UNIVERSITY OF CALIFORNIA

Los Angeles

# A Digital RF Transmitter with Background Nonlinearity Correction

A dissertation submitted in partial satisfaction

of the requirements for the degree

Doctor of Philosophy in Electrical and Computer Engineering

by

Seyed Mehrdad Babamir

2019

ABSTRACT OF THE DISSERTATION

A Digital RF Transmitter with Background Nonlinearity Correction

by

Seyed Mehrdad Babamir
Doctor of Philosophy in Electrical and Computer Engineering
University of California, Los Angeles, 2019
Professor Behzad Razavi, Chair

This dissertation describes a new digital transmitter architecture that automatically corrects static and dynamic nonlinearities with no need for digital predistortion or adaptation. We draw upon the Newton-Raphson method of solving equations and show that it leads to $\Delta\Sigma$ modulation as a special case and to a compact, efficient transmitter in the general case. A complete transmitter realized in 28-nm CMOS technology achieves an overall efficiency of 50% while delivering +24 dBm with an adjacent channel power of $-35.4$ dB and a receive-band noise of $-137$ dBc/Hz.

The dissertation of Seyed Mehrdad Babamir is approved.

Gregory J. Pottie

Danijela Cabric

Katsushi Arisaka

Behzad Razavi, Committee Chair

University of California, Los Angeles

2019

*To my lovely family*

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

# VITA

2009–2013 B.S. (Electrical Engineering), Sharif University of Technology.

2013–2015 M.S. (Electrical Engineering), Sharif University of Technology.

# PUBLICATIONS

A. Rezvanitabar, **S.M. Babamir**, B. Behmanesh, S.M. Atarodi, "A tunable-Q 4-path bandpass filter with Gm-C second-order baseband impedances," *Iranian Conference on Electrical Engineering,* pp. 244-248, May 2017.

F.S. Babamir, A. Hatamizadeh, **S.M. Babamir**, M. Dabbaghian, A. Norouzi, "Application of genetic algorithm in automatic software testing," *International Conference on Networked Digital Technologies,* pp. 545-552, July 2010.

F.S. Babamir, E. Amini, **S.M. Babamir**, A. Norouzi, B.B. Ustundag, "Genetic Algorithm and Software Testing based on Independent Path Concept.," *International Conference on Genetic and Evolutionary Methods,* pp. 180-186, 2010.

F.S. Babamir, A. Hatamizadeh, **S.M. Babamir**, M. Dabbaghian, V. Azadehgan, "Applying Critical Edge to Test Data Generator.," *Software Engineering Research and Practice*, pp. 266-269, 2010.

# CHAPTER 1

# Introduction

Digital radio-frequency (RF) transmitters (TX's) have gained popularity in recent years. The realization of transmit functions in the digital domain offers many advantages: (1) analog blocks, such as baseband filters, variable-gain amplifiers, offset-cancellation digital-to-analog converters (DACs), upconversion mixers, and predrivers, are omitted, thus contributing no noise in the receive-band, (2) carrier feedthrough and quadrature mismatches can be suppressed in the digital domain, (3) RF power control is readily implemented within the output DAC, and (4) while an analog transmitter must deal with nonlinearity-gain trade-offs at all points in the chain, a digital transmitter contains only a single analog port at the RF DAC output.

The greatest challenge facing RF transmitters, analog or digital, is the trade-off between linearity and efficiency, which in turn has led to many linearization techniques. Since the die temperature varies considerably with the TX output power, the linearization must continue in real time; i.e., foreground calibration techniques lose their efficacy if they attempt to correct a highly nonlinear output stage.

This dissertation introduces a new approach to TX linearization that corrects for both static and dynamic nonlinearity in the background. The correction's efficacy allows designing the DAC for maximum efficiency with almost arbitrary integral nonlinearity (INL). Targeting the wideband code-division multiple access (WCDMA) standard as an example, the simple, compact architecture affords the highest efficiency reported to date. Realized in 28-nm standard CMOS technology, the complete transmitter delivers +24.1 dBm with an adjacent channel power ratio (ACPR) of $-35.4$ dB and an overall efficiency of 50%.

The next section provides a brief background on nonlinearity calibration, and Section 1.2 deals

with the performance requirements. Section 3.1 describes the basic idea, and Section 3.2 presents the evolution of the TX architecture. Section 4 describes the design of the building blocks, and Section 5 summarizes the experimental results.

## 1.1 Background

In the past 15 years, extensive work has been dedicated to digital transmitters [2]-[8]. As expected, the RF DAC has been the focus of these developments since it limits the overall TX performance. Figure 1.1(a) shows an architecture example [9], where the baseband quadrature signals, $I_{in}$ and $Q_{in}$, are applied to a digital predistortion (DPD) block before reaching the RF DACs. The DPD can be viewed as the inverse of the DAC's characteristic. The DACs are clocked by the local oscillator (LO) phases, $LO_I$ and $LO_Q$, so as to upconvert the baseband signals. Figure 1.1(b) depicts a unit cell of the DACs, where $D_j$ denotes the cell's digital input [9]. In order to maximize the power efficiency, the transistors in the cell act as switches with a low on-resistance. The two DACs thus suffer from a low output resistance, affecting each other's signals. This interaction between the $I$ and $Q$ DACs requires a two-dimensional (2D) polynomial correction [9], hence the need for the 2D look-up table (LUT) in Fig. 1.1(a). The situation becomes even more challenging in the presence of dynamic nonlinearities, calling for "complex predistortion" (delayed polynomials) [10].

In addition to the complexity of the look-up tables, conventional digital transmitters also suffer from the drift of the DAC nonlinearity with the temperature. As an example, Fig. 1.2 plots the integral nonlinearity of our DAC (Section 3.2) at 0 °C and 75 °C, revealing a change of several percent and underscoring the need for background nonlinearity correction. For example, the power control loop between a WCDMA base station and a mobile user can adjust the mobile TX output power by 60 to 70 dB according to the path loss between the two. As a result, the TX die experiences a substantial temperature change during communication.

Prior work on background calibration of analog transmitters relies on FPGA-based or Matlab-based correction with off-the-shelf components [11], [12]. Also, the system in [12] takes half a second to adapt to new conditions, and corrects for only static nonlinearity. To our knowledge, no

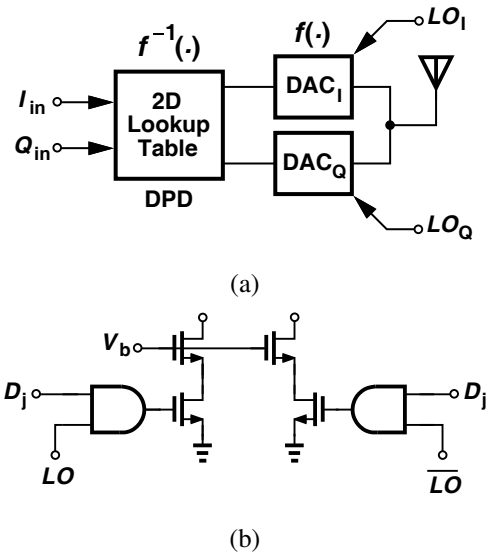



Figure 1.1. (a) Simplified TX architecture in [9], and (b) the DAC unit cell.

background nonlinearity correction has been reported for digital transmitters.

## 1.2 Performance Requirements

Before presenting our work, we summarize the performance requirements that a generic digital transmitter would need to satisfy. As an example, we consider the WCDMA specifications [13]: data rate: 3.84 Mcps; channel bandwidth: 5 MHz; carrier frequency: 1920 to 1980 MHz (band I); peak output power: +24 dBm (for "class C" as defined by the standard); ACPR at this power level: 33 dB; ACPR for alternate adjacent channel ($\mathrm{ACPR}_2$): 43 dB; output noise in the receive band: -125 dBm/Hz.

The RF DAC output quality is dictated by the ACPR [or the error vector magnitude (EVM)] and the receive-band noise (RXBN), the former imposing a maximum INL of 6.5% and the latter requiring a resolution of 13 bits and an output thermal noise floor below $-125$ dBm/Hz. These bounds must be met at a clock rate of 2 GHz for WCDMA with acceptably small differential nonlinearity (DNL) and output glitches, both of which tend to raise the noise floor.

A digital TX employing background calibration can downconvert the RF DAC output to baseband signals, digitize these signals, and provide real-time feedback. The analog-to-digital convert-

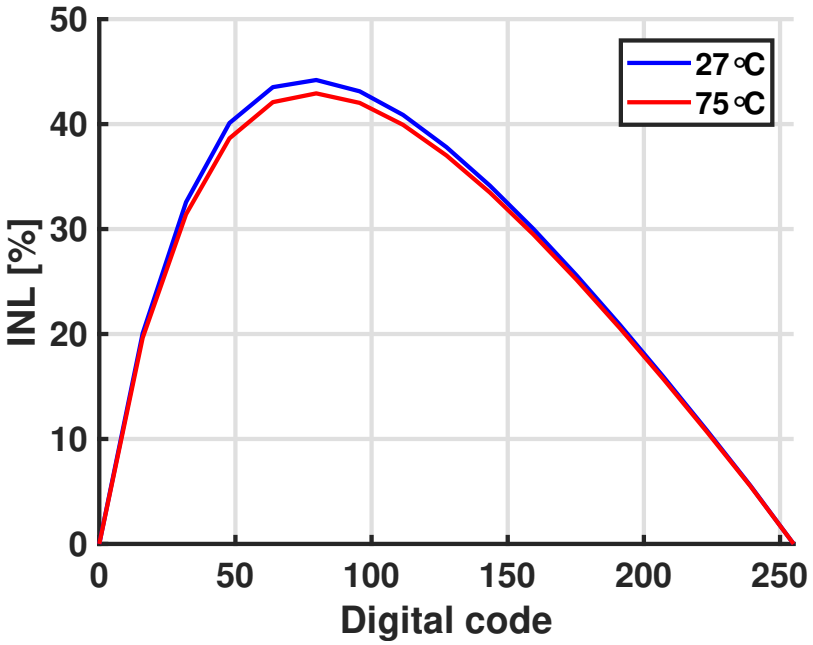



Figure 1.2. DAC INL as a function of input.

ers (ADCs) necessary in the loop must also satisfy certain conditions so as to negligibly corrupt the downconverted signals. We expect that the necessary INL and resolution of the ADCs are similar to those of the RF DAC.

# CHAPTER 2

# Predistortion Method

In this chapter, we study different approaches in modeling power amplifiers for predistortion purposes. Some of the data in this chapter are brought from [1].

## 2.1 Physical Models

Physical models are derived based on KVL, KCL, and equations governing different components in the circuit such as differential equations between inductors' and capacitors' voltage and current. Transistors as nonlinear components have sophisticated models; therefore, this makes the overall physical model of a Digital PA very complicated. We prefer simplest models since they will be implemented on-chip. Hence, physical models are not good candidates.

## 2.2 Empirical Models

Empirical models are classified into three categories and are listed in order of their complexity.

- Memoryless nonlinear models or static nonlinear models
- Nonlinear models with memory or dynamic nonlinear models
- Nonlinear models with non-uniform tap memory

In the subsequent sections, different static and dynamic models are introduced and compared. These models are tested using an off-the-shelf class-AB power amplifier. The input signal is WCDMA whose power is 5 dB more than the input referred 1-dB compression point of the power

amplifier. In other words,

$$P_{in,avg} = P_{1\ dB} + 5\ dB \tag{2.1}$$

This assures that the power amplifier is heavily saturated and it has a nonlinear model.

The comparison among different models is based on two metrics defined as follows.

1. Error in ACPR prediction or $\Delta$ACPR

   $\Delta$ACPR is the difference between the measured ACPR of the PA under test and the predicted ACPR by the model. That is to say,

   $$\Delta\mathrm{ACPR} = \mathrm{ACPR}_{\mathrm{measured}} - \mathrm{ACPR}_{\mathrm{model}} \tag{2.2}$$

   where

   $$\mathrm{ACPR} = \frac{\int_{adjacent\ channel} |Y(f)|^2 df)}{\int_{main\ channel} |Y(f)|^2 df} \tag{2.3}$$

   where $Y(f)$ is the output spectrum.

2. Normalized Mean-Square Error or NMSE

   The normalized mean-square error is defined as follows.

   $$\mathrm{NMSE} = \frac{\sum |y_{meas}(i) - y_{model}(i)|^2}{\sum |y_{meas}(i)|^2} \tag{2.4}$$

   where y represents the output samples.

### 2.2.1 Static Nonlinear Models

In static nonlinear models, the output is only a function of the current input samples whereas dynamic nonlinear models represent a relation based on the current and past input samples. Polynomial model is one the most used static nonlinear models. The polynomial order depends on how nonlinear the Digital PA is. Another static nonlinear model is "Saleh", which is used to be employed for Traveling Wave Tube Amplifiers (TWTAs). Nowadays, a modified version of Sale model is used for solid-state PAs (SSPAs).

Table 2.1. Performance of the polynomial model

| Polynomial order | NMSE (dB) | $\Delta$ACPR |
|---|---|---|
| 5 | -30.9 | 1.5 |
| 13 | -33.4 | 0.9 |

#### 2.2.1 1 Polynomial Model

The first model to examine is polynomial model. Since power amplifiers are usually differential, we eliminate the even order sentences. The coefficients of the polynomial, $\alpha_1$, $\alpha_3$, ... are found by curve fitting based the measured data. Based on these coefficients, the ACPR is predicted and compared with the measured ACPR.

$$y(t) = \alpha_1 x(t) + \alpha_1 x(t)^3 + \cdots \tag{2.5}$$

For this special system, having fifth order polynomial is enough as no significant improvement is observable when the order is raised to 13.

#### 2.2.1 2 Saleh Model

Prior to explaining the Saleh model, AM/AM and AM/PM conversion concepts should be clarified.

Suppose $x(t) = r(t)\cos(\omega_0 t)$ is input of the power amplifier. Then, the output will look like $y(t) = g(r)\cos(\omega_0 t + \phi + \Phi(r))$. In this relation, $g(r)$ is called AM/AM conversion function which represents how the envelope will be affected when the input signal passes through the nonlinear power amplifier. Moreover, $\Phi(r)$ denotes the AM/PM conversion function and implies how the input envelope affects the phase of the output signal.

In spite of the polynomial model, Saleh model defines relations for AM/AM and AM/PM conversions. The general formula for these relations is,

$$(\alpha r^{\eta})/(1 + \beta r^2)^{\nu} \tag{2.6}$$

where $\eta$ and $\nu$ will have different values for AM/AM and AM/PM conversions. For the power

amplifier under test,

$$g(r) = \alpha r/(1 + \beta r^2) \tag{2.7}$$

$$\Phi(r) = (\alpha r^2)/(1 + \beta r^2) \tag{2.8}$$

are assumed. We can observe that the performance of the Saleh model is poor.

Table 2.2. Performance of Saleh model

| Model | NMSE (dB) | $\Delta$ACPR |
|---|---|---|
| Saleh | -8.3 | 20.2 |
| Saleh (AM/AM only) | -27.5 | 1.3 |

The reason of the poor performance is that this model is designed for TWTAs whose AM/PM conversion function does not fit solid-state power amplifiers. If only the AM/AM conversion is considered, the result improves. Poor fitness of AM/PM function is clearly shown in Fig. 2.1.

It can be seen that AM/PM curve of TWTA (with the right y-axis) does not cross zero except at origin. As expected, $\Phi(r) = (\alpha r^\eta)/(1 + \beta r^2)^\nu$ does not have zero except at origin. In contrast, AM/PM curve of LDMOS (with the left y-axis) crosses zero at the input amplitude of one so when AM/PM Saleh model is fitted, poor performance is achieved. To amend this model to be usable for solid-state power amplifiers, a constant $\epsilon$ is subtracted from the Saleh AM/PM model.

$$\Phi(r) = \frac{\alpha r^\eta}{(1 + \beta r^2)^\nu} - \epsilon \tag{2.9}$$

The performance for the modified Saleh model is shown in Table 2.3.

