Integrated RC Frequency References in Standard CMOS

Gürleyük, C.

DOI
10.4233/uuid:140a9735-4956-4875-a334-ea80cd49f29f

Publication date
2024

Document Version
Final published version

Citation (APA)
Gürleyük, C. (2024). *Integrated RC Frequency References in Standard CMOS*. [Dissertation (TU Delft), Delft University of Technology]. https://doi.org/10.4233/uuid:140a9735-4956-4875-a334-ea80cd49f29f

Integrated RC Frequency References in Standard CMOS

Integrated RC Frequency References in Standard CMOS

Dissertation

for the purpose of obtaining the degree of doctor
at Delft University of Technology,
by the authority of the Rector Magnificus Prof. dr. ir. T.H.J.J. van der Hagen,
chair of the Board for Doctorates,
to be defended publicly on
Tuesday 22, October 2023 at 15:00 o'clock

by

Çağrı GÜRLEYÜK

Master of Science in Electronics Engineering,
İstanbul Technical University, İstanbul, Türkiye,
born in İstanbul, Türkiye.

This dissertation has been approved by the promotor.

Composition of the doctoral committee:

Rector Magnificus,	voorzitter
Prof. dr. K.A.A. Makinwa,	Delft University of Technology, promotor
Dr. Fabio Sebastiano,	Delft University of Technology, promotor
Independent members:	
Prof. dr. Ian Galton,	UC San Diego, US
Prof. dr. ing. L.C.N. de Vreede	Delft University of Technology
Prof. dr. ir. Klaas Bult,	Delft University of Technology
Prof. dr. Yuongcheol Chae,	Yonsei University, Korea
Dr. ir. A.J. Annema,	University of Twente

ISBN 978-94-93391-62-8

An electronic version of this dissertation is available at
http://repository.tudelft.nl/.

Contents

Summary

This thesis describes the design, implementation, and characterization of integrated RC frequency references. The primary focus of the work is a frequency-locked loop (FLL) architecture, realized in standard CMOS, that digitizes the phase shift of an integrated RC filter, processes the results in the digital domain, and uses it to control an output frequency to maintain its accuracy over temperature variations.

Chapter 1 introduces the thesis by discussing the progress of accuracy in frequency measurements and its impact on human development throughout history. Quartz crystals are presented as the primary frequency reference of the electronic era. While cheap to produce and widely available, quartz crystals pose a serious shortcoming: they are only available as discrete components. The economic trend of the past few decades is to integrate more functional components on a single monolithic die to achieve systems-on-chip (SoC). MEMS (microelectromechanical systems) and BAW (bulk acoustic wave) in the form of multi-chip modules present possible alternatives, even though they do not achieve full integration. This work's motivation is to realize a frequency reference compatible with a standard CMOS process and with accurate enough to enable wireline communication apllications without a quartz crystal.

Chapter 2 presents a literature survey of the state-of-the-art integrated frequency references. Integrated frequency references are studied in three categories: their on-chip time constant, oscillator architecture, and the type of temperature compensation. Time constants available in a CMOS process are limited to the few common choices of RC, LC, and thermal diffusivity (TD) and determine performance parameters such as accuracy and power consumption. Oscillator architectures such as harmonic, relaxation oscillators, and frequency-locked loops are discussed, as they generally degrade the accuracy of their time constant. Finally, the temperature sensitivity of both the time constant and the oscillator must be addressed through analog or digital temperature compensation. A survey of integrated frequency references published in the last two decades determines a research direction towards the motivation of the thesis: an FLL-based RC reference with digital temperature compensation.

Chapter 3 discusses the digital FLL-based RC frequency reference solu-

tion through analysis of its components. The Wien Bridge, an RC band-pass filter, is proposed as the time constant due to its high sensitivity and simple architecture. Synchronous demodulation, integrated into a phase-domain delta-sigma modulator (PDDSM), is proposed to digitize the phase shift of this RC filter. This digitized word is then used to control a digitally controlled oscillator (DCO) after being processed by a digital loop filter (DLF). The dynamics of this digital frequency control loop are studied, laying the groundwork for the following experimental work.

Chapter 4 introduces the Dual-RC frequency reference. Instead of a traditional digital temperature compensation loop, the Dual-RC reference combines the phase shift of two WBs made with different resistors with TCs of opposite signs to achieve a temperature-independent frequency. It features PDDSMs to digitize the phase shift of the two WBs accurately and features digital domain polynomial processing to linearize and combine the two channels. The prototype 7 MHz reference is implemented in a standard CMOS process and achieves ±162.5 ppm inaccuracy from -45 °C to 85 °C, showing a 5.5x improvement to the state of the art. It boasts a 330 ppb Allan Deviation floor in a 3s measurement time, displaying a 12x improvement over the state-of-the-art.

Chapter 5 builds on the experimental results of the Dual-RC reference to improve accuracy and increase the industrial applicability of the design. Utilizing a low-TC WB for the frequency-sensitive channel and a resistor-based temperature sensor, its structure resembles a traditional temperature-compensated digital FLL with simplified digital domain polynomial processing. A closed-loop trimming methodology is proposed to determine the coefficients of the polynomials very accurately. In addition to maintaining the noise performance of the previous design, the prototype 16 MHz reference implemented in a standard CMOS process achieves ±90 ppm inaccuracy -45 °C to 85 °C for 20 samples, a 2.5x improvement to the state-of-the-art Dual-RC design.

Chapter 6 summarizes the main findings of the thesis. RC-based FLLs with digital temperature compensation can achieve below ±300 ppm inaccuracy, enabling wireline communication applications in systems-on-chip. Additionally, this chapter features contemporary work that has been published during the work in this thesis. Finally, the future direction for the research is proposed. While this thesis primarily focused on temeprature compensation, mechanical stress and aging still pose limitiations on the accuracy of integrated frequency references, and work on addressing these issues is already ongoing.

Samenvatting

Deze proefschrift beschrijft het ontwerp, de implementatie en de karakterisering van geïntegreerde RC-frequentiereferenties. De primaire focus van dit werk ligt op een frequentievergrendelingslus (FLL)-architectuur, gerealiseerd in standaard CMOS, die de faseverschuiving van een geïntegreerd RC-filter digitaliseert. De resultaten hiervan worden digitaal verwerkt en deze worden gebruikt om een uitgangsfrequentie te controleren om de nauwkeurigheid te behouden bij temperatuurschommelingen.

Hoofdstuk 1 geeft een introductie van de vooruitgang in frequentiemetingen en de impact ervan op de menselijke ontwikkeling door de geschiedenis heen. Quartz-kristallen worden gepresenteerd als de primaire frequentiereferentie van het elektronische tijdperk. Hoewel goedkoop te produceren en breed beschikbaar, hebben quartz-kristallen een belangrijk nadeel: ze zijn alleen beschikbaar als discrete componenten. De economische trend van de afgelopen decennia is om meer functionele componenten op een enkele monolithische chip te integreren om zo system-on-chip (SoC) te realiseren. MEMS (micro-elektromechanische systemen) en BAW (bulk akoestische golven) in de vorm van multi-chip modules bieden mogelijke alternatieven, hoewel ze geen volledige integratie bereiken. De motivatie van dit werk is het realiseren van een frequentiereferentie die compatibel is met een standaard CMOS-proces en nauwkeurig genoeg is om communicatie via vaste lijnen mogelijk te maken zonder een quartz-kristal.

Hoofdstuk 2 biedt een literatuuronderzoek naar de huidige stand van geïntegreerde frequentiereferenties. Deze worden bestudeerd in drie opzichten: hun on-chip tijdconstante, de oscillatorarchitectuur en het type temperatuurcompensatie. Tijdconstanten die beschikbaar zijn in een CMOS-proces zijn beperkt tot enkele gemeenschappelijke keuzes zoals RC, LC en thermische diffusiviteit (TD) en bepalen prestatieparameters zoals nauwkeurigheid en energieverbruik. Oscillatorarchitecturen zoals harmonische, relaxatieoscillatoren en frequentievergrendelingslussen worden besproken, aangezien ze over het algemeen de nauwkeurigheid van hun tijdconstante verminderen. Ten slotte moet de temperatuurgevoeligheid van zowel de tijdconstante als de oscillator worden aangepakt via analoge of digitale temperatuurcompensatie. Een overzicht van geïntegreerde frequentiereferenties gepubliceerd in de afgelopen twee decennia bepaalt de onderzoeksrichting naar de

motivatie van de scriptie: een op FLL gebaseerde RC-referentie met digitale temperatuurcompensatie.

Hoofdstuk 3 bespreekt de digitale FLL-gebaseerde RC-frequentiereferentie-oplossing door analyse van zijn componenten. De Wien Bridge (WB), een RC-banddoorlaatfilter, wordt voorgesteld als de tijdconstante vanwege zijn hoge gevoeligheid en eenvoudige architectuur. Synchrone demodulatie, geïntegreerd in een fase-domein delta-sigma modulator (PDDSM), wordt voorgesteld om de faseverschuiving van dit RC-filter te digitaliseren. Dit gedigitaliseerde woord wordt vervolgens gebruikt om een digitaal gestuurde oscillator (DCO) te controleren, nadat het is verwerkt door een digitaal lusfilter (DLF). De dynamiek van deze digitale frequentiebesturingslus wordt bestudeerd, waarmee de basis wordt gelegd voor het volgende experimentele werk.

Hoofdstuk 4 introduceert de Dual-RC frequentiereferentie. In plaats van een traditionele digitale temperatuurcompensatielus, combineert de Dual-RC referentie de faseverschuiving van twee WBs gemaakt met verschillende weerstanden, met TC's van tegengestelde tekens, om een temperatuuronafhankelijke frequentie te bereiken. Het bevat PDDSM's om de faseverschuiving van de twee WBs nauwkeurig te digitaliseren en bevat digitale polynomiale verwerking om de twee kanalen te lineariseren en te combineren. Het prototype van de 7 MHz-referentie is geïmplementeerd in een standaard CMOS-proces en bereikt een onnauwkeurigheid van ±162,5 ppm van -45 °C tot 85 °C, wat een verbetering van 5,5x ten opzichte van de huidige stand van de techniek laat zien. Het prototype vertoont een Allan-deviatie van 330 ppb in een meettijd van 3 seconden, wat een verbetering van 12x is ten opzichte van de huidige stand van de techniek.

Hoofdstuk 5 bouwt voort op de experimentele resultaten van de Dual-RC referentie om de nauwkeurigheid te verbeteren en de industriële toepasbaarheid van het ontwerp te vergroten. Door gebruik te maken van een lage-TC WB voor het frequentiegevoelige kanaal en een op weerstanden gebaseerde temperatuursensor, lijkt de structuur op een traditionele temperatuur gecompenseerde digitale FLL met vereenvoudigde digitale polynomiale verwerking. Een gesloten-lus afstemmethodologie wordt voorgesteld om de coëfficiënten van de polynomen zeer nauwkeurig te bepalen. Naast het behouden van de ruisprestaties van het vorige ontwerp, bereikt het prototype van de 16 MHz-referentie, geïmplementeerd in een standaard CMOS-proces, een onnauwkeurigheid van ±90 ppm van -45 °C tot 85 °C voor 20 samples, wat een verbetering van 2,5x is ten opzichte van het Dual-RC ontwerp.

Hoofdstuk 6 vat de belangrijkste bevindingen van het proefschrift samen. RC-gebaseerde FLL's met digitale temperatuurcompensatie kunnen een on-

nauwkeurigheid van minder dan ±300 ppm bereiken, waardoor ze geschikt zijn voor toepassingen in draad gebonden communicatie in system-on-chip. Daarnaast wordt in dit hoofdstuk werk besproken dat tijdens dit onderzoek is gepubliceerd. Tot slot wordt de toekomstige richting van het onderzoek voorgesteld. Hoewel deze scriptie zich voornamelijk richtte op temperatuur-compensatie, vormen mechanische spanningen en veroudering nog steeds beperkingen voor de nauwkeurigheid van geïntegreerde frequentiereferenties, en er wordt al gewerkt aan het aanpakken van deze problemen.

1

Introduction

Time is a fundamental aspect of existence. Every event happens at a specific point in the time dimension. From our perspective, these events form a unidirectional flow of moments, from the past to the present, extending into the future. A ball allowed to fall freely from the top of a building might now be in the air but will undoubtedly hit the floor in the future. Such observations of ordered sequences of events and the causal relationships between them inspired efforts to quantitatively define these moments of interest and measure the intervals between them.

The primary method of measuring time intervals is by counting reliably repetitive events. The sun rises and sets once a *day*, a weight suspended by a one-meter wire swings about 30 times a *minute*, and the Cs-133 atom transitions between its two hyperfine ground states about nine billion times a *second*. These repeated events all define specific intervals or *periods*. Multiples or fractions of these periods can be used as references to quantify any arbitrary time difference. It is thus possible to quantify the time it takes for a ball to fall: dropped from the top of a 90-meter-high building 2 seconds (or 0.0033 days) ago, it will hit the ground precisely 2.284 seconds (or 0.0038 days) later.

However, all naturally periodic events suffer from variation. Time measurements from perfect observations of the sun's position in the sky can vary by up to 16 minutes in a year (or ±30 parts-per-million, ppm) just due to the eccentricity of the Earth's orbit [1]. A pendulum's rate of oscillation varies not only because of the nonidealities of its mechanism but also because of ambient temperature or humidity variations. These cause the period of even the best-constructed pendulums to vary by a few seconds in a day,

or about ±120ppm [2]. These variations add *inaccuracy* to the said time measurement.

The widely accepted International System of Units defines the standard period chosen for time measurements as the second, *s*. The second is a base unit: one of seven such units based on constants and used to define all other units encountered in nature. Unlike the other base units, its definition depends on the measurement of a natural phenomenon. In the SI system, a second is defined as the time it takes for 9 192 631 770 ground-state hyperfine transitions of the Cs-133 atom at rest. Many other units are derived from this definition, including the unit for frequency, the number of occurrences of a repeating event per second. The derived unit for frequency is Hertz, Hz, in the units of s^{-1}, the reciprocal of one second.

When the current definition of a second was accepted in 1967 [3], it represented the highest accuracy and stability of any measurement. The transition rate of Cs-133 is based on its fundamental physics: excited Cs-133 atoms will transition back to their base state at this rate at any location, unaffected by external factors such as gravity or temperature. A typical Caesium frequency standard has a relative accuracy of 5×10^{-13} [4]: two measurements taken from different units will thus vary by one second in approximately 63 thousand years. This represents a nine-orders-of-magnitude improvement in accuracy compared to time measurements made with a pendulum clock.

More accurate time measurements enable tighter scheduling of events and activities and increase the accuracy of any other measurement that involves temporal events. Consequently, accuracy in time measurements correlates well with humankind's technological and societal advancement. As human civilization advanced, systems required tighter synchronization between more and more users. These requirements resulted in an exponential increase in the accuracy of time measurements within a few millennia.

1.1. History of Timekeeping

Time measurements in ancient times required little accuracy. Hunter-gatherers only needed to know when the seasons would change to coordinate their movement from pastures to valleys. They could rely on inaccurate observations of the rhythmic patterns of stars in the night sky to deduce and predict the passing of seasons. The agriculturalists that followed a few thousand years later would need to know more precisely when to sow, how long to keep their crops in the soil, and when to reap. Inaccuracies in timekeeping could mean a loss of crop yield and an understocked winter.

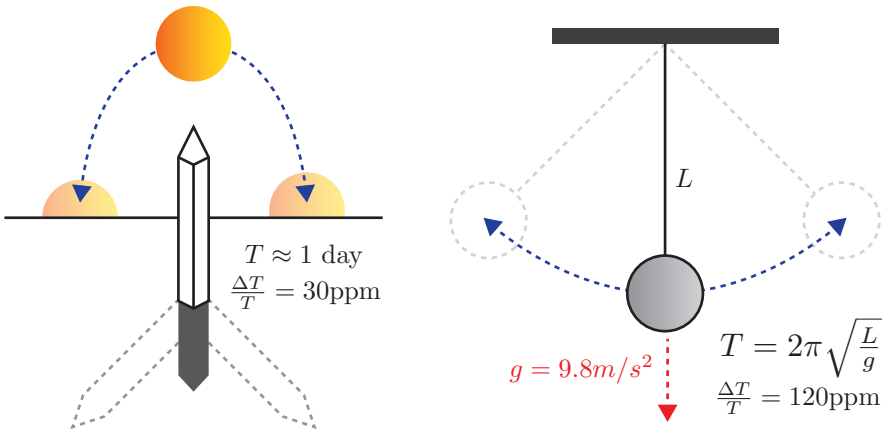


Figure 1.1: Two common ancient timing devices: a sundial and a pendulum, and their relative inaccuracy.

Consequently, it was the agriculturalists who developed the first primitive timing devices and methods for organizing days and seasons into calendars. The time of the day based on the sun's position in the sky, or the *apparent solar time*, would be the general basis of timekeeping. For a few thousand years, such devices remained the primary forms of time measurements.

The next significant improvement in timing accuracy occurred in the 16$^{\text{th}}$ century when Galileo Galilei observed that the rate of a pendulum's swing depends only on its length. Based on this principle, the Dutch scientist Christiaan Huygens developed the first pendulum clock in 1656. With the increased accuracy that this provided, he was able to calculate the error in the apparent solar time due to the eccentricity of the Earth's orbit [1]. Since then, the use of harmonic oscillators for conducting time measurements has been the norm.

The next advances in timing accuracy emerged from navigational requirements. The Sicily naval disaster of 1707 involved the loss of four English Royal Navy fleet ships, causing the death of 2,000 sailors. It occurred because the ships' navigators could not correctly estimate their longitude because they lacked accurate clocks. The pendulum clocks available at the time were affected by variations in the Earth's gravity, making them too inaccurate for use aboard ships. The British government offered a prize for a clock that could work on a ship with better than 3 seconds per day, or ±30 ppm, accuracy. In 1773, the prize was won by John Harrison, whose

H series of chronometers were able to keep time to about 5 seconds in a 10-day trip or to about 6ppm accuracy [5].

The advent of the industrial age required even more cooperation between people. Factories needed to operate on tight schedules to be profitable, and trains needed to run on time to carry workers to these factories. People became more aware of the current time as the rhythm of life started to run on a fixed reference to enable mass cooperation [6]. As a result, in 1880, the British established the Greenwich Mean Time to ensure that their vast empire had a reference time across the globe.

The most significant increase in timekeeping accuracy has occurred in the last century. The development of atomic clocks enabled the most accurate time measurements ever conducted and, in turn, has enabled previously undreamt applications. A great example is GPS, which is composed of satellites housing very accurate atomic clocks that broadcast their current time reading to the entire world [7]. Since they are in known orbits, a person receiving broadcasts from three GPS satellites can triangulate their position anywhere on Earth to an accuracy of about 5 meters, a 10 ppm measurement considering the Earth's circumference. This level of positional accuracy is now available to virtually anyone on Earth who can afford a smartphone. This widespread availability is the result of great strides in electronic engineering, allowing a vast amount of computation to be housed in a portable form factor. The mass-production economics of integrated circuit technology has given everyone easy access to this newfound power.

1.2. Electronic Timekeeping

Electronic systems require accurate timing devices for timing, communication, and synchronization. Synchronous digital circuits are designed with timing constraints: data transfer between a series of flip-flops via asynchronous logic elements has to occur in a specified time to ensure that the input of the receiving flip-flop is ready before the next cycle. Such constraints warrant the need for a stable frequency: if the frequency is too high, it could result in metastability and data corruption.

Frequency accuracy is of particular importance for the synchronization of electronic communication systems. As shown in Figure 1.2, these systems consist of a transmitter that sends data through a channel and a receiver that resamples the data stream. Each side of such a system has a local oscillator that establishes the local timing. Frequency errors between the transmitting and receiving sides result in loss of synchronization and corrupt data transmission. Generally, receivers rely on clock-data-recovery meth-

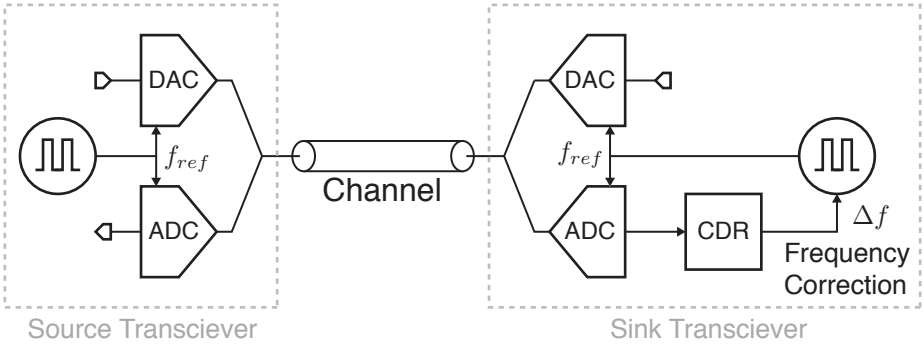


Figure 1.2: Block diagram of an electronic communication system where a sink transceiver uses a CDR algorithm to lock its local frequency to the source frequency

ods that lock their local oscillator to the transmission frequency; however, these algorithms can correct for only a limited range of frequency errors: the capture range. This defines a communication standard's tolerable frequency error. For example, WiFi can only tolerate a ±20 ppm error between transceivers [8].

Another application for accurate frequency references was created by the emergence of the Internet of Things (IoT), which consists of multiple sensor nodes that acquire and exchange data. These generally have very tight power consumption constraints, as they are supplied by battery or harvested power. A traditional solution to reduce the power consumption is to operate them in sleep-wake cycles, during which the device briefly senses and transmits data before going to sleep for a set amount of time [9]. The only block that must run constantly in such a system is the timing unit, bringing additional power consumption constraints. Its frequency inaccuracy, or a lack of long-term frequency stability, then causes increased power consumption and a shorter lifetime.

Fortunately, most electronic systems have access to an accurate frequency reference in a small form factor: the quartz crystal. Virtually all electronic systems today integrate one or more quartz crystals.

1.3. The Quartz Crystal

A quartz crystal consists of silicon and oxygen atoms arranged in a lattice structure. Quartz crystals possess piezoelectric properties: they produce an electric potential when subjected to mechanical forces and produce mechanical stress on the crystal when subjected to an electric potential. This energy

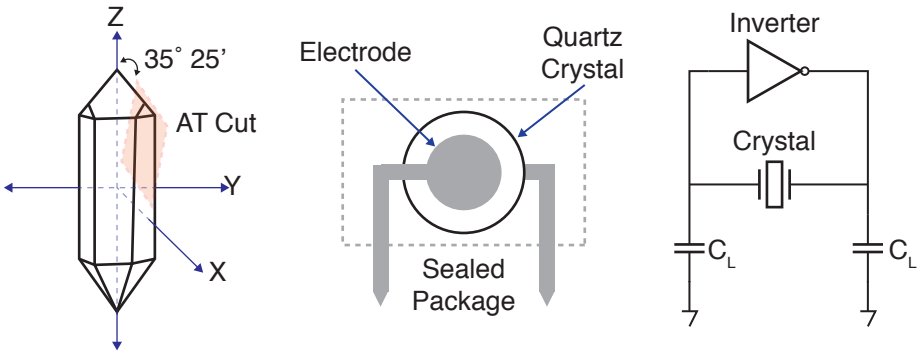


Figure 1.3: Quartz crystal and an AT cut profile, contacted with electrodes and mounted in a hermetically sealed package, finally used in a crystal oscillator circuit using an inverter.

transformation from the electrical to the mechanical domain can be used to excite a crystal into electromechanical resonance. The oscillation between the electrical and mechanical potentials happens at a specific frequency, also called the resonance frequency of the crystal. It is possible to build an oscillator with a quartz crystal by combining this frequency-selective property with an electrical circuit that compensates for the losses of the crystal, as in Figure 1.3.

Because the crystal lattice structure of quartz is very regular and consistent, the resonant frequency of a quartz crystal is very accurately determined by its size and shape. Specific crystal orientations can be adopted to lower the temperature or stress sensitivity of the resonance frequency. Quartz crystals can be mass-produced using methods of synthetic growth, resulting in a very low impurity density and a very accurate geometry, leading to time constants that are accurate to better than ±100 ppm.

However, quartz crystals suffer from nonidealities, which limit their accuracy and stability. They exhibit temperature dependency, which their cut can reduce to about ±10 ppm, but they require compensation to achieve better accuracy and stability levels. Temperature-compensated (TCXO) or oven-controlled (OCXO) oscillators use a temperature sensor or active temperature control of the crystal to increase temperature stability, but this comes at the expense of increased power consumption and production costs. Quartz crystals are also sensitive to mechanical stress. External forces, such as shocks and vibrations, can result in frequency shifts. Stress-compensated (SC) cuts can be adopted to reduce the sensitivity to mechanical stresses,

but this comes at the expense of lower temperature stability [10]. Finally, quartz crystals are prone to aging due to chemical processes that occur with time. Contaminants from the atmosphere can penetrate the crystal and behave like impurities. Similarly, their metal electrodes can react with the quartz, creating interface oxides [11]. Quartz crystals and oscillators need to be housed in hermetically sealed packages to avoid damping and slow down such aging processes.

The requirements for processing, mounting, contacting, and housing quartz crystals make them discrete components. Being only available in such a form is the main disadvantage of quartz crystals. Processes developed to build high-stability quartz oscillators are incompatible with modern integrated circuit processes, and the two cannot be produced on the same die. Electronic systems incorporating an external quartz crystal must sacrifice precious board space. Moreover, having to source an extra component incurs additional costs. These factors have driven the search for an integrated frequency reference to replace quartz crystals.

1.4. Integrated Frequency References

Integrated circuit technology enables the realization of numerous electronic devices on a single monolithic silicon wafer. These devices can be combined to build functional blocks, such as amplifiers, data converters, and central processing units, allowing for better performance, miniaturization, and economical fabrication of systems of immense scale. The lack of an integrated frequency reference as a functional circuit block motivates the search for quartz alternatives.

Two such alternatives that have achieved commercial success are BAW (bulk acoustic wave) resonators and MEMS (micro-electromechanical systems) (Figure 1.4). BAW resonators [12] are constructed by depositing thin piezoelectric materials on the surface of a silicon substrate and rely on the same piezoelectric properties that crystal oscillators possess to build frequency references. They are generally constructed as separate chips that can be stacked on top of a traditional integrated circuit. In contrast, MEMS [13] resonators are miniaturized mechanical tuning forks, generally constructed by etching silicon dies to build a physical structure that shows mechanical resonance. However, their manufacturing process is also not compatible directly with standardized IC processes. BAW and MEMS oscillators generally feature a two-chip system, one for the resonator and the other for the electronic circuits that sustain oscillation and provide temperature compensation [12, 13]. These systems can achieve a higher level of

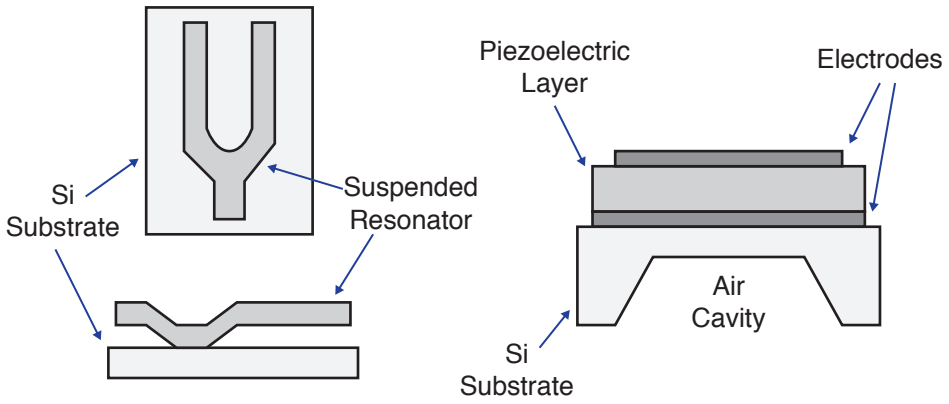


Figure 1.4: MEMS and BAW resonators, crystal replacements used for frequency references that enable multi-die integration.

integration than hermetically-sealed quartz crystal oscillators and so have been commercially successful, which is an affirmation of the market desire for further integration of frequency references. Despite their performance and commercial success, multi-die-based systems still need to achieve full integration. A fully integrated frequency reference must be compatible with a standard CMOS process, using only the limited devices available. Only then can the frequency reference block be co-integrated to realize a complete system-on-chip.

Unfortunately, integrated circuit processes have yet to yield a set of devices that match the stability and accuracy of MEMS or BAW resonators. Frequency references based on LC oscillators [14] made a short-lived commercial appearance for low-accuracy applications in the early 2010s. While they were the only example of fully integrated frequency references at the time, they displayed electromagnetic susceptibility, which caused them to be taken off the market. The accuracy of RC oscillators has generally been limited to about ±1000 ppm, which ruled them out as viable contenders for many applications as the integrated frequency reference solution. The problem of realizing fully integrated frequency references is still unsolved, and taking a step towards a solution forms the primary motivation of this thesis.

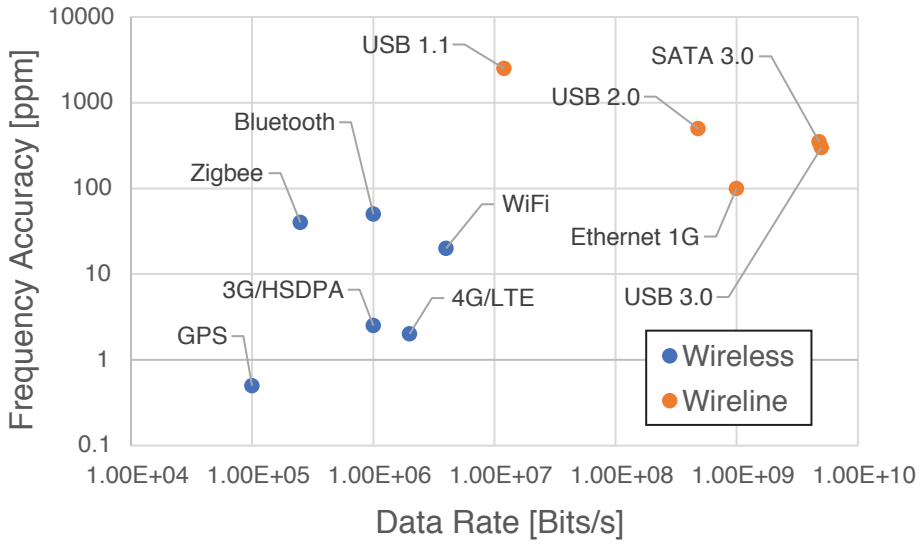


Figure 1.5: Frequency accuracy (absolute, DC) requirements of wireless and wired communication standards vs. their maximum bitrate

1.5. Motivation and Challenges

Frequency references that utilize multi-chip solutions have already achieved the goal of combining package-level integration with accuracy comparable to that of a crystal oscillator. These devices achieve better than ±100 ppm inaccuracy and are economically viable quartz replacements. However, they still require extra production steps, so the challenge of building low-cost, fully integrated frequency references remains.

It should be noted that many important applications do not require crystal-level performance. For instance, the frequency accuracy required by modern communication standards depends on the data rate and the losses of the communication medium. As shown in Figure 1.5, wireless communication standards, such as WiFi[8] or Bluetooth[15], generally require inaccuracies below ±10 ppm, which is readily attainable by relatively inexpensive crystals but cannot be matched by on-chip components. Wireline communication standards that operate in cabled environments, such as USB[16] and SATA[17], can tolerate more frequency inaccuracy, up to ±350 ppm.

The primary goal of this work is to design a frequency reference in standard CMOS that achieves better than ±300 ppm inaccuracy. This is quite challenging because on-chip elements suffer from fabrication tolerances and

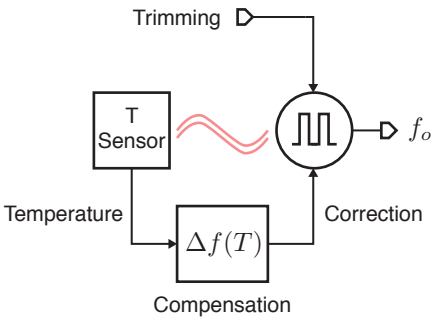


Figure 1.6: Idealized block diagram of a trimmed and temperature-compensated frequency reference

environmental variations. Due to fabrication tolerances, multiple samples of the same frequency reference will have different output frequencies while environmental factors, such as temperature, will cause the frequency to drift over its lifetime.

As shown in Figure 1.6, these issues can, in principle, be addressed by the use of trimming and temperature compensation. However, implementing these techniques comes with new challenges. The effectiveness of trimming is limited by the presence of system noise (frequency jitter), while the effectiveness of temperature compensation schemes is limited by the non-linearity and spread in the temperature dependency of on-chip elements. Furthermore, such schemes also need to be fast enough to track thermal transients.

1.6. Organization of the Thesis

This chapter briefly introduces time measurements, their place in history, and the challenges associated with designing (integrated) electronic frequency references. The rest of this thesis describes the implementation of accurate frequency references in a standard CMOS process using trimming and digital temperature compensation. The goal is to achieve better than ±300 ppm inaccuracy over a wide temperature range, thus enabling the implementation of wireline telecommunication protocols without the need for external components.

Chapter 2 presents a literature survey of the state-of-the-art in integrated frequency references at the conception of this work. Building an on-chip frequency reference is divided into three distinct steps: choosing an on-chip time constant, an oscillator architecture, and finally, implementing temper-

ature compensation. Various choices for each step are discussed, together with their advantages and limitations. A survey of the state-of-the-art motivates the exploration of frequency references in which the frequency of an oscillator is locked to an RC time constant and the use of digital techniques to compensate for their nonlinear temperature dependence.

Chapter 3 discusses such frequency-locked loops in more detail. Synchronous demodulation and delta-sigma conversion are proposed to digitize the phase shift of an RC filter. This information can then be used to set the frequency of a digitally controlled oscillator. The operation of such phase-domain digitizers is analyzed, and their main error sources are discussed. Then, the dynamics of the resulting frequency-locked loops are studied. This chapter lays the foundations for the architecture used in the experimental work of chapters 4 and 5.

Chapter 4 proposes the Dual-RC frequency reference. Instead of using a traditional temperature compensation loop, which typically requires a separate temperature sensor, the phase shift of two RC filters with different temperature dependencies is digitized. This information is then combined in the digital domain to achieve a low-TC frequency error signal in a digital frequency-locked loop. The Dual-RC reference implemented in a standard CMOS process achieves ±250 ppm inaccuracy from -45 °C to 85 °C for 12 samples with a two-point trim and higher-order digital polynomial correction. The work in this chapter represents the first step in reducing the inaccuracy of RC frequency references to significantly below 1000 ppm, showing a 5x improvement to the state-of-the-art.

Chapter 5 describes the implementation of RC frequency reference with improved inaccuracy. An architecture based on a single low-TC Wien Bridge compensated with a complementary temperature sensor is proposed. This architecture aims to address the shortcomings of the Dual-RC implementation to reduce inaccuracy and lower power consumption and area. Moreover, an industrially feasible method for temperature compensation is presented. The architecture is implemented in a standard CMOS process, and a new method for trimming the reference is proposed. The work achieves ±400 ppm inaccuracy from -45°C to 85 °C with a simple, industrially feasible, two-point trim over 20 samples. Moreover, adding a higher-order digital polynomial correction in the loop improves the inaccuracy to ±90 ppm. The work in this chapter represents the second step, showing a 2x improvement in residual inaccuracy compared to the Dual-RC.

Chapter 6 summarizes the main findings of the thesis. The unexplored aspects of RC frequency references are discussed, and a direction for future work on integrated RC frequency references is proposed.

1

References

[1] C. Huygens, *Kort Onderwys aengaende het gebruyck der Horologien tot het vinden der Lenghten van Oost en West,* (1665).

[2] Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character **85**, 505–526 (1911).

[3] M. A. Lombardi, T. P. Heavner, and S. R. Jefferts, *NIST primary frequency standards and the realization of the si second,* NCSLI Measure **2**, 74–89 (2007).

[4] *5071A Primary Frequency Standard Datasheet*, Microsemi (2019).

[5] E. Bruton, *The History of Clocks and Watches* (London: Little, Brown, 2000).

[6] Y. N. Harari, D. Vandermeulen, and D. Casanave, *Sapiens* (Albin Michel, 2020).

[7] D. W. A. N. A. C. C. Hodge, *The Science of Timekeeping*, Hewlett Packard (1997).

[8] A. Behzad, Z. M. Shi, S. Anand, L. Lin, K. Carter, M. Kappes, T.-H. Lin, T. Nguyen, D. Yuan, S. Wu, Y. Wong, V. Fong, and A. Rofougaran, *A 5-GHz direct-conversion CMOS transceiver utilizing automatic frequency control for the IEEE 802.11a wireless LAN standard,* IEEE Journal of Solid-State Circuits **38**, 2209 (2003).

[9] F. Sebastiano, L. Breems, and M. K. A. A., *Mobility-based time references for Wireless Sensor Networks* (Springer, 2013).

[10] J. R. Vig and A. Ballato, *Method of Making Miniature High Frequency SC-cut Quartz Crystal Resonators,* (1985).

[11] F. Iwasaki, A. H. Shinohara, H. Iwasaki, and C. K. Suzuki, *Effect of impurity segregation on crystal morphology of Y-bar synthetic quartz,* Japanese Journal of Applied Physics **29**, 1139 (1990).

[12] D. Griffith, P. T. Roine, T. Kallerud, B. Goodlin, Z. Hughes, and E. T.-T. Yen, *A ±10ppm −40 to 125°C BAW-based Frequency Reference System for Crystal-less Wireless Sensor Nodes,* in *2017 IEEE International Symposium on Circuits and Systems (ISCAS)* (2017) pp. 1–4.

[13] M. H. Perrott, J. C. Salvia, F. S. Lee, A. Partridge, S. Mukherjee, C. Arft, J. Kim, N. Arumugam, P. Gupta, S. Tabatabaei, S. Pamarti, H. Lee, and F. Assaderaghi, *A Temperature-to-Digital Converter for a MEMS-Based Programmable Oscillator With $< \pm 0.5$-ppm Frequency Stability and $<$ 1-ps Integrated Jitter,* IEEE Journal of Solid-State Circuits **48**, 276 (2013).

[14] M. S. McCorquodale, G. A. Carichner, J. D. O'Day, S. M. Pernia, S. Kubba, E. D. Marsman, J. J. Kuhn, and R. B. Brown, *A 25-MHz Self-Referenced Solid-State Frequency Source Suitable for XO-Replacement,* IEEE Transactions on Circuits and Systems I: Regular Papers **56**, 943 (2009).

[15] *IEEE Standard for Information technology– Local and metropolitan area networks– Specific requirements– Part 15.1a: Wireless Medium Access Control (MAC) and Physical Layer (PHY) specifications for Wireless Personal Area Networks (WPAN),* IEEE Std 802.15.1-2005 (Revision of IEEE Std 802.15.1-2002) , 1 (2005).

[16] *Universal Serial Bus 3.2 Specification*, USB Implementers Forum, Inc. (2022), rev. 1.1.

[17] *Long Term Frequency Accuracy and SSC Profile Tests for Transmitters*, Serial ATA International Organization (2007), rev. 1.0.

2

Integrated Frequency References

Economic motivations often drive innovation in circuit design. In the case of frequency references, the incentive results from increased system costs associated with quartz crystals. The recent success of MEMS-based frequency references confirms the economic value of increased integration.

The main goal of this thesis is to realize fully integrated, CMOS-compatible frequency references with reasonable accuracy. This goal has yet to be fully achieved due to the spread and temperature dependence of the components available in CMOS processes. This chapter will present a systematic study of these sources of inaccuracy in typical CMOS frequency references.

2.1. Introduction

Building a frequency reference involves three distinct steps. Firstly, a stable physical quantity defining the oscillation frequency must be selected. Then, a mechanism must be designed to maintain a stable oscillation at this frequency. Finally, the effect of external sources of frequency variation, such as temperature, humidity, and physical stress, must be corrected.

Building an *integrated* frequency reference means that these steps must be carried out within the confines of integrated circuit technology. Unfortunately, this limits the variety of physical mechanisms that can be used to define the oscillation frequency. On the flip side, integrated circuit technologies exhibit some of the tightest tolerances known to humanity. The accuracy of photolithography allows the definition of geometric structures with very high repeatability, resulting in components with low mismatch.

Also, the materials used in these processes are generally highly pure. The mono-crystalline silicon used in the substrate of integrated circuits has impurity densities of a few parts part-per-billion. So, it should be possible to exploit the strengths of integrated circuit technologies to realize very accurate frequency references.

Another advantage of integrated circuit technology is the availability of significant computational resources, especially in advanced CMOS technologies. The possibility of *going digital* enables complex solutions that are not feasible with analog techniques alone.

This chapter reviews the state-of-the-art in integrated frequency reference design at the conception of this work. First, integrated time constants commonly available in standard CMOS processes and the circuit elements used to build them will be discussed, focusing on their accuracy. Next, circuit architectures to realize oscillators will be presented. The primary error sources in these circuit architectures will be identified. Then, methods used to compensate for these error sources will be studied. Finally, a survey of state-of-the-art integrated frequency references in 2018 will be presented to discern a promising direction for the work conducted in this thesis.

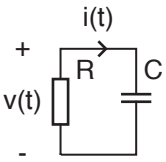
2.2. Integrated Time Constants

The first step in building an integrated frequency reference is the choice of an integrated time constant, i.e., a physical quantity with units of seconds that defines the oscillation frequency. A fully integrated frequency reference should derive its time constant using only the available integrated circuit components. References based on MEMS, BAW, or SAW resonators do not comply with this requirement as they require additional processing steps or package-level integration of multiple chips.

Unfortunately, searching for a standard CMOS-compatible time constant yields only a few options. Combinations of various circuit elements, resistors, capacitors, and inductors are the most common, resulting in resistor-capacitor (RC) and inductor-capacitor (LC) time constants. In addition, thermal time constants can be used to build frequency references [1, 2]. In the rest of this section, these time constants and the circuit elements that make them will be discussed.

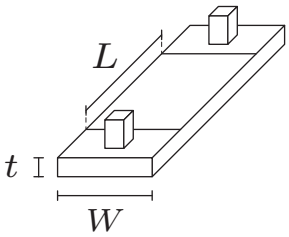
2.2.1. RC - Resistor & Capacitor

A resistor and a capacitor can be combined to realize an RC time constant. The capacitor stores energy as an accumulated charge over a potential difference, and the resistor then dissipates this energy by discharging the ca-



$$\omega_{RC} = \frac{1}{RC}$$

Figure 2.1: RC circuit



$$R = \frac{\rho \cdot L}{t \cdot W} = R_{\square} \frac{L}{W}$$

Figure 2.2: A typical resistor in a planar IC process

pacitor with a current proportional to the voltage across the capacitor and its resistance.

In Figure 2.1, the voltage across the simple RC network decays exponentially from its initial value by:

$$V(t) = V_0 e^{\frac{-t}{\tau_{RC}}} \quad \rightarrow \quad \tau_{RC} = RC \quad [s] \tag{2.1}$$

The time constant τ_{RC} is the characteristic property of this first-order, linear time-invariant system. As expected, the unit of RC is in seconds; thus, an RC network can define the frequency of an oscillator.

Integrated Resistors

Resistors in integrated circuit technologies generally consist of thin planar strips of doped material sandwiched between two insulating layers. The well-known expression in Figure 2.2 gives the resistance R of such a planar structure, with width W, length L, thickness t, and resistivity ρ. Since the thickness of the layer is not a variable parameter in a planar IC process, the parameter $R_{\square}$ defines the resistivity of a square of material with constant thickness.

