise floor which determines the lowest detectable concentration of the target molecules. Therefore, the best reported electrochemical biosensors so far only achieve their sensitivity on the order of 100nM for the target molecules, which is several orders of magnitude higher than the concentrations used in regular biochemical experiments. This greatly limits the application and popularity of this type of sensor.

2.2.3 CMOS Magnetic Biosensor

The limitation of the above two CMOS biosensor schemes has stimulated research efforts in searching for a new non-optical biosensing modality while maintaining a high sensitivity and a low back-ground noise level. As a result, biosensing based on magnetic micro/nano particle labels have been proposed as one of the promising candidates [16].

Basically, magnetic particles in either micron or nanometer size are first attached onto the target molecules by labeling technique, such as ELISA sandwich structure. Then the magnetic biosensor sense the presence of those labeled magnetic particles to infer the existence of the target molecules in the test sample, shown in Figure 2.7. In general, the magnetic particle based biosensing scheme therefore presents following advantages. First of all, it fundamentally eliminates the bulky and expensive optics, which potentially leads to a low form factor and low system cost. Secondly, magnetic signal avoids significant signal drifting and interface noise due to hybridization process experienced by electrochemical biosensing. Moreover, magnetic signal does not have the signal quenching or decaying problems encountered by fluorescent labels in optical detection systems. Furthermore, since most biosamples do not produce magnetic signals with a comparable strength with respect to the magnetic labeling particles, magnetic biosensing can potentially achieve a very high signal to background noise ratio.

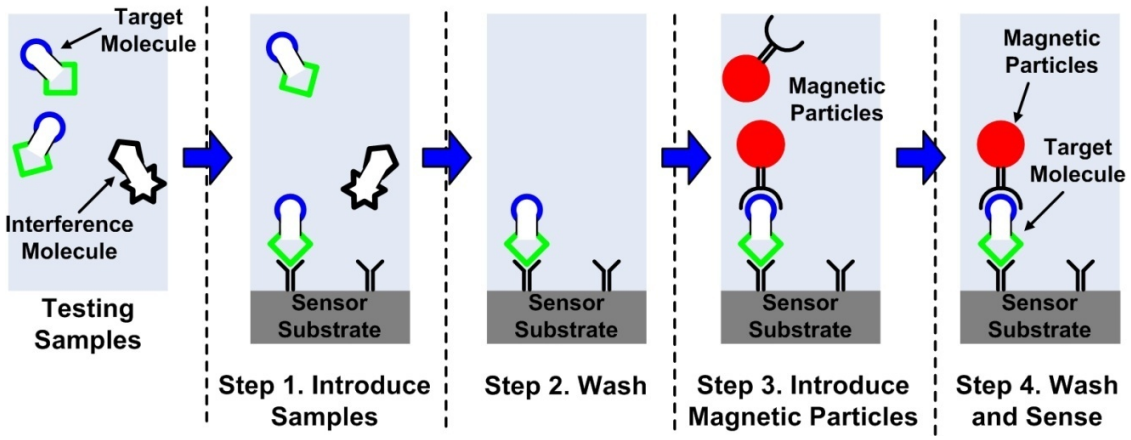


Figure 2.7: Magnetic label based biosensing

Current reported integrated magnetic sensor schemes include the Giant Magnetoresistance (GMR) biosensor [6][18][19] and the Hall Effect biosensor [9]. The former, also as a more sensitive sensing scheme, will be discussed here.

The GMR sensor utilizes the magnetoresistance property whose effective resistance changes when the magnetic labels are drawn close to the sensor surface. An implementation example is given as follows for the GMR sensor [6].

The sensor system adopts a differential scheme by using one active sensor and one reference sensor. To pattern the spin-valve structure on-chip, a 10-layer nanometer-thickness metal sputtering process is used. Then, a Ta/Au/Ta interconnect is used to link the spin-valve layer to the CMOS chip pad. Finally, Au patches are formed on top of the sensor for depositing probe molecules. The scanning electron microscope (SEM) photo and the cross-sectional view of the GMR sensor are shown in Figure 2.8. The spin-valve structures thus detect the immobilization of the magnetic particles and output it as the effective resistance change, which is eventually amplified by subsequent circuits.

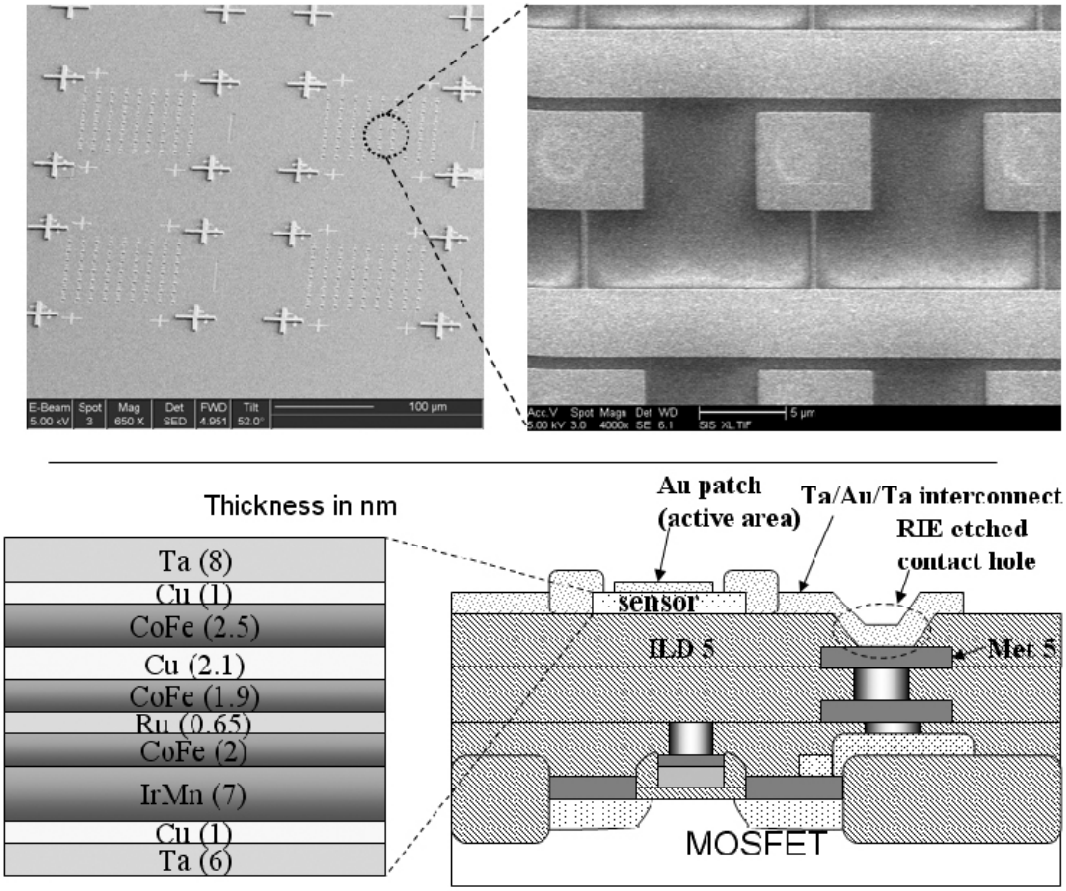


Figure 2.8: SEM photo and the cross-sectional view of the GMR sensor

As the experiment, 10nM complementary DNA samples labeled by magnetic nanoparticles are used as the sensing target. Meanwhile, the 100nM non-complementary DNA samples are used as the control experiment. The sensor response is shown in Figure 2.9.

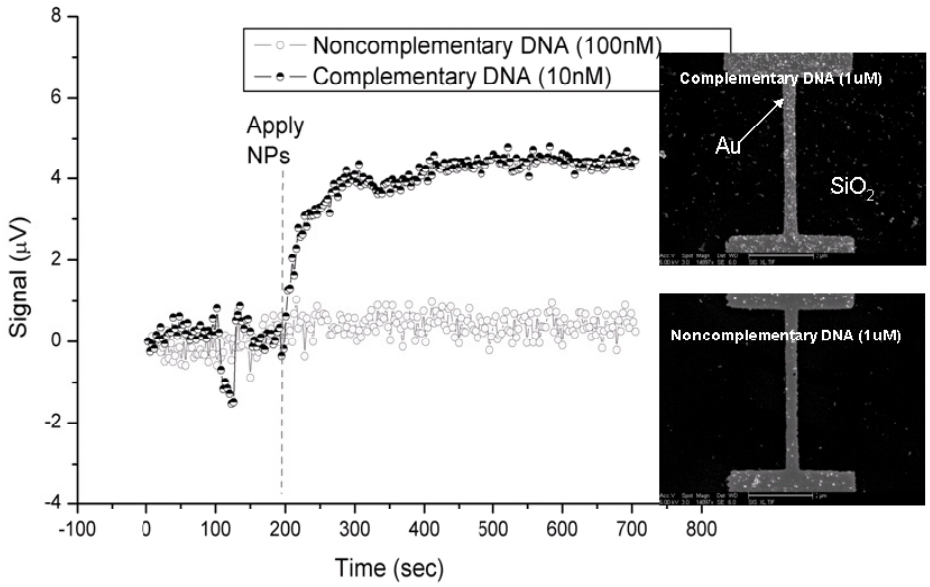


Figure 2.9: Measured signals for 10nM target DNA and 100nM control DNA with SEM images of particle coverage

GMR sensor, as a typical example of magnetic label based biosensing, achieves better sensitivity compared with the electrochemical biosensors, while not needing the expensive and complicated optical setups used for fluorescence biosensors. However, the GMR sensor and the Hall Effect magnetic sensor both need external magnetic field (AC and/or DC) for biasing purpose. This increases the system complexity and the form factor. Also, the mechanical calibration of the fields complicates the sensor module setups. Most importantly, expensive post-processing steps are required to fabricate the sensing structure, i.e., 10-layer spin-valve process for GMR sensor and deep passivation layer etching for Hall Effect sensor, which inevitably increase the total cost and lower the system yield.

2.3 Chapter Summary

In this chapter, we have briefly introduced the biosensor technology. The CMOS process, as a low-cost high-yield implementation tool is demonstrated as a promising platform of biosensor design.

The challenges and opportunities for molecular biosensor are also presented, which requires high sensitivity, good scalability and large parallel processing capabilities. Addressing this advanced sensing application, three CMOS based biosensor schemes, i.e., fluorescence biosensor, electrochemical biosensor and magnetic biosensor are presented with detailed implementation examples.

As a performance comparison among these sensing modalities, the fluorescence biosensor demonstrates the best sensitivity reported so far. However, it requires expensive post-processing to form the optical filters and external bulky excitation laser source. On the other hand, the electrochemical sensors can be realized as a complete compact system with no extra post-processing steps. But it typically suffers sensor sensitivity degradation due to external noise and interference. The magnetic sensor achieves a competitive sensitivity compared with the two modalities, while not needing any optical systems. However, magnetic sensors reported so far still require expensive and complicated post-processing steps to form the functional sensor structures and system. Moreover they need external biasing magnetic fields for sensor operation. These demands essentially defeat their initial purpose of low-cost and small form-factor and therefore significantly limit their applications in practical point-of-care (PoC) molecular sensing.

Chapter 3: Frequency-Shift-Based Magnetic Particle Sensing Scheme

The chapter is organized as follows. After a background introduction in Section 3.1, Section 3.2 introduces the sensing mechanism of our proposed frequency-shift scheme. Section 3.3 shows the line-width narrowing effect for the LC oscillator's frequency detection scheme compared with the impedance function sensing scheme based on the same LC tank. This accounts for the ultra-high sensitivity of our proposed scheme. In order to investigate the sensor signal response with the presence of the magnetic particles, the sensor transducer gain is defined and derived in Section 3.4. In addition, the sensor noise floor will be studied and analyzed in Section 3.5. Finally, Section 3.6 presents the sensor signal-to-noise (SNR), which is shown to be entirely determined by the sensing inductor design. Evaluation on the SNR for a wide range of inductor layouts indicates that a smaller sensing inductor gives a better SNR. Design-limiting factors other than the sensor SNR are also presented and discussed in Section 3.6.

3.1. Introduction

Future Point-of-Care (PoC) molecular-level diagnosis requires advanced biosensing systems that can achieve high sensitivity, ultra-portability and low power consumption, all within a low price-tag. Targeting on-site detection of biomolecules, such as DNA, RNA or protein, this type of system will play a crucial role in a variety of applications such as in-field medical diagnostics, epidemic disease control, biohazard detection and forensic analysis, in the near future.

Traditionally, microarray technology is used to provide both quantitative and qualitative information for biomolecular sensing [15]. However, to ensure their high sensitivity at the pico-molar level, the microarray system relies on optical detection on attached fluorescent labels, which requires bulky and expensive optical devices including multi-wavelength fluorescent microscopes. This fact fundamentally limits the overall size and cost of the microarray system, which eventually makes the technology unsuitable for PoC applications.

On the other hand, magnetic biosensors are proposed as a promising candidate for these PoC applications. The basic mechanism for magnetic detection and its advantages over fluorescence-based schemes has been discussed in detail in Section 2.2.3. However, in spite of the aforementioned apparent advantages for magnetic sensing, magnetic biosensors developed thus far require externally generated strong magnetic biasing fields and/or exotic post-fabrication processes. This still limits the ultimate form-factor of the system, total power consumption and cost, which unfortunately defeats the original purpose to use the magnetic sensing system for PoC applications [6][8][9][18][19].

To address these impediments, we propose a frequency-shift-based magnetic biosensing scheme implemented with an on-chip low-noise LC oscillator, which is fully planar and compatible with standard CMOS processes [20]. Moreover, this scheme can potentially provide high detection sensitivity without using any (electrical or permanent) external magnets, and thus presents itself as an ideal solution for a portable and low power PoC molecular detection system.

3.2. Frequency-Shift-Based Magnetic Sensor Mechanism

The core of our proposed sensing scheme is an on-chip LC resonator. If the test sample contains target molecules, magnetic particles will be immobilized onto the sensor surface after the hybridization procedures shown in Figure 2.7, Section 2.2.3.

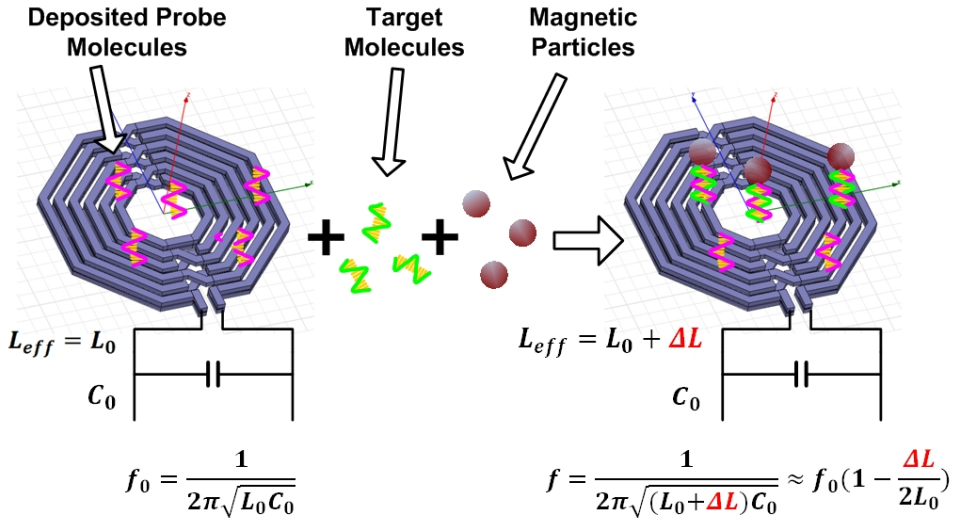


Figure 3.1: Proposed frequency-shift-based magnetic particle sensing scheme

The current through the onchip inductor generates a magnetic field and thus polarizes the magnetic particles which behave as superparamagnetic materials. This polarization then creates a magnetization for those magnetic particles and increases the total magnetic energy in the space. Consequently, this magnetic energy change leads to an increase of the effective inductance of the resonator, which directly results in a resonating frequency down-shift given by,

$$f = \frac{1}{2\pi\sqrt{LC}} = \frac{1}{2\pi\sqrt{(L_0 + \Delta L)C_0}} \approx f_0\left(1 - \frac{\Delta L}{2L_0}\right) \tag{3.1}$$

where L_0 and C_0 are the nominal inductance and capacitance of the onchip LC resonator.

Therefore, by detecting this frequency down-shift, one can infer the existence of the magnetic particles quantitatively. The detailed derivation on how this frequency down shift is related with the presence of the magnetic particles, i.e., the sensor transducer gain, will be derived in Section 3.4.

3.3. Oscillator Based Frequency-Shift Sensing

The mechanism for sensing the presence of magnetic objects by effective inductance change has long been recognized and also widely implemented, e.g., the metal detectors for security check used at airports. However, those traditional implementations are mainly based on sensing the impedance function of the resonator tank directly, which include sensing the change in the amplitude and/or the phase of the impedance function through high-precision circuits, such as a Wheatstone Bridge structure.

However, the impedance function linewidth of an LC resonator, determining the relative amplitude and/or phase shift for a given relative inductance change, is fundamentally limited by the quality factor Q of the resonator. For the on-chip implementation, the quality factor of the resonator is largely dictated by the inductor quality factor, which is very limited, e.g., in the range of 10~20, for a low cost and high substrate conductivity process, such as CMOS. However, the typical relative frequency shift for a single micron-size magnetic particle is in the range of several or sub-ppm (part per million, 10^{-6}). Therefore, this impedance sensing method results in a poor sensitivity, which is not suitable for our magnetic particle detection application.

On the other hand, if an onchip LC oscillator is built by using the same low-Q onchip LC tank, a significant line-width compression effect can be shown for the oscillator's phase

noise profile compared with the corresponding impedance function line-width, shown in Figure 3.2.

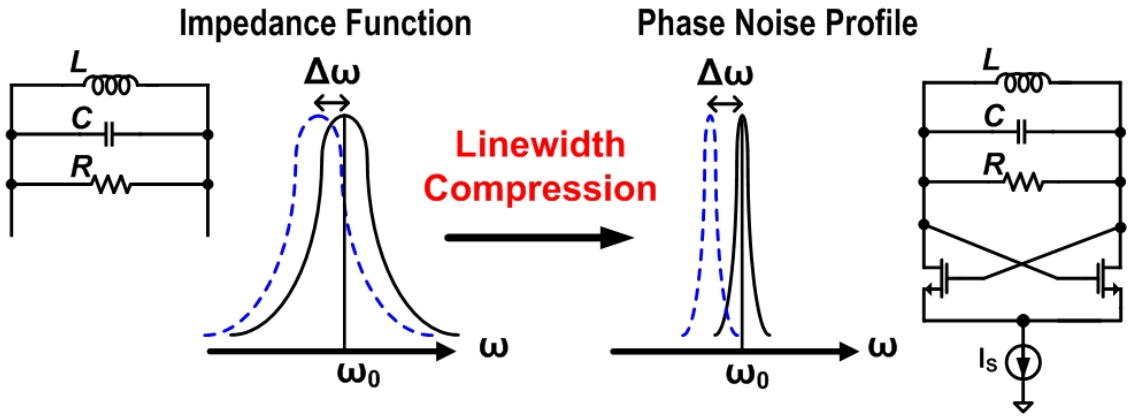


Figure 3.2: Line-width compression effect

This effect is due to the virtual damping phenomena in the active LC oscillators which results in the phase diffusion much slower than the amplitude damping in a normal passive LC resonator by tank dissipation [21]. This line-width narrowing ratio can be estimated as

$$r = \frac{\Delta_{osc}}{\Delta_{res}} \approx \frac{D}{1/2RC} = \frac{2Q}{\omega_0} \cdot D \tag{3.2}$$

where D and ω_0 are the virtual damping rate and the oscillation frequency for the oscillator and Q is the quality factor for the tank.

Moreover, assuming that the $1/f^2$ phase noise is dominant, D can be calculated as

$$D = \frac{L\{\Delta\omega\}}{2} \cdot (\Delta\omega)^2 \tag{3.3}$$

where $L\{\Delta\omega\}$ is the phase noise power spectrum density at the offset frequency of $\Delta\omega$ for the oscillator.

Let us consider an example of a 1GHz LC oscillator, whose phase noise at 600kHz offset is -121dB/Hz. And let us assume the tank Q is 10. Both of the spec numbers are

typical values for CMOS implementation. Based on equation 3.2 and 3.3, D ≈ 5.6Hz, and therefore r ≈ 1.8×10^{-8}.

This significant line-width compression effect suggests that implementing the sensor as an onchip oscillator, whose tank is composed of the sensing inductor and some appropriate capacitors, can result in an ultrasensitive magnetic particle sensor. The narrowed-down phase noise profile thus can easily discern a tiny relative frequency shift, which would be impossible to see with the conventional impedance sensing method with a low Q on-chip LC tank.

The discussion in this section essentially lays the fundamental basis for the high sensitivity performance of our proposed frequency-shift-based magnetic biosensor.

3.4. Sensor Signal Strength

To fully characterize the performance of a sensor system, one needs to model both its signal response and its noise behavior. The sensor signal response with respect to certain amount of sensing targets, i.e., the transducer gain, is derived with an approximate close-form solution in this section, while the sensor noise floor fundamentally limited by the oscillator phase noise will be given in the next section.

With the quasi-static assumption, a current I conducting in the sensor inductor, or any equivalent electromagnetic structure, generates a magnetic field $\overrightarrow{H_{ext}}$ at the coordinate (x, y, z) according to the Biot-Savart law,

$$\overrightarrow{H_{ext}}(x, y, z) = \frac{I}{4\pi}\oint \frac{d\vec{l'} \times \vec{R}}{R^3} \tag{3.4}$$

where the line integration is along the current conducting path and R is the vector pointing from the incremental line vector $d\vec{l'}$ towards the point (x, y, z).

Since most of the commercially available magnetic particles, such as micro/nano magnetic beads, are composed of superparamagnetic nanoparticles dispersed in a nonmagnetic matrix, e.g., polystyrene, its induced magnetization M can be expressed in a Langevin function form (3.5).

$$\vec{M}(\overrightarrow{H_{in}}) = M_{sat}\left[coth\left(\frac{u_0 m_p|\overrightarrow{H_{in}}|}{kT}\right) - \left(\frac{kT}{u_0 m_p H}\right)\right] \cdot \left(\frac{\overrightarrow{H_{in}}}{|\overrightarrow{H_{in}}|}\right) \tag{3.5}$$

where $\overrightarrow{H_{in}}$ is the total magnetic field inside the magnetic bead, instead of the external excitation magnetic field $\overrightarrow{H_{ext}}$. Here, we assume the magnetic material is isotropic. At Curie region, with a high temperature or an excitation magnetic field, the Langevin function can be approximated to the classical formula for magnetization which determines the effective susceptibility χ_{eff} of the superparamagnetic particle as

$$\vec{M}(\overrightarrow{H_{in}}) \approx \frac{M_{sat} u_0 m_p}{3kT}\overrightarrow{H_{in}} = \chi_{eff}\overrightarrow{H_{in}}. \tag{3.6}$$

The relative permeability is thus given by

$$\mu_{eff} = \chi_{eff} + 1. \tag{3.7}$$

Since the polarization is an open magnetic circuit problem, demagnetization effect needs to be taken into consideration to calculate the total magnetization $\vec{M}$ given the external excitation field $\overrightarrow{H_{ext}}$ [22]. By applying the demagnetization factor $\vec{D}$, whose general expression is a 2-dimensional tensor, the magnetic field inside of the bead $\overrightarrow{H_{in}}$ and the externally applied $\overrightarrow{H_{ext}}$ can be related by

$$\overrightarrow{H_{in}} = \overrightarrow{H_{ext}} - \vec{D}\vec{M}. \tag{3.8}$$

In general, the demagnetization factor $\vec{D}$ depends on the geometry of the magnetic subject under the excitation field. Assuming the magnetic bead is of spherical shape and by taking the average magnetic field $\overrightarrow{H_{in}}$ inside of the bead, $\vec{D}$ is reduced to a diagonal matrix of $\begin{bmatrix} D_{xx} & 0 & 0 \\ 0 & D_{yy} & 0 \\ 0 & 0 & D_{zz} \end{bmatrix} = \begin{bmatrix} \frac{1}{3} & 0 & 0 \\ 0 & \frac{1}{3} & 0 \\ 0 & 0 & \frac{1}{3} \end{bmatrix}$. Combining equation (3.6) and (3.8) yields the apparent magnetic susceptibility χ_{app} as,

$$\vec{M}\left(\overrightarrow{H_{in}}\right) = \chi_{eff}\overrightarrow{H_{in}} = \frac{\chi_{eff}}{1 + D_{xx}\chi_{eff}}\overrightarrow{H_{ext}} = \chi_{app}\overrightarrow{H_{ext}}. \tag{3.9}$$

Equation 6 demonstrates two important facts. First, χ_{app} is always smaller than χ_{eff}. Moreover, χ_{app} has its maximum value of $1/D_{xx}$ when χ_{eff} approaches infinity. This means that even if the magnetic particle is entirely made of ferromagnetic material with high susceptibility (χ_{eff} is on the order of hundreds or thousands), χ_{app} still remains as a small value ($1/D_{xx}$) which results in a small magnetic signal. This is essentially the fundamental reason why magnetic particle sensing is challenging.

Assume that the volume of the entire space is V; the volume and the apparent susceptibility for the magnetic particle are V_m and χ_{app}; and the $\vec{H}_{ext}$ for the ith magnetic particle is $\vec{H}_{ext,i}$. Then the total magnetic energy difference for the entire space with or without the presence of the magnetic particles can be calculated as

$$\Delta E_m = \frac{1}{2}\iiint_V \left(\vec{H}\cdot\vec{B}\right)_w dv - \frac{1}{2}\iiint_V \left(\vec{H}\cdot\vec{B}\right)_{wo} dv = \sum_i \frac{1}{2}\iiint_{V_m} \chi_{app}\mu_0 \left\|\vec{H}_{ext,i}\right\|^2 dv. \tag{3.10}$$

Thus, this magnetic energy increase in the space leads to the effective inductance increase [23],

$$\Delta L = \frac{2\Delta E_m}{I^2} = \frac{2\sum_i \frac{1}{2}\iiint_{V_m} \chi_{app}\mu_0 \left\|\vec{H}_{ext,i}\right\|^2 dv}{I^2} \approx \frac{\sum_i \chi_{app}\mu_0 \left\|\vec{H}_{ext,i}\right\|^2 V_m}{I^2}$$

$$= \frac{\left(\chi_{app}\overline{\left\|\vec{B}_{ext}\right\|^2}V_m\right) \times (\#\ of\ Particles)}{\mu_0 I^2}, \qquad (3.11)$$

where $\overline{\left\|\vec{B}_{ext}\right\|^2}$ is defined as the average magnetic flux density for all the magnetic particles.

Consequently, we can define the transducer gain of our magnetic senor as

$$Transducer\ Gain = \frac{Frequency\ Shift\ (\frac{\Delta f}{f_0})}{\#\ of\ Magnetic\ Particles} = \frac{\chi_{app}}{2\mu_0} \cdot \frac{\overline{\|B_{ext}\|^2}}{I^2 L_0} \cdot V_m. \qquad (3.12)$$

This transducer gain equation reveals two important properties of our sensor system. First of all, the sensor sensitivity is proportional to the magnitude square of the averaged excitation magnetic field generated by the sensing inductor. This suggests that the sensor transducer gain has its spatial dependence across the sensor inductor surface. Secondly, the transducer gain is composed of three factors. For a given type of magnetic particle, the first factor determined by the magnetic material property and the last factor of the particle size are both fixed values. However, the factor in the middle is essentially controlled by the sensing inductor, which implies that the sensor inductor design is crucial to achieve a desired transducer gain. These two issues will be discussed in details in the following subsections.

In addition to the magnetic energy based derivation, identical transducer gain results can be achieved through equivalent derivations based on magnetic flux and mutual inductance, which will not be shown here due to the limit of space.

3.4.1 Spatial Variation of the Sensor Transducer Gain

As indicated by equation (3.12), the sensor transducer gain is a function of the excitation magnetic field generated by the sensing inductor, which thereby presents a spatial variation. The transducer gain for a 6-turn inductor with its d_{out} of 140μm, width of 5μm and spacing of 3.5μm for a D=1μm magnetic bead is shown as follows. The passivation layer thickness is 0.9μm. The EM simulation is through Ansoft Maxwell® [24].

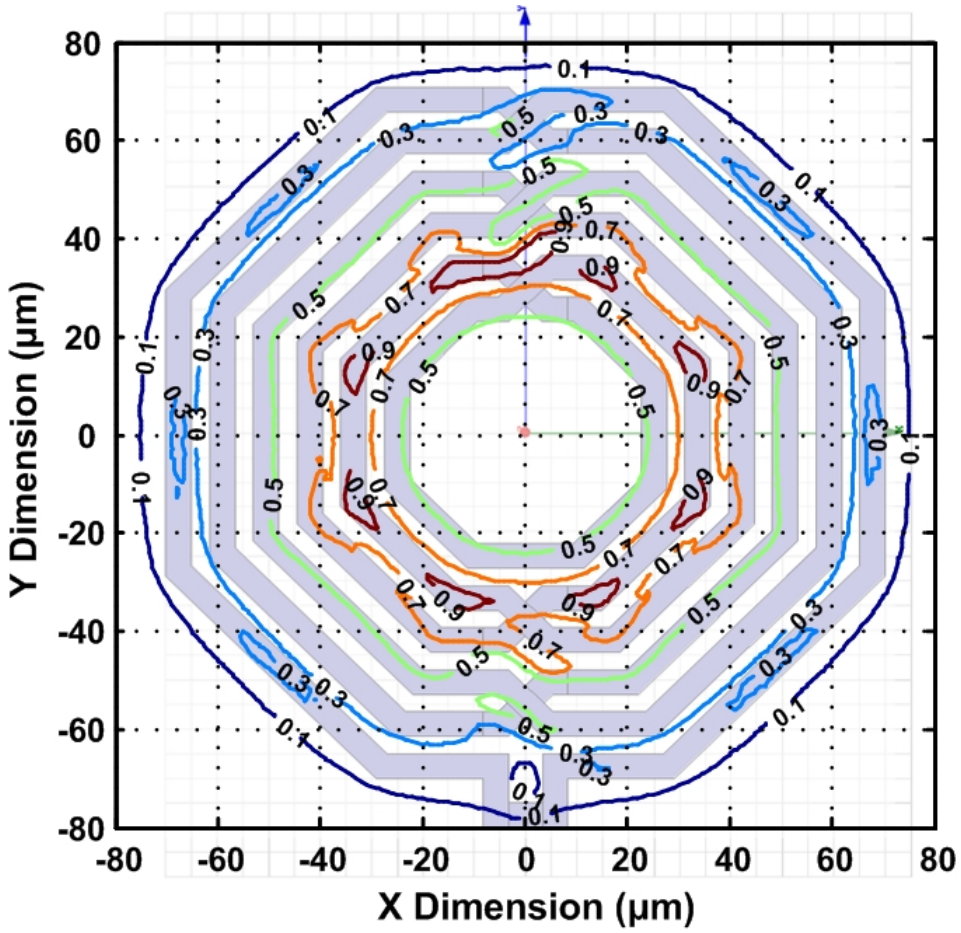


Figure 3.3: Spatial variation of sensor transducer gain ($\Delta f/f_0$ for one 1μm magnetic bead)

Starting from the center of the inductor, the relative large transducer gain value is because of the positive addition of the magnetic fields generated by the multiple turns. This

gain plateau is thus followed by a gain peak due to both the positive magnetic addition and the close spatial proximity to the inductor traces. Next, the sensor inductor presents a lower transducer gain region due to the addition/cancellation of the magnetic fields from different inductor turns. The transducer gain gradually tapers off outside of the sensor inductor region because of the magnetic field decaying.

For a magnetic particle sensor, a spatially homogeneous transducer gain is preferred. This can be achieved by depositing the probe molecules only onto the plateau region or by confining the microfluidic chamber within that region. Moreover, a homogeneous-gain inductor design can take this spatial gain variation effect into account based on the physical intuitive analysis we just presented.

As an example, a stacked spiral inductor, D_{out} of 60μm, is designed to achieve a much more homogeneous sensor transducer gain above the same 0.9μm passivation layer as the example shown in Figure 3.3. The multiple metal traces in different layers are engineering to imitate a half spherical shape. Note that the transducer gain shown in Figure 3.5 is significantly larger than the gain shown in the Figure 3.3. This transducer gain increase due to a smaller inductor footprint will be discussed in Section 3.4.2.

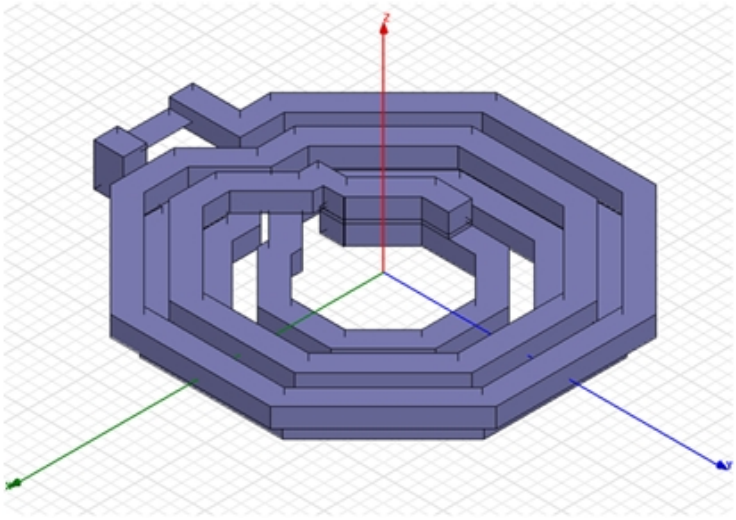


Figure 3.4: 3D view of the stacked inductor layout

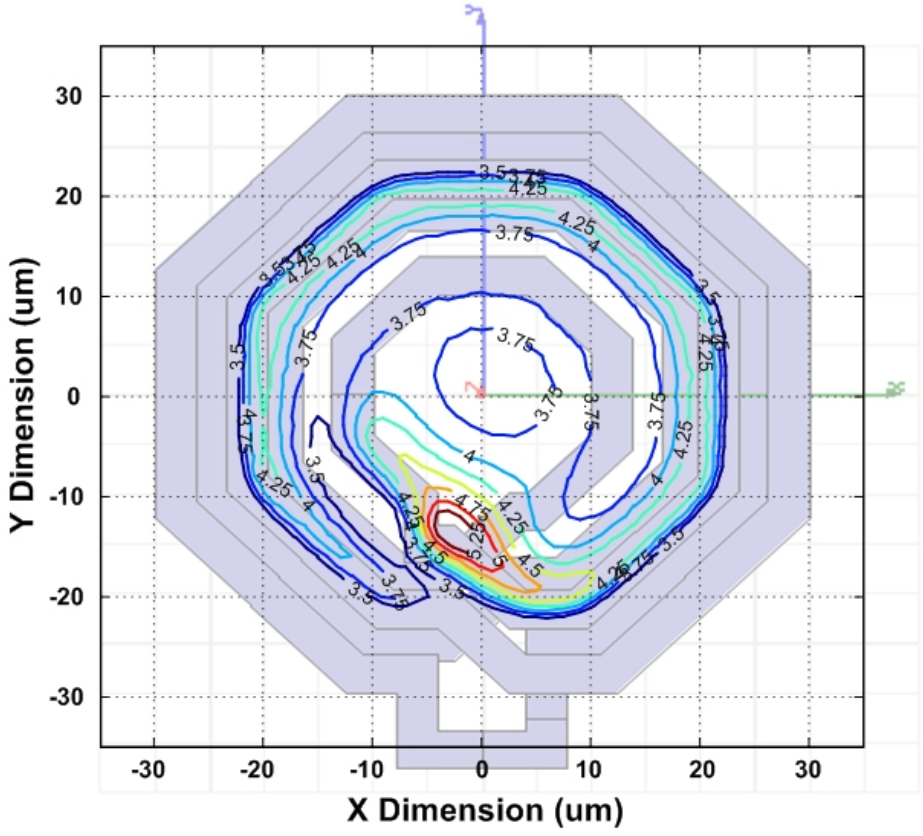


Figure 3.5: Spatial variation of sensor transducer gain for the stack inductor ($\Delta f/f_0$ for one 1μm magnetic bead)

On the other hand, this spatially varying transducer gain also provides extra location information of the present magnetic particles. For example, based on this idea, a "magnetic microscope" may be implemented by a sensing inductor layout with strong spatial gain variations, which detects the attached magnetic particle as well as determining its location.

3.4.2 Sensor Inductor Scaling Rule

As we mentioned in Section 3.4, for the transducer gain expression of equation (3.12), only the middle is within the designer's control, which is also entirely dependent on the sensing inductor design.

In order to achieve some intuitive design insight for the relationship between the transducer gain and the inductor layout, we can assume isomorphic scaling with a scaling factor l on a given inductor geometry. Based on equation (3.12), we can calculate the average transducer gain in the entire effective sensing volume by

$$\overline{Transducer\ Gain} = k_1 \cdot \frac{1}{I^2 L_0} \cdot \overline{\|B_{ext}\|^2} = k_1 \cdot \frac{1}{I^2 L_0} \cdot \frac{\iiint_{V_{sense}} \|B_{ext}\|^2 dv}{V_{sense}}, \quad (3.13)$$

where the effective sensing volume for the given inductor V_{sense} is chosen as the region where the magnetic field magnitude does not decay significantly with respect to the peak magnetic field and the factor k_1 represents the product of $\frac{\chi_{app}}{2\mu_0} \cdot V_m$, which is independent of the inductor design.

Note that for a given spiral inductor, in the x-y plane the magnetic field generally decreases significantly outside of the inductor, while in the z direction the magnetic field decreases sharply after certain distance which is roughly proportional to the inductor size.

Therefore, the V_{sense} is proportional to l^3. Also, due to the isomorphic scaling, we can approximate the magnetic energy stored in the volume V_{sense} is proportional to the total magnetic energy in space with some proportion constant k_2, as

$$\frac{1}{2\mu_0}\iiint_{V_{sense}} \|B_{ext}\|^2 dV \approx k_2 \cdot \frac{1}{2\mu_0}\iiint \|B_{ext}\|^2 dv. \tag{3.14}$$

Therefore, the averaged transducer gain across the sensing volume can be further simplified as,

$$\overline{Transducer\ Gain} \approx k_1 k_2 \mu_0 \cdot \frac{1}{\frac{1}{2}I^2 L_0} \cdot \frac{\frac{1}{2\mu_0}\iiint \|B_{ext}\|^2 dv}{V_{sense}} = \frac{k_1 k_2 \mu_0}{V_{sense}} \propto \frac{1}{l^3}, \tag{3.15}$$

where l is the scaling factor in the isomorphic scaling. The above equation indicates that the smaller the sensing inductor, the larger the average sensor transducer gain. Moreover, the gain is roughly inversely proportional to the 3rd power of the scaling factor.

More accurate results on the averaged sensor transducer gain can be computed through electromagnetic (EM) simulation for different sensing inductor geometry. The figure shown below is the average sensor transducer gain (ΔL/L per 1μm magnetic bead) with respect to different D_{out} and its proportional trace width for a one turn symmetric inductor, which directly confirms our derivations. The upper limit and the lower limit consider the numerical errors in the EM simulations.

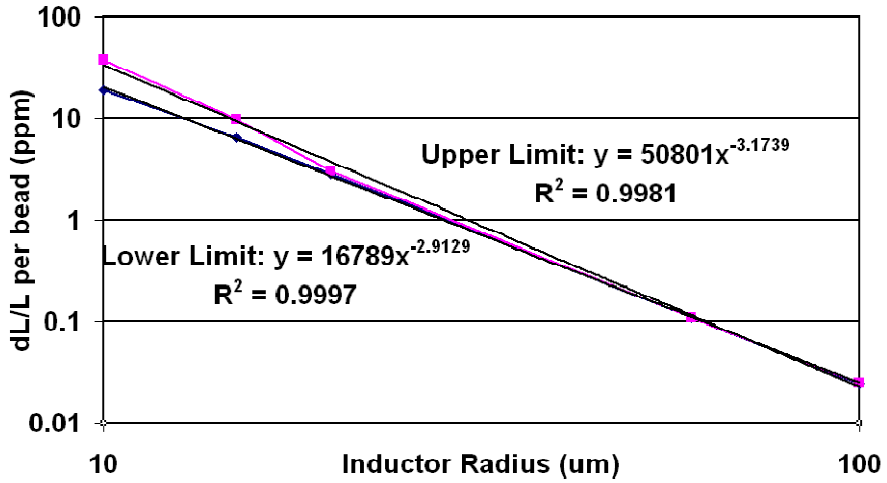


Figure 3.6: Averaged sensor transducer gain for different inductor sizes (ΔL/L per 1μm bead)

3.5 Sensor Noise Floor

The transducer gain introduced in the previous sections models how the sensor will respond to the presence of the sensing targets, i.e., magnetic particles for our study. On the other hand, the sensor's sensing limit is also determined by the sensor noise floor.

For our frequency-shift magnetic sensing scheme, the sensor noise sources can be classified into two categories. One is the intrinsic noise, which is mainly due to the phase noise of the sensing oscillator. This phase noise provides a limit on the total accumulated jitter which sets the uncertainty in our frequency counting result. The other type of noise is the extrinsic noise, which for our case accounts for the frequency drifting due to environmental effects, such as temperature variation and supply noise. Both types of the noise sources will be discussed in the following sections.

3.5.1 Oscillator Phase Noise

Phase noise represents the frequency instability of the oscillator. In general, a free-running oscillator's phase noise is composed of two parts. The $1/f^3$ phase noise at low offset frequencies (typically below kHz range) is mainly due to the up-conversion of the flicker noise power, while the up-conversion of the white noise results in the $1/f^2$ phase noise [25], shown as follows.