Table 2.3. Performance of the modified Saleh model

| Model | NMSE (dB) | $\Delta$ACPR |
|---|---|---|
| Saleh | -8.3 | 20.2 |
| Modified Saleh | -32 | 0.9 |

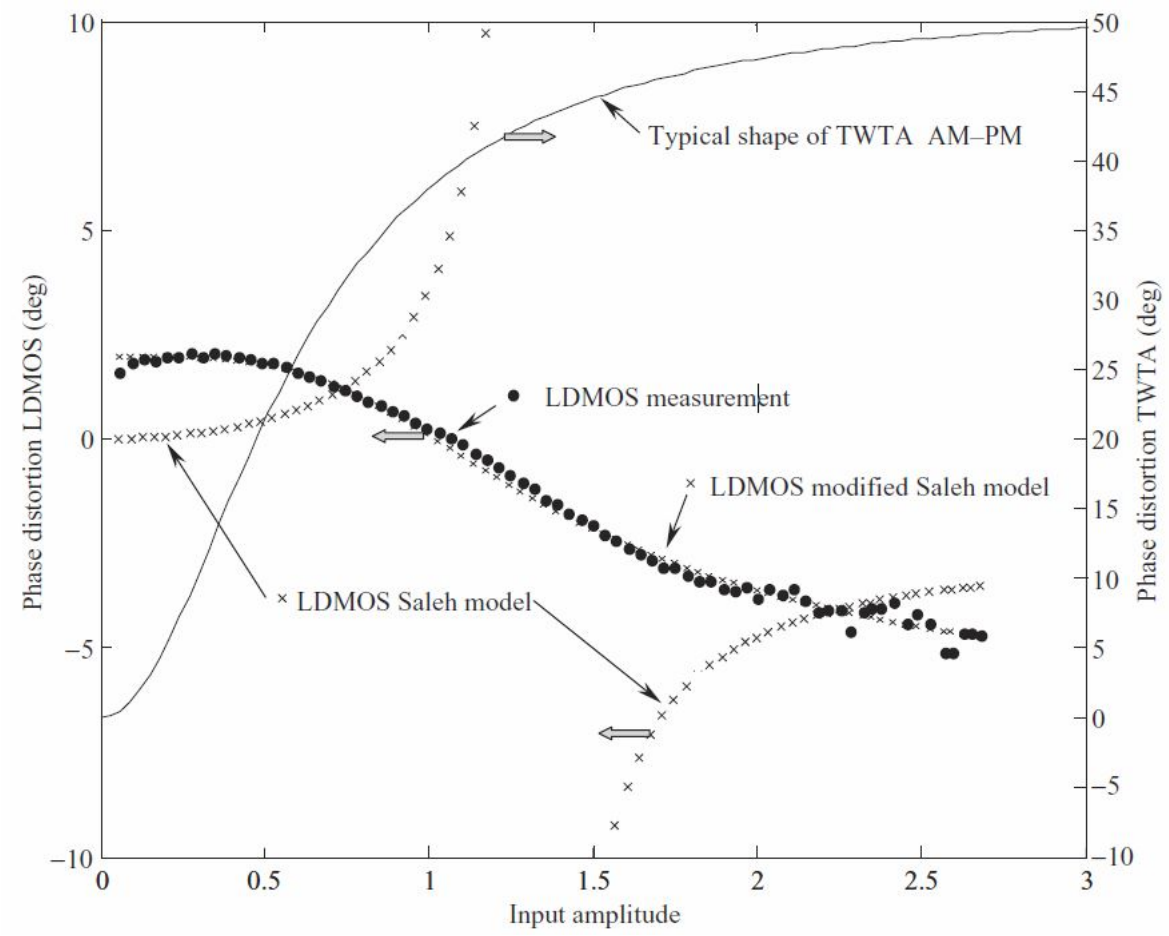



Figure 2.1. Phase distortion for TWTA and SSPA.

### 2.2.1 3 Berman and Mahle

Berman and Mahle proposed a relation solely for AM/PM distortion. Therefore, this model comes with another function for modeling the AM/AM distortion. The performance of this model along with other models is shown in Table 2.4.

Comparing performance and complexity of different models, we can conclude that among these static models, polynomial is preferable because of its simplicity in implementation and good performance.

Table 2.4. Performance of all the aforementioned models

| Model | NMSE (dB) | ΔACPR |
|---|---|---|
| 5th order polynomial | -30.9 | 1.5 |
| 13th order polynomial | -33.4 | 0.9 |
| Saleh | -8.3 | 20.2 |
| Saleh (AM/AM only) | -27.5 | 1.3 |
| Modified Saleh | -32 | 0.9 |
| AM/PM: Berman & Mahle<br>AM/AM: Power series | -32.8 | 0.6 |

### 2.2.2 Dynamic Models

The simplest and the most general dynamic model is two-box model. In this model, memory effect and nonlinearity are decomposed. For the memory effect, a simple FIR filter is employed while for nonlinearity, any static model can be used. If the nonlinear box follows the linear one, it is called Wiener model. On the other hand, if the linear box follows the nonlinear one, we call it Hammerstein model.

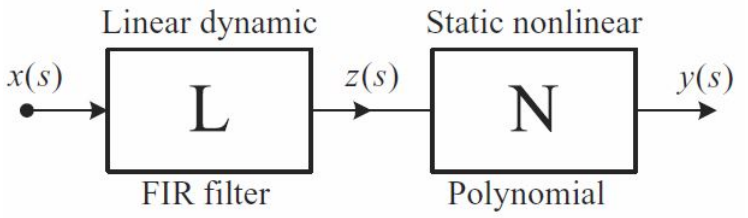



Figure 2.2. Wiener model.

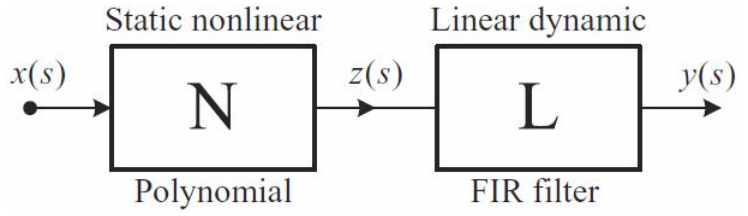



Figure 2.3. Hammerstein model.

Also, there are some more dynamic models such as Chebyshev and Hermite ones which are not appropriate for on-chip implementation due to their complexity.

## 2.3 Polynomial Implementation

To model the nonlinearity, the polynomial model is used. Polynomial is the most commonly used model and can be employed for any analytical function. Since the digital power amplifier model and its inverse are clearly analytical, polynomial model is a good option. On the other hand, polynomials can be easily implemented using Horner's method. Therefore, it suits on-chip implementation for a nonlinear block.

$$y_I[n] = \sum_k \alpha_k x_I[n]^k + \sum_{i,k} \beta_k x_Q[n]^k \tag{2.10}$$

$$y_Q[n] = \sum_k \gamma_k x_I[n]^k + \sum_{i,k} \zeta_k x_Q[n]^k \tag{2.11}$$

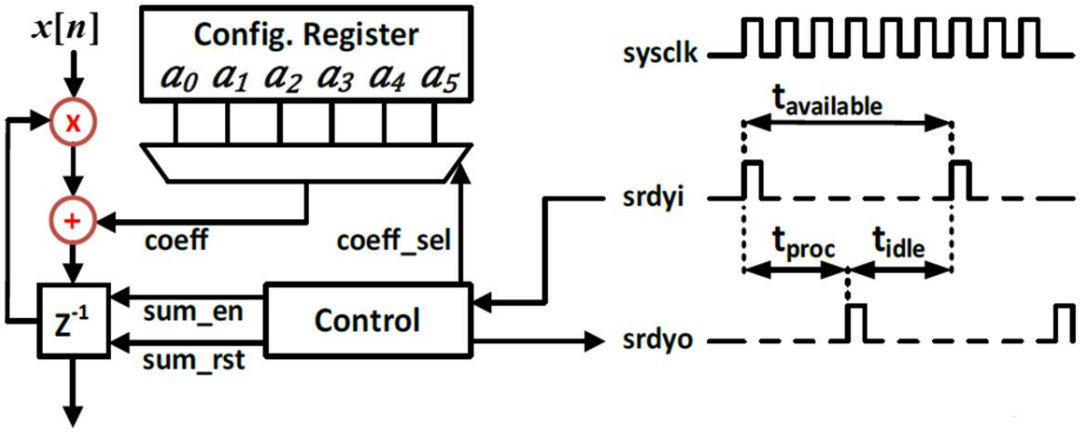



Figure 2.4. Horner's method.

$$\begin{aligned} y[n] &= \alpha_0 + \alpha_1 x + \alpha_2 x^2 + \alpha_3 x^3 + \cdots + \alpha_n x^n \\ &= \alpha_0 + x(\alpha_1 + x(\alpha_2 + (\alpha_3 + \cdots + x(\alpha_{n-1} + x\alpha_n)))) \end{aligned} \tag{2.12}$$

Figure 2.4 shows the Horner's method. This loop only uses one multiplier and one adder. $x[n]$ circulates in this loop and is multiplied by itself to generate all of the polynomial sentences. To

implement inverse of the digital power amplifier, $Y_I$ and $Y_Q$ must be replaced with $I_{in}$ and $Q_{in}$, respectively. Similarly, $X_I$ and $X_Q$ need to be replaced with $I_{out}$ and $Q_{out}$, respectively.

$$I_{in} \rightarrow y_I$$

$$Q_{in} \rightarrow y_q$$

$$I_{out} \rightarrow x_I$$

$$Q_{out} \rightarrow x_Q$$

## 2.4 Finding Coefficients Using Adaptation

Digital predistortion implements a mathematical relation such as polynomial, and the coefficients in this relation are unknown. After finding them, it is desirable to keep the coefficients updated on the fly since the temperature or any other condition may change. To achieve this goal, adaptive algorithms are the best candidates.

Adaptive algorithms are used to estimate the inverse model of a power amplifier. They collect the input and output samples of the power amplifier and estimate its inverse model. These algorithms are classified into two categories. One is Stochastic Gradient family and the other one is Recursive Least-Square (RLS) family.

Stochastic Gradient family includes Normal/Leaky/sign LMS, among which, sign-LMS is the simplest one. RLS family includes the RLS algorithm itself and its derivatives. As this adaptation is aimed to be on chip, the hardware complexity will be one of the major concerns. Therefore, it is reasonable to choose LMS family for our adaptation as it needs less hardware to implement.

To use any adaptive algorithm, first we need to assume a model for the digital predistorter that can be perfectly fit. This model should be chosen based on the behavior of the digital power amplifier. For example, if we want to model a system which compresses as the input grows, it's not reasonable to choose a model (such as an exponential one with positive exponent) which expands as the input rises. After choosing a proper model for the digital predistorter, we have to reshape the

model's relation in the form of $y = XA$ where y is the output sample, $X$ is a vector based on the input samples, and $A$ is the vector of coefficients which we wish to find. Adaptation is an iterative procedure through which, the vector $A$ shall be found. Every iteration contains 3 steps as follow.

$$\begin{aligned} \hat{y} &= X_n A_n \\ e_n &= y_n - \hat{y}_n = y_n - X_n A_n \\ A_{n+1} &= A_n + \mu X_n^T e_n = A_n + \mu X_n^T (y_n - X_n A_n) \end{aligned} \tag{2.13}$$

In this notation, $\hat{y}_n$ is the polynomial model output (model's prediction), and $y_n$ is the measurement power amplifier output. In the first two steps, the error of the model is found and in the third step, a new $A$ is calculated based on the old $A$ and the error, $e$. We can tune how fast the adaptive algorithm converges by factor $\mu$. This factor should be between 0 and 1. As $\mu$ is raised, adaptive algorithm converges faster but with higher probability of disconvergence.

To better understand this adaptive algorithm, assume that the model accurately predicts the power amplifier behavior. That is to say, $e_n$ equals zero. Consequently, the third step will be simplified to $A_{n+1} = A_n$. Therefore, coefficients have already converged to the optimum values. This usually happens once the model becomes mature enough through iterative adaptation.

## 2.5 Transmitter's Performance

In this section, the performance of the designed digital power amplifier is discussed. Since this digital power amplifier is very nonlinear, we need to choose an appropriate predistorter model to linearize the transmitter. Therefore, a two-dimensional polynomial is used. This polynomial is found by interpolation according to Fig. 2.5.

By changing the input amplitude, the transmitter's output power, efficiency and linearity change. Figure 2.6 (and Fig. 2.7) shows a comparison between $I_{in}$ and $I_{out}$ (and $Q_{in}$ and $Q_{out}$) waveforms. In addition, Fig. 2.8 and 2.9 show the digital power amplifier's output spectrum and ACPR.

Peak to Average Power Ratio (PAPR) is an important factor that affects the digital power am-

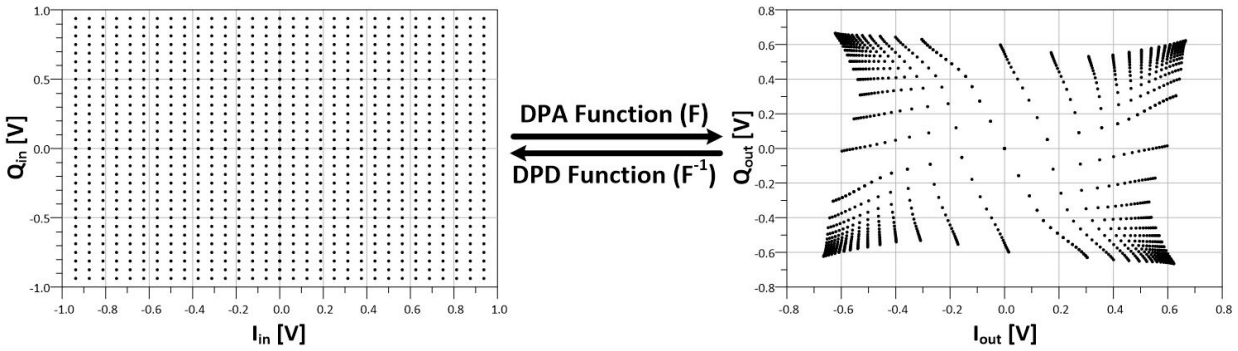



Figure 2.5. Applied input and the resulting output constellation for the digital power amplifier.

plifier efficiency.

$$\mathrm{PAPR} = \frac{\mathrm{P_{out,max}(t)}}{\mathrm{P_{out,avg}(t)}} \tag{2.14}$$

Assuming a constant maximum transmission power level, as PAPR increases, the average output power decreases, causing a decrease in efficiency. On the other hand, if constant average output power is assumed, as PAPR increases, the maximum transmission power level increases, resulting in worse ACPR. Therefore, achieving acceptable performance with higher PAPR is more difficult. The PAPR for this WCDMA standard is 3.3 dB. As it can be seen, with such a high-order polynomial, we barely meet WCDMA mask. This polynomial can potentially consume so much power that drastically affect the transmitter's efficiency. In the next section, we explain our proposed linearization method that has simpler hardware than a predistorter but is stronger in linearizing the digital power amplifier.

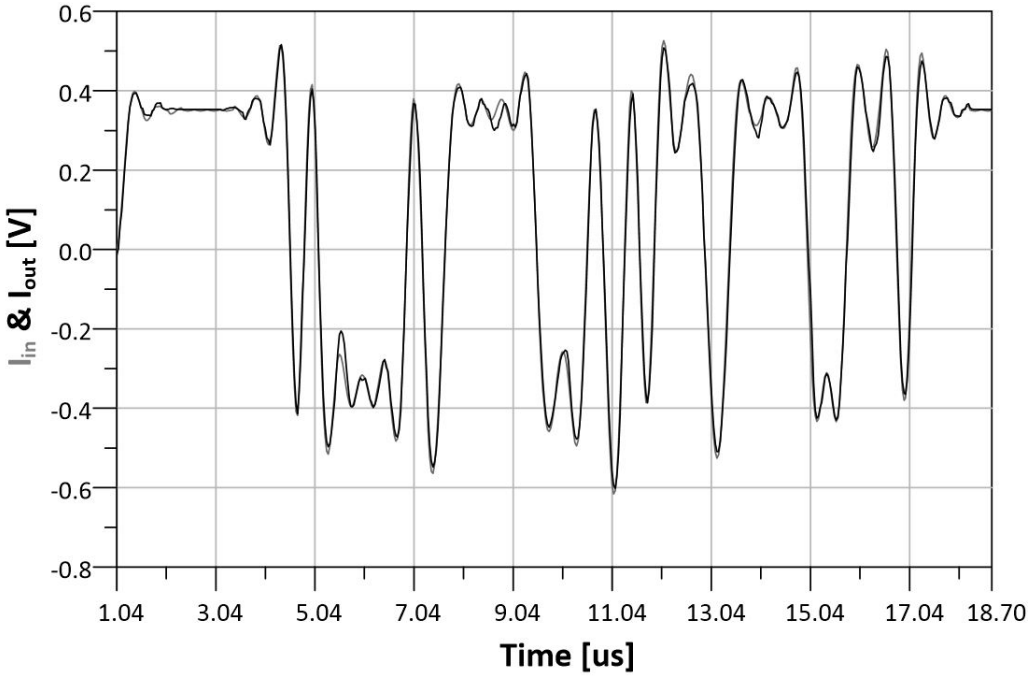



Figure 2.6. Input (gray) and output (black) in-phase waveforms.