Nonidealities in the production process limit the accuracy of an integrated resistor. Variations in its dimensions or resistivity will cause its resistance to spread. Relative variations in the lateral dimensions can be mitigated by increasing the size of the resistor. However, the thickness of the layer is less well controlled, showing significant variation. Moreover, resistivity is a function of the doping level and spreads significantly between wafers.

The accuracy of integrated resistors is also limited by their temperature and voltage dependence. A large voltage across a resistor causes charge carrier velocity to approach the maximum drift velocity, causing an increase in resistivity. Similarly, temperature variation can cause changes in charge carrier concentration and mobility. The modulation of resistance due to these external factors can be modeled by the temperature and voltage coefficients $TC_{1...N}$ and $VC_{1...N}$ as in:

$$R(T,V) = \left(R_{\square,0}\frac{L}{W}\right)\left(1 + TC_1(T-T_0) + TC_2(T-T_0)^2 + \ldots\right) \left(1 + VC_1(V-V_0) + VC_2(V-V_0)^2 + \ldots\right) \tag{2.2}$$

Types of Integrated Resistors

Various integrated resistors can be fabricated in a typical planar CMOS process. The most commonly used layers are diffusion, polysilicon, and metal (Figure 2.3). These have different trade-offs between resistivity, absolute and relative accuracy, and voltage and temperature coefficients. The properties and trade-offs of these commonly used integrated resistors are discussed below.

Diffusion Resistors Diffusion resistors consist of bulk silicon doped with positive or negative charge donors. The doping modulates the charge concentration of the material, modulating its resistivity. Highly doped diffusion layers (P^+ and N^+) are used to form the source/drain terminals of MOS transistors. They have relatively low resistivity ($R_\square$=10-100Ω) but have high absolute variation (±25%) due to their dependence on dopant concentration. Their temperature dependence is due to the reduced mobility of charge carriers with increasing temperature, resulting in positive TCs of about 1500 ppm/°C. Lightly doped diffusion layers (P^- and N^-) are used to form the bulk region of MOS transistors. Due to their low impurity count, they provide high resistivity ($R_\square$=1000-5000 Ω) but also have higher absolute variation (±40%). Moreover, they have a much higher voltage coefficient of 8000 ppm/V and a temperature coefficient of 4000 ppm/°C. Diffusion resistors have additional nonidealities, such as the voltage dependence

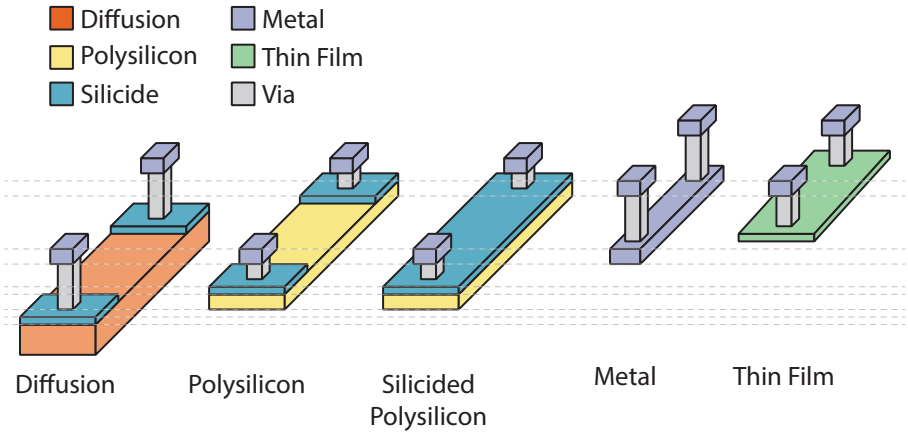


Figure 2.3: Different types of resistors in a standard CMOS process

of their parasitic capacitance, leakage effects, and high mechanical stress sensitivity. Due to these reasons, they are usually not preferred for high-accuracy applications.

Polysilicon Resistors Polysilicon (polycrystalline silicon) is often used as the gate material of short-channel (< 1µm) planar MOS devices. Depending on its doping levels, polysilicon can provide a relatively high resistivity ($R_{\square}$=100-1000 Ω) but with a high absolute resistivity variation (±25%).

Additionally, it is possible to modulate the TCs of polysilicon resistors by modulating their doping concentration [3]. Their TCs can be tuned to cover a broad range, from -1500 ppm/°C to 1500 ppm/°C, including values as low as 100 ppm/°C. Polysilicon resistors are also quite linear, with voltage coefficients of about 100 ppm/V. Being sandwiched between oxide layers, they display relatively low parasitic capacitance, which is desirable in high-speed applications. For these reasons, polysilicon resistors are the preferred choice for building frequency references, as evidenced by their widespread use [4–7]. Polysilicon resistors are used in the RC frequency references in chapters 4 and 5.

Silicided Resistors Doped silicon or polysilicon can be combined with a metal such as nickel, cobalt, titanium, or tungsten to form a silicide, which reduces their resistivity further. These metallic silicon compounds retain the properties of both their constituents. They have low resistivity ($R_{\square}$=5-10 Ω) compared to unsilicided polysilicon, but still much higher than that of

metals. They inherit the temperature sensitivity of the metal, resulting in significant positive temperature coefficients on the order of 3000 ppm/°C [8]. They also exhibit relatively low absolute variation since their resistivity is not dependent on dopant concentration. Their high TCs and low resistivity variation make them especially attractive for high-resolution temperature sensors [9]. Silicided polysilicon resistors are used in the Dual-RC frequency reference presented in chapter 4.

Metal Resistors Metal interconnects, generally used as electrical connections between circuit elements, can also be used to build integrated resistors. They have very low resistivity ($R_{\square}$=50-70 mΩ for thin metal layers, $R_{\square}$=10-40 mΩ for thick metal layers), which makes them suitable for use as interconnects. Their absolute accuracy is high, primarily limited by geometry effects. Metal resistors are also highly linear. However, they have high TCs of 3000 ppm/°C due to the reduction of electron mobility at high temperatures. Metal resistors are not commonly used in circuit design due to the difficulty of achieving high resistances in a feasible circuit area.

Thin Film Resistors Thin films of conductive material deposited onto insulators may be used as integrated resistors. Compounds such as nickel-chromium (NiCr) or silicon-chromium (SiCr) can be deposited with thicknesses ranging from 3-200 nm by sputtering them on an insulating silicon oxide layer. This allows for precise film thickness control, enabling resistors with a wide range of resistivity ($R_{\square}$=30-2000 Ω) [10] with very low absolute resistivity variation. Also, their temperature sensitivities can be finely tuned to near zero by controlling their annealing properties [11]. These properties make thin-film resistors attractive in high-accuracy circuit design, including frequency references [12]. However, their integration requires additional process steps, which makes them a non-standard option in planar CMOS processes. Polysilicon and diffusion resistors are impossible to realize in FinFET processes due to the quantized nature of the transistors. In such processes, thin-film resistors have become the sole option [13].

Integrated Capacitors

Capacitors implemented in integrated circuit technologies generally belong to one of two categories. A parallel plate capacitor can be built using two conducting layers separated by an insulator. The parasitics of active devices, such as diodes formed by PN junctions or MOS gates, can also be used as capacitors.

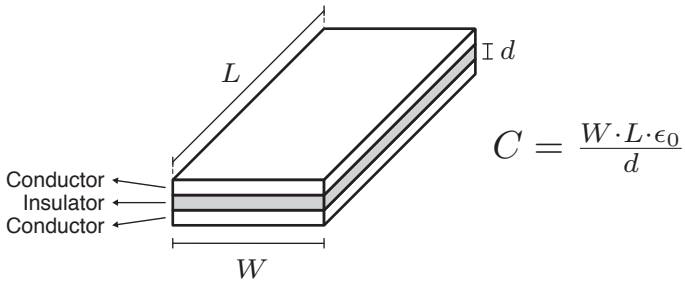


Figure 2.4: A typical parallel plate capacitor in a planar IC process

Capacitors based on the parasitics of active devices are available in all CMOS processes. Usually, these capacitors show good capacitance density (5 fF/μm^2). However, their values depend heavily on doping concentration and so spread significantly. Similarly, their capacitance is generally a function of the depletion region width, which, in turn, is a strong function of voltage and temperature. Their voltage dependence limits their use when large voltage swings occur, as may be expected from oscillators. Their temperature dependence and its variation represent another source of error. Additionally, reverse-biased PN junctions are prone to leakage. The leakage current is a highly temperature-dependent parasitic parameter and behaves like a parallel, nonlinear resistor across the capacitor.

Capacitors that consist of two conducting layers separated by an insulator do not suffer from most of the parasitic effects of device capacitors, making them the preferred choice for applications that require accuracy. They are traditionally constructed as parallel plate structures, as in Figure 2.4, but can be implemented as fingered wires or the more exotic woven or quasi-fractal structures [14]. The primary limit to the accuracy of such capacitors is the thickness variation of their insulating layers, which is usually around 10%. Typically, only metal-oxide-metal (MOM) capacitors are available, where interconnect metals are used as the conducting layers. MOM capacitors have generally high voltage ratings due to the thick oxide layer separating them but consequently have low capacitance density (> 1 fF/μm^2). For frequency reference applications, the capacitor's absolute value (thus the density) is also essential. Metal-insulator-metal (MIM) capacitors, where a thin layer of an insulator is deposited between an intermediate metal layer, are the preferred choice for this application. These capacitors benefit from the high capacitance density granted by the low-thickness insulator but come at the

cost of an additional mask and are generally available only as an extra option.

2.2.2. LC - Inductor & Capacitor

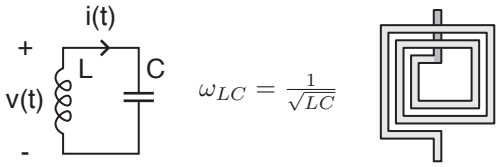


Figure 2.5: LC circuit and a planar inductor

Another commonly used electrical on-chip time constant is the LC tank. Like the capacitor, the inductor is a circuit element that stores energy. As the current flow through the inductor changes, a magnetic field is induced, which creates a potential difference across the inductor that opposes the current flow. The energy required to overcome this potential is stored in the magnetic field. Contrary to an RC circuit, in an LC tank, energy is not dissipated but transferred between the capacitor and the inductor.

The time constant for the LC tank in Figure 2.5 can be derived from its natural frequency: the frequency at which the energy transfer between the two components occurs.

$$i(t) = Ae^{j\frac{-t}{\sqrt{LC}}+j\phi_0} \quad \rightarrow \quad \tau_{LC} = \frac{1}{\omega_n} = \sqrt{LC} \qquad [s] \tag{2.3}$$

The LC time constant is again in the unit of seconds, implying that an LC tank can be used to define the frequency of an oscillator.

Integrated inductors can be realized as planar coils in metal interconnect layers of a circuit process. Generally, higher metal layers with low resistivity are preferred to build inductors, resulting in lower losses and higher quality factors. As the coil geometry defines the inductive properties of these coils, layer thickness variations that reduce the accuracy of integrated capacitors are not of great concern for inductors. Moreover, their temperature coefficients are low and related to the parasitic resistance of the coil and their quality factors [15]. These stable and well-defined properties of inductors make them attractive candidates for building frequency references.

Unfortunately, practical size limitations limit the viable inductance values achievable on-chip. A 200x200 µm multiple-turn coil inductor will generally

have 100s of pH inductance [16]. This limitation and practical area requirements for integrated capacitors result in resonance frequencies in the GHz range for integrated LC tanks. These GHz frequency outputs and the frequency dividers following them generally result in high power dissipation for LC frequency references [17, 18].

2.2.3. Thermal Diffusivity

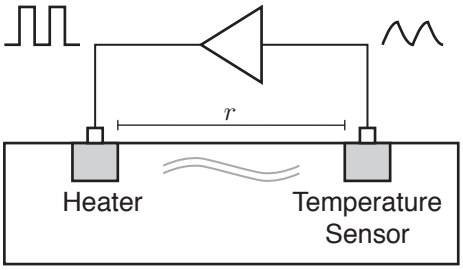


Figure 2.6: The thermal oscillator

On-chip time constants can also be created in the thermal domain. For example, a thermal oscillator can be made [19], which consists of a thermal delay line in an electrical feedback loop. As shown in Figure 2.6, a thermal delay line can be made by implementing a heater and a temperature sensor in close proximity on a silicon chip. Driving the heater generates heat waves that diffuse through the silicon substrate and cause temperature fluctuations. A temperature sensor placed a distance r away senses the temperature in its vicinity and drives the amplifier, closing the feedback loop. The power dissipated at the resistor will cause temperature fluctuations in the substrate that decrease as the distance from the resistor increases. Additionally, due to the thermal inertia of the substrate, it will take a specific time before the temperature change reaches the sensor. The oscillation frequency will be defined by the dynamics of this electrothermal system, known as an electrothermal filter (ETF).

The heat equation that describes the dynamic behavior of an ETF is a partial differential equation that relates the rate of temperature change to the temperature gradient:

$$\frac{\partial T}{\partial t} = \frac{k}{c\rho}\nabla^2 T \tag{2.4}$$

where

$$D = \frac{k}{c\rho} \qquad \left[\frac{m^2}{s}\right] \tag{2.5}$$

2

where k is thermal conductivity, c is the specific heat capacity, and ρ is density.

Electrothermal filters can be analyzed with an electrical analogy between the electrical and thermal domains [20]. The temperature T [K] can be regarded as an electrical potential, and heat flow [W/s] can be regarded as an electrical current. It is then possible to define a thermal impedance Z_{TH} as the ratio of the temperature fluctuations at the temperature sensor, ΔT, to the heat flow, P. For a simplified electrothermal filter with a point heat source and a point temperature sensor, the frequency-dependent thermal impedance is given by [21]:

$$Z_{TH}(\omega) = \frac{1}{2\pi kr} e^{-r\sqrt{\frac{\omega}{2D}}} e^{-jr\sqrt{\frac{\omega}{2D}}} \tag{2.6}$$

Assuming an amplitude control mechanism ensures unity loop gain, the oscillation frequency ω_{TH} for the thermal oscillator of Figure 2.6 is when the phase shift across the filter is 2π.

$$\angle Z_{TH} = -r\sqrt{\frac{\omega_{TH}}{2D}} = 2\pi \tag{2.7}$$

The thermal time constant of the electrothermal filter in seconds can then be derived as:

$$\tau_{TH} = \frac{r^2}{4\pi D} \qquad [s] \tag{2.8}$$

The electrothermal time constant depends on two parameters: the distance between the heater and the sensor and the thermal diffusivity of silicon. Both of these parameters are highly controlled in integrated circuit processes. The bulk material is a single silicon crystal (mono-crystalline silicon) and is one of the purest manufactured materials. Similarly, the parameter r, the distance between the heater and the sensor, is dependent on the accuracy of the photolithographic process and is controlled to a sub-nm level in modern CMOS processes. Relatively small electrothermal filters that promise high accuracy can be built by taking advantage of the high level of control [2].

The primary disadvantage of electrothermal filters is the high power consumption of the heater. The only way to increase the temperature signal

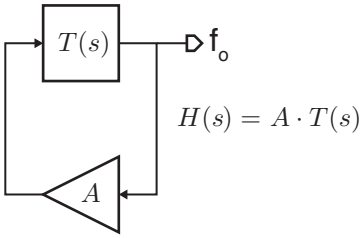


Figure 2.7: A feedback network representation of an oscillator with the frequency selective linear T(s) and the amplifier A

at the thermal sensor is to dissipate more power in the heating element, implying a low quality-factor for electrothermal filters. The implementation in [1] dissipates 2.1 mW of power to achieve a 48 ps jitter.

2.3. Oscillator Architectures

A stable oscillator needs to maintain its output frequency *and* amplitude over time. Losses in practical time constants will dampen oscillations and need to be compensated via an active circuit. In a practical oscillator, an active circuit injects energy into the system from a DC power source to ensure that the oscillation amplitude does not decay over time.

The architecture of such an active circuit may dramatically affect the oscillator's performance. Oscillators can be categorized into two groups depending on the linearity of their active element: harmonic and relaxation oscillators. The rest of this section discusses these categories of oscillators.

2.3.1. Harmonic Oscillators

Oscillators that output a spectrally pure sine wave are called harmonic oscillators. They generally consist of linear components whose dynamics can be readily analyzed with feedback theory.

An essential foundational theory for electrical oscillators is the Barkhausen stability criterion, which describes a mathematical condition for a feedback circuit to maintain a stable oscillation. A stable oscillation is defined as having a constant frequency and amplitude. Given a feedback network, as shown in Figure 2.7, where an amplifier A closes the loop around a linear network with the transfer function T(s), the criterion states that a stable oscillation will occur when:

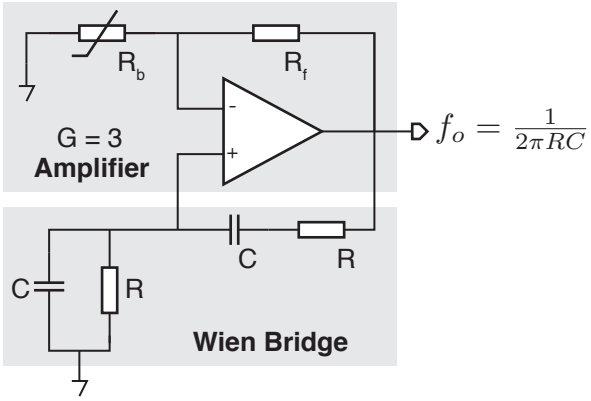


Figure 2.8: The Wien Bridge oscillator

$$\begin{aligned} \left|H(s)\right| &= 1 \\ \angle H(s) &= N \cdot 2\pi \end{aligned} \qquad N \in \mathbb{Z}$$

For stable oscillation, the phase shift around the loop must be an integer multiple of 2π. Additionally, the loop gain at the oscillation frequency must be unity. If the gain exceeds unity, the oscillation amplitude will grow, bounded by the range of the amplifier. If the gain is smaller than unity, the oscillation will decay, or the oscillator will not start. In reality, the amplifier needs to provide a gain greater than unity to ensure a fast start-up, and the gain needs to reduce to unity after the required amplitude level has been reached.

In the following, the characteristics of some common harmonic oscillators used in the literature are described.

Wien Bridge Oscillator

The Wien Bridge oscillator is a well-known harmonic oscillator circuit. Its simplified schematic is shown in Figure 2.8. The Wien Bridge is a passive network of two resistors and capacitors with a band-pass response. At its center frequency, defined by $1/2\pi$RC, the Wien Bridge has 0° phase shift and a gain of 1/3. The amplifier is configured to have a gain of R_f/R_b = 3 to compensate for the loss in the Wien Bridge. Although this gain is not enough to guarantee start-up, making the gain larger causes the oscillation amplitude to grow, saturating the amplifier and degrading the spectral purity of the output.

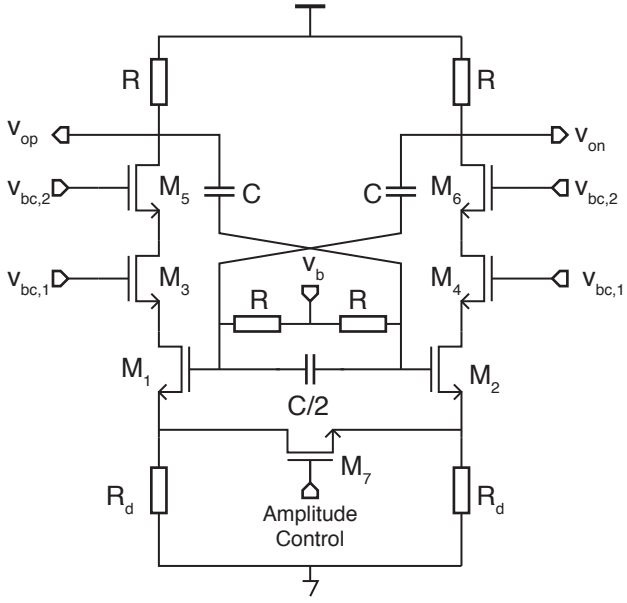


Figure 2.9: Integrated Wien Bridge oscillator implemented in [4]

In his 1939 patent, the problem of stabilizing the feedback gain was solved by Bill Hewlett (of Hewlett-Packard fame) using a positive TC thermistor for R_b[22]. Before the oscillator starts, R_f/R_b is greater than 3. As the oscillation builds up, the current flowing through R_b increases its resistance and reduces the amplifier gain. This mildly nonlinear effect stabilizes the gain, resulting in a high linearity sine wave at the output. It was this spectral purity that ensured the success of the very first product of Hewlett-Packard, the low-distortion audio oscillator, HP200A [23].

Recently, integrated frequency references based on Wien Bridge oscillators have been presented [4, 5]. Figure 2.9 shows the circuit diagram of the 6 MHz Wien Bridge oscillator in [4]. Its fully differential Wien Bridges incorporates complementary TC polysilicon resistors combined to lower their temperature coefficient and MIM capacitors for high accuracy. Degenerating $M_{1,2}$ by R_d provides the gain $R/R_d = 3$ needed to achieve unity loop gain. A separate amplitude control loop (not shown) detects the output amplitude and tunes the degeneration resistors via M_7 to regulate the gain. This design's primary challenge is to achieve an high output-impedance amplifier as it alters the Wien Bridge transfer function and causes a shift in the output frequency. Moreover, the output impedance is a function of MOS device pa-

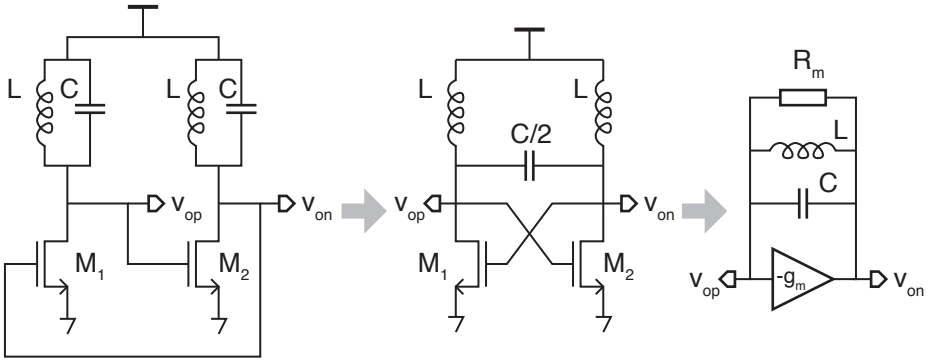


Figure 2.10: LC harmonic oscillator schematic in two-stage, cross-coupled pair, and negative resistance forms

rameters, which are less accurate than the Wien Bridge elements. In [4], the output impedance is increased by the double gain-boosted cascode transistors $M_{3\text{-}6}$. This design achieves a ±26400 ppm absolute inaccuracy for seven measured samples (3σ), with a temperature dependency of 86 ppm/°C.

LC Oscillators

LC circuits derive their time constant from the energy exchange between two storage elements at their resonance frequency. The resonant elements with a high quality-factor result in LC networks with excellent frequency selectivity that can be used to realize harmonic oscillators.

Figure 2.10 shows a typical integrated LC oscillator circuit in its three common forms: two cascaded stages, a cross-coupled pair, and simplified negative resistance [16]. The operation of the cross-coupled pair in a harmonic oscillator is easier to understand when seen as a derivative of the two-stage oscillator. In each stage, an inverting common-source transistor drives an LC load that exhibits high impedance and 0° phase shift at its resonance frequency. Due to the two cascaded inverting stages, the loop gain has a 360° phase shift. As long as the transistors have the transconductance needed to achieve larger than unity gain at the resonance frequency, oscillation will start.

The two stages must provide unity gain at the resonance frequency to maintain a stable oscillation amplitude. The loss the LC tank exhibits at its resonance frequency can be represented by the parallel resistor, R_m. The cross-coupled pair behaves as a *negative* transconductance: a positive voltage across the oscillator outputs results in a negative current to be drawn from the tank. The transistors $M_{1,2}$ can be sized or biased to provide

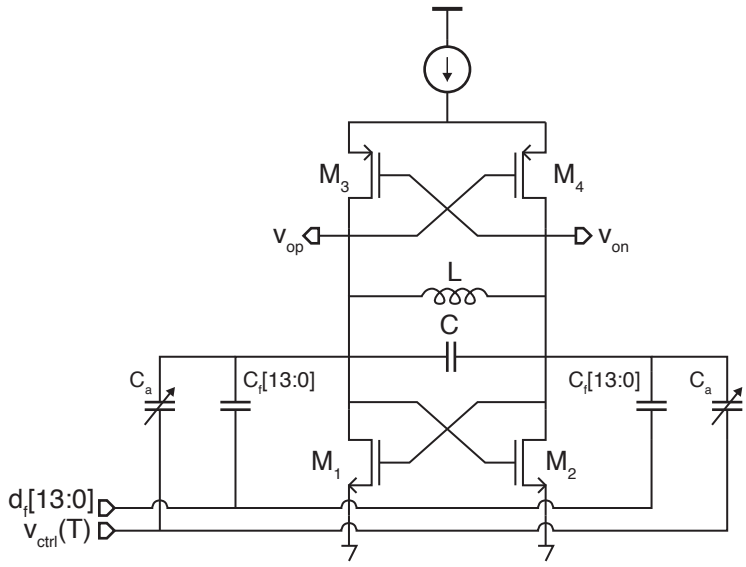


Figure 2.11: Temperature-independent LC oscillator used in [18]

a transconductance with a magnitude equal to $1/R_m$ at the oscillation frequency, which achieves unity gain and a stable oscillation frequency. The oscillator also provides some in-built amplitude control, as increasing oscillation amplitude reduces the headroom of $M_{1,2}$, pushing them out of saturation and reducing their excess gain.

Integrated frequency references based on LC oscillators were used in the early 2000s as crystal replacements [18]. Initially supporting better than ±500 ppm inaccuracy over PVT, they were targeted to replace external crystals used for wireline communication protocols such as USB 2.0 [17]. The design in [18], shown in Figure 2.11, uses complementary cross-coupled NMOS and PMOS pairs, biased with a temperature-independent constant current. The LC oscillator is designed to oscillate at 800 MHz, which is divided by 16 to provide a 25 MHz reference output. Two capacitor banks allow the tank frequency to be tuned to reduce the spread. The digitally programmable bank (C_f) consists of switched MOS capacitors and allows post-production trim of the tank frequency, covering ±10% with 20 ppm steps. The programmable analog bank (C_a) consists of accumulation-mode PMOS varactors controlled by a temperature-dependent bias voltage $v_c(T)$. This voltage is generated via a programmable resistor network that features selectable p-poly, n-poly, p-diffusion, and n-well resistors to fine-tune its TC

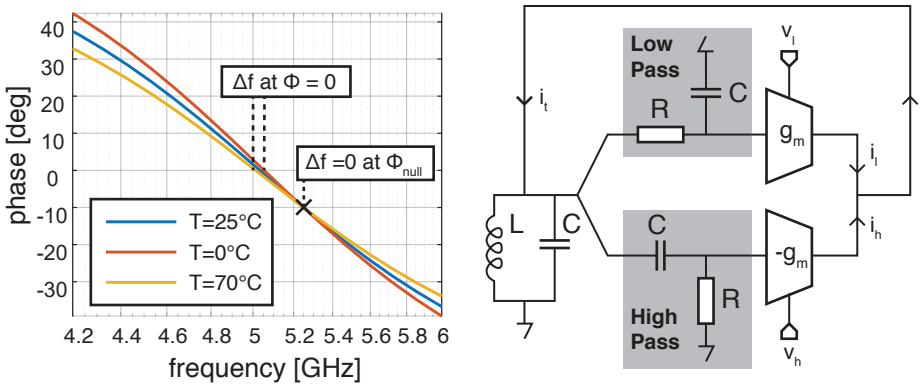


Figure 2.12: The LC oscillator in [25] based on achieving oscillation at Φ_{null}, where the TC of the tank is the lowest

to compensate for that of the oscillator itself and achieve frequency trimming in production [17]. The design achieves ±152 ppm absolute accuracy from -10 °C to 80 °C for ten compensated devices while dissipating 60 mW. The excellent inaccuracy achieved demonstrated, for the first time, that fully integrated frequency references could challenge the accuracy of crystal oscillators.

Later work [15, 24] improved upon the same concept of a temperature-compensated LC oscillator. Changing the oscillator to a Colpitts core reduces the uncompensated temperature dependence, relaxing the compensation circuitry. Additionally, the low jitter of the LC oscillator enables post-production trimming to better than 2 ppm [15]; however, this was done at 16 temperature points, making it too expensive for most applications. A recent example of a temperature-compensated Colpitts oscillator[24] achieves a 1 ppm/°C residual temperature coefficient from -50 °C to 170 °C for 48 samples with a single-point room temperature trim while dissipating 4.25 mW for a 1.38 GHz output frequency.

Another LC oscillator architecture that targets a low TC is presented in [25]. The LC oscillator loop gain has a 360° phase-shift due to the two inverting stages and thus forces the tank to work at its 0°-phase point. As the impedance of an LC tank changes with the temperature (due to parasitic resistors), a shift of the 0° phase occurs, causing the output frequency shift Δf. However, a specific phase exists, Φ_{null}, for which the different phase versus frequency curves cross, where Δf is zero. Thus, if some excess phase

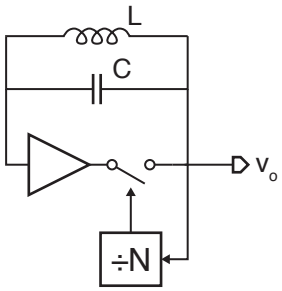


Figure 2.13: The low-power LC oscillator in [26] based on pulsed energy injection

shift in the loop forces the tank to exhibit Φ_{null} phase shift during oscillation, the temperature sensitivity of the tank will be zero. In [25], an all-pass filter adds this excess phase shift to the oscillator loop. The all-pass filter is constructed from parallel low-pass and high-pass sections that drive the transconductances whose strength can be adjusted via control voltages, $v_{l,h}$. The excess phase shift imparted in the oscillator loop is programmable, and the oscillator can thus be fine-tuned to force the tank to Φ_{null}. The design [25] achieves ±100 ppm absolute accuracy from -40 °C to 80 °C for ten devices while dissipating 24 mW.

A significant disadvantage of LC oscillators is their relatively high power consumption. This is because the values of practical on-chip inductors limit the achievable resonance frequencies to the GHz regime. Power-hungry amplifiers are then required to sustain such high-frequency oscillation. While not a fully integrated solution, since it uses a bond wire inductor, the work in [26] implements a pulsed-drive concept to reduce the amplifier's power consumption. As shown in Figure 2.13, the design relies on injecting energy into the tank once every N cycle and turning off the amplifier otherwise. This dramatically decreases power consumption at the cost of nonidealities due to the excess delay of the divider and injection circuitry. Without temperature compensation, the design achieved 0.76% absolute accuracy with a 92 ppm/°C TC while dissipating only 46 µW. This dramatic reduction in power consumption comes at the cost of reduced noise performance and larger temperature sensitivity compared to a standard LC oscillator.

Ring Oscillators

Ring oscillators are a class of circuits constructed by connecting several inverting gain stages in feedback. Their use is prevalent in phase-locked loop and clock and data recovery circuits due to their high oscillation frequencies,

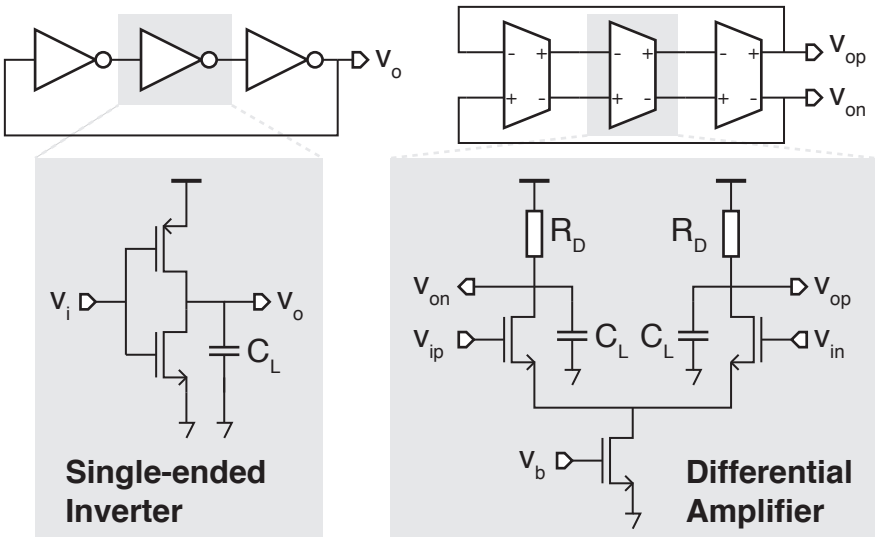


Figure 2.14: Two of the most common ring oscillator topologies: single-ended CMOS inverter-based and differential-amplifier-based

extensive tuning range, and ease of integration [16].

The most common ring oscillator topologies are shown in Figure 2.14. The single-ended version most commonly uses CMOS inverters. The oscillation frequency for a single-ended ring oscillator is:

$$f_o = \frac{1}{2Nt_d} \qquad \text{where} \qquad t_d \approx \frac{C_L V_{DD}}{\mu_n C_{ox} \frac{W}{L}(V_{DD} - V_t)^2} \sim \frac{C_L}{g_m} \tag{2.9}$$

where N is the number of delay stages used, and t_d is the propagation delay of the CMOS inverter. Each edge needs to propagate twice through the chain to complete one period of oscillation. The time constant of a ring oscillator is defined by the load capacitance, C_L, and the transconductance, g_m, of the inverter stage. Both strongly depend on process (μ_n, C_{ox}, V_t) and geometry (W, L) parameters. Designing these for accuracy is difficult. Moreover, the strong dependence of the delay on supply voltage usually necessitates tight supply regulation.

The differential ring oscillator addresses the supply sensitivity problem using fully symmetric amplifiers. Figure 2.14 shows the differential stage with an RC load driven by the differential pair biased with a constant current. The oscillation frequency for the differential ring can be evaluated from the

expression of the loop gain as:

$$
\begin{aligned}
H(s) &= \left(-g_m \frac{R}{RCs+1}\right)^N \\
\angle H(jw_0) &= \tan^{-1}(RC\omega_0) = \frac{180}{N} \\
f_0 &= \frac{1}{2\pi RC} tan\left(\frac{180}{N}\right)
\end{aligned}
\tag{2.10}
$$

To first order, RC-loaded ring oscillators will be more accurate than their single-ended counterparts as the dependence on the transconductance is absent. However, this comes at the cost of noise efficiency. A fully differential ring oscillator will have a lower phase noise figure-of-merit, primarily due to the reduced oscillation amplitude due to the headroom requirements of its current source [16].

Ring oscillators occupy the gray area between harmonic and relaxation oscillators. A ring oscillator with a few stages will behave more like a harmonic oscillator. It will also oscillate faster as its stages will provide less gain and operate more around their linear amplification region. As the number of stages increases, the edges get sharper relative to the oscillation period, and each stage starts behaving more like a comparator, keeping its output in the memory afforded by its output capacitance until it receives a new edge. This behavior is more akin to that of a relaxation oscillator. This duality usually causes ring oscillators to be wrongly classified as relaxation oscillators.

Unfortunately, ring oscillators make poor frequency references due to their low accuracy. The design in [27] uses a three-stage NMOS differential ring oscillator with a controllable PMOS load as its frequency source. The uncompensated measured accuracy for the ring is reported to be ±11.8%.

However, ring oscillators excel in applications that require an inaccurate but controllable oscillator. It is possible to tune a ring oscillator using several different techniques, as shown in Figure 2.15. The work in [28] degenerates every stage with a controlled current source to vary the ring frequency. Similarly, the design in [29] stacks a controlled NMOS device to reduce the short-circuit current, simultaneously achieving controllability and power optimization. The work in this thesis presented in chapters 4 and 5 utilizes ring oscillators with supply current control driven by current DACs. Another common technique is to regulate the ring oscillator supply voltage to control its frequency [30, 31]. The design in [31] regulates the supply voltage of the pseudo-differential stages of its ring oscillator to control its frequency.

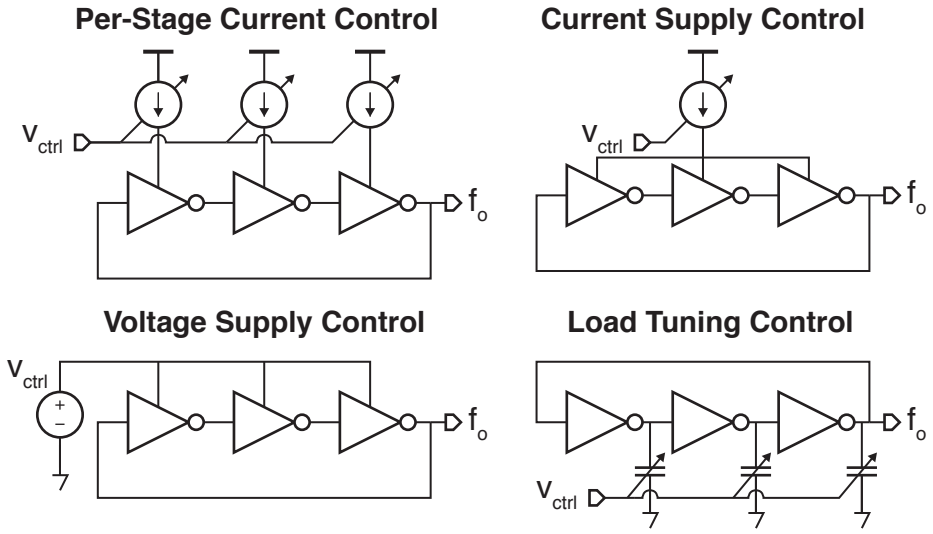


Figure 2.15: Methods for controlling the output frequency of ring oscillators: per-stage current control used in [28, 29], current supply control used in chapters 4 and 5, voltage supply control used in [30, 31], and load tuning control

2.3.2. Relaxation Oscillators

Relaxation oscillators constitute the second class of oscillators. They are differentiated by a *strongly* nonlinear element and usually feature a single energy tank. Nonlinear oscillators were studied by the Dutch physicist Balthasar van der Pol, who described their behavior by his namesake second-order differential equation:

$$\frac{d^2x}{dt^2} - \mu(1 - x^2)\frac{dx}{dt} + x = 0 \tag{2.11}$$

In 2.11, the van der Pol equation, the parameter μ controls the nonlinearity of the oscillatory system. For $\mu = 0$, the equation contains no nonlinearity, and its behavior is of a simple harmonic oscillator. As μ increases, the effect of nonlinearity increases, and the oscillations start to switch between two levels that look more discrete. The term relaxation oscillator describes the *tension* building up before the system *relaxes* as it switches between these discrete states.

A typical implementation for a relaxation oscillator is shown in Figure 2.16 [32]. In its initial state (Q = 0, !Q = 1), the switch M1 is open, and a current, i_{ref} (derived from a voltage reference v_{ref} and resistor R), charges the capacitor C_1. As soon as the voltage across C_1 exceeds v_{ref}, the comparator

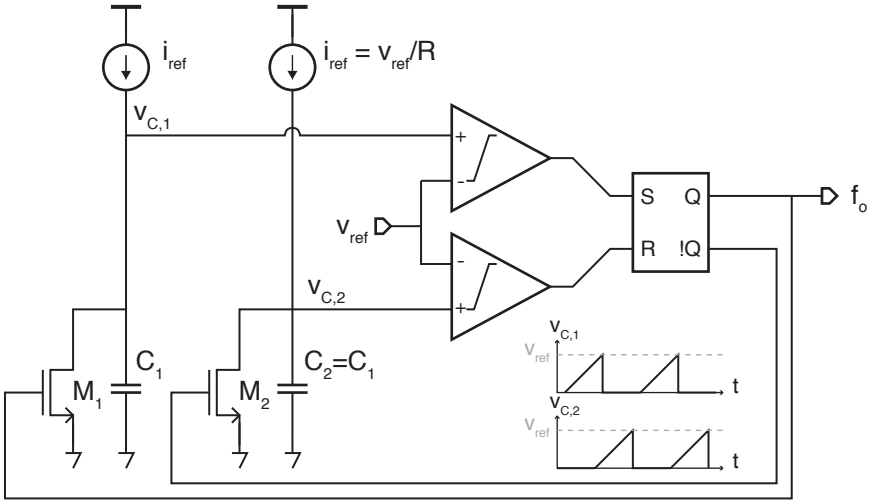


Figure 2.16: The schematic diagram for a typical relaxation oscillator implementation

drives the output of the SR latch to Q = 1. In this state, the switch M_1 is closed, and M_2 is opened, resetting the voltage across C_1 and starting the charging cycle for C_2. The voltage $v_{c,2}$ then rises until it exceeds v_{ref}, after which the circuit reverts to its original state. The circuit will oscillate between the two states with a period $t_o = 2RC$ if all other components are assumed to be ideal.

The primary error sources in this circuit are due to the nonlinear elements. The offset of the comparators directly affects the oscillation frequency, as errors in the voltage comparison behave the same way as errors in the time constant. Offset errors can be mitigated by using dynamic techniques such as chopping [33–35]. An effect that cannot be easily mitigated is the delay of the comparators and the latch. This increases the time before the state changes and directly affects the oscillation frequency. The real problem is that this delay depends on both process and temperature.A brute-force solution is to reduce the absolute delay of the comparators by increasing their power consumption. Alternatively, the delay variation can be compensated by using replica circuitry [35] or FLLs [36].

2.3.3. Frequency Locked Loops

The active elements required to sustain oscillation (either linear amplifiers or nonlinear comparators) may corrupt the time constant in oscillators and may dominate frequency inaccuracy. Therefore, architectures that can de-

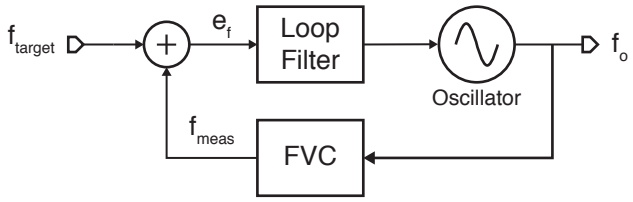


Figure 2.17: Generalized block diagram a frequency-locked loop, a closed-loop oscillator system

couple the time constant from the mechanism to sustain oscillation promise increased accuracy. One approach is to use a closed-loop architecture, in which the frequency error of an oscillator oscillator is measured and then compensated by a feedback loop. Such architectures are commonly referred to as frequency-locked loops (FLLs).

A generalized block diagram for an FLL is given in Figure 2.17. The block in the feedback path measures the oscillator's output frequency, f_{meas}, and outputs it in a domain that allows further processing. Generally, this will be the voltage domain, so this block will usually be a frequency-to-voltage converter (FVC). The measured frequency is then subtracted from the target frequency, f_{target}, yielding the frequency error signal, e_f, which drives a loop filter. The loop filter output closes the feedback loop by controlling the oscillator's output. The gain of the loop filter drives the error signal to zero, effectively locking f_{meas} to f_{target}.

Compared to conventional oscillators, the primary benefit of the FLLs is the relaxation in the oscillator's specifications. If the oscillator frequency varies due to external effects, the loop will maintain it at f_{target}. The only requirement is that the oscillator can be driven to f_{target} under these conditions. However, any measurement error in f_{meas} will be directly observed at the oscillator's output.

