Monolithic RF Frontends for Ubiquitous Wireless Connectivity
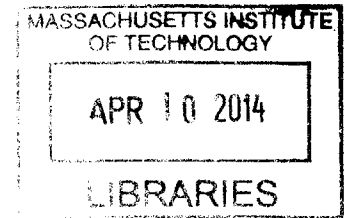
by

Sushmit Goswami

B. E., University of Delhi (2005)
M. S., Arizona State University (2007)

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Electrical Engineering and Computer Science

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

February 2014

Author . . . .
Department of Electrical Engineering and Computer Science
January 31, 2014

Certified by. .
Joel L. Dawson
Associate Professor of Electrical Engineering
Thesis Supervisor

Certified by . . . .
Hae-Seung Lee
Professor of Electrical Engineering
Thesis Supervisor

Accepted by . . . .
Leslie A. Kolodziejski
Professor and Chair, Department Committee on Graduate Students

Monolithic RF Frontends for Ubiquitous Wireless Connectivity

by

Sushmit Goswami

Submitted to the Department of Electrical Engineering and Computer Science
on January 31, 2014, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Electrical Engineering and Computer Science

Abstract

The desire for ubiquitous connectivity is pushing radios towards highly-integrated, multi-standard and multi-band implementations. This thesis explores architectures for next-generation RF frontends, which form the interface between the RF transceiver and antenna. RF frontend performance has important implications for the energy efficiency, frequency range and sensitivity of the radio.

Ubiquitous connectivity requires bringing online previously unconnected, closed-circuit systems. Case in point, the recently ratified 802.11p standard targets wireless access in vehicular environments. The first part of this thesis presents an RF frontend for 802.11p applications. Gallium Nitride is used as an enabling technology platform for monolithic integration of high-power RF functions. A number of architectural techniques are proposed to enhance energy efficiency.

Even in relatively mature use cases like smartphones, significant evolution is needed to address future needs. Emerging wireless standards specify dozens of bands covering several octaves for worldwide connectivity, which need to be supported with a single device. However, in current multi-band radio implementations, significant redundancy is still the norm in the RF frontend. In the second part of this thesis, an improved architecture for multi-band, time-division duplexed radios is introduced, which replaces multiple narrowband frontends with a frequency-agile solution, tunable over a wide frequency range. A highly digital architecture is adopted, leading to a fully integrated solution wherein both efficiency and achievable frequency range benefit from CMOS scaling.

Thesis Supervisor: Joel L. Dawson
Title: Associate Professor of Electrical Engineering

Thesis Supervisor: Hae-Seung Lee
Title: Professor of Electrical Engineering

Acknowledgements

I must begin by thanking my parents and sister for their unconditional love and support for all my endeavors. Over the years, they have sacrificed much to ensure I had the best shot at all that is worthwhile. Any achievement, however big or small, is as much theirs as it is my own.

I am grateful to Prof. Joel Dawson for offering me a position in his group back in 2010, and thereby facilitating my transition from industry to MIT. From the very beginning, he was incredibly supportive of my ideas, and let me define my own research problems. I consider the opportunity to work with him a privilege and a truly enriching experience.

I thank Prof. Harry Lee for supporting me for the latter part of my journey at MIT, and letting me continue my research in RF. His expertise in ADC design is truly phenomenal. I also had the pleasure of serving as a teaching assistant for him and learned a lot from his experience as an instructor.

Prof. Anantha Chandrakasan's contribution to the completion of my studies extends far beyond just being on my committee. His unwavering commitment to supporting students in any way possible, regardless of group affiliation is a source of inspiration to me. He facilitated my transition from one research group to another, and generously allowed me to test in his lab despite the busy schedule of his own students.

Prof. Li-Shiuan Peh provided financial support to start a new project with NTU Singapore, for which I am grateful. At NTU Singapore, Dr. Pilsoon Choi served as an active collaborator with me. I thoroughly enjoyed working with him.

Prof. Charlie Sodini generously gave me a desk in his lab area without me having any affiliation with his group, for which I am grateful. I also thank him for free Red Sox tickets.

The Advanced Concepts Committee at MIT Lincoln Lab (LL) provided financial support for one of my projects. In particular, I would like to thank Dr. Helen Kim for championing the research proposal and seeing it through the committee. Helen

Contents

List of Figures

12

List of Tables

Chapter 1

Introduction

By some estimates [1], the wireless industry reached a critical juncture in 2013 in that the number of mobile, internet-connected devices exceeded the world's population. This explosive growth in devices has been possible due to the maturation of several enabling technologies, principal among them being the highly integrated radio. Even after decades of work, radio design continues to be an exciting area of research due to the continual evolution in existing use-cases and emergence of new paradigms, both of which bring challenges and opportunities in equal measure.

In the last few years, mobile internet has come of age. Connectivity on the move is no longer a luxury, but a base feature in most mobile devices sold today. The next generation of mobile devices (smartphones, tablets etc.) is expected to be faster, lighter and have longer battery life. Increased mobility of users, as evinced by declining PC sales and a fast-growing mobile market [2], also necessitates that devices offer uninterrupted connectivity irrespective of location, while sustaining high data-rates to support cloud-based services. Further, as the industry approaches 100 % penetration in developed countries, user growth in the future will be mainly driven by developing countries, which are home to the majority of the 5 billion people worldwide currently without internet access [3], putting significant price pressure on this market.

Within the present decade, it is anticipated that internet connectivity will also extend to entirely new classes of devices. Vehicles, parking meters, industrial sensor nodes, home thermostats, continuous health monitors etc. are all examples of pre-

viously closed-circuit systems that are expected to come online. This proliferation of connectivity has been termed the *internet of things* (IoT) [4] [5] [6]. While the IoT paradigm will rely on both wired and wireless connectivity, the vast majority of devices which come online will be wireless, computationally lean, battery or self powered and extremely compact. From the radio designer's perspective, design for next-generation applications puts forth the following challenges:

- Ubiquitous connectivity cannot become a reality at current price points. Due to increased price pressure in the mature but still growing mobile device market, and the emergence of an entirely new one (i.e IoT), further commoditization of radio hardware is needed. The monolithic (single-chip) radio has been envisioned for many years, and represents the ideal solution from a cost and form-factor standpoint. However, full integration in low-cost technology remains elusive.

- Since the vast majority of future wireless systems will be battery or self powered, with continually shrinking form factors, large reservoirs of energy (high-capacity batteries or super-capacitors) are not feasible. Therefore, radio systems need to be more energy-efficient in order to maximize up-time for a given amount of stored or harvested energy.

- The coexistence of an exponentially growing number of wireless devices requires the radio system to be more flexible. At any given time, the optimal frequency band for communication is a function of geographical location, presence of other similar coexisting devices, data-rate needs and available wireless infrastructure (access points, routers). Clearly, radios should support communication over wide frequency range to enable location agnostic or ubiquitous connectivity. Radio operation over a wide frequency range conflicts directly with the first two objectives, and therefore presents a major challenge.

1.1 Overview of state-of-the-art mobile radios

Figure 1-1 shows the block diagram of a typical mobile radio system. When partitioned by power level, there are two main subsystems:

1.1.1 RF transceiver

In this work, the term *RF transceiver* is used to denote the low-power section of the radio, which includes all the mixed-signal processing that occurs between digital baseband data (commonly represented with IQ pairs) and modulated RF signals. On the transmitter (TX) side, this involves digital-analog conversion, filtering, up-conversion to RF carrier frequency and some amplification. On the receiver (RX) side, the signal path performs down-conversion, filtering and analog-digital conversion. Since the transceiver only comprises low-power circuitry, design has benefitted tremendously from the scaling of CMOS technology. Increasingly, more and more functionality has been integrated, while power consumption has gone down. For instance, multi-standard transceivers have been demonstrated on a single-chip, both in the form of multiple dedicated sections [7] and also more ambitious software-defined transceivers [8], which can be reconfigured to work in one of several different bands at a time.

1.1.2 RF frontend

The RF frontend represents the high-power section of the radio. As shown in figure 1-1, a typical RF frontend consists of three key components:

1. **Power amplifier (PA):** The TX signal from the transceiver is not strong enough to drive the antenna directly. The PA forms the last amplifying stage in the TX signal path. For instance, WLAN requires a maximum output power of at least 0.5 W or 27 dBm, which corresponds to 14 V_{PP} across a 50 Ω load. In addition to the stringent output power specification, PA design is further complicated by use of linear modulation techniques in modern wireless systems to maximize data-rates. Generally, it is easy to get high RF PA linearity or
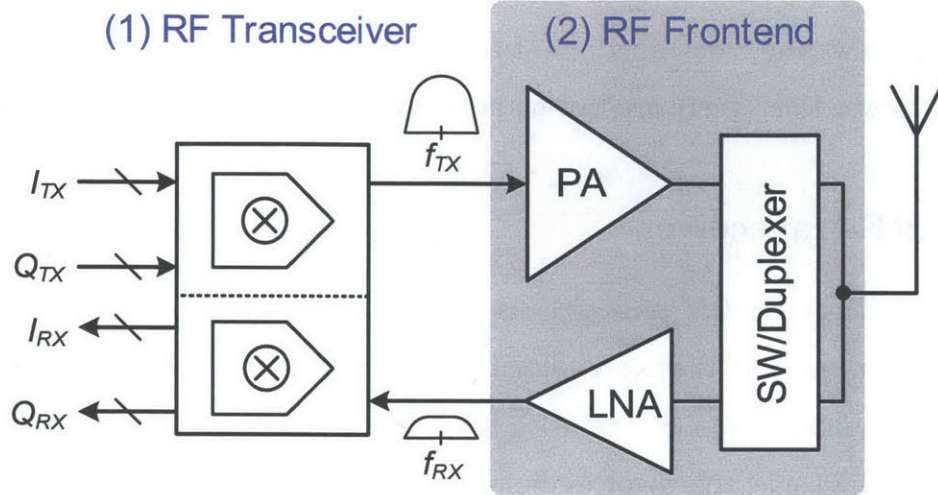
Figure 1-1: Key components of the RF frontend

efficiency independently, but very difficult to achieve them simultaneously [9]. Therefore, the PA is the most power hungry component of the radio and its efficiency largely determines the efficiency of the entire radio system. The choice of device technology plays a critical role in achievable performance. PA efficiency and therefore power consumption also suffers if wide bandwidths are desired, making multi-band design extremely challenging.

2. **Low noise amplifier (LNA):** The LNA forms the first stage in the RX signal path and provides enough gain with low noise so that sensitivity is not compromised by the noise of subsequent stages. High linearity is also important to preserve overall linearity of the RX path, and also for tolerating interfering signals without corrupting the desired signal.

3. **Switch or duplexer:** Radios can be divided into two classes. Time-division duplexing (TDD) radios (figure 1-2(a)) are half-duplex i.e. they time-share the same frequency for both transmit and receive. On the other hand, frequency-division duplexing (FDD) radios (figure 1-2(b)) utilize different frequencies to

enable simultaneous transmit and receive operation. For TDD, an RF switch (controlled by the baseband processor) is used to time-multiplex between TX and RX mode. For FDD, a duplexer is used to tailor the frequency responses of the TX and RX path to a pair of frequencies defined by the standard, while sharing the same antenna. Since these components lie directly in the RF signal path, they also need to handle very high output power, while their loss adversely impacts performance.
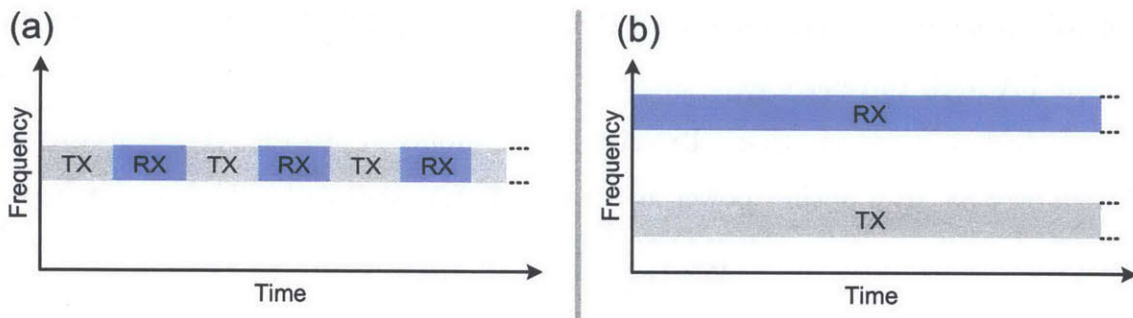
Figure 1-2: Radio classes: (a) Time division duplexing or TDD (b) Frequency division duplexing or FDD

CMOS technology evolution has not benefitted RF frontends to the same extent as transceivers because breakdown voltages decrease with scaling, making high-power compatibility more difficult rather than easier. Further, analog-intensive design techniques are still the norm for high-power RF circuits, and performance does not always benefit from scaling e.g. resonant LC circuits are used extensively, occupy large die area and do not scale. Discrete solutions using multiple device technologies are still common. Many of the current technologies in use are expensive and a barrier to radio commoditization. Finally, much hardware redundancy results from the lack of compelling multi-band solutions, further increasing area and cost.

To summarize, the key to addressing many of the challenges in radio design now lie in the RF frontend. In this thesis, new fully integrated RF frontend architectures suitable for the next generation of wirelessly connected devices are proposed, implemented and evaluated.

1.2 Research contributions

The first contribution of this work is a high power RF frontend architecture which leverages GaN technology. A vehicular connectivity application employing the new 802.11p standard is targeted. Traditionally, RF frontends for high power applications have relied on discrete components assembled into multi-chip modules. The proposed ultra-compact and fully integrated architecture is shown in figure 1-3. In addition to the superior efficiency offered by GaN technology, two architectural techniques are used to further boost performance. First, part of the TX/RX switch is absorbed into the PA itself to reduce loss. Second, the PA employs a dual-bias technique for transconductance linearization. To validate the proposed architecture, a prototype IC with over 33 dBm output power at 5.9 GHz is demonstrated in GaN technology.
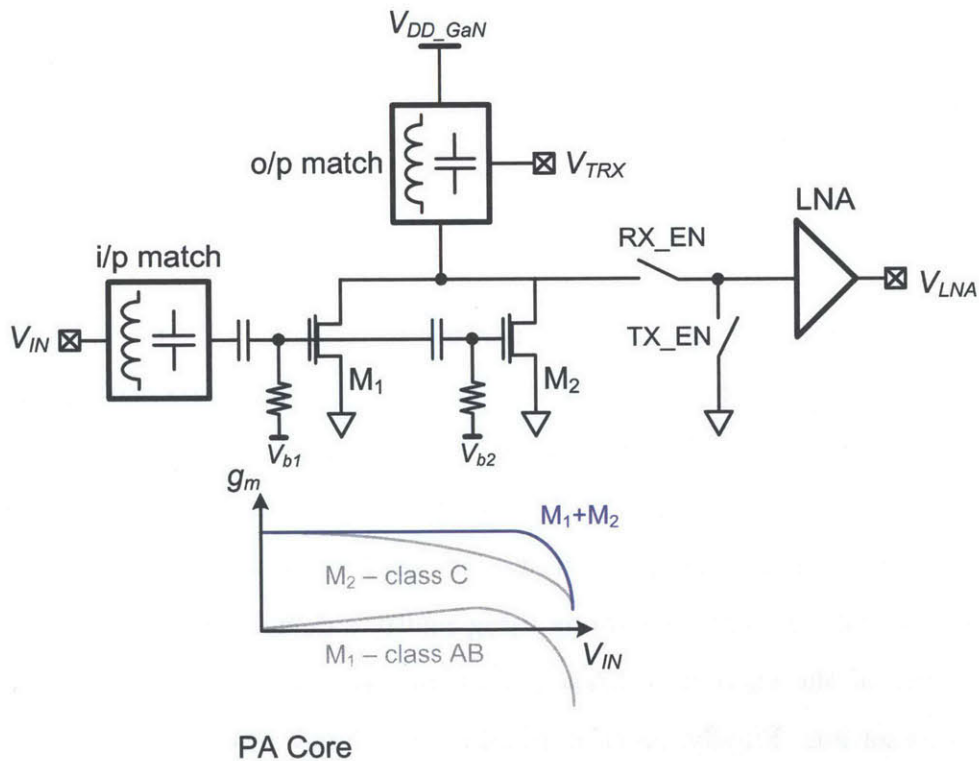
Figure 1-3: High-power GaN RF frontend architecture

The second contribution of this work is a frequency-agile RF frontend architecture for multi-band mobile radios. A unified solution for WLAN and TDD-LTE is targeted. The proposed architecture, which is shown in figure 1-4 eschews traditional narrow-

band analog RF circuits. Instead, both TX and RX signal paths employ broadband topologies inspired by the CMOS inverter. The frequency response of the system is digitally tunable over a wide range. Therefore, multiple narrowband frontend components can be replaced by this minimalistic and flexible solution. To validate the proposed architecture, a prototype IC with over 27 dBm output power and a target frequency range exceeding 1.8 - 3.6 GHz is demonstrated in SOI CMOS technology.
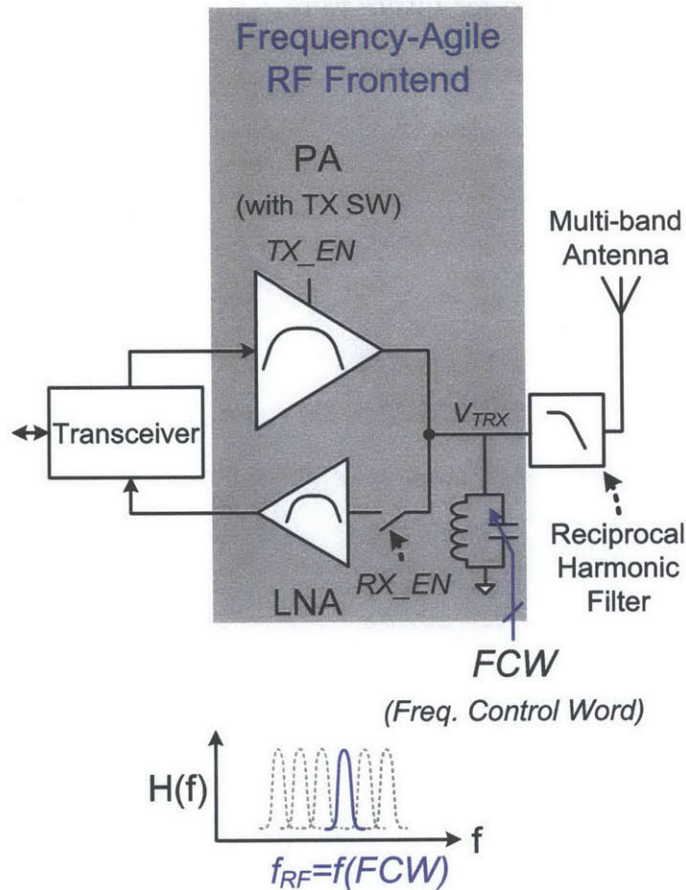
Figure 1-4: Frequency-agile RF frontend architecture for multi-band TDD radios

The RF frontend architectures presented in this thesis are specific to TDD standards. Therefore, this work only considers integration of the TX/RX switch. Nevertheless, it is worth noting that techniques introduced herein for PA efficiency enhancement and frequency-agile design are generally useful to any class of radio (TDD or FDD).

23

1.3 Thesis organization

This thesis is organized as follows. Chapter 2 discusses in more detail key challenges in the design of RF frontends for the next generation of wirelessly connected devices. Chapter 3 and 4 present the theory, design, implementation and measurement results of the two RF frontend prototypes introduced in the previous section. Finally, this work is put into perspective in chapter 5, wherein key contributions are summarized and a few directions are proposed for future work.

Chapter 2

Challenges in RF Frontend Design

As highlighted in chapter 1, evolution of the RF frontend is essential for realizing radios suitable for the next generation of wireless devices. In this chapter, principal design challenges are discussed in detail, and relevant prior art discussed to better differentiate the work detailed in chapters 3 and 4.

2.1   Monolithic integration

RF frontends require simultaneously high-power and high-frequency operation, thereby posing the one of the last hurdles towards monolithic integration. Traditionally, RF frontends have utilized heterogenous integration in the form of a multi-chip module, wherein each component is designed in the most optimal technology. Figure 2-1(a) shows a typical TDD example, while the desired evolution to a single-chip solution is shown in figure 2-1(b). Monolithic integration has the following potential benefits:

- **Cost:** Integration on a single-chip, particularly on a CMOS platform, is typically much cheaper than fabricating two or three different chips. Further, packaging cost is also minimized.

- **Form factor:** Single-chip solutions have smaller weight and volume than multichip modules.

25

- **Performance:** Tight integration of RF frontend components reduces interconnect loss.
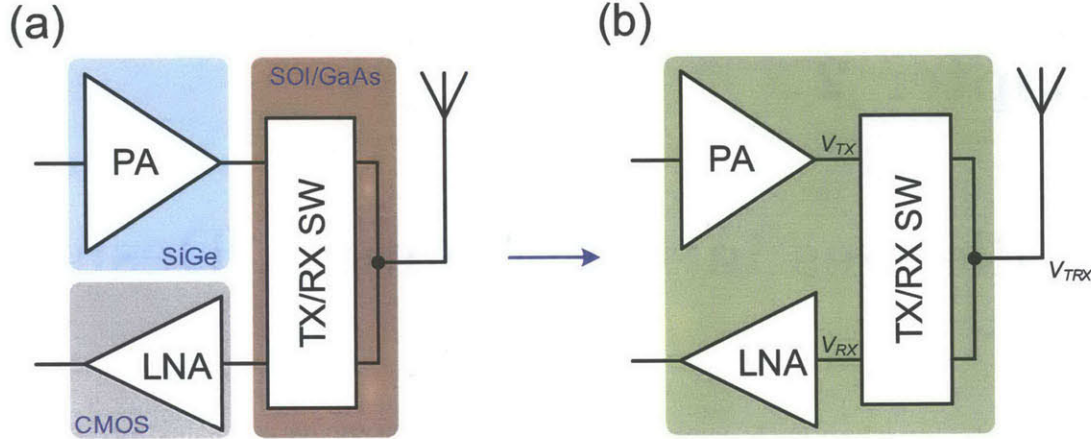
Figure 2-1: RF frontends employing (a) Heterogenous integration with multiple IC's (b) Monolithic or single-chip integration

For moderate power (up to 30 dBm) mobile applications, CMOS technology is especially attractive due to its low-cost, continued scaling and compatibility with mostly digital transceiver circuitry. However, CMOS integration of RF frontend components poses challenges, particularly for the PA and TX/RX switch.

Figure 2-2 shows an idealized PA topology. The drain inductor provides DC-feed and resonates out device capacitance at the desired RF frequency. The power extracted from the PA is maximized (saturated) when the RF component of device current causes the drain to swing over the entire available voltage headroom of 0 to $2V_{DD}$. This saturated power level is denoted as P_{SAT}. The function of the matching network is to perform impedance transformation from the antenna impedance Z_0 to R_{opt} for maximum power extraction. R_{opt} can be determined by:

$$P_{SAT} = \frac{V_{DD}^2}{2R_{opt}} = \frac{V_{TX}^2}{2Z_0} \tag{2.1}$$

Fine-line CMOS technology has a very low rated V_{DD} of about 1 V. At $P_{SAT} = 1\,\text{W}$ or 30 dBm, the required R_{opt} is 0.5 Ω, corresponding to a transformation ratio $\left(\frac{Z_0}{R_{opt}}\right)$
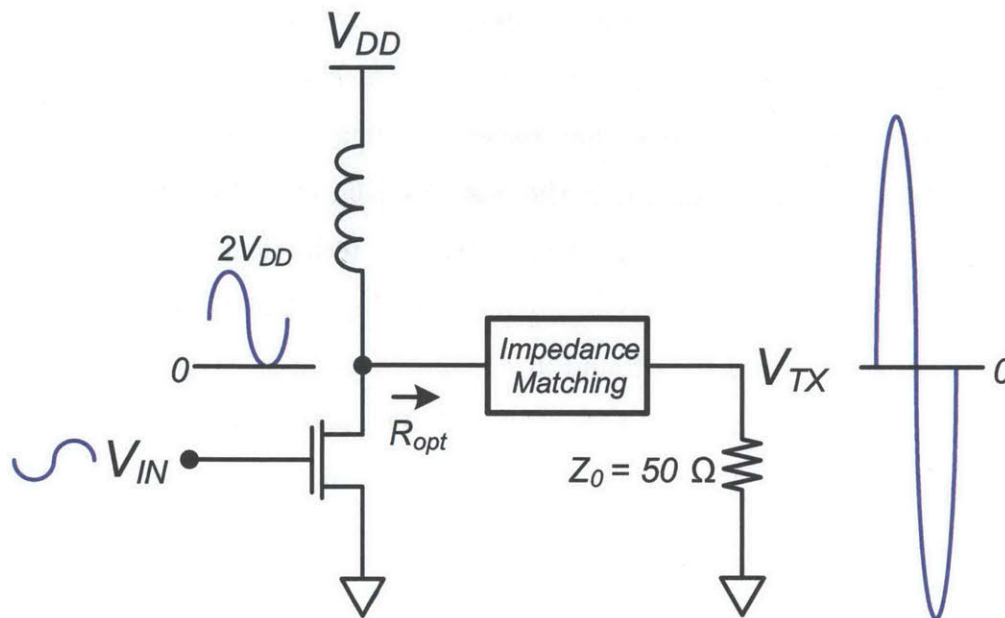
Figure 2-2: PA topology

of 100. Implementing such an extreme transformation is challenging. First, the circuit becomes very sensitive to parasitic resistance in interconnect, an issue exacerbated by the fact that the PA device is quite large. Second, when implementing such circuits with conventional L or π impedance matching topologies, component values become intractable and loss increases unacceptably. The search for alternate matching networks for CMOS RF power generation that alleviate some of these issues remains an active area of research, with power combining [10] [11] emerging as a prominent technique. Above 2 GHz, power levels exceeding 30 dBm have been demonstrated in CMOS [12] [13] [14].

At the same power level of 1W or 30 dBm, the voltage swing across at V_{TX} is 20 V_{PP}. Recall from figure 2-1(b) that the TX/RX switch is connected in between the antenna at the PA/LNA. Such high voltages make the design of the TX/RX switch also very challenging. In the TX mode, the PA transmits at power levels up to P_{SAT}. The TX branch switch is on and needs to conduct high RF voltages up to ± 10 V centered at 0 to the antenna Z_0. However, in bulk-CMOS technology, the body of the

transistor is always at ground, implying that for high power levels drain- and source-body diodes can have reverse breakdown during positive voltage excursions, and turn-on during negative voltage excursions. The RX branch switch has to block the same voltage levels, which leads to similar issues. Various body isolation techniques in bulk-CMOS have been attempted in the past [15] [16], but adequate power handling and reliability remains a concern. More advanced technologies like GaAs pHEMT provide inherent isolation but necessitate multiple chips. A more attractive solution is to use floating-body transistors found in silicon-on-insulator (SOI) technology. While SOI is marginally more expensive than bulk CMOS, it offers an excellent integration medium for RF frontends by combining the benefit of scaled CMOS with excellent RF isolation.

For high-power applications beyond 30 dBm, specially tailored high-power devices, which offer simultaneously high breakdown voltage, saturation current, transition frequency and RF isolation are essential. Gallium Nitride (GaN) technology has emerged as a compelling contender [17] but integration with the CMOS transceiver remains a challenge. Fortunately, recent research has opened up the possibility of GaN transistors as an add-on to CMOS wafers [18], which could enable monolithic integration in the near future.

2.2   Energy-efficient operation

For medium- and long-range communication, the TX section of the RF frontend, comprising the PA and TX branch switch, remains the most power hungry component of the radio. Therefore, maximizing TX efficiency goes a long way in maximizing battery life, or alternately, maximizing communication range on a given power budget.

PA architectures have been surveyed and analyzed extensively in many excellent references, including [9] [19]. Only a brief, high-level description is provided here for context. More details are presented in chapter 3 and 4 where appropriate. Figure 2-3 compares the asymptotic efficiency limit of various PA architectures versus power back-off from peak power. The desire for increased data-rates leads to non-constant

envelope or linear modulation schemes with ever increasing peak to average power ratios (PAPR), typically 6 dB or higher. Essentially, the dynamic range of the transmitted signal is increased in order to pack more information into a given frequency band. PA architectures fall into two main classes:

- **Classic linear architecture:** Class-A/B architectures [9] are inherently linear but suffer from poor efficiency at back-off. Their circuit topology is similar to the simplified example shown in figure 2-2. Distinction between Class-A and Class-B stems from the conduction angle of the active device over the RF cycle, which is 2π (always on) and π (on half the time) respectively, and controlled by the gate bias voltage. Class-B improves efficiency at the cost of linearity. As shown in figure 2-3, efficiency at 6 dB back-off is only 12.5 % and 39 % for ideal Class-A and Class-B amplifiers, respectively. In practical design, it is common to choose a conduction angle between 2π and π and refer to the amplifier as Class-AB.

- **Switch-mode based advanced architectures:** To improve efficiency under high PAPR signal drive, there is a strong push to move from Class-AB to more advanced PA architectures which promise better efficiency. Most such architectures employ nonlinear switch-mode amplifiers (e.g. Class-D/E/F etc.) at their core. The output power at any given time is proportional to both the square of the supply voltage and inversely proportional to the load impedance, but independent of the input amplitude. Therefore, output power is varied or backed-off either through supply or load modulation techniques. Outphasing [20] is one promising candidate which can be implemented with both lossy and lossless combiners. With lossy combiners, efficiency reduces at back-off in proportion to output power, just like a Class-A PA [21]. Lossless combiners are highly preferred as the asymptotic efficiency of the architecture is 100 % irrespective of power back-off due to load modulation, but do not offer impedance isolation and work best with low output impedance cores like voltage-mode Class-D [22]. Polar modulation [23] is another alternative based on supply modulation with

the same constant 100 % asymptotic efficiency characteristic. However, implementing the supply modulator is challenging, particularly for wide bandwidths. ML-LINC and AMO [24] are hybrids of polar and outphasing which simplify the supply modulator while improving average efficiency.
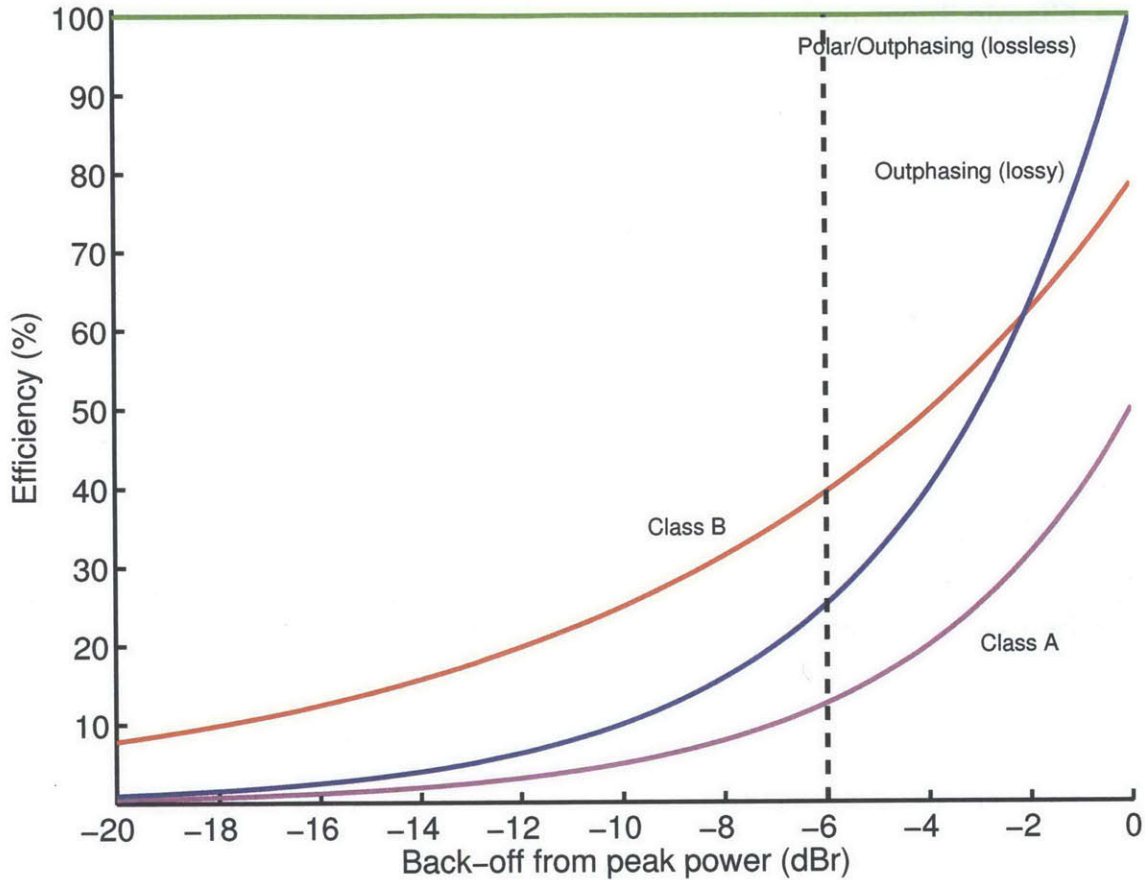
Figure 2-3: Asymptotic efficiency limits of different PA architectures

For moderate power levels up to 30 dBm, CMOS technology represents an attractive but very challenging medium for efficient and linear PA design. The extreme impedance transforms required to extract power from low voltages (see section 2.1) lead to high-Q matching networks or power combiners which are lossy when implemented on conductive deep sub-micron CMOS substrates [11]. On the other hand, the level of integration possible in CMOS is unrivaled and can enable new architectures with better efficiency. The key lies in exploring architectures that leverage the strengths of CMOS and overcome the weaknesses. Best-in-class GHz-range Class-AB

CMOS PA's report peak efficiencies around 40 % [12]. The theoretical promise of switching based architectures is tremendous, and there has been a lot of interest in moving to advanced architecture like polar, outphasing and variants thereof. Interestingly, the peak efficiencies of GHz-range switching CMOS PA's is degraded by practical issues such as matching network and switching loss, and generally not much better than their Class-AB counterparts [25] [26]. However, the real promise of these switching architectures is in boosting average efficiency when driven by modulated signals, where they outperform Class-AB solutions by about 1.5 - 2.0×.

For high power levels beyond 30 dBm, new materials such as GaN promise higher efficiency and power density but cannot support complicated efficiency enhancing architectures on a single-chip until hybrid GaN-CMOS platforms [18] become feasible. Therefore, more research into minimalistic architectures, which completely integrate all matching elements on-chip while achieving competitive efficiency and linearity is needed.

Finally, it is worth mentioning here that TX branch switch loss should be included when optimizing for overall TX efficiency. This aspect is given detailed treatment in section 3.3.1.

2.3 Multi-octave frequency coverage

The number of frequency bands and wireless standards supported by mobile devices has steadily risen in the last decade. By revenue, smartphones is the biggest market [2] and therefore an apt use-case to consider. Most smartphones sold today support at least the following standards, frequency bands and peak power levels:

- **LTE:** FDD and/or TDD, 5 - 10 out of over 40 possible bands, anywhere from 0.4 GHz to 3.8 GHz, 30 dBm

- **WLAN:** TDD, 2.4 - 2.5 GHz and 4.9 - 5.9 GHz, 27 dBm, possibly multiple antennas and radios for MIMO

- **Bluetooth:** TDD, 2.4 - 2.5 GHz, 20 dBm

Even though the amount of integration and connectivity offered by current devices is impressive, there are compelling drivers to support even more frequency bands:

- To enable universal phones, eliminating the need to build separate models for different carriers and countries. For the manufacturer, this implies consolidation and for the user, true worldwide connectivity.

- To make phones communicate with new classes of devices, such as those coming online through the IoT paradigm e.g. Medical implant communications services (MICS) band (402 - 405 MHz).

- To support MIMO in more bands with the same amount of hardware.

Unfortunately, the notion of RF circuits operating over a wide frequency contradicts well established design techniques. Figure 2-4(a) shows a generic multi-stage RF amplifier chain. Typically, each stage employs a resonant circuit to counteract device capacitance to achieve gain.
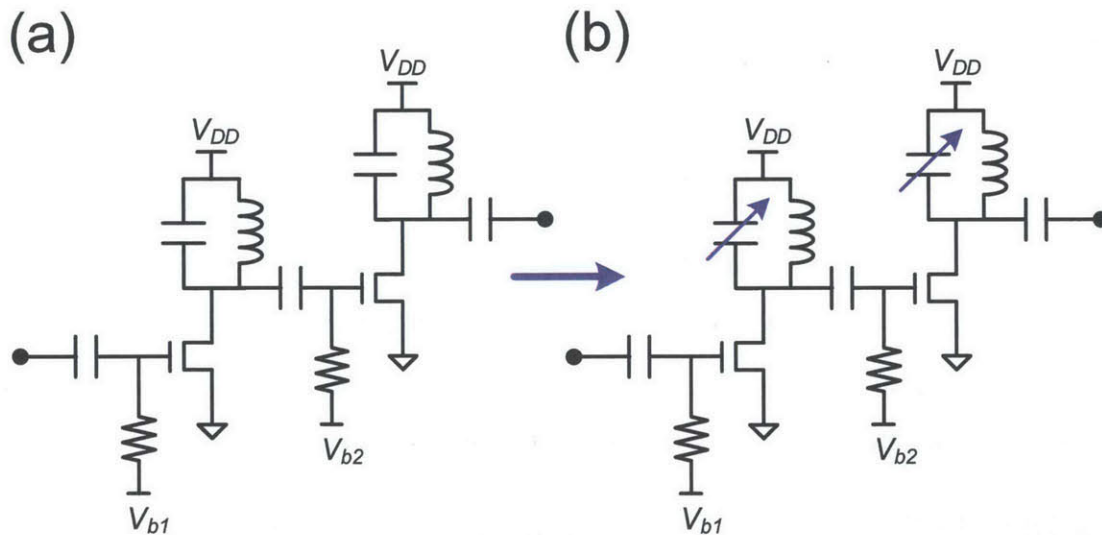
Figure 2-4: Resonant RF amplifier chain (a) Fixed (b) Tunable

A natural extension of the conventional approach leads to tunable RF circuits, as shown in figure 2-4(b). Such circuits are indeed feasible for low-power circuitry,

and make up building blocks for many multi-standard CMOS transceivers [27] and transceiver blocks [28] [29]. More recently, CMOS RX design has evolved to the point that a filter-less solution for multi-band applications now appears in sight [30].