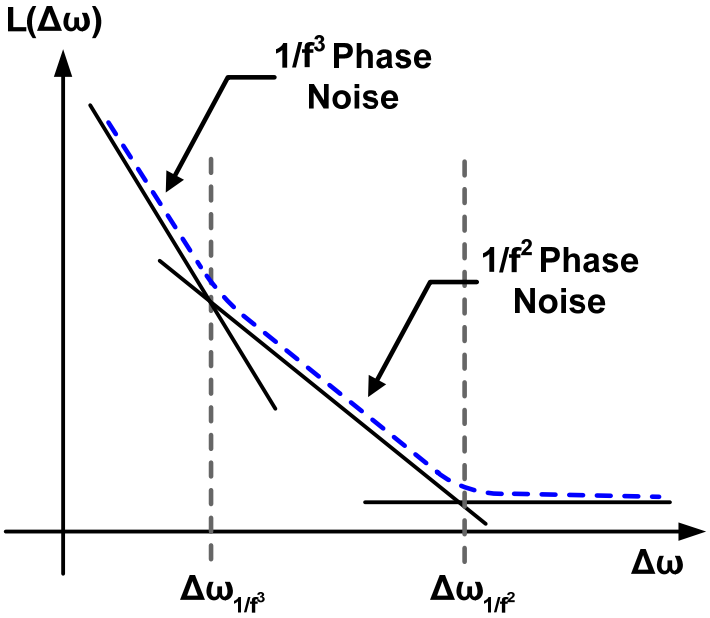


Figure 3.7: Typical oscillator phase noise profile

Generally, the oscillator frequency is determined through frequency counting, which registers the number of transitions of the oscillation waveform within a given time window. Thus the frequency measurement uncertainty is directly determined by the accumulated jitter within this measurement time as

$$\sigma^2_{\frac{\Delta f}{f_0}} = \frac{(\Delta f)^2}{f_0^2} = \frac{(\Delta T)^2}{T^2} = \frac{\sigma_T^2}{T^2} \tag{3.16}$$

where f_0 is the nominal frequency and T is the total time for the counting window. Assuming the phase noise $\phi(t)$experiences a stationary process, the frequency uncertainty can be derived as

$$\sigma^2_{\frac{\Delta f}{f_0}} = \frac{\sigma_T^2}{T^2} = \frac{1}{T^2}\cdot\frac{1}{\omega_0^2}E\{[\phi(t+T)-\phi(t)]^2\} = \frac{1}{T^2}\cdot\frac{2}{\omega_0^2}\big[R_\phi(0)-R_\phi(T)\big]. \quad (3.17)$$

By Wiener-Khinchin theorem, the above quantity can be related with the phase noise of the oscillator as

$$\sigma^2_{\frac{\Delta f}{f_0}} = \frac{2}{\omega_0^2T^2}\big[R_\phi(0)-R_\phi(T)\big] = \frac{2}{\pi\omega_0^2T^2}\int_{-\infty}^{+\infty} S_\phi(\omega')\big[1-e^{j\omega' T}\big]d\omega'$$

$$= \frac{4}{\pi\omega_0^2T^2}\int_0^{+\infty} S_\phi(\omega')\sin^2\frac{\omega' T}{2}d\omega', \quad (3.18)$$

where the $S_\phi(\omega)$ is the phase noise profile. These jitter/phase-noise equations will be revisited in Chapter 5 when low noise techniques are introduced.

We can further relate this noise quantity with the inductor design. Assuming a fixed biasing current density for the cross-coupled cores, at the maximum tank amplitude (limited by supply V_{DD}), the biasing current I_d, the transistor width W, and the current noise power spectrum density are given by

$$I_d \propto W \propto \frac{V_{DD}}{R_{p,tank}} = \frac{V_{DD}}{\omega_0 LQ} \propto i_n^2(\omega'), \quad (3.19)$$

where $i_n^2(\omega')$ denotes the transistor output current noise power spectrum density at the frequency of ω'. Note that at a low offset frequency ω', $i_n^2(\omega')$ is dominated by flicker noise. But at a large offset frequency, $i_n^2(\omega')$ is dominated by the channel thermal noise.

At a low offset frequency $\omega_0{}'$ where the up-converted flicker noise power is dominant, the oscillator phase noise profile can be determined by

$$S_\phi(\omega')_{\omega'=\omega_0'} = \frac{A_0^2}{q_{max}^2} \cdot \frac{i_n^2(\omega_0{}')}{2\omega_0{}'^2} \propto \frac{A_0^2}{V_{DD}^2 C^2} \cdot \frac{V_{DD}}{\omega_0 LQ} \propto \frac{L}{Q} \tag{3.20}$$

where $A_0$is the DC term of the oscillator's impulse sensitivity function (ISF) $\Gamma(t)$ [26]. This leads to the relation between the sensor noise $\sigma_{\frac{\Delta f}{f_0}}$ and the inductor L and Q as

$$\sigma_{\frac{\Delta f}{f_0}}^2 = \frac{4}{\pi\omega_0^2 T^2} \int_{-\infty}^{+\infty} S_\phi(\omega') \sin^2 \frac{\omega' T}{2} d\omega' \propto \frac{L}{Q}. \tag{3.21}$$

This result suggests that the sensor noise floor (due to the oscillator phase noise) also depends on the sensing inductor design.

3.5.2 Temperature and Supply Variations

As we mentioned in Section 3.5, the non-ideal operation environment of the sensor oscillator also leads to non-negligible oscillation frequency shift. Generally, this environmental effect is dominated by variations in the chip temperature and the supply voltage. Since these variations are not known a priori, they can be approximately treated as low-frequency noise/drifting for the oscillation frequency.

A generic complimentary cross-coupled LC oscillator is shown as follows.

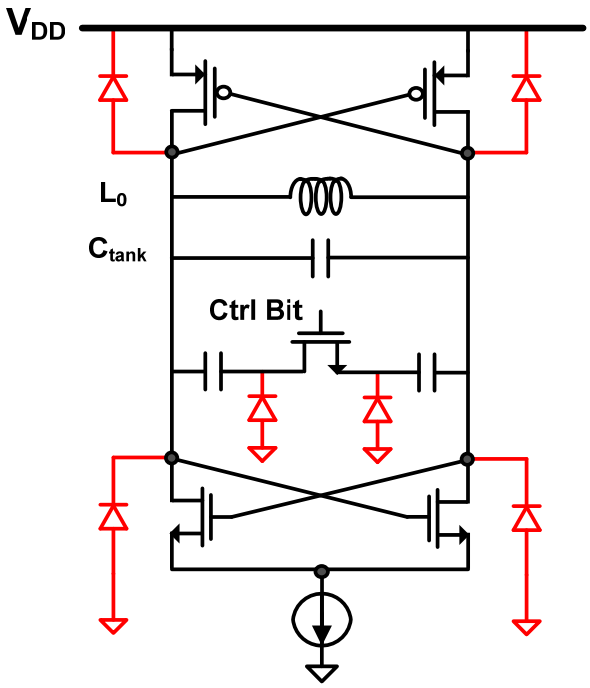


Figure 3.8: A generic complimentary cross-coupled LC oscillator schematic with parasitic junction diodes highlighted

The circuit elements which to the first order determine the oscillation frequency are the tank inductor (L_0), tank capacitor (C_{tank}), and the parasitic junction capacitors (C_j). The frequency-temperature sensitivity formula can be derived as

$$\frac{1}{f}\frac{\partial f}{\partial T} = -\frac{1}{2}\left(\frac{1}{L_0}\frac{\partial L_0}{\partial T} + \frac{C_{tank}}{C_{tank}+C_j}\cdot\frac{1}{C_{tank}}\frac{\partial C_{tank}}{\partial T} + \frac{C_j}{C_{tank}+C_j}\cdot\frac{1}{C_j}\frac{\partial C_j}{\partial T}\right). \quad (3.22)$$

Although the inductor's loss increases with the temperature, the inductance value generally has a negligible temperature dependency. The same temperature insensitivity also exists for the tank capacitors, since the temperature variation only causes a minor change for the relative permittivity of the dielectric material. However, the junction capacitances vary significantly with respect to the temperature, shown as

$$\frac{1}{C_j}\frac{\partial C_j}{\partial T} = \frac{1}{C_j}\frac{\partial}{\partial T}\left(\frac{C_{j0}}{\sqrt[n]{1-\frac{V_{Bias}}{\psi_0(T)}}}\right) = -\frac{1}{n}\frac{\frac{V_{Bias}}{\psi_0(T)}}{\psi_0(T)-V_{Bias}}\frac{\partial \psi_0(T)}{\partial T}, \tag{3.23}$$

where n is the doping profile coefficient, $\psi_0(T)$ is the built-in potential and V_{Bias} is the biasing voltage across the junction [27]. Depending on these parameters, a typical temperature coefficient of the junction capacitance is on the range of 250ppm~350ppm/°C. Based on the relative weight between the tank capacitor and the total junction capacitors, the overall temperature coefficient for the oscillation frequency, shown in equation (3.22) is thus typically smaller than tens of ppm/°C.

On the other hand, in terms of supply voltage variation, the frequency drift due to the parasitic junction capacitance change will be dominant. This is because the diode biasing voltage is directly related with the supply voltage [27], whose voltage sensitivity can be shown as

$$\frac{1}{C_j}\frac{\partial C_j}{\partial V_{Bias}} = \frac{1}{C_j}\frac{\partial}{\partial V_{Bias}}\left(\frac{C_{j0}}{\sqrt[n]{1-\frac{V_{Bias}}{\psi_0(T)}}}\right) = \frac{1}{n}\cdot\frac{1}{\psi_0(T)-V_{Bias}}. \tag{3.24}$$

So both the temperature and the supply variations have a large impact on the frequency stability, which sets a limit on the sensor sensitivity besides the oscillator phase noise. However, as will be shown in Chapter 5, these environmentally related factors can be largely suppressed in a practical sensor embodiment by adopting certain implementation schemes, so the fundamental determining factor for the noise floor is the phase noise of the oscillator. Therefore, for the noise discussion in the following part of the thesis, we will consider the intrinsic oscillator phase noise only.

3.6 Sensing Inductor Optimization for Maximizing Sensor Signal-to-Noise Ratio (SNR)

With the analysis results for both the sensor signal strength (transducer gain) and the sensor noise floor, we can therefore arrive at the sensor signal-to-noise ratio (SNR) as

$$SNR = \frac{\frac{\Delta f}{f_0}}{\sigma_{\frac{\Delta f}{f_0}}} \propto \frac{\overline{\|B_{ext}\|^2}}{I^2} \cdot \frac{1}{L_0} \cdot \left(\frac{Q}{L_0}\right)^{\frac{1}{2}}, \tag{3.25}$$

where $\overline{\|B_{ext}\|^2}$ is the average magnetic flux density across the inductor sensing area.

Thus, the sensor SNR is fully determined by the inductor design, which needs to be carefully optimized. Moreover, the inductor layout also limits the basic footprint of a sensor cell. Then choosing the inductor geometry is the key step in designing the frequency-shift based magnetic sensor.

As an example, the averaged relative SNR for a 6-turn symmetric inductor (separation=3um) using dual thick coppers is shown in Figure 3.9. The SNR at those black circled points are zero because those inductor configurations violate DRC rules.

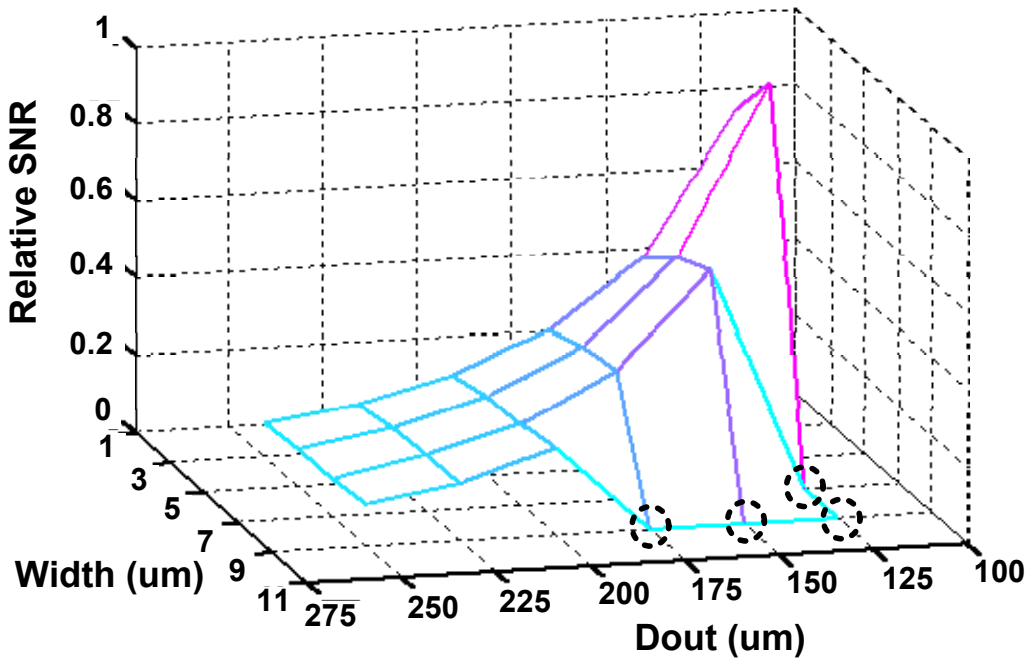


Figure 3.9: Averaged relative SNR for different inductors

As we can see, for a certain D_{out}, the SNR does not vary significantly with its width. This is because although inductors with a larger width will have a better Q and a smaller L, it has a smaller averaged $\overline{\|B_{ext}\|^2}$ which leads to an overall relatively constant SNR. On the other hand, for the same turn width, a larger D_{out} gives a much lower SNR. This is due to the fact that at a larger outer diameter, a smaller averaged $\overline{\|B_{ext}\|^2}$ together with a larger inductance value result in a significantly degraded SNR.

Figure 3.9 indicates that at the same operating frequency a smaller inductor size gives a better SNR. However, in reality an inductor with a very small size is undesirable for implementation. This is because the interconnections start to contribute non-negligible resistance and parasitic inductance, which degrade the sensor SNR according to equation 3.25. Furthermore, to sustain the same voltage amplitude (at a fixed V_{DD}), a low impedance

tank (for a smaller inductor) will conduct a large current, which may result in magnetic saturation for the target particles. This also decreases the actual sensor signal shown in equation 3.12. In addition, for a small size inductor (high peak-Q frequency), a high operating frequency increases its Q, and thus its SNR. But the aforementioned high frequency magnetic loss eventually degrades the magnetic signal and limits the operation frequency to be 1GHz [29] [30]. Besides the SNR consideration, other constrains also limit the design space, including sensor power consumption, hot electron degradation, sensor cell pixel size requirement (foot-print), etc.

3.7 Chapter Summary

In this chapter, we introduce an Effective-Inductance-Change based magnetic particle sensing scheme whose core part is an on-chip LC resonator. Utilizing the phase noise line-width narrowing effect in LC oscillators, we propose a novel sensing method which utilizes the relative frequency shift of the LC resonator based oscillator at the presence of the magnetic particles. This sensing approach can potentially achieve a significantly better sensitivity compared with conventional LC tank impedance sensing method, which is fundamentally limited by the quality factor of the tank.

To understand the sensor operating mechanism, we have derived a close-form formula for the sensor transducer gain based on the underlying physics, which captures the sensor signal behavior (relative frequency shift) when the magnetic particles are presented. Based on this transducer gain formula, we have proposed an approximate scaling rule as a guide line for the sensor inductor design. On the other hand, the sensor noise floor has been carefully investigated, including the intrinsic noise determined by the oscillator phase noise

and the extrinsic (environmental) noise dominated by temperature and supply voltage variations. Further derivation on the oscillator phase noise relates the intrinsic noise floor with the inductor design parameters.

Finally, the signal-to-noise (SNR) ratio of the sensor scheme is derived and evaluated for a wide range of inductor design geometry through electromagnetic (EM) simulation. The results suggest that a smaller sensing inductor gives a better SNR value. In practice, the sensing inductor design is also limited by other design-limiting-factors, which are extensively discussed in this chapter.

Chapter 4: Frequency-Shift Magnetic Particle Sensor Implementation

4.1. Introduction

In Chapter 3, we have proposed the frequency-shift based magnetic sensing scheme and studied its mechanism including the sensor SNR and sensor design scaling rule. In this chapter, we will present two sensor implementation examples based for this scheme.

Section 4.2 demonstrates our first version discrete implementation. The sensor circuit is composed of high-frequency low noise bipolar transistors and discrete capacitors. The spiral inductors for desired sensing functionalities are patterned on the thin film circuit board directly. Several experiments are performed, whose measurement results are compared with simulations to validate our modeling analyses shown in Chapter 3. Finally, the limitations of this design implementation are summarized.

In order to improve the sensor performance significantly, we designed and implemented our second version sensor embodiment presented in Section 4.3. The sensor is integrated on a standard 130nm CMOS technology with 8 differential sensing cells on chip. To overcome the impediments encountered in the first sensor embodiment, advanced schemes such as differential sensing and temperature controlling are implemented in this integrated version. The sensor has also been integrated with low cost Polydimethylsiloxane (PDMS) microfluidic device for sample delivery. The measurement results shows that this sensor implementation achieves an ultra-high sensitivity of one single magnetic particle (D=1um) and is able to reliably detect 1n-Molar actual DNA sample. This is the best

sensitivity for CMOS magnetic sensor reported so far. Furthermore, a significant SNR improvement can be achieved by averaging the data samples.

Finally, a summary will conclude this chapter in Section 4.4.

4.2. A Discrete Magnetic Particle Sensor based on Thin-Film Technology

To verify the proposed sensing scheme, the frequency-shift magnetic sensor is first implemented as a discrete design. The circuit board is based on a one-layer thin-film technology with gold plated traces with a line-width resolution of 20μm [31]. The low loss substrate is made by alumina. These specs are essential to implement a small footprint spiral inductors with a high quality factor. The thin-film board is designed in Protel® and shown as follows.

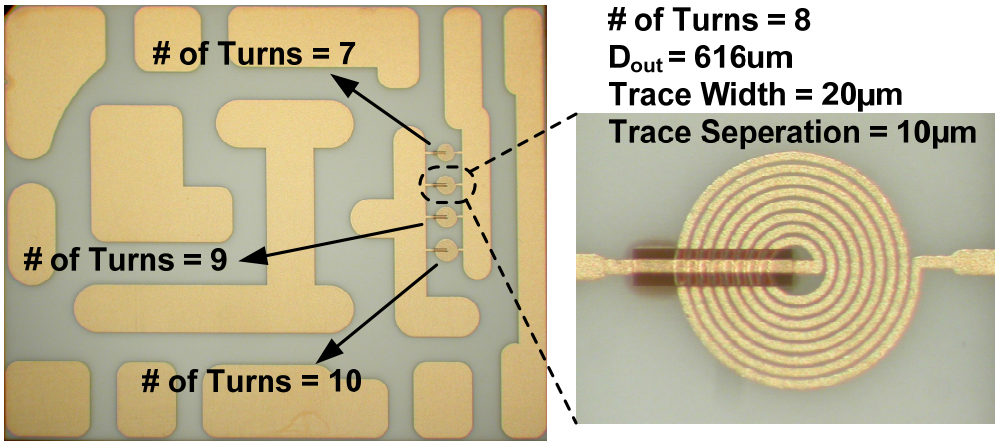


Figure 4.1: The thin-film board used in the 1st version magnetic sensor design

Note that the four inductors with different sizes are implemented in parallel on the board. During the experiment, three of them will be unused and trimmed-off by laser. An open space has been left on the right side of thin-film board which is reserved for sample

delivery and subsequent microfluidic channel integration. The inductors specs are shown in the following table.

Table 4.1 SPECS FOR THE FOUR THIN-FILM INDUCTORS				
	Inductor 1	Inductor 2	Inductor 3	Inductor 4
Outer Diameter (μm)	556	616	676	736
Trace Width/Separation (μm)	20/10	20/10	20/10	20/10
Number of Turns	7	8	9	10
Inductance (nH)	17.2	23.5	31.1	40.2
Quality Factor at (100MHz)	12.4	13.6	14.8	16.0

4.2.1 Sensor Circuit

The sensing oscillator, operating at a nominal frequency of 100MHz, adopts a Coplitts topology for an improved phase noise performance, shown as the following figure with all the main circuit elements denoted.

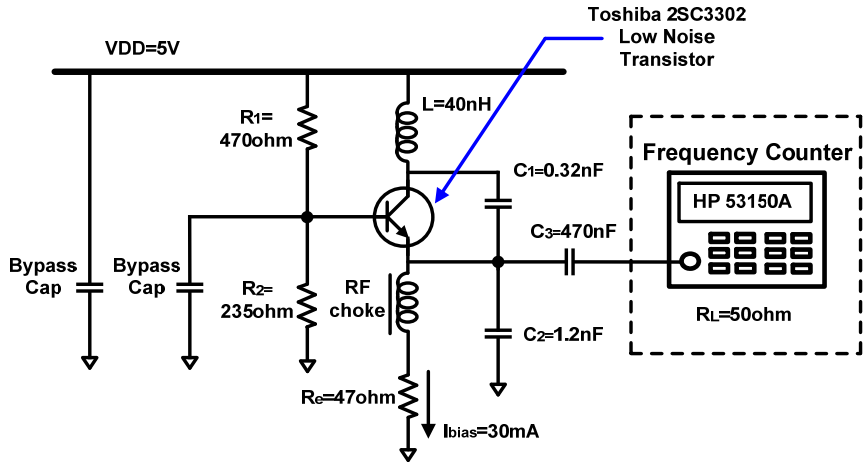


Figure 4.2: The discrete sensor oscillator schematic

4.2.2 Measurement Results

Two measurements are performed for this discrete type magnetic sensor. The first measurement is based on dry experiments which sense the deposited magnetic particles (Dynabeads® MyOne, D=1μm [16]) after the buffer solution evaporates. The second measurement is based on wet experiments, which test the magnetic particle samples in aqueous buffer solutions. In the following part of this section, the two measurements together with their results will be presented.

In the dry experiment, the thin-film sensor oscillator circuit has been coated with parylene (thickness of 3μm). This provides electrical isolation between the deposited magnetic particle samples with the thin-film sensing inductors. Then, magnetic beads solutions with different concentration are deposited on the sensing inductor via pipette (Rainin, D20). Due to the operation of the oscillator, the temperature on the sensing inductor surface increases and leads to a fast evaporation of the water content in the bead samples. A waiting time of 15 minutes is adopted to achieve the thermal steady state. The oscillator's oscillation frequency is counted by an off-the-shelf frequency counter HP 53150A. The summary on the experiment procedures is listed in Figure 4.3 and the sensing inductor (Inductor 2) with deposited magnetic particles is shown in Figure 4.4.

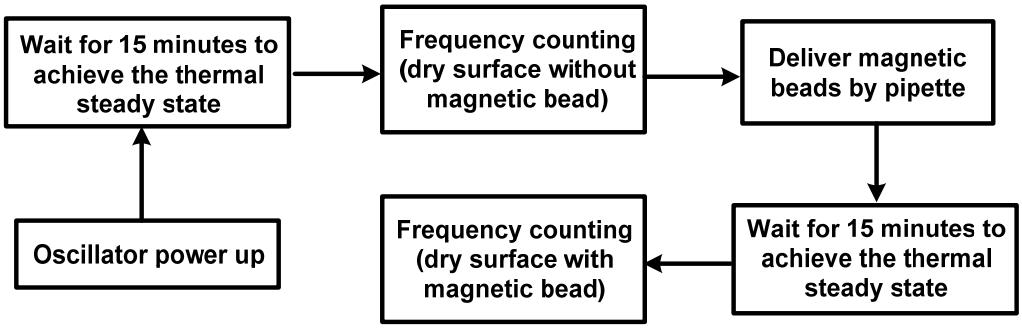


Figure 4.3: The experiment procedures for the dry experiment

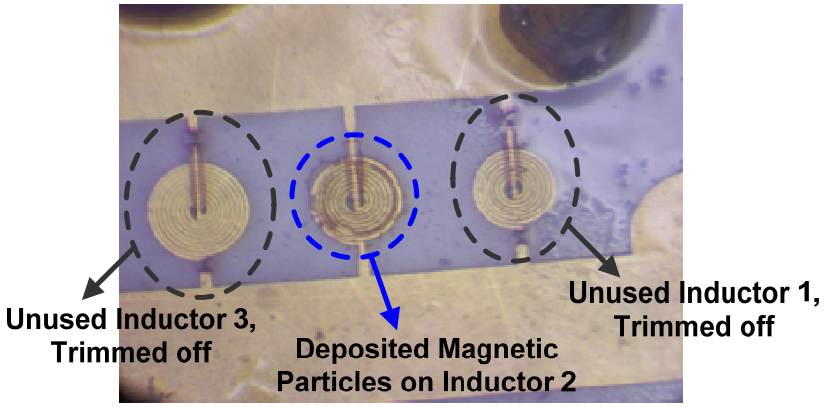


Figure 4.4: Sensing inductors deposited with target magnetic particles

Typical measurement results of frequency counting with respect to time are shown in Figure 4.5. For this frequency plot, the total number of magnetic beads is around 45000. Traces with different colors represent different measurements. Before applying the magnetic beads, the oscillator frequency is measured as the baseline frequency. After applying the magnetic beads, the frequency is again recorded. The difference in frequency indicates the detected signal due to the magnetic particles. The experiment is then repeated and a result summary is shown in Figure 4.6.

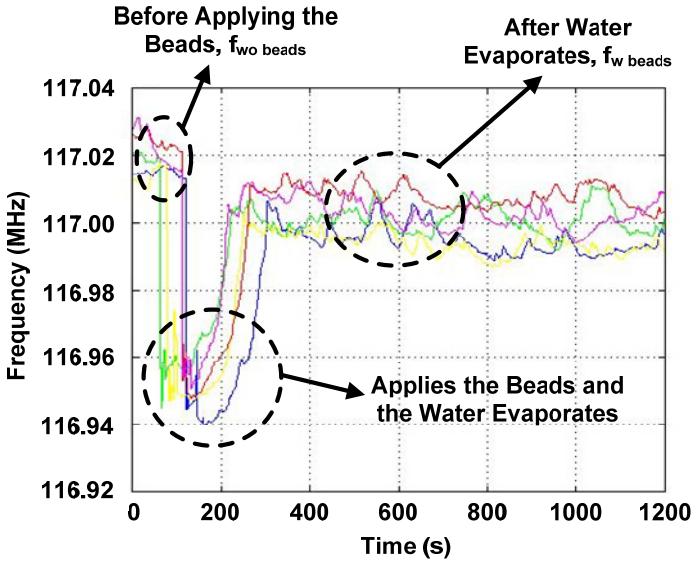


Figure 4.5: Typical frequency counting results with respect to time

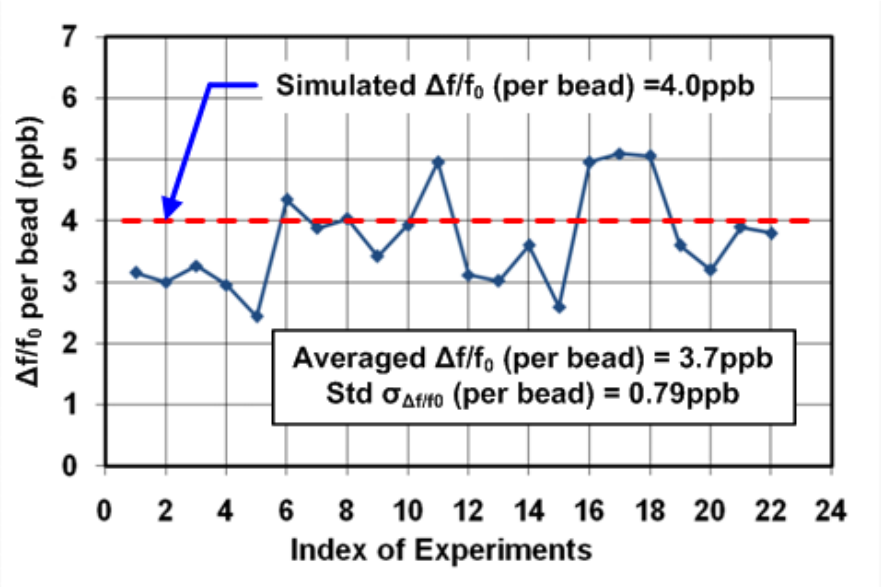


Figure 4.6: Measurement results summary on the dry experiment

The averaged $\Delta f/f_0$ per bead from the measurement results is 3.7ppb, which is in close agreement with the averaged $\Delta f/f_0$ per bead of 4.0ppb based on electromagnetic simulation. Although the error bar of 0.79ppb is not small, mostly due to the inaccuracy of sample delivery and oscillation frequency jittering due to thermal variation, the experiment still verifies that our sensor scheme is practically valid with the dry magnetic particle samples.

The second experiment, as the wet experiment, is to verify the sensor's functionality with magnetic particle samples in aqueous condition. As shown in the previous experiment, the sensor surface temperature increase due to the oscillator operation results in fast water evaporation. Therefore, a microfluidic channel must be used to prevent this evaporation and maintain an aqueous experiment condition. Polydimethylsiloxane (PDMS) is used to form the microfluidic structure, which contains a fluidic channel, a sensing chamber and a pair of pneumatic control valves, shown in Figure 4.7. DI (de-ionized) water is used as the buffer to dilute the magnetic particle samples and to clean the sensor surface. The sensor

module together with the PDMS microfluidic device is shown in Figure 4.8. The measurement procedure is shown in Figure 4.9.

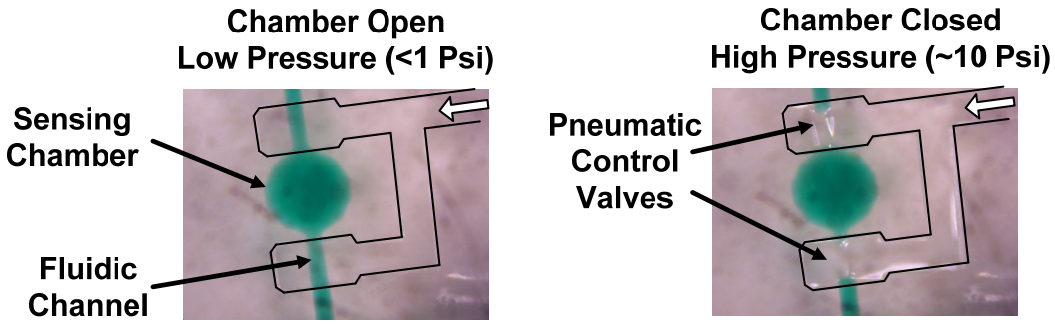


Figure 4.7: PDMS microfluidic structure

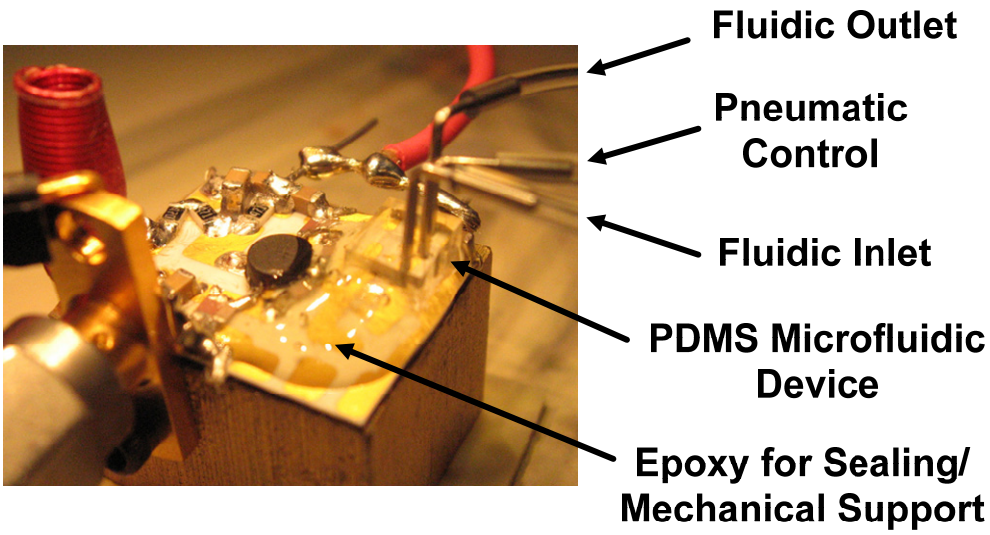


Figure 4.8: Sensor module with the PDMS microfluidic device

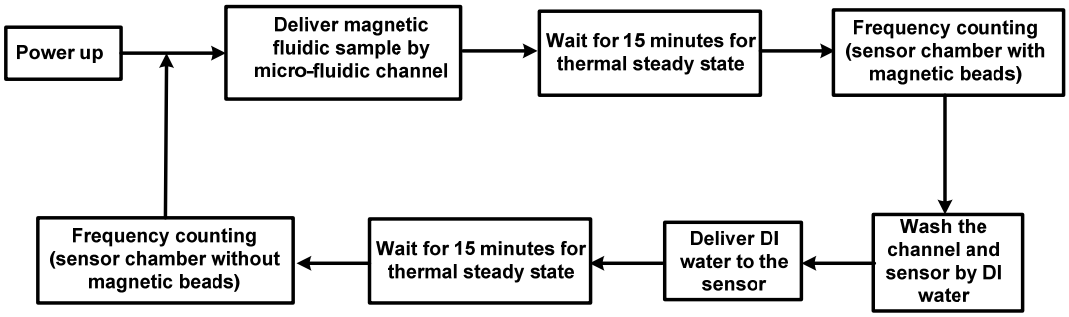


Figure 4.9: Measurement procedures for the aqueous experiment

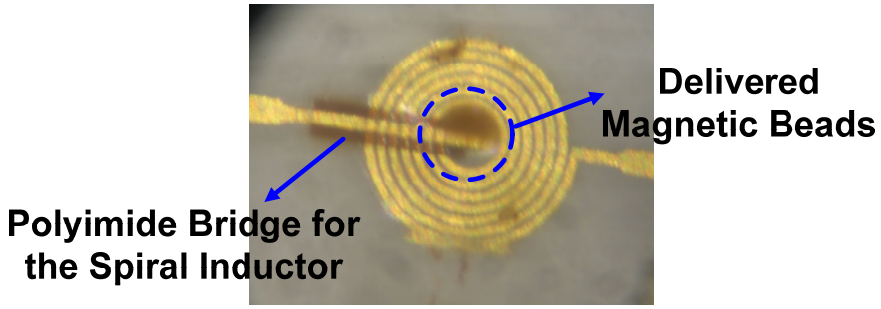


Figure 4.10: The delivered magnetic beads on the sensor inductor (Inductor 1)

A picture of magnetic beads delivered on the sensor inductor (Inductor 1) is shown in Figure 4.10. A summary on the measurement result is shown in Figure 4.11. As we can see the average $\Delta f/f_0$ per bead from the measurement results is 5.2ppb while the $\Delta f/f_0$ per bead by simulation is 4.8ppb. The increase on the average frequency shift per bead is due to the fact that the delivered magnetic particles tend to aggregate to the positions where a large magnetic field is present, shown in Figure 4.10. Overall, this aqueous experiment still shows that our sensor scheme is practically viable with the magnetic particle carried by aqueous buffer solutions.

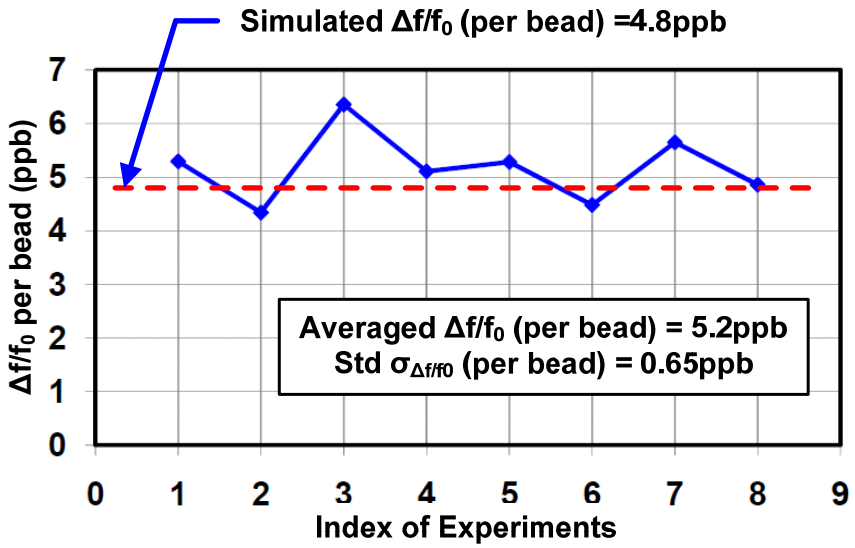


Figure 4.11: Measurement results for summary on the aqueous experiment

In summary, this discrete implementation of our proposed sensor scheme verifies that by using the relative frequency shift (inductance change) of the LC resonant tank, micron-size magnetic particles are indeed detectable.

4.2.3 Limitation and Potential Improvement

As presented in the previous sections, the discrete design verifies the functionality of our frequency-shift based magnetic sensor scheme. However, there are following limitations on this particular implementation.

First of all, the sensor transducer gain is low. As we have shown in the measurements, for Dynabeads® MyOne (D=1μm), the signal strength is typically 3~5ppb/bead. This is not a surprising result, since low sensitivity for large inductor footprint is predicted by our analysis in Chapter 3. However, due to the fact that the inductor trace width is limited by the thin-film process and the parasitic inductance for connecting the tank is prominent in the discrete design, inductors with large peripherals (large inductance) are still used for this implementation. Based on our studies in Chapter 3, the way to significantly increase the signal strength is by decreasing the size of the sensing inductor and therefore operating at a higher oscillation frequency. This is practically suitable for an integrated implementation.

Moreover, the sensor oscillator's frequency experiences a significant low-frequency drift, shown in Figure 4.5. This could be due to the supply noise and temperature variation, etc. Therefore, a differential sensing scheme can be implemented as a well-matched sensing oscillator pair, i.e., a sensing oscillator and a reference oscillator. If these two oscillators share the same operating environment, such as the supply, the biasing, the ground reference and the temperature, taking the frequency reading difference between the

two oscillators essentially suppresses the low-frequency drift due to all these environmental factors. This differential sensor scheme also prefers an integrated implementation, where matching within the sensor pair can be reliably achieved.

Furthermore, the temperature of this sensing oscillator may vary during the sensing operation, particularly when samples are delivered onto the sensor surface. This leads to significant frequency drift due to the non-zero temperature coefficient of the oscillator discussed in Chapter 3. In order to register the oscillation frequency at the thermal steady-state for a fair comparison between the target samples and the control samples, a long waiting time is unavoidable. Therefore, to facilitate fast data acquisition, a temperature controller implementation is required to actively stabilize the temperature of the sensor.

4.3. A Fully Integrated CMOS Sensor Array with PDMS Microfluidics Sample Delivery

Based on the measurement results and further discussions on our discrete sensor implementation in Section 4.2, we find that our sensor performance can be substantially improved if we implement the scheme in integrated form.

In this chapter, we will present our first-version sensor design, which is an 8-element sensor array system, based on a low-cost standard CMOS process [33]. First the system architecture will be described in Section 4.3.1. Then the design details, with emphasis on the sensing inductor, the sensor oscillator and the temperature controller will be given. PDMS microfluidic structures are designed to deliver the target bio-samples onto the specific sensor cell [34]. A low-cost technique to bond the PDMS microfluidic devices

onto the CMOS chip will be presented. The measurement of the integrated sensor implementation will eventually conclude this section.

4.3.1 Sensor Array System Architecture

To fully explore the strength of scaling and parallel processing in CMOS technology, we have implemented the sensor structure in an 8-cell array [35]. This potentially enables the sensor system to detect eight different types of bio-samples simultaneously. The sensor architecture is shown in Figure 4.12.

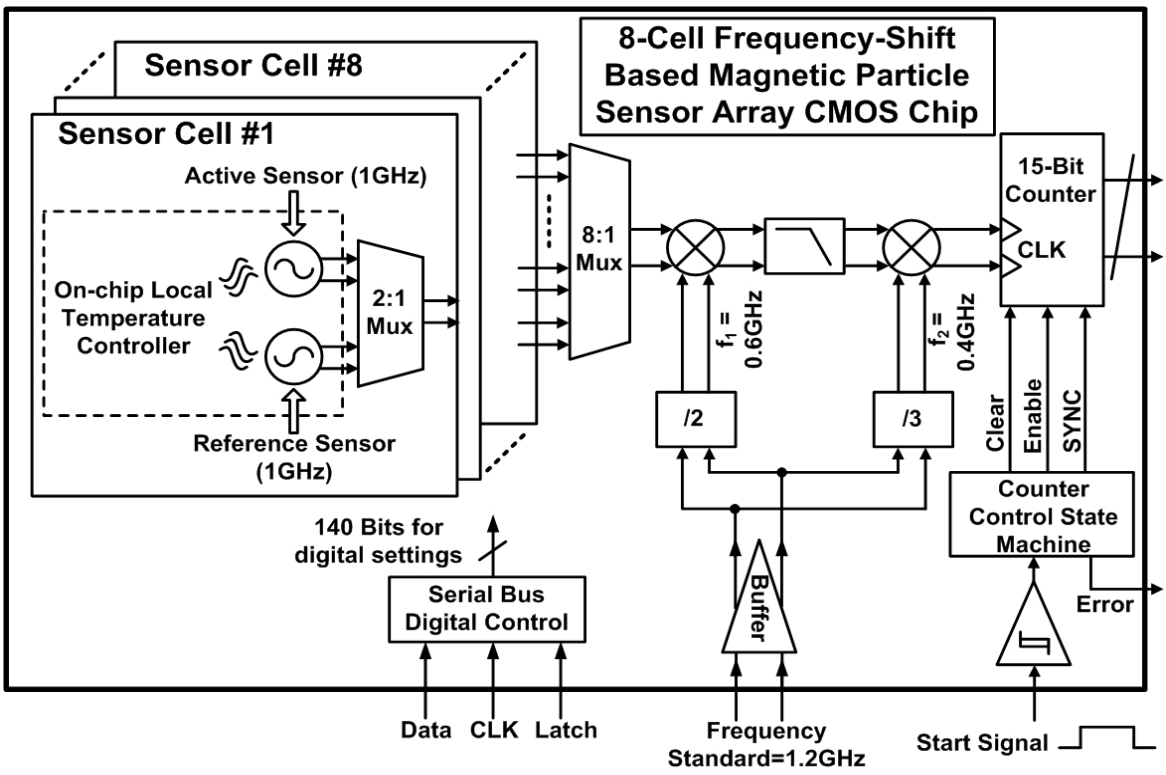


Figure 4.12: The 8-cell CMOS sensor array system architecture

The entire system contains eight independent sensor cells. Within each sensor cell, a differential sensing scheme has been implemented. It is composed of two well matched sensor oscillators, one as the active sensor and the other as the reference sensor, both

operating at a 1GHz nominal frequency. In the layout, the two oscillators are placed in close proximity to each other to improve matching and minimize their on-chip temperature difference. The oscillator pair also shares the power supply, biasing and ground line, which ensures that the differential operation will suppress the low-frequency common-mode noise and drifting described in Section 4.2.3. The detail design issues for sensing oscillator will be presented in Section 4.3.2. Also, based on Section 4.2.3, to improve the sensor sensitivity, in terms of the design, a smaller foot-print sensing inductor should be used, which leads to a higher operating frequency. However, the upper limit of the operating frequency is determined by the magnetic material under test. For micron-size magnetic beads composed of nano-magnetic particles dispersed in a nonmagnetic polystyrene matrix, they typically experience a significant magnetic loss for a frequency above 1GHz in that the real part of their permeability starts to decrease while the imaginary part increases dramatically. This is essentially non-ideal for our magnetic sensing. Therefore, we choose 1GHz as our sensor operating frequency.