Frequency-to-Voltage Converters

The accuracy of a FLL is mainly determined by the gain, inherent accuracy and the reference of its frequency-to-voltage converter (FVC). This seperates the tasks of maintaining oscillation and achieving accuracy. Precision circuit techniques generally incompatible with oscillators can be employed in the FVC, and more sophisticated temperature compensation schemes can be devised. The rest of this section discusses a few of the most common

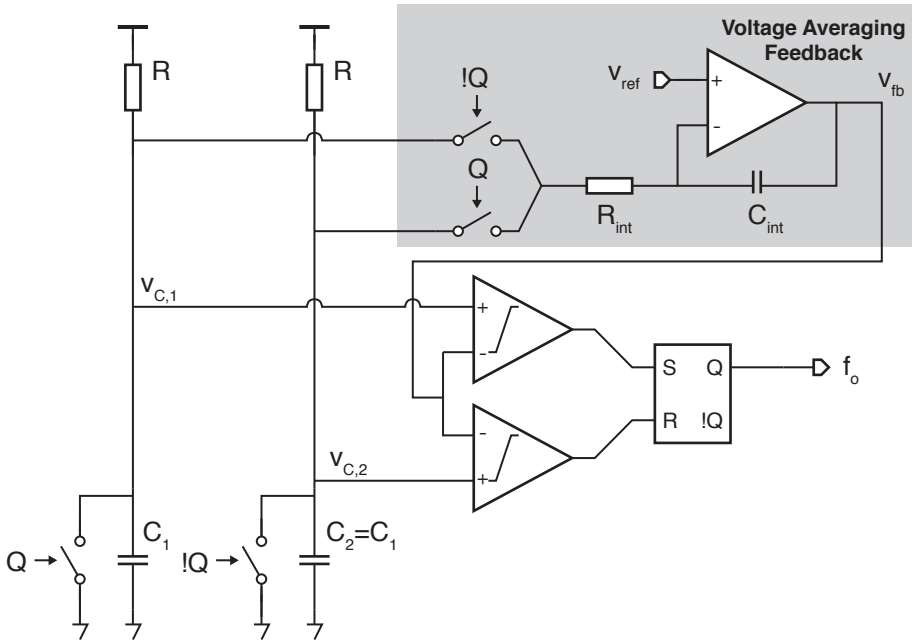


Figure 2.18: The schematic diagram for the voltage averaging feedback design in [36]

techniques used to build FVCs.

Voltage-Averaging Feedback An early example of an FLL (although not referred to as one in its publication) is presented in [36]. It implements a FVC by utilizing the frequency-to-voltage conversion achieved inside a relaxation oscillator. In the state Q! of C_1 is charged until its voltage equals v_{fb}, at which point its respective comparator flips and the circuit switches state to Q! to charge C_2. The voltage averaging feedback circuit features an analog integrator that integrates the final voltages on the capacitors, $v_{c,1}$ and $v_{c,2}$ in !Q and Q cycles, respectively. Comparing this voltage to the reference v_{ref} generates the frequency error signal, which is integrated to yield the trip voltage, v_{fb}. This voltage is used to tune the oscillator to the correct frequency. Once the loop settles, the average value of v_c is locked to v_{ref}, meaning the circuit is frequency-locked.

In this circuit, the time constant is achieved by $C_{1,2}$ and R. R_{int} and C_{int} do not affect the oscillation frequency but set the integrator's low-frequency pole. The delay of the comparators do not affect the frequency of oscillation, and their offset is removed via chopping. The design achieves a 23 ppm/°C

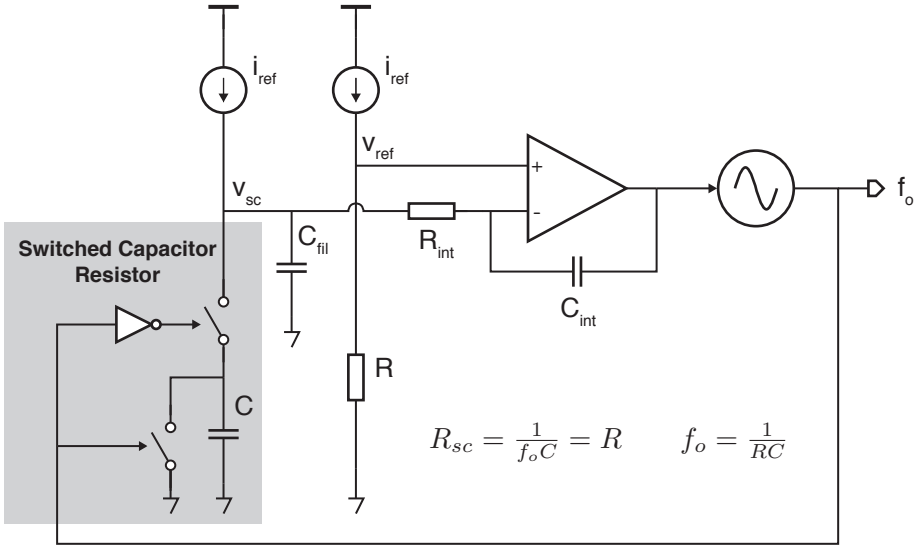


Figure 2.19: The schematic diagram for the switched-capacitor-resistor-based frequency reference in [29]

residual temperature coefficient from -40 °C to 125 °C. This state-of-the-art result inspired further interest in FLL architectures.

Switched-Capacitor Resistors Another common implementation of FVCs uses switched capacitor resistors (SCRs). A SCR behaves like an equivalent resistor with the value R_{sc}. When a reference current is passed through it, the average voltage across the switched-capacitor resistor is a function of its switching frequency.

The design in [29], Figure 2.19 uses the voltage difference generated by a SCR and a resistor driven by identical reference currents to build an FVC. An integrator formed by R_{int} and C_{int} integrates the error voltage between v_{sc} and v_{ref} and drives the controlled oscillator, closing the feedback loop. At steady-state v_{sc} will be equal to v_{ref}, meaning that the output frequency has been locked to $(RC)^{-1}$. The design achieves a 27.5 ppm/°C residual temperature coefficient from -40 °C to 80 °C and dissipates 99 µW for a 70 kHz output frequency.

The designs in [7, 31] utilize the same SCR but instead use current-domain processing. As shown in Figure 2.20, an error current generated from the difference between the reference resistor and the switched-capacitor resistor is integrated into a capacitor. Then, a ring oscillator is tuned by con-

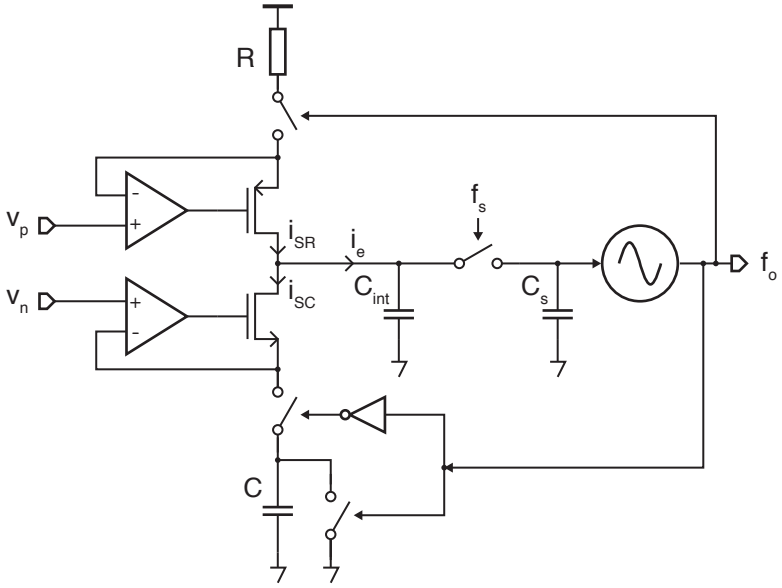


Figure 2.20: The schematic diagram for the switched capacitor resistor based frequency reference with a current-domain integrator in [31]

trolling its bias current by this integrated voltage. The design in [7] achieves a 90 ppm/°C residual temperature coefficient from -20 °C to 100 °C and has 2.7% absolute accuracy for 20 samples while dissipating 60 µW. The design in [31] improves upon the idea by eliminating the most power-hungry components of the circuit. The resistance is amplified by switching it with a low duty-cycle, and a very efficient switched-capacitor down converter generates the reference voltages for its current buffers. The feedback voltage is also sampled to reduce transient effects on the oscillator. It achieves a 13.8 ppm/°C residual temperature coefficient from -25 °C to 85 °C for one reported sample while dissipating 4.7 nW for a 3 kHz output frequency.

The designs discussed are limited in accuracy primarily due to the non-idealities of their integrating stages. The amplifier offset, shown in Figure 2.19, results in a frequency error proportional to the frequency-to-voltage converter sensitivity. The design in [29] addresses the problem using ping-pong auto-zeroed amplifiers, effectively eliminating the offset. The auto-zeroing scheme reduces the amplifier offset variation by a factor of 55. With such techniques, the inaccuracy originating from FVC elements can be negligible, resulting in about 10 ppm/°C residual temperature coefficient.

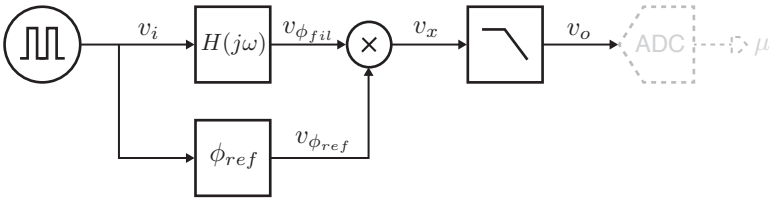


Figure 2.21: Block diagram of a synchronous demodulator for phase detection of the linear filter $H(j\omega)$

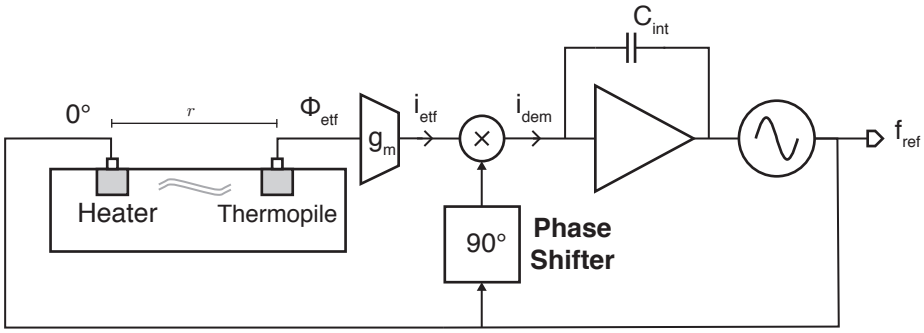


Figure 2.22: The schematic diagram for FLL in [21] based on a phase-domain FVC

Phase-domain FVCs A very early implementation of FLLs used phase-domain processing to determine the time constant of integrated electrothermal filters. A synchronous demodulator, shown in Figure 2.21, uses a reference frequency with a phase shift, ϕ_{ref}, to demodulate the output of a filter driven at the same frequency. The DC output voltage, v_o, represents the filter's relative phase shift, $\phi_{fil} - \phi ref$. This phase shift is a frequency-dependent signal and can be used in FLLs to control the frequency of oscillators. Also, since the measurement is conducted away from low frequencies, any corruption due to flicker noise or offset of the amplifiers can be avoided.

In the design of [21], shown in Figure 2.22, a reference oscillator (f_{ref}) drives an ETF at 0° phase. The ETF imparts a phase shift (Φ_{etf}) to the drive

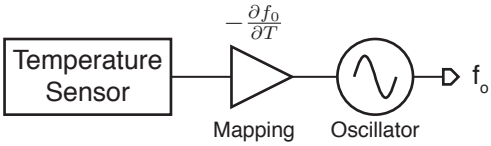


Figure 2.23: Generalized block diagram for the temperature compensation of an oscillator

signal, depending on the distance between its heater and thermopile. ETF's output voltage is converted into a current, i_{etf}, which is then demodulated by the 90° phase shifted f_{ref}. The demodulator output is a current proportional to the phase shift of the filter at the drive frequency. This current is used in a loop filter to control the frequency of the oscillator, locking the output frequency to the zero-phase shift frequency of the ETF.

The design of [21] functions as a temperature sensor due to the high (but repeatable) temperature sensitivity of the ETF. Converting such a system into a frequency reference requires addressing the ETF's temperature dependence via compensation, as discussed in the next section.

2.4. Temperature Compensation

Frequency references need to maintain a stable output frequency across all operating conditions. Temperature variations pose the first significant challenge to designing integrated frequency references. As discussed in Sections 2.2 and 2.3, integrated time constants exhibit inherent temperature dependencies exacerbated by the nonidealities of the active circuitry used to build oscillators. This temperature dependence can significantly contribute to the frequency reference's inaccuracy.

Methods to reduce this temperature dependence are known as temperature compensation. Temperature compensation is generally achieved by sensing the temperature via an auxiliary circuit and tuning the oscillator in the opposite direction of its sensitivity, as in Figure 2.23. An integrated temperature sensor senses the die temperature, which is then mapped to a frequency control signal. The mapping function replicates the *inverse* temperature sensitivity of the uncompensated oscillator, and the resulting frequency is not sensitive to temperature. Achieving this can be as simple as using circuit elements with opposite temperature coefficients together or as complex as using a digital signal processing circuit to map the output of a digital temperature sensor to tune the oscillator.

The simplest way to achieve temperature compensation for RC frequency

references is to use so-called zero-TC resistors. Suppose two types of resistors with opposite polarity temperature coefficients are available in the process. It is possible to build a composite resistor using a series combination of these temperature-dependent resistors to achieve a zero-TC resistor. If the two resistors combined are R_p with a positive TC_p and R_n with a negative TC_n, resulting in:

$$R_{ztc} = R_p + R_n = R_{p,0}(1 + TC_p \cdot T) + R_{n,0}(1 + TC_n \cdot T) \quad (2.12)$$

$$= (R_{p,0} + R_{n,0})(1 + TC_{ztc} \cdot T) \quad (2.13)$$

where the zero-TC resistor R_{ztc} has a temperature coefficient given by:

$$TC_{ztc} = \frac{R_{p,0}}{R_{p,0} + R_{n,0}} TC_p + \frac{R_{n,0}}{R_{p,0} + R_{n,0}} TC_n \quad (2.14)$$

Choosing $R_{p,0}/R_{n,0} = -TC_n/TC_p$ results in $TC_{ztc} = 0$. However, the accuracy of this method depends on the matching of two different types of resistors. If the resistors' values spread due to geometry effects, the error in the temperature coefficient will be:

$$\Delta TC_{ztc} = \frac{\partial TC_{ztc}}{\partial R_{p,0}} \Delta R_{p,0} + \frac{\partial TC_{ztc}}{\partial R_{n,0}} \Delta R_{n,0} \quad (2.15)$$

$$= \frac{R_{n,0}(TC_p - TC_n)}{(R_{p,0} + R_{n,0})^2} \Delta R_{p,0} + \frac{R_{p,0}(TC_p - TC_n)}{(R_{p,0} + R_{n,0})^2} \Delta R_{n,0} \quad (2.16)$$

Assuming two resistors with TC_n = 3000 ppm/°C and TC_n = -1500 ppm/°C with 15% worst-case spread (ΔR/R) for both results in a residual TC of 300 ppm/° C. Because of the TC_p-TC_n factor in 2.16, it is beneficial to use resistors with low TCs.

The design in [37] uses a ZTC resistor constructed from complementary TC polysilicon resistors in its relaxation oscillator core to achieve 64.3 ppm/°C residual temperature sensitivity. In [28], a ZTC resistor is formed by a single polysilicon resistor by tuning its sheet and end resistances to achieve 126 ppm/°C residual TC. The design in [29] uses a zero-TC resistor as part of its temperature compensation scheme, achieving 27.4 ppm/°C.

Temperature compensation can also take a form more akin to the generic structure of Figure 2.23, where a correction is applied to tune the oscillator. In conventional oscillators, the correction is directly applied to a tunable element and pulls the frequency to stabilize it over temperature. A

temperature-dependent bias generator may be used to derive a compensation signal. In the design of [32], a relaxation oscillator (similar to the schematic diagram in Figure 2.16) uses currents derived from a bandgap regulator, achieving 300 ppm/°C residual TC. In [27], a ring oscillator is loaded with a controllable PMOS device that sets the time constant for the delay cell. The control voltage is derived from temperature and process compensation circuitry using a bandgap regulator. The uncompensated measured accuracy for the ring is reported to be ±11.8% and was improved to ±2.64% after temperature compensation. The LC frequency reference design in [17] uses a current with programmable TC to control varactors that enable tuning the resonant frequency of the LC tank. The programmable TC is achieved by deriving current from a reference voltage and a resistor bank formed by four different types of resistors.

Unfortunately, the methods described above hit yet another barrier in accuracy: they cannot compensate for higher-order temperature coefficients. This limitation emerges from their analog nature, which makes generating arbitrary and programmable nonlinearities practically impossible. The following section describes how digital processing achieves temperature compensation that can further reduce the inaccuracy of frequency references.

2.4.1. Digital Nonlinear Temperature Compensation

Reducing the residual temperature coefficient of RC frequency references below the ±2000 ppm levels requires addressing the *nonlinear* temperature dependence of their resistors. Implementing the necessary nonlinear functions requires post-processing of temperature in the digital domain, where complex polynomial functions can be realized to mimic and compensate for the nonlinear temperature dependence of integrated time constants.

In its simplest form, digital compensation can be applied directly to tune an oscillator over temperature, as in Figure 2.24. The frequency reference design in [15] compensates for the LC oscillator's temperature dependence in the digital domain. The design achieves higher-order compensation using a look-up table to implement a nonlinear function. The output of a bandgap-based digital temperature sensor is mapped to a word that is converted back to voltage via a DAC and controls the oscillator frequency by tuning its varactors. The design in [15] achieved ±2.5 ppm inaccuracy from 0 °C to 80 °C for three devices while dissipating 14.25 mW. However, in the higher-order temperature compensation method, the designers used 17 temperature measurements to derive their look-up table, a number far beyond economic feasibility for production.

Correcting the frequency measurement provided by the frequency-to-

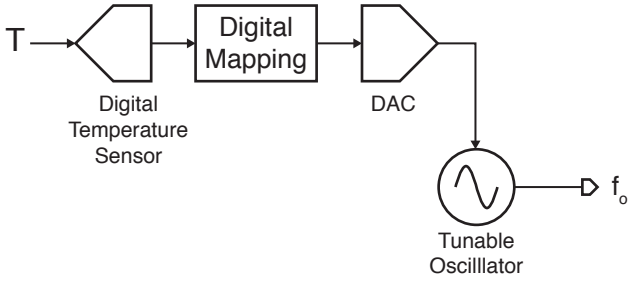


Figure 2.24: Block diagram for the LC frequency reference in [15] using digital nonlinear temperature compensation

voltage converter in an FLL is generally more feasible. In the design in Figure 2.25 [2], the challenge of digitizing an on-chip time constant is resolved by phase-domain delta-sigma modulators (PDDSMs). The PDDSM digitizes the frequency-dependent phase shift of an electrothermal filter (ETF). The strongly nonlinear temperature dependence of the phase shift is addressed via a digital nonlinear temperature compensation scheme. The output of a digital bandgap-based temperature sensor is mapped to the nonlinear temperature sensitivity of the phase shift using polynomial processing. The design achieved a state-of-the-art 11 ppm/°C residual temperature coefficient from -55 °C to 125 °C for 24 measured samples, showing less than 0.15% absolute inaccuracy. However, due to the ETF heaters' power-hungry nature, the design dissipates 2.1 mW.

Frequency references can also be built by digitizing the phase shift of an integrated RC filter. In such a frequency reference, the nonlinear temperature coefficient of resistors could be addressed in the digital domain to build accurate frequency references without the power penalty of ETFs. The motivation for the work in this thesis is to extend the architecture and methods used in [2] to build RC frequency references. The details on using phase-to-digital converters to build frequency-locked loops will be discussed in chapter 3. Chapters 4 and 5 will describe the implementation details for two frequency references realizing such digital nonlinear temperature compensation.

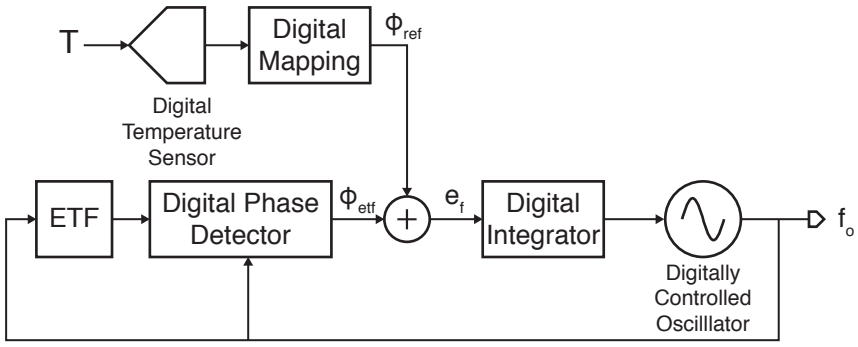


Figure 2.25: Block diagram for the ETF-based frequency reference in [2] using digital non-linear temperature compensation

2.5. Integrated Frequency Reference Survey

Given the many possible ways of realizing an integrated frequency reference, a survey of the state-of-the-art was carried out at the conception of this work to determine a promising approach for the design of a high-accuracy frequency reference.

Figure 2.26 shows the residual temperature coefficient and frequency-normalized power consumption of integrated frequency references published between 2000 and 2018. The work is grouped according to the type of time constant they use. LC-based frequency references exhibit the smallest residual inaccuracy but consume very high power. As expected, they achieve better accuracy than references that use doped devices such as resistors. However, the limited size of on-chip inductors makes their oscillation frequency much higher than that of RC references, implying high power consumption in the following circuitry. To address this drawback, the LC reference in [38] is unique and uses a pulsed oscillator architecture.

RC-based frequency references consume low power but have limited accuracy. By using large on-chip resistors, lower oscillation frequencies can be achieved, resulting in low power consumption. Also, low-frequency operation generally relaxes oscillator requirements, albeit at the cost of increased jitter. Unfortunately, RC frequency references are limited to 10 ppm/°C levels of residual temperature coefficient, almost an order of magnitude higher than LC references. While the most accurate RC references compensate for the first-order temperature coefficient of their resistors, they typically do not compensate for the higher-order temperature coefficients, which then limits

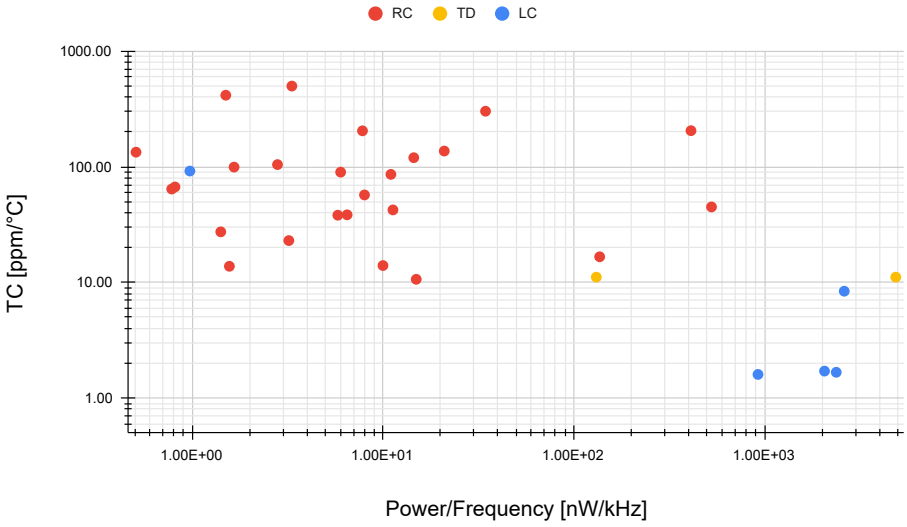


Figure 2.26: Temperature coefficient versus power consumption of integrated frequency references, grouped by choice of time constant

their ultimate accuracy.

Figure 2.27 shows the temperature coefficient vs frequency-normalized power consumption of RC references grouped by oscillator architecture. It can be seen that the use of FLL architectures consistently results in better accuracy. This can be attributed to the fact that in an FLL, the circuitry for maintaining oscillation is separated from the circuitry for maintaining accuracy.

2.6. Conclusion

This chapter presented a study of integrated frequency reference designs that constitute the state-of-the-art. The steps in building a frequency reference were described: choosing a time constant, building an oscillator, and implementing temperature compensation. Available choices and techniques for each of these steps were discussed and compared. The analysis of the landscape culminated in a survey of published integrated frequency references in the current century.

From the survey, a few observations can be made. LC references are already accurate, but they pose the challenge of reducing their power con-

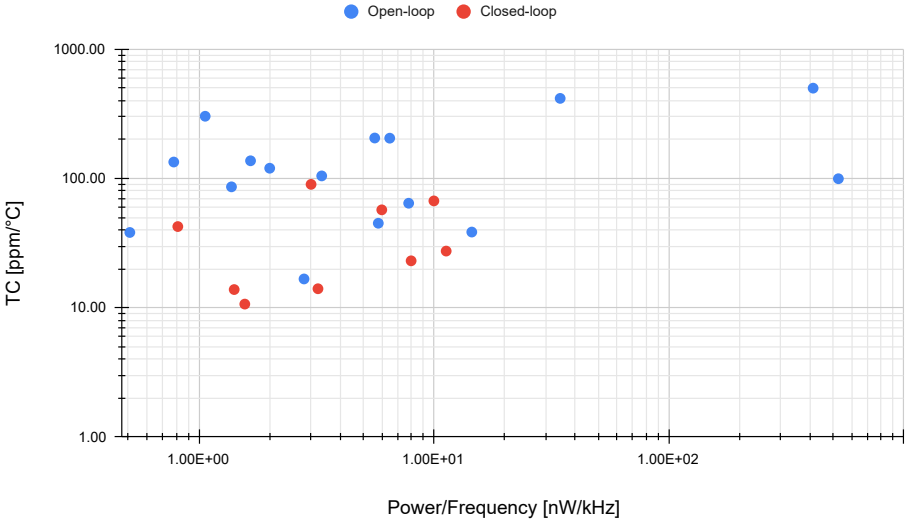


Figure 2.27: Temperature coefficient versus power consumption of integrated RC frequency references, grouped by choice of architecture

sumption. On the other hand, RC references can be built with several orders of magnitude lower power but are limited in their accuracy. FLL architectures can achieve residual TCs close to that of LC oscillators. However, RC references are limited by the inaccuracy of their time constant, specifically, the nonlinear temperature dependence of their resistors.

The nonlinear temperature dependence of resistors is not easy to compensate with traditional analog compensation methods. However, one benefit of working in a CMOS process is the availability of ready *digital* signal processing. The analysis and study in this chapter direct the work in this thesis to explore an FLL-based RC reference that will use digital temperature compensation to address the nonlinear temperature dependence of integrated resistors.

References

[1] S. M. Kashmiri, M. A. P. Pertijs, and K. A. A. Makinwa, *A Thermal-Diffusivity-Based Frequency Reference in Standard CMOS With an Absolute Inaccuracy of ±0.1% From -55 °C to 125 °C,* IEEE Journal of Solid-State Circuits **45**, 2510 (2010).

[2] S. M. Kashmiri, K. Souri, and K. A. A. Makinwa, *A Scaled Thermal-Diffusivity-Based 16 MHz Frequency Reference in 0.16 µm CMOS,* IEEE Journal of Solid-State Circuits **47**, 1535 (2012).

[3] M. Raman, T. Kifle, E. Bhattacharya, and K. Bhat, *Physical Model for the Resistivity and Temperature Coefficient of Resistivity in Heavily Doped Polysilicon,* IEEE Transactions on Electron Devices **53**, 1885 (2006).

[4] V. De Smedt, P. De Wit, W. Vereecken, and M. S. J. Steyaert, *A 66 µW 86 ppm/°C Fully-Integrated 6 MHz Wienbridge Oscillator With a 172 dB Phase Noise FOM,* IEEE Journal of Solid-State Circuits **44**, 1990 (2009).

[5] V. De Smedt, W. Dehaene, and G. Gielen, *A 0.4-1.4V 24MHz fully integrated 33µW, 104ppm/V supply-independent oscillator for RFIDs,* in *2009 Proceedings of ESSCIRC* (2009) pp. 396–399.

[6] J. Lee and S. Cho, *A 10MHz 80µW 67 ppm/°C CMOS reference clock oscillator with a temperature compensated feedback loop in 0.18µm CMOS,* in *2009 Symposium on VLSI Circuits* (2009) pp. 226–227.

[7] K. Ueno, T. Asai, and Y. Amemiya, *A 30-MHz, 90-ppm/°C fully-integrated clock reference generator with frequency-locked loop,* in *2009 Proceedings of ESSCIRC* (2009) pp. 392–395.

[8] A. A. Naem and L. Y. Chee, *Variation of cobalt silicide resistivity with temperature,* Journal of Applied Physics **79**, 9149 (1996), https://doi.org/10.1063/1.362586 .

[9] S. Pan, Y. Luo, S. Heidary Shalmany, and K. A. A. Makinwa, *A Resistor-Based Temperature Sensor With a 0.13 pJ · K2 Resolution FoM,* IEEE Journal of Solid-State Circuits **53**, 164 (2018).

[10] B. El-Kareh and L. N. Hutter, *Silicon Analog Components: Device design, Process Integration, characterization, and reliability* (Springer, 2020).

[11] D. Nachrodt, U. Paschen, A. ten Have, and H. Vogt, *Ti/Ni(80a Nearly Zero Temperature Coefficient of Resistance for Integration in a Standard CMOS Process,* IEEE Electron Device Letters **29**, 212 (2008).

[12] G. Zhang, K. Yayama, A. Katsushima, and T. Miki, *A 3.2 ppm/°C Second-Order Temperature Compensated CMOS On-Chip Oscillator Using Voltage Ratio Adjusting Technique,* IEEE Journal of Solid-State Circuits **53**, 1184 (2018).

[13] A. L. S. Loke, D. Yang, T. T. Wee, J. L. Holland, P. Isakanian, K. Rim, S. Yang, J. S. Schneider, G. Nallapati, S. Dundigal, H. Lakdawala, B. Amelifard, C. Lee, B. McGovern, P. S. Holdaway, X. Kong, and B. M. Leary, *Analog/Mixed-Signal Design Challenges in 7-nm CMOS and Beyond,* in *2019 IEEE Custom Integrated Circuits Conference (CICC)* (2019) pp. 1–8.

[14] R. Aparicio and A. Hajimiri, *Capacity limits and matching properties of integrated capacitors,* IEEE Journal of Solid-State Circuits **37**, 384 (2002).

[15] E. O. Ates, A. Ergul, and D. Y. Aksin, *Fully Integrated Frequency Reference With 1.7 ppm Temperature Accuracy Within 0–80°C,* IEEE Journal of Solid-State Circuits **48**, 2850 (2013).

[16] B. Razavi, *Design of CMOS phase-locked loops: From circuit level to architecture level* (Cambridge University Press, 2020).

[17] M. S. McCorquodale, J. D. O'Day, S. M. Pernia, G. A. Carichner, S. Kubba, and R. B. Brown, *A Monolithic and Self-Referenced RF LC Clock Generator Compliant With USB 2.0,* IEEE Journal of Solid-State Circuits **42**, 385 (2007).

[18] M. S. McCorquodale, G. A. Carichner, J. D. O'Day, S. M. Pernia, S. Kubba, E. D. Marsman, J. J. Kuhn, and R. B. Brown, *A 25-MHz Self-Referenced Solid-State Frequency Source Suitable for XO-Replacement,* IEEE Transactions on Circuits and Systems I: Regular Papers **56**, 943 (2009).

[19] G. Bosch, *A thermal oscillator using the thermo-electric (seebeck) effect in silicon,* Solid-State Electronics **15**, 849 (1972).

[20] S. M. Kashmiri, *Electrothermal frequency references in standard CMOS* (Springer, 2015).

[21] K. A. A. Makinwa and M. F. Snoeij, *A CMOS Temperature-to-Frequency Converter With an Inaccuracy of Less Than $\pm 0.5^{\circ}C$ (3σ) From $-40^{\circ}C$ to $105^{\circ}C$,* IEEE Journal of Solid-State Circuits **41**, 2992 (2006).

[22] W. Hewlett, *Variable frequency oscillation generator,* (1942).

[23] *Hewlett-Packard's first product: The 200A - HP history,* (2021).

[24] A. S. Delke, A.-J. Annema, M. S. O. Alink, Y. Jin, J. Verlinden, and B. Nauta, *A Single-Trim Frequency Reference Achieving ±120 ppm Accuracy From −50 °C to 170 °C,* IEEE Journal of Solid-State Circuits **56**, 3434 (2021).

[25] N. Sinoussi, A. Hamed, M. Essam, A. El-Kholy, A. Hassanein, M. Saeed, A. Helmy, and A. Ahmed, *A single LC tank self-compensated CMOS oscillator with frequency stability of ±100ppm from −40°C to 85°C,* in *2012 IEEE International Frequency Control Symposium Proceedings* (2012) pp. 1–5.

[26] V. De Smedt, G. G. E. Gielen, and W. Dehaene, *Transient Behavior and Phase Noise Performance of Pulsed-Harmonic Oscillators,* IEEE Transactions on Circuits and Systems I: Regular Papers **61**, 2119 (2014).

[27] K. Sundaresan, P. Allen, and F. Ayazi, *Process and temperature compensation in a 7-MHz CMOS clock oscillator,* IEEE Journal of Solid-State Circuits **41**, 433 (2006).

[28] Y.-C. Shih and B. Otis, *An On-Chip Tunable Frequency Generator for Crystal-Less Low-Power WBAN Radio,* IEEE Transactions on Circuits and Systems II: Express Briefs **60**, 187 (2013).

[29] M. Choi, S. Bang, T.-K. Jang, D. Blaauw, and D. Sylvester, *A 99nW 70.4kHz resistive frequency locking on-chip oscillator with 27.4ppm/ºC temperature stability,* in *2015 Symposium on VLSI Circuits (VLSI Circuits)* (2015) pp. C238–C239.

[30] J. Lee, P. Park, S. Cho, and M. Je, *5.10 A 4.7MHz 53µW fully differential CMOS reference clock oscillator with −22dB worst-case PSNR for miniaturized SoCs,* in *2015 IEEE International Solid-State Circuits Conference - (ISSCC) Digest of Technical Papers* (2015) pp. 1–3.

[31] T. Jang, M. Choi, S. Jeong, S. Bang, D. Sylvester, and D. Blaauw, *5.8 A 4.7nW 13.8ppm/°C self-biased wakeup timer using a switched-resistor scheme,* in *2016 IEEE International Solid-State Circuits Conference (ISSCC)* (2016) pp. 102–103.

[32] A. Vilas Boas and A. Olmos, *A temperature compensated digitally trimmable on-chip IC oscillator with low voltage inhibit capability,* in *2004 IEEE International Symposium on Circuits and Systems (ISCAS)*, Vol. 1 (2004) pp. I–501.

[33] K. Choe, O. D. Bernal, D. Nuttman, and M. Je, *A precision relaxation oscillator with a self-clocked offset-cancellation scheme for implantable biomedical SoCs,* in *2009 IEEE International Solid-State Circuits Conference - Digest of Technical Papers* (2009) pp. 402–403,403a.

[34] F. Sebastiano, L. J. Breems, K. A. A. Makinwa, S. Drago, D. M. W. Leenaerts, and B. Nauta, *A Low-Voltage Mobility-Based Frequency Reference for Crystal-Less ULP Radios,* IEEE Journal of Solid-State Circuits **44**, 2002 (2009).

[35] K.-J. Hsiao, *A 32.4 ppm/°C 3.2-1.6V self-chopped relaxation oscillator with adaptive supply generation,* in *2012 Symposium on VLSI Circuits (VLSIC)* (2012) pp. 14–15.

[36] Y. Tokunaga, S. Sakiyama, A. Matsumoto, and S. Dosho, *An On-Chip CMOS Relaxation Oscillator With Voltage Averaging Feedback,* IEEE Journal of Solid-State Circuits **45**, 1150 (2010).

[37] Y.-H. Chiang and S.-I. Liu, *A Submicrowatt 1.1-MHz CMOS Relaxation Oscillator With Temperature Compensation,* IEEE Transactions on Circuits and Systems II: Express Briefs **60**, 837 (2013).

[38] V. De Smedt, G. Gielen, and W. Dehaene, *A 0.6V to 1.6V, 46µW voltage and temperature independent 48 MHz pulsed LC oscillator for RFID tags,* in *IEEE Asian Solid-State Circuits Conference 2011* (2011) pp. 109–112.

3

RC References Based on Phase-to-Digital Converters

3.1. Introduction

The current trend in integrated frequency references is to design closed-loop systems in which integrated time constants are used to control the frequency of inaccurate oscillators. As discussed in chapter 2, such frequency-locked loops (FLLs) achieve the highest accuracy compared to harmonic or relaxation oscillators. FLLs circumvent the accuracy problem of traditional oscillators by transferring the accuracy constraints to their frequency-to-voltage converters (FVCs). However, FVCs still suffer from the high and nonlinear temperature coefficients of integrated components, necessitating nonlinear temperature compensation methods implemented in the digital domain.

Phase-domain FVCs with a high-resolution digital output can be made by employing a phase detector as the summing node of a delta-sigma modulator. The resulting phase-domain delta-sigma modulators (PDDSMs) have been used in temperature sensors to digitize the phase shift of electrothermal filters (ETFs) [1, 2] and RC filters [3]. Figure 3.1 shows the block diagram of a PDDSM for ETF readout. The phase-shifted output of the ETF is converted into a current, i_{etf}, which is then demodulated by one of two phase references ($\Phi_{0,1}$) at the same frequency, yielding an output current proportional to Φ_{etf}. Since it is part of a feedback loop, the gain of the loop filter drives i_{dem} to zero, on average. In this steady-state condition, the bitstream output average represents the phase shift of the ETF as a function of the phase references.

Since the ETF phase shift also depends on the input frequency, a PDDSM

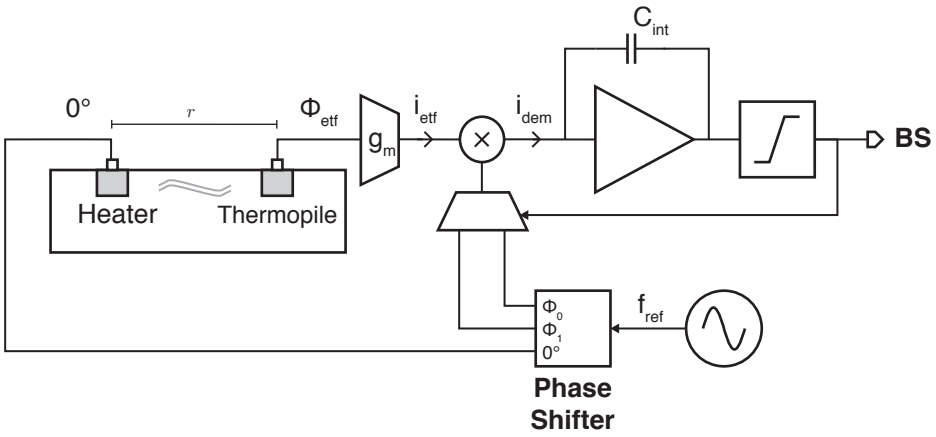


Figure 3.1: Schematic diagram of a phase-domain delta-sigma modulator used as an electrothermal temperature sensor[1], consisting of a synchronous demodulator embedded in a delta-sigma loop

can also be used to build a frequency-to-digital converter. In [4, 5] (Figure 2.25), the temperature dependent output of such a PDDSM is compensated by the output of a BJT-based temperature sensor. The compensated digital signal then controls the frequency of a digitally controlled oscillator, thus implementing a digital FLL that achieved an inaccuracy of 0.1% from -55 °C to 125 °C.

The work in this thesis aims to extend this architecture to the design of highly accurate RC frequency references. Phase shifts of RC filters will be digitized via PDDSMs, and their nonlinear temperature dependence will be accurately compensated in the digital domain to realize accurate frequency references.

The rest of this chapter will discuss the design of PDDSM-based RC frequency references. First, the choice of the RC filter will be discussed. Then, an analysis will be conducted on the time and frequency domain behavior of synchronous demodulators and PDDSMs. Nonidealities of these blocks and their effects on the resulting frequency reference will be discussed. Finally, an analysis of digital frequency-locked loop dynamics will be presented.

3.2. The RC Filter

Resistors and capacitors can be combined in many different ways to make filters. However, frequency reference applications have specific requirements which constrain the number of choices. Firstly, the filter order should be

minimized to reduce circuit complexity in applications that target accuracy. A lower-order filter generally features fewer internal nodes that can be corrupted by parasitics, which are detrimental to accuracy. A low-pass or band-pass structure should be chosen to impose a bandwidth limitation on the resistors' thermal noise. These combined constraints lead to two obvious candidates: a first-order low-pass filter or a second-order band-pass filter.

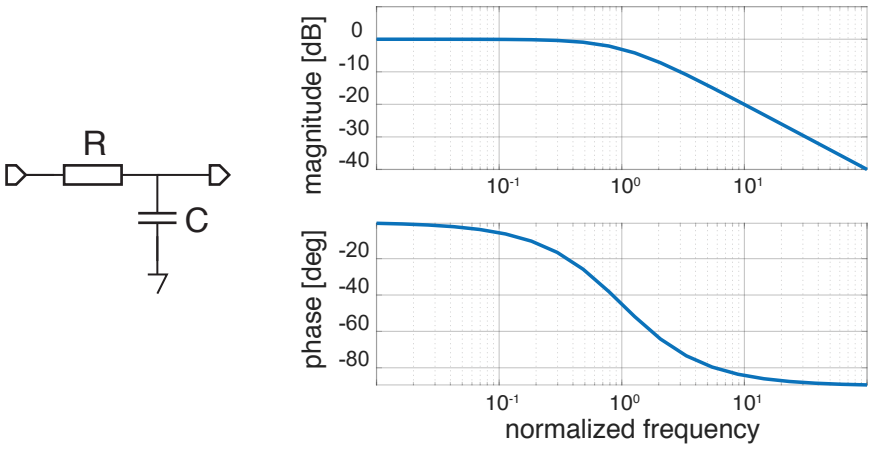


Figure 3.2: First-order low-pass filter schematic diagram and its magnitude and phase response

First Order Low-pass Filter The simplest filter that satisfies the above requirements is a first-order low-pass filter. Its magnitude and phase responses are:

$$|H(jw)| = \frac{1}{\sqrt{1+\omega^2R^2C^2}} \tag{3.1}$$

$$\angle H(jw) = tan^{-1}(-\omega RC) \tag{3.2}$$

As shown in Figure 3.2, the low-pass filter shows its highest phase sensitivity at its -3dB gain frequency (bandwidth) and a 45° phase shift. In a frequency control application, this is the point at which the filter would be utilized.

The Wien Bridge The Wien Bridge is a circuit initially created to measure capacitance with a frequency reference and a ratio of resistors. As shown

in Figure 3.3, its original form consists of two branches: a reference branch composed of two resistors (gray) and a measurement branch comprised of two resistors and two capacitors (black). The bridge is balanced when driven at the correct frequency; in this state, V_{meas} reads zero.

