RF Transmitter Design for Large Antenna Array Applications

Pengpeng Lu

Electrical Engineering and Computer Sciences
University of California at Berkeley

December 1, 2020

RF Transmitter Design for Large Antenna Array Applications

by

Pengpeng Lu

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Engineering - Electrical Engineering and Computer Science

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Elad Alon, Chair
Professor David Aldous
Professor Ali Niknejad

Fall 2018

RF Transmitter Design for Large Antenna Array Applications

Abstract

RF Transmitter Design for Large Antenna Array Applications

by

Pengpeng Lu

Doctor of Philosophy in Engineering - Electrical Engineering and Computer Science

University of California, Berkeley

Professor Elad Alon, Chair

Recent advances in wireless technologies have enabled fast increase of mobile data traffic. In fact, mobile data traffic has grown 18-fold over the past 5 years. The rapid growing demand imposes a big near-far problem on our wireless network. Beamforming technique provides an opportunity to efficiently solve the near-far problem, by combining elements in an antenna array in such a way that signals at particular angles experience constructive interference while others experience destructive interference.

This thesis focuses on the design of energy-efficient RF transmitters for large antenna array applications. In order to be energy-efficient, we should minimize the overhead power consumption. It makes the design challenging while minimizing out-of-band emissions and supporting multiple/reconfigurable bands as well as programmability in terms of performance (i.e., output power, noise figure, resolution, bandwidth, etc.), since any fixed frequency band-pass filters should be eliminated.

To address these challenges, a mixer-last TX architecture is proposed, which uses current DACs to deliver charge directly to the 50 Ohm output RF load. Out-of-band emissions are reduced by digital oversampling, and the charging operation of a baseband capacitor.

Using the proposed transmitter architectures, two chips were implemented. The first chip was fabricated in TSMC's 65nm CMOS technology. With a peak output power of 5.1dBm, the first chip consumes 49.2 mW and the measured noise floor at 40MHz offset is -155dBc/Hz.

Realizing it is impossible to build a single transmitter that is power efficient under all applications, we developed a generator for our proposed TX architecture using Berkeley Analog Generator (BAG) framework [1]. A generator captures the design methodology and is process portable. It could significantly lower the design cost so that different transmitters that are efficient for different design specifications can be generated.The second chip was generated by the generator and was fabricated in TSMC's 16nm CMOS technology. It consumes 5.14 mW at peak output power of -19.6 dBm. Compared with the first 65nm prototype, it is more power efficient in the output power range of less than -4 dBm.

To My Family

Contents

List of Figures

List of Tables

Acknowledgments

The past 6 years have been very challenging and stressful for me. Throughout this wild journey, I feel extremely fortunate to have received so much help and support from so many individuals, without whom I can never reach this finish line. I would like to take this opportunity to express my most sincere gratitude to every one of them.

First, I would like to express my sincere gratitude to my advisor, Professor Elad Alon, not only for his guidance in my research, but also for his understanding, patience, encouragement and continuous support over the years. After knowing him for more than 6 years, I am still impressed by how passionate he is about his work, and how diligently he works every day. There is no better advisor a student could hope to find than Elad.

Besides, I would like to thank Prof. Ali Niknejad, Prof. David Aldous, and Prof. Borivoje Nikolic for being on my thesis and qualification exam committee and for their valuable comments and encouragement.

My sincere thanks also goes to all BWRC staff, especially Candy Corpus, who made my experience at BWRC as smooth and stress-free as possible, and James Dunn and Brian Richards, who helped me wrestle with CAD tools and computers.

FADER is a huge project and I'm grateful to have worked with many talented friends on this project (in no particular order): Lingkai Kong, D. J. Seo, Antonio Puglielli, Amy Whitcombe, Eric Chang, Greg Lacaille, Nathan Narevsky, Zhongkai Wang, Kosta Trotskovsky, Marko Kosunen.

I would also like to thank my fellow BWRC colleagues (some of them have graduated) - Lucas Calderin, Jiashu Chen, Pi-Feng Chiu, Sijun Du, Seobin Jung, Nai-Chung Kuo, Benyuanyi Liu, Zhaokai Liu, Yue Lu, Alberto Puggelli, Sameet Ramakrishnan, Nicholas Sutardja, Angie Wang, Ruochen Wang, Meng Wei, Bonjern Yang, Luya Zhang, Qichen Zhang, Bo Zhao - for working with me on various projects, and having fun together.

Last but not the least, I would like to thank my parents, for their unconditional love and continuous support throughout my life. To my best friend Ye Tian, for being there through all the ups and downs. To my husband Guanyu, for taking care of me, and sharing all my laughs and tears. I can never get through my Ph.D journey without their love and support.

Chapter 1

Introduction

1.1 Data Traffic Challenge

In the past few years, the world has witnessed huge and fast increase in data traffic. In fact, mobile data traffic has grown 18-fold over the past 5 years, from 400 petabytes (1 petabyte $= 10^{15}$ bytes) per month in 2011, to 7.2 exabytes (1 exabyte $= 10^{18}$ bytes) per month at the end of 2016. Cisco Visual Networking Index predicts that by 2021, monthly global mobile data traffic will be 49 exabytes [2] (Fig 1.1).
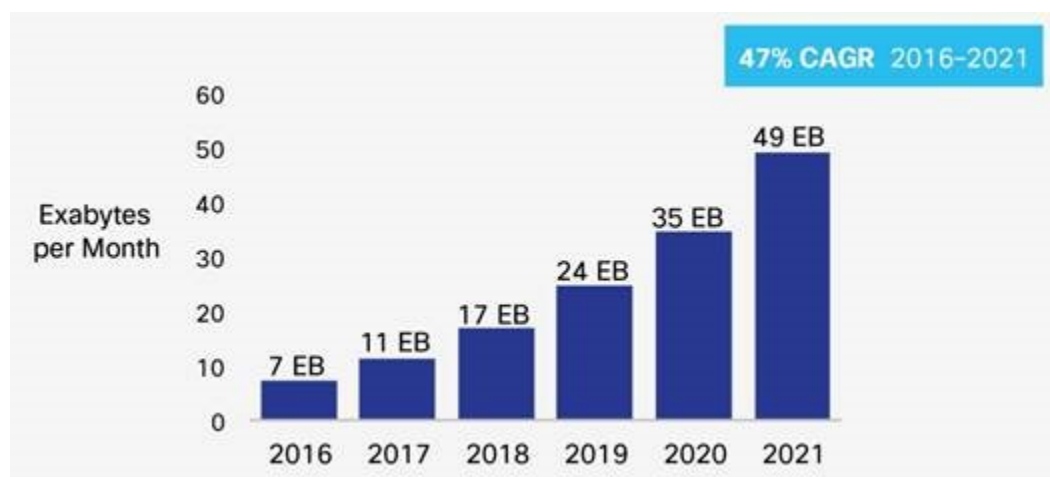
Figure 1.1: Cisco Forecasts 49 Exabytes per Month of Mobile Data Traffic by 2021. Adapted [reprinted] from "Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2016 —2021 White Paper", Cisco, 2017
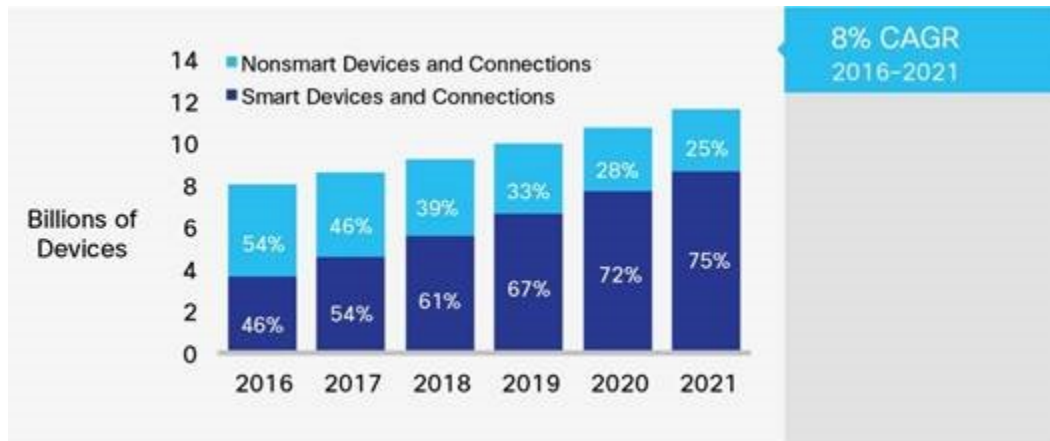
Figure 1.2: Global Growth of Smart Mobile Devices and Connections. Adapted [reprinted] from "Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2016 —2021 White Paper", Cisco, 2017

This trend is due to both the big increase in the data traffic per device and the number of devices worldwide. Both are expected to grow even faster in the next 5 years (Fig 1.2).

The rapid growing demand imposes a big near-far problem on our wireless network: Consider a receiver and two transmitters, one close to the receiver, the other far away. If both transmitters transmit simultaneously and at equal powers, then due to the inverse square law the receiver will receive more power from the nearer transmitter (Fig 1.3) [3] . Since one transmission's signal is the other's interferer, the signal-to-interference-plus-noise ratio (SINR) for the farther transmitter is much lower. This makes the farther transmitter more difficult, if not impossible, to understand.
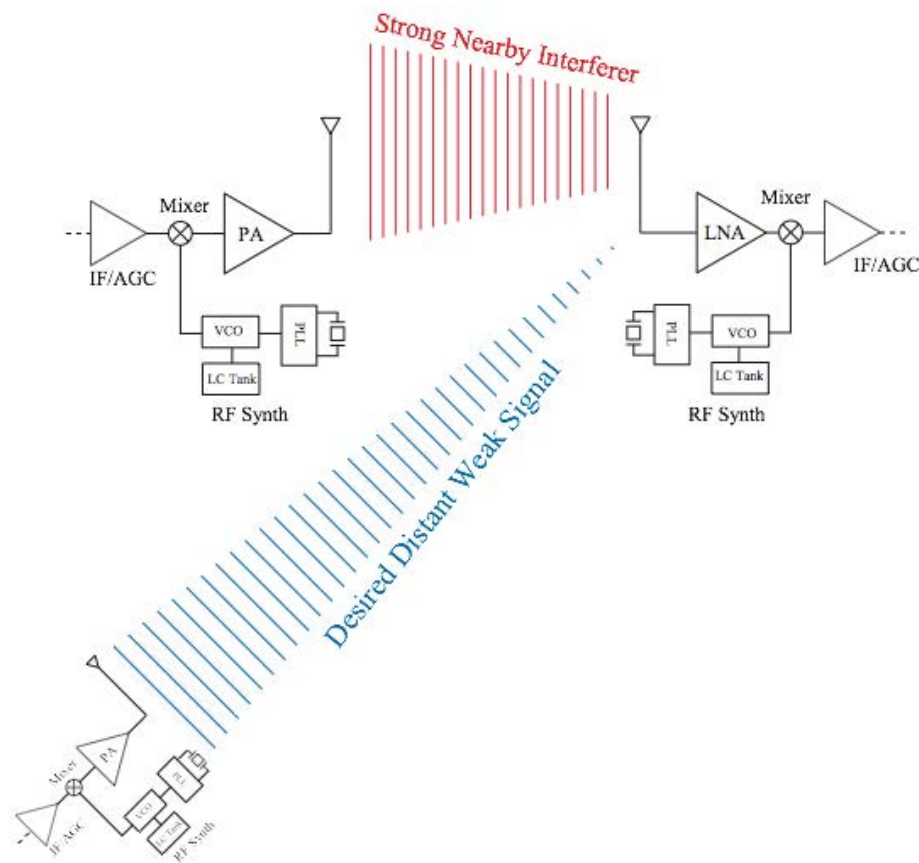
Figure 1.3: Near Far Problem. Adapted [reprinted] from "The xG Vision: Making the Internet truly wireless", Ali Niknejad, 2014

1.2 Beamforming

The near-far problem could be greatly relaxed if the transmitting and receiving are directional [4]. Beamforming is a technique for directional signal transmission or reception. This is achieved by combining elements in an antenna array in such a way that signals at particular angles experience constructive interference while others experience destructive interference [5–9].

To change the directionality of the array when transmitting, a beamformer controls the phase and relative amplitude of the signal at each transmitter. The simplest way is to put multiple antennas in a line and apply phase shifting only at each transmitter. Fig 1.4 illustrates the operating principle of a simple two-antenna phased array [10]. For simplicity, assume

each antenna transmits its own signal through an omnidirectional antenna. In an arbitrary direction θ in space, due to the physical distance between these two antennas, a phase difference between signals from these two elements occurs. Far away from the antennas, the phase difference can be approximated as:

$$\phi = \frac{2\pi d}{\lambda}cos\theta \tag{1.1}$$

Here d and θ denote the distance between antennas and the wavelength, respectively.
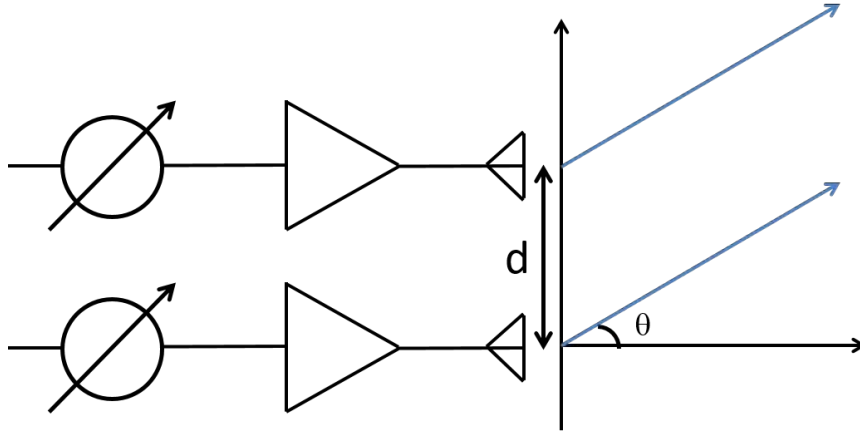
Figure 1.4: Operating Principle of a phased array transmitter

By phase shifting the signal from the first TX by ϕ relative to the second TX , the signals in the direction θ in space can be realigned, and add in voltage. It is worth noting that at the same time, in the direction θ' that satisfies

$$\frac{2\pi d}{\lambda}cos\theta' = \pi - \frac{2\pi d}{\lambda}cos\theta, \tag{1.2}$$

signals from these two TXs are out of phase. Therefore, the energy delivered in this direction is zero. Because energy is proportional to the square of voltage amplitude, the energy delivered in this direction is 4 times larger than a single TX. Generally, in a system with N elements, the transmitted energy is increased by N^2 in the desired direction. This enables the system to boost the energy in the direction of target users, and reduce interference of other nearby systems.

The spacing between two adjacent antennas in the array is typically $\lambda/2$ [11] to perform Nyquist sampling in space. For GHz applications, this is several centimeters (15cm@1GHz, 6.25cm@2.4GHz, 3cm@5GHz), which implies large array area. Luckily, we have enough "surface" areas around us - walls, tables, ceilings, etc. We could leverage the "surfaces" to build large antenna arrays. This concept leads us to our "xG" vision in BWRC.

1.3   xG Vision

The industry is expecting 5G networks by 2020, which is positioned for a future with "everything in the cloud"—for example, immersive video conferencing, internet of things, low latency and high reliability machine-centric communication, and so on [12, 13].