A temperature controller is implemented locally for every sensor cell. It regulates the on-chip temperature for the oscillator active cores through a thermal-electrical feedback loop to minimize the frequency drifting effect induced by the ambient temperature change. The details for temperature regulator design will be presented in Section 4.3.3.

As we have covered in Chapter 3 that a typical relative frequency shift for a micron-size magnetic bead is on the order of several ppm or sub-ppm level depending on the sensing inductor design. Therefore, to detect this tiny frequency shift, a frequency counter with large number of bits is required, e.g., 20 bits for a frequency resolution of 1ppm. Therefore, to facilitate accurate detection of such a small frequency shift, the two-step

down-conversion architecture is used to shift the frequency center tone of 1GHz to a baseband frequency below 10kHz. Unlike direct downconversion, this architecture guarantees that neither of the LO signals are close to the sensor free-running frequency nor its harmonics and hence prevents oscillator pulling or injection locking on the sensing oscillator pairs. By using a baseband 15-bit frequency counter, a frequency counting resolution of better than 0.3Hz (3×10^{-4}ppm) is thus achieved.

4.3.2 Sensor Core Design

As we have mentioned in Chapter 3, the first step for designing the frequency-shift magnetic sensor is to determine the inductor design. Based on the electromagnetic simulation on a TSMC 130nm CMOS process with a dual thick copper (3.3μm thickness) option, the sensing inductor design is chosen as a 6-turn symmetric inductor, with the d_{out} of 140μm, width of 5μm and trace separation of 3.5μm. Note that in real biology experiments, this sensor oscillator may need to operate with samples under aqueous condition. The water content may introduce extra loss to the inductor. Therefore, to have a conservative estimation on this Q degradation, we have added an extra layer of sea water, 60μm in height, on top of the SiN passivation layer with its bulk conductivity of 4siemens/m and a frequency-dependent dielectric loss tangent table to capture both the ionic loss and the dielectric loss across the simulated frequency.

The effective differential inductance and its quality factor plot are shown in Figure 4.13 for the cases of with/without the lossy sea water layer. A lateral view of the physical layers used in this electromagnetic model is shown in Figure 4.14.

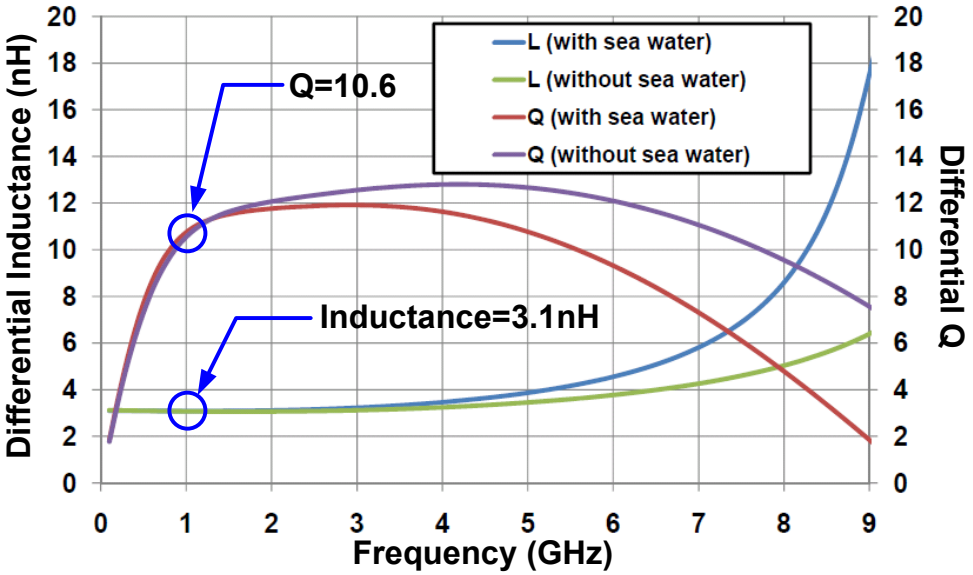


Figure 4.13: The effective differential inductance and its quality factor

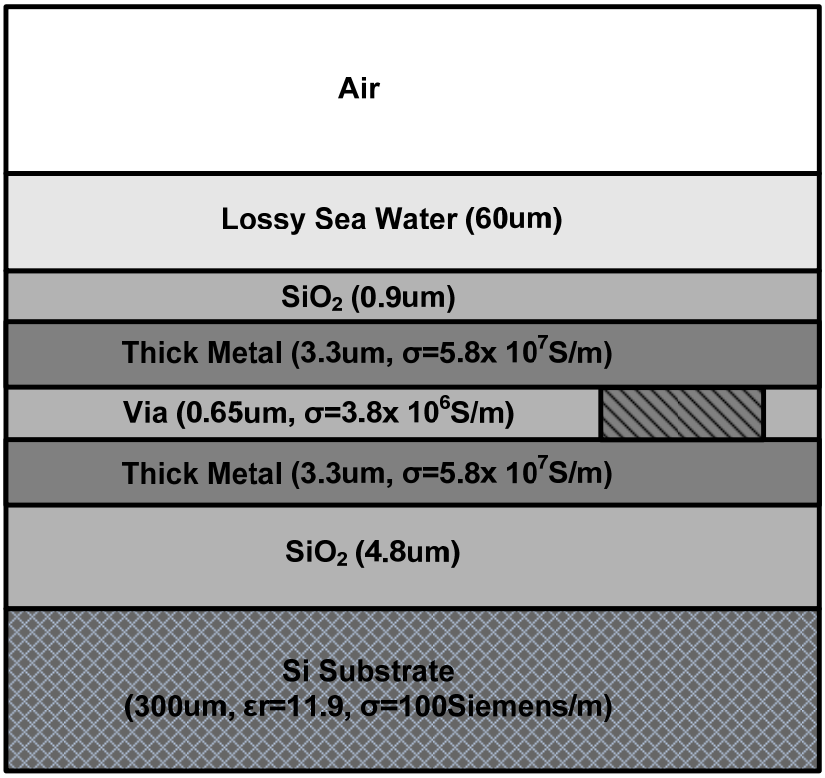


Figure 4.14: The lateral view of the inductor EM module (not on scale)

Figure 4.13 shows that even with the lossy sea-water layer, the inductor still achieves its Q of 10.6 at the 1GHz operating frequency, which is suitable for a high-quality low-noise integrated oscillator implementation.

Based on this inductor layout, the average sensor response to a single one D=1μm magnetic bead is around 0.45ppm/bead by numerical simulations on our close-form formula in Chapter 3.

A sensor oscillator can be designed by using this inductor geometry. We have adopted a differential complementary cross-coupled oscillator biased with an NMOS current source. The relative weight between the NMOS active pair and the PMOS active pair has been carefully optimized to minimize the flicker noise up-conversion from the NMOS current tail. Moreover, to improve the matching, a novel layout for a fully symmetric cross-coupled pair is adopted, which improves the intrinsic oscillator frequency stability and the robustness against process gradient.

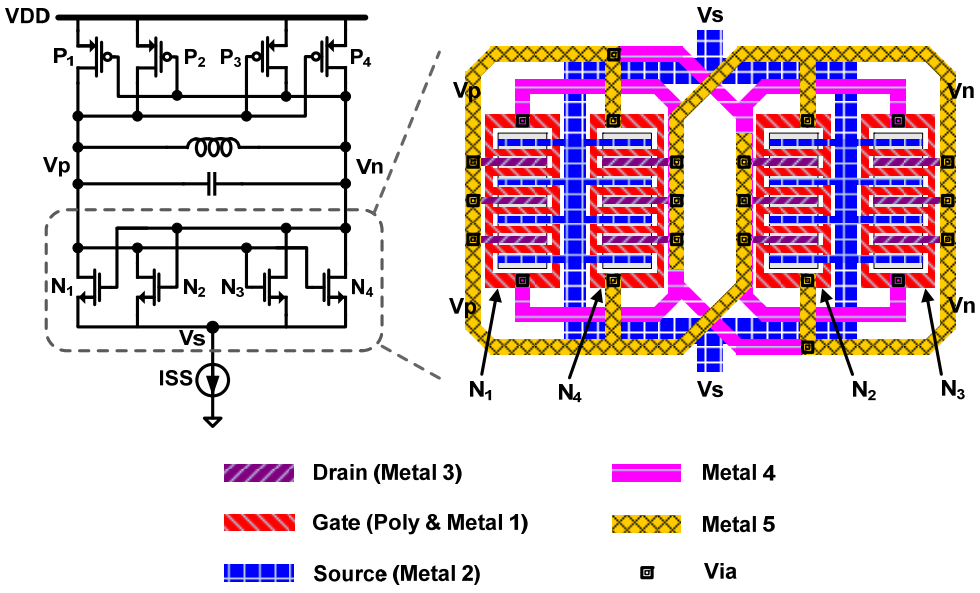


Figure 4.15: The layout for cross-coupled pair with the NMOS pair as an example

As shown in Chapter 3, the $1/f^3$ noise in the total phase noise determines the minimum noise floor of the relative frequency counting, and therefore the sensor minimum sensitivity. To lower down this $1/f^3$ noise content, non-minimum length is used for the transistors in the cross-coupled pair. To provide certain frequency tunability of the sensing oscillator, a switched capacitor bank has been adopted to choose the desired operating frequency of the sensor.

Overall, consuming 4mA from a 1.2V supply, the oscillator achieves a phase noise of -135dBc/Hz and -59dBc/Hz at the offset frequencies of 1MHz and 1kHz, respectively. The layout of the differential oscillator pair is shown in Figure 4.16.

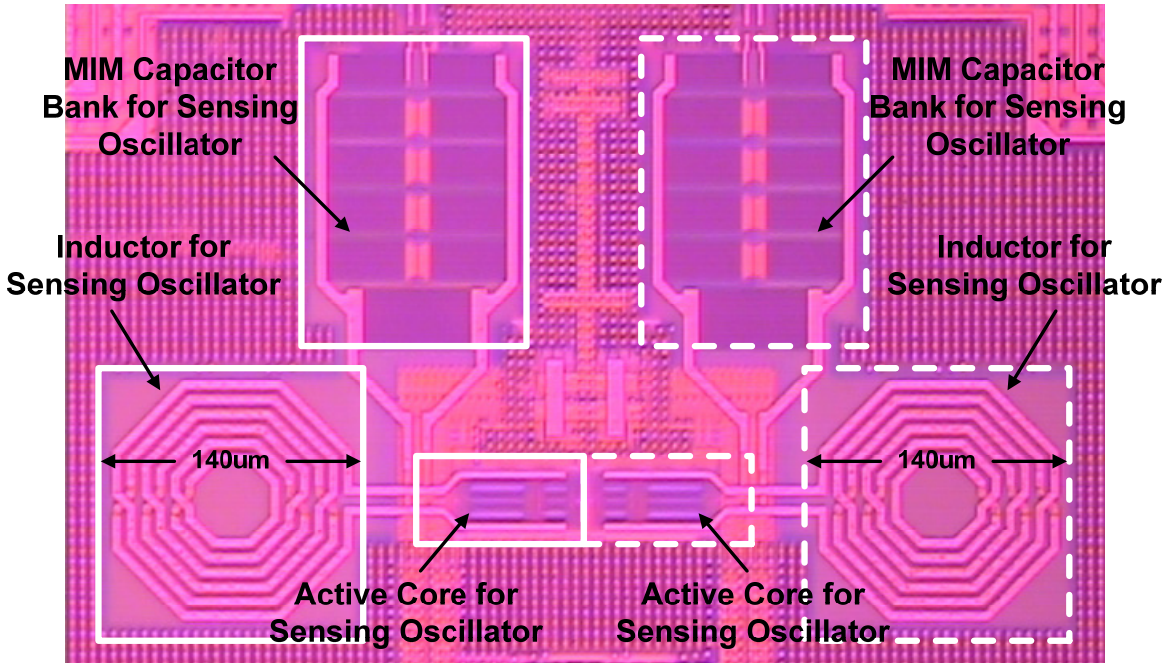


Figure 4.16: Layout for the differential sensing oscillator pair

4.3.3 Temperature Controller Design

An on-chip temperature controller is designed and implemented for every sensor cell. Generally a temperature controlling system contains the following five parts which form an electrical-thermal feedback loop [32], shown in Figure 4.17.

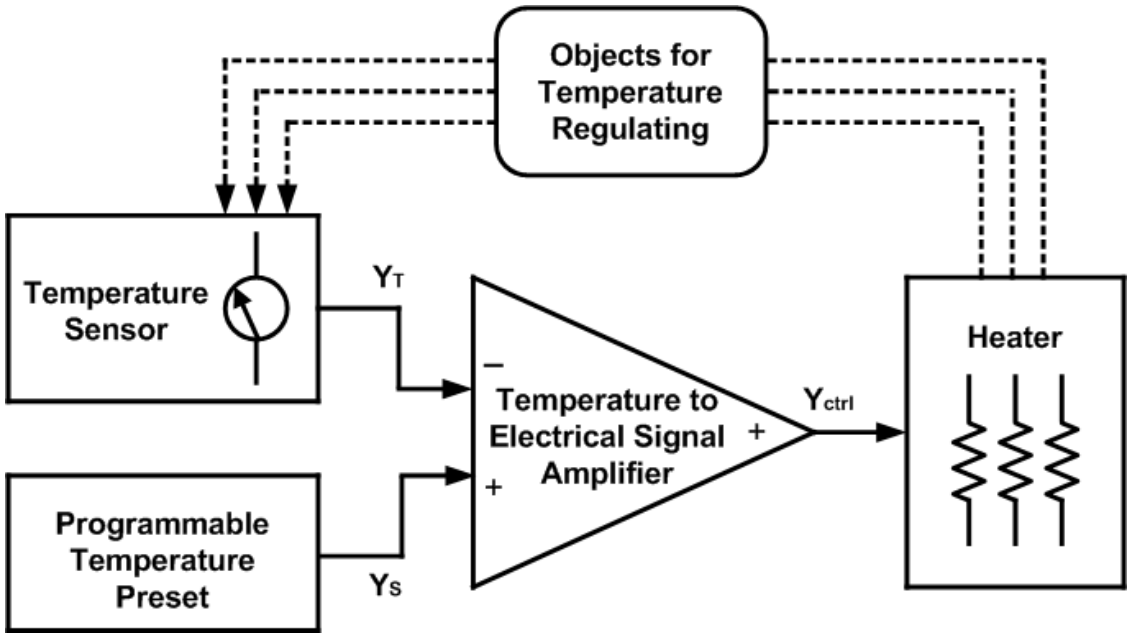


Figure 4.17: The generic temperature controller system schematic

A temperature sensor measures the temperature of the target object whose temperature needs to be regulated to yield a certain signal Y_T. This signal is then compared with a target temperature reference signal Y_S, which can be programmable, and further amplified by some electrical driver amplifier. The output signal of this driver amplifier, Y_{ctrl} thus controls the on/off state and the output power of a heater structure. This completes the electrical path of the temperature regulation. The thermal flow generator by the heater then affects the temperature of the target objects, and this temperature is then sensed by the

temperature sensor. This thermal path therefore completes the entire electrical-thermal feedback loop.

The simplified schematic of our temperature regulator circuit design is shown in Figure 4.18. The temperature sensor is implemented as a proportional-to-absolute-temperature (PTAT) voltage and a bandgap voltage is used as the temperature reference. The PTAT voltage is programmable with a 12-bit control on its output resistors which therefore shifts its output voltage. These two voltages are then fed into a two-stage buffer which amplifies the difference of these two temperature signals and drives a heater transistor array. A common-mode feedback circuit is implemented to lock the common-mode voltage of the driver circuit to the threshold voltage of a dummy heater transistor. This provides a reliable stand-by voltage of the driver output to prevent any false turning-on of the heater transistors due to process/voltage/temperature variations or modeling inaccuracy.

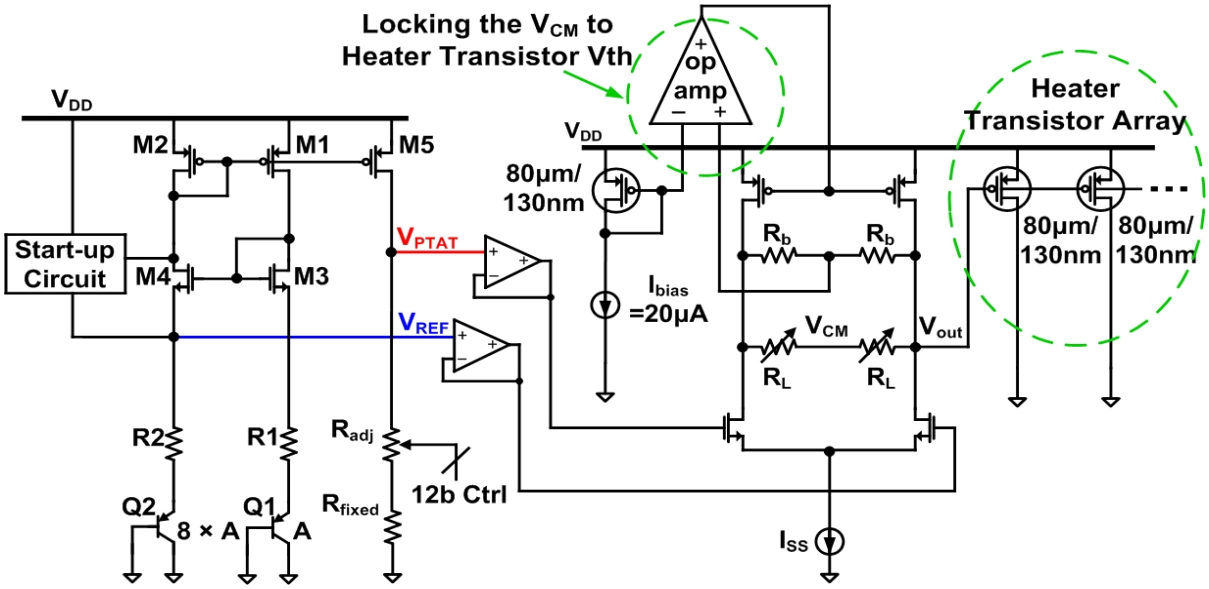


Figure 4.18: The simplified schematic of the temperature controller

What needs to be emphasized is the layout configuration of the heater structure. This is shown in Figure 4.19.

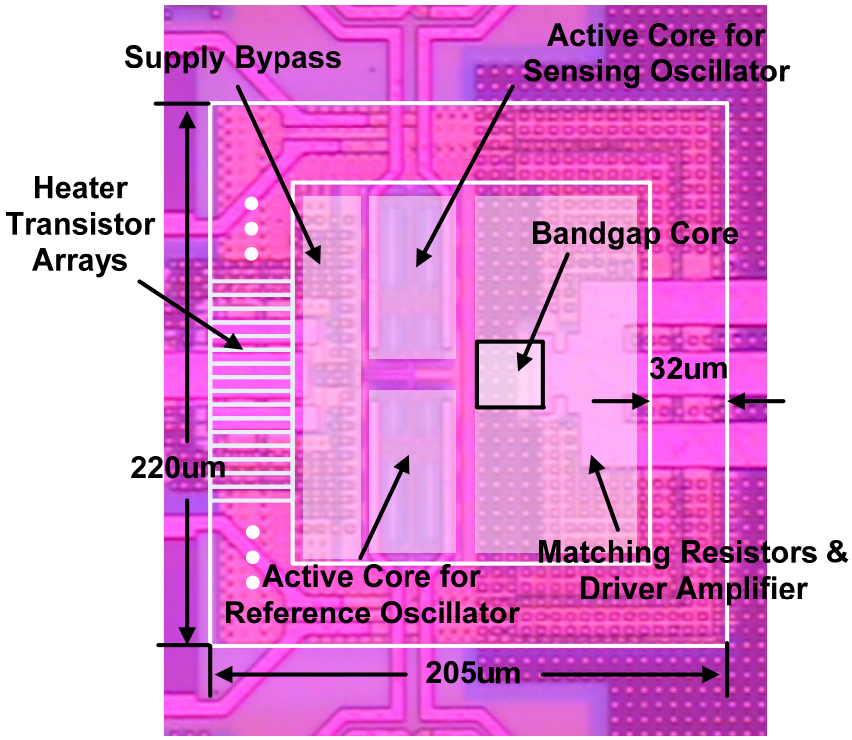


Figure 4.19: The layout configuration of the heater structure

Note that the bandgap core, which includes the reference and the PTAT voltage generation circuits, is placed in a close proximity of the two oscillator active devices for accurate temperature sensing. The power PMOS transistors, as the heater array, form a ring-shape structure and surround the oscillator cores to minimize the spatial temperature difference within the controller.

Overall, this thermal controller forms a 1st-order electrical-thermal feedback loop which has a typical 20.5dB DC gain and is stabilized by a dominant pole in the kHz range.

4.3.4 The Low-Cost Bonding Technique to Attach PDMS Microfluidic Structures to Integrated Circuit Chip

As we have shown in Chapter 2, integrated circuit (IC) technology nowadays presents itself as a promising and powerful implementation tool for biomedical and biochemical applications. On the other hand, to form a complete sensing and/or actuation system, low - cost polydimethylsiloxane (PDMS) devices are often used for functionalities such as sample handling and delivering. Therefore, reliable bounding between the PDMS devices and the IC chips remains as a key step for system integration. However, existing bounding methods involve complicated lithography steps and/or other polymer materials. Moreover, the existing methods are not compatible with conventional low-cost PDMS devices. This will inevitably increase the total system cost and integration difficulty.

To address this issue, we propose a novel low cost bounding technique to attach the PDMS devices onto the IC chip with high mechanical reliability [36]. Our proposed technique will be presented in steps in the following part of this section.

Step 1: The IC chip is immobilized by adhesives, such as silver epoxy, onto the module substrate, shown in Figure 4.20. The module substrate can be either brass or PCB board.

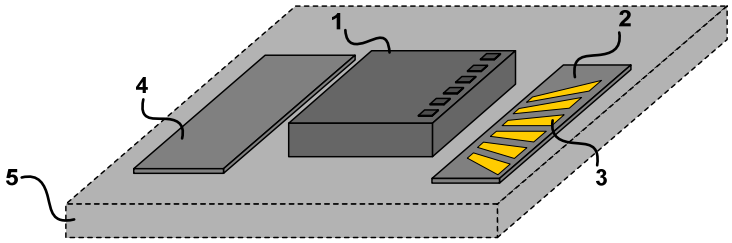


Figure 4.20: The IC chip (Block 1) is immobilized onto the substrate (Block 5). Block 2 and 4 are the substrate with electrical conductive traces (Block 3).

Step 2: Wire-bonds are used to connect the pads on the IC chip with the electrical conductive traces, shown in Figure 4.21.

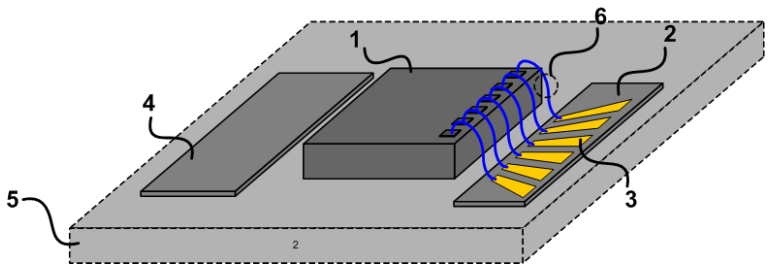


Figure 4.21: Wire-bonds (Block 6) form electrical connections between the IC pads and the electrical conductive traces

Step 3: Fully clean the IC chip surface area and place the PDMS device on it with critical features aligned, shown in Figure 4.22. Note that the PDMS device here can be larger than the IC chip in terms of area. This fact significantly saves the die area expense for the IC chip for a given PDMS device and enables the PDMS device to occupy a larger area for integrating more features and functions. Since the PDMS device needs to extend out of the chip, certain IC pad arrangements should be designed a priori to accommodate this extension.

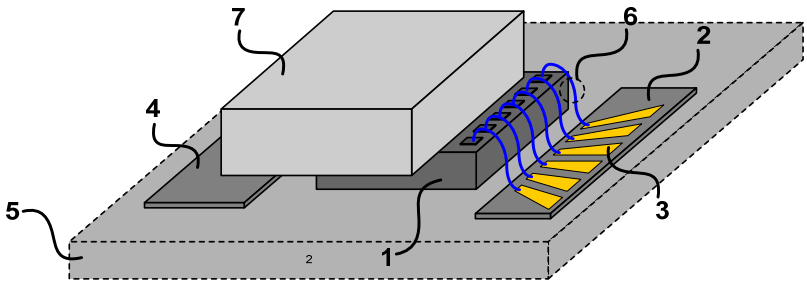
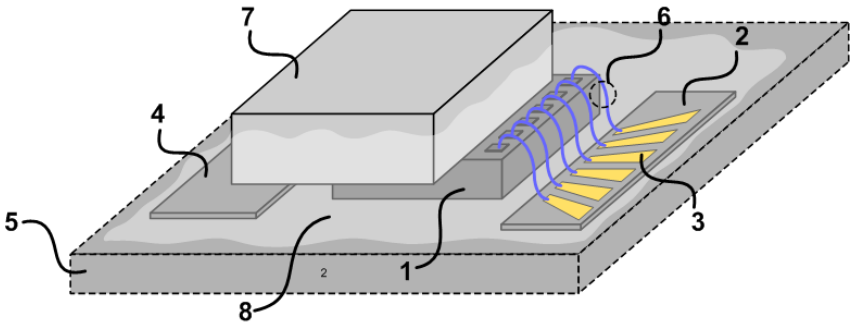
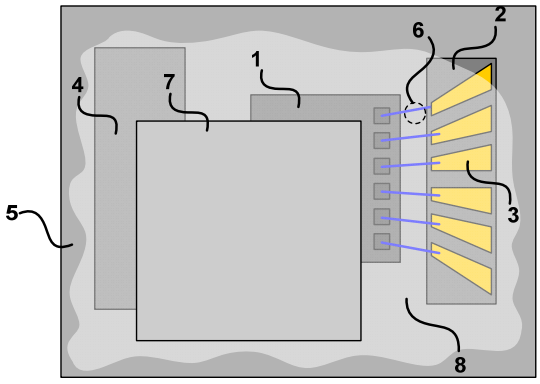


Figure 4.22: The PDMS device (Block 7) is placed on top of the IC chip

Step 4: Mix the polydimethylsiloxane (PDMS) part A and part B with a high weight ratio, such as 20:1. Then apply the mixture as the adhesive around the IC chip and the PDMS devices, shown in Figure 4.23 (a) and (b). The wire-bonds can be covered if needed. The adhesive mixture can flow beneath the PDMS device which is present outside of the IC chip. And the adhesive mixture can flow onto the module substrate and/or the substrates which support the electrical conductive traces. This adhesive mixture increases the total contact area between the PDMS device and the IC chip, which accounts for the significant improvement of the bounding strength.



(a)



(b)

Figure 4.23: The PDMS mixture (Part 8) is applied to surround the IC chip and the PDMS structure

Step 5: Curing the applied mixture. This can be achieved by preserving the module at around 70˚C for 30 minutes, or around 40˚C overnight, or room temperature for about 2 days. Note that the actual curing temperature and time can vary for specific applications. In general, a longer temperature is expected if a lower curing temperature is used.

What needs to be emphasized here is that our proposed bonding technique does not require any post processing lithography steps, such as patterning and etching. Therefore it saves the total module cost significantly, which aligns with our primary goal of achieving a low-cost sensor system.

4.3.5 Measurement Results

In this section, the measurement results of our integrated magnetic sensor will be presented.

The entire chip is fabricated in a standard TSMC 130nm CMOS process with a dual-thick copper option (thickness of 3.3μm). The entire chip occupies an area of 2.95mm×2.56mm. The chip microphotograph, with the eight differential sensing cells high-lighted, is shown in Figure 4.24. All the active pads are located at the top edge of the chip so that the PDMS device can be placed onto the CMOS chip with its extensions on the other three sides.

The measured and simulated phase noise profile of the sensor oscillator is shown in Figure 4.25. We can see a very close agreement between the measurement and the simulation results.

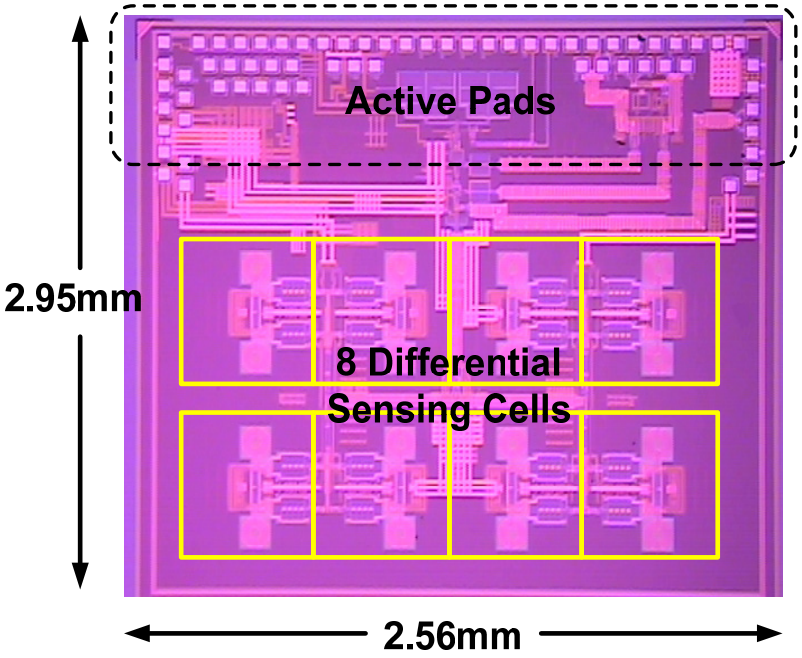


Figure 4.24: CMOS magnetic sensor array chip microphotograph

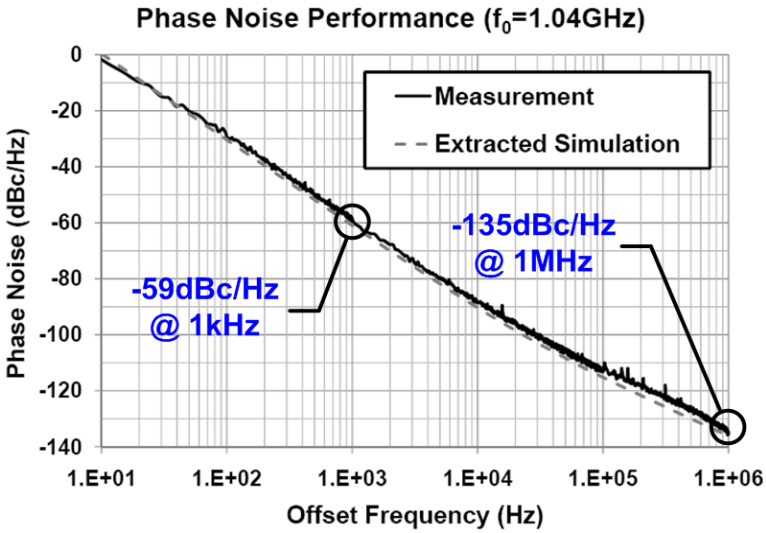


Figure 4.25: The phase noise plot for sensor oscillator (measurement and simulation)

The effect of differential scheme on frequency counting is shown in Figure 4.26. The blue trace and the red trace show the two oscillators' stand-alone frequency counting results. The black curve represents the differential frequency counting results on the

sensing oscillator pair. The low frequency common-mode drifting is greatly suppressed by differential scheme. Overall, the differential frequency uncertainty is 0.13ppm before averaging. Compared with the discrete design presented in Section 4.2, we can see this integrated design with differential scheme results in more than two orders of magnitude improvement on oscillation frequency stability.

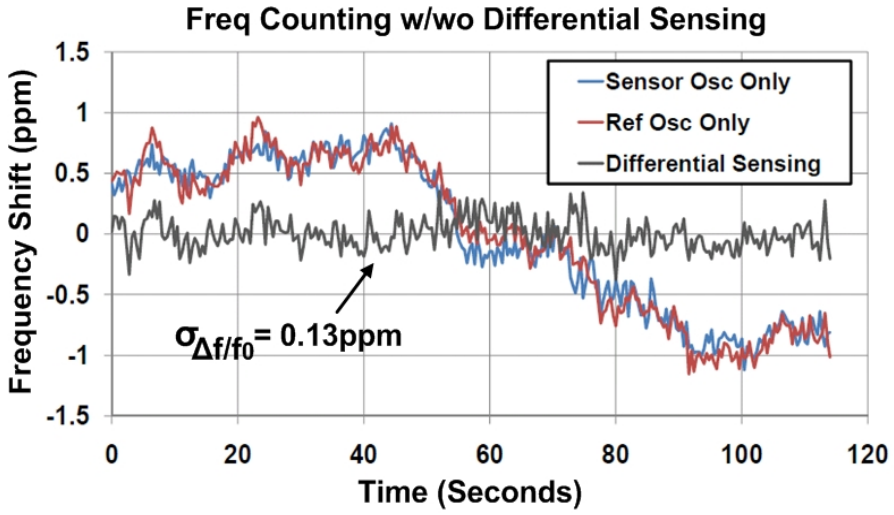


Figure 4.26: The frequency counting results with/without differential scheme

The heater response in terms of the total heater current versus on-chip temperature variations is depicted in Figure 4.27. When the on-chip temperature deviates from the target temperature, i.e., 29˚C for this measurement, the heater starts to draw a DC current from its supply and heat up the VCO active cores. The measured heater responses for three different loop-gain settings are shown, which have close agreements with the simulated responses.

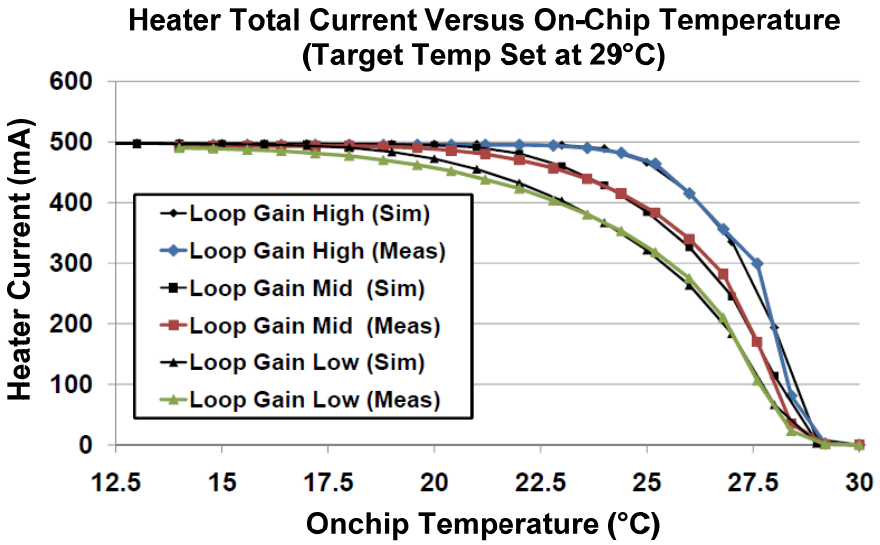


Figure 4.27: The heater response versus the on-chip temperature

The microphotograph for the sensor chip with the PDMS device is shown in Figure.4.28. The critical circuit blocks are high-lighted. The zoom-in view of a differential sensing cell is provided on the right. The sensing and reference inductors are covered by the sensing and reference microfluidic chambers controlled by pneumatic valves.

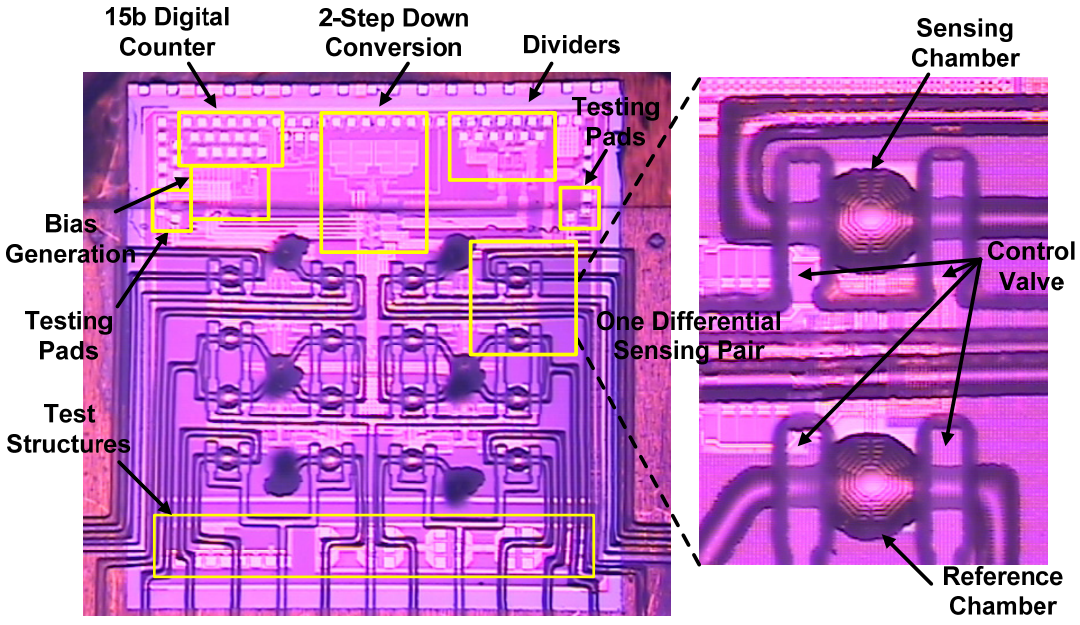


Figure 4.28: The Micrograph for the CMOS chip attached with PDMS microfluidics

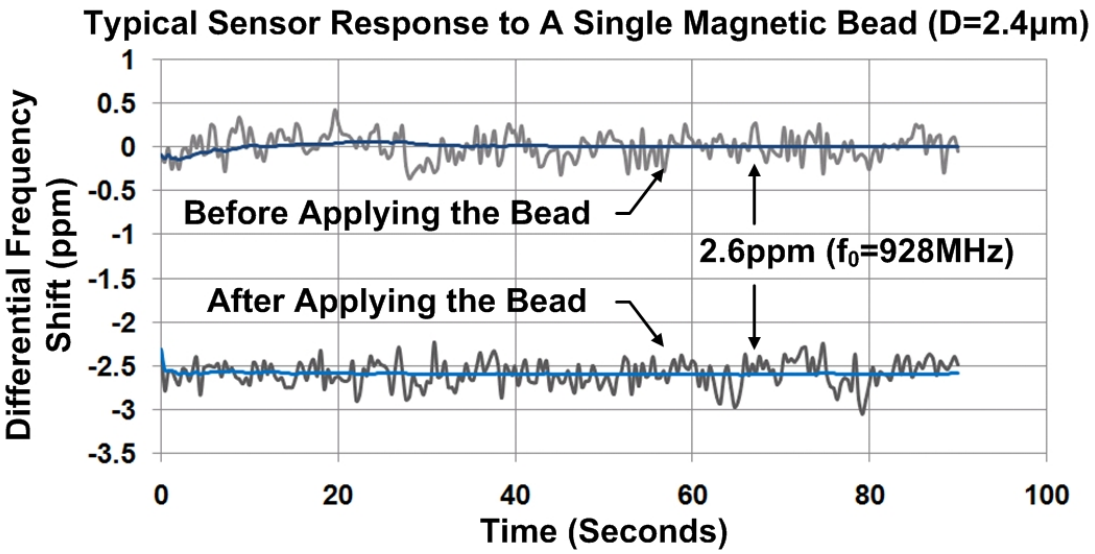
Two experiments are performed to verify the sensor's functionality. The first experiment uses micron-size magnetic particles, Dynabeads® products [16], directly as the test samples. Since the sizes of the magnetic beads are easily discerned under the optical microscopes without using any fluorescence labeling techniques, this experiment provides a straightforward way to test the sensor's basic responses. The second experiment is to verify the sensor's detection functionality for real magnetically labeled biosamples. In this experiment, actual DNA samples are used as target molecules, which are labeled by 50-nm diameter magnetic particles. The measurement results are presented as follows.

In the first experiment, three types of magnetic beads are used. They are Dynabeads® M-450 Epoxy, Dynabeads® Protein G and Dynabeads® MyOne™. The measurement results are summarized in Table 4.2. Averaging on sensing data is performed to achieve a high signal-to-noise ratio. For all the three particle types, one single magnetic bead can be detected with a large SNR within less than a 3-minute data averaging.

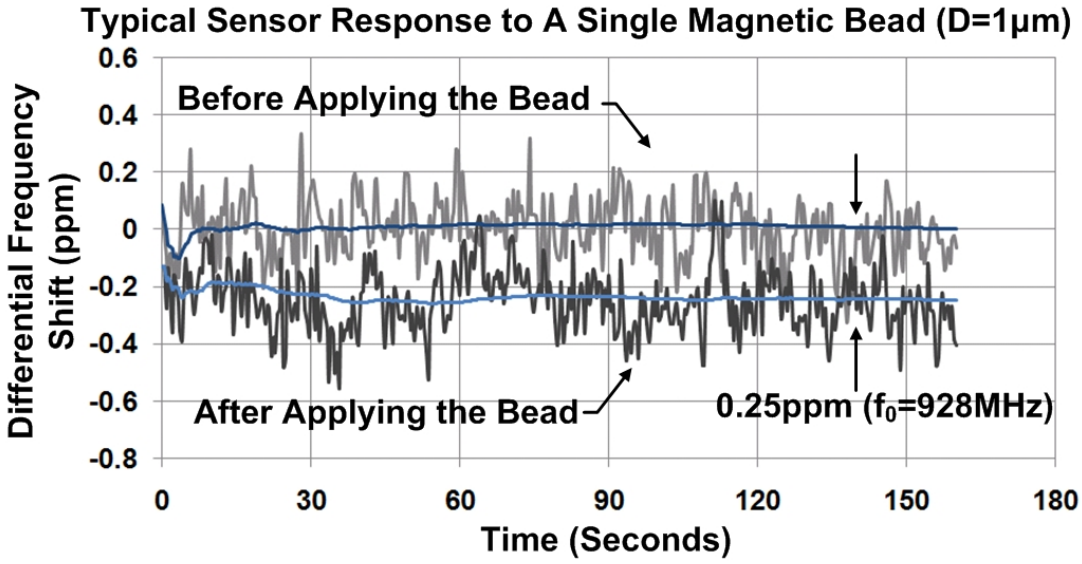
Table 4.2 SUMMARY ON SENSOR RESPONSE TO MAGNETIC BEADS

	Bead Size (Diameter)	Averaged Δf/f per Beads	Sensitivity (# of Beads)	SNR	Averaging Time (s)
DynaBeads® M-450 Epoxy	4.5μm	9.6ppm	1	57 dB	90
DynaBeads® Protein G	2.4μm	2.6ppm	1	46 dB	90
DynaBeads® MyOne™ Carboxylic Acid	1μm	0.23ppm	1	28 dB	160
Polystyrene Bead (Non-Magnetic)	1μm	0.0035ppm			

Typical measurement results for one single magnetic bead of 2.4μm and 1μm are shown in Figure 4.29 (a) and (b). The blue curves represent the data trace after averaging.



