Monolithic Wireless Transceiver Design

Filip Maksimovic
Kristofer Pister, Ed.
Ali Niknejad, Ed.

Electrical Engineering and Computer Sciences
University of California at Berkeley

May 1, 2020

Monolithic Wireless Transceiver Design

by

Filip Maksimovic

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Engineering - Electrical Engineering and Computer Science

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Kristofer S.J. Pister, Chair
Professor Ali M. Niknejad
Professor Stephen D. Glaser

Fall 2018

The dissertation of Filip Maksimovic, titled Monolithic Wireless Transceiver Design, is approved:

Chair   _____   Date   _____

            _____   Date   _____

            _____   Date   _____

University of California, Berkeley

Monolithic Wireless Transceiver Design

Abstract

Monolithic Wireless Transceiver Design

by

Filip Maksimovic

Doctor of Philosophy in Engineering - Electrical Engineering and Computer Science

University of California, Berkeley

Professor Kristofer S.J. Pister, Chair

Recently, there has been an increasing push to make everything wireless. In contrast to high-performance cellular communication, where the demand for enormous quantities of data is skyrocketing, these small wireless sensor and actuator nodes require low power, low cost, and a high degree of system integration. A typical CMOS system-on-chip requires a number of off-chip components for proper operation, namely, a crystal oscillator to act as an accurate frequency reference, and an antenna. The primary goal of this thesis is to address the hurdles associated with operating without these components at as low a power level as possible. This is a step towards the ubiquitous presence of wireless communication.

In this work, an evaluation of transceiver performance is performed with power, performance, and physical size in mind. Operation of a low-power standards compatible 2.4 GHz transmitter (TX) is demonstrated without the use of an off-chip frequency reference. These 2.4 GHz transceivers (TRX), called the single chip motes, operate at low power levels without an off-chip frequency reference. The first single chip mote demonstrated RF chip-to-chip communication in the presence of local oscillator drift caused by temperature variation. It used a free-running LC tank oscillator that was calibrated against drift with periodic network traffic. The next single chip mote was a 2.4 GHz, 802.15.4 TRX, BLE advertising TX system-on-chip with integrated digital baseband and a Cortex M0. Once again, the chip uses no off-chip frequency reference. Finally, a design of high frequency transceiver with integrated antenna is presented, paving the way for a fully on-chip solution.

To no one in particular

Contents

List of Figures

List of Tables

Acknowledgments

First and foremost, I would like to thank Kris and Ali, for being my advisors. Kris was always so enthusiastic about research and so genuinely excited when things worked. Ali would always help debug difficult circuits problems and explain complicated subjects to me. His mindset and group culture made me a better designer, and his patience and positive demeanor were traits that I will always try to aspire to. Other influential faculty were Elad Alon, who helped teach me circuits, and Anant Sahai, who was the instructor (with Ali) for EE 16A, one of the most enjoyable experiences of my PhD. I would also like to thank Steve Glaser for serving on my qualifying exam and dissertation committees, and for asking the hardest question in my qualifying exam.

Brad Wheeler and I worked together almost every day for the past five years, and it was an enormous priviledge. If I could put a second author on my dissertation, I would put his. I'd also like to thank the other members of the SCμM team, David Burnett, Osama Khan, and Sahar Mesri. The single chip motes were ambitious, by academic standards, and I'm not sure if it would have been possible with a different group of people. Towards the end of my PhD I also had the opportunity to work with Sashank Krishnamurthy who is a phenomenal designer.

Even though I never worked with this next group of people in a technical capacity, their importance to my graduate school experience cannot be understated. Joey Greenspun, Claire Lochner, and Ozzy LaCaille, thank you for being my very best friends. Joey, your oppressive optimism never went unappreciated. Claire, it was really fun to grow up with you (in the graduate student sense). And Ozzy, thanks for all of the venting and bar-napkin circuits. Go Nuggets! Thank you roommates, Luke Latimer, Daniel Gerber, Andrew Townley, Stephen Twigg, and Joe Corea, for always keeping it real, and for always being down for a beer after a rough day, which was almost every day. I also owe thanks to a number of students now far-removed from Berkeley: Matthew Spencer, Mike Lorek, Stephan Adams, Hanh Phuc Le, and Alberto Puggelli. Each one of you helped me find my bearings early on in grad school. And Alberto, thanks for getting me out of that bit of trouble in Italy. I owe you one. There are also members of the Pister and Niknejad groups who I have not thanked elsewhere. Lorenzo, Yi-An, Nima, and Bo from Ali's group, and Lydia, Alex, Dan and Daniel, Hani, Craig, Brian, and Nathan from Kris's group.

Legendary wizard Archibald Harris Jr. would never allow me to submit this dissertation without thanking his companions: Kody (Joey), Nara (Claire), Deeba (Daniel Contreras), Wakunda (Balthazar Lechene), and dungeon master Daniel Drew. The "weekly" pizza and escapism helped me a great deal. Over the year, Archibald became a part of me, and I have all of you to thank (or maybe curse) for it.

When you see a fence post with a turtle post balanced on top, you know he didn't get up there by himself. So, thank you, mom and dad, for putting me on the post. Thank you brothers, Nikola and Stevan, for always laughing at my terrible attempts at humor. And, of course, thank you, Rale and Djole. I'll see you both again soon. Konstantinovici (Duda, Neda, Veca, i Mile) volim vas sve.

Chapter 1

Introduction

1.1 Wireless Communication in the 21st Century

Wireless connectivity has evolved considerably from the first data transmissions of Hertz, Tesla, and Marconi. In 2017, 4.66 billion people own a mobile phone [1]. Cellular technology and its rapidly increasing throughput demands fuel technological innovation and development of new, high data rate, long range wireless communications. The proposed mm-wave backhaul network necessary for 5G is one example. Verizon and AT&T's purchase of significant swaths of licensed spectrum in the 28-39 GHz band confirms this trend. Lying in stark contrast to the high-performance cat-video-streaming wireless networks (like LTE, 5G, and modern WLAN), lurks the increasing presence of low power wireless devices. These devices, often referred to as part of the so-called "Internet of Things" consume minimal power, communicate with low data rates, should be physically small, and cost infinitesimally little. And, of course, there will be a lot of them. Over twenty billion by 2020, if you believe [2]. The enormous Mirai DDOS attack [3] that crippled Dyn DNS (that serves Amazon Cloud Services, among other popular web services and sites) in October of 2016 was fueled by a botnet of millions of compromised wireless devices. From an economic perspective, recent market analyses predict upwards of $300 million revenue for the IoT industry [4], which is comparable to the advanced energy industry in the United States [5].

Although current commercial applications of these wireless devices have gained little traction, industrial, urban, and medical applications show some promise. In the oil refining industry, for example, [6] has demonstrated automatic localization of gas leaks with 1 meter spacial resolution using a distributed network of gas sensors. In medicine and personal health, [7] has developed a wireless ingestible gas sensor that can estimate diets and assist with diagnosing gut disorders. Similarly, [8] has demonstrated a flexible, wireless, wearable medical device for constant monitoring of multiple vital statistics. And, finally, [9] describes the Padova Smart City project in which a network of wireless sensors attached to light poles throughout the municipality of Padova, Italy, collect sensor data (such as CO level, air temperature, humidity, vibrations, and sound levels). None of these technologies would

be possible without miniature (for ease of integration) and low-power (for increased battery life) wireless communication.

1.2 Radio Miniaturization

Gartner's hype assesment of IoT cites low-cost development boards as an integral technology for accelerating IoT beyond inflated expecations (I guess we're there now) to the trough of disillusionment [2]. If, instead of a development board, everything needed for a wireless transceiver were integrated on a single piece of silicon, the price would be negligible - as little as < \$0.05 per chip if produced in volume (approximated from [10]) - and the device could be interfaced easily with a sensor as part of a consumer product or industrial monitoring network. The ultimate goal of the work presented in this thesis is to build a single chip platform that could easily be used for any of the forementioned applications, and, potentially, unlock new, previously unexplored use cases.
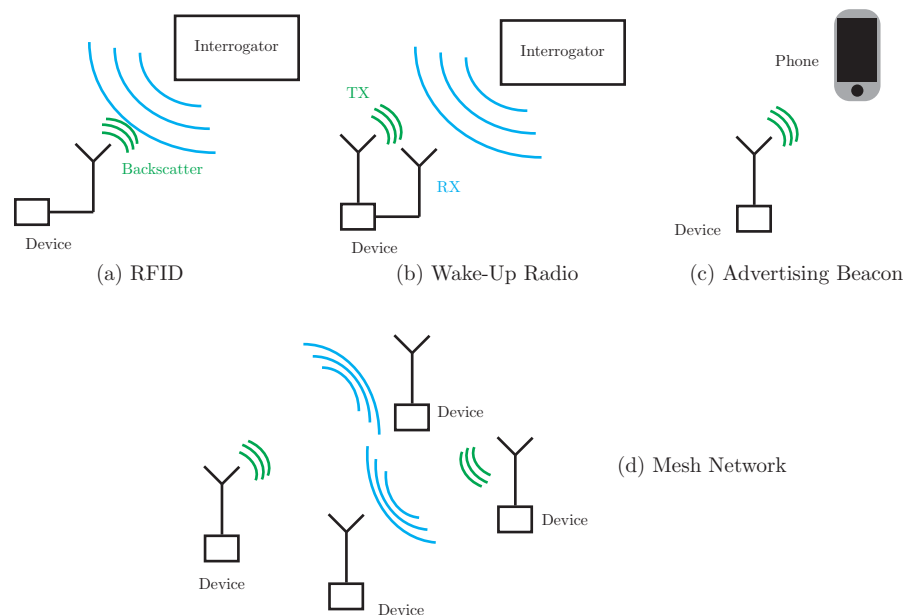
Figure 1.1: An illustration demonstrating various communication modalities for wireless sensor deployments

A number of potential types of wireless communication have been proposed for various applications, as shown in Fig. 1.1. RFID typically operates with either backscattering (modulation of the antenna impedance that creates a reflection) or by energy harvesting of the interrogator's RF transmission and subsequent side-channel transmission. The benefit is that the device can be completely passive, and consume zero power, when not in use. All

of the energy in the system is provided by the high-power interrogator. The downside, of course, is that the interrogator is necessary for any communication, and that the data rates are extremely low if a high degree of power transfer is needed [11]. In a similar vein, wake-up radios [12] and [13] can operate at extremely low power levels and provide a similar function to RFID tags, but with higher performance. The interrogator sends a coded sequence. If the wake-up portion of the receiver hears it, it activates the dormant higher-power transceiver to communicate to the interrogator. This has two significant disadvantages in addition to requiring an interrogator: non-zero sleep power (the receiver is always listening), and the wake-up receiver needs to have similar sensitivity to the high-power receiver to reduce the probability of either erroneous wake-up or a missed reception (a proverbial hitting of the snooze button). In the advertising beacon example, the device acts exclusively as a transmitter and periodically sends packets that can be detected either by a specialized receiver or, more commonly, existing commercial hardware. This device would not require specialized equipment, but, unless heavily duty cycled (a tradeoff against data rate) could burn significant power. And, if compatibility with cellular phones or routers is needed for the application, it must adhere to wireless standards.

The mesh network is more useful in broad sensor deployments, such as in [6] and, in addition to generally requiring compatibility with existing wireless standards (both for reliability and for integration with existing infrastructure) requires symmetrical communication. Every node in the mesh must be able to transmit and receive data from other identical nodes.

1.3 Wireless Standards

Fortunately, many standards already exist both for mesh networking and for beacon-style advertising. Typically, RFID is simple enough that standards are not needed, but the FCC does put limits on transmitted power in the unlicensed portions of the spectrum. Some wake-up standards have been proposed [14] but, if the radios do adhere to any particular standard, are generally built with IEEE 802.11 in mind. I will briefly summarize the requirements of two widely-used wireless communication standards, Bluetooth Low Energy, and 802.15.4, relevant to transmitter and frequency synthesizer design, as they are the dominant power consumers in a wireless transceiver. The decision to operate in accordance with existing standards was made for two reasons. The first was integration with existing hardware. On a fundamental level, the purpose of these transceivers was to make every sensor node tiny, low-power, and wireless. And that becomes a considerably more attractive proposition if these wireless sensors can exist in established networks and operate in conjunction with off-the-shelf hardware. The second reason is the stack that exists as part of many of these standards. Bluetooth, for instance, has become the standard for low datarate and low power interaction between wireless handsets and wireless peripherals. As a result, if the transceiver is compatible with Bluetooth, the device can communicate to phones and computers. IEEE 802.15.4 can be combined with various protocol stacks (such as WirelessHART, Zigbee, or OpenWSN) to connect a sensor to the internet.
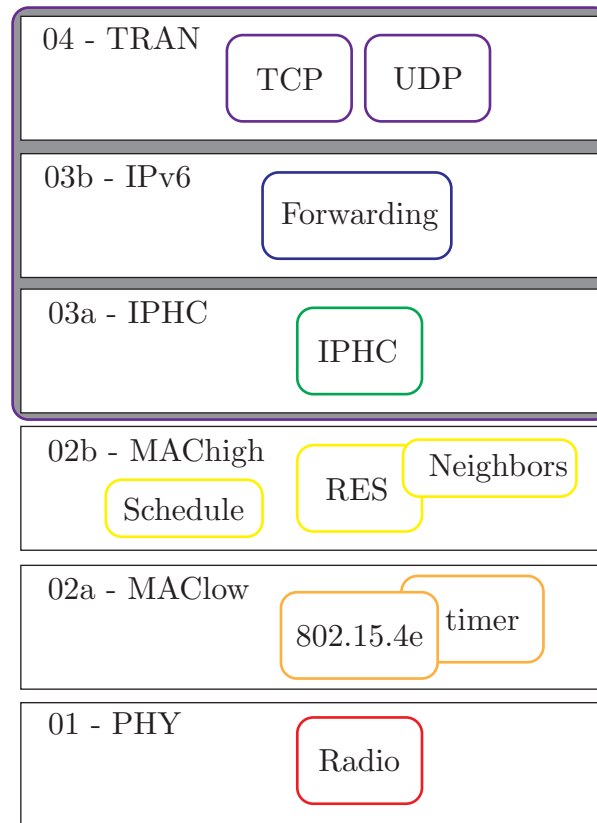
Figure 1.2: The OpenWSN stack [15]. As long as the wireless transceiver adheres to the requirements set by the 802.15.4e standard, and as long as on-chip digital hardware is capable of running the MAC and protocol layer packet assembly (TX) and disassembly (RX) and the scheduling, the device can operate in an internet-connected wireless sensor network

Bluetooth Low Energy

Bluetooth Low Energy (co-compatible with the more common but more power intensive Bluetooth) is a 1 Mbps data rate link that uses Gaussian pulse-shaped frequency shift key with a modulation index of between 0.45 and 0.55. The standard mandates at most ± 50 ppm error on the transmitter clock. However, most commercial Bluetooth receivers can demodulate with data rate errors of nearly 500 Hz. BLE channels are spaced equally across the 2.4 GHz ISM band (between 2.4 GHz and 2.485 GHz) with 2 MHz channel spacing. This is a summary of relevant PHY-layer specifications for BLE:

- Minimum Output power of -20 dBm, maximum output power of +20 dBm (depends on class of device, ranges from +20 dBm for a class 1 device to 0 dBm for a class 3

device)

- Modulation is Gaussian Frequency Shift Key (GFSK) with a modulation index between 0.45 and 0.55.

- The frequency deviation during a 1010 sequence will be within 80% of a 00001111 sequence.

- The minimum frequency deviation (at 1 Msym/s) is 185 kHz and the maximum is 370 kHz.

- Symbol timing accuracy is $\pm$ 50 ppm.

- Zero crossing error must be less than $\pm\ 1/8^{th}$ of a symbol period.

- FCC part 15.247 rules: the 6 dB bandwidth of the transmitter spectrum must be 500 kHz.

- At 1 Msym/s modulation, spurious power at 2 MHz away from the channel must be, at maximum, an absolute -20 dBm. At 3 MHz or further, it must be, at maximum, and absolute -30 dBm.

- At 2 Msym/s modulation (fast data rate), spurious power at 4 MHz away must be, at a maximum, an absolute -20 dBm. At 5 MHz, it must also be less than -20 dBm. At 6 MHz or greater, it must be -30 dBm.

- Deviation of center frequency cannot exceed $\pm$ 150 kHz including both starting frequency and drift. During a packet, the frequency cannot drift more than 50 kHz, and the drift rate must be less than 400 Hz/μs.

802.15.4, OpenWSN

802.15.4 is an IEEE-maintained standard designed for use in low data rate personal area networks. It is a 2 Mbps data rate link modulated by orthogonal quadrature phase shift key (equivalent to minimum shift key, a special case of frequency shift key with modulation index of 0.5 [16]). The standard mandates at most $\pm$40 ppm of error on the 2 MHz data clock. However, recent network-level innovations [17] have demonstrated a high degree of timing accuracy even without a high quality timing reference. 802.15.4 is also spread evenly across the 2.4 GHz ISM band with 5 MHz channel spacing.

Various network protocols exist on top of the 802.15.4 standard, primarily, WirelessHART, Zigbee, and OpenWSN [18]. These complete network stacks are designed specifically for multi-hop wide-area mesh networks. And, all three of them use time and frequency channel diversity in the form of a schedule. That way, many devices can communicate either simultaneously on different channels (FDMA) or at different times on the same channel (TDMA)

without interfering with one another. This schedule requires a high degree of channel selectivity in the presence of varying environmental conditions, and precise timekeeping over potentially long periods (network data rate dependent). Here I summarize relevant PHY-layer specifications for 802.15.4.

- Sixteen channels with channel spacing of 5 MHz.

- Modulation is offset-quadrature phase-shift key (equivalent to MSK)

- Modulation is performed at 2 Mchips/s (which is derived from an actual data rate of 250 kb/s)

- There is a relative limit of -20 dB (relative to carrier power) at a distance of 3.5 MHz from the channel

- The phase noise of the local oscillator (derived from receiver specifications) is -102 dBc/Hz at a 3.5 MHz offset, and -103.5 dBc/Hz at a 10 MHz offset [19]

Note that for the summaries of both BLE and 802.15.4, some of the requirements come directly from the specification document, and some are derived from specifications to translate more easily to circuit-level requirements.

WiFi

WLAN is currently adopting an offshoot of its wireless standard intended for compatibility with low-performance devices [20]. This is an appealing target, as WiFi is present in nearly all indoor environments. While [21] demonstrated a method to fake a minimalist 802.11a packet using exclusively FSK modulation, 802.11 is fundamentally based on a constellation of phase and amplitude modulation whose EVM specifications are difficult to meet without a precise timing or frequency reference. In addition, aside from the simple BPSK case, uses variable amplitude modulation, which considerably complicates transmitter and power amplifier design.

1.4   A Typical Mote

If the goal is to integrate a radio onto a single CMOS die, all of the external components associated with a typical mote must be either removed or worked around. The crystal is used as a precise time and frequency reference for phase-locked and frequency-locked loops. Unfortunately, because of the difficulty of performing energy harvesting to the degree needed to power a radio, an external power source is still needed. However, it is still important to reduce power as much as possible so that the lifetime of this power source is maximized. The antenna serves to transfer power from an electrical signal in a conductor to a propagating electromagnetic wave. Its size compared to the wavelength of the carrier frequency is an integral part of its efficiency.

2.4 GHz Antenna + match

Crystal

External
Battery

Radio

Passives
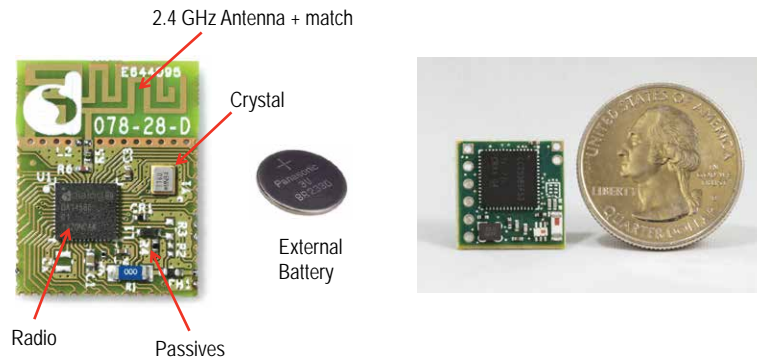
Figure 1.3: Board-level implementations of wireless communication nodes. Left is [22] a typical commercial wireless node. Right is [23] the smallest wireless node in research

Timing and Frequency Specificity

Removing the crystal means that the radio no longer has a timing reference that is insensitive to variations in temperature and voltage. In addition, no low-power CMOS timer exists that has jitter or frequency drift comparable to a crystal oscillator (this is a consequence of the crystal's high mechanical Q factor). This has dire consequences for scheduled mesh networks, as every node needs to have an accurate knowledge of the global network time. In addition, without an accurate data clock, transceiver performance is compromised (EVM is degraded for the transmitter, and sensitivity is degraded depending on the CDR bandwidth of the receiver). Some work has been done on network-level compensation of low frequency timers [17] and frequency synthesizers [24].

Removing the crystal reference has the added benefit of reducing system power consumption as it eliminates the phase locked loop (PLL). Without a reference of comparable or superior quality to the RF oscillator, it does not make sense to use a PLL or FLL. And without a PLL, there is no need for a high-power frequency divider. As a side note, because the RF oscillator is the highest quality frequency generated on chip, there could be some benefit to using a divider to generate a higher quality low frequency oscillation. However, to reiterate, the primary reasons for eliminating the crystal reference from a modern standards compliant transceiver are cost, and size (ease of integration).

It is interesting to note that, in non-military radios, crystal oscillators were not used for timing or as a frequency reference for channel selection. Before crystals were used, radios relied on the electrical resonance of an inductor and a capacitor. This can vary considerably with temperature, humidity, and the devices themselves age and the resonance changes with time. It was not unheard of for multiple AM radio stations to accidentally overlap in adverse conditions, and it was necessary for people to measure and tune their transmitter and receiver resonant frequencies. The transceivers presented in this thesis share this same fault. One of the most challenging parts of the single chip mote project (results presented in Chapters 4 and 5) was to ensure that the transceiver could automatically calibrate and compensate its

on-chip oscillators without the use of an absolute frequency reference.

Power

The vast majority of power is burned in the RF frequency synthesizer and in the power amplifier. Significant power is required in the oscillator because of the limited quality factor of on-chip magnetics (in the case of LC tank oscillators) or because of phase noise requirements (in the case of ring/relaxation oscillators). More detail on frequency synthesizer power consumption and performance is in chapter 2. For the power amplifier, suppose that the desired range of the transmitter is 100 meters to a receiver with -85 dBm sensitivity (the minimum requirement for a 802.15.4 receiver). Assuming perfect antennas and a 100% efficient transmitter, this would require 316μW of power. This is nearly double the power than a state-of-the-art 802.15.4 analog baseband chain (measured on the single chip mote v3), and approximately equal to the power consumption of a Cortex-M0 running at 5 MHz in 65nm CMOS (measured on the single chip mote v2). It is comparable to the filter and ADC receiver power in [25]. To complicate things further, due to MOSFET threshold and accessible power supply constraints, as well as the limited Q factor of on-chip inductors, efficiency of transmitters below 0dBm is usually less than 30% [26].

Antenna

The very first radios had no notion of a crystal reference. More recent radios can operate with extremely small power budgets (admittedly not standards compliant, but data rates and sensitivities are comparable to those dictated by standards) [13]. However, all of these radios operate with off-chip antennas. Attempts to operate transceivers with on-chip antennas in the 2.4 GHz ISM band (ISM stands for industrial, scientific, and medical - ISM bands are unlicensed, but not unregulated, bands of spectrum) have resulted in low efficiencies (and therefore small ranges) or operate by either inductive or capacitive coupling, not radiation [27]. The fundamental reason for the lack of efficiency is the Chu-Harrington-Wheeler limit [28] and [29] stated here:

$$Q \geq \frac{1}{(ka)^3} + \frac{1}{(ka)} \tag{1.1}$$

In Eq. 1.1 k is the wave number, $2\pi/\lambda$. An important consequence of this relation is that for electrically small antennas, the Q increases by the ratio of the wavelength λ to the radius encompassing the antenna structure a cubed. [30] has performed a literature survey on published antennas and has validated the claim in [28] and has derived an expression for bandwidth efficiency product of small antennas (assuming a VSWR of 2):

$$B\eta = \frac{1}{\sqrt{2}} \left(\frac{1}{ka} + \frac{1}{n(ka)^3} \right)^{-1} \tag{1.2}$$

In Eq. 1.2, η is the antenna efficiency, and $n = 1$ for linearly polarized antennas, and $n = 2$ for circularly polarized antennas. The bandwidth, B, is set by the Q factor (see equation 1.1) which, for electrically small antennas, is high. So for smaller values of ka (also sometimes referred to as aspect ratio) the bandwidth-efficiency product is reduced, and the bandwidth is increased. Efficiency drops sharply at aspect ratios less than 1. Increasing the dielectric constant of the material does increase the aspect ratio for an electrically small structure, but this generally offers minimal benefit as any improvement gained from having a higher electrical impedance antenna is lost in the dielectric-to-air coupling.

For on-chip antennas, doped silicon (the material of choice for the vast majority of CMOS processes) has a number of drawbacks. First, it is mildly conductive, so any currents in a metal structure will induce lossy conduction in the silicon substrate (eddy currents). However, it is not conductive enough to be used as a ground plane in a monopole (furthermore it would be parallel to currents in the antenna structure, rather than perpendicular, and will therefore destructively interact with desired radiative characteristics). In addition, the conductive return path that the substrate provides means that any antenna structure will have a significant, undesireable capacitive component [31]. At higher frequencies, there is also significant loss caused by radiative coupling to the substrate (the substrate acts as a waveguide).

There is one significant benefit of using an on-chip antenna. Because there is no need for any transmission lines, there is also no need to match either circuits or antennas to the (somewhat arbitrary) 50 Ω impedance.

1.5   Thesis Organization

This thesis can be considered in two major portions. The first portion pertains to the theory and design methodology of various transceiver components, namely, power amplifiers, local oscillators, and supply conditioning circuits. The second portion is a presentation of results from a number of the chips that were taped out to demonstrate the feasibility and function of a completely single-chip wireless transceiver.

Chapter 2 is about the design of RF frequency synthesizers. This includes a discussion of oscillator topologies, design of passives, and fundamental design tradeoffs with regards to power, noise, and area. I primarily focus on LC tank resonant oscillators.

Chapter 3 is about low-power, sinusoidally-driven power amplifiers.

Chapter 4 is about the design and measurement of a 2.4 (ish) GHz RF transceiver - the Single Chip Mote v1. The primary focus of this chip was to demonstrate the feasiblity of using a reference-free, LC tank oscillator-based transceiver in a standards compliant network.

Chapter 5 is about the design and measurement of a 2.4 GHz RF transceiver - the Sin-

gle Chip Mote v3. This system-on-chip was built with an integrated 802.15.4 transceiver and BLE transmitter and a Cortex M0 and digital baseband processing so that the device operated with digital bits-in and digital bits-out. The purpose of this chip was to, once again, demonstrate the feasibility of reference-free standards compliant communication. Two benefits of this chip, in addition to having a high degree of on-chip integration between RF and digital systems, were that it actually operated in the 2.4 GHz ISM band (unlike the single chip mote v1), and that very little of the testing actually required laboratory equipment - a compelling argument for the ease of wireless integration that I have emphasized in this introduction.

Chapter 6 focuses on the design and measurement of a fully integrated 24 GHz transceiver. This includes system-level implications of carrier frequency, design of the antenna-transceiver interface, techniques for taking advantage of the high antenna impedance, and some measured results.

Finally, Chapter 7 concludes the dissertation. I summarize the presented results and offer my take on potential future directions and applications for my work.

Chapter 2

On-chip Frequency Synthesis

The purpose of this chapter is to investigate the design tradeoffs associated with on-chip frequency synthesizers specifically in low-power applications. In Armstrong heterodyne and superheterodyne receivers, and in transmitters, it is necessary to locally synthesize the carrier frequency. This allows narrow-band transmission and reception of radio frequency signals. In this chapter I will investigate the power-performance tradeoff associated with oscillators and discuss the minimum possible power consumption.

2.1   Receiver and Transmitter Architectures

Why is an oscillator necessary? The vast majority of modern wireless communication occurs with carriers at radio frequencies (between 100 MHz and 10 GHz - although efficient communication above these frequencies has been demonstrated in the past few decades) and is not broadband. In other words, sampling the entire spectrum at the Nyquist rate will capture enormous quantities of noise. The noise depends on filtering - it is possible to filter at RF, either purely with electrical components or with off-chip SAW, BAW, or FBAR filters. In 1917 Armstrong developed the heterodyne receiver (and a bit later, the superheterodyne receiver) that works by using a frequency translator (mixer) to shift the carrier on which the signal is modulated to a low frequency before sampling it. Block diagrams of both receiver and transmitter are shown in Fig. 2.1.

A local oscillator at or near the carrier frequency is needed to perform the downconversion operation. Similarly, to transmit a narrow-band signal, it is either necessary to upconvert a baseband signal to the carrier, or to directly modulate the carrier itself. In either of these cases, the transmitter needs to locally synthesize a frequency at or near the carrier frequency.

Resonant Oscillator Basics

A basic schematic of a resonant LC tank oscillator is shown in Fig. 2.2.

(a)   (b)
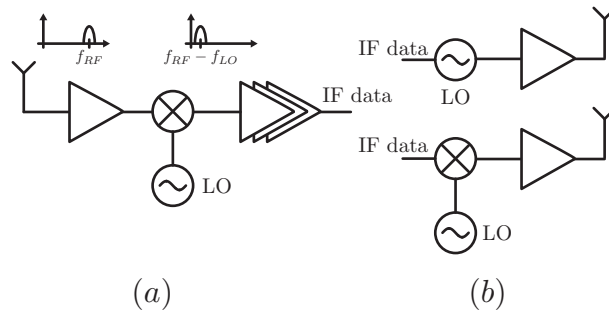
Figure 2.1: Block diagram of Armstrong's super-heterodyne receiver (a) a direct modulation and upconversion transmitter (b)
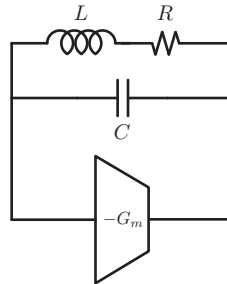
Figure 2.2: Simple circuit of a LC tank oscillator. The R represents the total series loss of the magnetic component. Capacitive losses are not shown because capacitor loss is significantly lower than inductor loss with on-chip components at frequencies below $\tilde{1}0$ GHz (this number is anecdotal and process dependent)

The fundamental requirement for a stable oscillation can be extracted from the Barkhausen stability criterion, which states that for a stable oscillation, the loop gain of a circuit must be equal to 1 for a phase shift of $0°$. In practice, this means that the magnitude of the transconductance G_m must be greater than the real parallel conductance of the tank circuit at the frequency of oscillation. Then, even if the small-signal transconductance is greater than the parallel conductance of the tank (which would nominally result in an unbounded oscillation), circuit nonlinearities will reduce the large-signal transconductance, and the loop gain will settle to 1. The frequency of the oscillation is dictated by the inductance and capacitance of the tank (Eq. 2.1).

$$f_{osc} = \frac{1}{2\pi\sqrt{L_p C}} \tag{2.1}$$

And the parallel resistance of the tank is given as a function of the inductor's quality factor

(Q) is given in Eq. 2.2. A higher inductor Q will result in a lower required transconductance. Note that I ignore the equivalent series resistance of the capacitor - this is generally fine for frequencies below 10 GHz where Q factors of on-chip capacitors are still substantially higher than Q factors of on-chip inductors.

$$R_p = R_s + \omega L Q \approx 2\pi f L Q \tag{2.2}$$

Phase Noise

In CMOS circuits, generating a large transconductance at low power is not really a problem. However, this comes at the penalty of noise. This manifests itself as a smearing of the oscillator's voltage spectrum, as shown in Fig. 2.3 (a). The consequences of phase noise are EVM errors in transmitter constellations, reduced interferer tolerance from reciprocal mixing, and degradation of high-SNR receiver performance. Generally speaking - the goal is to keep phase noise as low as possible for a given power consumption. Often times, phase noise is described in terms of single-sideband phase noise, which is shown in Fig. 2.3 (b).
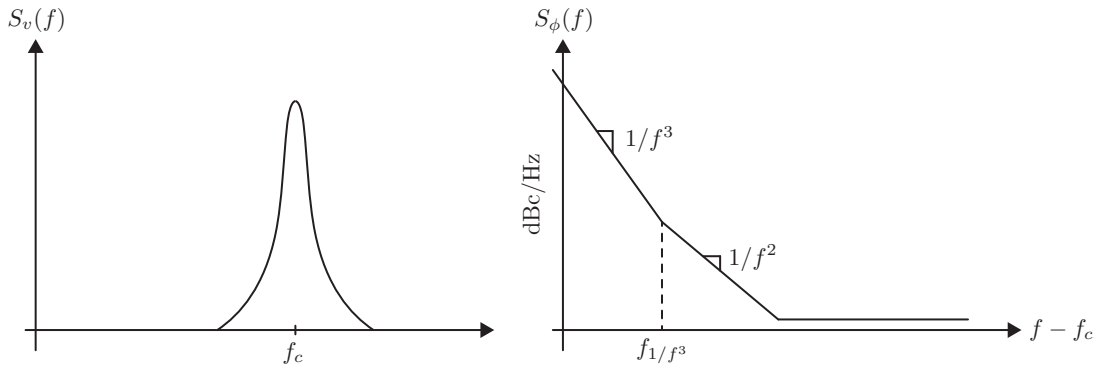
Figure 2.3: Phase noise in oscillators

An expression for an oscillator's phase noise is given in 2.3 [32].

$$L(f_m) = 10\log_{10}\left[\frac{2FkT}{P_{sig}} \left(1 + \left(\frac{\omega_0}{2Q\Delta\omega} \right)^2 \right) \left(1 + \frac{\Delta\omega_{1/f^3}}{\Delta\omega} \right) \right] \tag{2.3}$$

This formula is filled with largely empirical constants, and is not particularly helpful for design (besides saying that increasing Q is beneficial). In addition, the F, often referred to as the oscillator's noise factor, is generally computed after the oscillator has been built and measured. Later in this chapter there will be some discussion about computing F in an attempt to make a comparison between multiple oscillator topologies. In addition, Eq. 2.3 is used to generate the oscillator figure of merit (FoM) in Eq. 2.4. Much can be said about figures of merit, but for oscillators, it does tend to be a fairly good indication of how

well the oscillator was designed. This expression also completely ignores tuning - a critical component of any modern oscillator.

$$FOM = L(f_m) - 20\text{log}_{10}\left(\frac{f_0}{f_m}\right) + 10\text{log}_{10}(P_{dc}) \qquad (2.4)$$

2.2 Passive Design

The goal of this thesis is fully-on-chip wireless system integration, so all of the passives associated with the LC tank oscillator must be integrated on the CMOS die. Generally speaking, off-chip components do have higher Q factors at radio frequencies. But this comes at the penalty of potentially unpredictable chip-to-PCB parasitics. In addition, most of my commentary will seem anecdotal. This is intentional. These days, EM simulators are extremely advanced and, as long as the structure is simulated properly (accounting for ground return paths, guard rings, and correct port placement) the models are broadband and, from all evidence, seem correct. It is quite challenging to measure an inductance and Q directly, but power consumptions and attached circuits behave as expected.

Inductors

The inductor is the most important component to model correctly in an LC tank oscillator. The LQ product from Eq. 2.2 dominates the parallel tank impedance at resonance (for frequencies less than approximately 10 GHz). This impedance limits phase noise performance both because of noise generated from the tank and transconductance needed at a particular current consumption to meet the oscillation criterion. In addition, it dictates how much voltage amplitude swing is generated by the oscillator at a given current (this will be discussed in more detail when examining different topologies). Transformers will not be discussed in this section. The parallel capacitance in Fig. 2.4(c) is caused by both metal overlap from multiple inductor turns and from proximity of these turns. If there is only one turn, there will still be a capacitance from the differential feed. Inductance will not scale quadratically with number of turns because inner turns have less effective area. In addition, these additional turns will reduce the self-resonant frequency. As a rule-of-thumb: for a given area, a square inductor gives the highest inductance. And for a given inductance, an octagonal/circular inductor will have the highest Q, but will have a larger area. I can only say this with confidence for frequencies below 10 GHz. At any higher frequency, concerns about SRF will dominate. Empirical formulas for inductance of square and octagonal inductors are given in Eq. 2.5 from [34]. These are DC inductances. For a square inductor, $c_1 = 1.27$, $c_2 = 2.07$, $c_3 = 0.18$, and $c_4 = 0.13$. For an octogonal inductor, $c_1 = 1.07$, $c_2 = 2.29$, $c_3 = 0$, and $c_4 = 0.2$. Note from Fig. 2.4(c) as the structure approaches its SRF, the inductance will increase linearly, while the resistance increases by $1/\sqrt{f}$, so Q will increase as well.
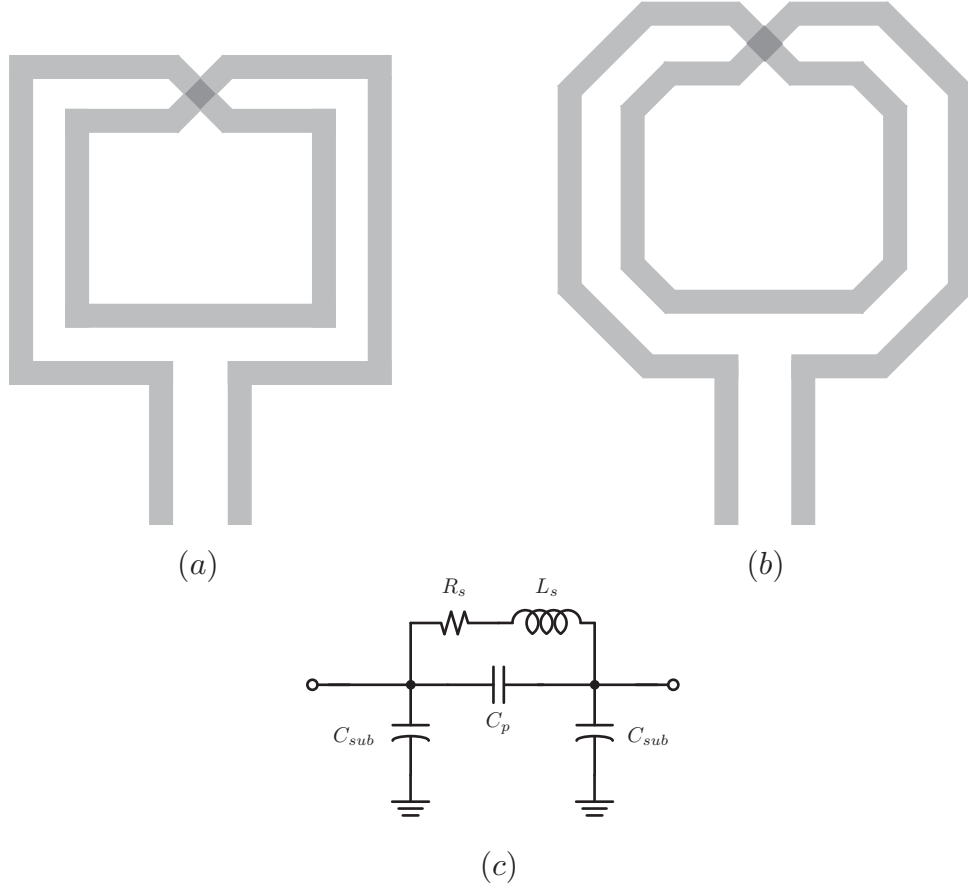
Figure 2.4: Examples of on-chip inductors and a simple equivalent circuit model [33]

$$L_{sq} = \frac{\mu n^2 d_{avg} c_1}{2} \left[\ln \left(\frac{c_2}{\kappa} \right) + c_3 \kappa + c_4 \kappa^2 \right] \tag{2.5}$$

$$\kappa = \frac{d_{out} - d_{in}}{d_{out} + d_{in}} \tag{2.6}$$

In the equation above, d_{out} is the diameter of the outermost turn, and d_{in} is the diameter of the innermost turn. An approximate expression for the series resistance R_s of the inductor is given in Eq. 2.7 where δ is the skin depth of the conductor, ρ is the resistivity of the metal, and w and d are the cross sectional width and height of the square conductor.

$$R_s = \frac{\rho l}{(wd - (w - \delta)(d - \delta))} \tag{2.7}$$

It is beneficial to either use thicker metal or strap metal layers together to trade SRF for Q. The SRF of the inductor should be around two to three times the operating frequency of the oscillator. This results in an optimal LQ product while leaving wiggle room for tuning capacitors and for transistor parasitics. It is also often beneficial to use patterned ground shields underneath the center of the inductor. This shield has strategically placed gaps to prevent Eddy currents induced in the mildly conductive substrate from de-Qing the inductor. A correctly designed ground shield can boost Q of an inductor by upwards of 10% while reducing the self-resonant frequency [35]. In addition, most process flavors offer a lightly doped region that should be used underneath the inductor.

Capacitive Tuning

What are the chances that the frequency of oscillation is exactly what the designer desired? Zero. This is a common argument against the FoM in 2.4 - it does not account for a necessary component of an oscillator in an actual system. Generally speaking, making modifications to an inductance is challenging (without de-Qing it substantially). Instead, it is much more common, and much simpler, to change the capacitance in Eq. 2.1.

A very loose rule of thumb for capacitor tuning is: parasitics will amount to approximately 20% - 25% of the tuning DAC's maximum capacitance. Parasitics come from: routing capacitance and inductance, and transistor parasitics. For a large capacitor, a correspondingly large transistor is needed to prevent de-Qing of the capacitor when it is "on". And, a larger DAC (in terms of both capacitance and number of bits) will have more routing parasitics. An example layout for a 5-bit DAC is shown in Fig. 2.5 (c). In the annotated layout I did not include the pull up/down transistors/inverters in Figs. 2.5 (a) and (b). Assume that the voltage swing at v_o+ and v_o- is large. When v_{ctrl} is high (the tuning capacitor is on) the AC coupled swing will be biased to a floating node and could turn off the pass transistor. If v_{ctrl} is low, that same AC coupled swing could turn the transistor on. To mitigate this, the nodes v_f+ and v_f- can be weakly pulled high (hence the inverter in Fig. 2.5 (b)). Note that routing both power and ground to every element of the capacitor DAC can be unpleasant. In cases where the amplitude of the swing is at or not far above the threshold of the transistor, it is often fine to use the pull down transistor only. This is described in detail in [36].

2.3   LC Tank Oscillator Topologies

For the most part, oscillator design is a solved problem. An oscillator designer needs to identify their power and phase noise requirements and understand how much area they have available. The most challenging part of oscillator design is practical tuning range and resolution. In this section I will investigate a number of different on-chip oscillators. Three major oscillator qualities should be considered: tuning (both range and resolution), voltage swing per DC current, and phase noise. Phase noise is considerably less important. It is possible to glean some degree of intuition from Leeson's formula in Eq. 2.3. For a given
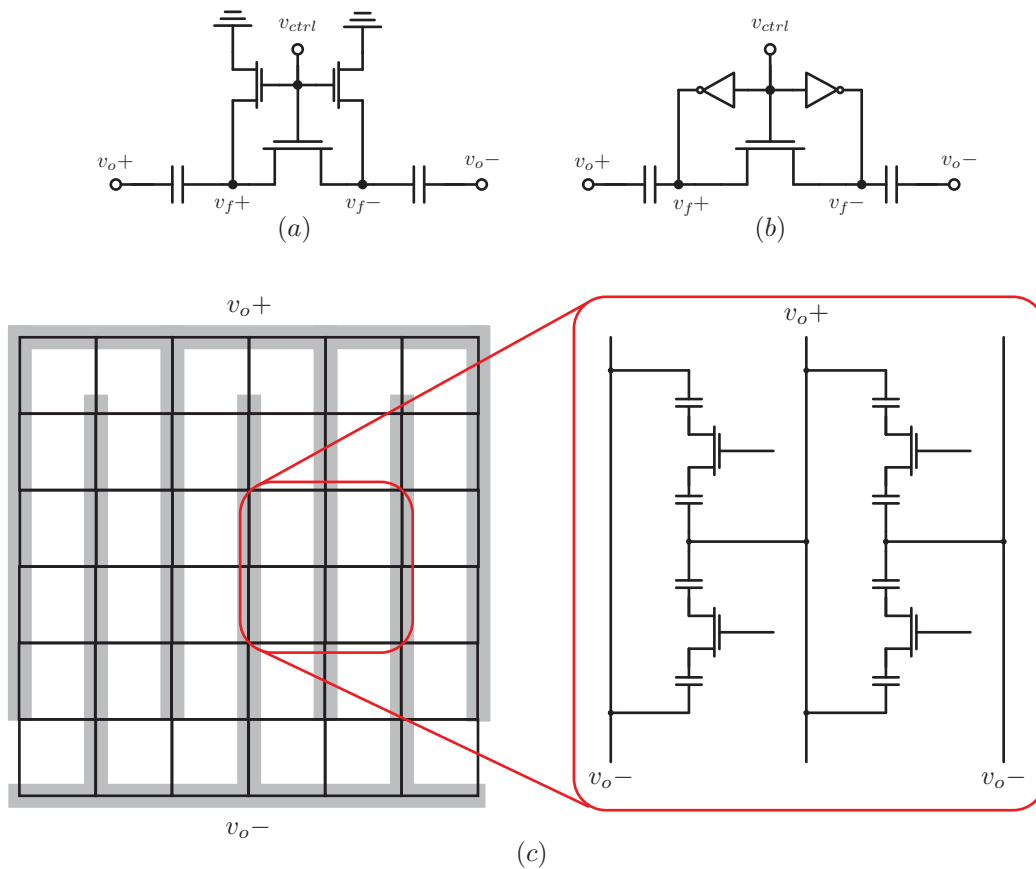
Figure 2.5: Capacitor tuning in an LC tank oscillator. (a) Shows an appropriate way to bias the floating node to prevent v_f from swinging low enough to turn off the pass device when its gate is high. (b) Shows a similar technique that also prevents the pass device from being turned on while its gate is low. (c) Shows a common method to lay out a capacitive DAC

current, to maximize swing, the LQ product of the tank should be maximized. However, to minimize phase noise, the inductance should be reduced as much as possible and the current increased as much as possible (for a fixed supply voltage) which is contrary to maximizing swing. The reason for this is that at some point, the supply voltage will limit the amplitude of oscillation because of nonlinearity in the active circuits.