Beyond 5G, researchers at BWRC (Berkeley Wireless Research Center) proposed the "xG" vision (Fig 1.5) [14–16]. In our "xG" vision, we will build a network that can grow organically with little to no user intervention or configuration. Access points (hubs) will be spreaded everywhere ,and allowed to talk to each other to form a mesh network. In order to reduce interference, antenna arrays (on "surfaces" around us) should be used to form beams and create point-to-point links. Small handsets do not need to spread energy out into a sphere of 1km radius, but simply a few meters, to the nearest "eWallpaper".
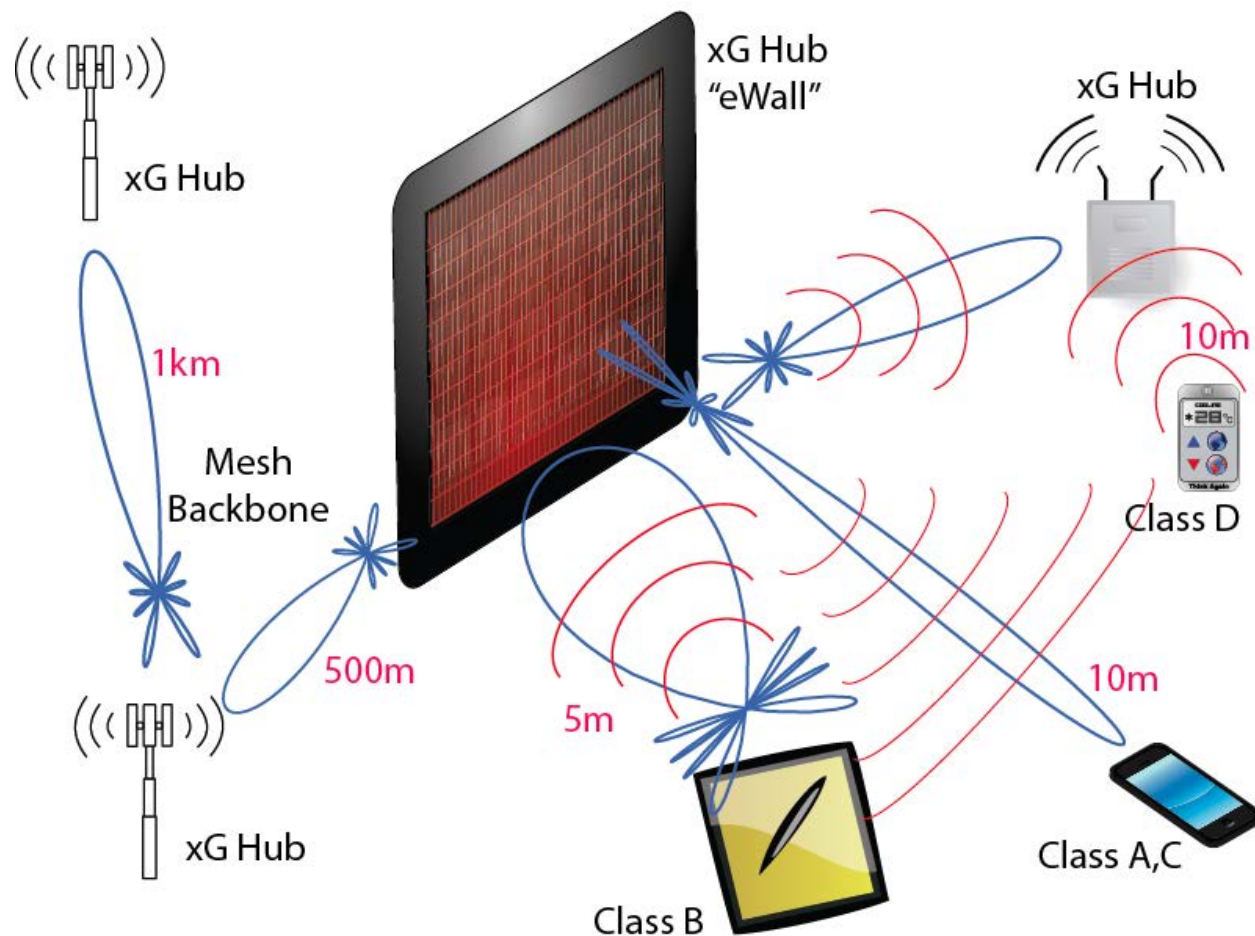
Figure 1.5: xG Vision. Adapted [reprinted] from "The xG Vision: Making the Internet truly wireless", Ali Niknejad, 2014

1.4   Thesis Organization

With the introduction of beamforming, and the concept of"xG" vision and eWallpaper in this chapter, the remaining focus of the thesis will then be the presentation of circuit design techniques to build an RF transmitter for multi-antenna applications. We first introduce the whole system –eWallpaper, and discuss high-level system design concerns in Chapter 2. In Chapter 3, we begin with a comparison of system power consumption of single-antenna systems and multi-antenna systems. Overhead power is identified as the bottleneck of the system efficiency. Then a transmitter architecture that minimizes overhead power consumption is proposed. Using the proposed transmitter architectures, two chips were implemented. The first chip was implemented with traditional CAD tools in TSMC 65nm CMOS technology. Its implementation details and measurement results are shown in Chapter 4. Realizing

the fact that it is impossible to build a single TX that is power efficient for all design specifications, we wrote a generator with BAG, a process-portable circuit generator framework [1] developed in BWRC. In Chapter 5, the development of the generator is discussed. The second chip was generated by the generator and fabricated in TSMC 16nm CMOS technology. Chapter 6 concludes the entire thesis, and discusses about future directions.

Chapter 2

eWallpaper

2.1 eWallpaper System

The core concept of our "xG" vision is to utilize the "surfaces" around us for large antenna arrays – eWallpaper. To fulfill what we envisioned, it should be an array of a large number of interconnected and controlled beam forming common module ASICs - with each of these ASICs including programmable RF front-ends, data-conversion, local digital signal processing, high-speed links and clock generation - assembled onto a flexible substrate containing printed antennas and other passive elements (Fig 2.1, 2.2).
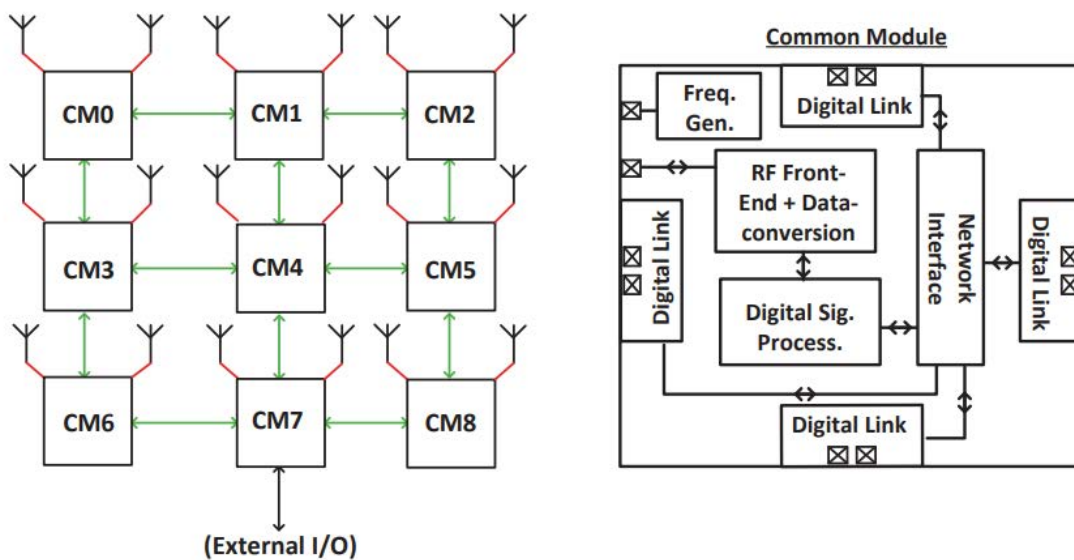
Figure 2.1: eWallpaper Block Diagram
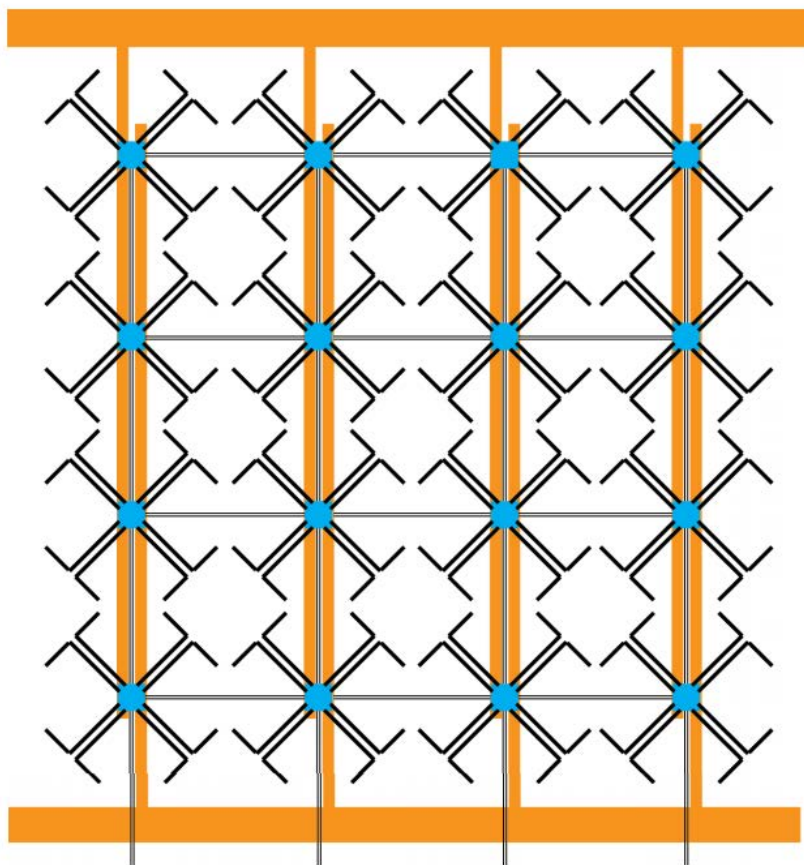
Figure 2.2: eWallpaper Vision

2.2 Choice of Number of Elements in eWallpaper

In Fig 2.1, for simplicity, it was shown that each ASIC supports two antenna. In reality, each ASIC could support different number of antennas. Therefore, it is necessary for us to understand the trade-offs and find the optimal number of antennas per ASIC first.

As shown in Fig 2.1, each ASIC consists of data-conversion, clock generation, local digital signal processing circuits, and digital high-speed links to enable it to talk to its neighbors. The advantage of having multiple antennas on one chip is that they could share part of the clock generation and local DSP circuits, and the number of serial links in the whole system is reduced, thus simplifying the system, and could possibly reduce the system power consumption. On the other hand, since the spacing between antennas is fixed ($\lambda/2$), the more antennas one chip supports, the further the routing distance from the chip to the antenna. The transmitter on the chip needs to output more power to compensate the higher

routing loss, which might increase the system power consumption. The physical routing could become very messy, if not impossible, when one chip supports a lot of antennas.

Consider a 2D array of N ASICs, each ASIC supports A antennas. From the above discussion, it is clear that the power consumption of all ASICs in the system can be grouped into 3 categories: 1) Local power consumption that doesn't change with the number of antennas per chip; 2) Power used to communicate with other ASICs; 3) Power consumed by RF front-end. We will next establish mathematical equations for the 3 categories respectively.

Local power

Local power consumption includes local DSP, clock generation circuits, etc. It doesn't change with the number of antennas per chip. Assume each ASIC consumes $P_{ov,ASIC}$ in its local DSP and clock generation circuits, then the whole antenna array system consumes $\frac{N}{A}P_{ov,ASIC}$. ($\frac{N}{A}$ is the number of ASICs in the array.)

High-speed links

To simplify our analysis, we assume each ASIC uses 4 transceivers to talk to all its nearest neighbors, and each transceiver costs P_{link}, the whole antenna array system consumes $4\frac{N}{A}P_{link}$. It is worth noting that in reality, not all ASICs need to use all 4 transceivers to talk to all its nearest neighbors.

RF front-end

Apparently, the total power consumption of RF front-ends is $NAP_{DC,TX}$, where $P_{DC,TX}$ is the DC power consumed by one RF transmitter. $P_{DC,TX}$ can be divided into two parts: 1) Overhead part, which grows linearly with the number of antennas, but doesn't change with output power. Overhead power ($P_{ov,ant}$) includes modulation, data conversion, phase shifting, etc.; 2) Power amplifier part, which changes with output power level. Power efficiency of a power amplifier is defined as:

$$\eta = \frac{P_{out,RF}}{P_{DC,PA}} \tag{2.1}$$

Then we could write $P_{DC,TX}$ in terms of RF output power and PA efficiency:

$$P_{DC,TX} = P_{ov,ant} + \frac{P_{out,RF}}{\eta} \tag{2.2}$$

Similar to cable losses, we could characterize the routing loss from the ASIC to the antenna in terms of decibels per unit length. Here we define the antenna loss to be k dB per half-wavelength, which is equivalent as:

$$P_{radiated} = P_{out,RF} \cdot 10^{-\frac{2kl}{\lambda}} \tag{2.3}$$

where l is the routing distance. Then

$$P_{DC,TX} = P_{ov,ant} + P_{radiated} \cdot 10^{\frac{2kl}{\lambda}} \cdot \frac{1}{\eta} \tag{2.4}$$

l increases with number of antennas each ASIC supports (A). We could approximate the l as $\frac{\sqrt{2}}{4}\lambda\sqrt{A}$.

As discussed in section 1.2, in the direction the antenna array points, the equivalent isotropically radiated power ($EIRP$) is N^2 times the radiated energy of one antenna. The radiated power of one antenna is:

$$P_{radiated} = \frac{EIRP}{(N)^2} \tag{2.5}$$

In summary, the total power consumption of RF front-end is:

$$P_{DC,TX} = P_{ov,ant} + \frac{EIRP}{(N)^2} \cdot 10^{\frac{2kl}{\lambda}} \cdot \frac{1}{\eta} \tag{2.6}$$

The antenna array's power consumption during RF transmission is:

$$P_{tot} = \frac{N}{A}P_{ov,ASIC} + 4\frac{N}{A}P_{link} + NP_{ov,ant} + \frac{EIRP}{N} \cdot 10^{\frac{2kl}{\lambda}} \cdot \frac{1}{\eta} \tag{2.7}$$

It is worth noting that in reality, $P_{ov,ASIC}$, $P_{ov,ant}$, P_{link} and η all depend on $EIRP$, N, A, and the chosen circuit architectures. However, we assumed them to be constant to get an insight into how we choose N and A in the first step. By taking partial derivatives with respect to A, the optimal number of antennas per ASIC (A) when number of antennas in the entire array (N) is fixed is given by the following equation:

$$\frac{\sqrt{2}\ln(10)k}{4} \cdot A^{\frac{3}{2}} \cdot 10^{\frac{\sqrt{2}}{2}k\sqrt{A}} = \frac{\eta N^2}{EIRP}(P_{ov,ASIC} + 4P_{link}) \tag{2.8}$$

Recall that $\frac{EIRP}{N^2}$ is power radiated per antenna, $\frac{EIRP}{\eta N^2}$ is the power amplifier power assuming zero routing loss. Equation 2.8 matches with our intuition - if the overhead power or high-speed link power consumed in each ASIC is large compared with the PA power, we would prefer to have more antennas per ASIC because it costs a lot of extra power to distribute antennas onto different ASICs. On the other hand, if routing is very lossy (k is big), we would prefer to have fewer antennas per ASIC in order to shorten the routing distance from the ASIC to the antennas.

In a real design scenario, if our design goal is to minimize the array power consumption, we could find the best (N, A) pair that minimizes the array power consumption for our targeted output power range. More often, N is chosen based on spatial filtering specs and beamforming algorithms, then we could find the best A that minimizes the array power consumption when N and output power range are given. In most cases, when N is large, having multiple antennas on one ASIC ($A > 1$) is more power efficient.
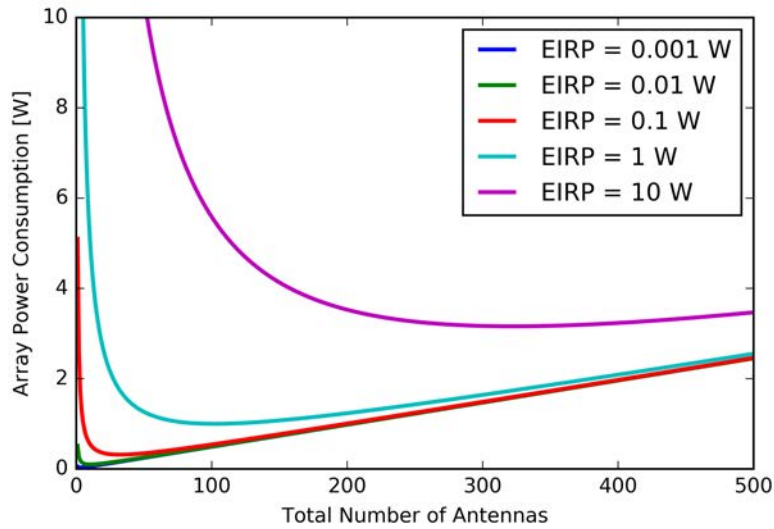
Figure 2.3: Total array power consumption as a function of total number of antennas for various EIRP. Number of antennas per ASIC is held constant at 4. ($P_{ov,ASIC} = 10mW, P_{ov,ant} = 1mW, P_{link} = 0.4mW$)
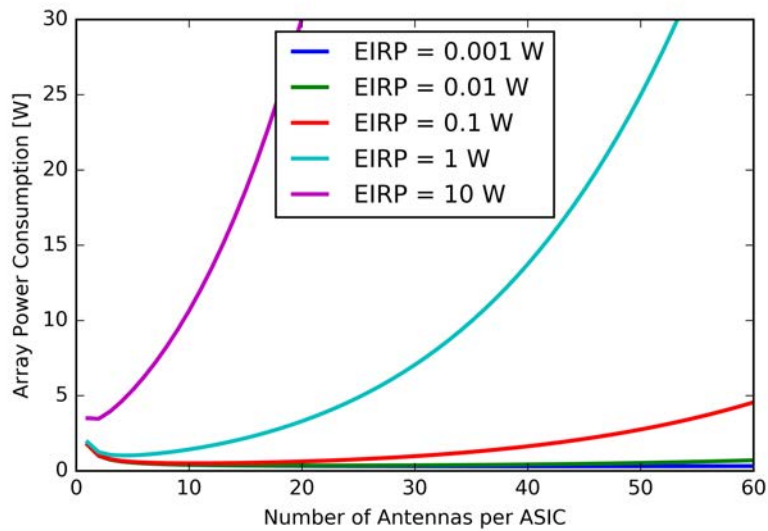
Figure 2.4: Total array power consumption as a function of number of antennas per ASIC for various EIRP. Total number of antennas is held constant at 128. ($P_{ov,ASIC} = 10mW, P_{ov,ant} = 1mW, P_{link} = 0.4mW$)