(a)



(b)

Figure 4.29: Typical sensor measurement results: (a) one single 2.4μm magnetic bead; (b) one single 1μm magnetic bead

In the second measurement, the actual DNA sample (1-kilo base-pair long, 1nano-Molar in concentration) labeled by magnetic nanoparticles (D=50nm) is tested. The PDMS bottom layer of the microfluidic channel is first modified by the biotin molecules, which thus immobilizes the DNA samples through biotin-streptavidin-biontin links [37]. The 50 nm magnetic particles are labeled onto the DNA samples through dig-antidig link [38]. Typical sensor measurement result (shown on the right) together with the binding chemistry (shown on the left) is demonstrated in Figure 4.30. As we can see a total relative frequency shift of 2.4ppm is registered for a DNA of 1nano-Molar concentration. The curves shown in blue represent the measurement data after performing averaging.

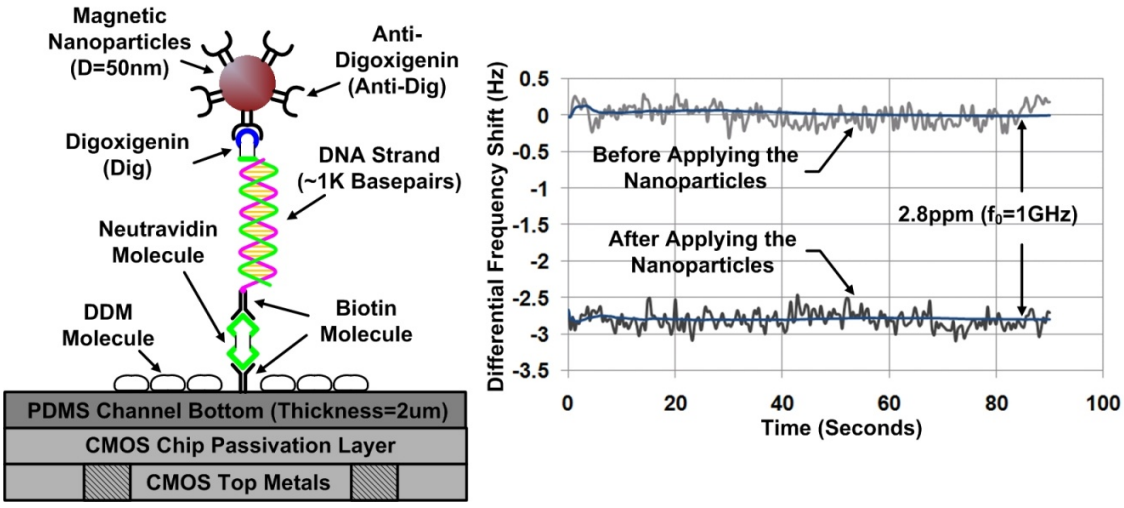


Figure 4.30: Typical sensor response to 1n-Moalr DNA samples labeled with 50-nm magnetic nanoparticles

To verify the functionality of the bounding chemistry, a fluorescent experiment is used. Two substrate surfaces have been prepared with the identical chemistry as the one used in the DNA sensing experiment. The target DNA is first delivered onto one of the surface, while only buffer solution is delivered onto the other. Then the magnetic particles labeled

by fluorescent groups, DyLight-488, are introduced onto the surface and washed after hybridization. The measurements are shown in Figure 4.31 with the surface chemistry on the left and the fluorescent images on the right. A significantly stronger fluorescent signals can be observed for the surface with the complementary DNA than that without the complementary DNA. This verifies that the magnetic nanoparticles are immobilized onto the sensor surface specifically due to the presence of complementary DNA.

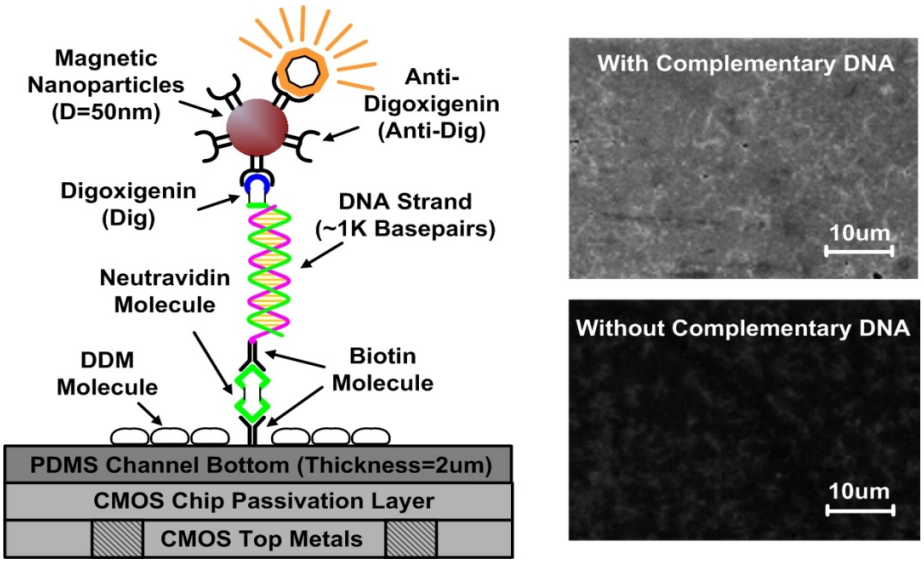


Figure 4.31: Fluorescent images of test surfaces with complementary DNA and without complementary DNA, shown on the right

4.4 Chapter Summary

This chapter focuses on the practical implementation side for the frequency-shift based magnetic sensing scheme proposed in Chapter 3.

A discrete sensor system is first built based on the thin-film technology with the spiral inductors directly patterned on the thin-film board. Both the dry experiment and the

aqueous measurement results achieve good matches with simulation results and demonstrate the functionality of detecting magnetic particles. With a frequency uncertainty of 10 ppm and an average frequency shift of 4 ppb/bead, an overall 2500 magnetic beads (D=1μm) can be detected. Moreover, the measurement results together with the analyses in Chapter 3 show that in order to improve the sensitivity, the sensor needs to be implemented with smaller sensing inductors, differential sensing scheme and temperature regulating. This naturally brings integrated circuit technology as a promising choice for our sensor implementation.

Therefore, as a second-version implementation, an eight-cell magnetic sensor array system is designed in a standard CMOS process. The sensing inductor is a six-turn symmetric spiral with an outer diameter of 140μm, more than four times smaller than the sensing inductor for the discrete design. Each sensor cell contains a differential sensing oscillator pair with a local-temperature regulation. This configuration greatly suppresses the low-frequency common-mode offset and noise, such as supply perturbation and temperature variations. As a result, this integrated design achieves a differential frequency uncertainty of 0.13ppm. This low noise floor guarantees the reliable detection of one single micron-size magnetic bead (D=4.5μm, 2.5μm and 1μm). Furthermore, the sensor has been tested with real DNA samples (1kb-long) labeled by magnetic nanoparticles (D=50nm). A DNA sample of 1nano-Molar concentration is reliably registered. This verifies the sensor's functionality for detecting magnetically labeled biosamples.

Chapter 5: Low Noise Techniques in Frequency-Shift Magnetic Particle Sensor

In order to discern the sensing target with low concentration or perform a conclusive test when the total sample quantity is limited, a sensor system with an ultra-high sensitivity is then desired. This suggests that one should either increase the sensor transducer gain or lower the sensor noise floor to achieve a high sensor signal-to-noise (SNR) ratio. As discussed in Section 3.4, the sensor transducer gain for our frequency-shift based magnetic sensor is entirely determined by the sensing inductor design, which encounters various practical constrains limiting the improvement on the sensor signal. On the other hand, Section 3.5 shows that the sensor noise floor is dominated by the intrinsic oscillator phase noise, which also experiences a strong trade-off with power consumption for conventional oscillator design.

Therefore, to further push the sensing limit of the frequency-shift based magnetic sensors, advanced techniques beyond the conventional inductor and oscillator optimization should be explored. In this chapter we will extend our previously demonstrated differential sensing scheme to propose a novel noise shaping and suppression technique without direct trading with sensor power consumption.

This chapter is organized as follows. Section 5.1 revisits the relationship between phase noise and timing jitter with an emphasis on both the $1/f^2$ and the $1/f^3$ phase noise. The effect due to different frequency counting time window is also introduced. This lays the foundation for various noise analyses and calculation in the subsequent sections. Section 5.2 presents a novel sensor noise suppression technique, namely Correlated Double

Frequency Counting technique (CDFC), with implementation examples. Based on the mathematical derivation on the phase noise suppression, a noise shaping function for CDFC is then proposed in Section 5.3. To further decrease the measurement noise floor, a novel interleaving-N CDFC technique together with potential implementation is presented in Section 5.4. Finally, a summary will conclude this chapter in Section 5.5.

5.1. Phase Noise, Jitter and Frequency Counting Window

As mentioned in Chapter 4, for a practical implementation, a differential sensing scheme is mandatory to suppress the common-mode supply noise, biasing noise and temperature variation residue. Therefore, by taking the frequency counting difference between the reference oscillator and the sensing oscillator, the total differential frequency uncertainty is given by,

$$\sigma^2_{\frac{\Delta f}{f_0},diff} = \frac{8}{\pi\omega_0^2 T^2}\int_0^{+\infty} S_\phi(\omega')\sin^2\frac{\omega' T}{2}d\omega' \tag{5.1}$$

where $S_\phi(\omega)$, T and ω_0 are the oscillator phase noise profile, the frequency counting time window and the center oscillation frequency.

At a large offset frequency, i.e., small frequency counting time window, where $1/f^2$ noise is dominant, this uncertainty can be calculated as,

$$\sigma^2_{\frac{\Delta f}{f_0},diff} = \frac{8}{\pi\omega_0^2 T^2}\int_0^{+\infty} \frac{A_2}{\omega'^2}\sin^2\frac{\omega' T}{2}d\omega' = \frac{2A_2}{\omega_0^2 T} = \frac{2k^2}{T} \tag{5.2}$$

where A_2 is the noise coefficient of the $1/f^2$ noise profile and k is the jitter coefficient for $1/f^2$ phase noise [26].

At a small offset frequency, i.e., large frequency counting time window, where $1/f^3$ noise is dominant, this uncertainty is calculated as,

$$\sigma^2_{\frac{\Delta f}{f_0},diff} = \frac{8}{\pi\omega_0^2 T^2}\int_0^{+\infty} S_\phi(\omega')\sin^2\frac{\omega' T}{2}d\omega' = \frac{\zeta^2 T^2}{T^2} = 2\zeta^2 \tag{5.3}$$

where ζ is the jitter coefficient of the $1/f^3$ phase noise [26]. Note that direct integration of (5.3) on an exact $1/f^3$ noise mathematical expression will result in a nonphysical unbounded result. This issue is normally remedied by adopting a modified $1/f^3$ phase noise profile [39].

Note that the factor of two in both equation (5.2) and (5.3) are due to the differential scheme, assuming the phase noise is independent and identical for both the sensing oscillator and the reference oscillator.

Therefore, the relative frequency counting uncertainty/noise (in power) can be plotted versus the frequency counting window T. At small T values, $1/f^2$ phase noise is dominant. Thus, the differential noise power decreases inversely proportional to T. However, at large T values, $1/f^3$ phase noise is dominant, which leads to a constant noise floor for this long counting window. This noise floor essentially determines the sensing limit of the sensing system, if frequency counting window T can be chosen arbitrarily.

In addition, a frequency resolution due to the uncertainty principle should be superimposed onto the derived phase-noise dependent frequency counting uncertainty. This uncertainty principle means a frequency resolution of 1Hz can be achieved only when the frequency counting window is equal or longer than 1 second. As we can show in the later section that this constraint is not fundamental and can be circumvented by special implementation techniques, such as fractional counting scheme. However, this frequency

uncertainty ($1/f_0^2T^2$ with f_0 as the oscillation frequency) is still listed here for completeness. The total frequency counting uncertainty is shown in Figure 5.1.

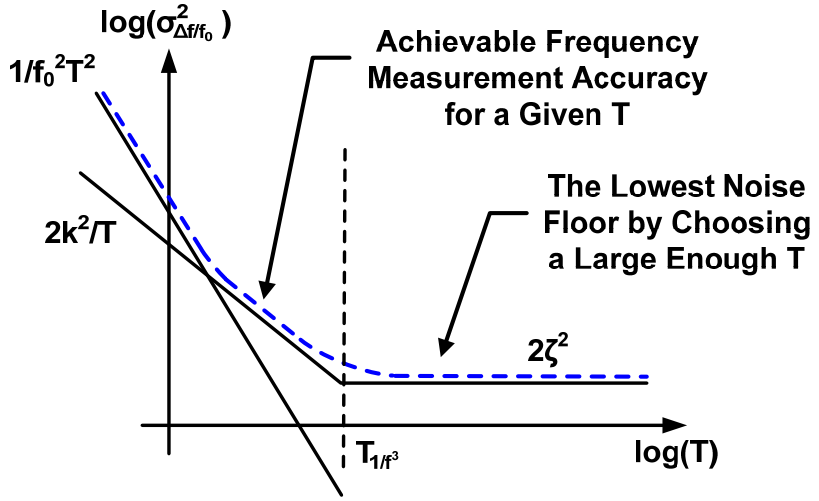


Figure 5.1: Frequency counting uncertainty for normal differential sensing scheme

First of all, this figure shows that a large uncertainty will be encountered if the frequency counting time is too short. Note that the relative strength between $1/f_0^2T^2$ and $2k^2/T$ determines which uncertainty will be dominant at that short counting time. Secondly, the minimum noise level is dictated by $2\zeta^2$ where ζ is the jitter coefficient of $1/f^3$ phase noise power.

Since the $2k^2/T$ jitter is from $1/f^2$ phase noise, which presents a random walk behavior in the phase noise while remains white in the frequency noise, a long time averaging, i.e., a longer counting time, on those uncorrelated noisy frequency measurements, decreases the noise power relative to the carrier power exactly by a factor of T, the total frequency counting time. This can be viewed as a special example of the classic averaging on i.i.d stochastic measurement results. On the other hand, the $1/f^3$ phase noise, behaving as 1/f noise in the frequency noise, is a correlated noise among adjacent frequency measurements,

which cannot be averaged out by a long counting time. Therefore, it results in a final noise floor of $2\zeta^2$ for the frequency counting measurement at a long measurement time.

Finally, the relative strengths among the $2k^2/T$ curve, $1/f_0^2T^2$ curve and the $2\zeta^2$ together decide the minimum counting time T_{min} above which the base-line noise level of $2\zeta^2$ can be achieved.

5.2. Correlated Double Frequency Counting (CDFC) Technique

Section 5.1 shows that for a differential frequency counting measurement, $2\zeta^2$, determined by the $1/f^3$ oscillator phase noise, appears as the fundamental noise floor. If a higher frequency counting resolution is desired, this flicker-related oscillator phase noise has to be minimized. For a typical oscillator design, when the DC term of the impulse sensitivity functions (responsible for flicker noise up-conversion) has been minimized, conventionally the only way to further decrease the oscillator's $1/f^3$ phase noise is by scaling up the power consumption to effectively increase the carrier-to-noise power ratio.

However, for a practical sensor implementation, other design constrains will set the upper limit of the oscillator power consumption. This includes the handheld sensing system's battery life, the electromigration effect on active/passive devices and magnetic saturation on the target particles due to excessive excitation current, etc. Therefore, it is desirable to explore noise reduction techniques with no or very limited power consumption overhead.

This section will propose a novel noise reduction method which satisfies the requirement, i.e., reducing the frequency counting uncertainty without increasing the power consumption [40].

5.2.1 Proposed Scheme and Circuit Topology

As discussed in Chapter 4, in order to minimize the common-mode perturbation/noise during sensing, a differential scheme needs to be implemented. And our output frequency-shift signal will be the frequency difference of this differential oscillator pair.

This differential scheme suggests that if we can relate the phase noise, in particular the $1/f^3$ phase noise (dominant at low frequency offset), to be correlated between the sensing oscillator and the reference oscillator, taking the frequency difference between these two will be able to suppress this correlated noise in a similar way for the aforementioned common-mode perturbation/noise suppression.

In general, for a typical cross-coupled electrical oscillator, the $1/f^3$ phase noise normally comes from the active devices including the cross-coupled cores and the tail current sources. The latter can be greatly suppressed by shaping its impulse-sensitivity-function, ISF, or by using the resistor as tail biasing. And the $1/f^2$ phase noise is mainly due to the LC tank loss and the cross-coupled core thermal drain noise. Therefore, one implementation approach suggested by this fact is that the oscillator active core should be shared between the sensing and the reference oscillator in order to correlate the $1/f^3$ phase noise. In addition, a practical disabling/enabling scheme, such as switches, is required for the two LC sensing tanks, so that they can be alternatively connected or disconnected from the shared active core to complete the differential sensing procedures.

We name this scheme as Correlated Double Frequency Counting (CDFC) scheme. One potential circuit implementation of this CDFC scheme is shown in Figure 5.2 as follows, where T is the total counting time for one oscillator and T_{set} is the reset time for a practical frequency counter implementation.

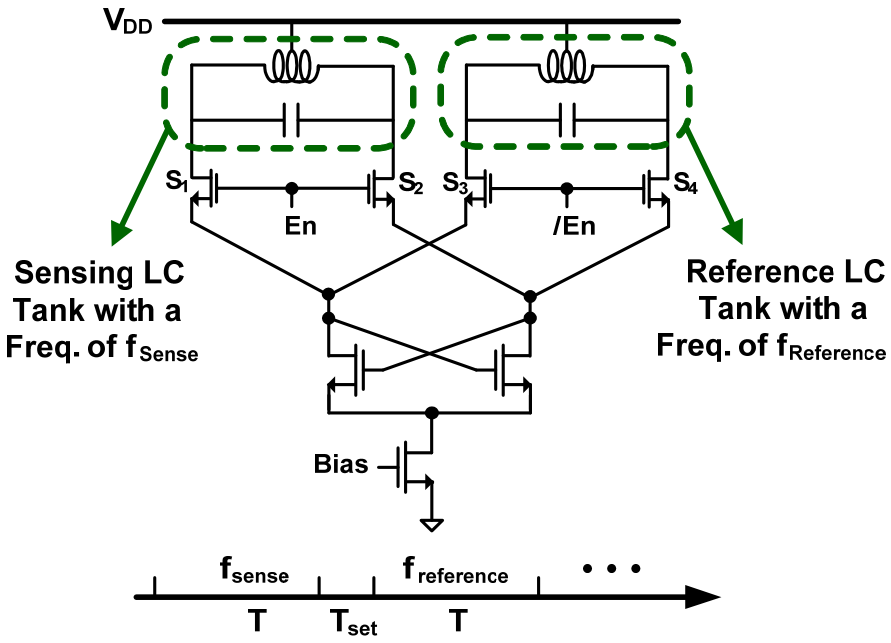


Figure 5.2: One simplified circuit schematic of the proposed CDFC scheme

The circuit in Figure 5.2 is based on an NMOS-only oscillator design. The two pairs of switches (S_1/S_2 and S_3/S_4) select either the sensing or the reference LC tanks to be connected into the oscillator circuit. Furthermore, the switches should be biased in deep triode region and connected in series with the parallel LC tank, which minimize its degradation on the quality factor of the tank. Moreover, since the flicker noise from the switches pair will be uncorrelated between the differential frequency counting samples, its $1/f^3$ phase noise contribution in the total phase noise should be minimized. In this section,

let us focus on the theoretical side of the CDFC scheme. A practical implementation example will be shown in the following section.

For a general CDFC implementation, various phase noise sources can be classified into the following four types.

Type 1: Correlated $1/f^3$ noise sources between differential sensing. These are the phase noise sources which will be suppressed through our CDFC scheme. They include the flicker noise of the cross-coupled core and the tail current source. Here we will mainly consider the former, since the latter can be significantly suppressed by using the current source device with large peripherals or optimizing its impulse sensitivity function.

Type 2: Correlated $1/f^2$ noise sources between differential sensing. These include the drain thermal noise from the cross-coupled core and the tail current source. They are denoted as "correlated $1/f^2$ noise sources" due to the fact that these noise sources are shared through the differential frequency counting scheme.

Type 3: Uncorrelated $1/f^3$ noise sources between differential sensing. The main contributor of this type is the flicker noise of the switch transistors. Because of the uncorrelated nature, this noise cannot be rejected through the differential counting scheme.

Type 4: Uncorrelated $1/f^2$ noise sources between differential sensing. This is mainly the thermal noise from the resonator tanks and the switch transistors.

Therefore, the total normalized jitter after one differential sensing sampling for the correlated noise $\phi_c(t)$ can be shown as,

$$\sigma^2_{\frac{\Delta f}{f_0}} = \frac{1}{T^2\omega_0^2} E\left\{\left[\phi_c(t+2T+T_{set}) - \phi_c(t+T+T_{set}) - [\phi_c(t+T_{set}) - \phi_c(t)]\right]^2\right\}$$

$$= \frac{1}{T^2} \cdot \frac{2}{\omega_0^2}\left[2R_{\phi,c}(0) - 2R_{\phi,c}(T) - 2R_{\phi,c}(T+T_{set}) + R_{\phi,c}(2T+T_{set}) + R_{\phi,c}(T_{set})\right]$$

(5.4)

where $\phi_c(t)$ includes both type 1 and type 2 noise mentioned above. By Wiener-Khinchin theorem, equation (5.4) can be related with the oscillator phase noise profile as,

$$\sigma^2_{\frac{\Delta f}{f_0}} = \frac{16}{\pi\omega_0^2 T^2}\int_0^{+\infty} S_{\phi,c}(\omega')\, sin^2\left[\frac{\omega' T}{2}\right]\sin^2\left[\frac{\omega'(T+T_{set})}{2}\right]d\omega' = 2\beta_{\zeta,c}\left(\frac{T_{set}}{T}\right)\zeta_c^2 + 2\beta_{k,c}\left(\frac{T_{set}}{T}\right)\frac{k_c^2}{T}$$

(5.5)

where $S_{\phi,c}(\omega)$ represent the total correlated phase noise profile and ζ_c and k_c are the jitter coefficients for the correlated $1/f^3$ and $1/f^2$ phase noise; $\beta_{\zeta,c}$ and $\beta_{k,c}$ are defined as the noise reduction factors for the type 1 and type 2 noise in the CDFC scheme, which can be computed and plotted as follows.

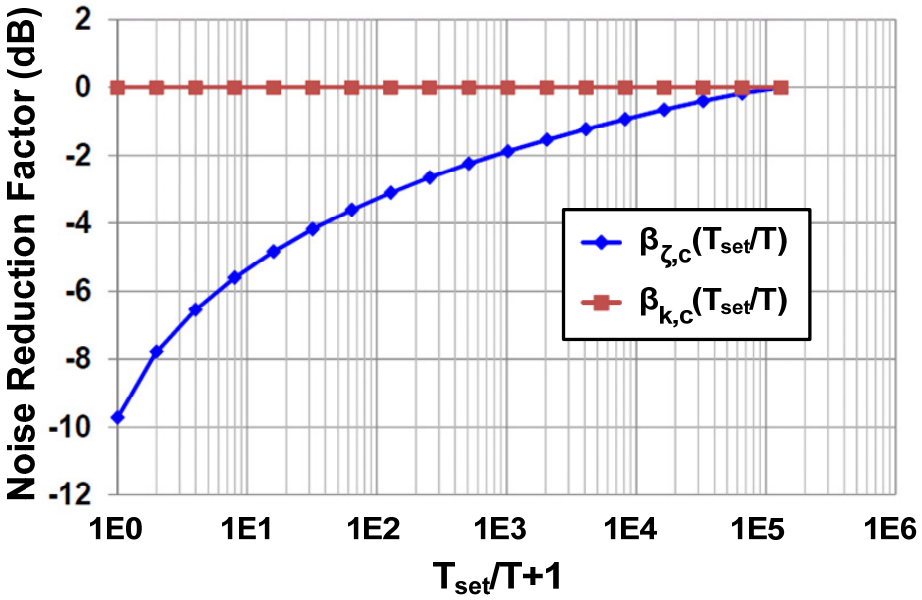


Figure 5.3: Noise reduction factor for correlated $1/f^3$ and $1/f^2$ phase noise

Note that the $\frac{T_{set}}{T}$ value specifies the ratio between the counting reset interval and the total counting time for the CDFC counting scheme.

As indicated in the Figure 5.4, $\beta_{k,c}$ holds a constant value of 1, independent of $\frac{T_{set}}{T}$. This result means although the $1/f^2$ noise is generated from the same source, i.e., the cross-coupled transistors, its noise power remains the same for the CDFC scheme compared with the normal differential scheme. Physically, this is because $1/f^2$ phase noise behaves as white frequency noise, and then the adjacent frequency counting samples are essentially uncorrelated in stochastic sense. Therefore, our CDFC scheme will not suppress this correlated $1/f^2$ phase noise, even it is from the shared cross-coupled cores.

On the other hand, in terms of $\beta_{\zeta,c}$, when $\frac{T_{set}}{T}$ is zero, negligible delay between the adjacent measurements, the best noise rejection factor of 9.8dB can be achieved. The rejection monotonically decreases when $\frac{T_{set}}{T}$ gets larger, with the worst-case value of 0dB, which means there is essentially no rejection when $\frac{T_{set}}{T}$ approaches infinity. This is because the $1/f^3$ phase noise, i.e. the 1/f frequency noise, has its autocorrelation function experiencing a relaxation behavior in the time domain, resulting in a decreasing correlation for a larger $\frac{T_{set}}{T}$. Therefore, minimizing the counting interval ensures maximizing the correlation between the adjacent differential frequency samples and therefore leads to the most noise power reduction.

On the other hand, the normalized jitter for the uncorrelated phase noise $S_{\phi,nc}(\omega)$ through normal differential scheme is given as

$$\sigma^2_{\frac{\Delta f}{f_0}} = \frac{8}{\pi\omega_0^2 T^2}\int_0^{+\infty} S_{\phi,nc}(\omega')\, sin^2\left[\frac{\omega' T}{2}\right] d\omega' = \frac{2k_{nc}^2}{T} + 2\zeta_{nc}^2 \tag{5.6}$$

where $S_{\phi,nc}(\omega)$ represents the total uncorrelated phase noise and k_{nc} and ζ_{nc} are the jitter coefficients for the $1/f^2$ and $1/f^3$ uncorrelated noise process for differential sensing.

In summary, the total frequency counting uncertainty can be expressed as,

$$\sigma^2_{\frac{\Delta f}{f_0}} = \frac{8}{\pi\omega_0^2 T^2}\left[2\int_0^{+\infty} S_{\phi,c}(\omega')\, sin^2\left[\frac{\omega' T}{2}\right] \sin^2\left[\frac{\omega'(T+T_{set})}{2}\right] d\omega' + \int_0^{+\infty} S_{\phi,nc}(\omega')\, sin^2\left[\frac{\omega' T}{2}\right] d\omega'\right]$$

$$= 2\beta_{\zeta,c}\left(\frac{T_{set}}{T}\right)\zeta_c^2 + 2\beta_{k,c}\left(\frac{T_{set}}{T}\right)\frac{k_c^2}{T} + \frac{2k_{nc}^2}{T} + 2\zeta_{nc}^2$$

$$= 2\beta_{\zeta,c}\left(\frac{T_{set}}{T}\right)\zeta_c^2 + \frac{2k_c^2}{T} + \frac{2k_{nc}^2}{T} + 2\zeta_{nc}^2. \qquad (5.7)$$

The plot on total frequency counting uncertainty versus counting time can be shown as follows:

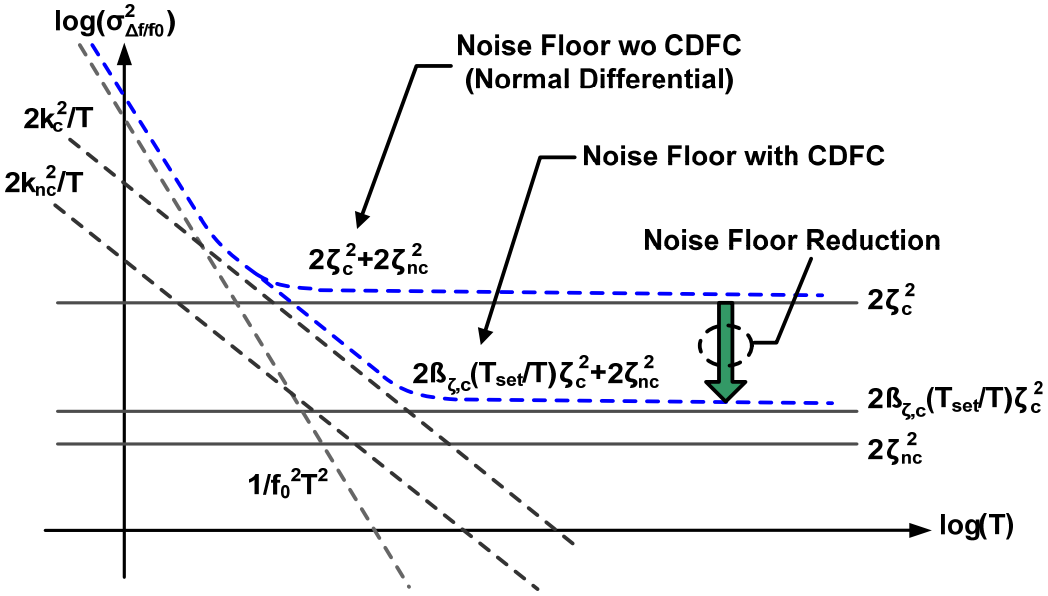


Figure 5.4: Frequency uncertainty for CDFC scheme and normal differential scheme

Figure 5.4 shows that with the same $\beta_{\zeta,c}$ value, the total noise floor reduction by the CDFC scheme is determined by the relative strength between the correlated and uncorrelated $1/f^3$ noise power for the specific circuit implementation. This agrees with our intuition that the uncorrelated $1/f^3$ noise, i.e., the flicker noise from the switch transistors, should be minimized.

Note that, for a practical implementation, the time reference of the frequency counting and the counter circuits also inevitably introduce noise sources which affect the total phase noise/jitter as measurement uncertainty. These noise sources can also be modeled as a linear combination of some $1/f^3$ and $1/f^2$ phase noise sources and superimposed onto the phase noise profile of the sensor oscillator system under counting. Moreover, in general these instrument-induced phase noises can be assumed to be correlated between differential sampling, since the same sampling circuit is in use. Therefore, the same noise reduction behaviors derived above for correlated $1/f^3$ and $1/f^2$ noise also apply to them.

5.2.2 A CDFC Circuit Implementation Example

In the actual implementation, to take the full advantage of this CDFC scheme, one needs to ensure the uncorrelated $1/f^3$ jitter power ζ_{nc}^2 is much smaller than the correlated jitter power ζ_c^2 before reduction. To be more specific, for the circuit schematic in Figure 5.2, the $1/f^3$ phase noise of the switches should be much smaller than the $1/f^3$ phase noise power of the cross-coupled active cores.

In this section, one design example will be presented to verify this implementation feasibility. The circuit schematic is shown in Figure 5.5 and the element values are shown in Table 5.1. The phase noise profiles based on ADS harmonic-balance simulator are shown in Figure 5.6.

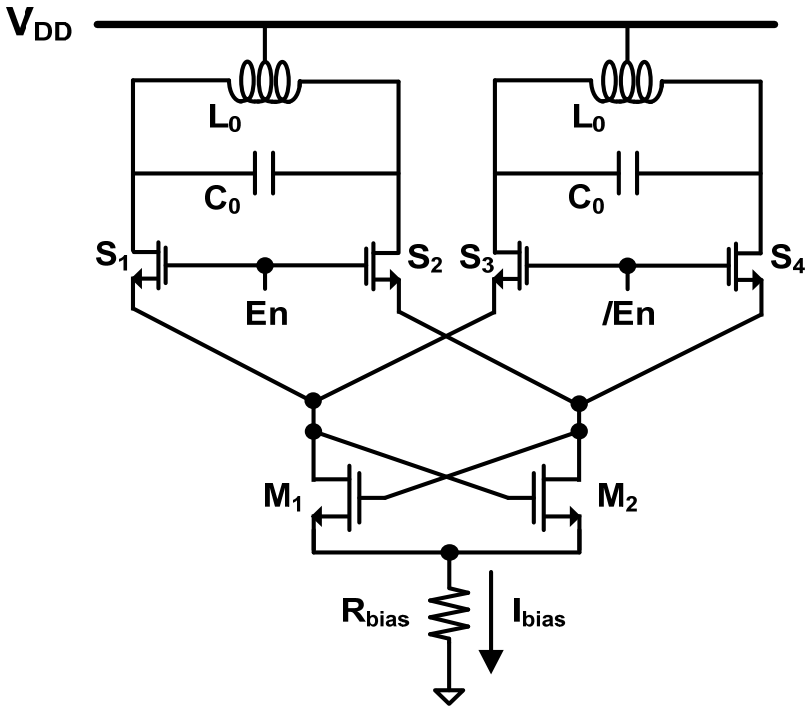


Figure 5.5: A simplified example schematic of CDFC implementation

Table 5.1 CIRCUIT ELEMENT DESIGN VALUES		
Cross Coupled Active Core	M_1 and M_2	W/L=480um/180nm, 1.8V Device
Switches	S_1, S_2, S_3 and S_4	W/L=96um/180nm, 1.1V Device
Biasing Resistor	R_{bias}	20 ohm
Tank Capacitor	C_0	7.02pF
Tank Inductor	L_0	2.65nH
Inductor Q (1GHz)/SRF		7.58/9.6GHz
Power Consumption	VDD/I	0.7V/11mA

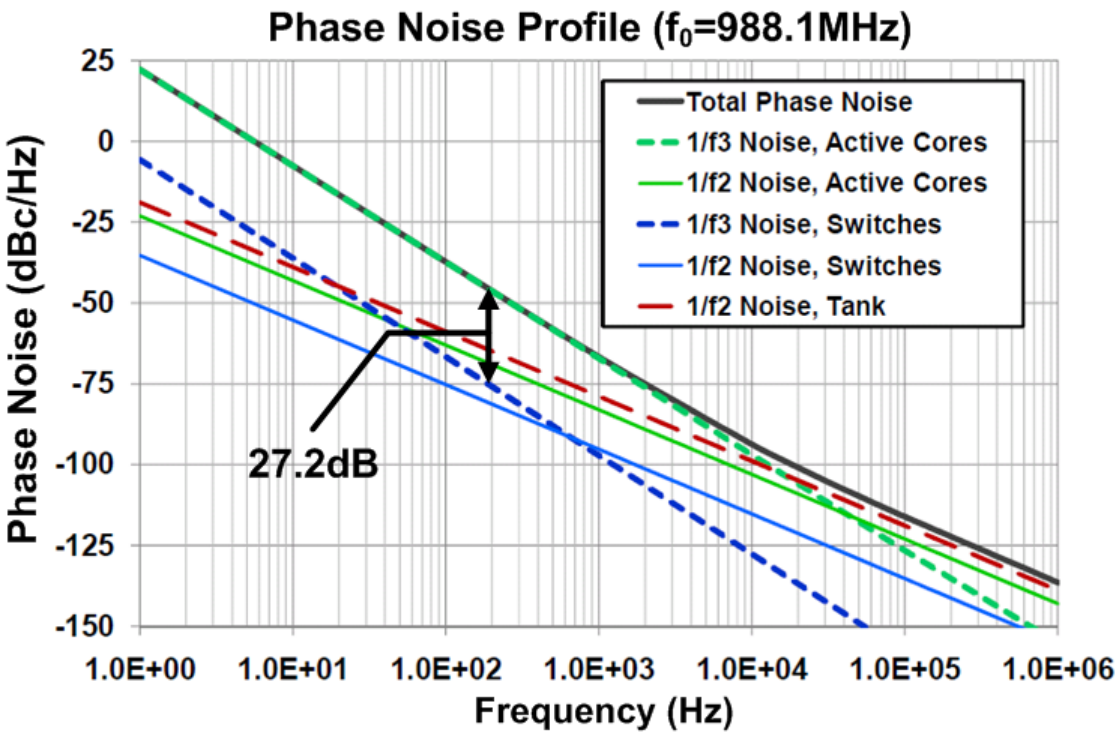


Figure 5.6: Simulated phase noise for the example circuit

Based on Figure 5.6, the total $1/f^3$ phase noise from the cross-coupled active core is 27.2dB less than the total $1/f^3$ phase noise from the switches. This means a maximum total noise reduction of 9.8dB predicted by CDFC calculation in Section 5.2.1 is indeed achievable.

5.3 Interleaving-N Correlated Double Frequency Counting and Fractional Frequency Counting

Section 5.2.1 shows that the maximum noise floor reduction by CDFC frequency counting scheme is around 9.8dB, when the counter reset delay is negligible compared with the total counting time. However, in Section 5.2.2, an example circuit implementation can

achieve 27.17dB noise power difference between the correlated 1/f^3 noise ($2\zeta_c^2$) and the uncorrelated 1/f^3 noise ($2\zeta_{nc}^2$). This suggests that, in typical practical implementations, our CDFC scheme may not be able to fully explore the potential of our proposed concept, i.e., suppression of the common-mode 1/f^3 noise by differential sensing. Therefore, to further increase the suppression effect as well as to push our sensor's sensitivity limit, an improved CDFC scheme needs to be investigated.

First of all, we need to understand the underlined physics which determines the fundamental limiting factor for the CDFC noise reduction ratio. Since the 1/f^3 phase noise is actually a 1/f noise in frequency, its autocorrelation function can be approximated as a relaxation process. A long reset interval between the two adjacent samples directly leads to a smaller correlation. This loss of correlation over time results in the imperfect cancellation of the 1/f^{3}noise. However, overlapping in operation time for the sensing oscillator and the reference oscillator is practically not achievable because of the direct coupling between the two tanks. Also, we cannot simply decrease the frequency counting time, since the relative frequency uncertainty and 1/f^3 noise reduction is essentially independent of counting time, shown in Figure 5.4.

However, one can actually divide the total frequency counting time of T (for both the sensing oscillator and the reference oscillator) into N sections and interleave them differentially in time. So we still have overall a measurement time of 2T, but it is distributed into N differential sensing pairs. Finally, the difference of the N differential sensing pairs can be added (averaged) together to yield the total frequency counting difference for this 2T measurement. This scheme is named as "Interleaving-N Correlated Double Frequency Counting", illustrated Figure 5.7. Note that a zero counter reset time is

assumed here for simplicity. Physically, in a much smaller time window (2T/N), this Interleaving-N CDFC scheme takes the differential sampling and adds up all the N differences as the overall differential result for the 2T measurement time. This leads to a significantly improved suppression effect on the low frequency noise.

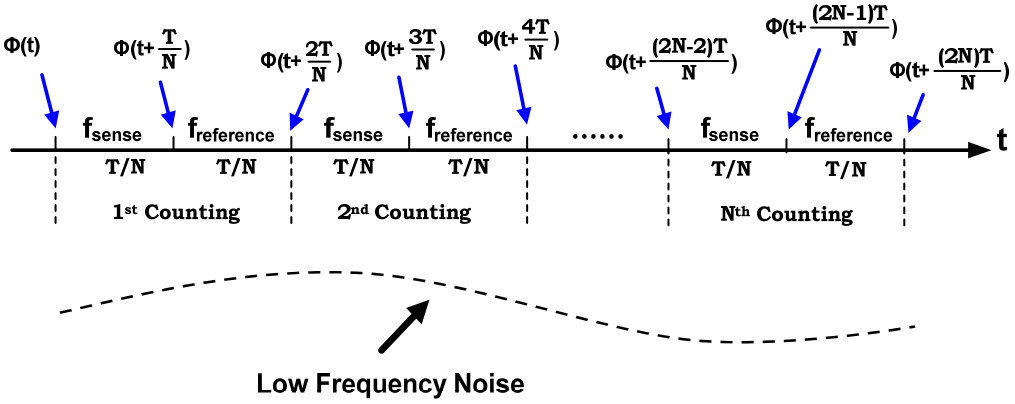


Figure 5.7: Interleaving-N CDFC scheme

We can model the correlated frequency-counting noise for interleaving-N CDFC scheme as,

$$\sigma^2_{\frac{\Delta f}{f_0},c} = \frac{1}{T^2\omega_0^2} E\left\{\left[\phi_{c,total}\right]^2\right\}$$

$$= \frac{1}{T^2\omega_0^2} E\left\{\left[\sum_{n=0}^{N-1} \phi_c\left(t + \frac{(2n+2)T}{N}\right) - 2\phi_c\left(t + \frac{(2n+1)T}{N}\right) + \phi_c\left(t + \frac{2nT}{N}\right)\right]^2\right\}$$

$$= \frac{16}{\pi\omega_0^2 T^2} \int_0^{+\infty} S_{\phi,c}(\omega') \cdot sin^4\left[\frac{\omega' T}{2N}\right] \cdot \frac{sin^2[\omega' T]}{sin^2\left[\frac{\omega' T}{N}\right]}\, d\omega' \qquad (5.8)$$

where $S_{\phi,c}(\omega')$ is the phase noise profile for the correlated noise.

In terms of the uncorrelated noise, its frequency-counting noise for interleaving-N CDFC scheme can be shown as,

$$\sigma^2_{\frac{\Delta f}{f_0},nc} = \frac{1}{T^2\omega_0^2} E\left\{\left[\phi_{nc,total}\right]^2\right\} = \frac{2}{T^2\omega_0^2} E\left\{\left[\sum_{n=0}^{N-1} \phi_{nc}\left(t + \frac{(2n+1)T}{N}\right) - \phi_{nc}\left(t + \frac{2nT}{N}\right)\right]^2\right\}$$

$$= \frac{2}{\pi\omega_0^2 T^2}\int_0^{+\infty} S_{\phi,nc}(\omega') \cdot \frac{sin^2[\omega' T]}{cos^2\left[\frac{\omega' T}{2N}\right]}\, d\omega' \tag{5.9}$$

where $S_{\phi,nc}(\omega')$ is the phase noise profile for the uncorrelated noise.