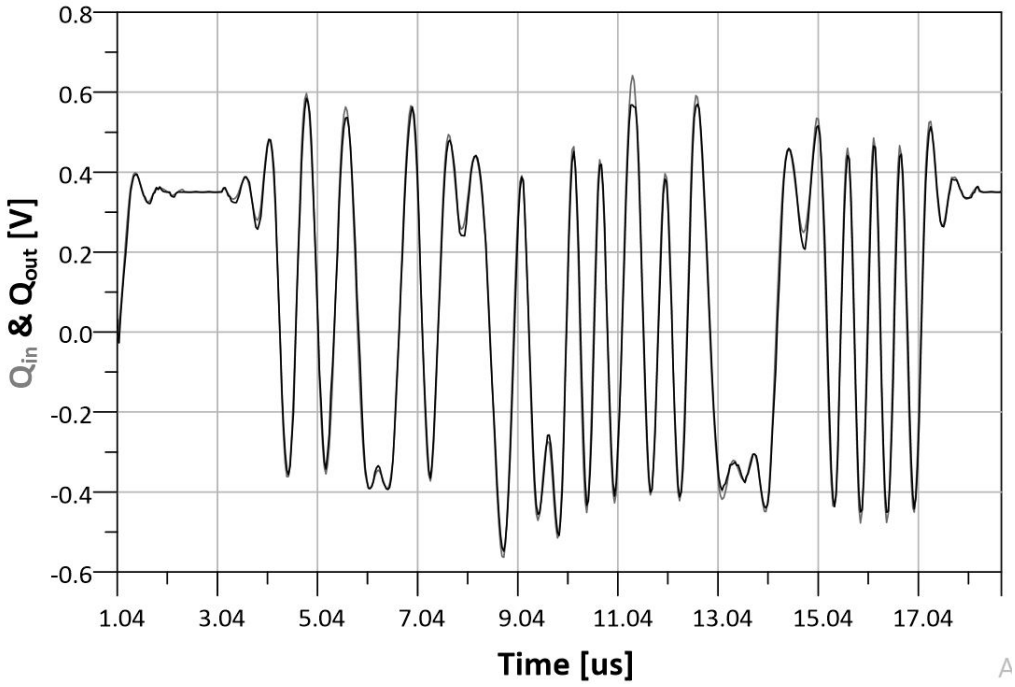



Figure 2.7. Input (gray) and output (black) quadrature waveforms.

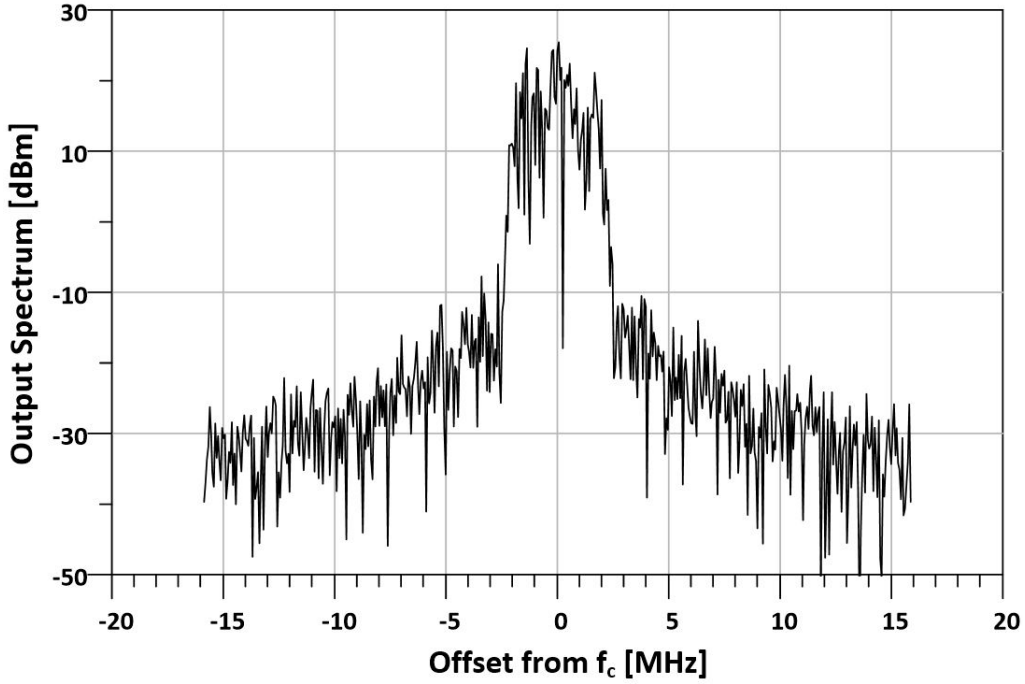



Figure 2.8. Output spectrum around carrier frequency.

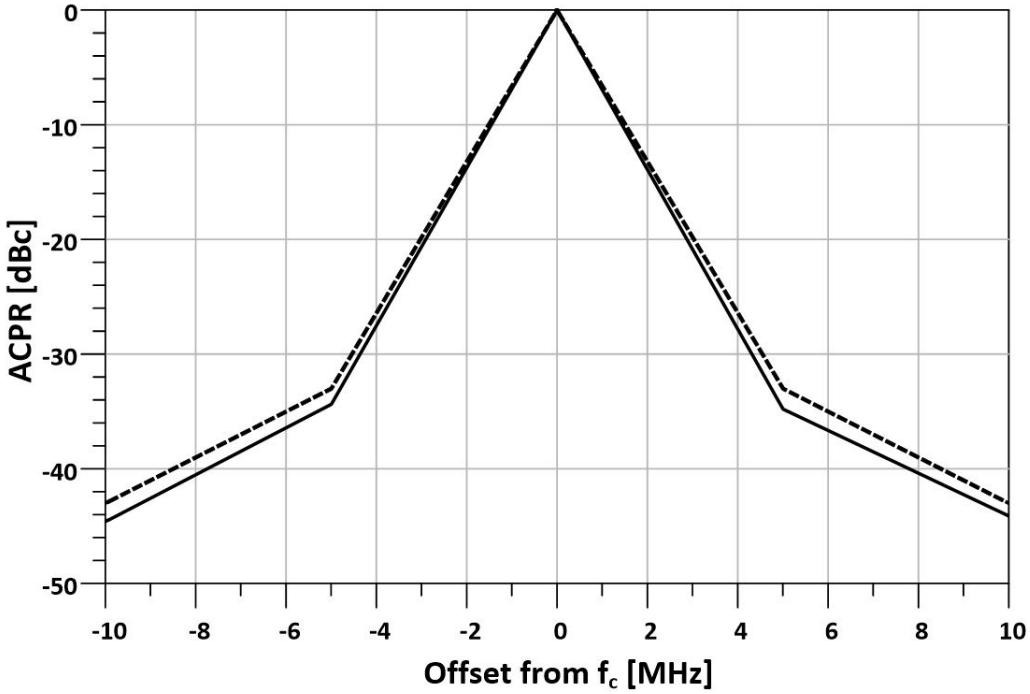



Figure 2.9. Required ACPR (dashed) and achieved ACPR (solid).

# CHAPTER 3

# Proposed Linearization Technique

## 3.1 Basic Idea

Figure 3.1(a) shows a DAC that, for now, is assumed to have only static nonlinearity, exhibiting an input-output characteristic $f(x)$. To linearize the DAC, we wish to precede it with a block whose behavior is to be determined [Fig. 3.1(b)]. Here, $w$ is the main input and eventually represents the TX baseband data. In conventional predistortion, the first block approximates $f^{-1}(\cdot)$ by a polynomial of the form $\alpha_1 w + \alpha_2 w^2 + ... + \alpha_n w^n$ with the coefficients $\alpha_j$ selected so as to make $y = f(x)$ a faithful replica of $w$. Equivalently, $f^{-1}(\cdot)$ maps each value of $w$ uniquely and statically to a value of $x$ once the coefficients are frozen.

Let us approach the problem from a different perspective by focusing on $x$. For $y = f(x)$ in Fig. 3.1(b) to become equal to $w$, we must force $f(x) - w$ to zero, where both $f(\cdot)$ and $x$ are known. We denote this difference (the "error function") by $g(x)$. For every value of the input, $w$, we wish to choose $x$ so that $g(x) \to 0$, i.e., so that $x$ is a root of $g(x)$. To ensure that $x$ satisfies $g(x) = 0$, we can utilize any equation solver, e.g., the Newton-Raphson technique. That is, to solve $g(x) = 0$, we iteratively select

$$x_{n+1} = x_n - \frac{g(x_n)}{g'(x_n)}. \tag{3.1}$$

The iteration relies on memory in the system, introducing a dynamic behavior. In our case, $g'(x_n) \approx f'(x_n)$ if $w$ changes slowly with time.

Before reducing these seemingly abstract concepts to practice, we make two remarks. First, the Newton-Raphson iteration must occur fast enough to keep up with the dynamics of the baseband signal, $w$. Second, the computation of $g(x_n)/g'(x_n)$ must be managed such that it can be realized

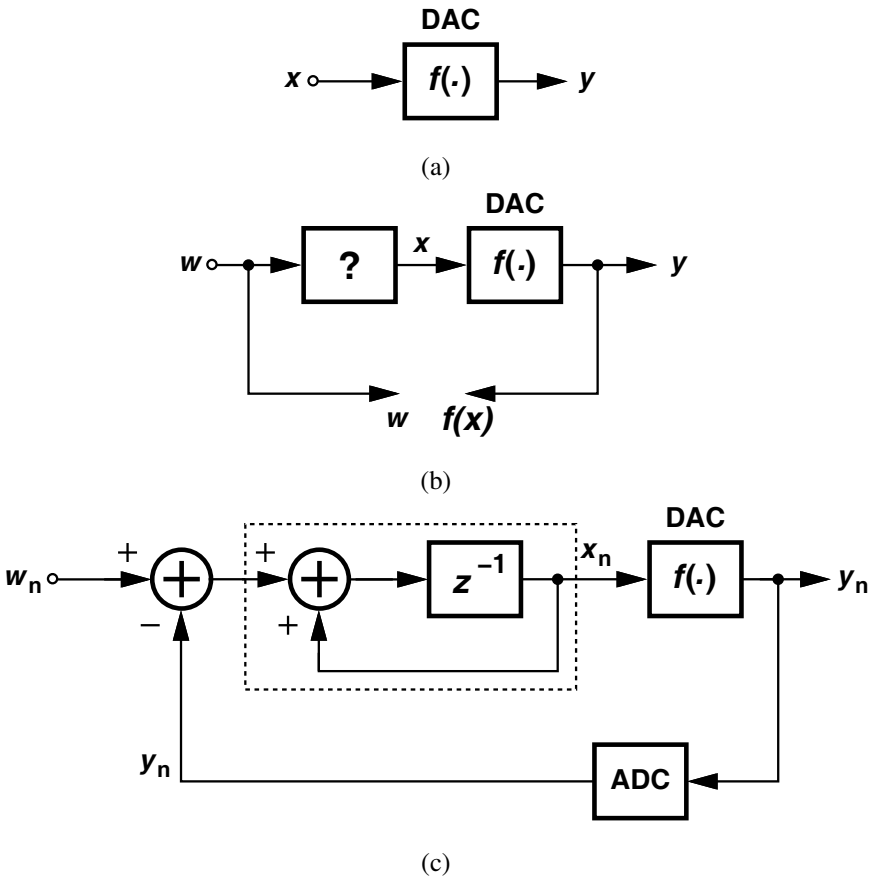



Figure 3.1. (a) Nonlinear DAC, (b) linearization by predistortion, and (c) linearization by $\Delta\Sigma$ loop.

efficiently. In this regard, let us, for now, make a rather coarse approximation and write $g'(x_n) \approx 1$. Since $f'(x_n) = g'(x_n)$, we have

$$\begin{aligned} x_{n+1} &= x_n - [f(x_n) - w_n] \\ &= x_n - (y_n - w_n). \end{aligned} \tag{3.2}$$

This result implies that the next value of $x$ can be obtained by subtracting the present error, $y_n - w_n$, from the present value of $x$, leading to the implementation shown in Fig. 3.1(c). Here, the one-cycle delay, $z^{-1}$, senses $x_n + (w_n - y_n)$ and generates $x_{n+1}$. Interestingly, the function within the dashed box is simply an integrator, thereby revealing that the overall system acts as a first-order $\Delta\Sigma$ modulator (DSM). That is, the DSM is a "poor man's" realization of the Newton-Raphson technique.

In retrospect, we could have directly conceived this idea: by placing the transmitter in a

$\Delta\Sigma$ loop, we can suppress the static and dynamic imperfections of the RF DAC. If the DAC errors are viewed as components appearing at its output, the high loop gain provided by the integrator substantially reduces them within some bandwidth. Nevertheless, our original idea, Eq. (3.1), in fact proves more powerful and allows us to further refine the transmitter architecture. Specifically, we will approximate $1/g'(x_n)$ by a function that readily lends itself to hardware implementation, thus reducing the receive-band noise (Section 3.2).

## 3.2 Proposed Transmitter Architecture

From the developments in the previous section emerges the conceptual TX architecture depicted in Fig. 3.2. Here, the $I$ and $Q$ paths and their integrators form two $\Delta\Sigma$ modulator loops. The TX output is downconverted, digitized, and subtracted from the input baseband signals, $I_{in}$ and $Q_{in}$. Unlike digital predistortion techniques, this architecture requires no look-up tables, digital multipliers, or finite impulse response (FIR) filters.

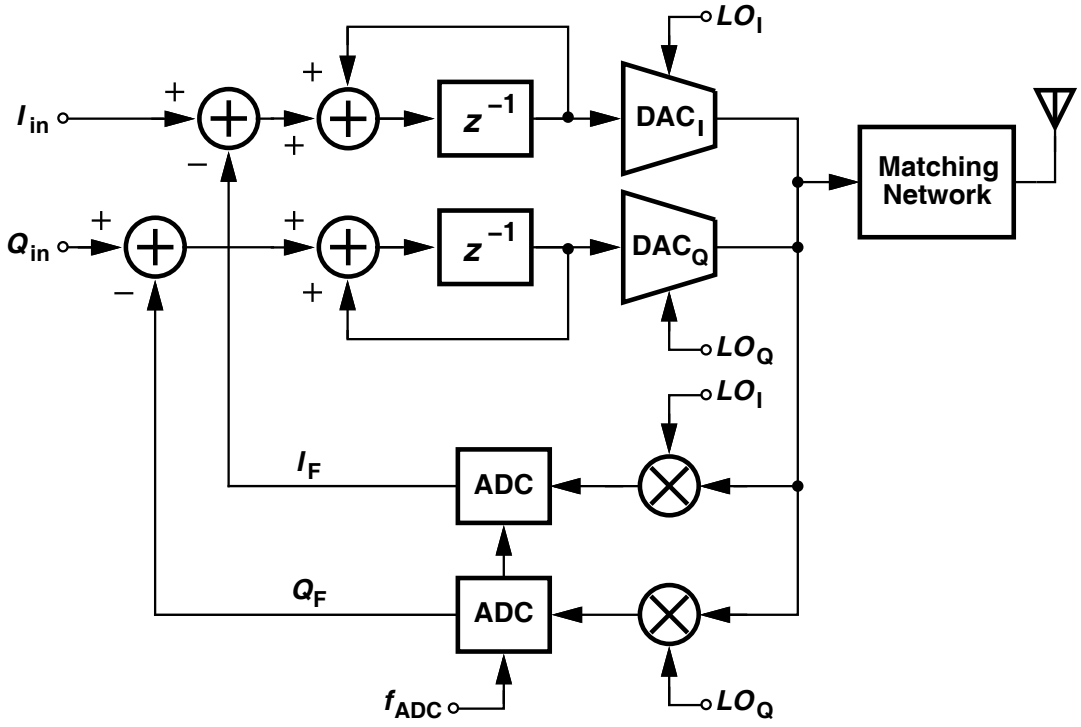



Figure 3.2. Conceptual TX architecture

### 3.2.1 Architecture Evolution

In this section, we describe a multitude of techniques that, step by step, simplify this transmitter, eventually offering a compact, efficient architecture. The final signal processing machine preceding the RF DAC contains only 512 flip flops.