Fig 2.3 shows the array power consumption as a function of total number of antennas in the array. Each curve corresponds to a different EIRP. As expected, for higher EIRP, the optimal number of antennas is bigger. Fig 2.4 shows the array power consumption as a function of number of antennas per ASIC. Again, each curve corresponds to a different EIRP. When EIRP is higher, lossy routing from ASIC to antenna dominates the power consumption, thus making it more power efficient to having fewer antennas per ASIC.
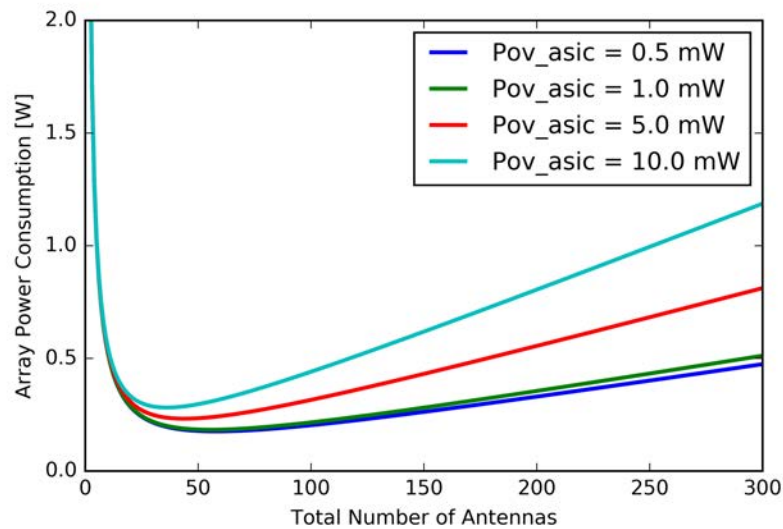
Figure 2.5: Total array power consumption as a function of total number of antennas for various $P_{ov,ASIC}$. Number of antennas per ASIC is held constant at 4. ($EIRP = 0.1W, P_{ov,ant} = 1mW, P_{link} = 0.4mW$)
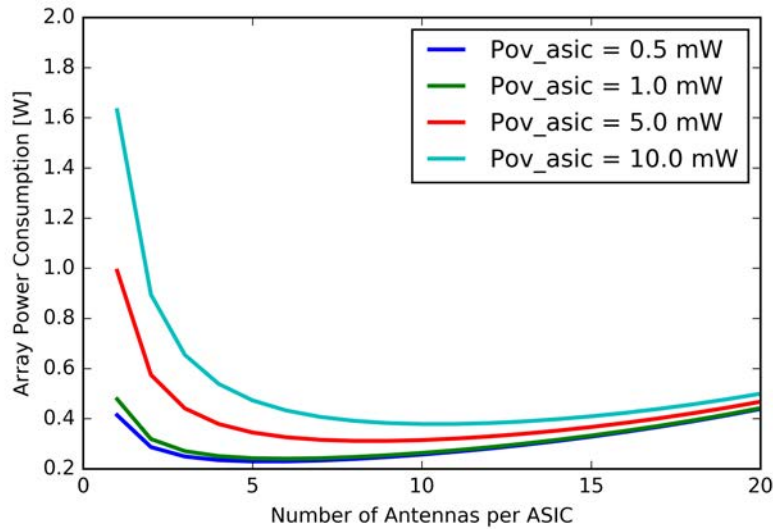
Figure 2.6: Total array power consumption as a function of number of antennas per ASIC for various $P_{ov,ASIC}$. Total number of antennas is held constant at 128. ($EIRP = 0.1W$, $P_{ov,ant} = 1mW$, $P_{link} = 0.4mW$)

Fig 2.5 and 2.6 show array power consumption as a function of total number of antennas in the array and number of antennas per ASIC, respectively. As expected, reducing overhead power reduces the array power consumption of the array. When overhead power is large, optimum N (total number of antennas in the array) is larger, optimum A (number of antennas per ASIC) is smaller. Again, we see the importance of minimizing overhead power consumption in our design.

2.3   Summary

In this chapter, we built a math model to find the optimal number of antennas in the array and on one ASIC. If targeting a large EIRP, it is preferable to have more antennas in the array and fewer antennas on one ASIC. If overhead power of one ASIC increases, the array power consumption increases, and it is preferable to have fewer antennas in the array and more antennas on one ASIC.

In the next chapters, we will zoom in one ASIC in the array, and focus on RF transmitter design. As we have seen in this chapter, it is important to reduce overhead power of the RF TX to minimize the array power consumption.

Chapter 3

RF Transmitter Design

3.1 RF Transmitter for Large Array Applications

Before we start to design the RF transmitter, we first need to think about the difference between designing a transmitter for a single-antenna system and for a multi-antenna system. We will then analyze power consumption of a multi-antenna system.

For simplicity, we group the power consumption of all circuits excluding that of the power amplifier (modulation, baseband, filtering, etc.) into a single term P_{ov}. This portion of power doesn't change with targeted output power. Then the power consumption of a single-antenna system is simply:

$$P_{tot} = P_{ov} + P_{PA} = P_{ov} + \frac{EIRP}{\eta} \tag{3.1}$$

where η is the efficiency of the power amplifier.

For an N-antenna system achieving the same $EIRP$, assume P_{ov} and η don't change with the antenna's output power, the system power consumption is:

$$P_{tot} = N \cdot (P_{ov} + \frac{EIRP}{N^2\eta}) = N \cdot P_{ov} + \frac{EIRP}{N\eta} \tag{3.2}$$

Comparing equation 3.1 and 3.2, while achieving the same $EIRP$, the power consumption of a multi-antenna system could possibly be smaller than that of a single-antenna system,

as long as

$$P_{ov} \ll \frac{EIRP}{\eta} \qquad (3.3)$$

While circuit designers spend most effort to maximize the power amplifier efficiency for a single-antenna system [17–20], from our analysis, when designing an RF transmitter for a multi-antenna system, it is crucial to minimize the overhead power consumption, ideally without sacrificing the power amplifier efficiency.
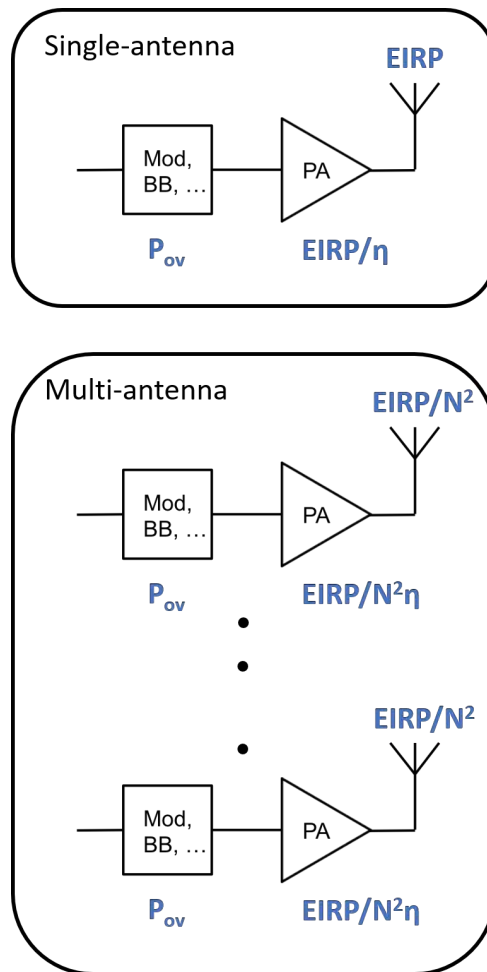
Figure 3.1: Power Consumption of Single-Antenna and Multi-Antenna System

3.2 Design Goal

Now let's consider the design specs of RF transmitters for eWallpaper applications.

While most RF transmitters for single antenna array applications are targeting high output power, with multiple-antenna array, each transmitter in the array only needs to transmit a relatively low output power. For example, with 64 antennas, each RF transmitter only needs to transmit -6dBm to achieve 30dBm $EIRP$.

To maximize the utility of a single common module ASIC design, the RF electronics on the module itself should be able to support multiple/reconfigurable bands - spanning carrier frequencies from 1GHz to 6GHz - as well as programmable in terms of performance (i.e., output power, noise figure, resolution, bandwidth, etc.).

At the same time, it is very important to minimize the out-of-band emissions of the transmitter, since the out-of-band emissions will become blockers and raise the noise floor of other users. To enable programmability of the transmitter, external nonreconfigurable blocks like the SAW filters should be removed.

In summary, our design goal is to design a programmable RF transmitter, while minimizing its out-of-band emissions and power consumption.

3.3 Challenge of Minimizing Out-of-band Emissions

Most conventional transmitter architectures [21–28] are capable of achieving stringent out-of-band emissions requirements by using high-order analog reconstruction filters after the DAC. However, these architectures consist of many analog blocks, which conflicts with our goal of minimizing absolute power consumption.
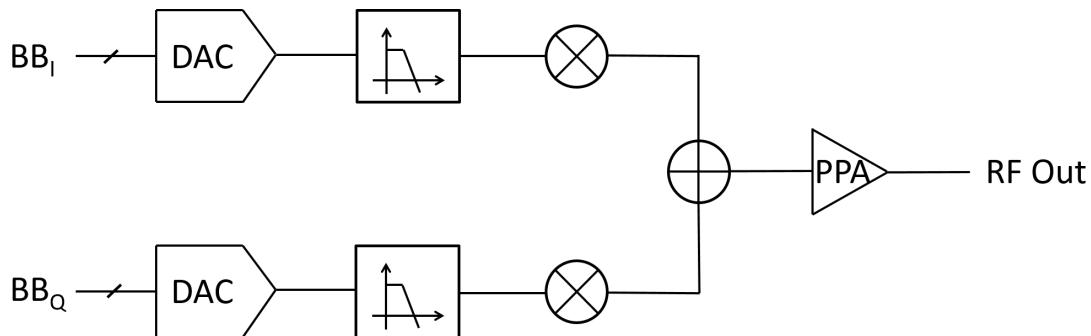
Figure 3.2: Typical Analog Transmitter Architecture

The direct digital RF transmitters [29–40] combines digital-to-analog conversion, up-conversion, and gain control into a single mixer unit cell array. Compared with conventional transmitter architectures, they are more robust and more adapt to low supply voltages. However, there is no baseband reconstruction filter in this architecture, thus both quantization noise and sampling aliases are directly upconverted to RF bands. The lack of filtering makes it difficult to meet the stringent specifications of out-of-band noise emissions.
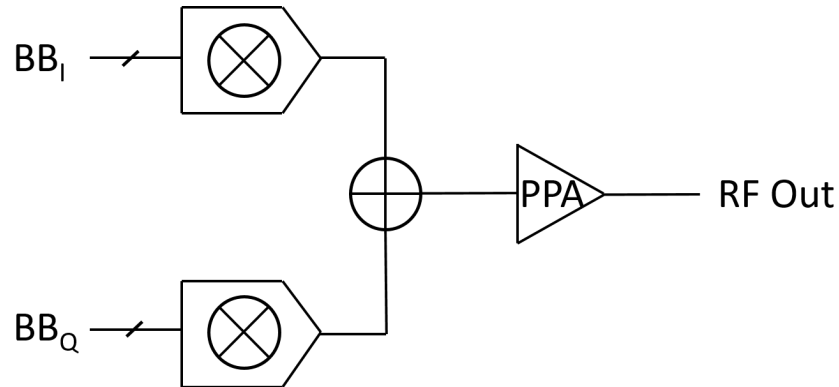
Figure 3.3: Typical Digital Transmitter Architecture

3.4 Proposed Architecture

We propose a mixer-last TX architecture shown in Fig 3.4. Since each TX transmits a low output power in a large antenna array, power amplifier is unnecessary. RF power is provided by the baseband current DACs. Design choices of each block will be discussed in the following two sections.
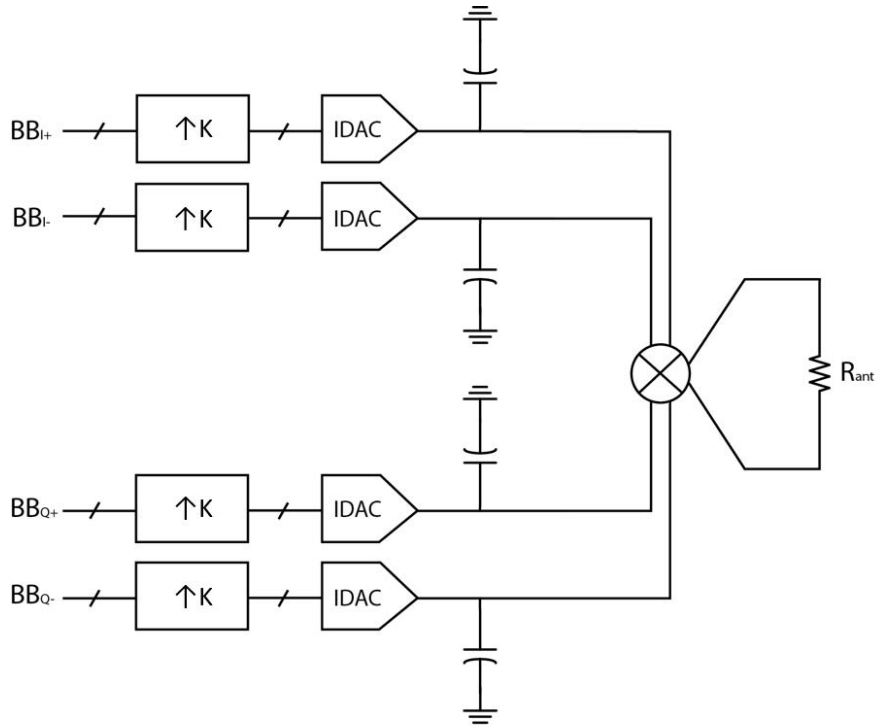
Figure 3.4: TX block diagram

Up-Conversion

In our design, we choose to use passive mixer for signal up-conversion. The major advantage of passive mixers is that they dissipate no DC power, except their clock generation circuits [21].

Conventional passive mixers are driven by 50% duty-cycle clocks. However, because at any given moment, one mixer from the I baseband and another mixer switch from the Q baseband are on, the two RF sides of the mixers need to be buffered by two transconductors, before wiring the up-converted I and Q together. To eliminate the need for the transconductors block, 25% duty cycle passive mixers are proposed [21, 27, 41]. With the 25% duty cycle passive mixers, at any given moment, only I or Q baseband is connected to the RF side. Therefore, there is no need for extra transconductors block [42].

Baseband

As discussed in section 3.3, the main disadvantage of the direct-digital RF modulators is the lack of baseband filtering. A common approach is to build a higher-order active recon-

struction filter at the output of DAC. However, this brings additional linearity and stability issues, and it costs extra power too. Passive filters dissipates zero power, and are excellent for linearity and stability concerns. But we could only afford simple passive filters since higher-order passive filters cost too much chip area. To achieve low out-of-band emissions with simple passive filters, we utilize oversampling and digital filtering techniques.