In low data-rate, low-performance applications, phase noise does not matter. It is nice to keep it low (as it simplifies transmitter and receiver design) but requirements are fairly trivial. As mentioned in the introduction, the phase noise requirement for 802.15.4 is -102 dBc/Hz at an offset of 3.5 MHz. For an on-chip LC tank oscillator, this requirement is trivial even at extremely low power levels (see results chapters 4 and 5). And, as mentioned earlier, operating a transconductor at high g_m/I_D tends to come at the price of noise (which has been deemed irrelevant). The limiting factors, then, are swing and tuning. The amplitude

of the oscillator waveform is important as it dictates the performance of circuits directly connected to the oscillator. A small swing will weakly drive transistors in triode, and, if LO buffers are used, will require larger current consumption. Therefore, in these designs it is beneficial to use a large inductor. However, a large inductance comes at the price of tuning range and resolution.

Topology Overview

For most (if not all) differential CMOS oscillators, the negative g_m required to overcome the loss in the tank is generated with a cross-coupled differential pair. Note the repeated usage of the cross-coupled pair in the four topologies in Fig. 2.6.
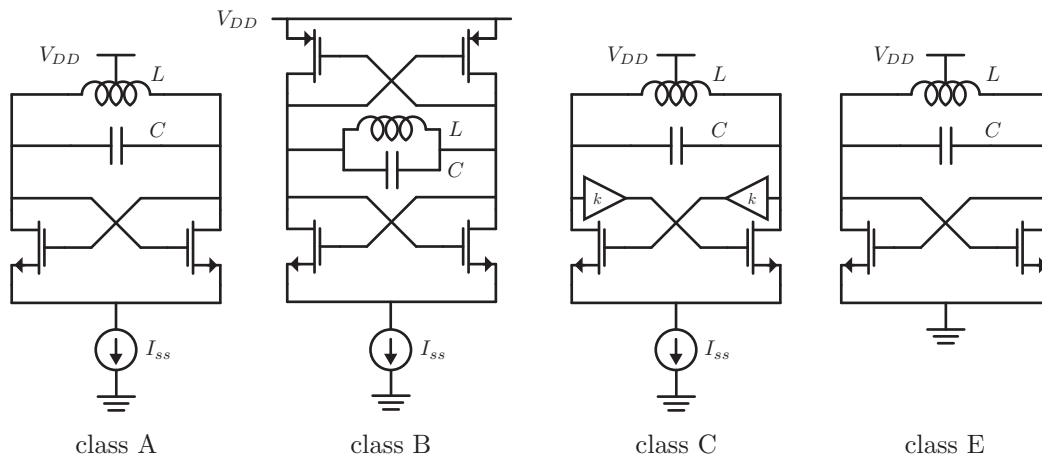
Figure 2.6: Four oscillator topologies

The class A oscillator is the simplest, and most closely mimics the structure in Fig. 2.2. The class B oscillator is identical, but has g_m contributions from both the NMOS and PMOS pairs, which means that the current required is halved. In addition, in the class A oscillator, the current only flows through half of the tank (because of the center tap), whereas in the class B oscillator the current flows through the entire tank impedance. This improvement is related to the "current efficiency" of the oscillator, i.e. how much of the bias current turns into amplitude in the first harmonic of the oscillation. This efficiency is correlated directly with swing, which was identified as a critical parameter for oscillator design. There are plenty of slight variations on the topologies shown. In most practical implementations, the voltage gain shown in the class-C oscillator is implemented with a transformer, and the inductances of the transformer turns are incorporated in the tank. In addition, it is certainly possible to combine class-C and class-D by adding voltage gain between the drain and source of the cross-coupled pair. Similarly, it is certainly possible to add PMOS devices (similar to the evolution from class-A to class-B) in either the class-C or class-D topologies. In addition,

many papers propose filtering techniques [37] to reduce tail phase noise and 1/f upconversion [38].

In the class A oscillator, at every cycle, I_{ss} flows through half of the tank. The voltage waveform is the first harmonic of the product of the current and the parallel resistance:

$$V_{a,A} = \frac{2}{\pi} I_{ss} R_p \tag{2.8}$$

Where R_p is the parallel tank impedance given in Eq. 2.2. For proper operation of this oscillator, the supply voltage must be at least one overdrive voltage (current source) plus one threshold.

$$V_{DD,A} = V_t + V_{od} \tag{2.9}$$

In the class B oscillator, the tank current will, once again, be a square wave. However, in one half of the cycle, it will have a positive $+I_{ss}$ flowing through it, and in the other half cycle, it will have $-I_{ss}$. So, the voltage waveform will have amplitude:

$$V_{a,B} = \frac{4}{\pi} I_{ss} R_p \tag{2.10}$$

For the same current, the amplitude of oscillation is doubled. The minimum supply voltage is:

$$V_{DD,A} = 2V_t + V_{od} \tag{2.11}$$

In the class C oscillator, the tank current waveform is a bit more complicated. Because of the increased loop gain (caused by the voltage amplification from drain to gate of the cross-coupled pair) the current through the tank has a very narrow conduction angle (hence the name class C).

$$V_{a,C} = I_{ss} R_p \tag{2.12}$$

There are a few problems with the class C oscillator that make it unattractive. First of all, for this equation to hold, the cross-coupled devices need to stay in the active region while they are conducting. If they enter triode, the amplitude drops (because the loop gain drops) and phase noise performance drops correspondingly [39]. The minimum supply voltage for an NMOS only class C oscillator is the same as for the class A. And the minimum supply voltage for a CMOS class C oscillator is the same as for the class B oscillator. Note that for the complimentary case, the amplitude of oscillation will be doubled for the same bias current.

Finally, the class D oscillator operates with intentional hard switching of the cross-coupled devices. In addition, it operates in the voltage mode, as opposed to the other oscillator topologies that operate in the current mode, and it's amplitude is approximately [40]:

$$V_{a,D} = 1.64 V_{DD} \tag{2.13}$$

The minimum supply voltage of the class D amplifier is V_t. The current consumption (assuming that the capacitor is lossless) is given by:

$$I_{dc,D} \approx 5.2 \frac{R_s V_{DD}}{\omega_0^2 L^2} \tag{2.14}$$

To summarize, if a given swing is required (and that swing is in the current-limited regime for the class A, B, and C oscillators), the power consumption of each topology can be calculated in terms of the desired swing, the transistor threshold voltages, and the tank's parallel impedance (which, to reiterate, can be maximized by having the highest possible LQ product). Assuming that the overdrive voltage of the current source is $V_t/2$, these minimum DC power consumptions are given in the following four equations:

$$P_{dc,A} = \frac{3}{4} \pi \frac{V_t V_{sw}}{R_p} = 2.356 \frac{V_t V_{sw}}{R_p} \tag{2.15}$$

$$P_{dc,B} = \frac{5}{16} \pi \frac{V_t V_{sw}}{R_p} = 0.982 \frac{V_t V_{sw}}{R_p} \tag{2.16}$$

$$P_{dc,C,n} = \frac{3}{2} \frac{V_t V_{sw}}{R_p} = 1.5 \frac{V_t V_{sw}}{R_p} \tag{2.17}$$

$$P_{dc,C,np} = \frac{5}{8} \frac{V_t V_{sw}}{R_p} = 0.625 \frac{V_t V_{sw}}{R_p} \tag{2.18}$$

$$P_{dc,D} = 1.93 \frac{V_t^2}{R_p} \tag{2.19}$$

In both chips presented in this thesis, the complimentary class B oscillator was used. This is in spite of the fact that the complimentary class C has superior performance. The reason that a class C was not used was because of complications involving the transformer design. Generating low-loss, high inductance, and high turns ratio transformer is considerably more challenging than generating a differential inductor with a high inductance and low loss, especially while keeping the structure's self-resonant frequency at least two times higher than an operating frequency of 2.4 GHz.

One intriguing possibility is to combine the class C and class D topologies and use a transformer in lieu of the inductor in Fig. 2.6. This boosts the voltage amplitude on the gates of the transistors, which means that the swing at the drains can be reduced. The supply voltage can be reduced as well, as long as the gates are DC biased to an appropriate voltage. This topology, with transformer, is shown in Fig. 2.7. A similar technique has been proposed in [41], although with a transformer tail choke (with additional coupling with a trifilar coil).
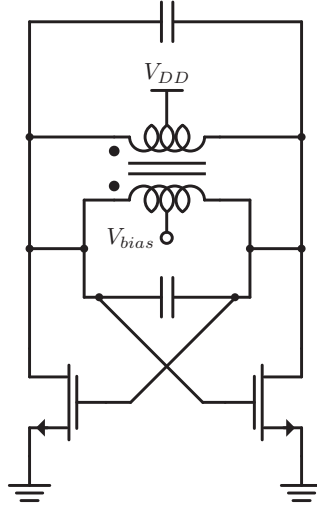
Figure 2.7: Class F style oscillator with transformer boosting of loop gain

Tuning Resolution

The three ways to perform fine tuning are: varactors and a voltage DAC, series capacitance, and source degeneration. Varactors will not be discussed because: they are not linear with tuning voltage (which complicates tuning and adds noise [42]) and voltage DAC noise couples directly to the tank, which is not desirable. All told, they are far more appropriate for voltage-controlled oscillators, not the digital code-controlled oscillators that are more prevalent in modern RF systems.

For low-power systems, having a large inductance with a large Q is desirable for practical system performance to reduce the oscillator's current consumption. The frequency of oscillation was given previously in Eq. 2.1. Of course, if the inductance is larger, the capacitor must be smaller for the same frequency. Furthermore, the change in capacitance required to create a particular Δf in the frequency is correspondingly smaller. To give a numerical example, for a 1 nH inductor, to change the oscillation frequency from 2.4 GHz to 2.4 GHz + 100 kHz is 366 aF. For the same change in frequency using a 10 nH inductance, the change in capacitance is 36.6 aF. Two different strategies for obtaining these tiny changes in capacitance are shown in Fig. 2.8.

The two strategies in Fig. 2.8 are series capacitive reduction (left) and capacitive degeneration (right). The change in effective capacitance from a change of ΔC on C_{fine} is:

$$C_{eff} = \frac{C_{big}\left(C_{fine} + \Delta C\right)}{C_{big} + \left(C_{fine} + \Delta C\right)} \tag{2.20}$$

The larger the capacitance C_{big} is, the smaller the effective change in capacitance will be. This is a viable solution, although having a larger capacitor C_{big} will result in bottom-plate
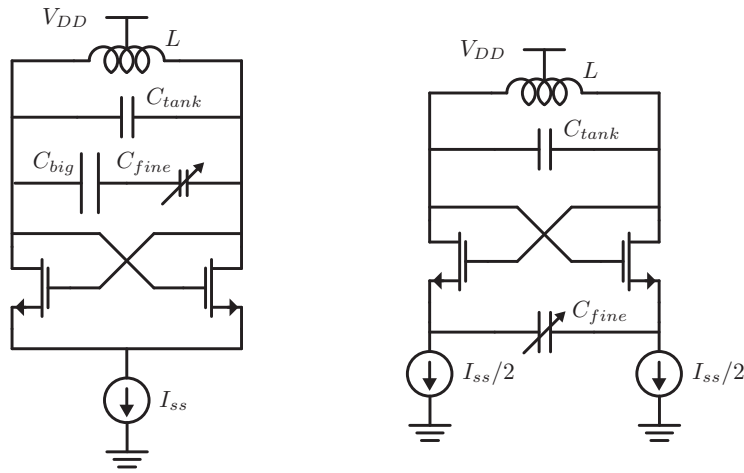
Figure 2.8: Two strategies for fine frequency tuning

parasitics. An alternative approach that avoids this issue is the capacitive degeneration technique in Fig. 2.8(b). The equivalent circuit is in Fig. 2.9.
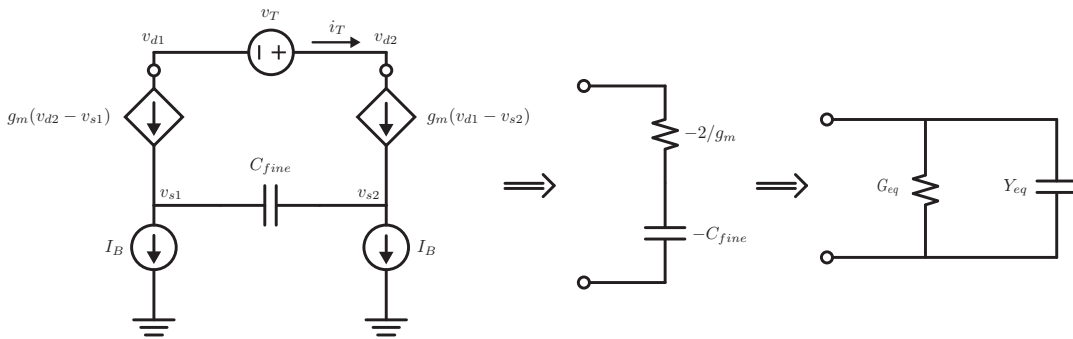
Figure 2.9: Equivalent circuit for the degenerated capacitor technique

The capacitance seen at the tank can be computed with a series-to-parallel transformation of the equivalent impedance. The equivalent capacitance is given in Eq. 2.21.

$$C_{eq} = -C_{fine}\frac{g_m^2}{g_m^2 + 4C_{fine}^2\omega^2} = -C_{fine}\left(\frac{1}{1 + \frac{1}{Q_{fine}^2}}\right) \tag{2.21}$$

In Eq. 2.21 the value Q_{fine} is the effective Q of the capacitor, and is given in Eq. 2.22. The equivalent parallel conductance is given in Eq. 2.3.

$$Q_{fine} = \frac{g_m}{2\omega C_{fine}} \tag{2.22}$$

$$G_{eq} = -\frac{g_m}{2} + \frac{g_m}{2}\left(\frac{Q_{fine}^2}{1+Q_{fine}^2}\right) \tag{2.23}$$

One important thing to note about the Q_{fine} of the capacitor: it is supposed to be small. In the limit that it approaches zero, the circuit approaches a standard LC tank topology, as expected (for the Q to reach zero, the capacitance should be infinitely large and therefore a short circuit at the frequency of oscillation ω). In addition, when the capacitor is large compared to g_m/ω, the effective capacitance seen at the tank will be:

$$C_{eq} = -CQ_{fine}^2 \tag{2.24}$$

In the equation , notice that there are two terms: the negative conductance $-g_m/2$ that is preserved from the standard cross-coupled topolology, and a positive conductance that arises from the finite Q of the series capacitor. This positive conductance effective de-Qs the LC tank, albeit only if Q_{fine} is large.

Phase Noise

This section is nearly identical to the analysis presented in [43]. It is repeated here with the presence of the degenerated capacitor, and is based on the impulse sensitivity function analysis [44]. Any individual noise source can be turned into a noise source in parallel with the tank. This noise at an offset frequency $\Delta\omega$ is given by Eq. 2.25.

$$L(\Delta\omega) = 10\log_{10}\left(\frac{\Gamma_{R_t,rms}^2}{q_{max}^2}\frac{\overline{i_{R_t}^2}/\Delta f}{2\Delta\omega^2}\right) \tag{2.25}$$

In Eq. 2.25, $\Gamma_{R_{rms}}^2$ is the impulse sensitivity function (ISF) of the noise source. q_{max} is the amount of charge that appears on the tank's capacitor. The ISF of a shunt resistor is derived in [44], and is restated here:

$$\Gamma_{R_t}(\phi) = \frac{\cos(\phi)}{2} \tag{2.26}$$

This assumes a sinusoidal oscillation at the tank that is in quadrature with the shunt resistor's ISF (i.e., $V(\phi) = A\sin(\phi)$). The rms ISF in Eq. 2.25 is calculated by integrating the ISF over the period of oscillation:

$$\Gamma_{R_t,rms}^2 = \int_{-\pi}^{\pi} \Gamma_{R_t}^2(\phi)d\phi \tag{2.27}$$

In the case of a cyclostationary noise source, an effective ISF can be calculated by multiplying a noise source's ISF by the oscillation-phase-dependent noise power. This will be relevant when analyzing the noise contribution of the commutating negative-g_m devices used

to sustain an oscillation. One powerful and intuitive use of the ISF-style analysis is that to calculate the ISF of a particular noise source, all that needs to be done is relate the noise current contribution of the noise source to the noise current contributed by a shunt resistor, as in Eq. 2.26. First, we can look into the phase noise added by the level restoring transistors. Because of the symmetry of the circuit, it is only necessary to analyze the noise addition from one of the two transistors (the other is identical in magnitude). There are also three different regimes. In the first, shown on the left in Fig. 2.10, M1, the noise contributor in question, is on, while M2 is in cutoff. In the second regime, shown on the right in Fig. 2.10, both transistors are on. The third regime, during which M1 is in cutoff, is trivial as the device contributes no noise.
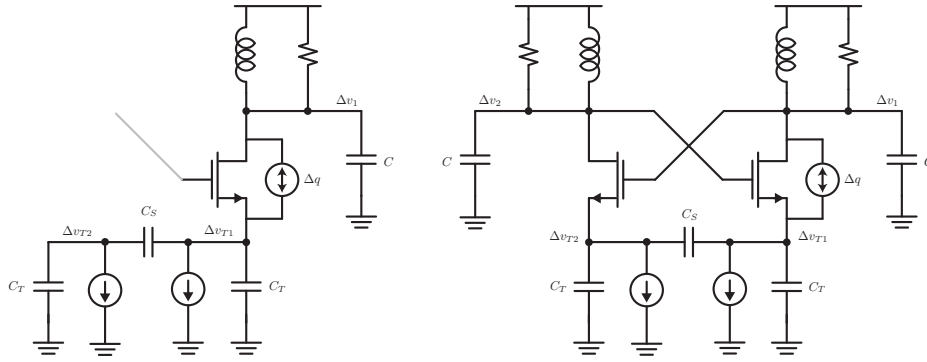
Figure 2.10: Effective circuits in various oscillator operating regions for calculating phase noise from one negative g_m device

When the transistor M2 is in cutoff, the noise generated by M1 sees a high impedance at its source. As long as the total capacitance at the source (which will be C_T as long as C_{fine} ¿¿ C_T) is smaller than $1/g_{m1}$, no noise from M1 will appear at the tank. The expression is nearly identical to the noise from a cascode device in a standard cascoded common source amplifier. When both transistors are conducting, the transfer function of M1's current noise seen at the output is given by Eq. 2.28

$$\Delta v_1 - \Delta v_2 = \frac{2g_{m2}}{\left(1 + \frac{C_T}{2C_{fine}}\right) g_{m1} + g_{m2}} \frac{\Delta q}{C}$$

(2.28)

Noise from the tail current source can be analyzed in a similar manner. The three different operating regimes are shown in Fig. 2.11.

When only M2 is conducting, the tail current's noise will all be seen at the tank, as shown in Eq. 2.29.

$$\Delta v_1 - \Delta v_2 = -\frac{\Delta q}{C}$$

(2.29)
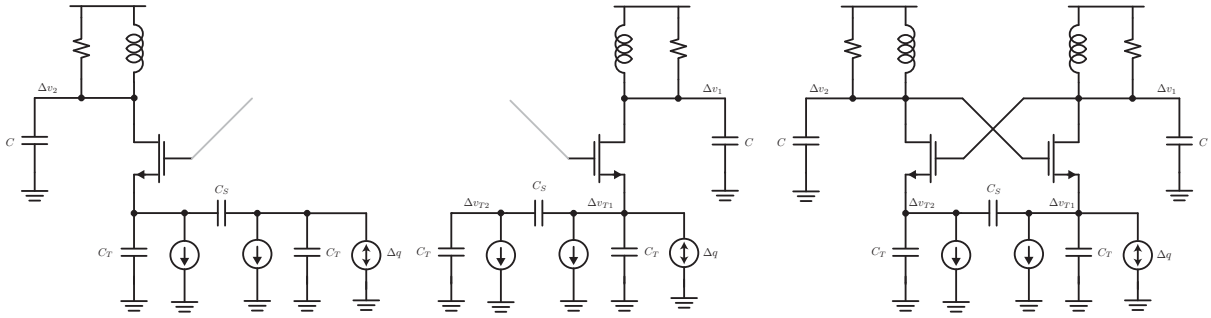
Figure 2.11: Effective circuits in various oscillator operating regions for calculating phase noise from the tail current source

And when only M1 is conducting, the tail current noise, once again, will all go to the tank.

$$\Delta v_1 - \Delta v_2 = \frac{\Delta q}{C} \tag{2.30}$$

When both are conducting, the noise current will split between M1 and M2. The final transfer function is given in Eq. 2.31.

$$\Delta v_1 - \Delta v_2 = \frac{g_{m2} - g_{m1}\left(1 + \frac{C_T}{2C_s}\right)}{g_{m2} + g_{m1}\left(1 + \frac{C_T}{2C_s}\right)} \frac{\Delta q}{C} \tag{2.31}$$

Unsurprisingly, in all of these noise expressions, as long as C_s is much greater than C_T, the phase noise from the switching devices and the tail current are identical to the expressions from [43]. If C_s is of similar magnitude compared to C_T, the noise from the switching devices is slightly reduced. However, this implies either a large C_T (which causes an increase in noise from M1 and M2 because the cascode assumption no longer holds) or a small C_s which degrades the tuning resolution and increased phase noise from the effective de-Qing of the tank from the conductance in Eq. 2.3.

To test the intuition behind these expressions, a 2.4 GHz LC tank oscillator was built using a 300 μA total bias current, an inductor with a differential inductance of 8 nH and a Q of 20 (high to reduce the influence of tank noise and emphasize transistor noise) a nominal g_m of 1.8 mS for the cross-coupled devices and 400 μS for each of the two tail devices. The differential degenerated capacitance and tail capacitance were swept, and the results are shown in Fig. 2.12. In the source capacitance sweep, the curve was shifted so that the phase noise is the same in the limit as C_s approaches a short circuit.

Even with the reduced g_m of the tail devices, increasing the tail capacitance actually reduces phase noise, in spite of claims in literature to the contrary. And, even with the high tank Q, the majority of the oscillator noise came from the resistance in the inductor. Any
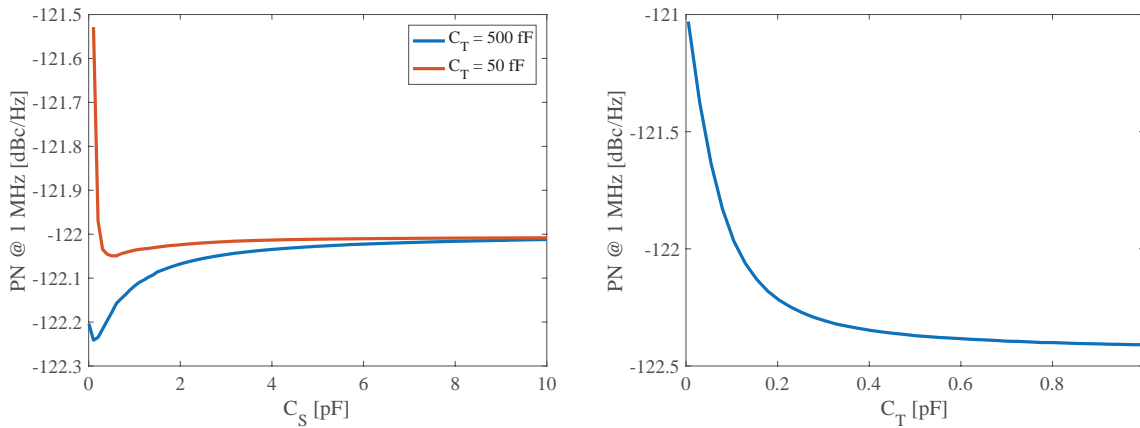
Figure 2.12: Phase noise at 1 MHz from carrier with varying tail capacitance and degenerated source capacitance

reductions in phase noise from filtering tail or transistor noise had a minimal influence on actual performance. The most effective way to reduce phase noise is to increase inductor Q.

Supply Noise

The impact of supply noise on phase noise is often ignored. Typically, this is appropriate, because if sufficient power can be burned in the supply regulation network, it can mitigate significant noise generators. However, when power minimization is a concern, any overhead from supply conditioning reduces battery life. In this section, two different supply parameters are varied: supply impedance and bandwidth. The supply noise is a fixed 500 kΩ resistor at room temperature. The supply resistance R_s in Fig. 2.13 is independent of the noise resistance. All simulations are performed using 65nm devices for the cross-coupled devices. Because of the tail current, the LC tank should present a current source load to the supply. Some simulations are performed with both a transistor bias and with a current source bias (with appropriate noise).

Results from three different oscillator topologies are shown: an NMOS-only class-A oscillator, a CMOS class-B oscillator without tail filtering, and an NMOS-only class-E oscillator. All use an 8 nH inductor with a Q of 20. The supply voltage of the class-A oscillator is 600 mV, and the supply of the class-B oscillator is 800 mV. To compensate for supply reduction from an increased source resistance, these supplies were slightly increased to keep the oscillator center frequency approximately constant. The purpose of these simulations is to confirm intuition and to understand what to keep in mind when designing low-dropout regulators (LDOs) and supply conditioning circuitry for on-chip LC oscillators.

In the class-A oscillator, when both devices are conducting, approximately no noise will appear at the tank. When one device is conducting and the other is off, as long as the
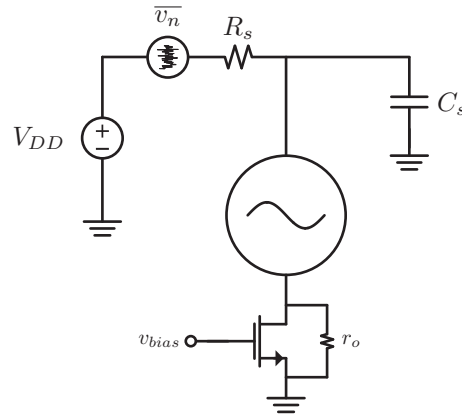
Figure 2.13: High-level schematic of supply noise simulation

conducting device is in saturation, both drains will present a high impedance, and very little differential noise will appear at the tank. The majority of supply-to-PN conversion in the class A oscillator occurs because of nonlinearity of capacitive parasitics and finite tail current resistance (thus causing a change in bias which, in turn, causes small fluctuations in frequency). These assumption do not apply if the commutating devices enter triode. So, to exacerbate the effects of supply noise, the devices were driven with excessive tail current.
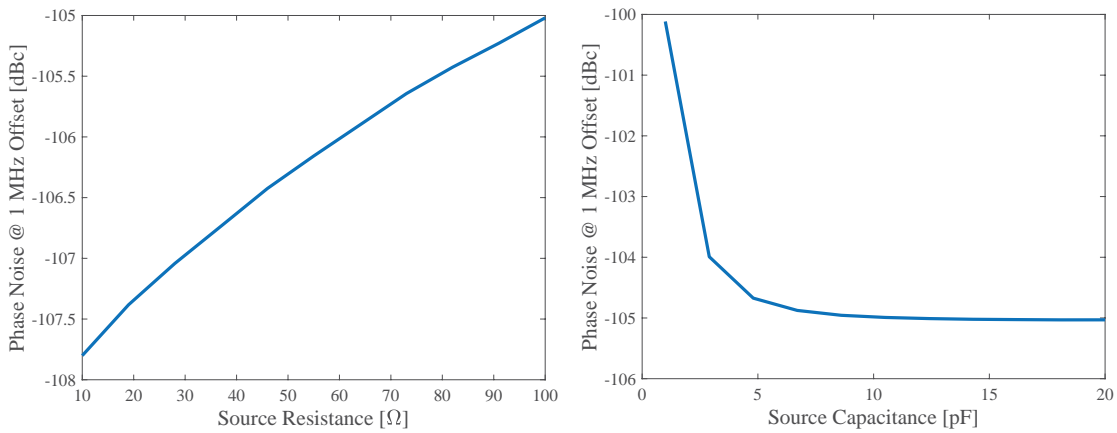
Figure 2.14: Influence of a fixed 500 kΩ noise resistance with varying source resistance (left) and source capacitance (right) in a class-A NMOS LC tank oscillator

As shown in Fig. 2.14, increasing the source resistance makes the finite resistance of the tail current look comparably worse, thus increasing the effect of the supply noise-current-frequency conversion. As expected, increasing capacitance filters the supply noise.

Because of the commutating PMOS devices, the class-B oscillator fundamentally has a DC-to-RF upconversion mechanism, much like current source noise in the class-A oscillator. In addition, if the PMOS and NMOS devices are not perfectly balanced, the oscillation waveform will have even harmonics, which will result in further added noise from the Groszkowski effect. For these reasons, unlike the class A oscillator, even if the bias current source is perfect, there will be some supply noise to phase noise conversion.
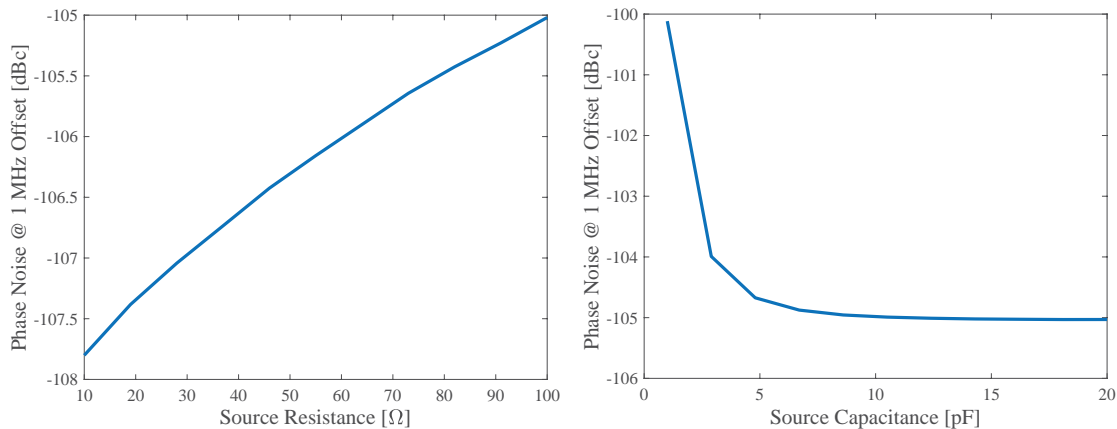
Figure 2.15: Influence of a fixed 500 kΩ noise resistance with varying source resistance (left) and source capacitance (right) in a class-B CMOS LC tank oscillator

In Fig. 2.15 an increased source resistance actually decreases the increase in phase noise from supply noise. Unlike in the class-A case, where the phase noise arose from a bias-dependent frequency, the noise effectively sees a resistive divider. The larger the source resistance, the less noise appears at the source of the PMOS devices. Once again, a larger capacitance is beneficial from a filtering perspective. To make the influence of supply noise more dramatic, the switching devices were driven into triode. When the devices remain in saturation, the influence of supply noise is minimal.

The class-E oscillator is interesting because, as noted earlier, a considerably lower supply voltage can be used, thus reducing power consumption while maintaining a high swing. However, the lack of current bias, and the fact that the oscillator devices inherently operate in triode for a large portion of the oscillation period (even when the oscillator is operating "correctly"). As shown in Fig. 2.16, even small quantities of broadband supply noise (class A and class B simulations used a 500 kΩ supply noise resistance) result in a significant degradation of phase noise performance.

The class-D oscillator (same as class-E but CMOS) has similar trends to the class-B oscillator. Increasing capacitance decreases the impact of supply noise, and increasing noise source resistance decreases the impact of supply noise. This implies that the commutating PMOS devices (this time in deep triode) present a low impedance, at least as far as noise is concerned. For brevity, simulation results were not included.
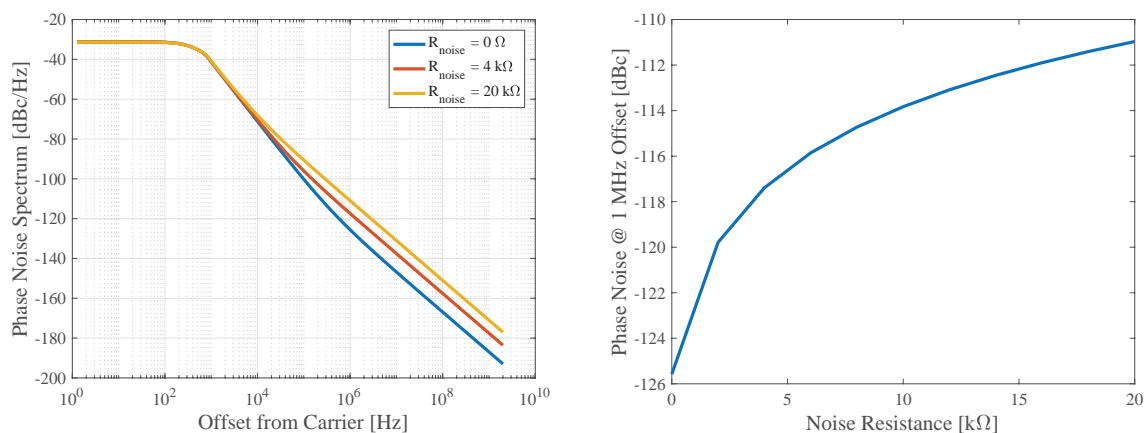
Figure 2.16: Phase noise degradation of a class-E oscillator with varying quantities of supply noise

Implications for LDO design

An oscillator's (or really, any analog load's) supply must be isolated from higher supply noise (line regulation). This is typically done with a low dropout regulator (LDO). LDOs have the added benefit of maintaining a constant output voltage even when there are variations in the load current (load regulation). A typical LDO schematic is shown in Fig. 2.17.
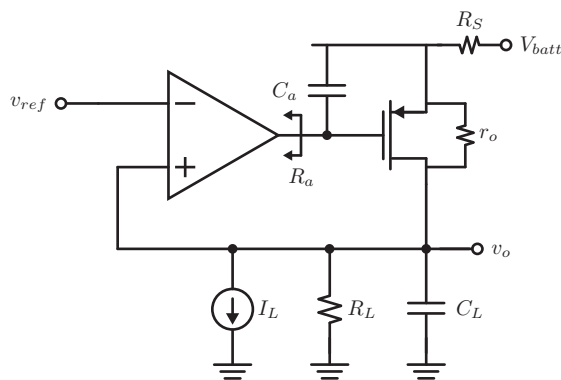
Figure 2.17: A typical LDO schematic used to derive line regulation, load regulation, and stablity criteria

The loop gain is given in Eq. 2.32. In this expression, ω_a is the frequency of the pole at the output of the amplifier, ω_o is the frequency of the pole at the regulated output voltage,

A is the DC gain of the amplifier, g_m is the transconductance of the output device, and R_o is the DC open loop resistance seen at the regulated node.

$$T(s) = \frac{A}{1 + s/\omega_a} \frac{g_m R_o}{1 + s/\omega_o} \tag{2.32}$$

An unstable LDO is not a particularly useful circuit. Unfortunately, the amplifier is often implemented with an OTA, which has a high output impedance. Because the load current is often relatively high, the pass device should be large. So, the amplifier pole ω_A is often low. In addition, the decoupling load capacitance C_L is often large as well. With no modifications, these poles often end up rather close to one another, which leads to instability. Because the output node is a low impedance (as shown in Eq. 2.33), adding capacitance to the output node is less effective, per unit of added capacitance, than adding it to the high impedance node (although, according to [45], and as shown in 2.33 and 5.10, explicitly limiting the natural bandwidth of the amplifier is also detrimental). A source follower between the amplifier and the pass device is not particularly useful. It would require either a large amount of current to keep its noise and output resistance low, or would need to be large, which defeats its purpose.

$$Z_{out}(s) = \frac{R_o}{1 + s/\omega_o} \frac{1}{1 + T(s)} \tag{2.33}$$

The output resistance has a zero at the amplifier pole frequency. If it is dominant, the regulator's load regulation will be worse than its DC regulation for frequencies between the amplifier pole and the unity gain frequency.

$$v_o/v_{noise,supply} = \frac{1 + s/\omega_a}{A} \frac{T(s)}{1 + T(s)} \tag{2.34}$$

As a quick aside: the expression in Eq. 2.34 assumes that the amplifier bandwidth is set by a capacitor to ground. If a dominant capacitance is added between the supply node and the gate of the pass device, the supply rejection will become dependent on the high supply's source impedance and a capacitive divider between this added cap and the load capacitance.

From Eq. 2.34, The LDO's ability to reject supply noise has a zero at the amplifier pole. If the amplifier pole is dominant, for frequencies between the amplifier pole and the unity gain frequency of the loop, supply rejection will be worse than if the output pole is dominant. Intuitively, if the amplifier pole is dominant, it degrades the loop's ability to act on discrepancies between the load voltage and the reference voltage. When the output pole is dominant, it attenuates noise at the output. Having an amplifier-pole dominant loop, from the perspective of attenuating supply noise, is undesirable.

The two reasons that someone might intentially make an amplifier pole dominant LDO are: it's easier to stabilize, because the amplifier pole is higher impedance, and because it can protect the battery voltage from load noise. As mentioned previously, if the amplifier pole is dominant, it will be unable to track changes in the load voltage. In the case of a noisy load, this is actually beneficial, because a changing load current would result in a changing source

current (through the pass device) that manifests itself as a changing high supply voltage if the battery voltage has a high source resistance R_S. So for digital switching loads (like the divider in chapter 5), in which line and load regulation are not particularly important (aside from the lurking risk of browning out the supply with a particularly large current draw and insufficient decoupling capacitance).

So far, only the regulator's line regulation, load regulation, and stability have been discussed. However, the LDO amplifier power consumption is a concern (it should be as low as possible), as well as the noise contribution to supply. The transfer function of an amplifier device's current noise to output voltage is given in Eq. 2.35.

$$v_o = \frac{1}{g_{m,A}} \left[\frac{T(s)}{1 - T(s)} \right] \overline{i_N} \tag{2.35}$$

This means that the input referred amplifier voltage noise appears directly at the load's supply with no attenuation - not good. There are essentially two ways to reduce amplifier noise: increase the tail current and/or increase the transistor width. If the LDO is output pole dominant, the transistor width is stability limited (from drain parasitics). A design example is given for the oscillator regulator in Ch. 5. The current draw of the oscillator (180 μA, in this case) and the desired dropout voltage (200 mV) sets the size of the pass device (because the V^* is effectively fixed). The load capacitance should be made as large as can physically fit on the chip (to keep the output pole dominant). The amplifier devices should then be made as large as possible (and current made correspondingly large) to keep the noise low. To summarize the design: if the goal is to make the output pole dominant, and if the pass device is sized to accomodate a particular dropout voltage and load current, and if the load capacitance is fixed by area constraints, there is very little leeway in the remainder of the design. Even if it were possible to burn infinite current in the amplifier, the bandwidth constraints (from stability) will limit load regulation, line regulation, and the amplifier noise impact on the output voltage. The oscillator's phase noise for three different LDO amplifier bias currents is shown in Fig. 2.18.

When the bias current was changed, the g_m/I_D of the amplifier devices was kept constant. The minimum phase margin (at the highest bias current) was 45°.

2.4   IQ Synthesis

Synthesizing a carrier with in-phase and quadrature components may be needed for complex modulation and demodulation. In a receiver, having I and Q LO generation allows direct conversion and complex baseband processing, in addition to a 3 dB improvement in noise figure (if image noise is completely canceled). In a transmitter, having I and Q available allows for quadrature modulation for high data rate communication (cartesian quadrature amplitude modulation). There are three different ways to generate in-phase and quadrature oscillation, as shown in Fig. 2.19.
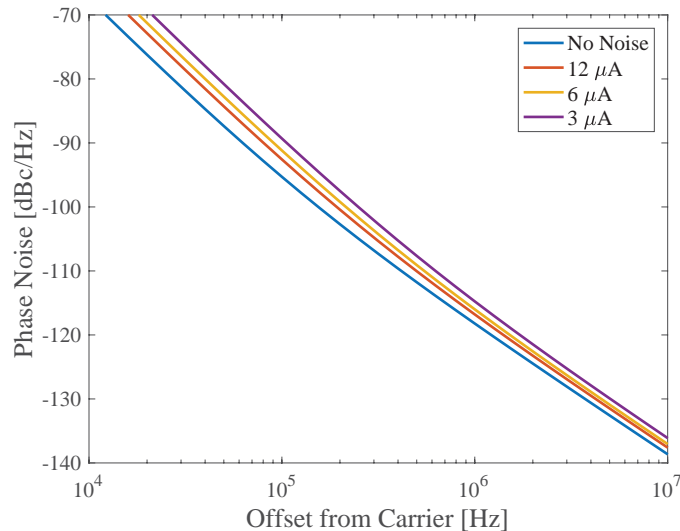
Figure 2.18: Impact of LDO amplifier noise on class-B oscillator phase noise

The quadrature oscillator has the benefit of reducing the phase noise (compared to a single oscillator) by 3 dB. However, this comes at a 2x power penalty, and, if the topology shown in Fig. 2.19(a) is used it also requires two inductors and two capacitor DACs, which is a significant increase in area.

The polyphase filter works by operating the RC pole frequency set by each resistor and capacitor at the frequency of operation ω_0. Under ideal conditions, in which the polyphase filter is un-loaded, it does not load the oscillator. However, if there is a capacitive load impedance, then the input resistance (appears in parallel with the tank impedance, thereby de-Qing the tank) is given in Eq. 2.36 [46].

$$Z_{in} = \frac{R}{1 + j} \tag{2.36}$$

So, to reduce the real loading caused by the polyphase filter, the resistor should be made as large as possible, and the capacitor should be scaled appropriately. This becomes impractical when the parasitic capacitance of the resistor becomes comparable, or larger than, the necessary capacitor. In addition, the polyphase filter is designed for one specific frequency. IQ mismatch, which limits the maximum possible image rejection ratio given in Eq. 2.37 [47] where ϵ is the voltage gain mismatch and θ is the phase mismatch in radians.

$$\text{IRR} = \frac{\epsilon^2 + \theta^2}{4} \tag{2.37}$$

The third way to generate I and Q is to operate the oscillator at $2\omega_0$, divide by two, and use logic to generate the appropriate phases, as shown in Fig. 2.19(b). This can come at a
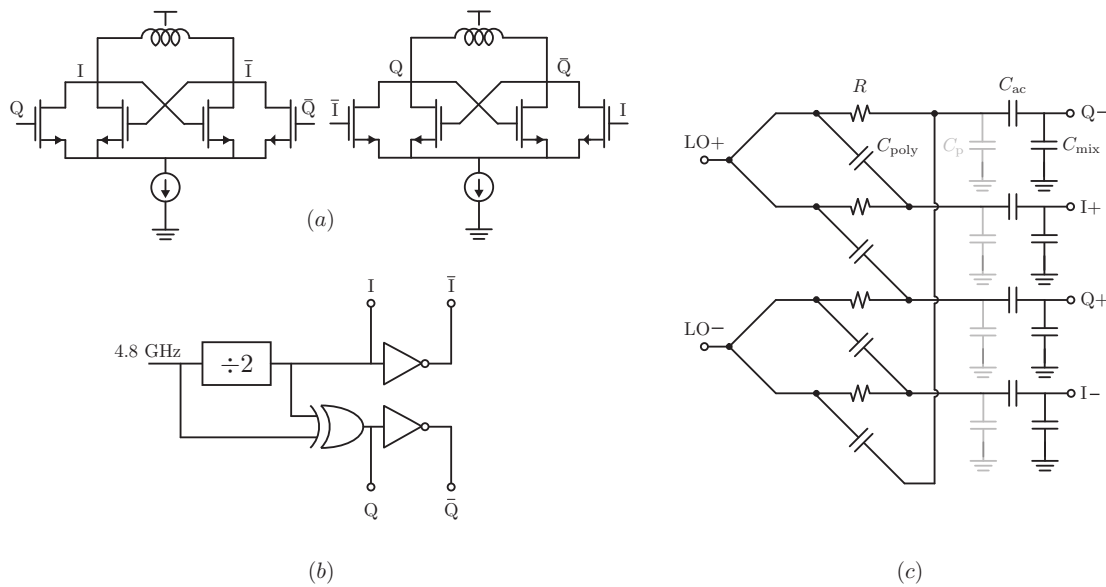
Figure 2.19: Various techniques for generating in-phase and quadrature oscillation

significant power penalty. The power-frequency tradeoff in an LC tank oscillator is not too straightforward because it depends heavily on the LQ product of the inductor (if the design is swing limited). LQ product stays approximately constant with frequency. The power-frequency tradeoff in a digital divider, however, is extremely straightforward. Presumably, the waveforms will be hard-switched, so the power consumption is proportional to $CV_{DD}^2 f$. This power consumption is highly process dependent.

Integer Dividers

Implementing dividers in these low-power settings is challenging primarily because if power is reduced in the oscillator, its swing will be low, and operating digital circuits with low swing degrades their performance. The simplest divide-by-two circuit is shown in Fig. 2.20.

Most divide-by-two circuits operate with the principle shown in Fig. 2.20, although the implementation of the latches may vary. In a traditional latch that might be found in a standard cell, if the input amplitude is small and sinusoidal, the setup and hold times are compromised by the low rise time of the clock. The rise time and small amplitude increases the on-resistance of the pass gates making it more challenging for the buffer to change the state of the latch. In a divider, this will cause cycle slipping. Alternative, faster strategies are shown in 2.21.

In 2.21 on the left is a TSPC flip flop. If the devices are sized properly (particularly the clocked gates) this divider can avoid swallowing pulses. However, as will be shown in Chapter 5, the cycle slipping is predictable and can be used to obtain larger divide ratios,
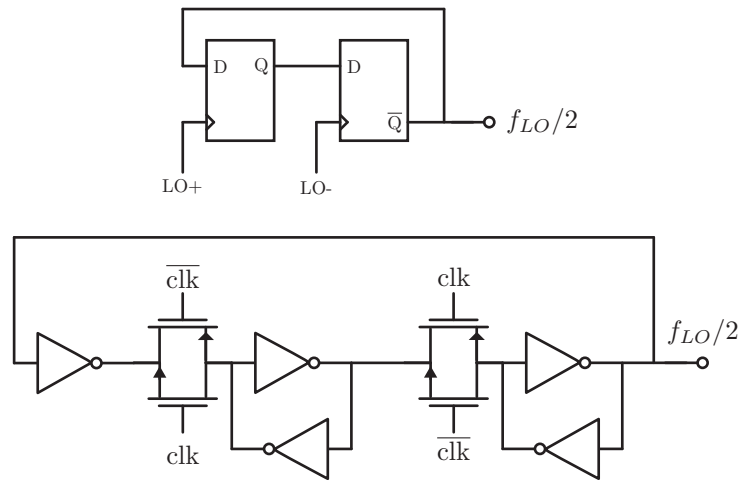
Figure 2.20: Latch-based frequency divider and traditional CMOS implementation

and operates more like an injection-locked oscillator than a traditional divider. In 2.21 on the right is a biased-CML style latch as presented in [48]. This circuit is intriguing because its current consumption is limited by the bias of the PMOS devices. Meaning, if the common mode of the input oscillation is high, the current will be quite low. And a low current actually makes this circuit less likely to slip cycles because it weakens the cross-coupled pair (this comes at the expense of duty cycle of the output oscillation). In addition, if sized in a particular way, this circuit will automatically generate 25% duty cycle waveforms (not guaranteed to be non-overlapping unless the PMOS devices maintain a low current) which is ideal for complex downconversion or upconversion mixers.