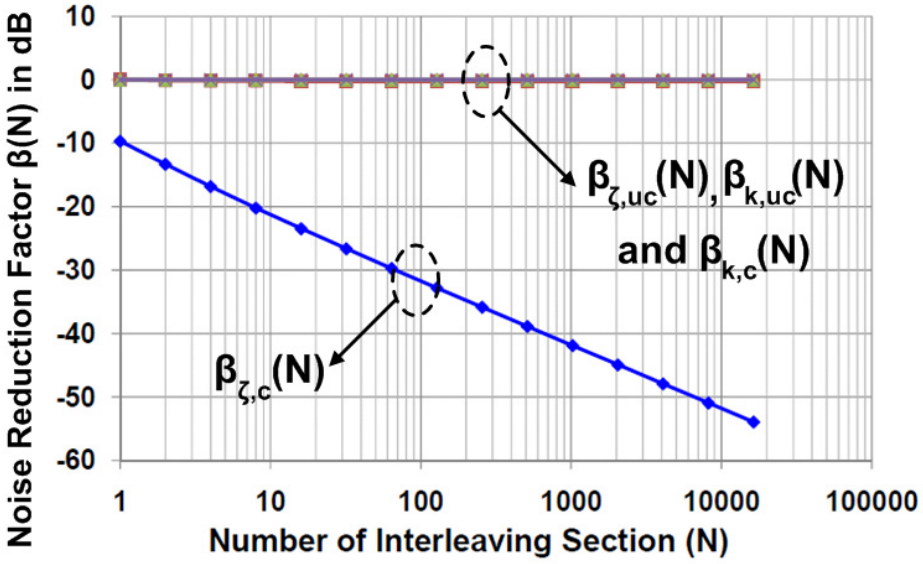
Therefore, the total frequency counting noise is given by,

$$\sigma^2_{\frac{\Delta f}{f_0},total} = \frac{2}{\pi\omega_0^2 T^2}[8\int_0^{+\infty}\left\{S_{\phi,c}(\omega') \cdot sin^4\left[\frac{\omega' T}{2N}\right] \cdot \frac{sin^2[\omega' T]}{sin^2\left[\frac{\omega' T}{N}\right]} + S_{\phi,nc}(\omega') \cdot \frac{sin^2[\omega' T]}{cos^2\left[\frac{\omega' T}{2N}\right]}\right\} d\omega'$$

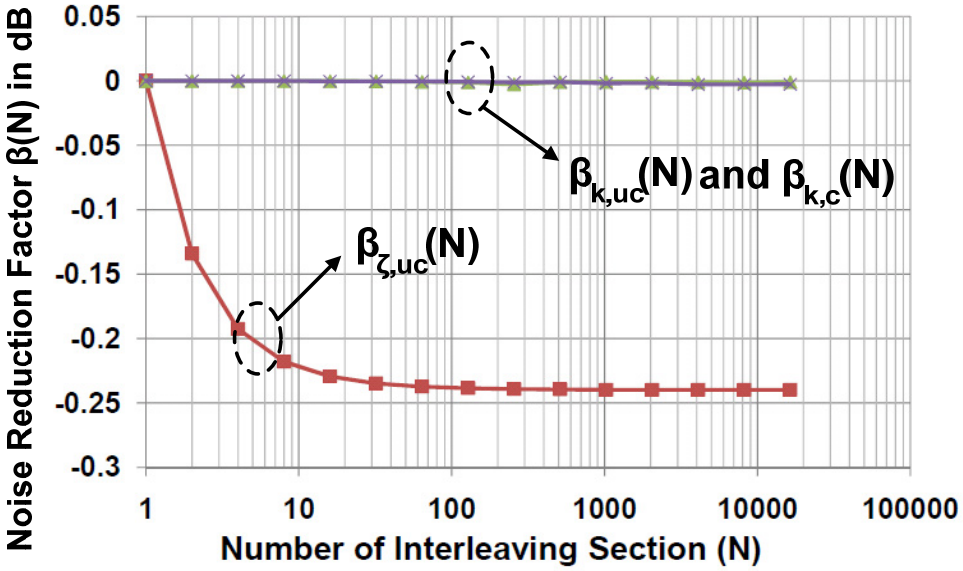
$$= 2\beta_{\zeta,c}(N)\zeta_c^2 + \frac{2\beta_{k,c}(N)k_c^2}{T} + \frac{2\beta_{k,nc}(N)k_{nc}^2}{T} + 2\beta_{\zeta,nc}(N)\zeta_{nc}^2 \tag{5.10}$$

where $\beta_{\zeta,c}(N)$, $\beta_{k,c}(N)$, $\beta_{k,nc}(N)$ and $\beta_{\zeta,nc}(N)$ are the noise reduction factor for the correlated $1/f^3$ phase noise, correlated $1/f^2$ phase noise, uncorrelated $1/f^2$ phase noise and uncorrelated $1/f^3$ phase noise, respectively.

These noise reduction factors is calculated and plotted in the Figures 5.8. Note that at N=1 Interleaving-N scheme is equivalent to the normal CDFC scheme as discussed in Section 5.2.1.



(a)



(b)

Figure 5.8: Noise reduction factor for four types of noise sources

As shown in the Figure 5.8 (a), noise reduction factor $\beta_{\zeta,c}(N)$ for the correlated $1/f^3$ noise continuously decreases with respect to the interleaving number N, which is within our expectation. Also, from both Figure 5.8 (a) and (b), the noise reduction factors $\beta_{k,c}(N)$ and $\beta_{k,nc}(N)$ for correlated and uncorrelated $1/f^2$ noise remains as 0dB. This is because for both cases, the $1/f^2$ phase noise is essentially white for frequency noise, which leads to uncorrelated frequency counting samples. Therefore, the Interleaving-N CDFC scheme will not result in any rejection on this noise. However, for $\beta_{\zeta,nc}(N)$, the noise reduction factor for uncorrelated $1/f^3$ noise, is interesting. It starts to decrease with number of intersections and settles at around -0.24dB. This is because by differential sensing scheme, the adjacent samples from the same $1/f^3$ noise source have an interval of T/N (one sampling period on the other oscillator). This time interval makes these $1/f^3$ noise samples slightly

uncorrelated. This small uncorrelated noise part is averaged out through summing all the N samples in the Interleaving-N scheme, which leads to this -0.24dB reduction.

Therefore, by interleaving-N CDFC scheme, the correlated $1/f^3$ is significantly suppressed by $\beta_{\zeta,c}(N)$. We can now plot the frequency counting uncertainty plot with respect to counting time T.

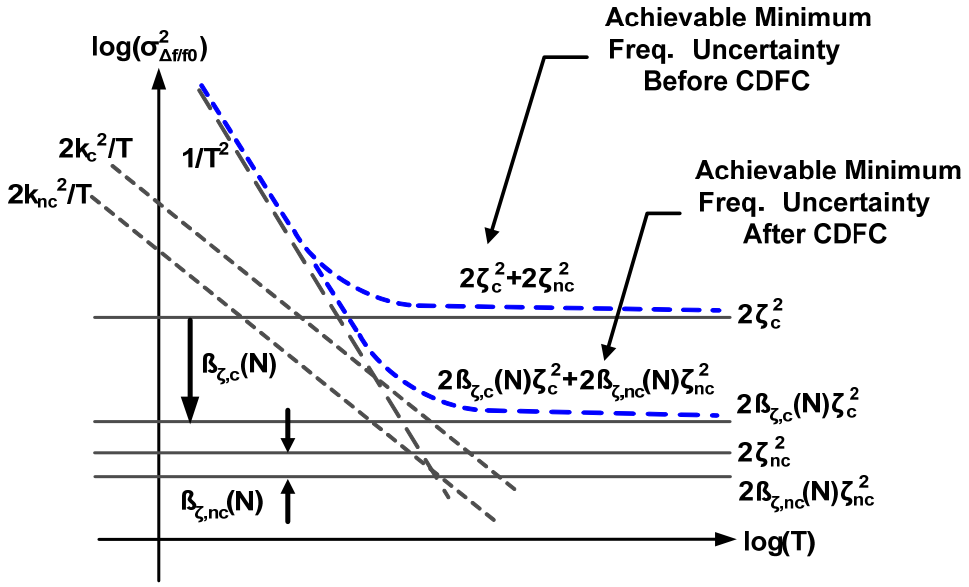


Figure 5.9: Frequency uncertainty for Interleaving-N CDFC scheme and normal differential scheme

Then by Interleaving-N CDFC scheme, the correlated $1/f^3$ noise power, as the dominant noise source for frequency counting uncertainty, can be suppressed for a factor larger than that of the normal CDFC scheme. The ultimate noise floor for this frequency counting scheme is limited by the uncorrelated $1/f^3$ noise power, when the correlated $1/f^3$ noise power is fully suppressed by a large enough N.

On the other hand, since every differential counting period is decreased by N the frequency counting error due to uncertainty principal is thus proportionally increased by N,

which means only N/T Hz can be resolved. To achieve a frequency resolution of 1/T Hz or even lower, a fractional frequency counter must be implemented, shown in Figure 5.10. The two examples are physically equivalent, since either the measurement signal or the counting time reference is delayed by T_D through a delay chain.

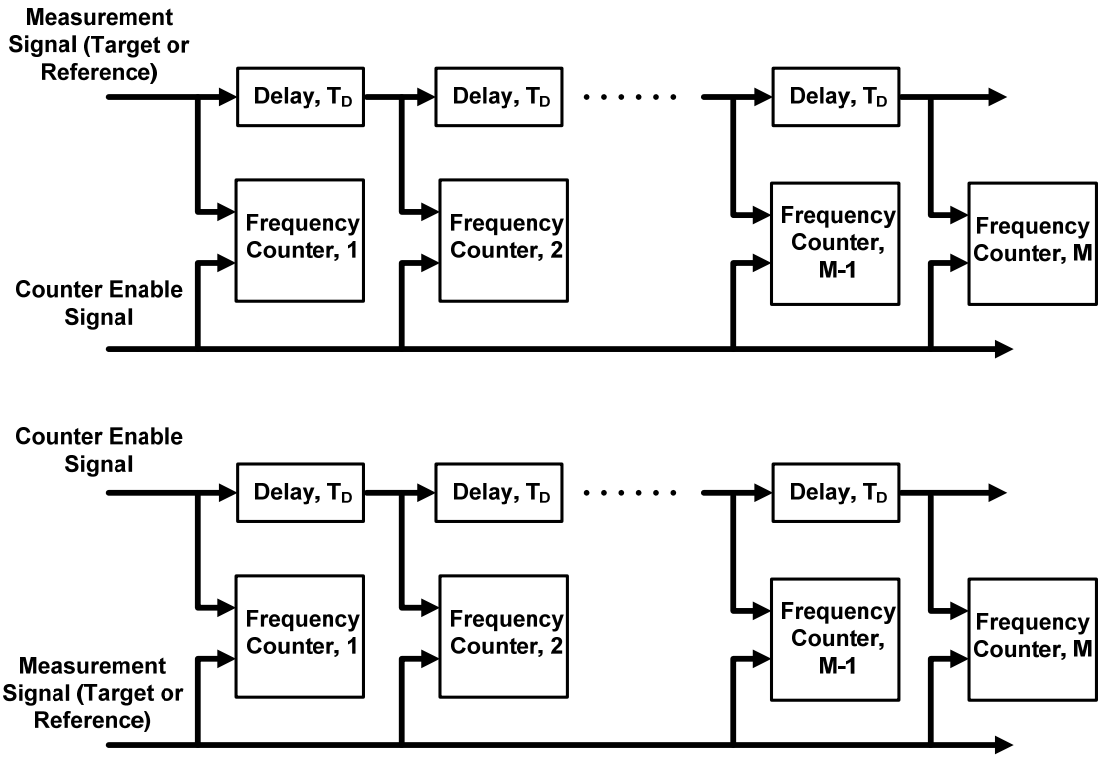


Figure 5.10: Two fractional frequency counter implementation examples

The example with delayed reference clock (counter enable signal) will be used to show the operation principle of a fractional frequency, depicted in Figure 5.11.

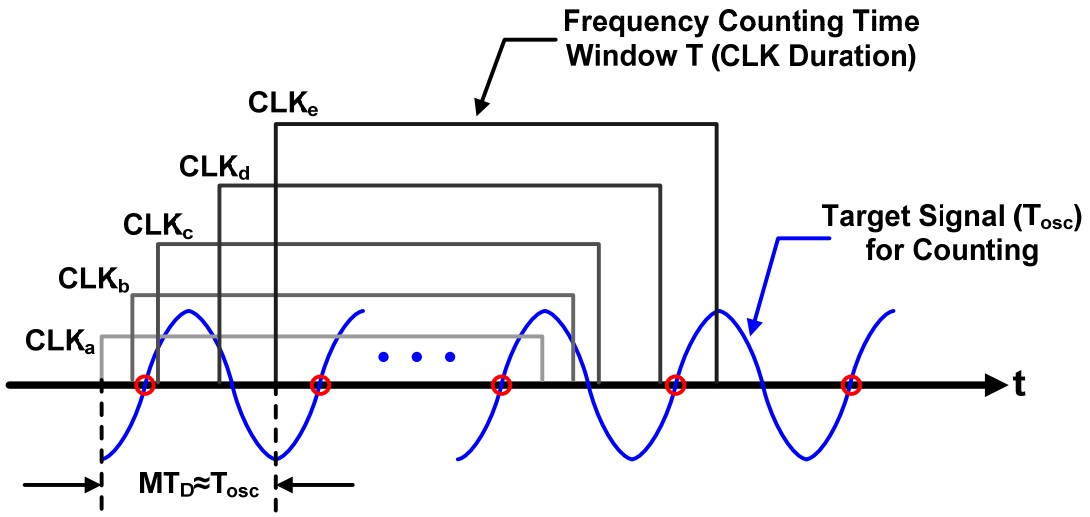


Figure 5.11: Fractional counter operation principle

If there are M counter cell, the total delay of the delay chain (MT_D) should be set at a value slightly larger than one oscillation period of the measurement signal (T_{osc}). This delay scheme essentially creates M reference clock or target signals with multiple phases for sampling. Assume CLK_a and CLK_e are the first clock and the last clock from the delay chain. And assume CLK_b and CLK_c are adjacent multi-phase clocks. Then for multi-phase clock with a counting period of T, the numbers of signal positive transitions (highlighted by red circle) are given by:

$$\#\ of\ transitions = \begin{cases} N+1, & from\ CLK_a\ to\ CLK_b \\ N, & from\ CLK_c\ to\ CLK_d \\ N+1, & from\ CLK_d to\ CLK_e \end{cases}. \qquad (5.11)$$

If we assume there are M_1 delayed clocks from CLK_a to CLK_b, M_2 delayed clocks from CLK_c to CLK_d and M_3 delayed clocks from CLK_d to CLK_e (M_1+M_2+M_3=M). Therefore, the averaged number of counted transitions is:

$$Averaged\ \#\ of\ transitions = \frac{(N+1)M_1 + (N)M_2 + (N+1)M_3}{M}. \qquad (5.12)$$

Equation (5.12) shows that within this T counting time, the minimum discernable transition number is 1/M. So the overall frequency resolution is 1/MT, which is M times smaller than the single-phase frequency counting system (resolution of 1/T).

In summary, by using enough parallel frequency counters (M), one can always keep the frequency resolution (1/MT) due to the uncertainty principle low enough. This essentially shows that only the relative frequency uncertainties introduced by $1/f^2$ and $1/f^3$ phase noises are actual fundamental limit for the frequency counting accuracy.

5.4 Noise Shaping Function in Correlated Double Frequency Counting

Through previous sections, we have derived the mathematical expressions on the relative frequency counting uncertainty (normalized accumulated jitter within a counting time T) for different proposed counting schemes.

These expressions show that for different counting schemes the total frequency counting uncertainty is always expressed as integrating (from 0 to +∞) the product of the correlated phase noise profile conditioned with a certain window function. The window function is determined by particular counting scheme in use. We can therefore define those window functions as noise shaping functions (NSF) and further summarize them in Table 5.2 and Figure 5.12.

Table 5.2 Noise Shaping Functions (NSF)

Normal Differential (No Correlation)	$8\,sin^2\frac{\omega T}{2}$

Normal CDFC (T_{set}/T) (Correlated Noise)	$16\,sin^2\left[\frac{\omega T}{2}\right]\sin^2\left[\frac{\omega T}{2}\left(1+\frac{T_{set}}{T}\right)\right]$
Interleaving-N CDFC (Correlated Noise)	$16\,sin^4\left[\frac{\omega T}{2N}\right]\cdot\frac{sin^2[\omega T]}{sin^2\left[\frac{\omega T}{N}\right]}$

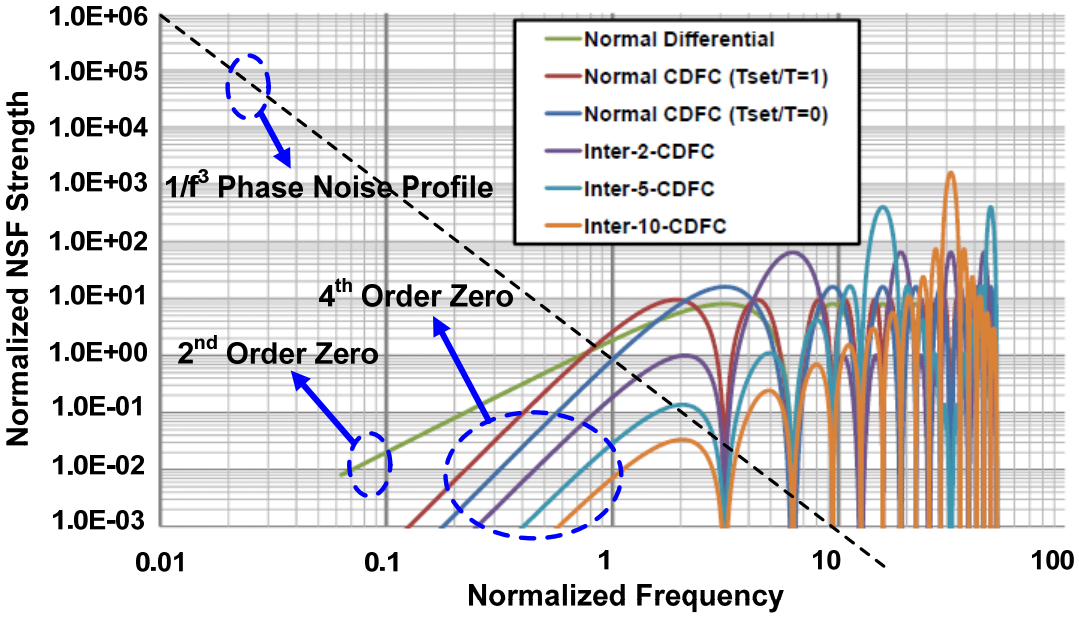


Figure 5.10: Noise Shaping Functions with $1/f^3$ phase noise shown as the dashed curve

The above plot shows the mechanism for a correlated double sampling (CDFC) scheme to yield a smaller frequency counting errors from a mathematical point of view. All the NSFs for CDFC schemes present a 4^{th}-order zero at zero frequency, while NSF for a normal differential sensing scheme only has a 2^{nd}-order zero. Since the major power content for a $1/f^3$ phase noise exists a low frequencies, this high zero order greatly attenuates its total integrated noise power. Moreover, a more efficient CDFC scheme, such as Interleaving-N CDFC with a larger N value, has a further more attenuated noise shaping

function strength at low frequencies. This therefore leads to a greater suppression on the total integrated noise power, shown in Figure.5.8 (a).

5.5 Chapter Summary

This chapter first formulates the relationship between the phase noise, accumulated jitter and the frequency counting uncertainty for our sensor scheme. Based on this, the minimum achievable frequency counting accuracy, i.e., the sensor sensitivity, is found to be limited by the $1/f^3$ noise of the measured sensor oscillator.

Since we have adopted the differential structure for the sensor embodiment, a new sensor circuit topology is proposed to achieve correlation between the $1/f^3$ noise from the sensing oscillator and the reference oscillator. This leads to a novel noise suppression technique, Correlated-Double-Frequency-Counting (CDFC) scheme, which significantly decreases the sensor noise floor, and therefore improves the sensor sensitivity without significant power consumption overhead. A practical sensor circuit implemented in a standard 65nm CMOS process is shown as a design example for this scheme. Moreover, the fundamental limit on correlated $1/f^3$ noise suppression by the proposed CDFC scheme is also studied.

In order to further increase the suppression on the correlated $1/f^3$ phase noise, a modified scheme, called Interleaving-N Correlated-Double-Frequency-Counting, is then proposed. A complete study on the noise reduction factors for different types of noise sources is presented, which verifies the viability of the Interleaving-N CDFC scheme. On the other hand, to improve the frequency counting resolution limited by uncertainty

principles, a fractional counting scheme with two implementation examples is proposed and presented.

Finally, the noise shaping functions for different frequency counting schemes are defined and plotted, which further verify the functionalities of the CDFC and Interleaving-N CDFC scheme from a mathematical perspective.

Chapter 6: Broadband Precision Phase and Amplitude Synthesis

6.1. Introduction

In the previous chapters, application of high-accuracy frequency synthesis in biosensors has been presented. In this chapter, we will cover the high-precision phase synthesis techniques designed for wireless communications and radar applications with an emphasis on integrated phased array system.

In general, a phased array system is defined as a group of antennas in which the relative phases of the received/transmitted signals are programmed in such a way that the effective radiation pattern of the array is reinforced in a desired direction and suppressed in the undesired directions [41] [42]. A conceptual view of a phased array receiver is shown in Figure 6.1.

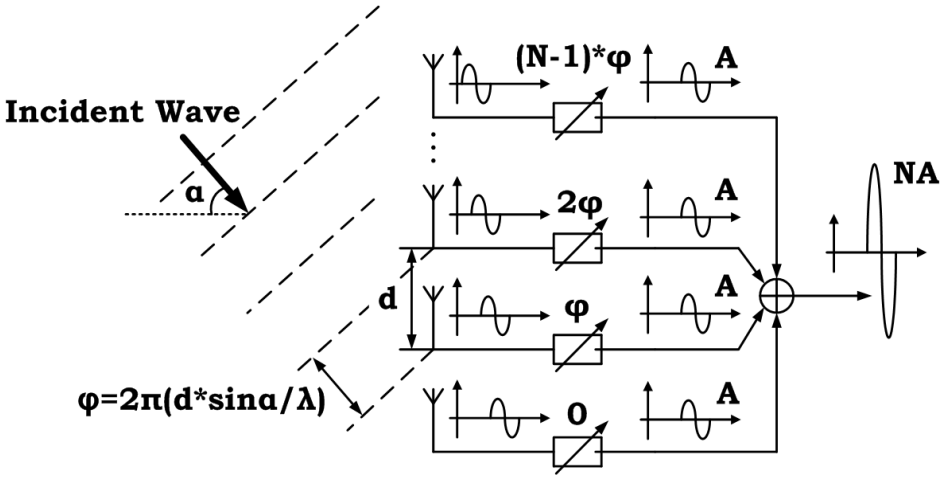


Figure 6.1: Schematic for a phased-array receiver system (constructive combining)

Based on reciprocity, this phased array concept, i.e., beam forming by adjusting a constant phase-offsets among array elements, can be used for both receiving and transmitting purpose.

Assuming the incident wave is coming toward the array with an incident angle of α, and the array elements are separated by a distance d, there exists a time-delay between the elements when receiving the incident signal. For a narrow-band signal, often used in wireless communication and radar applications, this delay difference can be approximated as a phase difference between each adjacent array elements, $\Delta\theta$, given as,

$$\Delta\theta = 2\pi \cdot \frac{dsin(\alpha)}{\lambda} = \pi sin(\alpha)_{d=\frac{\lambda}{2}}. \tag{6.1}$$

If we adjust the phase offset φ among the array elements to cancel this phase difference, signals received by N elements will be exactly in-phase and summed together as coherent addition. This is presented as a constructive beam-forming, which shows as a peak on the synthesized array pattern.

On the other hand, the same phase offset setting φ, if there is an unwanted signal incident at angle β for the array, this undesired signals are added out-of phase, shown in Figure 6.2. This thus forms a destructive beam-forming which effectively suppresses the unwanted signals. Depending on the values of β and φ, it can be presented as a null or an attenuation point compared with the main beam in the array pattern.

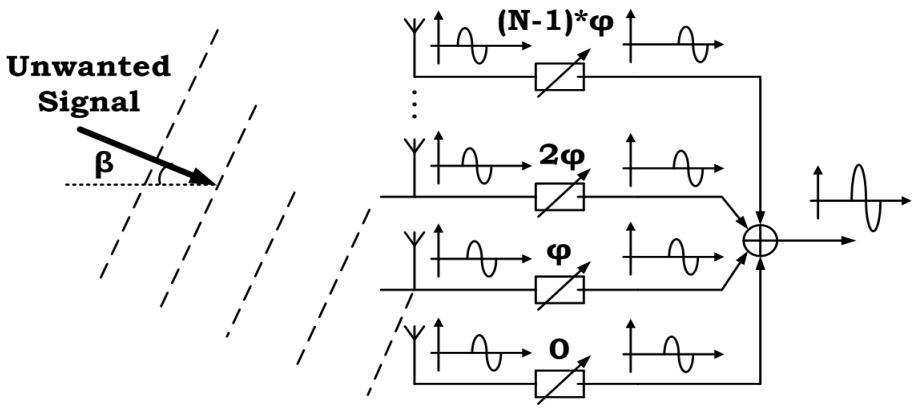


Figure 6.2: Schematic for a phased-array receiver system (destructive combining)

The array quality for spatial filtering is characterized by an array peak-to-null ratio (PNR), as

$$PNR = 20\log_{10}\left(\frac{Amp_{peak}}{Amp_{null}}\right). \tag{6.2}$$

Theoretically, at the null point, this array peak-to-null ratio is infinity, since ideal array pattern can achieve zero amplitude. However, due to many practical implementation issues, particularly mismatches both in phase and in amplitude among array elements, this zero null point can hardly be achieved in reality. Moreover, these mismatches also introduce distortions in the entire array pattern, which cannot be determined a priori.

This chapter is devoted to investigate this non-ideal array pattern in a practical implementation. First, the array degradation effect due to random phase/amplitude mismatches among array elements will be introduced. In particular, the phase mismatches will be discussed in details. To address this degradation issue, a high-resolution phase/amplitude synthesis scheme will be proposed and presented, which provides enough degree of freedom to compensate those mismatches. As an implementation example, a dual-band quad-beam CMOS phased array receiver system which covers a tritave

bandwidth of 6-to-18GHz will be presented. In the end, a summary section will conclude this chapter.

6.2. Array Degradation due to Random Phase/Amplitude Mismatches

A practical implementation of a phased array system presents inevitable mismatches for both phase and amplitude among the array elements. For a 4-element phased array system, without loss of generality, we can assume the phase mismatches have an identical independent Gaussian distribution among the array elements as $N(0,\sigma_\theta)$, while the amplitude mismatches also experience an identical independent Gaussian distribution among the array elements as $N(0,\sigma_X)$. If we take the 90° null-point, a cumulative distribution function (CDF) of PNR for different combination of $(\sigma_\theta, \sigma_X)$ can be calculated and plotted in Figure 6.3.

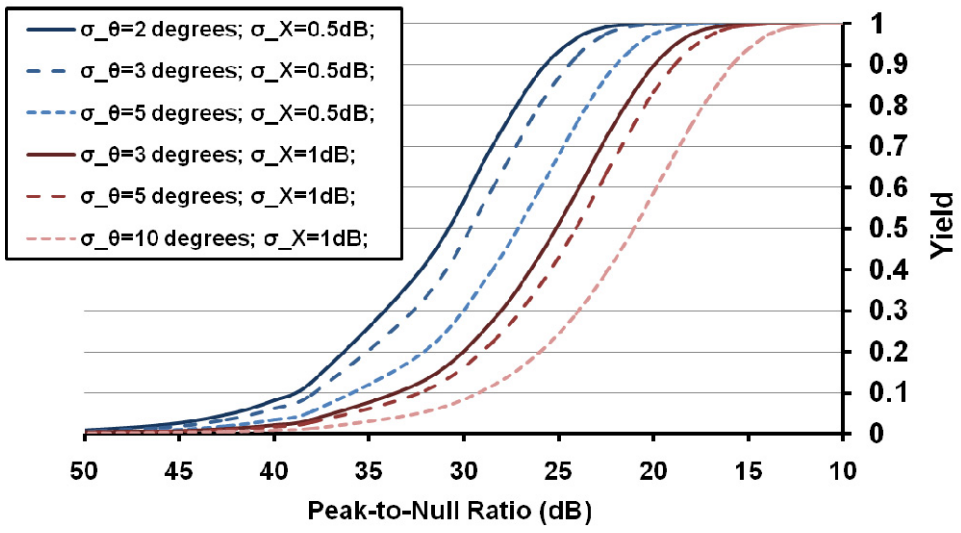


Figure 6.3: cumulative distribution function of PNR

From this figure we can see that if a peak-to-null ratio of 25dB is required with a 90% yield, for a σ_X of 0.5dB, σ_θ must be smaller than 3˚, which is a very challenging specification for a conventional phased array implementation.

A high quality array with high yield thus requires a tight control of those phase/amplitude errors among array elements. In reality, those errors are inevitable mainly due to the following three reasons. First of all, the actual circuits experience performance changes due to the process, voltage supply and temperature (P.V.T) variations. Secondly, there exist random mismatches due to fabrication even for well-matched circuits implemented on the same chip. Thirdly, the attenuation mismatches among antenna feed and the delay difference/clock skews in the clock reference signal distribution also directly affect the amplitude and phase matching among array elements. What needs to be emphasized is that the P.V.T. variations and the process mismatches are exacerbated in a deep sub-micron CMOS process, while the feed attenuation and reference delay issues are more problematic for a large array implementation where complicated distribution routings are often required.

Therefore, in a practical array implementation, to achieve a high-quality array pattern and a high array yield, we desire an array element with a reliable and independently-controllable precision phase/amplitude calibration capability for mismatch compensations.

6.3. High-Resolution Phase/Amplitude Synthesis Scheme

As presented in the previous section, the precision in phase and amplitude is crucial to achieve a desired array pattern. However, since the aforementioned mismatches are not known a priori, this suggests that in a robust phased array system design, the phase and

amplitude settings must have a high resolution and a high degree of freedom to compensate for those mismatches. Moreover, the phase/amplitude synthesis block must be intrinsically wideband to maintain its functionality in a broadband phased array. Also, the phase/amplitude synthesis block itself should be P.V.T. (Process, supply Voltage and Temperature) insensitive to provide a robust compensation performance. Furthermore, it is desired that the phase/amplitude synthesis scheme can be fully scalable for upgrading to a higher resolution implementation without excessive overheads. In addition, during the mismatch trimming by the phase/amplitude synthesis, the major receiver performance, such as the noise figure and the linearity, should not be severely compromised.

Considering all the facts described above, we propose a high-resolution phase/amplitude synthesis scheme shown in Figure 6.4. This can be considered as an extension of the LO phase shifting architecture, which instead of providing only limited discrete phase setting, supporting a full control on both phase and amplitude on the LO signal with a high resolution. Thus the amplitude and the phase of the baseband signal can be well tuned to achieve the optimum array pattern.

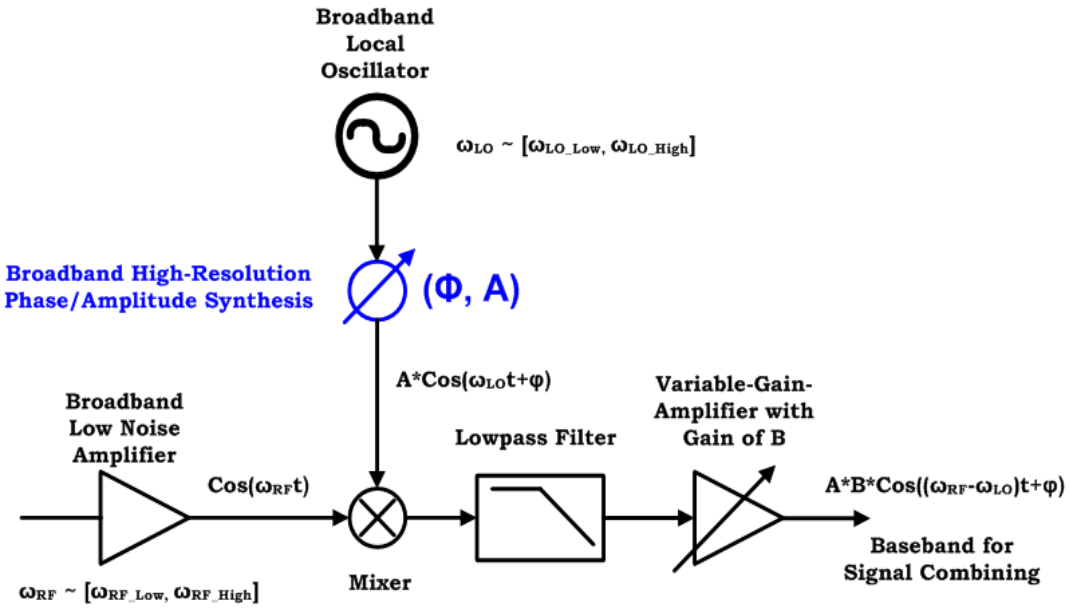


Figure 6.4: High-resolution phase and amplitude synthesis

One generic synthesis option is through Cartesian interpolation, where the quadrature LO signals are scaled by variable gain amplifiers (VGA) independently. And the summation of the scaled signals results in an LO signal with desired amplitude and phase, shown in Figure 6.5. The interpolation resolution is thus limited by the VGA's resolution.

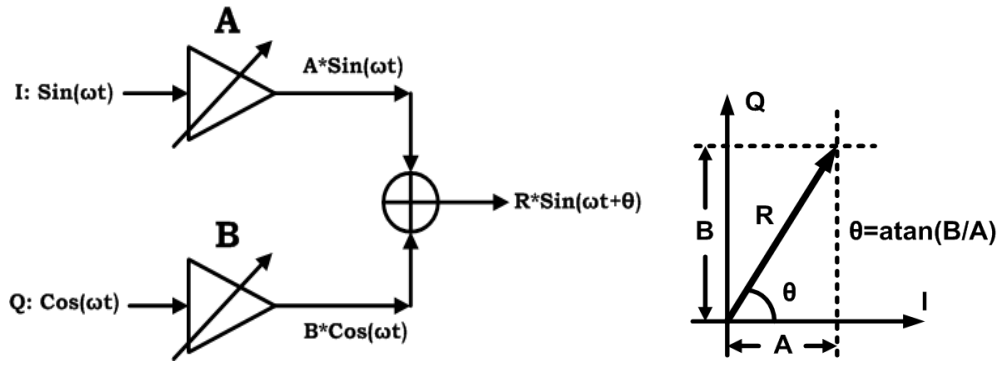


Figure 6.5: Cartesian interpolation for high-resolution phase and amplitude synthesis

However, conventional VGA designs have issues which greatly limit their applications in this scheme. For example, many VGA topologies have their AC bandwidth and/or

output DC values vary for different gain settings. Also, many VGA designs have their gain settings highly rely on modeling accuracy, and therefore are very P.V.T. dependent and difficulty for scaling. To overcome these impediments, we have proposed our VGA configuration as follows [43]:

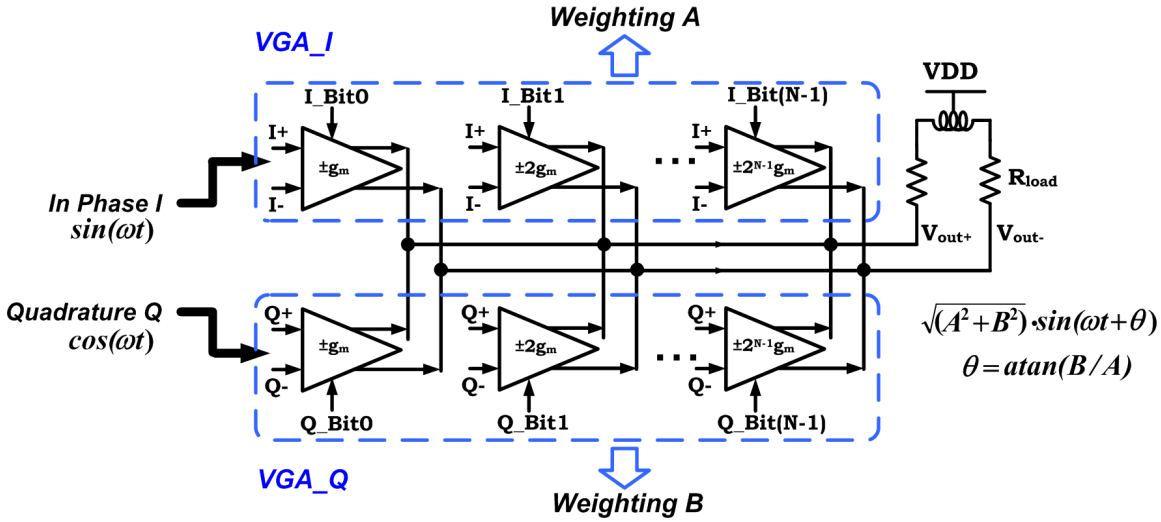


Figure 6.6: Proposed VGA and phase/amplitude synthesis circuits

The VGA circuit architecture is composed of N binary-weighted transconductance cells, which are individually functioning as a polarity selector. The digital programming bits thus control the polarities of the N cells, which then set the total VGA gain. The scaled in-phase (I) and the quadrature (Q) LO signals are summed together to form the desired phase and amplitude. Therefore, the interpolation accuracy of our proposed VGA is only dependent on the matching among the binary cells for P.V.T. insensitive operation, instead of relying on absolute value of the active/passive devices. The inter-cell matching can be significantly improved by well-known techniques, such as common-centroid layout. And this scheme can be simply scaled by adding more binary-weighted transconductance cells.

The transconductance unit cell (polarity-selector) is implemented as the circuit shown in Figure 6.7. Switches S and /S are used to control the output signal polarity. At either polarity setting, there will be one branch turned on and the other branch turned off. Therefore, both the previous stages and the following stages see exactly one ON differential pair and one OFF differential pair for every unit gain cells. This guarantees that for the entire VGA, the DC current consumption, the output DC biasing point, and the input/output parasitic capacitance (AC bandwidth) remain constant for arbitrary VGA gain settings.

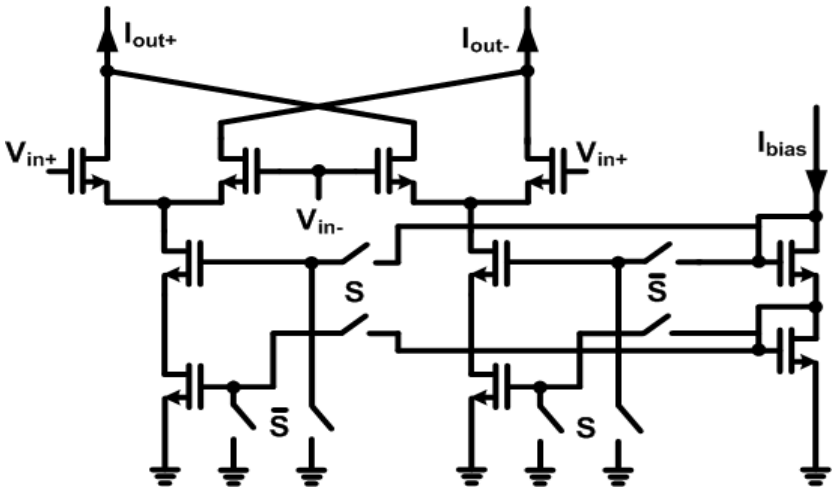


Figure 6.7: Polarity-selector transconductance unit cell

6.4. A 6-to-18 GHz Dual-band Quad-beam CMOS Phased Array Receiver System

In this section, as an implementation example, a 6-to-18 GHz dual-band quad-beam phased array receiver system designed in CMOS is presented. This study investigates the

possibility of realizing the entire multi-band multi-beam phased array receiver on a low-cost CMOS chip with its full potential to be scaled to a very large scaled array (more than 10^6 elements). To provide a full capability for mismatch compensation, this array implementation adopts the high-precision phase/amplitude synthesis scheme introduced in the previous sections, whose functionality will be shown with the measurement results.

6.4.1 System Architecture

In our phased array receiver system, antennas and broadband GaN LNAs receive the incoming signal in the horizontal (HP) and vertical (VP) polarizations, which are then fed into the CMOS receiver through two separate signal paths. The top-level array system architecture is shown in Figure 6.8.

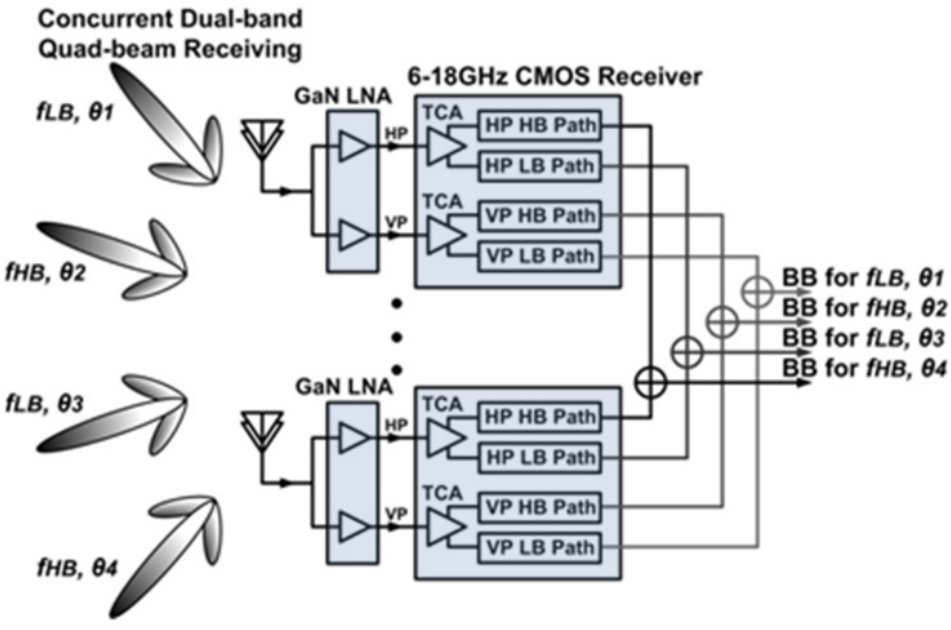


Figure 6.8: Proposed broadband multi-beam phased array system

The system architecture for the CMOS phased array receiver is shown in Figure 6.9. On the CMOS chip, a wideband tunable concurrent amplifier (TCA) first splits the RF signal (either HP or VP) into low-band (LB) from 6 to 10.4GHz and high-band (HB) from 10.4 GHz to 18GHz. The LB and HB RF signals are then separately down-converted by two mixers to IF and then baseband. Both LB and HB paths have independent on-chip frequency synthesizers, which provide two local oscillator signals (LO_1 and LO_2) each. The phase and amplitude synthesis scheme is implemented for the LO_2 signal achieved with a 10-bit Cartesian interpolation control [45].