Unfortunately, when RF frontend integration (including a high-power TX) is considered, power levels are high enough that tunable passives with good RF performance are not trivial to implement. With digital tuning schemes the switches used to implement tunable capacitors and/or inductors face the same issues as the TX/RX switch at high-power levels. Similarly, analog tuning elements such as varactors perform well only at low power levels. Indeed, *all* the CMOS PA research cited in sections 2.1 and 2.2 is focused on single-band operation. Due to limitations of current design techniques, most current multi-band RF frontends employ a "brute force" approach to multi-band coverage. Figure 2-5 shows a tri-band radio example, covering three bands centered at f_1, f_2 and f_3. For each band, dedicated circuitry is employed and switched in and out as needed. Since only one band per radio is active at a time, this architecture is not area and cost optimal. Tri- and dual-band PA arrays reported in [31] [32] [33] follow this concept, so do the more complete radio implementations found in [34] [35]. To the best of the author's knowledge, there are currently no monolithic high-power RF frontends tunable over multiple octaves in any CMOS technology. Clearly, alternate design techniques need to be explored to enable more minimalistic and flexible architectures.
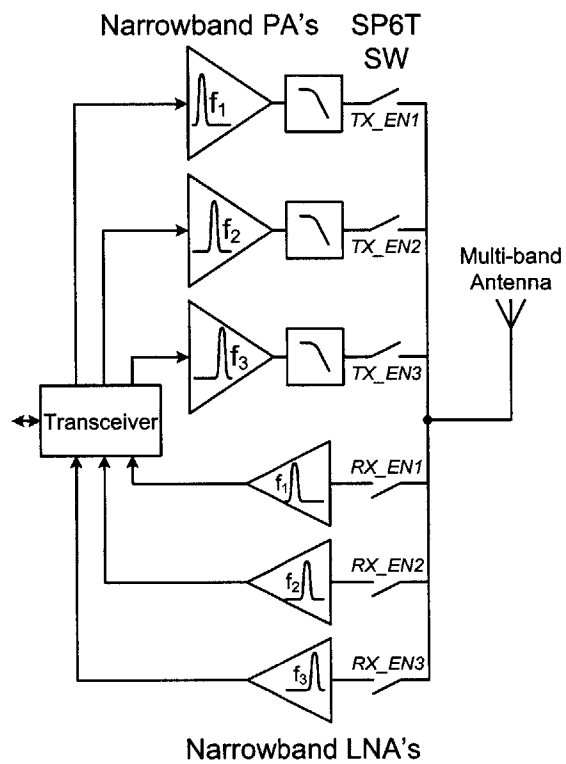
Figure 2-5: Conventional multi-band RF frontend employing hardware redundancy to cover multiple bands

Chapter 3

A High-power GaN RF Frontend for Vehicular Connectivity

The vision of ubiquitous connectivity requires bringing online previously isolated, closed-circuit systems. One of the more interesting use-cases pertains to vehicles. Modern automobiles include sophisticated electrical systems comprised of a vast array of sensors and actuators connected to a central computer. Sensors track vehicle location, speed, tire pressure, collision damage and image the environment in real-time; while actuators control locks, brakes, airbags and front/rear power distribution. Pushing sensor data into the cloud in real-time for contextual feedback enabled by tight integration with traffic signaling and map databases has the potential to fundamentally change transportation. Indeed, it is anticipated that vehicular connectivity will not only have implications for enhancing road navigation and safety [36], but will also be an integral part of future paradigms like self-driving cars [37].

This chapter presents the design of an RF frontend for the recently ratified 802.11p standard, which aims to bring dedicated wireless connectivity to vehicles. As discussed in chapter 2, high-power RF frontends have traditionally relied on multi-chip modules or discrete designs. The prevailing theme of this chapter is the interplay of emerging technology and architecture, which enables monolithic high-power RF integration. With a fixed power budget, maximizing energy efficiency is important to enhance communication range, which is especially important for vehicles. The im-

provement in baseline energy efficiency by leveraging GaN is theoretically quantified. In addition, system energy efficiency is further enhanced through two architectural techniques: (a) A modified TX/RX switching scheme (b) Dual-bias linearization.

Chapter organization is as follows: Sections 3.1 and 3.2 provide a brief overview of the 802.11p standard and GaN technology, respectively. Section 3.3 introduces the proposed architecture. Design of the transmitter (TX) and receiver (RX) paths is discussed in sections 3.4 and 3.5, respectively. Section 3.6 outlines physical design aspects. Finally, section 3.7 presents measurement results on the fully integrated GaN IC prototype.

3.1 IEEE 802.11p

802.11p [38], often referred to as wireless access in vehicular environments (WAVE), is an amendment to the well established and widely used 802.11 (WLAN) standard. 802.11p utilizes a dedicated 75 MHz band from 5.850 to 5.925 GHz in the form of seven 10 MHz channels with 64-QAM OFDM modulation to support data-rates up to 27 Mb/s. In the standard, average RF transmit power levels up to 28.8 dBm are supported.

(a) Vehicle-to-vehicle (V2V)   (b) Vehicle-to-basestation (V2B)

Figure 3-1: 802.11p use-cases

As shown in figure 3-1, both vehicle-to-vehicle (V2V) and vehicle-to-basestation (V2B) scenarios are possible. V2V communication can enhance road safety through

36

downstream multi-hop relay of adverse road conditions or accidents. V2B communication can improve navigation and enable traffic-optimized road signaling to reduce delay and congestion, in addition to providing dedicated, high-quality internet access on-the-move.

3.2 GaN: An enabling technology for monolithic high-power RF integration

The physical properties of several semiconductors including GaN are presented for comparison in table 3.1 [17].

Property	Si	GaAs	SiC	GaN
Bandgap (eV)	1.1	1.4	3.2	3.4
Critical e-field (MV/cm)	0.6	0.5	3.0	3.5
Charge density ($\times 10^{13}/cm^2$)	0.3	0.3	0.4	1.0
Thermal conductivity ($W/cm/K$)	1.5	0.5	4.9	1.5
Mobility ($cm^2/V/s$)	1300	6000	600	1500
Saturation velocity ($\times 10^7 cm/s$)	1.0	1.3	2.0	2.7

Table 3.1: Physical properties of various semiconductors

To enable monolithic integration of high-power RF frontends, the adopted device technology should have the following features:

1. **High power density:** The breakdown voltage of the technology, in conjunction with current density and thermal conductivity, determines the achievable power density (Watts/mm). Breakdown voltage also determines the voltage blocking limit of RF switches. The wide band gap and high critical field of GaN enable breakdown voltages as high as 150 V, nearly two orders of magnitude higher than fine-line CMOS. Further, the high charge density of GaN enables current densities as high as $0.8 - 1.5$ A/mm, very competitive with fine-line CMOS. GaN also features excellent thermal conductivity to facilitate

heat transfer away from high-power devices to keep junction temperatures reasonably low. Consequently, power densities achievable in GaN are typically $20 - 50\times$ higher than CMOS.

2. **High transition frequency (f_T):** As a rule-of-thumb, the RF operating frequency should be at least a $3 - 5\times$ lower than f_T. The competitive mobility and high saturation velocities in GaN enable an f_T of $20 - 40$ GHz for state-of-the-art devices at breakdown voltages of about 150 V [39], which is sufficiently high to make GaN an attractive candidate for many RF applications.

3.3 Proposed architecture

The first efficiency enhancement technique proposed in this work is a modified switching scheme that reduces TX mode switch loss. The performance impact of a conventional switching scheme is discussed first in section 3.3.1, followed by an introduction to the modified switching scheme used in the present work in section 3.3.2.

3.3.1 Performance impact of TX/RX switch
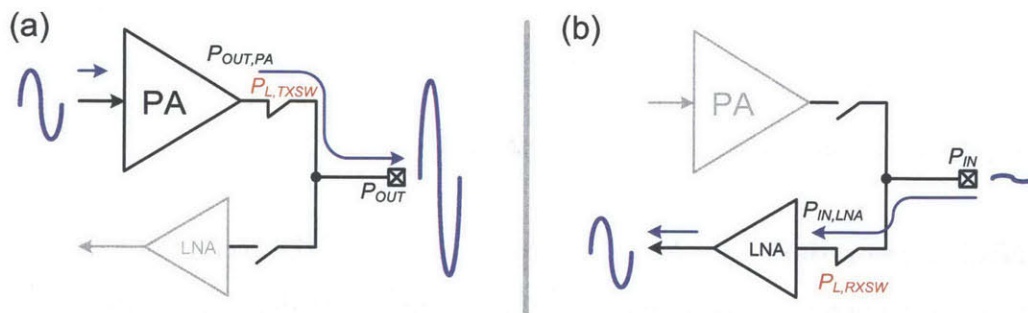
Figure 3-2: Performance impact of TX/RX switch (a) TX mode (b) RX mode

As shown in figure 3-2, an explicit series switch is usually employed in both the TX and RX branches in order to switch modes in a TDD radio. The design of TX branch switch (TXSW) is particularly challenging as it needs to meet three conflicting requirements simultaneously:

1. High RF power (both voltage and current) handling capability

2. Low insertion loss to minimize TX efficiency degradation

3. High linearity so as to not limit the overall linearity of the TX chain

While switches with adequate power handling and linearity have been demonstrated, they exhibit high insertion loss of 1 to 2 dB [40] and consume significant die area. Lossy series elements in the TX RF path (between PA and antenna) are highly undesirable due to the adverse impact on overall TX performance, which can be quantified with the aid of figure 3-2. If the PA produces an output power of $P_{OUT,PA}$, and the TXSW has a loss of $P_{L,TXSW}$ (in dB), then the achievable TX output power and efficiency are:

$$P_{OUT} = P_{OUT,PA} \times 10^{-\left(\frac{P_{L,TXSW}}{10}\right)}$$ (3.1)

$$\eta = \eta_{PA} \times 10^{-\left(\frac{P_{L,TXSW}}{10}\right)}$$ (3.2)

Since the PA dominates TX power consumption, the energy consumed by it is:

$$E_{TX} = \frac{P_{OUT,PA}}{\eta_{PA}} \times t$$ (3.3)

Typically, the total output power from the TX is dictated by the standard, implying that the PA output power has to be boosted to compensate for TXSW loss; its impact on TX energy consumption can be obtained from equations 3.1 and 3.3:

$$\frac{\Delta E_{TX}}{E_{TX}} = 10^{\left(\frac{P_{L,TXSW}}{10}\right)} - 1$$ (3.4)

3.3.2 Modified switching scheme

The proposed architecture focused on co-design of the TX and RX paths to essentially eliminate the explicit switch from the TX branch in order to improve efficiency and save energy. In mathematical terms, the goal is:

$$P_{L,TXSW} = 0 \tag{3.5}$$

Equation 3.4 is plotted in figure 3-3, showing that overcoming even 1 dB of loss *dramatically* lowers TX energy consumption by 20 %.
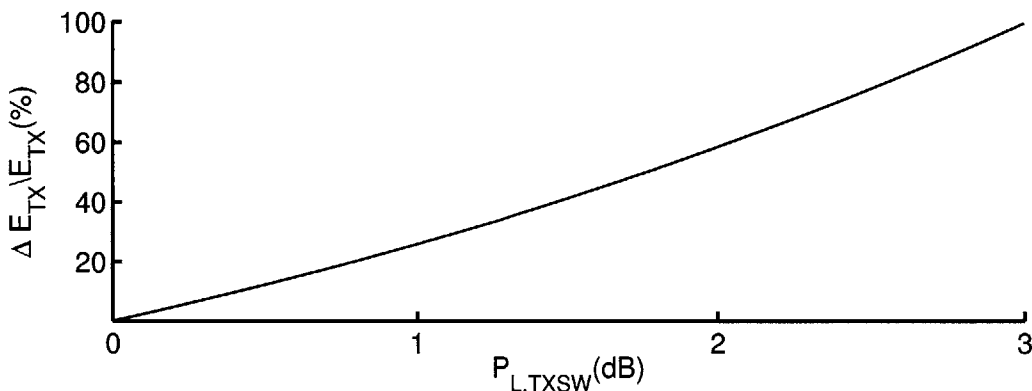
Figure 3-3: Impact of TXSW loss on energy consumption

Figure 3-4 shows the modified GaN frontend architecture in its two modes of operation. No explicit switch is included in the TX branch. Instead, the drain of the PA transistors is directly connected to the RX section. Frontend operation can be understood by considering the two modes of operation in turn:

1. **TX mode:** In the TX mode, the PA is active. The output matching network is designed to present the optimum impedance $Z_L(\omega)$ (when loaded by a 50 Ω antenna) which extracts maximum power from the GaN power transistor. Under this condition, the drain of the PA device sees large voltage swings up to approximately $2V_{DD}$. The RX branch series-shunt switch (RXSW) isolates and protects the inactive LNA from this voltage stress.

2. **RX mode:** In the RX mode, the PA transistors are turned off by employing a sufficiently large negative gate bias. Under this condition, the PA presents a small but finite capacitance $C_{O,PA}$ at the drain node. Since the output matching network is passive, its characteristics do not depend on the operating mode. The LNA is power matched to the parallel combination of Z_L and $C_{O,PA}$ at the

40

RF frequency ω to maximize gain and minimize noise figure. In mathematical terms:

$$Z_{RX}(\omega) = Z_{TX}(\omega)^* = \left(Z_L(\omega) \parallel \frac{1}{j\omega C_{O,PA}} \right)^* \tag{3.6}$$
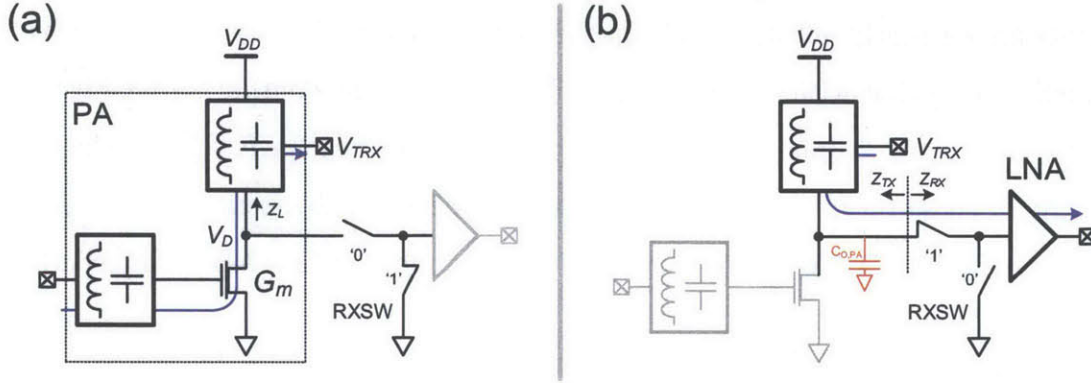
Figure 3-4: Proposed GaN RF frontend (a) TX mode (b) RX mode

3.4 TX design

The TX needs to satisfy the conflicting requirements of high output power, efficiency and linearity simultaneously. Suitability of GaN for high-power applications has already been discussed in section 3.2. In terms of overall TX efficiency, the modified architecture outlined in section 3.3 provides an advantage. The focus of this section is the design of an appropriate core PA topology with simultaneously high linearity and good efficiency suitable for an ultra-compact monolithic solution.

As previously discussed in section 2.2, PA design for non-constant envelope modulation schemes has two major flavors [9]: (a) Class-AB amplifiers which are inherently linear but somewhat inefficient (b) Switching amplifiers which are inherently nonlinear but very efficient in conjunction with more sophisticated architectures (e.g. EER, outphasing) to synthesize non-constant envelope waveforms.

In this work, a Class-AB design is chosen because it offers the most compact and fully integrable solution for a linear PA. While the peak efficiency of Class-AB ampli-

41

fiers under linear modulation is acknowledged to be somewhat lower than switching solutions, the features of GaN technology impact Class-AB architectures in a favorable way for the desired power levels. Efficiency of Class-AB amplifiers in GaN technology is theoretically quantified and compared to other technologies in section 3.4.1. On the other hand, switching amplifier based architectures require additional circuitry. For instance, polar modulation [23] requires a high-voltage supply modulator. Such circuits are currently not integrable on the same chip as GaN transistors. However, it is worth noting that with the emergence of platforms for monolithic GaN-CMOS integration [18], more complicated architectures may indeed become feasible in the future. Finally, in these future technology platforms, efficiency enhancement in the form of envelope tracking [19] will also be possible on the chosen Class-AB architecture.

3.4.1 GaN Class-AB efficiency analysis

Since Class-AB amplifiers are usually biased quite close to device threshold, the following analysis treats the amplifier as Class-B to arrive at an estimate for GaN Class-AB PA performance relative to other device technologies. Figure 3-5 shows the simulated DC I-V characteristics of a reference GaN power transistor. There are primarily two main loss mechanisms responsible for deviation of Class-B drain efficiency from the theoretical maximum, denoted here by $\eta_B = 78.5$ %:

1. **Finite device knee voltage:** Linear Class-AB operation requires the transistor to operate as a high output impedance current source. Therefore, the V_{DS} across the transistor cannot fall below a certain voltage, denoted as the device knee voltage V_{DSAT}. The impact of finite knee voltage on efficiency is quantified by the factor:

$$\eta_K = 1 - \frac{V_{DSAT}}{V_{DD}} \qquad (3.7)$$

2. **Output matching network loss:** In order to extract maximum *instrinsic*
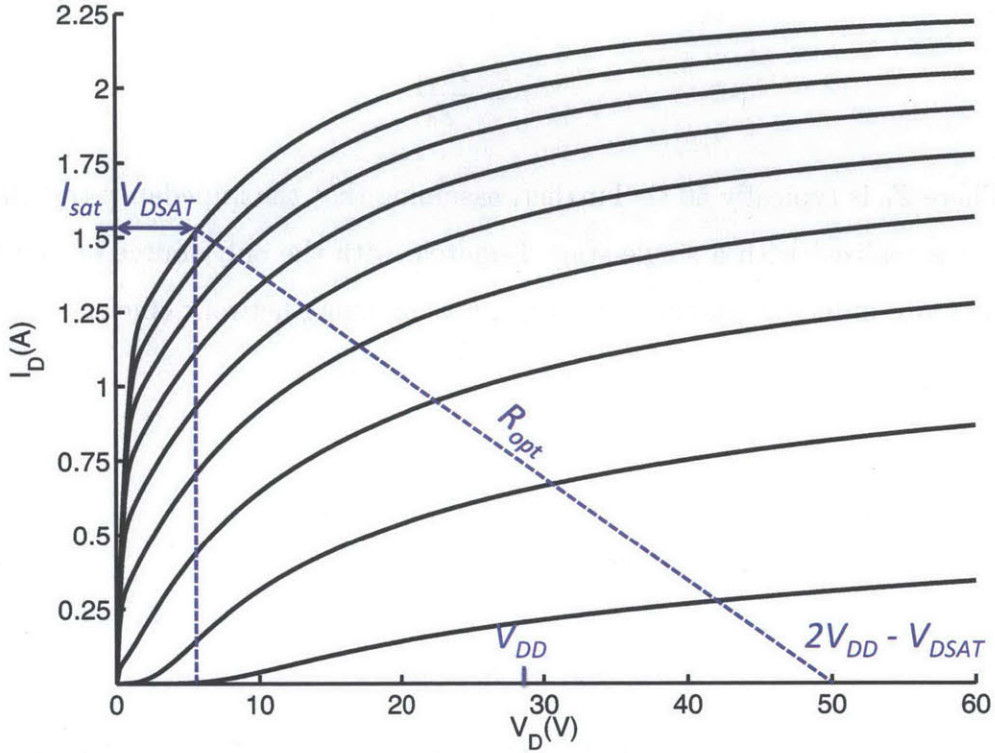
Figure 3-5: IV characteristics of a reference GaN power transistor

power $P_{SAT,i}$[1] from the transistor, the antenna impedance Z_0 has to be transformed to an optimal value R_{opt} given by:

$$P_{SAT,i} = \frac{(V_{DD} - V_{DSAT})^2}{R_{opt}} \tag{3.8}$$

From load-line theory [9], an alternate expression for R_{opt} can also be derived.

$$R_{opt} = 2\frac{V_{DD} - V_{DSAT}}{I_{sat}} \tag{3.9}$$

Substituting R_{opt} from equation 3.8 gives I_{sat}, which can be used to size the power transistor appropriately. The impedance transformation required to ex-

---

[1]Maximum power extractable with device staying in saturation region, assuming matching network loss is zero.

tract maximum power is therefore[2]:

$$r = \frac{R_{opt}}{Z_0} \tag{3.10}$$

Where Z_0 is typically 50 Ω. Further, assuming that the impedance transformation is realized with a single-stage L-match, with the only source of loss being the finite inductor quality factor Q_L[3], the matching network efficiency is [11]:

$$\eta_M = \frac{1}{1 + \frac{r}{Q_L^2}} \tag{3.11}$$

The overall PA efficiency is therefore:

$$\eta_{SAT} = \eta_B \eta_K \eta_M \tag{3.12}$$

Due to matching network loss, not all of the power extracted from the device reaches the output. Accounting for this loss, the actual output power realized (in watts) is:

$$P_{SAT} = \eta_M P_{SAT,i} \tag{3.13}$$

Representative process parameters for three different device technologies are used to compare the achievable efficiency over a range of power levels. η_K is independent of power level and computed from equation 3.7. η_M is determined by a two-step process. First, the required r is determined from equations 3.8 and 3.10. Second, a typical Q_L value of 15 is assumed to calculate η_M from equation 3.11.

The achievable efficiency limit is plotted in figure 3-6. The low breakdown voltage of CMOS and SiGe processes results in extreme downward impedance transforms at watt-level output power. Consequently, η_M suffers[4] and lowers the overall efficiency.

---

[2]Usually, Z_0 will be matched to $Z_{opt}(\omega) = R_{opt} \parallel \frac{1}{j\omega C_{O,PA}}$ but the reactance is omitted here for simplicity

[3]Up to 6 GHz, quality factor of integrated capacitors is typically much higher than inductor and can be neglected

[4]Some alternate matching techniques do exist that relax the tradeoff between r and η_M [11], but the argument remains generally valid.

In contrast, GaN only requires a moderate transformation for watt-level output power. At the target power level of 36 dBm (4 W), GaN can offer up to 65 % drain efficiency, which represents a very favorable starting point for the present design.
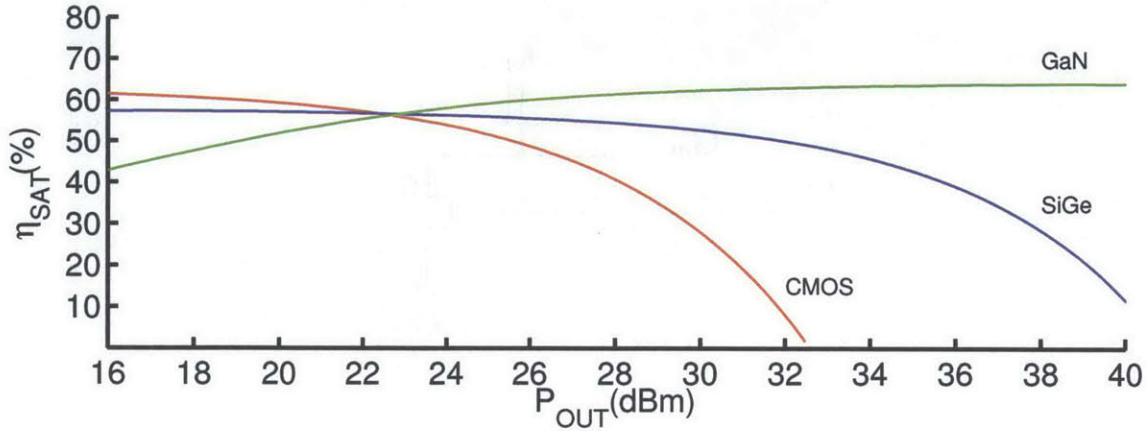
Figure 3-6: Class-AB efficiency limits compared for CMOS, SiGe and GaN processes

3.4.2 Dual-bias linearization

While the preceding section illustrates that Class-AB amplifiers in GaN technology are capable of respectable efficiency, nothing has been said so far about linearity. Unsurprisingly, the principal source of Class-AB PA nonlinearity is the active device, which can essentially be modeled as a large-signal nonlinear transconductance which is dependent on frequency, bias and drive conditions. To balance efficiency and linearity, it is standard practice to bias Class-AB amplifiers close to device threshold [9]. Under such bias conditions, the device exhibits significant compressive nonlinearity at high drive levels, leading to a lower than desired P_{1dB} point. Since Class-AB amplifiers have to be backed-off from their P_{1dB} point to meet system linearity and/or EVM requirements, a low P_{1dB} point corresponds to lower average efficiency. In this work, the second key efficiency enhancement technique is to increase the P_{1dB} point of the PA through cancellation of the compressive nonlinearity with an expansive counterpart. As previously demonstrated in [32] [41], compressive nonlinearity in Class-AB amplifiers can be cancelled by introducing an appropriately sized auxiliary power transistor biased Class-C with its output current combined in phase with the

45

main Class-AB device. This arrangement is shown in figure 3-7. Figure 3-8 shows the simulated individual transconductances G_{m1}/G_{m2} (Class-AB/C) over a range of drive voltages, as well as the linearized composite transconductor G_m.
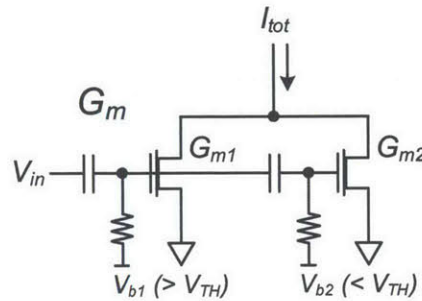
Figure 3-7: Composite Class-AB/C transconductor
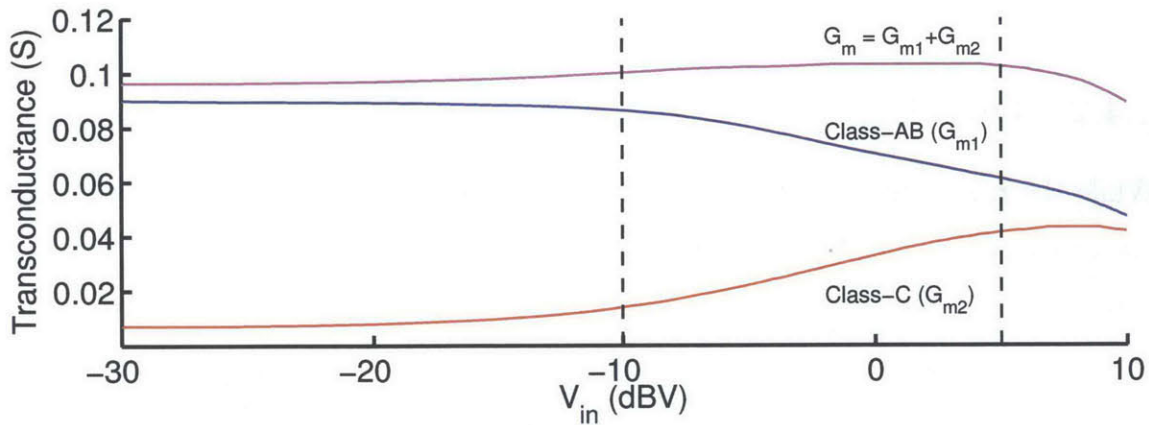
Figure 3-8: Dual-bias linearization of large-signal transconductance

3.4.3 Matching networks and circuit implementation

The optimum load impedance derived in equation 3.9 provides intuition but is a bit simplistic. In addition to neglecting device output capacitance, there are several effects which are not included. For instance, device capacitances are voltage-dependent, implying that under large-signal drive, the impedance seen will vary dynamically. Further, actual power transistors are not unilateral due to internal feedback, implying that both output and input impedances affect output power and efficiency.

46

Large-signal models [42] provided by the foundry try to capture most of these effects to enable computer-aided design. Under large-signal drive, active devices produce power at the desired fundamental frequency and its harmonics. Given a specific active device model, bias condition and supply voltage, PA behavior can be described as a function of source (gate) and load (drain) impedance vectors:

$$\eta_{PA} = f_1 \langle [Z_L(\omega) \ Z_L(2\omega) \ ...], [Z_S(\omega) \ Z_S(2\omega) \ ...] \rangle \qquad (3.14)$$

$$P_{OUT} = f_2 \langle [Z_L(\omega) \ Z_L(2\omega) \ ...], [Z_S(\omega) \ Z_S(2\omega) \ ...] \rangle \qquad (3.15)$$

PA design optimization can be partially automated through large-signal load-pull simulation. As shown in figure 3-9, the impedance tuners are controlled by an algorithm that searches for optimal impedance vectors which maximize either efficiency or output power. Generally, the impedance vectors corresponding to maximum output power and efficiency lie close to each other. In the present work, the design is optimized for efficiency.
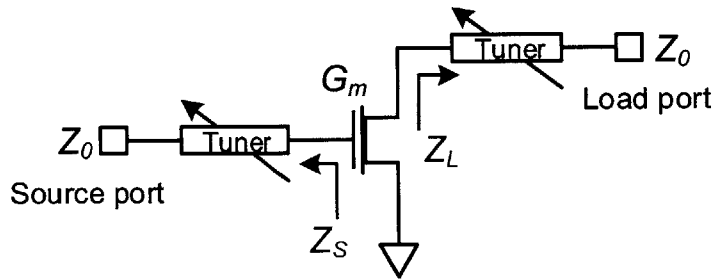
Figure 3-9: Load-pull setup

With load-pull based designs, it is common to specify impedance vectors up to the third harmonic 3ω. However, accurate control of harmonic impedances is more appropriate for distributed designs and not always possible in lumped, integrated designs such as the present one. Further, with lumped designs the efficiency enhancement resulting from optimized harmonic impedances is often negated by losses arising

47

from the additional passive components needed. Since the device capacitance tends to rotate towards a short at higher frequencies, $Z_L(2\omega)/Z_L(3\omega)$ and $Z_S(2\omega)/Z_S(3\omega)$ are set to 0 leaving $Z_L(\omega)$ and $Z_S(\omega)$ as the only degrees of freedom. Augmenting equation 3.9 to include device capacitance, the theoretical $Z_{opt}(\omega)$ serves as a good starting point:

$$Z_{opt}(\omega) = R_{opt} \parallel \frac{1}{j\omega C_{O,PA}} \tag{3.16}$$

Initially, the source impedance is set to Z_0 to determine $Z_L(\omega)$ with load-pull, and then a source-pull is performed to find $Z_S(\omega)$. Since efficiency is more sensitive to load impedance, the optimal $Z_L(\omega)$ is used, while a somewhat sub-optimal $Z_S(\omega)$ is used in order to improve stability. The final load and source impedances used are presented in table 3.2.

Impedance	Value
$Z_L(\omega)$	$Z_0(0.57 + j1.33)$
$Z_S(\omega)$	$Z_0(0.66 + j0.36)$

Table 3.2: PA impedances at fundamental frequency

The complete PA schematic is shown in figure 3-10, while the corresponding component values are listed in table 3.3. At the drain, the shunt-L series-C match provides impedance matching from Z_0 to $Z_L(\omega)$, DC bias and RF signal coupling with only two elements. A series-C shunt-L match is chosen for the gate as it performs the transformation from Z_0 to $Z_S(\omega)$ while also a symmetrical split of the RF signal to feed the two gates in phase for the dual-bias topology.
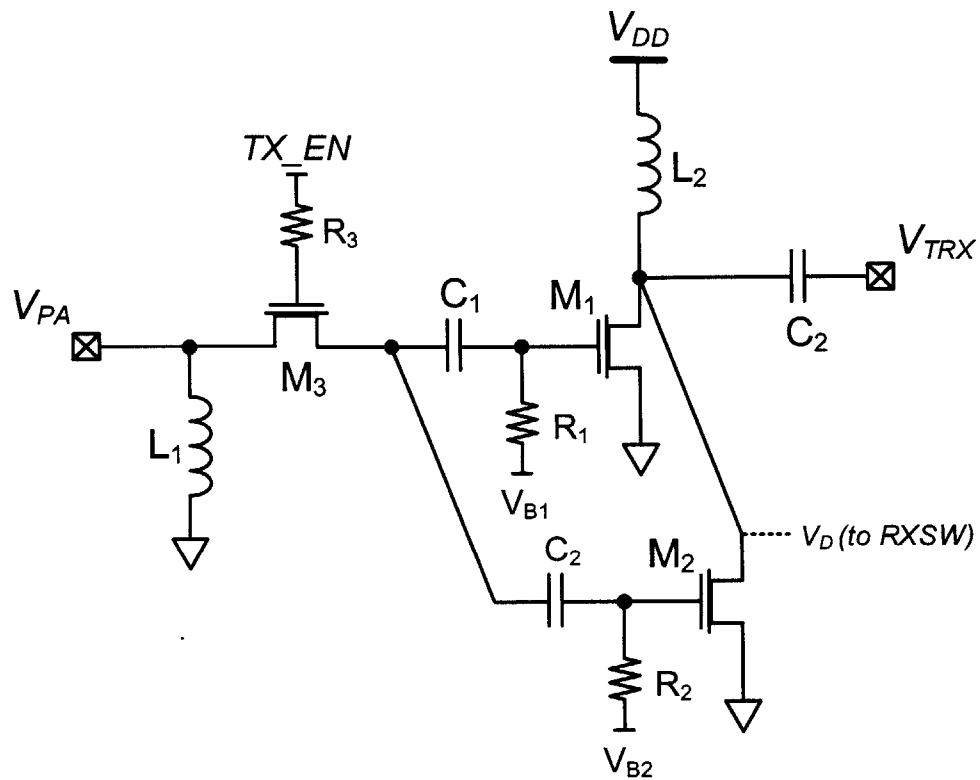
Figure 3-10: PA schematic

Component	Value
M_1, M_2	$180{\times}2$ μm
M_3	$80{\times}4$ μm
L_1	1.22 nH
L_2	1.25 nH
C_1, C_2	2.20 pF
C_3	0.33 pF
R_1, R_2	300 Ω
R_3	2 KΩ

Table 3.3: PA component values

3.5 RX design

3.5.1 RX switch

Since the RX incident power is quite low the RX switch is not required to conduct a lot of current. However, it must meet two other requirements:

1. In TX mode, isolate and protect the LNA from the high-power PA which is transmitting.

2. In RX mode, provide sufficiently low insertion loss to not significantly degrade the noise figure of the RX path.

The series-shunt configuration is the preferred switch topology to improve isolation and power handling [43]. In general, GaN is an excellent technology to implement high-voltage RF switches. The high breakdown voltage of GaN devices allows use of just a single device in the series path, which reduces on resistance and therefore insertion loss in the RX mode.

The RX switch schematic is shown in figure 3-11, with waveforms corresponding to the TX mode, which represents the high voltage stress case. Recall that in this modified frontend architecture (see figure 3-4(a)), the RX switch is directly connected to the drain of the PA device denoted here by V_D. As shown in figure 3-5, the maximum RF voltage amplitude at V_D in the TX mode is simply:

$$V_{max} = V_{DD} - V_{DSAT} \tag{3.17}$$

The capacitor C_1 can be assumed to be an ideal DC block for the present analysis. The LNA input V_{LNA} is grounded through M_2, which is on. The voltage stress at V_D is shared equally in the divider formed by the gate-source/drain parasitic capacitors C_P of M_1, which is off. The control voltage V_{OFF} should be chosen such that the RF voltage stress does not accidentally turn on M_1. Therefore, V_{OFF} should satisfy:

$$V_{TH} \gg max(V_{CP}) = V_{OFF} + \frac{V_{max}}{2} \tag{3.18}$$

50
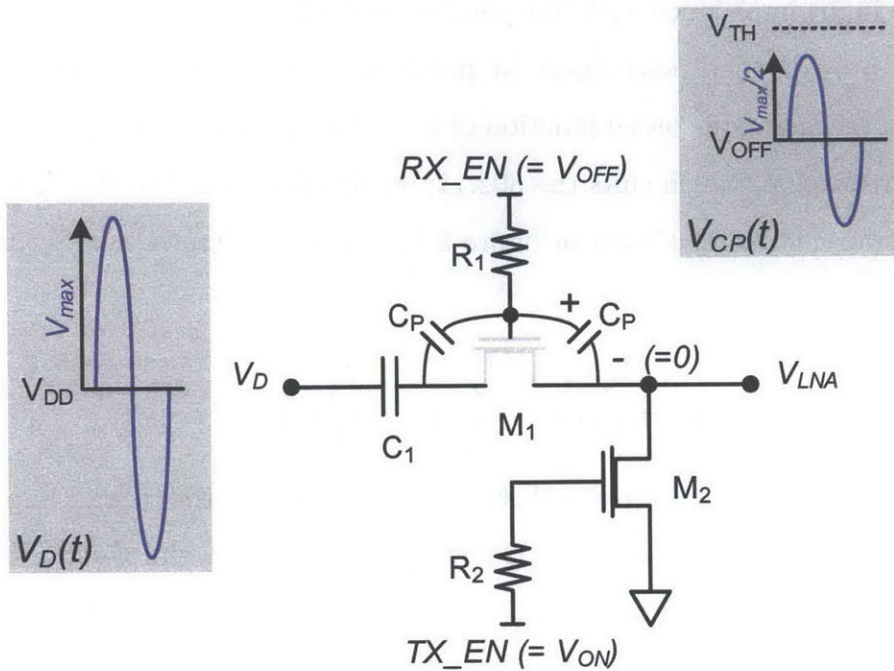
Figure 3-11: RX switch schematic showing TX mode waveforms

Where it should be noted that the threshold voltage V_{TH} is negative in this depletion-mode technology. Using the process parameters with the maximum expected V_{DD}, $V_{OFF} = -25$ V is found to provide adequate margin for this design. In choosing the switch device size, insertion loss trades off against isolation and area. The switch is sized to keep loss under 1 dB.