In the first step of architecture evolution, we note the DACs in Fig. 3.2 must be configured as a segmented topology (using nominally equal units) and be driven by a thermometer code. Segmentation proves essential here as it minimizes both the DNL and the output glitch energy [14]. We must then decide whether, in Fig. 3.2, binary-to-thermometer decoders should be interposed between the integrators and the DACs or the integrators themselves should be so realized as to generate a thermometer-code input. Figure 3.3 depicts the situation for the $I$ path (the $Q$ path is similar). An important observation comes to our aid here: a simple shift register holding a

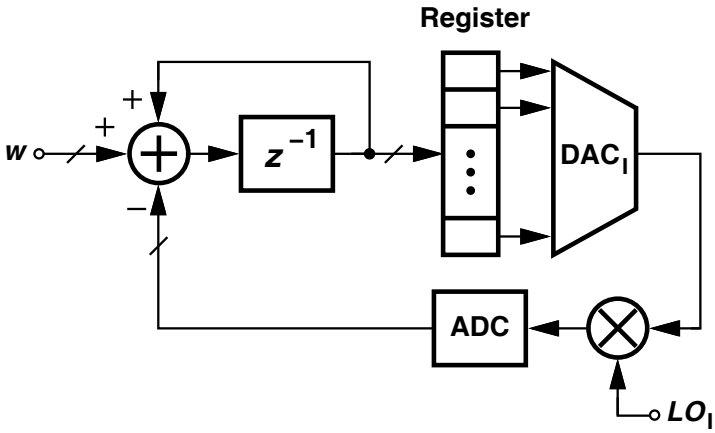



Figure 3.3. The $I$ branch showing the thermometer-code register

thermometer code can, in fact, act as an integrator if it receives a 1-bit input: when the input is $+1$, we shift the code up by one and insert another 1 at the bottom, and when the input is (equivalently) $-1$, we shift the code down by one. Thus, the explicit integrator in Fig. 3.3 can be omitted if the data is available in 1-bit form., i.e., as an oversampled sequence.

These thoughts lead to the conceptual arrangement shown in Fig. 3.4, where the baseband data is applied to a parallel-to-serial converter (e.g., a multiplexer) and the feedback signal is digitized by means of an oversampling ADC. We now have 1-bit representations in both paths. Note that

the overall TX feedback loop still operates as a $\Delta\Sigma$ modulator. Further simplification is possible if we recognize that the parallel-to-serial converter need not be a memoryless multiplexer and can alternatively be realized as an oversampling converter. For example, a digital $\Delta\Sigma$ modulator can convert the multi-bit data at $I_{in}$ to a 1-bit sequence. We can therefore "factor out" the oversampling modulators from the input and feedback paths and employ only one after the subtractor [Fig. 3.5].

The architecture illustrated in Fig. 3.5 merits five remarks. First, the system requires no adap-

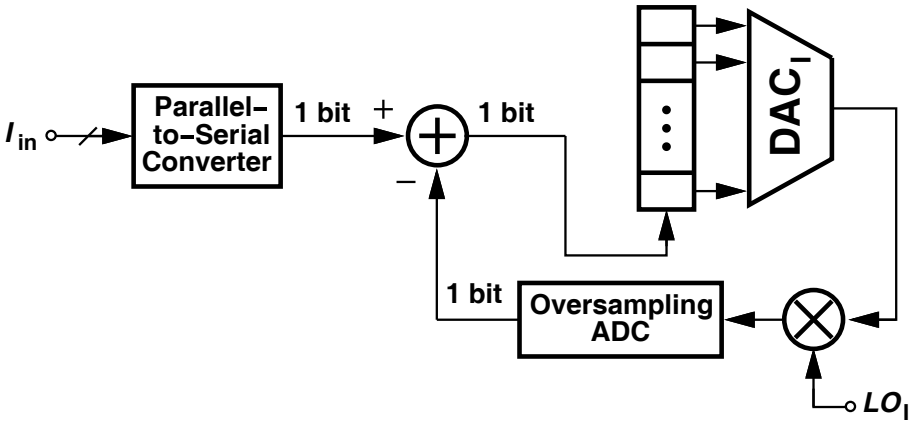



Figure 3.4. Use of register as integrator

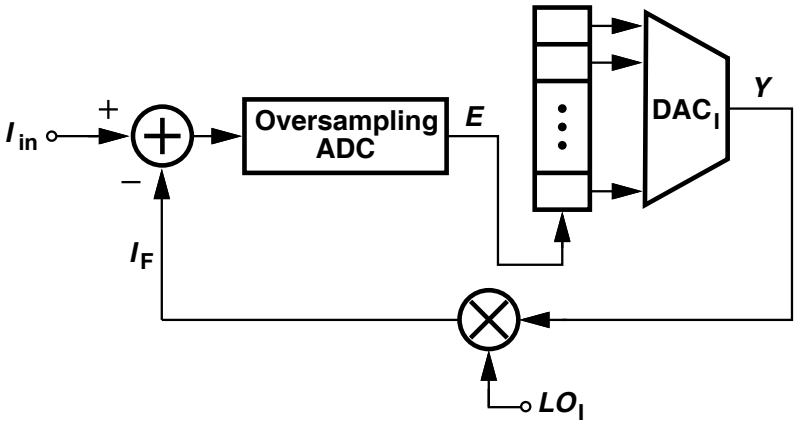



Figure 3.5. Simplified architecture

tation or training. After an initial settling time (about 50 ns in our work), the negative-feedback loop automatically corrects the DAC imperfections. Second, the oversampling converter digitizes the *difference* between $I_{in}$ and $I_F$ and hence its dynamic range can be much narrower than that of ADCs in Fig. 3.2. In our design, this transformation is equivalent to relaxing the ADC resolution

by about 7 bits. Third, since $I_F$ is an analog signal, so should be $I_{in}$, requiring proper implementation (Section 4). Fourth, the loop oversampling ratio must be high enough to suppress the DAC imperfections in the adjacent channels and in the RX band. Similarly, the ADC resolution and oversampling ratio must be chosen as to ensure acceptable quantization noise. Fifth in more demanding applications, the system in Fig. 3.5 can accommodate digital predistortion in the form of a look-up table interposed between the register and the DAC. One issue in the TX architecture of Fig. 3.5 is the low loop gain, an effect that can substantially degrade the performance. In Section 4, we explain the cause of this low gain and introduce a simple method of compensating it.

### 3.2.2 Choice of Loop Oversampling Ratio

We expect trade-offs among various parameters in the architecture of Fig. 3.5, e.g., the DAC resolution and raw nonlinearity, the loop oversampling clock rate, $f_{CK}$, and the ADC's oversampling ratio. Based on practical DAC design issues (Section 4), we assume for it a resolution of 8 bits and the INL profile depicted in Fig. 1.2. Note the very large nonlinearity. Higher resolutions lead to excessive complexity in the TX layout, and design for higher linearity degrades the power efficiency. We wish to determine the minimum acceptable $f_{CK}$. Our analysis employs a transistor-level model for the DAC and is performed in Agilent's ADS for its efficient frequency-domain analysis. For now, we assume the oversampling ADC has an infinite resolution.

Figure 3.6 plots the simulated TX output spectra for both open-loop and closed-loop operation. We observe that the latter exhibits an adjacent channel power of $-29$ dB, $-31$ dB, and $-33$ dB as $f_{CK}$ rises from 1 GHz to 2 GHz to 4 GHz, respectively. From simulations, we also arrive at the corresponding results for the RX band noise, obtaining $-110$ dBm/Hz, $-113$ dBm/Hz, and $-116$ dBm/Hz, respectively.

This analysis suggests that even with $f_{CK}$ = 4 GHz, we do not meet the WCDMA specifications of RXBN = $-125$ dBm/Hz. We must also revisit these results after we include the oversampling ADC's nonidealities.

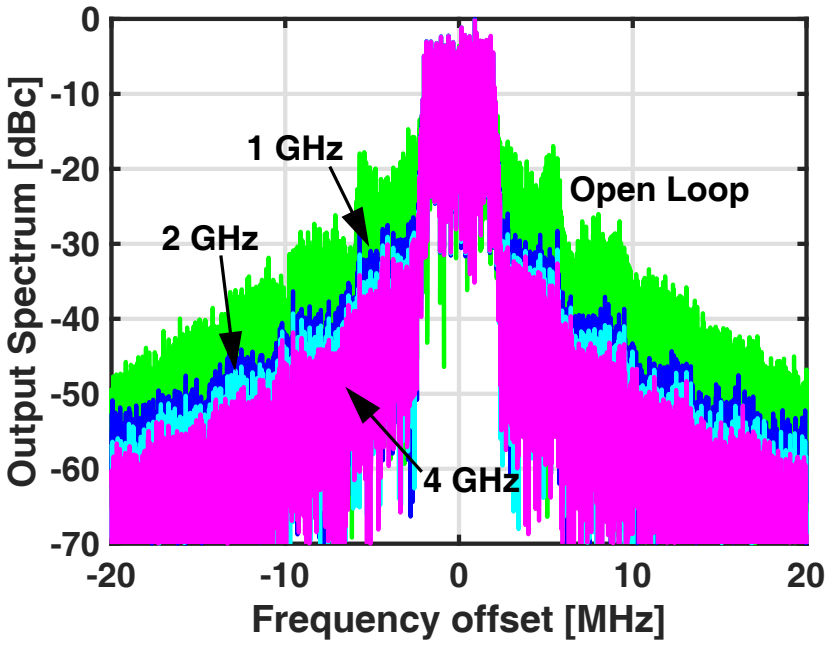



Figure 3.6. Output spectra for different values $f_{CK}$.

### 3.2.3 Architecture Refinement

In this section, we introduce a refinement to the TX architecture of Fig. 3.5 that substantially improves the performance, thus easing the design of the building blocks. Our focus is on the ACPR and the RX-band noise. Let us return to the Newton-Raphson method expressed by Eq. (3.1) and ask whether the approximation for $g'(x_n)$ can be improved from a constant value to a function that still lends itself to efficient implementation. The Newton-Raphson method requires that the error, $g(x) = f(x) - w$, be multiplied by $1/g'$ in each iteration, leading to the architecture depicted in Fig. 3.7(a). We factor out this coefficient and insert it in the input and output paths [Fig. 3.7(b)]. The output is now equal to $y/g'$ rather than the desired output, $y$, but if we multiply the main input by $g'$, the output changes back to $y$. That is, the $1/g'$ factor after $w$ should be removed.

We now turn to the $1/g'$ factor following the DAC and seek a hardware-efficient implementation for it. Specifically, we explore the possibility of merging the two. Denoting the transfer characteristic of the cascade by $P(x)$, we make two observations: (1) a 1-LSB increase in $x$ produces a change of $P(x+1\,\text{LSB})-P(x)$ at the TX output, which can be considered the derivative of $P$ with respect to $x$; and (2) according to simulations, the derivative of $P(x)$ behaves as shown in

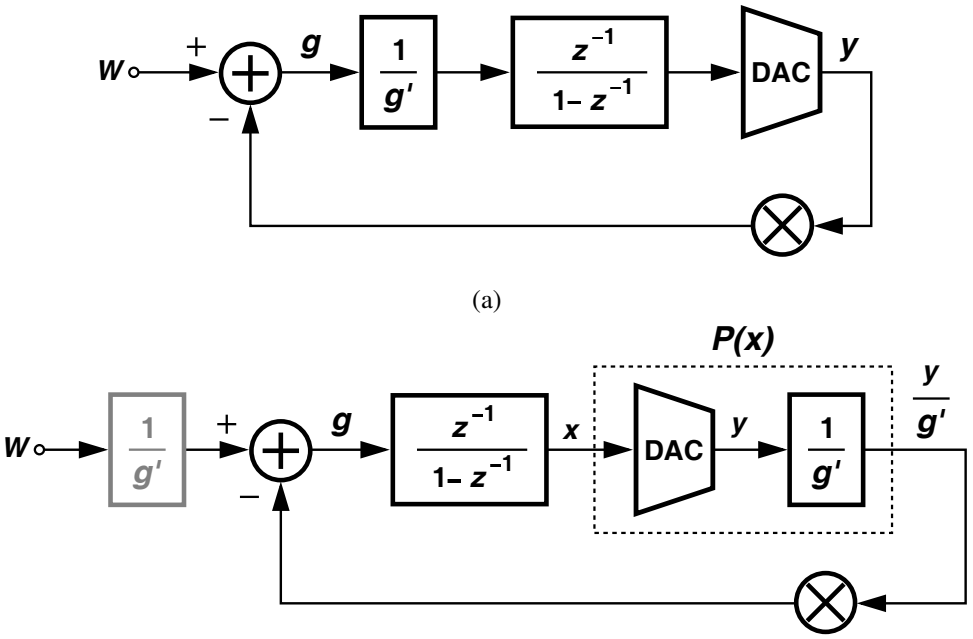



Figure 3.7. (a) Inclusion of $1/g'(x)$ to improve the accuracy of Newton-Raphson technique, and (b) transformation showing that $1/g'(x)$ can be included in the DAC.

Fig. 3.8. [For a given DAC design, $P(x) = f(x)/f'(x)$ is known.] Let us see how the $1/g'$ block can be absorbed by the DAC. The first observation made above must apply to the new DAC as well: in response to a 1-LSB increment in $x$, its output must change by an amount equal to $P'(x)$. This change is created by turning on one more DAC unit. With the $P'$ shape depicted in Fig. 3.8, we predict that the DAC output increment should be smaller for low $x$ values and larger for high $x$ values. Correspondingly, the DAC unit cells should be "weaker" for low $x$ values and "stronger" for high $x$ values. In principle, we can taper the units according to the shape of $P'(x)$, but the DAC design is greatly simplified if we approximate $P'(x)$ by a staircase function [Fig. 3.8]. Specifically, we choose scaling factors equal to 0.25, 0.375, 0.5, 0.75, 1, 2, and 4 for $1 \leq x \leq 64$, $64 < x \leq 80$, $80 < x \leq 160$, $160 < x \leq 192$, $192 < x \leq 208$, $208 < x \leq 240$, and $240 < x \leq 256$, respectively.

Derived from the Newton-Raphson technique, this free modification of the RF DAC reduces the ACPR by 4 dB and the RX-band noise by 3 dB. Figure 3.9 plots the TX output spectrum before and after nonuniform sizing of the DAC units.

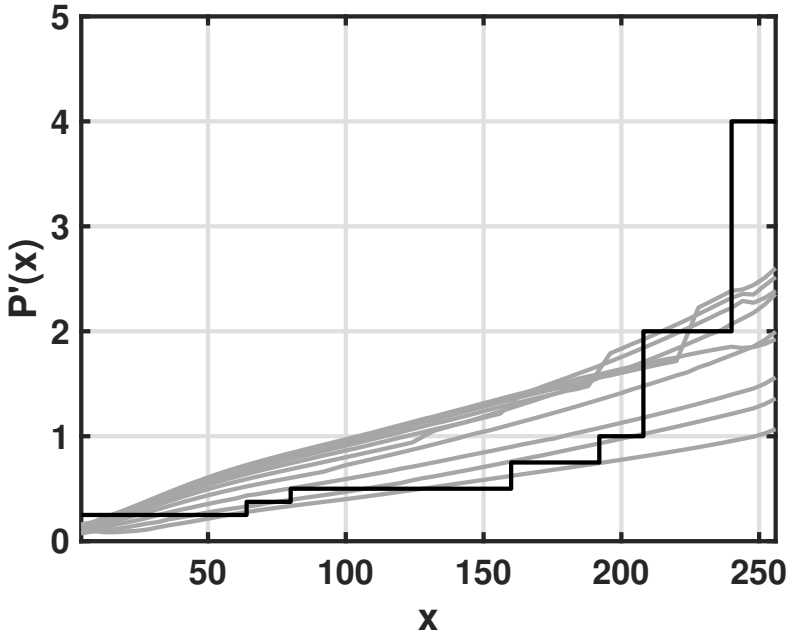



Figure 3.8. Behavior of $P'(x)$ and its staircase approximation

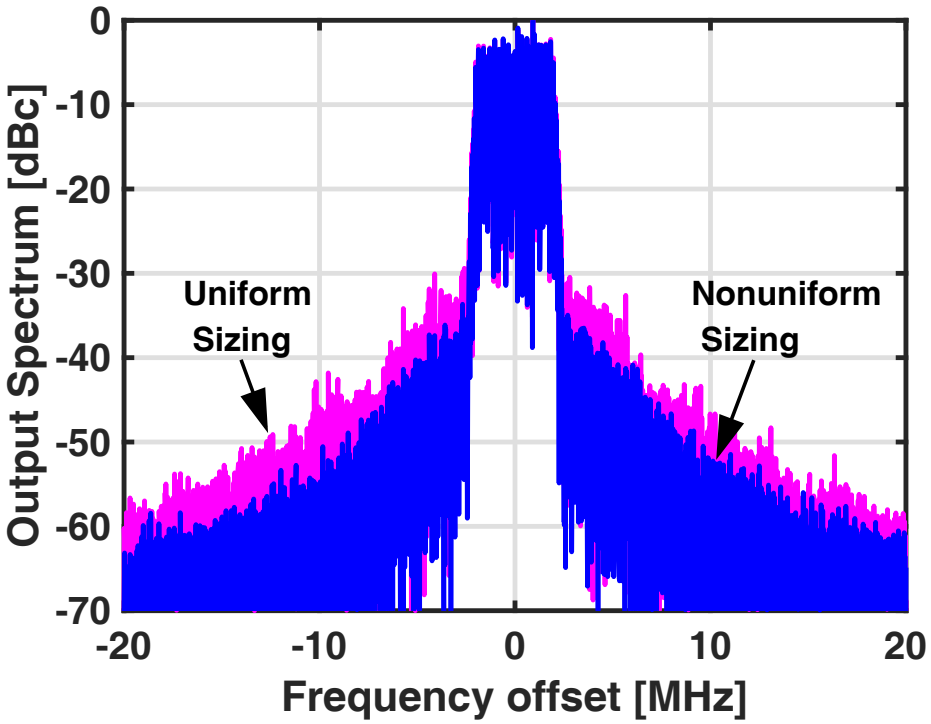



Figure 3.9. Simulated output spectra showing reduction of ACPR.