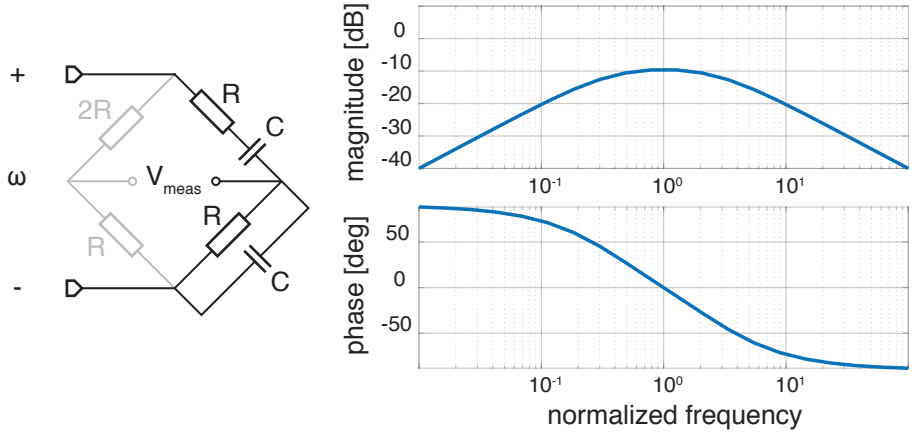


Figure 3.3: The Wien Bridge schematic diagram and its magnitude and phase response

The measurement branch of the Wien Bridge forms the frequency-dependent network that defines the time constant, and it is the circuit of interest. With the (-) input grounded, its magnitude and phase response are:

$$|H(jw)| = \frac{\omega RC}{\sqrt{(\omega RC)^4 + 7(\omega RC)^2 + 1}} \tag{3.3}$$

$$\angle H(jw) = \phi_{WB}(w) = tan^{-1}\left(\frac{1-(\omega RC)^2}{3\omega RC}\right) \tag{3.4}$$

Comparing the sensitivities of the Wien Bridge to the 1st-order low-pass filter reveals its advantages. Firstly, the Wien Bridge has a 0° phase shift, whereas the low-pass filter has a 45° phase shift at $\omega = (RC)^{-}1$. ϕ_{ref} for the synchronous demodulator of Figure 2.21 has to compensate for the extra phase shift of the low-pass filter, bringing additional complexity to generating the phase references. Additionally, at their center frequencies, the low-pass filter has less attenuation (-3 dB) than the Wien Bridge (≈ -9.5 dB). The larger output amplitude for the low-pass filter is better for achieving a larger SNR but requires a larger input dynamic range for the following stage.

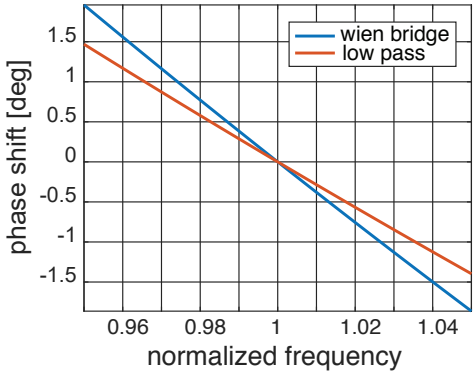


Figure 3.4: Phase shift vs. frequency characteristic of the Wien Bridge and the 1st-order low-pass filter. The LPF phase shift has been shifted by 45° for easy comparison.

However, the deciding factor between the two circuits is their phase shift's frequency sensitivity around their center frequency. Therefore, errors originating inside the loop manifest as smaller output frequency errors. As shown in Figure 3.4, the Wien Bridge has larger frequency sensitivity and is thus preferred over the first-order low-pass filter.

3.3. Synchronous Demodulator as a Phase Detector

Figure 2.21 shows the block diagram of a synchronous demodulator used for the phase detection of a linear filter. The filter input is a square wave with the amplitude, A, and the angular frequency, ω with the Fourier expansion:

$$v_i(t) = \frac{4A}{\pi} \sum_{k=1}^{\infty} \frac{\sin((2k-1)\omega t)}{2k-1} \tag{3.5}$$

The first input to the mixer is the filtered square wave, where each component is phase shifted by $\phi(k\omega)$ and attenuated by $|H(j\omega)|$:

$$v_{fil}(t) = \frac{4A}{\pi}\left(|H(j\omega)|\sin(\omega t + \phi(\omega)) + |H(j3\omega)|\frac{1}{3}\sin(3\omega t + \phi(3\omega)) + \ldots\right) \tag{3.6}$$

The second input to the mixer is a square wave with a ϕ_{ref} phase shift:

$$v_{\phi_{ref}}(t) = \frac{4A}{\pi}\left(\sin(wt + \phi_{ref}) + \frac{1}{3}\sin(3(wt + \phi_{ref})) + \ldots\right) \tag{3.7}$$

The mixer output, ignoring all but the fundamental, can then be written as:

$$v_x(t) = v_{fil}(t) \cdot v_{ref}(t) \tag{3.8}$$

$$= \frac{16A^2}{\pi^2}\left(|H(j\omega)|\sin(\omega t + \phi(\omega))\sin(\omega t + \phi_{ref}) + \ldots\right) \tag{3.9}$$

$$= \frac{8A^2|H(j\omega)|}{\pi^2}\left(\cos(\phi_{ref} - \phi(\omega)) - \cos(2\omega t + \phi(\omega) + \phi_{ref})\right) + \ldots \tag{3.10}$$

The second term in the above equation is the undesirable second harmonic. An ideal low-pass filter following the mixer removes this and any other harmonic content, leaving the final output as the frequency-dependent voltage:

$$v_o(\omega) = \frac{8A^2|H(j\omega)|}{\pi^2} cos\left(\phi_{ref} - \phi(\omega)\right) \tag{3.11}$$

This voltage represents the phase shift of the filter when driven at ω. If $\phi_{ref} = \pi/2$ then the output simplifies to:

$$v_o(\omega) = \frac{8A^2|H(j\omega)|}{\pi^2}\sin\left(\phi(\omega)\right) \tag{3.12}$$

When driven with a well-known frequency, this voltage can be used to determine the filter's phase shift. Conversely, if the zero-phase frequency of the filter is well-known, this voltage can be used as a frequency error signal. The output voltage of the mixer is a measure of the input frequency deviation from the well-known center frequency of the filter.

The issue with 3.12 is the sensitivity to input amplitude, A. The phase-domain delta-sigma modulator addresses this by achieving a ratiometric measurement independent of the amplitude.

3.4. Phase-Domain Delta-Sigma Modulator

In the phase domain DSM, as shown in Figure 3.5, two reference phases $\phi_{0,1} = \pi/2 - \phi'_{0,1}$ are used to demodulate the filter output. The demodulator output voltage is integrated and then quantized to a single-bit stream. This bitstream selects the reference to demodulate the signal in the next cycle.

Extending 3.12, the two possible output voltages of the mixer are:

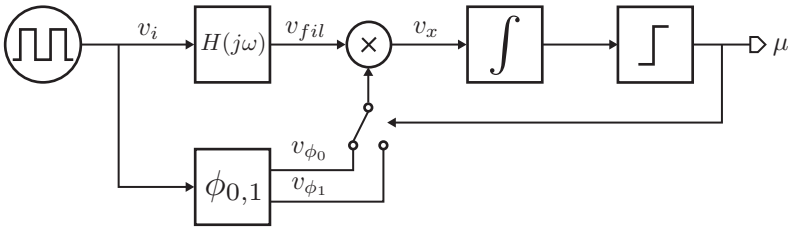


Figure 3.5: Block diagram of a PDDSM, a synchronous demodulator combined with a delta-sigma modulator loop

$$v_{x_{0,1}}(\omega) = \frac{8A^2|H(j\omega)|}{\pi^2}\sin\left(\phi(\omega) - \phi'_{0,1}\right) \tag{3.13}$$

The loop gain provided by the integrator in the delta-sigma modulator ensures that its input on average is zero, which defines a relationship between the average value of the bitstream output, μ, and integrator input as:

$$v_x = 0 \tag{3.14}$$
$$= \mu v_{x_0} + (1-\mu)v_{x_1} \tag{3.15}$$

It is then possible to derive the value of the bitstream average, μ, as a function of the phase references and the filter's phase shift.

$$\mu = \frac{v_{x_0}}{v_{x_0} - v_{x_1}} = \frac{\sin(\phi(\omega) - \phi'_0)}{\sin(\phi(\omega) - \phi'_0) - \sin(\phi(\omega) - \phi'_1)} \tag{3.16}$$

The coefficient that is dependent on amplitude in 3.12 disappears in 3.16, removing the sensitivity to the drive amplitude.

3.4.1. Cosine Nonlinearity

The preliminary analysis of the phase-domain delta-sigma modulator relies on the assumption that the integrator input is zero. However, this is true only *on average*. Equation 3.13 is nonlinear, which manifests as an error in the output bitstream, as shown in Figure 3.6.

The magnitude of this *cosine-nonlinearity* error is a function of the phase range of the PDDSM. As shown in Figure 3.6, for $\phi'_{0,1} = \pm\pi/4$, the error that

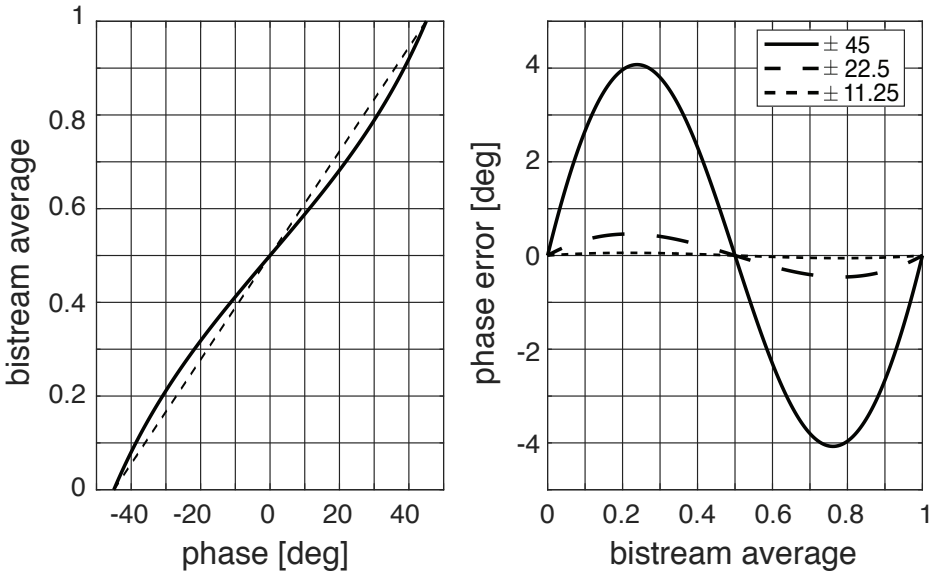


Figure 3.6: The ideal (dashed) and nonlinear (solid) output of an idealized PDDSM with $\phi'_{0,1} = \pm\pi/4$ and phase error for different phase reference ranges

originates from cosine nonlinearity can reach ±10% of the full range but reduces with smaller ranges.

Thankfully, cosine nonlinearity is a systematic error mechanism. 3.16 can be inverted analytically as in 3.17 for known phase references.

$$\phi(\omega) = \tan^{-1}\left(\frac{\mu\cos(\phi_1 - \phi_0) - \mu + 1}{\mu\sin(\phi_1 - \phi_0)}\right) - \phi_0 \tag{3.17}$$

Cosine nonlinearity correction fits nicely into a digital implementation. However, the trigonometric functions involved in the inverse function bring significant computational overhead. For this reason, a better approach is to use a polynomial approximation of 3.17, as demonstrated in chapter 4.

3.4.2. Effect of Higher-Order Harmonics

An assumption made in 3.10 was that only the fundamental of the square wave contributes to the output voltage. In reality, the odd harmonics of the square wave also contribute to the output voltage. Multiple DC components that are a function of the attenuation and phase shift of the filter at these higher frequencies appear. Considering the effect of the higher-order harmonics and assuming the same low-pass filter at the output yields the

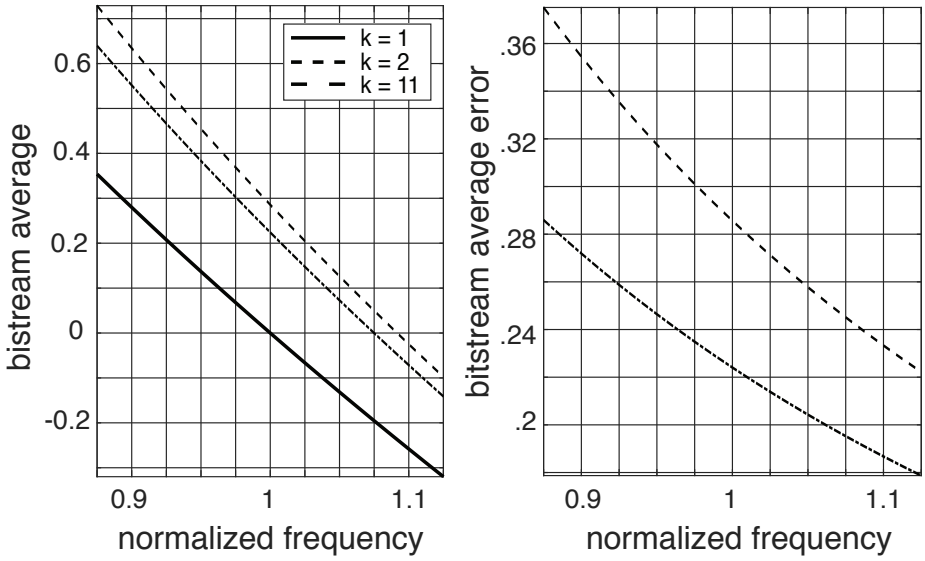


Figure 3.7: Bitstream output of a PDDSM and error with contributions from higher harmonics of a square wave

following equation for the demodulator output voltage:

$$v_x(t) = v_f(t) \cdot v_{ref}(t) \tag{3.18}$$

$$= \frac{8A^2}{\pi^2} \sum_{k=0}^{\infty} \frac{\left|H(j(2k-1)\omega\right|}{(2k-1)!} \cos\left((2k-1)\phi_{ref} - \phi((2k-1)\omega)\right) \tag{3.19}$$

Components that arise from the higher-order harmonics change the output bitstream of a PDDSM in two ways. Firstly, the output shifts from the expected zero value at the drive frequency, which causes a significant offset. Moreover, the higher-order harmonics show phase sensitivity across a more extensive range of the filter. These harmonics are centered around nonlinear portions of the filter's phase response and make the demodulator output more nonlinear.

Figure 3.7 shows the bitstream output of a PDDSM with a Wien Bridge filter when the higher harmonics of the square wave are considered. With only the third harmonic, the output shows a significant change in the bitstream average, and the error shows that nonlinearity over frequency is introduced.

Moreover, as the number of harmonics increases, the bitstream deviates further from its expected value, and its nonlinear characteristic changes.

Similar to the cosine nonlinearity, the effect of higher-order harmonics on the PDDSM's output is systematic. However, reversing this effect is more challenging, as the output depends on both the filter characteristics and the harmonic content of the driving square waves. Designing the system for a given center frequency is only possible through numerical simulation, as inverting 3.19 does not yield a closed-form expression. Meanwhile, the characteristics of the driving square wave are highly dependent on circuit parasitics. Generally, circuit simulations that include the parasitics of the drive circuitry are required to accurately characterize the contribution of higher-order harmonics to synchronous demodulators.

3.5. Frequency-Locked Loops

A frequency-locked loop (FLL) is a control system that locks the frequency of a controllable oscillator to the frequency defined by a reference.

All frequency-locked loops consist of three primary components. The first component is the frequency error detector, which converts the difference between the reference and the feedback frequency into a quantity the following components can process. The second component, the loop filter, generally takes the form of an integrator and provides gain for the feedback loop. The final component is the controlled oscillator, which provides the system's output and the feedback input to the frequency error detector.

FLLs can be realized either in the analog or the digital domain. In analog FLLs, the loop filter is generally a g_m-C or an RC feedback integrator. The frequency error detector in the analog domain is generally called a frequency-to-voltage converter (FVC). The controlled oscillator is referred to as a voltage-controlled oscillator (VCO). In digital FLLs, the loop filter input is a digital word, and the loop filter can be implemented as a simple accumulator. The frequency error detector in the digital domain is called a frequency-to-digital converter (FDC), and the controlled element is a digitally controlled oscillator (DCO).

3.5.1. Voltage Controlled Oscillator Dynamics

The instantaneous frequency of a VCO is defined by its center frequency, f_0, and its sensitivity, K_{vco}, and the value of its control input v_x as in 3.20. The phase of the VCO is the time integral of its instantaneous frequency, given in 3.21.

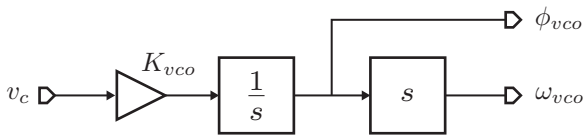


Figure 3.8: The linear model of a voltage-controlled oscillator

$$\omega_{vco}(t) = 2\pi(f_0 + K_{vco}v_c(t)) \tag{3.20}$$

$$\phi_{vco}(t) = \int_0^t \omega_{vco}(\tau)d\tau = 2\pi f_0 t + K_{vco}\int_0^t v_c(\tau)d\tau \tag{3.21}$$

The Laplace transform of its transfer function can be derived as:

$$\mathcal{L}\left\{\frac{\omega_{vco}(t)}{v_c(t)}\right\} = \frac{\omega_{vco}(s)}{v_c(s)} = K_{vco} \tag{3.22}$$

$$\mathcal{L}\left\{\frac{\phi_{vco}(t)}{v_c(t)}\right\} = \frac{\phi_{vco}(s)}{v_c(s)} = \frac{K_{vco}}{s} \tag{3.23}$$

Figure 3.8 shows the block diagram for 3.23. An implicit integrator between the VCO's input voltage and output phase exists, and the phase variable can be accessed when the VCO is used in a system with a phase detector. Then, this implicit integrator can be used as part of a control loop. For example, in a PLL, the integrator in the VCO provides the infinite loop gain for phase at DC.

In contrast, for systems where a frequency detector is utilized, the integrator is coupled with an implicit differentiator. This differentiator cancels the implicit integrator, and the VCO behaves like a gain with the value K_{vco}.

3.5.2. Analog FLL Dynamics

Figure 3.9 shows the block diagram of the analog frequency-locked loop based on the synchronous demodulator. The filter (Wien Bridge) and the synchronous demodulator comprise the frequency-to-voltage converter. The filter's phase shift, ϕ_{wb}, represents the deviation from its center frequency. This phase shift is converted to the error voltage, v_e, via synchronous demodulation. The loop filter integrates v_e, and its output, v_c, drives the

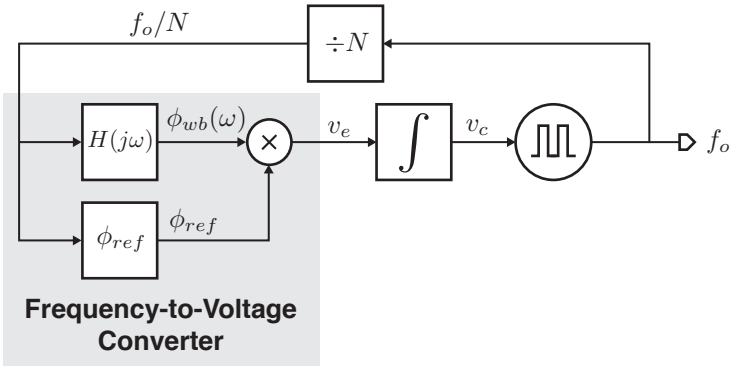


Figure 3.9: Block diagram of an analog frequency-locked loop based on the synchronous demodulator

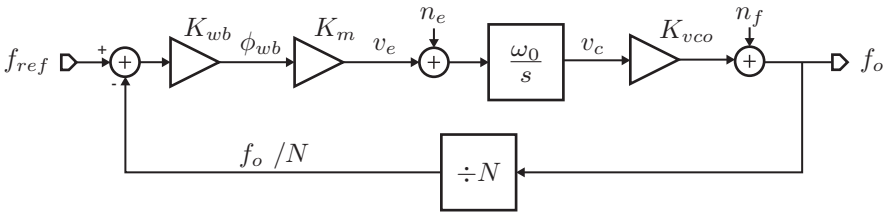


Figure 3.10: The linear model of the analog frequency-locked loop based on the synchronous demodulator

voltage-controlled oscillator. The oscillator output provides the system's output frequency and is divided before providing the input to the Wien Bridge to close the feedback loop.

To analyze the dynamics of the model, each of these components can be linearized about their steady-state operating points at the locked condition. Figure 3.10 shows the block diagram of the linearized FLL model. In the locked state, v_e is zero, and the system's output will have settled to the center frequency of the WB multiplied by N, $f_o = N f_{ref}$. The f_{ref} input of the linear FVC is a virtual one, integrated into the time constant of the FVC. K_{wb} is the frequency-to-phase sensitivity of the Wien Bridge and is the derivative of 3.4 at the zero-phase condition, as in:

$$K_{wb} = \left.\frac{\partial \phi_{wb}}{\partial f}\right|_{f=f_{ref}} = -\frac{4\pi}{3}RC \tag{3.24}$$

K_m is the phase-to-voltage sensitivity of the synchronous demodulator, and with the higher order harmonics ignored. It replaces the multiplication in Figure 3.9 and is the derivative of 3.12 at the zero phase condition, as in:

$$K_m = \left.\frac{\partial v_o}{\partial \phi}\right|_{\phi_{wb}=0} = \frac{8A^2}{3\pi^2} \tag{3.25}$$

With K_{vco} being the voltage-to-frequency sensitivity of the VCO, the open loop gain of the feedback loop is then:

$$T(s) = \frac{K_{wb}K_m K_{vco}\omega_0}{N}\frac{1}{s} \tag{3.26}$$

The analog FLL is a single-pole system. This pole is at zero, and the transfer function has a unity gain frequency of $K_{wb}K_m K_{vco}\omega_0/N$. The system is unconditionally stable with a maximal phase shift of 90° as frequency tends to infinity. Varying any sensitivity parameters does not drive the system to instability but changes the system's unity gain frequency and settling time.

The linear model of the FLL can also be utilized to analyze the behavior of error signals injected into different points in the loop. The first point of injection of interest is before the loop filter, n_e, which represents noise injected from the FVC components and the loop integrator. The second point is at the VCO output, n_f, which represents the noise of the VCO. These two signals have the transfer functions to the output frequency as the following:

$$\frac{f_o}{n_e} = \frac{K_{vco}\omega_0}{s + \frac{K_{wb}K_m K_{vco}\omega_0}{N}} \tag{3.27}$$

$$\frac{f_o}{n_f} = \frac{s}{s + \frac{K_{wb}K_m K_{vco}\omega_0}{N}} \tag{3.28}$$

Figure 3.11 shows the frequency response of the error transfer functions of the analog FLL. f_o/n_e shows a low-pass characteristic with a 20 dB/dec roll-off after the cut-off frequency at the unity-gain frequency of the open-loop transfer function. At DC, it has a gain of $N/(K_m K_{wb})$. A high frequency-to-phase sensitivity (K_{WB}) is desirable to reduce the noise contribution of elements that follow the Wien Bridge. However, a higher frequency-to-phase

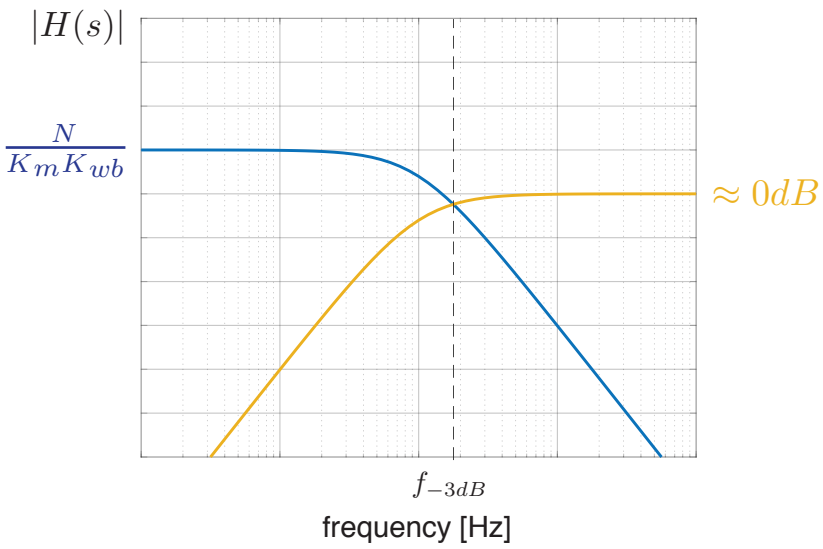


Figure 3.11: Frequency response of the analog FLL error transfer functions

sensitivity also implies a higher low-pass bandwidth, which requires adjusting other parameters to keep the bandwidth the same.

f_o/n_f, however, shows a high-pass characteristic and a gain of 1 at its passband. The loop attenuates the VCO's low-frequency noise; its transfer function has a zero at DC, implying that the static frequency error of the VCO is corrected.

One of the primary disadvantages of the analog FLL is that the loop bandwidth is a strong function of analog circuit parameters. While the filter sensitivity is usually well controlled due to its accuracy requirements (around ±10%), the bandwidth of the analog accumulator, ω_0, and K_{vco} traditionally spread more (±50%). These parameter variations can cause the bandwidth of the loop to vary significantly and cause the settling or noise behavior to change. Another disadvantage of the analog FLL is that the size of the devices can be significant. Usually, low bandwidth is required to filter out the noise of the FVC and attenuate any high-frequency signal content at the input of the VCO. Large passive devices (usually capacitors) are required to implement a low ω_0.

Both issues are analogous to analog phase-locked loops (PLLs). These have been addressed by the advent of all-digital PLLs, where digital phase detection allows moving the loop filter into the digital domain [6]. Digital loop filters will enable the implementation of a low-bandwidth loop without

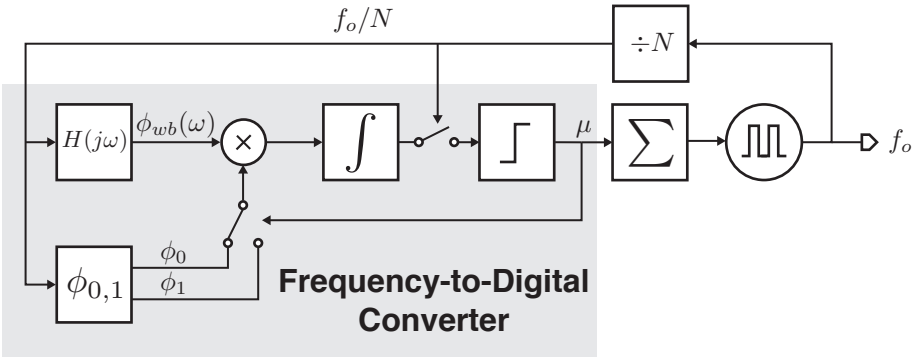


Figure 3.12: Block diagram of a digital frequency-locked loop based on the phase-domain delta-sigma modulator

resorting to large passive components. A programmable digital gain in the loop can ease the tuning of the loop bandwidth to counter the effect on heavily process-dependent K_{vco} [6]. Similar benefits can be achieved in FLLs by moving the loop filter to the digital domain, as discussed in the next section.

3.5.3. Digital FLL Dynamics

Figure 3.12 shows the block diagram of the digital frequency-locked loop based on the phase-domain delta-sigma modulator. The delta-sigma modulator's local feedback loop converts the filter's phase shift, ϕ_{wb}, into a bitstream with the average μ. This bitstream is accumulated, which then drives the digitally controlled oscillator. Similar to the analog FLL, the output of the digitally controlled oscillator provides the ultimate output of the system. It is fed back after division to both drive the Wien Bridge and provide the sampling clock for the phase-domain delta-sigma modulator.

While the analog and digital FLLs are analogous, they differ in one major aspect: the digital FLL is a sampled system. The sampling operation is implemented at the phase-domain delta-sigma modulator's comparator and sets the system's sampling frequency to $f_s = f_o/N$. Conversely, a zero-order hold is implied at the DCO input as the frequency of the oscillator changes at the sample rate. This sampled operation makes the digital FLL susceptible to aliasing, where the high-frequency noise of blocks can fold back and corrupt the DC signal. Thankfully, the integrator of the continuous-time delta-sigma

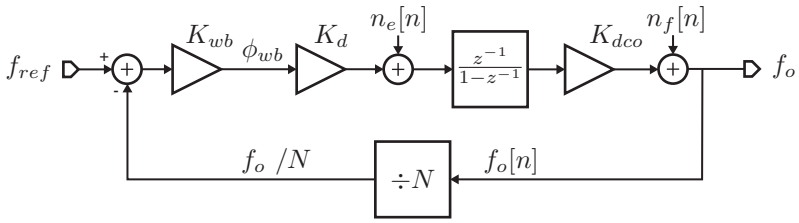


Figure 3.13: The linear model of the digital frequency-locked loop based on the phase-domain delta-sigma modulator

loop provides the required anti-alias filtering before sampling. With anti-aliasing, the system can be treated as a discrete-time feedback loop with the sample frequency f_s.

Figure 3.13 shows the block diagram of the linearized digital FLL. The parameters K_{wb} and K_{dco} are similar to their analog counterparts, with K_{dco} being the digital-to-frequency sensitivity (i.e., LSB in frequency) of the DCO. K_d is the phase-to-digital sensitivity of the delta-sigma modulator and is the derivative of 3.16. With $\phi_{0,1}$ defined as $\pm\phi_{ref}$, the gain can then be defined as:

$$K_d = \left.\frac{\partial \mu}{\partial \phi}\right|_{\phi_{wb}=0} = \frac{1}{\phi_{ref}} \tag{3.29}$$

The open-loop gain of the feedback loop can then be written as:

$$T(z) = \frac{K_{wb}K_dK_{dco}}{N}\frac{z^{-1}}{1-z^{-1}} \tag{3.30}$$

It is possible to derive the transfer functions of error signals to the frequency output using 3.30 as:

$$\frac{f_o}{n_e}(z) = \frac{K_{dco}z^{-1}}{1-(\frac{K_{wb}K_dK_{dco}}{N})z^{-1}} \tag{3.31}$$

$$\frac{f_o}{n_f}(z) = \frac{1-z^{-1}}{1-(\frac{K_{wb}K_dK_{dco}}{N})z^{-1}} \tag{3.32}$$

Figure 3.14 shows the frequency response of the error transfer functions of the digital FLL. n_e represents errors preceding the accumulator, such as

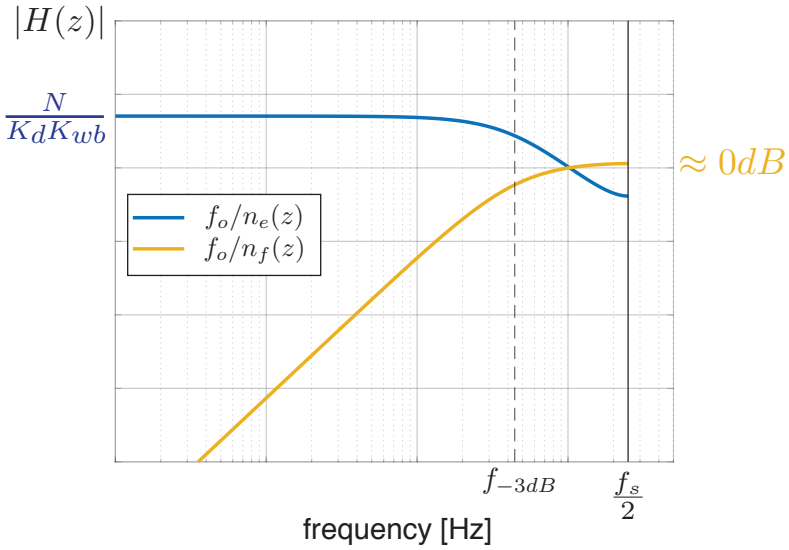


Figure 3.14: Frequency response of the digital FLL error transfer functions

the quantization noise of the delta-sigma modulator or the accumulator itself, and is low-pass filtered to the output. n_f represents the errors after the accumulator, primarily the quantization and random noise of the digitally-controlled oscillator, and is high-pass filtered to the output. One crucial difference of the digital FLL is that the cut-off frequency combines digital (K_d) and analog (K_{wb} and K_{dco}) sensitivities. While the digital sensitivities are well defined, the analog still suffers from process spread. The unity gain of the open-loop transfer function can get close to or even exceed the sampling frequency, which usually results in unstable behavior. However, this problem can be addressed in the digital domain by adding a digital gain around the accumulator to tune the bandwidth of the loop to ensure stability.

3.6. Conclusion

This chapter presented a study of frequency references based on synchronous demodulation of the phase shift of an RC filter. The various components of such a reference were analyzed, from the RC filter at the very core to the synchronous-demodulator-based phase-domain delta-sigma modulator and the larger digital frequency-locked loop. The analysis in this chapter lays the groundwork for the experimental work conducted in chapters 4 and 5.

A few crucial points that direct the choices in the following work origi-

nate from the analysis in this chapter. Firstly, due to its higher frequency-to-phase sensitivity, the Wien Bridge shows advantages in frequency control applications compared to a more straightforward first-order low-pass filter. Thus, both experimental works in this thesis utilize the Wien Bridge as their frequency-sensitive element. Additionally, the systematic nonidealities originating from synchronous demodulation were identified. In chapter 4, cosine nonlinearity correction is implemented as a separate digital correction stage to address one such systematic nonideality. Finally, the expressions used in analyzing the dynamics of frequency control loops are utilized in both chapters 4 and 5.

References

[1] S. M. Kashmiri, S. Xia, and K. A. A. Makinwa, *A Temperature-to-Digital Converter Based on an Optimized Electrothermal Filter,* IEEE Journal of Solid-State Circuits **44**, 2026 (2009).

[2] C. P. van Vroonhoven, D. d'Aquino, and K. A. Makinwa, *A thermal-diffusivity-based temperature sensor with an untrimmed inaccuracy of ±0.2°c (3s) from −55°c to 125°c,* in *2010 IEEE International Solid-State Circuits Conference - (ISSCC)* (2010) pp. 314–315.

[3] M. Shahmohammadi, K. Souri, and K. A. Makinwa, *A resistor-based temperature sensor for MEMS frequency references,* in *2013 Proceedings of the ESSCIRC (ESSCIRC)* (2013) pp. 225–228.

[4] S. M. Kashmiri, M. A. P. Pertijs, and K. A. A. Makinwa, *A Thermal-Diffusivity-Based Frequency Reference in Standard CMOS With an Absolute Inaccuracy of ±0.1% From -55 °C to 125 °C,* IEEE Journal of Solid-State Circuits **45**, 2510 (2010).

[5] S. M. Kashmiri, K. Souri, and K. A. A. Makinwa, *A Scaled Thermal-Diffusivity-Based 16 MHz Frequency Reference in 0.16 µm CMOS,* IEEE Journal of Solid-State Circuits **47**, 1535 (2012).

[6] R. Staszewski and P. Balsara, *Phase-domain all-digital phase-locked loop,* IEEE Transactions on Circuits and Systems II: Express Briefs **52**, 159 (2005).

4

A ±200 ppm CMOS Dual-RC Frequency Reference

This chapter proposes a dual-RC architecture that achieves higher-order temperature compensation by combining the phase responses of two RC networks with complementary temperature dependencies in the digital domain. After a 2-point trim, the compensated result emulates a zero-TC RC network. The implemented 7 MHz frequency reference achieves ±250 ppm from -45 °C to 85 °C with an open-loop trimming scheme, which improves to ±165 ppm after an additional closed-loop trimming step [1].

4.1. Architecture Proposal

Figure 4.1 shows the block diagram of the proposed Dual-RC frequency reference, which can be seen as a digital frequency-locked loop (FLL). The output frequency, f_{ref}, of a digitally-controlled oscillator (DCO) is controlled by a temperature-independent frequency error signal e_T. f_{ref} is divided by 16 to f_{drv} and drives the two Wien Bridge filters. The relationship between the passive values and the two frequencies is:

$$f_{ref} = N f_{drv} = \frac{N}{2\pi R_{\{p,n\}} C} \tag{4.1}$$

The filters impart a phase shift on f_{drv}, which is digitized via phase-domain delta-sigma modulators. Their output bitstreams are decimated to high-resolution digital phase words, $\phi_{p,n}$. The phase words have a large and nonlinear temperature dependence due to the higher-order TCs of $R_{p,n}$. The

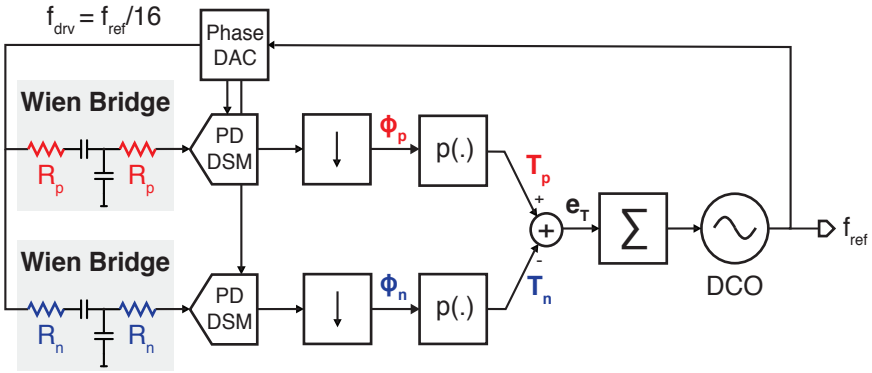


Figure 4.1: Block diagram of the Dual-RC frequency reference, showing the two-channel digital frequency-locked loop.

phase words are first subjected to further processing in the digital domain via two different polynomials, $p_{p,n}(.)$, to map them into linear functions of temperature. Following the polynomial processing, they are combined to obtain the temperature-independent frequency error signal e_T. A digital accumulator closes the feedback loop by accumulating e_T and driving the DCO's frequency control word. The temperature-dependence of $\phi_{p,n}$ are selected to be complementary, which, as will be discussed later, is essential to achieving good cancellation.

The Dual-RC architecture differs from the traditional FLL in one significant aspect. In conventional FLL architectures, the frequency error and temperature sensing channels are distinct. The frequency error channel extracts the frequency error, and the temperature sensor measures the die temperature. This distinction between the frequency detector and the temperature sensor is blurred in the Dual-RC architecture. Both channels are sensitive to both quantities and simultaneously serve as frequency and temperature sensors. Having some frequency sensitivity in both channels proves beneficial in terms of accuracy, as will be discussed later in this section.

The rest of this section provides further analysis of the Dual-RC architecture. First, the Wien Bridge and the Dual-RC temperature compensation scheme will be discussed. Then, the analysis will be used to develop a linear model of the Dual-RC FLL. Finally, inaccuracy and dynamic analysis of the Dual-RC reference will be presented.

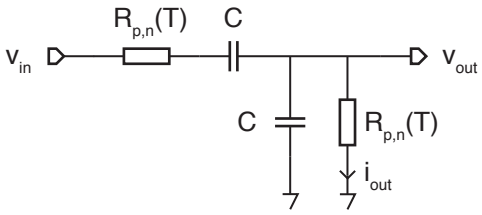


Figure 4.2: Single-ended schematic for the Wien Bridge filter

4.1.1. Wien Bridge

The Wien Bridge (WB) at the core of the frequency reference is shown in Figure 4.2. It is a second-order RC bandpass filter comprised of resistors, $R_{p,n}$, and capacitors, C. The WB capacitors are considered temperature-independent due to their much smaller magnitude TC. However, $R_{p,n}$ are significantly temperature-dependent, with their resistance given by:

$$R_{\{p,n\}}(T) = R_0(1 + TC_{1,\{p,n\}}(T - T_0) + TC_{2,\{p,n\}}(T - T_0) + ...) \tag{4.2}$$

When the divided output of the DCO drives the WB at f_{drv}, it imparts a phase shift to the signal at its output given by:

$$\phi_{\{p,n\}}(f, T) = -\tan^{-1}\left(\frac{R_{\{p,n\}}(T)^2 C^2 (2\pi f)^2 - 1}{3R_{\{p,n\}}(T) C (2\pi f)}\right) \tag{4.3}$$

Thus, the phase shift in 4.3 is a function of two independent variables: frequency and temperature. Figure 4.3 shows the dependence of the output phase on frequency and temperature as a function of two variables. The desirable frequency dependence is used to create a frequency error signal for an FLL to track. The undesirable temperature dependency corrupts the frequency error signal with temperature and needs to be compensated. This is addressed with the proposed Dual-RC temperature compensation.

4.1.2. Dual-RC Temperature Compensation

The Dual-RC temperature compensation scheme uses two WB filters built with resistors with complementary TCs to achieve temperature compensation. As described in the previous section, the WB output phase is a function of two independent variables: frequency and temperature. Extracting information about these two variables from the output phase, a function of both, requires minimally two independent outputs. Such a two-channel linearized

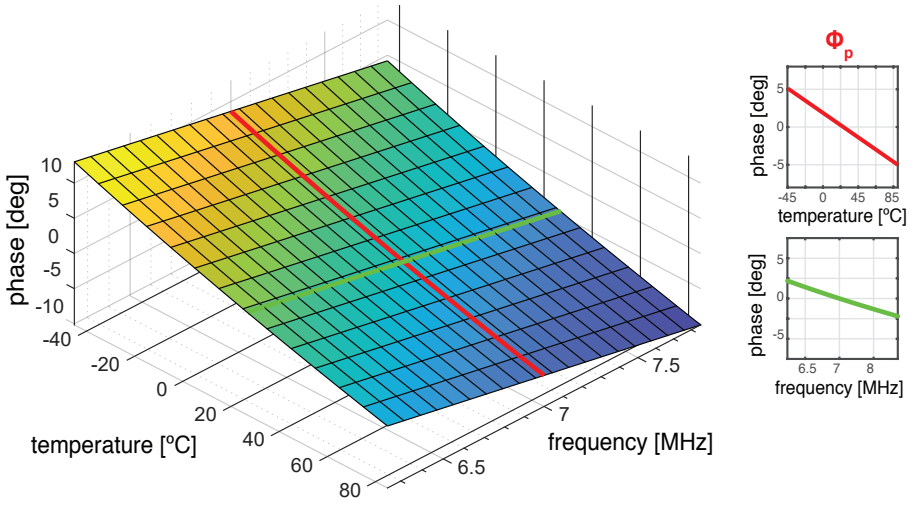


Figure 4.3: Frequency and temperature dependence of WB_p phase, additionally showing the temperature dependence at $f = f_{drv}$ and frequency dependence at $T = T_0$.

system is described in matrix form in 4.4. $K_{T,\{p,n\}}$ and $K_{f,\{p,n\}}$ are the two channels' linear temperature-to-phase and frequency-to-phase sensitivities.

$$\begin{bmatrix}\phi_p\\ \phi_n\end{bmatrix} = \begin{bmatrix}K_{T,p} & K_{f,p}\\ K_{T,n} & K_{f,n}\end{bmatrix} \cdot \begin{bmatrix}T\\ f\end{bmatrix} + \begin{bmatrix}\phi_{p,0}\\ \phi_{n,0}\end{bmatrix} \tag{4.4}$$

The matrix must be invertible for this system of equations to have a solution. Since the frequency sensitivity of two WBs with the same center frequency is the same, this leaves $K_{T,\{p,n\}}$ for the designer. In short, the two WBs should be constructed with resistors with different TCs to ensure a solution. In this design, one WB employs silicided p-polysilicon resistors, R_n with $TC_{1,n}$ = 0.33 %/°C, and the other uses unsilicided n-polysilicon resistors R_p with $TC_{1,p}$ = -0.15 %/°C.