3.5.2   Low noise amplifier (LNA)

The primary difference between the required LNA and most other implementations is the need for a reactive input impedance Z_{LNA} which forms the conjugate power match with Z_{TX} and the RX switch. In the RX mode, the RX switch is simply a small series resistance R_{ON}, therefore the matching condition is:

$$Z_{RX}(\omega) = Re(Z_{LNA}(\omega)) + R_{ON} + jIm(Z_{LNA}(\omega)) = Z_{TX}(\omega)^* \qquad (3.19)$$

51

It is well known that inductively degenerated LNA's [44] are capable of synthesizing a reactive input impedance with independent control of real and imaginary parts. Further, they provide good noise figure by providing some resonant voltage gain before the signal reaches gate-source junction of the active device. This topology is also inherently narrowband, which suits the 802.11p application well. For the inductively degenerated transconductor shown in figure 3-12, the input impedance Z_{in} is given as:

$$Z_{in} = \omega_T L_S + j \left(\omega L_S - \frac{1}{\omega C_{gs}} \right) \qquad (3.20)$$

Depending on the technology and desired real part of Z_{in}, the overall impedance may be either capacitive or inductive. However, it is relatively simple to rotate the impedance to the correct half of the smith chart by simply adding a compensating reactance $\pm X_C$ such that:

$$Z_{LNA} = \pm jX_C + \omega_T L_S + j \left(\omega L_S - \frac{1}{\omega C_{gs}} \right) \qquad (3.21)$$
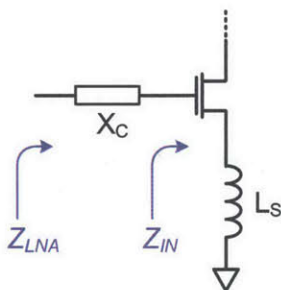
Figure 3-12: Transconductor with inductive degeneration [44]

Figure 3-13 shows the complete LNA schematic, while the corresponding component values are listed in table 3.4. The LNA active device is not downsized aggressively to maintain correlation with device models, which are not fitted to very small geometries. For the input matching network, some co-optimization is performed with the RX switch to improve performance. The output matching network performs an upward impedance transformation similar to the PA in order to extract more gain

from the device, in addition to conveniently providing both DC bias and RF signal coupling.
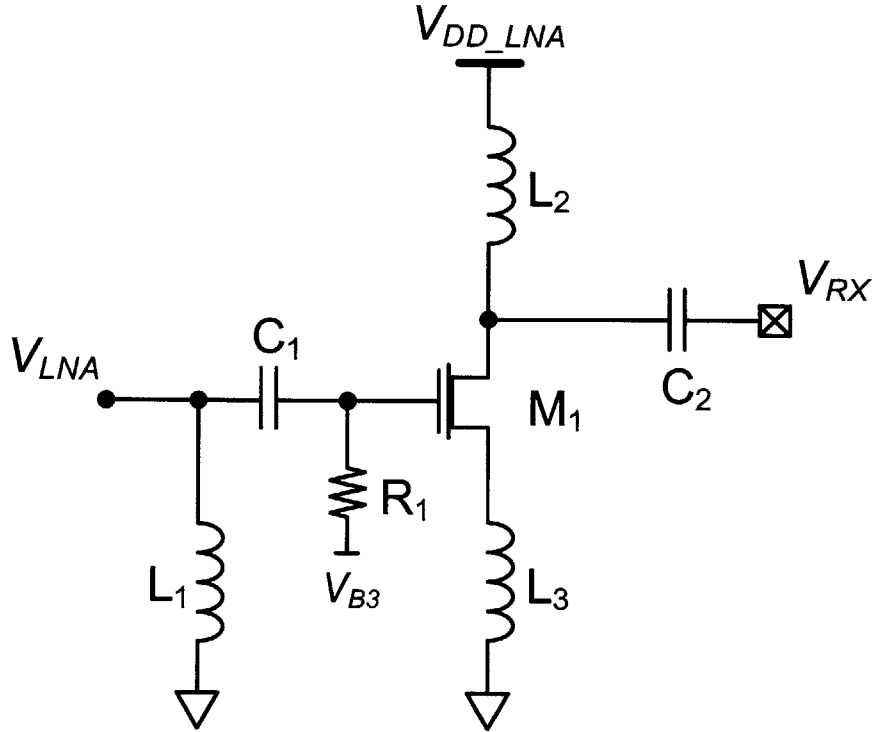
Figure 3-13: LNA schematic

Component	Value
M_1	100×2 μm
L_1	2.60 nH
L_2	2.40 nH
L_3	0.37 nH
C_1	2.20 pF
C_2	0.30 pF
R_1	2 KΩ

Table 3.4: LNA component values

3.6 Physical design

A prototype based on the proposed architecture is designed in commercially accessible GaN technology [39]. The die micrograph is shown in figure 3-14. Total die area is only 2.0 mm × 1.2 mm. The area split between the TX and RX is about 50/50 owing to the large number of passives in the RX path. Spiral inductors and metal-insulator-metal capacitors are used to design all matching networks. Co-simulation with foundry supplied device models and EM simulation-based models[5] is used for final verification of the design.

Package inductance is a major concern at 6 GHz. However, flip-chip packing is not an option as the metallized die backside must make good thermal and electrical connection to achieve specified performance. As a compromise, a combination of eutectic die attachment followed by wire bonding directly on board is selected. The PCB is designed on a high dielectric constant, low-loss material (Rogers 3010).

---

[5]for interconnect in the signal path and passive components
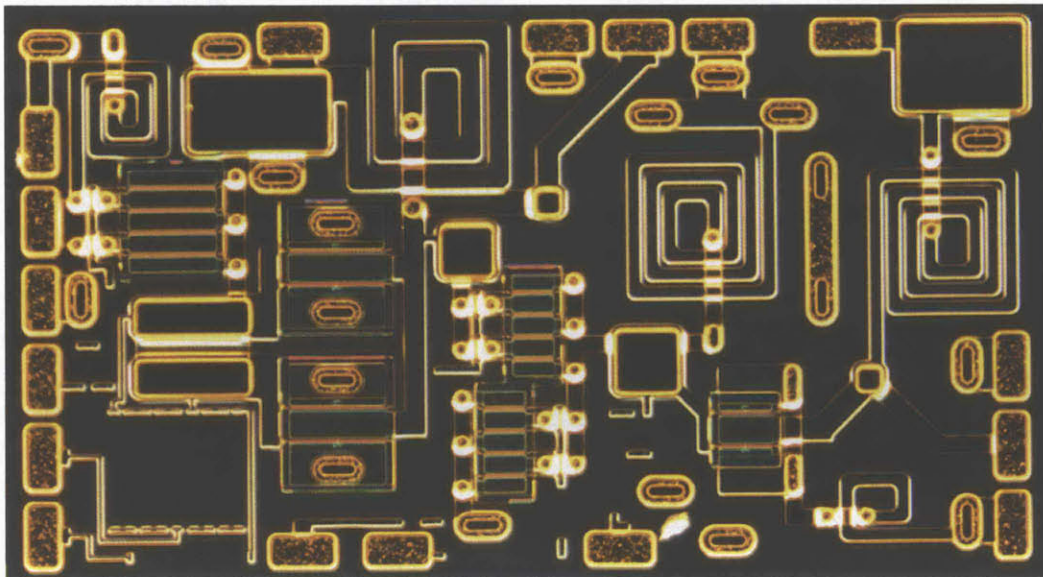
Figure 3-14: Die micrograph

3.7 Measurement results

3.7.1 Small-signal response

Figure 3-15(a) shows the small-signal response of the RF frontend in the TX mode, with the RX port terminated. Input return loss (s_{11}) is better than -20 dB, while power gain (s_{21}) is around 10 dB for the band of interest (5.850 - 5.925 GHz). Changing the port configuration and control voltages gives the small-signal response of the RX mode, which is shown in figure 3-15(b). Again good alignment with the band of interest is observed, with input return loss (s_{11}) better than -13 dB and power gain (s_{21}) around 8 dB. Noise figure is measured with a calibrated noise source, and is found to be 3.7 - 4.0 dB in-band. TX/RX path isolation (not shown) is better than -45 dB in band.

(a) TX mode   (b) RX mode
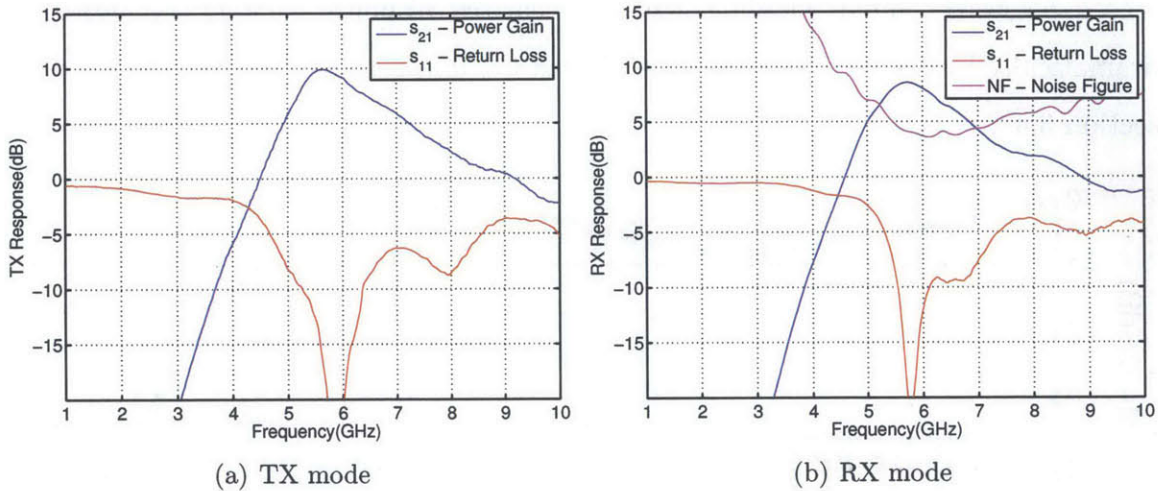
Figure 3-15: RF frontend frequency response

3.7.2 TX large-signal response

The TX path is first characterized with singe-tone excitation. High-power high-frequency measurements require some special considerations. RF signal sources do not generate enough output power to drive the TX directly. A highly linear off-chip GaN pre-amplifier with 10 dB of gain is used between the signal source and the TX.

The pre-amplifier is separately measured to have a P_{1dB} over 36 dBm, and therefore does not limit the linearity of the amplifier chain. Accurate RF power measurements are essential to benchmark TX performance. Since only in-band output power should be measured (excluding harmonics), a spectrum analyzer is used. The absolute power accuracy of the spectrum analyzer is not adequate to measure efficiency correctly. Therefore spectrum analyzer readings are calibrated using a power meter, which has its own accurate power reference built-in. All cable and trace losses are also de-embedded from the measurements.

The dual-bias for the PA transistors is tuned to maximize the P_{1dB}. Figure 3-16 shows the AM-AM response for two bias configurations consuming the same total quiescent current I_Q of 9 mA. In the Class-AB/AB case, the current partitioned equally amongst the two transistors, yielding a strongly nonlinear response. For the Class-AB/C case, one transistor carries most of I_Q while the other transistor is biased below threshold, yielding a more linearized response with IP_{1dB} improving dramatically by over 6 dB, thereby validating the linearization scheme introduced in section 3.4.2.
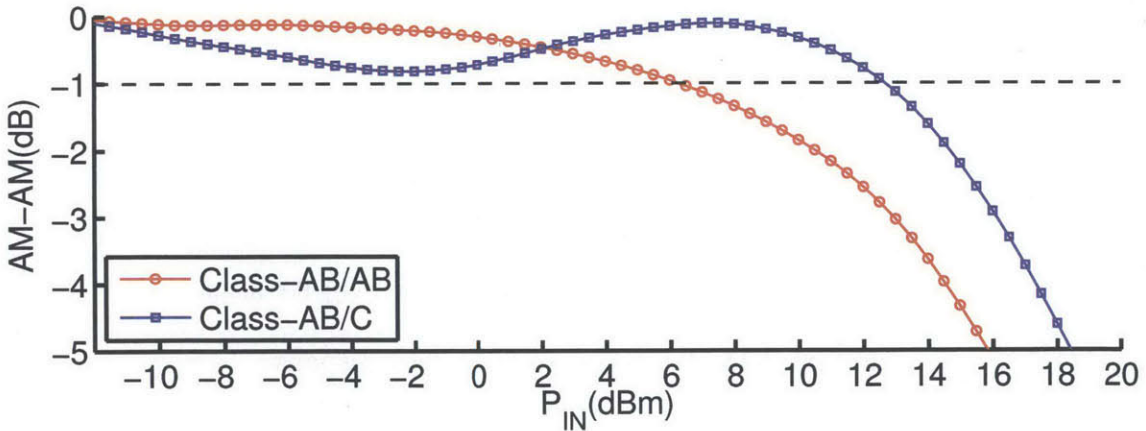
Figure 3-16: Dual-bias linearization

Figure 3-17 shows TX power sweeps with the optimized Class-AB/C bias. In additional to the nominal V_{DD} of 28 V, the TX is also tested at 37 V for direct rail connection in power-over-ethernet (PoE) scenarios. Key performance metrics are tabulated in table 3.5.
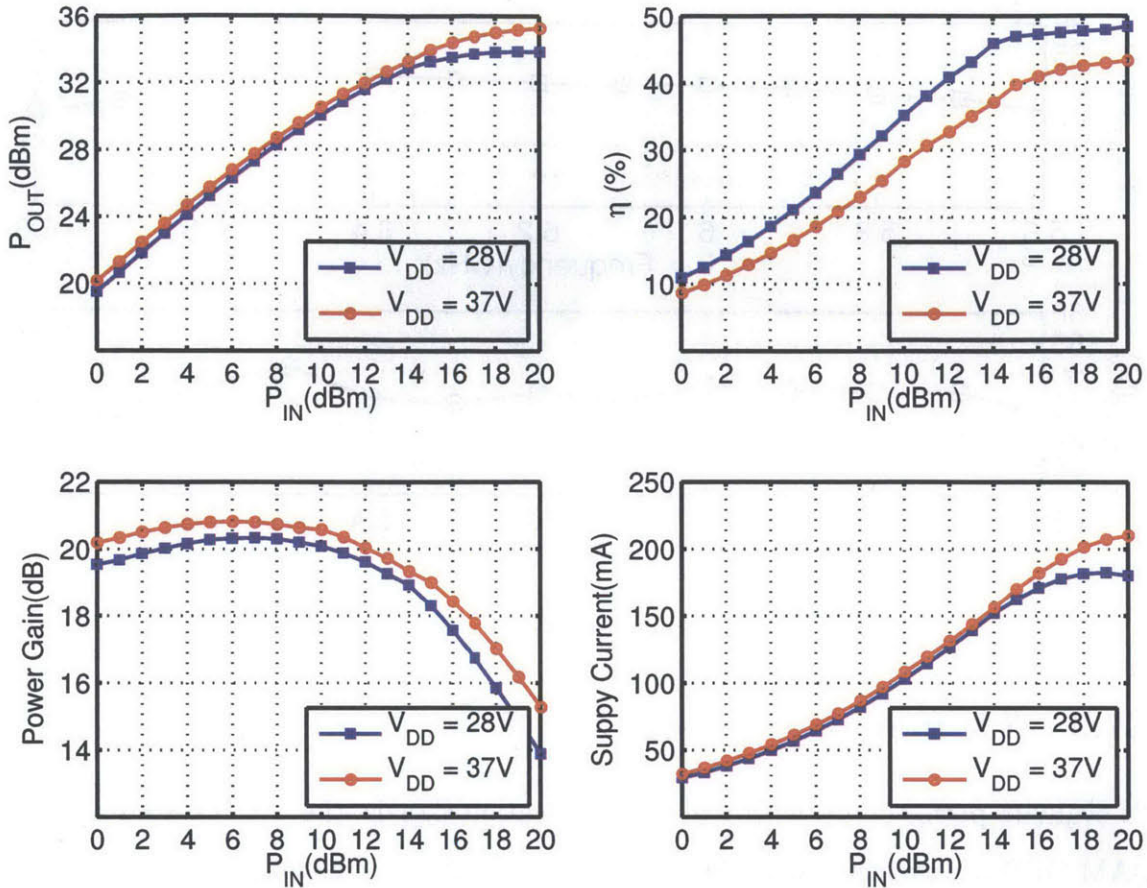
Figure 3-17: TX power sweep

V_{DD} (V)	28.0	37.0
I_Q (mA)	9.0	9.0
OP_{1dB} (dBm)	33.1	33.7
P_{SAT} (dBm)	33.9	35.3
η_{SAT} (%)	48.5	43.8

Table 3.5: TX CW performance

TX performance is also investigated over frequency. The TX is driven into saturation to characterize P_{SAT} and η_{SAT}. Results shown in figure 3-18 reveal a 1 dB BW of 1 GHz (5.7 - 6.7 GHz) or about 15 %. The best η_{SAT} and P_{SAT} of 50.8 % and 34.5 dBm are measured at 6.5 GHz. Since the design is optimized for 5.9 GHz, this indicates that the fabricated value of passives in the output match is probably lower than nominal.
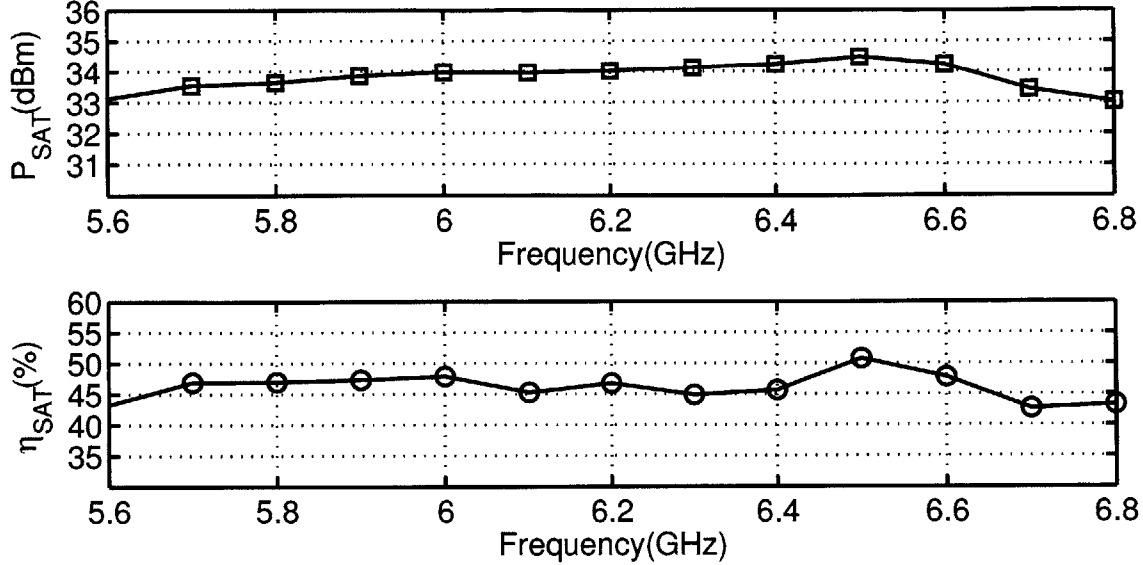
Figure 3-18: TX saturated power and efficiency over frequency

3.7.3 TX modulated tests

To evaluate performance with real communication signals, the TX is tested with 64-QAM OFDM waveforms with 20 MHz BW[6] and PAPR exceeding 7 dB. The average power level at the input is gradually raised to drive the PA into compression. As shown in figure 3-19, output power and efficiency improve with increased drive power, while EVM degrades due to nonlinearities that arise from large-signal operation. Figure 3-20 shows TX performance at the 802.11 EVM limit of -25 dB. The TX complies with the 802.11 spectral mask, while achieving an average efficiency of 30 % with 27.8 dBm output power, without applying any digital predistortion.

3.7.4 RX large-signal response

The RX is also tested for linearity. Figures 3-21(a) and 3-21(b) show the large-signal single-tone and two-tone response, respectively with $V_{DD\_LNA} = 12$ V and $I_Q = 5.5$ mA. The output referred 1 dB compression point (OP_{1dB}) and third-order intercept (OIP_3) lie at 10.8 dBm and 22.0 dBm, respectively.

---

[6]Even though 802.11p has a fixed BW of 10 MHz, tests are performed at 20 MHz to explore possibility of co-existence with 802.11n systems.
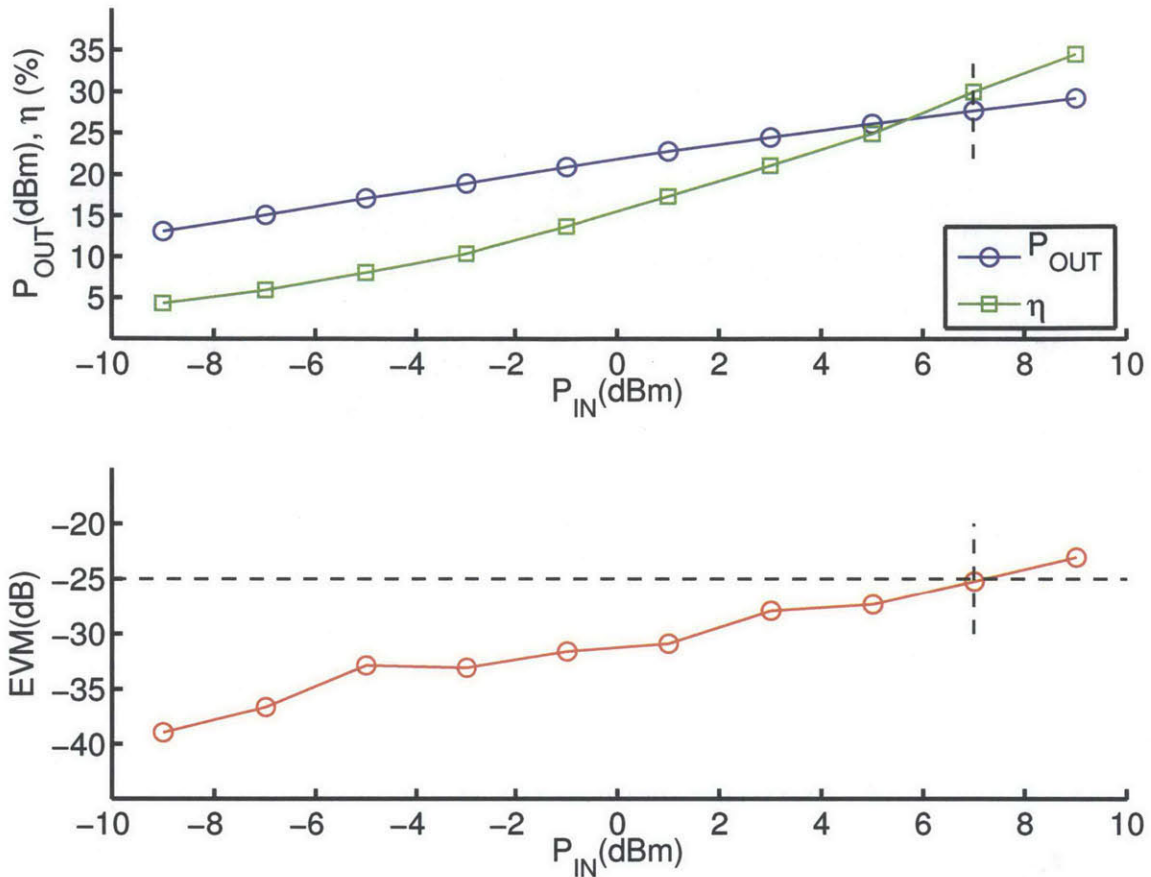
Figure 3-19: Power sweep with 20MHz BW OFDM signal at 5.875 GHz
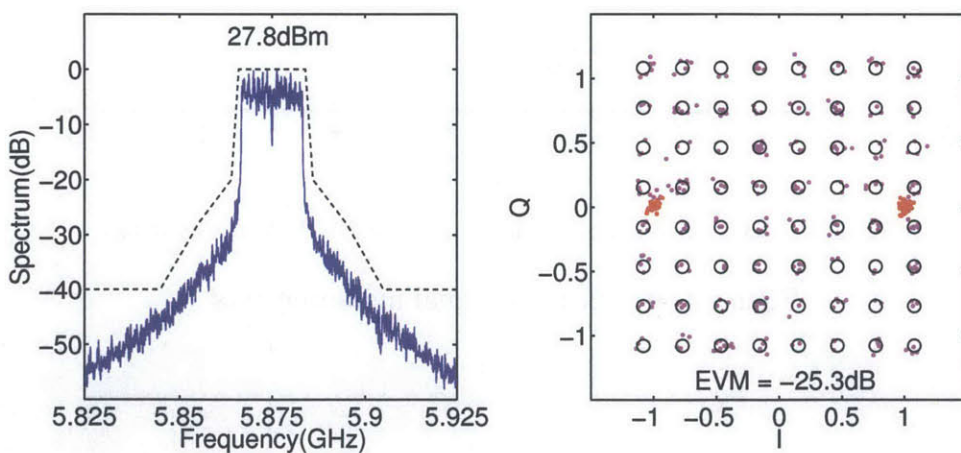
Figure 3-20: TX performance with OFDM signal at -25 dB EVM
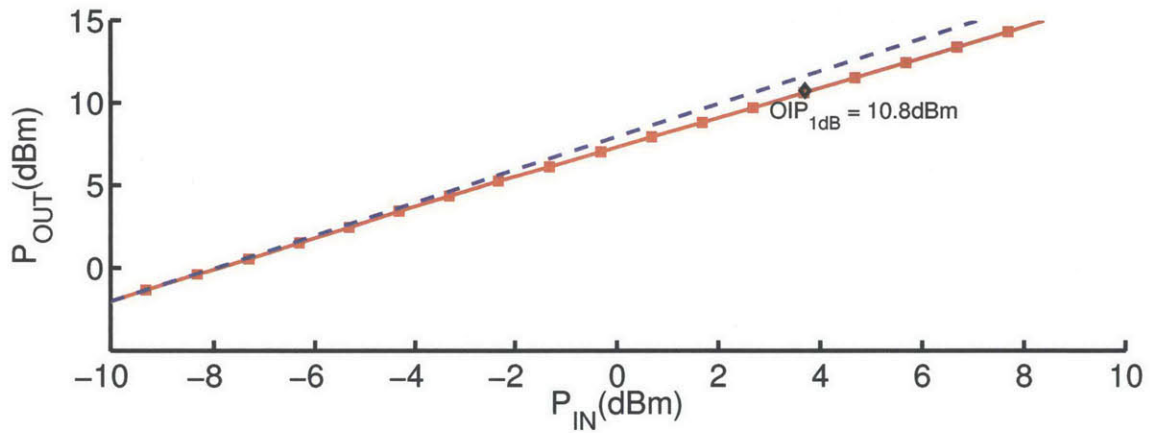
(a) OP_{1dB} - Output 1dB compression point
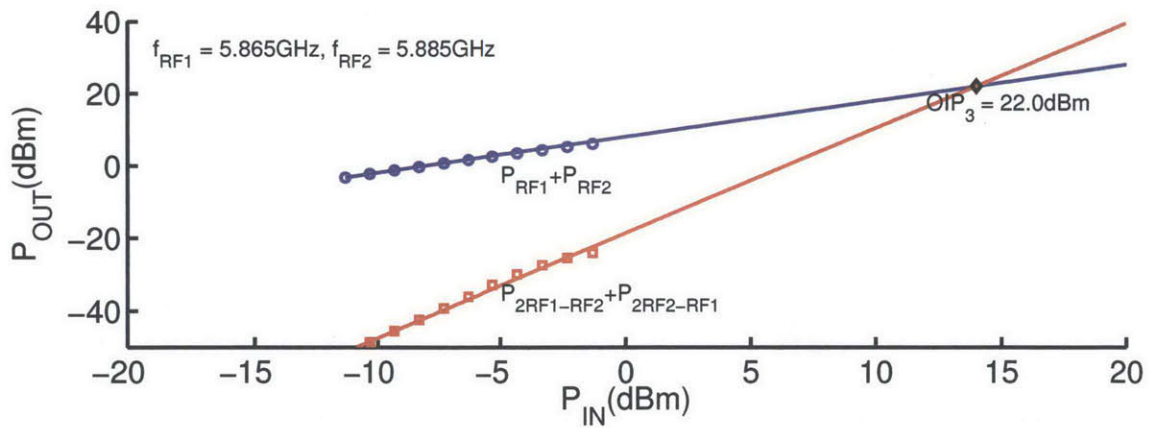
(b) OIP_3 - Output third-order intercept with 20 MHz tone spacing

Figure 3-21: Rx large-signal measurements

3.8 Comparison with other published work

Table 3.6 compares the achieved TX performance with some other recent work. Since the focus here is on a fully integrated and compact solution, discrete designs are not considered, as they occupy areas that are two to three orders of magnitude larger. [34] and [45] represent recently reported 802.11a (4.9 - 5.9 GHz) PA's with highest saturated efficiency and output power of 32.1 % (at 26 dBm) and and 30.3 dbm (at 19.4 %), respectively in CMOS technology. [46] reports a switch-based digital GaN TX with higher output power. It is seen that the this work easily outperforms CMOS with a large margin even without predistortion due to the superior device properties of GaN, and offers significantly better average efficiency than the switch-based GaN architecture, thereby validating the proposed G_m linearization technique. Finally, it is worth noting that the output power and efficiency numbers reported in this work include the TX switch, unlike the other publications which do not.

Parameter	[34]	[45]	[46]	This work
Application	WLAN	WLAN	Generic QAM	802.11p/n
Process	45-nm CMOS	65-nm CMOS	200-nm GaN	250-nm GaN
Architecture	Class-AB	Class-AB	Switching	Class-AB/C
Integrated TX/RX SW	No	No	No	Yes
Freq. band range (GHz)	4.9 - 5.9	4.9 - 5.9	7	5.7 - 6.7
$V_{DD}(V)$	-	6.35/4.80	28	28
P_{SAT} (dBm)	26.0	30.3/28.2	37.0	33.9
η_{SAT} (%)	32.1	19.4/24.1	-	48.5
P_{OUT} (dBm)	18.7	-	33.2	27.8
η (%)	-	-	19.1	30.0
IQ-BW (MHz) / PAPR (dB)	20 / -	-	20 / 6.2	20 / 7.3
EVM (dB)	-25.0	-	-28.0	-25.3
Predistortion	Yes	No	Yes	No

Table 3.6: TX performance comparison with other fully integrated solutions

Chapter 4

A Frequency-agile RF Frontend for Multi-band TDD Radios in 45-nm SOI CMOS

Mobile devices targeting ubiquitous connectivity need to support a growing number of frequency bands spanning multiple octaves. In the past, time division duplexing (TDD) was prevalent for internet connectivity (WLAN, Bluetooth etc.), while frequency division duplexing (FDD) was primarily used for voice (GSM, WCDMA etc.) applications. The ongoing adoption of the LTE shows a trend of convergence. The current LTE specification [47] lists 40 bands for worldwide operation, ranging from 0.7 to 3.8 GHz. While radio transceivers have evolved towards the software-defined radio vision of minimalistic and reconfigurable hardware, state-of-the-art RF frontends still employ dedicated, redundant hardware for each band; an approach which is not scalable to future needs. Therefore, it is imperative that reconfigurable RF frontend architectures be developed to cover the maximum number of bands with minimal hardware while achieving competitive performance. This chapter presents a new *frequency-agile* RF frontend architecture for multi-band TDD applications to meet this challenge.

This chapter is organized as follows: section 4.1 discusses prior art and identifies key shortcomings in the context of multi-band operation; section 4.2 presents the

proposed architecture, and outlines key innovations which address these shortcomings. Section 4.3 provides a brief overview of TDD operation and how the proposed architecture supports it. Sections 4.5 and 4.6 detail transmitter (TX) and receiver (RX) design, respectively. In section 4.7, custom design of RF passive components is detailed. The impact of CMOS scaling in improving the performance of the proposed architecture is theoretically analyzed in section 4.8. Physical design aspects such as layout and ESD protection are briefly discussed in section 4.9. Sections 4.10 and 4.11 present measurement results obtained from the prototype IC. Finally, in section 4.12 the achieved results are put into perspective through comparisons with prior work reported in prominent publications.

4.1 Conventional multi-band TDD architecture

Figure 4-1 shows a conventional multi-band mobile TDD radio. It consists of a software-defined CMOS transceiver (i.e. baseband and up/down-conversion circuitry) such as those recently reported [27], in conjunction with multiple narrowband power amplifier (PA) and low noise amplifier (LNA) modules, which are switched in and out depending on the desired frequency band. In the context of upcoming wireless standards like LTE, this approach has several shortcomings:

1. **Redundant components:** The use of multiple PA's and LNA's (typically one each per band) arises from the narrowband behavior inherent in resonant RF circuits. However, this approach is not scalable to dozens of bands required for future systems.

2. **Use of multiple device technologies:** As high-power PA's and RF switches are challenging to implement in CMOS, such implementations often combine heterogeneous device technologies like CMOS, SiGe and GaAs in a multi-chip module. Clearly, this is not a cost- or area-optimal architecture.

3. **Analog intensive design:** Analog RF circuits have limited reconfigurability and do not always benefit from CMOS scaling.
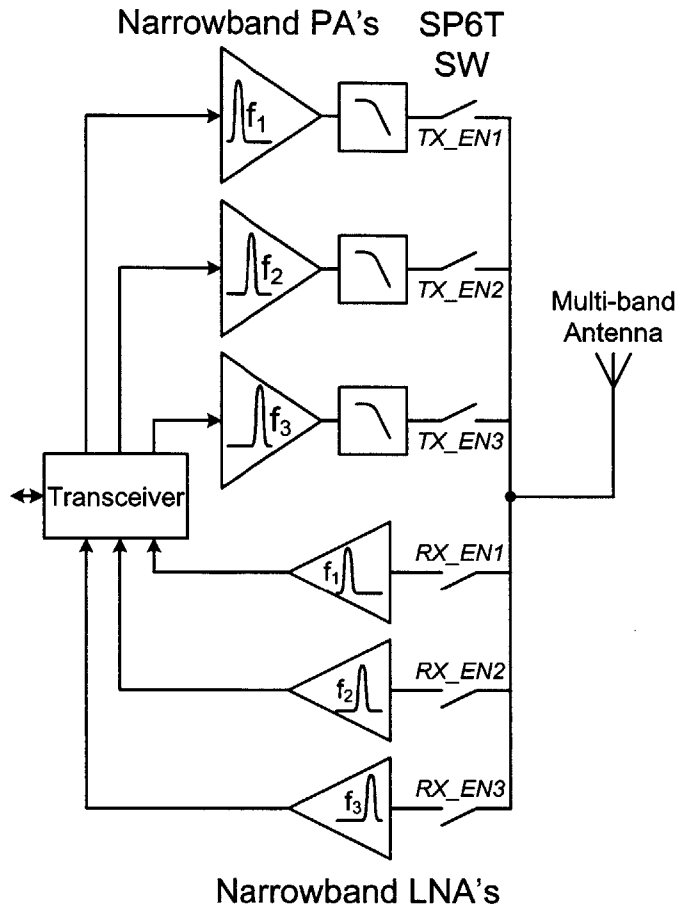
Figure 4-1: Conventional multi-band RF frontend architecture

4. **Complicated switching scheme:** To select among multiple PA's and LNA's, the RF frontend requires a multi-port RF switch which consumes significant die area and adversely impacts TX output power, efficiency and linearity.

4.2   Proposed frequency-agile architecture

The proposed architecture shown in figure 4-2 takes a more minimalistic and flexible approach to address shortcomings of the conventional architecture with *four key innovations*:

1. **Frequency-agile architecture:** Only a single PA and LNA, whose frequency response can be digitally tuned over a wide frequency range are used to cover
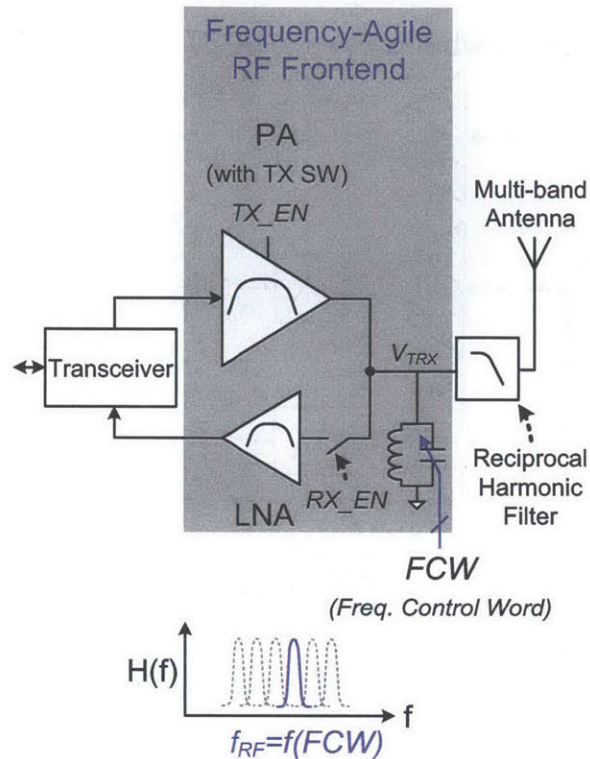
Figure 4-2: Proposed RF frontend architecture

multiple bands.