A Brief Comment on PLLs and FLLs

Recall from chapter 1 that the eventual goal of the work presented in this dissertation is to integrate every component of a wireless transceiver on a chip. This includes the frequency synthesizer. Often times, a phase- or frequency- locked loop is used to reduce close-in phase noise (meaning phase noise near the carrier frequency). However, this requires a reference that has phase noise performance that is N times better than the phase noise of the high frequency oscillator (where N is the ratio of oscillator frequency to reference frequency). Otherwise, the noise within the bandwidth of the loop will actually be inferior than the free-running case. Because of the high Q of on-chip inductors at radio frequencies, an LC tank oscillator is often the highest quality frequency reference on chip. The only exception, and the only case in which a PLL or an FLL might be used, is if the low frequency oscillator has superior resiliency to deterministic effects, like supply noise or temperature variation. If this is the case, the loop bandwidth should be extremely low so that the inferior noise
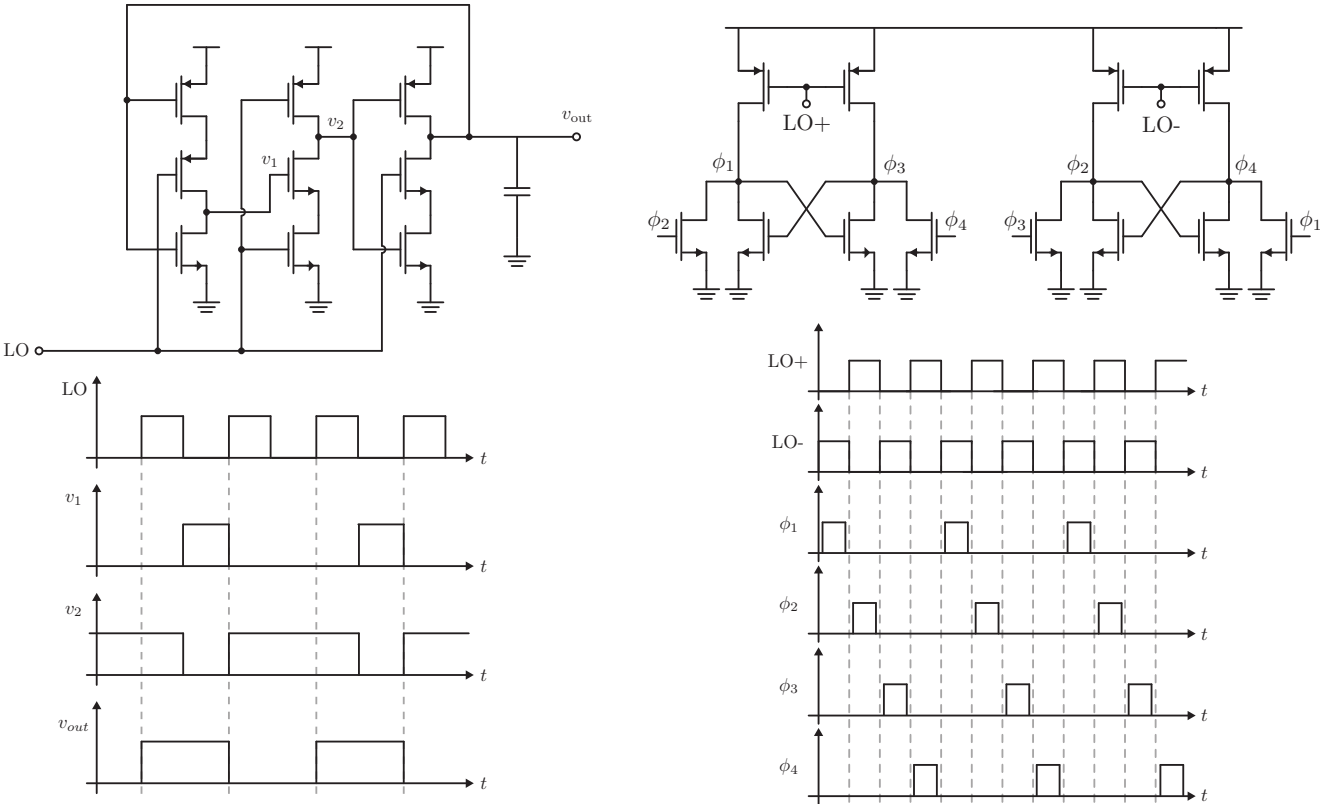
Figure 2.21: Faster divider topologies

performance of the reference does not influence the oscillator's noise.

Chapter 3

Transmitters

Much like the frequency synthesizers subsystems described in the previous chapter, transmitter system efficiencies are limited by the quality of passive components. In the case of power amplifiers - the circuit that drives the antenna - an additional constraint is the on-resistance and drain capacitance of transistors. Contrary to LC tank oscillators, switching power amplifier performance does improve with scaling technology as it is dependent on thresholds (which do not scale) and transistor on-resistances and parasitic drain capacitance, both of which do scale. While efficiency is of paramount importance, it is also important to keep the total system power budget in mind. Because transmit power is directly proportional to a wireless link's range, even a 100% efficient transmitter could be a significant - if not the single largest - power consumer in a wireless transceiver.

For that reason, all analysis done in this chapter focuses on high efficiency, low output power power amplifiers. The focus is on constant envelope modulation, or modulation schemes in which amplitude does not vary. These modulation schemes (examples include frequency modulation and basic phase modulations like BPSK and OQPSK) greatly simplify power amplifier design as they allow for the use of nonlinear switching amplifier classes that have significantly higher theoretical efficiencies compared to their linear counterparts. This increase in efficiency comes at the expense of linearity and spurious harmonic emission. Simple IoT physical-layer standards, like Bluetooth, Bluetooth Low Energy, and 802.15.4, all use constant envelope modulations.

In addition, all power amplifiers discussed in this chapter are designed for low output power. This puts significant stress on the matching network because it requires a transformation from 50 Ω to as large an impedance as is possible. In addition, the circuits required to drive the power amplifier become a significant percentage of the total system power consumption. So, minimizing the power of those blocks is critical. There are many benefits of operating at low power levels. Ground and supply bounce can be ignored as the swing at RF is always small. As long as a nonlinear amplifier is used, stability is a secondary concern.

3.1 Matching Networks

Fig. 3.1 shows an output stage driving a 50 Ω impedance with a matching network with voltage transformation ratio N (and impedance transformation ratio N^2).
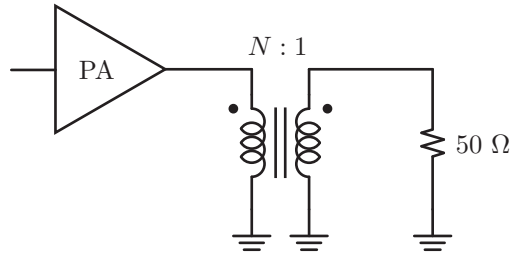
Figure 3.1: A generic amplifier and broadband matching network. The transformer steps down the voltage for low power operation, making the load impedance seen by the PA larger

The goal is to efficiently output a relatively small amount of power (somewhere between -20 dBm and 0 dBm), so the impedance Z_{out} should be made large. The voltage at the output of the power amplifier will be rail-to-rail or higher for high drain efficiency (topology dependent) and for the output power to remain low, the impedance should be very high. However, to transform 50 Ω to a large value incurs insertion loss in the matching network. Insertion loss at radio frequencies is typically dominated by the Q of the inductor, and is given in Eq. 3.1.

$$IL = \frac{1}{1 + \frac{Q}{Q_C}} \tag{3.1}$$

The Q in Eq. 3.1 is $\sqrt{m-1}$ where m is the impedance ratio. And Q_C is the component Q. For small transformation ratios, lossy components are acceptable. Suppose that a class-D output stage is used with a 1 V supply, and that the target output power is -10 dBm. That would require an impedance transformation ratio of 25. And if the Q_C of the inductor used in the match is 10 (a reasonable value for an on-chip inductor), the power lost in the match would be -13 dBm. Even if the drain efficiency of the PA were 100%, the matching network would incur a maximum efficiency of 67.2%. And the only solutions are to obtain a higher Q_C, which may not always be available, or lower the swing at the output of the power amplifier.

A transformer can also be used to perform the impedance transformation. A transformer can be made narrowband by resonating the coil inductance (thus achieving the same filtering effect as an L-match). In an L-match, it is important to consider the full system to determine which component orientation to use. For instance, if a series inductor is used at the output and the intent is to drive an off-chip load, is the wire bond inductance included in the match?

Will the output of the PA be AC coupled or DC coupled? Does the load need to be shared between transmitter and receiver? Often times external concerns like these will matching network design decisions.

3.2 Power Amplifier Topologies

Multiple dissertations and textbook chapters can (and have) been written about power amplifier topologies and their variations. The purpose of this brief overview is to discuss circuit operation from the perspective of a low power designer. The linear power amplifier topologies (classes A and B, in particular) will not be discussed in great detail. More information is available in [49]. In fact, linearity will not be discussed at all because all applications in this thesis involve constant-envelope frequency modulation. Neither AM-AM nor AM-PM distortion are relevant. For a discussion of an effective solution to linearity in switching amplifiers, albeit still at high power levels, see [50]. Similarly, because the carrier is being generated on-chip, power added efficiency is not a relevant design parameter. More information about PAE can also be found in [49]. All of these topologies also look quite similar. The differences often lie in exactly how the amplifier device(s) are driven. In addition, all of these topologies have been analyzed half to death.

Class A, B, and C

Schematics for typical class A and class B power amplifiers are shown in Fig. 3.2.
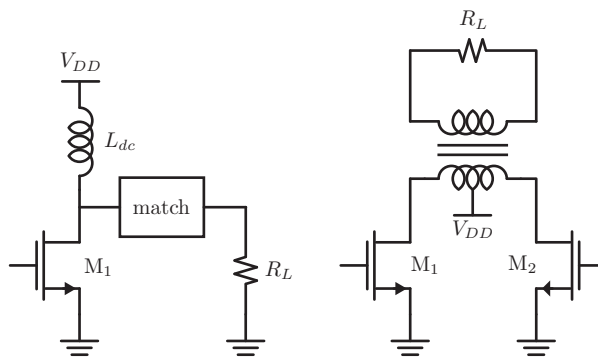
Figure 3.2: Class A and B linear power amplifiers

In the class A amplifier, the transistor M_1 is always conducting with a current exactly out of phase with the load. In addition, to obtain the maximum efficiency of 50%, the drain node of M_1 should swing between 0 V and $2V_{DD}$. In the class B amplifier, each one of the two transistors conducts for half of the switching cycle. In one half-cycle, M_1 is on and delivers current to the load, and in the other half-cycle, M_2 is on. Because of the reduced period

during which the transistors conduct, the maximum efficiency is $\pi/4$ once again at a drain swing of 0 V to $2V_{DD}$.

In both the class A and class B power amplifiers, there is a constant bias current through the amplifier. In the class A case, the single device is always conducting. In the class B case, each device is only conducting for half of the cycle. The class C power amplifier [51], [52] is, once again, identical to the others, but the conduction angle of the switch is less than half of the cycle. The narrower the conduction angle, the higher the drain efficiency, up to a theoretical limit of 100%. The drain voltage and drain currents of the three amplifiers are shown in Fig. 3.5.
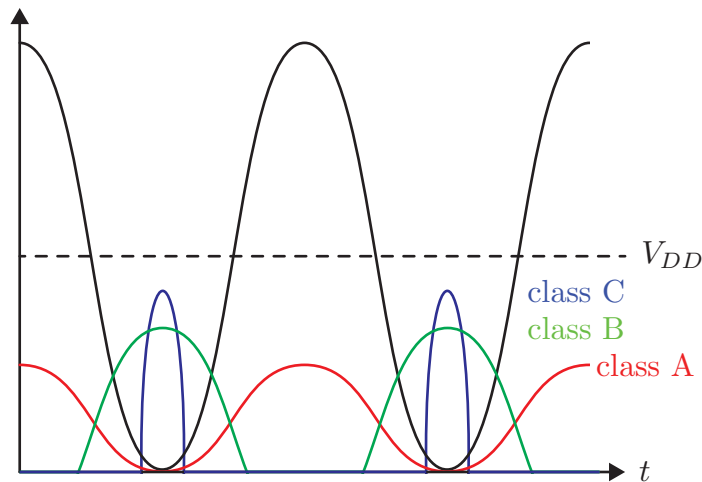
Figure 3.3: Transistor drain waveforms. The voltage waveform is in black, and the current waveforms are colored and correspond to the three different power amplifiers

Class D and E

Class D and E power amplifiers, often called switching power amplifiers, use transistors operating as switches. In proper operation, when the transistors are switching, the voltage across them is zero. This requires that the switches themselves have zero on-resistance and that the drain waveforms are correct. All of these switching power amplifiers have a theoretical efficiency of 100%. Schematics of the class D and class E are shown in Fig. 3.4.

In the class D PA, the drain voltage switches hard between 0 and V_{DD}. The matching network acts as a filter so that only the first harmonic of the waveform reaches the load. In the class E PA, the switch conducts for a half period. At the beginning of conduction, the drain voltage should be zero. The matching network will then resonate bringing the drain voltage above V_{DD} [53] [54]. In both PAs, switching should only be performed when $V_{ds} = 0$
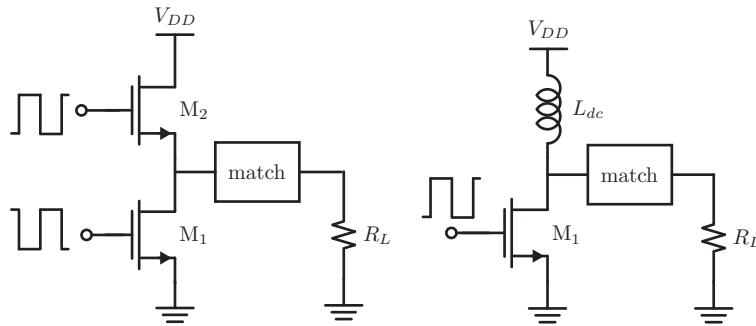
Figure 3.4: Schematics of the class D (left) and class E (right) power amplifiers

V. Sources of loss are finite switch conductance, matching network loss, and, in the case of the class D power amplifier, power lost charging and discharging the drain capacitance. In the class E power amplifier, this can be absorbed into the match.
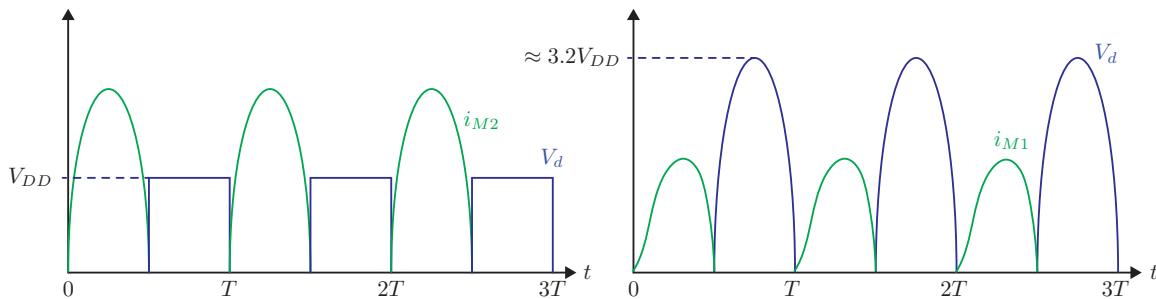
Figure 3.5: Ideal class D (left) and class E (right) waveforms

The superior ideal efficiency of switching power amplifiers makes them desirable for constant-envelope modulations. From the perspective of a low output power PA, a number of system-level concerns drive the selection of topology. For harmonics, the FCC mandates a maximum harmonic emission of -41 dBm. Depending on the output power, this could rule out the class D topology entirely unless additional harmonic filtering is performed (at the cost of efficiency) [55]. And often times, to reduce second order harmonic distortion caused by nonidealities, a feedback mechanism is required that reduces efficiency [56].

For area (number of elements), the class E uses two inductors, and the class D uses one. As far as voltage swing is concerned, the class E amplifier swings very high above the supply voltage which can place stress on the transistor and requires a higher matching network transformation ratio to obtain the same output power as the class D [26].

3.3 Switching Transients

The four other sources of loss are switch on-resistance, drain capacitance, gate drive parasitics, and poorly designed switching waveforms. Switch on-resistance and gate drive are a pretty flat tradeoff. The larger the switches, the smaller the on-resistance, but the larger the power needed to drive them (assuming CV^2 loss for hard-driven switches). Larger switches also come at the cost of increased drain parasitics, but in the class E power amplifier, this parasitic can be absorbed into the matching network. High gate drive power can be avoided, to an extent, by resonant gate drive.
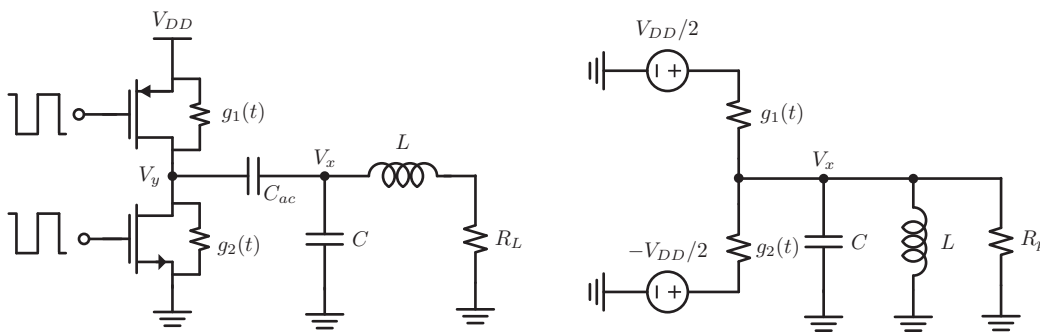
Figure 3.6: Schematic of Class D power amplifier with annotated switch conductances (left) and Thevenin equivalent model (right)

A class D power amplifier with a transformed impedance is shown in Fig. 3.6. Note that with this match, it is possible to absorb the drain parasitics of the amplifier, assuming that the AC coupling capacitor is large. Under ideal switching conditions, one of the two switches will be condcting at any given time, so the source conductance of the Thevenin equivalent circuit is a constant $g(t)$. At the first harmonic, efficiency will be limited by a voltage division between the load and the switch condutance (and by herein ignored matching network losses). However, this hard-switching comes at the expense of gate drive power. Switching waveforms are shown in Fig. 3.7.

Clearly, if the two switches are on at the same time (if $g_1(t)$ and $g_2(t)$ are both non-zero at the same time) there is a direct path from V_{DD} to ground that does not go to the load, which is disastrous for efficiency. It is preferable, then, to operate in the sinusoidal or low duty-cycle mode in Fig. 3.7. However, operation in either of these modes has negative consequences both for harmonics and for efficiency. With low duty-cycle switching, the dead time will incur class E style operation. A major difference is that there is no inductor to maintain a constant source current, and the zero voltage switching condition will be violated because the drain node will, on average, settle to $V_{DD}/2$. At every cycle approximately $C_D V_{DD}^2/4$ Joules will be expended charging or discharging that capacitor. In the case of a low input swing sinusoidal drive, the waveforms look suspiciously like class B operation from Fig. 3.5. That's because, with sinusoidal drive, the switches aren't commutating at all, they
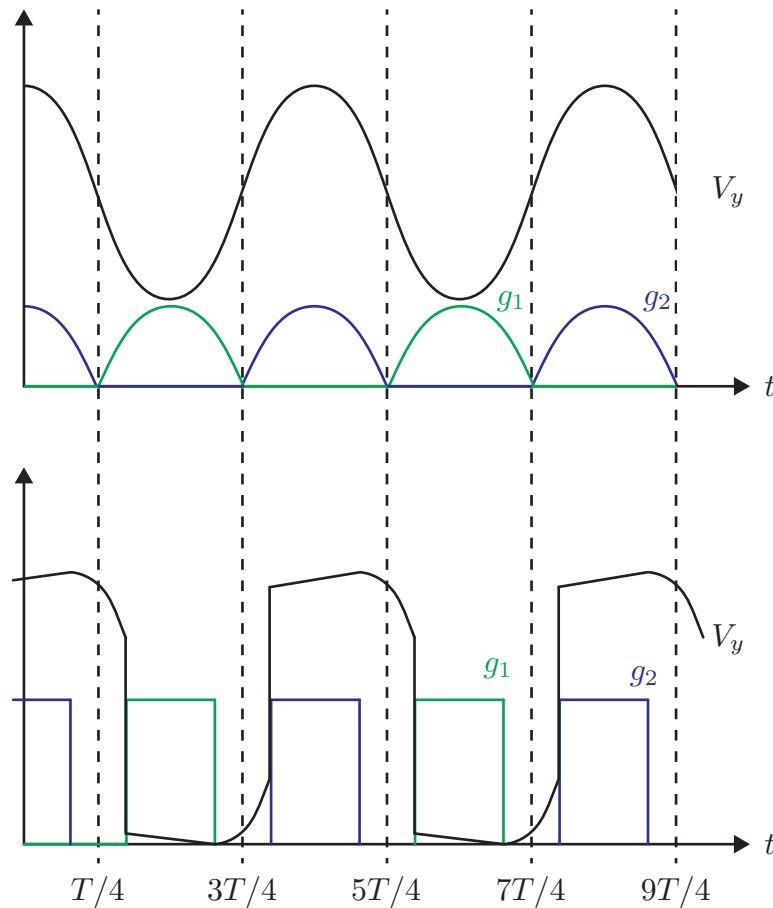
Figure 3.7: "Class D" waveforms with weak sinusoidal drive and low duty-cycle drive. The lack of inductor to maintain a fixed current results in undesirable performance both when the gates are resonant-driven (top) and square-driven with reduced duty cycle (bottom).

are biased. If this section had a single takeaway, it would be: do not operate a class B or C amplifier without a choke inductor to keep the current constant.

As will be shown in Chapter 5, a drain efficiency of approximately 25% was achieved with sinusoidal drive, and approximately 40% efficient with strong resonant drive with an output power of approximately -10 dBm and a simulated insertion loss of -1.2 dB. Unfortunately, from a system efficiency perspective, for low output power amplifiers, using sinusoidal drive is practically necessary. PA switches are inherently large to avoid switch conduction loss, and the power required to drive them with the sharply transitioning waveforms needed to achieve very high drain efficiencies is prohibitive even in deep submicron processes.

From a drain efficiency standpoint, either a class E or class C PA is better suited to be

driven sinusoidally. Unfortunately, because the drain swing will be approximately twice as high in the class C case, and nearly 1.8 times as high in the class E case. This requires a higher matching network transformation ratio, which degrades PA efficiency. Take the example from earlier. With a 1 V supply, a desired output power of -10 dBm on 50 Ω, and an inductor Q_C of 10, a class D power amplifier match will have 50 μW of insertion loss, a class C power amplifier match will have 100 μW of insertion loss, and a class E amplifier match will have 90 μW of insertion loss. In other words, the efficiency of the class D amplifier could be 1.8 times worse than an equivalent class E amplifier. This would not be a concern if the supply voltage could be reduced efficiently.

Chapter 4

The Single Chip Mote v1

This chapter summarizes the results of the first of the 2.4 GHz ISM band transceivers. We called these chips the "Single Chip Micro Motes" (or SCμM, for short). This particular iteration was a bare-bones transceiver that had a local oscillator, mixer, power amplifier, and simple analog baseband processing. The third iteration (presented in Chapter 5) is a complete bits-in bits-out 802.15.4 compliant transceiver (with some degree of Bluetooth Low Energy compliance). Functional portions of the second iteration will also be presented in this chapter. That section should only be read as a warning so that future designers do not repeat the same mistakes that I made.

The ultimate goal of these chips was to design and build a monolithically integrated transceiver that could communicate both with another version of itself and with commercial off-the-shelf transceiver SoCs. The only off-chip components needed are a power source (either from energy harvesting or from a battery) and a single antenna (on-chip antennas operating in the 2.4 GHz ISM have high loss). SCμM v1 had two separate ports for TX and RX.

4.1 Chip Overview

The first single chip mote was designed and fabricated in TSMC 65 nm GP CMOS to verify the feasibility of using a free-running, fully on-chip LC tank oscillator as the RF frequency synthesizer in a symmetrical communication system. The chip, while not quite being compliant with off-the-shelf transceivers (due to designer error, the LC tank's frequency, SCM v1's oscillator ran at a rip-roaring 2.6 GHz), was still used to demonstrate robust chip-to-chip narrowband communication without the use of an invariant frequency reference in the presence of varying temperature [24]. The block diagram and the die photo of the chip are shown in Fig. 4.1.

The receiver (schematic shown in Fig. 4.2) uses an approximately 1:5 voltage transforming matching network (with simulated insertion loss of 1.8 dB at 2.4 GHz) that correspondingly boosts the antenna impedance to approximately 1.2 kΩ. The mixer employs small
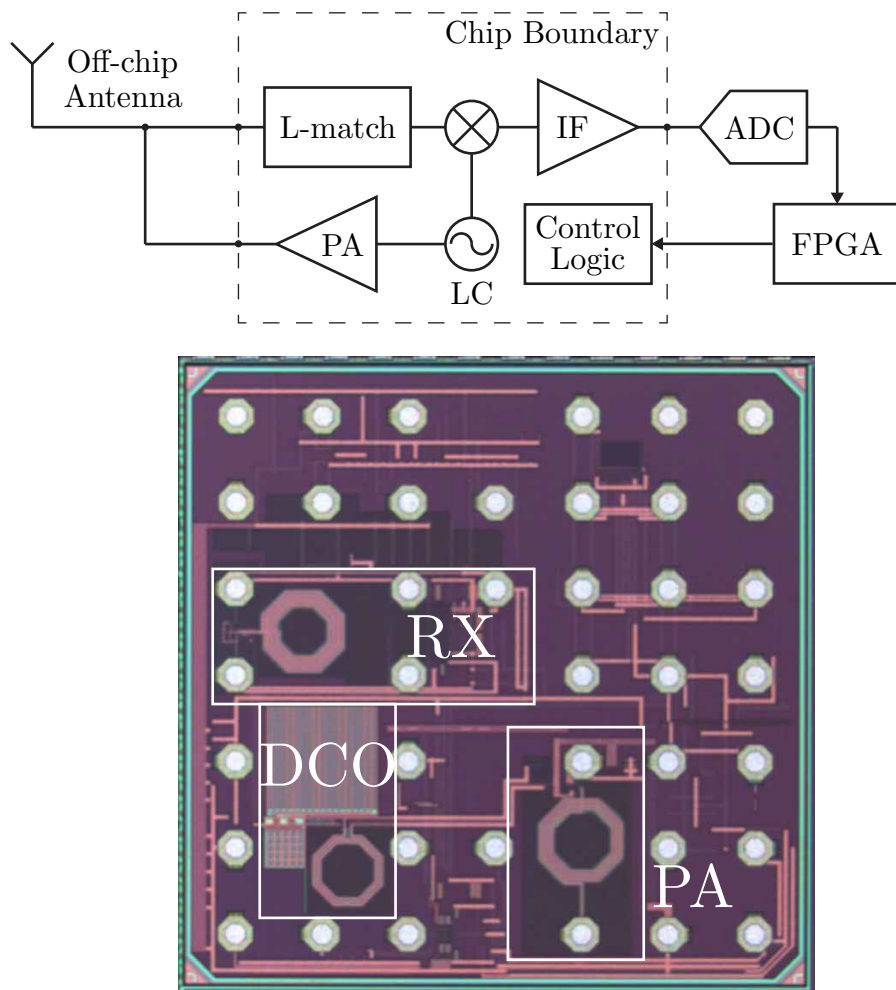
Figure 4.1: Single Chip Mote v1 Block Diagram

devices (minimum length, 3 μm width) to match this high source impedance.

The transmitter (schematic shown in Fig. 4.3) uses a class-D amplifier with a nominal output power of -5 dBm to a single-ended 50 Ω load and burns a static power of 1.58 mW for a drain efficiency of 19.8%. This power amplifier is not sinusoidally driven directly from the local oscillator - it has pre-amplifier stages to de-couple the PA and LO designs.

Overall, the system used a direct-modulation transmitter with modulation applied directly to the local oscillator. The target was 802.15.4: a 1 MHz FSK tone spacing and a data rate of 2 Mbps. The direct modulation eliminated the need for an upconversion mixer or DAC. The receiver was an in-phase only low-IF superheterodyne architecture. The single channel reduced the complexity and power of the receiver (at the cost of 3 dB sensitivity from noise in the image band), as it does not require generation of a quadrature oscillation.
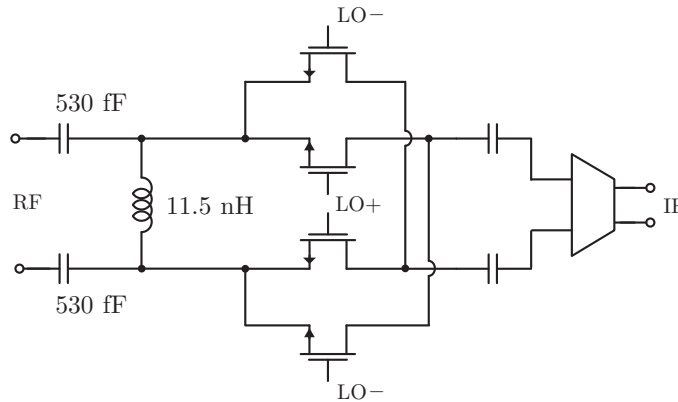
Figure 4.2: Single Chip Mote v1 Mixer-First Receiver
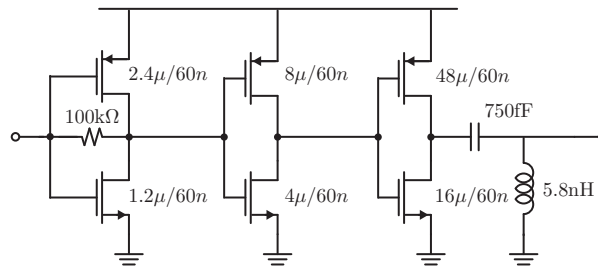
Figure 4.3: Single Chip Mote v1 Class-D Power Amplifier

The low-IF architecture eliminates the need for offset cancellation from LO self-mixing and avoids receiver flicker noise.

Local Oscillator Design and Measurements

The local oscillator is a CMOS LC tank that uses a degenerated capacitive DAC for fine resolution frequency tuning (as first proposed in [57]). The oscillator consumed 1 mW of power (with 1 V supply) and has phase noise of -92.1 dBc/Hz at 100 kHz offset for a figure of merit of -182.1 dB. The tank has a 3.2 nH inductor with a Q of 9.2. The total tuning range of the oscillator is 2.6 GHz to 3.1 GHz. While this did fall outside of the 2.4 GHz ISM band, it could still be used to demonstrate narrowband crystal-free communication between two single chip motes. Fig. 4.4 summarizes these results.

In Fig. 4.4, (a) shows the schematic including where the direct modulated MSK is performed, and (b) shows the tuning characteristics of both the coarse (drain) and fine (source) capacitive DACs. (c) is a plot of the measured and simulated phase noise of the oscillator up to the noise floor of the instrument, and (d) is a histogram of the oscillator's frequency (measured via zero crossings) after being downconverted to a 2.06 MHz intermediate frequency. Data in Fig. 4.4 was collected over a period of 4 ms, which is slightly less than
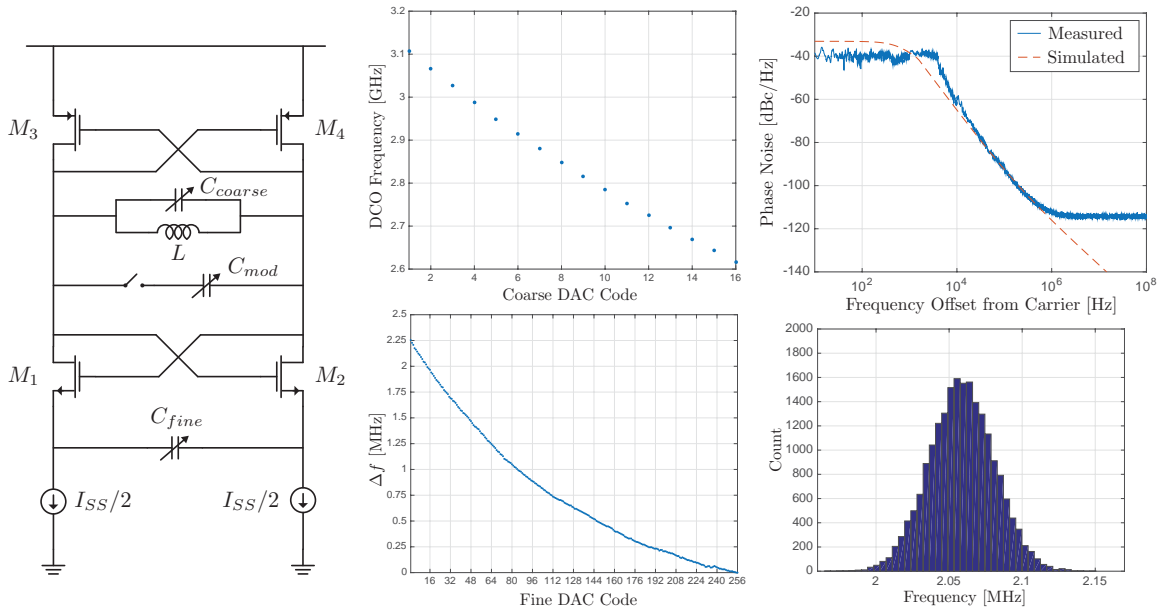
Figure 4.4: Overview of the Single Chip Mote v1's digitally controlled oscillator

the duration of the longest possible 802.15.4 packet (the longest packet is 4.256 ms). The frequency error allowed by the standard is ±40 ppm, or ±104 kHz on a 2.6 GHz oscillation. The rms frequency error of this oscillator was 22 kHz over the duration of the packet - well under the requirement set by the standard. This is confirmed by the transmitter's EVM measurement shown in Fig. 4.5
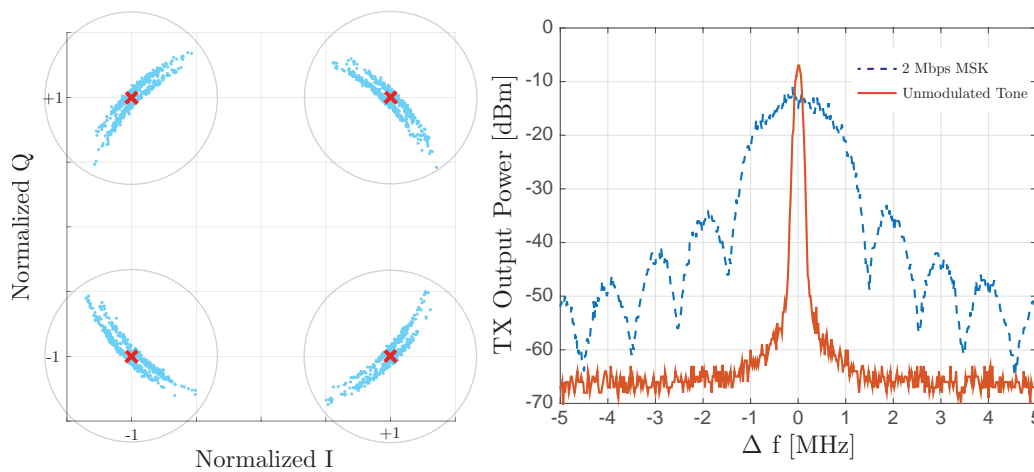
Figure 4.5: Single Chip Mote v1 transmitter performance. The transmitter EVM is 3.4%. The required transmitter EVM for 802.15.4 is 35%

In Fig. 4.5 the left figure shows the transmitter's output downconverted and demodulated as OQPSK. The red x's represent the four ideal constellation points, the blue dots represent the demodulated chips, and the grey circles are the 35% maximum EVM allowed by the 802.15.4 specification. This plot was generated using a receiver carrier recovery loop operating with a modest 100 kHz loop bandwidth. The right figure shows the unmodulated and modulated spectra of the transmitter. Notice the 1.5 × datarate notches in the modulated spectrum indicative of a modulation index of precisely 0.5.

While Fig. 4.5 demonstrates the feasibility of using a free-running oscillator over the duration of a single packet, it fails to capture effects of long-term drift and variation of frequency with external conditions, like temperature. Naturally, if the oscillator were compensated in a PLL or FLL with a PVT invariant reference (like, say, a crystal), this would not be a problem.

Temperature Compensation

However, a crystal frequency reference is a luxury that has no place on a transceiver with no external components. Fig. 4.6 shows the effect of temperature on the open loop oscillator's frequency.
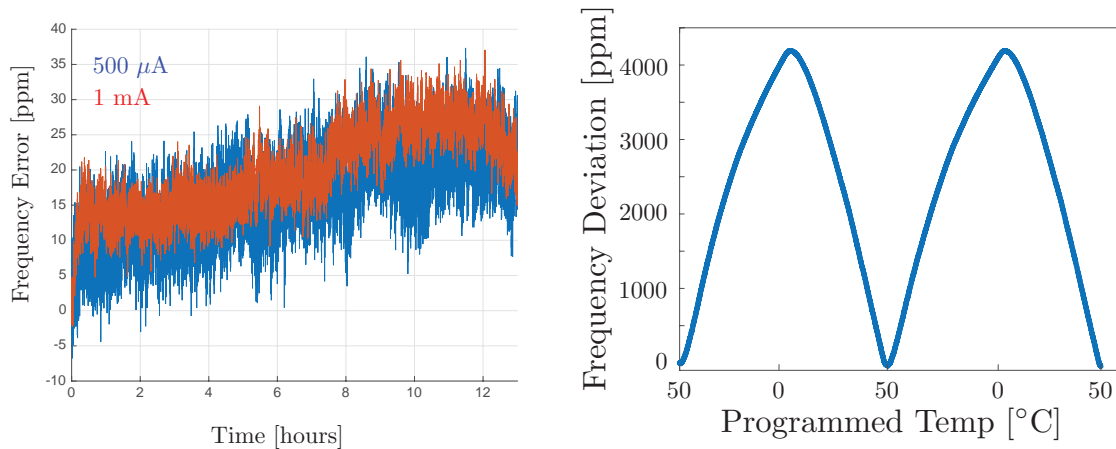
Figure 4.6: Variation in the Single Chip Mote v1's oscillator frequency in a temperature controlled environment (left) and in a programmed temperature chamber (right)

The oscillator varies with a temperature coefficient of approximately 100 ppm/^{o}C (compared to a simulated temperature coefficient of 25 ppm/^{o}C) which means that it would take a temperature change of less than 1°C for the oscillator to fail to meet specifications. While the vast majority of off-the-shelf receivers would still be able to track this transmitter as long as its frequency does not drift out of channel, this frequency error has other implications. In a scheduled frequency-hopping mesh network, for instance, if temperature changes untracked

(the industrial standard sets a maximum temperature change rate of 7°C/minute) [58] the radio could send a packet at the wrong time or on the wrong channel which could cause a collision (two packets interfere) or a missed transmission (no receiver is listening). Both consequences cause a packet error or an increase in the latency of the transmission.

Previous efforts in crystal-free radios have either suffered reduced sensitivity due to large receiver bandwidth [13] or utilize non-standard modulation approaches [59]. On this chip, we used the incoming RF tone from another chip to track oscillator frequency error. To test the feasibility of this approach, we injected an invariant 2.6 GHz RF tone into the receiver. We downconverted the tone and "de-modulated" it with a zero-crossing counter. The demodulation was effectively a one-bit oversampled comparator output and included clock and data recovery on FPGA. More information can be found in [60]. This counter ran for a prescribed amount of time and, after that time, was compared to an expected count. If the count was low, the local oscillator was sped up using the fine DAC. This was only possible within the limited range of the fine DAC. Results of this test are shown in Fig. 4.7.
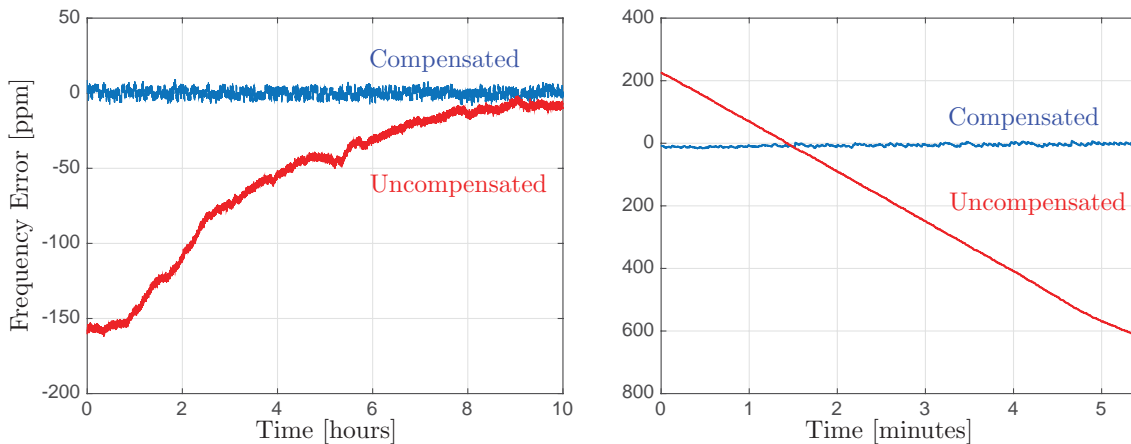
Figure 4.7: Variation in the Single Chip Mote v1's oscillator frequency in a temperature controlled environment (left) and in a programmed temperature chamber (right)

Updates in the oscillator's frequency were at a 10 Hz rate. Fig. 4.7 demonstrates that a crystal-free oscillator could run in a mesh network where other radios in the network have perfect frequency references (like off-the-shelf transceivers). But the bigger question is: can a mesh network comprised of single chip motes maintain a concept of channels and time? To test this, a transmitter mote was placed in a temperature chamber and subject to a 2^oC/min. temperature ramp, and the receiver was subject to standard lab conditions outside of the chamber.

There's one additional caveat: the transmitter is sending 802.15.4 formatted packets (preamble, start symbol, data, and CRC) at a rate of 5 packets/second, and the receiver only does a count-value comparison when it receives a valid packet. Therefore, baked into
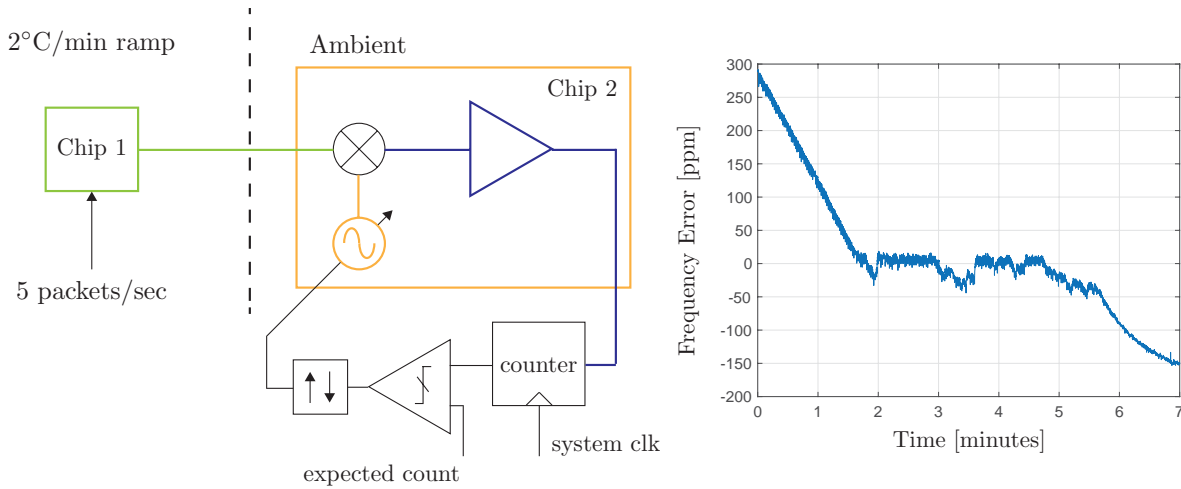
Figure 4.8: Block diagram of packet-level temperature compensation (left) and resulting receiver frequency error (right)

this more realistic experiment is a test of resilience against packet errors. The block diagram of the feedback and the results of the packet-based compensation are shown in Fig. 4.8. Note that temperature compensation is only performed between (approximately) minute 2 and minute 6 of the experiment. That is because of the limited range of the fine DAC used to execute adjustments in the oscillator frequency. The oversampling system clock was 100 MHz and was source from a function generator. Jitter was added with the generator's modulation function (it was modulated with noise). The IF was counter for the duration of an 802.15.4 packet.

In conclusion, the single chip mote version 1 was used to test and demonstrate the feasiblity of using free-running LC tank oscillators in an IEEE 802.15.4 transceiver. The frequency drift from noise alone 4.6 (left) was found to be less than the specified 40 ppm over a 13 hour time period. Temperature variations caused significant frequency shifts, but temperature varies relatively slowly and can be compensated over-the-air by maintaining a constant frequency difference between the incoming RF and on-chip LO. We demonstrated that a free-running transmitter drifting due to a 2°C/min. temperature ramp can be tracked by a receiver to within a standard deviation of 25 kHz. Furthermore, an open loop direct modulation transmitter was shown to achieve 3.4% EVM for MSK demodulation with a 1 MHz tone spacing at a data rate of 2 Mbps, as shown in Fig. 4.5. The use of fully integrated free-running RF oscillators enables a significant reduction in power by eliminating the need for a PLL, reduces the system cost by removing the off-chip crystal oscillator, and is a step towards fully on-chip integration of wireless transceivers.