Additionally, the nonlinear dependence of $\phi_{p,n}$ on temperature and frequency requires a slightly more complicated solution to the compensation problem. Solving a generic system of nonlinear equations requires iterative methods. For an FLL system, correct solutions are only required in the locked state, where one variable (frequency) is well-defined. Thus, a simpler solution can be achieved by linearizing the phase outputs about their

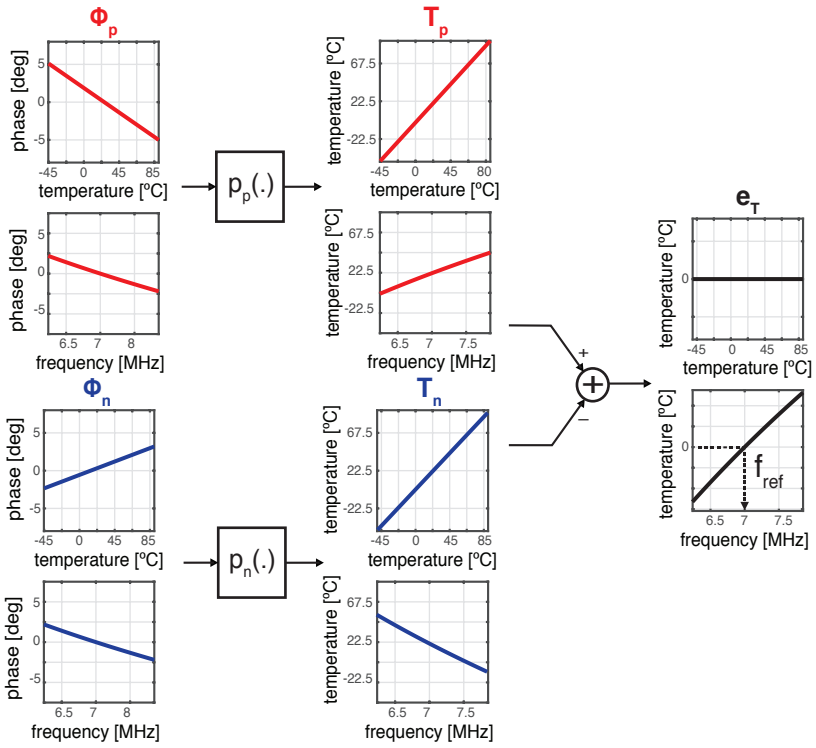


Figure 4.4: Temperature compensation principle for the proposed Dual-RC frequency reference

intended operating point of $f_{drv} = f_{ref}/N$ and combining them to realize the loop error signal.

Figure 4.4 demonstrates the operating principle of the Dual-RC compensation scheme. The phase outputs $\phi_{p,n}$ show similar frequency dependence but opposite temperature dependence. Two polynomial functions $p_{p,n}(\phi)$ map the phase outputs to digital words representing the die temperature $T_{p,n}$ as in 4.5. It is important to note that $T_{p,n}$ are sensitive to frequency and represent the correct temperature only when $f = f_{ref}$.

$$p_{\{p,n\}}(\phi_{WB,\{p,n\}}(f,T)) = T_{\{p,n\}}(f,T) \tag{4.5}$$

$$T_{\{p,n\}}(f,T)\big|_{f=f_{ref}} = T \tag{4.6}$$

The loop error signal, e_T, can be computed by the difference of $T_{p,n}$ as in 4.8.

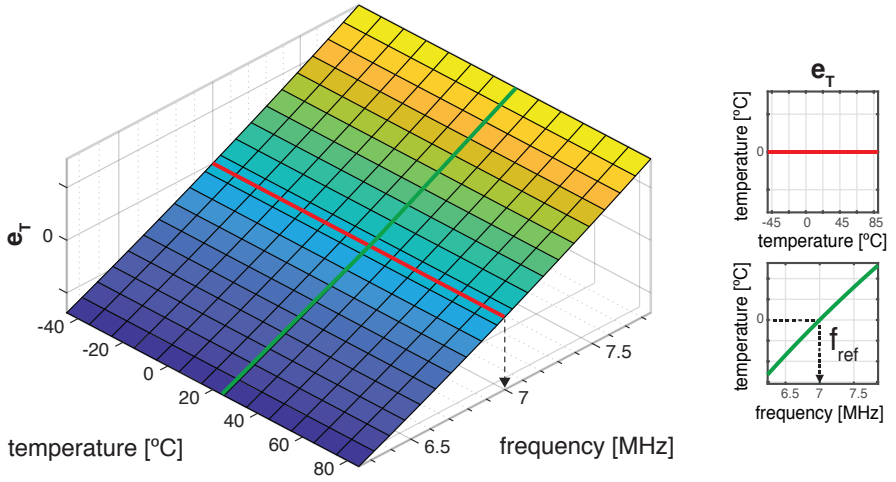


Figure 4.5: Frequency and temperature dependence of e_T, additionally showing the temperature dependency at f = f_{ref} and frequency dependence at T = T_0

$$e_T(f,T) = T_p(f,T) - T_n(f,T) \tag{4.7}$$
$$= p_p(\phi_{WB,p}(f,T)) - p_n(\phi_{WB,n}(f,T)) \tag{4.8}$$

$$e_T(f,T)\big|_{f=f_{ref}} = T - T = 0 \tag{4.9}$$

As shown in Fiugre 4.4, $_T$ crosses zero only at f = f_{ref}; additionally, it has zero temperature sensitivity about this operating point. In steady-state, the gain of the loop accumulator drives e_T to zero, ensuring that the DCO's output frequency is locked to f_{ref}.

For the frequency reference to operate correctly, its sensitivity to frequency should be maintained across all operating points. Moreover, it should not change sign to ensure loop stability. Figure 4.5 shows the frequency and temperature dependence of e_T. It shows that the temperature dependence at f = f_{ref} is zero, and additionally, it maintains a relatively constant sensitivity to frequency across the entire design space. The monotonicity of e_T ensures that convergence to the desired reference frequency will always be possible.

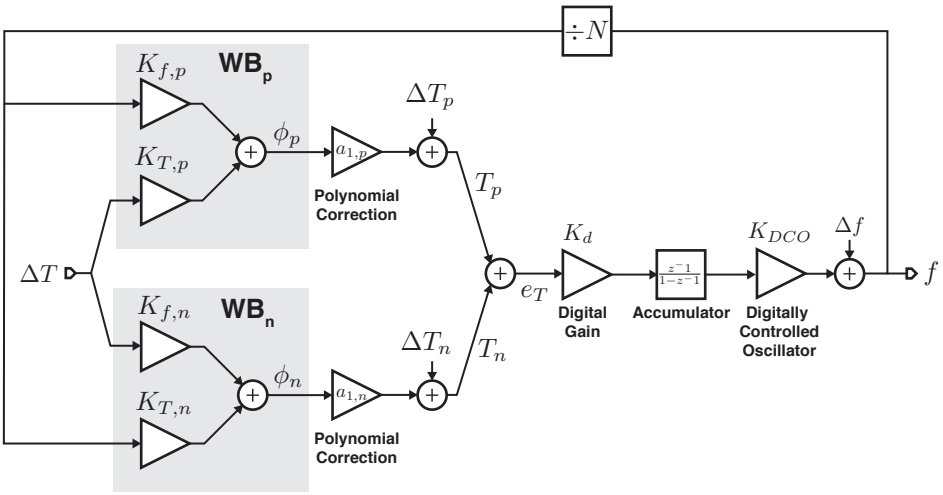


Figure 4.6: The linear model of the Dual-RC frequency reference

4.1.3. Linear Model

A linear model of the Dual-RC frequency reference was developed to study its inaccuracy and dynamics, as shown in Figure 4.6. The two WBs and their readouts are modeled as frequency-to-phase converters with two inputs: frequency and temperature. They output the phase signals $\phi_{WB,\{p,n\}}$, related to the two inputs via the frequency-to-phase sensitivities $K_{f,\{p,n\}}$ as in 4.10 and the temperature-to-phase sensitivities $K_{T,\{p,n\}}$ as in 4.11. These simplifications allow the two-channel system to be linearized as formulated in 4.4.

$$K_{f,p} = K_{f,n} = \left.\frac{\partial \phi}{\partial f}\right|_{\substack{f=f_{ref}/N \\ T=T_0}} = -\frac{4\pi}{3} R_0 C \tag{4.10}$$

$$K_{T,\{p,n\}} = \left.\frac{\partial \phi_{\{p,n\}}}{\partial T}\right|_{\substack{f=f_{ref}/N \\ T=T_0}} = -\frac{2}{3} T C_{1,\{p,n\}} \tag{4.11}$$

The linearization polynomials $p_{\{p,n\}}$ are modeled as gain blocks that map the phase outputs to temperature per their definition in 4.5. Values of these gain blocks are approximated as the linear coefficient of these polynomials:

$$\frac{\partial}{\partial\phi}(p_{\{p,n\}}(\phi)) = \frac{\partial}{\partial\phi}(a_{0,\{p,n\}} + a_{1,\{p,n\}}\phi + a_{2,\{p,n\}}\phi^2 + \ldots) \approx a_{1,\{p,n\}} \quad (4.12)$$

Furthermore, since the linearizing polynomials normalize the temperature sensitivity of their respective bridges as in 4.6, the value for the first order coefficient can be derived by:

$$\frac{\partial}{\partial T}\left(p_{\{p,n\}}\left(\phi_{\{p,n\}}(f,T)\right)\right)\Big|_{f=f_{ref}} = 1 \quad (4.13)$$

$$\frac{\partial p_{\{p,n\}}}{\partial \phi_{\{p,n\}}}\frac{\partial \phi_{\{p,n\}}}{\partial T}\Big|_{f=f_{ref}} = 1 \quad (4.14)$$

$$a_{1,\{p,n\}} = \frac{\partial \phi_{\{p,n\}}}{\partial T}^{-1} = K_{T,\{p,n\}}^{-1} \quad (4.15)$$

The loop error signal e_T is achieved by subtracting the outputs of the polynomial correction blocks. The loop accumulator is modeled by its z-domain transfer function, with the sample rate T_s. A gain block represents the programmable adjustment of the loop gain in the digital domain. Finally, the DCO is modeled as a digital-to-frequency converter with the gain K_{DCO}. The output of the DCO is the output frequency of the reference. The output frequency closes the feedback by driving the two WBs after being divided by N, representing the frequency division in the phase DAC.

4.1.4. Inaccuracy Analysis

The inaccuracy of the Dual-RC frequency reference is analyzed by referring its primary error sources, the phase inaccuracy of the WBs, to the frequency output. It is convenient to model these as temperature sensing errors, ΔT_p and ΔT_n, for comparison with prior work on RC-based temperature sensors [2, 3]. These temperature errors are referred to the frequency output by:

$$
\begin{aligned}
e_T = 0 &= \Delta f(K_{f,p}a_{1,p} - K_{f,n}a_{1,n}) + \Delta T_{p,n} \\
\Delta f &= \Delta T_{p,n}\frac{1}{K_{f,n}a_{1,n} - K_{f,p}a_{1,p}} \\
&= \Delta T_{p,n}\frac{1}{\frac{\partial\phi_{WB,n}}{\partial f}\left(\frac{\partial\phi_{WB,n}}{\partial T}\right)^{-1} - \frac{\partial\phi_{WB,p}}{\partial f}\left(\frac{\partial\phi_{WB,p}}{\partial T}\right)^{-1}} \\
&= \Delta T_{p,n}\frac{1}{\frac{2\pi R_0 C}{N}\left(\frac{1}{TC_{1,n}} - \frac{1}{TC_{1,p}}\right)}
\end{aligned}
\tag{4.16}
$$

Using 4.16 and 4.1, the standard deviation of the output frequency σ_f due to the standard deviation of temperature sensing errors of the two channels, $\sigma_{T,p}$ and $\sigma_{T,n}$ can be derived as:

$$
\frac{\sigma_f}{f_0} = \sqrt{\sigma_{T,p}^2 + \sigma_{T,n}^2}\left|\frac{1}{\frac{1}{TC_{1,n}} - \frac{1}{TC_{1,p}}}\right| \tag{4.17}
$$

The denominator of 4.17 reveals that temperature sensing errors are transferred to normalized frequency error with an effective TC, TC_{eff}, that is the combination of the TCs of the two channels, defined as:

$$
TC_{eff} = \frac{1}{\frac{1}{TC_{1,n}} - \frac{1}{TC_{1,p}}} \tag{4.18}
$$

Figure 4.7 shows the normalized TC_{eff} for a Dual-RC system where $TC_{1,n}$ is normalized to unity, and $TC_{1,p}$ is varied. This shows that both the relative sign and the magnitude of TCs are essential in determining the nature of the transfer. TC_{eff} is always smaller than unity if resistors with TCs of different signs are used ($TC_{1,p} < -TC_{1,n}$). This reduction implies an accuracy improvement compared to using one of the channels as a dedicated temperature sensor. TC_{eff} reduces to half when the magnitudes are equal ($TC_{1,p} = -TC_{1,n}$). The smallest normalized TC_{eff} is 0.5, and it represents the maximum gain in accuracy.

On the other hand, if resistors of the same sign are used ($TC_{1,p} > TC_{1,n}$), the TC_{eff} is always larger than unity. This increase implies that the frequency error is degraded compared to using one of the channels as a dedicated

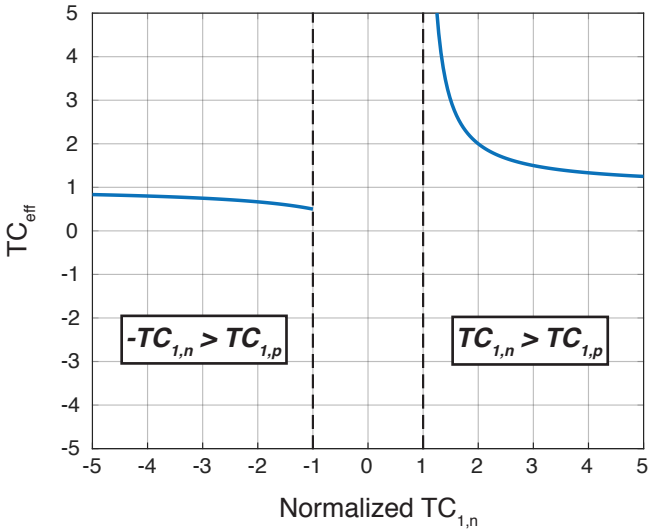


Figure 4.7: Normalized effective TC for a Dual-RC system

temperature sensor. When the two TCs have the same sign and magnitude, the TC_{eff} is infinity, meaning that the frequency error is infinite and the system does not work.

The analysis in this section provides guidelines for the choice of resistor types for the Dual-RC architecture. Ideally, the designer should choose two resistors with TCs of opposite sign and equal magnitude. If resistors with complementary TCs are unavailable, choosing the traditional temperature-compensated architecture (discussed in chapter 5) is more beneficial.

4.1.5. Dynamic Analysis

The linearized FLL is a single-pole system due to the dominant pole introduced by the digital loop accumulator. The poles of the PDDSMs and DCO are ignored since they are much higher in frequency than the dominant pole.

Using 4.16, the readout of the system can be simplified to a single gain element as in:

$$K_{f,eff} = K_{f,p}a_{1,p} - K_{f,n}a_{1,n} \tag{4.19}$$

Their frequency-dependent counterparts can replace $\Delta T_{\{p,n\}}$ and Δf to assess the dynamic behavior of error sources. $n_{T,\{p,n\}}(z)$ is the noise injected by the WB readouts at the output of the polynomials, and $n_f(z)$ is the noise

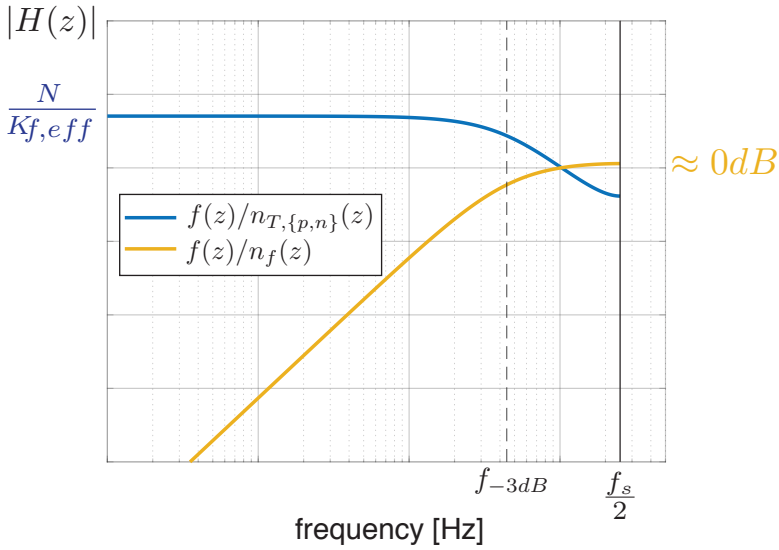


Figure 4.8: Frequency response of noise transfer functions' magnitude for the Dual-RC FLL

injected by the DCO. The transfer function for noise originating from the two WBs will both have the same low-pass characteristic:

$$\frac{f(z)}{n_{T,\{p,n\}}(z)} = \frac{K_d K_{DCO} z^{-1}}{1 + \left(\frac{K_{f,eff} K_d K_{DCO}}{N} - 1\right) z^{-1}} \tag{4.20}$$

On the other hand, the transfer function for noise originating from the DCO will have a high-pass characteristic as:

$$\frac{f(z)}{n_f(z)} = \frac{1 - z^{-1}}{1 + \left(\frac{K_{f,eff} K_d K_{DCO}}{N} - 1\right) z^{-1}} \tag{4.21}$$

Figure 4.8 shows the frequency response of the magnitude of the noise transfer functions 4.20 and 4.21. The bandwidth of the transfer functions at f_{-3dB} is defined by the factor $K_{f,eff}K_dK_{DCO}/N$. While having the digitally programmable gain factor, K_d, allows for the bandwidth to be easily set after production, K_T appearing in this expression creates an issue similar to the inaccuracy transfer analyzed in the previous section. If the two channels have TCs with the same sign, the resulting $K_{f,eff}$ is smaller, and to obtain the same bandwidth, K_d needs to be increased. Additionally, the noise originating from sources before the loop accumulator will be attenuated less to

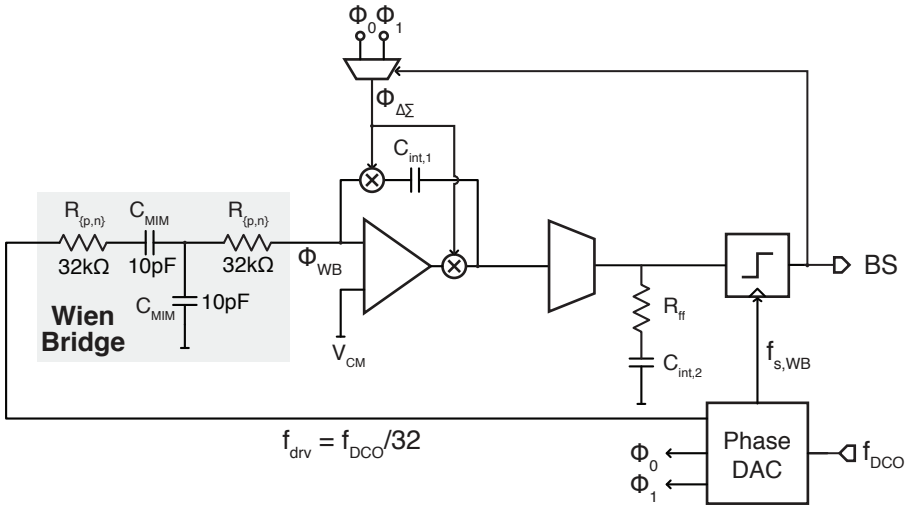


Figure 4.9: The simplified single-ended schematic diagram of the Wien Bridge and the Dual-RC Phase-Domain DSM readout

the output frequency. Instead, if they have TCs with opposite signs, the resulting $K_{f,eff}$ is higher, and the noise originating from sources before the loop accumulator is attenuated more to the output. It can also be observed that having equal signs with the same magnitude sets $K_{f,eff} = 0$, effectively opening the loop and causing the system not to function.

4.2. Circuit Design

4.2.1. Wien Bridge Readout: Phase-Domain DSM

The first step in building a digital FLL is to digitize the DCO frequency referenced to the time constant of an on-chip RC filter with very high resolution. On the other hand, since the speed of temperature compensation depends on the desired tracking speed of thermal transients, the bandwidth required is usually in the order of 10s of Hz. A delta-sigma architecture is the best choice to tackle the combined requirements of such a system. Thus, in the Dual-RC FLL, a phase-domain delta-sigma modulator (PDDSM) is used for digitizing the DCO frequency error.

Figure 4.10 shows the simplified single-ended schematic diagram of the PDDSM [2]. The PDDSM comprises the WB and a synchronous demodulator embedded in the first-stage integrator of a continuous-time DSM. The WBs in this design feature the two types of resistors with $R_{\{p,n\}} = 32\ k\Omega$ and

MIM capacitors with values C_{MIM} = 10pF. The WB is driven at $f_{drv} = f_{ref}/N$ = 437.5 kHz. Its output voltage is converted to a current through its shunt resistor and flows into the low-impedance input of the first-stage integrator. A chopper demodulator is driven by one of $\phi_{\{0,1\}}$ = 90±22.5°, selected by the bitstream output. The demodulators' output is then integrated into the capacitors of the first stage. In steady-state, gain in the delta-sigma loop filter drives the output of the demodulator to zero on average. The bitstream encodes the phase shift of ϕ_{WB} by ensuring that $\phi_{\Delta\Sigma}$ has an average 90° phase shift.

The DSM loop filter is comprised of two integrating stages. The first stage is built with a two-stage operational amplifier and features the large integration capacitor $C_{int,1}$ = 200 pF. The unity gain frequency of the first stage is determined by its integration capacitor and $R_{\{p,n\}}$ and is about 25kHz. Chopper switches that implement the synchronous demodulator serve the additional purpose of modulating the offset and 1/f noise of the first-stage integrator to f_{drive}. The loop filter of the DSM filters the up-modulated offset and 1/f noise. The second stage features a degenerated transconductance amplifier with g_m = 2 µS and $C_{int,2}$ = 8 pF, with a unity-gain frequency of 40 kHz. Additionally, the second stage implements the feed-forward coefficient required for DSM loop stability via a series resistor R_{ff} = 250 kΩ. For each WB's readout, degeneration and feed-forward resistors are built from the same material as the WB resistor to improve the robustness of the DSM loop filter to process spread and temperature [2].

The references $\phi_{\{0,1\}}$ are derived from the high-speed DCO clock in the phase DAC, where it is divided down to the WB center frequency in cascaded ripple dividers. Then, a delay chain generates the phase references. The DCO clock finally resynchronizes the references in a final layer of FFs to reduce the delay imparted by the logic. In this design, the phase references were chosen to be $\phi_{\{0,1\}}$ = 90±22.5° to cover the significant phase shift expected from the high-TC resistor. Such a broad phase range for the references causes additional cosine nonlinearity, which can reduce trimming accuracy. Thus, this is handled in the digital domain via correction polynomials, as explained in section 4.3.

4.2.2. Digitally Controlled Oscillator

The digitally controlled oscillator provides the digitally programmable frequency reference output. It closes the feedback loop by driving the two channels and clocks all other digital components of the system. Since the DCO operates within a digital feedback loop, it poses unique design constraints regarding its control range, resolution, and jitter.

Firstly, the DCO must be designed to provide the expected frequency over process, voltage, and temperature variations. Process variations are static after fabrication but have the most significant effect on the center frequency of the DCO, causing a ±50% shift. On the other hand, voltage and temperature variations are dynamic during operation but cause a ±5% shift.

Moreover, the DCO must have a high enough resolution to maintain the fidelity of the output frequency. Since the output frequencies of a DCO are quantized, if the system noise is low enough, the DCO will toggle between two quantization levels after settling. This switching will manifest as output jitter. Worse yet, if large enough, it drives the front-end readout to nonlinearity, which can cause accuracy degradation.

Finally, DCO resolution needs to be achieved while the transfer function is monotonic. Analysis in section 4.1.5 treats the DCO as a gain block with a specific digital-to-frequency gain corresponding to its LSB. If the output is not monotonic, it is implied that the DCO gain will change sign, and the feedback polarity will be inverted. In cases where the nonmonotonicity is constrained to a few codes that are in amplitude lower than the noise level, this might only degrade the noise behavior. However, if the nonmonotonicity affects many codes, it can cause the FLL to oscillate and create tones that ruin its operation entirely.

Given these design constraints, a coarse/fine architecture was chosen. Figure 4.10 shows the schematic diagram of the DCO. At its core, a 9-stage single-ended current-starved ring oscillator provides the output frequency after level shifters recover the signal to supply levels. The current is provided via two paths, the coarse DAC path providing an i_{cs} and the fine DAC path providing an i_{fn}. These currents are summed and steered into the ring oscillator core to tune the output frequency.

The coarse DAC is a 5-bit current steering DAC, providing a 180 µA variation on i_{cs} that covers a ±50% static variation on the DCO center frequency. The coarse DAC LSB corresponds to 125 kHz, defining the residual frequency variation that needs to be covered by the phase DAC range. The fine DAC is a 13-bit R-2R DAC, providing a 30 µA variation on i_{fn} that covers a ±7.5% variation on the DCO center frequency. The fine DAC LSB is 120 Hz in frequency, corresponding to 17 ppm. This current segmentation allows an extensive tuning range to be covered, and only the fine DAC needs to be designed to be monotonic.

The fine DAC is further segmented into a 5-bit unary and 8-bit binary R-2R structure to ensure its monotonicity. It uses highly resistive polysilicon unit resistors and standard DAC layout techniques to reduce the mismatch

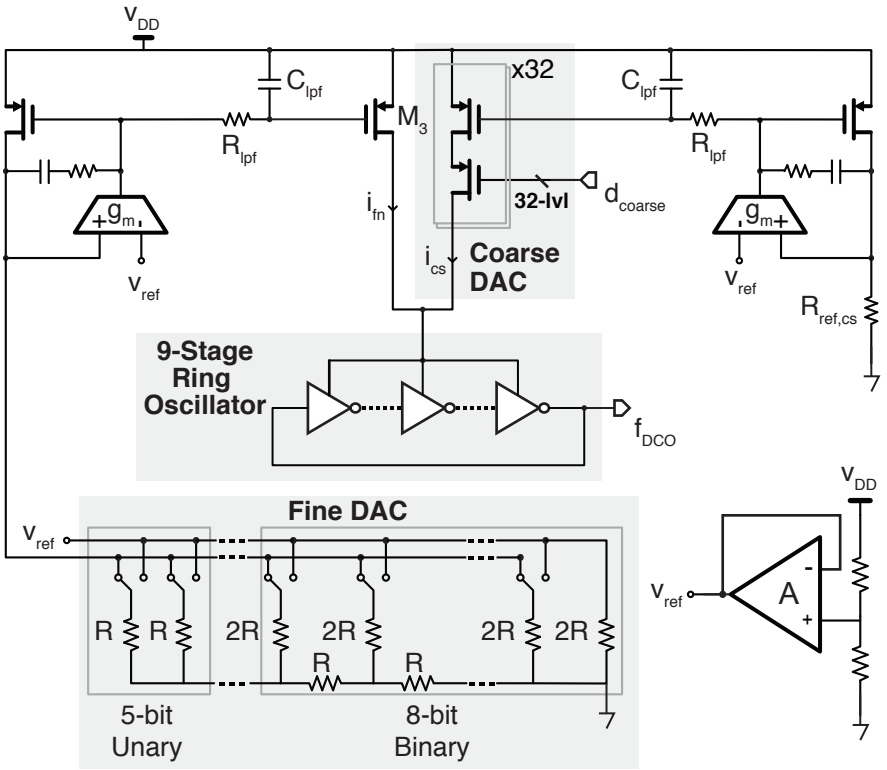


Figure 4.10: The schematic diagram of the digitally controlled oscillator

between unit resistors.

For each coarse and fine path, the references for the current steering devices $M_{1,3}$ are generated via an on-chip resistive divider that generates v_{ref}. For the coarse path, v_{ref} is the reference for the regulated current mirror composed by the g_m cell and M_4. Due to negative feedback, v_{ref} is developed across R_{ref} and is converted into a current that generates the bias voltage for the coarse DAC current cells. The reference voltage is directly supplied to one of the current outputs via a two-stage opamp buffer for the fine DAC. The regulated current mirrors also feature a single-stage RC low-pass filter with highly resistive polysilicon R_{lpf} and MOS capacitors C_{lpf}. The wide-band noise originating from v_{ref} and the regulated current mirrors is filtered by this filter. However, it is also undesirable for this RC filter to interact with the FLL loop, so it was designed with a cut-off frequency of 4 kHz, approximately ten times higher than the update rate of the loop.

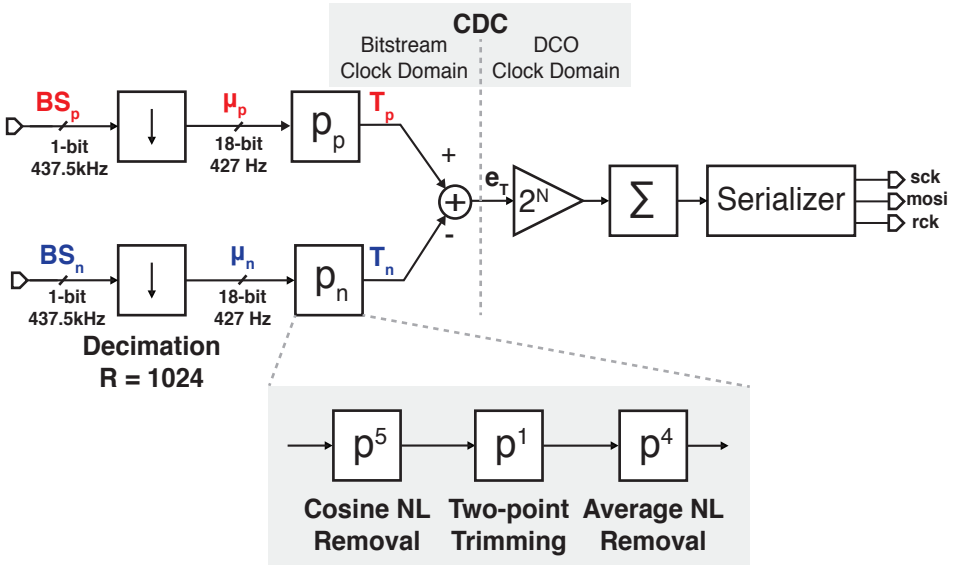


Figure 4.11: Block diagram of the Dual-RC digital comprising the temperature compensation, loop accumulator, and the serializer driving the DCO frequency word shift register

4.3. Digital Temperature Compensation & Loop Filter

The digital temperature compensation & loop filter glues the front-end Wien Bridge and its PDDSM readout to the back-end DCO of the system. It processes the outputs of the two PDDSMs to produce the temperature-independent loop control signal, e_T, accumulates it, and drives the DCO.

The first stage in digital processing is decimation. 2-stage cascaded integrate-comb (CIC) filters with a decimation factor of R = 1024 and differential delay of M = 2 convert the bitstream outputs to high-resolution 22-bit words. The 500 kHz sample rate of the input bitstreams is divided by 1024, which brings the rate of the following blocks to 427 Hz. CIC output words are quantized to 18 bits before polynomial processing, ensuring that multipliers in the subsequent polynomial processing can be limited to 36 bits.

Polynomials for each channel are implemented as a cascade of three stages. The first, 5^{th}-order polynomial, p^5, removes cosine nonlinearity originating from the PDDSM. Its implementation takes advantage of the oddness property of the cosine nonlinearity and only uses six multipliers and two adders. At the output of the first polynomial, the bitstream average is converted to phase, and the remaining two stages implement the phase-to-temperature conversion. The second, 1^{st}-order polynomial, p^1, implements

the two-point trimming directly on the WB phase, using one multiplier and adder. Finally, the average systematic nonlinearity of the trimmed WB phase is removed by a 4^{th}-order polynomial, p^4. It is the most expensive stage in terms of area, using seven multipliers and four adders.

Coefficients of p^5 are constant among all samples and are derived from fitting a polynomial to the analytical results obtained in chapter 3. On the other hand, coefficients of p^1 are unique per sample and are extracted after characterizing each sample at two unique temperature points. Coefficients of p^4 are the same for all samples but are extracted by characterizing a statistically significant set of devices.

After polynomial correction, a simple subtraction of the outputs of the two channels achieves the temperature-independent frequency error word, e_T. Then, a 24-bit accumulator running at the loop sample rate accumulates e_T, and its 13-bit MSBs drive the DCO to close the feedback loop. A programmable bit shifter preceding the accumulator allows for the modulation of the FLL loop gain, implementing K_d defined in Section 4.1.5. This programmability allows for the flexible positioning of the dominant pole of the FLL loop.

While the DCO and bitstream clocks are synchronous, their phase relationship is not always well-defined. A reliable transfer is required between the two clock domains, achieved by the clock-domain-crossing (CDC) block. The CDC captures the decimation pulse transition at 427 Hz by the 7 MHz DCO clock and converts it into a one-cycle-wide pulse, clocking the rest of the system.

The DCO provides an independent SPI interface to program its 13-bit fine control word. This interface serializes the 13-bit word into a 1-bit data bus, mosi, and sends it with a clock, sclk, through the SPI bus to the chip shift register. After the transfer, a register clock, rclk, latches the shift register contents to the DCO input.

The digital temperature compensation & loop filter block complete all operations in less than a single-cycle delay of the loop sample rate, $f_{DCO}/16/1024$. It imparts about 25 cycles of the DCO period as an excess delay due to the CDC and serialization. This extra delay does not significantly impact operation due to the substantial ratio between the reference frequency and the loop rate. However, it limits the lowest decimation rate achievable and, hence, the highest frequency that can be chosen for the loop bandwidth.

4.4. Experimental Results

The dual-RC frequency reference was implemented in a TSMC 0.18 µm standard CMOS process [1]. The chip has an active area of 1.59 mm^2, 0.62 mm^2 for each WB and PDDSM combination, and 0.35 mm^2 for the DCO.

Each PDDSM, including the WB, driver, and phase DAC, dissipates 162 µW from a 1.8 V supply, and the DCO dissipates 450 µW from a separate 1.8 V supply. The total analog power consumption of the frequency reference is 775 µW.

Twelve samples of the frequency reference were packaged in ceramic DIL28 packages before characterization. These samples all originate from a single production batch.

Since the nonlinear behavior of the two channels after production cannot be accurately extracted from simulations, the design requires some flexibility in determining trimming polynomial orders. Thus, the digital components of the frequency reference were realized off-chip in an FPGA evaluation board. This also allows for the flexible refinement of the digital blocks after tape out. Additionally, the abundance of DSP resources on the FPGA and their high speed allowed for a combinational implementation for most heavy-duty signal processing blocks. The power estimation from post-synthesis simulations in the used 180nm process showed that the digital logic implemented would consume approximately 50 µW, the CIC decimators dissipating the most significant power of 20 µW each while running at 437.5 kHz.

The characterization of the devices was carried out in two steps. First, the analog sub-blocks of the system were characterized in an open loop configuration, with an externally provided frequency reference and over temperature. Temperature compensation polynomials were derived from the characteristics derived from the open-loop characterization. Finally, these polynomials were loaded into the digital engine, and the dual-RC frequency reference was characterized again in a closed-loop configuration. During this step, another closed-loop trimming operation was applied to increase the accuracy of the frequency reference.

4.4.1. Open-Loop Block Characterization

The open-loop characterization involves an extensive temperature sweep of the WBs and their temperature and frequency sensitivity characterization. Characterization of the DCO range ensures its correct operation in the FLL.

Wien Bridge and PDDSM Characterization

In their open-loop characterization, the PDDSMs were driven by an external frequency reference at the target reference frequency of 7 MHz. This refer-

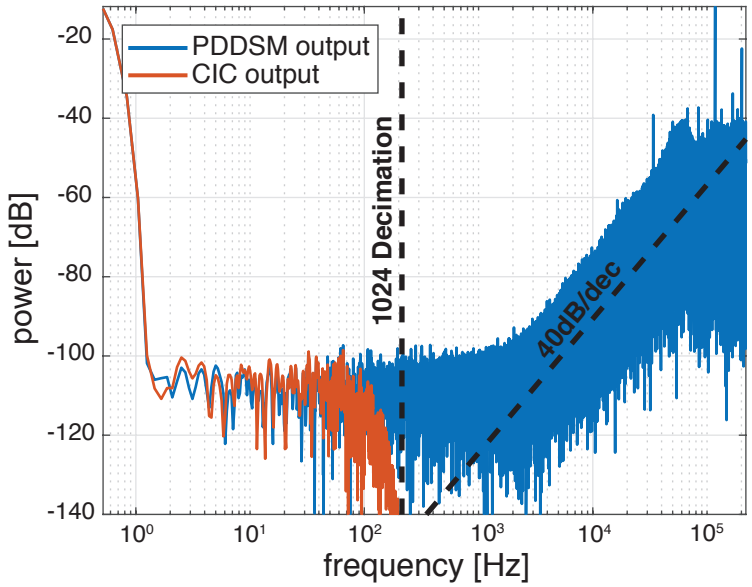


Figure 4.12: Power spectral density of the PDDSM output before and after the decimation filter

ence frequency gets divided through the phase DAC on the chip and generates the WB drive, phase references, and the DSM sampling frequency at 437.5 kHz. Then, the bitstream was captured for 2^{20} samples and externally decimated to arrive at the average bitstream value. Figure 4.12 shows the power spectral density of the PDDSM output at room temperature before and after the decimation filter. As expected, the quantization noise is 2^{nd}-order shaped with a 40 dB/dec slope and is effectively filtered out after the CIC. The resulting decimated word has a 0.013 m° resolution at the loop sample rate of 427 Hz.

WB phase characterization is achieved by recording the decimated bitstream outputs while the independent variables of frequency and temperature were swept across their operating range. Bitstream averages are subjected to the cosine nonlinearity removal polynomial as discussed in Section 3.4.1. WBs were stabilized at a fixed temperature of 27 °C, and the drive frequency was varied from 6.5 MHz to 8 MHz to characterize the phase vs. frequency characteristics. Then, the chips were placed in a temperature-controlled oven, driven by a fixed frequency of 7 MHz. The temperature was varied from -45 °C to 85 °C with steps of 10 °C to characterize the phase vs. temperature characteristics. Temperature stability was achieved

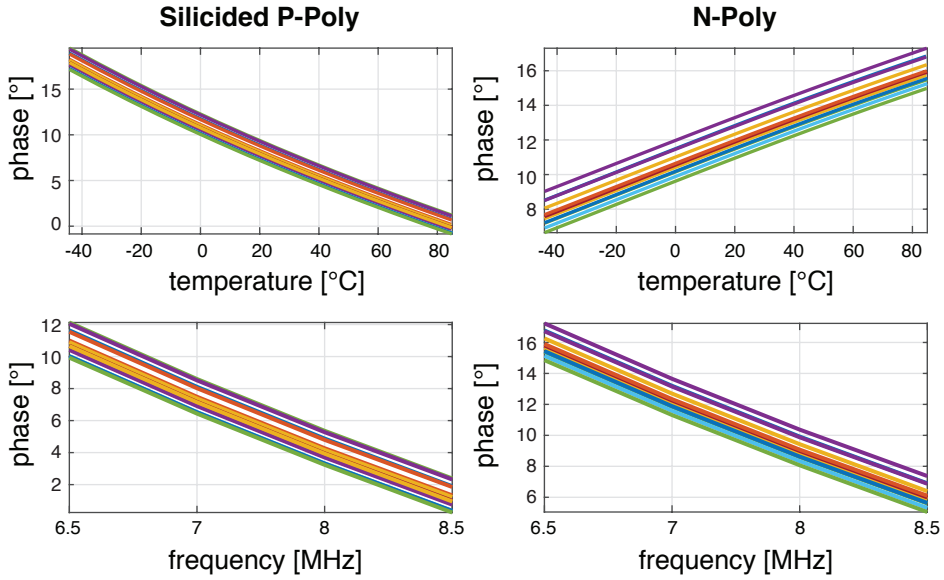


Figure 4.13: Phase versus temperature and phase versus frequency characteristics of the WBs with the silicided p-poly and n-poly resistors

by mounting the chips inside an aluminum block that behaves as a thermal capacitor, filtering the fast temperature variation caused by the oven's heating system. Additionally, the characterization temperature was measured via a Pt-100 reference resistor mounted inside the aluminum block close to the chips. Four samples were loaded onto the characterization board for each oven run, and the cycle was repeated three times to characterize the 12 samples.

Figure 4.13 shows the phase vs. temperature and phase vs. frequency characteristics of the WBs with silicided p-poly and unsilicided n-poly resistors for all 12 characterized samples. The sp-poly WB shows an average -18° phase shift across the temperature range, whereas the n-poly WB shows an 8° phase shift, consistent with their TCs. Both WBs display a 9° phase shift over the 1.5 MHz frequency range, corresponding to a 93.75 kHz range for f_{drv}.

DCO Characterization

Section 4.2.2 presents design constraints for a DCO that operates in a control loop, such as the Dual-RC FLL. In particular, it was stressed that the monotonicity of the DCO's digital-to-frequency transfer characteristic is paramount in maintaining loop stability. To verify that the designed DCO satisfies this

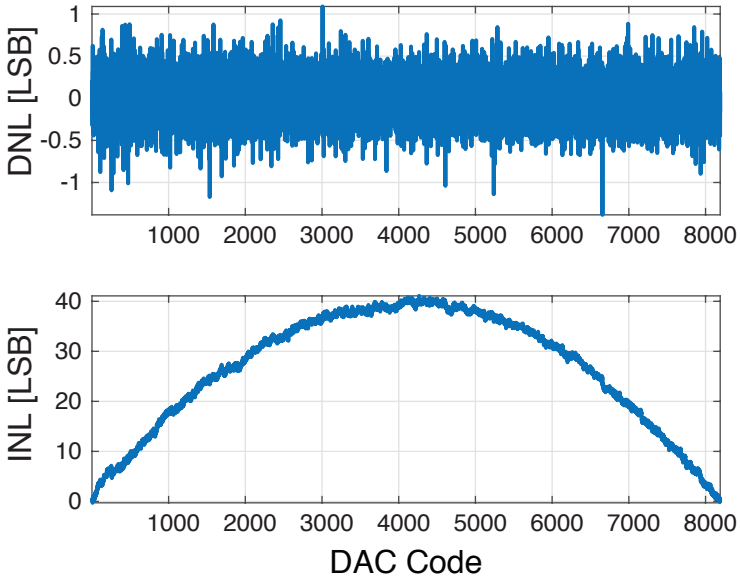


Figure 4.14: DNL and INL of the DCO output frequency with the Fine DAC control word

property, the DCO was characterized in open-loop by measuring the frequency output as the digital frequency control word was swept through the entire control range of the fine DAC.

Figure 4.14 shows that the DCO achieves 12-bit DNL. In the DNL plot, a few codes can be observed where the magnitude of the DNL is greater than one. The digital-to-frequency transfer characteristic is nonmonotonic around these codes and can create problems for the larger FLL loop. However, these codes are few and far between, and the inherent noise of the DCO results in a jitter larger than a single code. Thus, these scarce nonmonotonic incursions are dithered and do not cause a severe problem for the current design but should ideally be eliminated.