# CHAPTER 4

# Design of Building Blocks

In this chapter, we describe the design of the RF DACs, the oversampling ADC, and the downconversion mixers. A number of new ideas are proposed that improve the TX performance.

## 4.1 RF DAC

As the most power-hungry block in a transceiver, the RF DAC merits extensive design iteration and optimization. In contrast to conventional PA design, our methodology maximizes the efficiency for the peak desired output power with no emphasis on the DAC nonlinearity; the $\Delta\Sigma$ modulator loop effectively removes the resulting static and dynamic distortion.

Each DAC cell in the $I$ and $Q$ paths must translate the baseband data to RF and convert the resulting voltage to current. Figure 4.1(a) depicts a simple implementation where the two paths operate with 25%-duty-cycle LO waveforms and meet in the current domain. But we can also view the output current combining operation as an OR function and, since only one output transistor is on at a time, we move this function to the digital domain [Fig. 4.1(b)]. This merging of $I$ and $Q$ DACs halves the area and the output capacitance.

The merged DAC consists of 256 differential cells, each implemented as shown in Fig. 4.2. To maximize the efficiency, the DAC output stage and the off-chip matching network are designed for class-E operation, but at the cost of a single-ended drain voltage swing of 4.1 $\mathrm{V_{pp}}$. Thus, the unit cells employ triple cascodes, with $M_3$ and $M_6$ realized by thick-oxide transistors. Note that the 1.8-V supply tied to gates need not deliver any dc current and can be generated on chip by a charge pump. (Our experimental prototype uses an external 1.8-V supply.)

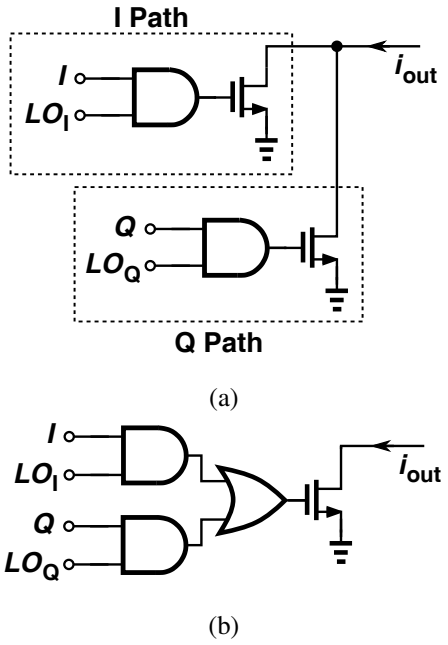



Figure 4.1. (a) A conceptual slice of RF DAC cell, and (b) merging of output transistors.

Beyond class-E operation, the series resistance of $M_1$-$M_3$ and $M_4$-$M_6$ in Fig. 4.2 determines the efficiency, as these devices must act as switches rather than current sources. On the other hand, wider transistors translate to greater input and output capacitances, with the former directly reducing the efficiency. Viewing the three NAND gates preceding $M_1$ (and $M_4$) as a "predriver," we note that their power dissipation, given by $fCV_{DD}^2$, rises with $W_1$. In this work, the total width of the input transistor on each side is about 5.65 mm, leading to a total predriver power of about 12 mW at 2 GHz.

Recall from Section 3.2 that the DAC units are scaled by factors of 0.25, 0.375, 0.5, etc. A width of 6.25 $\mu$m is chosen for the scaling factor of 0.25 and is used in cells number 1 to number 64.

## 4.2 Oversampling ADC

The oversampling ADC in Fig. 3.5 plays a critical role in the overall TX performance in terms of ACPR and the receive-band noise. We propose the use of a highly-oversampled $\Delta$ modulator as

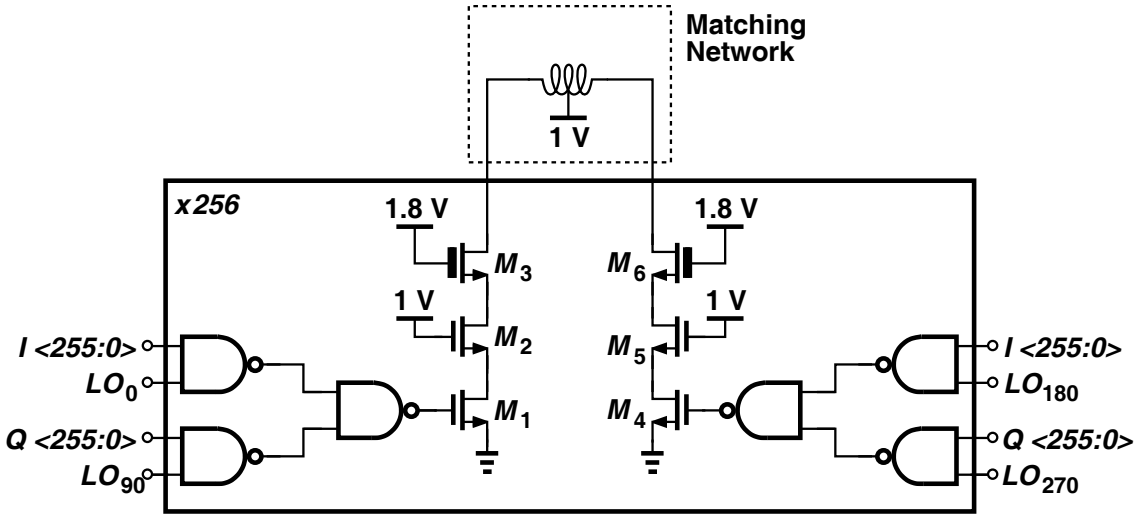



Figure 4.2. Complete RF DAC cell.

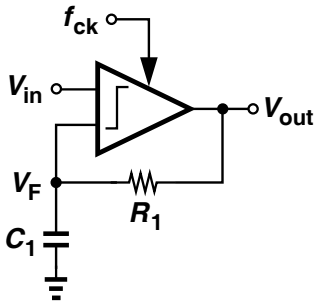



Figure 4.3. Simple $\Delta$ modulator.

an ADC. As explained below, the simplicity of the circuit, along with several new ideas, affords an efficient solution.

Figure 4.3 shows a simple $\Delta$ modulator, where the high gain of the comparator ensures that the output's running average, produced by the feedback RC network, tracks the analog input. For our subsequent design work, we must formulate the quantization noise spectrum of the output. The comparator itself acts as a 1-bit quantizer, exhibiting a total quantization noise power equal to $\Delta^2/12$, where $\Delta = V_{DD}$ (Appendix I).

The low-pass feedback loop around the comparator creates a high-pass shaping function for the noise. Figure 4.4 depicts a linear model for the modulator, indicating that the quantization noise,

$Q$, is shaped by

$$\begin{aligned}\frac{V_{out}}{Q} &= \frac{1}{1+\dfrac{A_0}{R_1C_1s+1}} \\ &= \frac{R_1C_1s+1}{R_1C_1s+1+A_0}.\end{aligned} \tag{4.1}$$

We thus expect a suppression factor of $1+A_0$ for noise frequencies below $1/(2\pi R_1C_1)$.

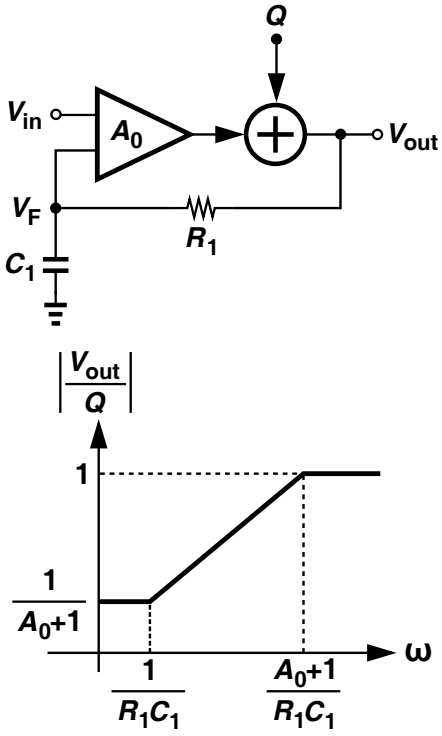



Figure 4.4. Quantization noise model.

We must now address two questions: (1) how much is $A_0$? and (2) how do we select $R_1C_1$? For the former, we first note that the gain of a 1-bit quantizer depends on its input amplitude. Let us observe that the high loop gain produces a small difference between $V_{in}$ and the running average that appears in $V_F$ in Fig. 4.3 if the input frequency is sufficiently smaller than $f_{CK}$. That is, the comparator does not see a significant differential voltage related to $V_{in}$. However, $V_F$ still experiences moderate changes due to the output rail-to-rail swings (Fig. 4.5). The triangular waveform at $V_F$ exhibits a peak swing of approximately $[V_{DD}/(4R_1C_1)](T_{CK})$, which we assume much greater than the difference seen by the comparator due to the analog input. To find the gain, we view the comparator as an amplifier that senses this triangular waveform and generates

an output first harmonic amplitude equal to $2V_{DD}/\pi$. Finding the first harmonic of the triangular waveform, we define the gain as the ratio of the output and input fundamental amplitudes:

$$\begin{aligned} A_0 &\approx \frac{2V_{DD}/\pi}{(8/\pi^2)V_{DD}/(4R_1C_1f_{CK})} \\ &\approx \pi R_1 C_1 f_{CK}. \end{aligned} \tag{4.2}$$

While intuitive, the forgoing calculation of $A_0$ tends to overestimate the gain. In Appendix I, we formulate $A_0$ using a different approach and observe that $A_0$ is closer to $R_1C_1f_{CK}$. In practice, $A_0$ depends on the input signal statistics and lies in this range. We hereafter conservatively assume that $A_0 \approx R_1C_1f_{CK}$.

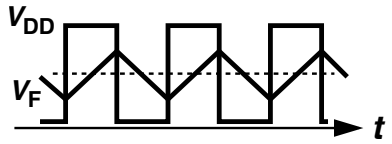



Figure 4.5. Circuit's waveform.

The comparator's quantization noise spectrum, $\Delta^2/(12f_{CK})$, is divided by $(1 + A_0)^2$ up to a frequency of $1/(2\pi R_1C_1)$, emerging as

$$S_Q(f) \approx \frac{V_{DD}^2}{12R_1^2C_1^2f_{CK}^3} \text{ for } f < \frac{1}{2\pi R_1C_1}. \tag{4.3}$$

For example, if $1/(2\pi R_1C_1) = 100$ MHz and $f_{CK} = 4$ GHz, we have $S_Q(f) \approx 3.5\times10^{-13}$ V$^2$/Hz $\equiv -94$ dBm/Hz. This is the quantization noise in the $\Delta$ modulator output. To refer this noise to the TX output in Fig. **??**(d), we must divide it by the gain from $Y$ to $E$ through the feedback path. We return to this point in Section 4.4.

Equation (4.3) makes it desirable to maximize the value of $R_1C_1$, but an excessively low corner frequency in the feedback path attenuates the signal of interest in $V_F$ (the running average), affecting the information carried by $V_{out}$. Since in our TX environment, both the baseband signal and the RX-band noise are of interest, we choose $1/(2\pi R_1C_1) \approx 100$ MHz.

## 4.3 Circuit Refinements

The architecture and circuit developments in the previous sections have assumed an 8-bit RF DAC with an INL of about 40% and a 4-GHz $\Delta$ modulator acting as the oversampling ADC in Fig. 3.5. To meet the performance specifications described in Section 1.2, the TX loop must reduce the DAC INL to about 6.5%. Moreover, the $\Delta$ modulator quantization noise spectral density must be further lowered by about 6 dB. In this section, we introduce a multitude of new circuit techniques that dramatically improve the performance.

In order to reduce the $\Delta$ modulator quantization noise, we double the effective oversampling rate by interleaving two $\Delta$ modulators. Illustrated in Fig. 4.6, the circuit employs two StrongArm comparators while running with $f_{CK} = 4$ GHz. Each comparator draws only 250 $\mu$W. The quantization noise drops by 1.2 dB in this case.

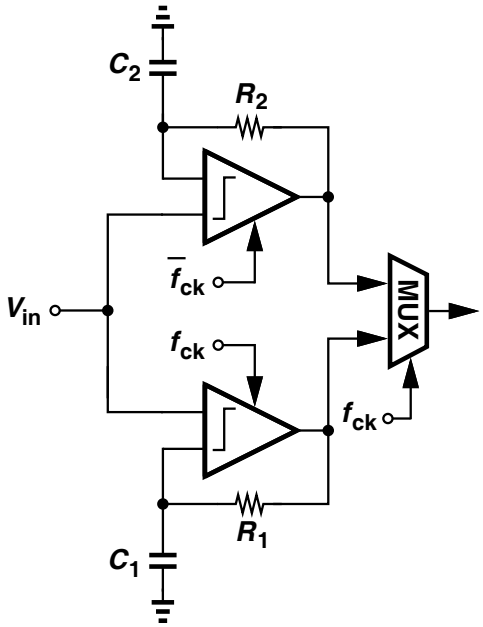
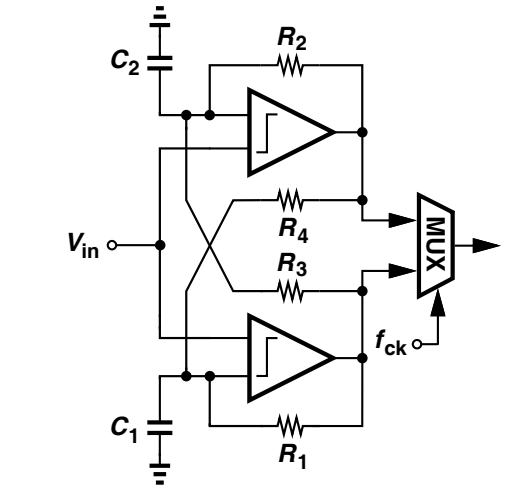


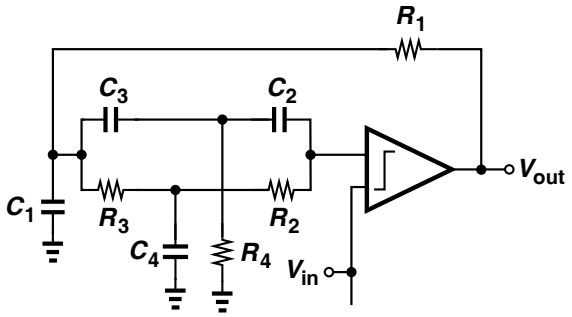

Figure 4.6. Time-interleaved $\Delta$ modulators.

As observed in the derivation leading to Eq. (4.2), the open-loop gain of the $\Delta$ modulator can potentially increase if we attenuate the clock swing sensed by the comparator. We now present three methods for this purpose. In the interleaved circuit of Fig. 4.6, we recognize that the two comparator outputs carry the first clock harmonic with opposite signs and the signal of interest

with the same sign. We therefore feed the output of each comparator to the input of the other [Fig. 4.7(a)] thereby reducing the clock swings at their inputs. With $R_1 = \cdots = R_4$, simulations indicate that the quantization noise falls by another 5.3 dB. We also interpose a passive notch filter between the feedback network and each comparator's input, with the notch frequency chosen equal to 4 GHz [Fig. 4.7(b)]. With this change, the $\Delta$ modulator's quantization noise drops by another 0.8 dB, reaching $6.5 \times 10^{-14}\ \mathrm{V^2/Hz}$.




(a)




(b)

Figure 4.7. (a) addition of $R_3$ and $R_4$ to attenuate clock swings, and (b) use of notch filter to attenuate clock swings.