The basic concept of an oversampling/interpolating DAC is shown in Fig 3.5 [43]. The N-bit input data are received at a rate of f_c, then the digital interpolation filter which is clocked at an oversampling frequency of Kf_c, inserts the extra data points. In the Nyquist case (A), the closest signal aliases are at f_c, which makes the requirements on the analog reconstruction filter quite severe. By oversampling and interpolating, the closest signal aliases are moved to a much higher frequency Kf_c, and the requirements on the filter are greatly relaxed as shown in (B).
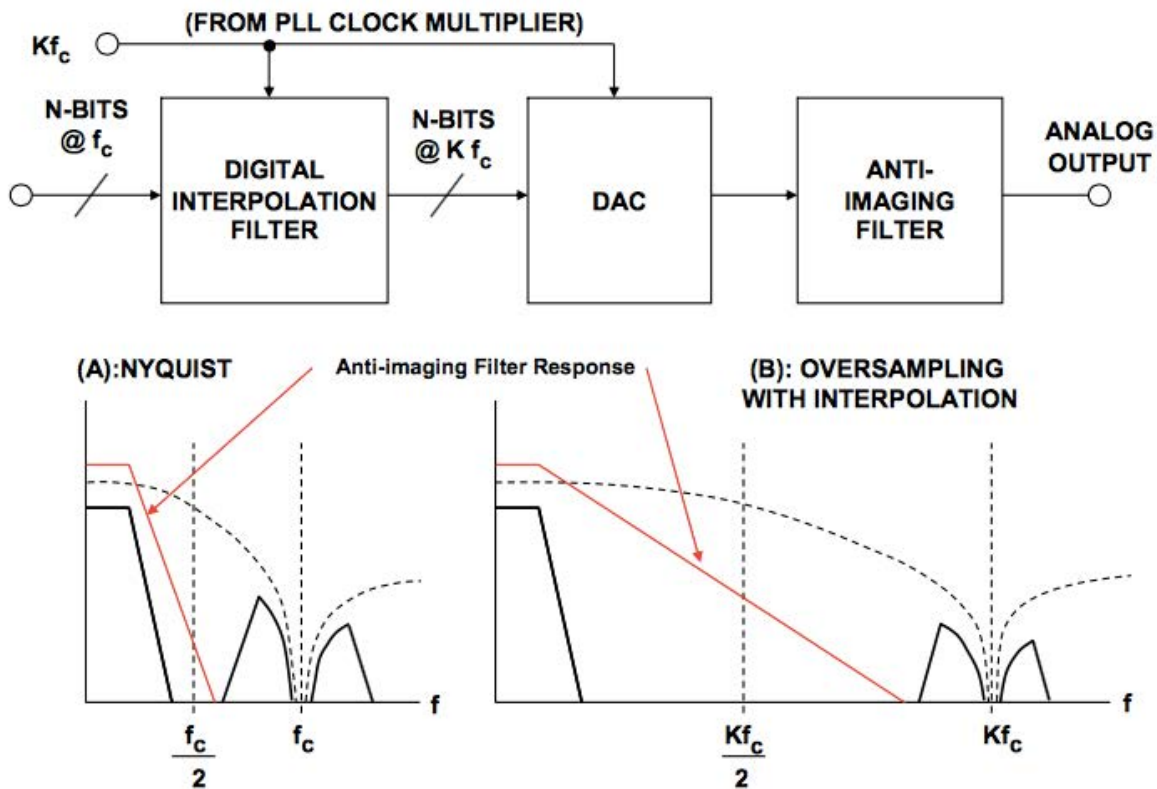
Figure 3.5: Oversampling Interpolating DAC. Adapted [reprinted] from "Oversampling Interpolating DACs", Walt Kester, Analog Devices, 2009

TX Operating Principle

The TX front-end operation could be split into two phases. As illustrated in Fig 3.6, in Phase I, LO1 is on, the capacitor in the Q channel is charged/discharged by the Q DAC. In Phase II, LO2 is on, then the RF charge - and hence the output power - is provided by both the charge stored on the Q capacitor as well as by the Q DAC. When LO3/LO4 is on, it just repeats Phase I/II, with opposite signal direction. Thus, even though each DAC is only directly connected to the antenna 50% of the time (I DAC when LO1 or LO3 is on, Q DAC when LO2 or LO4 is on), charge delivered by the DAC during the other 50% of the time also contributes to the RF output power, thus improving efficiency.
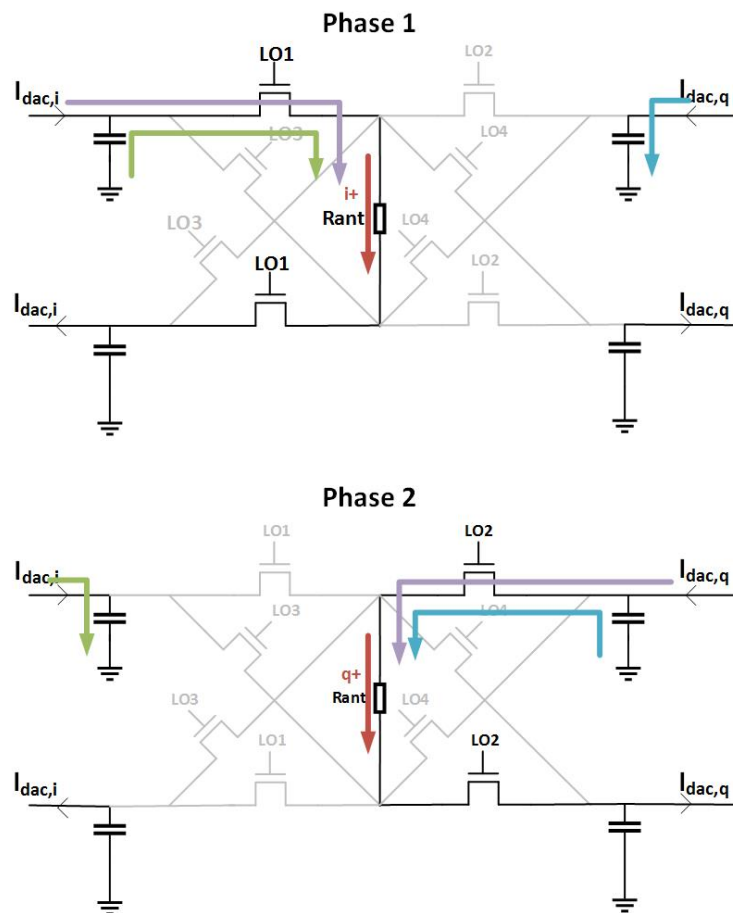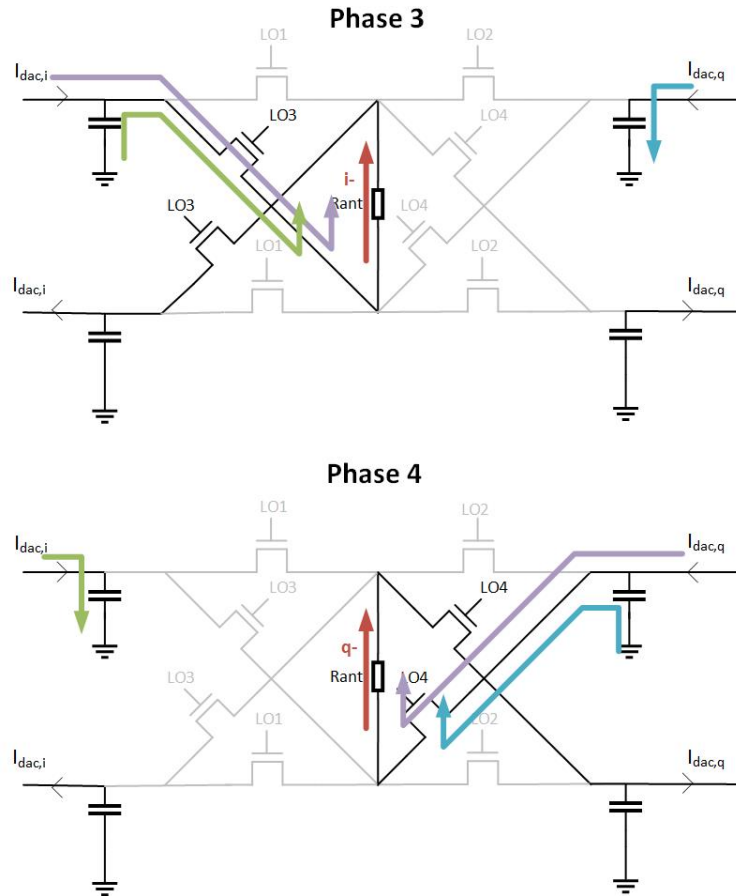
Figure 3.6: TX Operating Principle

We could easily derive the steady-state *rms* voltage. Let I_I and I_Q be the current output of the I DAC and Q DAC. In Phase I, Q DAC dumps charge onto the baseband capacitor.

$$Q_{cap} = I_Q \cdot T_{LO}/4 \tag{3.4}$$

In Phase II, the amount of charge flowed into the antenna resistance is the charge stored on the baseband capacitor Q_{cap}, and the charge from the Q DAC, which also equals to $I_Q \cdot T_{LO}/4$. The average current into the antenna resistance is:

$$I_{avg} = \frac{Q}{T} = 2I_Q \tag{3.5}$$

The *rms* voltage can be calculated as:

$$V_{rms} = \sqrt{\frac{1}{T_{LO}} \int_0^{T_{LO}} I(t)^2 R_{ant}^2 dt} = \sqrt{2}\sqrt{I_{I,rms}^2 + I_{Q,rms}^2} R_{ant} \tag{3.6}$$

In addition to this charge-storage mechanism, the baseband capacitor (in combination with the mixer switch resistance and the antenna resistance) introduces an intrinsic 1st-order low-pass filter in the signal path. This filter further reduces the TX's out-of-band emission and aliases.
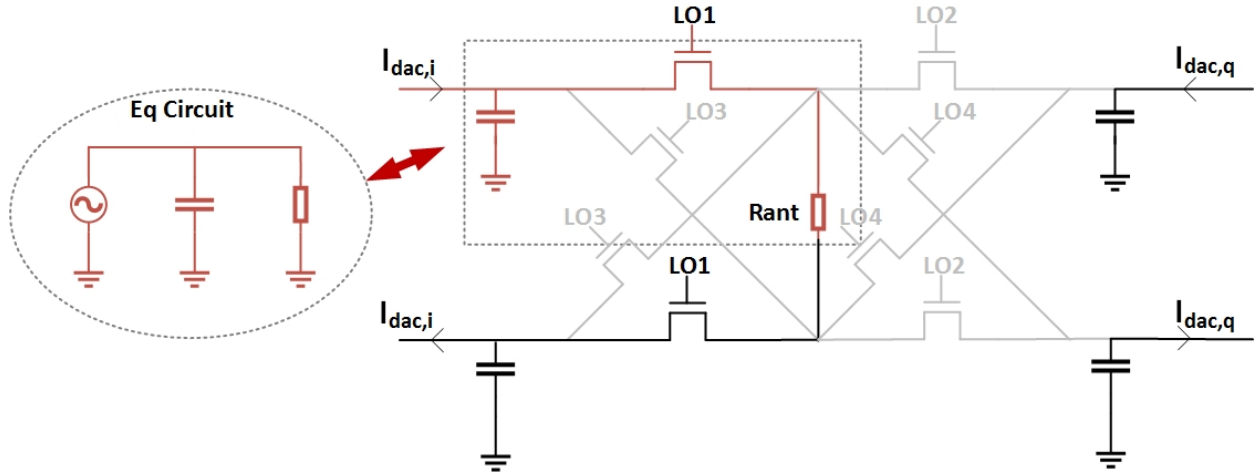
Figure 3.7: Intrinsic Filter

Since the mixer switch for I/Q is on for half of the time, the cut-off frequency of this low-pass filter is:

$$f_{-3dB} = \frac{1}{2\pi R_{eq} C_{bb}} = \frac{1}{2\pi \cdot 2(R_{mixer} + 0.5R_{ant})C_{bb}} \tag{3.7}$$

3.5 Summary

In this chapter, we first compared power consumption of a single antenna system and a large antenna array. It was shown that when designing RF transmitter for large antenna arrays, it is crucial to minimize the absolute power consumption.

Then we established our design goals: a configurable transmitter that transmits low output power, produces low out-of-band emissions and consumes low power consumption. We proposed a mixer-last TX architecture, and calculated its output power and bandwidth.

In Chapter 4, a first prototype using this architecture is implemented in TSMC 65nm CMOS technology. In Chapter 5, we write a generator with BAG framework. It takes system-level specifications as inputs, and generates a design in TSMC's 16nm CMOS technology. Analysis of output power and bandwidth in this chapter will be utilized to write the generator.

Chapter 4

Manual Implementation

4.1 Implementation

CMFB

The DAC was implemented in NMOS/PMOS complementary architecture (Fig 4.1). There are many reasons that could cause the currents of NMOS and PMOS to be unequal, such as accumulated mismatch of multiple current mirror stages (static error), dynamic output voltage (dynamic error), and so on. Since it is important to keep their currents equal at any time in order to maintain stable common mode, a feedback circuit is necessary to compensate for both static errors and dynamic errors real time.
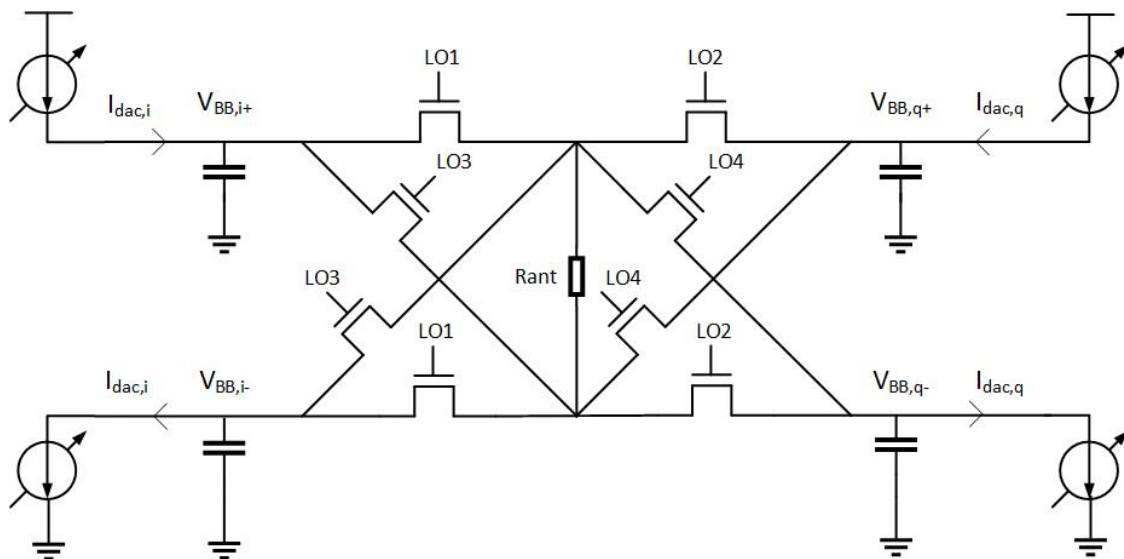
Figure 4.1: TX Front-End showing NMOS/PMOS current DACs

It is obvious that the common mode output voltage increases when PMOS unit current is larger than NMOS unit current, and decreases vice versa. We could then build the common-mode feedback (CMFB) circuit based on this observation.

The operation principle of the negative feedback is straightforward: when PMOS unit current is larger than NMOS unit current, the common mode output voltage (V_f) becomes larger than $0.5V_{DD}$, as a result, $I_{ref,p}$ decreases and $I_{ref,n}$ increases. A capacitor is placed at the gate to set the dominant pole of the feedback circuit.
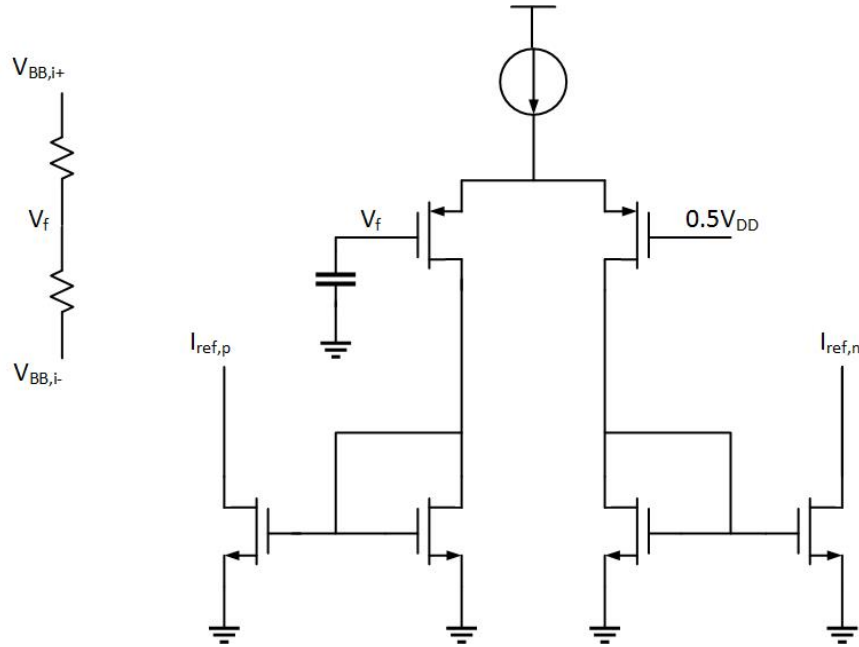
Figure 4.2: CMFB Circuit

Programmability

When implementing RF TX, one of our design goals is to make it programmable in terms of bandwidth and output power.

Recall that we have calculated the bandwidth of the TX in Chapter 3 (Equation 4.1), it is dependent on the baseband capacitor and mixer switch resistance. To make bandwidth programmable, the baseband capacitor is controlled by 7 digital bits (1 LSB is 2 pF).

$$f_{-3dB} = \frac{1}{2\pi R_{eq} C_{bb}} = \frac{1}{2\pi \cdot 2(R_{mixer} + 0.5R_{ant})C_{bb}} \tag{4.1}$$

The output power can be reduced simply by reducing DAC output current. However, at low output power levels (either due to back-off or e.g. for larger arrays), the efficiency will be very low. To improve efficiency at low output power levels, the DAC is divided into 4 identical segments that can be turned on/off independently.
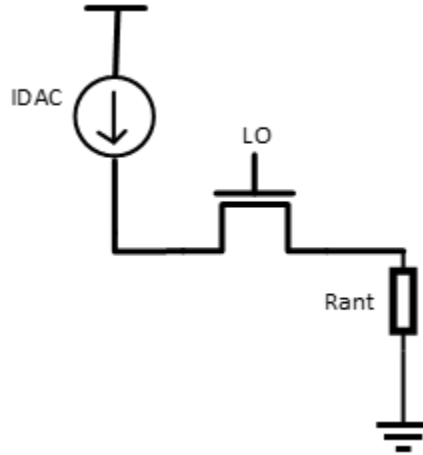
Figure 4.3: Simplified TX front-end to show headroom issue

Fig 4.3 shows a very simplified TX front-end. It is easy to see that when mixer switch is on,

$$V_{dd} = V_{ds,DAC} + V_{mixer} + V_{ant} \geq V_{dsat,DAC} + V_{mixer} + V_{ant} \tag{4.2}$$

$$R_{mixer} \leq \frac{V_{dd} - V_{dsat,DAC} - V_{ant}}{I_{mixer}} \tag{4.3}$$

The above equation shows the upper limit of mixer resistance. When transmitting lower output power, output current is reduced, the headroom of the DAC is relaxed, and hence the resistance of the mixer is relaxed. LO power could also be reduced at low output power levels. In this implementation, the LO mixer consists of 16 independent switches in parallel (each switch is sized to have 350 Ωon-resistance). At low output power levels, some of the mixer switches could be turned off to reduce LO power and improve system efficiency.