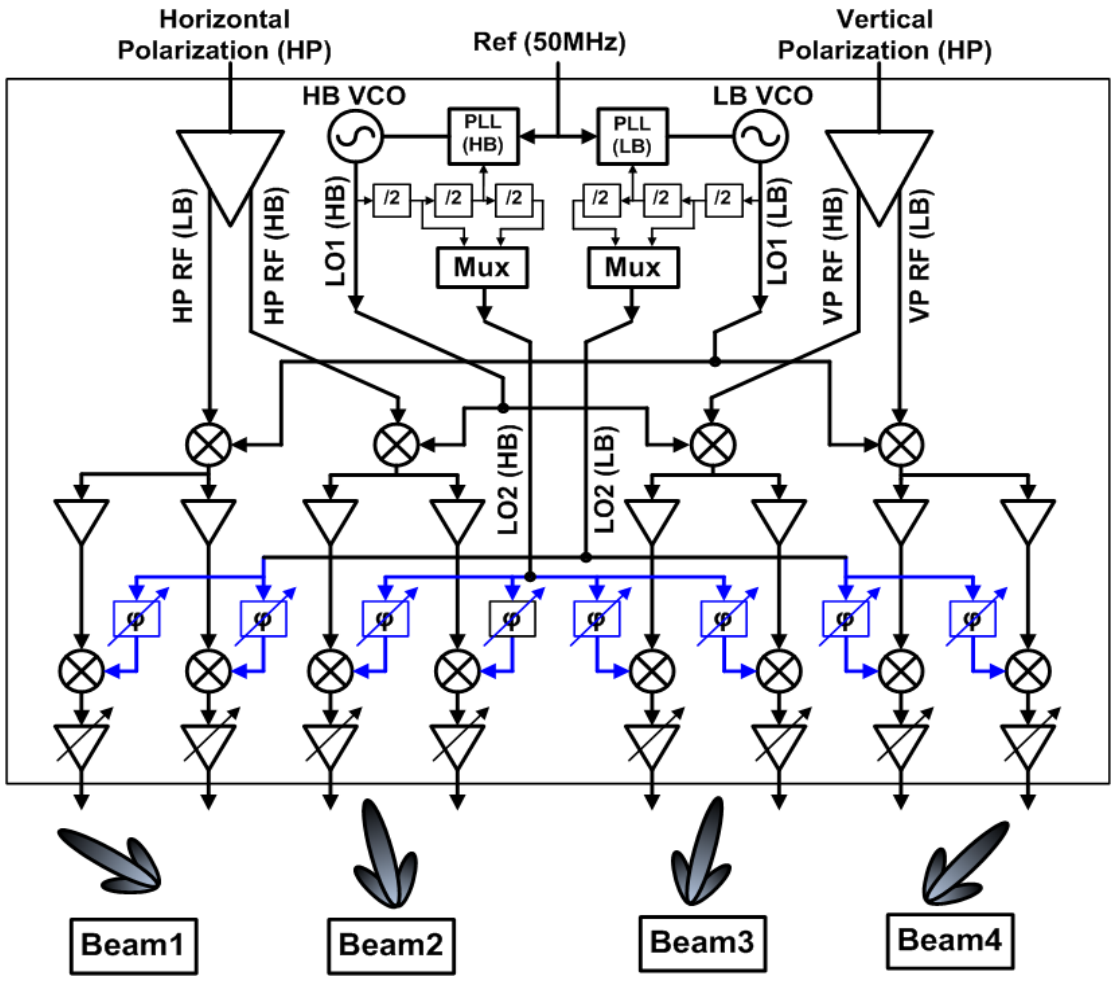


Figure 6.9: CMOS broadband multi-beam phased-array receiver architecture

By this system architecture, the receiver is capable of forming four independent beams simultaneously at two different frequencies in the 6 to 18GHz bandwidth. The CMOS chip is implemented in the IBM8RF 130nm CMOS process with a chip area of 3.0mm×5.2mm. Its microphotograph is shown in Figure 6.10, with the four signal paths and the dual-band LO path highlighted.

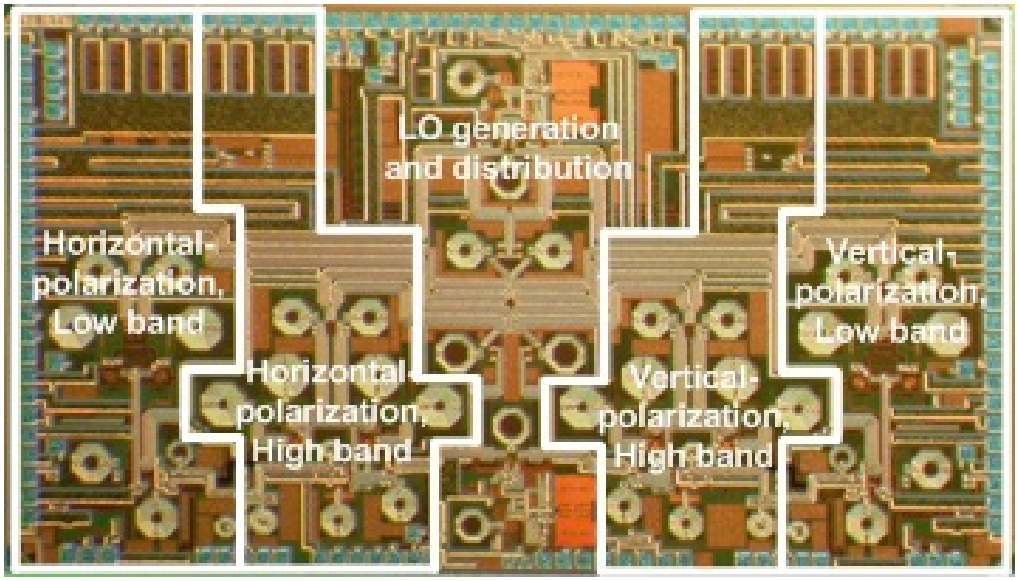


Figure 6.10: Microphotograph of the CMOS phased array receiver

6.4.3 Phased Array Measurement Results

The continuous phase amplitude interpolation has been measured for RF frequencies of 6GHz, 10.4GHz, 14GHz, and 18GHz, summarized in Table 6.1. Figure 6.11 shows the measured phase and amplitude interpolation results for the baseband signal at an RF frequency of 18GHz.

Table 6.1. Phase/Amplitude Interpolation Summary				
RF Freq.	**Phase Error RMS**	**Phase Error Max.**	**Amp. Variation RMS**	**Amp. Variation Max**
6GHz	0.5°	2.5°	0.4dB	1.3dB
10.4GHz	0.2°	1.3°	0.2dB	1.0dB
14GHz	0.3°	1.4°	0.2dB	1.3dB
18GHz	0.2°	1.5°	0.5dB	1.5dB

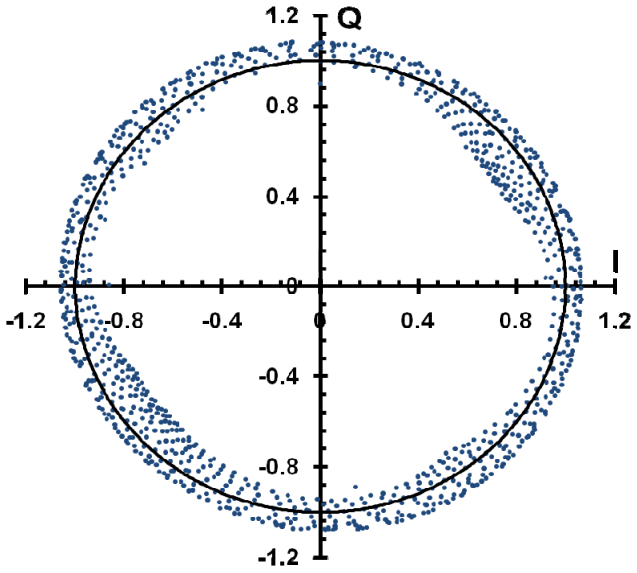


Figure 6.11: Measured baseband constellations at the RF frequency of 18GHz.

For the continuous interpolation case shown in Figure 6.11, the incident beam is assumed to come for any direction with respect to the array. Therefore, its phase error value is half of the phase interpolation step size. As a result, our phased array receiver chip achieves a maximum RMS phase error of 0.5˚ within a maximum RMS amplitude variation of 1.5dB across the 6 to 18 GHz bandwidth. This dense phase interpolation, together with

the amplitude adjusting possibilities ensures the full mismatch compensation capability of our receiver system.

The aforementioned phase/amplitude mismatches can be classified into two types, i.e., the mismatch within the receiver element and the mismatch among array elements.

The former mismatches contain two types of interpolation errors for our scheme.

The first type of error arises from practical limits on the LO I/Q signal phase and amplitude matching. This error is exacerbated by inevitable mismatches in the LO networks. The other type of error is the zero crossing distortion of the phase-shifted LO signal with excessive harmonics. I/Q- interpolating phase rotators are inherently dispersive systems, which offset the input by a constant phase shift instead of a constant group delay, shown in Figure 6.12. Note that $3\omega_0$ is assumed as the dominant harmonic for differential circuits. The resulting zero crossing errors in the dispersed LO waveforms lead to baseband phase errors through downconversions by switching mixers.

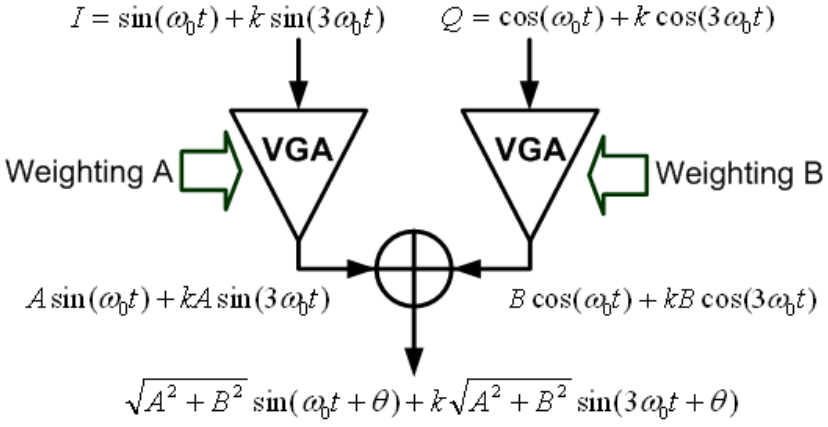


Figure 6.12: The phase rotator causes waveform dispersion for the input quadrature signals with multiple frequency contents.

However, dense phase interpolation can compensate both errors by choosing the appropriate I/Q weightings. The measured phase errors before and after compensations are shown in Figure 6.13 and Figure 6.14 for the case of LO I/Q mismatch and the case of non-sinusoidal LO, respectively. Both cases are based on 360º full range interpolation with a phase step of 11.25º.

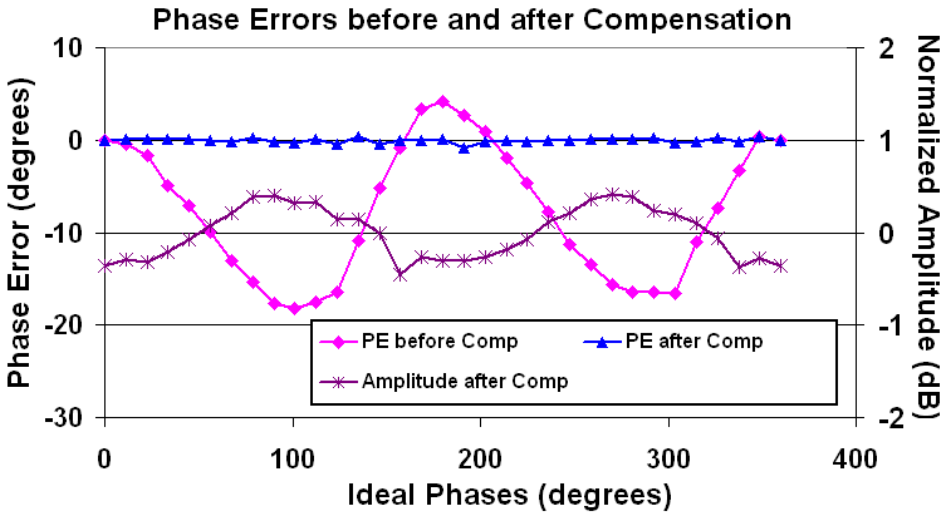


Figure 6.13: Phase errors before/after compensation for LO I/Q mismatch (f_{RF} = 10.4GHz). Within a 0.45dB amplitude variation, a 0.9º maximum phase error is achieved.

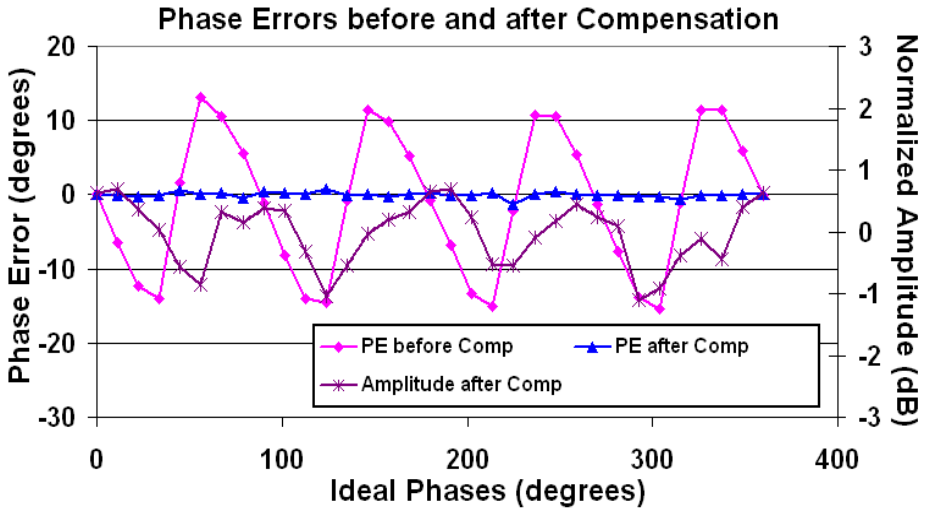


Figure 6.14: Phase errors before/after compensation for the non-sinusoidal LO (f_{RF} = 10.4GHz). Within a 1.1dB amplitude variation, a maximum 1.4˚ phase error is achieved.

On the other hand, the mismatches across the elements are mainly due to delay and gain/loss offsets in the RF feed paths and the reference clock distribution network. Although deterministic, these offsets are generally hard to predict a priori and compensated for off-chip, particularly in a very-large-scale array system. However, dense on-chip phase interpolation can easily compensate for these two offsets by providing a phase shift to cancel the summation effect of the phase offsets while providing adjustment on the amplitude mismatches. Figure 6.15 shows the 4-element array electrical pattern measured with/without offset calibration to verify the capability for compensating this type of mismatch.

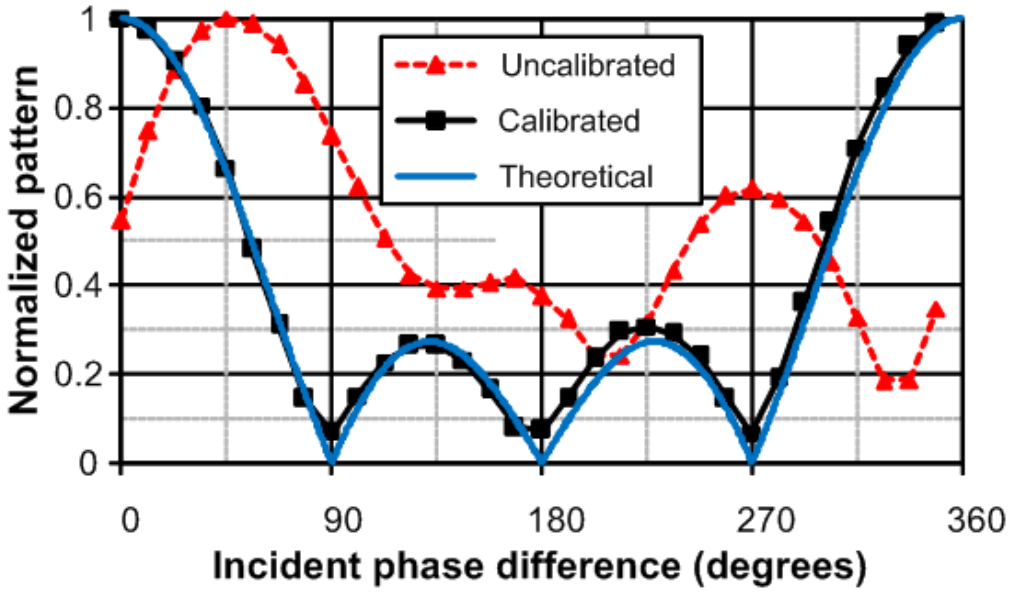


Figure 6.15: Array patterns with/without calibration (f_{RF}=10.4GHz)

Our 4-element electrical phased-array system measurement setup is shown in Figure 6.16. A 4-way power divider distributes the input signal into four RF feed paths. Discrete

phase shifters are used to form the effective input wave-front. A 50MHz synthesizer reference from off-chip crystal oscillator is sent to every element. The baseband output signals and their sum are monitored by a digital oscilloscope.

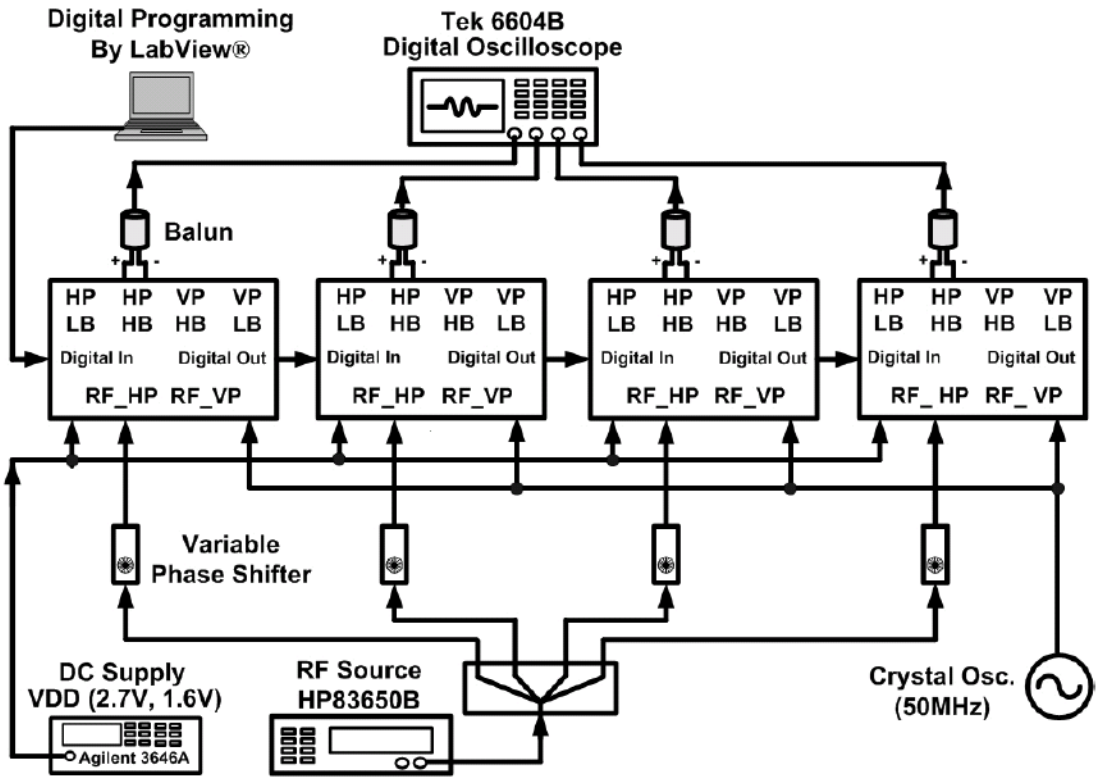


Figure 6.16: Measurement setup for electrical array performance characterization

The normalized electrical array patterns for beam forming at different incident angles have been measured for the RF frequencies of 6GHz, 10.4GHz, and 18GHz, respectively, shown in Figure 6.17. The worst case peak-to-null ratio is 21.5dB. The measured array patterns closely match the ideal ones due to the aforementioned compensations and calibrations facilitated by dense phase/amplitude interpolations.

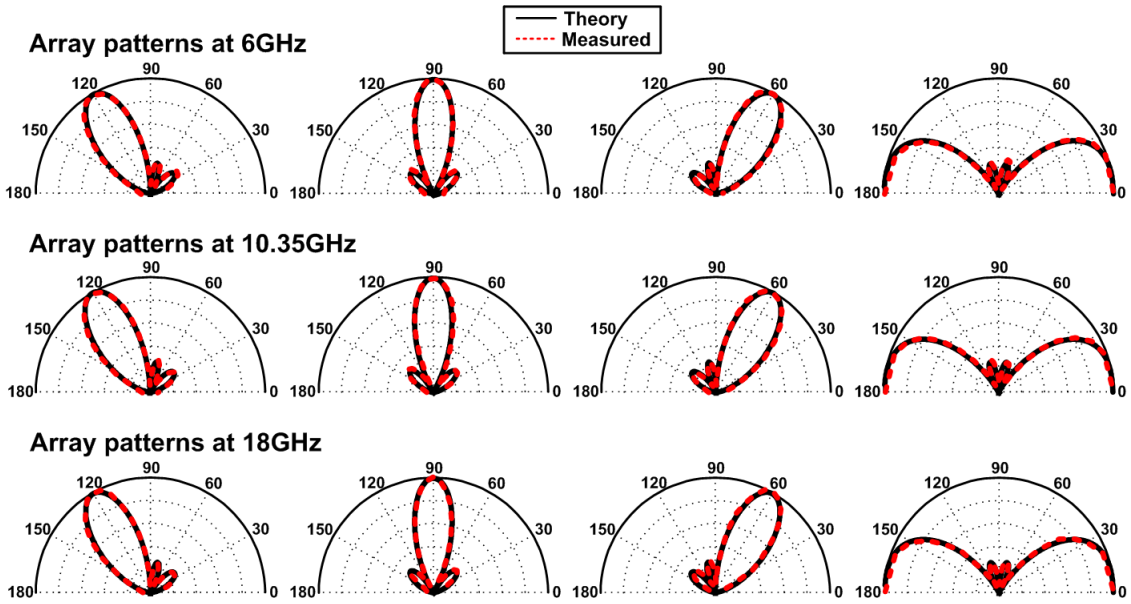


Figure 6.17: The measured and the ideal electrical array patterns

Based on the beam-forming capability, a phased-array receiver system has improved EVM results compared with a single receiver element mainly for the following three reasons. First of all, since the wanted signals are summed in phase (combining in amplitude) and the noise signals are added incoherently (combining in power), the signal-to-noise-ratio (SNR) for the array increases by a factor of N (the number of array elements). Secondly, unwanted interference signals with a different incident angle are attenuated due to the array spatial filtering property as long as the receiver system will not be saturated at their power levels. Thirdly, EVM caused by any uncorrelated errors from the on-chip LO, such as the phase noises out of the phase-lock-loop bandwidth, in a receiver element will decrease by √N in the array.

To demonstrate this EVM improvement capability, a 4-element receiver array has been measured and its EVM is compared to the EVMs of the individual elements at different

symbol rates, shown in Figure 6.18. The apparently decreased EVM response for the array operation validates the aforementioned improvements.

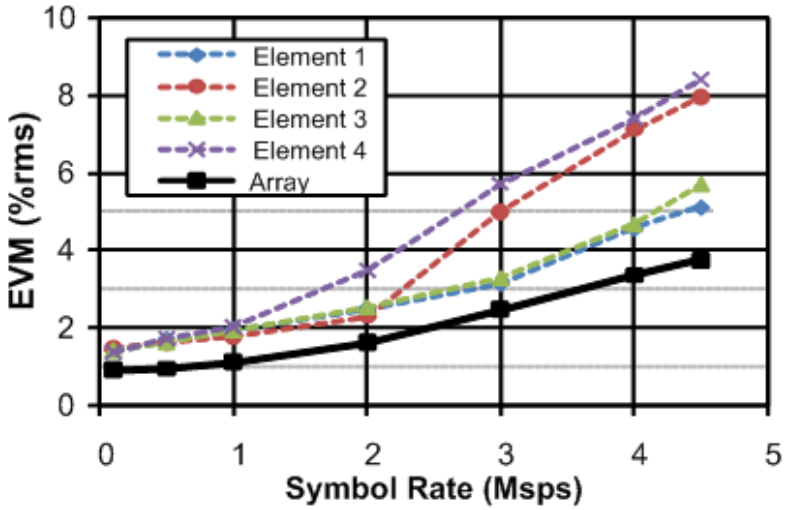


Figure 6.18: Measured EVM for the receiver elements and the 4-element phased-array

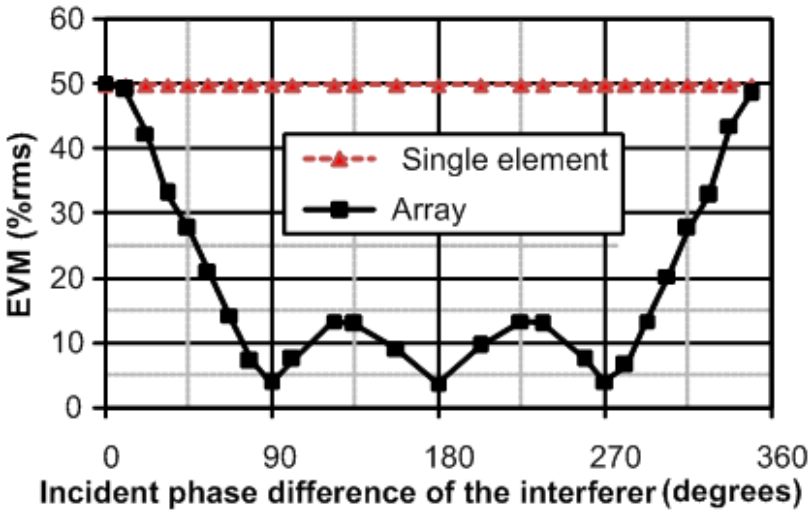


Figure 6.19: Measured EVM for the phased-array compared with a receiver element when the interference is incident at different angles

Figure 6.19 shows the measured array EVM when an in-band FM-modulated interference is applied at different incident angles. Significant rejection is achieved compared to a single receiver element when the interference is incident at an angle away

from the main lobe. Figure 6.20 shows the measured phase noise power spectral density decreases by 6dB in the 4-element array operation [47].

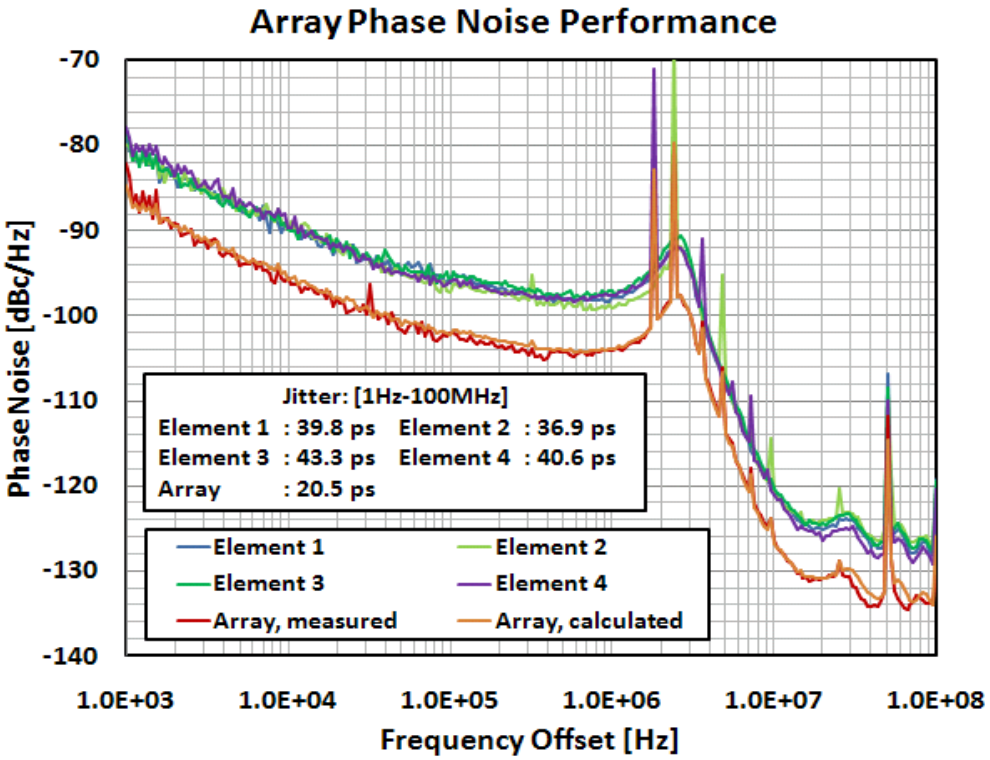


Figure 6.20: Measured phase noise performance (f_{RF}=7.5GHz)

6.5 Chapter Summary

In this chapter, we first introduce the operation mechanism of a phased array system, in particular the array pattern and the spatial filtering properties. We find from a yield point of view, for a practical array implementation, it is crucial to maintain good phase and amplitude matching among the array elements.

The major sources of the phase and amplitude mismatch in a practical array implementation have been identified. These include the gain/loss mismatches in the signal path, the errors for phase synthesis within the array cell, and the delay and clock skew

mismatches for the reference clock distribution among the array elements. In a practical implementation, these mismatch effects significantly degrade the formed array pattern, and thus the spatial filtering capability.

However, all of these mismatch effects are unknown a priori, and a different array element could have distinct mismatch values. This leads to formidable design challenges, particularly for a very-large-scale array. Therefore, it is desirable to implement the individual array element with full phase/amplitude synthesis capability to compensate these mismatches. To address these issues, we have proposed a precision phase and amplitude synthesis scheme. This scheme is itself inherently broadband and P.V.T. independent. Moreover, this scheme can be easily scaled to higher-resolution if needed.

As an implementation example, a 6-to-18 GHz dual-band quad-beam CMOS phased array receiver system is presented. The array is capable to form 4 spatially independent beams at two arbitrary frequencies across the 6-to-18 GHz bandwidth, which explores the spatial diversity and frequency diversity functionalities of a phased array system.

Based on the proposed phase/amplitude synthesis technique, the individual phased array element has achieved a maximum RMS phase error of 0.5˚ within a worst case RMS amplitude variation of 1.5dB for a continuous 360˚ interpolation across the 6-to-18 GHz bandwidth. A 4-element phased array receiver system is implemented based on the designed CMOS chip. The calibration capability is fully demonstrated with compensation on individual array element mismatches as well as inter-array mismatches to restore the desired array pattern. The complete electrical array pattern is measured at 6GHz, 10.35GHz and 18GHz. The array spatial filtering capability is also demonstrated by EVM and phase noise measurements.

Chapter 7: A Broadband Active Peaking Technique

7.1. Introduction

The insatiable hunger for a higher data rate in modern wireline/wireless communication systems has stimulated the development and implementation of various broadband circuits. Among current IC fabrication processes, CMOS is believed to be the most promising platform for those broadband communication circuits due to its unparalleled advantages, such as high integration level, versatile digital assistance, low cost, and remarkable reliability. However, CMOS has inferior high frequency performance compared to its compound counterparts due to low transconductance and excessive parasitic capacitances. Therefore, developing bandwidth extension techniques have attracted continuous research attention in this field.

Conventionally, shunting peaking by passive inductors is often used as the first choice for bandwidth boosting. Providing a complex pole pair and a left half plane transmission zero, this topology can effectively extend the circuit bandwidth by a maximum factor of 85% [48].

A more exotic solution for bandwidth extension has also been proposed based on filter synthesis theory [49] [50]. In this method, the entire network, including the output loadings at the first stage, the interconnections and the input terminations at the next stage, is optimized together to achieve a broadband transimpedance response with a maximum bandwidth enhancement ratio of 4.3. Due to the two-port nature, this inter-

stage network fundamentally outperforms the inductive peaking case which is simply a driving impedance topology.

Finally, the ultimate bandwidth enhancement can be achieved by using a distributed structure. Here, series inductors are added between adjacent stages' parasitic capacitances to form an artificial transmission line. The bandwidth of the distributed amplifier is normally limited by the cutoff frequency of the synthetic line, which can be optimized to approach the cutoff frequency of the transistors (f_t).

Nonetheless, all the aforementioned bandwidth extension techniques require lumped inductors which consume large amounts of expensive silicon area. Moreover, the space reserved for the inductor complicates the layout of its nearby blocks by imposing longer signal routing length, which adds extra capacitive loading and area consumption. This detrimental effect actually increases the broadband design difficulty. In addition, the lumped inductors can also pick up noises through high conductivity substrate (~1ohm·cm) in modern CMOS process.

The 6-to-18GHz dual-band quad-beam phased array receiver chip presented in Chapter 6 can be taken as an example. To achieve the broadband operation, overall 30 spiral inductors are used for bandwidth extension purpose, which occupies a huge amount of chip area. The chip microphotograph with shunt-peaking inductors highlighted is shown in Figure 7.1, where the yellow blocks highlight the inductors for bandwidth extension purpose. In addition, the blue blocks indicate the inductors used for resonant tank, matching and filtering.

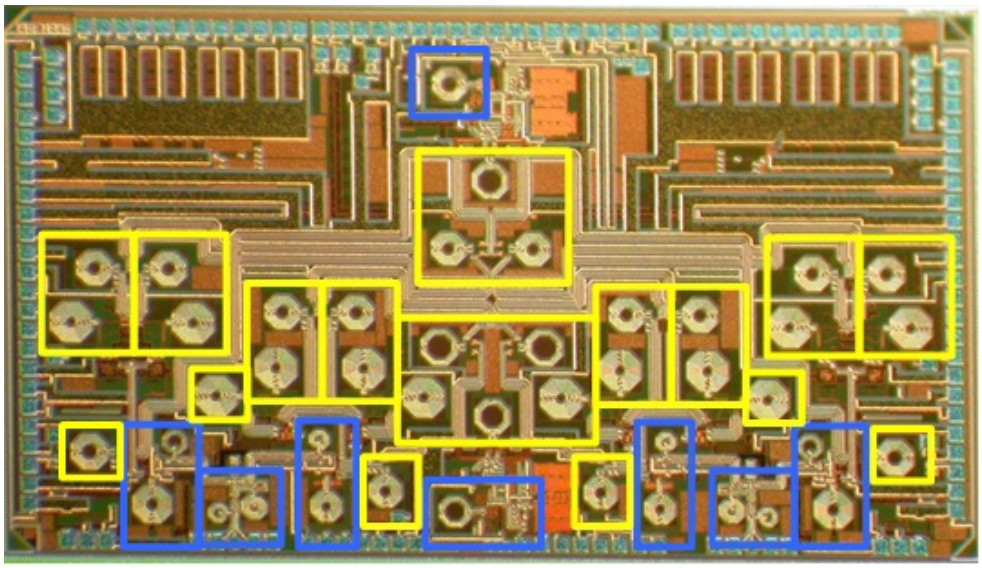


Figure 7.1: Microphotograph of the 6-to-18GHz dual-band quad-beam phased array receiver chip with all the passive inductors denoted

Therefore, it is desirable to achieve the gain-bandwidth extension without using lumped inductors, particularly in a broadband system demanding high integration levels.

Cherry-Hooper amplifiers have the advantages of complex-pole bandwidth extension and minimum inter-stage loadings [52]. Its modified topologies with active feedback have also been recognized for the advantage of the unilateral feedback path [53]–[55]. Traditionally, Cherry-Hooper amplifiers have been viably used in limiting amplifier chains. New applications such as LO buffers, analog transversal equalizer [56], and LNAs [57] [58] are also reported recently. However, to the best of the authors' knowledge, there is no research work reported so far which focuses on design optimization of Cherry-Hooper amplifiers or its derived topologies.

In this paper, we will present a novel design methodology, which minimizes the power consumption of a Cherry-Hooper amplifier under the constraints of gain-bandwidth, IIP3, noise performance, and gain peaking [59].

This chapter is organized as follows. Section 7.2 will review active-feedback Cherry-Hooper amplifiers with linear and weak nonlinear design equations derived. Then, the design methodology will be introduced in Section 7.3. Finally, Section 7.4 demonstrates several design examples with measurement results to validate our methodology. And the chapter will be concluded in Section 7.5.

7.2. General Cherry-Hooper Amplifier with Active Feedback

7.2.1 General Linear Transfer Function Analysis

The classic Cherry-Hooper amplifier is built based on alternately cascading series feedback (tranconductance) and shunt feedback (transimpedance) stages to minimize the loading between adjacent stages and undesired miller effects [25][60]. This brings the advantages of decoupled design for individual block and extended bandwidth for the entire amplifier chain.

Several modified Cherry-Hooper amplifiers with active feedback are depicted in Figure 7.2. By dividing the amplifier into transconductance stage and transimpedance stage, one can arrive at a general circuit topology shown in Figure 7.3. gm_1 block is the transconductance stage which acts as a voltage meter with high input impedance. The rest of the circuit functions as a transimpedance stage which converts the output current of the gm_1 stage to a voltage output. The shunt-shunt feedback lowers both the input and the output impedance of the transimpedance stage making them close to a current sensor and a voltage source, respectively.

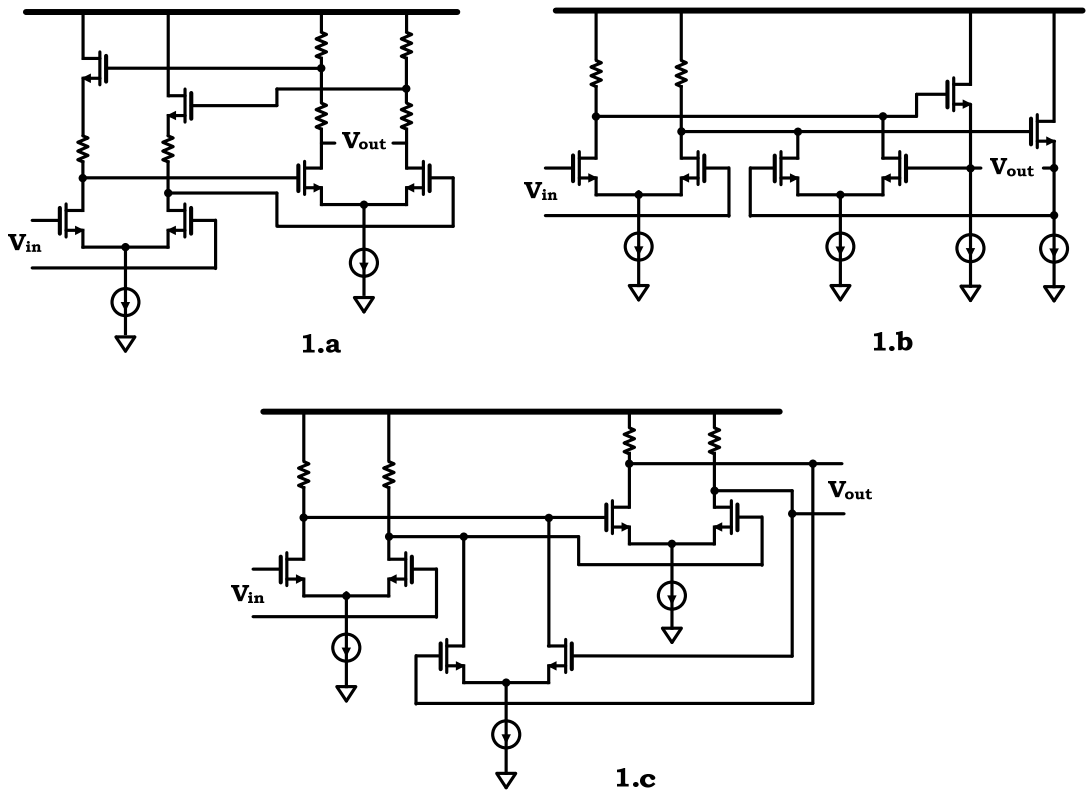


Figure 7.2. Cherry-Hooper amplifiers with active feedback (3 different types in differential configuration).

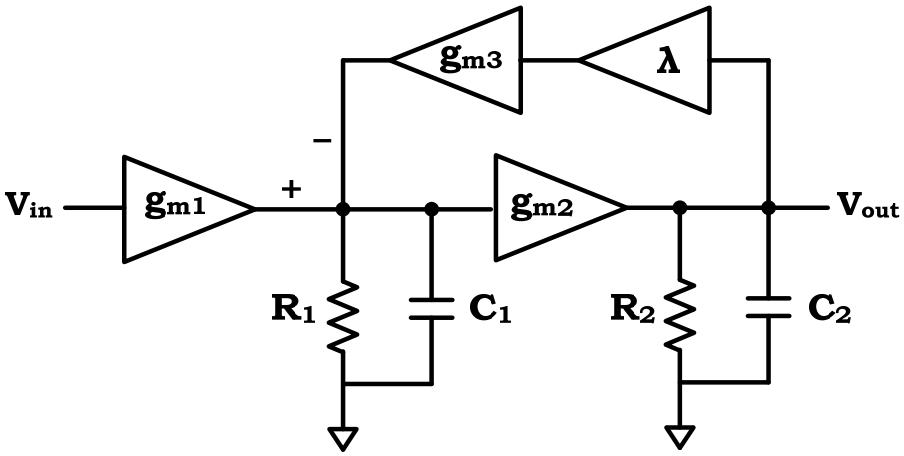


Figure 7.3: General Cherry-Hooper amplifier topology (single-ended type). λ represents any passive network gain from the output to the active feedback input.

The voltage transfer function for the general Cherry-Hooper topology is derived as

$$A_V(s) = \frac{\frac{g_{m1}g_{m2}}{C_1C_2}}{s^2 + \frac{R_1C_1 + R_2C_2}{R_1C_1R_2C_2}s + \frac{1 + \lambda g_{m2}g_{m3}R_1R_2}{R_1C_1R_2C_2}} = \frac{A_{V_{dc}}\omega_n^2}{s^2 + 2\xi\omega_n s + \omega_n^2}, \quad (7.1)$$

where $A_{V_{dc}}$ is the DC voltage gain and ω_n is the characteristic frequency of the 2nd-order system. λ represents any passive voltage dividing structure, such as resistive divider. The gain-bandwidth product is given by:

$$A_{V_{dc}}\omega_n^2 = \frac{g_{m1}g_{m2}}{C_1C_2}, \quad (7.2)$$

which only depends on the two forward transconductances and their capacitances. Note that $A_{V_{dc}}\omega_n^2$ is used instead of $A_{V_{dc}}\omega_n$ for the gain-bandwidth product because the system is of 2nd-order. If we define a factor K (unit in Ω^2) as

$$K = \frac{R_1R_2}{1 + \lambda g_{m2}g_{m3}R_1R_2}, \quad (7.3)$$

we can express the DC voltage gain and the characteristic frequency, respectively, by

$$A_{V_{dc}} = \frac{g_{m1}g_{m2}R_1R_2}{1 + \lambda g_{m2}g_{m3}R_1R_2} = g_{m1}g_{m2}K \quad (7.4)$$

and

$$\omega_n^2 = \frac{1 + \lambda g_{m2}g_{m3}R_1R_2}{R_1R_2C_1C_2} = \frac{1}{KC_1C_2}. \quad (7.5)$$

Therefore, as a 2nd-order system, the transfer function of the Cherry-Hooper amplifier can be fully determined by fixing the two forward stages, the factor K, and the damping factor ξ. Mathematically, a larger $A_{V_{dc}}\omega_n^2$ product represents a higher asymptotic curve of the 2nd-order transfer function providing a fundamentally better gain-bandwidth performance. For a given asymptotic curve determined by the two g_m stages, factor K thus trades $A_{V_{dc}}$with ω_n^2, and factor ξ determines the gain peaking value to extend the

bandwidth within the gain flatness limit.