The INL plot in Figure 4.14 reveals that the DCO transfer characteristic suffers from a large nonlinearity of 40 LSBs. Compared to the 12-bit DNL, the INL is only 8-bit, showing 2^{nd}-order nonlinear behavior caused by the finite output impedances of the regulated current mirror formed by M_1, M_2, and the g_m cell on the fine DAC path. However, this nonlinearity is not detrimental to the loop operation. In the steady state, there will be a very narrow range along this nonlinear curve where the reference will operate, which can be considered linear. The only impact is a slight variation in the loop bandwidth depending on the operation point, which is negligible.

4.4.2. Derivation of the Trimming Polynomials

After measuring phase-to-temperature characteristics for the two WBs in Section 4.4.1, the coefficients for the temperature compensation polynomials are derived. The two sets of polynomial coefficients p^1 and p^4 are derived concurrently to achieve the bitstream average to temperature conversion with as high an accuracy as possible.

A first-order polynomial is fit to the inverse of the phase-to-temperature characteristic to derive p^1. Only the -35 °C and 75 °C measurement points are used for this operation. A unique p^1 for each channel is determined for each sample, constituting the two-point trim step of the temperature compensation.

A trimming residue can be defined from the difference between the temperature output of p^1 and the measured value. This trimming residue is nonlinear and crosses zero at the two-point trim points. The nonlinear residue is averaged over all samples for the sp-poly and n-poly WBs, and a 4^{th}-order polynomial is used to fit this average residue. These two polynomials are the same for all samples and constitute the average systematic nonlinearity removal step of temperature compensation.

Once both p^1 and p^4 are derived, it is possible to estimate the final inaccuracy of the frequency reference. A final residue for samples can be defined as the difference between the output of p^4 and the measured temperature. Then, using the analysis in Section 4.1.4, the residue can be transferred to a frequency error estimate. The expected contribution of each channel to the total closed-loop frequency inaccuracy is derived via 4.17. Figure 4.15 shows that the WB with silicided p-poly resistors contributes ±50 ppm inaccuracy, whereas the WB with the n-poly resistors contributes ±140 ppm. By combining the two results, it can be estimated that the frequency reference inaccuracy will be ±150 ppm, which is dominated by the inaccuracy of the WB with n-poly resistors.

4.4.3. Final Inaccuracy Characterization

The final step after deriving the trimming polynomials is to characterize the inaccuracy of the frequency reference. Another temperature sweep is conducted with the FLL placed in the closed-loop mode and the derived polynomials loaded into the polynomial engine. The samples are then characterized in groups of four chips per run, and the temperature is swept from -45 °C to 85 °C with steps of 10 °C.

Figure 4.16 shows that frequency reference achieves ±250 ppm frequency inaccuracy for 12 samples. The box method yields a 3.8 ppm/°C residual temperature coefficient.

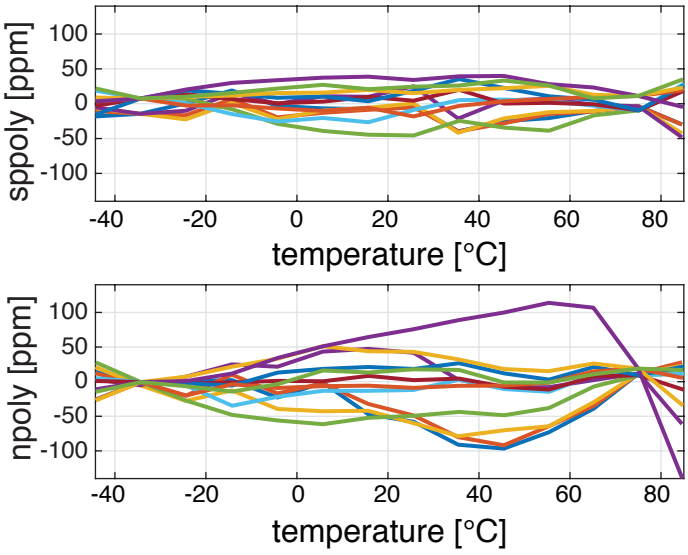


Figure 4.15: Frequency inaccuracy contribution of the WBs with silicided p-poly and n-poly resistors

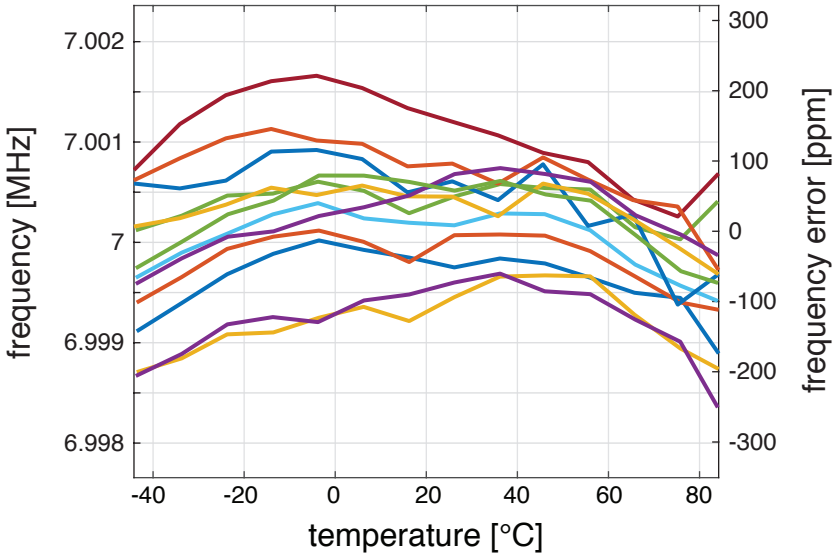


Figure 4.16: Inaccuracy of the open-loop temperature compensated frequency reference

The final reference inaccuracy is significantly worse than the estimated inaccuracy in the previous section from the open-loop characterization residuals. Observing the trimming temperatures -35 °C and 75 °C reveals high frequency-inaccuracy spread at these points. The inaccuracy at the trimming points should be zero if the calibration procedure is ideally executed.

Several factors in the open-loop trimming scheme adversely affect the inaccuracy. The first factor is that the open-loop characterization operating condition is inevitably different than the closed-loop frequency reference characterization condition. In the open-loop characterization setup, the DCO is turned off, the fine DAC SPI bus is quiet, and most of the DCO to phase DAC clock channel is inactive. Electrical activity on specific components of the device will couple to the WB through parasitic channels and corrupt the measurement. The second factor is related to the accuracy of the temperature measurement during characterization. Chips are characterized in open-loop groups of four and are referenced to a single temperature reference, the Pt-100. In each run, the actual temperature reached for a characterization point differs for each group of four chips. This difference necessitates that the reference points and readouts are interpolated for the average nonlinearity extraction for all devices, which introduces additional errors in the measurement. Even more significantly, temperature measurement accuracy relies on high temperature-stability. Although an aluminum block is used to increase stability, it is still challenging to get to m°C levels in thermal stability with a temperature-controlled oven. Such a thermal drift of 10 m°C coupled with the silicided p-poly TC_1 results in a 30 ppm error in the measurement, leading to a worse spread at -35 °C as the stability of the oven is worse when it is cooling.

Closed-loop Trimming

A better approach is to trim the reference in a closed-loop configuration. This ensures that the electrical operating condition of the system during trimming is its final operating condition. Also, since the target frequency is the same across all temperatures, the closed-loop approach does not require accurate temperature measurement.

Closed-loop trimming starts with the polynomial engine initialized with open-loop derived polynomials. The a_0 coefficient of the trimming polynomial p^1 is adjusted at the trimming temperatures of -35 °C and 75 °C to bring the output to 7 MHz. Finally, the devices are characterized again across temperature to assess their final inaccuracy.

Figure 4.17 shows that closed-loop trimmed frequency reference achieves ±165 ppm frequency inaccuracy for eight samples. The box method results

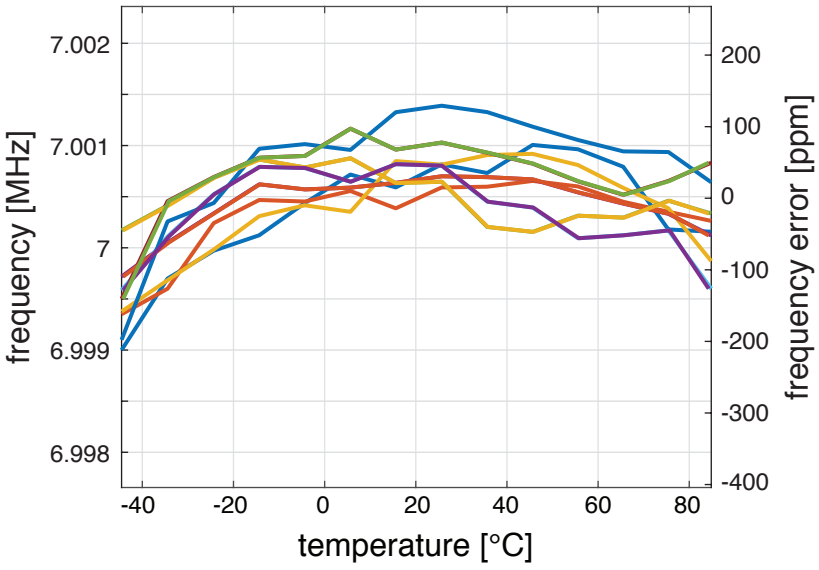


Figure 4.17: Inaccuracy of the closed-loop temperature compensated frequency reference

in a 2.5 ppm/°C residual temperature coefficient.

While the closed-loop trimming improves the accuracy as expected, it is clear that there is still room for further improvement. About 150 ppm of frequency spread remains at the trimming temperatures. In addition, there is still some residual curvature in the frequency error, which implies an erroneous estimation of p^4 coefficients. Since the nonlinearity estimation was not conducted in a closed-loop configuration, it is expected that the issues of open-loop trimming are still observable in this polynomial. However, the improvement in accuracy with closed-loop trimming shows that it is the best way to trim frequency references.

Supply Sensitivity

In addition to temperature, supply voltage variations can degrade the accuracy of the frequency reference. The supply voltages for the DCO and the PDDSMs varied between 1.7 V and 2 V to characterize the sensitivity of the output frequency to supply variation. Figure 4.18 shows the inaccuracy of the compensated frequency reference for 12 samples with power supply variation.

This measurement also includes the devices' spread at the nominal power supply, which is comparable to the total supply sensitivity. The variation from the nominal value for each sample is considered to remove the baseline

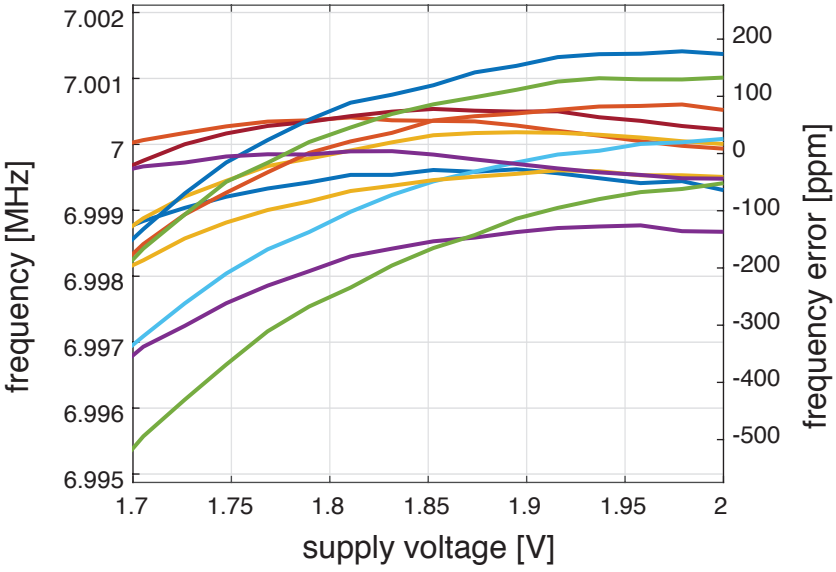


Figure 4.18: Inaccuracy of the compensated frequency reference with power supply variation

spread. The worst-case sample shows 540 ppm_{pp} variation across the 1.7 V to 2.0 V range, resulting in a 0.18 %/V supply sensitivity.

The residual supply sensitivity originates from the supply dependency of the WB readout. Since the DCO follows the loop integrators, its static error sources, such as supply, are nulled. On the other hand, for the WB, the drivers and the demodulating chopper switches are driven at supply voltage. Variations in the supply voltage vary the impedance of these switches, changing both the driving and the loading impedance seen by the WB, altering its phase shift.

4.4.4. Noise & Long-Term Stability

Noise measurements on the frequency reference are conducted to characterize its short-term and long-term noise behavior. A high-speed oscilloscope is used to analyze the short-term noise behavior. Long-term noise behavior is captured via an Allan Deviation measurement with a gapless frequency counter.

Figure 4.19 shows the period jitter measurement for both the open-loop DCO and the closed-loop temperature-compensated Dual-RC frequency reference. The DCO was brought to the reference frequency for the open loop case by adjusting the 13-bit frequency control word to obtain the expected

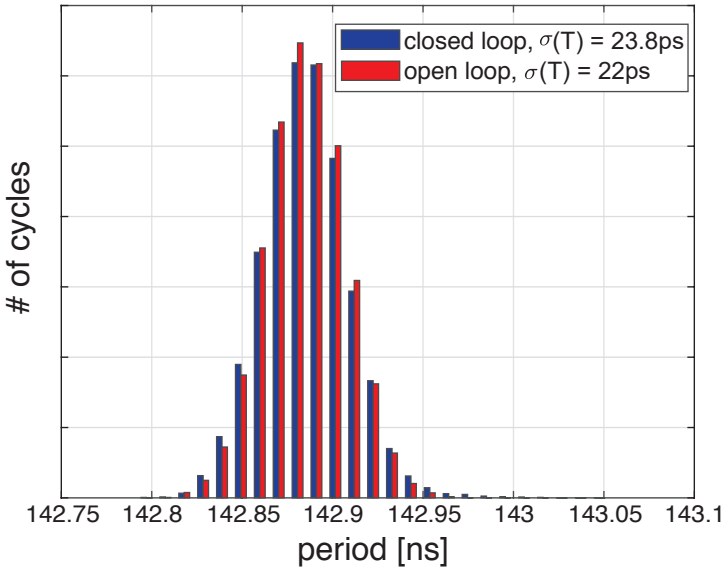


Figure 4.19: Period jitter for the closed-loop Dual RC frequency reference and the open-loop DCO

reference frequency. The open-loop period jitter of the DCO is 22 ps, and the closed-loop period jitter is 23.8 ps. This increase from open-loop to close-loop points to a 9 ps excess jitter introduced by the FLL loop.

Figure 4.20 shows the results of an Allan Deviation measurement conducted on the open-loop DCO and the closed-loop frequency reference. The DCO is programmed in the same manner as the period jitter measurement for the open-loop case. The open-loop DCO is limited to a noise floor of 4 ppm. Closing the loop for the trimmed frequency reference improves the noise floor to 330 ppb in a 3s measurement time, significantly improving the open-loop case.

Figure 4.20 reveals that the Allan Deviation is worse for short measurement times. The frequency quantization effect of the DCO is not considered in the open-loop case, but this effect can only account for 2.5 ps jitter due to the 17 ppm step size of the DCO. Careful observation of the DCO noise transfer function in Section 4.1.5 reveals that the DCO noise shows some peaking in its transfer function. This peaking is due to the proximity of the loop dominant pole (around 50 Hz for this design) to the loop Nyquist frequency and the non-dominant pole caused by the DCO filter. An experiment where the loop bandwidth was varied via the K_d programmable digital gain

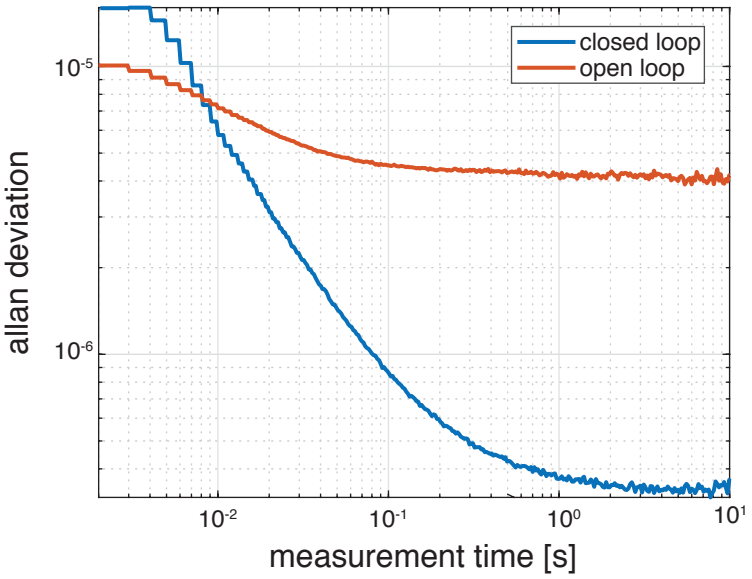


Figure 4.20: Allan deviation measurement for the closed-loop Dual RC frequency reference and the open-loop DCO

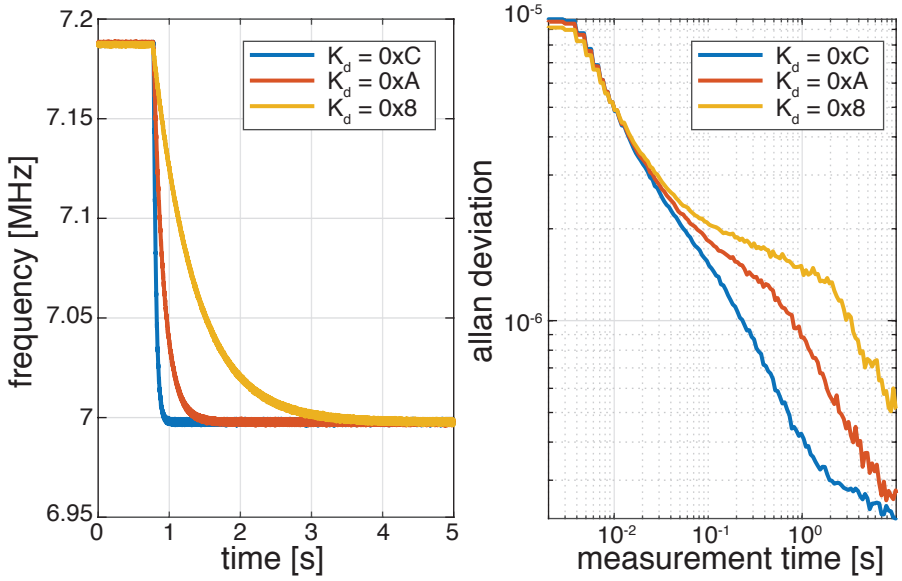


Figure 4.21: Transient settling measurement of the reference output frequency and Allan deviation measurement conducted after settling

was devised to verify this claim. Figure 4.21 shows the transient settling measurement of the output frequency after the loop is enabled at around 0.75 s, and an Allan deviation measurement is conducted after the loop has settled. For all cases, the loop shows first-order settling behavior due to the single low-frequency pole of the accumulator. Each of the three values of K_d corresponds to a factor-of-four decrease in the dominant loop frequency. For K_d = 0xC, the loop settles the fastest, in less than 100 ms, corresponding to the highest loop dominant pole frequency. In this regime, the ADEV floor is reached in the shortest measurement time, but inspection of the short-term noise shows the highest peaking. Conversely, for K_d = 0x8, the loop settling is the slowest, taking almost 4 s. In this case, the ADEV floor is not reached for a 10 s measurement time, but the short measurement times show the lowest noise. These observations confirm that the loop noise behavior is closely related to the loop bandwidth setting and can degrade short-term noise behavior by amplifying the noise of the DCO.

4.5. Conclusion

In this chapter, the Dual-RC frequency reference architecture has been proposed. The architecture combines the phase shifts of two WBs made from resistors with TCs of opposite signs to obtain a temperature-independent frequency error signal. The frequency of an inaccurate but digitally controllable oscillator is locked to this frequency error signal. Digitization of the on-chip RC time constants with high resolution via phase-domain DSMs allows further digital processing to remove the nonlinear effects that traditionally limit inaccuracy. The experimental 7 MHz reference implemented in a standard CMOS process achieves ±250 ppm inaccuracy from -45 °C to 85 °C for 12 samples when subject to an open-loop trim. Inaccuracy improves to ±162.5 ppm for eight samples when a secondary closed-loop trim is applied to the references.

Table 4.1 shows the performance summary of the Dual-RC frequency reference compared to other state-of-the-art RC frequency references. The proposed reference shows a 5.5x improvement in inaccuracy to the best-reported RC reference over a more significant number of samples. This improvement is achieved by using high-resolution and accurate digital conversion in the front end, allowing further digital processing of the on-chip time constants. Moreover, the proposed two-channel RC filter architecture presents an inherent improvement in accuracy compared to traditional temperature-compensated systems. Compared to the best-reported Allan Deviation floor [6], the dual-RC frequency reference shows a 12x improve-

	This Work	[4],2018	[5],2016	[6],2014
Process [nm]	180	180	180	65
Frequency [Hz]	7M	24M	3k	32.768K
Power [W]	775µ	200µ	4n	0.19µ
# of Trim Points	2 + Batch	2	2	2
Inaccuracy [ppm]	±250 | ±162.5	1150	±897	±2522
Temp. Range [°C]	-45 to 85	-40 to 150	-25 to 85	-20 to 90
# of Samples	12 | 8	>200	1	5
ADEV [ppm]	0.33	-	63	4

Table 4.1: Performance summary and comparison to prior art

ment. The Allan deviation floor represents excellent long-term stability. Adopting a high-resolution phase readout optimized for low flicker noise as the digitizer for the on-chip time constant brings the Allan deviation floor to the drift of the passives. However, the performance is still 10x worse than the typical 10 ppb floor of quartz-crystal oscillators.

References

[1] C. Gürleyük, L. Pedalà, S. Pan, F. Sebastiano, and K. A. A. Makinwa, *A CMOS Dual-RC Frequency Reference With ±200-ppm Inaccuracy From −45 °C to 85 °C,* IEEE Journal of Solid-State Circuits **53**, 3386 (2018).

[2] S. Pan, Y. Luo, S. Heidary Shalmany, and K. A. A. Makinwa, *A Resistor-Based Temperature Sensor With a 0.13 pJ · K2 Resolution FoM,* IEEE Journal of Solid-State Circuits **53**, 164 (2018).

[3] S. Pan, C. Gürleyük, M. F. Pimenta, and K. A. A. Makinwa, *10.3 A 0.12mm2 Wien-Bridge Temperature Sensor with 0.1°C (3σ) Inaccuracy from -40°C to 180°C,* in *2019 IEEE International Solid- State Circuits Conference - (ISSCC)* (2019) pp. 184–186.

[4] G. Zhang, K. Yayama, A. Katsushima, and T. Miki, *A 3.2 ppm/°C Second-Order Temperature Compensated CMOS On-Chip Oscillator Using Voltage Ratio Adjusting Technique,* IEEE Journal of Solid-State Circuits **53**, 1184 (2018).

[5] T. Jang, M. Choi, S. Jeong, S. Bang, D. Sylvester, and D. Blaauw, *5.8 A 4.7nW 13.8ppm/°C self-biased wakeup timer using a switched-resistor scheme,* in *2016 IEEE International Solid-State Circuits Conference (ISSCC)* (2016) pp. 102–103.

[6] D. Griffith, P. T. Roine, J. Murdock, and R. Smith, *17.8 A 190nw 33khz RC oscillator with ±0.21% temperature stability and 4ppm long-term stability,* 2014 IEEE International Solid-State Circuits Conference Digest of Technical Papers (ISSCC) (2014), 10.1109/isscc.2014.6757443.

5

A ±90 ppm CMOS RC Frequency Reference

Improving accuracy dictates minimizing the number of parameters degrading it. This approach leads to a design based on a *single* on-chip RC time constant. Thus, in this design [1], an RC filter with a low-TC resistor is used instead of the two-filter approach of the Dual-RC architecture. Temperature compensation is achieved via a separate, resistor-based Wheatstone bridge (WhB) temperature sensor, enabling a simpler digital temperature compensation scheme. A simple linear temperature compensation scheme achieves ±400 ppm inaccuracy from -45 °C to 85 °C after a 2-point trim. Inaccuracy improves to ±90 ppm with a nonlinear temperature compensation scheme based on a fixed 6^{th}-order correction polynomial.

5.1. Introduction

The Dual-RC architecture dictates the use of two types of resistors with complementary temperature coefficients (TCs) (section 4.1.4). Multiple kinds of resistors come at the cost of additional masks and processing steps. Moreover, the benefits of the Dual-RC scheme diminish quickly via any discrepancy in the magnitude of resistor TCs. Overall, the architecture's applicability is limited in processes that only feature a low-TC resistor.

Another problem with the Dual-RC architecture is the complexity of its measurement and trimming setup. This significantly increases characterization time, translating to infeasible post-production costs. Additionally, a temperature reference with high accuracy is required, which is not generally available in standard automated test equipment.

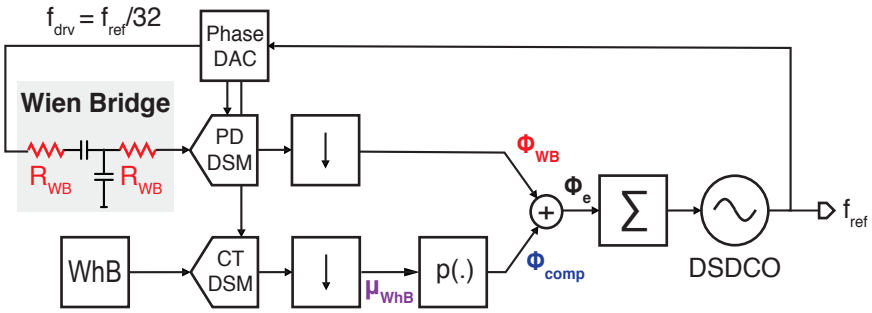


Figure 5.1: The block diagram of the proposed RC frequency reference architecture

5

Finally, both channels of the Dual-RC architecture required a polynomial engine to linearize their temperature dependencies. These engines involve high bit-width adders and multipliers. Since the reference is implemented in a 0.18μ process, the cost of the digital area is not negligible. While there might be potential to address this in the context of the Dual-RC system, the solution is not readily apparent. It is desirable to reduce the digital footprint as much as possible to increase the applicability of the frequency reference, which is only possible by reducing complexity.

In the rest of this chapter, a new architecture that addresses these shortcomings of the Dual-RC will be analyzed. Then, circuit improvements will be discussed. Finally, measurement results will be presented.

5.2. Architecture Proposal

Figure 5.1 shows the block diagram of the proposed improved RC frequency reference. It is a digital frequency-locked loop (FLL) in which the output frequency f_{ref} of a delta-sigma digitally-controlled oscillator (DSDCO) is locked to a frequency-dependent but temperature-independent phase shift, ϕ_e. ϕ_e is realized by compensating for the temperature dependency of the phase shift of an on-chip WB filter (ϕ_{WB}) when driven at $f_{drv} = f_{ref}/32$. The WB provides the RC time constant to which the DCO is locked. It is configured to have a center frequency at $f_{drv} = f_{ref}/N$, where N = 32 in this design.

$$f_{drv} = \frac{f_{ref}}{N} = \frac{1}{2\pi R_{WB} C} \tag{5.1}$$

The Phase DAC performs the frequency division by 32 and generates the reference signals to the phase-domain delta-sigma modulator (PDDSM), which digitizes ϕ_{WB}. A Wheatstone Bridge (WhB) temperature sensor determines the die temperature, and its output is digitized by a continuous-time delta-sigma modulator (CTDSM). Both modulator outputs are decimated to a lower frequency through two decimation filters. Then polynomial mapping is applied to the decimated WhB output to map it to ϕ_{WB}, and via simple addition, ϕ_e is obtained, achieving temperature compensation. The loop locks to ϕ_e, at which point the output will have locked to f_{ref} as intended.

The architecture is a traditional temperature-compensated frequency reference. Similar structures have been implemented for thermal-diffusivity [2], LC [3], and MEMS [4] time-constant references. Inherently, it is simpler than the Dual-RC because only one of the channels is sensitive to frequency. The lack of frequency sensitivity in the temperature sensor ensures a decoupling of the two variables, which allows temperature compensation to be implemented without affecting FLL loop dynamics.

5.2.1. Temperature Compensation Principle

Figure 5.2 shows that due to the TC of R_{WB}, ϕ_{WB} has an undesired temperature dependency in addition to its frequency dependency. Conversely, the output of the WhB temperature sensor μ_{WhB} is a temperature-dependent but frequency-independent signal. A polynomial mapping function's coefficients are obtained via trimming μ_{WhB} to map it to the inverse of ϕ_{WB} yielding ϕ_{comp}. $p(\cdot)$ can be derived from measurements of the two DSMs at different temperatures while the WB is being driven at f_{drv} to map $\mu_{WhB}(T)$ to $\mu_{WB}(T)$. This results in the compensating phase signal ϕ_{comp} as in 5.2.

$$\phi_{comp}(T)|_{f=f_{drv}} = p(\mu_{WhB}(T)) = -\phi_{WB}(T)|_{f=f_{drv}} \tag{5.2}$$

ϕ_{comp} is then added to ϕ_{WB} to yield ϕ_e that is the desired frequency-dependent but temperature-independent signal.

$$\phi_e(T,f) = \phi_{WB}(T,f) - \phi_{comp}(T,f) \tag{5.3}$$

$$\phi_e(T)|_{f=f_{drv}} = 0 \tag{5.4}$$

Over frequency, ϕ_e crosses zero at the intended target frequency of $f_{drv} = f_{ref}/N$. Since the infinite DC gain of the digital accumulator drives ϕ_e to zero, it is ensured that f_{DCO} will lock to the reference frequency, f_{ref}. Since ϕ_e has no temperature sensitivity, f_{ref} will be stable across temperature.

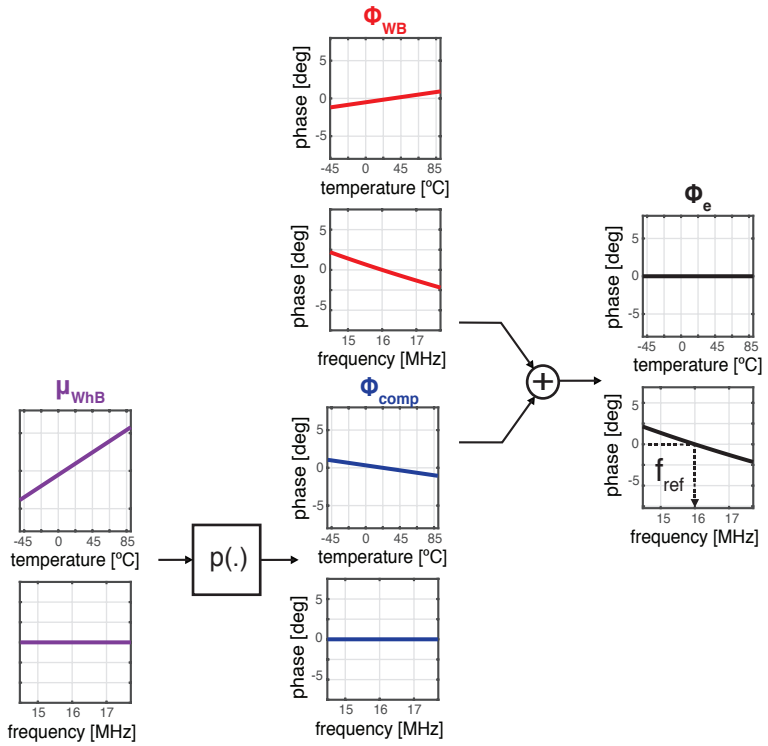


Figure 5.2: Temperature compensation principle for the proposed frequency reference

5.2.2. Linear Model

The simple linear model of Figure 5.3 is developed to gain insight into the dynamic behavior of the proposed temperature-compensated frequency reference. The sensitivities of the WB channel are identical to those developed for the model in chapter 4. The frequency input is converted to phase via the frequency-to-phase sensitivity of ϕ_{WB}, defined as K_{phi}, and derived from the WB phase equation as in 4.10. Similarly, the temperature input is converted to phase via the temperature-to-phase sensitivity of ϕ_{WB}, defined as K_T as in 4.11. The sources of inaccuracy originating from the Wien Bridge, primarily due to the spread of its components, are modeled as additive errors, $\Delta\phi_{WB}$.

The Wheatstone bridge is modeled as a temperature-to-digital converter. An input temperature change is converted to a digital signal through the temperature sensitivity of the Wheatstone Bridge and its readout $\partial\mu_{WhB}/\partial T$. The sources of inaccuracy originating from the Wheatstone Bridge are modeled as a lumped temperature error at its input to simplify the following

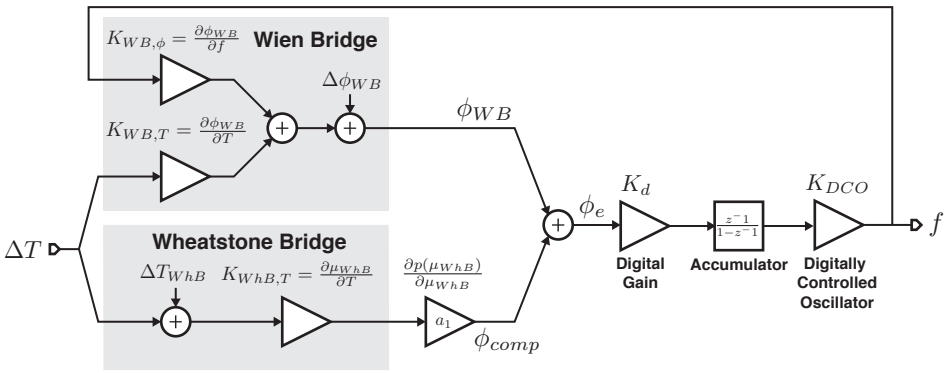


Figure 5.3: The linear model of the temperature-compensated frequency reference

inaccuracy analysis.

The coefficients polynomial $p(\cdot)$ are obtained through a two-point trimming procedure where both μ_{WhB} and ϕ_{WB} are measured at two temperature points. From these measurements, the coefficient a_1 is obtained so that the temperature sensitivity of ϕ_e is zero, as in 5.5.

$$\begin{aligned} 0 = \left.\frac{\partial \phi_e}{\partial T}\right|_{f=f_0} = \frac{\partial}{\partial T}\left(\phi_{WB} - p(\mu_{WhB})\right) = \frac{\partial}{\partial T}\left(\phi_{WB} - a_1\mu_{WhB} - a_0\right) \\ = \frac{\partial \phi_{WB}}{\partial T} - a_1\frac{\partial \mu_{Whb}}{\partial T} \\ a_1 = \frac{\partial \phi_{WB}}{\partial T}\left(\frac{\partial T}{\partial \mu_{WhB}}\right)^{-1} \end{aligned} \tag{5.5}$$

Similar to the Dual-RC case, the DCO is modeled as a block with the gain K_{DCO} with a control input and frequency output.

5.2.3. Inaccuracy Analysis

The inaccuracy of the frequency reference can be determined by transferring the error sources to the frequency output. The linear model developed in the previous section is utilized to derive these transfer functions.

The contribution of the $\Delta\phi_{WB}$ to frequency error, Δf in the steady-state condition is:

$$\phi_e = 0 = \Delta f \frac{\partial \phi_{WB}}{\partial f} + \Delta \phi_{WB}$$
$$\Delta f = \Delta \phi_{WB} \left(\frac{\partial \phi_{WB}}{\partial f} \right)^{-1} = \Delta \phi_{WB} \frac{-3N}{4\pi R_0 C} \tag{5.6}$$

Using 5.6 and 5.1, the variance of the output frequency due to the variance of phase, σ_ϕ is:

$$\frac{\sigma_f}{f_{ref}} = \left| \frac{\Delta f}{\Delta \phi_{WB}} \right| \frac{1}{f_{ref}} \sigma_\phi = \frac{3}{2} \sigma_\phi \tag{5.7}$$

5.7 indicates that the phase spread in the WB will manifest directly as output frequency errors. Selecting a filter with the highest phase-frequency sensitivity and the lowest number of components that spread and contribute to inaccuracy is imperative.

Similarly, the contribution of temperature sensing errors, ΔT, to the output frequency in the steady-state condition can be found by:

$$\phi_e = 0 = \Delta f \frac{\partial \phi_{WB}}{\partial f} - \Delta T a_1 \frac{\partial \mu_{WhB}}{\partial T}$$
$$= \Delta f \frac{\partial \phi_{WB}}{\partial f} - \Delta T \frac{\partial \phi_{WB}}{\partial T} \left(\frac{\partial T}{\partial \mu_{WhB}} \right)^{-1} \frac{\partial \mu_{WhB}}{\partial T}$$
$$\Delta f = \Delta T \frac{\partial \phi_{WB}}{\partial T} \left(\frac{\partial \phi_{WB}}{\partial f} \right)^{-1} = \Delta T \frac{TC_1 N}{2\pi R_0 C} \tag{5.8}$$

The variance of the output frequency, σ_f^2, due to the variance of temperature error σ_T^2 is then given by:

$$\frac{\sigma_f}{f_0} = \left| \frac{\Delta f}{\Delta T} \right| \frac{1}{f_0} \sigma_T = |TC_1| \sigma_T \tag{5.9}$$

5.9 shows that the TC_1 of the WB resistor should be minimized to suppress the contribution of temperature inaccuracy to frequency error. It also indicates that the temperature sensitivity of the WhB does not contribute to the error transfer. However, maximizing the WhB temperature sensitivity is still beneficial for two reasons. A more sensitive bridge yields a more energy-efficient temperature sensor. In addition, errors originating from the WhB readout will manifest as more significant temperature errors with a lower

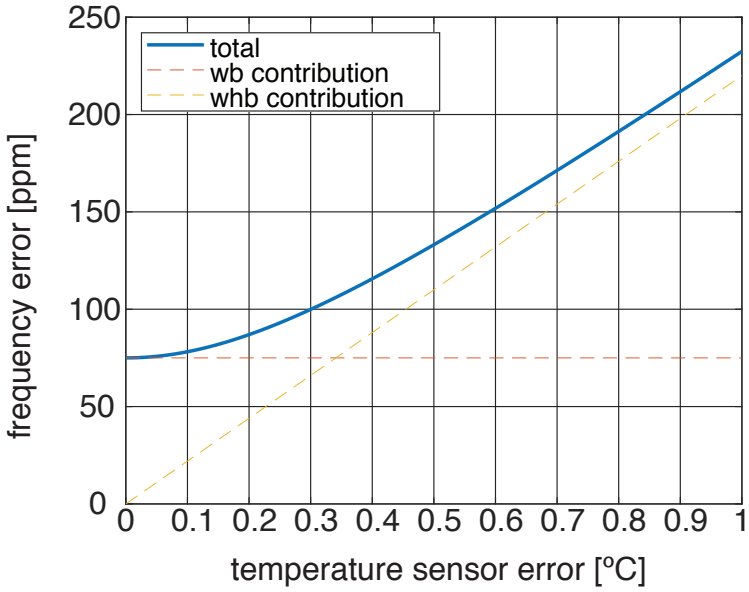


Figure 5.4: The simulated error of a frequency reference with a -220 ppm/°C WB resistor for increasing temperature error

bridge sensitivity. The total inaccuracy for the reference can be calculated as in 5.10

$$\frac{\sigma_f}{f_0} = \sqrt{|TC_1|^2\sigma_T^2 + \sigma_\phi^2} \tag{5.10}$$

Figure 5.4 shows the simulated frequency error of a frequency reference with an R_{WB} with TC_1 = -220 ppm/°C for varying temperature error. Additionally, the WB contributes a constant 75 ppm frequency error due to its phase spread. The magnitude of both error sources is assumed after trimming and systematic error correction. TC_1 of R_{WB} determines the slope of the error contributed by the WhB temperature sensor, and it completely dominates the total error for σ_T larger than 0.3 °C. Using a low-TC resistor relaxes the temperature sensor accuracy specifications. For a reasonably accurate 0.1 °C temperature sensor, it is possible to target a smaller-than-100 ppm inaccuracy frequency reference.

Comparison to the Dual-RC Architecture Comparing 5.10 and 4.17 gives insight into the differences between the two architectures regarding

accuracy. In the Dual-RC architecture, a pair of resistors with TCs of different signs consistently results in a more accurate frequency reference than in the traditional case. However, the benefits diminish as the magnitudes start to diverge, and the Dual-RC converges to the same inaccuracy as the traditional system. Conversely, when resistors with TCs of the same sign are utilized, the resulting frequency error is always higher, and the traditional architecture is more accurate than the Dual-RC.

5.2.4. Second-order Nonlinearity Compensation

The residual inaccuracy of first-order temperature compensation is expected to be dominated by the 2nd-order temperature coefficient of the resistors. Under the constraint that the first and second-order TCs of the two bridge outputs are related, it is possible to cancel the second-order dependency only with first-order compensation.

$$\begin{aligned}\phi_{WB}(T) &= \phi_{WB,0} + \phi'_{WB}\cdot T + \phi''_{WB}\cdot T^2 + \ldots \\ \mu_{WhB}(T) &= \mu_{WhB,0} + \mu'_{WhB}\cdot T + \mu''_{WhB}\cdot T^2 + \ldots\end{aligned} \tag{5.11}$$

$a_{1,0}$ of the ideal first-order fit, $p(x) = a_1 \cdot x + a_0$, is derived by equating the first-order sensitivity of the residual of trimming to zero as in 5.12.

$$\begin{aligned}\phi_e &= \phi_{WB}(T) - p(\mu_{WhB}(T)) \\ &= (\phi_{WB,0} - a_1\phi_{WhB,0} - a_0) + (\phi'_{WB} - a_1\mu'_{WhB})\cdot T \\ &\quad + (\phi''_{WB} - a_1\mu''_{WhB})\cdot T^2\end{aligned}$$

$$\begin{aligned}\phi_e = 0 \longrightarrow a_1 &= \frac{\phi'_{WB}}{\mu'_{WhB}} \\ a_0 &= \phi_{WB,0} - \mu_{WhB,0}\frac{\phi'_{WB}}{\mu'_{WhB}}\end{aligned} \tag{5.12}$$

The residual's second-order sensitivity can be zeroed if 5.13 is satisfied.

$$\frac{\phi''_{WB}}{\mu''_{WhB}} = \frac{\phi'_{WB}}{\mu'_{WhB}} \tag{5.13}$$

Given a specific process, the resistors used in the WB and WhB will present a fixed $TC_{1,2}$ determined by their physical properties. The ability to tune $TC_{1,2}$ independently is not within the circuit designer's toolset. However, the processes of resistance-to-phase conversion of the WB and the

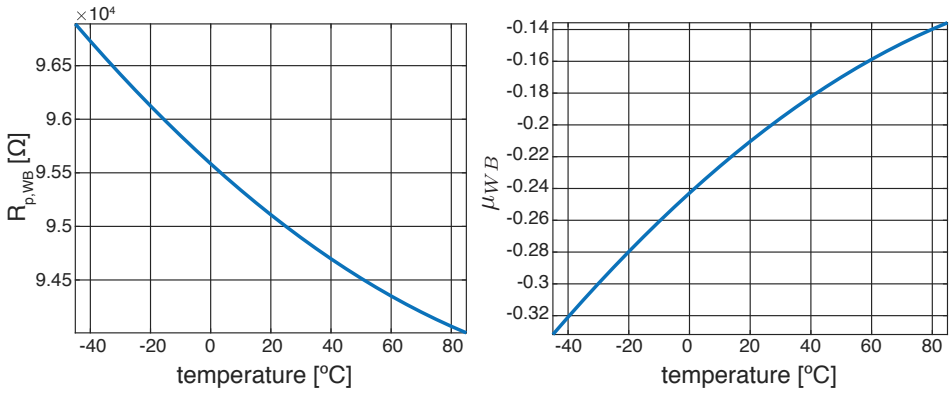


Figure 5.5: The simulated nonlinear WB resistance and the resulting PDDSM bitstream

temperature determination via the ratio of two resistors are nonlinear. The designer can choose specific parameters (i.e., phase range for the WB and the resistor ratios for the WhB) to try and modulate the TC_2 of the resulting bitstreams to cancel each other.

