POST-COMPENSATION OF TRANSMITTER'S NONLINEAR DISTORTIONS AT THE RECEIVER USING COMPRESSED SENSING TECHNIQUES FOR OFDM APPLICATIONS

BY

DAMILOLA SADIQ OWODUNNI

A Thesis Presented to the
DEANSHIP OF GRADUATE STUDIES

KING FAHD UNIVERSITY OF PETROLEUM & MINERALS

DHAHRAN, SAUDI ARABIA

In Partial Fulfillment of the
Requirements for the Degree of

MASTER OF SCIENCE

In

ELECTRICAL ENGINEERING

JULY 2012

KING FAHD UNIVERSITY OF PETROLEUM & MINERALS
DHAHRAN 31261, SAUDI ARABIA

DEANSHIP OF GRADUATE STUDIES

This thesis, written by **DAMILOLA SADIQ OWODUNNI** under the direction of his thesis adviser and approved by his thesis committee, has been presented to and accepted by the Dean of Graduate Studies, in partial fulfillment of the requirements for the degree of **MASTER OF SCIENCE IN ELECTRICAL ENGINEERING**.
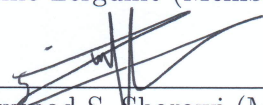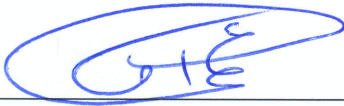
Thesis Committee
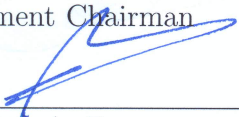
_____
Dr. Tareq Y. Al-Naffouri (Adviser)

_____
Dr. Oualid Hammi (Co-Adviser)

_____
Dr. Wajih A. Abu-Al-Saud (Member)

_____
Dr. Azzedine Zerguine (Member)

_____
Dr. Mohammad S. Sharawi (Member)

_____
Dr. Ali A. Al-Shaikhi
Department Chairman

_____
Dr. Salam A. Zummo
Dean of Graduate Studies

_____1(8)12_____
Date

To my $\{darling\ Mum\}^3$ & my wonderful Dad ...

ACKNOWLEDGMENTS

Alhamdulillaah, first and foremost. All praise is due to Allaah. And may He raise the rank of our Noble Prophet, his family and companions and grant them peace.

Before all else, I praise Allaah and thank Him for His unceasing, innumerable blessings upon me. I thank Him for blessing me to attend this great university, study under proficient faculty, work with terrific advisers, and complete this thesis work.

Then, after that, I would like thank the two greatest people in my life—my wonderful parents—for their constant support and encouragement throughout the years, in every aspect of my life and their great concern to see that I succeed in life. I thank them particularly for their encouragement and giving importance to my education as well as that of my brothers. Indeed, I owe them so much that words can never suffice and I don't think I can ever thank them enough. I can only ask Allaah that He rewards them abundantly.

My deep appreciation also goes to my adviser, Dr. Tareq Al-Naffouri who I decided to work with after I noticed his excellent teaching skills during a course I took with him. In this class, I realized that he has been blessed with two characteristics essential for any teacher to have: deep understanding of the material and a concern for the students' comprehension. I wouldn't be over exaggerating when I say his class was the best I have ever taken. I thank him for his guidance in my thesis work where his

keem, who, despite being thousands of miles away, are constantly concerned about my success. Their unwavering support, advice and encouragement through the years is much appreciated.

With the academic load at KFUPM, one could almost forget the *Deen* and life could sometimes get boring and somewhat depressing. But the blessing of great brothers and friends like Akram AbdulCader (the *'key'* to the others), Moosaa Abu Zaid, Jarullah, Yousef, Bilaal, a wonderful roommate ('Abdul-Hafeedh), a special companion ('Abdul-Hakeem) and many others alleviated this. I also thank the KFUPM *Tawjeeh* committee for the religious services they provide the students as well as the whole KFUPM Nigerian community (lead by Dr. Balarabe Yushau) for their hospitality and support and in particular, Dr. Aadam Olatunji for his forthcoming pieces of advice.