Figure 7.4 shows the three Cherry-Hooper topologies with their dominant parasitic capacitances denoted. Their equivalent half-circuits are presented for simplicity. Note that the dashed line in Figure 3.C denotes the feedback is from the opposite polarity of the differential outputs to maintain negative sign. Capacitance C_1 includes all the capacitances at the feedback summing node except for the miller-multiplied capacitance C_{gd2}. And C_L includes the parasitic capacitances and the load capacitance at the output node. The design equations for $A_{V_{dc}}$, ω_n^2, $A_{V_{dc}}\omega_n^2$, K, and ξ for each topology in Figure 7.4 are summarized in Table 7.1. Clearly the gain-bandwidth product of $A_{V_{dc}}\omega_n^2$ is independent of the feedback path.

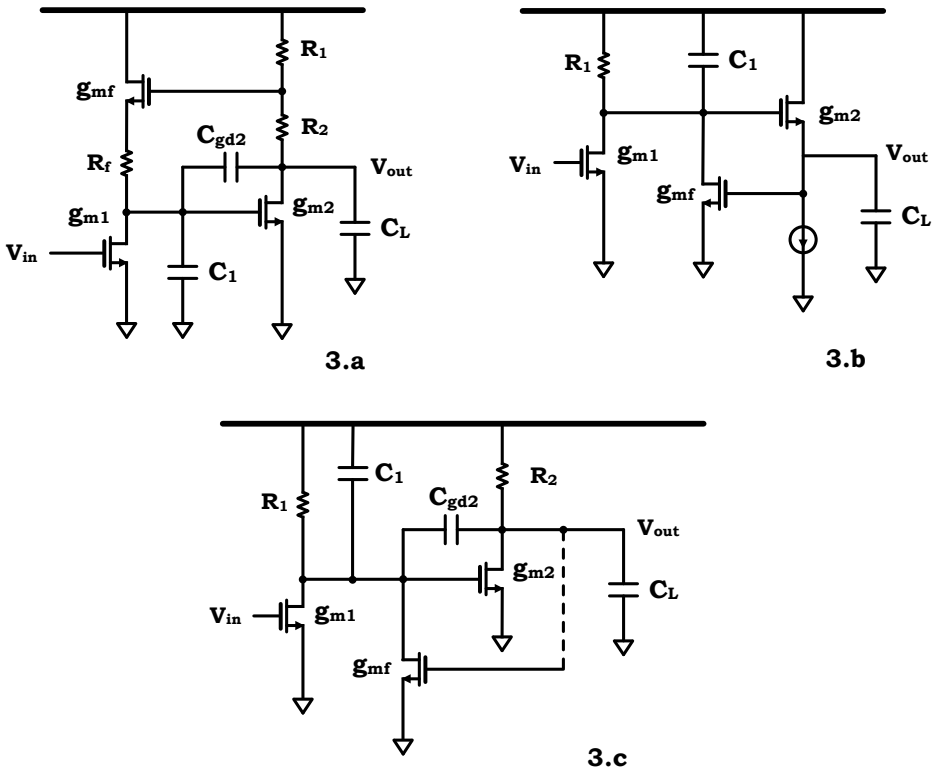


Figure 7.4: Different Cherry-Hooper amplifier topologies with dominant parasitics

TABLE 7.1 LINEAR DESIGN EQUATIONS FOR DIFFERENT CHERRY-HOOPER TOPOLOGIES

Topology 1	$\boldsymbol{A_{V_{dc}}\omega_n^2}$	$\frac{g_{m1}g_{m2}}{C_1C_L + C_1C_{gd2} + C_LC_{gd2}}$
	$\boldsymbol{A_{V_{dc}}}$	$\frac{g_{m1}g_{m2}(R_1+R_2)(\frac{1}{g_{mf}}+R_f)}{1+g_{m2}R_1} = g_{m1}g_{m2}K$
	$\boldsymbol{\omega_n^2}$	$\frac{1+g_{m2}R_1}{(R_1+R_2)\left(\frac{1}{g_{mf}}+R_f\right)\left(C_1C_L + C_1C_{gd2} + C_LC_{gd2}\right)} = \frac{1}{K\left(C_1C_L + C_1C_{gd2} + C_LC_{gd2}\right)}$
	$\boldsymbol{K}$	$\frac{(R_1+R_2)(\frac{1}{g_{mf}}+R_f)}{1+g_{m2}R_1}$
	$\boldsymbol{\xi}$	$\xi = \frac{\omega_n K}{2}\left[\frac{C_1+C_{gd2}}{R_1+R_2} + \frac{C_L + \frac{R_2}{R_1+R_2}C_{gd2}}{\frac{1}{g_{mf}}+R_f} + C_{gd2}g_{m2}\right]$
Topology 2	$\boldsymbol{A_{V_{dc}}\omega_n^2}$	$\frac{g_{m1}g_{m2}}{C_1C_L}$
	$\boldsymbol{A_{V_{dc}}}$	$\frac{g_{m1}R_1}{1+g_{mf}R_1} = g_{m1}g_{m2}K$
	$\boldsymbol{\omega_n^2}$	$\frac{g_{m2}(1+g_{mf}R_1)}{R_1C_1C_L} = \frac{1}{KC_1C_L}$
	$\boldsymbol{K}$	$\frac{R_1}{g_{m2}(1+g_{mf}R_1)}$
	$\boldsymbol{\xi}$	$\xi = \frac{\omega_n K}{2}\left[\frac{C_1}{R_2} + \frac{C_L}{R_1}\right]$
ogy 3	$\boldsymbol{A_{V_{dc}}\omega_n^2}$	$\frac{g_{m1}g_{m2}}{C_1C_L + C_1C_{gd2} + C_LC_{gd2}}$

	$\boldsymbol{A_{V_{dc}}}$	$\frac{g_{m1}g_{m2}R_1R_2}{1+g_{mf}g_{m2}R_1R_2} = g_{m1}g_{m2}K$
	$\boldsymbol{\omega_n^2}$	$\frac{1+g_{mf}g_{m2}R_1R_2}{R_1R_2\left(C_1C_L+C_1C_{gd2}+C_LC_{gd2}\right)} = \frac{1}{K\left(C_1C_L+C_1C_{gd2}+C_LC_{gd2}\right)}$
	$\boldsymbol{K}$	$\frac{R_1R_2}{1+g_{mf}g_{m2}R_1R_2}$
	$\boldsymbol{\xi}$	$\xi = \frac{\omega_n K}{2}\left[\frac{C_L+C_{gd2}}{R_1}+\frac{C_L+C_{gd2}}{R_2}+C_{gd2}g_{m2}-C_{gd2}g_{mf}\right]$

Note: The right-half plane high frequency transmission zero ($s = \frac{g_{m2}}{C_{gd2}}$) is neglected for all the three topologies.

7.2.2 Weak Nonlinearity Performance Analysis

Weak nonlinear effects in CMOS devices are typically caused by nonlinear transconductance (gm), nonlinear parasitic capacitance (C_{gs}, C_{db}, and C_{sb}), and nonlinear output resistance (r_o) [61]. When the load impedance is linear and much smaller than the r_o, and the operating frequency is low enough to neglect the parasitic capacitances, the nonlinear performance of the device will be dominated by the gm nonlinearity [62]. By using the Taylor series expansion, the nonlinear output current of a single device can be approximated by

$$I_d(V_{od}+v) \approx \alpha_0 + \alpha_1 v + \alpha_2 v^2 + \alpha_3 v^3$$

$$= I_d + \frac{\left(1+\frac{M}{2}V_{od}\right)\mu_n C_{ox} W V_{od}}{(1+MV_{od})^2 L_{eff}} v + \frac{\mu_n C_{ox} W}{2(1+MV_{od})^3 L_{eff}} v^2 - \frac{\mu_n C_{ox} MW}{2(1+MV_{od})^4 L_{eff}} v^3$$

$$= I_d + g_{m1}v + g_{m2}v^2 + g_{m3}v^3, \quad (7.6)$$

where V_{od} is the overdrive voltage, v is the small signal input voltage, and M is given by

$$M = \frac{1}{L_{eff}E_{sat}}\,. \tag{7.7}$$

The valid range for the V_{od} of the above Taylor series is normally between 100mV and 300mV, where the lower bound is limited by the near/sub-threshold effect, and the upper one by the vertical-field mobility degradation effect [25]. Note that equations (7.6) and (7.7) suggest that the Taylor coefficients can be calculated analytically by knowing the process parameters $\mu_n C_{ox}$ and $L_{eff}E_{sat}$ for a given transistor with a prescribed V_{od}.

Therefore, applying equations (7.6) and (7.7) to all three nonlinear transconductances of the general topology in Figure 7.3, the IIP3 of the entire Cherry-Hooper amplifier can be derived as

$$\begin{aligned}\frac{1}{AIIP_{3_{total}}^2} \approx{}& \frac{3g_{m1,3}}{4g_{m1,1}} + \frac{3g_{m2,3}}{4g_{m2,1}} \frac{g_{m1,1}^2 R_1^2}{\left(1 + g_{m2,1}R_1 g_{m3,1}R_2\lambda\right)^2} \\ &+ \frac{3g_{m3,3}}{4g_{m3,1}} \frac{g_{m1,1}^2 R_1^2 g_{m2,1}^2 R_2^2 \lambda^2}{\left(1 + g_{m2,1}R_1 g_{m3,1}R_2\lambda\right)^2} \\ ={}& \frac{1}{AIIP_{3_{M1}}^2} + \frac{1}{AIIP_{3_{M2}}^2}\left(\frac{V_{in_{M2}}}{V_{in_{M1}}}\right)^2 + \frac{1}{AIIP_{3_{M3}}^2}\left(\frac{V_{in_{M3}}}{V_{in_{M1}}}\right)^2,\end{aligned} \tag{7.8}$$

where $AIIP_{3_{Mi}}$ stands for the input referred 3rd-order intercept voltage of the ith transistor, $g_{mi,j}$ for the jth Taylor coefficient for the ith transistor, and $V_{in_{Mi}}$ for the input voltage at the ith transistor. Therefore, the ratio of $\frac{V_{in_{Mj}}}{V_{in_{Mi}}}$ represents the voltage gain from the input of the ith transistor to that of the jth transistor.

The derived $AIIP_3$ equations for all three topologies in Figure 7.2 are summarized in Table 7.2. For the 2nd and the 3rd topology, the common-source active feedback stage normally dominates the nonlinearity performance of the entire Cherry-Hooper amplifier

due to its high voltage gain. This is one of the reasons for these two topologies to be extensively used in the limiting amplifier chain or LO buffers. However, if a source follower is used in the feedback-path, as in the 1st topology, its nonlinearity contribution can be neglected. Moreover, it forms a load of $(\frac{1}{g_{mf}} + R_f)$ at the output of the first transistor with an effective negative nonlinearity which can cancel the nonlinear terms generated from g_{m1}. This explains those advanced applications such as LNA and FIR filters based on the 1st topology with a high $AIIP_3$. The dependence of total $AIIP_3$ on the $AIIP_3$ of each transistor for the three different topologies is also summarized in Table II.

In addition, one can further simplify the $AIIP_3$ for the 1st topology to get more design insights. For this topology, it is a common practice to design the voltage gain from the current summing node to the input to be close to unity $(\frac{V_{in_{M2}}}{V_{in_{M1}}} = \frac{g_{m1}R_1}{1+\lambda g_{m2}R_2 g_{mf}R_2} \approx 1)$ while letting the second stage contribute most of the gain. Moreover, in typical designs $g_{mf} \approx g_{m1} \approx 1/R_f$. Upon applying these two constraints to the $AIIP_3$ formula of the 1st topology in Table II, we obtain

$$\frac{1}{AIIP_{3_{total}}^2} = \frac{1}{AIIP_{3_{M1}}^2} - \frac{1}{AIIP_{3_{Mf}}^2(1 + R_f g_{mf})}\left(\frac{g_{m1}}{g_{mf}}\right)^2 + \frac{1}{AIIP_{3_{M2}}^2}\left(\frac{V_{in_{M2}}}{V_{in_{M1}}}\right)^2$$

$$\approx \frac{1}{AIIP_{3_{M1}}^2} - \frac{1}{2AIIP_{3_{M1}}^2}\left(\frac{g_{m1}}{g_{m1}}\right)^2 + \frac{1}{AIIP_{3_{M2}}^2}$$

$$\approx \frac{1}{2AIIP_{3_{M1}}^2} + \frac{1}{AIIP_{3_{M2}}^2}. \tag{7.9}$$

Equation (7.9) shows that the total $AIIP_3$ for a Cherry-Hooper amplifier of the 1st kind can be fully determined by the two forward transimpedance stages. However, as will be shown in the next section, with the above two constraints the design problem becomes

over-determined. However, for a Butterworth response, the final design parameters typically will normally stay close with those constraints as shown in the design example. Therefore, equation 7.9 still serves as a valid approximation for total $AIIP_3$.

TABLE 7.2 $AIIP_3$ DESIGN EQUATIONS FOR DIFFERENT CHERRY-HOOPER TOPOLOGIES

Topology 1	$$\frac{1}{AIIP_{3_{total}}^2} = \frac{1}{AIIP_{3_{M1}}^2} - \frac{1}{AIIP_{3_{Mf}}^2(1+R_f g_{mf})}\left(\frac{g_{m1}}{g_{mf}}\right)^2 + \frac{1}{AIIP_{3_{M2}}^2}\left(\frac{V_{in_{M2}}}{V_{in_{M1}}}\right)^2$$
	$AIIP_{3_{total}}^2$ Dependence on Different Stages **M_1: Medium M_2: Medium M_f: Negative**
Topology 2	$$\frac{1}{AIIP_{3_{total}}^2} = \frac{1}{AIIP_{3_{M1}}^2} + \frac{1}{AIIP_{3_{Mf}}^2}\left(A_{V_{dc}}\right)^2$$
	$AIIP_{3_{total}}^2$ Dependence on Different Stages **M_1: Medium M_2: None M_f: High**
Topology 3	$$\frac{1}{AIIP_{3_{total}}^2} = \frac{1}{AIIP_{3_{M1}}^2} + \frac{1}{AIIP_{3_{M2}}^2}\left(\frac{V_{in_{M2}}}{V_{in_{M1}}}\right)^2 + \frac{1}{AIIP_{3_{Mf}}^2}\left(A_{V_{dc}}\right)^2$$
	$AIIP_{3_{total}}^2$ Dependence on Different Stages **M_1: Medium M_2: Medium M_f: High**

7.2.3 Input Referred Noise Voltage PSD

The power spectrum density (PSD) of the total input referred noise voltage can be derived for the general topology as

$$\frac{\overline{V_{n,in}^2}}{\Delta f} = \frac{4KT\gamma}{g_{m1}} + \frac{4KT}{R_1 g_{m1}^2} + \frac{4KT\gamma g_{mf}}{g_{m1}^2} + \frac{4KT\gamma}{g_{m2} g_{m1}^2 R_1^2} + \frac{4KTR_2}{g_{m1}^2 R_1^2 g_{m2}^2 R_2^2}, \qquad (7.10)$$

where only resistive thermal noise and transistor channel noise are considered. Normally, the noise from R_2 can be neglected due to high gain nature of the amplifier, and the terms due to g_{mf} and R_1 can be related to g_{m1} or neglected for specific topology to achieve an approximated expression for $\frac{\overline{V_{n,in}^2}}{\Delta f}$ determined only by g_{m1} and g_{m2}. The detailed $\frac{\overline{V_{n,in}^2}}{\Delta f}$ design equations can be derived for different topologies and are summarized in Table 7.3. Again, one can see that the total input referred noise PSD density can be approximately determined by the two forward transconductance stages.

TABLE 7.3 $\frac{\overline{V_{n,in}^2}}{\Delta f}$ DESIGN EQUATIONS FOR DIFFERENT CHERRY-HOOPER TOPOLOGIES

Topology 1	$\frac{\overline{V_{n,in}^2}}{\Delta f} \approx \frac{4KT\gamma}{g_{m1}} + \frac{4KTg_{mf}^2}{g_{m1}^2(1+g_{mf}R_f)^2}\left(R_f + \frac{\gamma}{g_{m2}} + \frac{\gamma}{g_{mf}}\right) \approx \frac{4KT}{g_{m1}}\left(\frac{5\gamma}{4} + \frac{1}{4}\right) + \frac{KT\gamma}{g_{m2}}$ Constraint Used: $g_{mf} \approx g_{m1} \approx 1/R_f$ and $A_{V_{dc}} \gg 1$
Topology 2	$\frac{\overline{V_{n,in}^2}}{\Delta f} = \frac{4KT\gamma}{g_{m1}} + \frac{4KT}{R_1 g_{m1}^2} + \frac{4KT\gamma g_{mf}}{g_{m1}^2} + \frac{4KT\gamma}{g_{m2} R_1^2 g_{m1}^2} + \frac{\frac{\overline{I_{n,Iss_1}^2}}{\Delta f}}{g_{m2}^2 R_1^2 g_{m1}^2} \approx \frac{4KT\gamma}{g_{m1}}$ Constraint Used: $g_{m1}R_1 \gg 1$ and $g_{mf} \ll g_{m1}$

Topology 3	$\frac{\overline{V_{n,in}^2}}{\Delta f} = \frac{4KT\gamma}{g_{m1}} + \frac{4KT}{R_1 g_{m1}^2} + \frac{4KT\gamma g_{mf}}{g_{m1}^2} + \frac{4KT\gamma}{g_{m2} g_{m1}^2 R_1^2} + \frac{4KTR_2}{g_{m1}^2 R_1^2 g_{m2}^2 R_2^2}$ $\approx \frac{4KT\gamma}{g_{m1}} + \frac{4KT}{g_{m1}} + \frac{4KT\gamma}{g_{m2}} = \frac{4KT(\gamma+1)}{g_{m1}} + \frac{4KT\gamma}{g_{m2}}$ Constraint Used: $g_{mf} \approx g_{m1} \approx 1/R_f$, $g_{mf} \ll g_{m1}$, and $A_{V_{dc}} \gg 1$

7.3. Proposed Design Method for Cherry-Hooper Amplifier with Active Feedback

Based on the above analysis, we propose a novel design methodology for the active-feedback Cherry-Hooper amplifier which optimizes the DC power consumption under the constraints of gain-bandwidth, IIP_3, input-referred noise and gain peaking tolerance. The methodology will be introduced in a step-by-step fashion, with a buffer design example implemented in the 1st type of Cherry-Hooper topology.

With the design equations in Table 7.1, 7.2, and 7.3, one can see that the product of $A_{V_{dc}}\omega_n^2$, the $AIIP_3$ (if its topology is of the 1st kind), and the input referred noise for a Cherry-Hooper amplifier can be fully determined by parameters only associated with the two forward transconductance stages. Moreover, normally the power consumption of a Cherry-Hooper amplifier is also dominated by these two stages, since the feedback either reuses their DC currents (as in topology 1) or consumes very little current (as in topology 2 and 3). These two important properties suggest that one can decouple the designs of the forward stages and the feedback loop, which essentially avoids the iteration process in a conventional design methodology.

Therefore, the two gain transistors M_1 and M_2 can be optimized for a minimum power consumption (e.g., the DC biasing currents for a fixed supply voltage), while satisfying the prescribed constraints of gain-bandwidth, $AIIP_3$, and input-referred noise, according to equations shown in Table 7.1, 7.2, and 7.3. Once the optimum transistor sizing and minimum power consumption are determined, the parameters regarding the two forward stages are known. Then based on additional small-signal equations for $A_{V_{dc}}$, ω_n^2, K, and ξ in Table I, the rest of the circuit parameters can be easily calculated to complete the design.

As an example, the design procedure is carried out in steps as follows. The design target is a differential Cherry-Hooper buffer with a maximally flat voltage gain of 9dB, a -3dB bandwidth of 12GHz, a single-ended capacitive loading of 50fF, and a differential AIIP3 point of -1dBV, implemented in an 130nm CMOS process.

1) Calculate the characteristic frequency ω_n by required ω_{-3dB} and damping factor ξ

Since for a given $\omega_{-3\mathrm{dB}}$, less damping yields a smaller value for the ω_n, we should always use the maximum tolerable peaking ξ to alleviate the requirement on ω_n calculated by

$$\omega_n = \frac{\omega_{-3dB}}{\left(\sqrt{(2\xi^2-1)^2+1}+1-2\xi^2\right)^{\frac{1}{2}}}. \tag{7.11}$$

Here, in our design example, the maximally flat response in the design target requires $\xi = 1/\sqrt{2}$, which means $\omega_n$equals ω_{-3dB} ($2\pi \times 12$GHz).

2) Calculate the target gain-bandwidth product

Based on the required DC voltage gain and ω_n, we can calculate the necessary gain-bandwidth product of $A_{V_{dc}}\omega_n^2$, which defines the asymptotic curve of the transfer

function as mentioned in Section II.A.

Therefore, our design example requires

$$A_{V_{dc}}\omega_n^2 = 2.82 \times (2\pi \times 12 \times 10^9)^2 = 1.6 \times 10^{22}\,(rad/s)^2. \quad (7.12)$$

3) Compute the gain-bandwidth product for a given power consumption

With the total power consumption dominated by the two forward transistors, a prescribed current distribution between the two transimpedance stages determines the total power consumption. Under this DC current allocation, one can sweep the sizes of the two forward transistors and calculate the resulting $A_{V_{dc}}\omega_n^2$ products. This *calculation* and maximum searching of $A_{V_{dc}}\omega_n^2$ can be obtained in an analytical fashion, using the extracted parameters of $\mu_n C_{ox}$, $L_{eff}E_{sat}$, and the parasitic capacitance coefficients [63].

In our design example, after the parameter extraction on the IBM8RF 130nm process, we can calculate and plot the achievable $A_{V_{dc}}\omega_n^2$ for various combinations of I_{d1} and I_{d2}. For example, the case of I_{d1}=1mA and I_{d2}=1.25mA is shown in Figure.7.5. Note that the contour unit is $10^{22}(rad/s)^2$. The maximum $A_{V_{dc}}\omega_n^2$ is $1.95 \times 10^{22}\,(rad/s)^2$ with $Width_1$ =20μm and $Width_2$ =30μm.

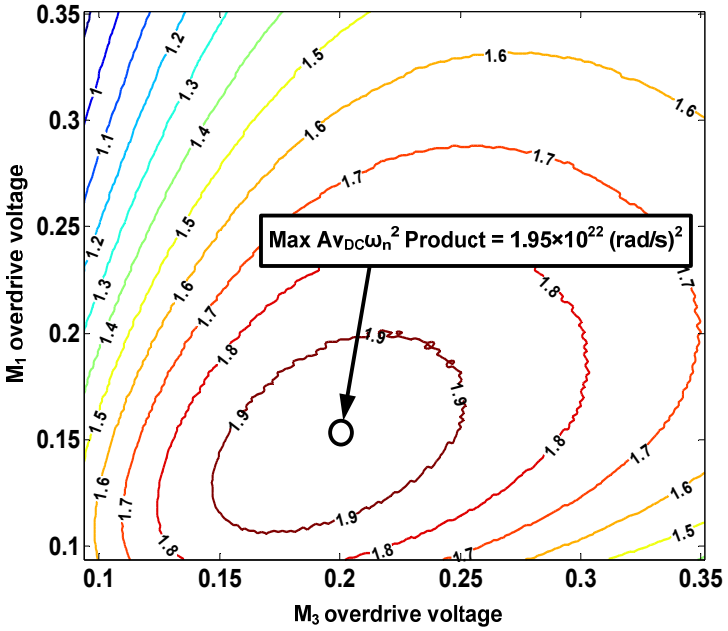


Figure 7.5: Gain-bandwidth plot for I_{d1}=1mA and I_{d3}=1.25mA

4) Determine the AIIP$_3$ and/or input referred noise constraints for a given power consumption

At a given DC current, sweeping the transistor width changes its overdrive voltage and g_m simultaneously. These two factors directly affect the nonlinearity and the noise performance of the entire Cherry-Hooper amplifier as shown in Table 7.2 and 7.3. Equal AIIP$_3$ and equal $\frac{\overline{V_{n,in}^2}}{\Delta f}$ curves can be calculated and superimposed on the $A_{V_{dc}}\omega_n^2$ contour plot made in step (3) shown in Figure 7.5. In general, at a fixed DC current, a smaller transistor size gives a larger overdrive voltage which leads to a better linearity, while its corresponding smaller gm will degrade the input referred noise, and vice versa. Therefore, the available region to select the maximum $A_{V_{dc}}\omega_n^2$ is the intersection of the AIIP$_3$ and noise constrain curves, which may exclude the maximum $A_{V_{dc}}\omega_n^2$ for the prescribed

power consumption. If these two curves do not intersect, shown as dashed lines in Figure 7.6, a larger power consumption has to be used to recalculate this available $A_{V_{dc}}\omega_n^2$ plot. Intuitively, on the $A_{V_{dc}}\omega_n^2$ plot with V_{od} values as independent variables, with a larger DC current, g_m will increase for any (V_{od1}, V_{od3}) point, which pushes the noise constrain curve towards the upper right corner. The $A_{V_{dc}}\omega_n^2$ product value also increases for every (V_{od1}, V_{od3}) point, since a larger self loading reduces the effect of the external capacitive load. However, fully determined by V_{od}, the $AIIP_3$ curve does not move. Therefore, a larger power consumption yields an increased intersection area of the $AIIP_3$/noise constraint curves and a larger $A_{V_{dc}}\omega_n^2$ value for every (V_{od1}, V_{od3}) point, which both help the design solution converge.

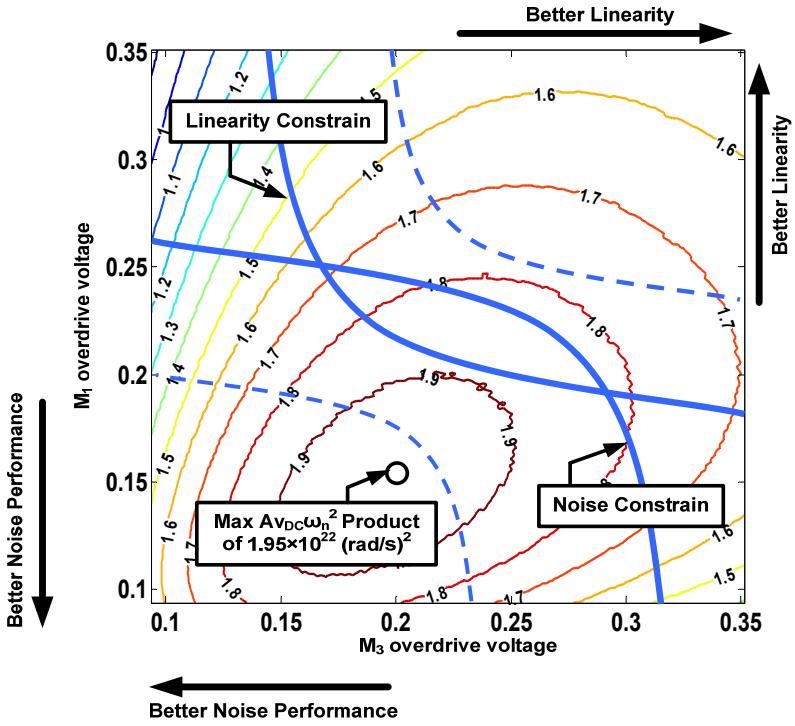


Figure 7.6: Linearity and noise trade-off on the gain-bandwidth product contour plot (unit $10^{22}(rad/s)^2$)

For the buffer design example, linearity instead of noise is a concern. Therefore, only the $AIIP_3$ is used as a constraint here. The $A_{V_{dc}}\omega_n^2$ plot with equal $AIIP_3$ curves are calculated and plotted for the case of I_{d1}=1mA and I_{d3}=1.25mA in Figure7.7. With $AIIP_3$ larger than -1dBV, the maximum $A_{V_{dc}}\omega_n^2$ is $1.76 \times 10^{22}\,(rad/s)^2$ at $V_{overdrive\_M1}$=240mV and $V_{overdrive\_M3}$=300mV.

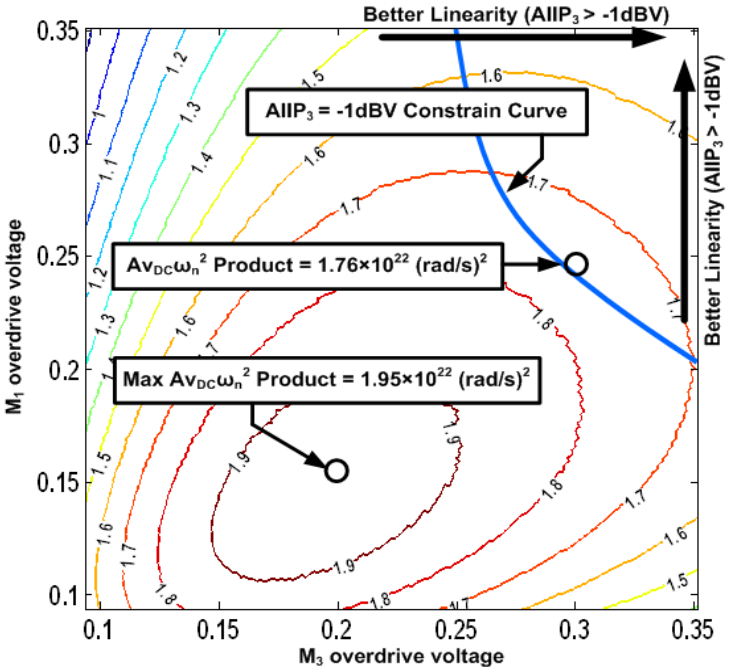


Figure 7.7: Gain-bandwidth product plot for I_{d1}=1mA and I_{d3}=1.25mA with $AIIP_3$= -1dBVcurve superimposed (contour unit of $10^{22}(rad/s)^2$)

5) Find the minimum power consumption satisfying the gain-bandwidth requirement

Steps 3) and 4) are repeated to obtain the maximum $A_{V_{dc}}\omega_n^2$ under the linearity/noise constraints for various power consumptions. Then one can find the minimum power consumption choice to satisfy the required gain-bandwidth product $A_{V_{dc}}\omega_n^2$. Again, this

entire searching process can be done analytically by using device models mentioned above.

In our design example, the achievable maximum $A_{V_{dc}}\omega_n^2$ products for various power consumptions with an $AIIP_3$ larger than -1dBV are plotted in Figure 7.8. Either I_{d1}=1.25mA/I_{d3}=1mA or I_{d1}=1mA/I_{d3}=1.25mA can satisfy the design goal. The latter is chosen here.

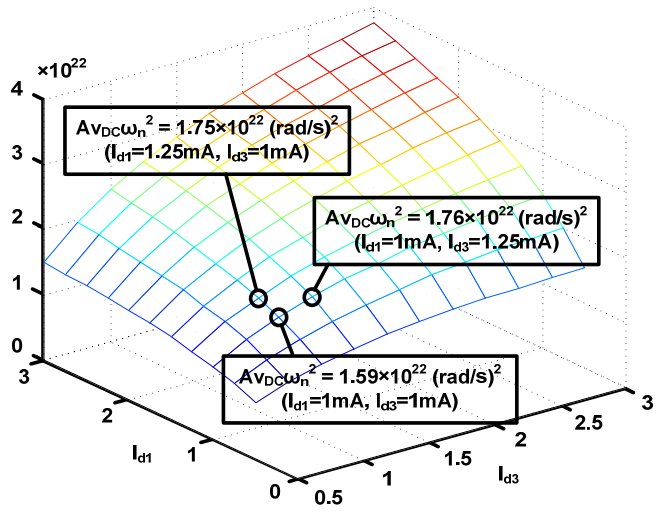


Figure 7.8. Achievable maximum gain-bandwidth product plot for various power consumptions with the constraint of $AIIP_3 \geq$ -1dBV

6) Compute the values for the other circuit elements

Once the sizes and the current consumptions of the two forward-gain transistors are determined, their transconductances and parasitic capacitances are known. The values of the other circuit elements can be calculated.

For our buffer example, the required design equations from Table I are restated here as

$$K = \frac{(R_1 + R_2)(\frac{1}{g_{mf}} + R_f)}{(1 + g_{m2}R_1)} = \frac{A_{V_{dc}}}{g_{m1}g_{m2}}, \tag{7.12}$$

$$\xi = \frac{\omega_n K}{2}\left[\frac{C_1 + C_{gd2}}{R_1 + R_2} + \frac{C_L + \frac{R_2}{R_1 + R_2}C_{gd2}}{\frac{1}{g_{mf}} + R_f} + C_{gd2}g_{m2}\right], \tag{7.13}$$

$$R_{in\_closeloop} = \frac{\frac{1}{g_{mf}} + R_f}{1 + g_{m2}R_1}. \tag{7.14}$$

We have three equations for the four variables R_1, R_2, $\frac{1}{g_{mf}}$, and R_f, so that one more degree of freedom exists in choosing individual $\frac{1}{g_{mf}}$ and R_f values while maintaining a fixed sum. This freedom can used to help set the DC bias point. Therefore, the assumption $g_{mf} \approx g_{m1} \approx 1/R_f$ for the $AIIP_3$ expression in section II.B leaves the problem overdetermined. The calculated design values for our example are listed in Table 7.4. It can be seen that although overdetermined, the solutions still stay close to all the design equations and assumptions in this particular example.

TABLE 7.4 DESIGN VALUES FOR BUFFER EXAMPLE IN SECTION III

W_1	$2I_1$	gm_1	W_2	$2I_2$	gm_2	W_f	gm_f
11um	2mA	6.0ms	10um	2.5mA	6.1	11um	6.0ms
R_f	R_1	R_2	C_L				
226ohm	230ohm	250ohm	50fF				

Our first pass design achieves its $A_{V_{dc}}$ of 7.6dB, ω_{-3dB} of $2\pi \times 11.5$GHz and a single-

ended $AIIP_3$ of -2.1dBV, which are close to our design targets of $A_{V_{dc}} = 9dB$, $\omega_{-3dB} = 2\pi \times 12$GHz, and $AIIP_3 = -1dB$. The small discrepancies are due to neglecting channel length modulations/body effect, using analytical models with extracted model parameters, and approximations in the nonlinearity expression. Including these higher-order effects masks the design insights and significantly complicates the design process. Therefore, the goal of our methodology is to arrive at an approximate solution, based on which the exact design goals can be achieved by further simple fine tunings.

What needs to be emphasized is that the above discussion does not include provisions for mismatches of both active and passive elements, since it is a strong function of available fabrication process and specific layout techniques which is beyond our discussion here. Taking into account these effects, one may be forced to somewhat increase the transistor sizes with corresponding power penalty. In addition, other circuit techniques, such as sweet-point IIP_3 biasing, gm boosting, and resistive degeneration, etc., are not included here, since they can be potentially superimposed to a standard Cherry-Hooper circuit achieved by this design methodology.

7.4. Design Examples

In this section, three Cherry-Hooper circuit implementations in standard 130nm CMOS process are demonstrated to validate the effectiveness of the proposed design methodology. The critical design values will be provided. Note that although only standard transistors are used here, further power saving is possible by designs with low-threshold voltage devices.

7.4.1 A DC-19GHz Broadband Buffer Amplifier with 10dB Gain

The circuit topology is shown in Figure 7.9. To achieve better output matching during measurement, the amplifier cascades with a source follower buffer whose effect is de-embedded from the reported results. The chip microphotograph and the measurement setup are shown in Figure.7.10. The differential RF inputs and outputs are measured by coplanar S-G-S probes. Discrete capacitors together with on-chip bypass capacitors are used to eliminate supply resonances in this broad bandwidth operation. The extracted frequency response of the voltage gain is shown in Figure 7.11. The simulated and measured differential $AIIP_3$ values at 1GHz are 0.7dBV and -0.8dBV, respectively.

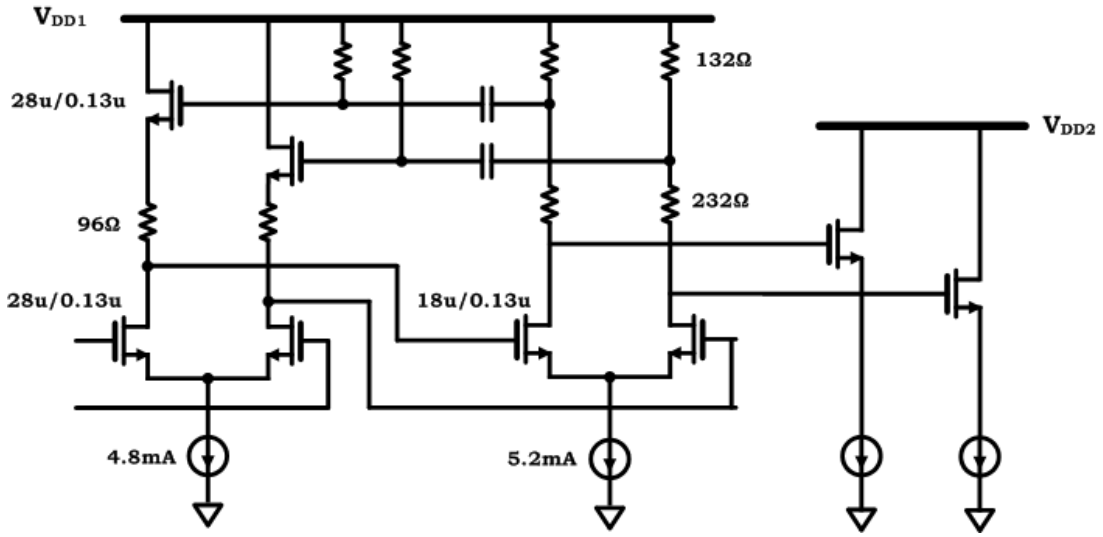
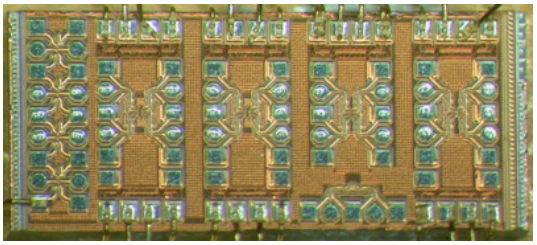
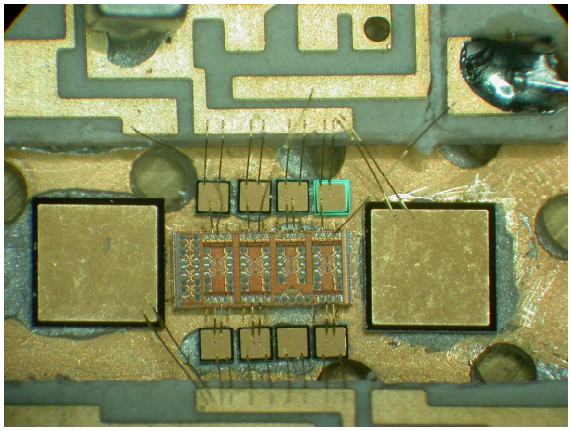


Figure 7.9: Schematic of the broadband Cherry-Hooper buffer



11.a



11.b

Figure 7.10: Chip Microphotograph (11.a) and Measurement Setup (11.b) for the broadband Cherry-Hooper buffer

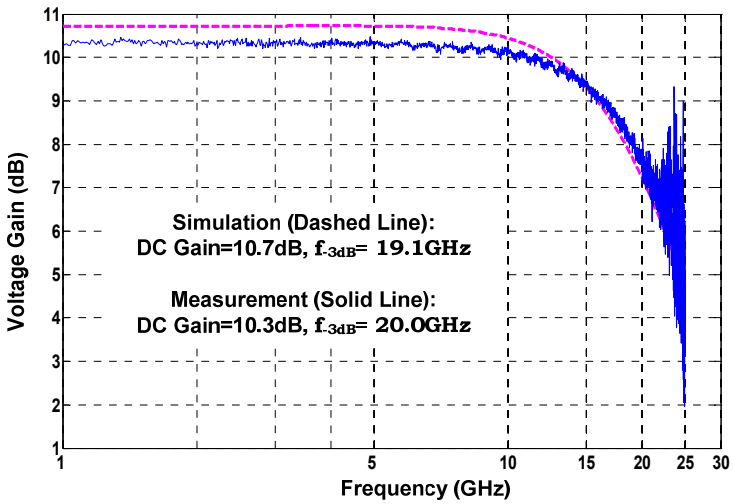


Figure 7.11: Simulated and measured voltage gain of the broadband Cherry-Hooper buffer loaded by the source follower buffer

7.4.2 A DC-12GHz Broadband Phase Rotator with a 10-Bit Resolution

The phase rotator topology introduced in [43] is able to achieve high resolution and immunity to P.V.T. variations. However, the employed current-commutating topology inevitably doubles the capacitances at the current summing node, which needs inductive peaking for bandwidth extension. This peaking inductor has occupied a large percentage of the LO layout area in the 6-to-18GHz dual-band quad-beam phased array receiver presented in Chapter 6. Here we will solve this problem by a modified Cherry-Hooper topology and implement it through our proposed design methodology.

The new phase rotator circuit schematic is shown in Figure 7.12. The two current-commutating VGAs convert the input quadrature LO (I/Q) signals into currents each scaled by a 5-bit digital weighting. Instead of a shunt-peaking load, the current summing node is loaded by a transimpedance stage using the 1st Cherry-Hooper topology introduced earlier. This Cherry-Hooper load presents a synthetic inductor by the gyrator effect together with a small resistive load due to the shunt-shunt feedback. Both facts are preferred for broadband current summation, which accounts for the significant bandwidth extension of this design. In addition, the modified load converts the summed current into the voltage domain at its output just like a standard Cherry-Hooper amplifier. Another Cherry-Hooper stage is cascaded to drive the differential 100ohm load.

The chip microphotograph and the measurement setup are shown in Figure 7.13. The DC supply paths are connected with short wire-bonds. Again, discrete chip capacitors and integrated on-chip capacitors are used to prevent supply resonances in the broad

operation bandwidth.