Bridges for the current design were simulated with process parameters for the passives obtained from the PDK, as shown in 5.1. For the WB, the simulation of the bitstream consisted of calculating the phase response at the drive frequency and its harmonics and the subsequent conversion to voltage through synchronous demodulation and filtering. Then, the bitstream value for each temperature was derived as the ratio of the two voltages generated with the phase references. Figure 5.5 shows the WB resistance and the resulting bitstream. Similarly, for the WhB, the simulation derives the current generated through each $R_{p,n,WhB}$ and calculates the bitstream required to balance the bridge. Figure 5.6 shows the two WhB resistors and the resulting bitstream.

The TCs for the two bitstream outputs are shown in Table 5.1. The TC ratios do not precisely satisfy 5.13, but their relative values imply a considerable reduction in the magnitude of the second-order residual. The two bitstreams are subjected to the trimming operation to verify the cancellation, and their final residual is converted to frequency error via 4.10. The residual frequency error can be seen in Figure 5.7. The residual error shows a reduced 2^{nd}-order nonlinear component and is primarily dominated by its 3^{rd}-order nonlinear component.

Fine-tuning the system parameters to achieve reasonably good cancellation of the second-order nonlinearity shows a 3x improvement. Unfortu-

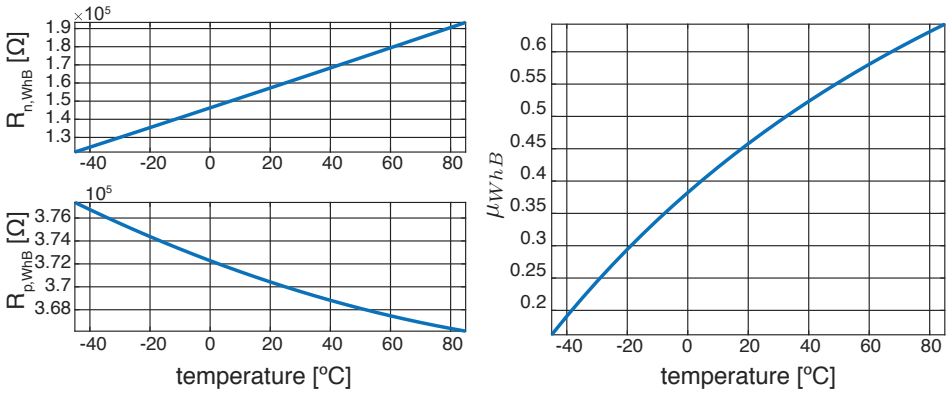


Figure 5.6: The simulated nonlinear WhB resistances and the resulting DSM bitstream

Component	Average Value	TC_1 [ppm/°C]	TC_2 [ppb/°C^2]
$R_{p,WB}$	95 KΩ	-220	845
C_{WB}	4.25 pF	15	-
$R_{p,WhB}$	370 KΩ	-220	845
$R_{n,WhB}$	160 KΩ	3300	602
μ_{WB}		2290	-7223
μ_{WhB}		2941	-11101
TC Ratio		=0.778	=0.650

Table 5.1: Process parameters for the components used in the second-order nonlinearity cancellation simulation and the TCs of the simulation outputs

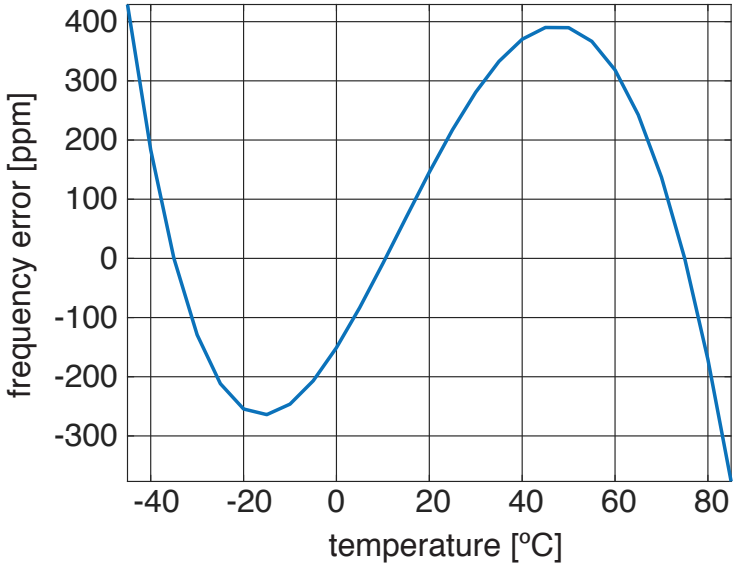


Figure 5.7: The simulated residual frequency error after two-point trimming, showing a reduced 2nd-order nonlinear component.

nately, from the designer's perspective, only a few free variables can be tuned to achieve the cancellation. If a different process does have considerably different behavior in the $TC_{1,2}$ of their passives, the effect might require a change in the ranges used or might even be impossible to achieve.

5.2.5. Dynamic Analysis

The dynamic analysis of the improved frequency reference is very similar to the digital FLL system, as discussed in chapter 3. An additional source that needs to be considered is the WhB noise. The transfer function of the WhB noise has a low pass characteristic as in 5.14. The temperature sensitivity of the WhB drops off due to the polynomial 5.5. The gain factor implies that it is also beneficial to reduce the temperature sensitivity of the WB to minimize the jitter contribution of the temperature sensor.

$$\frac{f(z)}{n_T(z)} = K_{WB,T}\frac{K_d K_{DCO} z^{-1}}{1+\left(\frac{K_{WB,\phi} K_d K_{DCO}}{N}-1\right)z^{-1}} \tag{5.14}$$

Figure 5.8 shows the frequency response of the magnitude of the noise transfer functions of the WB, WhB, and DCO error sources. These three

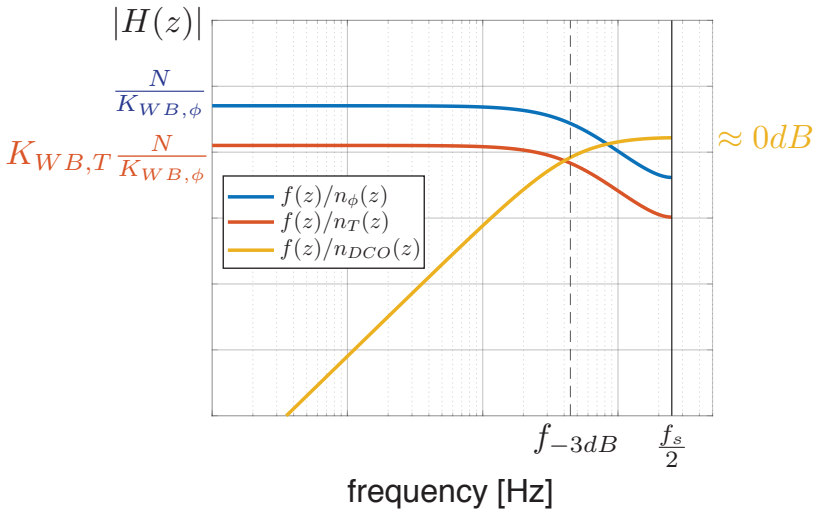


Figure 5.8: Frequency response of noise transfer functions' magnitude for the FLL

transfer functions' bandwidth is dictated by the factor $(K_{WB,\phi}K_dK_{DCO})/N$. The programmable gain K_d in the digital engine allows the bandwidth to be flexibly set after production. In this design, the loop bandwidth is set around 50Hz, enough to compensate for thermal transients, but more is needed to filter out most of the oscillator's wide band noise. Thus, system noise is dominated by the oscillator.

5.3. Circuit Design

5.3.1. Bridge Design

From the findings of the linear analysis of section 5.2.3, it is clear that the TC_1 of the WB should be minimized to minimize the inaccuracy and noise contribution of the temperature sensor. To this end, the WB was realized with the resistor with the lowest TC available in the process: p-type polysilicon resistors with a TC_1 = 200 ppm/°C and MIM capacitors with TC_1 = 30 ppm/°C). As a result, the inaccuracy of the WhB temperature sensor, typically less than 0.2°C, is expected to cause frequency errors in the order of 40ppm.

The temperature sensitivity of the WhB should be maximized to optimize its energy efficiency. In [5], this was achieved by combining silicided diffusion resistors with a TC_1 = 3300 ppm/°C and n-type polysilicon resistors with a TC_1 = -1500 ppm/°C. However, in this work, the primary goal is

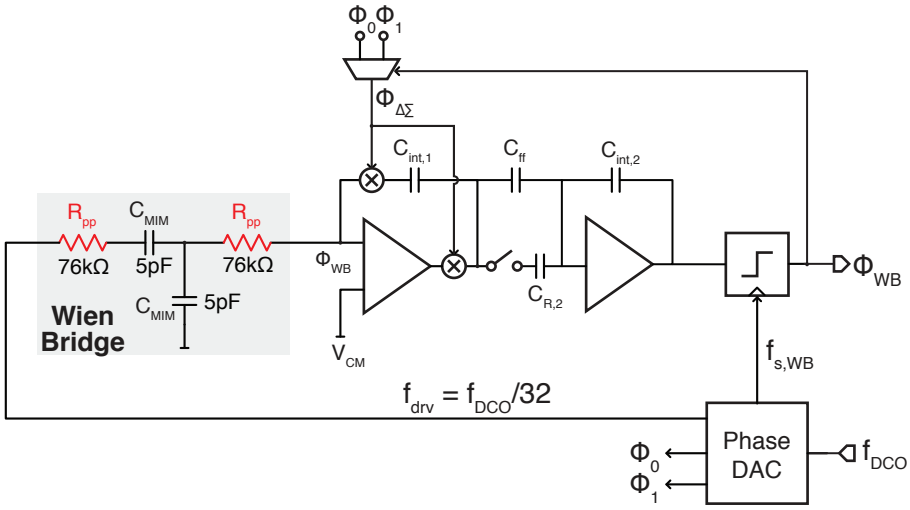


Figure 5.9: The simplified single-ended schematic diagram for the Wien Bridge and the Phase-Domain DSM readout

to optimize accuracy. As a result, the n-poly was replaced with the p-type polysilicon resistors used in the WB. In this design, the WhB consists of the resistors R_{sd} = 160kΩ, a silicided diffusion resistor with TC_1 = 3300 ppm/°C, and five units of the same p-polysilicon used in the WB, R_{pp} = 370kΩ with TC_1 = 200 ppm/°C. The choice ensures that the accuracy of the frequency reference is limited only by the spread of two types of resistors at the cost of some energy efficiency.

5.3.2. Wien Bridge Readout: Improved PDDSM

While architecturally very similar to its Dual-RC counterpart, multiple improvements to the PDDSM circuitry reduce its area significantly and achieve better performance [5]. This section describes these improvements.

Figure 5.9 shows the circuit diagram of the fully-differential Wien Bridge and the PDDSM readout. Its operation principle is identical to the Dual-RC design discussed in 4.2.1.

An imperative constraint for the DSM is that its 1st-stage integrator does not clip, which constrains $C_{int,1}$ to 180pF in the Dual-RC design. This capacitor occupies more than half the sensor area and almost 40% of the total frequency reference area.

Increasing the WB resistance to reduce the current into the integrator is the first step in reducing C_{int}. This work increases R_{WB} to 76kΩ, provid-

ing a 2x reduction of the 1st-stage input current. The reduction is at the expense of sensor resolution; thus, a further increase in the R_{WB} is undesirable. Another approach is to increase the output voltage swing capability of the amplifier. The amplifier is replaced with a two-stage Miller-compensated opamp based on current-reuse amplifiers to achieve a higher output voltage swing as in [5]. Overall, these improvements allow for a more than 7x reduction of C_{int}: from 180pF to 23pF.

To further reduce the area, the 2nd-stage integrator of the DSM is replaced with a switched-capacitor circuit. $C_{R,2}$, being switched at f_s, and $C_{int,2}$ determine the unity-gain bandwidth of the integrator. The primary benefit is gained by avoiding using the area-expensive R_{ff} resistor used in the Dual-RC design. Instead, the feed-forward coefficient is implemented by C_{ff}. The capacitive implementation comes at the cost of the two integrators not responding to temperature changes similarly. However, altering the noise-shaping characteristic over temperature has a negligible effect on system performance due to the high decimation factor used.

Combining these improvements reduces the area of the WB and the PDSDSM readout 6x compared to the Dual-RC design. With the modifications to the DCO, these improvements significantly reduce the total area cost of the frequency reference.

5.3.3. Wheatstone Bridge Readout: CTDSM

The second critical component of the temperature-compensated frequency reference is the temperature sensor. As shown in Section 5.2.3, the temperature sensor's accuracy is crucial for the reference. For its improved frequency reference implementation, a reference design [6] was modified for this accuracy target.

Figure 5.10 shows the circuit diagram of the used WhB temperature sensor. The WhB consists of the resistors R_{sd} = 160kΩ, a silicided diffusion resistor with TC_1 = 3300 ppm/°C and five units of the same p-polysilicon used in the WB, R_{pp} = 370kΩ with TC_1 = 200 ppm/°C. One unit is not switched, whereas the other four are controlled between the supply and V_{cm} by the bitstream. A change in temperature causes an imbalance in the bridge, creating a temperature-dependent output current, i_T into the 1st stage integrator. In steady-state, the loop gain of the DSM integrators zeroes this current by adjusting the density of the bitstream, balancing the bridge. Thus, the bitstream average μ_{WhB} is a digital representation of the temperature but is independent of the sampling frequency of the DSM.

The primary challenge in the design of the WhB readout lies in sizing the 1st-stage integration capacitor, $C_{int,1}$. A multi-bit DAC reduces the input

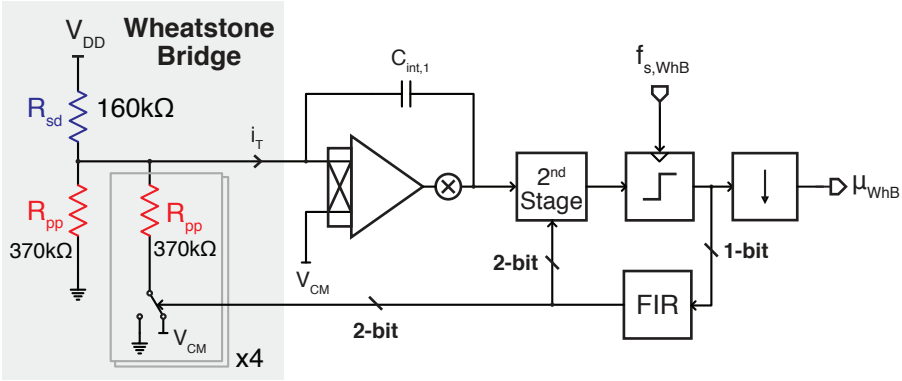


Figure 5.10: The simplified single-ended schematic diagram for the Wien Bridge and the Continuous-Time DSM readout

current of the integrator, allowing for the reduction of $C_{int,1}$, and is very feasible to implement for the WhB. However, driving this DAC with a 2-bit signal requires a multi-level quantizer, which is not desirable in a system where the digital logic is externally implemented. It was shown [6] that the FIR-DAC addresses this issue, enabling a 2x lower i_T while providing inherent dynamic-element matching.

The integrators used in the WhB CTDSM readout are the same as in the PDDSM. An additional FIR filter is implemented in the second stage to compensate for the delay introduced by the FIR-DAC.

The primary differences between the WhB temperature sensor and its reference design [6] originate due to changes introduced to achieve the highest accuracy possible. The WhB in this implementation is not trimmed, getting rid of transistor switches in the bridge, which contribute additional on-resistance to R_{sd} and potentially reduce accuracy. On the other hand, removing the ability to trim increases the process variation that needs to be covered by the DSM and reduces the energy efficiency of the temperature sensor. Additionally, the R_{pp} elements are switched between V_{CM} and 0 instead of V_{DD} and 0. The DSM only adjusts the value of the parallel combination of R_{pp} instead of varying the overall bridge sensitivity.

The WhB and its readout are optimized for accuracy at the cost of some energy efficiency. From a system perspective, the power cost of the WhB is insignificant compared to that of the DCO, and this trade-off is highly beneficial.

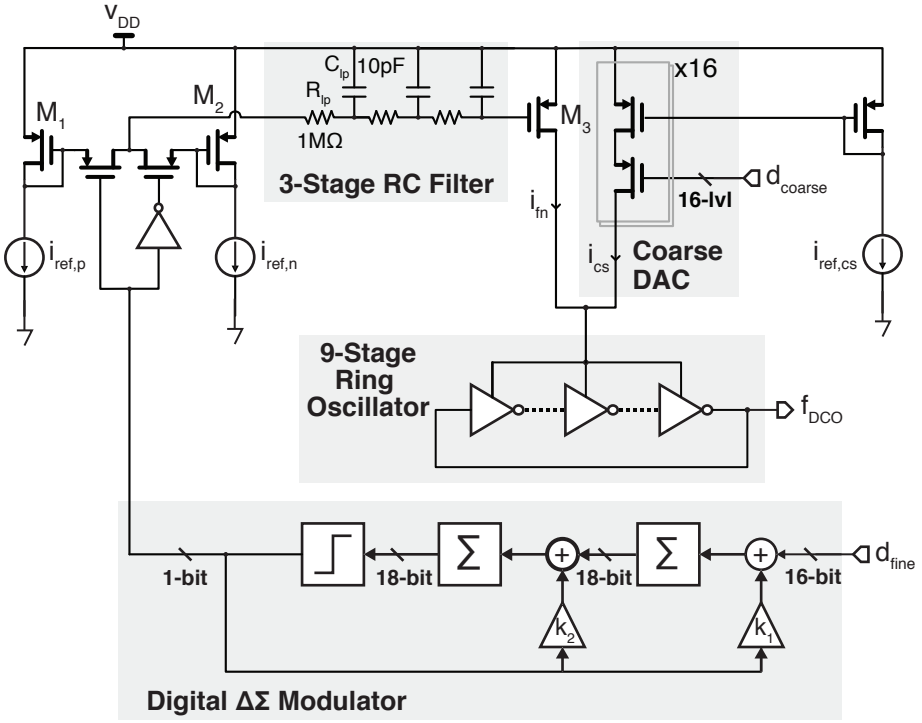


Figure 5.11: Schematic diagram of the Delta-Sigma DCO

5.3.4. Delta-Sigma Digitally Controlled Oscillator

Like the digitally controlled oscillator (DCO) used in the Dual-RC design, the DCO output frequency closes the feedback path by driving the RC filter and clocks all other system components. The high resolution and requirements led the previous design to a coarse/fine structure consisting of a 5-bit coarse current-steering DAC and a 13-bit fine R-2R DAC. However, the R-2R DAC occupied a significant silicon area (0.35 mm^2 [7]).

In the improved frequency reference, the fine DAC of the DCO is replaced with a delta-sigma DAC, as shown in Figure 5.11. A 2^{nd}-order digital delta-sigma modulator generates a 1-bit bitstream from the 16-bit loop filter output. The loop filter output (operating at f_{drive}/R = 500 kHz/1024 ≈488 Hz) is first oversampled by a factor of 8192 to the digital DSM operating frequency of 4 MHz. Then, it is subtracted from the 1-bit modulator output (scaled with 2^{16} in feedback) to generate the digital error signal. In closed-loop operation, this error signal is driven to zero with the gain provided by the two accumulators. The bitstream average represents the 16-bit

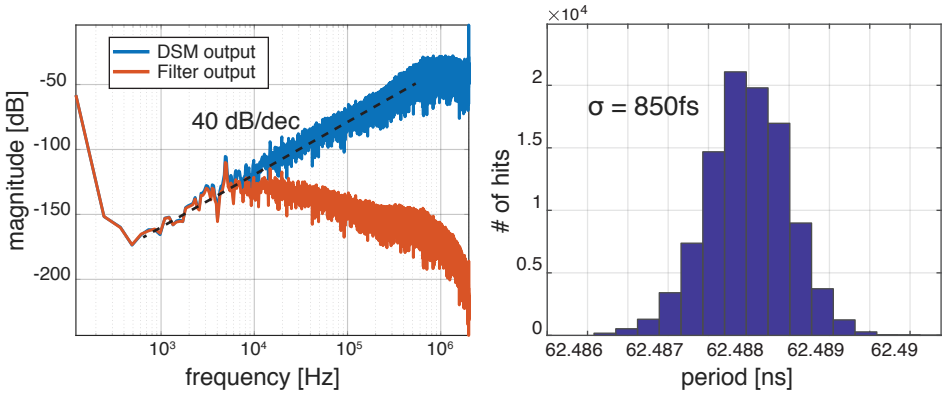


Figure 5.12: The simulated spectrum of the second-order delta-sigma modulator output and after the cascaded RC filter and output period jitter of the DCO

digital loop filter output code.

The DCO core features a 9-stage current-starved ring oscillator, similar to the Dual-RC DCO design. A 4-bit current-starved DAC with a nominal average i_{coarse} = 70µA provides coarse control of the frequency intended to cover the ±40% process variation and remains static during operation. The 1-bit DAC with a nominal average i_{fine} = 18µA is controlled by the delta-sigma modulator output and covers the ±7.5% temperature and voltage variation.

Two reference voltages, $v_{ref,0}$ and $v_{ref,1}$, are generated by setting the current of two diode-connected devices, $M_{1,2}$, to those generated from a reference. A constant-g_m circuit generates reference currents with temperature dependency opposite to that of the ring oscillator core, which relaxes the fine DAC's range. Depending on the bitstream output value, one of these voltages is selected to drive the input of the RC filter. The filter is a 3rd-order cascaded RC low-pass filter with a corner frequency of 4 kHz. The resistors of the filter, R_{lp}, are made from high-resistivity polysilicon, and for the capacitors, C_{lp} MOS capacitors are used to save chip area. The average output voltage of the filter is then converted to the current i_{fn} by M_3 and steered into the ring oscillator core. While this final step of voltage-to-current conversion is highly nonlinear, the monotonicity of the current is guaranteed by design.

An event-based simulation was run with the DSM, a simplified CCO core model, and the 3rd-order cascaded RC filter to verify the DSDCO. Figure 5.12 shows the spectrum of the digital DSM output and the output of the cascaded RC filter when a small constant code drives the DSM. The DSM output shows the expected 2nd-order noise shaping. The cascaded RC filter

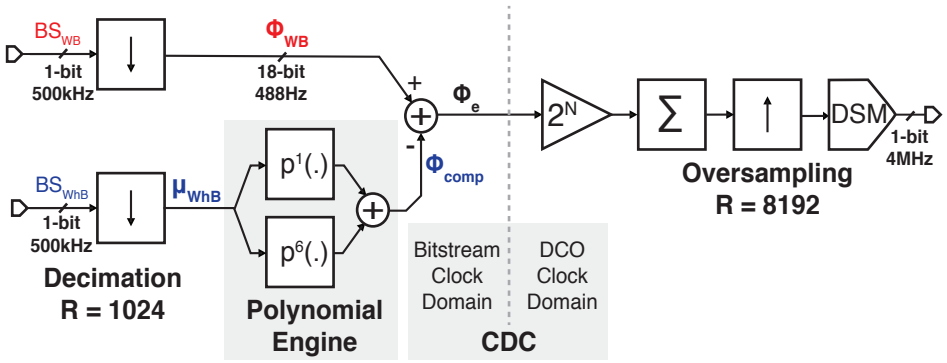


Figure 5.13: Block diagram of the digital loop filter and temperature compensation implemented in an FPGA

removes most of the quantization noise with a nominal cutoff frequency of 4 kHz. The filter output drives the CCO, which converts the output voltage to a current with its gain extracted from the simulation. The resulting output period jitter due to the DSM is 850 fs, which is negligible compared to the expected jitter due to white noise.

The delta-sigma DCO presents a significant improvement to the Dual-RC design. The hefty fine DAC of [7] is replaced by an inherently monotonic delta-sigma DAC that provides high resolution, and its additional noise contribution is negligible. These improvements reflect an area improvement to the frequency reference, in line with improvements to the other blocks.

5.4. Digital Loop Filter and Temperature Compensation

The digital components of the improved frequency reference reuse most of its blocks from the previous design in a simplified configuration. Similarly, it is realized off-chip in an FPGA. The primary changes are to the polynomial engine used, which is significantly simplified, and the addition of the digital DSM. Figure 5.13 shows the block diagram of the digital.

2-stage CIC decimation filters first decimate the two bitstreams. The 500 kHz sample rate of the input bitstreams is divided by a factor of 1024, which brings the sampling rate of the following blocks to 488 Hz. CIC output words are quantized to 18 bits, which simplifies the following logic and is high enough to ensure that any noise originating from digital quantization is well below the noise floor of the readout.

The polynomial engine processes the decimated output of the WhB DSM at 18 bits and realizes the polynomial mapping function p(·). The block implements a 6^{th}-order polynomial with 20-bit programmable coefficients. Its coefficients are calculated via trimming after production and programmed via an SPI interface to on-chip registers. Only p^1, which uses $a_{0,1}$, is utilized for the two-point trimming scenario, and $a_{2..5}$ are set to zero. For the two-point + batch scenario, a parallel p^6 branch is derived, but before programming, the coefficients of p^1 and p^6 are combined to get a single 6^{th}-order polynomial. Since the sample rate of the loop filter is low, the polynomial engine is implemented entirely with combinational logic.

As discussed in Section 4.3, there is a requirement for clock-domain crossing in the digital loop filter. The implementation in the improved design is the same, where a delayed decimation pulse clocks all the blocks.

The loop accumulator is a 32-bit register, and it is programmed to be wider than its 18-bit input to account for the range of dominant pole frequency that may be required. The input to the loop accumulator is scaled by a programmable left-shift operation that implements a 2^N gain. The left shift constitutes K_d and can be programmed to modulate the loop gain of the FLL.

Finally, the loop accumulator output is again quantized to 16 bits and oversampled by a factor of 8192. Since this oversampling clock at 4 MHz is also generated by dividers inside the digital block, there is no need for a CDC here. As described in Section 5.3.4, the loop accumulator output is converted into a 1-bit signal with shaped quantization noise that drives the DCO.

5.5. Experimental Results

The improved RC frequency reference was implemented in a TSMC 0.18 µm standard CMOS process [1]. The total active area is 0.3 mm^2, where WB + PDDSM and WhB + CTDSM each occupy 0.125 mm^2, and the ΔΣDCO occupies 0.05 mm^2. Each of the WB + PDΔΣM and WhB + CTΔΣM dissipate 50 µW from a 1.8 V supply. The ΔΣDCO (excluding the digital ΔΣM) dissipates 120 µW from a 1.8 V supply. The total analog power consumption of the frequency reference is 220 µW.

Digital components of the DFLL (decimation filters, temperature compensation, loop filter, and the digital DSM) were realized off-chip in an FPGA. However, later, an area/power-optimized version of the digital logic was implemented in the same process. This post-place-and-route implementation is estimated to consume 0.1 mm^2 area. The power consumption of the

simulated (post-synthesis estimation, without activity vectors) was 60 µW.

The measurements were carried out initially on a batch of 20 samples packaged in ceramic DIL24 packages. Afterward, another eight samples from a separate run were characterized to assess the architecture's feasibility and the stability of compensation polynomials over multiple batches.

The samples were subjected to a three-step trimming and characterization procedure. Firstly, a temperature characterization is conducted to determine the values of the two ΔΣM outputs while the FLL is in the closed-loop configuration. Next, the trimming polynomials are calculated for the two trimming scenarios. The first is a two-point-only trim, where for each sample, only the linear coefficients of the polynomials are derived from two temperature measurements. The second is a two-point + batch trim, where the average nonlinear residual error of the first step over all samples is removed by deriving the higher-order coefficients of the polynomial engine. In the final characterization, the calculated polynomials are loaded, and the inaccuracy of the frequency reference is measured over temperature.

5.5.1. Closed-loop Trimming Methodology

Trimming a frequency reference is prone to creating a chicken-and-egg problem. Outputs of the readout need to be determined accurately at the temperature of interest. Accurately reading the outputs is only possible if the circuits are driven at the target reference frequency. Unfortunately, without the proper trimming coefficients already in place, the DCO frequency will not be well-determined. Also, the DCO will be very sensitive to ambient temperature changes, so maintaining high thermal stability during trimming becomes necessary.

The traditional way to address this problem is to open the feedback loop and characterize the readout under thermal stability while driving it with an external frequency reference. The disconnect between the trimmed state of the system versus its normal operating condition contaminates the measurement. The errors made due to this disconnect will manifest as additional inaccuracy of the final frequency reference, as was observed in chapter 4.

For the improved reference, a new trimming methodology was developed that tries to avoid these problems altogether. The method relies on configuring the FLL in its normal closed-loop operating condition, as shown in Figure 5.14, which only lacks the correct trimming coefficients. Then, the coefficient a_0 is adjusted via a binary search algorithm, where the output frequency is directed to converge to the final reference. The closed-loop search ensures that the FLL is brought to its eventual operating condition. Then, polynomial coefficients can be derived from the last value of a_0, which

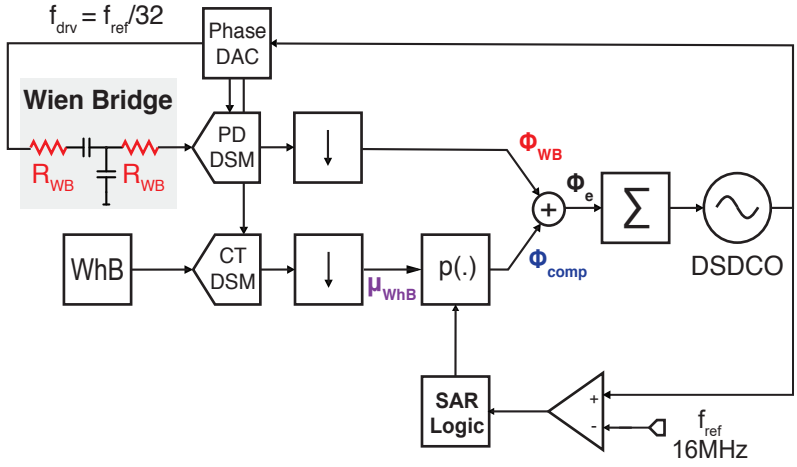


Figure 5.14: Block diagram of the RC frequency reference under the closed-loop trimming setup using the binary search algorithm

contains information about μ_{WB} and a readout of μ_{WhB}. Trimming coefficients derived through this procedure are expected to recreate the final operating condition accurately where the DCO output is at the reference frequency.

The procedure starts with the feedback closed and all polynomial coefficients set to 0. The closed-loop FLL locks to a frequency where the WB shows 0 phase shift, $f_{DCO}[0]$, where $0=\mu_{WB}|_{f_{drive}=f_{DCO}[0]}$. $f_{DCO}[0]$ is compared to f_{ref} = 16MHz, and depending on the result, the external SAR logic updates the MSB of a_0. This process is repeated after the FLL locks to the new $f_{DCO}[1]$, and MSB-1 of a_0 is updated. After predetermined N+1 steps of this binary-search algorithm, the $f_{DCO}[N]$ will have converged to f_{ref}, and the final values $a_0[N]$ and $\mu_{WhB}[N]$ are recorded.

Figure 5.15 shows the convergence of the frequency error and a_0 for a single sample. The initial frequency error is on the order of 10000 ppm, but after a few steps of the search tends to zero. Meanwhile, a_0 approaches to $\mu_{WB}|_{f_{drv}=f_{ref}}$. From measurements, it can be seen that 18 steps are sufficient to ensure that the residual frequency error after convergence is < 10 ppm. Steps taken after this are under the system's noise and will not contribute to increasing accuracy.

From these final recorded values of a_0 and μ_{WhB}, it is possible to derive the outputs of the two readouts as in 5.15 and 5.16.

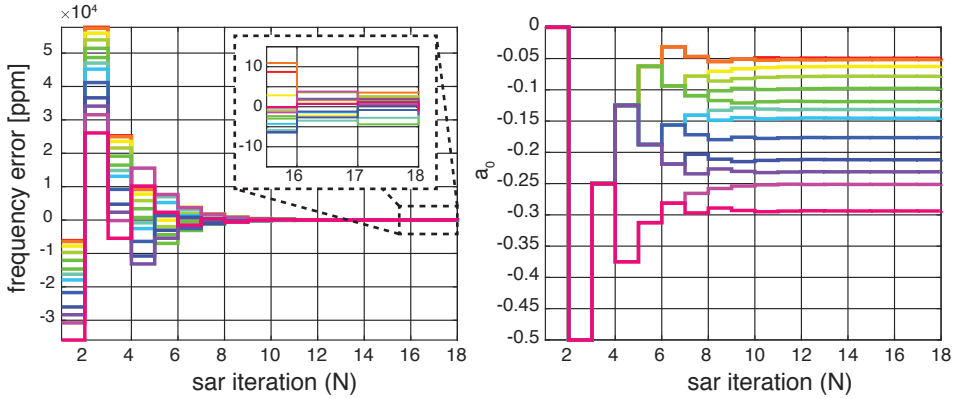


Figure 5.15: Measured convergence of frequency error and a_0 with the closed-loop trimming methodology for a single sample at different temperatures

$$\mu_{WB}(T)|_{f_{drv}=f_{ref}} = a_0[N] \tag{5.15}$$

$$\mu_{WhB}(T) = \mu_{WhB}[N] \tag{5.16}$$

A simple but effective improvement to the trimming procedure can be achieved by preloading the a_1 coefficient of the p^1 with its expected average value. a_1 represents the WB to WhB temperature sensitivities' ratio and can be estimated before measurement either from simulation or a coarse characterization. Having the average a_1 value in place achieves some rudimentary temperature compensation that reduces the sensitivity of f_{DCO} to ambient temperature changes. Thus, the requirement for thermal stability subsides during the SAR operation, and the measurement can be carried out faster. The only difference to the above procedure is the calculation of $\mu_{WB}(T)$, which is modified as in 5.17.

$$\mu_{WB}(T)|_{f_{drv}=f_{ref}} = a_1 \cdot \mu_{WhB}[N] + a_0[N] \tag{5.17}$$

$$\mu_{WhB}(T) = \mu_{WhB}[N] \tag{5.18}$$

Figure 5.16 shows the measured bitstream averages for the WB, $\mu_{WB}(T)$, and the WhB, $\mu_{WhB}(T)$, for 20 samples derived from the closed-loop trimming step. Over the -45 °C to 85 °C range, μ_{WB} variation corresponds to a 0.5° phase shift with $\phi_{0,1}$ = 90°, 95° and for the 200 ppm/°C TC_1 of the p-poly resistor. Similarly, the μ_{WhB} variation is found to be as expected with

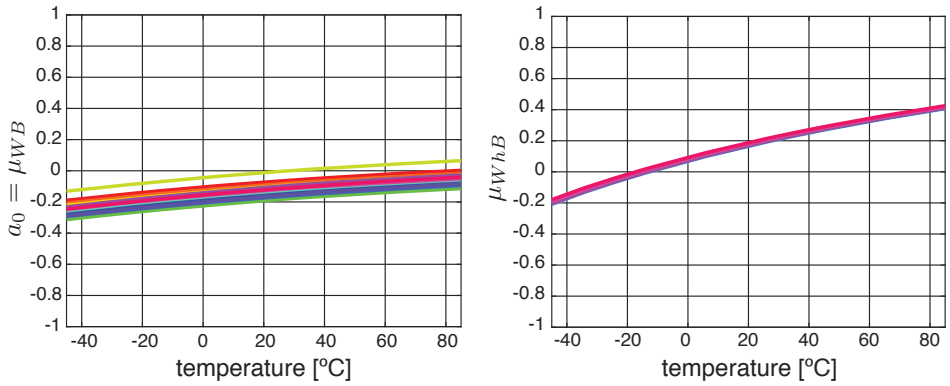


Figure 5.16: Measured bitstream averages for μ_{WB} and μ_{WhB} for 20 samples.

the ratio of silicided n-diffusion and p-poly resistors used in the bridge. Interestingly, the two outputs show similar curvature over this temperature range. The similarity in curvature is due to the second-order nonlinearity cancellation proposed in 5.13.

5.5.2. Derivation of the Trimming Polynomials

After temperature characterization of the readout outputs, polynomial coefficients that map the output of μ_{WhB} to μ_{WB} can be derived. Values obtained in the previous step for μ_{WB} are plotted against μ_{WhB} as the markers in Figure 5.17 (a) to visualize the mapping.

As mentioned before, the first scenario is the two-point trim. Only the highlighted points in Figure 5.17 corresponding to the two temperature targets -35°C and 75°C are used to derive the linear fit coefficients $a_{0,1}$ where $p^1(\mu_{WhB}) = a_1 \cdot \mu_{WhB} + a_0$. The solid lines in Figure 5.17 (a) are the first-order fits to each sample obtained in this step passing through the highlighted points. Figure 5.17 (b) shows the different $a_{0,1}$ coefficients obtained for each sample. Unlike the a_0 (offset) coefficient, a_1 displays a minimal spread. The low TC spread implies that the accuracy of the bridges is maintained through the readout, and the primary spread is due to bridge passive geometry.

Figure 5.18 shows the residual outputs of the two-point trim operation and their average across samples. Interestingly, even though the previous step only removes linear errors, i.e., the TC_1 and offset, the resulting residuals have a weak second-order component and a strong third-order component. Matching the second-order nonlinearity of the two bridges and

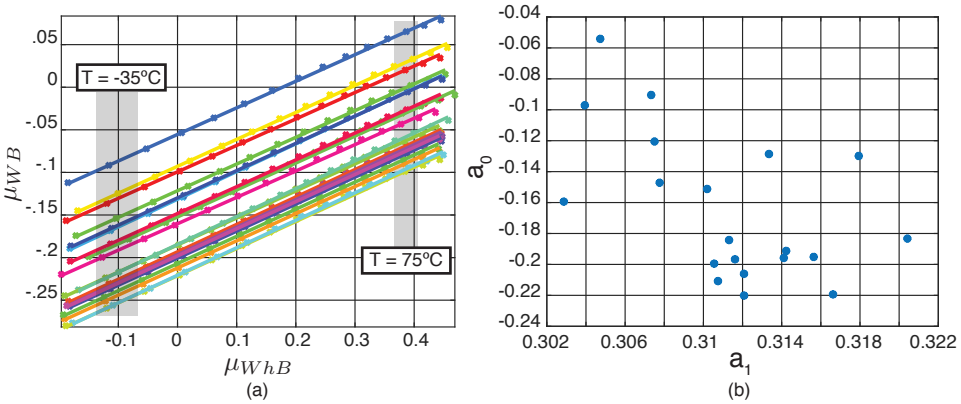


Figure 5.17: (a) Measured bitstream average for the WB, μ_{WB}, mapped to the bitstream average of WhB, μ_{WhB} (markers) and the first-order fit (solid line) using the two highlighted trimming points at -35°C and 75°C. (b) $a_{0,1}$ coefficients of the first-order fit.

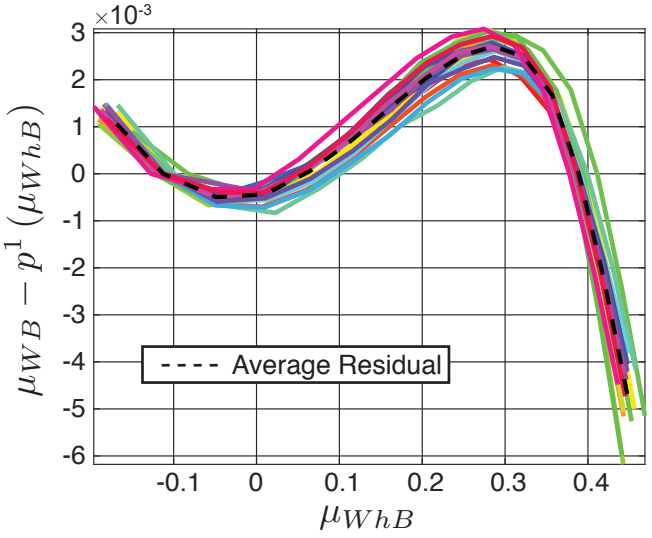


Figure 5.18: Residual outputs of the two-point trim and its average across samples

their readouts is explained in Section 5.2.4.

Implementation of the two-trim + batch calibration scenario is required to increase the accuracy further. A sixth-order polynomial p^6 fits the average residual of the two-point trim and constitutes the batch component. This polynomial is the same for every sample. Still, it requires a statistically significant set of samples in a batch to be characterized before an accurate average residual can be obtained. All 20 samples were used to derive the average polynomial for this work.

5.5.3. Final Inaccuracy Characterization

With the polynomials for both temperature compensation scenarios derived, the remaining step is verifying the frequency reference's final inaccuracy. To this end, another temperature sweep was conducted where the FLL was closed, and the polynomials derived in the previous trimming step were loaded into the polynomial engine.

For the two-point temperature compensation scenario, only p^1 was used. Figure 5.19 shows that the two-point compensated frequency reference achieves ±385 ppm frequency inaccuracy over 20 samples. The box method results in a residual temperature coefficient of 5.7 ppm/°C. The similarity in the nonlinear characteristics of the residuals from Figure 5.18 and the frequency inaccuracy of Figure 5.19 is expected. When only p^1 is used, the residuals from Figure 5.18 will have to be compensated via a deviation of the DCO frequency.

For the two-point + batch temperature compensation scenario, p^1, which is unique to each sample, and p^6, which is the same for all samples, were used. p^1 for each sample was substituted as the argument to p^6 to simplify the hardware requirements. This process yields a new sixth-order polynomial that is unique to each sample. Figure 5.20 shows that the two-point + batch temperature-compensated frequency reference achieves ±90 ppm frequency accuracy over 20 samples. A residual temperature coefficient of 1.3 ppm/°C is derived with the box method.

Batch-to-batch Stability

Characterizing the robustness of the temperature compensation methodology against process variation is only possible with samples obtained from multiple production batches. Second-order nonlinearity compensation with the two-point trim might vanish if the first and second-order TCs of the resistors vary considerably. Thus, it is essential to characterize samples from multiple batches to ensure that the trimming methodology can maintain the inaccuracy levels over significant process variations.