In summary, TX's bandwidth is programmable by digital controlled baseband capacitor. When targeting low output power, segments of DAC and mixer switch could be turned off to reduce power consumption (Fig. 4.4).
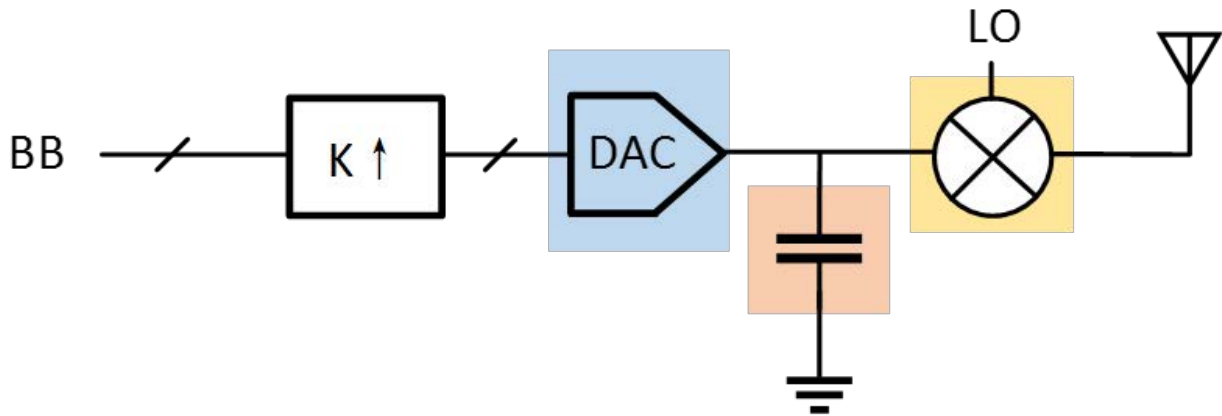
Figure 4.4: Scalable Blocks

4.2 Measurement Results

A test-chip (Fig 4.5) with the proposed TX architecture was implemented in a 65nm CMOS process. The TX occupies $0.83mm^2$, of which $0.29mm^2$ corresponds to the TX core, with the baseband capacitor occupying $0.54mm^2$. [1] In this prototype, the maximum LO frequency was limited to 1.7GHz by the LO frequency divider (which receives an externally generated 4x clock).

---

[1]The baseband capacitor is shared with an RF RX on the same chip.

2.1 mm

2.3 mm
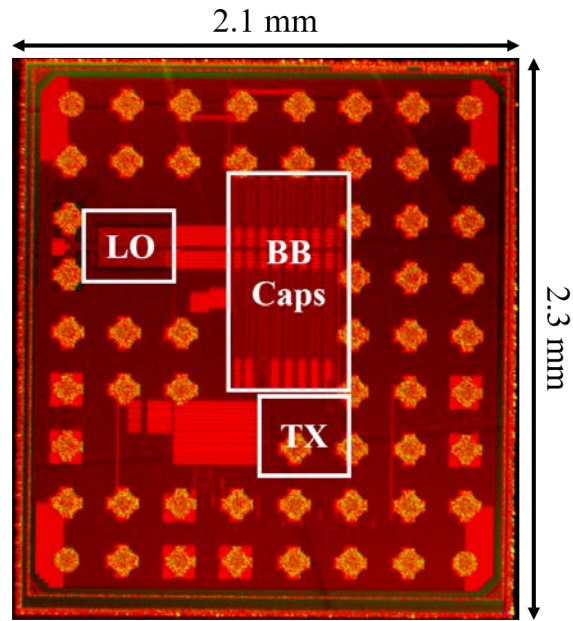
Figure 4.5: Die Photo of 65nm TX Design

Using 1.2V/1V analog/digital supplies, at a sample rate of 20 MS/s and oversampling rate of 320 MS/s, the TX achieves a peak output power of 5.1 dBm, and a peak total efficiency (i.e., output power divided by the total chip power) of 6.64%.

Pout [dBm]	DAC [mW]	LO [mW]	dig [mW]	Other [mW]	Total [mW]	Sys Eff [%]
-6.58	23.58	2.3	3.3	2.1	31.3	0.70
-0.66	27.16	4.5	3.3	2.1	37.1	2.32
2.83	30.62	6.7	3.3	2.1	42.7	4.49
5.15	33.92	9.9	3.3	2.1	49.2	6.64

Table 4.1: Power Breakdown at different Pout of 65nm TX Design

Table 4.1 shows the power breakdown as Pout backs off (4 levels correspond to 4 segments of DAC). As expected, both LO mixer and DAC power are reduced due to the TX scalable configuration. It greatly reduces power consumption in the case when the antenna array is large, and each antenna is configured to send a very low output power. As shown in Fig. 4.6, utilizing these features can reduce power by up to ~35% in large arrays with fixed (independent of array size) EIRP.
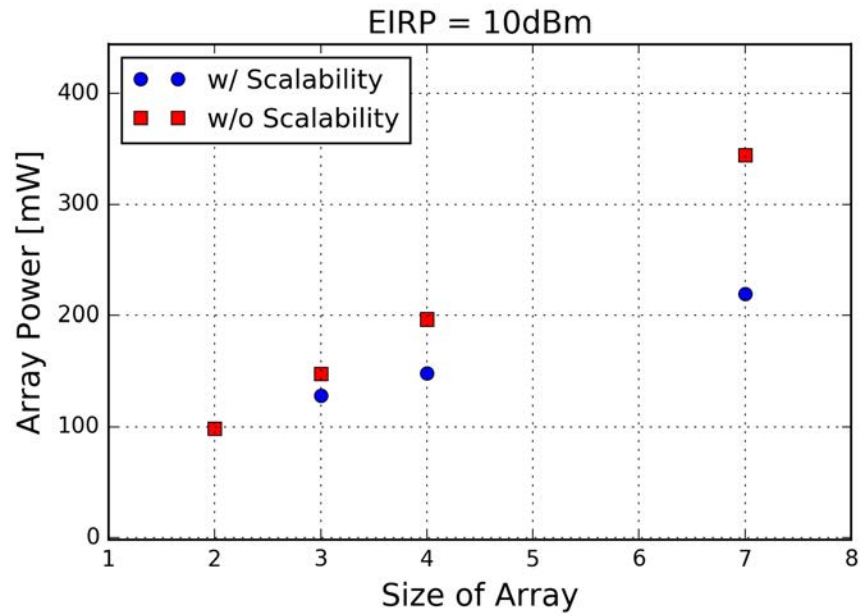
Figure 4.6: Array Power Consumption w/ and w/o scalability features

The measured output spectrum of a 4.2MHz single-tone is shown in Fig 4.7. At 4.7dB backoff from the 5.1dBm peak power, measured image and LO feedthrough are -50dBc and -60dBc, respectively. The measured noise floor at 40MHz offset is -155dBc/Hz.
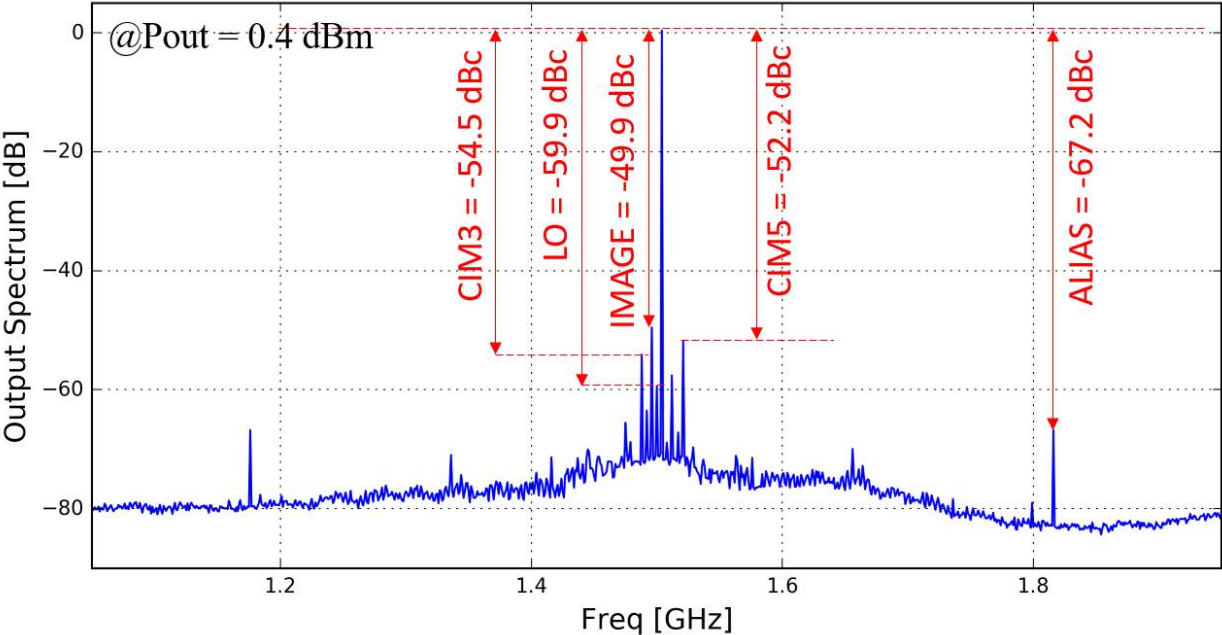
Figure 4.7: Measured Spectrum of 4.2MHz single-tone

Fig 4.9 shows the measured alias rejection and calculated alias rejection with sinc filter at different input frequency.
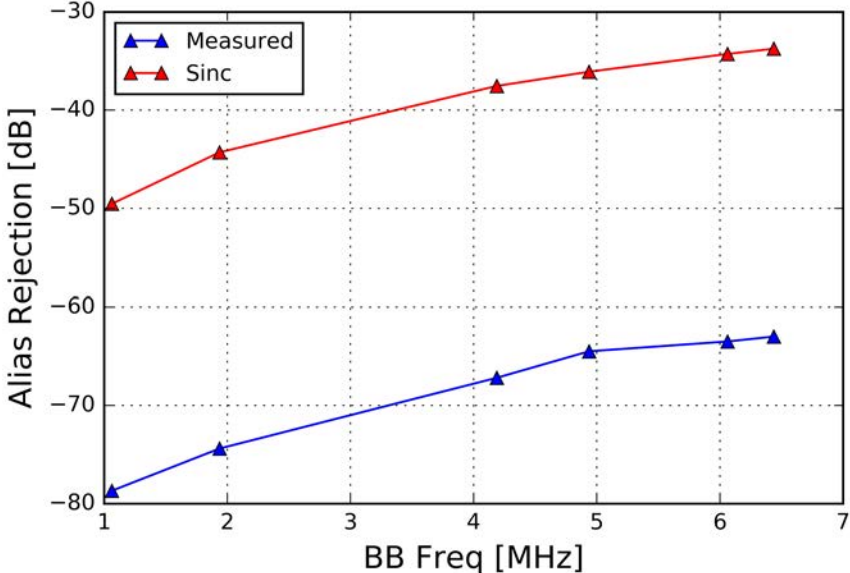
Figure 4.8: Measured alias rejection and calculated alias rejection with sinc filter
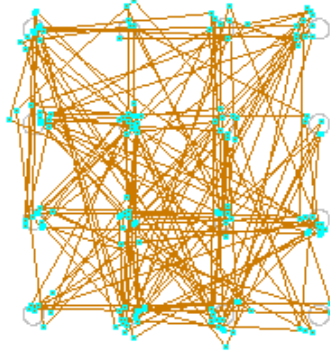
Figure 4.9: Measured 16QAM constellation

The design achieves a raw EVM of 6.21% for a 20MS/s 16QAM constellation. This EVM is mostly due to I/Q mismatch accumulated in multiple current mirror stages, and could be corrected by independent I/Q DAC biases.

Ref	[23]	[44]	[45]	This Work
Architecture	Analog	Analog	QDAC	IDAC
RF bandwidth [MHz]	20	20	20	20
LO Frequency [GHz]	1.9	1.9	1	1.5
Noise [dBc/Hz]	-164	-155	-155	-155
@Offset	@48M	@80M	@45M	@40M
Power	-0.3dBm	0dBm	1dBm	0.4dBm
Max Pout [dBm]	4	6	1	5.1
C-IM3 [dBc]	-67	-57.1	<-50	-54.5
	@0dBm	@2.3dBm	@1dBm	@0.4dBm
EVM [%]	1.05	1	NA	6.2
@Power	@4dBm	@4dBm	@1dBm	@5.1dBm @1dBm
Consumption [mW]	199	98	41.3	49.2 37.1
Total Efficiency [%]	1.26	2.56	3.05	6.64 3.39
Supply [V]	1.5/2.7	1.8	0.9/1.8	1/1.2
Active Area [mm^2]	5.06	1.3	0.25	0.83
Technology [nm]	90	55	28	65

Table 4.2: Comparison Table

As shown in Table 4.2, the proposed TX achieves the lowest absolute power level, even as compared to [45] (which is implemented in 28nm). This low absolute power combined with the power scalability features make this design well suited for massive MIMO applications.

4.3 Reflection

This chip was designed manually with traditional CAD tools. Looking back on the design process, there were a few problems we spent a lot of time and effort on. Firstly, we had to resize the transistors after the first layout is completed and post-layout simulation is done to include the extracted parasitics. We had no other choice but to redraw a large portion of the layout. Secondly, we tried to reuse another designer's differential-pair in our design, we failed since we had different design specifications, and we ended up designing our own circuit and drawing the layout. We spent weeks on these two problems, and wish we could improve the design efficiency in the future.

More importantly, although our TX has lowest power consumption in the output power range of -6dBm to 5dBm, it is not proved to be power efficient in other output power ranges. For example, to transmit -20dBm, it still consumes 31.3mW (only 1 segment of DAC is turned on). Actually it is impossible to build a single transmitter that is power efficient under all applications.

The problems described above lead us to write a generator with BAG framework [1]. It greatly shortens the analog circuit design cycle, and it captures the design procedure in an executable way, which makes it easy to reuse the design. The design cost can perhaps be reduced enough that different transmitters that are efficient for different design specifications can be generated. We'll discuss the implementation of such a generator in the next chapter.

Chapter 5

BAG Generator Implementation

5.1   Introduction to BAG

Motivation

As have mentioned in last chapter, unless the designer has very accurate initial layout parasitics, it is usually the case that after the first-pass layout is completed and post-layout simulation is done to include the extracted parasitics, the whole design sizing and layout must be performed again. Often, each iteration it can require just as much work as the first-pass layout.

In addition, compared to digital circuit design flows, analog/mixed-signal circuit design lacks a common set of well-defined steps in the design flow. The design procedure is highly dependent on the designer's knowledge and experience, and could only be captured in a document, which makes it very difficult to reuse a design under different design specifications or different processes.

To shorten the analog circuit design cycle, and to capture the design procedure in an executable way, Crossley et al. introduced the Berkeley Analog Generator (BAG) framework in 2013 [46, 47]. In 2018, Chang et al. presented BAG2, an evolved and updated version of BAG that enables development of process-portable circuit generators [1]. With this framework, instead of designing one circuit instance, the designers capture their methodology as an executable circuit generator. With these generators, designers can easily produce many circuit instances of the same architecture with different specifications, which makes design reuse practically possible.

For RF TX design, it is impossible to build a single transmitter that is power efficient under all applications. A generator could significantly lower the design cost so that different transmitters that are efficient for different design specifications can be generated with almost

zero marginal cost.

Generator Design Flow

There are three pieces in a circuit generator: 1) A schematic generator; 2) A layout generator; 3) A design script that translates system-level design specifications to lower level input parameters for the schematic and layout generator.

The process of writing a schematic generator is as follows. 1) Create a normal schematic of the circuit, which serves as a template for the schematic generator. 2) Choose input parameters, such as transistor dimensions, for the corresponding schematic. 3) Implement the schematic generation in a Python class to transform the schematic template into a specific circuit instance. There are various BAG2 functions available to simplify this process, including adding or modifying pins, instances and instance connections.