## 4.4 $\Delta$ Modulator with High Gain

The next modification of the $\Delta$ modulator deals with the low loop gain of the overall TX architecture. We first describe the cause of the low gain. To measure the loop transmission, we break the

loop at the mixer output in Fig. 3.5, apply a 1-mV step to the subtractor input, and examine the DAC output voltage sensed by the mixer. For simplicity, we assume a mixer conversion gain of unity. The oversampling ADC (the $\Delta$ modulator) generates a periodic sequence consisting of one pulse of height $\mathrm{V_{DD}}$ (=1 V) and 999 low levels so as to deliver an average value equal to 1 mV. The register thus increments by 1 LSB every 1000 clock cycles, producing at the DAC output a staircase voltage with a slope of $V_{LSB,DAC}/(1000T_{CK})$, where $V_{LSB,DAC}$ denotes the DAC output voltage LSB size. More generally, for a step of $\Delta V$ at the subtractor input in Fig. 3.2, the DAC output has a slope of $(V_{LSB,DAC}/V_{DD})(\Delta V/T_{CK})$. The discrete-time loop transmission is therefore given by $(V_{LSB,DAC}/V_{DD})[z^{-1}/(1-z^{-1})]$, implying a gain of $V_{LSB,DAC}/V_{DD}$ for the integrator. If the DAC output full-scale voltage is comparable to $V_{DD}$, then this factor is around 1/256 for an 8-bit DAC, degrading the TX loop's ability to correct the DAC distortion.

Since the overall loop consists of the $\Delta$ modulator, the register, the DAC, and the downconversion mixer, we have few options for introducing a gain of 200-300 to compensate for the integrator loss. If realized by a conventional amplifier, such a high gain would entail severe nonlinearity and noise issues. We thus propose a new amplification method that simply draws upon the $\Delta$ modulator's comparator.

Illustrated in Fig. 4.8, the idea is to view the comparator as a high-gain amplifier and place a resistive network around it to obtain a low-frequency closed-loop gain of $1 + R_1/R_M$. We select $(R_1||R_M)C_1$ according to the desired corner frequency and $R_1/R_M \approx 200$ to achieve a high closed-loop gain.

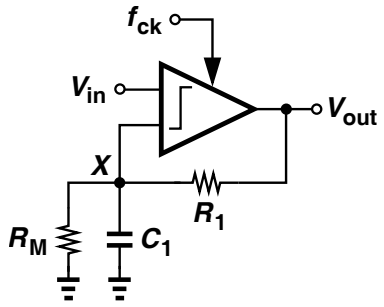



Figure 4.8. $\Delta$ modulator having a closed-loop gain of $1 + R_1/R_M$.

The topology resembles a high-gain feedback amplifier except that the comparator acts as a

discrete-time circuit running at a high oversampling ratio. Figure 4.9 plots the simulated input and output spectra of the high-gain $\Delta$ modulator with $V_{in} = (2\ \text{mV})\cos(2\pi \times 15.875\ \text{MHz})$, $R_1 = 300\ \text{k}\Omega$, $R_M = 1.5\ \text{k}\Omega$, and $C_1 = 1\ \text{pF}$. We observe a gain of about 46 dB. Simulations also indicate little change in the harmonic distortion at the $\Delta$ modulator output when dc gain is raised from 1 to 200.

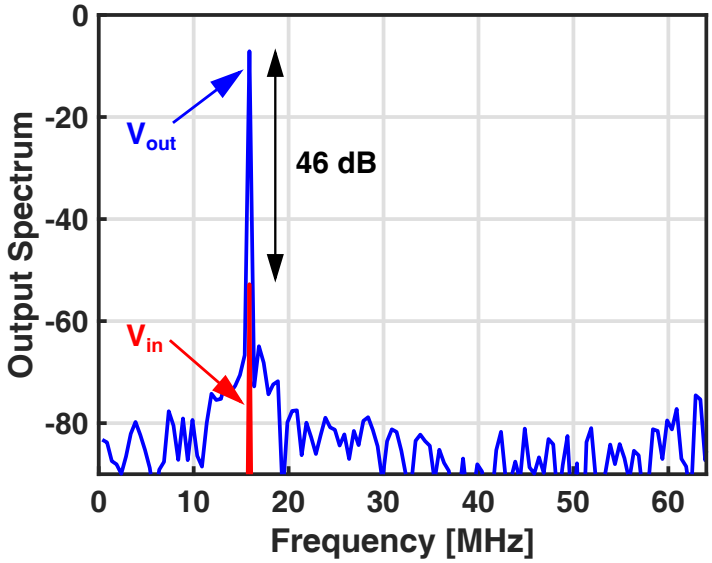



Figure 4.9. Simulated output spectrum showing the gain.

We should mention that the comparator offset, $V_{OS}$, is compensated by the overall TX loop because $V_{OS}$ appears before the integrator. For the integrator output to remain bounded, the feedback path in Fig. 3.5 must bring to the subtractor an offset exactly equal to $V_{OS}$.

The foregoing modifications dramatically reduce the $\Delta$ modulator's quantization noise as it is referred to the TX output. With a gain of 200 and a mixer loss of 15 dB, the $-102$-dBm/Hz noise calculated in the previous section is attenuated by approximately 31 dB when referred to the TX output, well exceeding WCDMA specification of $-125$ dBm/Hz.

## 4.5 $\Delta$ Modulator with Baseband Inputs

We must still address an issue raised in Section 3.2: in Fig. 3.5, we must somehow subtract the analog output of the mixer from the digital baseband data. To this end, we convert node $X$ in Fig. 4.8 to a virtual ground by grounding the other input of the comparator, utilize a binary-weighted resistor DAC to convert $I_{in}$ to an analog current, and inject the result into $X$ (Fig. 4.10). Similarly, the mixer output, $V_{mix}$, is summed with $I_{in}$ in the current domain. In this work, the unit resistor, $R = 2.9\ \mathrm{k}\Omega$, is chosen large enough so that resistor mismatches still allow a monotonic behavior for the DAC. Note that $R||(2R)||\cdots||(2^6R)||(2^7R)||(2^7R)||R_{mix}$ acts as $R_M$ in Fig. 4.8 and, along with $R_1$, defines the closed-loop gain.

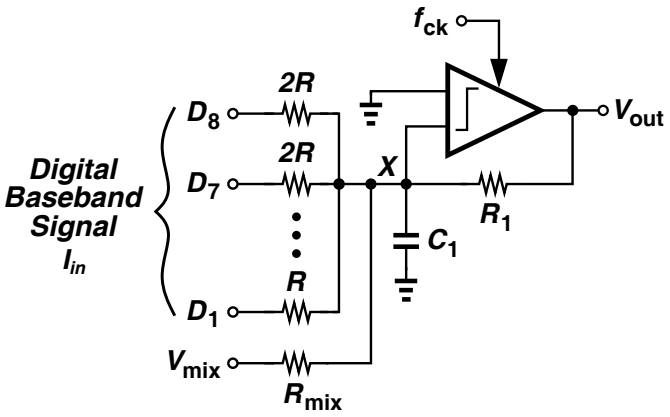



Figure 4.10. Conversion of digital baseband signal, $I_{in}$, to analog form.

We now describe another simplification in the $\Delta$ modulator design. The $\Delta$ modulator/subtractor studied thus far has single-ended inputs. Shown in Fig. 4.11 is the fully-differential, interleaved circuit. For simplicity, the baseband input is denoted by BB and its DAC by a single resistor in a dashed box. An interesting question that arises here is whether we can short the virtual ground nodes $X_1$ and $X_2$. These nodes carry the desired signal with the same polarity and the clock waveform with opposite polarities. Thus, shorting $X_1$ to $X_2$ and $Y_1$ to $Y_2$ removes the odd harmonics of the clock, reduces the clock swings at these nodes, and hence *increases* the open-loop gain of the comparators. According to simulations, this method raises the signal-to-quantization-noise ratio (SQNR) at the output by 6 dB. The topology can be further simplified if the multiplexer is inserted within the feedback loop, as shown in Fig. 4.12.

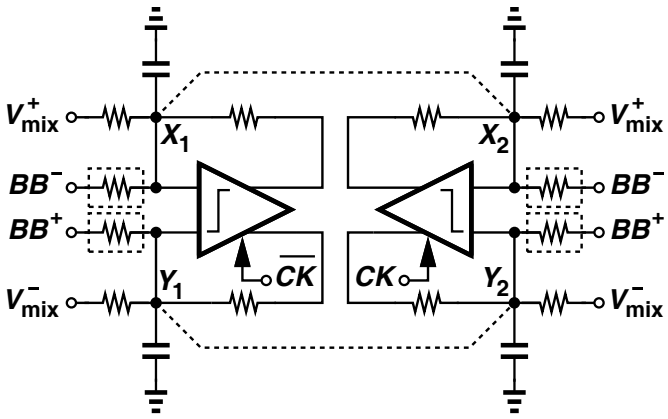



Figure 4.11. Differential interleaved $\Delta$ modulators.

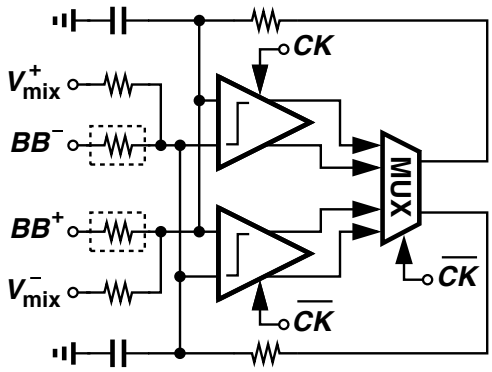



Figure 4.12. Simplified topology.

## 4.6 Mixer Implementation

Implementing Newton-Raphson technique requires a quadrature demodulator since Newton-Raphson technique is applied on the base band signal but the transmitter's output signal is at RF. Therefore, a demodulator is needed to extract the base band information to be used by Newton-Raphson technique.

Figure 4.13 shows how any distortion added by downconversion affects the transmitter's characteristic. Based on (4.4), any distortion added by the mixer directly appears at the transmitter's output. Therefore, linearity of the mixer is one of the key factors affecting the transmitter's linearity.

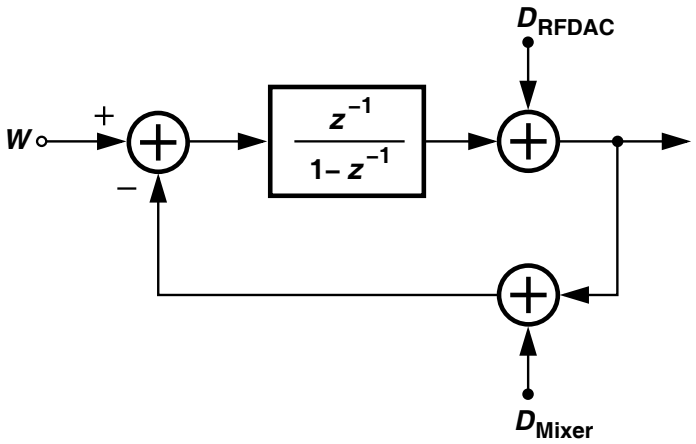
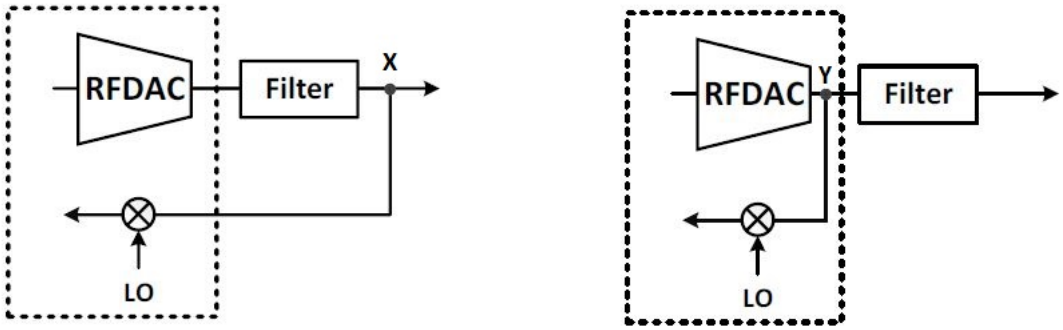



Figure 4.13. Effect of different imperfections on the transmitter's output.

$$Y = Xz^{-1} + D_{DAC}(1 - z^{-1}) + D_{Mixer}z^{-1} \tag{4.4}$$

Due to the importance of mixer's linearity, passive mixers are preferred over the active ones. Moreover, active mixers have transistors operating in saturation whereas passive ones have transistors acting only as switches. Hence, passive mixers are better choices for an all-digital transmitter. Finally, passive mixers are better options as downconverters because of no power consumption (except the power consumed to drive LOs).




(a) (b)

Figure 4.14. Feedback from the output node.

As for linearity, larger swing of input RF signal deteriorates the passive mixer's linearity. It is due the fact that large input swing causes significant variation of $\mathrm{V_{gs}} - \mathrm{V_{th}}$. Change of $\mathrm{V_{gs}} - \mathrm{V_{th}}$ of mixer transistors causes variation in their channel resistance and therefore, the gain of the mixer.

Another important point is shown in Fig. 4.14(a) where a demodulator receives the transmitter's output signal and extracts its in-phase and quadrature components. However, in this configuration the demodulator is connected to a node (i.e. $X$) which does not exist on the chip. This is not desirable as it needs extra routings on PCB. Moreover, the bond wire and PCB also affect the transmitter's output signal, distorting the Newton-Raphson operation as the demodulated signal is not the same as the transmitter's output signal. To solve this problem, Newton-Raphson is applied on the RF DAC's output signal rather than the transmitter's output signal (Fig. 4.14(b)). It helps ensure the RF DAC's output meets the WCDMA requirements. This means that the transmitter's output signal also meets the WCDMA requirements since the output matching circuit passes the in-band signal with only some voltage gain. Although using node $Y$ has the benefit of avoiding off-chip routing, this node has stronger higher harmonics compared to node $X$ as the signal at node $Y$ is unfiltered. Using the unfiltered signal results in a number of issues. First, the higher harmonics raise the amplitude of the mixer's input signal and thus, the mixer becomes nonlinear. Second, assuming a square-wave LO, it's 3rd harmonic downconverts the 3rd harmonic present at the RF DAC's output, resulting in an unexpected base band signal added to the main one. To avoid 3rd harmonic downconversion, we can use interpolated LOs for mixers which have weaker 3rd harmonic. However, this method requires eight-phase LO which needs extra hardware to generate. Figure 4.15 shows the ACPR of the transmitter in a single-tone test after attenuating the mixer's input.

To solve these issues, we filter the 3rd harmonic before downconversion. An on-chip low-pass filter cannot attenuate the 3rd harmonic without attenuating the fundamental one well. Another option to attenuate higher harmonics is to notch them. An LC notch filter at $3f_{LO}$ has limited attenuation due to low quality factor of on-chip inductors. In addition, on-chip inductors are bulky and occupy a wide area. Due to the aforementioned problems, a twin-T notch filter is a better option (Fig. 4.16(a)). The notch frequency of twin-T filter changes by process and temperature variation. Assuming a change of 10% in $f_n = \frac{1}{RC}$, a twin-T still attenuates the input 3rd harmonic by more than 20 dB (Fig. 4.16(b)). To compare this filter by a first order low-pass filter, 10 dB attenuation of $3f_{LO}$ is achieved by attenuating the main signal by 3 dB.