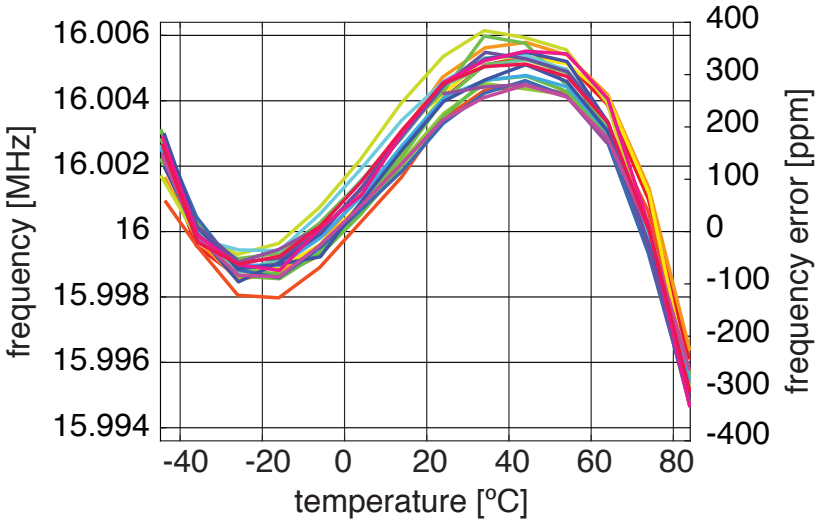


Figure 5.19: Inaccuracy of the two-point temperature-compensated frequency reference

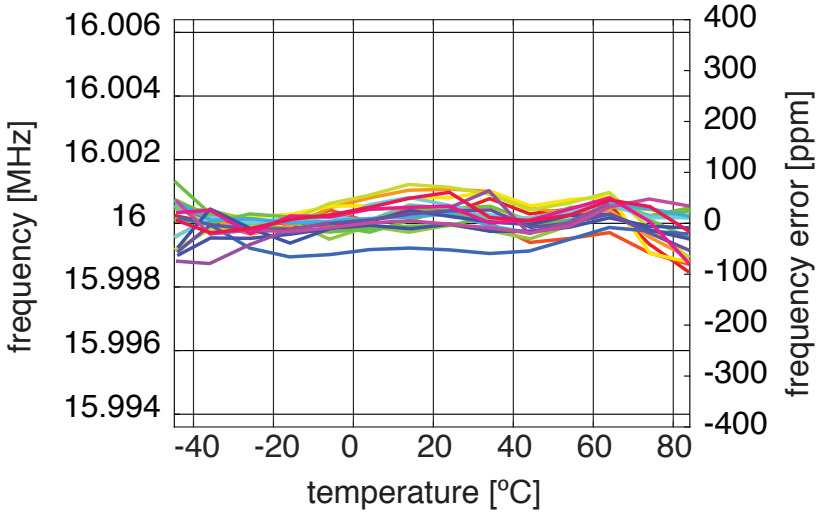


Figure 5.20: Inaccuracy of the two-point and higher-order batch temperature-compensated frequency reference.

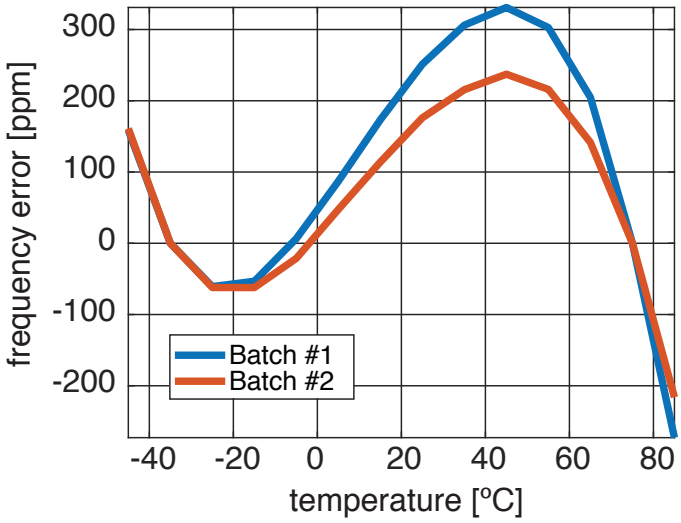


Figure 5.21: After two-point trimming from two sample batches, average frequency error estimations show the shifted nonlinear residual frequency error curves.

Samples from another batch were characterized to verify the robustness of the design and the trimming methodology to process variation. Figure 5.21 shows the average frequency error for the two batches derived from the two-point trimming residuals of Figure 5.18. It can be observed that the two residuals maintain the same overall shape. Second-order dependence on temperature is very low in both cases, verifying the robustness of the nonlinearity cancellation achieved in the design. Over multiple batches, the two-point temperature-compensated frequency reference can maintain the < ±400 ppm inaccuracy level achieved. However, the magnitude of the third-order dependence is different for the two curves. If the p^6 from the first batch were to be used in the second, additional systematic inaccuracy would approximately degrade the inaccuracy to ±200 ppm. Each batch would require a different p^6 to be characterized from a statistically significant number of samples to maintain the < ±100 ppm inaccuracy of the two-point + batch temperature compensation.

Supply Sensitivity

1.8V supply for both the DCO and the DSMs was varied between 1.6V and 2V to characterize sensitivity to supply voltage. Figure 5.22 shows the output frequency inaccuracy with supply voltage variation. The frequency error is normalized to the output frequency for each sample at 1.8V to remove the

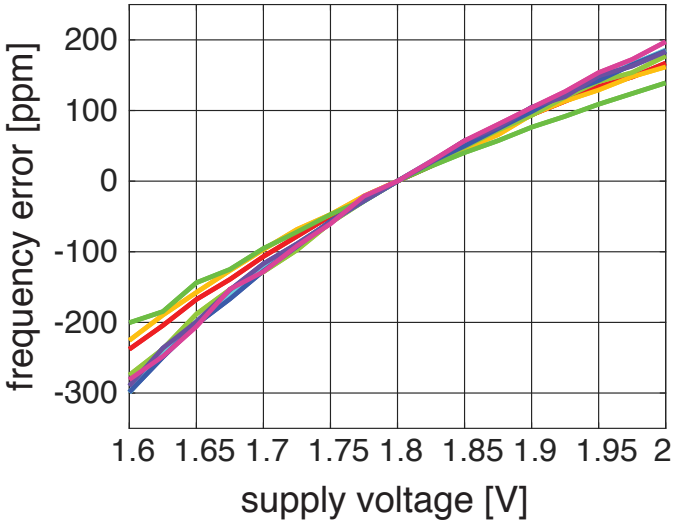


Figure 5.22: Frequency inaccuracy with supply voltage variation, normalized to the output frequency for each sample at 1.8V

residual spread of different samples at the characterization temperature (25 °C).

The worst-case sample shows a 500 ppm_{pp} frequency inaccuracy, resulting in a voltage sensitivity coefficient of 0.12 %/V. Since the DCO is frequency locked, the source of the inaccuracy is attributed to the voltage sensitivity of the two bridges. The WB supply sensitivity mechanism is explained in chapter 4 and is the same for this design. With its differential architecture, the WhB is expected to be more resistant to supply voltage variation in comparison, as the voltage sensitivity is limited to the mismatch of its DAC switches.

When used in a system with a standard ±5% supply variation, the inaccuracy caused by the supply sensitivity is two orders of magnitude higher than temperature inaccuracy. An LDO could be used to reduce the supply sensitivity by three orders of magnitude and render the inaccuracy caused by supply variation negligible.

5.5.4. Noise & Long-Term Stability

Like the previous design, the frequency reference's short-term (high-frequency) noise is characterized via a period jitter measurement. Its long-term stability is characterized via an Allan Deviation measurement. Figure 5.23 shows the

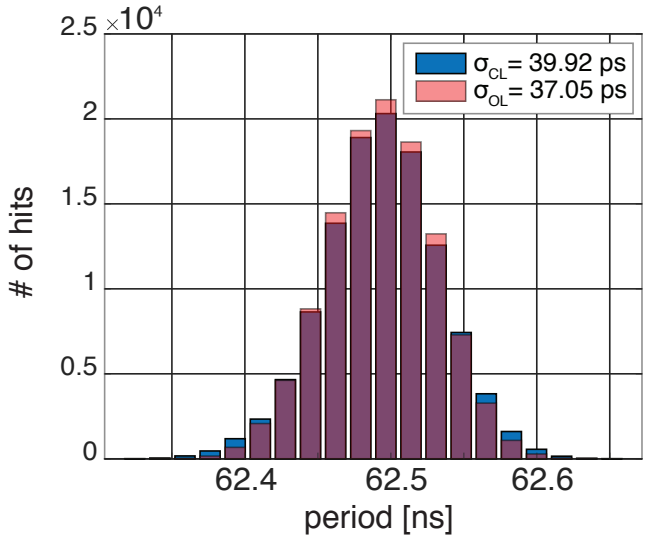


Figure 5.23: Period jitter of the open-loop DCO (with a non-zero DSM input) and the closed-loop frequency reference

period jitter measurement for both the open-loop DCO and the closed-loop trimmed frequency reference. The DCO was brought to the reference frequency for the open-loop case by adjusting the frequency control word to obtain the expected reference frequency. Thus, it includes the noise contribution of the DSM. The open-loop period jitter of the DCO is 37.05 ps_{rms}, and the closed-loop period jitter is 39.92 ps_{rms}. The increase in the period jitter implies that closing the loop adds approximately a 15 ps_{rms} jitter to the system.

Figure 5.24 shows the results of an Allan Deviation measurement conducted on the open-loop DCO, programmed similarly to the period jitter measurement and the closed-loop frequency reference. The open-loop DCO is limited to a noise floor of 20 ppm. Closing the loop for the trimmed frequency reference improves the noise floor to 320 ppb in an 8 s measurement time, showing significant improvement to the open-loop case. In the closed-loop case, it can be observed that ADEV for short measurement times is degraded, consistent with the increased period jitter observed. After a measurement time of 20 ms, consistent with a loop bandwidth of 50 Hz, ADEV improves with a -20 dB/dec slope.

From both measurements, it can be inferred that the closed-loop operation shows slight degradation in the noise performance of the system. Ob-

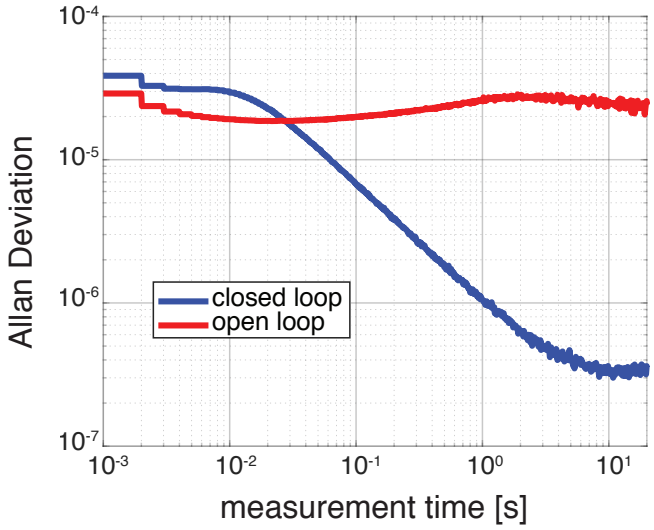


Figure 5.24: Allan deviation of the open-loop DCO (with a non-zero DSM input) and the closed-loop frequency reference

serving the noise transfer function of the DCO reveals that around a band just above the loop bandwidth, the DCO transfer function shows peaking. The peaking amplifies the DCO open-loop noise in this band when the loop is closed. It is possible to mitigate this effect by decreasing the loop bandwidth further; however, this would cause the ADEV floor to be reached with even longer measurement times.

5.6. Conclusion

In this chapter, an RC frequency architecture has been proposed to improve the accuracy of the previous Dual-RC frequency reference. The implemented 16 MHz reference is an FLL, which is locked to the phase shift of a low-TC WB filter, and is compensated by a resistor-based temperature sensor. A closed-loop trimming method is applied, allowing the temperature compensation polynomial coefficients to be accurately determined. An industrially feasible two-point trimming scheme achieves ±400 ppm inaccuracy from -45 °C to 85 °C. This improvement is due to the two bridges' inherent 2^{nd}-order nonlinearity cancellation. The inaccuracy improves to ±90 ppm with a nonlinear temperature compensation scheme based on a batch 6^{th}-order correction polynomial.

Compared to the Dual-RC reference and other state-of-the-art in Table

	This Work		[1]	Dual-RC [7]	[8]
Process [nm]	180			180	65
Frequency [MHz]	16			7	32
Power [W]	220µ			775µ	34µ
# of Trim Points	2	2 + Batch		2 + Batch	2
Inaccuracy [ppm]	±385	±90		±170	±530
Temp. Range [°C]	-45 to 85			-45 to 85	-45 to 85
# of Samples	20			12	6
ADEV [ppm]	0.32			0.33	2.5

Table 5.2: Performance summary and comparison to prior art

5.2, improvements in accuracy can be observed. In both two-point and two-point + batch cases, the frequency reference outperforms comparable work, with the most significant progress achieved by the two-point + batch compensation. It is the first reported sub-100 ppm inaccuracy RC frequency reference.

Another significant contribution of the work is the proposed trimming methodology. During the trimming operation, a temperature measurement is not utilized. Any error in the measurement of this reference temperature would add errors to the final trimming polynomials, which decreased the absolute reference accuracy; however, temperature sensors in industrial automated test equipment show about ±1 °C inaccuracy. In the improved closed-loop trimming methodology, the only measurement required is a comparison of the output frequency of the DCO to a reference, implemented simply via a frequency counter. Alleviating the requirement for an accurate temperature reference increases the industrial applicability of the proposed RC frequency reference.

References

[1] C. Gürleyük, S. Pan, and K. A. A. Makinwa, *A 16 MHz CMOS RC Frequency Reference With ±90 ppm Inaccuracy From −45 °C to 85 °C,* IEEE Journal of Solid-State Circuits **57**, 2429 (2022).

[2] S. M. Kashmiri, M. A. P. Pertijs, and K. A. A. Makinwa, *A Thermal-Diffusivity-Based Frequency Reference in Standard CMOS With an Absolute Inaccuracy of ±0.1% From -55 °C to 125 °C,* IEEE Journal of Solid-State Circuits **45**, 2510 (2010).

[3] E. O. Ates, A. Ergul, and D. Y. Aksin, *Fully Integrated Frequency Reference With 1.7 ppm Temperature Accuracy Within 0–80°C,* IEEE Journal of Solid-State Circuits **48**, 2850 (2013).

[4] M. H. Perrott, J. C. Salvia, F. S. Lee, A. Partridge, S. Mukherjee, C. Arft, J. Kim, N. Arumugam, P. Gupta, S. Tabatabaei, S. Pamarti, H. Lee, and F. Assaderaghi, *A Temperature-to-Digital Converter for a MEMS-Based Programmable Oscillator With < ±0.5-ppm Frequency Stability and < 1-ps Integrated Jitter,* IEEE Journal of Solid-State Circuits **48**, 276 (2013).

[5] S. Pan, C. Gürleyük, M. F. Pimenta, and K. A. A. Makinwa, *10.3 A 0.12mm2 Wien-Bridge Temperature Sensor with 0.1°C (3σ) Inaccuracy from -40°C to 180°C,* in *2019 IEEE International Solid- State Circuits Conference - (ISSCC)* (2019) pp. 184–186.

[6] S. Pan, Y. Luo, S. Heidary Shalmany, and K. A. A. Makinwa, *A Resistor-Based Temperature Sensor With a 0.13 pJ · K2 Resolution FoM,* IEEE Journal of Solid-State Circuits **53**, 164 (2018).

[7] C. Gürleyük, L. Pedalà, S. Pan, F. Sebastiano, and K. A. A. Makinwa, *A CMOS Dual-RC Frequency Reference With ±200-ppm Inaccuracy From −45 °C to 85 °C,* IEEE Journal of Solid-State Circuits **53**, 3386 (2018).

[8] A. Khashaba, J. Zhu, M. Ahmed, N. Pal, and P. K. Hanumolu, *3.5 A 34μW 32MHz RC Oscillator with ±530ppm Inaccuracy from −40°C to 85°C and 80ppm/V Supply Sensitivity Enabled by Pulse-Density Modulated Resistors,* in *2020 IEEE International Solid- State Circuits Conference - (ISSCC)* (2020) pp. 66–68.

6

Conclusion

This thesis has described the design, implementation, and characterization of integrated frequency references based on RC time constants. The two presented designs each achieved a significant stride in reducing the state-of-the-art inaccuracy in RC frequency references. These results were possible by implementing a digital frequency-locked loop that extracts, linearizes, and locks to the time constant of an RC filter. This chapter summarizes the main findings of the work and proposes directions for future work to improve the accuracy of integrated frequency references.

6.1. Main Findings

Traditional oscillators based on RC time constants do not reach the accuracy limit of their passive components. Typical harmonic and relaxation oscillators have inaccuracies of ±10000 ppm in the industrial temperature range. The origin of this inaccuracy is generally the circuitry required to sustain oscillation with the lossy RC time constant.

FLLs with accurate frequency-to-digital converters based on synchronous demodulators enable high-accuracy RC frequency references. Frequency-to-digital converters can extract the frequency error of an input signal and can serve as the basis for building digital frequency-locked loops. Digitizing the frequency error signal enables the use of digital correction techniques. These can address both systematic effects, such as cosine nonlinearity (chapter 3), and enable trimming and correction of the higher-order temperature coefficients of resistors (chapters 4 & 5).

A high-accuracy frequency reference can be built with a combination of high-TC resistors with the Dual-RC configuration, where the outputs of two PDDSMs are combined in the digital domain to achieve a zero-TC frequency error signal. High temperature-coefficient passives are generally undesirable in frequency reference applications as they place a stringent requirement on the temperature sensor's accuracy. The Dual-RC frequency reference, however, combines the time constant of two complementary high-TC RC Wien Bridges in the digital domain, alleviating the need for a separate temperature sensor. This allows high-TC resistors to be used to build an accurate frequency reference. The Dual-RC system was implemented in a standard 0.18 μm CMOS process. It achieved an inaccuracy of ±200 ppm from -45 °C to 85 °C, representing a significant step in reducing the inaccuracy of RC frequency references to below ±0.1% (Chapter 4).

A reasonably accurate temperature sensor can achieve high-accuracy temperature compensation when a low-TC resistor is available to build an RC time constant. The relaxed requirements on the temperature sensor allow the use of a more straightforward temperature compensation architecture, similar to [1]. The temperature-compensated frequency reference was implemented in a standard 0.18 μm CMOS process (chapter 5). The simplified linear temperature compensation configuration achieves an inaccuracy of ±400 ppm from -45 °C to 85 °C without using a polynomial engine, enabling a significant reduction in the digital overhead of the design in chapter 4. With the addition of a higher-order polynomial temperature compensation, the achieved inaccuracy improves to ±90 ppm. This design represents the first RC frequency reference to achieve inaccuracy less than ±100 ppm.

Improving the accuracy of RC frequency references requires the use of improved techniques in measurement and trimming. Methodologies, where frequency references are trimmed using open-loop measurements (as conducted in the reference in Chapter 4), fall short as the target inaccuracy levels go below ±100 ppm from -45 °C to 85 °C. Replicating the operating condition becomes crucial in achieving an accurate measurement that can used to trim the devices. A closed-loop trimming methodology, as used in the design in Chapter 5, is required to achieve better accuracy.

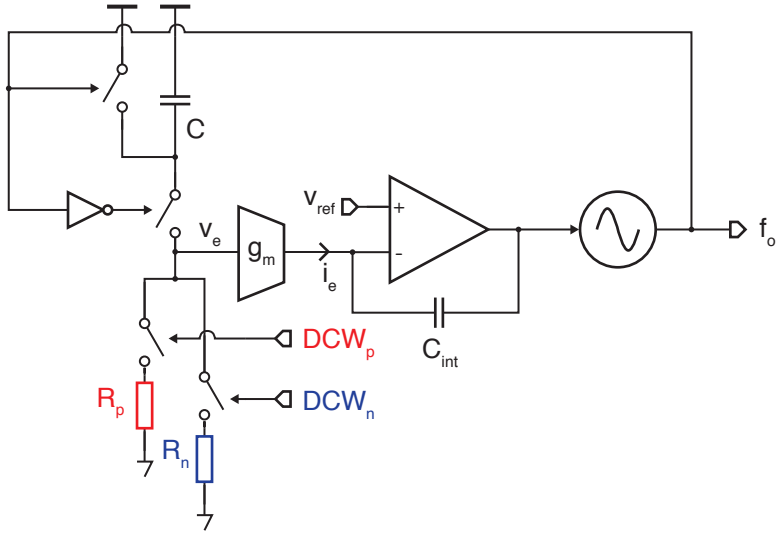


Figure 6.1: Switched-capacitor FLL featuring pulse density modulated resistors used in [2]

6.2. Contemporary Work

The interest in integrated frequency references has been widespread. During the course of this thesis, much contemporary work has been published that has improved the accuracy of RC frequency references. This section discusses two such bodies of work.

The design in [2] demonstrates the limits of traditional temperature compensation by integrating near-perfect digital trimming capability into a switched-capacitor FLL. As shown in Figure 6.1, the design features a switched-capacitor FLL, whose reference resistor is implemented by the parallel combination of two pulse density modulated resistors with complementary temperature coefficients. The digital control words (DCWs), obtained after post-production trimming, are encoded into single-bit pulse density streams via digital delta-sigma modulators. These streams define the ratio with which the two resistors are combined, allowing for precise trimming to achieve a ZTC resistor post-production. Effectively, the design implements digital *linear* temperature compensation, a per-sample trimmed ZTC resistor, and achieves 8.4 ppm/°C residual TC for six samples. As expected, the design is limited in accuracy by a strong second-order dependency on temperature, which is expected due to the linear temperature compensation scheme. Currently, this design represents the limits on ZTC resistors

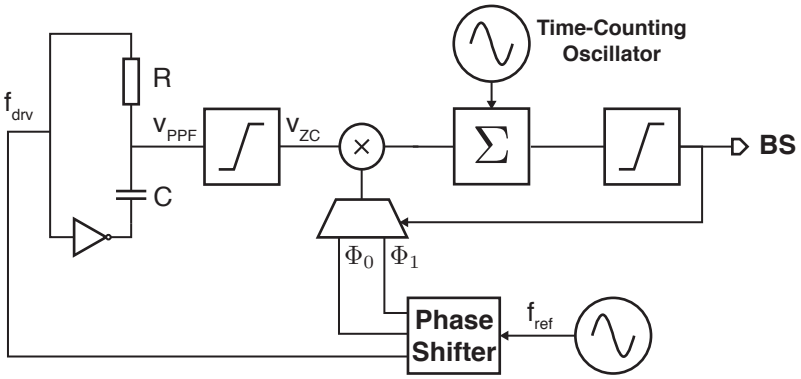


Figure 6.2: Poly-phase filter based phase-domain delta-sigma modulator used in [3]

implemented with polysilicon resistors.

FLLs, articularly those that implement complicated nonlinear temperature compensation schemes, show higher power consumption than their open-loop counterparts. The high power consumption can be attributed to the added power requirements of frequency-to-voltage converters and, in the latter case, the cost of performing high-resolution analog-to-digital conversion. A design that improved upon the Dual-RC reference in this respect [3] replaces the Wien Bridge with a poly-phase filter (PPF) and implements a highly digital readout to improve power consumption. The PPF, compared to the WB, shows higher frequency-to-phase sensitivity, which relaxes the specifications of the following readout circuitry. As shown in Figure 6.3, the design adopts a simple zero-crossing detector followed by a highly digitized version of the phase-domain delta-sigma modulator. In this highly digital DSM, the chopper demodulators are replaced by a digital XOR gate, and the analog integrator is replaced by a time-counting oscillator and digital accumulator. Moreover, the DCO control is implemented via a $\Delta\Sigma$ DAC, similar to the structure presented in 5, reducing its footprint. Replacing most analog components with their digital counterparts results in a dramatic reduction in area and power consumption for the design. Implemented in a 65 nm process, the design occupies only 0.06 mm^2 area and consumes 142 µW. While it shows great improvement in area and power, it does not show an improvement in accuracy: a residual TC of 2.56 ppm/°C is presented, a very similar level to the original dual-RC reference, as the complementary polysilicon resistors available in the 65nm process show very similar characteristics to the ones used in the work in this thesis.

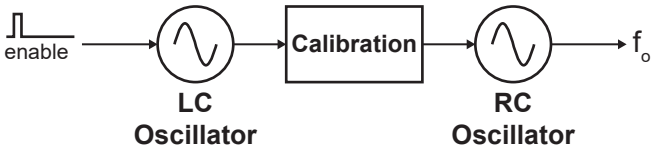


Figure 6.3: LC trained RC frequency reference [4]

A different aproach to reducing the inaccuracy of RC frequency refrences is presented in [4]. In this work, a temperature-compensated current-controlled ring oscillator is periodically calibrated to the frequency output of a co-integrated LC Colpitts oscillator. Duty cycling the LC oscillator mitigates the primary issue of its high power-consumption. The system consumes an always-on power of 210 µW, in line with RC frequency references. Its accuracy, on the other hand, is determined by the LC oscillator. The work achieves a residual TC of 0.7 ppm/°C with a 1-point trim, a result that is 2x better than the best RC reference with 2-point trim.

6.3. Future Work

This section discusses the unexplored issues in this thesis and possible future directions in the field of integrated frequency references.

6.3.1. Stress and Aging

The RC references presented in this thesis primarily addressed the inaccuracy caused by temperature variations. However, these are not the only external variations an electronic system faces in operation. The characterization of the devices should be extended to cover all possible environmental variations to accomplish practical industrial applicability of the solutions presented.

A well-known effect that causes a degradation in electrical components' accuracy is mechanical stress. Mechanical stresses develop on integrated circuits during the fabrication, packaging, and assembly processes, and these strains can cause shifts in the electrical parameters of integrated passives and actives. Since this work is based on resistors and capacitors, the effects of mechanical stress will undoubtedly impact the frequency accuracy of the

developed frequency references.

While it is not covered widely in literature, polysilicon resistors suffer from aging, which refers to a resistance shift from its designed value as the device remains in operation. Aging can be caused due to both electrical or thermal stress on the integrated passives. While thermal effects are demonstrated for more extensive ranges than typically experienced by consumer integrated circuits, the work in [5] implies that the accuracy levels required by this work might be difficult to maintain through the device's lifetime under thermal stress. A shift in resistance on the order of 0.1% that would be expected would dominate the inaccuracy and renders production trimming futile.

Accelerated aging tests conducted on frequency references based on p/n-poly resistors reveal a degradation of 5000 ppm and 2000 ppm variation, respectively [6]. Diffusion resistors perform significantly better, showing about 600 ppm variation [6]. Compensating for resistor aging is shown to be possible via duty cycling replica resistors to limit their current [7]. Further study should characterize the long-term aging effects on polysilicon resistors to assess their repeatability, which will enable further solutions.

Inspiration can be gathered to compensate for these environmental variations from the Dual-RC architecture. In the Dual-RC architecture, two channels with different sensitivities to temperature were combined to build a temperature-independent but frequency-dependent network. Bringing in a third free variable, such as stress, renders the Dual-RC incapable of compensating for it. Effectively, the Dual-RC system formed a system with two unknowns (frequency and temperature) and two equations (two Wien Bridge readouts with different TCs). A system that has to solve for three unknowns (i.e., frequency, temperature, and stress) would need three equations, requiring the addition of a third channel with its sensitivities adequately tuned to ensure cancellation. Such a design presents challenges in sensor design, where multiple channels now need to be tuned for multiple sensitivities. Moreover, in a three-channel system, ensuring loop stability becomes more challenging, as the coefficients used to cancel parasitic effects will change the frequency sensitivity of the system. However, a Tri-RC system might be successful in addressing the effects of stress and aging.

6.3.2. Switched-Capacitor FDCs

The foremost prerequisite for implementing nonlinear temperature compensation is to digitize a time constant. In this thesis, phase-domain delta-sigma modulators were adopted to achieve the digitization of the phase shift of integrated RC filters. However, due to their simpler architectures, it would be better to realize a frequency-to-digital converter with using a switched-

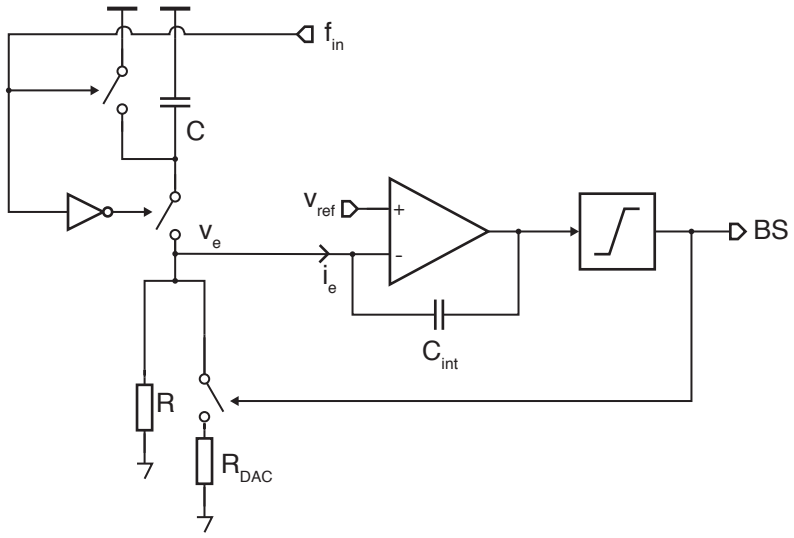


Figure 6.4: Conceptual circuit diagram for a switched-capacitor frequency-to-digital converter

capacitor core.

A conceptual circuit diagram for a switched-capacitor frequency-to-digital converter is given in Figure 6.4. The switched-capacitor is driven by the input frequency and its current is balanced by the resistor. The resistor current is a combination of a static and a controlled resistor (R_{DAC}) to cover the variation. The error current (i_e) is integrated and quantized into a bitstream. Since the error current must be zero in steady-state, the bitstream represents the switched-capacitor resistor's current, which is a function of the input frequency. Thus, frequency-to-digital conversion is implemented, and such a circuit can be used in a digital FLL.

Realizing such a compact frequency-to-digital converter might be beneficial to extend the work in [2] with nonlinear temperature compensation. A switched-capacitor frequency-to-digital converter can reduce the area and power cost of nonlinear temperature compensated frequency references.

6.3.3. LC Frequency References

A promising direction can be found in replacing the RC filter with an LC filter in a similar loop where the phase shift of the LC network is extracted. Traditional usage of an LC time constant is embedded in an oscillator, which degrades its frequency accuracy. Using an LC time constant in a frequency-

to-voltage conversion system could yield higher accuracy, but it presents challenges. Due to the difficulty in implementing high-value inductors on-chip, the resonance frequencies for LC tanks are generally in the GHz range. This frequency would imply a requirement for an even faster oscillator in the back end of an FLL system. While it still might not result in a low-power system due to the high-frequency oscillator requirement, such FVCs might be the next step in the evolution of integrated frequency references.

References

[1] S. M. Kashmiri, M. A. P. Pertijs, and K. A. A. Makinwa, *A Thermal-Diffusivity-Based Frequency Reference in Standard CMOS With an Absolute Inaccuracy of ±0.1% From -55 °C to 125 °C,* IEEE Journal of Solid-State Circuits **45**, 2510 (2010).

[2] A. Khashaba, J. Zhu, N. Pal, M. G. Ahmed, and P. K. Hanumolu, *A 32-MHz, 34-µW Temperature-Compensated RC Oscillator Using Pulse Density Modulated Resistors,* IEEE Journal of Solid-State Circuits **57**, 1470 (2022).

[3] W. Choi, J. Angevare, I. Park, K. A. A. Makinwa, and Y. Chae, *A 0.9-V 28-MHz Highly Digital CMOS Dual-RC Frequency Reference With ±200 ppm Inaccuracy From −40 °C to 85 °C,* IEEE Journal of Solid-State Circuits **57**, 2418 (2022).

[4] A. S. Delke, T. J. Hoen, A.-J. Annema, Y. Jin, J. Verlinden, and B. Nauta, *A Single-Trim Frequency Reference System With 0.7 ppm/°C From −63 °C to 165 °C Consuming 210 µW at 70 MHz,* IEEE Journal of Solid-State Circuits **58**, 2585 (2023).

[5] M. Ehmann, F. Schubert, P. Ruther, and O. Paul, *Thermally activated ageing of polysilicon,* in *SENSORS, 2002 IEEE*, Vol. 1 (2002) pp. 602–606 vol.1.

[6] S. Pan, X. An, Z. Yu, H. Jiang, and K. A. A. Makinwa, *A Compact 10-MHz RC Frequency Reference With a Versatile Temperature Compensation Scheme,* IEEE Journal of Solid-State Circuits **58**, 3450 (2023).

[7] K.-S. Park, N. Pal, Y. Li, R. Xia, T. Wang, A. Abdelrahman, and P. K. Hanumolu, *A Temperature- and Aging-Compensated RC Oscillator With ±1030-ppm Inaccuracy From -40 °C to 85 °C After Accelerated Aging for 500 h at 125 °C,* IEEE Journal of Solid-State Circuits **58**, 3459 (2023).

A

Integrated Frequency Reference Survey

The integrated frequency reference survey introduced in section 2.5 aimed to quantify published work until 2017 to justify the architectural choices that gave direction to this thesis. In this section, the data for the survey is extended to include designs up to 2024, the year in which the work concluded. The performance of 52 published frequency references (including the designs discussed in this chapter) have been tabulated. The data can be accessed from `ttps://docs.google.com/spreadsheets/d/1FfJPQdaqo3QUu59OJcdIPdSXEw_x8txaP-fsyEmuip8`.

In this survey, certain normalized parameters have been derived from the published results to achieve a fair comparison of the featured designs. One such parameter is the residual temperature coefficient. Since most designs do not share a standard temperature range in which they report accuracy, the reported values need to be normalized. The residual temperature coefficient uses the *box method*; it is derived by dividing the relative peak-to-peak frequency variation by the entire temperature range.

Figure A.1 shows the residual temperature coefficient and frequency-normalized power consumption of integrated frequency references grouped by choice of time constant. Comparison to Figure 2.26 reveals that recent development has been focused on RC, and recent work has been targeting the high-accuracy regime. These recent RC references were successful in reducing inaccuracy to levels previously attainable only by LC. However, it is worth noting that LC still maintains the highest accuracy in integrated frequency references, albeit with a smaller margin.

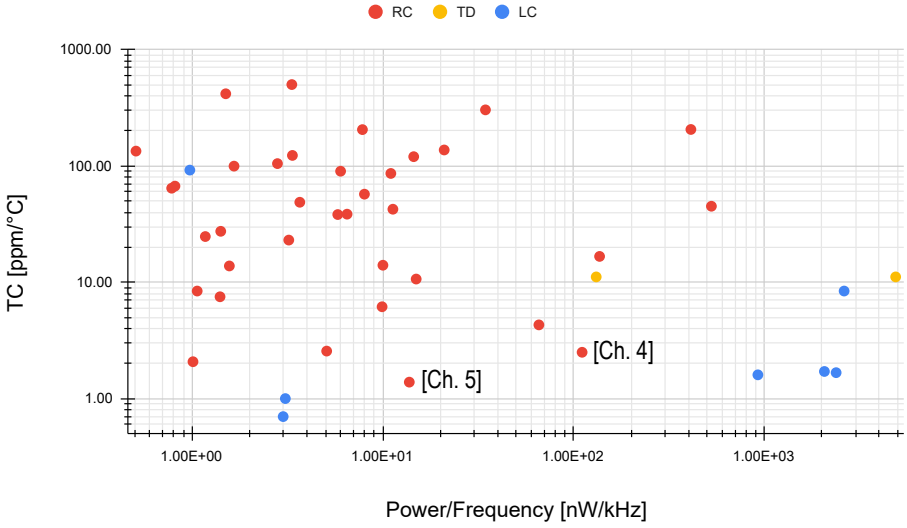


Figure A.1: Temperature coefficient versus power consumption of integrated frequency references, grouped by choice of time constant

The improvement in the accuracy of integrated RC references can be attributed to two main factors. The first is the widespread adoption of FLLs. Figure 2.27 had predicted that FLLs would show better accuracy, due to their ability to separate the problem of maintaining oscillation from achieving accuracy. Adding recent work to the same graph (Figure A.2) reveals this prediction to be correct, as most designs that contributed to improving the accuracy of RC references have adopted closed-loop architectures. The second factor is nonlinear temperature compensation. A virtual barrier of 5 ppm/°C divides the RC frequency references into two groups: linear (ZTC) and nonlinear temperature compensated. This ZTC barrier emerges from the nonlinear temperature dependency of integrated polysilicon resistors, which linear temperature compensation cannot address. Digital nonlinear temperature compensation has been they key factor in improving accuracy beyond this virtual barrier.

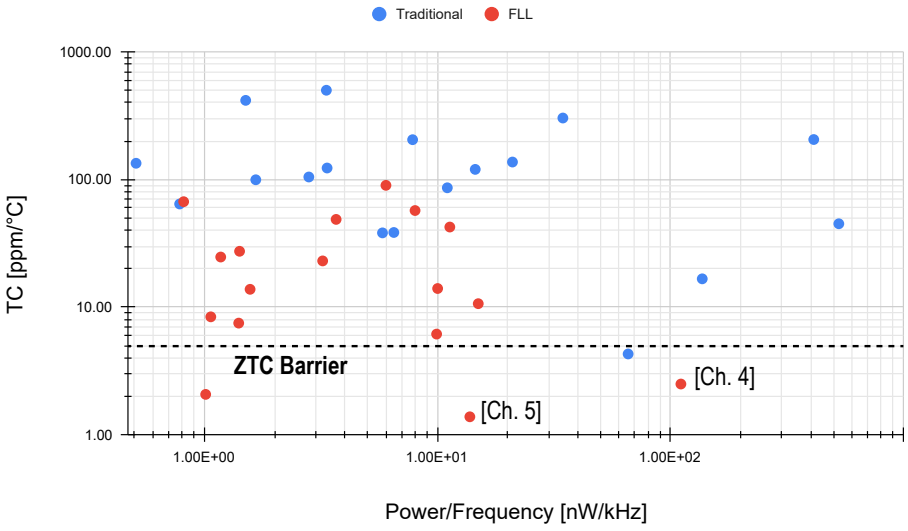


Figure A.2: Temperature coefficient versus power consumption of integrated RC frequency references, grouped by choice of architecture

List of Publications

Journal Papers

- **Ç. Gürleyük**, L. Pedalà, S. Pan, F. Sebastiano and K. A. A. Makinwa, *A CMOS Dual-RC Frequency Reference With ±200-ppm Inaccuracy From −45 °C to 85 °C*, IEEE Journal of Solid-State Circuits, vol. 53, no. 12, pp. 3386-3395, Dec. 2018.
- M. Pimenta, **Ç. Gürleyük**, P. Walsh, D. O'Keeffe, M. Babaie and K. A. A. Makinwa, *A 200-µW Interface for High-Resolution Eddy-Current Displacement Sensors*, IEEE Journal of Solid-State Circuits, vol. 56, no. 4, pp. 1036-1045, April 2021.
- **Ç. Gürleyük**, S. Pan and K. A. A. Makinwa, *A 16 MHz CMOS RC Frequency Reference With ±90 ppm Inaccuracy From −45 °C to 85 °C*, IEEE Journal of Solid-State Circuits, vol. 57, no. 8, pp. 2429-2437, Aug. 2022.

Conference Papers

- L. Pedalà, **Ç. Gürleyük**, S. Pan, F. Sebastiano and K. A. A. Makinwa, *A frequency-locked loop based on an oxide electrothermal filter in standard CMOS*, ESSCIRC 2017 - 43rd IEEE European Solid State Circuits Conference, Leuven, Belgium, 2017.
- **Ç. Gürleyük**, L. Pedala, F. Sebastiano and K. A. A. Makinwa, *A CMOS Dual-RC frequency reference with ±250ppm inaccuracy from −45°C to 85°C*, 2018 IEEE International Solid-State Circuits Conference - (ISSCC), San Francisco, CA, USA, 2018.
- S. Pan, **Ç. Gürleyük**, M. F. Pimenta and K. A. A. Makinwa, *10.3 A 0.12mm2 Wien-Bridge Temperature Sensor with 0.1°C (3σ) Inaccuracy from -40°C to 180°C*, 2019 IEEE International Solid-State Circuits Conference - (ISSCC), San Francisco, CA, USA, 2019.
- **Ç. Gürleyük**, S. Pan and K. A. A. Makinwa, *3.4 A 16MHz CMOS RC Frequency Reference with ±400ppm Inaccuracy from −45°C to 85°C After Digital Linear Temperature Compensation,* 2020 IEEE International Solid-State Circuits Conference - (ISSCC), San Francisco, CA, USA, 2020.
- M. Pimenta, **Ç. Gürleyük**, P. Walsh, D. O'Keeffe, M. Babaie and K. Makinwa, *A 200µW Eddy Current Displacement Sensor with 6.7nmRMS Resolution,* 2020 IEEE Symposium on VLSI Circuits, Honolulu, HI, USA, 2020.

- H. Jiang, S. Pan, **Ç. Gürleyük** and K. A. A. Makinwa, *31.3 A 0.14mm2 16MHz CMOS RC Frequency Reference with a 1-Point Trimmed Inaccuracy of ±400ppm from -45°C to 85°C,* 2021 IEEE International Solid-State Circuits Conference (ISSCC), San Francisco, CA, USA, 2021.

Patents

- P. M. Walsh, D. MacSweeney, D. O'Keeffe, K. Makinwa, M. Pimenta, D. Seguine, **Ç. Gürleyük**, *High performance inductive sensing all digital phase locked loop*, US Patent US11552635B2, 2023.

Acknowledgements

The Ph.D. journey that culminated in this thesis lasted almost nine years, and it involved more people than myself. I would like to express my gratitude to those who have directly or indirectly contributed to this work.

Foremost, I would like to thank my promotor, Kofi Makinwa. His high standards and exceptional direction have driven this work to success. His insight into electronic systems and his exceptional sieve of ideas have kept me inspired throughout my tenure with him and beyond. I left his group a much better engineer than when I joined.

Secondly, I would like to thank those whom I collaborated with directly on the frequency references in this thesis. Fabio Sebastiano provided much-needed guidance during my first couple of years in the EI group. Lorenzo Pedalà and I made a great complementary team for the few years we worked together. His exceptional analytical mindset enabled us to figure out the intricacies of the Dual-RC system. Sining Pan's exceptional work on resistor-based temperature sensors served as the stepping stone in building frequency references. I feel very lucky that my time coincided with his in the EI group. Zu-Yao and Lukasz have provided invaluable lab support. Also, thanks to Roberto Nonis and Infineon Technologies for sponsoring this work and for their hospitality during my time in Villach.

Finally, I would like to thank my colleagues in the EI group. Jan Angevare completed the triumvirate of the 14th floor, and he was an inspiring colleague and a great friend. Burak Gönen instigated me joining the group, and has been a good friend and collague during our time there and beyond. My colleagues Pierluigi, Bahman, Ugur, Thije, Jeroen, Shoubik, Matheus, and everyone else in the EI group that I fail to mention here have made my time in EWI very enjoyable.

Lastly, I'm very grateful for my family: my partner Nergis, and our little boy Volkan.

About the Author

Çağrı Gürleyük was born in Istanbul, Türkiye, in 1987. He received his B.Sc. and M.Sc. degrees in electronics engineering from Istanbul Technical University, Istanbul, Türkiye, in 2011 and 2015, respectively. He pursued his Ph.D. degree with the Electronic Instrumentation group at Delft University of Technology, Delft, The Netherlands, from 2016 to 2020.

From 2014 to 2016, he was a Design Engineer with Analog Devices, Istanbul, Türkiye, where he worked in the high data-rate group on low-noise transimpedance amplifiers for LIDAR applications. Between 2020 and 2024, he was with Ethernovia, Zeist, The Netherlands, as a Senior Member of Technical Staff, working on automotive multi-gigabit ethernet transceivers. In 2024, he joined Optiver, Amsterdam, The Netherlands, working on low-latency trading systems.

His research interests include (but are not limited to) wireline communication, frequency generation, and mixed-signal control systems. He serves as a reviewer for IEEE JSSC and TCAS. He was a recipient of the ADI Outstanding Student Designer award in 2020.