The measured and the simulated S_{21} are shown in Figure 7.14. The measured I/Q VGAs INL and DNL performance is given in Table 7.5. In a broad bandwidth, it is difficult to generate perfectly matched differential quadrature inputs, and any input I/Q mismatches will cause setup-based artifacts to degrade phase interpolation results. However, if we characterize the staircase gain curve of the I/Q VGAs separately, we can construct the phase interpolation results assuming quadrature inputs. The results at 12GHz (highest frequency) are depicted in Figure 7.15 with the I/Q weighting of -16 omitted for the purpose of symmetry.

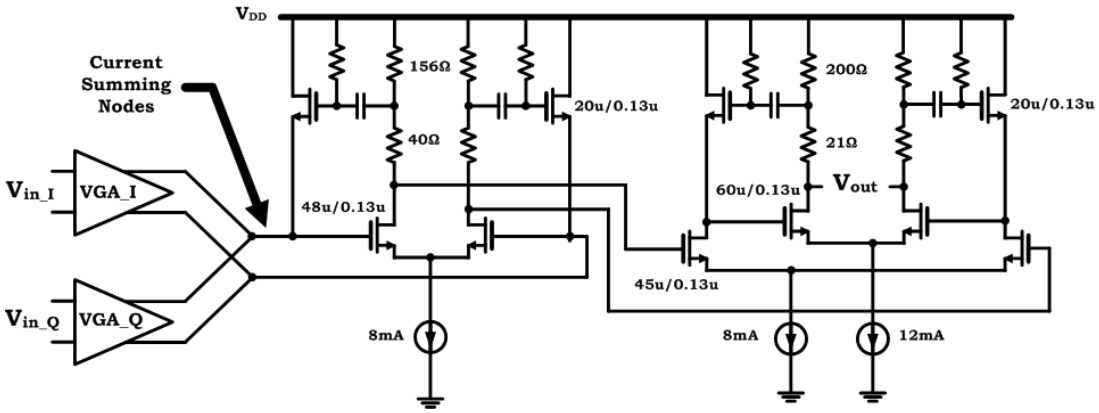
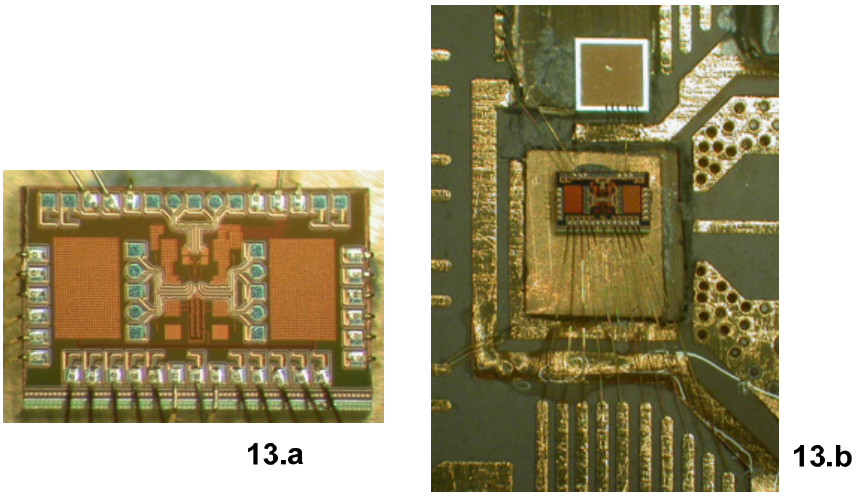


Figure 7.12: The broadband phase rotator with modified Cherry-Hooper topology



13.a 13.b

Figure 7.13: Chip (a) and Module microphotograph (b) for the broadband phase rotator

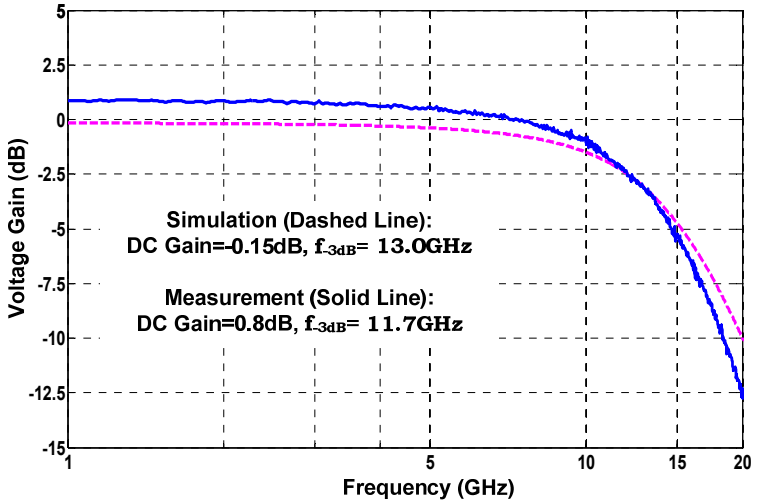


Figure 7.14: Simulated and measured S21 of the broadband Cherry-Hooper phase rotator

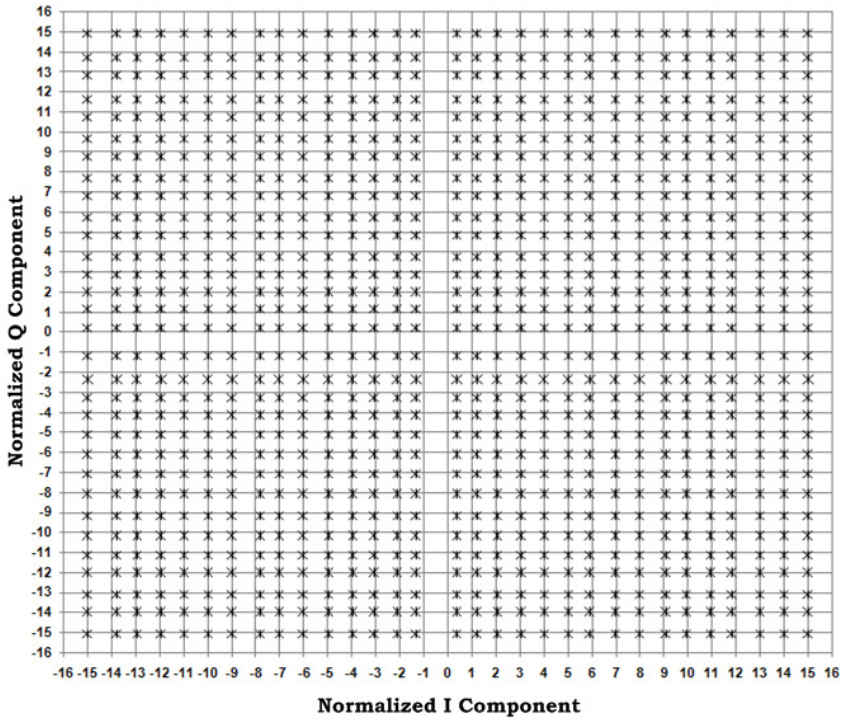


Figure 7.15: Phase interpolation capability of the broadband phase rotator at 12GHz

Note: Each grid represents an ideal interpolation point for the its I/Q weightings, while each cross indicates an interpolation based on measurement.)

TABLE 7.5 INL/DNL SUMMARY FOR PHASE ROTATOR VGAS

VGA I	***2GHz***	**6GHz**	**10GHz**	**12GHz**
INL (LSB)	0.39	0.17	0.30	0.33
VGA I	***2GHz***	**6GHz**	**10GHz**	**12GHz**
DNL (LSB)	0.11	0.07	0.41	0.23
VGA Q	***2GHz***	**6GHz**	**10GHz**	**12GHz**
INL (LSB)	0.17	0.21	0.31	0.36
VGA Q	***2GHz***	**6GHz**	**10GHz**	**12GHz**

DNL (LSB)	0.07	0.13	0.14	0.21

Note: These INL/DNL evaluations exclude the zero-weighting point, whose mismatch errors are mostly caused by measurement setup artifacts.

7.4.3 Dual Beam-Forming Network in Phased Array with LO Phase Shifting

CMOS based phased array systems are used in advanced communication and radar systems, where low-cost, scalability, and concurrent multi-beam operation are needed [45]. An LO path phase shifting scheme is proposed to circumvent design trade-offs among power, noise, and linearity encountered by RF path phase shifters [64]. With this scheme, multi-beam forming can be achieved by concurrent operations of several beam-forming networks, shown in Figure 7.16 as an example. To form N independent beams, the IF and the LO signals are split to 2N paths to achieve N quadrature, phase-shifted baseband signals, which are summed across the array elements to form the beam. However, this topology presents high capacitive loadings for both the IF and the LO buffers, which drive 2N mixers and 2N phase rotators respectively. Moreover, since phase rotators require quadrature inputs, the I/Q components of LO signals need to be delivered to the 2N phase rotator separately, which exacerbates the design challenge. Then, if inductors are used for the buffers and phase rotators, the resulting routings will lead to design issues, such as long signal traces, high parasitic capacitances, and large coupling effects. Therefore, we propose a novel inductorless multi-beam forming

network based on the modified Cherry-Hooper topology, whose implementation is demonstrated here as a design example.

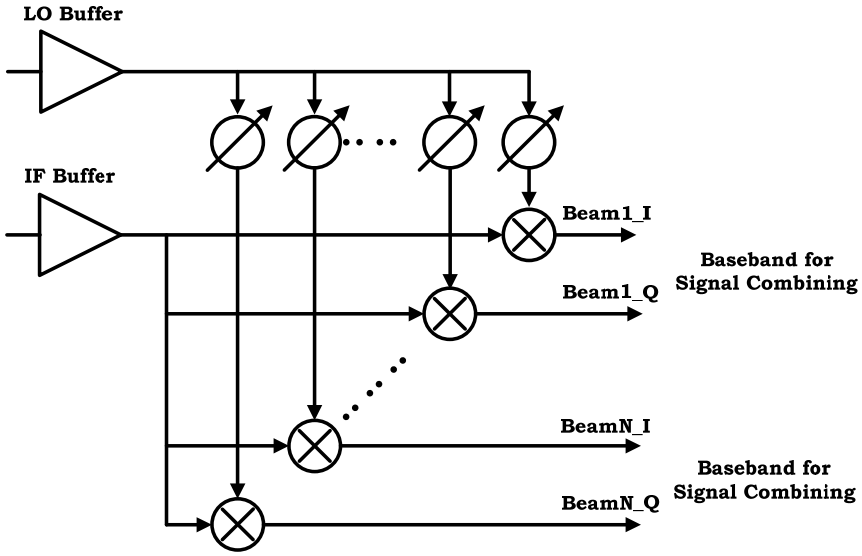


Figure 7.16: Concurrent multi-beam forming network with phase shift at IF mixer

This network covers an IF bandwidth from 1.25 to 6GHz for a 10.4~18GHz broadband phased array receiver, with concurrent dual beam functionality [65]. The LO buffer adopts the 3rd type Cherry-Hooper topology with common-source feedback. Its circuit schematic is shown in Figure 7.17 including the LO distribution lines. Note that the feedback strength can be adjusted by changing the feedback transistor gain through its DC current. The IF buffer uses a modification of the 1st type Cherry-Hooper topology by employing a folded cascode to reduce voltage headroom constraints, as shown in Figure 7.18 with the IF signal distribution lines. The 10-bit phase rotator adopts a similar topology as in example B but with a 5.5dB gain to drive a set of switching-type mixers. The simulated frequency response of this beam-forming network is shown in Figure 7.19. The measured 360º full range phase interpolation is depicted in Figure 7.20. A 4-element phased array system is built based on this receiver, whose dual-beam electrical array

pattern is demonstrated in Figure 7.21.

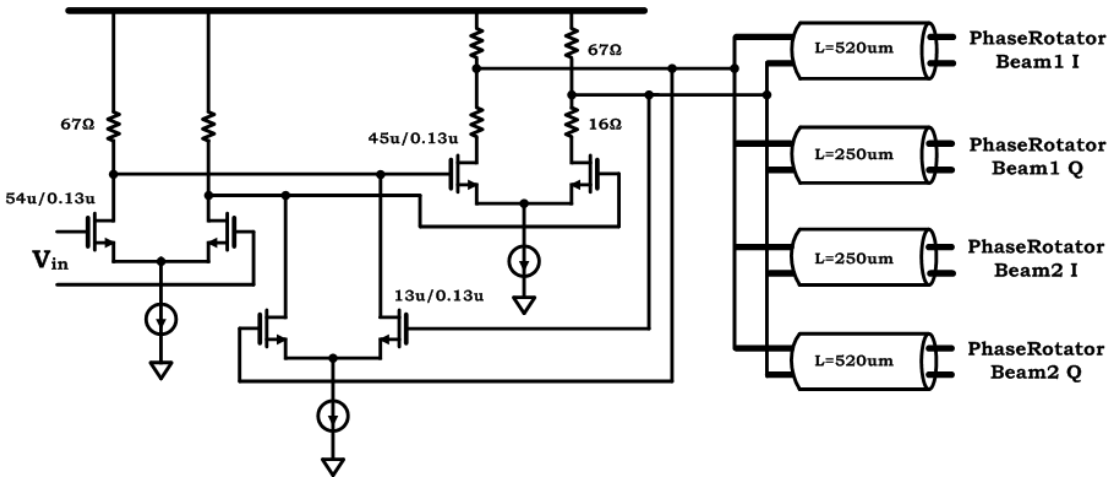


Figure 7.17: Schematic of LO buffer together with its distribution transmission lines

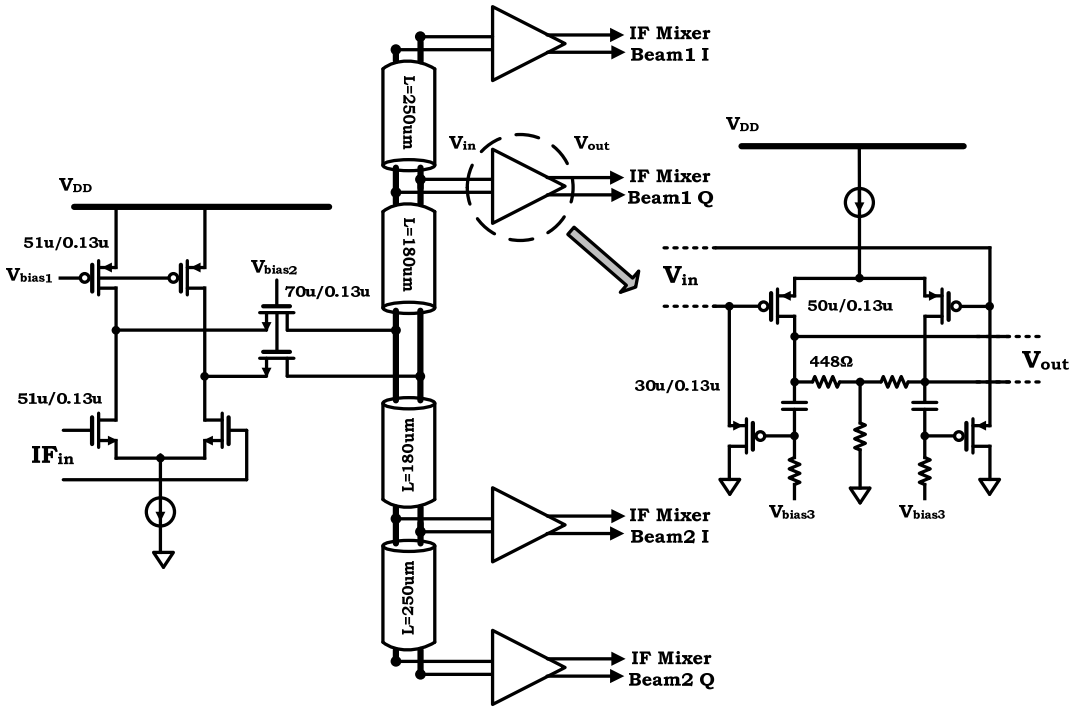


Figure 7.18: Schematic of IF buffer together with its distribution transmission lines

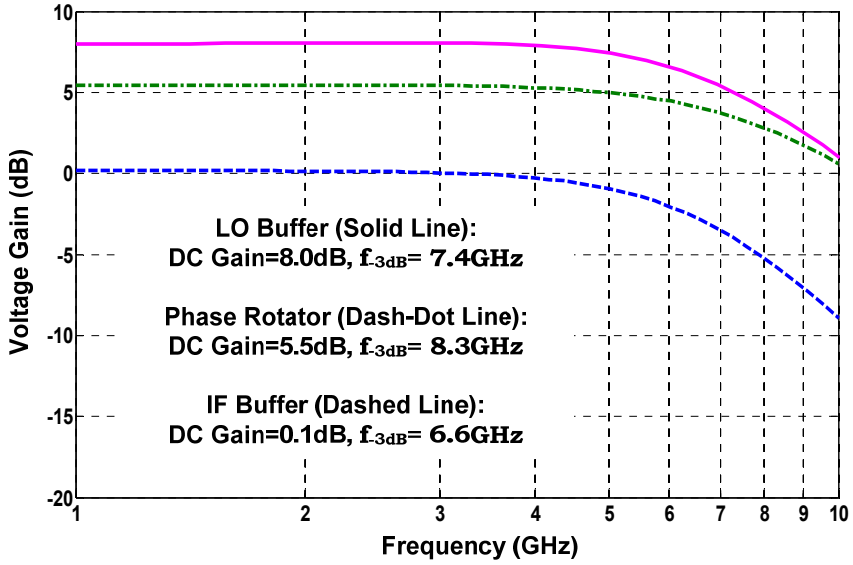


Figure 7.19: Extracted frequency response simulation of the LO buffer, phase rotator and IF buffer with their corresponding distribution networks and loads

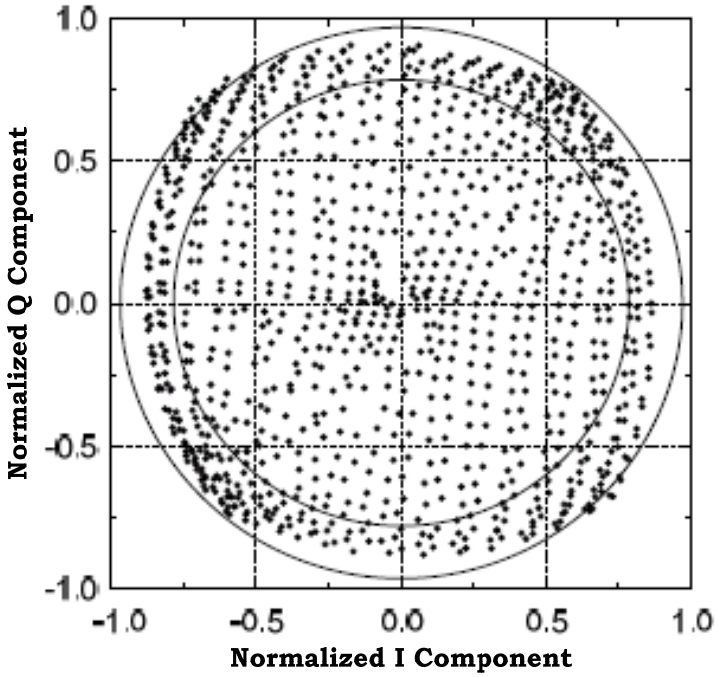


Figure 7.20: Measured 360º full range constellation of the baseband output with 1024 (10^{24}) interpolation points at f_{RF}=18GHz (f_{IF}=6GHz)

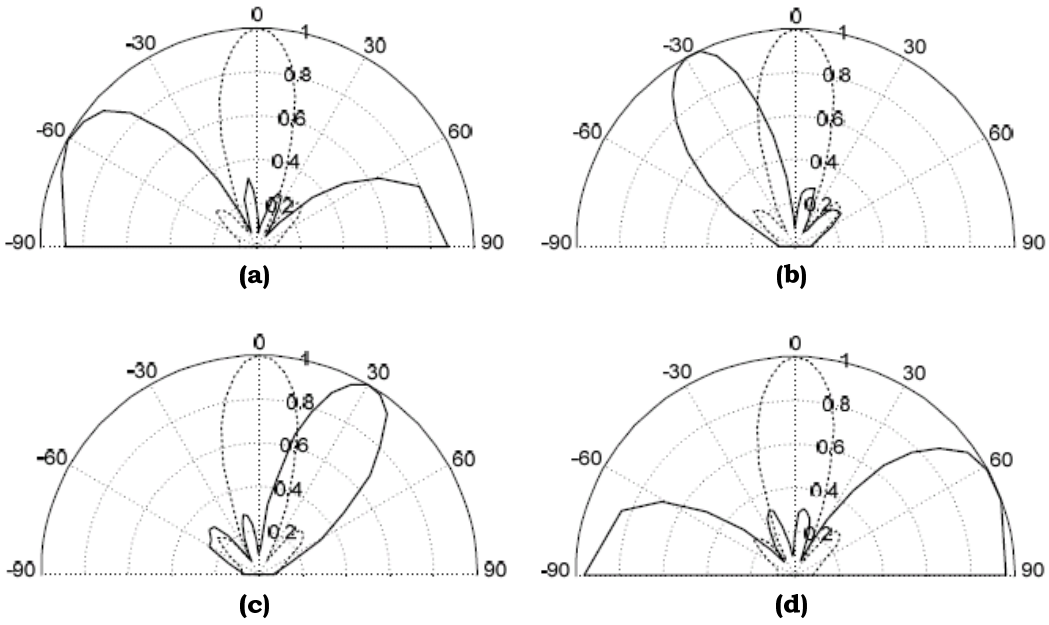


Figure 7.21: Measured concurrent dual-beam array pattern (f_{RF}=18GHz). With beam 1 (dotted line) at 0º incident angle, beam 2 is set to (a) -60º, (b) -30º, (c) 30º, and (d) 60º.

The chip microphotograph is shown in Figure 7.22. The entire network consumes 106.3mA from a 2.5V supply and occupies an area of 380μm×1080μm. In a previously published CMOS phased array receiver, dual beam forming in the same IF frequency range is achieved by extensive shunt peaking [45]. In comparison, the proposed Cherry-Hooper solution achieves 81% of area saving (from 2.16mm^2 to 0.41mm^2) and 28% of power saving (from 328mW to 266mW). Therefore, by carefully choosing and modifying Cherry-Hooper topology, one can achieve broadband multi-beam forming without using inductors in the various broad-band gain stages. The resulting compact layout further simplifies the broadband design due to interconnections length minimization.

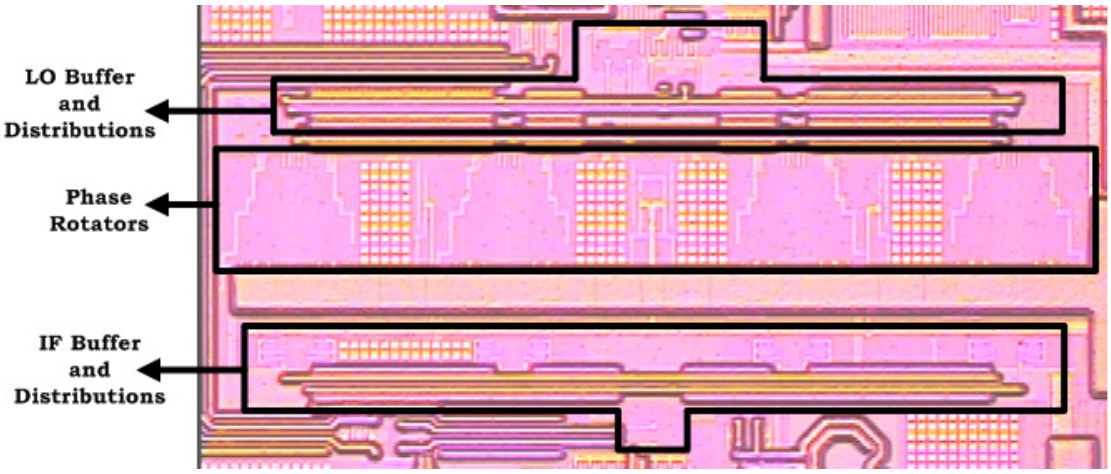


Figure 7.22: Chip microphotograph of the dual-beam forming network

7.5 Chapter Summary

In this chapter, we first reviewed the conventional bandwidth extension techniques, which rely on passive inductor structures. However, those inductors consume large chip area and are not scalable with the process technology. Since in bandwidth extension applications, inductor quality factor generally is not a stringent requirement, we propose to use Cherry-Hooper amplifier topology with active feedback for gain-bandwidth peaking purpose.

Based on the circuit analysis, we propose a design methodology to completely decouple the designs of the forward path and the feedback loop, which greatly simplifies the design procedure. Moreover, further analysis directly reveals the tradeoff among gain-bandwidth, linearity and noise performance. Furthermore, for a specific topology, given specs, such as gain-bandwidth, peaking tolerance, output capacitive loadings, IIP3, and noise performance, an amplifier design can be achieved with minimum power consumption by using our proposed optimization algorithm.

As implementation examples, three Cherry-Hooper based broadband circuit designs are presented together with measurement results. For all the three cases, desired broadband performance has been achieved though this active bandwidth extension strategy. This chapter demonstrates that by eliminating those on-chip passive inductors, broadband systems based on Cherry-Hooper topologies and its modifications can achieve low power designs with very compact layouts, which save substantial amount of die area and significantly increase the system integration level.

Chapter 8: Conclusion

Ultra-high-quality frequency and phase synthesis techniques suitable for practical integrated circuits implementation play a crucial role in novel circuit and system level applications. In this dissertation, a novel biosensing scheme and an advanced phased-array system are presented as two application examples for these precision techniques.

As the first demonstration, a frequency-shift-based magnetic biosensing scheme is introduced. This scheme is to address the PoC biomolecular diagnosis which requires high-sensitivity, ultra-portability and low cost. Compared with existing biosensing schemes, our proposed scheme achieves a competitive sensitivity with no optical devices, no external biasing fields and no expensive post-processing steps. A discrete implementation is first presented to verify the basic sensing mechanism and reveal some important design insights. And an integrated version is designed in a standard 130nm CMOS process, including differential sensing and temperature controlling schemes. Overall, the measured differential sensor noise floor ($\Delta f/f_0$) is 0.13ppm, which ensures reliable detection of one single micron-size magnetic particle (D=4.5μm, 2.4μm and 1μm). Furthermore, the sensor successfully detects real 1n-Molar DNA samples labeled by magnetic nanoparticles.

In the second part, a high-resolution amplitude/phase synthesis technique is proposed to address the mismatch and offset issues encountered by a practical phased array system. It employs a dense Cartesian interpolation scheme with an easily scalable architecture, which achieves a wide operation bandwidth and a constant AC/DC performance for different digital interpolation settings. As an implementation example, a 6-to-18GHz dual-band quad-beam phased array CMOS receiver is presented, which is capable of forming 4

spatially independent beams at two arbitrary frequencies across the 6-to-18 GHz tritave bandwidth. By enabling the phase/amplitude synthesis functions for mismatch compensations, the array element has achieved a maximum RMS phase error of 0.5˚ within a worst-case RMS amplitude variation of 1.5dB for a continuous 360˚ interpolation across the entire 6-to-18 GHz bandwidth. A 4-element phased array receiver system is implemented based on the designed CMOS chip. With the calibration function, the array pattern is measured at 6GHz, 10.35GHz and 18GHz, with the worst case peak-to-null ratio of 21.5dB.

In the third part of the work, a broadband circuit design methodology based on Cherry-Hooper topology is proposed to extend the operation bandwidth without using passive inductors. By applying this technique, significant chip area used for inductive peaking can be saved. As implementation examples, we have shown a DC to 19GHz 10dB gain broadband buffer amplifier, a DC to 12GHz broadband phase rotator with a 10-bit resolution and a beam-forming network in a 10.4GH to 18GHz phased array receiver chip with dual-beam capability. The measurement results thus verify the viability of the proposed bandwidth extension functionality.

8.1 Future Work

As a continuation of this topic, future work would possibly be focused on the following areas.

In terms of the frequency-shift based magnetic biosensor, the sensing inductor layout can be further optimized to result in more spatially homogeneous sensitivity. This effectively increases the sensing area and also helps improve the sensor linearity when a

large number of magnetic particles are attached for high target molecule concentration. Furthermore, novel inductor layout and circuit design techniques can be explored to minimize the sensor footprint while maintaining a stable sensing oscillator operation. This leads to a higher sensor integration level on the same chip area and has its potential for applications such as advanced microarray technology, which targets sequencing the complete human genome on a single sensor array chip in the future.

On the side for wireless communications and phased array radars, although the high-resolution phase and amplitude synthesis enables compensation of those offsets and mismatches, the actual calibration procedures are often time consuming and eventually impractical for a very-large-scaled array. An automatic calibration algorithm based on low circuit/system overhead can be studied which potentially leads to a phased array system with “self-healing” capabilities to adjust its beam forming against any random and time-varying mismatches.

Bibliography

[1] International Technology Roadmap for Semiconductors, http://www.itrs.net/

[2] B. Jang, P. Cao, A. Chevalier, A. Ellington, and A. Hassibi, "A CMOS Fluorescent-Based Biosensor Microarray," *ISSCC Dig. Tech. Papers*, pp. 436–437, Feb. 2009.

[3] B. Rae, C. Griffin, K. R. Muir, J. M. Girkin, E. Gu, D. R. Renshaw, E. Charbon, M. D. Dawson, R. K. Henderson, "A Microsystem for Time-Resolved Fluorescence Analysis using CMOS Single-Photon Avalanche Diodes and Micro-LEDs," *ISSCC Dig. Tech. Papers*, pp. 166–167, Feb. 2008.

[4] M. Schienle, C. Paulus, A. Frey, F. Hofmann, B. Holzapfl, P. Schindler-Bauer, and R. Thewes, "A Fully Electronic DNA Sensor With 128 Positions and In-Pixel A/D Conversion," *IEEE J. Solid-State Circuits*, vol. 39, no. 12, pp. 2438–2445, Dec. 2004.

[5] F. Heer, M. Keller, G. Yu, J. Janata, M. Josowicz, and A. Hierlemann, "CMOS Electro-Chemical DNA-Detection Array with On-Chip ADC," *ISSCC Dig. Tech. Papers*, pp. 168–169, Feb. 2008.

[6] S. Han et al., "A High-Density Magnetoresistive Biosensor Array with Drift-Compensation Mechanism," *ISSCC Dig. Tech. Papers*, pp. 168–169, Feb. 2007.

[7] Y. Liu et al., "CMOS Mini Nuclear Magnetic Resonance System and its Application for Biomolecular Sensing," *ISSCC Dig. Tech. Papers*, pp. 140–141, Feb. 2008.

[8] T. Aytur, P.R. Beatty, and B. Boser, "An Immunoassay Platform Based on CMOS Hall Sensors", *Solid-State Sensor, Actuator and Microsystems Workshop*, Jun. 2002.

[9] P. Besse et al., "Detection of a Single Magnetic Microbead Using a Miniaturized Silicon Hall Sensor," *Appl. Phys.Letters*, vol 80, no. 22, pp. 4199–4201, Jun. 2002.

[10] A. Babakhani, X. Guan, A. Komijani, A. Natarajan, and A. Hajimiri, "A 77 GHz Phased Array Transceiver With On-Chip Dipole Antennas: Receiver And Antennas," *IEEE J. Solid-State Circuits*, vol. 41, no. 12, pp. 2795–2806, Dec. 2006.

[11] A. Natarajan, A. Komijani, X. Guan, A. Babakhani, and A. Hajimiri, "A 77 GHz Phased Array Transceiver With On-Chip Dipole Antennas: Transmitter And Local LO-Path Phase Shifting," *IEEE J. Solid-State Circuits*, vol. 41, no. 12, pp. 2807–2819, Dec. 2006.

[12] A. Babakhani, D. B. Rutledge, and A. Hajimiri, "Transmitter Architectures Based on Near-Field Direct Antenna Modulation (NFDAM)," *IEEE J. Solid-State Circuits*, vol. 43, no. 12, pp. 2674–2692, Dec. 2008.

[13] International Union of Pure and Applied Chemistry, *"Compendium of Chemical Terminology,"* http://www.iupac.org/.

[14] B. Alberts, D. Bray, K. Hopkin, A. Johnson, J. Lewis, M. Raff, K. Roberts, and P. Walter, *Essential Cell Biology*, Garland Science/Taylor & Francis Group, 2003.

[15] U. Nuber, *DNA Microarrays (Bios Advanced Methods),* Taylor and Francis Group, 2005.

[16] www.invitrogen.com.

[17] A. Hassibi, H. Vikalo and A. Hajimiri, "On noise processes and limits of performance in biosensors," *Appl. Phys.Letters*, vol 102, no. 1, pp. 014909–12, Jul. 2007.

[18] G. Li et al., "Detection of Single Micron-Sized Magnetic Bead and Magnetic Nanoparticles Using Spin Valve Sensors for Biological Applications," *J. Appl. Phys.*, vol. 93, no. 10, pp. 7557–7559, May 2003.

[19] D.R. Baselt, G. U. Lee, M. Natesan, S.W. Metzger, P.E. Sheehan, and R.J. Colton, "A Biosensor based on Magnetoresistance Technology", *Biosensor and Bioelectronics*, vol.13, issue 7-8, pp. 731–739, Oct. 1998.

[20] H. Wang, A. Hajimiri, and Y. Chen, *"Effective-Inductance-Change Based Magnetic Particle Sensing,"* U.S. Provisional patent CIT-5108-P, Mar 7th, 2008.

[21] D. Ham and A. Hajimiri, "Virtual Damping and Einstein Relation in Oscillators," *IEEE J. Solid-State Circuits*, vol. 38, no. 3, pp. 407–418, Mar. 2003.

[22] D. E. Bray and R. K. Stanley, *Nondistructive Evaluation*, Taylor and Francis Group, 1997.

[23] D. Griffiths, *Introduction to Electrodynamics*, Benjamin Cummings, 1999.

[24] www.ansoft.com.

[25] T. H. Lee, *The Design of CMOS Radio-Frequency Integrated Circuits,* Cambridge University Press, 2003.

[26] A. Hajimiri and T. H. Lee, *The Design of Low Noise Oscillators*, Springer, 1999.

[27] P. R. Gray, P. J. Hurst, S. H. Lewis, and R. G. Meyer, *Analysis and Design of Analog Integrated Circuits*, John Wiley & Sons, 2001.

[28] Y. Wu and V. Aparin, "A Temperature Stabilized CMOS VCO for Zero-IF Cellular CDMA Receivers*," VLSI Dig. Tech. Papers*, pp. 398-401, 2005.

[29] P. Talbot, A.M.Konn, C. Brosseau, "Electromagnetic characterization of fine-scale particulate composite materials," *J. of Magnetism and Magnetic Materials*, 249, 2002, pp. 481-485.

[30] C. Brosseau, J.B. Youssef, P. Talbot, and A. M. Konn, "Electromagnetic and magnetic properties of multicomponent metal oxides heterostructures: Nanometer versus micrometer-sized particles," *J. of Applied Physics*, Vol 93, No.11, Jun. 2003, pp. 9243–9256.

[31] Applied Thin-Film Product (ATP), http://www.thinfilm.com/.

[32] W. V. Petegem, "Electrothermal Simulation and Design of Integrated Circuit," *IEEE J. Solid-State Circuits*, vol. 29, no. 2, pp. 2795–2806, Feb. 1994.

[33] H. Wang, Y. Chen, A. Hassibi, A. Scherer, and A. Hajimiri, "A Frequency-Shift CMOS Magnetic Biosensor Array with Single-Bead Sensitivity and No External Magnet," *IEEE ISSCC Dig. Tech. Papers*, pp. 438–439, Feb. 2009.

[34] H. Wang and A. Hajimiri, *"Fully Integrated Temperature Regulator for Biochemical Application,"* U.S. Provisional patent CIT-5109-P, Mar 7th, 2008.

[35] H. Wang and A. Hajimiri, *"Ultrasensitive Magnetic Particle Sensor System,"* U.S. Provisional patent CIT-5224-P, Sept. 15th, 2008.

[36] H. Wang and A. Hajimiri, "*Low Cost Bounding Technique for Integrated Circuit Chips and PDMS Devices*," U.S. Provisional patent CIT-5320-P, Feb. 25th, 2009.

[37] B. Huang, et al, "Phospholipid Biotinylation of Polydimethylsiloxane (PDMS) for Protein Immobilization," Lab on a Chip, vol. 6, no. 3, pp. 369–373, Jun. 2006.

[38] Ocean NanoTech, LLC, http://www.oceannanotech.com/.

[39] A. Demir, A.Mehrotra, and J. Roychowdhury, "Phase Noise in Oscillators: A Unifying Theory and Numerical Methods for Characterization," *IEEE Trans. Circuits Syst.*, vol. 47, No.5, pp. 655–674, May 2000.

[40] H. Wang, A. Hajimiri, S. Kousai, "*Noise Suprression Techniques in High-Precision Long-Term Frequency/Timing Measurements*," U.S. Provisional patent CIT-5318, Feb. 20th, 2009.

[41] A. Bhattacharyya, *Phased Array Antennas: Floquet Analysis, Synthesis, BFNs, and Active Array Systems*, New York: J. Wiley & Sons, 2006.

[42] N. Fourikis, *Phased Array-Based Systems and Applications*, New York: J. Wiley & Sons, 1997.

[43] H. Wang and A. Hajimiri, "A wideband CMOS linear digital phase rotator," *IEEE Proc. CICC,* pp. 671–674, Sept. 2007.

[44] S. Jeon, Y.-J. Wang, H. Wang, F. Bohn, A. Natarajan, A. Babakhani, and A. Hajimiri, "A Scalable 6-to-18 GHz Concurrent Dual-Band Quad-Beam Phased-Array Receiver in CMOS," in *IEEE ISSCC Dig. Tech. Papers*, Feb. 2008, pp. 186–187.

[45] S. Jeon, Y.-J. Wang, H. Wang, F. Bohn, A. Natarajan, A. Babakhani, and A. Hajimiri, "A scalable 6-to-18GHz concurrent dual-band quad-beam phased-array receiver in CMOS," *IEEE J. of Solid-State Circuits*, vol. 43, no. 12, pp. 2660–2673, Dec. 2008.

[46] H. Wang, S. Jeon, Y.-J. Wang, F. Bohn, A. Natarajan, A. Babakhani, and A. Hajimiri, "A tunable concurrent 6-to-18 GHz phased array system in CMOS," *IEEE MTT-S Int. Microwave Symp. Dig.*, pp.687–90, Jun. 2008.

[47] F. Bohn, H. Wang, A. Natarajan, S. Jeon, and A. Hajimiri, "Fully integrated frequency and phase generation for a 6–18GHz tunable multi-band phased-array receiver in CMOS," *IEEE RFIC Dig.*, pp. 439–442, Jun. 2008.

[48] S. S. Mohan, M. D. M. Hershenson, S. P. Boyd, and T. H. Lee, "Bandwidth extension in CMOS with optimized on-chip inductors," *IEEE J. Solid-State Circuits*, vol. 35, no. 3, pp. 2311–2320, Dec. 2004.

[49] D. Pi, B. K. Chun, and P. Heydari, "A Synthesis-based Bandwidth Enhancing Technique for CML Buffers/Amplifiers," in *IEEE Proc. CICC*, pp. 467-470, Sep. 2007.

[50] H. Bode, *Network Analysis and Feedback Amplifier Design, Princeton*, NJ: Van Nostrand, 1945.

[51] D. M. Pozar, *Microwave Engineering*, 3rd ed. New York: John Wiley & Sons, 2005.

[52] E.M. Cherry and D.E. Hooper, "The design of wide-band transistor feedback amplifiers," *Proceedings of the I.E.E.*, vol. 110, no. 2, pp. 375, 1963.

[53] C. D. Holdenried, M. W. Lynch, and J. W. Haslett, "Modified CMOS Cherry-Hooper Amplifier with Source Follower Feedback in 0.35μm Technology," in *Proc. of IEEE European Solid-State Circuits Conference (ESSCIRC)*, Sep. 2003, pp. 553-556.

[54] J. Abbott, C. Plett, and J. W. M. Rogers, "A 15GHz, 1.8V, Variable-Gain, Modified Cherry-Hooper Amplifier," in *IEEE Proc. CICC*, Sep. 2005, pp. 645-648.

[55] S. Galal and B. Razavi, "10Gb/s Limiting Amplifier and Laser/Modulator Driver in 0.18μm CMOS Technology," *IEEE J. Solid-State Circuits*, vol. 38, no. 12, pp. 2138–2146, Dec. 2003.

[56] D. H. Garduno and J. S. Martinez, "A CMOS 1Gb/s 5-Taps Transversal Equalizer Based on Inductorless 3rd-order delay cells," in *IEEE ISSCC Dig. Tech. Papers*, Feb. 2007, pp.232-233.

[57] J. Borremans, P. Wambacq, and D. Linten, "An ESD-Protected DC-to-6GHz 9.7mW LNA in 90nm Digital CMOS," in *IEEE ISSCC Dig. Tech. Papers*, Feb. 2007, pp. 422–423.

[58] R. Ramzan, S. Andersson, J. Dabrowski, and C. Svensson, "A 1.4V 25mW Inductorless Wideband LNA in 0.13μm CMOS," in *IEEE ISSCC Dig. Tech. Papers*, Feb. 2007, pp. 424–425.

[59] H. Wang, S. Jeon, Y.-J. Wang, and Ali Hajimiri, "Design methodology for Cherry-Hooper amplifiers with active feedback," in preparation for *IEEE J. of Solid-State Circuits*.

[60] B. Razavi, Design of Analog CMOS Circuits, McGraw Hill Higher Education, 2003.

[61] S. A. Maas, *Nonlinear Microwave and RF Circuits*, 2nd ed. Artech House Publisher, 2003.

[62] S. Kang, B. Choi, and B. Kim, "Linearity Analysis of CMOS for RF Application", *IEEE Proc. Int. Microwave Symp.*, Jun. 2002, pp. 279–282.

[63] A. F. Tong, K. S. Yeo, L. Jia, C. Q. Geng, J. G. Ma, and M. A. Do, "Simple and Accurate Extraction Methodology for RF MOSFET valid up to 20GHz", *IEE. Proc. – Circuits Devices Syst.*, vol. 151, no. 6, pp. 587-592, Dec. 2004.

[64] X. Guan, H. Hashemi, and A. Hajimiri, "A fully integrated 24-GHz eight-element phased-array receiver in silicon," *IEEE J. Solid-State Circuits*, vol. 39, no. 12, pp. 2311–2320, Dec. 2004.

[65] S. Jeon, H. Wang, Y.-J. Wang, F. Bohn, A. Natarajan, A. Babakhani, and A. Hajimiri, "Wideband Multi-Beam Phased-Array Systems in CMOS," to be submitted to *IEEE Trans. Microw. Theory Tech.*