Test Setup

SCμM v1 had a low degree of integration. Most of the baseband processing was performed off-chip with a Tow-Thomas filter and comparator. The receiver itself was a zero-crossing counter. Recall that 802.15.4 uses a specific type of FSK modulation, so if the receiver downconverts to a low intermediate frequency, a "1" bit will have a different number of zero crossings in a bit period than a "0" bit. In addition, all of the local oscillator biasing and all power conditioning was performed off-chip. All LDOs were discrete components, and the LO's bias current was mirrored on chip, but was originally generated with a TI LM334 current source with temperature compensation. The temperature coefficient of the oscillator was considerably higher than simulated, and the going theory was that this chip had a higher temperature coefficient than advertised (likely due to user error, not because of any inherent problem with the chip).

4.2 The Single Chip Mote v2

For the sake of completeness, I will include a small subsection on the second chip in the Single Chip Mote lineage. From a radio transceiver perspective, the chip was unsuccessful. One benefit was that the radio state machines and Cortex M0 were tested in silicon for the first time. This led to the more successful Single Chip Mote v3, which will be presented in the following chapter of this work. A block diagram of the chip is shown in Fig. 4.9. Not shown was feedback to automatically perform the carrier recovery that was demonstrated on SCM v1.

For the receiver, the local oscillator was a current-starved ring oscillator. Bias generation and supply conditioning was performed on chip. A schematic of the receiver is shown in Fig. 4.10. Unfortunately, the phase noise and drift of the ring oscillator was so large that the RF downconversion was not functional. This was caused by a combination of LDO noise (the LDO amplifier burned 500 nA) and substrate noise (there was no "bottom" device in the ring oscillator). [61] uses a very similar topology but had the foresight to use a integrated PLL and external crystal reference to reign in the ring's considerable jitter. The latching devices are a common delay-cell technique, as described in [62] and in [63].

For the separate transmitter port, the oscillator was a separate LC tank oscillator. It burned 250 μA from a 1V supply (bias mirrored from off-chip). Modulation (once again, FSK directly modulated onto the oscillator itself) was performed with a varactor and an R2R DAC. The same extraction mistake was made on this chip as on SCM v1 (gate level vs. transistor level resulted in double-counting of MOM capacitors in the layout) so the oscillator was tunable between 2.7 GHz and 3.2 GHz. The power amplifier was constructed with a somewhat interesting topology. The goal was to efficiently generated a very low output power (between -20 dBm and -10 dBm) from a 1 V supply. This was performed both by stacking two Class-D PAs on top of one another (inspired by [64]) and by using a 2:1 step-down switched capacitor DC-DC converter to reduce the PA's supply from 1 V to 500
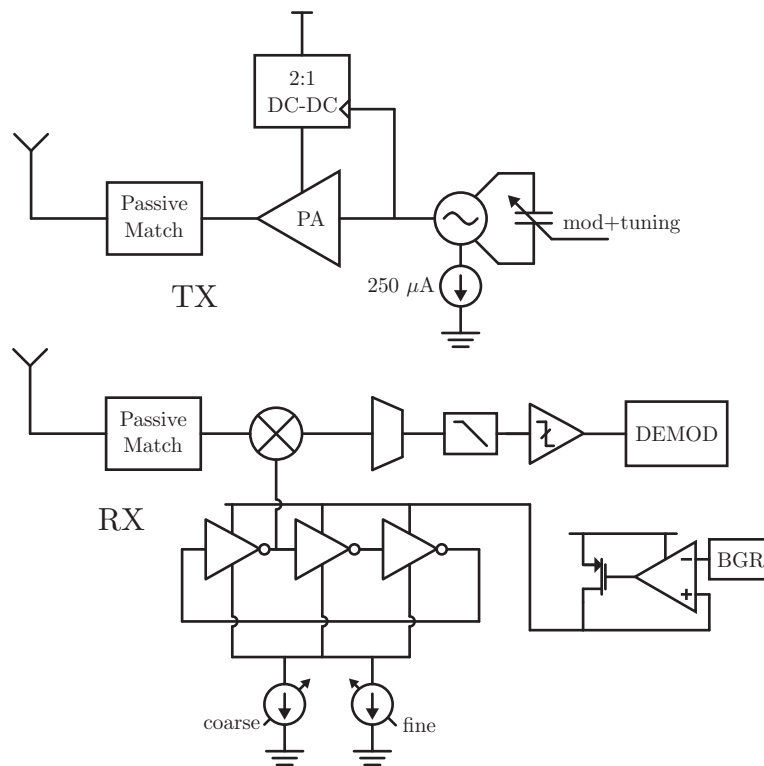
Figure 4.9: Top-level schematic of the SCM v2 TRX

mV. The transmitter schematic is shown in Fig. 4.12. To eliminate the effect of switching spurs from the DC-DC converter, the converter operated at the RF transmit frequency, and the large switches were resonated as part of the oscillator's resonant load.

Although many of the ideas on this chip were interesting, the execution was lacking. The power amplifier did operate with a 500 mV rail delivering -15 dBm of output power with approximately 20% drain efficiency. However, the 2:1 step-down DC-DC converter did not function, as shoot-through current through the power amplifier completely discharged the load capacitance at every cycle. In fact, in this process, due to transistor drain parasitics, it is extremely difficult (if not impossible) to have a DC-DC converter switching at RF with an efficiency $> 50\%$. I still think that it is an interesting idea that should be investigated in a more advanced process node. As mentioned earlier, the ring oscillator's noise was too high to perform meaningful downconversion. In addition, due to capacitive parasitics, the ring oscillator burned approximately 150 μA from a 1 V supply, which is actually more power than a LC tank oscillator with superior noise performance (see Chapter 5). In addition, because of the wide tuning range and narrow current resolution needed, the area of the oscillator+tuning was approximately 200 μm by 300 μm. Some things about this chip did work and were subsequently used in SCM v3. All of the bias generation including band gap
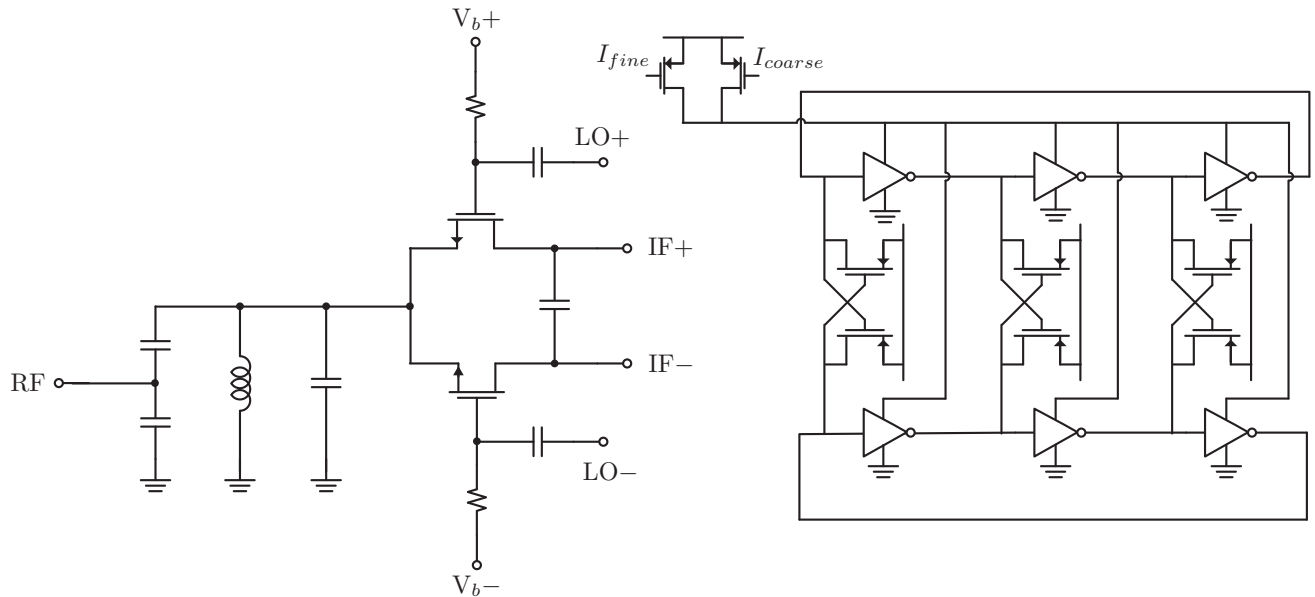
Figure 4.10: Schematic of SCM v2 front-end and oscillator

references and current sources were functional. In addition, the digital baseband and the synthesized Cortex M0 were verified.

There are two arguments for using a ring oscillator over an LC tank oscillator: tuning range, and area. However, the penalty in performance (both phase noise and power) is dramatic. Furthermore, and the real damning factor regarding ring oscillators, is how challenging they are to simulate accurately. Integrated circuit simulators are extraordinarily advanced, and will account for every factor that they can. But ring oscillators are so sensitive to the slightest variations in environmental conditions (temperature and voltage, primarily) that are difficult to include in models, that their performance is often substantially worse than the simulators predict. To have any chance of success, it is necessary to heavily regulate supply voltages, maintain low impedance paths to mitigate effects of external interference, and, likely, have some degree of frequency correcting loop. It is possible to make a ring oscillator radio work [61], [13], [65], but, any benefits are outweighed by the support structures that the oscillator needs to function. On SCM v2, all bias was generated on chip, all supplies were regulated with 30 dB of simulated supply rejection, and the phase noise was still two orders of magnitude greater than expected.
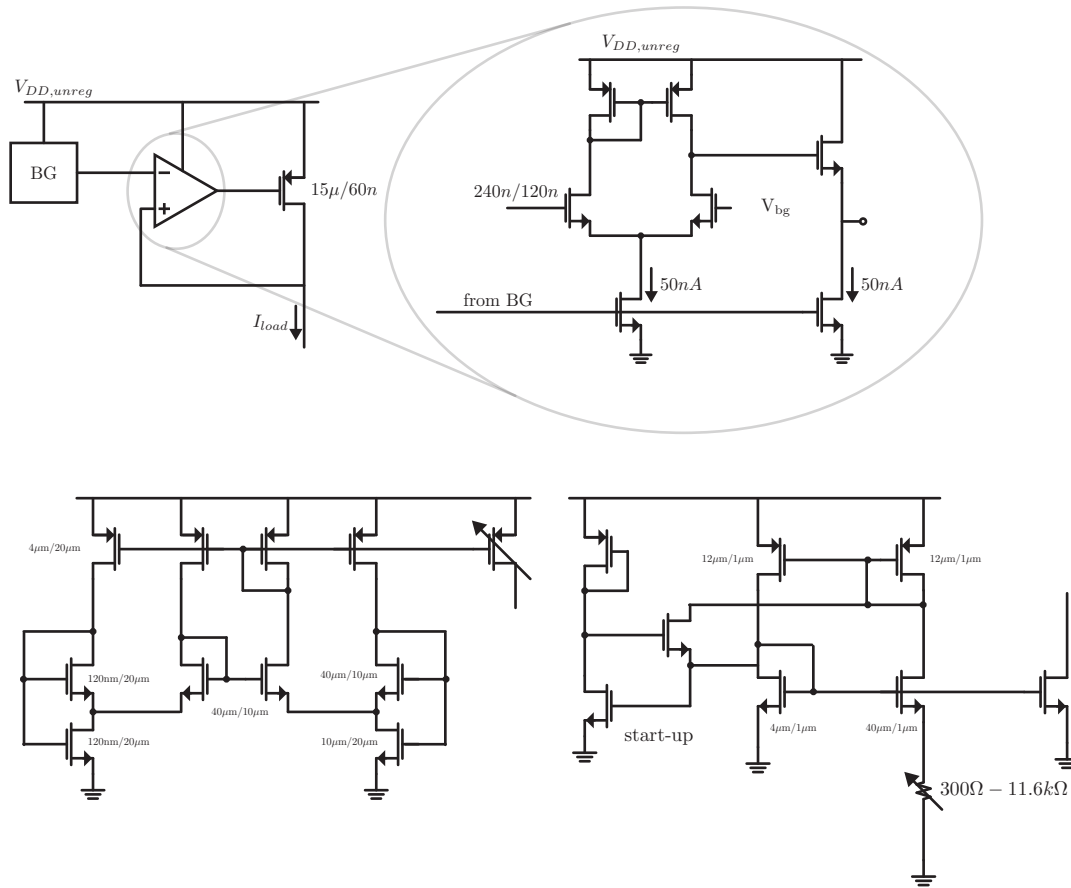
Figure 4.11: Detailed schematic of SCM v2 biasing and supply conditioning circuits (LDO, fine and coarse current sources). The coarse current source (bottom right) is a standard constant-g_m current reference. The fine current source (bottom left) is similar to a constant-g_m but with additional degeneration to significantly reduce the generated reference current.
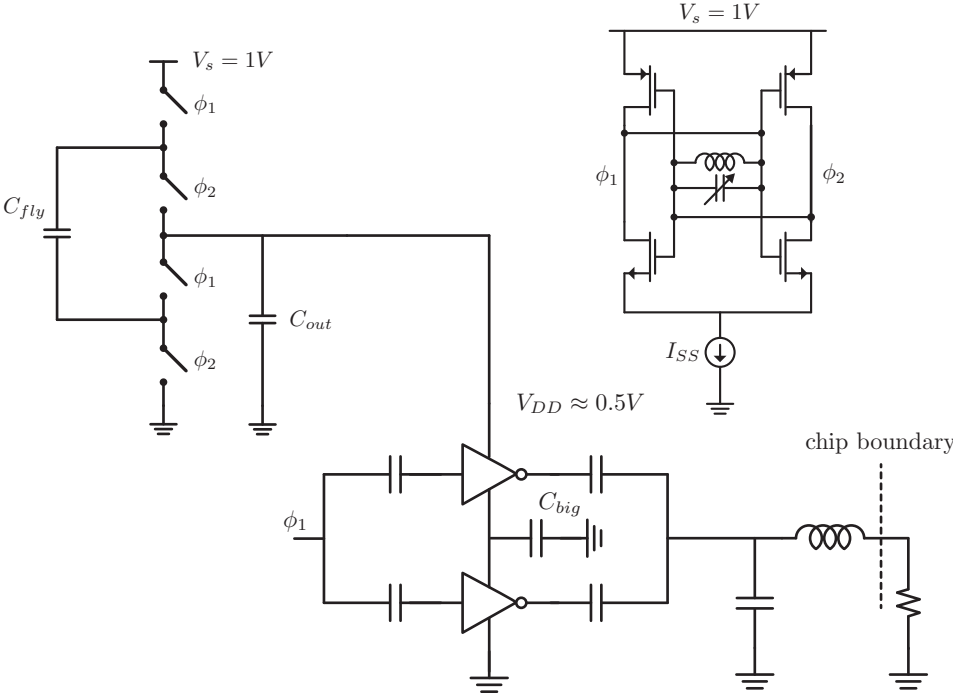
Figure 4.12: SCM v2 transmitter schematic

Chapter 5

The Single Chip Mote v3

5.1 Chip Overview

The third iteration of the single chip mote was designed and fabricated in TSMC 65nm LP CMOS and was intended to be a fully-integrated 802.15.4 transceiver with baseband processing, radio state machine [66], Cortex M0, and an ADC to interface to an external sensor or internal on-chip temperature monitor. The transmitter (but not the receiver) was also designed with Bluetooth Low Energy compatibility in mind. As mentioned in the introduction, there is a high degree of similarity between 802.15.4 and BLE modulation schemes (a reminder: MSK, equivalent to 802.15.4's half-sine shaped OQPSK, versus Gaussian FSK). And, in contrast to 802.15.4, most mobile phones have a receiver capable of receiving BLE transmissions and advertisements, which makes BLE an attractive target for a chip that people may actually use in the future. An annotated die photo of the chip is shown in Fig. 5.1. The Cortex M0 (with 128 kB of SRAM) occupies an area of approximately 1.5 mm by 0.9 mm. The analog portions of the radio occupy 1 mm by 0.96 mm, and the digital baseband, transmit state machine, and scan chain occupy 930 μm by 560 μm. This chip was a tremendous multi-student and postdoc undertaking, and I encourage you to read the acknowledgements for more details.

One benefit of having a full system-on-chip was the ability to test blocks and demonstrate functionality with software. The software to use the transmitter, oscillator, and divider on the chip is documented in Appendix A. In addition, the high degree with which the radio can be controlled directly by the software is quite unique. Most wireless systems-on-chip (besides possibly Nordic Semiconductor SoCs) obfuscate much of the radio's low-level operation. The versaility of the single chip mote as a wireless platform allows for significant system-level experimentation.

The block diagram of the chip is shown in Fig. 5.2. The highlighted portions are the ones that will be discussed most in this chapter - each one having its own subsection. The first subsection, which is not highlighted at all in the figure, is about supply conditioning and bias generation. A major goal of this chip was to operate with only two off-chip connections:
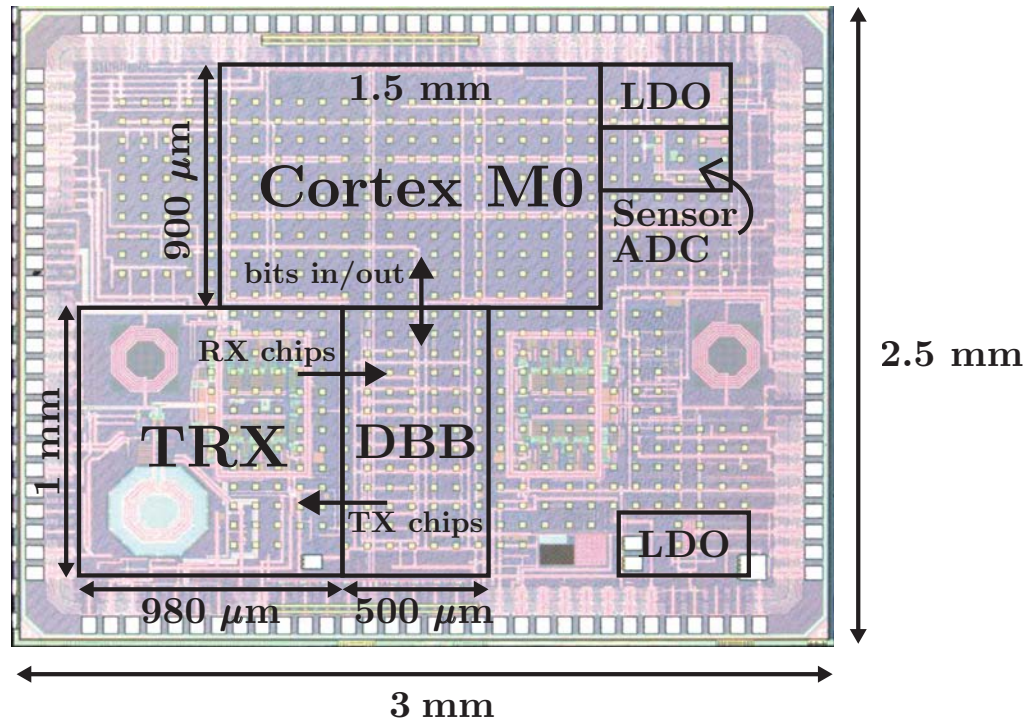
Figure 5.1: Annotated SCM v3 die photo. The entire chip was 7.5 mm^2. The transceiver active area was approximately 0.98 mm^2. The total active area (counting LDO decoupling capacitance, but not global supply decoupling capacitance or sensor ADC) was approximately 3.06 mm^2

battery and antenna. All bias circuits were required to maintain consistent analog and RF performance in the presence of adverse conditions, namely temperature and battery voltage variation. In addition, because the chip has a significant digital component, the supply circuitry needed to be designed to decouple large digital switching transients from the radio. At best, these transients could cause spurs and reduce performance, and at worst, current spikes could cause catastrophic failures in the analog components. The next section is about the design of the local oscillator. Primary design considerations were tuning range and resolution, minimizing power while maintaining good phase noise performance, and having an oscillator that can appropriately drive both transmitter and receiver circuitry. All load circuits are built-in as part of the resonant load of the oscillator; there is no buffer. One additional consideration was the direct modulation (effectively demonstrated on the Single Chip Mote v1, see previous chapter) of both 802.15.4 MSK and BLE GFSK. The subsequent section is about the performance of the power amplifier, design decisions made for appropriate co-operation with the shared-antenna receiver interface, and the consequences of those design decisions. Next, I discuss the integer-N divider attached to the LO. The divider itself is relatively high power and should only be used sporadically in actual operation. It was built

for low duty-cycle frequency compensation of the local oscillator to account for drift. It was implemented with a structure similar to an injection-locked ring oscillator. It is a standard TSPC latch divider that is intentionally sized to cause predictable cycle slipping. Finally, at the end of the chapter, I discuss the crux of the system-level problem: how to use this local oscillator and transmitter in an actual network without a crystal oscillator. Demonstrations of operation as a 802.15.4 transceiver and BLE advertising transmitter are presented as pudding-level proof.
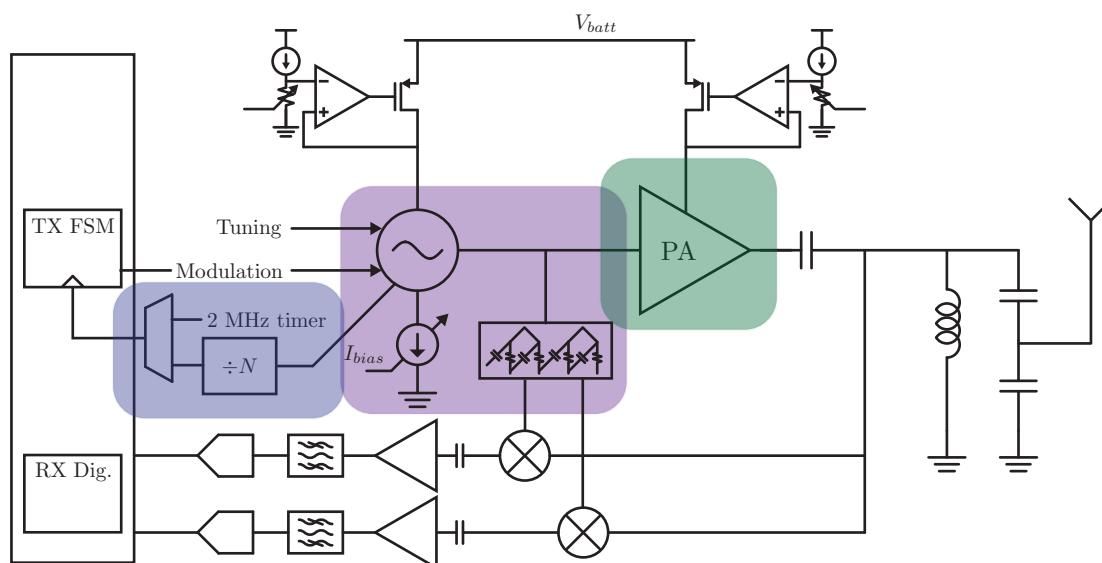
Figure 5.2: Block diagram of the transceiver

5.2   Supply Conditioning and Bias Generation

A high-level schematic of the bias networks for the RF portion of the chip are shown in Fig. 5.3. In this section I will discuss each circuit in some level of detail. The blocks were: a fractional band gap voltage reference, two 2T reference voltages, a constant g_m current reference, and an individual LDO for each of the three major RF blocks. These circuits were designed to operate with battery voltages between 1.2 V and 1.5 V, although with calibration it is likely possible that the RF system could work with battery voltages between 1 V and 1.8 V, although this has not been verified.

(a)

(b)
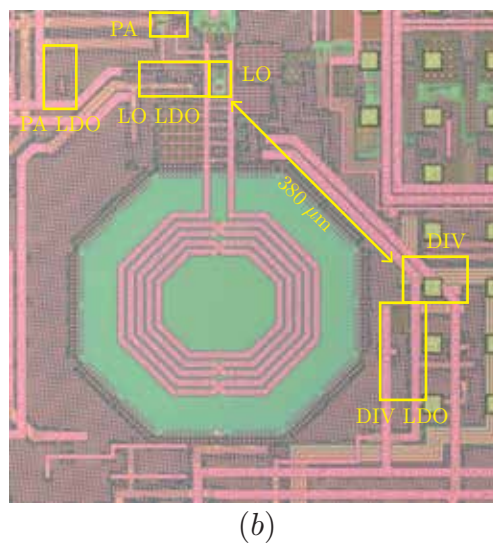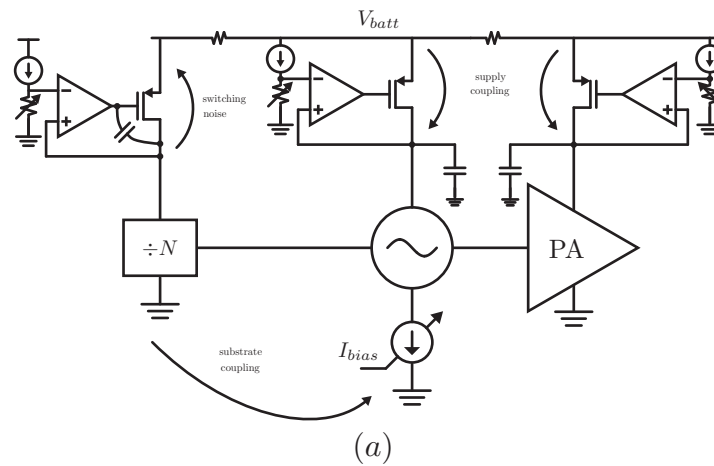
Figure 5.3: Power supply network showing the most likely sources of coupling (a) and the physical layout of the transmitter (b)

Bias generation

Every bias voltage and every bias current must be generated from a single 1.5 V supply. 1.5 V was chosen because it is a relatively standard alkaline battery voltage, and because many of the physically smallest available coin cell batteries are at or near 1.5 V. There were three separate reference voltages generated on chip. The first was a fractional band gap used as a reference voltage for each on-chip LDO. It is based on the circuit from [67], and a schematic with annotated transistor dimensions is shown in Fig. 5.4.

Each branch of the circuit in Fig. 5.4 drew 1 μA. The simulated DC supply rejection was
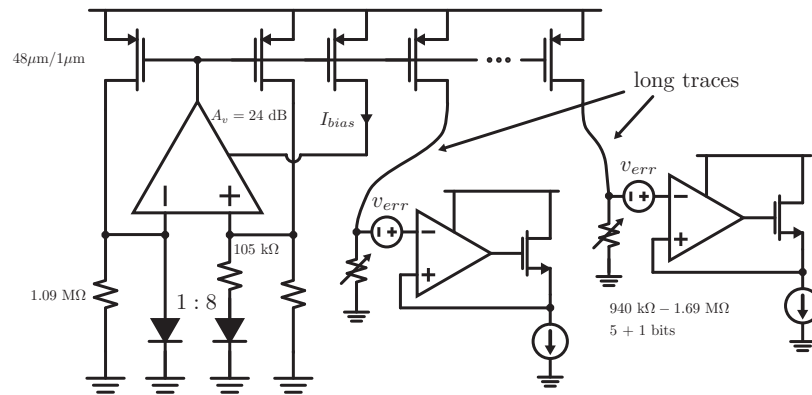
Figure 5.4: Reference routing network. The bias current for the band gap generator amplifier is mirrored from the mirrored PMOS current (shown in the figure as I_{bias})

28.6 dB at 1.5 V battery voltage, and 26.7 dB at 1.2 V battery voltage. The intent of this design was to mirror the reference current many times and distribute the current to multiple distant and disparate locations. The current sources were co-located for matching and for ease of layout, and the output currents were routed to distant locations on the chip. Then, each LDO that used the BG as a reference had its own resistor DAC. The primary reason for this was so that the core band gap circuit did not need to be replicated in multiple places on the chip. Note that there are multiple large resistors - the total area of the band gap reference core was approximately 100 μm by 85 μm (primarily resistors) which is a significant chunk of real estate. One additional benefit of distributing the current to each LDO was that the input-referred offset of every individual LDO, which could vary, could be tuned out. This resistor DAC could be used to tune the reference voltage between 750 mV and 850 mV (or approximately between 1.05 V and 1.15 V in a so-called "panic mode" that proved to be extremely useful). This tuning was performed manually via scan chain settings. The panic mode was implemented by an extended range resistor DAC. In addition, different blocks of the chip could operate on slightly different supply domains. One negative consequence of routing this current is that there is no real way to mitigate external signals that could couple onto the long traces. In fact, because the gain of the loop and the currents through the transistors are so low, it is a relatively high impedance even at the BG side of the circuit. In addition, the large transistors and small currents mean that the noise contribution of the reference is quite high. And, because the output voltage follows the reference voltage exactly (within the LDO's loop bandwidth), reference noise cannot be tolerated without compromising RF performance. To mitigate this, large, bandwidth-limiting capacitors were used (approximately 45.2 pF for each of the LO and PA LDOs), once again at a significant area penalty. The tradeoff between noise and capacitance (effectively area) is shown in Fig. 5.5. At this narrow bandwidth, the noise, predictably, becomes dominated by the flicker

noise of the current mirror devices. Spending area to reduce reference noise as opposed to using the capacitance to reduce LDO noise was primarily an ease-of-layout decision. And, adding filtering capacitance to LDOs often has negative stability implications. No startup circuits or diode trim were included in the band gap. There was sufficient leakage in the diodes (total area of 35 μm by 35 μm) for the loop to start up. Startup conditions were only tested at room temperature.
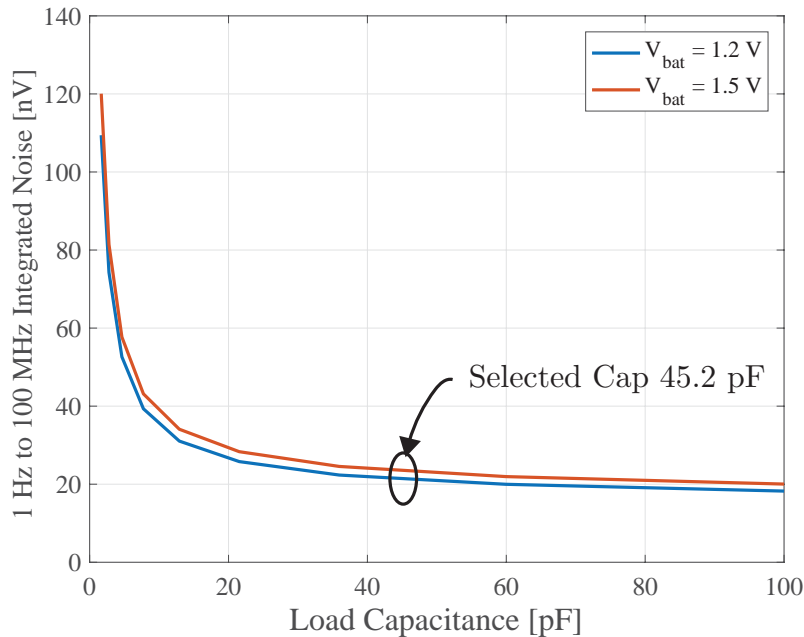
Figure 5.5: Band gap reference noise/area tradeoff

The second and third voltage reference generated were so-called two transistor bias circuits first proposed in [68]. The schematic of the two circuits is shown in Fig. 5.6. One of these was used to generate a PTAT (proportional to absolute temperature) voltage that could be sampled by the ADC and used as a temperature measurement. The other was intended to be temperature invariant so that it could be used as the bias voltage for the ADC.

Simulation results of the two-transistor circuits are shown in Fig. 5.7. These results demonstrate both operation under designed conditions (800 mV supply voltage) and with varying supply and process. Unlike [68] the circuits used on this chip did not use a native device. Instead, two NMOS devices with different threshold flavors (lvt and hvt, in this case) were used. The output voltage is still approximately the difference in thresholds between the two devices. Threshold variation has impact on the output voltage. The mismatch plot in Fig. 5.7 was performed by applying a static DC voltage of approximately appropriate level to the gate of the smaller of the two transistors in each circuit. Because of noise
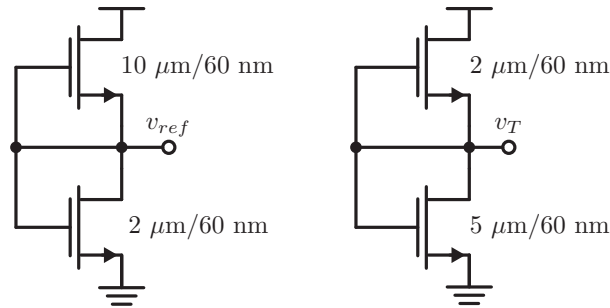
Figure 5.6: Schematic of temperature invariant v_{ref} and PTAT v_T circuits with annotated transistor dimensions

issues in the ADC and PGA, these voltages could not be measured on the chip. Because of nonlinear effects, primarily gate leakage, the temperature characteristic of the PTAT is slightly quadratic between 0°C and 80°C. In nominal conditions, the slope of the voltage varies by 4 μV/°C across the temperature range. For the reference circuit, the maximum simulated difference between voltages was 320 μV at a nominal 800 mV supply. The reference consumes 11 nA at 25°C and 120 nA at 100°C, and the PTAT consumes 2.7 nA at 25°C and 34 nA at 100°C, all at 800 mV supply. Although this seems to offer little benefit over a resistor divider, certainly in terms of circuit performance, it occupies very little area and burns a tiny amount of power. To burn the same amount of power a resistor divider would be approximately 330 × 330 μm^2 with a total resistance of 300 MΩ (estimate made using the highest density poly resistor in 65nm).

The last major bias circuit that was designed on this chip was the current source for the radio's local oscillator. This circuit's design boiled down to a tradeoff between turn-on time (plus area, though area was consistently sacrificed to ensure healthy bias network performance) and noise. In theory, the majority of oscillator phase noise should come from the finite-Q load and not from supply or tail current noise. But, if the noise from these sources is large, it will negatively influence LO performance. To maintain consistent operating conditions in spite of any supply voltage or temperature variation, a constant g_m circuit was used. Similar to the fractional bandgap bias generator, the current source design essentially boiled down to a tradeoff between power consumption and area against noise and startup time. A filter was placed at the gate of the local oscillator current source. This 11 pF capacitor reduced the noise impact of the low-power constant g_m current reference, but occupied significant area and increased the startup/settling time of the LO, which is shown in Fig. 5.8. The current source could also be tuned with a resistor DAC.

Note here that the LDO itself turns on relatively quickly (less than 100 ns) but it takes upwards of 50 μs before the oscillator settles to its steady-state frequency. This is almost certainly caused by settling in the current source. As the bias changes, the parasitic capacitance
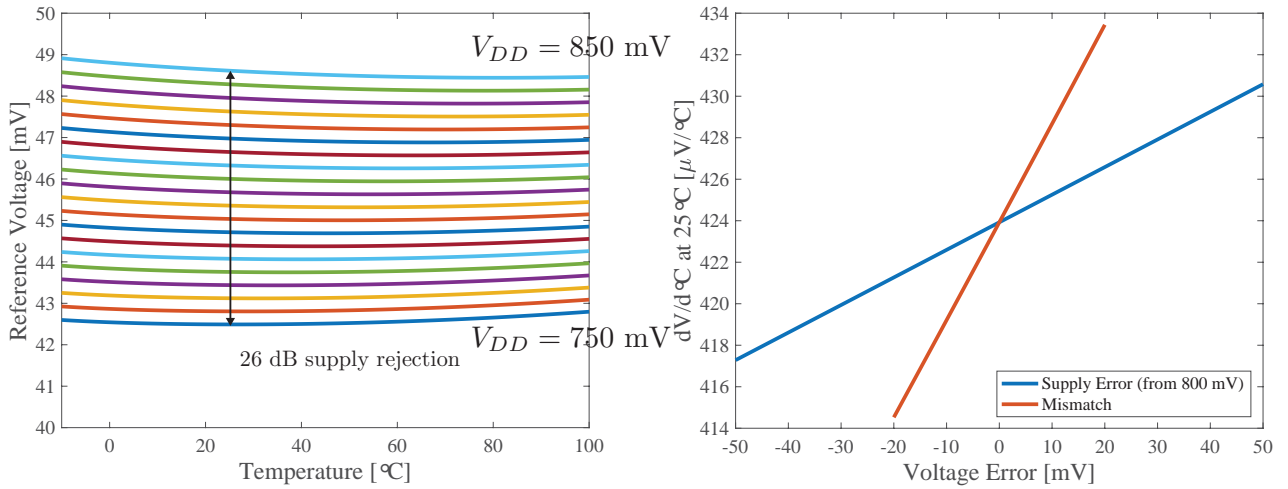
Figure 5.7: Simulated performance of 2T reference circuits with supply variation from 750 mV to 850 mV (in 5 mV steps) and with ±10 mV of threshold mismatch
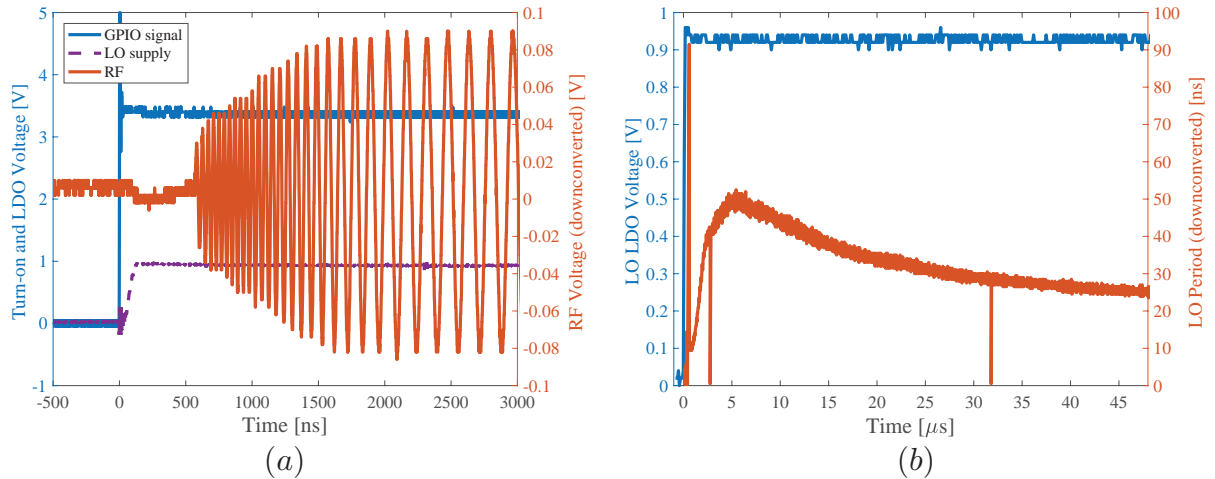
Figure 5.8: Local oscillator LDO and current source startup transients - the LO and PA are cold-started at $t = 0$ to 800 mV settings

of the cross-coupled transistors in the oscillator change. If it were a startup caused by the high Q and low loop gain of the LC tank, there would be amplitude settling in the envelope of the oscillation, and not in the frequency (it would look more like a second-order system's response to a step function). The measurement in Fig. 5.8 was made by downconverting the output of the LO (with power amplifier on) with an off-chip passive mixer. The RF port of the mixer was driven by an arbitrary RF signal generator emitting a +5 dBm tone 40 MHz

above the chip's LO frequency.

LDO design

Fundamentally, the requirements for the LDO supplying analog/narrowband RF loads versus the divider's digital switching load are different. For the analog blocks, it was critical to prevent any noise on the battery supply from reaching the rail of the sensitive loads. For the digital block, it is more important to prevent any switching transients from reaching the shared battery supply. This concept is presented in Fig. 5.3. In this schematic, the dominant poles of each of the three LDOs are marked with capacitors. Observe, in the die photo in Fig. 5.3, that some effort was made to keep the large switching transients from the divider physically distant from the sensitive LO electronics. In spite of this, there were still significant spurs.

The divider LDO primarily existed to isolate the battery voltage (and any sensitive circuits attached to said battery voltage) from digital transients. In addition, unlike constant-current analog loads, reducing the supply voltage of the divider does reduce power consumption (assuming that the majority of consumption comes from leakage and from switching loss, both of which are supply dependent). The divider LDO amplifier was a PMOS input folded cascode to achieve high loop gain at the expense of bandwidth. For this particular LDO, load noise is of no consequence because the load is digital. And any load noise (either deterministic switching noise or LDO supply noise) is either attenuated by the loop gain (within the bandwidth) or filtered. The only complicated tradeoff for this LDO was maintaining stability while having enough decoupling capacitance to prevent significant current spikes from browning out the supply (the amplifier pole is dominant, so the output/load pole location dictates stability margin). Any poles and zeros from the biasing network are outside of the unity gain bandwidth of the loop. The amplifier pole was made dominant with minimal area by attaching an explicit Miller capacitor between the gate and drain of the PMOS pass device.

The power amplifier LDO was more for output power tuning than anything else. Noise on the power amplifier's supply is irrelevant because both 802.15.4 and BLE use constant-envelope modulation. As long as there is no AM-FM distortion, noise on the PA's LDO is not particularly important. The PA's supply dependent parasitics (that cause LO load pulling) caused less than 10 Hz of LO frequency error for a 100 mV change in PA supply voltage. Because the PA's current transients happen at 2.4 GHz (and corresponding harmonics) it will be heavily attenuated even by minimal decoupling capacitance. Therefore, an output-pole dominant LDO with a simple NMOS common source amplifier with active load was used. The gain of the amplifier was low and the bandwidth was high so that less load capacitance needed to be used for stability.

The local oscillator LDO, however, had the added requirement of adding minimal noise so that it would not affect RF performance. As stated in the previous paragraph about the PA's LDO design, if only amplitude noise is considered, the LDO's noise doesn't matter. But, an oscillator is highly nonlinear and is, fundamentally, a DC to RF converter. So of

course any amplitude noise will cause some degree of frequency and phase noise. These mechanisms are described in a bit more detail in Chapter 2. As was the case for the PA, the load is, in principle, a static current, so there is no needed to "defend" the high supply from noise. Instead it is more important to prevent any incidental high supply noise from coupling onto the output. Therefore, an output pole dominant loop was used. The difficulty here was having low amplifier power consumption and low noise while having a higher amplifier pole bandwidth than load bandwidth. A simulated summary of loop stability of all three LDOs under nominal conditions is shown in Fig. 5.9. Simulated supply rejections are shown in Fig. 5.10. For the divider LDO's attenuation of load transients, a 20 Ω source resistance was used, and the dBV measured at the high supply is measured using a 1 mA small signal load perturbation. The small signal simulations match closely to the results from a transient simulation in which the divider load was a 1 mA, 100 ns long spike with 500 ns period.

Under actual operating conditions, the oscillator LDO will only be turned on during transmission or reception of packets, and the PA LDO will only be on during transmission. Power switching and LDO turn-on transients are shown in Fig. 5.11. So, the two circuits will be heavily duty cycled. This is good because the total energy consumption of the radio can be scaled back by reducing the TRX duty cycle. However, a negative consequence of using output-pole dominant LDOs for both of these circuits is that the capacitance needs to be charged up every time the circuit is turned on. This results in an estimated energy penalty of 184 pJ to receive and 332 pJ to transmit (assuming a regulated supply of 800 mV and minimal capacitive parasitics from everything besides the decoupling capacitance). The LO LDO amplifier burns 5.98 μA of current, the PA LDO amplifier burns 4.84 μA of current, and the divider LDO amplifier burns 760 nA of current. All of these numbers were simulated, and include biasing overhead for the LDO amplifiers.