My deep appreciation also goes to my fellow colleagues at the KFUPM Electrical Engineering department, particularly brothers Ahmad Salim, Mohammad Tamim AlKhodary, Sulaiman Lanre Taiwo, Yanal Al-Faouri, Nuruddeen Iya, and others for their support and advice during my program.

Lastly, I thank the entire KFUPM community in general and I ask Allaah that He rewards all the Muslims.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

THESIS ABSTRACT

NAME: Damilola Sadiq Owodunni

TITLE OF STUDY: Post-compensation of Transmitter's Nonlinear Distortions

at the Receiver Using Compressed Sensing Techniques for

OFDM Applications

MAJOR FIELD: Electrical Engineering

DATE OF DEGREE: MAY 2012

Correcting the power amplifier's (PA) distortion of orthogonal frequency division multiplexing (OFDM) signals has been the target of a lot of research. However, to meet spectrum mask requirements, these attempts have almost exclusively focused on the transmitter. In this thesis, we consider the PA non-linearities as a sparse phenomenon and then use four algorithms based on the novel field of compressed sensing (CS) to estimate and cancel these distortions at the receiver side. The first two algorithms are the CS and weighted CS algorithms which use frequency-domain free-carriers for measurements. We improve upon these with the third—an intelligent data-aided algorithm which does not require free-carriers and thus allows full bandwidth use. The last approach combines pre-compensation at the transmitter with CS-based post-

compensation at the receiver. All four algorithms are applied to measurement-based

amplifier models and results show their effectiveness in correcting the amplifier's dis-

tortion.

ملخص الرسالة

الاسم الكامل:		داميلولا صادق أودوني

عنوان الرسالة: تعويض تشوهات الإرسال الغير الخطية في الجانب المستقبل باستخدام تقنيات الإستشعار المضغوط لتطبيقات تقسيم التردد المتعامد المتداخل (OFDM).

التخصص: الهندسة الكهربائية

تاريخ الدرجة العلمية: يوليو 2012

إن تصحيح التشويه الناتج عن مضخم الطاقة لإشارات تقسيم التردد المتعامد المتداخل (OFDM) كانت و لا زالت محل دراسة عدد كبير من الباحثين، و لتلبية متطلبات الطيف الكهرومغناطيسي المتاح ركزت هذه المحاولات بشكل عام على طرف الارسال فقط.

في هذه الأطروحة، قمنا بدراسة تشوهات الإرسال الغير الخطية بإعتبارها ظاهرة متناثرة حيث تم إستخدام خوارزميات قائمة على تقنيات الإستشعار المضغوط لتقدير وإلغاء هذه التشوهات في الجانب المستقبل. و قد تم دراسة أربع خوارزميات ،الأولى مبنية على الإستشعار المضغوط و الثانية مبنية على الإستشعار المضغوط الموزون و كلتاهما تستخدم ترددات خالية للقياس. أمّا الخوارزمية الثالثة فهي تطوير لما سبق حيث تشكل خوارزمية ذكية مبنية على البيانات، و تتميز بعدم إستخدام ترددات خالية للقياس مما يسمح بإستخدام كامل للنطاق الترددي. تقوم الخوارزمية الرابعة بالجمع بين التعويض المسبق للجانب المرسل و التعويض المبني على الإستشعار المضغوط التشوهات في الجانب المستقبل.

أخيراً تم تطبيق هذه الخوارزميات على نماذج لمضخمات الطاقة لأغراض القياس، وقد أظهرت النتائج فعالية هذه الخوارزميات في تصحيح التشويه لمضخمات الطاقة.