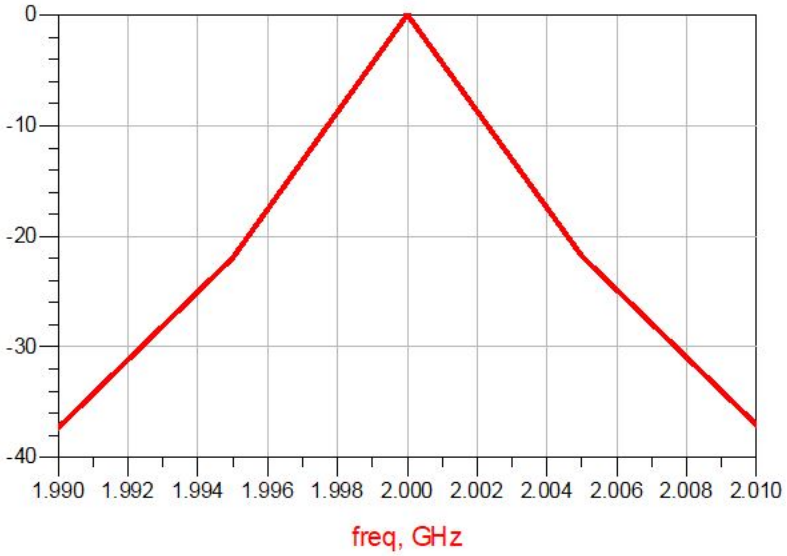
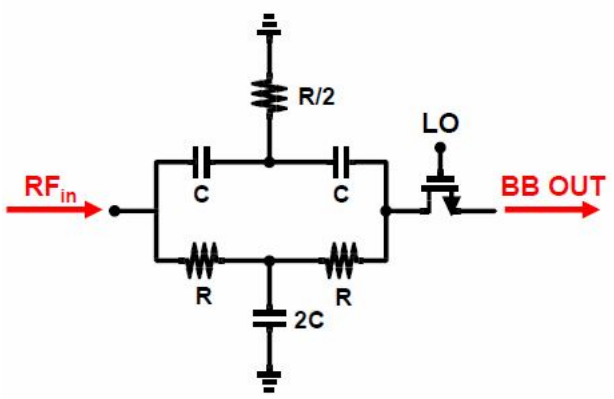


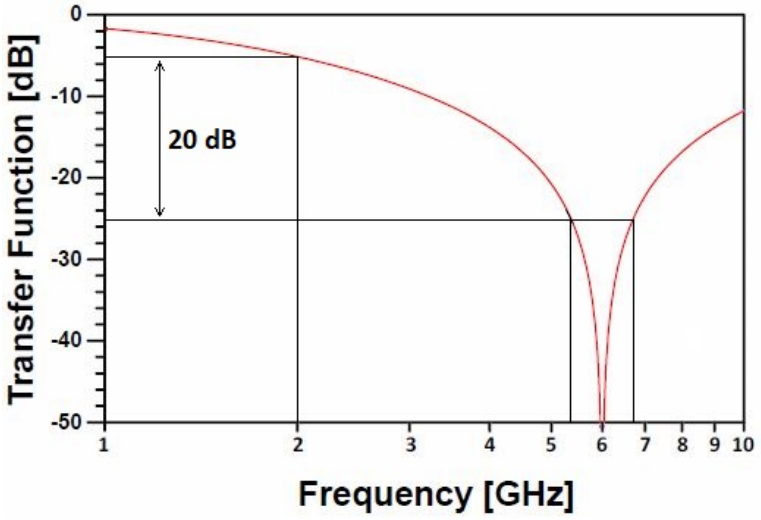

Figure 4.15. Transmitter's output ACPR.

Figure 4.17 shows the 3rd harmonic amplitude of the mixer's output (normalized to the fundamental one) when input RF amplitude varies. In other words, this figure shows the transmitter's ACPR in a single-tone test. Based on this figure, the input RF amplitude of this downconverter must not exceed $0.72\ \mathrm{V_{pp}}$ . This ensures that the ACPR criterion is met. Based on our simulations, the amplitude of the RF signal at the RF DAC's output exceeds this value; therefore, attenuating this signal before reaching the mixer is necessary. However, it is important to accommodate the largest possible input swing for the mixer since its pre-attenuation decreases the entire loop gain and consequently, deteriorates the linearity. Moreover, the downconverted signal is later processed by an analog-to-digital converter. Smaller amplitude of the downconverted signal worsens the data converter's SQNR; therefore, the output noise floor will be higher. To realize the attenuator, a resistive divider is not a good candidate for RF signal attenuation due to its noise and nonlinearity. Rather, a capacitive divider can serve the same purpose without aforementioned disadvantages. To realize such a divider, a capacitor is added before the twin-T filter (Fig. 4.18). This capacitor attenuates the signal by $\frac{Z_{in}}{Z_{in}+1/jC_1\omega}$.

Moreover, it is observed that doubling W and L improves the downconverter's linearity. This is explained by the short-channel effect, drain-induced barrier lowering. Doubling L reduces this effect and helps improve the mixer's linearity (Fig. 4.19).




(a)




(b)

Figure 4.16. (a) Twin-T filter preceding the mixer switch, (b) transfer function of twin-T filter.

Another important aspect is the trade-off between the large and small value of the load capacitance ($C_2$). Its larger value helps better attenuate the upconverted signal at the output of the mixer but deteriorates the loop's phase margin and raises the chance of instability. In contrast, smaller $C_2$ attenuates the upconverted signal insufficiently. This insufficient attenuation raises the swing of the mixer's output. This larger swing, however, exacerbates the mixer's linearity (due to larger $\mathrm{V_{gs}} - \mathrm{V_{th}}$ variation). Therefore, to make a compromise a value of $C_2 = 300$ fF is chosen.

Another important point in designing a mixer for this system is its LO duty cycle. The RF DAC has 25% inphase and quadrature LOs. Moreover, this RF DAC is functioning as a mixer, upconverting the base band input signal to RF frequencies. The downconverter used in this sys-

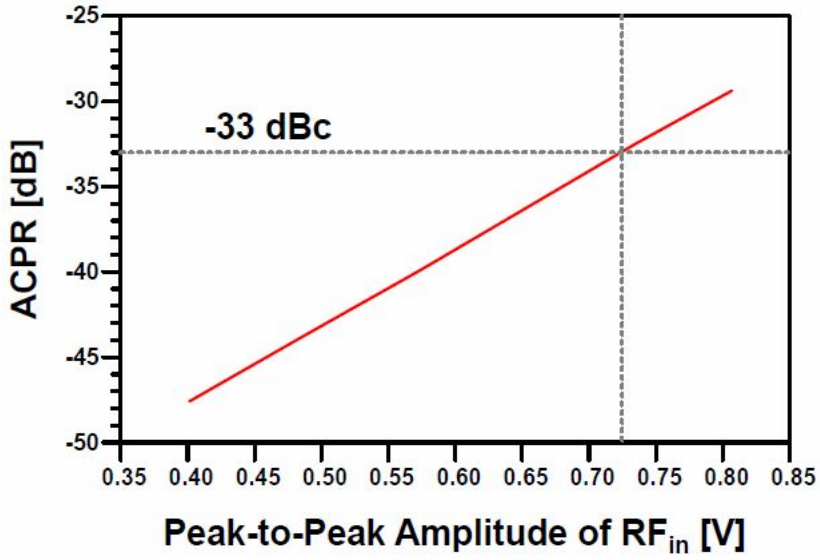



Figure 4.17. ACPR as a function of input swing.

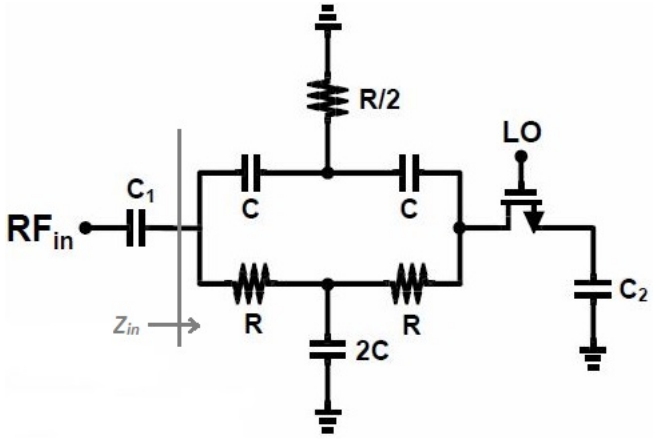



Figure 4.18. RF signal division.

tem functions inversely, making the output (upconverted) signal revert to base band. Hence if the same LOs are used, the difference between the RF DAC's base band input signal and the down-converted base band signal must be explained by the RF DAC's nonlinearity since the RF DAC is quite nonlinear but mixer is sufficiently linear. Based on this rationale, 25% LOs are used for the downconverter.

Lastly, it is essential to set the dc level of the downconverted signal. Based on Fig. 4.20, node $A$ is connected to supply, ground and input through capacitors. Therefore, this node is dc-wise float. However, it is later connected to a data converter whose input dc level is 0.5 V, dictating the

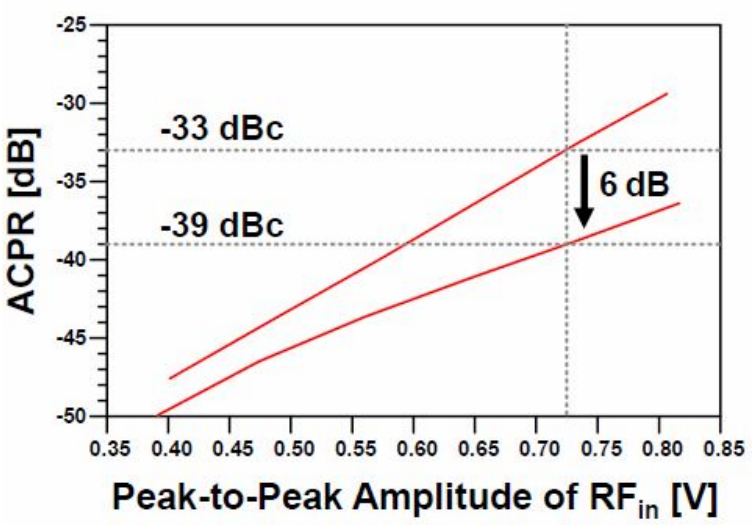



Figure 4.19. ACPR as a function of input swing for doubled width and length.

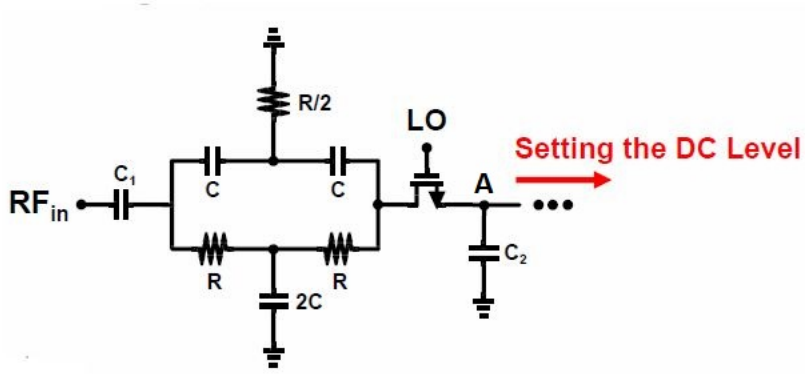
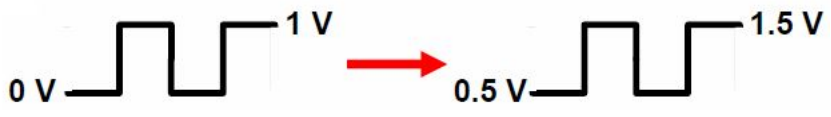
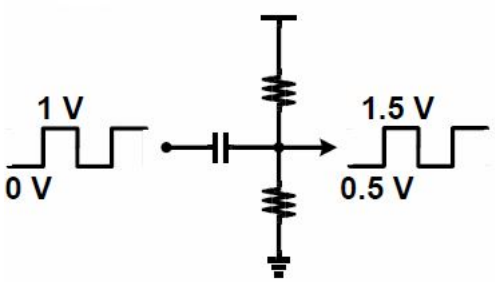



Figure 4.20. Dc level of the mixer's output.

dc level of node $A$ to be 0.5 V too. This dc level causes very poor linearity for this mixer. Let's assume that the mixer's output swing is 0.6 $V_{pp}$; therefore, its output changes from 0.2 V to 0.8 V. Consequently, $V_{gs} - V_{th}$ changes from -0.2 V to 0.4 V when $V_g = 1$ V (assuming $V_{th} \approx 0.4$ V). This means that the switch is sometimes off when it is supposed to be on. Therefore, this dc level is too high for downconverter to be sufficiently linear. To solve this problem, we use clock boosting technique in which the dc level of the mixer's LO is shifted accordingly [Fig. 4.21(a)]. Raising LO's dc level helps increase the switch gate voltage when it is supposed to be on. Figure 4.21(b) shows an HPF filter resetting the dc level to 750 mV.




(a)




(b)

Figure 4.21. (a) Clock boosting, (b) resetting the dc level.

In this new circuit, when $\mathrm{V_g} = 1.5\ V$, $\mathrm{V_{gs}} - \mathrm{V_{th}}$ changes from 0.3 V to 0.9 V which implies that this switch is always on. When $\mathrm{V_g} = 0.5\ V$, $\mathrm{V_{gs}} - \mathrm{V_{th}}$ changes from -0.7 V to -0.1 V, ensuring the switch is off. Figure 4.22 shows the mixer's output spectrum in a single-tone test after the clock boosting technique is applied.

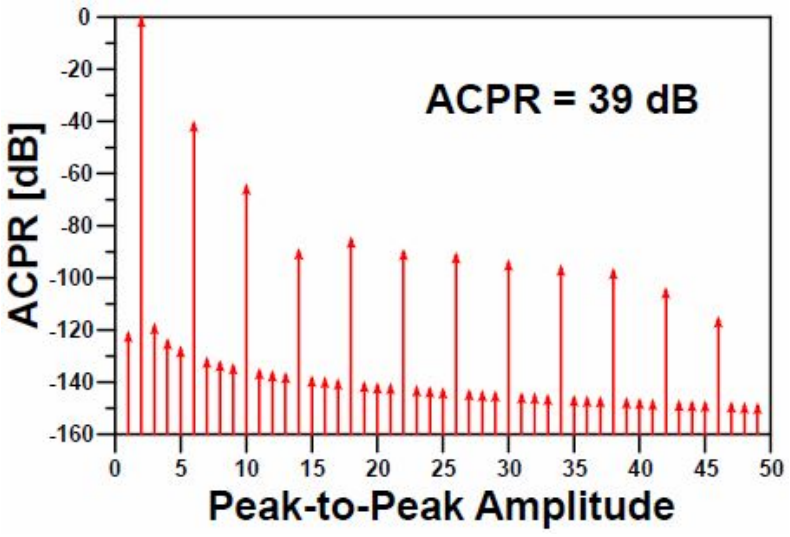



Figure 4.22. Mixer's output spectrum.

# CHAPTER 5

# Experimental Results

The complete transmitter has been fabricated in TSMC's 28-nm CMOS technology. It consists of four $\Delta$ modulators, two registers, each containing 256 flipflops, the merged RF DAC, two downconversion mixers, and I/Q clock generation with 25% duty cycle. Receiving 8-bit baseband digital $I$ and $Q$ inputs, the TX operates with a 1-V supply, except for the gates of the thick-oxide devices in Fig. 4.2, which are tied to 1.8 V. Figure 5.1 shows the die photograph with an active area of 0.35 mm $\times$ 0.32 mm.

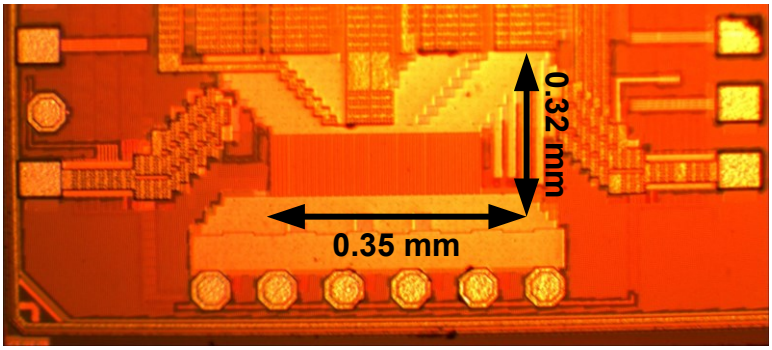



Figure 5.1. Chip photograph.

The TX generates differential outputs, which travel through multiple bond wires and tapered transmission lines on the printed-circuit board for matching and differential to single-ended conversion.

The TX is measured with digital baseband data that has been subjected to root-raised-cosine filtering in an FPGA. All of the clock phases are generated on-chip by means of frequency dividers and logic from an external 8-GHz input.

The results reported here are in the context of WCDMA, but the highly-flexible architecture readily lends itself to other standards as well. Figure 5.2 shows the measured output spectrum. The output power is 24.1 dBm (including 2.7 dB loss due to cables, connectors and bond wires). The ACPRs in the adjacent and alternate adjacent channels are equal to $-35.4$ dB and $-44.6$ dB, receptively, exceeding the WCDMA specifications. The receive-band noise is $-137$ dBc/Hz at 50-MHz offset. Under these conditions, the efficiency is 50%.