2. **Monolithic integration:** All blocks are integrated on a single-chip in a silicon-on-insulator (SOI) CMOS technology.

3. **Maximally digital design philosophy:** Both PA and LNA unit cells utilize broadband topologies inspired by the CMOS inverter. As a result, key performance parameters like PA efficiency and tuning range directly benefit from the continual scaling of CMOS technology.

4. **Simplified switching scheme:** The switching scheme is simplified two-fold. First, the number of ports is reduced to just two as only one PA and LNA are used. Second, the TX branch RF switch is absorbed into the PA itself without any power loss penalty.

4.3 TDD operation and proposed TX/RX switching scheme

Time division duplexing or TDD, illustrated in figure 4-3(a) implies that the radio works in half-duplex mode i.e. at a given time, it either transmits or receives. Therefore, the same frequency can used for both transmit and receive. This arrangement is often called unpaired spectrum usage to differentiate it from frequency division duplexing or FDD, illustrated in figure 4-3(b), which uses a pair of frequency bands to transmit and receive simultaneously.

To share a common antenna for both TX and RX while supporting multiple bands, TDD radios use a single-pole-multi-throw switch. It is worth noting that the switch shown in figure 4-1 performs both band-selection and time-multiplexing, while in the proposed architecture of figure 4-2 it does only the latter, since a single PA and LNA cover all bands.
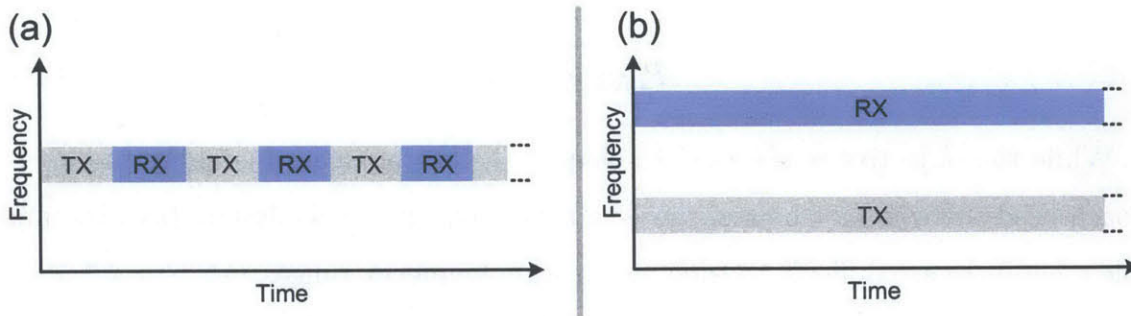
Figure 4-3: Radio duplexing schemes (a) Time division duplexing or TDD (b) Frequency division duplexing or FDD

Non-idealities in TX/RX switching circuitry adversely impact overall performance and have already been given detailed treatment in section 3.3.1. Recall that the TX branch switch (TXSW) is particularly challenging to design, if its loss is denoted by $P_{L,TXSW}$ (in dB), then the impact on overall TX output power and efficiency is given as:

$$P_{OUT} = P_{OUT,PA} \times 10^{-(\frac{P_{L,TXSW}}{10})} \qquad (4.1)$$

$$\eta = \eta_{PA} \times 10^{-(\frac{P_{L,TXSW}}{10})} \tag{4.2}$$

Where $P_{OUT,PA}$ and η_{PA} denote output power and efficiency of the standalone PA. Despite the design challenges mentioned in section 3.3.1, such switches have been demonstrated in SOI technology. Generally, power handling and insertion loss trade off against each other. Reported designs in SOI achieve sufficient linearity with insertion loss in the region of 0.5 - 1.0 dB at GHz frequencies [43] [48]. Loss also increases with the number of branches (throws) required, which is a direct consequence of the number of PA's that need to be switched in and out to support all required bands, again reinforcing the assertion that the concept of figure 4-1 is not scalable to dozens of bands.

Recall from section 3.3.1 that even overcoming even 1 dB of loss can *dramatically* lower TX energy consumption by 20 % while maintaining the same P_{OUT}. Therefore, like chapter 3, this work also aims for:

$$P_{L,TXSW} = 0 \tag{4.3}$$

While the objective is identical to chapter 3, the conjugate matching based approach used there cannot be transposed to this frequency-agile design, because multiple reactances are difficult to tune over a wide frequency range.

The modified switching scheme used in this work can be understood by considering the two modes of operation, shown in figure 4-4. Both the PA and LNA cores have a broadband response. The LC resonator at the common TX/RX port V_{TRX} determines the frequency response of the system, and can be digitally tuned for band selection by the frequency control word (FCW).

1. **TX mode:** V_{TRX} serves as the RF output. The PA transmits data with high output power, leading to very high voltage swings. For instance, V_{TRX} when loaded by a 50 Ω antenna would swing 14 V_{PP} at 27 dBm output power. The RX path is inactive and isolated from V_{TRX} by the RX switch, which is specifically designed to be tolerant of high voltages in its off state.

68

2. **RX mode:** V_{TRX} serves as the RF input. The PA is switched off, and is designed (without adding extra elements in the signal path) such that it presents a high output impedance $Z_{O,PA}$ looking back from V_{TRX}. The RX branch switch is on, and the incident RX signal from the antenna passes to the RX signal path, which is matched to 50 Ω to facilitate maximum power transfer.
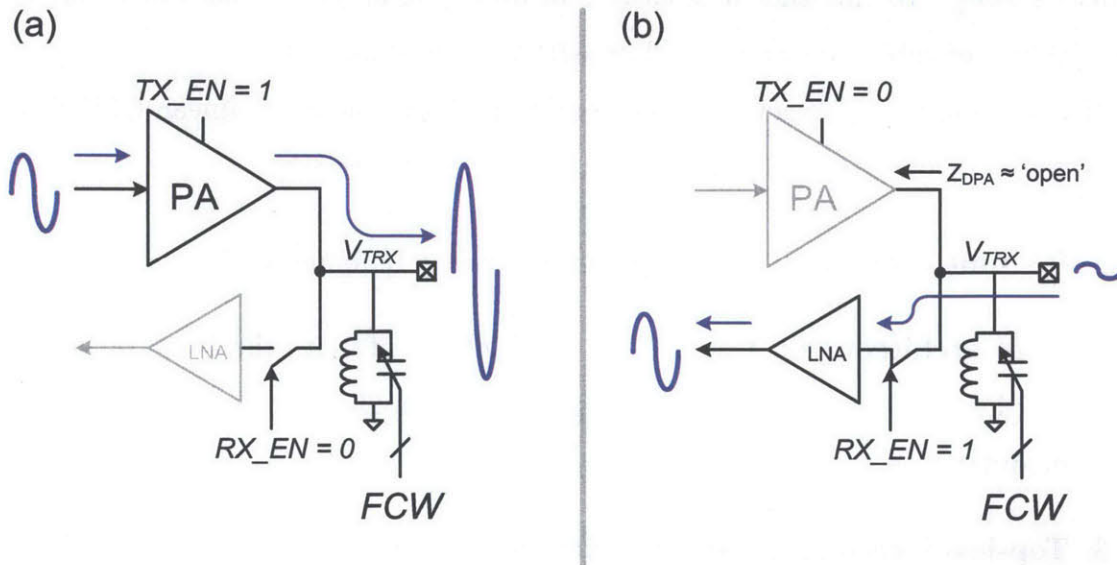
Figure 4-4: Proposed TX/RX switching scheme (a) TX mode (b) RX mode

Since the architecture requires no explicit TXSW, it overcomes the loss attributed to it, thereby improving TX output power, efficiency and linearity.

4.4 Target frequency range

The primary goal of this design is to cover the maximum number of TDD-LTE bands. From the LTE specification [47], it is observed that 10 out of 12 TDD-LTE bands lie in the 1.8 - 3.6 GHz range. This range also conveniently covers WLAN (802.11g) (2.4 - 2.5 GHz) and WiMax (2.3 - 2.7 GHz and partially 3.3 - 3.8 GHz). It therefore makes sense to center the design at the geometric mean of the 1.8 - 3.6 GHz range i.e. $f_C = 2.5$ GHz, and then attempt to maximize the frequency range around the center frequency with architectural innovation and careful design.

4.5 TX design

The primary design goal for the TX is to maximize frequency range while maintaining competitive efficiency, linearity and output power. Simultaneously satisfying all four requirements makes the design particularly challenging. Further, to be future-proof, the architecture should be digital friendly with performance directly benefiting from CMOS scaling. To this end, it is clear that switching amplifier based architectures are highly preferable over analog (Class-AB) style implementations.

If a bottom-up approach is taken, outlining a switch-based, linear CMOS TX architecture involves design decisions at three levels of hierarchy:

1. **PA unit cell:** The basic, RF power generating building block.

2. **Core architecture:** The smallest combination of unit cells (with or without supplemental circuitry) that yields a linear amplifier e.g. outphasing, polar modulation etc.

3. **Top-level architecture:** Matching network, power control and combining scheme.

In the following sections, a detailed and clear rationale for design choices made at each of these levels is presented. Mathematical analysis is presented where appropriate to develop design insight.

4.5.1 PA unit cell

For the PA unit cell, the voltage-mode Class-D structure is chosen for three main reasons:

1. It is inherently digital, with a CMOS inverter-like topology that only relies on switching performance of transistors.

2. A simple drive scheme can be implemented by smaller replica stages to design a broadband, high-gain RF signal chain.

3. The natural turn-off (tri-state) ability can be exploited to eliminate the explicit
   TX switch in a TDD frontend, this feature is detailed in section 4.5.4.
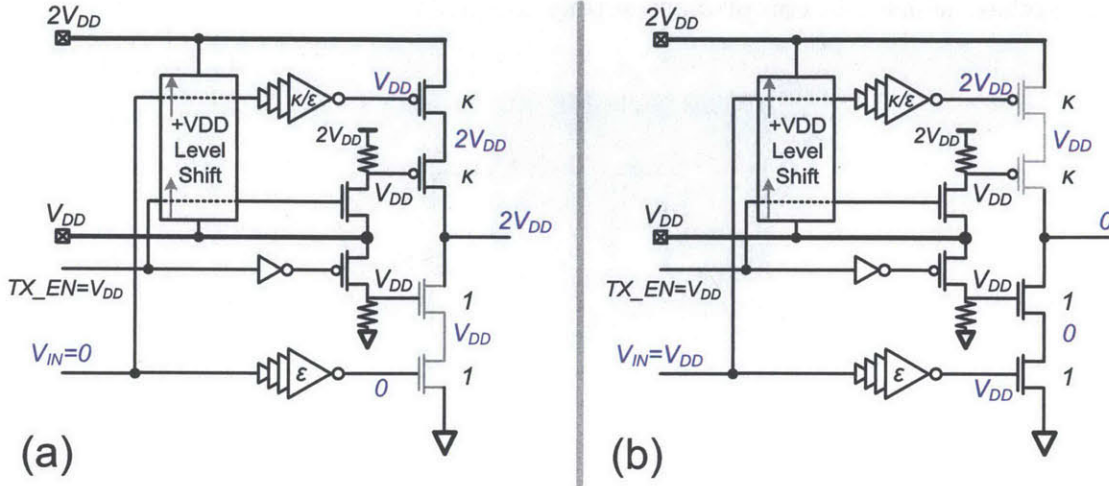
Figure 4-5: Class-D PA unit cell in both TX mode RF switching states

Since a high-power PA (> 27 dBm) is targeted, it is prudent to maximize the output power available from each unit cell. A modified version of the stacked Class-D topology reported in [22] [49] [50] is used in this work to enable RF switching between 0 to $2V_{DD}$ at the output, thereby boosting the unit cell output power by 6 dB compared to conventional CMOS inverter. Given their superior switching performance, only thin-gate transistors are used. Figure 4-5 shows the two steady state conditions, where no gate-source or gate-drain junction sees voltages beyond V_{DD} while operating at $2V_{DD}$. Intermediate devices in the stack which are connected to the output see a slightly higher drain-source voltage for a short time during switching transients, but this does not pose a reliability issue. To aid the analysis which follows, device sizing is normalized to the final stage NMOS device. A DC level-shifted replica of the RF input V_{IN} feeds the P-side driver logic, which operates between V_{DD} and $2V_{DD}$. Both N- and P-side drivers include multiple stages of tapered CMOS inverters, with a conservative post-layout fanout ($\frac{1}{\epsilon}$) of 2.5. The last stage in both the N- and P-side driver chains is sized asymmetrically (skewed) to reduce crowbar current [51], which results from both N- and P-side transistors being simultaneously on for a short

time during switching transients. Minimizing crowbar current is important for two reasons. First, crowbar current reduces PA efficiency as it simply flows from supply to ground without contributing any RF output power and second, it results in large current spikes, which can compromise supply integrity.
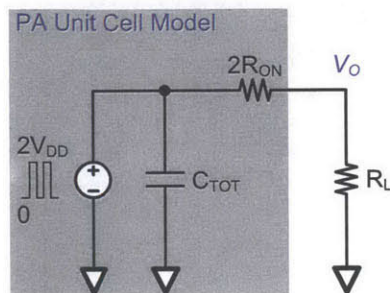
Figure 4-6: Simplified electrical model of the PA unit cell

The operation of the PA unit cell can be analyzed by studying the simplified electrical model shown in figure 4-6. Assuming that each unit cell is loaded by an impedance R_L, the output power P_O can be calculated. The output waveform is a square-wave with amplitude V_{DD}. From fourier-series analysis, the voltage amplitude produced at the fundamental frequency can be calculated as $\frac{4}{\pi}V_{DD}$. The devices are sized in such a way that the on resistance of the PA unit cell is a small fraction ν of the load impedance:

$$\nu = \frac{2R_{ON}}{R_L} \tag{4.4}$$

The voltage appearing across the load is attenuated by the factor $\frac{1}{1+\nu}$ because of resistive voltage division:

$$V_O = \frac{4}{\pi}V_{DD}\frac{1}{1+\nu} \tag{4.5}$$

The RF output power, as a function of R_L is:

$$P_O(R_L) = \frac{V_O^2}{2R_L} = 8\left(\frac{V_{DD}}{\pi}\right)^2 \frac{1}{R_L(1+\nu)^2} \tag{4.6}$$

Assuming that crowbar current is minimized by design, there are two remaining sources of loss that limit PA unit cell efficiency:

1. **Conduction loss** due to finite on resistance of the PA unit cell switches. Since the same current flows through the switch on resistance and the load, equation 4.4 can be used to write the conduction loss as:

$$P_{L,res} = \nu P_O \tag{4.7}$$

2. **Switching loss** due to finite capacitances that need to be charged and discharged once every RF cycle. From figure 4-5, it is clear that gate-source capacitances are charged from 0 to V_{DD}, while gate-drain capacitances are charged from $-V_{DD}$ to $+V_{DD}$, which makes them effectively $4\times$ larger. In SOI processes, capacitances to bulk are small and are neglected in this analysis. If the ratio of NMOS to PMOS strength is denoted as κ, then the total equivalent capacitance discharged from 0 to V_{DD} in the PA unit cell final stage is simply:

$$C_{DPA} = 2(1 + \kappa)(C_{GS} + 4C_{GD}) \tag{4.8}$$

There is additional capacitance attributed to the N- and P-side driver chains. Relative to the final stage, each driver stage is sized successively smaller by a factor ϵ (< 1). Since several tapered stages are usually employed, a capacitance multiplication factor for all the driver stages can be computed by using the Maclaurin series:

$$\frac{\epsilon}{1 - \epsilon} = \epsilon + \epsilon^2 + ... \tag{4.9}$$

Additional driver capacitance is therefore:

$$C_{DRV} = \frac{\epsilon}{1 - \epsilon}(1 + \kappa)(C_{GS} + 4C_{GD}) \tag{4.10}$$

73

Adding equations 4.8 and 4.10, the total equivalent capacitance is:

$$C_{TOT} = C_{DPA} + C_{DRV} = \left(\frac{2-\epsilon}{1-\epsilon}\right)(1+\kappa)(C_{GS} + 4C_{GD}) \tag{4.11}$$

The switching loss can be written as:

$$P_{L,cap} = C_{TOT}V_{DD}^2 f_{RF} \tag{4.12}$$

Considering the above losses, the *instrinsic*[1] efficiency of the PA unit cell is:

$$\eta_D = \frac{P_O}{P_O + P_{L,res} + P_{L,cap}} \tag{4.13}$$

4.5.2 Core architecture

Switching amplifiers such as Class-D can only support constant envelope modulation schemes (e.g BPSK) when deployed in a standalone configuration [9]. Since modern communication standards almost always employ non-constant envelope signals (e.g. 2^X-QAM, OFDM), a more elaborate architecture which enables linear amplification with switching amplifiers should be used. The PA core is the simplest combination of unit cells that can handle both amplitude and phase modulation. The two most promising architectures to form the PA core are outphasing [20] and polar modulation [23]. For the present design, outphasing is chosen over polar modulation because:

1. Outphasing offers the potential for wider modulation bandwidths compared to polar architectures, as the latter typically require simultaneously high bandwidth and linear envelope amplifiers[2].

2. Outphasing works particularly well with Class-D unit cells. Specifically, the low output impedance of the Class-D architecture renders it relatively immune to

---

[1]Limited mainly by device technology used for circuit implementation i.e. neglecting losses in matching network and other non-idealities.

[2]It is noteworthy that recently polar modulation with Class-D cells [52] without an explicit envelope amplifier has emerged as a interesting technique. However, difficulty in implementing tunable, high-power floating capacitors makes this architecture less attractive for frequency-agile design, which is the present focus.

reactive loading, thereby enabling the use of non-isolating and ideally lossless power combiners [22] [53], resulting in higher average efficiency.

3. Significant research has been conducted into all-digital phase modulators, making outphasing a promising approach for fully digital radios [54] [55] [56].

Figure 4-7(a) shows the implementation of the outphasing PA core, which consists of two Class-D PA unit cells and a transformer. The transformer performs the amplitude vector combining necessary for outphasing. For the purpose of this section it is adequate to assume an ideal, 1:1 transformer.
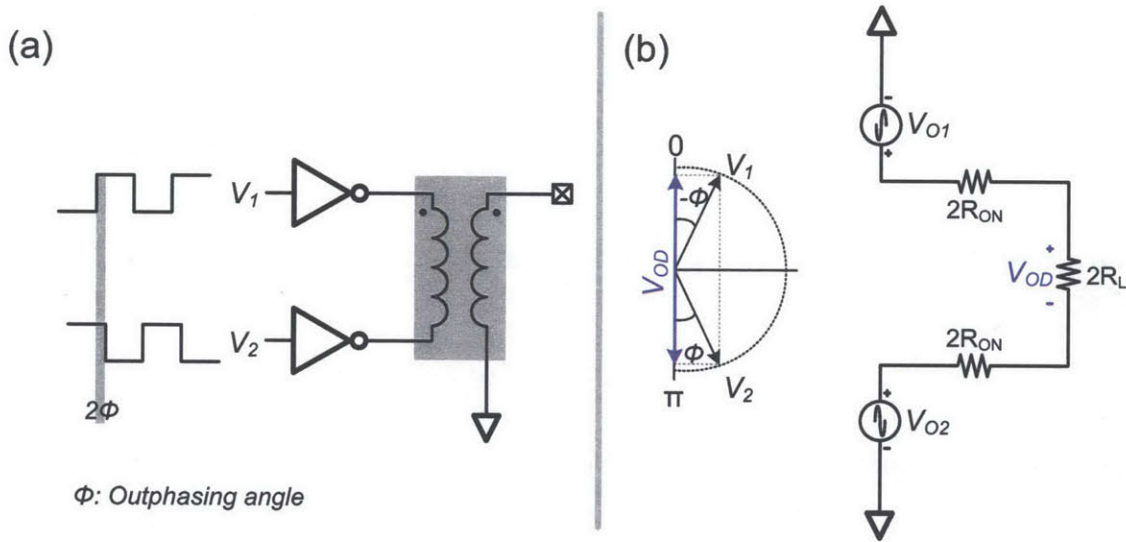
Figure 4-7: Outphasing PA core (a) Implementation (b) Equivalent electrical model at fundamental frequency

The two PA unit cells are driven by rail-to-rail CMOS signals which contain only phase information:

$$V_1 = \frac{V_{DD}}{2}[1 + sgn(cos(\omega t + \theta - \phi))] \tag{4.14}$$

$$V_2 = \frac{V_{DD}}{2}[1 - sgn(cos(\omega t + \theta + \phi))] \tag{4.15}$$

Where sgn is the signum function, θ is the phase of the incoming RF signal and ϕ is the outphasing angle.

75

Outphasing action can be understood by studying the simplified electrical model of figure 4-7(b), which is valid at the fundamental frequency ($f = \frac{\omega}{2\pi}$). The two Class-D PA unit cells produce square-wave output waveforms, whose fundamental frequency representation is simply given by:

$$V_{O1} = \frac{4}{\pi}V_{DD}cos(\omega t + \theta - \phi) = \frac{4}{\pi}V_{DD}[cos(\phi)cos(\omega t + \theta) + sin(\phi)sin(\omega t + \theta)] \quad (4.16)$$

$$V_{O2} = -\frac{4}{\pi}V_{DD}cos(\omega t + \theta + \phi) = \frac{4}{\pi}V_{DD}[-cos(\phi)cos(\omega t + \theta) + sin(\phi)sin(\omega t + \theta)] \quad (4.17)$$

Since the load resistor $2R_L$ appears differentially between the two PA unit cells, the common-mode voltage component (second term) in equations 4.16 and 4.17 produces no current. The output voltage waveform is simply:

$$V_{O1} - V_{O2} = \frac{8}{\pi}V_{DD}cos(\phi)cos(\omega t + \theta) \quad (4.18)$$

The waveform amplitude across the load is:

$$V_{OD} = \frac{8}{\pi}\frac{1}{1+\nu}V_{DD}cos(\phi) \quad (4.19)$$

The output power can be calculated in a manner similar to the previous section:

$$P_{OD} = \frac{V_{OD}^2}{4R_L} = 16\left(\frac{V_{DD}}{\pi}\right)^2\frac{cos^2(\phi)}{R_L(1+\nu)^2} \quad (4.20)$$

Comparing equations 4.6 and 4.20, P_{OD} may be re-written as:

$$P_{OD} = 2P_O(R_{L,mod}) \quad (4.21)$$

$R_{L,mod}$ is simply the real part of the effective load impedance seen by the two unit cells that comprise the outphasing PA core, modulated by the outphasing angle ϕ in response to RF signal amplitude:

76

$$R_{L,mod} = \frac{R_L}{cos^2(\phi)} \tag{4.22}$$

From equation 4.22, it is clear that this implementation of outphasing relies on *load modulation*. Clearly, peak power occurs at $\phi = 0$, and can be modulated all the way to zero with $\phi = \frac{\pi}{2}$. Details of how the ϕ is mathematically derived from the original RF signal can be found in appendix C.

4.5.3 Top-level architecture

PA impedance transformation

Achieving high PA output power in a deep submicron CMOS process is challenging due to the large impedance transform required from the antenna impedance Z_0 ($= 50\ \Omega$) to the optimal PA impedance R_{opt} needed to extract this power from the low supply voltage. R_{opt} can be calculated from equation 4.6 by substituting $R_L = R_{opt}$ and $P_O = P_{SAT}$[3]. Assuming $V_{DD} = 1$ V, $P_{SAT} = 30$ dBm[4] and $\nu = 0.15$, we obtain $R_{opt} = 0.61\ \Omega$. The required impedance transformation ratio is therefore:

$$\chi = \frac{Z_0}{R_{opt}} = 83 \tag{4.23}$$

Tunable matching network

It is well known that with traditional L and π impedance matching topologies, a tradeoff exists between transformation ratio χ and network efficiency [11]. Large values of χ, as required here to extract high output power from a low supply adversely impact efficiency. Further, considering that the ultimate goal is a *tunable* matching network, L and π topologies are unattractive because they would require multiple,

---

[3]This calculation does not imply that a single PA unit cell (section 4.5.1) is required to produce all of the output power. As previously discussed (section 4.5.2), a minimum of two unit cells are needed to implement outphasing. R_{opt} is simply the effective impedance presented to the supply by all the unit cells that comprise the PA working in unison; with $2n$ PA unit cells operating in parallel, each cell will see an impedance $2n$ times larger ($= 2nR_{opt}$).

[4]Target saturated output power of 28 dBm + 2 dB margin for on-chip losses

floating and tunable high quality factor (Q) RF inductors and capacitors, which are very difficult to implement.

In this work, a tunable version of the transformer series-combining topology [57] [11], compatible with high output power is proposed for three main reasons:

1. Being transformer based, it is inherently compatible with the outphasing PA core presented in section 4.5.2

2. The number of tunable passives in minimized to just one ground referenced capacitor

3. It has been analytically shown that unlike L and π topologies, the efficiency of this approach is independent of χ [11]
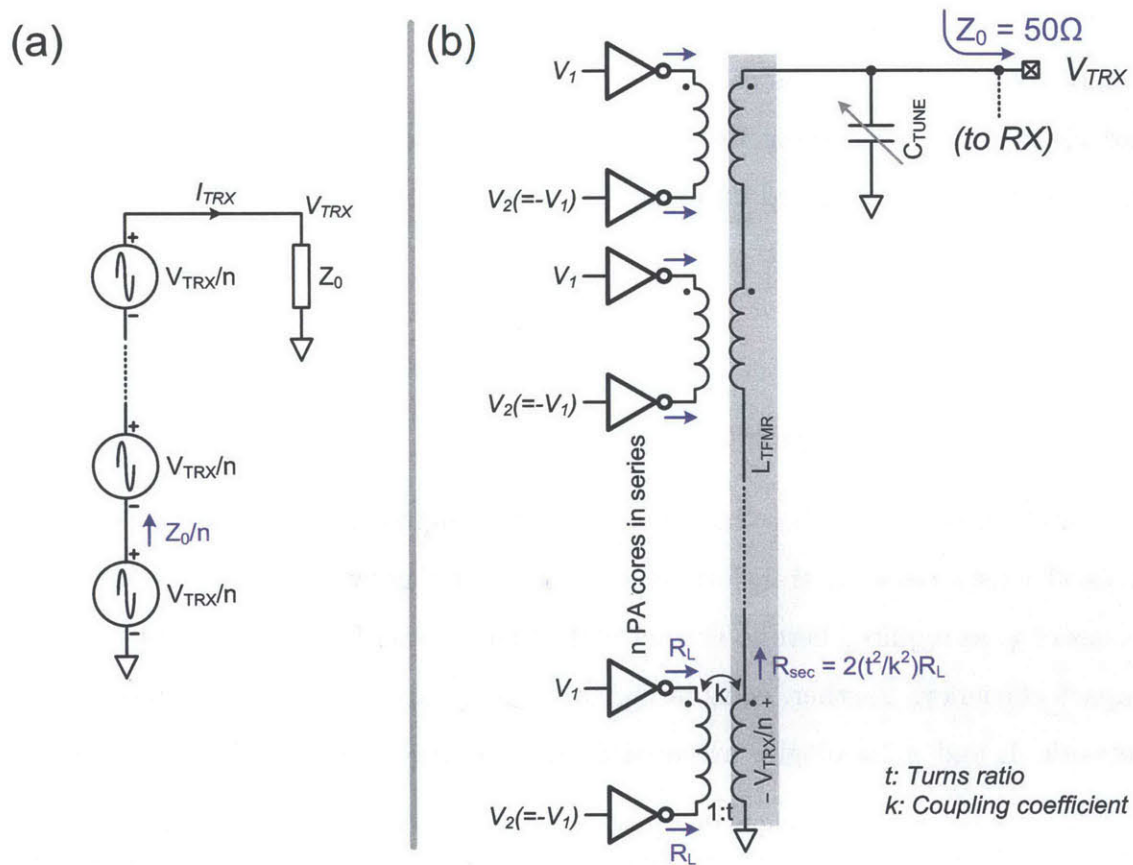
Figure 4-8: PA tunable matching network (a) Simplified electrical model (b) Actual implementation

78

Figure 4-8(a) shows a simplified electrical model of the transformer series-combining topology used in this work. n identical voltage sources are stacked in series to sum up to V_{TRX}. As the circuit is symmetrical, the voltage produced by each source is n times smaller than at the load ($\frac{V_{TRX}}{n}$), while the current is the same as the load (I_{TRX}). Therefore, the impedance seen by each of n voltage sources is simply $\frac{Z_0}{n}$.

Figure 4-8(b) shows the actual topology used. Each of the n voltage sources corresponds to one PA core. The transformer secondary windings are connected in series for voltage summation. The transformers are composed of coupled inductors with coupling coefficient k and turns ratio t. The total magnetizing inductance appearing at V_{TRX} is denoted by L_{TFMR} $(= nL_2)$. L_{TFMR} is resonated out by the capacitor bank C_{TUNE}, which can be reprogrammed to support different bands. The resonant frequency f_{RF} is given by:

$$f_{RF} = \frac{1}{2\pi\sqrt{L_{TFMR} \times C_{TUNE}}} \tag{4.24}$$

Recall that at peak output power, the outphasing angle $\phi = 0$, which implies that the two PA unit cells in each core are exactly out of phase ($V_2 = -V_1$). Consistent with section 4.5.1, the single-ended impedance seen by each of the $2n$ PA unit cells is R_L. Two such impedances in series form the primary side impedance $2R_L$, which is approximately reflected to the secondary side (see appendix B) as:

$$R_{sec} = 2\left(\frac{t^2}{k^2}\right) R_L \tag{4.25}$$

At resonance, figure 4-8(a) and (b) are equivalent, therefore:

$$R_L = \frac{Z_0 k^2}{2nt^2} \tag{4.26}$$

Since there are a total of $2n$ PA unit cells that see R_L, the effective PA load impedance is:

$$R_{opt} = \frac{R_L}{2n} = \frac{Z_0 k^2}{4n^2 t^2} \tag{4.27}$$

79

Finally, the effective impedance transformation ratio is:

$$\chi = \frac{Z_0}{R_{opt}} = \frac{4n^2t^2}{k^2} \tag{4.28}$$

Equating 4.23 and 4.28 and assuming that the transformers are symmetrical with $t = 1$ and $k = 0.8$, it is seen that to meet the targeted output power: $n \geq 3.64$. The closest integer value of $n = 4$ is used in this work.

Full TX output power
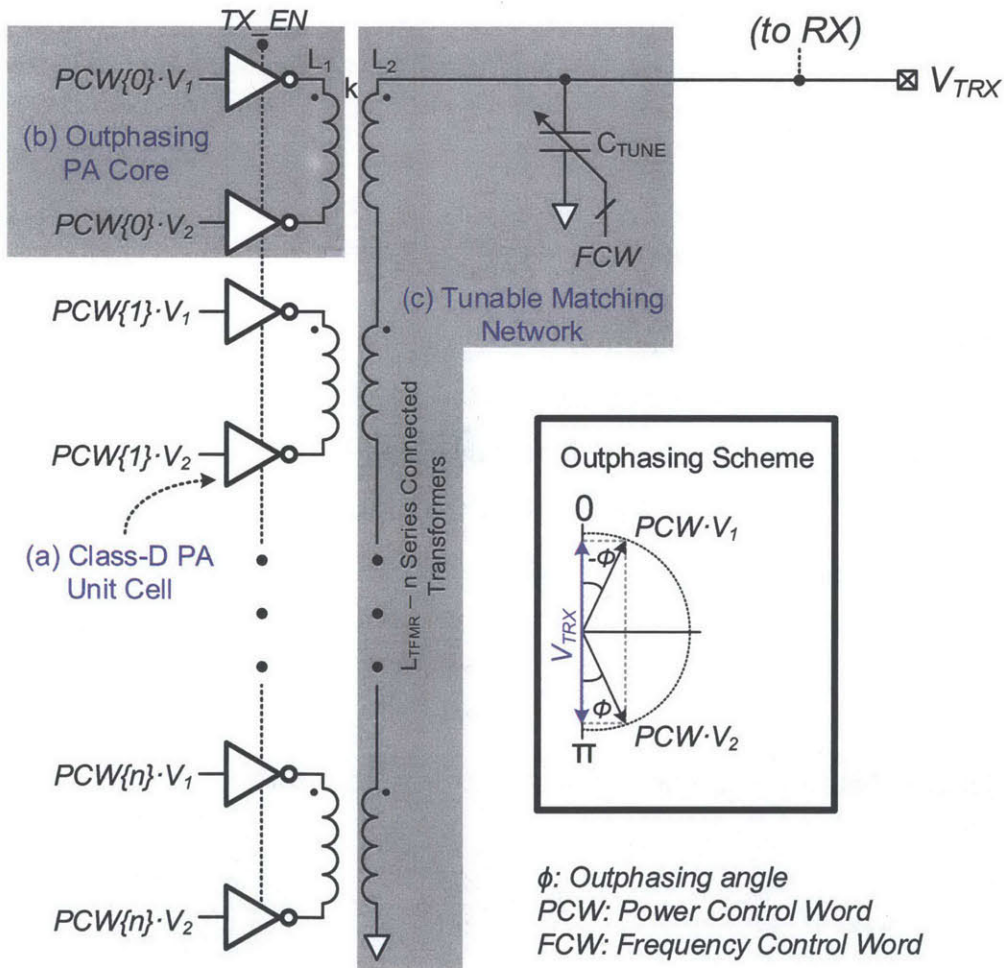
Figure 4-9: TX architecture: (a) Class-D PA unit cell (b) Outphasing PA core (c) Tunable matching network

The full TX subsystem architecture is shown in figure 4-9. At this level, an additional

80

degree of freedom is introduced for static back-off of output power. Each of the n outphasing PA cores is gated with an n bit, thermometer coded power control word (PCW). If the i_{th} bit is 0, the corresponding outphasing PA core produces no RF voltage. Using equation 4.19 and applying superposition, the total RF voltage generated by n cores as a function of PCW is:

$$V_{TRX} = PCW \frac{t}{1+\nu} \frac{8}{\pi} V_{DD} cos(\phi) \tag{4.29}$$

It is worth noting that this stage that the tunable matching network is not ideal. The efficiency of this passive network is denoted here by η_M, and is obviously independent of output power. Since the TX output port V_{TRX} is connected to the antenna with impedance Z_0, the PA output power is therefore:

$$P_{OUT} = \eta_M \frac{V_{TRX}^2}{2Z_0} = \eta_M PCW^2 \left(\frac{t}{1+\nu}\right)^2 \frac{32V_{DD}^2}{\pi^2 Z_0} cos^2(\phi) \tag{4.30}$$

Saturated TX output power and efficiency

To achieve maximum (saturated) output power, PCW $= n$ and $\phi = 0$:

$$P_{SAT} = \eta_M n^2 \left(\frac{t}{1+\nu}\right)^2 \frac{32V_{DD}^2}{\pi^2 Z_0} \tag{4.31}$$

Further, the TX efficiency at saturated output power is given by:

$$\eta_{SAT} = \eta_D \times \eta_M \tag{4.32}$$

4.5.4   Embedded TX switching

Figure 4-10 shows how the TX branch switch is absorbed into the PA itself by repurposing transistors M_2 and M_3 in RX mode.

Recall that in TX mode ($TX\_EN = 1$), the gates of M_2 and M_3 are connected to V_{DD} through M_5 and M_6 (see figure 4-5). Therefore, M_2 and M_3 toggle between linear and cutoff region with their corresponding rail devices M_1 and M_4, preventing them from seeing drain-source voltages in excess of V_{DD} and ensuring device reliability.

81

In RX mode ($TX\_EN = 0$), M_2 and M_3 function as the TX branch switch by turning off through large pull-down and pull-up resistors R_1 and R_2. The off-state PA presents an output capacitance $C_{O,PA}$, which is dominated by device parasitics at V_{OUT}. Two such capacitors in series form a second, undesired resonant circuit with the transformer primary inductance L_1 at frequency f_{res} given as:

$$f_{res} = \frac{1}{2\pi}\sqrt{\frac{2}{L_1 C_{O,PA}}} \tag{4.33}$$
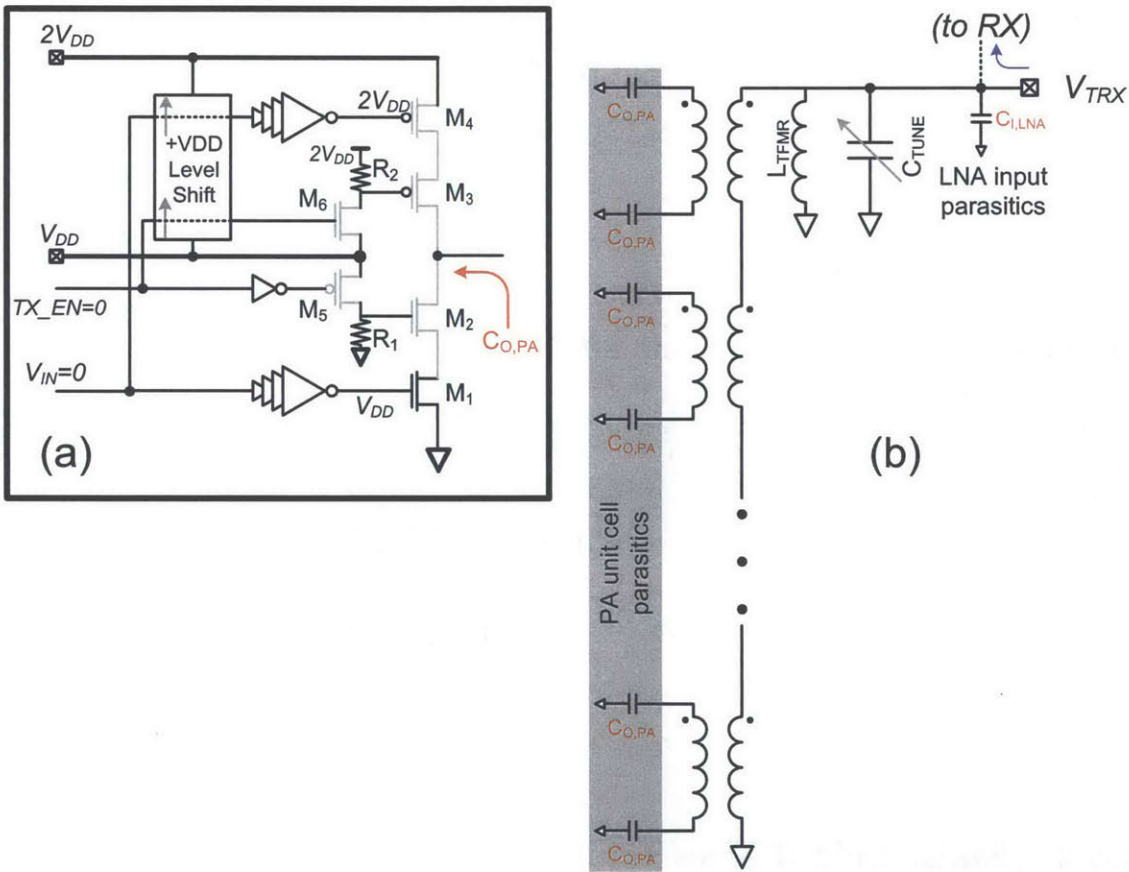
Figure 4-10: Embedded TX switching (a) PA unit cell in RX mode (b) Impact of parasitic capacitance loading on RX signal path

Since TDD standards work on unpaired spectrum i.e. the frequency for TX and RX is the same, the frequency range covered in RX mode should match the TX mode. To make sure that the parasitic loading does not significantly alter the frequency response, this resonant frequency should much higher than the maximum frequency

of interest, namely:

$$f_{res} \gg max(f_{RF}) \tag{4.34}$$

If equation 4.34 is indeed satisfied, the original desired resonance at V_{TRX} (see equation 4.24) is maintained in the RX mode, allowing for the incoming signal to propagate to the RX signal path.