5.3 Local Oscillator

A schematic of the local oscillator is shown in Fig. 5.12. A key result from the first iteration of the single chip mote was that a very low power and free-running oscillator could be used to synthesize a pure enough frequency to transmit and receive 802.15.4 and BLE packets. The local oscillator on the third iteration was a similar design (again, a CMOS class-B LC tank oscillator). There were a number of important differences. The thick top metals offered in TSMC's 65nm LPRF process resulted in a higher inductor Q factor (simulated Q of 18), which allowed considerably lower oscillator power consumption. The high LQ product of the inductor, however, required an extremely small capacitance change to directly modulate the frequency (1 MHz spacing for 802.15.4, and 62.5 kHz steps for BLE, more on the modulation later). Therefore, modulation was performed at the degenerated source of the NMOS cross-coupled pair [57]. This resulted in the modulation spacing having a somewhat awkward dependence both on the bias current (which changed the impedance of the current source as well as the g_m of the cross-coupled pair, both of which contribute to the change in frequency seen at the tank) and the fine tuning code (an effective change of
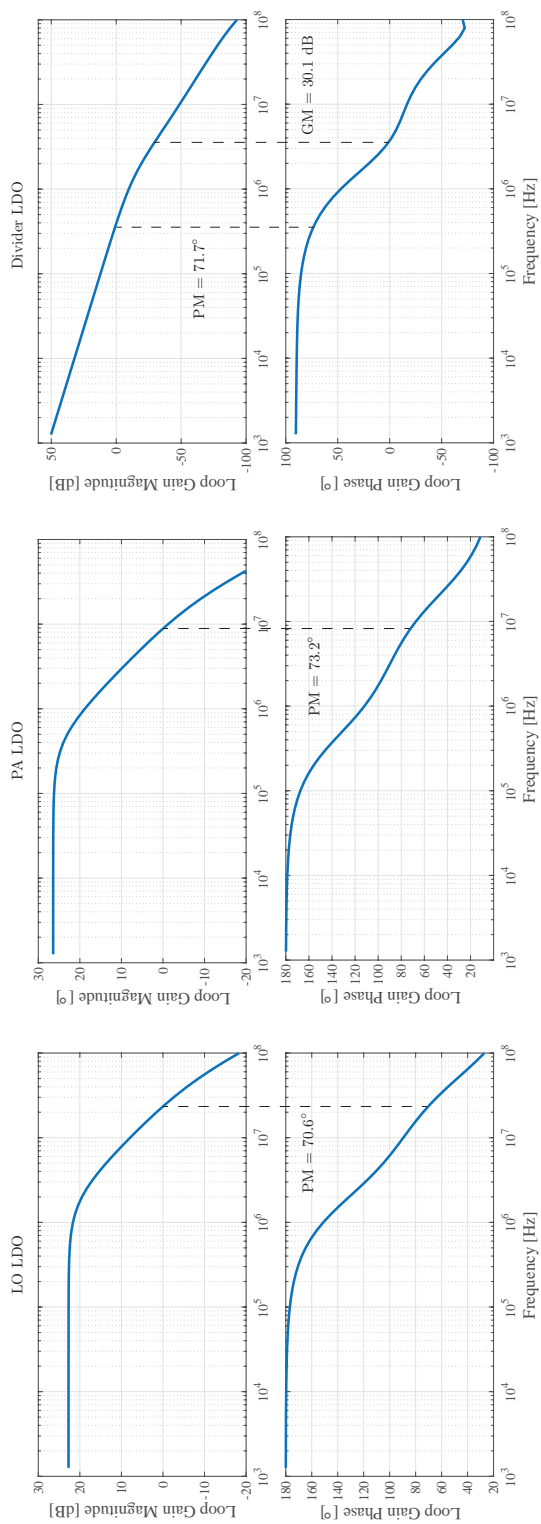
Figure 5.9: LDO stability summary
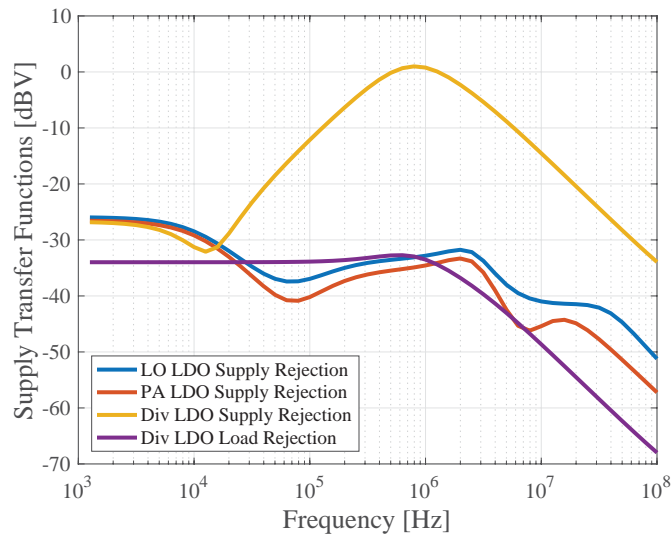
Figure 5.10: Simulated LDO supply and load rejections
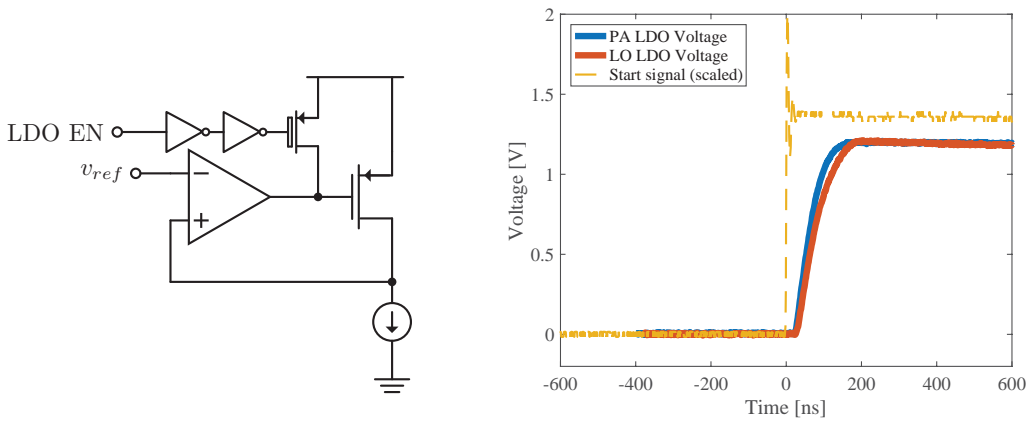
Figure 5.11: Measured LDO startup transients for LO and PA regulators. The LDO EN signal that pulls the LDO pass device gate to the source also disables the LDO amplifier's bias

$\Delta C/C$). There is no static non-tunable capacitance in the tank. There is, however, a fixed 13.7 fF capacitor in parallel with the fine DAC and modulation (at the source) to "bias" the capacitance in a region that is conducive for the desired tuning range and resolution and modulation spacing.

In addition to significantly reducing the power of the local oscillator, three additional improvements were made. First, a passive polyphase filter was added (at a relatively minor
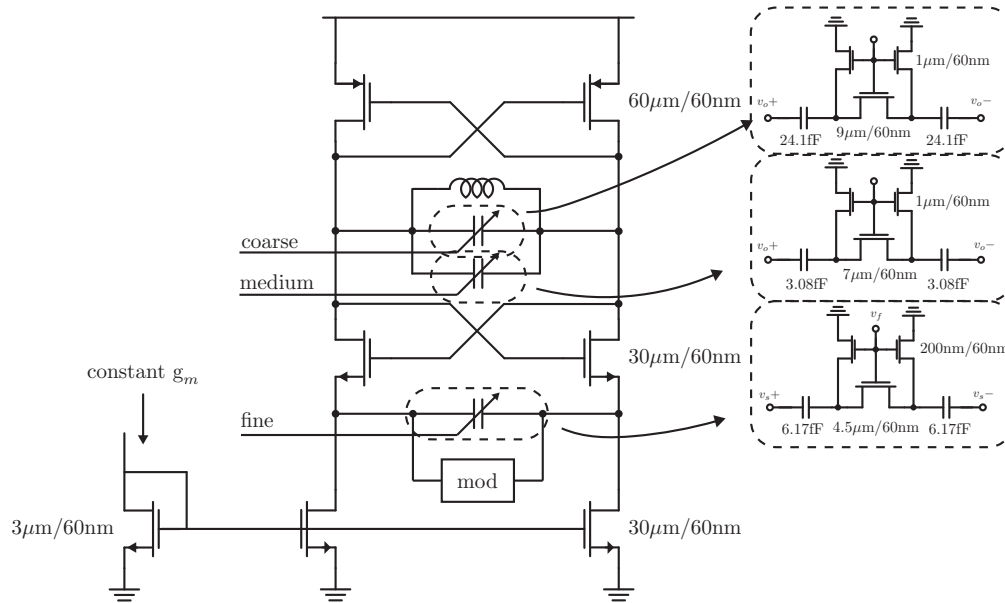
Figure 5.12: Local Oscillator Schematic

power penalty) to enable complex baseband processing. Second, the oscillator was tunable to within the 2.4 GHz ISM band, which immediately enabled communication to off-the-shelf hardware. And third, the three frequency tuning DACs (fine resolution with low range, middle resolution with middle range, and coarse resolution with large range) had overlapping dynamic ranges. This meant that there were genuinely no gaps in the tuning across the entire ISM band. More information about the tuning and eventual calibration (against chip-to-chip and temperature variation) is given later in this chapter.

The simulated inductor performance is shown in Fig. 5.13. At 2.4 GHz, the inductance is 7.4 nH with a Q of approximately 18, which results in a parallel impedance of 2 kΩ. In reality, the parallel impedance of the tank is halved, from approximately 2 kΩ to 1 kΩ by various loading effects (primarily the polyphase filter, when active, and finite Q of capacitive loading from the tuning DAC). However, having a large LQ product for the inductor with a safe SRF does simplify the design and allows the oscillator to have a high swing at low current levels.

Tuning and Modulation Tone Spacing

The downside of using a large inductance is that parasitic capacitance has a larger effect (in relative Δf). In addition, to achieve high tuning resolution, the effective change in capacitance must be small. The requirement for frequency resolution derived from the $\pm$ 40 ppm specificity dictated by both the 802.15.4 and Bluetooth standards is a tuning resolution of 190 kHz or less. This, in turn, with a 7.4 nH inductor, requires an effective change in
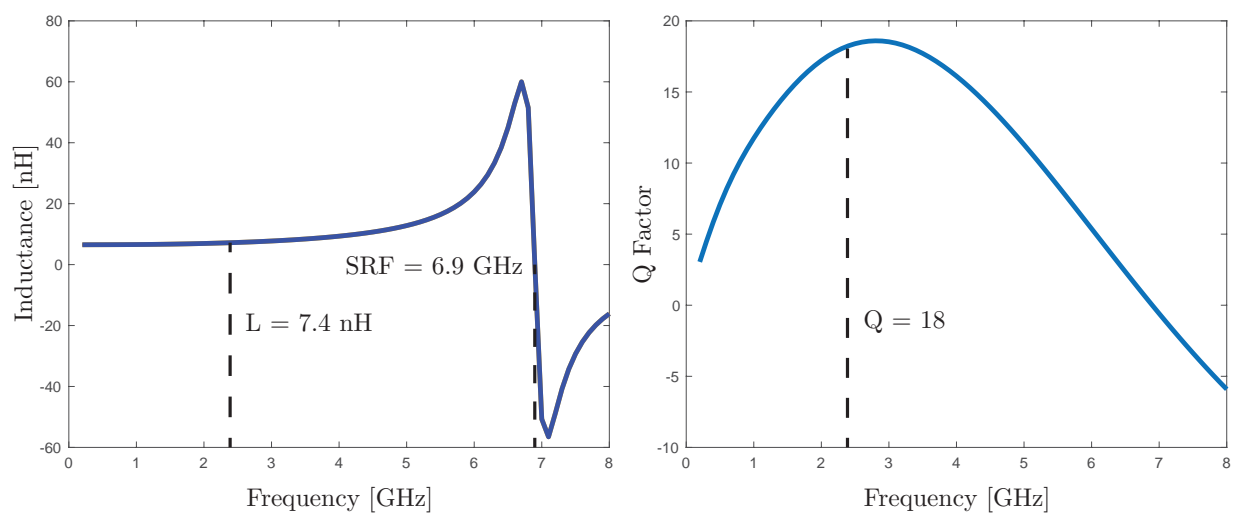
Figure 5.13: Simulated inductance and Q of LO inductor. Simulation was performed using Integrand EMX

capacitance of 9.4 aF. To achieve this tiny capacitance change, the same differential capacitive degeneration that was used on the single chip mote v1 was used once again. And, to maintain both range and resolution, a "middle" resolution/range capacitive DAC was added. In total, there were three separate, overlapping, 5-bit DACs.

In measurement, the oscillator was tunable between 2.1 GHz and 2.65 GHz (a relative tuning range of 23.1%) in 100 kHz increments (in the ISM band, 2.4 GHz - 2.485 GHz) without dithering and without a varactor. More details about the tuning will be presented in this chapter's section on calibration.

As was the case on the single chip mote v1, FSK modulation was performed by directly changing the frequency of the local oscillator at a specific datarate (1 Mbps for BLE, and 2 Mbps for 802.15.4). As was the case with the fine tuning, modulation was performed by varying the degenerated source capacitance. However, varying the tunable LC bias current will, in turn change the change in frequency. This happens because the effective change in capacitance at the drain varies with the transconductance of the cross-coupled pair. A plot of the nominally 1 MHz frequency spacing as a function of LO bias current is shown in Fig. 5.14. The modulation spacing was designed to be tunable to overcome this (as the current could vary somewhat from chip to chip, and it all typically comes down to mismatch in the LDO amplifier input pair). The modulation spacing was tunable by gating the modulation signal to a 3-bit capacitor DAC.
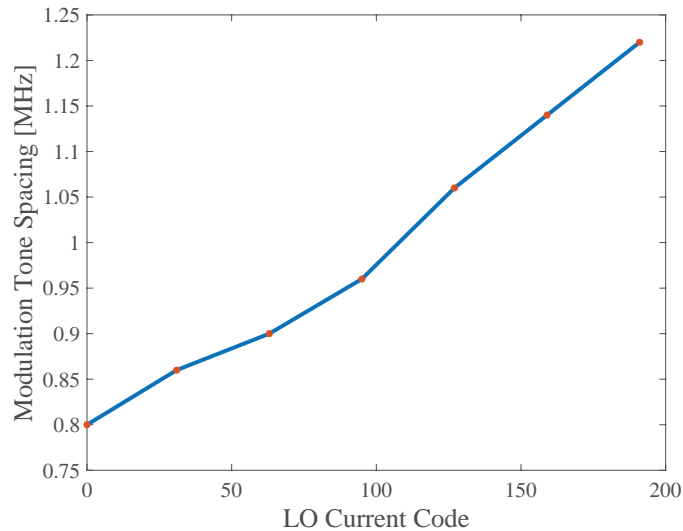
Figure 5.14: Tone spacing vs. current code for 802.15.4 modulation. The current varies from 170 μA to 850 μA over this range

Polyphase Filter

To generate in-phase and quadrature oscillation from the 2.4 GHz LC tank oscillator, a polyphase filter was used, as shown in Fig. 5.15(a). This feature was built into the receiver to allow for complex baseband processing and to improve the receiver noise figure by 3 dB by eliminating any noise or interference in the image band. In a perfect world, a differential polyphase filter would not load the oscillator at all. However, because of capacitive parasitics of the resistors and because, ultimately, the load of the polyphase filter is a biased mixer (primarily capacitive load) the tank, which is driving the filter directly, is loaded by the input impedance of the filter (approximately $R/\sqrt{2}$). To prevent this from effectively de-Qing the tank, the resistors should be made as large as possible. A large resistor, however, has significant capacitive parasitics (lower cut-off frequency) and the correspondingly small capacitor is susceptible to mismatch and requires careful layout. Ultimately, the resistors in the filter were chosen to be 6.4 kΩ and the capacitors were 10.18 fF - a corner frequency of nearly 2.44 GHz. The polyphase filter could be turned off in transmit mode with biased NMOS pass gates.

There is intrinsic systematic mismatch caused by asymmetries in the layout. However, the theoretical image rejection limit is approximately 23.7 dB [47] with 3° of phase error and 1 dB of amplitude error. This was the maximum measured error in the ISM band (phase error varied between +1° and -3°). Amplitude error can be tuned to within 1 dB by separately changing the gain of the I and Q baseband paths.
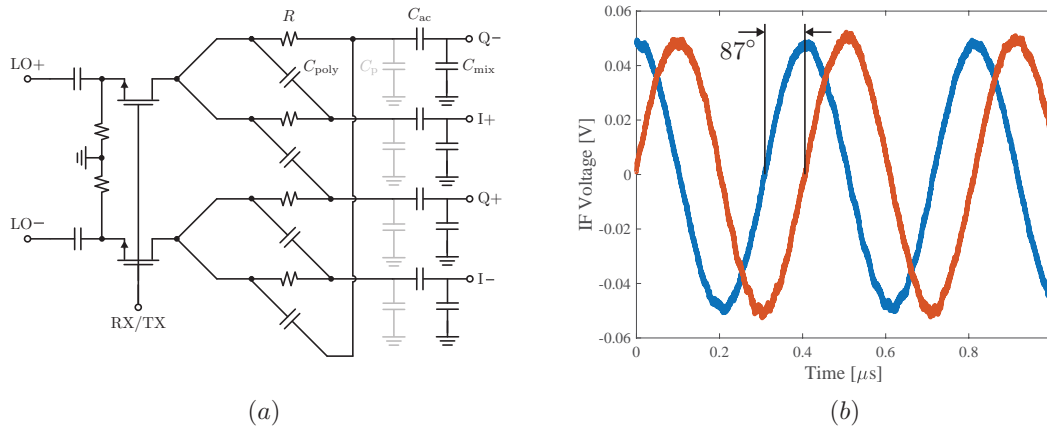
Figure 5.15: Polyphase filter schematic and measured downconverted phase error

Phase Noise

The phase noise of the oscillator in maximum FoM conditions is shown in Fig. 5.16. The maximum FoM was measured at an LC bias current of 170 μA. This plot uses exclusively on-chip bias generation and regulation. Because of the high degree of agreement between simulation and measurement at frequencies about 700 kHz, further simulations were used to justify the design decisions regarding the LDO. The discrepancies between measurement and simulation at below 700 kHz are caused by coupling from a low frequency oscillator running on chip, and what appears to be poorly modeled flicker noise in the process. The phase noise at 1 MHz spacing from the carrier is measured to be -111.9 dBc/Hz, which corresponds to a FoM of -188.6 dBc (just good enough to not be embarassing). The figure of merit deviates by less than 1 dB across the tuning range at this particular current level, but can be as poor as -183.2 dBc at the maximum current setting. The simulated phase noise at 1 MHz with both bandgap reference noise and LDO noise was -108.9 dBc/Hz. With no LDO noise whatsoever, it was -118.2 dBc/Hz at 1 MHz in simulation. The primary noise contributors were the input pair of the LDO amplifier and the LDO pass device. One could certainly argue that more power should have been burned in the LDO. However, this would make the loop more difficult to stabilize (with a dominant output pole) and the phase noise is certainly good enough anyways.

Notice in Fig. 5.16 that the measured data is slightly better than simulation. To speed up the simulation time, an estimate of the LDO decoupling capacitance was used that was likely pessimistic. In addition, there is some degree of filtering at the source node of the cross-coupled pair (the fine tuning and modulation capacitors) that were not included in the simulation. There is also a significant close-in spur in the measured phase noise centered at almost exactly 100 kHz. The source of this disturbance is still unknown.
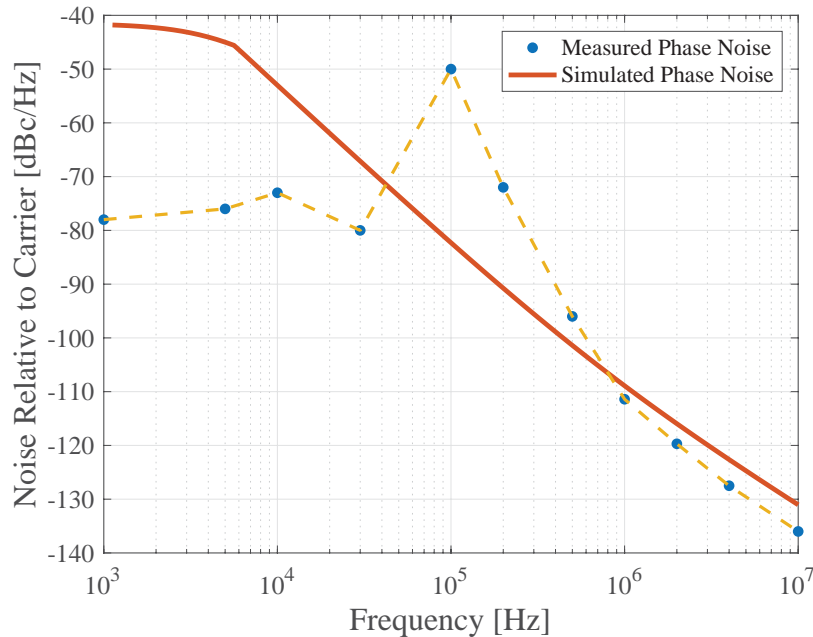
Figure 5.16: Measured and simulated LO phase noise (need to re-take this data)

5.4   Power Amplifier and RF Modulation

The direct modulation of the local oscillator (for FSK) simplifies power amplifier design. And, the constant-envelope modulation of both 802.15.4 and BLE allows for the use of efficient but nonlinear power amplifiers. The PA chosen was a class-D amplifier with no LO buffers. The PA's gate capacitance was treated as a part of the oscillator's tank. A dummy replica was used to approximately balance the differential loading on the oscillator (the replica is an identical copy of the PA, but with a floating output and grounded supply). The main challenges of the PA design were high system efficiency at low output power, and co-existence with the receiver. Transmitter duty cycles do tend to be fairly low, but maintaining low power consumption reduces supply droop in high-ESR batteries, and relaxes energy requirements if the chip is powered by energy harvesting. The targeted power consumption of the PA and LO was 1 mW, and the goal was to have as high an efficiency as possible.

A detailed schematic of the transmitter, mixer, and front-end is shown in Fig. 5.17. The PA was disabled by turning off its LDO. Because the receiver uses a low-IF architecture with AC coupling, any LO leakage through the PA has no impact on receiver performance.

Matching Network

Note in Fig. 5.17 that the transmitter and receiver share a single-ended antenna port. Although receiver requirements are generally lax (-70 dBm sensitivity for BLE, for instance)
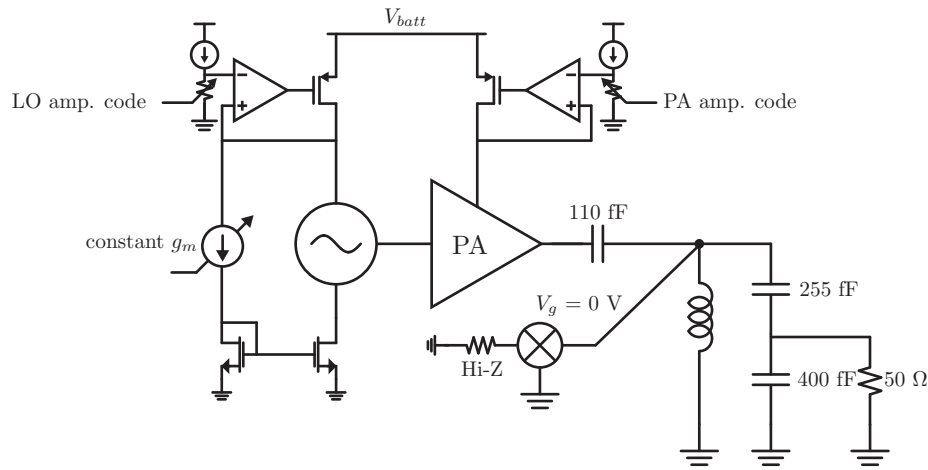
Figure 5.17: Radio Front-End Schematic

the vast majority of commercial and academic BLE and 802.15.4 transceivers have receiver sensitivities at or below -85 dBm (-85 dBm is the requirement for 802.15.4 sensitivity). The matching network was designed primarily with the receiver performance in mind, and it boosts the source impedance with an equivalent voltage step-up ratio of approximately 4.5. Boosting the source impedance and voltage is beneficial for receiver noise figure.

One negative consequence of this particular matching network is that the inductor is a DC path to ground. This particular L-match configuration was selected so that, in receive, the mixer would be AC coupled from the antenna (and, more realistically, from any potentially DC coupled test equipment). The DC path to ground meant that the output of the PA needed to be AC coupled from the top of the inductor. This capacitor is not part of the match. Mostly, all this does is shift the phase a bit, which is irrelevant. However, bottom plate parasitics of this capacitor are not resonated and cause a loss of efficiency. It is possible to have a switch in series with the inductor. However, in this particular process, the switch needed to be gargantuan (>1 mm width, minimum length) to not incur a significant Q penalty (and therefore add considerable match insertion loss). The AC coupling and high swing of the PA mean that the input to the mixer does swing below ground, and can turn on mixer switches (gates biased to ground). No current can flow from source to drain because the IF stages present a high impedance. However, current can flow through the forward biased source-body diode in the NMOS mixer switches. To reduce this current, the mixer switches were placed in a well and the bodies were connected with large resistors to ground.

A second negative consequence is that drain capacitance of the PA is not a part of the matching network, and any hard-switching PA will incur significant CV^2 power loss. This leads to an uncomfortable tradeoff between switch size to balance this switching loss against loss caused by on-resistance of smaller switches. There is likely a more clever way to share the

matching network between receiver and transmitter. However, this particular configuration had been used successfully with the receiver on SCM v1.

Power Amplifier

Because linearity is irrelevant to the modulation, a switching PA was a logical choice. In addition, because of the low power consumption (and therefore low output power) harmonic distortion is less of a concern. The FCC mandates a spurious harmonic emission of less than -41 dBm. Because harmonics were not particularly important, a class D PA was a natural choice of topology. A clever possible alternative would have been to use a PMOS class E PA, although this would have been fairly challenging to co-integrate with the mixer devices as the shared node would swing well below ground (more than the class D). In addition, the local oscillator swings mid-rail, which is quite convenient for driving a CMOS inverter. It is not convenient for driving a PMOS device (which requires gate biasing) and an additional shunt capacitor is needed as part of the match. A schematic of the class D power amplifier and associated simulated transient waveforms under nominal conditions are shown in Fig. 5.18.
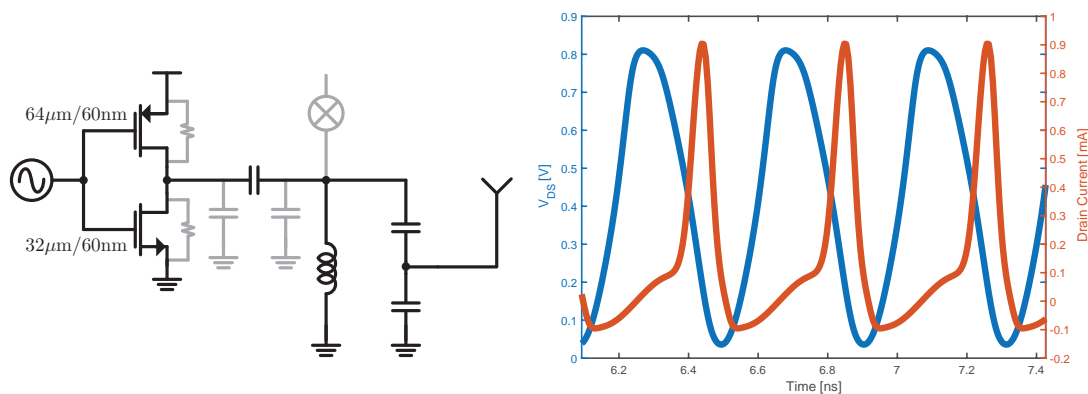
Figure 5.18: PA schematic (sources of efficiency loss shown in grey) and simulated transient drain voltage and drain current of the NMOS device

Efficiency

Two parameters could be tweaked to change output power and efficiency: the input power (dictated by LO current, and therefore swing) and the regulated supply of the power amplifier. Plots of output power, PA drain efficiency, and system efficiency are shown in Fig. 5.19. The absolute maximum output power is approximately -7.5 dBm, but is not shown in Fig. 5.19. At higher LO swing, the PA drain efficiency is increased. This makes sense because the PA gates are being driven harder. However, this comes at the price of increased

LO power. An optimal system efficiency of approximately 15% comes at an output power of -9 dBm with a DC PA drain current of 504 μA and an LO current of 340 μA. Fig. 5.19 (a) shows output power with varying PA supply voltage, (b) shows drain efficiency with varying PA supply voltage, and (c) shows system efficiency (LO + PA power) with varying PA supply voltage. In each plot, the LO supply is fixed at 800 mV and each individual trace represents a different LO current level. All of these efficiencies are calculated from the regulated supply voltages of the LO and the PA. If taken from a 1.5 V supply, the maximum system efficiency drops to 11% again at -9 dBm output power.
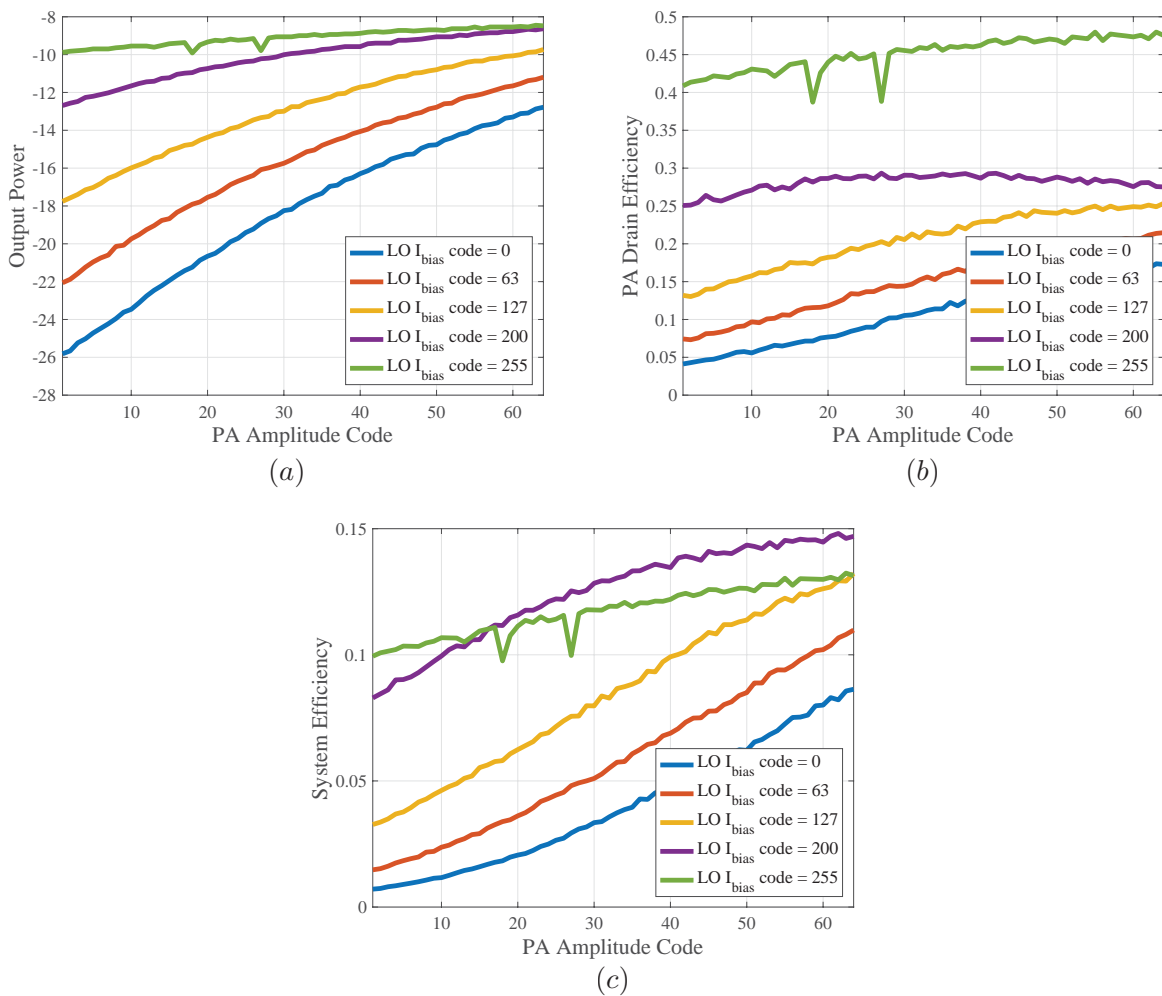
Figure 5.19: Measured PA output power, PA drain efficiency, and TX system efficiency

A consequence of changing the LO current was that bias point changed along with it. There is no gate bias at the input of the PA. If the DC level of the oscillation is not at the optimal bias of the PA, it reduces output power, efficiency, and causes significant second

harmonic distortion because the NMOS and PMOS are no longer driving with equal strength. Plots of second and third order harmonic distortion are shown in Fig. 5.20. The primary cause of increased HD2 at low LO current is the bias level.

(a)   (b)
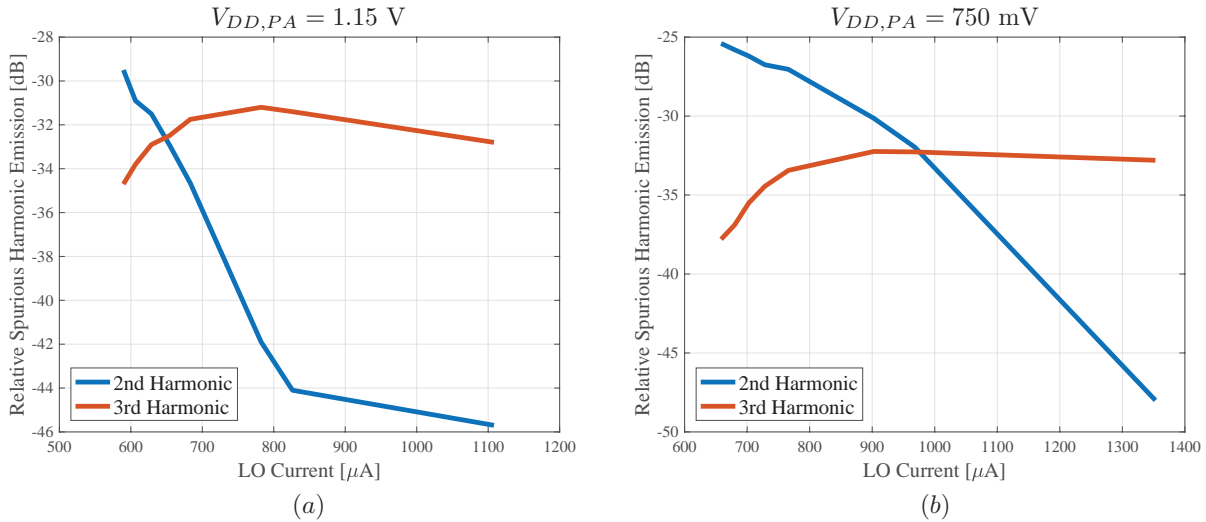
Figure 5.20: Measured 2nd and 3rd order harmonic distortion at varying LO swing levels at (a) 750 mV LO supply and (b) 900 mV LO supply

These two plots in 5.20 are primarily for illustrative purposes. The power amplifier rail was set to its maximum regulated value (1.15 V) and the local oscillator was set to 0.75 V in Fig. 5.20 (a) and 0.85 V in Fig. 5.20 (b) so that the oscillator is intentionally biased poorly. As the swing and the PMOS device is driven harder, second order harmonic distortion is reduced as the duty cycle at the output of the PA approaches 50%.
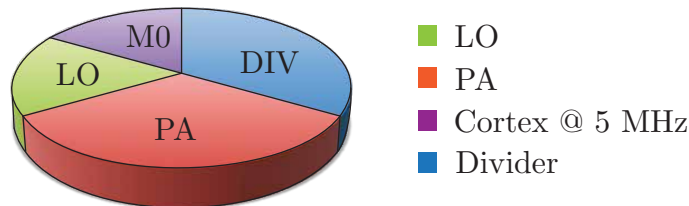
Figure 5.21: Measured power consumption

A performance comparison against similar previously published BLE and 802.15.4 transceivers is shown in Table 5.1. An emphasis is placed on the transmitters. Very few designs exist

at this power level. Takeaways from this brief literature survey are: all efficiencies at this power level are quite similar, efficiency is dictated primarily by supply voltage and matching network quality. The standard topologies used here and in all of the cited papers are essentially equivalent in terms of performance. Unless someone comes up with a really clever idea, any improvements in efficiency at these power levels will continue to be marginal. It was mentioned earlier that the match was optimized for RX. The only downsides were the AC coupling capacitor, which could not be made a part of the L-match, and any current loss in the off-state mixer switches. Assuming the same match network Q, the presence of the receiver has minimal impact on transmitter efficiency. However, a class-E PA could be used, which would result in both higher output power and higher efficiency (likely similar to [26]). In Table 5.1, [61] never actually mentions their system efficiency, the 20% efficiency at -10 dBm assumes that the local oscillator burns approximately 400 μA (the number is not explictly stated, it is estimated from the summary table). The 37.5 MHz reference is not included in this calculation.

5.5   Divider

In this particular system, an RF divider is not strictly necessary. There is no PLL or FLL, and even if there were, there is no reference to compare against. An integer divider was still included for system flexibility. It was used to generate high-precision baseband clocks, such as the clock that drives the GFSK module ($\approx$ 20 MHz), the clocked $g_m - C$ switched capacitor baseband amplifiers ($\approx$ 64 MHz), and the baseband chipping clocks (exactly 1 MHz for BLE and exactly 2 MHz for 802.15.4). The LC tank oscillator is the highest quality (in terms of jitter) frequency synthesis on chip (and could be comparable in SSB phase noise performance to a crystal oscillator as long as there is no temperature or supply variation, albeit at substantially higher power consumption - [72] has -125.1 dBc/Hz at 1 MHz spacing from a carrier of 4.51 GHz, and the phase noise from [73] is -155 dBc/Hz, worst case, at 1 MHz spacing and a carrier of 25 MHz). One specific application of the divider that drove the design was use as a baseband clock. Then, during modulation, the LC tank can essentially clock itself. There is one caveat - the divide ratio needs to change when transmitting at different channels. To maintain compatibility with both BLE and 802.15.4, this requires divide ratios shown in Fig. 5.22. The channel spacing for 802.15.4 is 5 MHz, so, to generate a 1 MHz clock (which is multiplied by 2 to generate the 2 MHz 802.15.4 chipping clock) the divider ratios must be multiples of 5. And, the channel spacing for BLE is 2 MHz. To generate the 1 MHz chipping clock the divide ratios must be multiples of 2.

In addition, running a divide-by-N at 2.4 GHz is power intensive. A pre-scaler was put before the large counter-based programmable divider to reduce overall power consumption (at a minor jitter penalty). BLE and 802.15.4 require different pre-scaler divide ratios, so selectability of pre-scaler divide ratio was built in. The full divider block diagram is shown

	This Work	[55] 2015	[56] 2015	[25] 2015	[26] 2016	[69] 2018	[70] 2018	[71] 2018	[61] 2018
Technology	65 nm	55 nm	40 nm	40 nm	28 nm	65 nm	40 nm	28 nm	40 nm
Supply Voltage	0.8 V/1 V	1.2 V	1.1 V	1 V	1 V	1 V	0.8 V	0.2 V	0.6 V
On-Chip Match	Yes	Yes	Yes	Yes/No	Yes	Yes	Yes	Yes	No
P_{out}	-8.8 dBm	2.3 dBm	0 dBm	-2 dBm/0 dBm	0 dBm	-3 dBm	1.8 dBm	0 dBm	-10 dBm
PA Efficiency	27%	50 %			41%	25%	30%	30.6%	41 %
Sys. Efficiency	14.7%	10 %	13%	14.3%/22.7%	27.8%	16.1%	24.8%	25%	20 %
Area	0.96 mm²	0.64 mm²	1.1 mm²	1.3 mm²	0.65 mm²	1.64 mm²	0.8 mm²	0.53 mm²	0.0166 mm²
Standards	BLE TX, 15.4 TRX	BLE TRX	BLE TRX	BLE/15.4 TRX	BLE TX	BLE TRX	BT5/BLE TRX	BLE TRX	BLE TX

Table 5.1: Performance summary of 802.15.4 and BLE transceivers and transmitters operating in the sub-10 mW regime
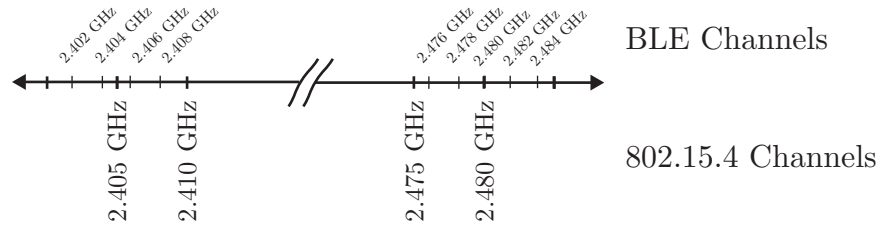
Figure 5.22: BLE and 802.15.4 channel frequencies

in Fig. 5.23. This includes the static blocks that generate the $\approx$20 MHz clock for the GFSK module and the $\approx$64 MHz clock for baseband filters. Because of the frequency-doubling circuit implementation, the jitter of the 2 MHz clock is higher (the output of the delay line has large rise time, so amplitude noise from the XOR turns into phase noise).
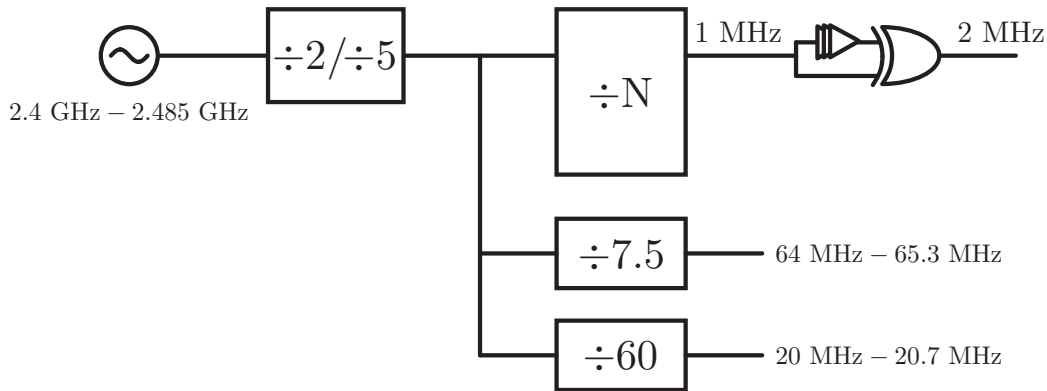
Figure 5.23: Top-level divider block diagram

One negative consequence of using a divider is that it generates LO spurs. Part of the goal of the LDO network was to mitigate these spurs as much as possible. The spurs when generating both 1 MHz and 2 MHz divide products are shown in Fig. 5.24. One interesting thing to note is that the largest spur in each case does not come from the output frequency of the divider, but at two times further (i.e. 2 MHz for a 1 MHz divider output, and 4 MHz for a 2 MHz divider output). This happens because current spikes from the supply will come at every edge transition of the divider's switching.
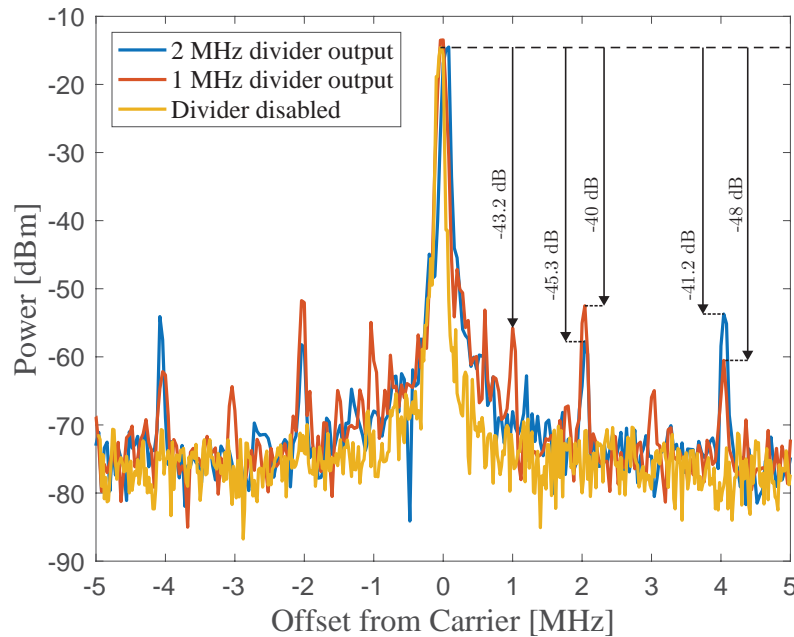
Figure 5.24: Measured output spectra with dividers running

Tunable Dynamic Pre-Scaler

One mildly interesting circuit used in the divider was a dynamic injection-locked pre-scaler. The block diagram illustration is in Fig. 5.25 (c) and the transistor-level schematic is in (d). The first stage of the division is the most power hungry because switching power scales linearly with frequency. Therefore, any potential power savings in the pre-scaler can result in a significant reduction in overall divider power.

The pre-scaler divide ratio was tunable on-chip between 2 (measured power of 224 μW) and 5 (measured power of 108 μW). For the divide-by-5 case, this is significantly less than their standard CMOS pre-scaler counterparts (231 μW at $\div 2$, Fig. 5.25(a), 175 μW at $\div 5$, Fig. 5.25(b)). Note that in the measurement setup it is not possible to isolate the pre-scalers themselves, so these numbers include the power of the pre-scaler output buffers (which is quite significant - see Fig. 5.3 - the pre-scaler is next to the oscillator, whereas the programmable counter-based integer divider is approximately 380 μm away). All of the pre-scalers were in parallel, and were enabled/disabled by power gating. This does not influence LO power because they present a small capacitive load that is absorbed as part of the tank capacitance.

The jitter was nearly equal in the $\div 2$ version but, unsurprisingly, because the last stage of the division in the dynamic pre-scaler is clocked by the RF oscillator, the jitter in the dynamic $\div 5$ divider is superior to its static counterpart by a factor of approximately 2.

In Fig. 5.25 (c) the dynamic pre-scaler is essentially a current-starved injection-locked
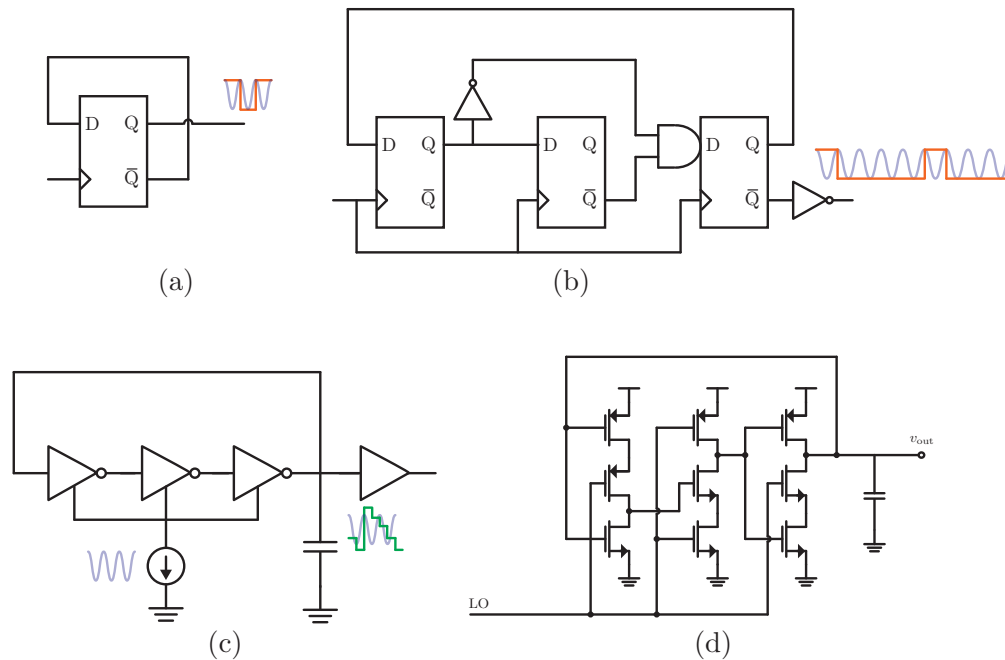
Figure 5.25: Divider pre-scaler schematics. (a) Shows a static divide-by-2 circuit using a flip flop. (b) Shows a static divide-by-5 circuit using flip flops and combinational logic. (c) Shows an intentionally slowed injection-locked divider. The tail current dictates how much charge is removed from the output capacitor, which sets the frequency divide ratio. It works a bit like a resetting charge-domain counter. (d) Shows the transistor-level implementation of the "charge counter" from (c)

ring oscillator. It is implemented with a TSPC flip flop (Fig. 5.25(d))that is sized to operate in a similar fashion to a ring oscillator, but with additional devices that reduce dynamic power consumption. In Fig. 5.26 the divide ratio was set by the divider input power. This circuit offers two advantages over traditional divider topologies. The first is that it can operate at input levels significantly lower than CMOS rail-to-rail swings. It is also relatively widely tunable and, unlike a counter-based integer divider, requires fewer gates and therefore lower CV^2 power consumption. It does not automatically generate 50% duty-cycle waveforms.