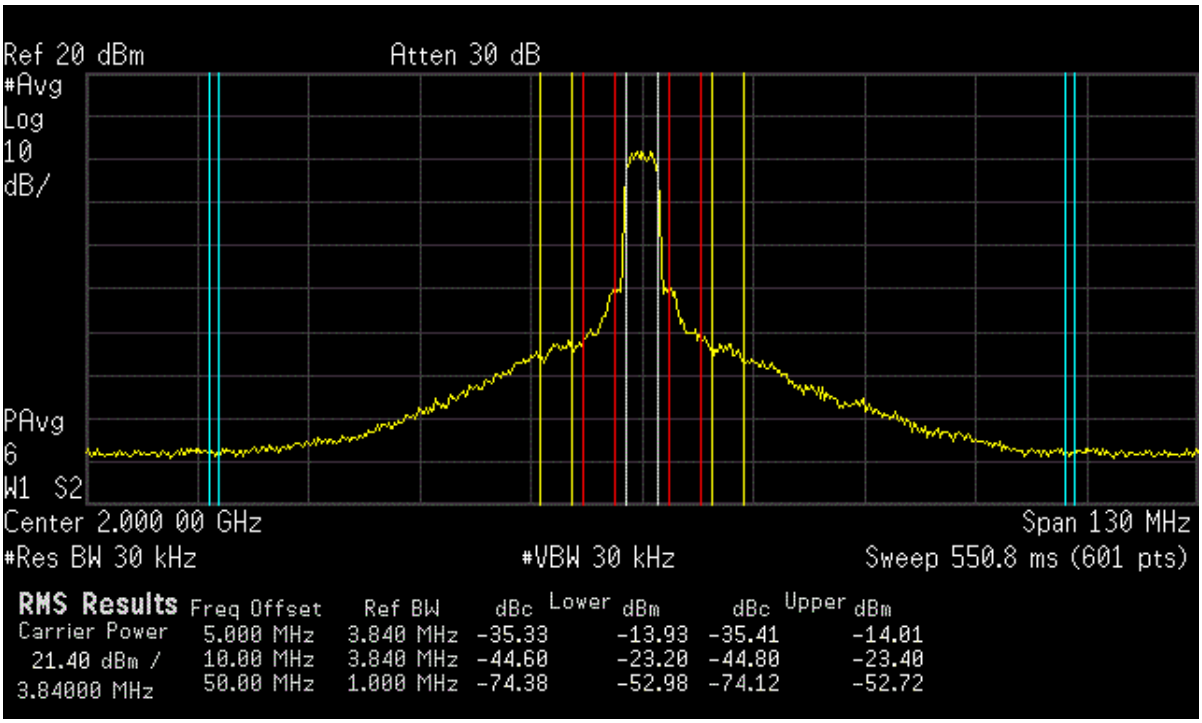



Figure 5.2. Measured output spectrum, ACPR and RXBN

Figures 5.3, 5.4, and 5.5 respectively plot the EVM, ACPR, and efficiency as a function of the output power. The EVM remains much less than the WCDMA specification, and $\mathrm{ACPR}_1$ and $\mathrm{ACPR}_2$ reach a minimum of $-53$ dB and $-58$ dB, respectively. The efficiency begins from about 4% at $\mathrm{P_{out}} = +5\ \mathrm{dBm}$ and sharply rises to 50%.

Table 5.1 summarizes the performance of our TX and compares it with prior-art transmitters in the range of 0.8 GHz to 2 GHz. We have achieved the highest efficiency and the smallest die area.

Table 5.2 compares our RX-band noise with those of transmitters in the vicinity of 2 GHz. A far higher efficiency and smaller area can be observed for our prototype. Note that only our work

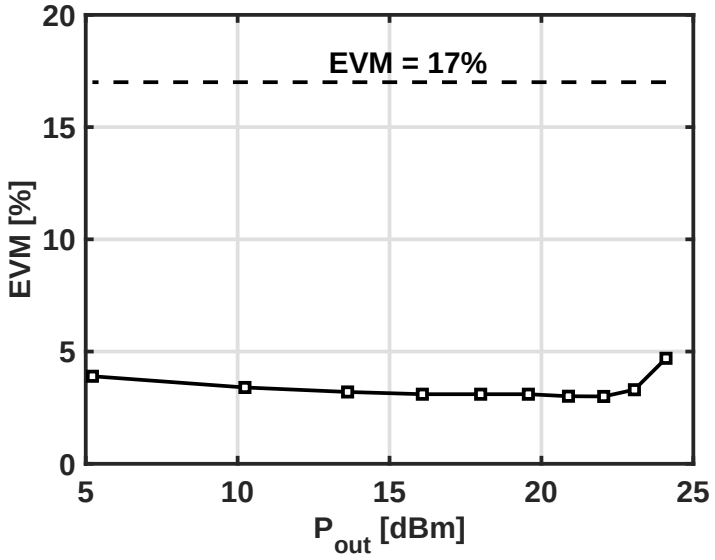



Figure 5.3. EVM as a function of output power.

corrects the nonlinearity in the background.

Table 5.1. Performance summary and comparison to prior art

| Reference | Oishi JSSC Dec. 2014 | Yamamoto MTT Mar. 2016 | Ko ISSCC 2017 | Yin ISSCC 2018 | This Work |
|---|---|---|---|---|---|
| Frequency (GHz) | 1.95 | 1.98 | 1.88 | 0.8 | 2 |
| $P_{out}$ (dBm) | 28 | 28 | 29.8 | 22.9 | 24.1 |
| Efficiency (%) | 39 | 35.1 | 40.8 | 26.1 | 50 |
| Supply (v) | 1.2/3.7 | 3.4 | 3.4 | 1.2/2.4 | 1/1.8 |
| Standard | WCDMA | WCDMA | WCDMA | WIFI | WCDMA |
| Technology | 90–nm CMOS | CMOS/GaAs hybrid | 153–nm CMOS | 55–nm CMOS | 28–nm CMOS |
| Area ($mm^2$) | 14 | 27.5* | 3.01 | 1.11 | 0.11 |
| Calibration | None | None | None | Foreground | Background |

* Glass epoxy laminate

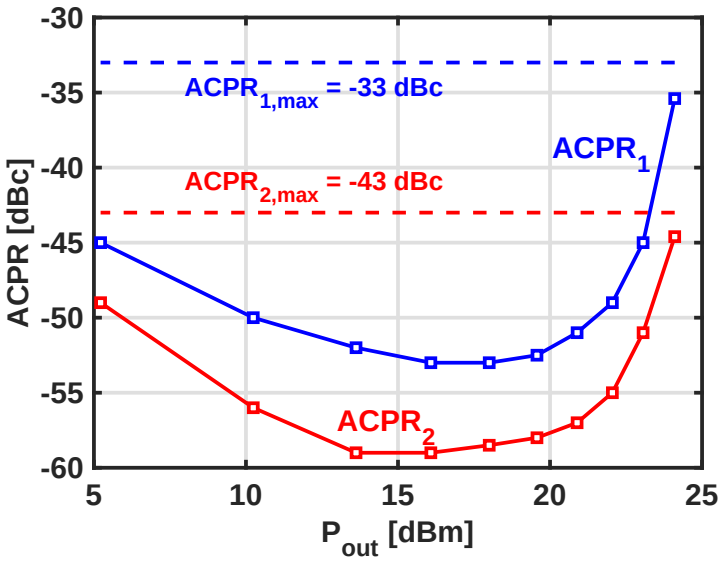



Figure 5.4. ACPR as a function of output power.

Table 5.2. Performance summary and comparison to prior art

| Reference | Yuan RFIC 2015 | Bhat ISSCC 2017 | This Work |
|---|---|---|---|
| Frequency (GHz) | 2 | 2.2 | 2 |
| $P_{out}$ (dBm) | 14.5 | 24 | 24.1 |
| Efficiency (%) | 12.2 | 16 | 50 |
| Resolution (bit) | 7 | 9 | 8 |
| CMOS Technology | 65 nm | 65 nm | 28 nm |
| Noise Floor (dBc/Hz) | −123 @ 140 MHz offset | −149 @ 140 MHz offset | −137 @ 50 MHz offset |
| Area ($mm^2$) | 1.75 | 3 | 0.11 |
| Calibration | Foreground | Foreground | Background |

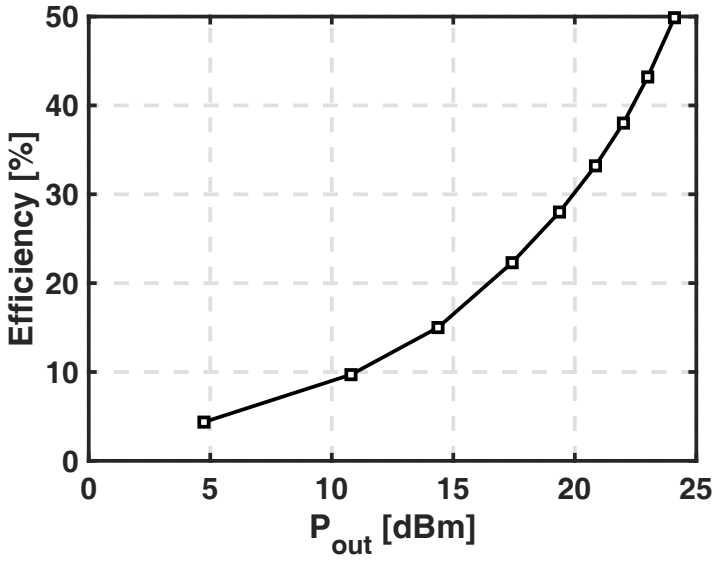



Figure 5.5. Efficiency as a function of output power.

# CHAPTER 6

# Conclusion

A simplified implementation of the Newton-Raphson equation solver leads to a TX embedded in a $\Delta\Sigma$ modulator loop. Moreover, the high-speed feedback ADCs necessary for background calibration can be realized as 1-bit $\Delta$ modulators. With other simplifications, a compact, highly efficient architecture emerges that can serve in multiple standards.

# APPENDIX I

In this appendix, we describe another approach to computing the gain of the comparator in the $\Delta$ modulator of Fig. 4.3. The gain can be defined so as to ensure a zero correlation between the output quantization noise, $Q$, and the comparator input differential voltage, $X$ [21]. Expressing the comparator output as

$$Y = A_0 X + Q, \tag{6.1}$$

we find the correlation by writing

$$\begin{aligned} E[XQ] &= E[X(Y - A_0 X)] \\ &= E[XY] - A_0 E[X^2]. \end{aligned} \tag{6.2}$$

Thus, $E[XQ] = 0$ if

$$A_0 = \frac{E[XY]}{E[X^2]}. \tag{6.3}$$

For simplicity, let us assume that the comparator output swings between $-V_{DD}/2$ and $+V_{DD}/2$. With $V_{in} = 0$ in Fig. 4.3, $Y$ toggles at a rate of $f_{CK}/2$, creating a triangular wave at the other input with a peak amplitude of $[V_{DD}/(4R_1C_1)]T_{CK}$. Note that $X = V_{in} - V_F = -V_F$. Since the comparator acts as a discrete-time circuit, the correlation between its input and output must be calculated at the sampling points, namely, at the positive and negative peaks of the triangular wave. We thus have

$$E[XY] = \frac{V_{DD}}{4R_1C_1} T_{CK} \frac{V_{DD}}{2} + \frac{-V_{DD}}{4R_1C_1} T_{CK} \frac{-V_{DD}}{2}. \tag{6.4}$$

Also,

$$E[X^2] = (\frac{V_{DD}}{4R_1C_1} T_{CK})^2 + (-\frac{V_{DD}}{4R_1C_1} T_{CK})^2. \tag{6.5}$$

It follows that

$$A_0 = 2R_1C_1 f_{CK}. \tag{6.6}$$

The forgoing calculation assumes that in Fig. 4.3, $V_{in} = 0$. If, for example, $V_{in} = V_{DD}/4$, then $A_0 \approx R_1C_1 f_{CK}$. With a time-varying input, and depending on its statistics, $A_0$ has an average value in the range of $R_1C_1 f_{CK}$ and $2R_1C_1 f_{CK}$.

According to [22], the quantization noise of a 1-bit quantizer can be approximated by $\Delta^2/12$, where $\Delta$ denotes the total height of the quantizer's bang-bang characteristic. In our case, $\Delta = V_{DD}$, yielding $\Delta^2/12 = V_{DD}^2/12$.

REFERENCES


[1] D. Schreurs *et al*, *RF Power Amplifier Behavioral Modeling*, Cambridge University Press, 2010.

[2] R.B. Staszewski *et al.*, "All-digital PLL and transmitter for mobile phones," *IEEE J. Solid-State Circuits*, Vol. 40, pp. 2469-2482, Dec. 2005.

[3] A. Kavousian *et al.*, "A Digitally modulated polar CMOS power amplifier with a 20-MHz channel bandwidth," *IEEE J. Solid-State Circuits*, Vol. 43, pp. 2251-2258, Oct. 2008.

[4] M. Collados *et al.*, "High-power digital envelope modulator for a polar transmitter in 65nm CMOS," *Proc. IEEE Custom Integrated Circuits Conference*, pp. 733-736, Sept. 2008.

[5] A.D. Presti *et al.*, "A high-resolution 24-dBm digitally-controlled CMOS PA for multi-standard RF polar transmitters," *European Solid-State Circuits Conference*, pp 482-485, Sept. 2008.

[6] D. Chowdhury *et al.*, "A 2.4GHz mixed-signal polar power amplifier with low-power integrated filtering in 65nm CMOS," *Proc. IEEE Custom Integrated Circuits Conference*, pp. 1-4, Sept. 2010.

[7] K. Seah *et al*, "A digital polar amplifier for ultra-wideband with dynamic element matching," *Proc. IEEE International Conference on Ultra-Wideband*, pp 1-4, Sept. 2010.

[8] D. Chowdhury *et al*, "An efficient mixed-signal 2.4-GHz polar power amplifier in 65-nm CMOS technology," *IEEE J. Solid-State Circuits*, Vol. 46, pp. 1796-1809, June 2011.

[9] C. Lu *et al.*, "A 24.7dBm all-digital RF transmitter for multimode broadband applications in 40nm CMOS," *IEEE ISSCC Dig. Tech. Papers*, pp. 332-333, Feb. 2013.

[10] F. Roger, "A 200mW 100MHz-to-4GHz 11th-order complex analog memory polynomial predistorter for wireless infrastructure RF amplifiers," *IEEE ISSCC Dig. Tech. Papers*, pp. 94-95, Feb. 2013.

[11] Y. Guo *et al.*, "Power adaptive digital predistortion for wideband RF power amplifiers with dynamic power transmission," *IEEE Trans. MTT*, vol. 63, no. 11, pp. 3595-3607, Nov. 2015.

[12] S. Chung *et al.*, "Digital predistortion using quadrature $\Delta\Sigma$ modulation with fast adaptation for WLAN power amplifiers," *IEEE MTT-S Int. Microw. Symp. Dig.*, pp. 1-4, June 2011.

[13] 3GPP TS 25.101 V16.0.0: 3rd Generation Partnership Project; Technical Specification Group Radio Access Network; User Equipment (UE) radio transmission and reception (FDD) (Release 16)

[14] B. Razavi, *Principles of Data Conversion System Design*, 1995.

[15] K. Oishi *et al.*, "A 1.95 GHz fully integrated envelope elimination and restoration CMOS power amplifier using timing alignment technique for WCDMA and LTE," *IEEE J. Solid-State Circuits*, Vol. 49, pp. 2915-2924, Dec. 2014.

[16] K. Yamamoto *et al.*, "A WCDMA multiband power amplifier module with Si-CMOS/GaAs-HBT hybrid power-stage configuration," *IEEE Trans. MTT*, Vol. 64, no. 3, pp. 810-825, March 2016.

[17] J. Ko *et al.*, "A high-efficiency multiband class-F power amplifier in 0.153$\mu$m bulk CMOS for WCDMA/LTE applications," *IEEE ISSCC Dig. Tech. Papers*, pp. 40-41, Feb. 2017.

[18] Y. Yin *et al.*, "A compact dual-band digital doherty power amplifier using parallel-combining transformer for cellular NB-IoT applications," *IEEE ISSCC Dig. Tech. Papers*, pp. 408-410, Feb. 2018.

[19] W. Yuan *et al.*, "A quadrature switched capacitor power amplifier in 65nm CMOS," *IEEE Radio Freq. Integr. Circuits Symp. Dig.*, pp. 135-138, May 2015.

[20] R. Bhat *et al.*, "A >1W 2.2GHz switched-capacitor digital power amplifier with wideband mixed-domain multi-tap FIR filtering of OOB noise floor," *IEEE ISSCC Dig. Tech. Papers*, pp. 234-235, Feb. 2017.

[21] S.H. Ardalan *et al.*, "An analysis of nonlinear behavior in Delta-Sigma modulators," *IEEE Trans. Circuits Syst.*, Vol. 34, no. 6, pp. 593-603, June 1987.

[22] R. Schreier *et al*, *Understanding delta-sigma data converters*, Piscataway, NJ: IEEE Press, 2005.