The corresponding layout generator takes both schematic parameters and layout parameters, such as wire width/spacing, as its input. BAG2 provides two frameworks – XBase and Laygo for layout generation. Since we used XBase in our design, we'll only briefly introduce XBase here. It is an abstract class specialized in drawing layout floorplans for analog circuits. Its critical methods include draw_base(), to draw rows of transistors and substrate taps, and draw_mos_conn() and connect_to_substrate(), to connect transistors and substrate taps to the routing grid.

In the next sections, we will first show the design process of the DAC, and implement its layout generator.

Design Script

Fig 5.1 shows the design script workflow. It first translates TX design specifications (out-of-band emissions and output power) to DAC design specifications (INL, DNL, output current), and then finds lower level parameters (DAC segmentation, transistor dimensions of DAC unit cell) that satisfy DAC design specifications.
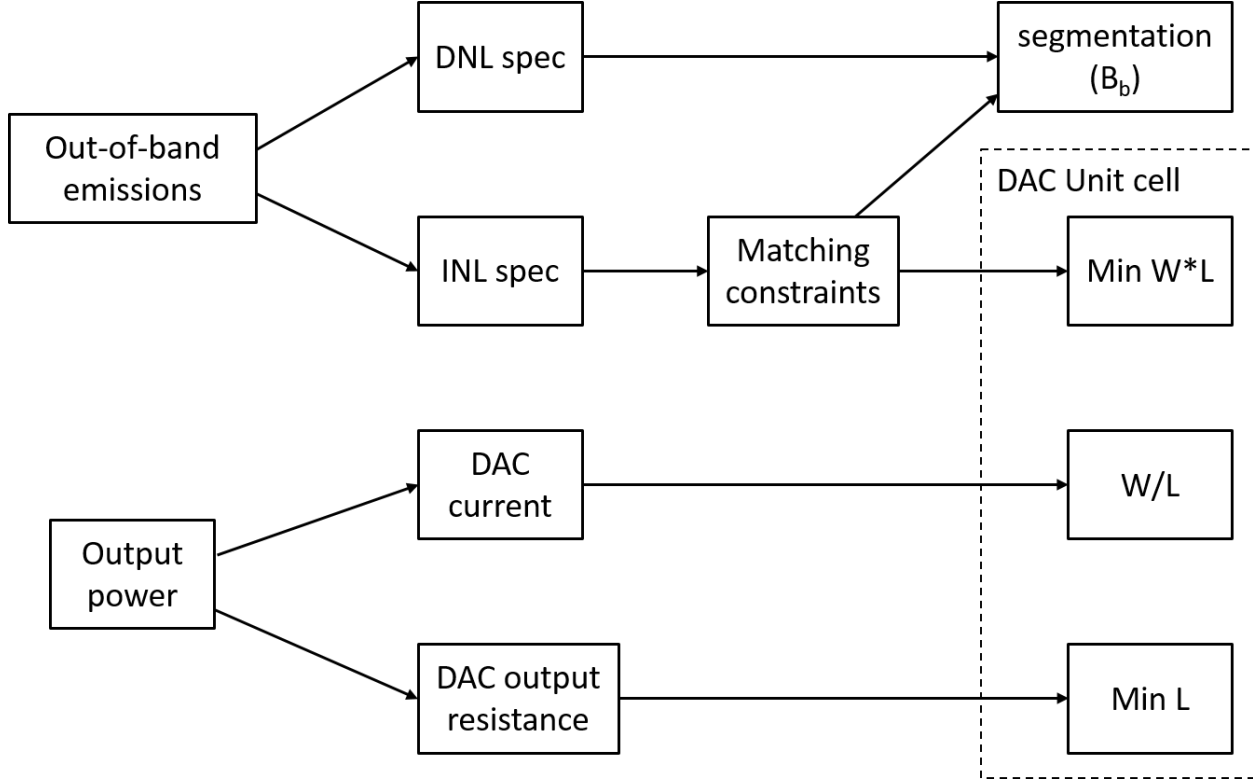
Figure 5.1: DAC Design Flow

TX specs to DAC specs

It is straightforward to calculate the required DAC output current to achieve a given TX output power. We have derived the TX output voltage in Chapter 3:

$$V_{rms} = \sqrt{\frac{1}{T_{LO}} \int_0^{T_{LO}} I(t)^2 R_{ant}^2 dt} = \sqrt{2}\sqrt{I_{I,rms}^2 + I_{Q,rms}^2} R_{ant} \qquad (5.1)$$

The peak output current of the DAC is:

$$I_{peak} = \sqrt{2}I_{I,rms} = \sqrt{2}I_{Q,rms} = \frac{\sqrt{2}}{2} \cdot \frac{V_{rms}}{R_{ant}} = \frac{\sqrt{2}}{2} \cdot \sqrt{\frac{P_{out}}{R_{ant}}} \qquad (5.2)$$

DAC performance specifications are divided into two basic categories: static and dynamic.

Static specifications are behaviors at a steady state, while dynamic specifications refer to behaviors observed during a code-to-code transition.

The most important static specification for current-steering DACs is linearity. It is typically measured by integral and differential nonlinearity (INL and DNL). The differential nonlinearity (DNL) is defined as the difference between an actual step height and the ideal value of 1 LSB. Note that $DNL = -1LSB$ implies a missing code. The integral nonlinearity (INL) is a measure of the deviation between the ideal output value and the actual measured output value for a certain input code.

Dynamic specifications include spurious free dynamic range (SFDR), signal-to-noise ratio (SNR), and so on.

TX specifications of out-of-band emissions can be translated into dynamic specifications such as SNR, SFDR directly. At low input frequencies, SFDR is mostly dependent on INL. A general equation is

$$SFDR \cong 20log_{10}(2^B/INL) \tag{5.3}$$

where B is the resolution of the DAC [48].

DNL error could add additional noise beyond the effects of quantization in the spectrum. Assume that DNL error is uniformly distributed over $\pm 1/2LSB$, which is the same as quantization noise. Hence, the joint PDF for two uniform PDFs is a triangular shape [49].

Quantization error power in this case is:

$$P_q = \int_{-1LSB}^{+1LSB} P(x) \cdot x^2 dx = \frac{LSB^2}{6} \tag{5.4}$$

Signal-to-noise ratio is therefore:

$$SNR = \frac{P_s}{P_q} = 10 \cdot log_{10}(\frac{3}{4} \cdot 2^{2N}) = (6.02N - 1.25)dB \tag{5.5}$$

This indicates that the uniformly distributed $\pm 1/2LSB$ DNL error reduces overall SNR by 3dB. According to this effect, we set the maximum DNL error to be $\pm 1/2LSB$.

In summary, we could translate TX specification of output power into DAC output current by equation 5.2. TX specification of out-of-band emissions is first translated into SFDR, and DAC's INL specification is calculated by equation 5.3. Based on the observation that non-zero DNL across many codes reduces overall SNR, we set maximum DNL error to be $\pm 1/2LSB$.

DAC specs to lower level parameters

In this section, we first discuss the trade-offs of different decoding schemes, and then describe how we translate DNL/INL specifications into the right decoding scheme to use and DAC unit cell transistor dimensions.

Segmentation

Three decoding schemes are generally used, resulting in three DAC architectures: binary-weighted, thermometer-decoded and segmented DACs. Decoding schemes also affect the linearity of the DAC.

In a binary DAC, every bit switches a current to the output that is twice as large as the next least significant bit. The maximum DNL occurs at half-scale transition, 2^{N-1} unit sources are switched on/off and $2^{N-1} - 1$ other independent sources are switched off/on. Assuming a normal distribution for the unit current sources with a standard deviation σ_u, and that the mismatches between the unit current sources are uncorrelated, the standard deviation of the DNL of this transition is:

$$\sigma_{DNL} = \sqrt{(2^{B-1} - 1)\sigma_u^2 + (2^{B-1})\sigma_u^2} = \sqrt{2^B - 1}\sigma_u \qquad (5.6)$$

In contrast, in a thermometer decoded DAC, every unit current source is addressed separately. The digital input code is converted to a thermometer code that controls the switches. In this architecture, the DAC has a guaranteed monotonic behavior since one additional current source has to be switched to the output for one extra LSB. The advantage of this architecture is its good DNL performance. The standard deviation of its DNL is: $\sigma_{DNL} = \sigma_u$.
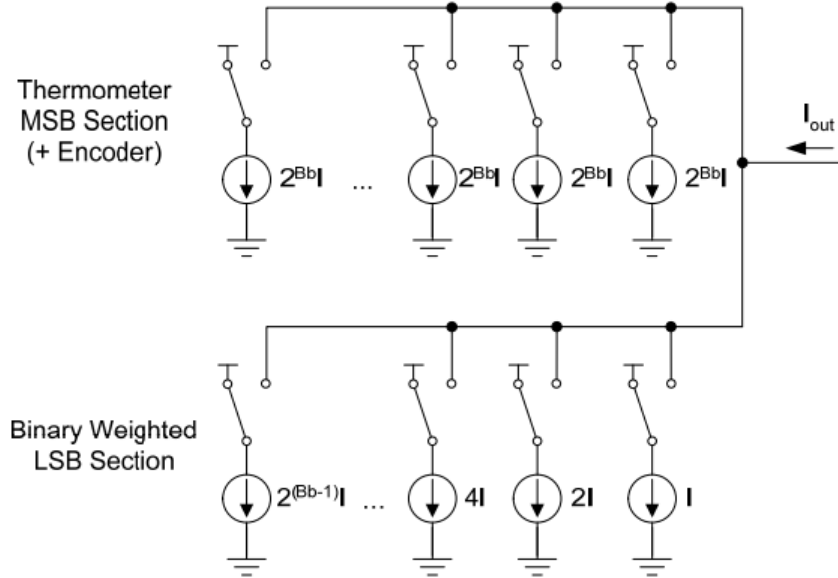
Figure 5.2: Segmented DAC Design. Adapted [reprinted] from "Analog-Digital Interface Integrated Circuits" (Lecture Notes), Bernhard Boser, 2014

Most current-steering DAC are implemented with a segmented architecture (Fig 5.2). In this case, B_b LSBs of the DAC are implemented using a binary architecture, while the $B - B_b$ MSBs are implemented in a thermometer array. The worst case DNL is:

$$\sigma_{DNL} = \sigma_u \sqrt{2^{B_b+1} - 1} \tag{5.7}$$

When $B_b = 0$, it is equivalent to a thermometer DAC; when $B_b = B - 1$, it is equivalent to a binary DAC.

It is worth mentioning that INL is the same for all three decoding schemes. The standard deviation of INL is maximum at mid-scale ($k = N/2$). N is number of unit elements, B is number of bits [50].

$$\sigma_{INL} \cong \sigma_u \sqrt{\frac{N}{2}\left(1 - \frac{\frac{N}{2}}{N}\right)} = \frac{1}{2}\sigma_u\sqrt{N} \cong \frac{1}{2}\sigma_u\sqrt{2^B} \tag{5.8}$$

The key take-aways from the discussion of INL/DNL of three different decoding scheme are: 1) Binary and thermometer schemes are just special cases of segmented scheme. Equation 5.7 applies to all schemes. With this equation, we could find B_b given DNL specification and the standard deviation of unit current sources σ_u. 2) INL is not dependent on B_b. It only depends on matching accuracy of unit current sources. With equation 5.8, we could derive σ_u from INL specification.

Transistor Dimensions

We will discuss implications of σ_u and output impedance on transistor dimensions respectively in this section.

Due to random variations during each fabrication step and environment variations, mismatch exists between any identically designed transistors [51]. Random mismatches are determined by the inherent matching properties of the technology. The standard deviation of unit current sources σ_u is inversely proportional to the gate area of the transistor [48].

$$\sigma_u \propto \frac{1}{\sqrt{WL}} \tag{5.9}$$

After we find the standard deviation of unit current sources σ_u from INL specification, we could find the minimum WL that satisfies the matching accuracy constraint.

Output resistance of a transistor increases when its channel length increases. We could find the minimum L that satisfies the output resistance requirement of the DAC. We'll explain the output resistance requirement of the DAC in the following paragraphs.

When input code is k, k PMOS unit current sources and $N - k$ NMOS unit current sources are switched to the positive output terminal, while $N - k$ PMOS unit current sources and k NMOS unit current sources are switched to the negative output terminal. $N = 2^B$ is the total number of unit current sources. B is number of bits of the DAC. The output current is:

$$I_{out}(k) = \frac{(N - 2k)I}{1 + \frac{Nr_L}{r_o}} \tag{5.10}$$

where r_L is the load resistance at DAC's output, and r_o is the output resistance of a unit current source. In our design, $r_L = 2r_{mixer} + r_{ant}$. Extra power is consumed by current flowing into DAC's output resistance, so we need

$$r_o \gg Nr_L \tag{5.11}$$

to minimize power consumption.

In this design, we used the minimum overdrive voltage V^* that enables transistors to operate in strong inversion region. This maximizes the voltage swing at the DAC output, and enables using large mixer switch resistance without hurting the headroom of DAC. Mixer switch size is minimized and so is mixer driver power consumption. W/L is then directly dependent on DAC output current specification given a fixed V^*.

With the above analysis, sizing the DAC unit cell is quite straightforward. Given INL specification, we could first calculate the required matching accuracy of current sources σ_u, then derive the minimum WL of the transistor. Requirements of output resistance sets the minimum L, and output current I is proportional to $\frac{W}{L}$. Given these constraints, we could find W and L that meet our design specifications.

Layout Generator

Unit Cell Layout Generator

To improve matching of current mirrors, the diode-connected device is placed in the middle of the layout, and the output device is split into two equal number of fingers and placed symmetrically on two sides of the diode-connected device. The dummy devices are also placed symmetrically. This simple layout trick is easily captured by the layout generator.
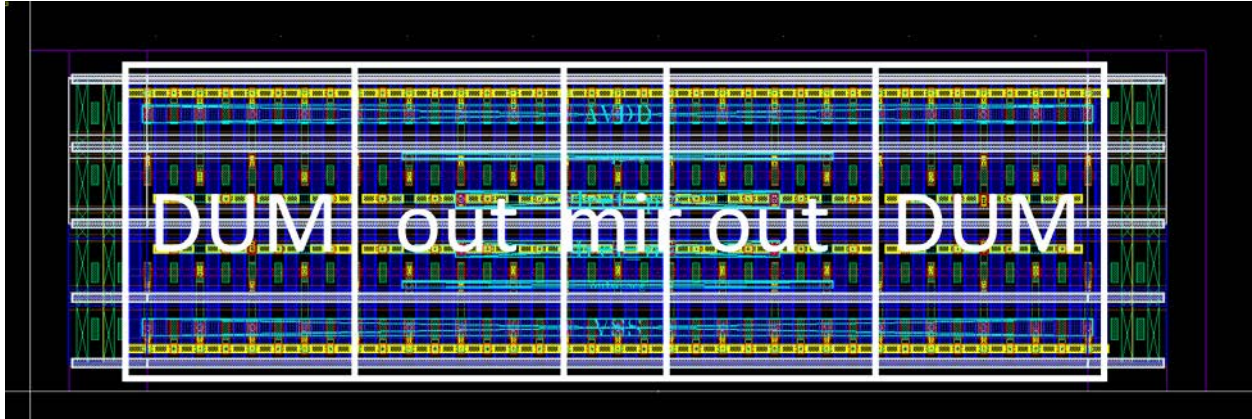
Figure 5.3: DAC Unit Cell Layout

Array Layout

There are essentially two steps to generate the DAC array layout: placing all binary/thermometer cells in an array, connecting their terminals according to a layout floorplan that optimizes its performance.

Many fabrication processes (spread of doping and oxide thickness over the wafer, temperature gradients, die stress, etc.) may cause gradients over a current source array. The gradient errors in these large arrays can become very significant and introduce large systematic errors [52, 53]. Double centroid switching schemes could compensate the systematic errors and optimize DAC's performance [54, 55].

In this design, we use a 4b/5b binary/thermometer segmentation (calculated by design script). The current source of the thermometer array is 8 times the LSB current. In a double centroid switching scheme, each current source is then divided into four current sources, each delivering twice the LSB current. In every quadrant, an 8 x 4 array is used, the remaining places per quadrant are occupied by the binary bits and the same binary bit are arranged symmetrically to compensate for systematic errors. Dummy rows and columns are added to avoid edge effects. A graphical representation of the double centroid structure is shown in Fig 5.4. Different colors denote different quadrants. The shaded area represents the dummy cells.