One important caveat: the $\div$ 5 in particular requires both a tuning of the width of the divider's device, and a tuning of the LO current. So realistically this pre-scaler, as taped-out on this chip, is not particularly useful. However, the circuit and design principles could be adopted in a more practical way (perhaps with tuning of the load capacitor, or with a tunable-supply buffer) to create a widely tunable and reliable integer or fractional pre-scaler. By varying the local oscillator's supply and current, it is possible to change the pre-scaler's divide ratio to every integer between two and seven inclusive. The original goal was to have a single pre-scaler and tune the strength of the pull-down devices whose gates
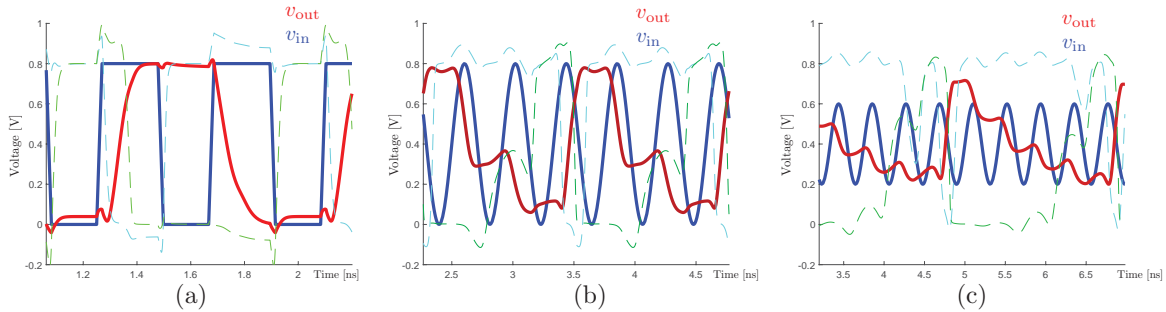
Figure 5.26: Simulated output voltage of the dynamic pre-scaler at different input power levels

are connected to the LO (in Fig. 5.25(d) that essentially looks like a resistor DAC. However, the implementation of the DAC signifcnatly loaded the LO, and the impending tapeout deadline necessitated its removal.

5.6   Local Oscillator Calibration

So far, every circuit presented is not particularly innovative nor cutting-edge, in terms of performance. The real challenge lies in the lack of absolute frequency reference. In this section I will discuss open loop calibration results for the on-chip RF local oscillator as well as propose a number of potential network-level calibration techniques.

Tuning

The fundamental challenge to make a simultaneously high resolution, large tuning range DAC is from capacitive and resistive parasitics. In the case of this particular oscillator, another difficulty arises from the high inductance of 7.4 nH (simulated with a 2.5D EM simulator). To meet the requirements of either 802.15.4 or BLE, the channel accuracy specification requires a tuning resolution of approximately 180 kHz (plus or minus 40ppm at 2.4 GHz). Which, in the worst-case scenario (at the high-end of the 2.4 GHz ISM band) requires a capacitive LSB of 80 aF, which is quite challenging to achieve.

A high tuning range is required so that the oscillator can tune to the 2.4 GHz ISM band across process, voltage, and temperature (mostly process variation, as it can result in as much as $\pm$ 15% variation in capacitance (based on an approximation of 3σ variation), or approximately $\pm$ 200 MHz of frequency variation with a 7.4 nH inductor).

To simplify the DAC design, three separate, overlapping, 5-bit DACs were used (called fine, mid, and coarse, referring to the varying resolutions of each individual DACs). The fine DAC has approximately 100 kHz (16 aF) resolution, the mid DAC has approximately 700 kHz (0.11 fF) resolution, and the coarse DAC has approximately 11 MHz (1.73 fF) resolution.

The reason that three separate DACs are used, as opposed to the more traditional two, is that the fine DAC is implemented with a degenerated capacitor as described in [57] and analyzed in Chapter 2 of this dissertation. A plot of the overlapping frequency tuning characteristic is shown in Fig. 5.27. Note that in this figure, only coarse codes relevant to the ISM band were shown. In addition, because the drop in frequency with mid code overlap was so large, it was mitigated slightly for the sake of illustration. A full sweep of all 2^{15} codes took a prohibitively long time to measure.
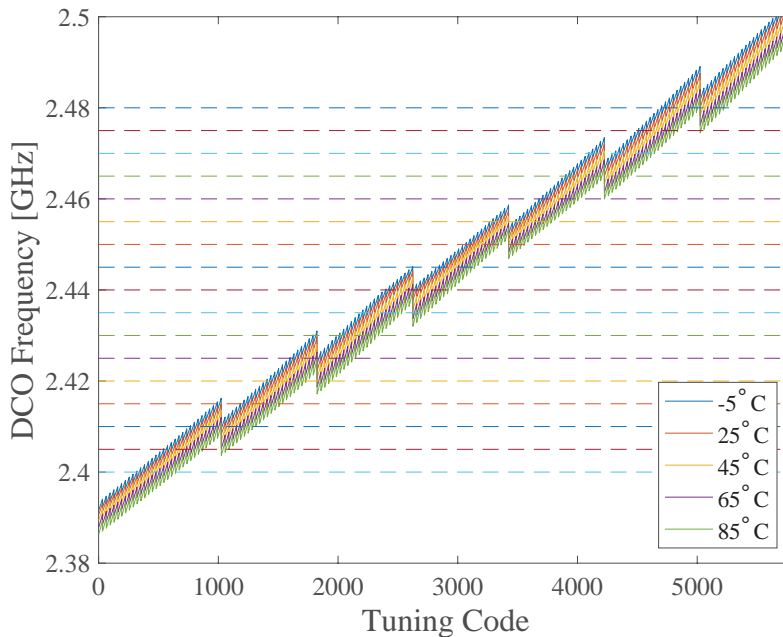
Figure 5.27: Overlapping tuning characteristic across the industrial temperature range. Horizontal lines indicate 802.15.4 channel frequencies

Monotonic Tuning Characteristic

A monotonic frequency tuning characteristic is helpful for two reasons. First, if the oscillator is ever put in a PLL, a non-monotonic characteristic makes control challenging. Second, having monotonic codes significantly simplifies any oscillator calibration. And, if the resulting tuning characteristic is sufficiently linear, it is possible to perform a simple two-point calibration against an absolute frequency reference. Furthermore, assuming a linear variation with temperature, this calibration only needs to be performed once in the chip's lifetime (assuming that the power in the oscillator, specifically, is low enough that aging effects are not an issue). Then each other channel is simply a linear extrapolation of the now-monotonic tuning characteristic.

The most obvious strategy is to simply measure every code in or near the ISM band and post-process the information to generate the monotonic codes. However, given the resolution of the fine DAC this would require an absolute minimum of 850 codes, one for each 100 kHz LSB in the 85 MHz ISM band (and that is assuming that the algorithm magically guesses codes perfectly). And, once again because of the high resolution, to measure a frequency difference of 100 kHz requires an accumulating counter to run for at least 10 μs. Realistically, because of the divide ratio and because of the need to average out any jitter in either the LC tank oscillator or in the counters reset timer, this number could be as high as 100 ms. And, as a consequence, a full sweep of the DAC codes could take upwards of multiple minutes, which, in circuits terms, is an eternity. Depending on application, the power penalty could also be fairly significant; both the local oscillator and divider consume non-negligible amounts of power.

A simple alternative is to use the following algorithm to break the monotonic code n into three 5-bit codes representing the values of the fine (f), mid (m), and coarse (c) settings.

```
c = n/coarse_divs + 19;
n = n % coarse_divs;
m = n / mid_divs;
f = n % mid_divs;
```

The variables coarse_divs and mid_divs represent the number of fine codes in a single coarse DAC step, and the number of fine codes in a fine DAC step, respectively. The 19 is simply a scalar addition to get close to the low-end of the ISM band. Variation from PVT in the lot of received chips (possibly all from the same wafer) is less than a coarse code, so the 19 does not need to be changed from chip to chip. Two issues arose. First, the fine and mid codes did not scale in the same way with increasing frequency. This is expected, and can be summarized by the two equations that dictate how a small change in capacitance at the source (Eq. 5.1) vs. at the drain (Eq. 5.2) changes frequency as a function of the starting frequency.

$$f_{osc} = f_0 \sqrt{1 + \frac{\bar{g_m}^2 L}{4C_{fine}}} \tag{5.1}$$

$$f_{osc} = f_0 \sqrt{\frac{C_d}{C_d + C_{coarse}}} \tag{5.2}$$

The slopes of these two tuning characteristics vary across the ISM band. This was solved by modifying the algorithm slightly to have three separate modulo functions. In essence, the coarse_divs and mid_divs from the pseudocode were changed across the ISM band. This resulted in monotonic tuning characteristics with minimal DNL for all chips tested. The numbers corresponding to number of LSBs in a mid code and number of LSBs in a coarse code do need to be tweaked for each individual chip, but frequency measurements only need

to be performed at the codes where the DACs overlap, and it can be performed with a low-accuracy oscillator. The process could easily be automated for each individual chip.

Process, Voltage, and Temperature

This is not intended to be a statistical analysis, only a series of observations regarding a small number of chips from (likely) the same lot. Any plots are for illustrative purposes only, and are not made based on either measured variation or on data from the fab.

Capacitors themselves, particularly larger ones, have little variation on the same die. However, any capacitive parasitics of transistors can exhibit comparably significant variation. In addition, because transistor thresholds vary significantly in deep submicron processes, bias voltages and currents will be unique on every chip. Generally speaking, an LC tank oscillators frequency will have some dependence on small variations in bias current or supply variation. However, the degenerated capacitor tuning used to implement the extremely fine frequency resolution is highly dependent on the transconductance of the cross-coupled NMOS devices, which will vary with the randomly changing bias.

Another somewhat unexpected issue arose regarding frequency pulling caused by various circuits used in receive or transmit mode. The pulling itself is not unexpected. Recall that every circuit that the oscillator drives is integrated directly into the tank capacitance. The part that was somewhat unexpected was that it appeared that the receive (polyphase filter active) and transmit (power amplifier active) were frequency and, potentially, swing, dependent. In transmit mode, because the PA itself is an inverter, the Miller effect will present an input-swing dependent capacitance to the oscillator. And, in receive mode, the AC coupled NMOS switch to enable and disable the polyphase filter can exhibit awkward behavior at high swing. And of course, at the same current setting, the tank impedance will be higher at higher frequency, so the swing of the oscillators voltage will be greater. This effect is measured in Fig. 5.28.

Two-point Calibration

The simplest strategy is to perform a one-time, two-point calibration to determine an individual oscillator's 2.4 GHz code and the oscillator's slope across the ISM band. Because of the somewhat unpredictable pulling caused by the mixer loading and the PA loading, this calibration would need to be performed in both transmit mode and receive mode. All measurements here were performed by measuring the output frequency of the divider. Frequency pulling from the divider itself is less than a fine LSB, and can be ignored. This is because the divider input gates were small (nearly minimum size). In addition, the divider pre-scaler can be power gated without changing the bias of the LO buffers with less than 5 μA of current penalty. The results of using a "perfect" N-point calibration are compared to a two-point linear calibration (red) and a three-point quadratic calibration (green) in Fig. 5.29
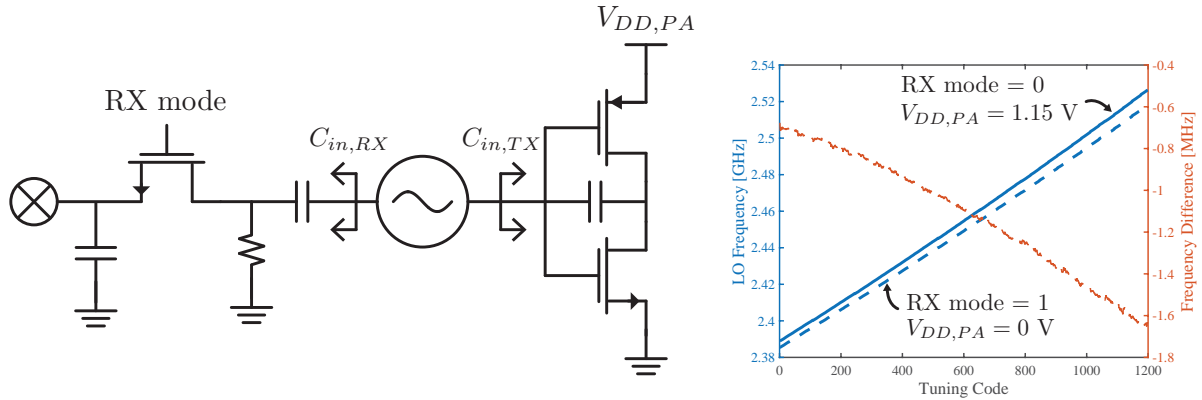
Figure 5.28: Effects of receiver and transmitter load pulling on TX and RX DCO tuning characteristics. These are static frequency offsets caused by nonlinear amplitude-dependent capacitances. Large RF signals at the antenna port cause injection pulling (at around -30 dBm) and injection locking (at around -20 dBm)

In Fig. 5.29, to generate the "perfect" blue calibration, frequency at every code was measured, and for each channel, the code corresponding to the lowest frequency error was chosen. To generate the linear fit, the frequency of the oscillator was measured at two codes (in this experiment, codes 400 and 1400) and from those points a linear code vs. channel function was generated. Similarly, for the quadratic fit, the frequency was measured at codes 400, 900, and 1400, and a quadratic code vs. channel function was generated. These procedures were performed for three different chips, both in transmit and receive. The points represent the mean of all six cases, and the error bars represent the standard deviation.

Temperature Calibration

At a given temperature, it is possible to get relatively close to a given channel (within 500 kHz) with a two-point calibration and linear extrapolation. The caveat is, once temperature changes, the calibration falls apart. Fig. 5.30 shows the local oscillator frequency variation under typical lab bench conditions (measurements taken ever 100 ms) and subject to varying temperature in a temperature chamber. The frequency varies with a slope of -100 kHz/°C (40 ppm/°C).

One way to correct for this is to make a measurement of temperature on chip, and adjust the oscillator's frequency accordingly. A technique inspired by [74] was used to make an estimate of the chip's temperature. Two oscillators, one at 2 MHz and the other at 32 kHz [75] were measured. The ratio of one count to the other is approximately linear with temperature between 0°C and 85°C, as shown in Fig. 5.31.

The plot in Fig. 5.32 of calibrated frequency versus temperature is more of a gold standard than a practical result. To generate it, the oscillator was first subjected to temperatures
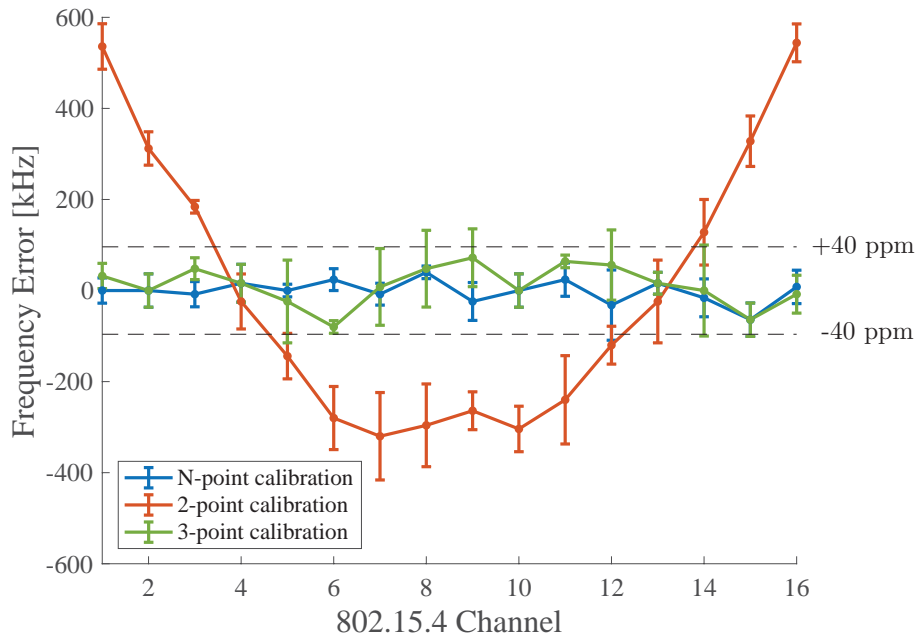
Figure 5.29: Frequency error at 802.15.4 channels with various calibration techniques

between -5°C and 90°C, and was placed in a feedback loop with a perfect frequency reference. The codes required for the oscillator to run at 2.44 GHz were stored in a look-up table at every temperature. Then, in a second temperature sweep, rather than using a perfect reference, an estimate of temperature was made with the linear fit of the ratio of the 2 MHz and 32 kHz oscillators in Fig. 5.31(b). That temperature estimate was subsequently used to determine the oscillator's frequency setting (from the look-up table). Basically, it is a massive N-point calibration. This was also not performed with the counters available on chip, but, rather, with an off-chip frequency counter. Feedback was then implemented through a MATLAB program that programmed the scan chain.

Calibration in an 802.15.4 Network

Thus far, all strategies for frequency selectivity require a large amount of calibration. This is not intrinsically negative, but in commercial settings, any calibration against temperature requires significant overhead that can increase the per-die price of the chip. In the end, it could require an LC code for every channel at every temperature. Assuming that there are 11 bits of tuning needed in the ISM band (with some extra codes) and assuming that a new set of codes needs to be stored for every 1°C, a calibration look-up table would occupy 3.74 kB of memory for 802.15.4 and an additional 9.35 kB for BLE. This assumes a temperature range of 0°C to 85°C, and codes both for transmit and receive. This iteration of the mote
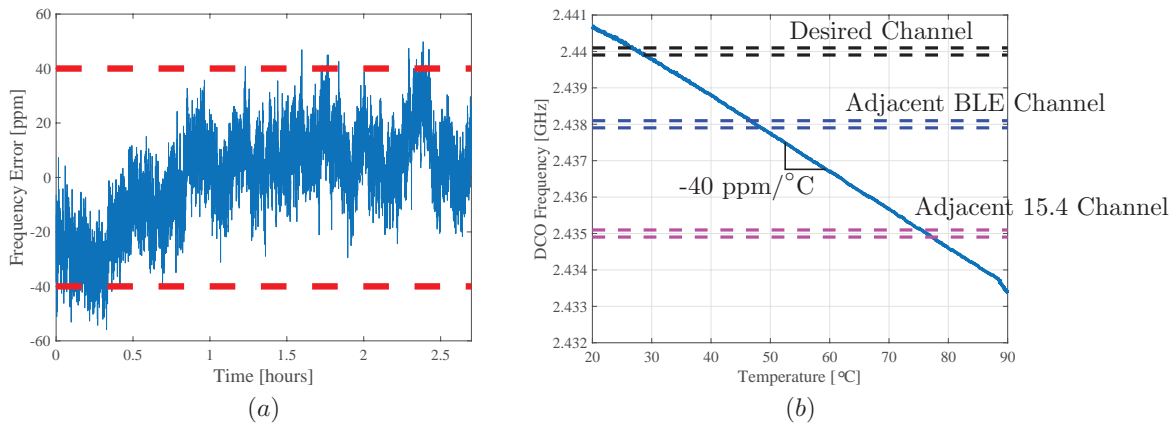
Figure 5.30: Open-loop oscillator frequency on the lab bench (a) and with varying temperature (b)
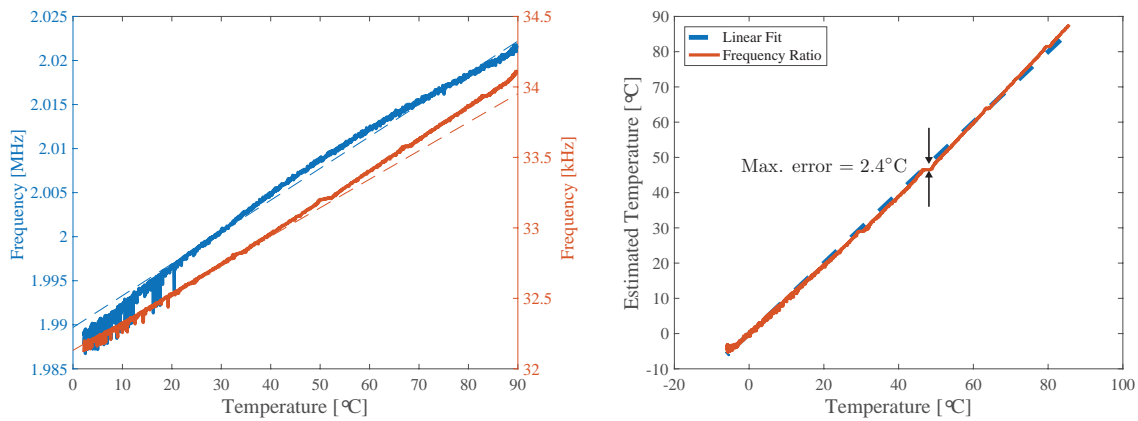
Figure 5.31: Low frequency temperature response of the on-chip 32 kHz and 2 MHz oscillators and a linear fit of their ratio

has a total of 64 kB of data memory and 64 kB of instruction memory, and no flash. So a grand N-calibration across every channel and every temperature is within the realm of feasibility, but would also require a herculean effort for each individual chip.

Eventually, the single chip mote will exist within a network. And the network itself can be relied on to provide some degree of compensation. All of these network strategies assume that the two-point calibration from the previous section is performed, and that there is a temperature measurement on chip that is accurate to within 2 °C. The strategies listed here have essentially been demonstrated already (albeit with higher quality oscillators) in [76].

The first strategy assumes that the single chip mote is used in a network with a large amount of traffic or a network in which the other agents wish to integrate SCM. SCM can
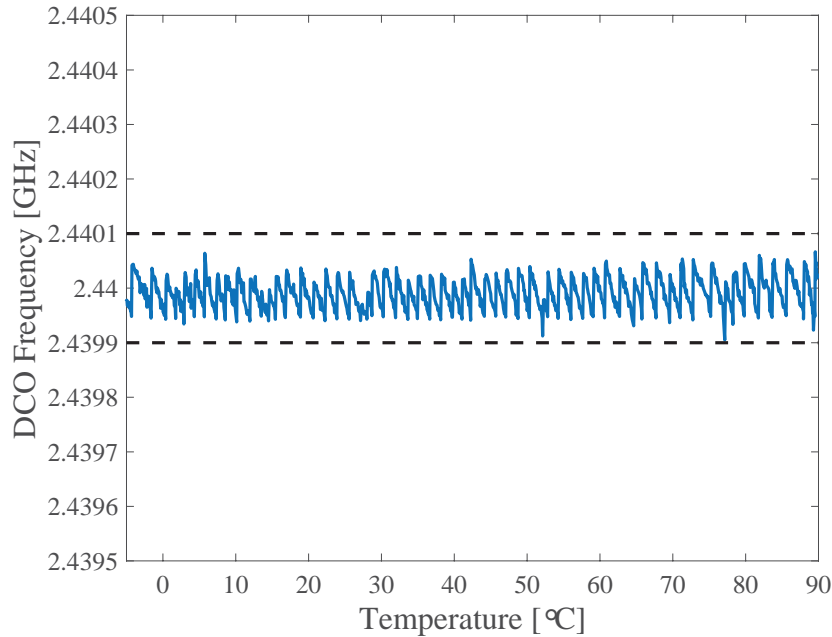
Figure 5.32: Results of idealized calibration conditions

use another radio in the network as a "beacon". This beacon transmits packets at a fixed rate and on a fixed channel. The single chip mote, in receive mode, can then search for this beacon over a narrow frequency range (limited by the calibration error). Once it receives a packet, it knows its exact LO frequency. And, after an uncertain local low-frequency clock is compared against the timing of the incoming packets [77], it can be used as a frequency reference in the future. This process is illustrated in Fig. 5.33. This beaconing approach is not unique, and is used in Zigbee, OpenWSN, and in WiFi.

The next questions are: how much time and how much energy does the mote require to acquire a lock? And, how much energy needs to be expended to make a transmission in the future? To answer the first question, assume that there is a fixed probability PDR that one of the beacon packets is received. In addition, assume that from Fig. 5.29 the maximum error that the two-point calibration is 500 kHz, or 5 LSBs of the tuning characteristic. To simplify things, assume that the single chip mote receiver's "beacon search" starts 5 LSBs below it's estimate, so that the maximum distance in frequency between this initial guess and the correct receiver frequency is 1 MHz, or 10 LSBs. If the beacons are transmitting at a rate of $1/T_B$, it will require a time-to-receiver lock of approximately:

$$T_{lock} = \frac{T_B d}{\text{PDR}} \qquad (5.3)$$

Where d is the number of LSBs between the initial guess and the actual frequency. Because the receiver must be on this entire time, this will consume $T_{lock} \times P_{RX}$.   Note
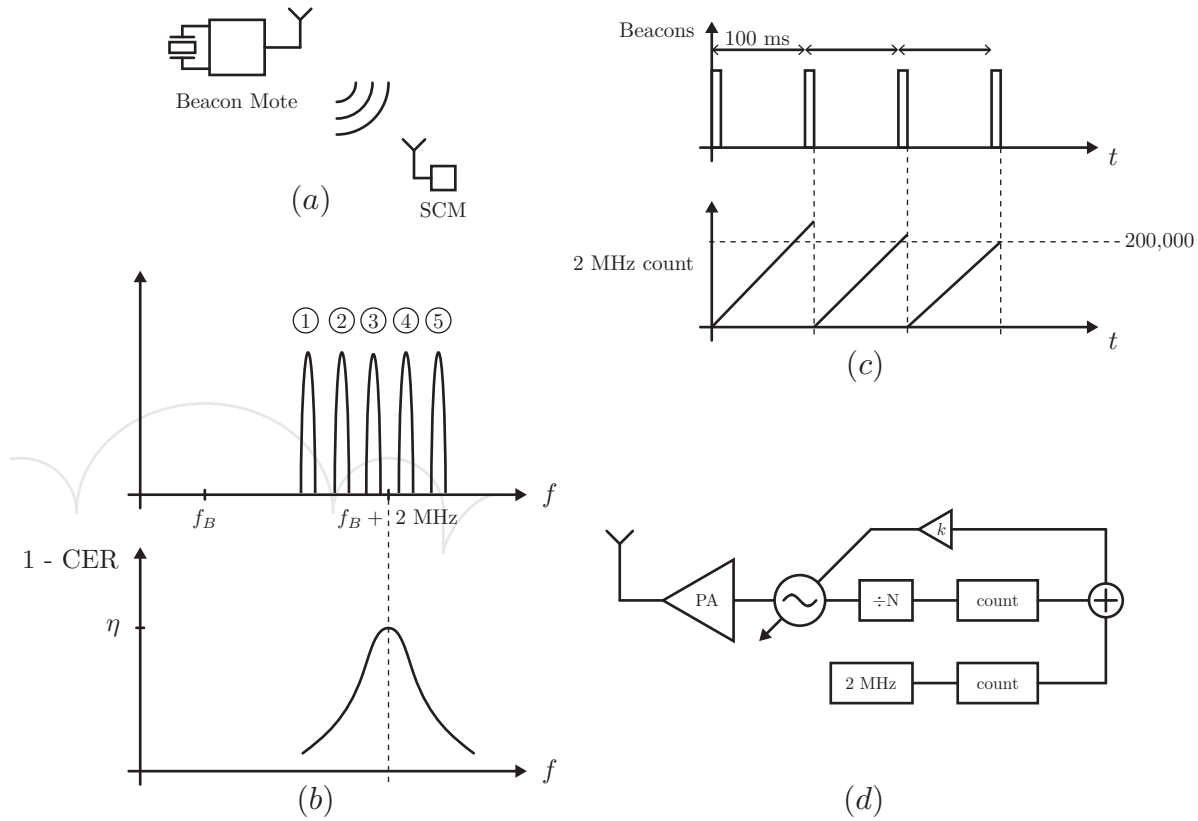
Figure 5.33: Illustration of potential network calibration strategy

that running the divider is optional. The low-frequency oscillator is compensated with packet timing. One alternative, potentially power-hungry, approach is to divide the LC tank frequency (which is now known) and count this result against the low frequency oscillator. To transmit a packet later on (loop shown in Fig. 5.33(d)), it is first necessary to wait approximately 50 μs (see Fig. 5.8) to allow the LO to settle. To obtain the frequency resolution specified by the 802.15.4 standard, there is a tradeoff between divider power and count time. The lower the divide ratio, the higher the divide power. But, the lower divide ratio, the lower the amount of time the divider's output needs to be counted to obtain resolution at high frequency.

The second strategy assumes that SCM exists in a network with little or no traffic. This could be in an application in which the mote is needed purely as a transmitter, as would likely be the case if it were used as a wireless sensor node. Under these conditions, SCM could either be required to transmit a packet when polled, or it could be required to transmit data whenever it measures something interesting. Whether it is polled at a schedule or unknown time, it can use the same technique as in the busy network where there

is a dedicated beacon (the interrogator could operate as a beacon). However, if the mote needs to make an unsolicited transmission, the only possible way to learn anything aside from relying on the two-point calibration is to attempt to transmit at some interval and try repeatedly at multiple different frequencies close to a channel and hope. It is possible to wait for acknowledgement, but this adds complications from the different frequency pulling associated with operating in transmit or receive modes. Unfortunately, if the calibration has gone wrong, nobody will hear SCM scream.

5.7   System Demonstrations

Two factors ultimately determined the success or failure of the transceiver: compatibility with off-the-shelf 802.15.4 and BLE transceivers, and the ability for the transceiver to communicate with others of its kind. This section shows that binary success as well as some of the peripheral hardware required to obtain it.

Bluetooth Low Energy

Because the radio receiver state machine is only compatible with 802.15.4 packets, fully-fledged Bluetooth or BLE communication is not possible. However, BLE does allow for unsolicited advertising on particular channels. And, because the transmitter itself is compatible with the BLE PHY, the single chip mote can send these unsolicited packets on BLE channels 37, 38, and 39 (corresponding to center frequencies of 2.402 GHz, 2.426 GHz, and 2.48 GHz). As a quick reminder, the BLE chipping rate is 1 MHz $\pm$ 40 ppm, which is relatively strict. Therefore, a FIFO must be used to isolate the data-to-chip converter (in this case, the on-chip Cortex M0) from the physical modulation. A somewhat similar FIFO is used for 802.15.4 packet transmission, and is described in [66].
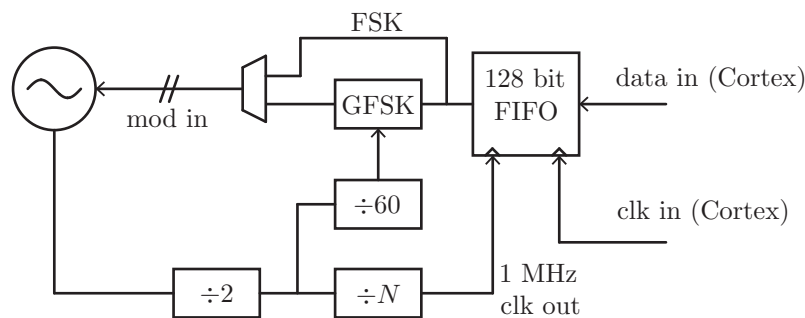
Figure 5.34: On-chip BLE modulation schematic

There was one small error with the asynchronous FIFO on the chip - the output clock is gated off of the input clock. As a result, only 128 bits of BLE data can be sent at the

appropriate data rate. This is still sufficient for one-way communication. The structure of an advertising packet is shown in the following figure (adapted from [78]):

			Payload		
Preamble	AA	Header	AdvA	AdvData	CRC
1 byte	4 bytes	2 bytes	6 bytes	0-31 bytes	3 bytes

			Payload	
0x55	0x6B7D9171	0xA060	AdvA	CRC
1 byte	4 bytes	2 bytes	6 bytes	3 bytes
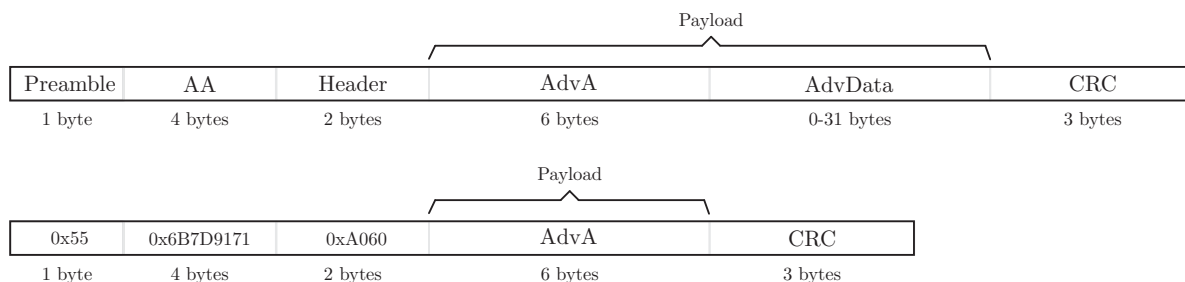
Figure 5.35: Bluetooth Low-Energy general advertising packet structure (top) and minimum-number-of-bits packet structure (bottom)

If the payload field has a size of zero, the valid advertising packet has a length of exactly 128 bits. It is still possible to transmit a small amount of data and a device identifier instead of the advertising address. There are some rules in the BLE specification regarding advertising addresses [79] but they are beyond the scope of a chip in an academic setting. The effective data rate is extremely low, but this does allow the chip to transmit a minimal amount of data asymmetrically. To overcome the limitations of the on-chip buffer size, a significantly larger off-chip shift register was synthesized on an FPGA. This shift register will be included in silicon on a future re-spin of this chip. One additional benefit of this implementation is the ability to send a full 256 bytes of arbitrary 802.15.4 formatted data (data can be clocked out of the FIFO at either 1 Mbps or 2 Mbps).

In addition, BLE does technically require Gaussian Frequency Shift Key (GFSK) modulation. This is approximated by using oversampling a 3-bit binary weighted capacitor DAC at approximately 20 MHz (approximately because this 20 MHz oversampling clock is derived from a divided version of the LC tank oscillator).

Antenna considerations

As a parting note for this chapter, some minimal experimentation was performed with the single chip mote's transmitter and a number of different antennas. Two different antennas were attached to the chip's RF port while transmitting BLE packets at 10 ms intervals. The resulting advertising packet is successfully received by a bluetooth sniffing application on an Android telephone. Note that in this particular experiment, the GFSK block is not actually in use (the local oscillator is being modulation by ordinary FSK with 500 kHz tone spacing). Two different antennas were used: a 12.5 cm rubber ducky with advertised resonance at 2.4 GHz, and a wirebond from the RF port to a capacitance on the far side of the chip. Photographs of the two antennas are shown in Fig. 5.38. In addition, a plot of the measured
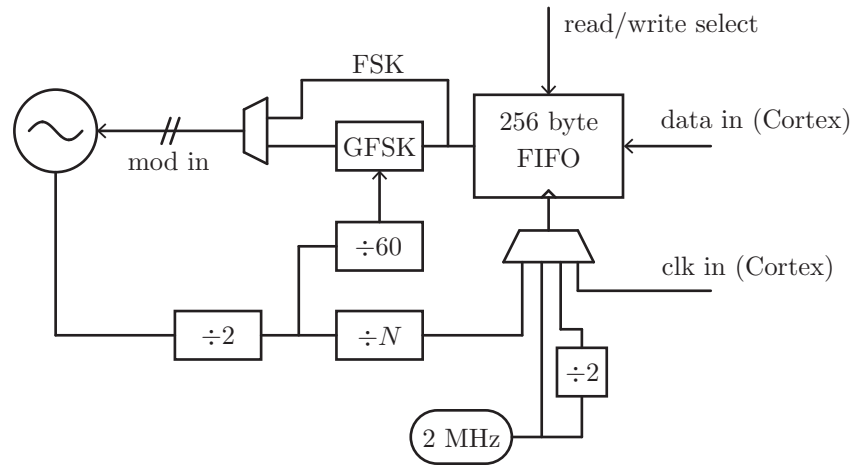
Figure 5.36: Off-chip BLE modulation schematic. In this diagram, only the FIFO itself is off chip. All clocks and control signals are generated on chip

RSSI as a function of approximate distance from the mote. Measurements were taken with a telephone, and similar trends were observed when transmitting a tone from the single chip mote and measuring received power with a spectrum analyzer. These measurements were performed with a transmitter output power of -11 dBm (approximately 3 dB off of the maximum). Assuming a receiver with -90 dBm of sensitivity, the mote with the wire bond antenna has a range of approximately 1.7 m, potentially sufficient for asymmetric short-range sensor applications.
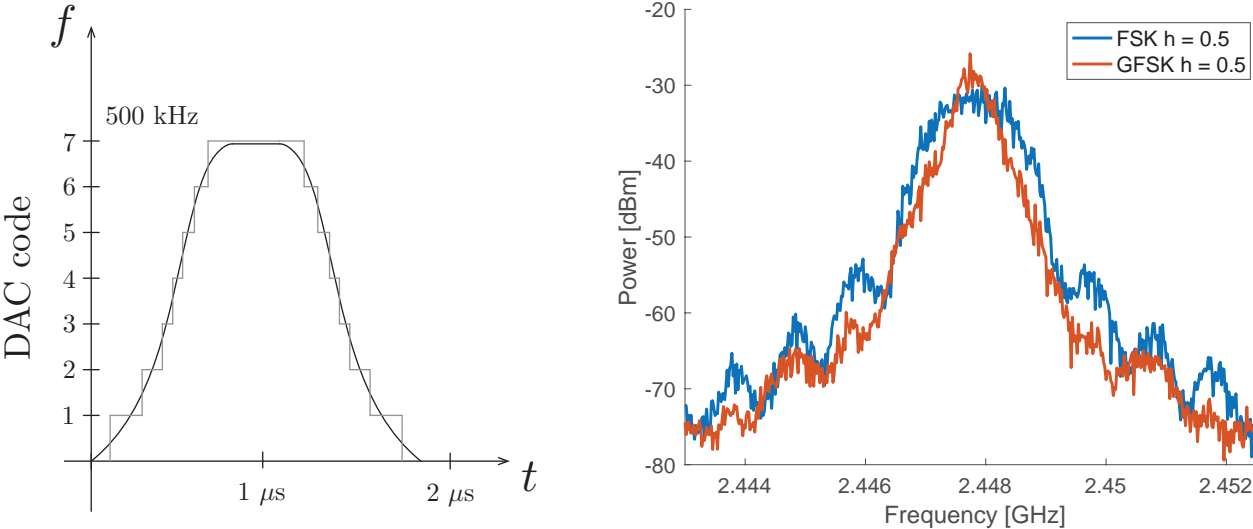
Figure 5.37: Comparison of modulated spectrum with a frequency spacing of 500 kHz, data rate of 1 Mbps, with both generic FSK and Gaussian FSK
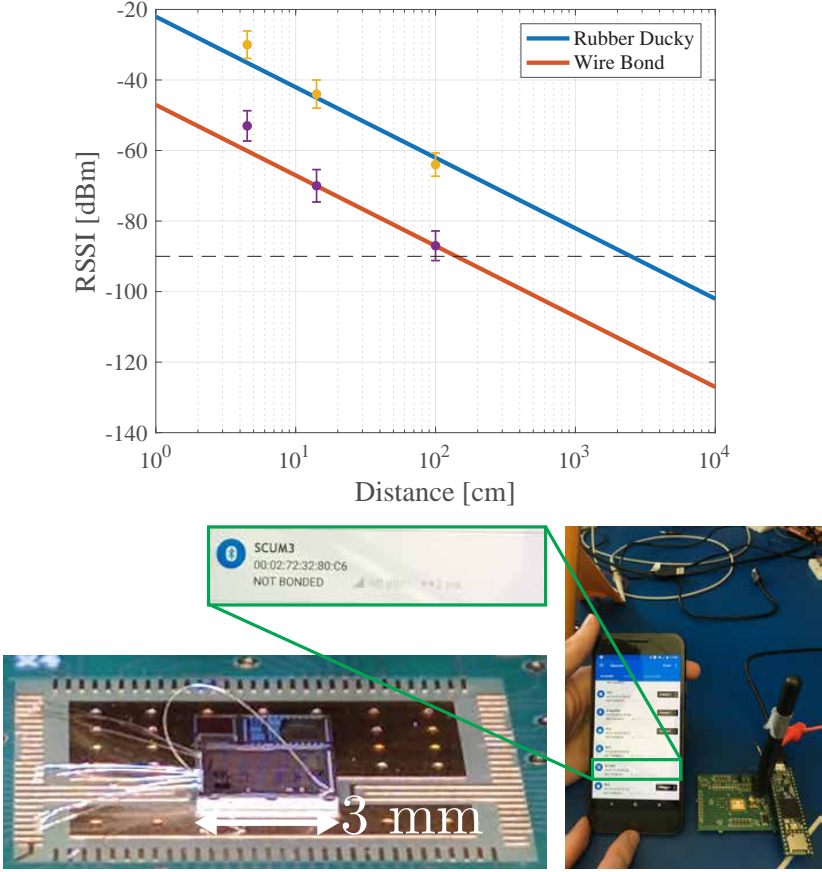
Figure 5.38: Single chip mote transmitter with wirebond and rubber ducky antenna; demonstration of BLE advertising capability

Chapter 6

Monolithic Transceiver Integration

This chapter details the efforts to integrate every component of a transceiver onto a single silicon die in a standard CMOS process. The chip includes the frequency synthesizer, the radio's analog signal processing, and the antenna. Baseband digital processing was not included on the chip, but the work summarized in the previous chapter demonstrated that inclusion of digital processing is possible. Unlike [80] energy harvesting and storage are not included on this chip. However, the power budget of the transceiver is small. And, while it is difficult to characterize the energy/bit performance of the receiver as the majority of power is used to drive the off-chip baseband processing and associated parasitics, the transmitter burns a DC power of 800 μW from a 1V supply (transmitter output power of -3dBm). And, in simulation, the transmitter can transmit OOK at a rate of 1 Gb/s and the frequency has four bits of tuning for FSK modulation, resulting in a combined data rate of 4 Gb/s. The energy efficiency of the transmitter is therefore simulated to be approximately 200 fJ/bit, which makes it competitive with state-of-the-art optical transmitters [81].

6.1 Selection of Carrier Frequency

The choice of carrier frequency heavily influences the design of associated circuitry. As carrier frequency increases, antenna radiation resistance increases, reactance decreases, and efficiency increases. In addition, inductor quality factor increases, but this turns out to be insignificant because SRF concerns limit the inductance and therefore the LQ product. In addition, allowable fractional bandwidth at higher frequencies is generally larger. However, operating at higher frequencies and bandwidths requires higher peak power consumption, and assuming equal transmit power and sensitivity, higher carrier frequency links have lower range. The goal of the analysis is to maximize range and data rate while minimizing power and maintaining a form factor that is feasible in CMOS. Starting with the Friis transmission formula, and the model for the wireless link shown in Fig. 6.1.

$$P_{r,dBm} = P_{t,dBm} + 2G_a + 20\log_{10}\left(\frac{\lambda}{4\pi d}\right) \tag{6.1}$$
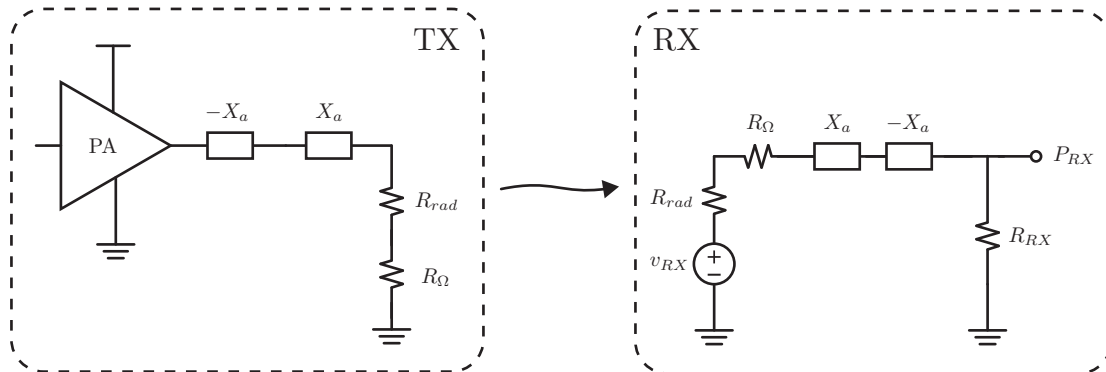
Figure 6.1: Illustration of Link Budget

Where P_r is the received power, P_t is the transmitted power, G_a is the antenna gain (the product of efficiency and directivity), $\lambda = c/f$ is the carrier wavelength, and d is the range of the link. The receiver impedance is perfectly matched to the receive antenna. For simplicity, assume that the antennas have 1.7 dB of directivity ($G_a = 1.7$ dB). To start, a brief analysis of the transmitter and antenna is shown to demonstrate the tradeoff between carrier frequency and available power at the receiver. Notice in Fig. 6.1. If the total DC power in the transmitter is P_{DC}, then the radiative power delivered to the antenna is:

$$P_t = \eta P_{DC} \left(\frac{R_{rad}}{R_{rad} + R_\Omega} \right)^2 \tag{6.2}$$

In Eq. 6.2, η is the transmitter's efficiency, R_{rad} is the antenna's radiation resistance, and R_Ω is the ohmic, or loss, resistance of the antenna. The Friis equation from Eq. 6.1 then becomes:

$$P_r = 10\log_{10}\left(\eta P_{DC} \left(\frac{R_{rad}}{R_{rad} + R_\Omega} \right)^2 \right) + 2G_a + 20\log_{10}\left(\frac{\lambda}{4\pi d} \right) + 20\log_{10}\left(\frac{R_{rad}}{R_{rad} + R_\Omega} \right) \tag{6.3}$$

Notice that the antenna power loss is counted twice, once for transmit and once for receive. Electrically small antennas inherently have low efficiency, as stated in the Chu-Harrington-Wheeler theorem (originally in Chapter 1, restated here):

$$Q \geq \frac{1}{(ka)^3} + \frac{1}{ka} \tag{6.4}$$

In this equation, Q is the ratio of antenna reactance to radiation resistance, k is the wave number, and a is the largest radial dimension of a sphere that contains the antenna [28]. In this equation, the Q only accounts for radiation resistance, so the radiation resistance is small. That means that efficiency is low because the ohmic resistance of the structure

will be significantly larger than the radiation resistance. To further summarize this brief overview of wireless communication: equation 6.1 states that higher transmitter power and higher antenna efficiency result in higher received power (directivity also helps, but is often difficult to attain in practice particularly with electrically small antennas, which are nearly isotropic). The Q in Eq. 6.4 assumes no ohmic loss. Intuitively, a high Q means that the radiation resistance is low, and therefore, the power transmitted to the receiver will be small. So, assuming CMOS chip size scales of mms, operating at higher frequencies is attractive. However, increasing the operating frequency increases path loss. To analyze this further, the radiation resistance, loss resistance, and reactance of two antennas are shown:
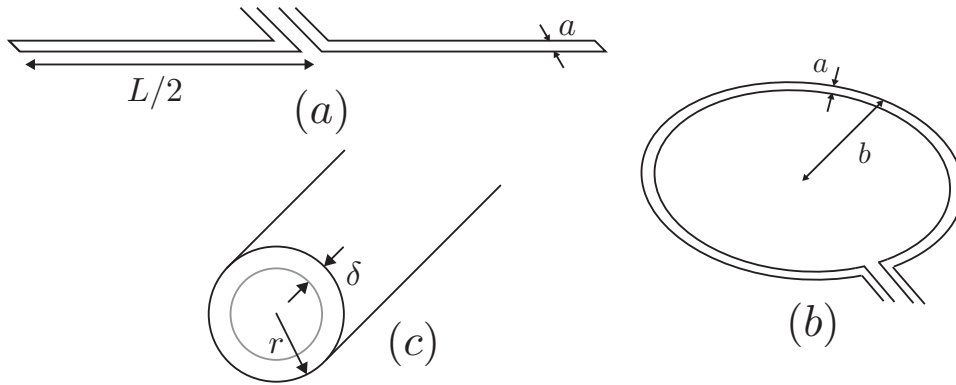
Figure 6.2: Antenna dimensions for (a) a dipole and (b) a loop; and (c) the cross-section of the antenna conductor

$$R_{rad,dipole} = 20\pi^2 \left(\frac{L}{\lambda}\right)^2 \quad (6.5) \qquad R_{rad,loop} = 31171 \left(\frac{\pi b^2}{\lambda^2}\right)^2 \quad (6.6)$$

$$X_{dipole} = -\frac{377}{\pi^2}\frac{\lambda}{L}\ln(L/r) \quad (6.7) \qquad X_{loop} = +377\frac{2\pi b}{\lambda}\ln\left(\frac{b}{a}\right) \quad (6.8)$$

$$R_{\Omega,dipole} = \frac{L}{6\pi r}\sqrt{\frac{\pi f \mu}{\sigma}} \quad (6.9) \qquad R_{\Omega,loop} = \frac{b}{r}\sqrt{\frac{\pi f \mu}{\sigma}} \quad (6.10)$$

These expressions are directly taken or derived from chapters 4 and 5 of [82] and chapter 13 of [83] and use the dimensions shown in Fig. 6.2. There are a number of significant assumptions built into the derivations of these impedances, particularly that the antennas are electrically small (largest dimension $<< \lambda$). In addition, the cross-section to calculate the ohmic loss is shown in Fig. 6.2 with an estimated skin depth.