4.6   RX design

Figure 4-11 illustrates the RX section, which consists of the RX switch and the LNA.
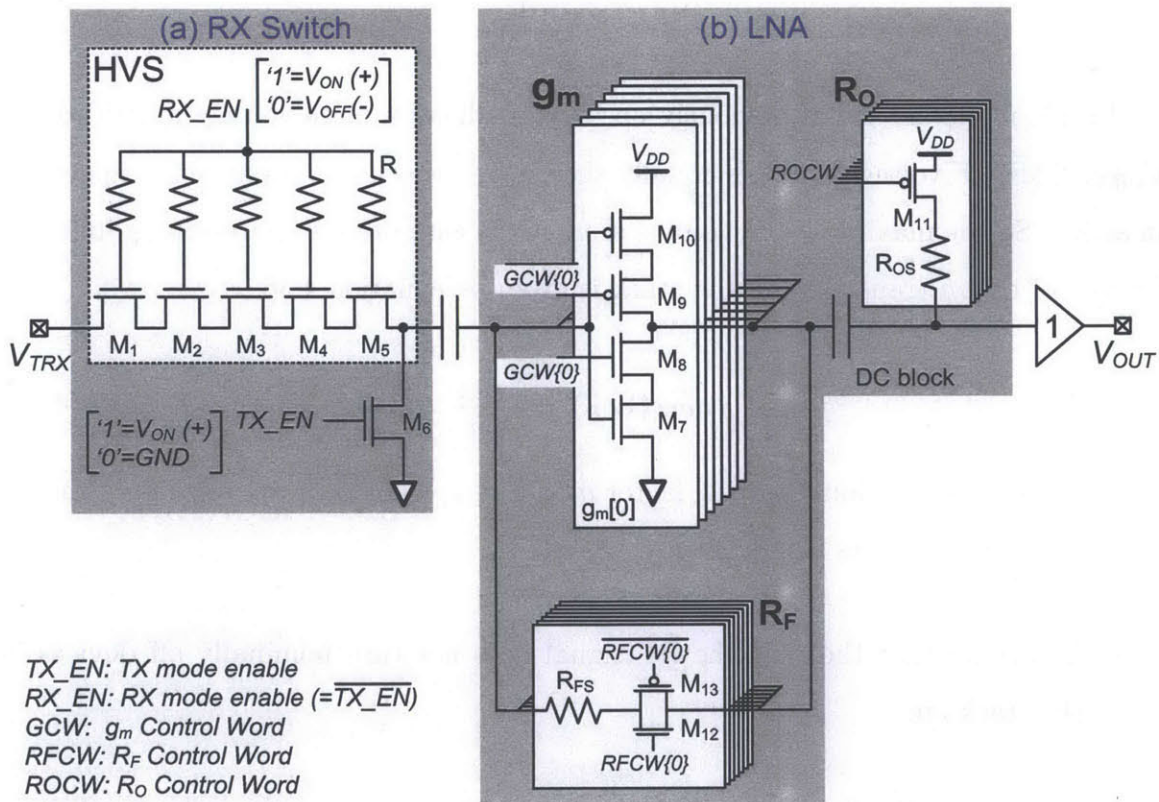
Figure 4-11: RX architecture (a) RX switch (b) Gain and power scalable LNA

4.6.1 RX switch

The RX switch, shown in figure 4-11(a) employs a series-shunt topology to maximize TX/RX isolation [48]. The series path uses a high voltage switch (HVS) with m stacked floating-body SOI NMOS's.

The TX mode ($TX\_EN = \overline{RX\_EN} = 1$) corresponds to series-off shunt-on configuration. The LNA input is grounded, and isolated from the high voltage swing at V_{TRX} by the HVS, which is off. Operation of the HVS can be easily understood by a superposition of DC and RF voltages, and is detailed in figure 4-12 . The RX_EN signal uses a negative DC logic-0 voltage V_{OFF}, applied through gate resistors R. Since no DC current flows through the switch, all the intermediate drain-source nodes have a DC level of 0. The gate resistors are sized large such that they appear as open circuits at RF. This is ensured by satisfying equation 4.35.

$$f_{RF} >> \frac{1}{2\pi R C_P} \tag{4.35}$$

In TX mode the PA is obviously on, and produces a large voltage amplitude at V_{TRX}. This RF voltage stress is equally shared by a stack of $2m$ parasitic capacitors in series. So the maximum RF voltage seen at the gate-drain or gate-source junction for any of the m transistors in the stack is given by equation 4.36.

$$max(V_{CP}) = \frac{V_{TRX}}{2m} \tag{4.36}$$

For reliable operation, stacking factor m and V_{OFF} should be appropriately chosen to satisfy two equations simultaneously:

1. To ensure that the that the RF signal does not turn nominally off devices in the stack on:

$$V_{TH} > V_{OFF} + \frac{V_{TRX}}{2m} \tag{4.37}$$

2. To avoid oxide breakdown:
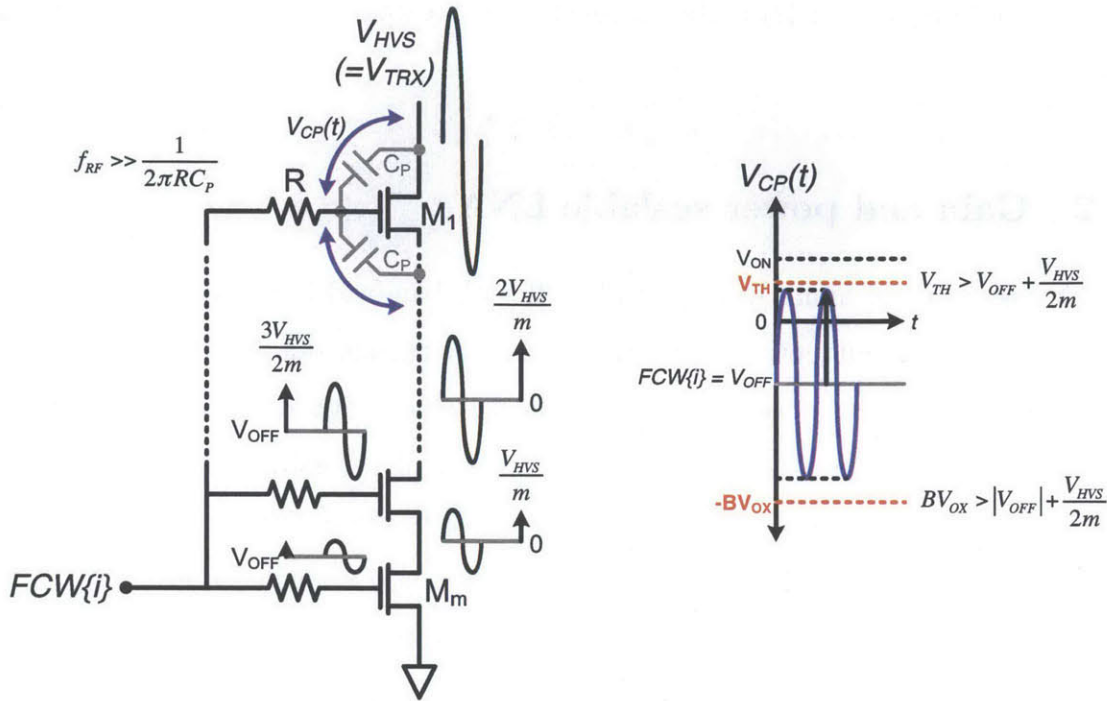
Figure 4-12: HVS off state operation details

$$BV_{OX} > | V_{OFF} | + \frac{V_{TRX}}{2m} \tag{4.38}$$

As before, a maximum P_{SAT} of 30 dBm (1 W) is assumed. It follows that the worst case voltage amplitude at V_{TRX} is 10 V. For the technology and device type used, the design choice of $m = 5$, $V_{OFF} = $ -1.4 V satisfies both equations with adequate margin.

The RX mode ($TX\_EN = \overline{RX\_EN} = 0$) corresponds to series-on shunt-off configuration. $V_{ON} = V_{DD} = 1$ V for low on-resistance. The RX signal passes through m series-connected on devices (in triode region) that form the HVS to the LNA. The insertion loss of the switch is given by:

$$P_{L,RXSW}(dB) = 20log_{10}\left(\frac{Z_0}{mR_{ON} + Z_0}\right) \tag{4.39}$$

85

Z_0 (= 50 Ω) is the reference impedance to which both the LNA and antenna are matched, while R_{ON} (< 1 Ω) is the on resistance of a single NMOS transistor in the stack.

4.6.2 Gain and power scalable LNA

In LNA design, noise figure usually trades off with bandwidth and power consumption. Over the years, inductive degeneration [44] has remained a popular topology as it is capable of excellent noise figure and input matching with low power dissipation. However, this technique is inherently narrowband, and necessitates multiple LNA's as shown in figure 4-1 for multi-band coverage. Due to the desire for multi-standard compatibility, broadband designs in CMOS have recently attracted a lot of research attention. Common gate [58] and shunt feedback [59] topologies are effective ways of achieving input matching, but compromise somewhat on noise figure and power dissipation. More recently, broadband, noise canceling LNAs have become popular [60][5]. With broadband designs, resilience to out-of-band interferers is especially important to make sure that the RX frontend does not saturate due to blockers. Some innovative techniques like impedance mixing [29] have been recently reported to address these issues.

In this work, a shunt feedback LNA is employed as it offers acceptable noise figure while being inherently broadband. If even lower noise figure and better out-of-band rejection is required, the proposed RF frontend architecture does not preclude employing more sophisticated architectural techniques like noise canceling and impedance mixing when implementing the full RX chain.

The LNA implementation shown in figure 4-11(b) also follows a maximally-digital design approach. A CMOS inverter based transconductance (g_m) cell is employed with resistive output load ($R_L = R_O \parallel r_{on} \parallel r_{op}$) and shunt feedback ($R_F$) for low voltage, broadband operation. The choice of an inverter based g_m cell has two key

---

[5]Noise canceling is intuitively understood as a hybrid of an unmatched, low noise figure amplifier and a matched noisy amplifier whose noise is sensed and cancelled by the former to relax the tradeoff between noise figure and bandwidth.

benefits:

1. The complimentary topology reuses current, boosting the achievable transconductance per unit current (g_m/I_{ds}) to reduce power.

2. N/PMOS g_m superposition [61] and a nearly rail-rail output voltage swing offer improved linearity at low supply voltage (V_{DD}).

The simulated response of the g_m cell, which illustrates these two benefits is shown is figure 4-13.
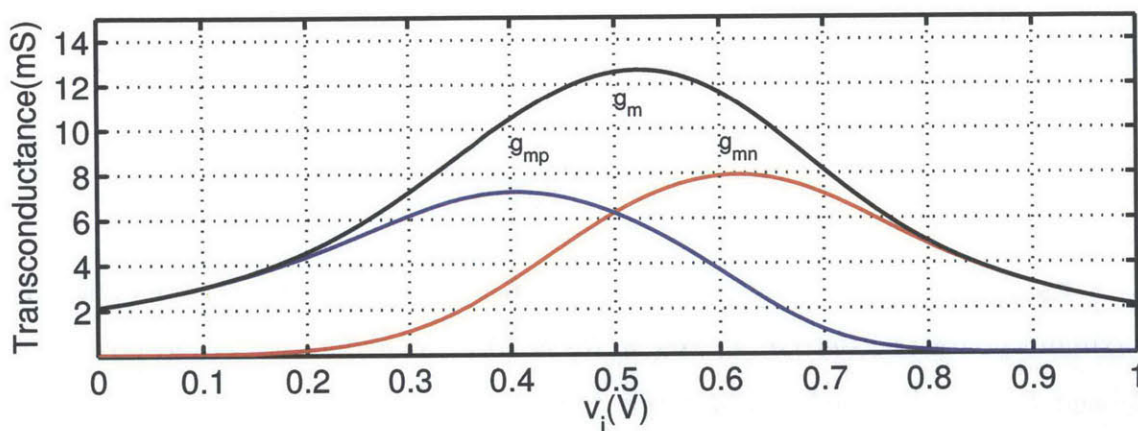
Figure 4-13: Complimentary g_m cell response showing superposition of g_{mn} and g_{mp}

To calibrate for PVT variation and implement moderate gain control, the LNA design incorporates digital tuning. The g_m cell has 7 slices that can be selectively turned off with a 3b g_m control word GCW. Within each g_m cell, M_7 and M_{10} are the main signal transistors, while M_8 and M_9 are simply used to tri-state (turn-off) the cell through the relevant GCW bit when not in use. R_F and R_O are also tunable through 3b words RFCW and ROCW respectively. The tunable LNA design parameters are given by:

$$g_m = GCW \times (g_{mn} + g_{mp}) = GCW \times (g_{m7} + g_{m10}) \qquad (4.40)$$

$$R_F = \frac{R_{FS}}{RFCW} \qquad (4.41)$$

87

$$R_O = \frac{R_{OS}}{R_{OCW}} \tag{4.42}$$

The small-signal performance of the LNA can be analyzed using the linearized model presented in figure 4-14. Noise sources are also shown in grey. The presented analysis is only valid for frequencies significantly lower than the 3 dB BW of the LNA. From node analysis, the voltage gain of the LNA (A_V) is given as:

$$A_V = \frac{v_o}{v_i} = \left(-g_m + \frac{1}{R_F}\right)(R_L \parallel R_F) \approx -g_m(R_L \parallel R_F) \tag{4.43}$$

While the LNA input impedance $(Z_{I,LNA})$ is given as:

$$Z_{I,LNA} = \frac{R_F + R_L}{1 + g_m R_L} \tag{4.44}$$

For maximum A_V, both g_m and R_F are set to maximum while R_O is open ($R_L = r_{on} \parallel r_{op}$). The design parameters are chosen such that the input is matched for maximum power transfer i.e. $Z_{I,LNA} = R_S = 50 \ \Omega$. To reduce gain (and power) g_m, R_F and R_O are all scaled appropriately to maintain input match at lower gain.

The LNA output is not inherently matched to 50 Ω. Since LNA's typically drive the capacitive input of a mixer, 50 Ω output matching is indeed not required in a radio SoC implementation. However, if the LNA is to be tested individually, as is the case here, a matched RF output is indeed preferred. Therefore, a highly linear source follower based voltage buffer is cascaded at the LNA output to provide this match. Assuming the output buffer gain is unity, the measured voltage gain at V_{OUT} will be 6 dB lower than the LNA gain, and should be de-embedded post measurement.

4.6.3 Noise analysis

The noise figure of the switch is simply equivalent to its loss:

$$NF_{RXSW}(dB) = P_{L,RXSW} \tag{4.45}$$

The noise factor of the LNA can be calculated by doing a straightforward noise
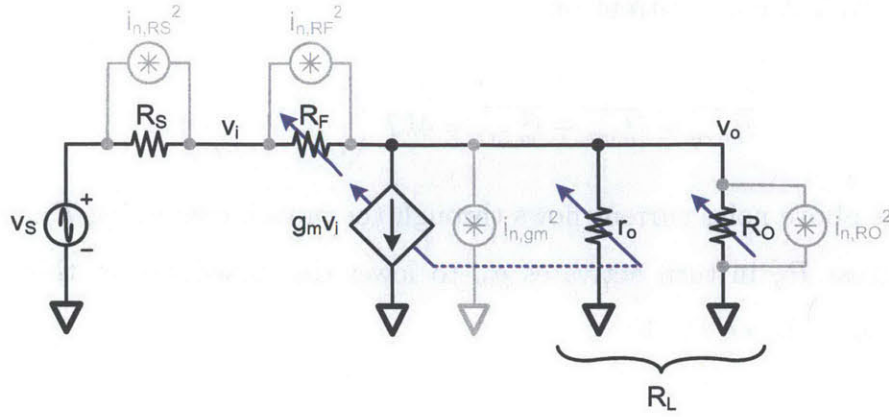
Figure 4-14: LNA small-signal model including noise sources

analysis by referring again to figure 4-14. Only thermal noise from physical resistors and channel noise from NMOS/PMOS devices is considered. The latter can be modeled as a single composite noise source:

$$\overline{i_{n,gm}^2} = 4kT \left(\frac{\gamma}{\alpha}\right) g_m \Delta f \tag{4.46}$$

$$\frac{\gamma}{\alpha} \geq \frac{2}{3} \tag{4.47}$$

$\frac{\gamma}{\alpha}$ is a process dependent constant. its value is bounded on the lower end to $\frac{2}{3}$ for long-channel devices, but can be much higher for short-channel devices [44].

The noise from R_S and R_F simply shows up in parallel if current noise representation is used. Since the input is matched i.e. $Z_{I,LNA} = R_S$, the current divides equally between the two branches. The input noise currents can be converted to an output noise voltages as follows:

$$\overline{v_{no,RS}^2} = \frac{A_V^2}{4} \overline{i_{n,RS}^2} R_S^2 = A_V^2 kT R_S \Delta f \tag{4.48}$$

$$\overline{v_{no,RF}^2} = \frac{A_V^2}{4} \overline{i_{n,RF}^2} R_S^2 = A_V^2 kT \frac{R_S^2}{R_F} \Delta f \tag{4.49}$$

The noise from the g_m rendering active device and the output resistor R_O results in

89

a total equivalent noise current of:

$$\overline{i^2_{no,eq}} = \overline{i^2_{n,gm}} + \overline{i^2_{n,RO}} = 4kT \left(\frac{\gamma}{\alpha} g_m + \frac{1}{R_O} \right) \Delta f \qquad (4.50)$$

Part of the above noise current flows through the branch containing R_F and R_S, the voltage across R_S in turn activates g_m to lower the impedance at this node. The equivalent impedance R_{eq} is:

$$R_{eq} = R_L \parallel R_F + R_S \parallel \left(1 + \frac{R_F}{R_S} \right) \frac{1}{g_m} \approx R_L \parallel R_F \parallel \left(1 + \frac{R_F}{R_S} \right) \frac{1}{g_m} \qquad (4.51)$$

The approximation made above is $R_F \gg R_S$, which is reasonable because the open loop gain $g_m R_L$ is high. $i^2_{no,eq}$ can now be converted to voltage noise:

$$\overline{v^2_{no,eq}} = \overline{i^2_{no,eq}} R^2_{eq} \qquad (4.52)$$

The noise factor, using output referred noise voltages is simply given as:

$$F = \frac{\overline{v^2_{no,RS}} + \overline{v^2_{no,RF}} + \overline{v^2_{no,eq}}}{\overline{v^2_{no,RS}}} \qquad (4.53)$$

Substituting values from equations 4.48, 4.49 and 4.52 into 4.53:

$$F = 1 + \frac{R_S}{R_F} + \frac{4}{g_m R_S} \left(\frac{\gamma}{\alpha} + \frac{1}{g_m R_O} \right) \left(\frac{\frac{1}{R_L} + \frac{1}{R_F}}{\frac{1}{R_L} + \frac{1}{R_F} + \left(1 + \frac{R_F}{R_S} \right) \frac{1}{g_m}} \right)^2 \qquad (4.54)$$

It is clear that in order to minimize noise figure g_m should be maximized. Low noise figure clearly comes at the expense of power consumption, but the g_m superposition principle previously discussed helps in relaxing the power and noise figure tradeoff to some extent. R_F and R_O should also be maximized for low noise, but the values are constrained by gain and matching requirements, as previously outlined in equations 4.43 and 4.44.

The overall RX path noise figure, with all quantities expressed in decibels, is then:

$$NF = NF_{RXSW} + 10log_{10}(F) \qquad (4.55)$$

4.7 Design of RF passives

The passive components in the RF signal path must be designed with great care for three primary reasons:

1. The shunt resonator formed by L_{TFMR} and C_{TUNE} determines the frequency response of the system. Therefore, any variation in component values from the nominal will shift the frequency range. Limitations associated with the practical implementation of these passives also limits the tuning range of the frontend.

2. Losses in RF passive components, which are quantified by their finite quality factor (Q), limit the efficiency of the PA in TX mode, and the overall noise figure of the signal chain in the RX mode.

3. Large RF voltage and current swings encountered in the TX mode must be tolerated without any reliability issues.

4.7.1 Choice of component values

Recall equation 4.24, where it is clear that the resonant frequency of the tank formed by total transformer inductance L_{TFMR} and tunable capacitance C_{TUNE} determines the frequency response of the system:

$$\omega = 2\pi f_{RF} = \frac{1}{\sqrt{L_{TFMR} \times C_{TUNE}}} \qquad (4.56)$$

As will become evident in section 4.7.3 , only a finite C_{TUNE} tuning ratio is practically achievable, and limits what range of frequencies can be covered. Nevertheless, it is clear that the problem is under constrained i.e. an infinite number of combinations

of numerical values for L_{TFMR} and C_{TUNE} can satisfy equation 4.56 at any given f_{RF}.

Losses in passive components degrade performance both in the TX and RX mode. Priority is given to the TX mode, since it is the high-power mode of operation. Therefore, optimal values of passive components are determined with the aim of maximizing efficiency η_M (see section 4.5.3) of the PA matching network.

The transformer based power combiner was introduced in section 4.5.3. Recall that R_{sec} is the differential impedance seen by each of n series-connected secondary windings terminated in Z_0. Revisiting figure 4-8:

$$R_{sec} = \frac{Z_0}{n} \tag{4.57}$$

For a standard antenna impedance of $Z_0 = 50 \ \Omega$ and $n = 4$, which is dictated by the output power requirement (derived in section 4.5.3), we have $R_{sec} = 12.5 \ \Omega$. The efficiency of transformer based power combiners have already been studied extensively in prior work [11] [62]. The analysis is not reiterated here but the key result is restated and applied to the current design. The optimal primary side inductance ($L_{1,opt}$) which maximizes the efficiency when resonated with a shunt capacitor (see equation 4.24) is given by:

$$\omega L_{1,opt} = \frac{R_{sec}}{t^2} \frac{A}{1 + A^2} \tag{4.58}$$

Where t is simply the turns ratio:

$$t = \sqrt{\frac{L_2}{L_1}} \tag{4.59}$$

Further, A is given by:

$$A = \frac{1}{\sqrt{\frac{1}{Q_2^2} + \frac{Q_1}{Q_2}k^2}} \tag{4.60}$$

For this optimal inductance value, the maximum combiner efficiency is:

92

$$\eta_{M,opt} = \cfrac{1}{1 + \frac{2}{Q_1 Q_2 k^2} + 2\sqrt{\frac{1}{Q_1 Q_2 k^2}\left(1 + \frac{1}{Q_1 Q_2 k^2}\right)}} \qquad (4.61)$$

Where L_{TFMR} is related to $L_{1,opt}$ by:

$$L_{TFMR} = nL_{2,opt} = nt^2 L_{1,opt} \qquad (4.62)$$

As mentioned previously in section 4.4, the design is centered at $f_C = 2.5$ GHz. Equation 4.58 implies that for frequencies above f_C, the fixed chosen inductance will be higher than optimal, while for frequencies below f_C, it will be lower than optimal. Clearly, implementing this frequency-agile architecture requires some compromise in efficiency. From prior work [57] [62], reasonable approximations can be made about transformer parameters in a sub-100nm CMOS process even before any design is done. Since a symmetrical ($t = 1$) design is targeted, $Q_1 = Q_2 = 10$ at $f_C = 2.5$ GHz is assumed, with a $k = 0.8$. Substituting these values into equations 4.24, 4.58 and 4.60 gives $L_{1,opt} = 390$ pH. Further, equation 4.24 with $f_C = 2.5$ GHz gives $C_{TUNE}(f_C) = 2.6$ pF. These values are used as starting points in transformer and capacitor bank design, which is detailed in the next two sections.

4.7.2 Transformer based power combiner

Transformers are designed with the aid of electromagnetic (EM) simulation. Figure 4-15 shows the top and cross-sectional view of the transformer combiner. The combiner is composed of four individual transformers, which have their secondary windings connected in series. An octagonal geometry is employed to minimize undesired cross-coupling from one transformer to another [62]. Since the digital stackup in this process offers no ultra-thick metal, the top three metal layers are stitched in parallel with vias to form a low resistance, composite RF conductor. Primary (PRI) and secondary (SEC) inductors are composed of two windings each, which are interleaved (PRI-1/SEC-1/PRI-2/SEC-2) in a manner similar to [63]. There are three main benefits of this arrangement:
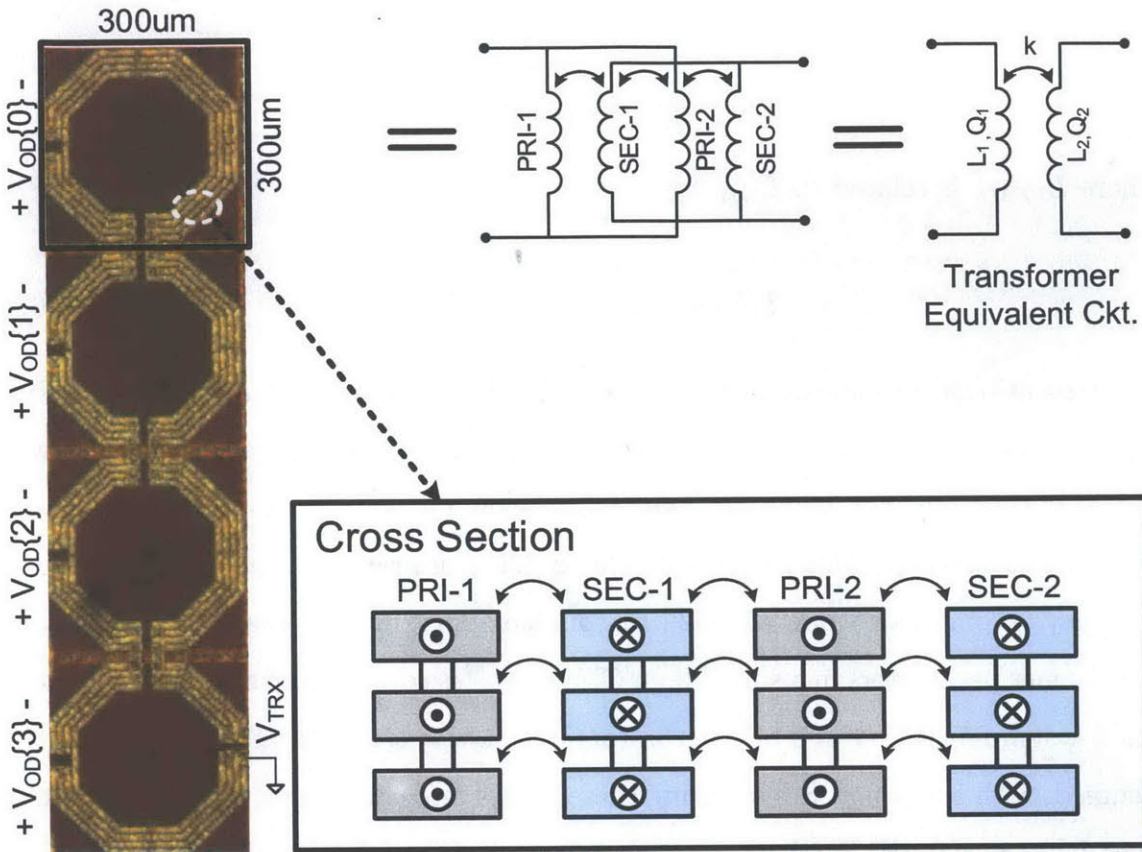
Figure 4-15: Transformer based power combiner details

1. This transformer arrangement enhances lateral coupling between the two windings, because the two central conductors (SEC-1/PRI-2) achieve flux linkage on both sides.

2. Wide metal traces are preferred for low DC resistance and reliable operation with high currents. However, thick top conductors have maximum width DRC rules. With multiple interleaved conductors both requirements can be satisfied simultaneously.

3. Skin effect becomes important at higher frequencies, with the current becoming more concentrated at the edges of wide conductors [64]. The current density at higher frequencies can be modeled as having a exponentially decaying characteristic:

94

$$J = J_S e^{-\frac{d}{\rho}} \tag{4.63}$$

The J_S is the sheet current density at the surface and d is the distance from it. The skin depth ρ is given by:

$$\rho = \sqrt{\frac{2\sigma}{\omega\mu}} \tag{4.64}$$

The skin depth at 2 GHz for Cu conductors is about 2 μm. Due to skin effect, 95 % of the current will be concentrated within 3ρ of the edges, implying that conductors thicker than $6\rho = 9$ μm will not significantly improve the inductor Q. As seen from equation 4.64 this phenomenon becomes more significant at higher frequencies due to reduction in skin depth as $f^{-\frac{1}{2}}$. Having two thinner conductors uses the physical metal width 2× more effectively, reducing the effective winding resistance (R_1, R_2) and enhancing the quality factor (Q_1, Q_2) for higher frequencies.

A method-of-moments 2.5D EM simulator is used to simulate the transformer structure. A few iterations are performed through geometry tweaks to approach the optimal value of inductance i.e. $L_{1,opt} = 390$ pH, while maximizing the winding quality factors (Q_1, Q_2) and mutual coupling coefficient (k) for high efficiency (see equation 4.61). The s-parameters obtained from the EM simulation can be converted to z-parameters [65], from which the relevant transformer parameters can be extracted. These parameters correspond to the transformer equivalent circuit model [66] shown in figure 4-15, which is only valid at frequencies significantly lower than transformer self-resonance. The mathematical framework for this extraction procedure is presented in appendix B. Results from the extraction for the optimized structure are plotted in figure 4-16. The impact of non-ideal on-chip passives on efficiency across the multi-octave frequency tuning range is quantified later in section 4.7.4.

95
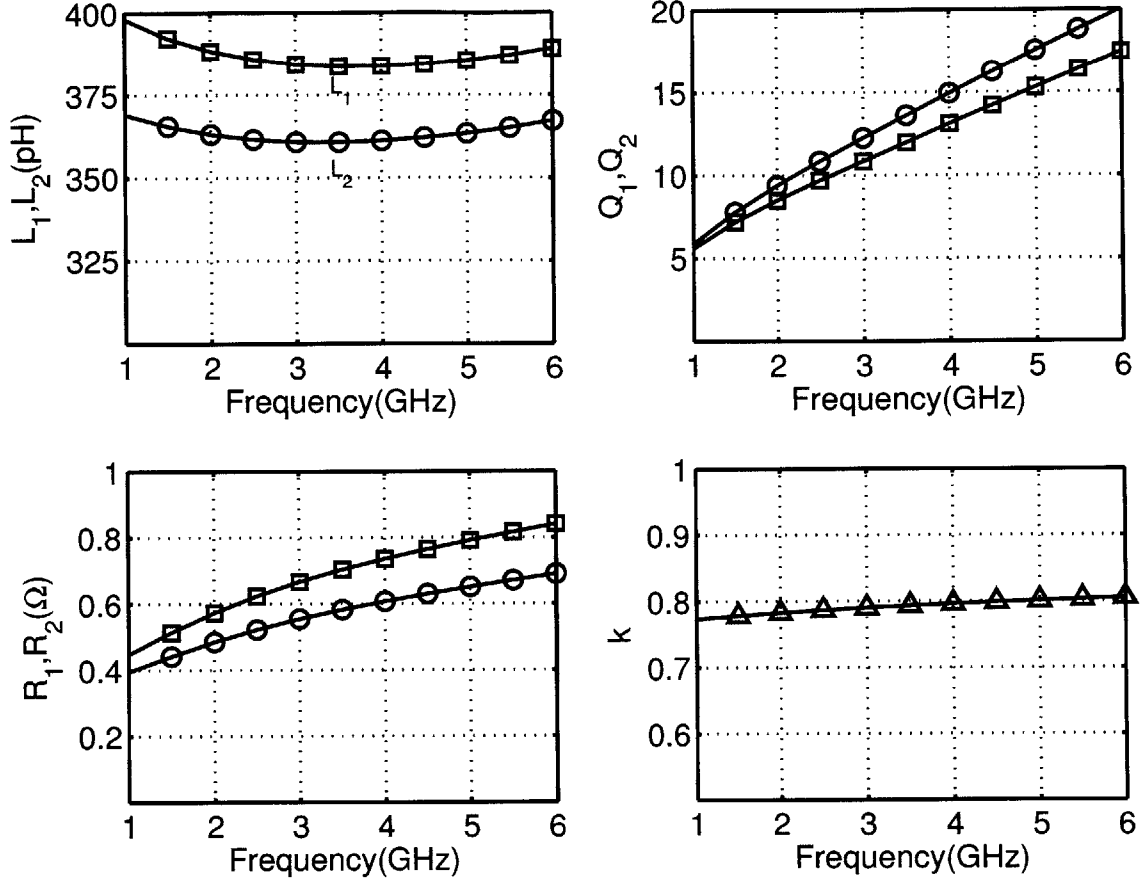
Figure 4-16: Transformer parameters extracted from EM simulation

4.7.3 Digitally tunable capacitor bank

Figure 4-17 shows the implementation of the capacitor bank. All the capacitor unit slices are unary weighted. To keep the capacitor layout symmetrical about the V_{TRX} trace, non-binary bit weighing of 1,2,4,7 is used. The capacitance C_U (= 280 fF) is implemented with metal-on-metal (MoM) capacitors, which rely on the capacitive coupling in closely spaced inter-digitated metal fingers constructed with multiple interconnect layers. A square geometry maximizes the achievable Q [67]. In this design, 6 (metal-3 to metal-8) stacked metals are used. The bottom two metal layers are omitted to minimize bottom-plate parasitics.

For compatibility with high voltages at V_{TRX} in the TX mode, a high voltage switch (HVS), employing stacking with m (= 4) SOI floating body NMOS's is used. The design of this switch is similar to the one presented in section 4.6.1. In the on

96
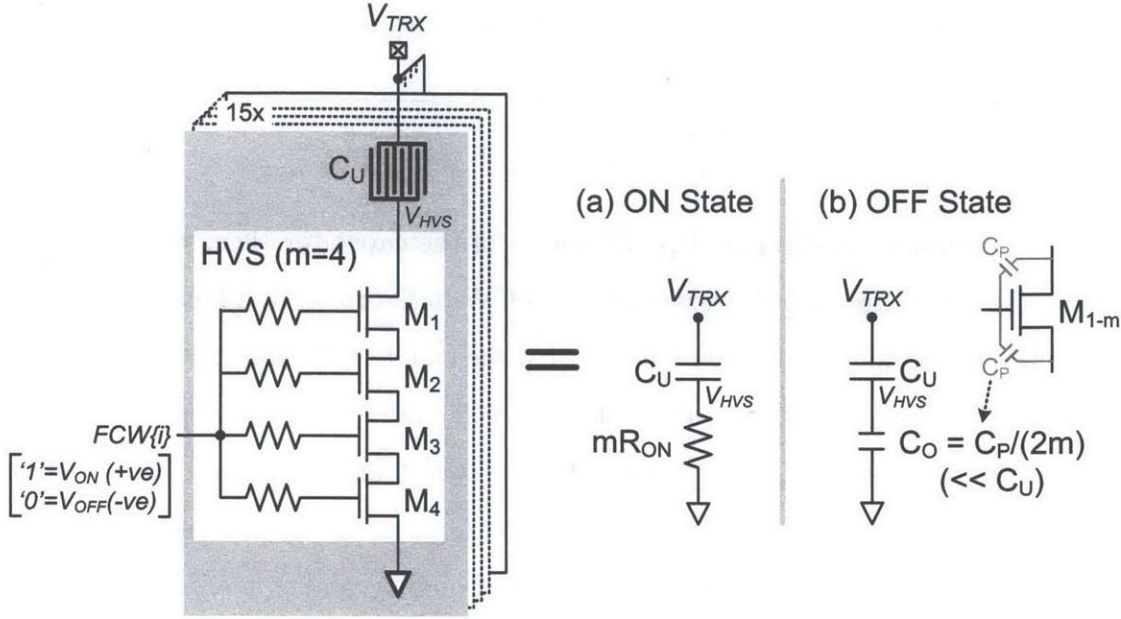
Figure 4-17: C_{TUNE} implementation details

state, the Q of the capacitor is limited by the on resistance of the HVS ($= m \times R_{ON}$). The on state capacitor parameters are:

$$C_{ON} = C_U \tag{4.65}$$

$$Q_{ON}(\omega) = \frac{1}{\omega C_U \times m R_{ON}} \tag{4.66}$$

In the off state, the bottom plate parasitic (ψC_U), in addition to the off-state parasitic capacitance of the HVS ($C_O = \frac{C_P}{2m}$) show up in series with C_U and result in a much smaller, but finite capacitance value C_{OFF} given as:

$$C_{OFF} = \frac{C_U(C_O + \psi C_U)}{C_U + C_O + \psi C_U} \approx C_O + \psi C_U = \frac{C_P}{2m} + \psi C_U \tag{4.67}$$

The value of the total capacitance C_{TUNE} at any given value of FCW is:

$$C_{TUNE} = FCW \times C_{ON} + \overline{FCW} \times C_{OFF} \tag{4.68}$$

Referring to equations 4.24 and 4.68, it becomes clear that the implementation of

C_{TUNE} determines the frequency ratio r that can be achieved by the RF frontend:

$$r = \frac{max(f_{RF})}{min(f_{RF})} = \sqrt{\frac{C_{ON}}{C_{OFF}}} = \sqrt{\frac{C_U}{C_O + \psi C_U}} \qquad (4.69)$$

Table 4.1 shows the simulated performance of the capacitor slice, while the realized C_{TUNE} values are plotted versus the 4b FCW in figure 4-18. The simulated $r = 3.24$. It should be noted that the simulated value of r is a theoretical upper bound. In practice, r is limited by several other factors like additional routing parasitic capacitance. In this design, Q_{ON} is limited by the switch on resistance. The capacitor switch is sized with the goal of achieving a target Q_{ON} at the design center frequency f_C (= 2.5 GHz) of about 25. Hence, the product of Q_{ON} and ω is a design constant (see equation 4.66), denoted by β.