	30	26	25	28	28	25	26	30	
	14	9	11	12	12	11	9	14	
	17	6	5	19	19	5	6	17	
	22	0	2	20	20	2	0	22	
	21	3	1	23	23	1	3	21	
	18	4	7	16	16	7	4	18	
	13	10	8	15	15	8	10	13	
	29	24	27	B3 B0	B2 B3	27	24	29	
	29	24	27	B3 B2	B1 B3	27	24	29	
	13	10	8	15	15	8	10	13	
	18	4	7	16	16	7	4	18	
	21	3	1	23	23	1	3	21	
	22	0	2	20	20	2	0	22	
	17	6	5	19	19	5	6	17	
	14	9	11	12	12	11	9	14	
	30	26	25	28	28	25	26	30	

Figure 5.4: Double Centroid Switching Scheme

It is worthwhile to re-emphasize that the layout generator itself only knows that the layout has four symmetric quadrants. It doesn't know about the switching scheme at all – the scheme is generated in another piece of code, following the concept of starting from the center and growing out to the edge, and passed on to the layout generator as a 2-dimensional array. The element at (i, j) represents the thermometer cell at row i, column j in the upper-left quadrant. Number of rows ($Nrows$) and columns ($Ncols$) are also input parameters of the layout generator. They are calculated to make the aspect ratio of the DAC layout close to 1. Binary cells are all placed in the middle two rows, also following the concept of starting from the center and growing out to the edge. Total area of binary cells in terms of area of 1 LSB is:

$$A = 1 + 1 + 2 + ... + 2^{B_b-2} = 2^{B_b-1} A_{LSB} \tag{5.12}$$

It equals to area of 1 thermometer cell.

Horizontal and vertical wires are drawn to connect the cells (Fig 5.5). Shared terminals of the cells are supplies (VDD, GND), gate voltage inputs, outputs, and enable signal of the DAC. There are $Ncols$ data inputs per row, except the middle two rows, where there are $Ncols - 1$ thermometer cells and $B_b - 1$ binary cells. Similarly, there are $Nrows - 1$ data inputs per column, except the middle two columns, where there are $Nrows - 1$ thermometer cells and $B_b - 1$ binary cells. To make layout of the cells as identical as possible, $Ncols + B_b - 2$ wires are drawn per row, and $Nrows + B_b - 2$ wires are drawn per column. Redundant wires are connected to GND as dummy wires.

Spacing between two adjacent rows in the array is $max(Width_{wiresPerRow}, Height_{cell})$, and spacing between two adjacent columns in the array is $max(Width_{wiresPerCol}, Width_{cell})$, where $Width_{wiresPerRow}$ and $Width_{wiresPerCol}$ denote total width of all horizontal wires per row, and total width of all vertical wires per column, respectively. As discussed in the last paragraph, they are both functions of $Ncols$, $Nrows$ and B_b.
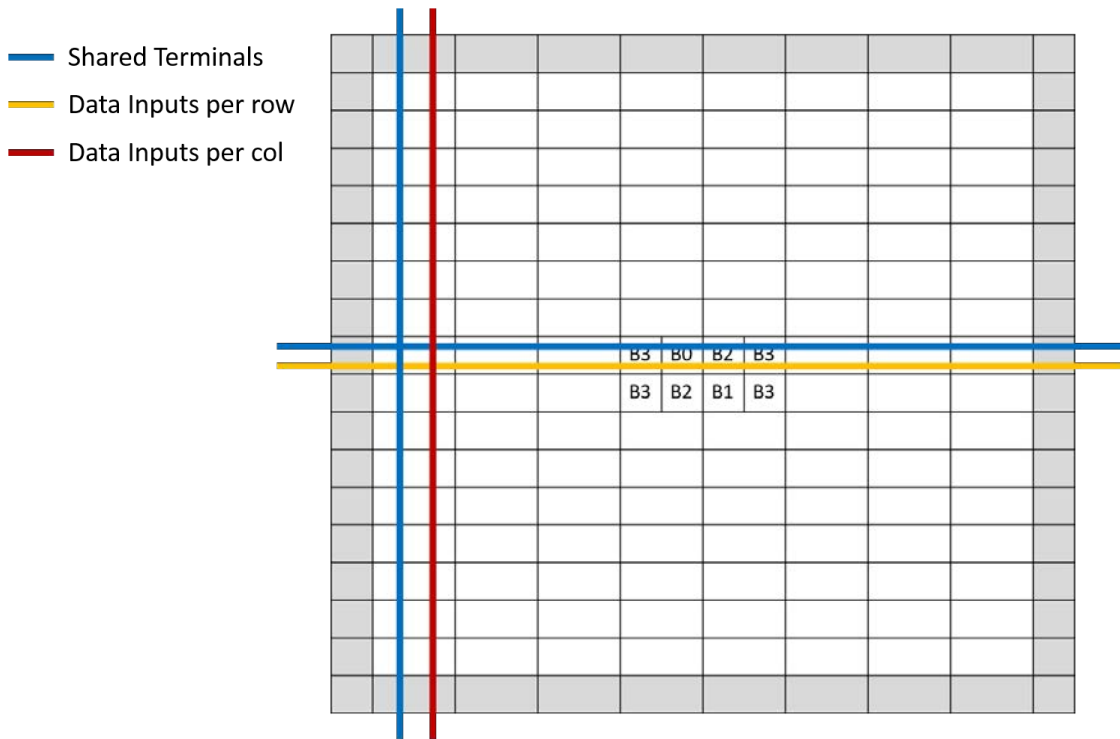
Figure 5.5: Layout Routing

5.2   Measurement Results

A test-chip (Fig 5.6) with the analog blocks of the proposed TX architecture was generated with BAG2, and fabricated in a 16nm CMOS process. There are 4 TRX on the chip, and

each TX occupies $0.075mm^2$, of which $0.034mm^2$ corresponds to the DAC and its peripheral circuits, $0.003mm^2$ corresponds to the mixer, and the baseband capacitor occupies $0.038mm^2$. [1]
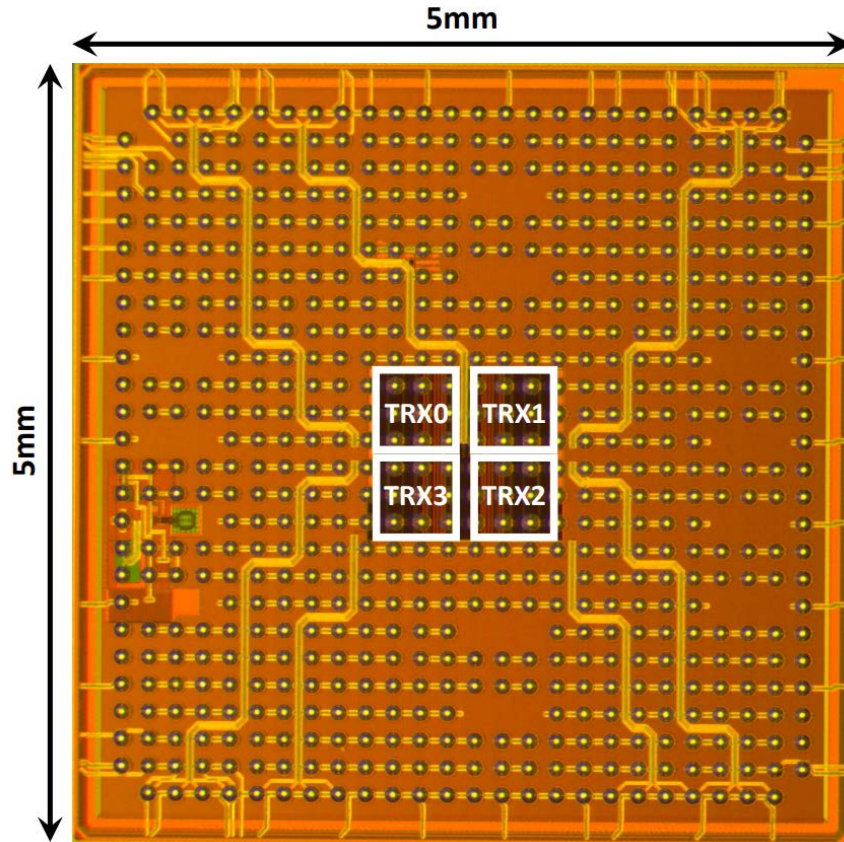
Figure 5.6: Die Photo of 16nm TX Design

Using 0.9V/0.9V analog/digital supplies, at a sample rate of 160 MS/s, a single TX achieves a peak output power of -19.6 dBm, and consumes 5.14 mW. Different from our first prototype, there is no digital oversampling and interpolation filter on this chip. This implementation targets a much lower output power.

Pout [dBm]	DAC [mW]	LO [mW]	Other [mW]	Total [mW]	Sys Eff [%]
-19.6	2.130	1.184(@1GHz)	1.827	5.141	0.21

Table 5.1: Power Breakdown of 16nm TX Design

---

[1]The baseband capacitor is shared with the corresponding RF RX.

Fig 5.7 compares the power consumption of the two chips more clearly. Red line shows measured power consumption of the first 65nm prototype at different power levels. At lower output power levels, some segments of the DAC are turned off to reduce power consumption. Blue line shows calculated array power consumption with this 16nm prototype achieving different EIRP. The 16nm prototype has lower power consumption in the range of less than -4 dBm, while the 65nm prototype has lower power consumption in higher output power range. With the same generator, we can generate designs that are power efficient for other output power ranges and other design specifications.
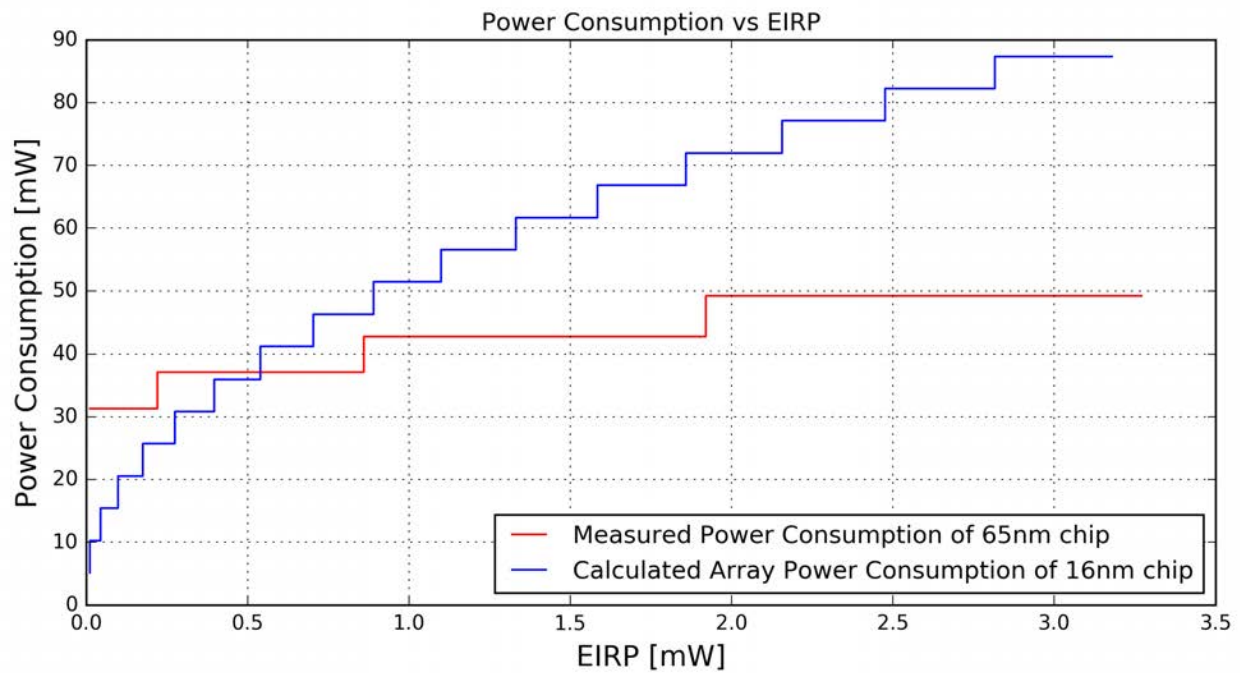
Figure 5.7: Power consumption comparison of the 16nm and 65nm TX designs

Fig 5.8 shows the measured DAC output spectrum of a 3.75MHz single-tone sampled at 20MS/s. SQNR and SFDR of the DAC is 40.8dB and 48.2dB, respectively.
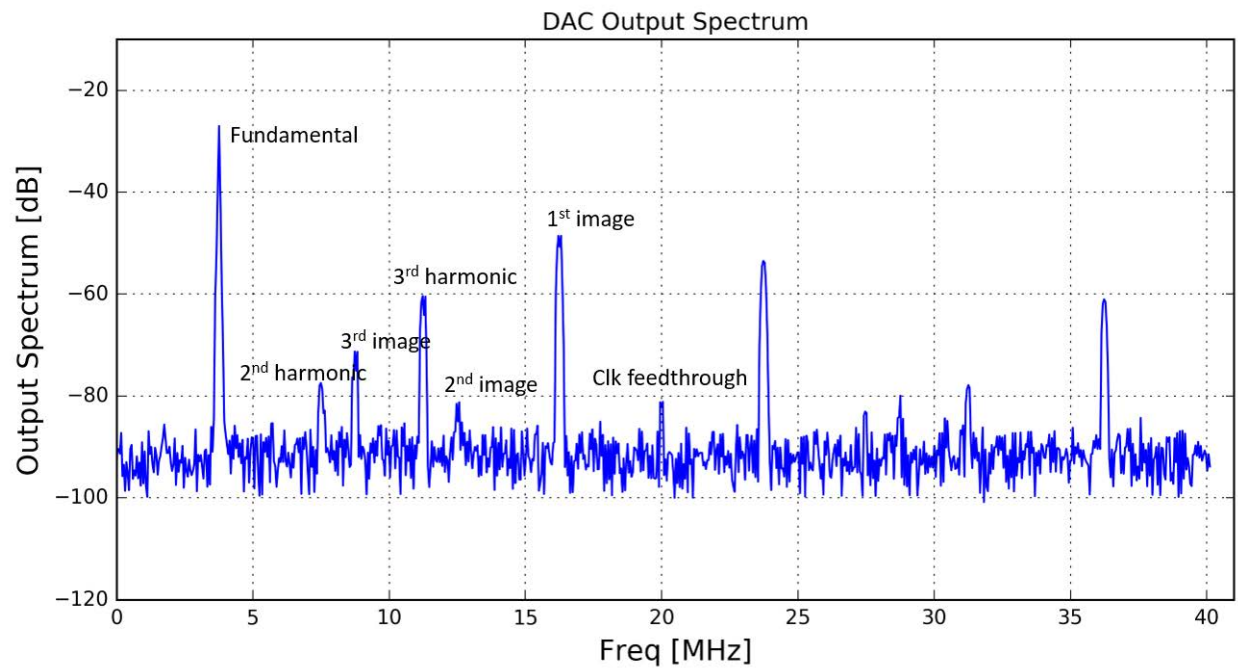
Figure 5.8: Measured DAC Spectrum of 3.75MHz single-tone sampled at 20MS/s

The measured output spectrum of a 10MHz single-tone, sampled at 160MS/s is shown in Fig 5.9. At 3.7dB backoff from the -19.6dBm peak power, measured image and LO feedthrough are -44.0dBc and -55.6dBc, respectively.
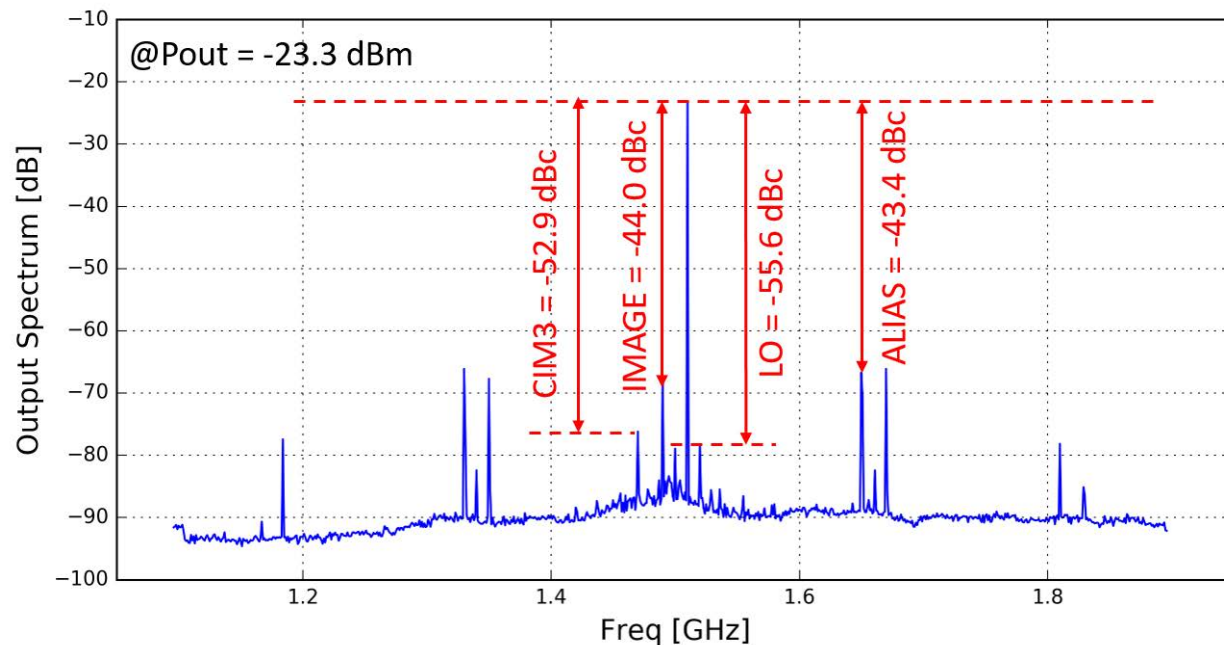
Figure 5.9: Measured Spectrum of 10MHz single-tone

5.3 Summary

In this Chapter, we wrote a generator of our proposed TX architecture with the framework of Berkeley Analog Generator (BAG). It takes system-level specifications (out-of-band emissions and targeted output power) as inputs, and generates a design in TSMC's 16nm CMOS technology. This implementation targets a low output power of -19.6dBm, and consumes 5.14 mW. Compared with the first 65nm prototype, it is more power efficient in the output power range of less than -4 dBm.