Fig. 6.3 graphically shows the range/frequency tradeoff in Eq. 6.3. In Fig. 6.3, loop antennas with diameters of 1 mm and 5 mm are used (with corresponding resistances and reactances in Eqns. 6.6, 6.10, and 6.8), and the received power is plotted as a function of
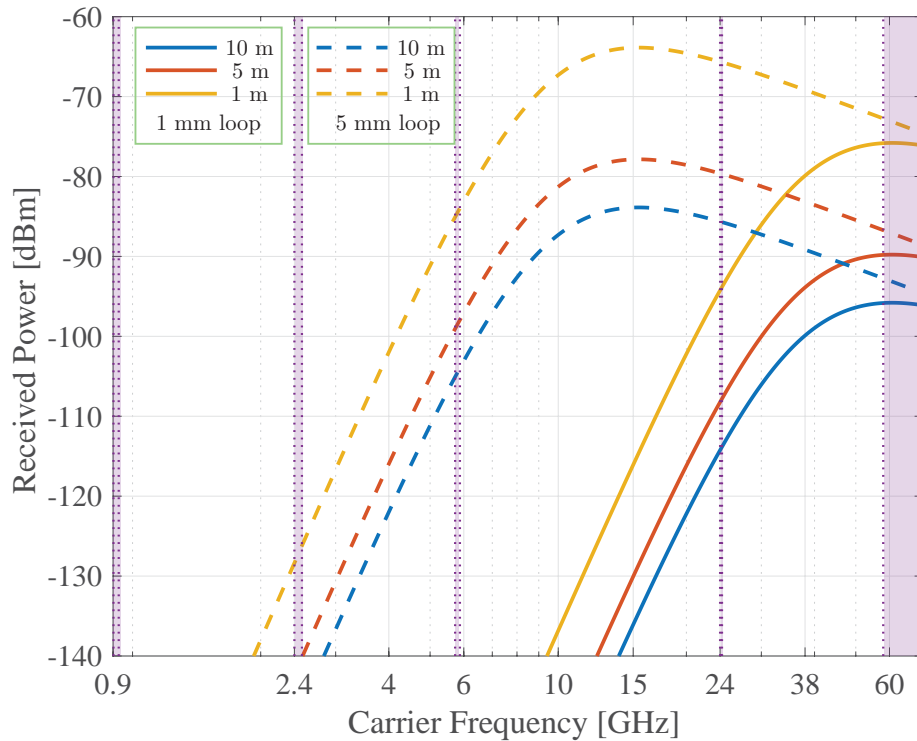
Figure 6.3: Received power in the symmetrical link

carrier frequency at ranges of 1 m, 5 m, and 10 m. The ηP_{DC} product is assumed to be 1 mW. The purple bands represent the ISM (industrial, scientific, medical) bands that are allocated for free usage by the FCC. The assumption in the derivation of electrically small antenna parameters is that the current distribution throughout the structure is constant for a loop, or linearly tapering in a dipole. This assumption breaks down, and the structure begins to exhibit capacitively or open-circuit loaded transmission line behavior, when the dimension of the structure approaches $\lambda/10$. This occurs at a frequency of approximately 30 GHz for the 1 mm diameter loop, and at 6 GHz for the 5 mm loop. Fig. 6.3 does not include these effects, and is slightly pessimistic.

This relation is great for using an electrically small on-chip antenna, but this component must be integrated with CMOS transmitters and receivers. And, while f_T and f_{max} of deep submicron processes are well above 200 GHz, careful design is required at the frequencies at which on-chip antennas are not catastrophically lossy. Design up to 10 GHz is basically narrow-band analog. Above this, gate resistances, via inductances, substrate capacitor Qs, and routing parasitics begin to limit performance.

So far, communication of actual data has been ignored. A wireless link transmitting a tone is great, but the sending of actual data is limited by the Shannon limit:

$$I < B\log_2\left(1 + P_r/N\right) \tag{6.11}$$

Where I is the effective *reliable* information rate, B is the channel bandwidth (which is limited by the FCC, and is related to the data rate) and the SNR of the received signal within the bandwidth of the receiver. Note that the antenna and corresponding matching network (if any) will limit the bandwidth of the received signal. The second equation determines the bit error rate that depends on the spectral efficiency of the modulation:

$$P_r/N = E_b/N_0 \frac{r}{B} \tag{6.12}$$

In this equation, r is the data rate. The quantity E_b/N_0 defines the bit error rate (BER) depending on the modulation. As E_b/N_0 increases, the bit error rate decreases. One additional equation that effectively sets the noise figure requirement of the receiver is:

$$S = -174 + 10\log_{10}(B) + \text{NF} + \text{SNR}_{req} \tag{6.13}$$

In this equation, S is the receiver's sensitivity, and is dependent on NF, the receiver chain's noise figure, and SNR_{req} is the minimum required SNR at the output of the receiver to detect a signal. The -174 is the noise power in 1 Hz of bandwidth at room temperature. This expression completely ignores linearity. Ideally, a receiver should digitize the signal as close to the antenna interface as is possible. However, it would require an outrageously high ADC resolution, so some degree of gain is necessary. In addition, downconversion to a low frequency is also necessary in a heterodyne or superheterodyne architecture, and the mixing operation contributes some level of noise. To calculate the receiver's noise figure, it is necessary to assume a particular topology. In this case, assume a mixer-first architecture with a shunt matching reactance (for voltage gain and to increase the antenna's source resistance). The mixer is then followed by an OTA in resistive feedback. This type of architecture is analyzed in [84]. A mixer-first receiver with inductive source impedance is also analyzed in [85]. The schematic of the receiver with noise sources is shown in Fig. 6.4.

An approximate noise factor for the expression (adapted from [84]) is:

$$F = 1 + \frac{R_{sw}}{R'_a} + \frac{(R'_a + R_{sw})}{R'_a}\frac{\pi^2 - 8}{8} + \frac{2}{\pi^2}\frac{R_F}{R'_a}\left(\frac{2(R'_a + R_{sw})}{\pi^2 R_F}\right)^2 + \frac{2}{\pi^2}\frac{\overline{v_{nA}^2}}{4kTR'_a} \tag{6.14}$$

Noise figure is: $NF = 10\log_{10}(F)$. In this expression, R_{sw} is the switch resistance, R'_a is the transformed antenna impedance (transformed by the match X_a and X_r), R_F is the OTA feedback resistance, and $\overline{v_{nA}^2}$ is the input referred noise of the amplifier per Hz of bandwidth. The takeaways from this formula are: the transformed antenna impedance should be as high as possible. The consequence of that is: using a high-Q reactive antenna and matching its reactance is actually beneficial for receiver noise figure performance (impact of circuits' noise is less). In addition, it is beneficial to make R_F as large as possible, thus making the input impedance of the amplifier effectively an open circuit at low frequency. Then, the
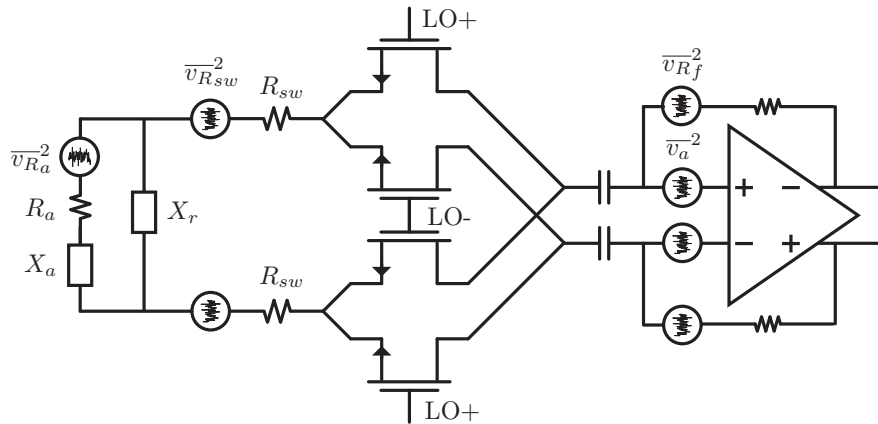
Figure 6.4: Simulated antenna performance

noise of the amplifier boils down to a straightforward power-gain-bandwidth tradeoff. In addition, if the antenna resistance is very large, it is possible to make the switch resistances correspondingly large. This, in turn, increases the input impedance. A similar conclusion was reached in [86]. And, according to [87], using a low duty-cycle sinusoidal drive increases the input impedance (and noise), but also increases the conversion gain of the mixer from -3.9 dB to as high as 0 dB, although this case requires extremely low duty cycle and suffers from poor linearity, so it is not practical. In the case that the source resistance is still larger than the switch impedance, this improvement in gain reduces the power requirements on the proceeding baseband amplifiers.

Ultimately, a carrier frequency of 24 GHz was chosen as the carrier frequency. Even with an electrically small antenna with high reactive impedance it is still possible to have a good noise figure because of the potential for passive gain in the front-end. And it can transmit with reasonable radiated efficiency. The main reason for choosing 24 GHz over 60 GHz was that modestly sized mixer switches could still be resonated in an oscillator with a relatively high LQ product inductor.

6.2 Antenna Design

The analysis in the previous section relied on a number of assumptions and simplifications that, while realistic, are not particularly conducive to design of the transceiver itself. While the dipole antenna has superior efficiency compared to the loop antenna, particularly if meandered to increase its effective length [88] it is important to remember that a sub-resonant dipole will be capacitive, which means that it must be matched using an inductive element. As discussed in chapter 2 of this dissertation, inductors can be quite lossy, even at high frequencies, and it is difficult to generate the high impedances needed to match to an

on-chip dipole antenna. In addition, input impedances of receiver circuits will be capacitive which means that the inductor would need to resonate the reactance of both the antenna and the receiver circuitry. So, in spite of its inferior efficiency, a loop antenna was used to simplify the design of the matching network and corresponding receive and transmit circuits.

At a 24 GHz carrier with an area limit of 1 mm^2, the antenna will be electrically small, and thus inherently somewhat inefficient. However, one benefit of using an electrically small antenna is that the source impedance in receive mode (and load impedance in transmit mode) is both reactive and relatively high Q (see Eq. 6.4). In receive, a high source impedance intrinsically improves the receiver's noise figure, as the majority of noise will be generated by the source. And, while a high load impedance (for power matching) can be difficult to generate, a mixer-first architecture with small switches presents a high impedance even at high frequencies. The use of an inductive antenna is challenging because the self-resonant frequency of a loop of any meaningful size is compromised by the proximity of the grounded substrate (in spite of this being a relatively high impedance connection, it still adds a significant parallel capacitance). This ultimately limited the area and efficiency of the on-chip antenna. HFSS simulations of the antenna over a silicon substrate are shown in Fig. 6.5. This is compared to an EMX simulation of the same structure. HFSS estimated a radiation efficiency of 8% (-11 dB). This is slightly higher than the discrepancy between HFSS and EMX assuming that all of the additional resistance in the HFSS model at 24 GHz is radiative.
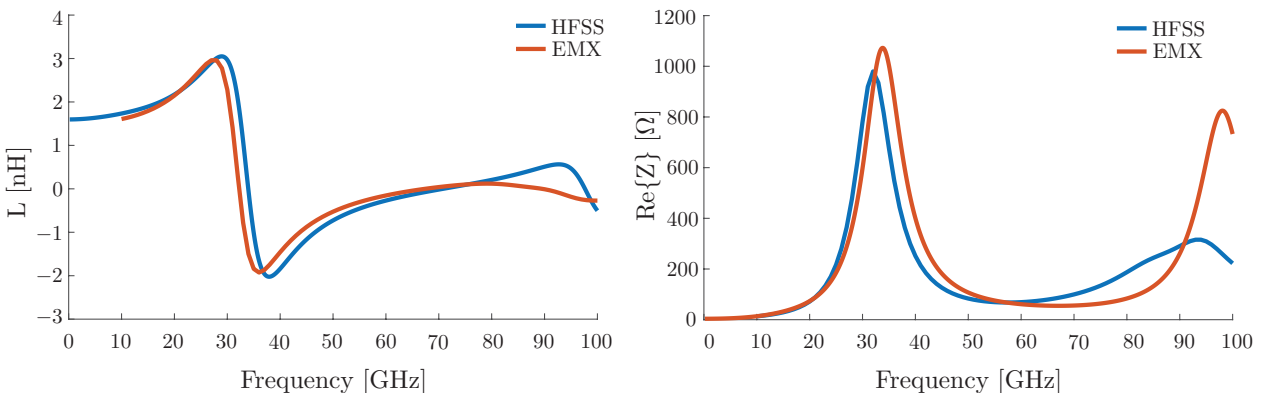
Figure 6.5: Simulated antenna performance

6.3 Transceiver Design

System Architecture

The schematic of the full transciever system is shown in Fig. 6.6.
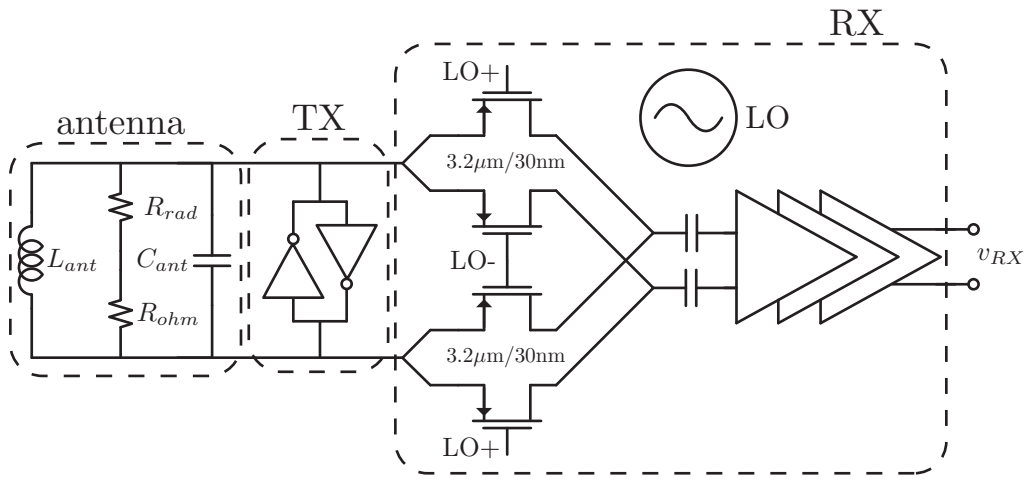
Figure 6.6: 24 GHz transceiver schematic with shared antenna interface. R_{rad} and R_{ohm} are transformed parallel resistances.

Transmitter

The transmitter takes advantage of the inductive antenna impedance as the tank of a CMOS class-B oscillator. Because the load impedance is high, high drain efficiencies are possible at low output powers because the inductance that would normally be part of a matching network is built into the load. Another advantage of using this topology for the transmitter is that the antenna will operate near its own self-resonant frequency (shifted slightly due to parasitic capacitance of the active components). The differential impedance of the antenna itself is both large and inductive below the self-resonant frequency. That allows the use of a cross-coupled Class C or Class D oscillator in lieu of a more traditional power amplifier output stage. The additional benefit is that there is no carrier synthesis overhead. All power spent in the entire transmitter system is delivered directly to the load. If a constant-envelope modulation is used, then amplitude noise of the transmitter has no impact on the link. Phase noise, on the other hand, could reduce performance. Because there is not a demodulator built into the chip, it is difficult to quantitatively determine the degradation of receiver sensitivity caused by phase noise.

The schematic of the transmitter is shown in Fig. 6.7.

A selection of simulated modulated waveforms from the transmitter are shown in Fig. 6.8 and Fig. 6.9. The maximum transmitter speed for on/off key was 1 Gbps. The amplitude modulation was not linear, so only one bit is actually useful.
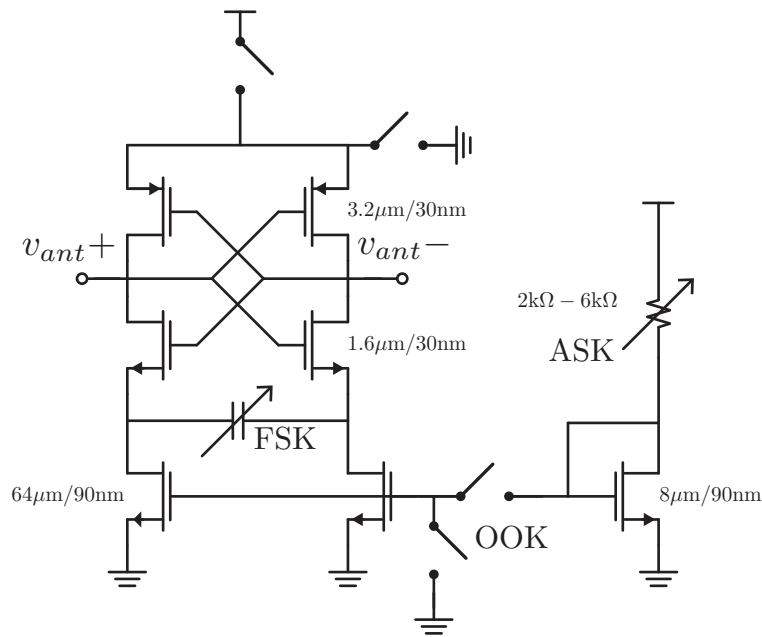
Figure 6.7: Transmitter schematic (antenna not shown) with annotated modulation
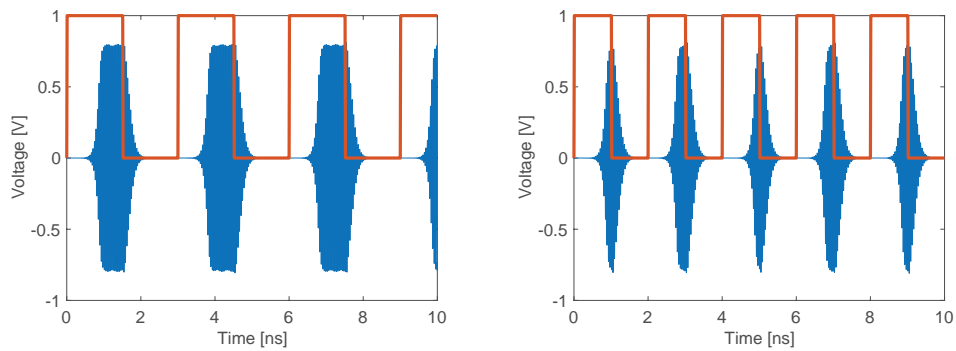
Figure 6.8: Simulated OOK modulation of the transmitter at 667 Mbps (left) and 1 Gbps (right)

Receiver

As previously mentioned in this chapter, a mixer-first receiver was selected because, if the switches are small, it can present a high impedance to the antenna. The first baseband stage following the mixer is a simple circuit from [89] that has large feedback resistors so that the input impedance is large. This is done in a further attempt to mimic the observations
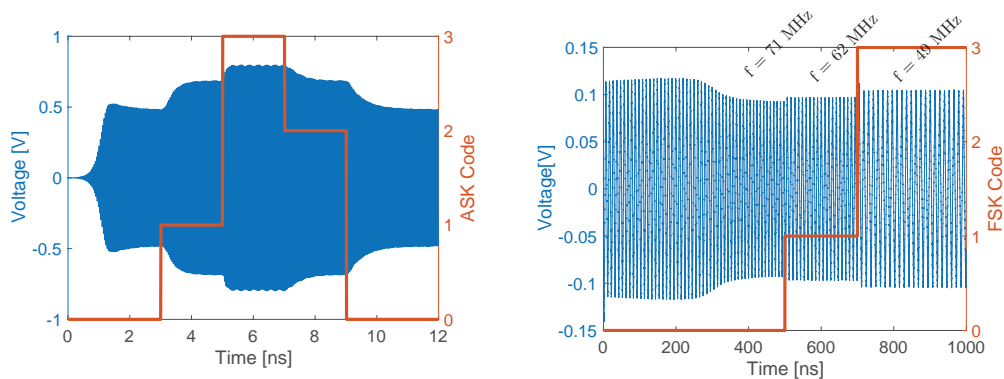
Figure 6.9: Simulated ASK (left) and FSK (right) modulation of the transmitter. The FSK shows the output of the mixer in a loop-back test

in [86] regarding optimal receiver impedance. This means that the real component of the antenna impedance (after matching) is less than the input impedance of the receiver. The baseband chain is shown in Fig. 6.10.

The local oscillator uses the same class-B topology that has been presented multiple times in this dissertation. It uses a PMOS current source so that the DC level of the oscillation is more appropriate for driving the mixer switches. As duty cycle of the gate drive waveform decreases, mixer gain and mixer noise both increase. However, if there is ever overlap in mixer drive waveforms, conversion gain rapidly degrades. The oscillator is shown in Fig. 6.11. The inductor at 24 GHz had an inductance of 780 pH and a Q of 13.9. It occupied an area of 220 μm by 220 μm, including the guard ring. The tuning characteristic is also shown in Fig. 6.11. Because the TX and LO center frequencies were based on two different inductors, it was important that the receive LO was tunable to a frequency within the IF bandwidth of the transmitter frequency.

Simulations of receiver noise figure are shown in Fig. 6.12 (a). And, simulations of baseband gain and bandwidth are shown in Fig. 6.12 (b). The noise figure was simulated with a differential source impedance of 1 kΩ, which was estimated from the 1-port HFSS model with parasitic capacitive loading from the transceiver. Linearity was ignored for a number of reasons. First, at these power levels, it is quite challenging to maintain linear circuits. Second, an interferer of significant power will cause more issues than baseband linearity. It will pull the local oscillator frequency both through the antenna port and via radiation to the oscillator's inductor (which is also basically an antenna).

RX-TX co-integration

There is no AC coupling between the antenna and the mixer. So, in transmit mode, the DC level of the antenna is raised approximately to mid-rail. The baseband stages are power gated, so the IF side of the mixer is pulled to ground with a high impedance connec-
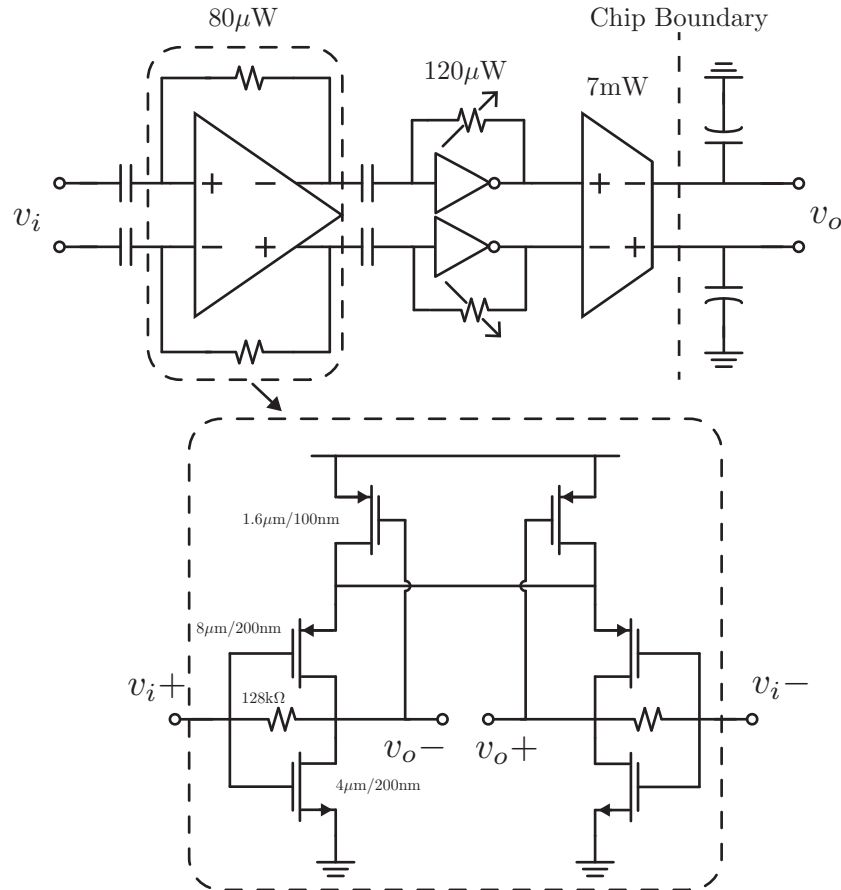
Figure 6.10: Baseband stage schematics and estimated power consumption

tion. This is the primary reason why the oscillator running makes little to no difference in the transmitter functionality. In receive mode, when the transmitter is off, the rail of the transmitter is pulled to ground and the gate of the current source is also pulled to ground.

6.4 Measured Results

Because of designer error, it was not possible to turn the receiver LO off (a switch from NMOS current source to PMOS current source was made late in the game, and the corresponding on/off circuitry was not modified). The local oscillator was measured to consume approximately 800 μW of power. The unmodulated transmitter oscillator circuit was measured to burn approximately 1.1 mW. The baseband stages in total burned 240 μW, and the baseband output driver burned 7 mW. All of these numbers were measured from a board-regulated 1 V supply. The first test that was performed was a loop-back test,
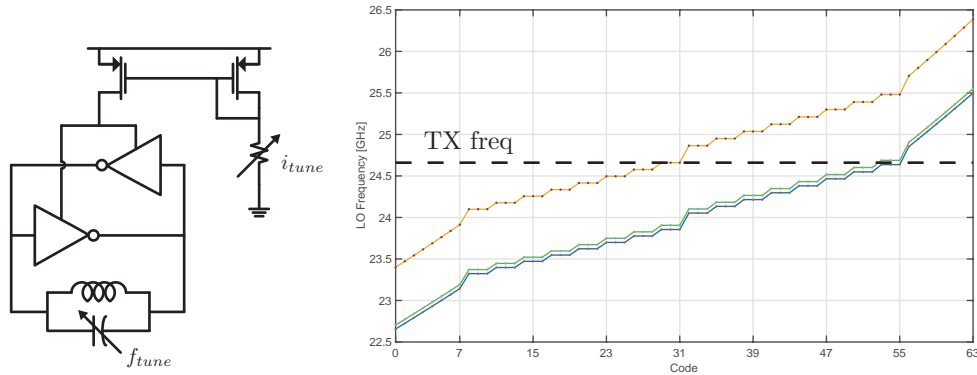
Figure 6.11: Local oscillator and tuning characteristic
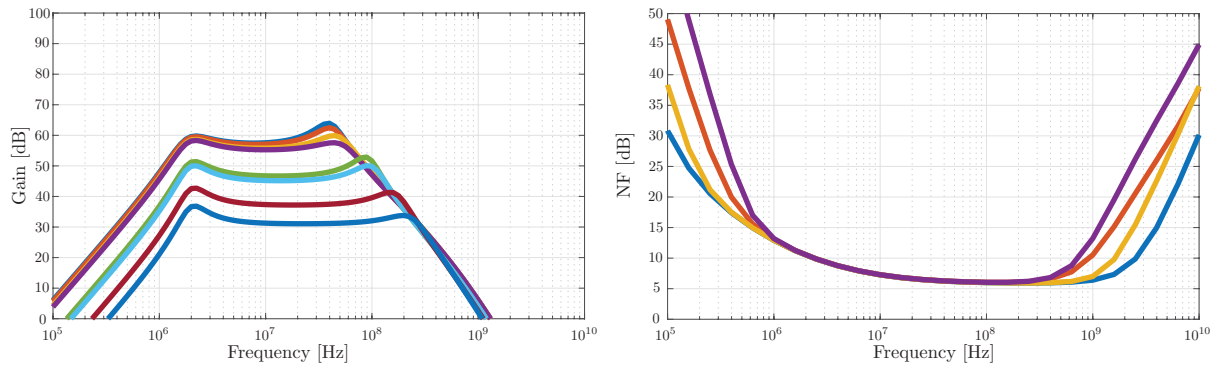
Figure 6.12: Simulated receiver transfer function and noise figure (with mixer) at various gain settings

in which the transmitter and receiver on the same chip were activated. Assuming minimal chip-to-chip variation, the receiver LO could be tuned to receive the transmitted signal from another chip.

Notice that at a high LO frequency (coarse code = 0) decreasing the LO frequency decreased the intermediate frequency. At a low LO frequency (coarse code = 1) decreasing the LO frequency has the opposite effect. This strongly suggests both that the transmitter is, in fact, running, and that it is generating a tone at a frequency somewhere between the two coarse codes. However, there was never successful transmission from one chip to another. Suspected reasons are: antenna loss was higher than simulated, and receiver sensitivity was worse than expected because of incorrectly simulated antenna-mixer interface.
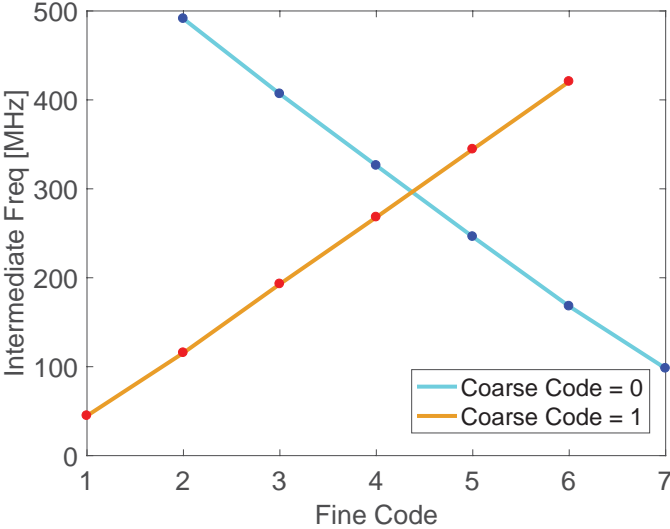
Figure 6.13: Intermediate Frequency versus LO code in a loopback test

Chapter 7

Conclusions and Future Work

7.1 Conclusions

This work discussed the challenges of a fully-integrated, single-chip, radio transceiver. This included hefty discussion of the details of necessary circuit components as well as measured silicon results and some demonstrations of system-level functionality. Ultimately, the most successful contribution was the demonstration of a crystal-free radio that could communicate with off-the-shelf wireless transceivers using existing network protocols. The removal of the crystal oscillator as a necessary component in a low-datarate radio greatly reduces the size and cost of a wireless sensor node. This, in turn, unlocks a myriad of new applications in which either of these factors were prohibitively limiting. In addition, the chip consumes around ten times less peak power than existing commercial hardware, which extends battery life and could lead to standards compliant transceivers that use energy harvested from their environment.

7.2 Future Work

There are a number of opportunities for improvement on the Single Chip Mote v3. This section will discuss some of the feasibility and merits of a number of simple circuit and system changes that could improve performance or unlock new potential applications for the transceiver.

Switching Supply Regulation

A significant quantity of the power consumption of SCM v3 was in series regulation of constant-current loads, such as the power amplifier and local oscillator. A simple solution to this problem is to use on-chip DC-DC converters that can efficiency higher than the inverse of the conversion ratio (which is the maximum possible efficiency of a series or shunt regulator). This idea was briefly discussed at the end of Chapter 4, but in minimal detail because it did

not function properly. As an example, [90] has demonstrated 3:1 step-down conversion in an area of 0.378 mm^2 with over 60% efficiency. The area is considerable, but this is primarily because of the high current draw that the DC-DC converter needed to accomodate. In addition, if the single chip mote is ever re-designed in a more advanced CMOS technology node, all of the processor and digital baseband area will be significantly reduced, thus leaving room for DC DC converters. [61] has demonstrated high drain efficiency from a Class-D PA operating at a supply voltage of 600 mV, and [72] has demonstrated best-in-class oscillator figure of merit using a 600 mV supply, but it is feasible to reduce this voltage considerably further - potentially lower than 200 mV, using either the class D or class F oscillator from Chapter 2. Assuming that the circuits consumed the same current that they do on SCM v3, and assuming a DC-DC conversion efficiency of 60%, this could reduce the transmit power from 1 mW to less than 500 μW. The real improvement, however, would come in receive mode, in which the local oscillator is the block with the highest power consumption. Its active power could be reduced to around 50 μW, less than the power consumption of the baseband chain and demodulator on SCM v3.

Low-Latency Wireless

Low-power wireless networks can be applied to industrial sensing and automation, particularly in circumstances in which wired connections can be cumbersome or expensive. In these wireless control applications, operating with extremely low-latency (sub-ms or even sub-μs) is important. High-reliability, low-latency wireless communication has been proposed in [91], and the conclusion of the work is that the key is spectral diversity. This has been experimentally confirmed to some degree in an 802.15.4 context in [92]. However, many wireless standards, OpenWSN being one example, explicity trade away latency for improved reliability. One way to create a high reliability transmitter is to have a single radio that transmits the same data on multiple different channels simultaneously, as shown in Fig. 7.1.

If the goal is to suppress image upconversion, the mixers must be quadrature. However, with this scheme, any in-band image tones will be located on 802.15.4 channels and will be in-phase with the signal that was there originally. Alternatively, complex baseband could be generated at DC and upconverted the same way without any issues with out-of-band emission. Note that in this particular configuration, the output stage would need to be fairly linear to limit spurious emissions from intermodulation products.

Duty-cycled PLL

As mentioned in Chapter 5, LC tank oscillators are the highest quality frequency references available on chip. However, as discussed in detail in Chapters 2 and 4, tuning range and resolution can be challenging. On the other hand, a ring oscillator can have very wide tuning range, but suffers from inferior phase noise. A ring oscillator frequency locked to an LC tank could have high performance close-in phase noise while maintaining the wide tuning range. This has potential applications for a crystal-free software defined radio, or in
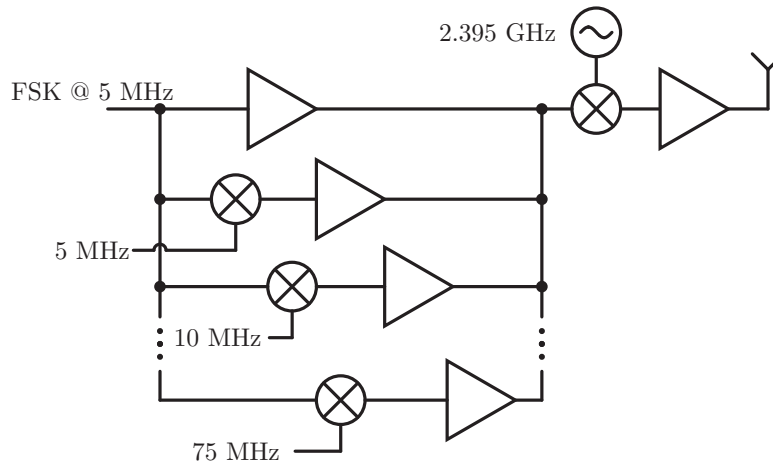
Figure 7.1: Block diagram of multi-band transmitter

a scenario in which multiple separate ring oscillators need to be locked to a single on-chip oscillator. Any area benefits from using a ring oscillator are sacrificed to use the LC tank.

The general idea is to use a low divide ratio divider for both the LC tank and the ring oscillator, compare them using either a phase or frequency detector, and adjust the ring oscillator accordingly. Then, after a lock is made, the loop and the LC tank can be turned off. The duration over the which loop is on dictates the bandwidth below which the ring oscillator's phase noise should follow the higher quality LC tank. The loop bandwidth of the PLL or FLL must be higher to allow the loop to settle. A proposed block diagram is shown in Fig. 7.2.

A similar technique is proposed in [93].

7.3 Parting Words

Although crystal-free radio does have some potential niche applications, and could re-invigorate the waning enthusiasm in the internet of things and, perhaps more generally, wireless sensor networks and nodes, it does lack a certain broad appeal. Even the slightest supply or temperature variation can completely knock the local oscillator off channel, distort the motes' sense of time, and not just degrade performance, but knock the transceiver out of operation completely. Re-establishing a wireless connection could then take a significant amount of overhead, both in energy and in time. In most applications, having an absolute frequency reference just makes sense. One potential application of using free running oscillators is in mm-wave and THz frequency synthesis. Although these oscillators will be equally susceptible to supply and temperature variation, using a PLL is actually detrimental to performance. It takes an enormous quantity of power to divide or manipulate the waveform.
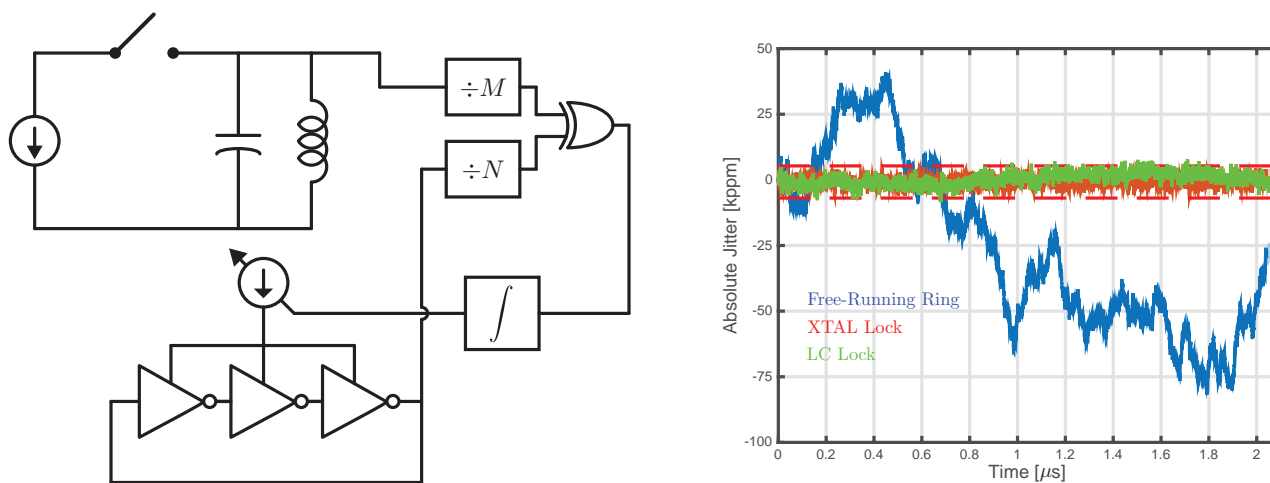
Figure 7.2: Schematic and simulated result of a ring oscillator with and without period correction with LC tank. In this simulation, the LC tank is continuously running.

And, because close-in phase noise is limited by the reference phase noise multiplied a large divide ratio, close-in phase noise is actually degraded.

The centerpiece of this dissertation, the Single Chip Mote v3, could become a useful research tool in the future, both for sensor integration and applications of tiny, cheap radios, and for research into timing and communication in wireless network protocols. In addition, the chip's low power consumption and high degree of integation allow it to be the brains and means of communication for micro robots.

Bibliography

[1]   Statista. *Number of mobile phone users worldwide from 2013 to 2019 (in billions)*. 2018. URL: https://www.statista.com/statistics/274774/forecast-of-mobile-phone-users-worldwide/ (visited on 03/29/2018).

[2]   K. Panetta and A. Velosa. *Hype Cycle for the Internet of Things, 2016*. 2016. URL: https://www.gartner.com/smarterwithgartner/7-technologies-underpin-the-hype-cycle-for-the-internet-of-things-2016/ (visited on 03/25/2018).

[3]   M. Antonakakis, T. April, M. Bailey, M. Bernhard, E. Bursztein, J. Cochran, Z. Durumeric, J.A. Halderman, L. Invernizzi, M. Kallitsis, D. Kumar, C. Lever, Z. Ma, J. Mason, D. Menscher, C. Seaman, N. Sullivan, K. Thomas, and Y. Zhou. "Understanding the Mirai Botnet". In: *Proceedings of the 26th USENIX Security Symposium*. Aug. 2017, pp. 1093–1110.

[4]   L. Columbus. *2017 Roundup of Internet of Things Forecasts*. 2017. URL: https://www.forbes.com/sites/louiscolumbus/2017/12/10/2017-roundup-of-internet-of-things-forecasts/#1eec8a761480 (visited on 03/11/2018).

[5]   Navigant Research. "Advanced Energy 2017 Market Report, Global and U.S. Market Revenue 2011-2016 and Key Trends in Advanced Energy Growth". In: (2017).

[6]   F. Chraim, Y. B. Erol, and K. Pister. "Wireless Gas Leak Detection and Localization". In: *IEEE Transactions on Industrial Informatics* 12.2 (Apr. 2016), pp. 768–779. ISSN: 1551-3203. DOI: 10.1109/TII.2015.2397879.

[7]   K. Kalantar-Zadeh, K.J. Berean, N. Ha, A.F. Chrimes, K. Xu, D. Grando, J.Z. Ou, N. Pillai, J.L. Campbell, R. Brkljaca, K.M. Taylor, R.E. Burgell, C.K. Yao, S.A. Ward, C.S. McSweeney, J.G. Muir, and P.R. Gibson. "A human pilot trial of ingestible electronic capsules capable of sensing different gases in the gut". In: *Nature Electronics* 1 (Jan. 2018), pp. 79–87.

[8]   Y. Khan, A.E. Ostfeld, C.M. Lochner, A. Pierre, and A.C. Arias. "Monitoring of Vital Signs with Flexible and Wearable Medical Devices". In: *Advanced Materials* 28.22 (), pp. 4373–4395. DOI: 10.1002/adma.201504366.

[9]   A. Zanella, N. Bui, A. Castellani, L. Vangelista, and M. Zorzi. "Internet of Things for Smart Cities". In: *IEEE Internet of Things Journal* 1.1 (Feb. 2014), pp. 22–32. ISSN: 2327-4662. DOI: 10.1109/JIOT.2014.2306328.

[10] Global Semiconductor Alliance. *Wafer Fabrication & Assembly Pricing.* 2017. URL: https : / / www . gsaglobal . org / gsa - resources / reports / wafer - fabrication - assembly-pricing/ (visited on 03/12/2018).

[11] P. Grover and A. Sahai. "Shannon meets Tesla: Wireless information and power transfer". In: *2010 IEEE International Symposium on Information Theory.* June 2010, pp. 2363–2367. DOI: 10.1109/ISIT.2010.5513714.

[12] N. M. Pletcher, S. Gambini, and J. Rabaey. "A 52 μW Wake-Up Receiver With -72 dBm Sensitivity Using an Uncertain-IF Architecture". In: *IEEE Journal of Solid-State Circuits* 44.1 (Jan. 2009), pp. 269–280. ISSN: 0018-9200. DOI: 10.1109/JSSC.2008.2007438.

[13] C. Salazar, A. Cathelin, A. Kaiser, and J. Rabaey. "A 2.4 GHz Interferer-Resilient Wake-Up Receiver Using A Dual-IF Multi-Stage N-Path Architecture". In: *IEEE Journal of Solid-State Circuits* 51.9 (Sept. 2016), pp. 2091–2105. ISSN: 0018-9200. DOI: 10.1109/JSSC.2016.2582509.

[14] D. K. McCormick. "IEEE Technology Report on Wake-Up Radio: An Application, Market, and Technology Impact Analysis of Low-Power/Low-Latency 802.11 Wireless LAN Interfaces". In: *802.11ba Battery Life Improvement: IEEE Technology Report on Wake-Up Radio* (Nov. 2017), pp. 1–56. DOI: 10.1109/IEEESTD.2017.8055459.