Parameter	Value
C_{ON} (fF)	280.0
C_{OFF} (fF)	26.7
mR_{ON} (fF)	9.5
$\beta = \omega Q_{ON}$ (rad/s)	3.77×10^{11}

Table 4.1: Capacitor slice simulation results
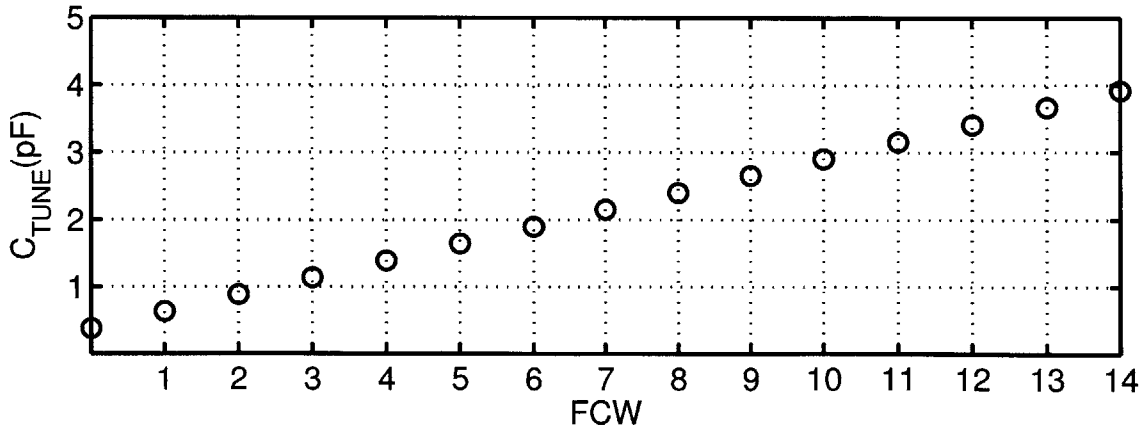
Figure 4-18: Designed value of C_{TUNE} vs. FCW

4.7.4 Tunable matching network efficiency

The overall matching network is simulated with models extracted for both the transformer combiner and the digitally tunable capacitor bank. The frequency response is plotted by tuning the FCW from minimum to maximum in an AC sweep. Figure 4-19 shows the results from four individual transformers connected in a series combination, essentially conforming to the model parameters presented in figure 4-16. The maximum achievable efficiency at each FCW setting and corresponding resonant frequency (f_{eff}) are also highlighted. Recall that the transformer inductance is designed to be optimal around the center frequency of $f_C = 2.5$ GHz. Equation 4.61 predicts 78 % theoretical maximum efficiency, which agrees well with the simulated value of 75 %. The slight difference is easily explained by finite capacitor Q included in the simulation, which is neglected in the equation to keep the derivation simple.
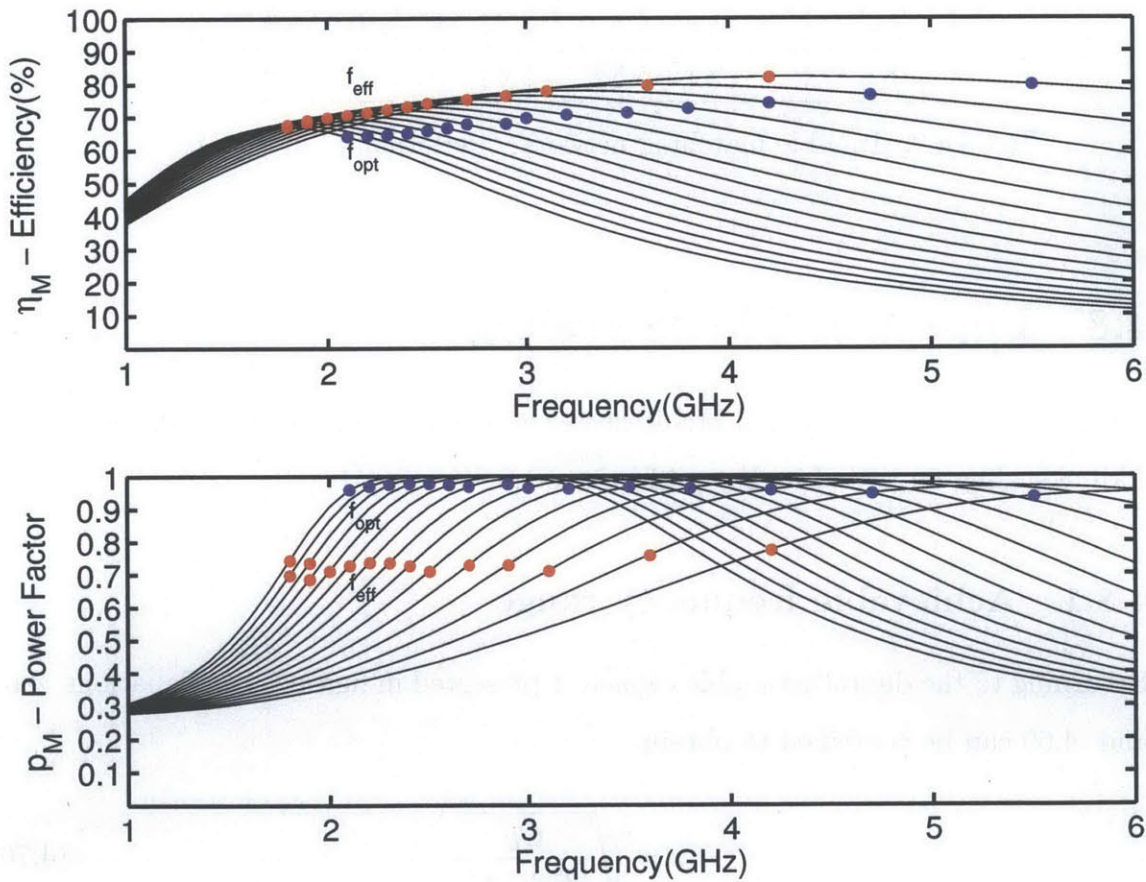
Figure 4-19: Tunable matching network - Efficiency and power factor vs. FCW

It is important to note that for non-ideal transformer coupling i.e. $k < 1$, maximum efficiency and power factor do not simultaneously occur at the same frequency. As discussed in section 4.5.1, Class-D unit cells are somewhat immune to reactive loads (i.e. poor power factor), but advantage of this should be taken mostly when the PA is operating in outphasing mode, which will inevitably cause power factor degradation because the transformer is used a non-isolating combiner. This phenomenon is well analyzed in [53]. For best overall linearity, the power factor should still be close to unity at peak power. Therefore, for a given value of tunable capacitance C_{TUNE}, a better choice than f_{eff} is to operate the PA at the frequency f_{opt} where the product of efficiency and power factor is maximized. The range of achievable efficiency and power factor with f_{eff} and f_{opt} are compared in table 4.2. Clearly, the improvement in power factor trades off with efficiency.

Setting	f_{low} - f_{high} (GHz)	η_M (%)	Power factor
f_{eff}	1.8 - 4.2	67 - 82	0.69 - 0.77
f_{opt}	2.1 - 5.5	64 - 80	0.94 - 0.98

Table 4.2: Tunable matching network - Performance at f_{eff} vs. f_{opt}

4.8 Impact of CMOS scaling

Having outlined the proposed architecture, it is worthwhile to study the impact of CMOS scaling on two of its key performance metrics:

4.8.1 Achievable frequency range

Returning to the digitally tunable capacitor presented in figure 4-17. Equations 4.67 and 4.69 can be combined to obtain:

$$r = \sqrt{\frac{C_U}{\frac{C_P}{2m} + \psi C_U}} \tag{4.70}$$

Recall from section 4.7.3 that the product of Q_{ON} and ω is a constant given by β:

100

$$\beta = \frac{1}{mR_{ON}C_U} \qquad (4.71)$$

Further, a common switch figure-of-merit is the time constant formed by the on resistance and off state parasitic capacitance between the two switch terminals:

$$\tau = \frac{R_{ON}C_P}{2} \qquad (4.72)$$

Combining equations 4.70, 4.71 and 4.72 gives the achievable frequency ratio r as:

$$r = \sqrt{\frac{1}{\tau\beta + \psi}} \qquad (4.73)$$

ψ is a metric of capacitor performance and dependent on the process metal stack and type of capacitor. As it is not really related to gate length to a first order, it is also assumed to be constant. r is plotted for different SOI device technologies in figure 4-20.
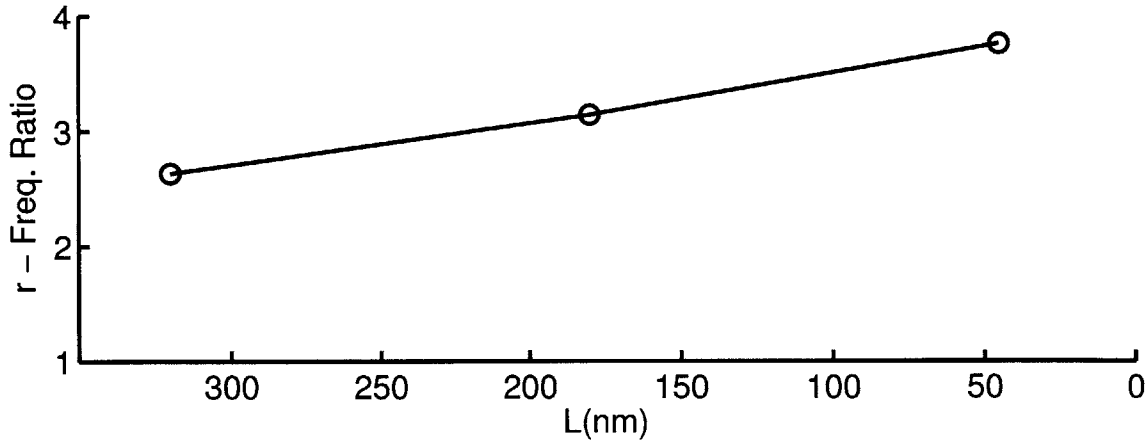
Figure 4-20: Impact of CMOS scaling on achievable frequency ratio ($\beta = 3.77 \times 10^{11}$)

4.8.2 PA instrinsic efficiency

Returning to the simplified electrical model of the Class-D PA unit cell presented in figure 4-6, from equations 4.4 and 4.13 we have:

101

$$\eta_D = \frac{1}{1 + \nu + \frac{P_{L,cap}}{P_O}} \tag{4.74}$$

Substituting calculated expressions for P_O, C_{TOT} and $P_{L,cap}$ from equations 4.6, 4.11 and 4.12 respectively, we obtain:

$$\eta_D = \frac{1}{1 + \nu + \frac{\pi^2(1+\nu)^2}{4\nu} \left(\frac{2-\epsilon}{1-\epsilon}\right)(1+\kappa)R_{ON}(C_{GS} + 4C_{GD})f_{RF}} \tag{4.75}$$

There is only one strongly process dependent term in equation 4.75; the time constant given by $R_{ON}(C_{GS}+4C_{GD})$. This time constant can be re-written with more familiar quantities. In addition to τ given in equation 4.72 with $C_{GD} \approx C_P$, another common transistor parameter is the transition frequency f_t, given by the expression:

$$f_t = \frac{g_m}{2\pi(C_{GS} + C_{GD})} \tag{4.76}$$

Further, if the transistor behavior is approximated by square-law physics and early effect ignored, then:

$$g_m = \frac{1}{R_{ON}} \tag{4.77}$$

Substituting values from equations 4.72, 4.76 and 4.77 into 4.75, the final expression for PA intrinsic efficiency η_D is:

$$\eta_D = \frac{1}{1 + \nu + \frac{\pi^2(1+\nu)^2}{4\nu} \left(\frac{2-\epsilon}{1-\epsilon}\right)(1+\kappa)\left(\frac{1}{2\pi f_t} + \frac{3}{2}\tau\right)f_{RF}} \tag{4.78}$$

Model predicted efficiency is plotted for several different device technologies vs. frequency in figure 4-21. It is also noteworthy that for the same gate length, SOI technology will perform significantly better than its bulk counterpart due to reduced parasitic capacitance.

As CMOS technology scales, gate length L decreases, f_t increases, while τ decreases. Therefore, equations 4.69 and 4.78 prove that achievable frequency range and PA intrinsic efficiency benefit from CMOS scaling. Figures 4-20 and 4-21 quan-
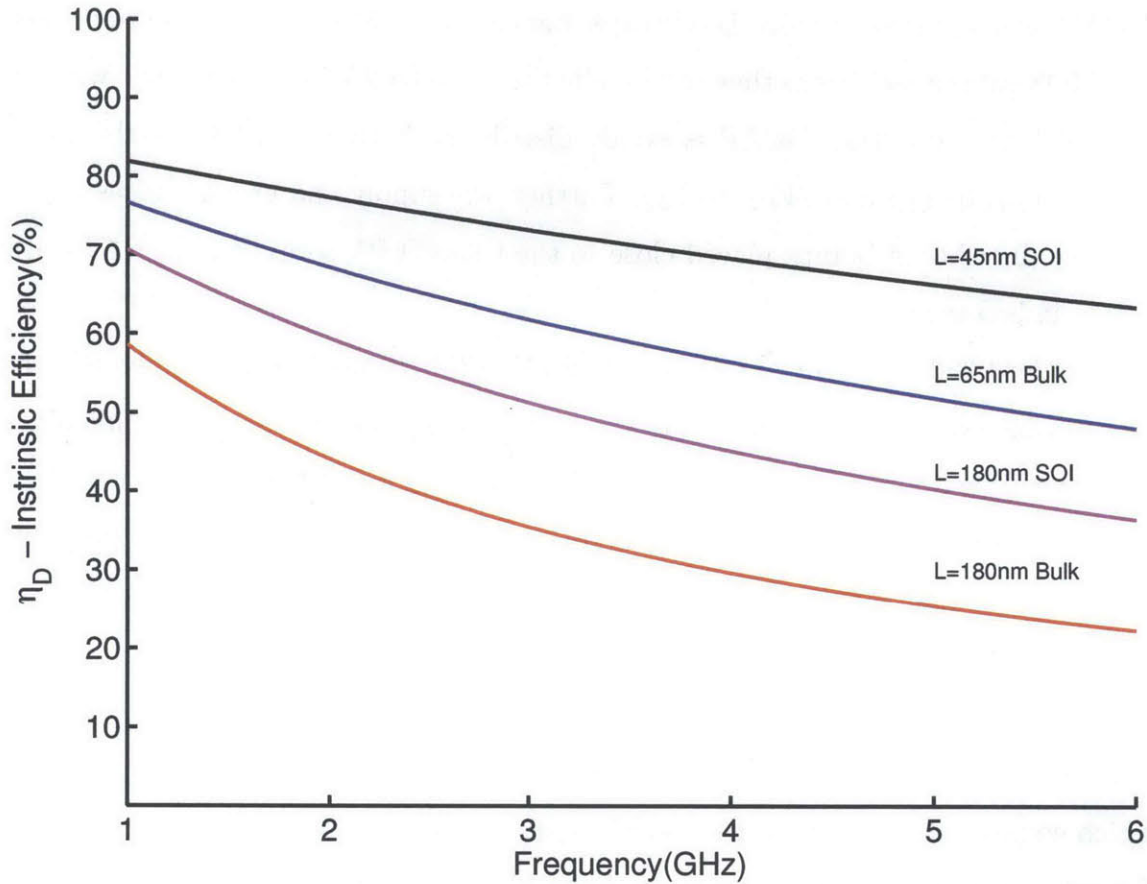
Figure 4-21: Impact of CMOS scaling on PA unit cell efficiency ($\nu = 0.15$)

tify the improvement by plotting model predicted values for representative CMOS technologies.

4.9 Layout, ESD and packaging

The prototype is integrated in a 45-nm SOI CMOS process. The die micrograph is shown in Figure 4-22. The RF frontend occupies an active area of only 1.5 mm × 1.4 mm, no larger than a single CMOS PA with comparable output power [31] [53] . Flip-chip interconnect is used for superior RF performance. The Class-D PA produces current transients at the switching frequency f_{RF} akin to clocked digital circuits. Current transients, if not addressed, can cause supply droops and severely degrade PA performance. To maintain supply integrity, 400 pF of decoupling capacitance

(DCAP) is integrated on-chip. Lower capacitance density MoM capacitors are favored over MOS gate capacitors as they can handle the $2V_{DD}$ (= 2 V) supply voltage without any reliability concerns. DCAP is evenly distributed between all three rails: $2V_{DD}$ to V_{DD}, V_{DD} to V_{SS} and $2V_{DD}$ to V_{SS}. Further, the supply and ground buses utilize multiple distributed bumps placed close to the Class-D PA cores to reduce parasitic supply inductance.

Standard ESD structures, robust to > 2 kV HBM are used on all supply, RF and control pins except V_{TRX}. Protecting V_{TRX} with clamps or diodes is very difficult, if not impossible due to the simultaneously high-frequency and high-voltage waveform encountered in the TX mode. Fortunately, another elegant feature of the proposed architecture is that the grounded transformer secondary inductor at this node provides inherent ESD protection [68] [69] at no capacitance or area penalty.

To facilitate testing, a two-stage package is designed. As shown in figure 4-23(a), the prototype IC is first flipped on to a custom ceramic chip-scale thin-film package which employs Alumina substrate which 6 μm gold metallization. Alumina can handle extreme thermal temperatures of above 350 °C needed to perform the flip-chip attachment. The package also has space to mount 0201-size ceramic decoupling capacitors close to the IC to guarantee supply integrity. Once assembled, the Alumina package is then mounted to a 4-layer Rogers RO4350B PCB shown in figure 4-23(b), and electrical connections are made through 15-mil gold ribbons. The PCB has additional 0402-, 0603- and 1206-size decoupling capacitors, as well as wideband baluns to convert the incoming RF drive signals to differential before they are fed in to the IC.
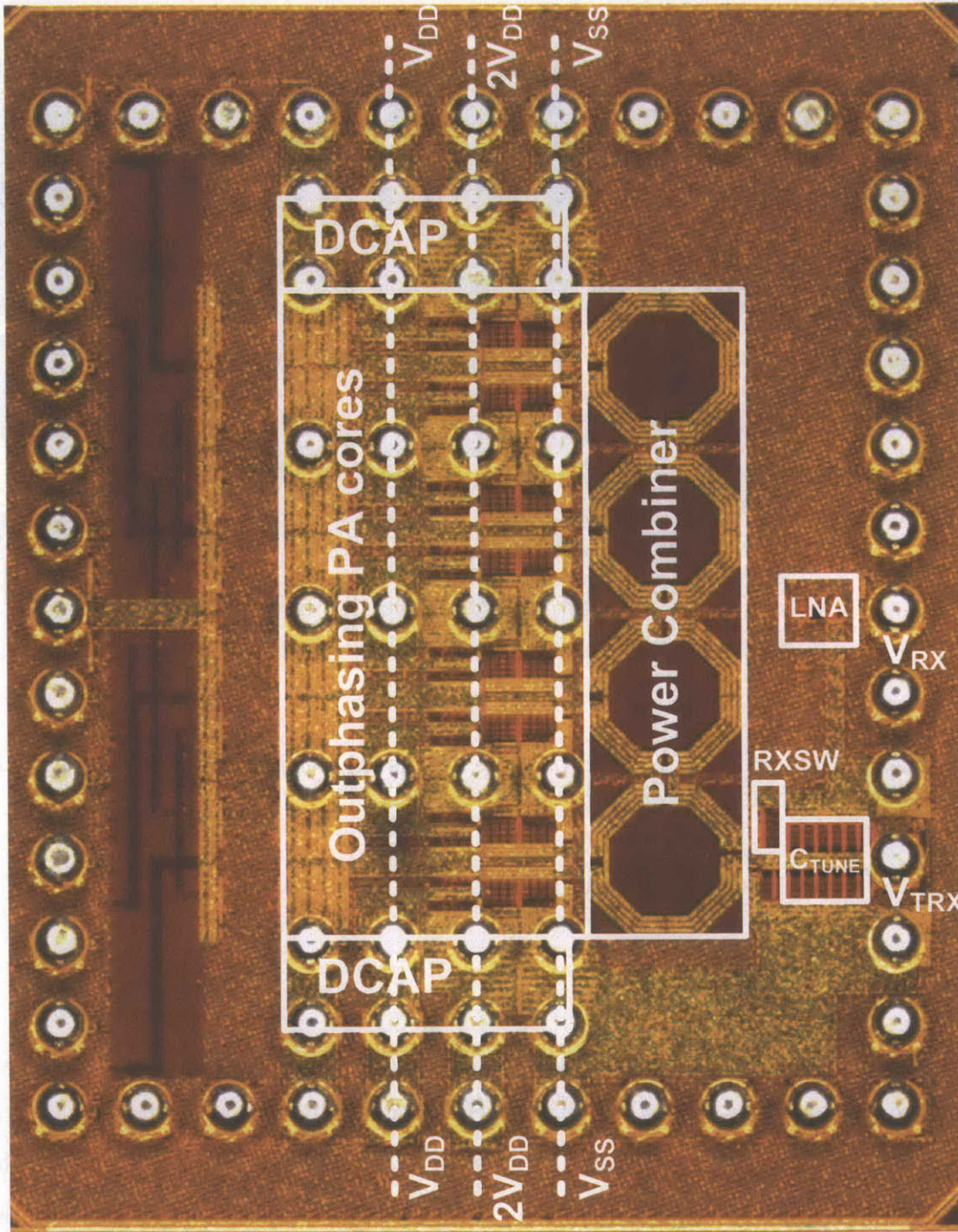
Figure 4-22: Die micrograph

105
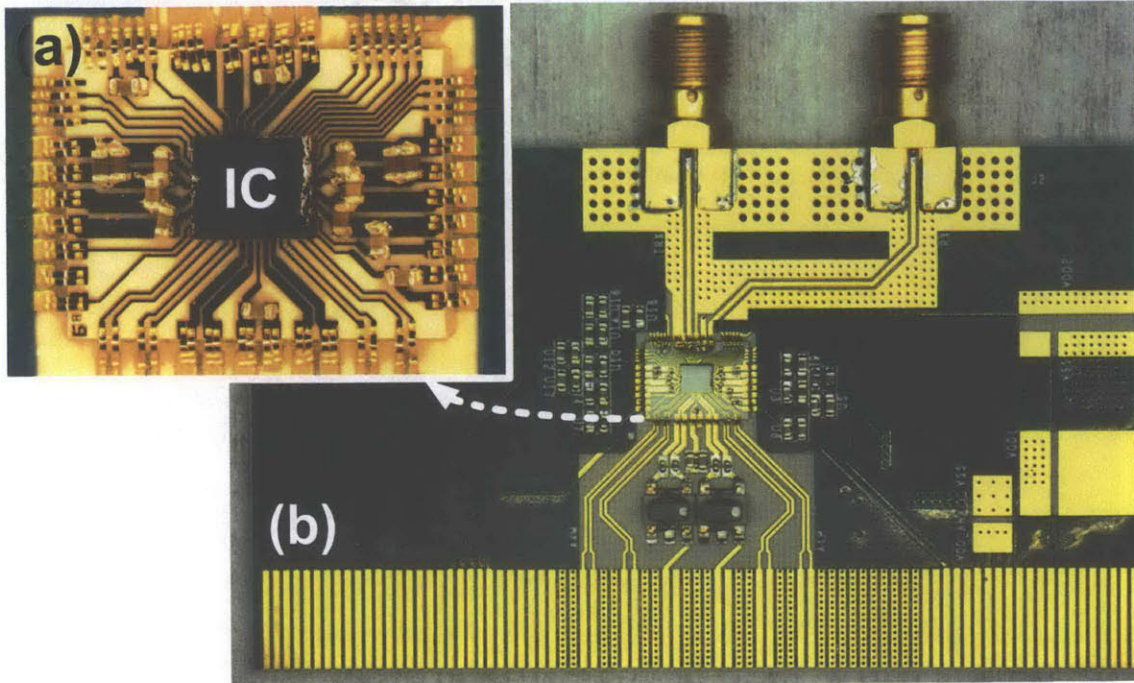
Figure 4-23: Custom two-stage RF package details (a) Ceramic (Al_2O_3) package with flip-chip IC attachment (b) Rogers RO4350B PCB

4.10 TX measurements

4.10.1 TX mode setup

The TX mode experimental setup is shown in figure 4-24. The entire setup is controlled via a computer through multiple USB connections. A custom motherboard is designed to support the testing and is based on a commercially available FPGA platform. Two commercially available direct up-conversion transmitters, which combine dual IQ-DAC's and complex modulators are used to generate RF input signal vectors V_1 and V_2. The motherboard supplies the baseband data (I_1, Q_1, I_2, Q_2) to the IQ modulators in real-time. It should be noted that the IQ modulators are in fact used only to modulate phase in real-time, the amplitude information being coded into the phase itself via the outphasing control law (see appendix C). θ is the phase of the original RF signal, while $\pm\phi$ is the outphasing angle. A low phase-noise signal gener-

106
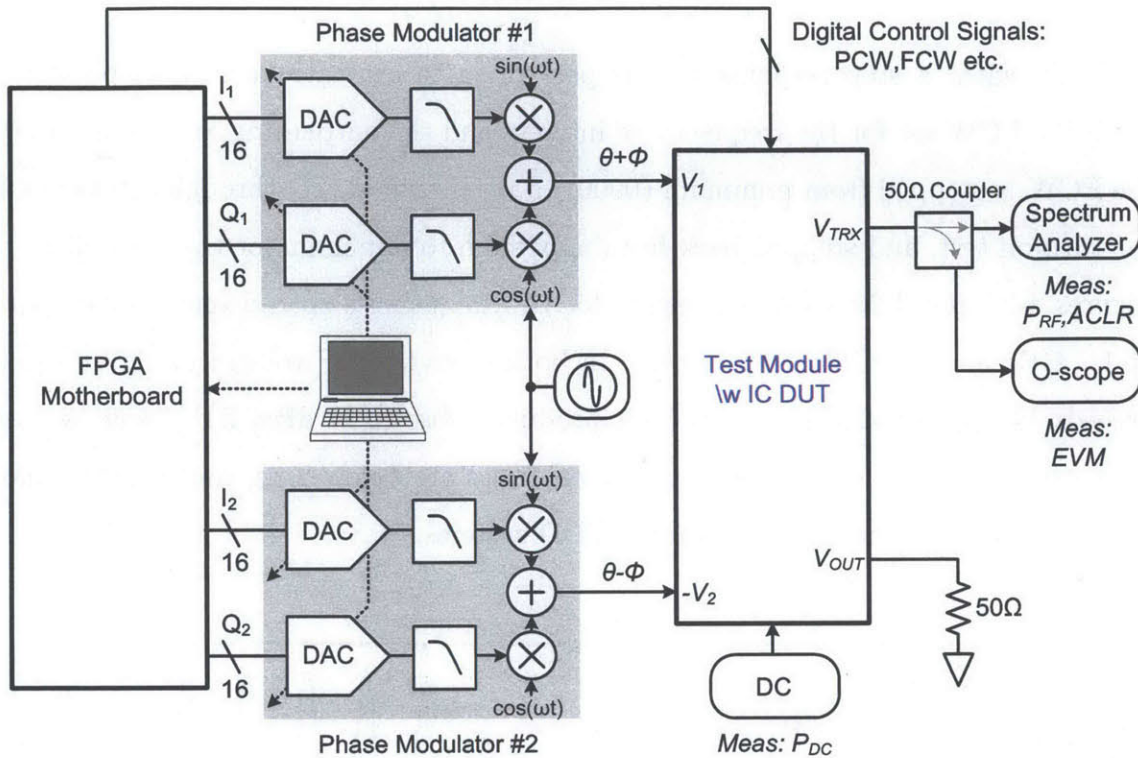
Figure 4-24: TX mode experimental setup

ator is used as the RF source. The specific instruments used for testing are listed in appendix D. The two outphasing vector V_1 and V_2 are fed from the IQ modulators into the test module. Single-ended to differential conversion occurs on-board before being the signal is fed into the IC, which contains internal differential 50 Ω terminations. The port V_{TRX} is connected to the spectrum analyzer and oscilloscope simultaneously through low loss RF cables and a power splitter to monitor the PA output power and EVM respectively. No off-chip filtering of any kind is applied between V_{TRX} and the measurement equipment. All losses from V_{TRX} to the test equipment are carefully de-embedded to get accurate RF power measurements. Accurate measurement of DC power is equally important for efficiency measurements. A highly accurate sourcemeter with 4-wire supply sensing is used to de-embed losses due to DC resistance in the supply traces. The RX path output V_{OUT} is terminated to 50 Ω.

4.10.2 Turn on/off transients

To begin with, a step response test is performed to evaluate basic PA operation. With the FCW set for the frequency of interest and the outphasing angle ϕ set to 0, the PCW is stepped from minimum (0000) to maximum (1111) through software for the turn-on test, and stepped back from maximum to minimum for the turn-off test. Figures 4-25 and 4-26 show the results for two frequencies spaced one octave apart at 1.7 GHz and 3.4 GHz, respectively. In both cases the RF voltage swing at V_{TRX} exceeds 14 V_{PP}, validating that the PA produces close to 27 dBm RF power over at least one octave. The turn on and turn off times are below 2 ns, much faster than the μs range switching periods of most TDD systems.
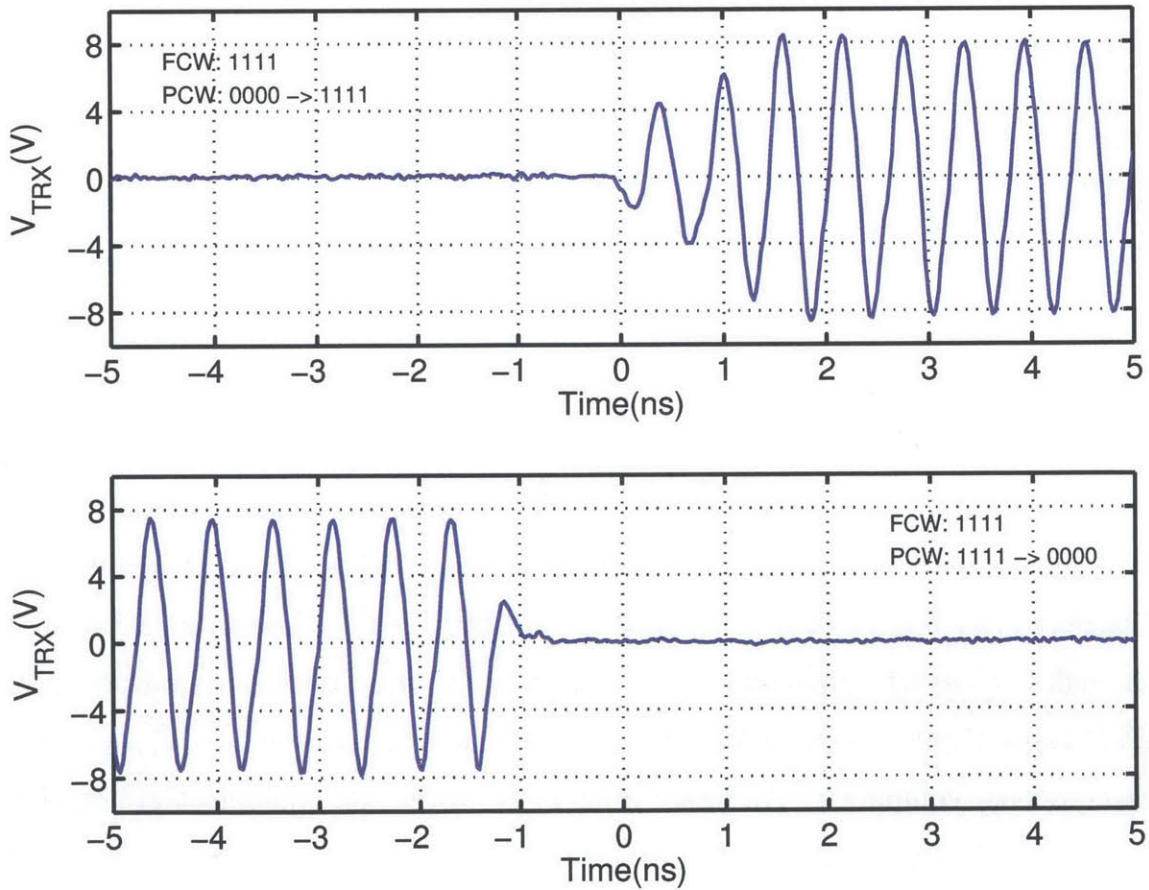
Figure 4-25: TX step response at $f_{RF} = 1.7$ GHz
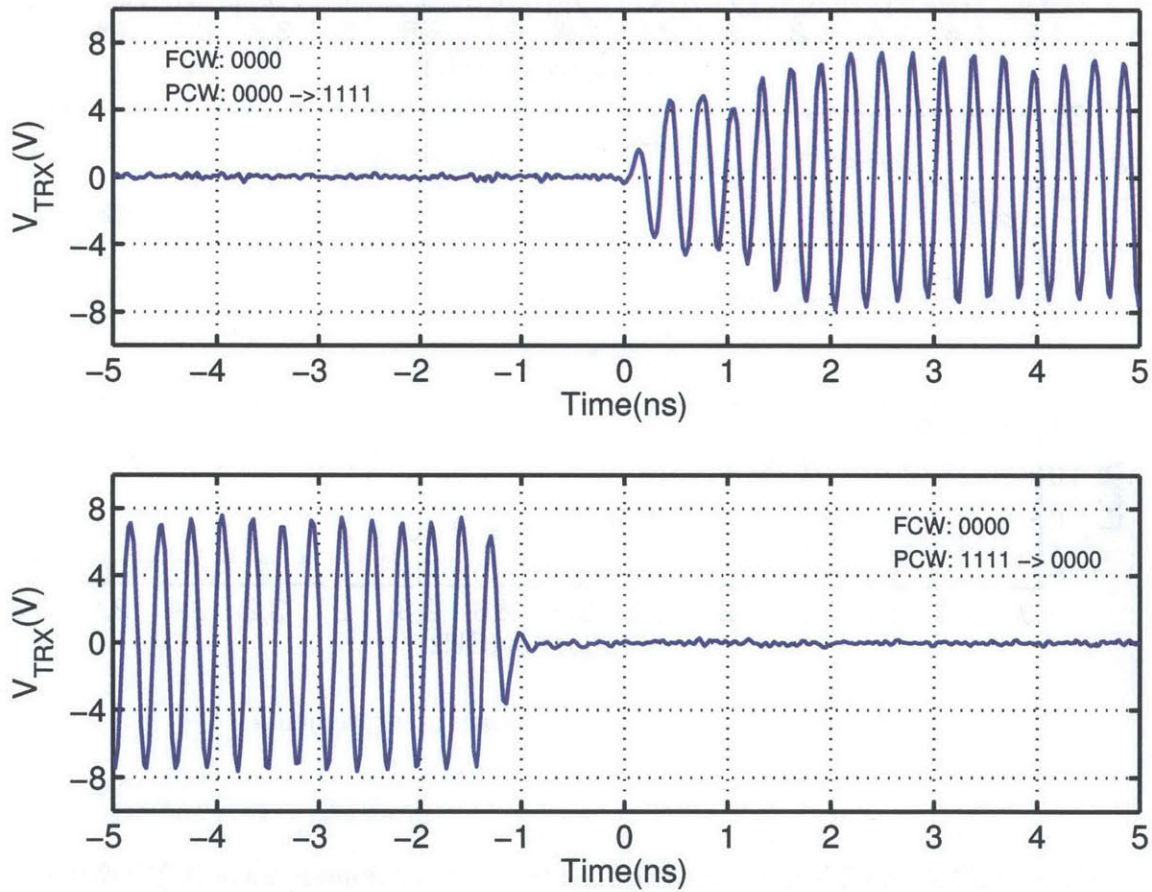
Figure 4-26: TX step response at $f_{RF} = 3.4$ GHz

4.10.3 Continuous-wave (CW) performance
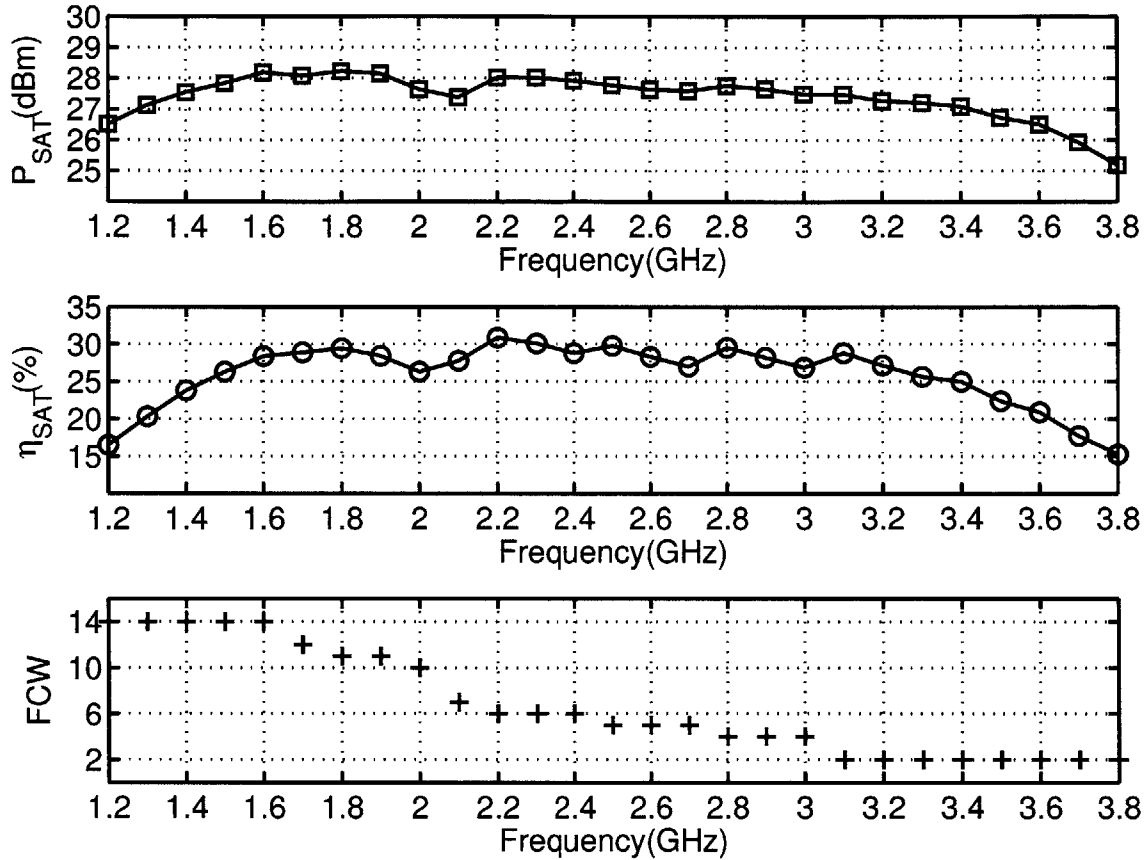
Figure 4-27: TX continuos-wave measurements

Figure 4-27 shows TX mode measurements under continuous-wave (CW) drive. At each frequency, the FCW is optimized to maximize saturated output power (P_{SAT}), followed by fine adjustments in the *static* outphasing angle (ϕ) to calibrate path delay mismatches. Only fundamental[6] output power is included in the measurements. As shown, nearly constant P_{SAT} of 27.7 ± 0.5 dBm is measured from 1.3 to 3.3 GHz. Best-case total efficiency ($\eta_{SAT} = \frac{P_{SAT}}{P_{DC}}$, including all 6 driver stages) is 30 % at 2.2 GHz.