Chapter 6

Conclusions

6.1 Thesis Summary

Recent advances in wireless technologies have enabled fast increase of mobile data traffic. The rapid growing demand imposes a big near-far problem on our wireless network. Large antenna arrays could mitigate near-far problem by combining elements in an antenna array in such a way that signals at particular angles experience constructive interference while others experience destructive interference.

This work focuses on RF Transmitter design for large antenna array applications. A mathematical model of the array power consumption is built. It could be used to find the optimal number of antennas in the array and on one chip. It trades off the radiated power and the overhead power incurred by adding more elements to the array.

Power consumption of a single antenna system and a large antenna array is compared. While improving efficiency of the power amplifier is important for single antenna system designs, it is critical to minimize the absolute power consumption when designing a transmitter for a large antenna array. To maximize the utility of a single common module ASIC design, the RF transmitter should be able to support multiple/reconfigurable bands, as well as programmable in terms of performance (i.e., output power, bandwidth, etc.). It also requires the RF transmitter to minimize its out-of-band emissions.

A mixer-last TX architecture is proposed in this work, which uses current DACs to deliver charge directly to the 50Ωoutput RF load. Out-of-band emissions are reduced by high oversampling ratio in digital domain and a 1st-order passive reconstruction filter.

A first chip was fabricated in TSMC's 65nm CMOS technology. With a peak output power of 5.1dBm, the first chip achieves a peak system efficiency of 6.58% and -155dBc/Hz at

40MHz offset. It consumes lowest power compared with previous works that have similar out-of-band emissions performance.

A generator was written with the framework of Berkeley Analog Generator (BAG). It takes system-level specifications (out-of-band emissions and targeted output power) as inputs, and generates a design in TSMC's 16nm CMOS technology. This implementation targets a low output power of -19.6dBm, and consumes 5.14 mW. Compared with the first 65nm prototype, it is more power efficient in the output power range of less than -4 dBm.

6.2 Future Directions

In this thesis, we wrote a BAG2 generator for the TX architecture we proposed. The design choices we made are based on our knowledge and experience. With BAG2 framework, it becomes possible to investigate many design choices. It will be quite interesting and intriguing to build a framework that enables researchers to compare many design choices (different architectures, different layout floorplans, different transistor dimensions, etc.) quantitatively.

Taking a bigger step forward, after we build many different TX generators, we could potentially write some machine learning algorithm, and let it write a generator for an unseen set of design specifications and limitations. Essentially we could leverage knowledge in machine learning area to build a generator of generators.

Bibliography

[1] E. Chang, J. Han, W. Bae, Z. Wang, N. Narevsky, B. Nikolić, and E. Alon. "BAG2: A process-portable framework for generator-based AMS circuit design". In: *IEEE Custom Integrated Circuits Conference (CICC)* (2018), pp. 1–8.

[2] Cisco. *Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2016 —2021 White Paper.* 2017. URL: https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/mobile-white-paper-c11-520862.html.

[3] Wikipedia. *Inverse-square law.* 2018. URL: https://en.wikipedia.org/wiki/Inverse-square_law.

[4] N. H. Dawod, R. Hafez, and I. Marsland. "A Multiuser Zeroforcing System with Reduced Near-Far Problem and MIMO Channel Correlations". In: *IEEE CCECE* (May 2006), pp. 306–309.

[5] B. V. Veen and K. Buckley. "Beamforming: A versatile approach to spatial filtering". In: *IEEE ASSP Mag.* 2 (1988).

[6] A. J. Paulraj and T. Kailath. *Increasing capacity in wireless broadcast systems using Distributed Transmission/Directional Reception (DTDR).* 1994. URL: http://www.google.com/patents/US5345599.

[7] S. Talwar, M. Viberg, and A. J. Paulraj. "Blind estimation of multiple co-channel digital signals using an antenna array". In: *IEEE Signal Process. Lett.* 2 (Feb. 1994), pp. 29–31.

[8] G. J. Foschini and M. J. Gans. "On limits of wireless communications in a fading environment when using multiple antennas". In: *Wireless Pers. Commun.* (Mar. 1998), pp. 311–335.

[9] H. Huh, G. Caire, H. C. Papadopoulos, and S. A. Ramprashad. "Achieving Massive MIMO Spectral Efficiency with a Not-so-Large Number of Antennas". In: *IEEE Trans. Wireless Communications* 9 (Sept. 2012), pp. 3226–3239.

[10] Lingkai Kong. "Energy—Efficient 60GHz Phased-Array Design for Multi-Gb/s Communication Systems". PhD dissertation. University of California, Berkeley, 2014.

[11]   R. J. Mailloux. *Phased array antenna handbook*. Artech House Boston, 2005.

[12]   *Framework and overall objectives of the future development of IMT for 2020 and beyond*. 18th Meeting of ITU Working Party 5D, 2015.

[13]   C. J. Zhang. "Realizing Massive MIMO in LTE-Advanced and 5G". In: *Brooklyn 5G Summit*. 2015.

[14]   Ali M. Niknejad. *The xG Vision: Making the Internet truly wireless*. 2014. URL: http://files-eu.clickdimensions.com/ericssoncom-anlg4/documents/xg_vision_5g_public_.pdf.

[15]   A. Puglielli, N. Narevsky, P. Lu, T. Courtade, G. Wright, B. Nikolić, and E. Alon. "A scalable massive MIMO array architecture based on common modules". In: *IEEE International Conference on Communication Workshop (ICCW)*. June 2015, pp. 1310–1315.

[16]   A. Puglielli, A. Townley, G. Lacaille, V. Milovanović, P. Lu, K. Trotskovsky, A. Whitcombe, N. Narevsky, G. Wright, T. Courtade, E. Alon, B. Nikolić, and A. M. Niknejad. "Design of Energy- and Cost-Efficient Massive MIMO Arrays". In: *Proceedings of the IEEE*. Vol. 104. 3. 2016, pp. 586–606.

[17]   N.-C. Kuo, B. Yang, A. Wang, L. Kong, C. Wu, V. P. Srini, Elad Alon, B. Nikolic, and A. M. Niknejad. "A Wideband All-Digital CMOS RF Transmitter on HDI Interposers With High Power and Efficiency". In: *IEEE Transactions on Microwave Theory and Techniques* (Nov. 2017).

[18]   H. Xu, Y. Palaskas, A. Ravi, M. Sajadieh, M. A. El-Tanani, and K. Soumyanath. "A flip-chip-packaged 25.3 dBm class-D outphasing power amplifier in 32 nm CMOS for WLAN application". In: *IEEE J. Solid-State Circuits* 7 (July 2011), pp. 1596–1605.

[19]   Z. Deng, E. Lu, E. Rostami, D. Sieh, D. Papadopoulos, B. Huang, R. Chen, H. Wang, WH Hsu, CH Wu, and O. Shanaa. "A dual-band digital-WiFi 802.11a/b/g/n transmitter SoC with digital I/Q combining and diamond profile mapping for compact die area and improved efficiency in 40 nm CMOS". In: *IEEE ISSCC Dig. Tech. Papers* (2016), pp. 172–173.

[20]   S. Hu, S. Kousai, and H. Wang. "A broadband mixed-signal CMOS power amplifier with a hybrid class-G Doherty efficiency enhancement technique". In: *IEEE J. Solid-State Circuits* 3 (Mar. 2016), pp. 598–613.

[21]   X. He, J. van Sinderen, and R. Rutten. "A 45nm WCDMA transmitter using direct quadrature voltage modulator with high oversampling digital front–end". In: *IEEE ISSCC Dig. Tech. Papers* (2010), pp. 62–63.

[22]   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 ISSCC Dig. Tech. Papers* (2011), pp. 374–376.

[23] O. Oliaei, M. Kirschenmann, D. Newman, K. Hausmann, H. Xie, P. Rakers, M. Rahman, M. Gomez, C. Yu, B. Gilsdorf, and K. Sakamoto. "A multiband multimode transmitter without driver amplifier". In: *IEEE ISSCC Dig. Tech. Papers* (2012), pp. 164–166.

[24] C. Jones, B. Tenbroek, P. Fowers, C. Beghein, J. Strange, F. Beffa, and D. Nalbantis. "Direct–conversion WCDMA transmitter with –163dBc/Hz noise at 190MHz offset". In: *IEEE ISSCC Dig. Tech. Papers* (2007), pp. 336–607.

[25] T. Sowlati et al. "Single–chip multiband WCDMA/HSDPA/HSUPA/EGPRS transceiver with diversity receiver and 3G DigRF interface without SAW filters in transmitter/3G receiver paths". In: *IEEE ISSCC Dig. Tech. Papers* (2009), pp. 116–117.

[26] M. Cassia, A. Hadjichristos, H. S. Kim, J. S. Ko, J. Yang, S. O. Lee, and G. Sahota. "A low–power CMOS SAW–less quad band WCDMA/HSPA/HSPA+/1X/EGPRS transmitter". In: *IEEE J. Solid-State Circuits* (2009).

[27] M. Collados, H. Zhang, B. Tenbroek, and H.-H Chang. "A low–current digitally predistorted direct–conversion transmitter with 25% duty–cycle passive mixer". In: *IEEE Trans. Microw. Theory Tech.* (2014).

[28] P. Rossi, N. Codega, D. Gerna, A. Liscidini, D. Ottini, Y. He, A. Pirola, E. Sacchi, G. Uehara, C. Yang, and R. Castello. "An LTE transmitter using a class–A/B power mixer". In: *IEEE ISSCC Dig. Tech. Papers* (2013), pp. 340–341.

[29] 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 broadband mixed-signal CMOS power amplifier with a hybrid class-G Doherty efficiency enhancement technique". In: *IEEE J. Solid-State Circuits* 3 (Mar. 2016), pp. 598–613.

[30] P. Eloranta, P. Seppine, S. Kallioinen, T. Saarela, and A. Parssinen. "A multimode transmitter in 0.13um CMOS using direct-digital RF modulator". In: *IEEE J. Solid-State Circuits* 12 (Dec. 2007), pp. 2774–2784.

[31] M. Ingels, X. Zhang, K. Raczkowski, S. Cha, P. Palmers, and J. Craninckx. "A linear 28 nm CMOS digital transmitter with 2 x 12 bit up to LO baseband sampling and –58 dBc C–IM3". In: *Proc. Eur. Solid State Circuits Conf. (ESSCIRC)* (2014), pp. 379–382.

[32] Z. Boos, A. Menkhoff, F. Kuttner, M. Schimper, J. Moreira, H. Geltinger, T. Gossmann, P. Pfann, A. Belitzer, and T. Bauernfeind. "A fully digital multimode polar transmitter employing 17b RF DAC in 3G mode". In: *IEEE ISSCC Tech. Dig. Papers* (2011), pp. 376–378.

[33] C. Lu, H. Wang, CH Peng, A. Goel, S. Son, P. Liang, A. M. Niknejad, HC Hwang, and G. Chien. "A 24.7 dBm all-digital RF transmitter for multimode broadband applications in 40 nm CMOS". In: *IEEE ISSCC Dig. Tech. Papers* (Feb. 2013), pp. 422–423.

[34] M. S. Alavi, R. B. Staszewski, L. C. N. de Vreede, and J. R. Long. "A wideband 2 x 13bit all–digital I/Q RF–DAC". In: *IEEE Trans. Microw. Theory Tech.* 4 (Apr. 2014), pp. 732–752.

[35] S.– M. Yoo, J. S. Walling, O. Degani, B. Jann, R. Sadhwani, J. C. Rudell, and D. J. Allstot. "A class–G switched–capacitor RF power amplifier". In: *IEEE J. Solid-State Circuits* 5 (May 2013), pp. 1212–1224.

[36] A. Jerng and C. G. Sodini. "A wideband ΔΣdigital–RF modulator for high data rate transmitters". In: *IEEE J. Solid-State Circuits* 8 (Aug. 2007), pp. 1710–1722.

[37] A. Pozsgay, T. Zounes, R. Hossain, M. Boulemnakher, V. Knopik, and S. Grange. "A fully digital 65nm CMOS transmitter for the 2.4–to–2.7 GHz WiFi/WiMAX bands using 5.4 GHz ΔΣRF DACs". In: *IEEE ISSCC Dig. Tech. Papers* (2008), pp. 360–619.

[38] S. Fukuda, S. Miya, M. I. K. Hamashita, and B. Nauta. "Direct-digital modulation (DIDIMO) transmitter with -156 dBc/Hz Rx-band noise using FIR structure". In: *Proc. Eur. Solid State Circuit Conf. (ESSCIRC)* (2012), pp. 53–56.

[39] W. M. Gaber, P. Wambacq, J. Craninckx, and M. Ingels. "A CMOS IQ direct digital RF modulator with embedded RF FIR-based quantization noise filter". In: *Proc. Eur. Solid State Circuit Conf. (ESSCIRC)* (2011), pp. 139–142.

[40] A. Mirzaei and H. Darabi. "A low-power WCDMA transmitter with an integrated notch filter". In: *IEEE J. Solid-State Circuits* 12 (Dec. 2008), pp. 2868–2881.

[41] H. Wang, P. Jiang, T. Mo, and J. Zhou. "A low–WCDMA transmitter with 25% du-ty–cycle LO generator in 65nm CMOS". In: *IEEE International Conference on ASIC* (Oct. 2011).

[42] A. Mirzaei, D. Murphy, and H. Darabi. "Analysis of direct–conversion IQ transmitter with 25% duty–passive mixers". In: *IEEE Transactions on Circuits and Systems* 10 (Oct. 2011), pp. 2318–2331.

[43] Kester, Walt. *Oversampling Interpolating DACs*. 2009. URL: https://www.analog.com/media/ru/training-seminars/tutorials/MT-017.pdf.

[44] N. Codega, P. Rossi, A. Pirola, A. Liscidini, and R. Castello. "A current-mode, low out-of-band noise LTE transmitter with a class-A/B power mixer". In: *IEEE J. Solid-State Circuits* (2014).

[45] P. E. P. Filho, M. Ingels, P. Wambacq, and J. Craninckx. "A transmitter with 10b 128MS/s incremental-charge-based DAC achieving -155dBc/Hz out-of-band noise". In: *IEEE ISSCC Dig. Tech. Papers* (2015), pp. 1–3.

[46] J. Crossley, A. Puggelli, H.-P. Le, B. Yang, R. Nancollas, K. Jung, L. Kong, N. Narevsky, Y. Lu, N. Sutardja, E. J. An, A. L. Sangiovanni-Vincentelli, and E. Alon. "BAG: A design-oriented integrated framework for the development of AMS circuit generators". In: *IEEE/ACM International Conference on Computer-Aided Design (IC-CAD)* (2013), pp. 74–81.

[47] John Crossley. "BAG: A designer-oriented framework for the development of AMS circuit generators". PhD dissertation. University of California, Berkeley, 2014.

[48] B. Boser. *Lecture Note: Analog-Digital Interface Integrated Circuits.*

[49] Opris, Ion E. *Challenges in A/D Design and Practical Understanding of A/D Specifications.* 2003.

[50] S. Kuboki, K. Kato, N. Miyakawa, and K. Matsubara. "Nonlinearity analysis of resistor string A/D converters". In: *IEEE Transactions on Circuits and Systems* (1982), pp. 383–390.

[51] M. J. M. Pelgrom, A. C. J. Duinmaijer, and A. P. G. Welbers. "Matching properties of MOS transistors". In: *IEEE J. Solid-State Circuits* (1989), pp. 1433–1439.

[52] G. Van der Plas, J. Vandenbussche, W. Sansen, M. Steyaert, and G. Gielen. "A 14-bit intrinsic accuracy Q random walk CMOS DAC". In: *IEEE J. Solid-State Circuits* (Dec. 1999), pp. 1708–1718.

[53] J. Bastos, A. Marques, M Steyaert, and W. Sansen. "A 12-bit intrinsic accuracy high-speed CMOS DAC". In: *IEEE J. Solid-State Circuits* (Dec. 1998), pp. 1959–1969.

[54] A. Van den Bosch, M. Borremans, J. Vandenbussche, G. Van der Plas, A. Marques, J. Bastos, M. Steyaert, G. Gielen, and W. Sansen. "A 12 bit 200 MHz low glitch C-MOS D/A converter". In: *IEEE Custom Integrated Circuits Conference (CICC)* (1998), pp. 249–252.

[55] C.-H. Huang, T.-T. Hou, and H.-H. Wang. "A 12-bit 250-MHz Current-Steering DAC". In: *International Conference on ASIC* (2005).