[15] T. Watteyne, X. Vilajosana, B. Kerkez, F. Chraim, K. Weekly, Q. Wang, S. Glaser, and K. Pister. "OpenWSN: a standards-based low-power wireless development environment". In: *Transactions on Emerging Telecommunications Technologies* 23.5 (), pp. 480–493. DOI: 10.1002/ett.2558. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/ett.2558. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/ett.2558.

[16] J. Notor, A. Caviglia, and G. Levy. "CMOS RFIC architecture for IEEE 802.15.4 networks". In: *Cadence Design Systems, Inv. 41* (2003).

[17] O. Khan, B. Wheeler, D. Burnett, F. Maksimovic, S. Mesri, K. Pister, and A. Niknejad. "Frequency reference for crystal free radio". In: *2016 IEEE International Frequency Control Symposium (IFCS).* May 2016, pp. 1–2. DOI: 10.1109/FCS.2016.7546749.

[18] T. Watteyne, X. Vilajosana, B. Kerkez, F. Chraim, K. Weekly, Q. Wang, S. Glaser, and K. Pister. "OpenWSN: a standards based low power wireless development environment". In: *Transactions on Emerging Telecommunications Technologies* 23.5 (), pp. 480–493. DOI: 10.1002/ett.2558. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/ett.2558. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/ett.2558.

[19] N.J. Oh and S.G. Lee. "Building a 2.4-GHZ radio transceiver using IEEE 802.15.4". In: *IEEE Circuits and Devices Magazine* 21.6 (Jan. 2005), pp. 43–51. ISSN: 8755-3996. DOI: 10.1109/MCD.2005.1578587.

[20]   S. Aust, R. V. Prasad, and I. G. M. M. Niemegeers. "IEEE 802.11ah: Advantages in standards and further challenges for sub 1 GHz Wi-Fi". In: *2012 IEEE International Conference on Communications (ICC)*. June 2012, pp. 6885–6889. DOI: 10.1109/ICC.2012.6364903.

[21]   J. Im, H. S. Kim, and D. D. Wentzloff. "A 335 μW 72dBm receiver for FSK back-channel embedded in 5.8GHz Wi-Fi OFDM packets". In: *2017 IEEE Radio Frequency Integrated Circuits Symposium (RFIC)*. June 2017, pp. 176–179. DOI: 10.1109/RFIC.2017.7969046.

[22]   Dialog Semiconductor. "Bluetooth Low Energy 4.2 SoC". In: *DA14580 Data Sheet* (Nov. 2016).

[23]   Craig Schindler. private communication. June 2018.

[24]   B. Wheeler, F. Maksimovic, N. Baniasadi, S. Mesri, O. Khan, D. Burnett, A. Niknejad, and K. Pister. "Crystal-free narrow-band radios for low-cost IoT". In: *2017 IEEE Radio Frequency Integrated Circuits Symposium (RFIC)*. June 2017, pp. 228–231. DOI: 10.1109/RFIC.2017.7969059.

[25]   Y. H. Liu, C. Bachmann, X. Wang, Y. Zhang, A. Ba, B. Busze, M. Ding, P. Harpe, G. J. van Schaik, G. Selimis, H. Giesen, J. Gloudemans, A. Sbai, L. Huang, H. Kato, G. Dolmans, K. Philips, and H. de Groot. "13.2 A 3.7mW-RX 4.4mW-TX fully integrated Bluetooth Low-Energy/IEEE802.15.4/proprietary SoC with an ADPLL-based fast frequency offset compensation in 40nm CMOS". In: *2015 IEEE International Solid-State Circuits Conference - (ISSCC) Digest of Technical Papers*. Feb. 2015, pp. 1–3. DOI: 10.1109/ISSCC.2015.7063013.

[26]   M. Babaie, F. W. Kuo, H. N. R. Chen, L. C. Cho, C. P. Jou, F. L. Hsueh, M. Shah-mohammadi, and R. B. Staszewski. "A Fully Integrated Bluetooth Low-Energy Transmitter in 28 nm CMOS With 36% System Efficiency at 3 dBm". In: *IEEE Journal of Solid-State Circuits* 51.7 (July 2016), pp. 1547–1565. ISSN: 0018-9200. DOI: 10.1109/JSSC.2016.2551738.

[27]   S. O'Driscoll, S. Korhummel, P. Cong, Y. Zou, K. Sankaragomathi, J. Zhu, T. Deyle, A. Dastgheib, B. Lu, M. Tierney, J. Shao, C. Gutierrez, S. Jones, and H. Yao. "A 200 μm $\times$ 200 μm $\times$ 100 μm, 63nW, 2.4GHz injectable fully-monolithic wireless bio-sensing system". In: *2017 IEEE Radio Frequency Integrated Circuits Symposium (RFIC)*. June 2017, pp. 256–259. DOI: 10.1109/RFIC.2017.7969066.

[28]   L. J. Chu. "Physical Limitations of Omni-Direction Antennas". In: *Journal of Applied Physics* 19 (1948), pp. 1163–1175.

[29]   H. A. Wheeler. "Fundamental Limitations of Small Antennas". In: *Proceedings of the IRE* 35.12 (Dec. 1947), pp. 1479–1484. ISSN: 0096-8390. DOI: 10.1109/JRPROC.1947.226199.

[30] D. F. Sievenpiper, D. C. Dawson, M. M. Jacob, T. Kanar, S. Kim, J. Long, and R. G. Quarfoth. "Experimental Validation of Performance Limits and Design Guidelines for Small Antennas". In: *IEEE Transactions on Antennas and Propagation* 60.1 (Jan. 2012), pp. 8–19. ISSN: 0018-926X. DOI: 10.1109/TAP.2011.2167938.

[31] H. M. Cheema and A. Shamim. "The last barrier: on-chip antennas". In: *IEEE Microwave Magazine* 14.1 (Jan. 2013), pp. 79–91. ISSN: 1527-3342. DOI: 10.1109/MMM.2012.2226542.

[32] D. B. Leeson. "A simple model of feedback oscillator noise spectrum". In: *Proceedings of the IEEE* 54.2 (Feb. 1966), pp. 329–330. ISSN: 0018-9219. DOI: 10.1109/PROC.1966.4682.

[33] Yu Cao, R. A. Groves, Xuejue Huang, N. D. Zamdmer, J. O. Plouchart, R. A. Wachnik, Tsu-Jae King, and Chenming Hu. "Frequency-independent equivalent-circuit model for on-chip spiral inductors". In: *IEEE Journal of Solid-State Circuits* 38.3 (Mar. 2003), pp. 419–426. ISSN: 0018-9200. DOI: 10.1109/JSSC.2002.808285.

[34] Thomas H. Lee. *Planar Microwave Engineering. A Practical Guide to Theory, Measurements and Circuits.* Cambridge University Press, 2004.

[35] C. P. Yue and S. S. Wong. "On-chip spiral inductors with patterned ground shields for Si-based RF ICs". In: *IEEE Journal of Solid-State Circuits* 33.5 (May 1998), pp. 743–752. ISSN: 0018-9200. DOI: 10.1109/4.668989.

[36] A.D. Berny, A.M. Niknejad, and R.G. Meyer. "A 1.8-GHz LC VCO With 1.3-GHz Tuning Range and Digital Amplitude Calibration". In: *IEEE Journal of Solid-State Circuits* 40.4 (Apr. 2005), pp. 909–916. ISSN: 0018-9200. DOI: 10.1109/JSSC.1975.1050582.

[37] M. Garampazzi, P. M. Mendes, N. Codega, D. Manstretta, and R. Castello. "Analysis and Design of a 195.6 dBc/Hz Peak FoM P-N Class-B Oscillator With Transformer-Based Tail Filtering". In: *IEEE Journal of Solid-State Circuits* 50.7 (July 2015), pp. 1657–1668. ISSN: 0018-9200. DOI: 10.1109/JSSC.2015.2413851.

[38] M. Shahmohammadi, M. Babaie, and R. B. Staszewski. "A 1/f Noise Upconversion Reduction Technique for Voltage-Biased RF CMOS Oscillators". In: *IEEE Journal of Solid-State Circuits* 51.11 (Nov. 2016), pp. 2610–2624. ISSN: 0018-9200. DOI: 10.1109/JSSC.2016.2602214.

[39] A. Mazzanti and P. Andreani. "Class-C Harmonic CMOS VCOs, With a General Result on Phase Noise". In: *IEEE Journal of Solid-State Circuits* 43.12 (Dec. 2008), pp. 2716–2729. ISSN: 0018-9200. DOI: 10.1109/JSSC.2008.2004867.

[40] L. Fanori and P. Andreani. "Class-D CMOS Oscillators". In: *IEEE Journal of Solid-State Circuits* 48.12 (Dec. 2013), pp. 3105–3119. ISSN: 0018-9200. DOI: 10.1109/JSSC.2013.2271531.

[41] C. Li, M. Yuan, C. Chang, Y. Lin, C. Liao, K. Hsieh, M. Chen, and R. B. Staszewski. "19.6 A 0.2V trifilar-coil DCO with DC-DC converter in 16nm FinFET CMOS with 188dB FOM, 1.3kHz resolution, and frequency pushing of 38MHz/V for energy harvesting applications". In: *2017 IEEE International Solid-State Circuits Conference (ISSCC)*. Feb. 2017, pp. 332–333. DOI: 10.1109/ISSCC.2017.7870396.

[42] S. Levantino, C. Samori, A. Bonfanti, S. L. J. Gierkink, A. L. Lacaita, and V. Boccuzzi. "Frequency dependence on bias current in 5 GHz CMOS VCOs: impact on tuning range and flicker noise upconversion". In: *IEEE Journal of Solid-State Circuits* 37.8 (Aug. 2002), pp. 1003–1011. ISSN: 0018-9200. DOI: 10.1109/JSSC.2002.800969.

[43] P. Andreani, Xiaoyan Wang, L. Vandi, and A. Fard. "A study of phase noise in colpitts and LC-tank CMOS oscillators". In: *IEEE Journal of Solid-State Circuits* 40.5 (May 2005), pp. 1107–1118. ISSN: 0018-9200. DOI: 10.1109/JSSC.2005.845991.

[44] A. Hajimiri and T. H. Lee. "A general theory of phase noise in electrical oscillators". In: *IEEE Journal of Solid-State Circuits* 33.2 (Feb. 1998), pp. 179–194. ISSN: 0018-9200. DOI: 10.1109/4.658619.

[45] E. Alon, V. Stojanovic, and M. A. Horowitz. "Circuits and techniques for high-resolution measurement of on-chip power supply noise". In: *IEEE Journal of Solid-State Circuits* 40.4 (Apr. 2005), pp. 820–828. ISSN: 0018-9200. DOI: 10.1109/JSSC.2004.842853.

[46] F. Behbahani, Y. Kishigami, J. Leete, and A. A. Abidi. "CMOS mixers and polyphase filters for large image rejection". In: *IEEE Journal of Solid-State Circuits* 36.6 (June 2001), pp. 873–887. ISSN: 0018-9200. DOI: 10.1109/4.924850.

[47] B. Razavi. "Design considerations for direct-conversion receivers". In: *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing* 44.6 (June 1997), pp. 428–435. ISSN: 1057-7130. DOI: 10.1109/82.592569.

[48] B. Razavi, K. F. Lee, and R. H. Yan. "Design of high-speed, low-power frequency dividers and phase-locked loops in deep submicron CMOS". In: *IEEE Journal of Solid-State Circuits* 30.2 (Feb. 1995), pp. 101–109. ISSN: 0018-9200. DOI: 10.1109/4.341736.

[49] Behzad Razavi. *RF Microelectronics. A Practical Guide to Theory, Measurements and Circuits*. Pearson Education Inc., 2012.

[50] D. Chowdhury, L. Ye, E. Alon, and A. M. Niknejad. "An Efficient Mixed-Signal 2.4-GHz Polar Power Amplifier in 65-nm CMOS Technology". In: *IEEE Journal of Solid-State Circuits* 46.8 (Aug. 2011), pp. 1796–1809. ISSN: 0018-9200. DOI: 10.1109/JSSC.2011.2155790.

[51] B. W. Cook, A. Berny, A. Molnar, S. Lanzisera, and K. S. J. Pister. "Low-Power 2.4-GHz Transceiver With Passive RX Front-End and 400-mV Supply". In: *IEEE Journal of Solid-State Circuits* 41.12 (Dec. 2006), pp. 2757–2766. ISSN: 0018-9200. DOI: 10.1109/JSSC.2006.884801.

[52] K. Natarajan, J. S. Walling, and D. J. Allstot. "A class-C power amplifier/antenna interface for wireless sensor applications". In: *2011 IEEE Radio Frequency Integrated Circuits Symposium.* June 2011, pp. 1–4. DOI: 10.1109/RFIC.2011.5940713.

[53] N. O. Sokal and A. D. Sokal. "Class E-A new class of high-efficiency tuned single-ended switching power amplifiers". In: *IEEE Journal of Solid-State Circuits* 10.3 (June 1975), pp. 168–176. ISSN: 0018-9200. DOI: 10.1109/JSSC.1975.1050582.

[54] R. Meshkin, A. Saberkari, and M. Niaboli-Guilani. "A novel 2.4 GHz CMOS class-E power amplifier with efficient power control for wireless communications". In: *2010 17th IEEE International Conference on Electronics, Circuits and Systems.* Dec. 2010, pp. 599–602. DOI: 10.1109/ICECS.2010.5724583.

[55] J. Prummel, M. Papamichail, J. Willms, R. Todi, W. Aartsen, W. Kruiskamp, J. Haanstra, E. Opbroek, S. Rievers, P. Seesink, J. van Gorsel, H. Woering, and C. Smit. "A 10 mW Bluetooth Low-Energy Transceiver With On-Chip Matching". In: *IEEE Journal of Solid-State Circuits* 50.12 (Dec. 2015), pp. 3077–3088. ISSN: 0018-9200. DOI: 10.1109/JSSC.2015.2469674.

[56] T. Sano, M. Mizokami, H. Matsui, K. Ueda, K. Shibata, K. Toyota, T. Saitou, H. Sato, K. Yahagi, and Y. Hayashi. "13.4 A 6.3mW BLE transceiver embedded RX image-rejection filter and TX harmonic-suppression filter reusing on-chip matching network". In: *2015 IEEE International Solid-State Circuits Conference - (ISSCC) Digest of Technical Papers.* Feb. 2015, pp. 1–3. DOI: 10.1109/ISSCC.2015.7063015.

[57] L. Fanori, A. Liscidini, and R. Castello. "Capacitive Degeneration in LC-Tank Oscillator for DCO Fine-Frequency Tuning". In: *IEEE Journal of Solid-State Circuits* 45.12 (Dec. 2010), pp. 2737–2745. ISSN: 0018-9200. DOI: 10.1109/JSSC.2010.2077190.

[58] Dust Networks. "LTC5800-IPM SmartMesh IP Node 2.4GHz 802.15.4e Wireless Mote-on-Chip". In: *LTC5800-IPM Data Sheet* ().

[59] S. Drago, D. M. W. Leenaerts, F. Sebastiano, L. J. Breems, K. A. A. Makinwa, and B. Nauta. "A 2.4GHz 830pJ/bit duty-cycled wake-up receiver with -82dBm sensitivity for crystal-less wireless sensor nodes". In: *2010 IEEE International Solid-State Circuits Conference - (ISSCC).* Feb. 2010, pp. 224–225. DOI: 10.1109/ISSCC.2010.5433955.

[60] B. Wheeler. "Low Power, Crystal-Free Design for Monolithic Receivers". MA thesis. EECS Department, University of California, Berkeley, Dec. 2018.

[61] X. Chen, J. Breiholz, F. Yahya, C. Lukas, H.S. Kim, B. Calhoun, and D. Wentzloff. "A 486 μW All-Digital Bluetooth Low Energy Transmitter with Ring Oscillator Based ADPLL for IoT Applications". In: *RFIC 2018.* (June 2018).

[62] Z.Z. Chen and T.C. Lee. "The Design and Analysis of Dual-Delay-Path Ring Oscillators". In: *IEEE Transactions on Circuits and Systems I: Regular Papers* 58.3 (Mar. 2011), pp. 470–478. ISSN: 0018-9200. DOI: 10.1109/JSSC.1975.1050582.

[63] D. Burnett, B. Wheeler, F. Maksimovic, O. Khan, A.M. Niknejad, and K. Pister. "Narrowband Communication with Free Running 2.4 GHz Ring Oscillators". In: *2017 International Conference on Performance Evaluation and Modeling in Wired and Wireless Networks (PEMWN)*. Nov. 2017. DOI: 10.1109/ISSCC.2017.7870396.

[64] A. Molnar, B. Lu, S. Lanzisera, B. W. Cook, and K. S. J. Pister. "An ultra-low power 900 MHz RF transceiver for wireless sensor networks". In: *Proceedings of the IEEE 2004 Custom Integrated Circuits Conference (IEEE Cat. No.04CH37571)*. Oct. 2004, pp. 401–404. DOI: 10.1109/CICC.2004.1358833.

[65] E. Alpman, A. Khairi, R. Dorrance, M. Park, V. S. Somayazulu, J. R. Foerster, A. Ravi, J. Paramesh, and S. Pellerano. "802.11g/n Compliant Fully Integrated Wake-Up Receiver With 72-dBm Sensitivity in 14-nm FinFET CMOS". In: 53.5 (May 2018), pp. 1411–1422. ISSN: 0018-9200. DOI: 10.1109/JSSC.2018.2817603.

[66] S. Mesri. "Design and User Guide for the Single Chip Mote Digital System". MA thesis. EECS Department, University of California, Berkeley, May 2016. URL: http://www2.eecs.berkeley.edu/Pubs/TechRpts/2016/EECS-2016-71.html.

[67] H. Banba, H. Shiga, A. Umezawa, T. Miyaba, T. Tanzawa, S. Atsumi, and K. Sakui. "A CMOS bandgap reference circuit with sub-1-V operation". In: *IEEE Journal of Solid-State Circuits* 34.5 (May 1999), pp. 670–674. ISSN: 0018-9200. DOI: 10.1109/4.760378.

[68] M. Seok, G. Kim, D. Blaauw, and D. Sylvester. "A Portable 2-Transistor Picowatt Temperature-Compensated Voltage Reference Operating at 0.5 V". In: *IEEE Journal of Solid-State Circuits* 47.10 (Oct. 2012), pp. 2534–2545. ISSN: 0018-9200. DOI: 10.1109/JSSC.2012.2206683.

[69] H. Liu, Z. Sun, D. Tang, H. Huang, T. Kaneko, W. Deng, R. Wu, K. Okada, and A. Matsuzawa. "An ADPLL-centric bluetooth low-energy transceiver with 2.3mW interference-tolerant hybrid-loop receiver and 2.9mW single-point polar transmitter in 65nm CMOS". In: *2018 IEEE International Solid - State Circuits Conference - (ISSCC)*. Feb. 2018, pp. 444–446. DOI: 10.1109/ISSCC.2018.8310375.

[70] M. Ding, X. Wang, P. Zhang, Y. He, S. Traferro, K. Shibata, M. Song, H. Korpela, K. Ueda, Y. H. Liu, C. Bachmann, and K. Philips. "A 0.8V 0.8mm2 bluetooth 5/BLE digital-intensive transceiver with a 2.3mW phase-tracking RX utilizing a hybrid loop filter for interference resilience in 40nm CMOS". In: *2018 IEEE International Solid - State Circuits Conference - (ISSCC)*. Feb. 2018, pp. 446–448. DOI: 10.1109/ISSCC.2018.8310376.

[71] J. Yin, S. Yang, H. Yi, W. H. Yu, P. I. Mak, and R. P. Martins. "A 0.2V energy-harvesting BLE transmitter with a micropower manager achieving 25% system efficiency at 0dBm output and 5.2nW sleep power in 28nm CMOS". In: *2018 IEEE International Solid - State Circuits Conference - (ISSCC)*. Feb. 2018, pp. 450–452. DOI: 10.1109/ISSCC.2018.8310378.

[72] C.C. Lim, J. Yin, P.I. Mak, H. Ramiah, and R.P. Martins. "An Inverse-Class-F CMOS VCO with Intrinsic-High-Q 1st and 2nd Harmonic Resonances for 1/f2 to 1/f3 Phase-Noise Suppression Achieving 196.2 dBc/Hz FOM". In: *2018 IEEE International Solid-State Circuits Conference - (ISSCC)*. Feb. 2018, pp. 374–375. DOI: 10.1109/ISSCC.2010.5433955.

[73] Julian Hagedorn. "Crystal Oscillator Performance of the CDCLVC1310. Application Report SCAA119". In: (Sept. 2012).

[74] S. Jeong, Z. Foo, Y. Lee, J. Y. Sim, D. Blaauw, and D. Sylvester. "A Fully-Integrated 71 nW CMOS Temperature Sensor for Low Power Wireless Sensor Nodes". In: *IEEE Journal of Solid-State Circuits* 49.8 (Aug. 2014), pp. 1682–1693. ISSN: 0018-9200. DOI: 10.1109/JSSC.2014.2325574.

[75] 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". In: *2014 IEEE International Solid-State Circuits Conference Digest of Technical Papers (ISSCC)*. Feb. 2014, pp. 300–301. DOI: 10.1109/ISSCC.2014.6757443.

[76] D. Stanislowski, X. Vilajosana, Q. Wang, T. Watteyne, and K. S. J. Pister. "Adaptive Synchronization in IEEE802.15.4e Networks". In: *IEEE Transactions on Industrial Informatics* 10.1 (Feb. 2014), pp. 795–802. ISSN: 1551-3203. DOI: 10.1109/TII.2013.2255062.

[77] O.U. Khan, D. Burnett, F. Maksimovic, B. Wheeler, S. Mesri, A. Sundarajan, B.L. Zhou, A. Niknejad, and K. Pister. "Time Keeping Ability of Crystal Free Oscillators". In: *IEEE IoT Journal in submission* (2018).

[78] *Bluetooth Low Energy Beacons*. SWRA475A. Revised Oct. 2016. Texas Instruments. Jan. 2015.

[79] *Bluetooth Core Specification*. v5.0. Bluetooth SiG. Dec. 2016.

[80] M. Tabesh, M. Rangwala, A. M. Niknejad, and A. Arbabian. "A power-harvesting padless mm-sized 24/60GHz passive radio with on-chip antennas". In: *2014 Symposium on VLSI Circuits Digest of Technical Papers*. June 2014, pp. 1–2. DOI: 10.1109/VLSIC.2014.6858380.

[81] P. Moser, W. Hofmann, P. Wolf, J. A. Lott, G. Larisch, A. Payusov, N. N. Ledentsov, and D. Bimberg. "81 fJ/bit energy-to-data ratio of 850 nm vertical-cavity surface-emitting lasers for optical interconnects". In: *Applied Physics Letters* 98.23 (2011), p. 231106. DOI: 10.1063/1.3597799. eprint: https://doi.org/10.1063/1.3597799. URL: https://doi.org/10.1063/1.3597799.

[82] Constantine A. Balanis. *Antenna Theory. Analysis and Design*. John Wiley and Sons, 1982.

[83] Matthew N.O. Sadiku. *Elements of Electromagnetics. 5th edition*. Oxford University Press, 2010.

[84] C. Andrews and A. C. Molnar. "Implications of Passive Mixer Transparency for Impedance Matching and Noise Figure in Passive Mixer-First Receivers". In: *IEEE Transactions on Circuits and Systems I: Regular Papers* 57.12 (Dec. 2010), pp. 3092–3103. ISSN: 1549-8328. DOI: 10.1109/TCSI.2010.2052513.

[85] Y. Liu and W. A. Serdijn. "Analysis and Design of a Passive Receiver Front-End Using an Inductive Antenna Impedance". In: *IEEE Transactions on Circuits and Systems I: Regular Papers* 65.2 (Feb. 2018), pp. 733–744. ISSN: 1549-8328. DOI: 10.1109/TCSI.2017.2734851.

[86] M. Stoopman, Y. Liu, H. J. Visser, K. Philips, and W. A. Serdijn. "Codesign of Electrically Short Antenna Electronics Interfaces in the Receiving Mode". In: *IEEE Transactions on Circuits and Systems II: Express Briefs* 62.7 (July 2015), pp. 711–715. ISSN: 1549-7747. DOI: 10.1109/TCSII.2015.2406371.

[87] A. R. Shahani, D. K. Shaeffer, and T. H. Lee. "A 12-mW wide dynamic range CMOS front-end for a portable GPS receiver". In: *IEEE Journal of Solid-State Circuits* 32.12 (Dec. 1997), pp. 2061–2070. ISSN: 0018-9200. DOI: 10.1109/4.643664.

[88] Y. Su, J. Jau Lin, and K.K. O. "A 20 GHz CMOS RF down-converter with an on-chip antenna". In: *ISSCC. 2005 IEEE International Digest of Technical Papers. Solid-State Circuits Conference, 2005.* Feb. 2005, 270–597 Vol. 1. DOI: 10.1109/ISSCC.2005.1493973.

[89] Y. Lien, E. Klumperink, B. Tenbroek, J. Strange, and B. Nauta. "A mixer-first receiver with enhanced selectivity by capacitive positive feedback achieving +39dBm IIP3 and 3dB noise figure for SAW-less LTE Radio". In: *2017 IEEE Radio Frequency Integrated Circuits Symposium (RFIC)*. June 2017, pp. 280–283. DOI: 10.1109/RFIC.2017.7969072.

[90] H. P. Le, S. R. Sanders, and E. Alon. "Design Techniques for Fully Integrated Switched-Capacitor DC-DC Converters". In: *IEEE Journal of Solid-State Circuits* 46.9 (Sept. 2011), pp. 2120–2131. ISSN: 0018-9200. DOI: 10.1109/JSSC.2011.2159054.

[91] V. N. Swamy, S. Suri, P. Rigge, M. Weiner, G. Ranade, A. Sahai, and B. Nikolic. "Cooperative communication for high-reliability low-latency wireless control". In: *2015 IEEE International Conference on Communications (ICC)*. June 2015, pp. 4380–4386. DOI: 10.1109/ICC.2015.7249012.

[92] B. Kilberg, C. B. Schindler, A. Sundarajan, A. Yang, and K.S.J. Pister. "Experimental Evaluation of Low-Latency Diversity Modes in IEEE 802.15.4 Networks". In: *2018 IEEE International Conference on Emerging Technologies and Factory Automation.* Sept. 2018.

[93] Y. A. Li, M. Mar, B. Nikolic, and A. M. Niknejad. "On-chip spur and phase noise cancellation techniques". In: *2017 IEEE Asian Solid-State Circuits Conference (A-SSCC)*. Nov. 2017, pp. 109–112. DOI: 10.1109/ASSCC.2017.8240228.

Appendix A

SCM v3 Documentation

In this appendix I will describe software that I wrote for the single chip mote and operation of the scan chain for the local oscillator, transmitter, and divider portions of the chip. The software is available on the git. It is highly encouraged to read [66] for more information about setting up the Keil environment. For the sake of brevity, this assumes that you have installed: the Keil compiler (for ARM programming), an Arduino environment (for programming the Teensy microcontroller - used for scan chain programming, three wire bus (3WB) bootloading, and for serial communication to the chip, which also requires a UART-to-RS232 translator chip), and the Matlab or Python scan and 3WB programming scripts. UART is only needed for debugging. The vast majority of information here is for the single chip mote v3. Modifications made to v3b are explicitly mentioned, although exact procedures for programming and debugging v3b are not included. It is impossible to predict what gremlins lurk in a chip that has not returned from the fab at the time of writing.

A.1   Chip diagrams

A Cadence screenshot of SCM v3, annotated with relevant pad labels, is shown in Fig. A.1.

Some of the pin annotations are abbreviated from their names in the Cadence layout view. For SCM v3, the relevant pad connections are:

- **VBAT** supply voltage. The radio was designed to run between 1.2 V and 1.5 V. It is theoretically possible to run at higher voltages for brief periods of time.

- **ASC_Phi1** first ASC clock. The scan chain is a two-clock shift register.

- **ASC_Phi2** second ASC clock.

- **ASC_load** active high - when enabled, this loads the clocked bits into the registers that drive the output of the scan chain.
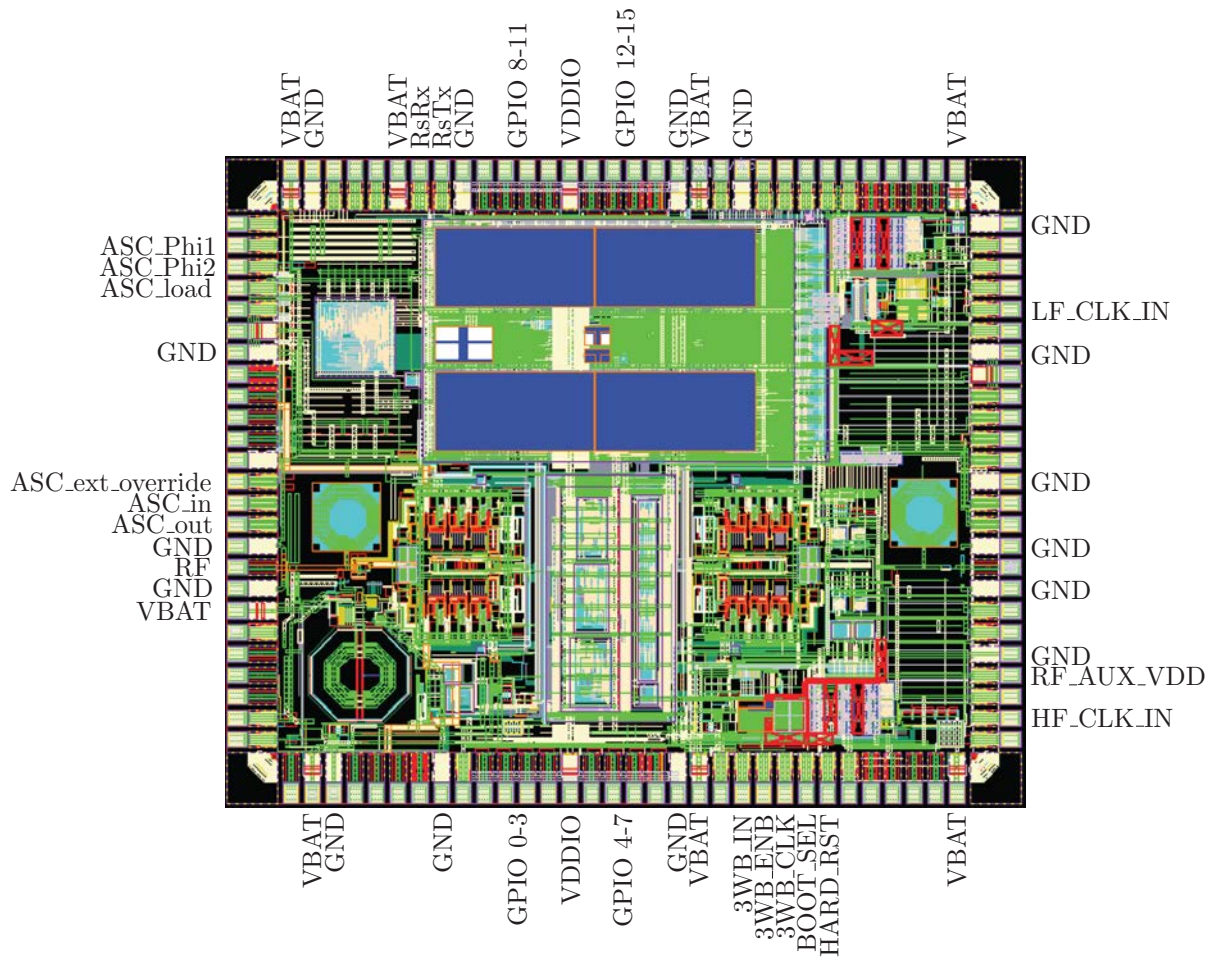
Figure A.1: Pad diagram of the single chip mote v3

- **ASC_ext_override** Set to 0 if scan chain is being programmed by the Cortex. Set to 1 if scan chain is being programmed externally (from the pads).

- **ASC_in** input to the scan chain

- **ASC_out** output of the scan chain. Not strictly necessary, but useful for debugging.

- **RF** antenna connection

- **GPIO** labeled from left to right in each one of the chunks. These are not strictly necessary, but are useful both for debugging and necessary for interfacing with external sensors/actuators.

- **VDDIO** GPIO supply voltage. Has been tested at 1.5V and at 3.3V.

- `3WB_IN` 3WB programmer input

- `3WB_ENB` 3WB enable (active low)

- `3WB_CLK` 3WB programmer clock

- `BOOT_SEL` set to 0 (default) to bootload from the on-chip optical programmer. This is not operational on SCM v3. It should work on SCM v3b. Set to 1 to bootload from off chip via 3WB. So, set it to 1.

- `HF_CLK_IN` this is an input to an on-chip counter that is necessary for estimating the LC tank frequency (the on-chip connection was not functional).

- `RF_AUX_VDD` buffers that need to be powered. Set to 0.8 V.

- `LF_CLK_EXT` external Cortex M0 clock. The on-chip clock does not work under default conditions, so an external clock should be supplied. This can come from a variety of sources, either the Teensy, the divided LC tank, or one of the 2 MHz clocks.

- `RsTx and RsRx` UART debugger. Can connect to a computer if a UART-to-RS232 translating chip is used.

Many of these pins are only necessary due to various bugs/errors made on SCM v3. If all of these issues are resolved on SCM v3b, the only necessary connections are VBAT, GND, and RF. Programming will be performed optically, and the scan chain will be programmed by the Cortex M0.

A.2 Programming Procedure

This assumes that a bin file has already been generated to run code on the Cortex. First, run the matlab code scm3_ASC_func_v2, which should be available in the git repo. The inputs to this function are shown in Fig. A.2.

Descriptions of what these inputs do is given in the Scan Chain section of this Appendix. For unknown reasons, on powerup, the scan chain needs to be programmed twice. Sometimes. After programming the scan chain, run the script called matlab_3wb_inft_v1. Ensure that the binfile path points to the file generated by the Keil compiler. Keil does not work on macs. To write code, I have successfully used VMWare Fusion (which should be available to Berkeley EECS students for free).

Even if the eventual intent is to program the scan chain registers from the Cortex (which is recommended), it is necessary to program the scan chain externally to enable clocking the Cortex from off-chip. After this is done, scan chain control is returned to the Cortex. There are three functions: initialize_ASC (sets the default scan configuration for start-up in a global array), analog_scan_chain_write (clocks in the bits), and analog_scan_chain_load (programs the scan chain). The Matlab scripts automatically perform the necessary toggling of external pins on all SCM v3 PCBs that we have made.

```
function [outvalid] = scm3_ASC_func_v2( lo_enable,...
                                        pa_enable,...
                                        div_enable,...
                                        lo_fine_tune,...
                                        lo_mid_tune,...
                                        lo_coarse_tune,...
                                        lo_current,...
                                        pa_panic,...
                                        pa_ldo,...
                                        lo_panic,...
                                        lo_ldo,...
                                        div_panic,...
                                        div_ldo,...
                                        sel_12,...
                                        en_backup_2,...
                                        en_backup_5,...
                                        pre_dyn_sel, ...
                                        en_64mhz,...
                                        en_20mhz,...
                                        static_code,...
                                        static_en,...
                                        bg_panic,...
                                        bg)
```

Figure A.2: Inputs to the Matlab scan function

A.3   Scan Chain

In this section I will attempt to describe the functionality of the chain bits. The scan chain is nominally controlled by functions from the Cortex M0, so I will not go into excessive detail regarding the exact bits in the array that need to be changed to perform certain functions. Rather, I will describe them in chunks.

- `fine_code` 5 bit control of the LC tank's fine tuning DAC f0, ... ,f4, fd

- `mid_code` 5 bit control of the LC tank's mid tuning DAC m0, ... m4, md

- `coarse_code` 5 bit control of the LC tank's coarse tuning DAC c0, ... c4, cd

- `lo_tune_select` set to 0 for LC control from the Cortex `analog_cfg`, 1 for scan

- `polyphase_enable` set to 0 to disable the polyphase filter. 1 to enable.

- `lo_current_tune` 8 bit control of LC tank's current between 180 μA and 800μA LSB to MSB.

- `test_bg` 7 bit control of the test band gap (connected to pad) from 0.75 V to 0.85 V if MSB = 0, and from 1.05 V to 1.12 V if MSB = 1 (panic).

- `pa_ldo_rdac` see `test_bg`, controls power amplifier supply.

- `lo_ldo_rdac` see `test_bg`, controls LO supply.

- `div_ldo_rdac` see `test_bg`, controls divider supply.

- `mod_logic` controls the source of modulation for the 802.15.4 capacitor. The MSB determines whether the modulation is inverted or not (1 for inversion). The next bit selects cortex or pad modulation (0 for cortex, 1 for pad). The next two bits can be used to test the modulation, setting the control bit to VDD or to ground. Note: for 802.15.4 modulation, the bits should be set to 1000. For BLE modulation from pad using the 802.15.4 capacitor, it should be set to 0111. A schematic of this control is shown in Fig. A.3.

- `mod_15_4_tune` tunes the frequency spacing of the 802.15.4 modulation capacitor. 4 bits, but the LSB is a dummy (not binary weighted).

- `sel_1mhz_2mhz` 0 uses the x2 XOR multiplier. 1 is pass throgh.

- `pre_2_backup_en` enables the static flip flop div-by-2 prescaler.

- `pre_5_backup_en` enables the static flip flop div-by-5 prescaler. This is recommended for standard operation.

- `pre_dyn` is a 3-bit, one-hot selection of three different injection-lock TSPC-esque pre-scalers. It needs to be inverted (so all 1s will disable all three dynamic pre-scalers). The second of the three is the strongest, and will result in a div-by-2 for almost all LO and divider settings. The first can consistently give a div-by-5 for most settings. The third one is the weakest, and it is possible to divide by up to 7 using it.

- `div_64mhz_enable` enables a 64 MHz output frequency divider to clock the baseband stages and receiver ADC.

- `div_20mhz_enable` enables a 20 MHz output frequency divider to clock the BLE GFSK module. This is also the clock output that is connected to the LC counter.

- `div_static_code` sets the static divider ratio. This can also be controlled from the Cortex's analog_config, and generally should be. More on that later.

- `div_static_reset_b` active low reset of the static divider.

- `dyn_div_N` there is another, theoretically lower power, divider on the chip as well. It has never been tested.

- `div_tune_select` set to 0 to have the divider controlled from the cortex. Set to 1 for the divider to be controlled from scan chain.

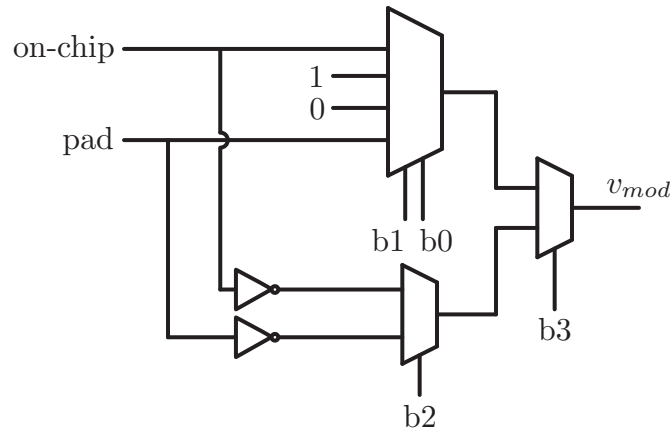- `BLE_module_settings` will be obfuscated in future versions of the chip.

Figure A.3: 802.15.4 Modulation Logic Schematic

A.4 Cortex Code

All of the cortex code is built into the SCM repo at

`https://repo.eecs.berkeley.edu/git/projects/pistergroup/scm-digital.git`

I will now describe some of the very low-level functions for directly controlling parts of the transmitter from the cortex. Start with the relevant config and rdata registers:

```
#define ACFG_DIV__ADDR   *(unsigned int *)(APB_ANALOG_CFG_BASE + 0x00140000)
#define ACFG_DIV__ADDR_2 *(unsigned int *)(APB_ANALOG_CFG_BASE + 0x00180000)
#define ACFG_LO_ADDR     *(unsigned int *)(APB_ANALOG_CFG_BASE + 0x001C0000)
#define ACFG_LO__ADDR_2  *(unsigned int *)(APB_ANALOG_CFG_BASE + 0x00200000)

#define ASYNC_FIFO__ADDR *(unsigned int *)(APB_ANALOG_CFG_BASE + 0x00680000)
```

`ACFG_DIV__ADDR` contains two control bits to set the RF divider divide ratio. `ACFG_DIV__ADDR_2` has the remaining control bits, as well as enable and reset signals for the divider. Quick note here: this divider struggles at low supply voltages, and will not work for odd divide ratios if the input frequency is high (around 1.2 GHz). The code to control these divider registers is called `digProgram(div_ratio, reset, enable)`. Reset is active low. A diagram of these two registers is shown in Fig. A.4. The pre-scaler must be enabled for this divider to have an output (see scan chain for details).

`ACFG_LO__ADDR_2` has the fine frequency control LSB and fine frequency control dummy bit. `ACFG_LO__ADDR` has the remaining control bits. The function `LC_FREQCHANGE(coarse, mid, fine)` controls these overlapping capacitor DACs. This function obfuscates the dummy bits. The function `LC_monotonic(LC_code, mid_divs, coarse_divs)` implements the func-

`ACFG_DIV_ADDR`

d11	d10														

`ACFG_DIV_ADDR_2`

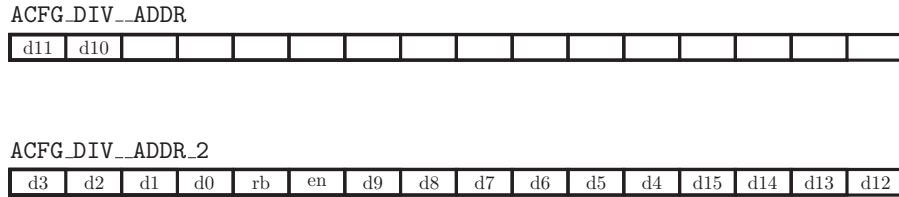| d3	d2	d1	d0	rb	en	d9	d8	d7	d6	d5	d4	d15	d14	d13	d12

Figure A.4: Divider registers, bit by bit

tion in Chapter 6. If you are tuning the LC oscillator, this you will probably use `LC_monotonic`. In any case, the diagram of the two LC control registers is shown in Fig. A.5. The bits labeled with a "d" are the dummy bits, and will change the frequency by approximately five LSBs of the given DAC.

`ACFG_LO_ADDR`

f1	f2	f3	f4	md	m0	m1	m2	m3	m4	cd	c0	c1	c2	c3	c4

`ACFG_LO_ADDR_2`

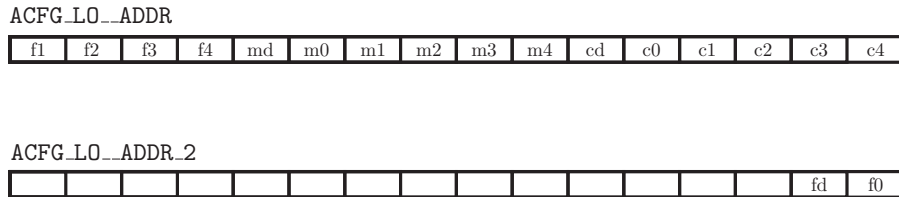|														fd	f0

Figure A.5: Oscillator frequency tune registers, bit by bit

There are a number of functions that control the scan chain, but they are relatively straightforward, especially considering the exposition given in the previous section. There are also a functions written to generate BLE advertising packets (`gen_ble_packet`) and to use the on-chip asynchronous FIFO to transmit BLE packets (`transmit_ble_packet`). At the moment, they are designed for transmitting 128 bit packets but they can easily be repurposed to transmit larger payloads. The transmit function in particular can be repurposed with completely arbitrary data for potential experimentation with 802.15.4 chipping sequences.

A.5   Common Configurations

This section describes the necessary scan/Cortex procedure to use the radio in various modes.

Receive Mode (RF only)

1. Enable the LC tank LDO and current source. This can be done either through GPIO (for fast start) or with the scan chain. Enabling the LC tank can be configured to start

automatically via the radio state machine on SCM v3b.

2. Set the LC tank current to an appropriate level (most people have used a current level of 127 - this appears to offer a good tradeoff between PA efficiency, LO current, and phase noise).

3. Enable the polyphase filter.

4. Program the LC tank so that it is 2.5 MHz ABOVE the receive channel frequency.

5. Enable the IF chain and digital baseband. This is documented elsewhere.

Transmit Mode - 802.15.4

1. Enable the LC tank LDO and current source.

2. Enable the PA LDO. Again, on SCM v3 this can be done with either GPIO (fast) or scan chain.

3. Ensure that the polyphase filter is disabled.

4. Tune the LC tank frequency to 500 kHz ABOVE the desired channel frequency.

5. If data is transmitted from the on-chip state machine, set the four `mod_logic` bits to 1000. If data is transmitted from a pad, set `mod_logic` to 1000. The bit-bang direct modulation is inverted from the bits coming from the state machine.

6. Ensure that the transmit clock (source can be chosen) is within 40 ppm of 2 MHz.

7. Use instructions in [66] to load and transmit a packet.

Transmit Mode - BLE

1. Enable the LC tank LDO and current source.

2. Enable the PA LDO.

3. Enable the divider LDO. The divider does not need to run, BUT the data buffers run off of the divider supply.

4. Ensure that the polyphase filter is disabled.

5. Tune the LC tank frequency to approximately 250 kHz BELOW the desired channel frequency.

6. Disable the 802.15.4 DAC by setting `mod_logic` to 0010 or 0001.

7. If data is transmitted from a pad, set the `mod_logic` bits to 0010. If data is transmitted from the Cortex (requires some trickery here: the Cortex clock needs to be an exact integer multiple of 1 MHz, and bits go straight from memory mapped IO to the LC tank. This also requires assembly code. Fortunately Keil allows in-line assembly).

8. Bypass both the FIFO and the GFSK module.

9. Run the BLE transmitting assembly code.

It is also possible to use the 802.15.4 capacitor DAC to transmit BLE. However, this is only possible from off-chip. The only DAC that can be controlled directly from the Cortex is the set of 500 kHz BLE capacitors. On SCM v3b, this procedure will be significantly different for two reasons. First, there is a 2048-bit FIFO so that the Cortex does not need to run at an exact multiple of 1 MHz. Second, there is a multiplexer to select whether the data goes to the 802.15.4 capacitors or the BLE capacitors.