---

[6]At the RF frequency, excluding all harmonics

4.10.4 Static outphasing response

The static outphasing response of the TX is validated by sweeping the outphasing angle ϕ over the entire $(+\frac{\pi}{2} -\frac{\pi}{2})$ range, for all four PCW settings. Figures 4-28(a) - 4-28(c) show the results over a frequency range of 1.44 - 3.41 GHz. Both output power and efficiency are normalized to the maximum (saturated) values reported in figure 4-27. The test frequencies are aligned to lie in relevant LTE bands. Good agreement with the outphasing control law is observed. Note that maximum output power does not occur at exactly $\phi = 0$ due to delay mismatch between the two phase paths. This mismatch is calibrated out for all the other TX measurements including the CW performance already presented in figure 4-27.

The efficiency trend also shows the benefit of non-isolated transformer based combining. For instance, at the mid-range frequency of 2.31 GHz shown in figure 4-28(b), the normalized efficiency at 6 dB back-off ($\frac{P_{OUT}}{P_{SAT}} = 0.25$) with the PCW set to maximum (1111) is 0.43, as compared to the expected value of 0.25 for the isolating case, a 1.7× improvement.

The results also show the promise of further efficiency enhancement with *dynamic or signal-dependent* PCW modulation, which would make the TX a multi-level outphasing system. For instance, as shown in figure 4-28(b) realizing 6 dB back-off with a lower PCW of 0011 rather than 1111 would yield a normalized efficiency of 0.59, another 1.4× improvement. However, this technique comes at the cost of linearity has already been validated recently (albeit for a single band design) in [53], so the work will not be repeated here.
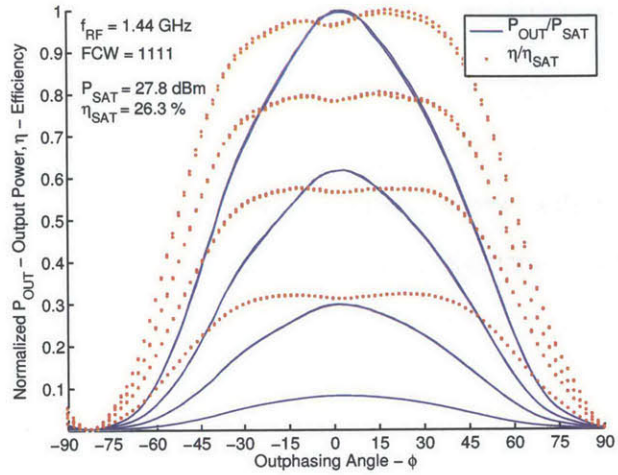
(a) $f_{RF} = 1.44$ GHz (LTE Band 11)
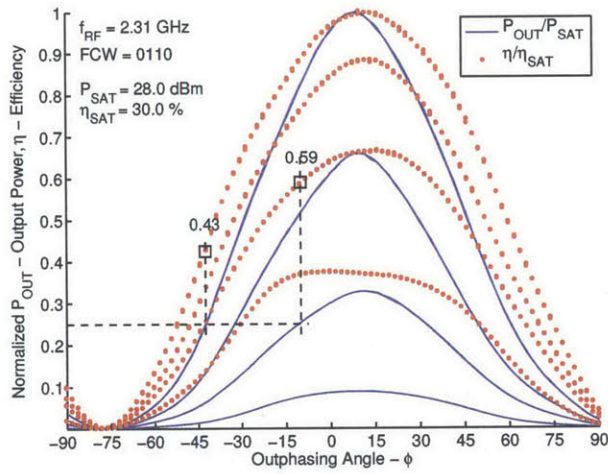
(b) $f_{RF} = 2.31$ GHz (LTE Band 40)
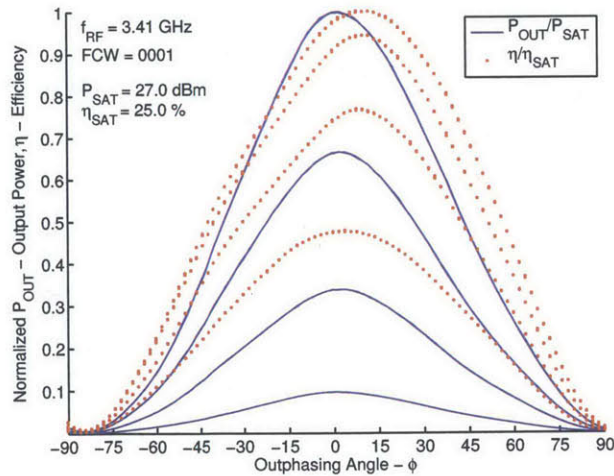
(c) $f_{RF} = 3.41$ GHz (LTE Band 42)

Figure 4-28: Normalized output power and efficiency vs. power control word (PCW)

112

4.10.5 Modulated signal performance

Tests are performed with 64-QAM, 20 MHz modulated signals with a PAPR of approximately 5.2 dB after performing some crest-factor reduction. Various test frequencies corresponding to LTE bands spanning more than one octave are used to thoroughly characterize the TX. Lookup table based predistortion [24] is applied to improve performance. For each measurement, the optimal FCW is used. In-band output power (P_{OUT}, integrated over 20 MHz) is measured as before with the calibrated spectrum analyzer. Adjacent channel leakage ratio (ACLR) is also measured with channel BW and offsets for E-UTRA and UTRA scenarios [47]. Error vector magnitude (EVM) measurements are performed by capturing the data with a high-speed oscilloscope.

For 64-QAM modulation, the measured output spectra and demodulated IQ constellations for selected frequencies are shown in figures 4-29(a) to 4-29(c), while a detailed summary is presented in table 4.3. P_{OUT} of 22.5 $\pm$ 0.9 dBm is measured from 1.44 to 3.41 GHz. Best-case average efficiency is 17.4 % at 1.72 GHz with 23.4 dBm P_{OUT}.

Further tests with 802.11g (64-QAM OFDM, 20 MHz) signals at 2.41 GHz show that the PA also meets the WLAN mask at 21.6 dBm P_{OUT} and -25.5 dB EVM.

f_{RF} (GHz)	LTE TDD	LTE FDD	FCW	P_{OUT} (dBm)	Eff. (%)	E-UTRA ACLR1 (dBc)	E-UTRA ACLR2 (dBc)	UTRA ACLR1 (dBc)	EVM (dB)
1.44		11(21)	1111	22.7	17.2	-32.3	-39.3	-36.6	-31.5
1.72		3,4,10(9)	1100	23.4	17.4	-32.9	-37.9	-37.2	-31.8
1.91	33,39(35-37)	(2,25)	1011	23.1	14.9	-31.9	-34.2	-36.4	-32.4
2.31	40		0110	22.7	14.4	-35.3	-38.8	-39.8	-35.7
2.60	38,41	(7)	0101	22.8	13.2	-33.9	-38.8	-38.3	-35.6
3.41	42		0001	21.6	9.8	-30.9	-35.6	-35.2	-30.6

Table 4.3: TX performance in multiple LTE bands for 64-QAM, 20 MHz signals with PAPR=5.2 dB
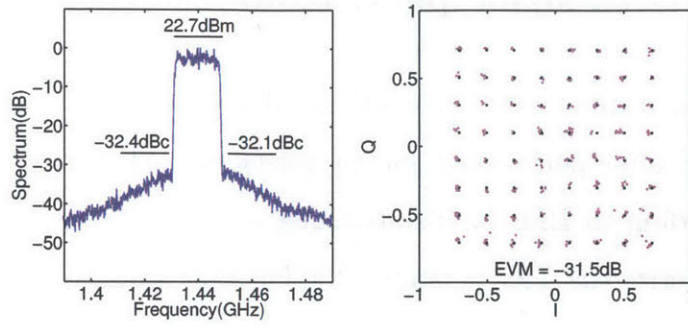
(a) $f_{RF} = 1.44$ GHz (LTE Band 11)
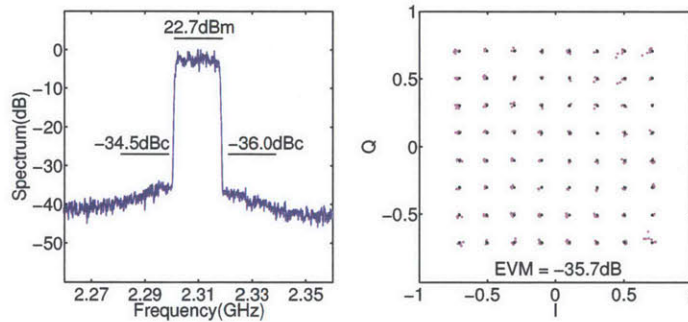
(b) $f_{RF} = 2.31$ GHz (LTE Band 40)
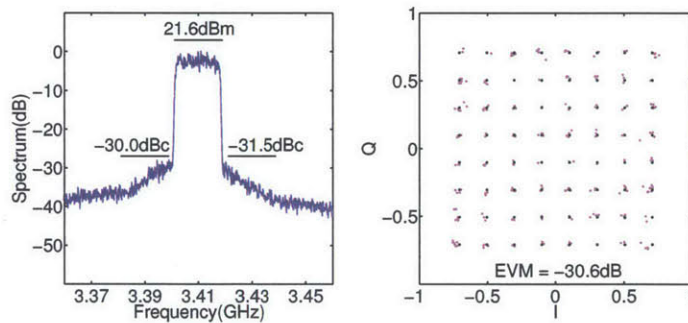
(c) $f_{RF} = 3.41$ GHz (LTE Band 42)
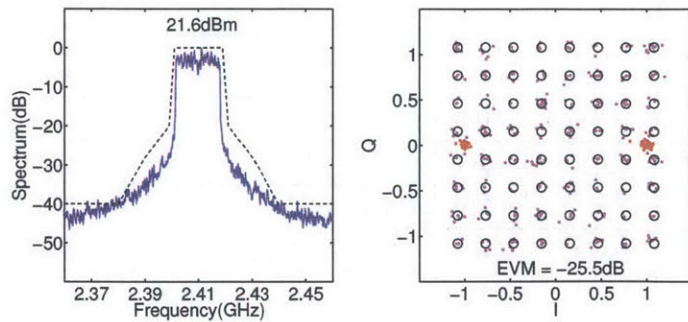
(d) $f_{RF} = 2.41$ GHz (802.11g Band)

Figure 4-29: Outphasing TX performance with 20 MHz, 64-QAM modulated signals with PAPR = 5.2 dB

114

4.11 RX measurements

4.11.1 RX mode setup
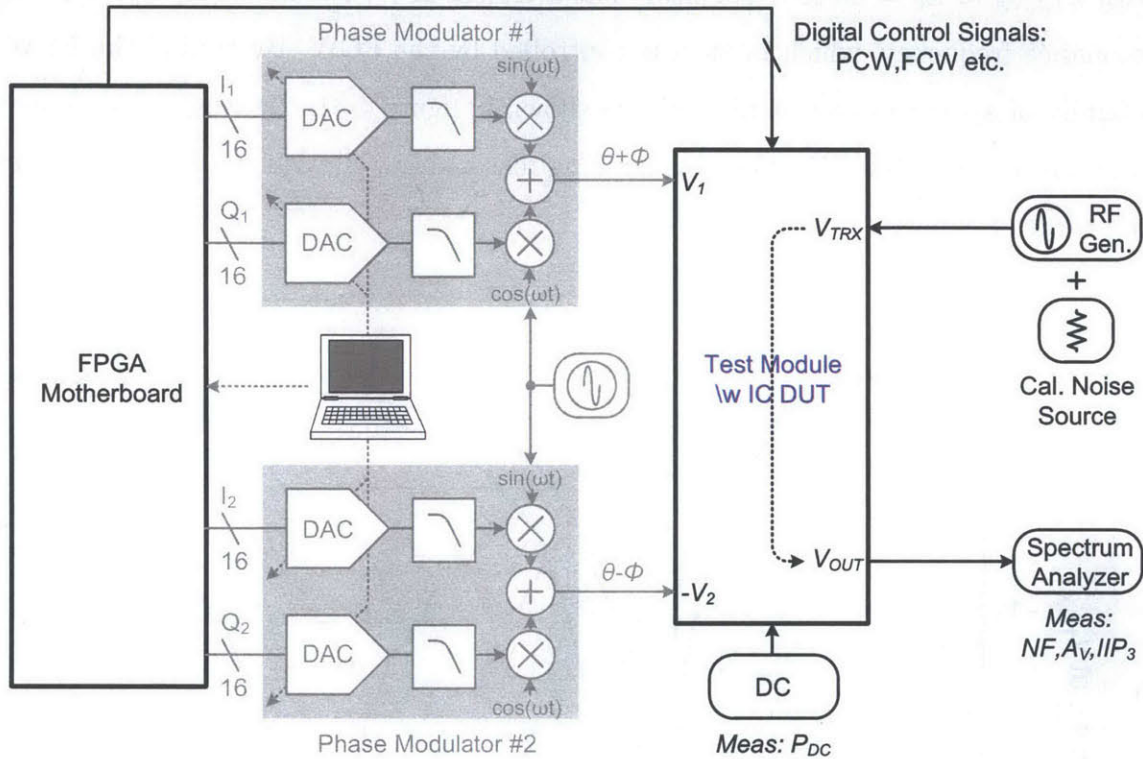
Figure 4-30: RX mode experimental setup

The RX mode experimental setup is shown in figure 4-30. IQ modulators are not required and deactivated through software. V_{TRX} now serves as input for the RX signal while V_{OUT} is the output, observed with a spectrum or network analyzer. A calibrated noise source is used to measure noise figure. For IIP_3 measurements, the two-tone function of the RF signal generator is used.

4.11.2 Small-signal measurements

Testing the RX mode allows observing the resonance of the tunable matching network directly, which largely determines the frequency response of the RX path. With all

other ports terminated, the RF I/O port V_{TRX} can be connected to the network analyzer to observe the return loss, with the LNA control words (GCW, RFCW and ROCW) set to nominal values to power match the RX. Under this condition, the V_{TRX} is essentially loaded by the tunable matching network (L_{TFMR}, C_{TUNE}) in parallel with $Z_{I,LNA} \approx Z_0 = 50 \ \Omega$ [7]. Its clear that a trough in s_{11} should be observed at the resonance frequency, which in turn is controlled by the FCW. By tuning the FCW, a family of s_{11} traces can be plotted. As shown in figure 4-31, the tunable resonance ranges from 1.6 to 3.5 GHz, thereby agreeing reasonably well with the designed values and target frequency range.
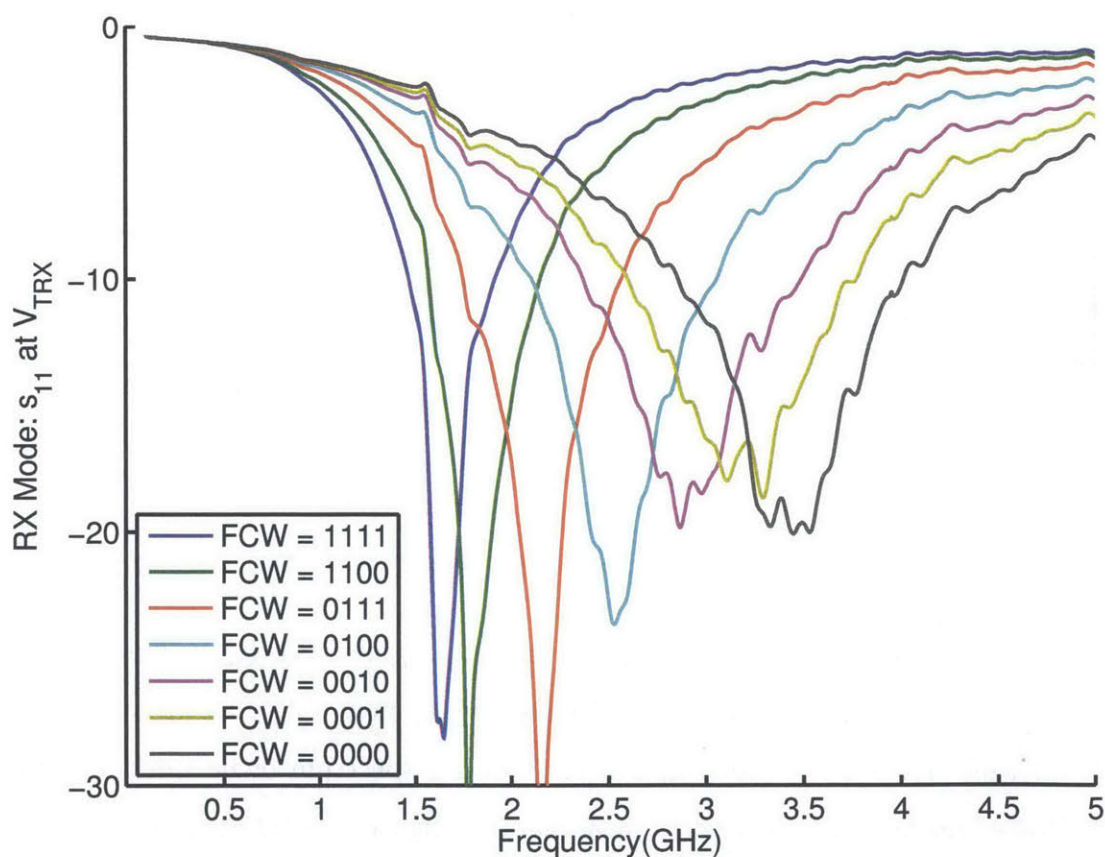
Figure 4-31: Frequency Response

Figure 4-32 shows RX mode LNA small-signal measurements. The LNA control words are set to maximize the voltage gain (A_V) and minimize noise figure (NF),

---

[7]Assuming the off-state PA unit cell loading is negligible, as discussed in section 4.5.4

while the FCW is optimized with frequency to match the input using the data of figure 4-31. $A_V > 14$ dB, NF = 4.4 ± 1.6 dB are measured from 1.3 to 3.3 GHz, with only 6 mA current drawn from $V_{DD} = 1$ V. The power consumption of the voltage buffer is not included in the measurement. RX branch switch loss $(P_{L,RXSW})$ is also individually measured with a test structure to be 2 dB.
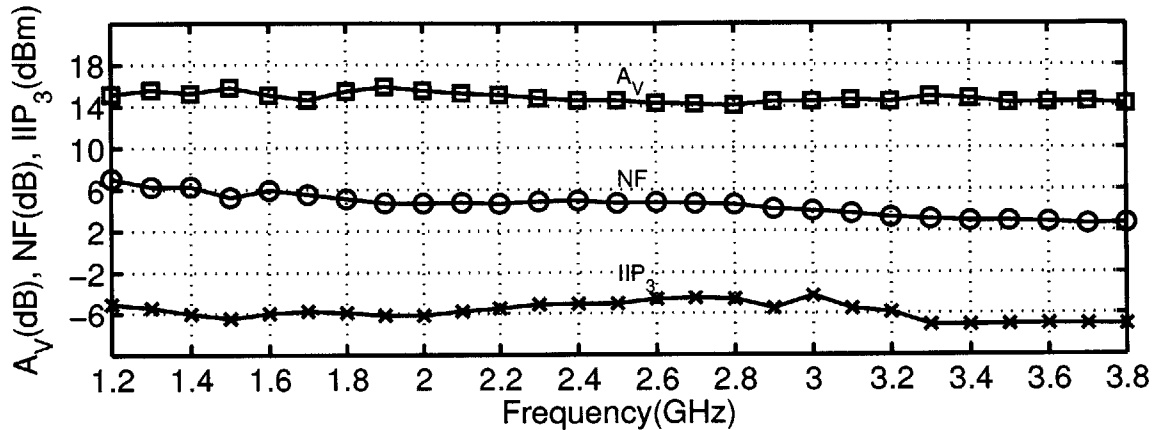
Figure 4-32: LNA small-signal measurements - A_V: Voltage gain, NF: Noise figure, IIP_3: Input-referred third order intercept

4.11.3 Third-order intercept test

To evaluate linearity, the input-referred third-order intercept (IIP_3) is measured across a range of frequencies. A tone spacing of 20 MHz is used. Figure 4-32 summarizes the results. The LNA achieves $IIP_3 > -7$ dBm across the entire frequency range.

Figure 4-33 shows details of how the measurement is performed for one frequency point centered at 1.8 GHz. The optimal FCW is set as with all other measurements. A signal generator is used to synthesize the two tones 20 MHz apart (f_{RF1}, f_{RF2}) with equal power. The total power in these two tones (P_{IN}) is then swept, and the fundamental $(P_{RF1} + P_{RF2})$ and third-order inter modulation power $(P_{2RF1-RF2} + P_{2RF2-RF1})$ is measured. Linear extrapolation to the hypothetical point where the two power levels meet gives the uncalibrated, input-referred third-order intercept $(P_{L,RXSW} + IIP_3)$, from which the RX switch loss should be subtracted to the get

117

the IIP_3 referred to the input of the LNA.
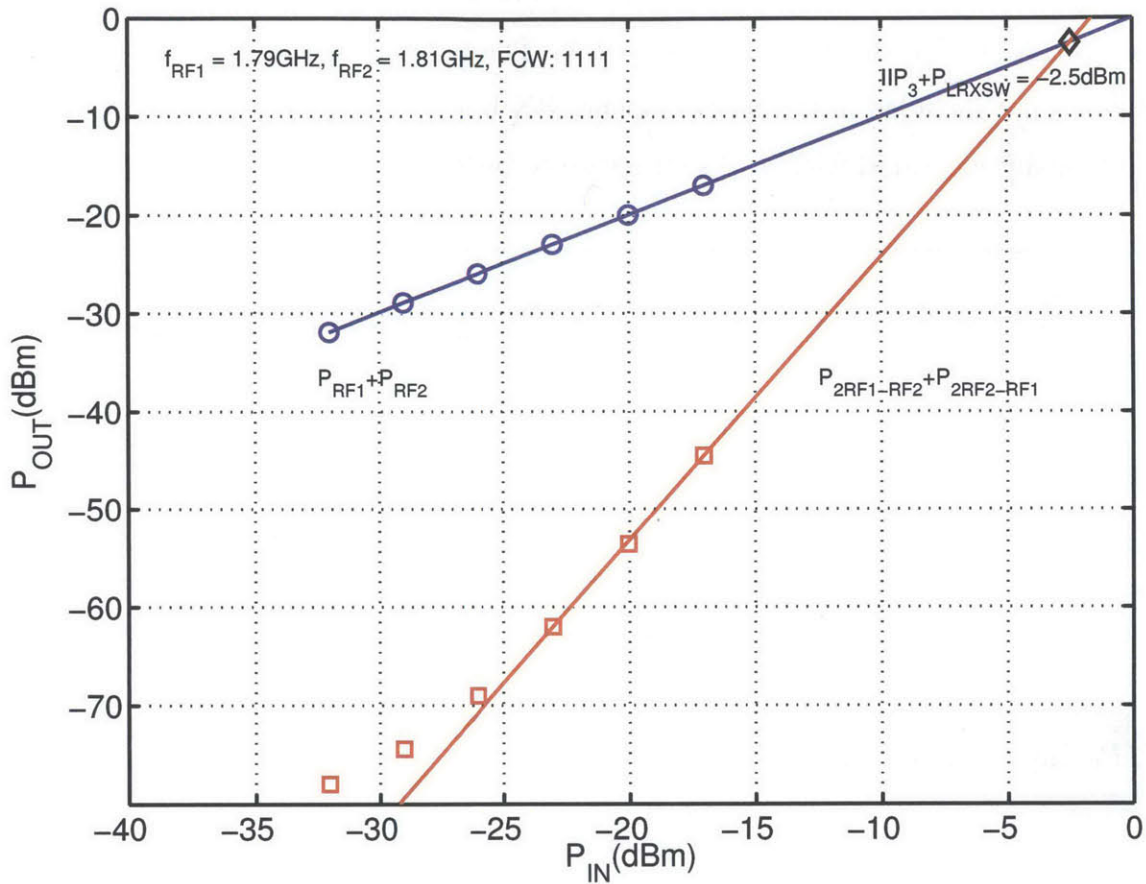
Figure 4-33: LNA two-tone power sweep for IIP_3 - $f_{RF1} = 1.79$ GHz, $f_{RF2} = 1.81$ GHz

4.12 Performance comparison and conclusions

4.12.1 Comparison with state-of-the-art CMOS PA's

For a fair comparison, only silicon based designs (bulk CMOS, SOI, SOS) at frequencies up to 6 GHz that achieve a P_{SAT} greater than 20 dBm are considered. All of the designs are fully integrated and include driver circuits.

Achieving simultaneously high PA bandwidth (BW), output power and efficiency is very challenging due a number of reasons. First, impedance transformation networks are usually BW limiting. Second, loss usually trades off against BW, lowering

efficiency of broadband designs. Third, harmonic tuning techniques to enhance efficiency tend to be inherently narrowband. To evaluate the present work in this context, a comparison with prior work in the (BW,P_{SAT}) and (BW,η_{SAT}) space shown in figures 4-34(a) and 4-34(b), respectively. To decouple the comparison from the absolute value of the center frequency f_C, a percentage measure of bandwidth is used: $BW = \frac{f_{high} - f_{low}}{f_C}$. A few relevant observations can be made about prior work. First, Class-AB PA designs [13] [14] [45] [70], which rely on multiple tuned stages with *fixed* matching networks do not exceed 30-35 % 3 dB BW[8]. Second, switching architectures [22] [52] [53] [63], greatly cut down on the number of BW limiting reactive components, achieving up to 42 % 1 dB BW. Finally, an exception among switching architectures is [71], which approaches nearly one octave 1dB BW of 64 %.

This work, based on a digital PA core with a *tunable* high-power matching network compatible achieves the widest 1 dB BW of 87 % reported to date.

It is noteworthy that the above comparison is still not completely fair since this work includes a TXSW while the standalone PA's referred do not. *Equivalent* PA output power and efficiency for the present work assuming $P_{L,TXSW} = 1$ dB is also plotted in figures 4-34(a) and 4-34(b). If needed, output power can be further boosted through more aggressive power combining [53] or custom high-voltage device engineering [77].

This PA performance comparison is not complete without consideration of linearity. In this regard, it should be noted that this work demonstrates linearity that is competitive with other published architectures with much lower BW. Specifically, the three closest competitors to the present work [22] [53] [71] in figure 4-34(b) utilize modulated test signals which are either less stringent or approximately equivalent to those used here, yielding comparable test results in terms of ACLR and spectral mask compliance. It is also worth noting that [22] [53] only report linear test results at one frequency. Even for the closest competitor [71], linear test results are only available for a relatively narrow range of 11 % compared to reported 1 dB BW of 64 %. The present work for the first time demonstrates consistent linearity over a

---

[8]These publications only report 3 dB BW, so 1 dB BW numbers are not available.

multi-octave frequency range. Further, there is still room for improvement. Specifically, this work utilizes a relatively simple lookup table based static predistortion algorithm [24]. In future work, adoption of more sophisticated linearization schemes which account for memory effects will further improve performance, truly breaking the bandwidth-efficiency and bandwidth-linearity tradeoffs.
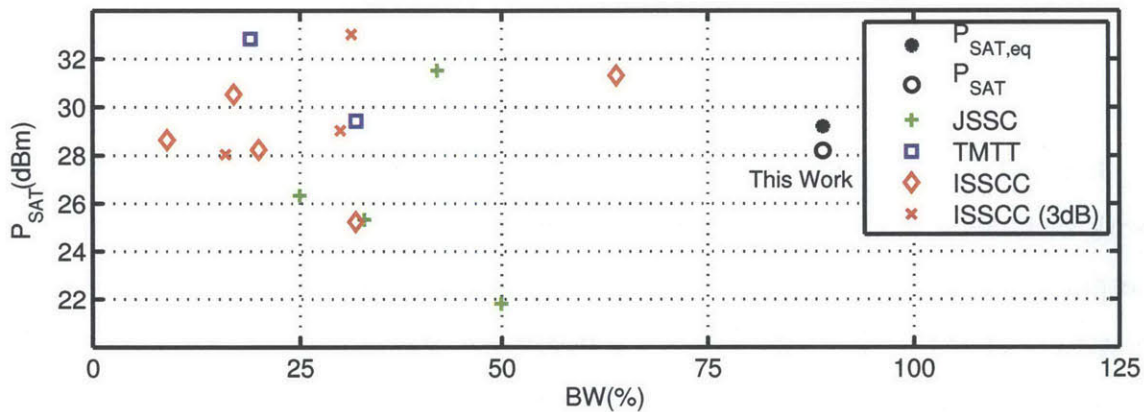
(a) Output power vs. bandwidth
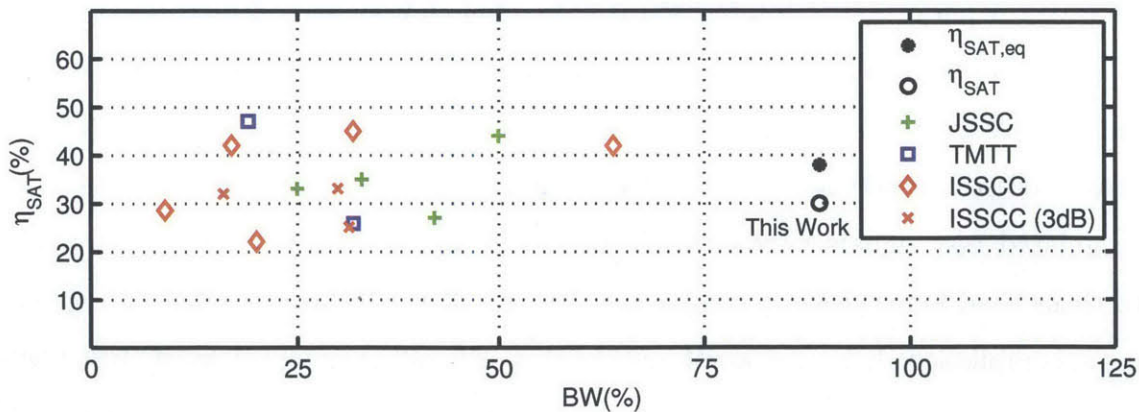
(b) Efficiency vs. bandwidth

Figure 4-34: Comparison with state-of-the-art CMOS PA's; JSSC - [22] [53] [63] [72]; TMTT - [73] [74]; ISSCC - [45] [52] [71] [75] [76]; ISSCC (3dB) - [13] [14] [70] (these papers report 3 dB BW instead of 1dB)

4.12.2 Comparison with state-of-the-art CMOS RF frontends

Finally, the design is compared with prior work on TDD RF frontends in CMOS. The performance comparison is shown in table 4.4[9]. While prior work targets one or two bands, the modulated frequency range of 1.44 - 3.41 GHz achieved through the proposed frequency-agile architecture in this work covers 10 (out of 12 currently defined) TDD-LTE bands[10], as well as the unlicensed band used for WLAN (802.11g) / Bluetooth (2.4 - 2.5 GHz) and WiMAX (2.3 - 2.7 GHz and partially 3.3 - 3.8 GHz), thereby offering an unprecedented level of RF frontend integration for TDD scenarios.

Parameter	[34]	[78]	[35]	This work
Application	WLAN	WLAN	WLAN	WLAN/TDD-LTE
CMOS Process	45-nm	32-nm	65-nm	45-nm SOI
Integrated TX/RX SW	Yes/No	Yes	No	Yes
Integrated balun	Yes/Yes	No	Yes	Yes
Freq. band range (GHz)	2.4-2.5/4.9-5.9	2.4-2.5	2.4-2.5	1.44-3.41
PA's+LNA's	2+2	1+1	1+1	1+1
No. bands	2	1	1	12
P_{SAT} (dBm)	29.0/26.0	27.1	27.5	27.6±0.6
η_{SAT} (%)	31.3/32.1	29.0	35.3	25.0-30.0
P_{OUT} (dBm)	22.3/18.7	20.3	22.4	22.5±0.9
η (%)	-	14.6	18.0	9.8-17.4
Predistortion	Yes	No	Yes	Yes

Table 4.4: Performance comparison with other published TDD RF frontends.

---

[9]For [78], TX efficiency is calculated from PA efficiency using 1.3 dB loss reported for TXSW as TX numbers in this work include power dissipation of other blocks like the LO generator.

[10]In addition, 10 (out of 28 currently defined) FDD-LTE bands are also within the frequency range, but would require a different, system-level implementation involving duplexers, which is beyond the scope of this work.

Chapter 5

Conclusion

5.1 Summary of results

The results presented in this work represent several compelling advances in the area of RF frontend design for next-generation applications.

The RF frontend architecture presented in chapter 3 shows the promise of emerging GaN technology in enabling high-power RF integration. The efficiency advantage offered by GaN compared to other technologies for watt-level transmitters was theoretically quantified. A modified switching architecture was proposed to improve overall transmitter efficiency. A simple but effective transconductance superposition technique was demonstrated for GaN transistors to improve linearity. All active and passive components were fully integrated on one chip. The prototype IC was fabricated in a 250-nm GaN HEMT process targeting 802.11p, a recently ratified standard for vehicular connectivity. Measurements showed excellent performance in the 802.11p band with best-in-class transmitter efficiency for a monolithic solution.

The frequency-agile RF frontend architecture presented in chapter 4 represents a radical departure from conventional RF design techniques. Starting with the observation that CMOS processes have become increasingly more digital friendly, the transmit chain was re-architected with a fully switch-based architecture, for which key performance metrics were theoretically shown to directly benefit from CMOS scaling. The key design goal was not to have the best performance in any one frequency

band, but rather to demonstrate a more minimalistic and flexible architecture that performs competitively over a wide frequency range. The prototype IC, fabricated in 45-nm SOI CMOS showed competitive efficiency, output power and linearity over a multi-octave frequency range, for the first time opening up the possibility of using a single RF frontend for dozens of bands in next-generation mobile wireless systems.

Recall that in chapter 2, three principal design challenges were identified in RF frontends: (a) Monolithic integration (b) Energy-efficient operation and (c) Multi-octave frequency coverage. The architectures proposed in this work address all three challenges across a range of RF power levels. Since RF frontend design has a disproportionate impact on the mobile radio system as a whole, the advances presented herein go a long way towards enabling the vision of *ubiquitous* connectivity.

5.2 Future work

There are exciting opportunities for future work in the sphere of GaN for RF applications at both device and system level:

1. The intrinsic performance of GaN transistors has ample room for improvement. The prototype IC in this work utilized a commercial, 250-nm GaN process. There is ongoing research into scaling gate lengths to improve f_T [79] [80], enhancement of breakdown voltages with novel field structures, and enhancement mode devices which do not require negative control voltages [81]. As technology improves, GaN transistors will become intrinsically more energy-efficient, operate at higher RF frequencies and become suitable for more complicated circuit architectures with higher transistor count. With improved devices and architectural creativity, more efficient switch based architectures could be monolithically integrated in GaN. Further, the frequency-agile approach presented in chapter 4 could be extended to much higher power levels.

2. While GaN has showed promising results for high-power RF frontend integration, a complete vehicular connectivity radio based on 802.11p requires integra-

tion with the medium-access control, baseband and transceiver subsystems. In fact, the work presented herein is part of a much bigger project with the goal of delivering such a solution. These other radio subsystems are highly digital and do not need the high-power capabilities of GaN. Clearly, CMOS remains the technology of choice to implement them. In the short term, a complete solution would be a two-chip GaN-CMOS module in a single package. In the long term, there is tremendous interest in growing GaN alongside CMOS transistors. Indeed, some work has already been done in this direction [18] [82]. The ultimate solution would therefore be monolithic integration of an entire 802.11p radio on a GaN-enhanced CMOS process.

The frequency-agile architecture presented in this work for mobile applications is intended for TDD standards. While there are a growing number of RF applications and standards that can benefit from the techniques presented, *universal* RF frontends that impact *all* wireless applications will require more research into FDD scenarios to enable dual-mode (TDD+FDD) solutions. There are several unique challenges in FDD systems. Principal among them are:

1. Since the transmitter and receiver are operational simultaneously, the noise power spectral density of the transmitter in the receive band should be sufficiently low so as not to de-sensitize the receiver [83]. Employing dedicated filtering at the PA output or improving duplexer performance can help, but this usually comes at the cost of efficiency. This is one of the principal reasons why in spite of nearly a decade of work, CMOS based switching architectures have not yet made it into any FDD products.

2. Frequency-agile FDD architectures will require tunable RF duplexers and filters compatible with high output power. This has been a very active and challenging area of research extending well beyond solid-state circuit innovation, but a compelling solution that meets all system-level requirements remains elusive. MEMS [84] and liquid-RF [85] are two exciting areas which might result in a future breakthrough.

Appendix A

List of Abbreviations

ACLR: Adjacent Channel Leakage Ratio

ADC: Analog to Digital Convertor

BW: Bandwidth

C4: Controlled Collapse Chip Connection

CMOS: Complimentary Metal Oxide Semiconductor

CW: Continuous Wave

DAC: Digital to Analog Convertor

DRC: Design Rule Check

DUT: Design Under Test

ESD: Electrostatic Discharge

E-UTRA: Evolved Universal Terrestrial Radio Access

EVM: Error Vector Magnitude

GaAs: Gallium Arsenide

GaN: Gallium Nitride

HBM: Human Body Model

HEMT: High Electron Mobility Transistor

IC: Integrated Circuit

IoT: Internet of Things

FDD: Frequency Division Duplexing

LINC: Linear Amplification with Nonlinear Components

LNA: Low Noise Amplifier

LTE: Long Term Evolution

MIMO: Multiple Input and Multiple Output

NF: Noise Figure

OFDM: Orthogonal Frequency Domain Multiplexing

PA: Power Amplifier

PAPR: Peak to Average Power Ratio

PC: Personal Computer

PVT: Process, Voltage and Temperature

RF: Radio Frequency

RX: Receiver/Receive

SiGe: Silicon Germanium

SoC: System on a Chip

SOI: Silicon On Insulator

TDD: Time Division Duplexing

TX: Transmitter/Transmit

WLAN: Wireless Local Area Network

Appendix B

Transformer Analysis and Modeling

B.1   Transformer equivalent circuits
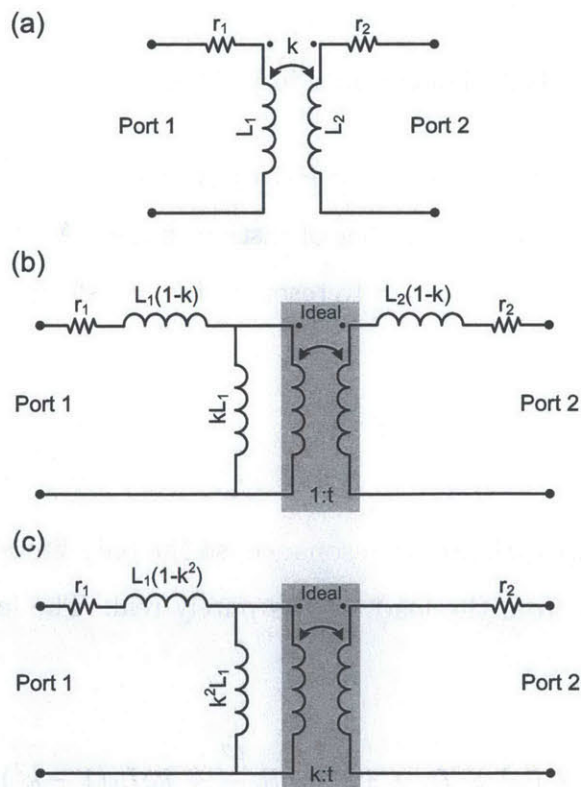
Figure B-1: Transformer equivalent circuit models

B.2 Impedance transformation

The coupling coefficient for on-chip transformers is typically well below unity. Imperfect coupling between primary and secondary windings impacts the impedance transformation.

To obtain design insight without excessive complexity, the finite quality factor of inductors is ignored in this analysis. This simplification is justified for two reasons. First, the objective here is to study impedance transformation, not to calculate efficiency. Second, for moderately high quality factors (≥ 10), this assumption only introduces minor error in the impedance calculation.
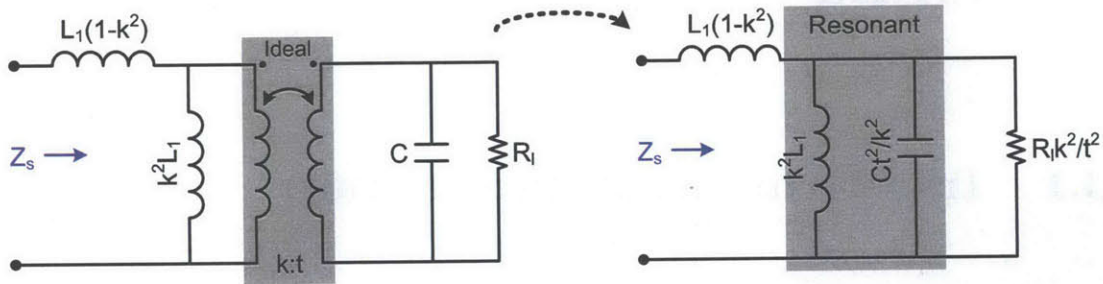
Figure B-2: Transformer impedance transformation

The non-symmetric model of figure B-1(c) is most convenient to use here, as it yields a network with the least number of distinct nodes. A single shunt capacitor is used in parallel with the load resistor to resonate the transformer inductance. Clearly, for resonance:

$$L_2\omega = \frac{1}{C\omega} \tag{B.1}$$

The resonator is an open-circuit at resonance, so the only finite admittance reflected to the primary stems from the load, and is purely real. The impedance seen at the input port is simply:

$$Z_s(\omega) = R_s + jX_s = R_l\frac{k^2}{t^2} + j\omega L_1(1 - k^2) \tag{B.2}$$

For moderate coupling, the first term dominates at low frequencies, yielding an

impedance transformation ratio of:

$$\frac{R_l}{R_s}(\omega_-) = \frac{t^2}{k^2} \tag{B.3}$$

If the primary is driven by a voltage source, then admittance is more useful to calculate extracted power:

$$Y_s(\omega) = \frac{R_l \frac{k^2}{t^2} - j\omega L_1(1-k^2)}{\left(R_l \frac{k^2}{t^2}\right)^2 + \omega^2 L_1^2(1-k^2)^2} \tag{B.4}$$

Equation B.4 indicates that as the frequency increases, the real part of the admittance will decrease due to frequency dependent term in the denominator, resulting in a reduction in extracted power. Nevertheless, equation B.3 is a useful starting point for design. Finally, the load power factor is:

$$p = \frac{R_l \frac{k^2}{t^2}}{\sqrt{\left(R_l \frac{k^2}{t^2}\right)^2 + \omega^2 L_1^2(1-k^2)^2}} \tag{B.5}$$

B.3   Lumped model extraction

The equations corresponding to the transformer model shown in figure B-1 are:

$$V_1 = (j\omega L_1 + r_1)I_1 + j\omega M I_2 \tag{B.6}$$

$$V_2 = j\omega M I_1 + (j\omega L_2 + r_2)I_2 \tag{B.7}$$

While the standard z-parameter representation is:

$$V_1 = Z_{11}I_1 + Z_{12}I_2 \tag{B.8}$$

$$V_2 = Z_{21}I_1 + Z_{22}I_2 \tag{B.9}$$

If the model is valid, the two representations are equivalent, therefore:

131

$$L_1 = \frac{imag(Z_{11})}{\omega} \tag{B.10}$$

$$r_1 = real(Z_{11}) \tag{B.11}$$

$$L_2 = \frac{imag(Z_{22})}{\omega} \tag{B.12}$$

$$r_2 = real(Z_{22}) \tag{B.13}$$

$$M = \frac{imag(Z_{12})}{\omega} \tag{B.14}$$

$$k = \frac{M}{\sqrt{L_1 L_2}} \tag{B.15}$$

Appendix C

Outphasing Control Law

(a)

(b)
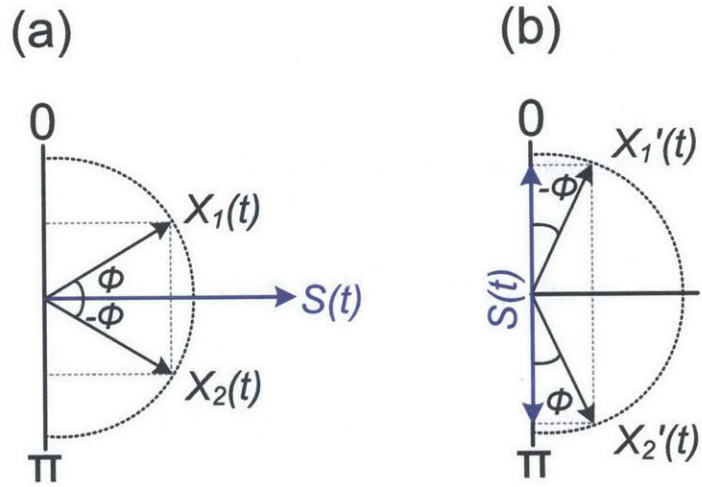
Figure C-1: Outphasing vector representation

$$S(t) = A(t)cos[\omega t + \theta(t)] \tag{C.1}$$

In an outphasing configuration, the amplitude information A(t) is represented by the outphasing angle $\phi(t)$:

$$\phi(t) = acos[\frac{A(t)}{max \mid A(t) \mid}] \tag{C.2}$$

The two outphasing vectors, assuming an arbitrary amplitude $R(t)$ are:

$$X_1(t) = R(t)cos[\omega t + \theta(t) + \phi(t)] \qquad (C.3)$$

$$X_2(t) = R(t)cos[\omega t + \theta(t) - \phi(t)] \qquad (C.4)$$

For the outphasing representation given by equation C.3 and C.4 to be equivalent to the original signal given in equation C.1 i.e. no linear scaling of output power:

$$S(t) = X_1(t) + X_2(t) \qquad (C.5)$$

$$R(t) = \frac{1}{2}max \mid A(t) \mid \qquad (C.6)$$

It is worth noting that outphasing works equally if the original signal is constructed with a difference of two vectors. In this alternate configuration, the original signal and corresponding outphasing vectors are given as:

$$S(t) = X_1'(t) - X_2'(t) \qquad (C.7)$$

$$X_1'(t) = R(t)cos[\omega t + \theta(t) - \phi(t)] \qquad (C.8)$$

$$X_2'(t) = R(t)cos[\omega t + \theta(t) + \phi(t) + \pi] \qquad (C.9)$$

Appendix D

List of RF Test Equipment

Calibrated noise source: Agilent Technologies N4000A

FPGA platform: Opal Kelly XEM5010

High-speed oscilloscope: Agilent Technologies DSA80000B

Network analyzer: Agilent Technologies N5242A

Phase modulator: Analog Devices AD9779A-DPG2-EBZ

Power supplies: Keithley Instruments 2602A, Agilent Technologies 3646A

RF signal generator: Agilent Technologies E8267C

Spectrum analyzer: Agilent Technologies N9020A

Bibliography

[1] *Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2012 - 2017.*

[2] "Tablet Shipments Forecast to Top Total PC Shipments in the Fourth Quarter of 2013 and Annually by 2015," *IDC Press Release*, 11th September 2013.

[3] M. P. du Rausas, J. Manyika, E. Hazan, J. Bughin, M. Chui, and R. Said, "Internet matters: The Net's sweeping impact on growth, jobs, and prosperity," *McKinsey Global Institute*, May 2011.

[4] J. Manyika and M. Chui, "All Things Online," *Foreign Affairs*, 13th September 2013.

[5] M. Chui, M. Loffler, and R. Roberts, "The Internet of Things," *McKinsey Quarterly*, March 2010.

[6] L. Atzori, A. Iera, and G. Morabito, "The internet of things: A survey," *Computer Networks*, vol. 54, no. 15, pp. 2787 – 2805, 2010.

[7] A. Behzad, K. Carter, E. Chien, S. Wu, M. Pan, C. Lee, T. Li, J. Leete, S. Au, M. Kappes, Z. Zhou, D. Ojo, L. Zhang, A. Zolfaghari, J. Castanada, H. Darabi, B. Yeung, R. Rofougaran, M. Rofougaran, J. Trachewsky, T. Moorti, R. Gaikwad, A. Bagchi, J. Rael, and B. Marholev, "A Fully Integrated MIMO Multi-Band Direct-Conversion CMOS Transceiver for WLAN Applications (802.11n)," in *IEEE International Solid-State Circuits Conference*, 2007, pp. 560–622.

[8] M. Ingels, V. Giannini, J. Borremans, G. Mandal, B. Debaillie, P. Van Wesemael, T. Sano, T. Yamamoto, D. Hauspie, J. Van Driessche, and J. Craninckx, "A $5mm^2$ 40nm LP CMOS 0.1-to-3GHz multistandard transceiver," in *IEEE International Solid-State Circuits Conference*, 2010, pp. 458–459.

[9] S. C. Cripps, *RF Power Amplifiers for Wireless Communications*, 2nd ed. Norwood, MA: Artech House, 2006.

[10] P. Reynaert and M. Steyaert, "A 2.45-GHz 0.13-μm CMOS PA With Parallel Amplification," *IEEE Journal of Solid-State Circuits*, vol. 42, no. 3, pp. 551–562, 2007.

[11] I. Aoki, S. Kee, D. Rutledge, and A. Hajimiri, "Distributed active transformer-a new power-combining and impedance-transformation technique," *IEEE Transactions on Microwave Theory and Techniques*, vol. 50, no. 1, pp. 316–331, 2002.

[12] A. Afsahi and L. Larson, "An integrated 33.5dBm linear 2.4GHz power amplifier in 65nm CMOS for WLAN applications," in *IEEE Custom Integrated Circuits Conference*, 2010, pp. 1–4.

[13] A. Afsahi, A. Behzad, and L. Larson, "A 65nm CMOS 2.4GHz 31.5dBm power amplifier with a distributed LC power-combining network and improved linearization for WLAN applications," in *IEEE International Solid-State Circuits Conference*, 2010, pp. 452–453.

[14] D. Chowdhury, C. Hull, O. Degani, P. Goyal, Y. Wang, and A. Niknejad, "A single-chip highly linear 2.4GHz 30dBm power amplifier in 90nm CMOS," in *IEEE International Solid-State Circuits Conference*, 2009, pp. 378–379,379a.

[15] A. Madan, M. McPartlin, Z.-F. Zhou, C.-H. Huang, C. Masse, and J. Cressler, "Fully Integrated Switch-LNA Front-End IC Design in CMOS: A Systematic Approach for WLAN," *IEEE Journal of Solid-State Circuits*, vol. 46, no. 11, pp. 2613–2622, 2011.

[16] A. Kidwai, C.-T. Fu, J. Jensen, and S. Taylor, "A Fully Integrated Ultra-Low Insertion Loss T/R Switch for 802.11b/g/n Application in 90 nm CMOS Process," *IEEE Journal of Solid-State Circuits*, vol. 44, no. 5, pp. 1352–1360, 2009.

[17] U. K. Mishra, P. Parikh, and Y.-F. Wu, "AlGaN/GaN HEMTs-an overview of device operation and applications," *Proceedings of the IEEE*, vol. 90, no. 6, pp. 1022–1031, 2002.

[18] J. W. Chung, B. Lu, and T. Palacios, "On-Wafer Seamless Integration of GaN and Si (100) Electronics," in *IEEE Compound Semiconductor Integrated Circuit Symposium*, 2009, pp. 1–4.

[19] F. Raab, P. Asbeck, S. Cripps, P. Kenington, Z. Popovic, N. Pothecary, J. Sevic, and N. Sokal, "Power amplifiers and transmitters for RF and microwave," *IEEE Transactions on Microwave Theory and Techniques*, vol. 50, no. 3, pp. 814–826, 2002.

[20] D. C. Cox, "Linear amplification with nonlinear components," *IEEE Transactions on Communications*, pp. 1942–1945, Dec. 1974.

[21] A. D. Pham, "Outphase Power Amplifiers in OFDM Systems," Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge, MA, 2005.

[22] H. Xu, Y. Palaskas, A. Ravi, M. Sajadieh, M. El-Tanani, and K. Soumyanath, "A Flip-Chip-Packaged 25.3 dBm Class-D Outphasing Power Amplifier in 32 nm CMOS for WLAN Application," *IEEE Journal of Solid-State Circuits*, vol. 46, no. 7, pp. 1596–1605, 2011.

[23] L. R. Kahn, "Single-sideband transmission by envelope elimination and restoration," *Proc. of the IRE*, vol. 40, no. 7, pp. 803–806, Jul. 1952.

[24] P. A. Godoy, "Techniques for High-Efficiency Outphasing Power Amplifiers," Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge, MA, 2011.

[25] P. Godoy, S. Chung, T. Barton, D. Perreault, and J. Dawson, "A 2.4-GHz, 27-dBm Asymmetric Multilevel Outphasing Power Amplifier in 65-nm CMOS," *IEEE Journal of Solid-State Circuits*, vol. 47, no. 10, pp. 2372–2384, 2012.

[26] S.-M. Yoo, J. Walling, E.-C. Woo, B. Jann, and D. Allstot, "A Switched-Capacitor RF Power Amplifier," *IEEE Journal of Solid-State Circuits*, vol. 46, no. 12, pp. 2977–2987, 2011.

[27] M. Ingels, V. Giannini, J. Borremans, G. Mandal, B. Debaillie, P. Van Wesemael, T. Sano, T. Yamamoto, D. Hauspie, J. Van Driessche, and J. Craninckx, "A 5 mm^2 40 nm LP CMOS Transceiver for a Software-Defined Radio Platform," *IEEE Journal of Solid-State Circuits*, vol. 45, no. 12, pp. 2794–2806, 2010.

[28] V. Giannini, M. Ingels, T. Sano, B. Debaillie, J. Borremans, and J. Craninckx, "A multiband LTE SAW-less modulator with -160dBc/Hz RX-band noise in 40nm LP CMOS," in *IEEE International Solid-State Circuits Conference*, 2011, pp. 374–376.

[29] J. Borremans, G. Mandal, V. Giannini, B. Debaillie, M. Ingels, T. Sano, B. Verbruggen, and J. Craninckx, "A 40nm CMOS 0.4-6GHz Receiver Resilient to Out-of-Band Blockers," *IEEE Journal of Solid-State Circuits*, vol. 46, no. 7, pp. 1659–1671, 2011.

[30] D. Murphy, H. Darabi, A. Abidi, A. Hafez, A. Mirzaei, M. Mikhemar, and M.-C. Chang, "A Blocker-Tolerant, Noise-Cancelling Receiver Suitable for Wideband Wireless Applications," *IEEE Journal of Solid-State Circuits*, vol. 47, no. 12, pp. 2943–2963, 2012.

[31] K. Kanda, Y. Kawano, T. Sasaki, N. Shirai, T. Tamura, S. Kawai, M. Kudo, T. Murakami, H. Nakamoto, N. Hasegawa, H. Kano, N. Shimazui, A. Mineyama, K. Oishi, M. Shima, N. Tamura, T. Suzuki, T. Mori, K. Niratsuka, and S. Yamaura, "A fully integrated triple-band CMOS power amplifier for WCDMA mobile handsets," in *IEEE International Solid-State Circuits Conference*, 2012, pp. 86–88.

[32] A. Afsahi, A. Behzad, V. Magoon, and L. Larson, "Linearized Dual-Band Power Amplifiers With Integrated Baluns in 65 nm CMOS for a 2×2 802.11n MIMO WLAN SoC," *IEEE Journal of Solid-State Circuits*, vol. 45, no. 5, pp. 955–966, 2010.

[33] H.-H. Liao, H. Jiang, P. Shanjani, J. King, and A. Behzad, "A Fully Integrated 2×2 Power Amplifier for Dual Band MIMO 802.11n WLAN Application Using SiGe HBT Technology," *IEEE Journal of Solid-State Circuits*, vol. 44, no. 5, pp. 1361–1371, 2009.

[34] R. Kumar, T. Krishnaswamy, G. Rajendran, D. Sahu, A. Sivadas, M. Nandigam, S. Ganeshan, S. Datla, A. Kudari, H. Bhasin, M. Agrawal, S. Narayan, Y. Dharwekar, R. Garg, V. Edayath, T. Suseela, V. Jayaram, S. Ram, V. Murugan, A. Kumar, S. Mukherjee, N. Dixit, E. Nussbaum, J. Dror, N. Ginzburg, A. EvenChen, A. Maruani, S. Sankaran, V. Srinivasan, and V. Rentala, "A fully integrated 2×2 b/g and 1×2 a-band MIMO WLAN SoC in 45nm CMOS for multi-radio IC," in *IEEE International Solid-State Circuits Conference*, 2013, pp. 328–329.

[35] C. Lee, A. Behzad, B. Marholev, V. Magoon, I. Bhatti, D. Li, S. Bothra, A. Afsahi, D. Ojo, R. Roufoogaran, T. Li, Y. Chang, K. Rao, S. Au, P. Seetharam, K. Carter, J. Rael, M. MacIntosh, B. Lee, M. Rofougaran, R. Rofougaran, A. Hadji-Abdolhamid, M. Nariman, S. Khorram, S. Anand, E. Chien, S. Wu, C. Barrett, L. Zhang, A. Zolfaghari, H. Darabi, A. Sarfaraz, B. Ibrahim, M. Gonikberg, M. Forbes, C. Fraser, L. Gutierrez, Y. Gonikberg, M. Hafizi, S. Mak, J. Castaneda, K. Kim, Z. Liu, S. Bouras, K. Chien, V. Chandrasekhar, P. Chang, E. Li, and Z. Zhao, "A multistandard, multiband SoC with integrated BT, FM, WLAN radios and integrated power amplifier," in *IEEE International Solid-State Circuits Conference*, 2010, pp. 454–455.

[36] D. Waters, "Connected cars promise safer roads," *BBC News*, vol. 2007, 2007.

[37] J. Markoff, "Google cars drive themselves, in traffic," *The New York Times*, vol. 10, p. A1, 2010.

[38] "IEEE Std 802.11p-2010 - Amendment to IEEE Std 802.11-2007," pp. 1–51, 2010.

[39] R. Smith, S. Sheppard, Y.-F. Wu, S. Heikman, S. Wood, W. Pribble, and J. Milligan, "AlGaN/GaN-on-SiC HEMT Technology Status," in *IEEE Compound Semiconductor Integrated Circuit Symposium*, 2008, pp. 1–4.

[40] W. Ciccognani, M. De Dominicis, M. Ferrari, E. Limiti, M. Peroni, and P. Romanini, "High-power monolithic AlGaN/GaN HEMT switch for X-band applications," *Electronics Letters*, vol. 44, no. 15, pp. 911–912, 2008.

[41] T. W. Kim, B. Kim, and K. Lee, "Highly linear receiver front-end adopting MOSFET transconductance linearization by multiple gated transistors," *IEEE Journal of Solid-State Circuits*, vol. 39, no. 1, pp. 223–229, 2004.

[42] I. Angelov, H. Zirath, and N. Rosman, "A new empirical nonlinear model for HEMT and MESFET devices," *IEEE Transactions on Microwave Theory and Techniques*, vol. 40, no. 12, pp. 2258–2266, 1992.

[43] C. Tinella, J. Fournier, D. Belot, and V. Knopik, "A high-performance CMOS-SOI antenna switch for the 2.5-5-GHz band," *IEEE Journal of Solid-State Circuits*, vol. 38, no. 7, pp. 1279–1283, 2003.

[44] D. Shaeffer and T. Lee, "A 1.5-V, 1.5-GHz CMOS low noise amplifier," *IEEE Journal of Solid-State Circuits*, vol. 32, no. 5, pp. 745–759, 1997.

[45] M. Fathi, D. Su, and B. Wooley, "A 30.3dBm 1.9GHz-bandwidth 2×4-array stacked 5.3GHz CMOS power amplifier," in *IEEE International Solid-State Circuits Conference*, 2013, pp. 88–89.

[46] M. Watanabe, R. Snyder, and T. LaRocca, "Simultaneous linearity and efficiency enhancement of a digitally-assisted GaN power amplifier for 64-QAM," in *IEEE RFIC Symposium*, 2013, pp. 427–430.

[47] *3GPP TS 36.101 V11.1.0, E-UTRA user equipment transmission and reception.*

[48] A. Scuderi, C. Presti, F. Carrara, B. Rauber, and G. Palmisano, "A stage-bypass SOI-CMOS switch for multi-mode multi-band applications," in *IEEE RFIC Symposium*, 2008, pp. 325–328.

[49] H. Xu, Y. Palaskas, A. Ravi, M. Sajadieh, M. Elmala, and K. Soumyanath, "A 28.1dBm class-D outphasing power amplifier in 45nm LP digital CMOS," in *Symposium on VLSI Circuits*, 2009, pp. 206–207.

[50] B. Serneels, T. Piessens, M. Steyaert, and W. Dehaene, "A high-voltage output driver in a 2.5-V 0.25-μm CMOS technology," *IEEE Journal of Solid-State Circuits*, vol. 40, no. 3, pp. 576–583, 2005.

[51] T.-P. Hung, D. Choi, L. Larson, and P. Asbeck, "CMOS Outphasing Class-D Amplifier With Chireix Combiner," *IEEE Microwave and Wireless Components Letters*, vol. 17, no. 8, pp. 619–621, 2007.

[52] S.-M. Yoo, J. Walling, E. C. Woo, and D. Allstot, "A switched-capacitor power amplifier for EER/polar transmitters," in *IEEE International Solid-State Circuits Conference*, 2011, pp. 428–430.

[53] W. Tai, H. Xu, A. Ravi, H. Lakdawala, O. Bochobza-Degani, L. Carley, and Y. Palaskas, "A Transformer-Combined 31.5 dBm Outphasing Power Amplifier in 45 nm LP CMOS With Dynamic Power Control for Back-Off Power Efficiency Enhancement," *IEEE Journal of Solid-State Circuits*, vol. 47, no. 7, pp. 1646–1658, 2012.

[54] A. Ravi, P. Madoglio, H. Xu, K. Chandrashekar, M. Verhelst, S. Pellerano, L. Cuellar, M. Aguirre-Hernandez, M. Sajadieh, J. Zarate-Roldan, O. Bochobza-Degani, H. Lakdawala, and Y. Palaskas, "A 2.4-GHz 20-40-MHz Channel WLAN Digital Outphasing Transmitter Utilizing a Delay-Based Wideband Phase Modulator in 32-nm CMOS," *IEEE Journal of Solid-State Circuits*, vol. 47, no. 12, pp. 3184–3196, 2012.

[55] P. Madoglio, A. Ravi, H. Xu, K. Chandrashekar, M. Verhelst, S. Pellerano, L. Cuellar, M. Aguirre, M. Sajadieh, O. Degani, H. Lakdawala, and Y. Palaskas, "A 20dBm 2.4GHz digital outphasing transmitter for WLAN application in 32nm CMOS," in *IEEE International Solid-State Circuits Conference*, 2012, pp. 168–170.

[56] M. Heidari, M. Lee, and A. Abidi, "All-Digital Outphasing Modulator for a Software-Defined Transmitter," *IEEE Journal of Solid-State Circuits*, vol. 44, no. 4, pp. 1260–1271, 2009.

[57] P. Haldi, D. Chowdhury, P. Reynaert, G. Liu, and A. Niknejad, "A 5.8 GHz 1 V Linear Power Amplifier Using a Novel On-Chip Transformer Power Combiner in Standard 90 nm CMOS," *IEEE Journal of Solid-State Circuits*, vol. 43, no. 5, pp. 1054–1063, 2008.

[58] S. Woo, W. Kim, C.-H. Lee, K. Lim, and J. Laskar, "A 3.6mW differential common-gate CMOS LNA with positive-negative feedback," in *IEEE International Solid-State Circuits Conference*, 2009, pp. 218–219,219a.

[59] J. Borremans, P. Wambacq, C. Soens, Y. Rolain, and M. Kuijk, "Low-Area Active-Feedback Low-Noise Amplifier Design in Scaled Digital CMOS," *IEEE Journal of Solid-State Circuits*, vol. 43, no. 11, pp. 2422–2433, 2008.

[60] F. Bruccoleri, E. A. M. Klumperink, and B. Nauta, "Wide-band CMOS low-noise amplifier exploiting thermal noise canceling," *IEEE Journal of Solid-State Circuits*, vol. 39, no. 2, pp. 275–282, 2004.

[61] C.-T. Fu, H. Lakdawala, S. Taylor, and K. Soumyanath, "A 2.5GHz 32nm $0.35mm^2$ 3.5dB NF -5dBm P_{1dB} fully differential CMOS push-pull LNA with integrated 34dBm T/R switch and ESD protection," in *IEEE International Solid-State Circuits Conference*, 2011, pp. 56–58.

[62] G. Liu, P. Haldi, T.-J. K. Liu, and A. Niknejad, "Fully Integrated CMOS Power Amplifier With Efficiency Enhancement at Power Back-Off," *IEEE Journal of Solid-State Circuits*, vol. 43, no. 3, pp. 600–609, 2008.

[63] D. Chowdhury, S. Thyagarajan, L. Ye, E. Alon, and A. Niknejad, "A Fully-Integrated Efficient CMOS Inverse Class-D Power Amplifier for Digital Polar Transmitters," *IEEE Journal of Solid-State Circuits*, vol. 47, no. 5, pp. 1113–1122, 2012.

[64] T. H. Lee, *Planar Microwave Engineering*. Cambridge University Press, 2004.

[65] D. Frickey, "Conversions between S, Z, Y, H, ABCD, and T parameters which are valid for complex source and load impedances," *IEEE Transactions on Microwave Theory and Techniques*, vol. 42, no. 2, pp. 205–211, 1994.

[66] J. Long, "Monolithic transformers for silicon RF IC design," *IEEE Journal of Solid-State Circuits*, vol. 35, no. 9, pp. 1368–1382, 2000.

[67] I. M. Kang, S.-J. Jung, T.-H. Choi, J.-H. Jung, C. Chung, H.-S. Kim, K. Park, H. Oh, H.-W. Lee, G. Jo, Y.-K. Kim, H.-G. Kim, and K.-M. Choi, "RF Model of BEOL Vertical Natural Capacitor (VNCAP) Fabricated by 45-nm RF CMOS Technology and Its Verification," *IEEE Electron Device Letters*, vol. 30, no. 5, pp. 538–540, 2009.

[68] J. Borremans, S. Thijs, P. Wambacq, Y. Rolain, D. Linten, and M. Kuijk, "A Fully Integrated 7.3 kV HBM ESD-Protected Transformer-Based 4.5-6 GHz CMOS LNA," *IEEE Journal of Solid-State Circuits*, vol. 44, no. 2, pp. 344–353, 2009.

[69] D. Linten, S. Thijs, M. Natarajan, P. Wambacq, W. Jeamsaksiri, J. Ramos, A. Mercha, S. Jenei, S. Donnay, and S. Decoutere, "A 5-GHz fully integrated ESD-protected low-noise amplifier in 90-nm RF CMOS," *IEEE Journal of Solid-State Circuits*, vol. 40, no. 7, pp. 1434–1442, 2005.

[70] Y. Tan, H. Xu, M. El-Tanani, S. Taylor, and H. Lakdawala, "A flip-chip-packaged 1.8V 28dBm class-AB power amplifier with shielded concentric transformers in 32nm SoC CMOS," in *IEEE International Solid-State Circuits Conference*, 2011, pp. 426–428.

[71] S. Kousai and A. Hajimiri, "An octave-range watt-level fully integrated CMOS switching power mixer array for linearization and back-off efficiency improvement," in *IEEE International Solid-State Circuits Conference*, 2009, pp. 376–377,377a.

[72] E. Kaymaksut and P. Reynaert, "Transformer-Based Uneven Doherty Power Amplifier in 90 nm CMOS for WLAN Applications," *IEEE Journal of Solid-State Circuits*, vol. 47, no. 7, pp. 1659–1671, 2012.

[73] B. Francois and P. Reynaert, "A Fully Integrated Watt-Level Linear 900-MHz CMOS RF Power Amplifier for LTE-Applications," *IEEE Transactions on Microwave Theory and Techniques*, vol. 60, no. 6, pp. 1878–1885, 2012.

[74] S. Pornpromlikit, J. Jeong, C. Presti, A. Scuderi, and P. Asbeck, "A Watt-Level Stacked-FET Linear Power Amplifier in Silicon-on-Insulator CMOS," *IEEE Transactions on Microwave Theory and Techniques*, vol. 58, no. 1, pp. 57–64, 2010.

[75] B. Koo, T. Joo, Y. Na, and S. Hong, "A fully integrated dual-mode CMOS power amplifier for WCDMA applications," in *IEEE International Solid-State Circuits Conference*, 2012, pp. 82–84.

[76] J. Walling, H. Lakdawala, Y. Palaskas, A. Ravi, O. Degani, K. Soumyanath, and D. Allstot, "A 28.6dBm 65nm Class-E PA with Envelope Restoration by Pulse-Width and Pulse-Position Modulation," in *IEEE International Solid-State Circuits Conference*, 2008, pp. 566–636.

[77] J. Sonsky, A. Heringa, J. Perez-Gonzalez, J. Benson, P. Y. Chiang, S. Bardy, and I. Volokhine, "Innovative High Voltage transistors for complex HV/RF SoCs in baseline CMOS," in *Symp. VLSI Technology Dig. Tech. Papers*, 2008, pp. 115–116.

[78] Y. Tan, J. Duster, C.-T. Fu, E. Alpman, A. Balankutty, C. Lee, A. Ravi, S. Pellerano, K. Chandrashekar, H. Kim, B. Carlton, S. Suzuki, M. Shafi, Y. Palaskas, and H. Lakdawala, "A 2.4GHz WLAN transceiver with fully-integrated highly-linear 1.8V 28.4dBm PA, 34dBm T/R switch, 240MS/s DAC, 320MS/s ADC, and DPLL in 32nm SoC CMOS," in *Symposium on VLSI Circuits*, 2012, pp. 76–77.

[79] D. S. Lee, O. Laboutin, Y. Cao, W. Johnson, E. Beam, A. Ketterson, M. Schuette, P. Saunier, and T. Palacios, "Impact of Al_2O_3 Passivation Thickness in Highly Scaled GaN HEMTs," *IEEE Electron Device Letters*, vol. 33, no. 7, pp. 976–978, 2012.

[80] K. Shinohara, D. Regan, A. Corrion, D. Brown, S. Burnham, P. Willadsen, I. Alvarado-Rodriguez, M. Cunningham, C. Butler, A. Schmitz, S. Kim, B. Holden, D. Chang, V. Lee, A. Ohoka, P. Asbeck, and M. Micovic, "Deeply-scaled self-aligned-gate GaN DH-HEMTs with ultrahigh cutoff frequency," in *IEEE International Electron Devices Meeting*, 2011, pp. 19.1.1–19.1.4.

[81] B. Lu, O. Saadat, and T. Palacios, "High-Performance Integrated Dual-Gate AlGaN/GaN Enhancement-Mode Transistor," *IEEE Electron Device Letters*, vol. 31, no. 9, pp. 990–992, 2010.

[82] H.-S. Lee, K. Ryu, M. Sun, and T. Palacios, "Wafer-Level Heterogeneous Integration of GaN HEMTs and Si (100) MOSFETs," *IEEE Electron Device Letters*, vol. 33, no. 2, pp. 200–202, 2012.

[83] B. Razavi, "RF transmitter architectures and circuits," in *IEEE Custom Integrated Circuits Conference*, 1999, pp. 197–204.

[84] E. Brown, "RF-MEMS switches for reconfigurable integrated circuits," *IEEE Transactions on Microwave Theory and Techniques*, vol. 46, no. 11, pp. 1868–1880, 1998.

[85] C.-H. Chen and D. Peroulis, "Liquid RF MEMS Wideband Reflective and Absorptive Switches," *IEEE Transactions on Microwave Theory and Techniques*, vol. 55, no. 12, pp. 2919–2929, 2007.