درجة الماجستر في العلوم

جامعة الملك للبترول والمعادن

الظهران ــ المملكة العربية السعودية

رمضان 1433هـ

CHAPTER 1

INTRODUCTION

1.1 Motivation

1.1.1 Orthogonal Frequency Division Multiplexing (OFDM)

Today's world is one filled with a plethora of wireless communication devices. From the ubiquitous mobile phone and laptop to the more recent but already popular tablet computer, there is an on-going explosion in the use of wireless devices in every sphere of people's lives—be it school, business or pleasure. The result of this is a huge demand for high data rates to enable the much needed swift and rapid voice, video and data transfer. This, however, proves to be a challenge especially in the face of limited available bandwidth.

As a solution to this problem, orthogonal frequency division multiplexing (OFDM), which is a modulation scheme that allows high data transfer rates within a restricted bandwidth, was proposed. In an OFDM system, the data to be transmitted is split into a large number of streams each of which is modulated (at a low symbol rate)

onto one of an equally large number of sub–carriers distributed over particular frequencies in such a way that these sub-carriers are orthogonal to each other. This gives OFDM its high spectral density while at the same time allowing robustness to multipath fading, inter-symbol interference and other channel conditions. These benefits of OFDM have earned it an adoption in the standards of many of today's most important communication systems including digital subscriber line (DSL), wireless local area networks (WLAN), worldwide interoperability for microwave access (WiMAX), long term evolution (LTE), and digital audio broadcasting (DAB) [1–3].

However, since OFDM is a combination of a large number of sub-carriers, the output signal is essentially a sum of many sine waves and at the points where these waves add constructively, high peaks occur resulting in an OFDM output signal with a high peak-to-average power ratio (PAPR) as can be seen in the sample time domain OFDM signal in Figure 1.1. This high PAPR of the OFDM signal however, causes a problem at the transmitter's power amplifier (PA). This problem is described in the following section.

1.1.2 The Problem at the Power Amplifier: Efficiency vs. Linearity

Considering the effect of power consumption in wireless communication systems on the environment and on the cost of operation of wireless network infrastructure, as well as the battery life of mobile terminals, substantial research has been carried out with an aim of reducing the energy used by both the base station and the mobile
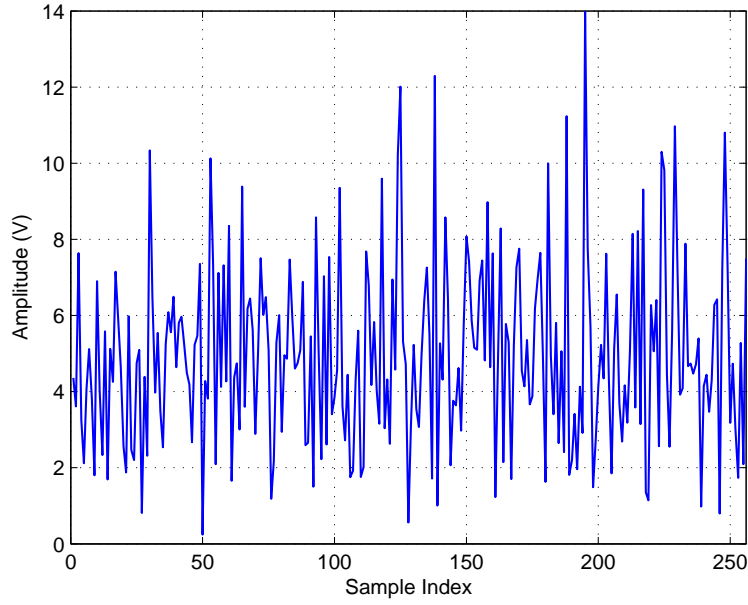
Figure 1.1: Time domain OFDM signal with high PAPR

devices.

In terms of power consumption, the key component of wireless systems is the power amplifier (PA) and with wireless communication consuming about 60 TWh or 0.33% of global electricity consumption annually [4], operating the PA at high efficiency is of primary importance in order to avoid energy wastage.

Besides its efficiency, another important consideration with regards to the power amplifier is its linearity. To avoid loss of information due to high error levels (in terms of bit error rate and error vector magnitude, for example) and to meet spectrum mask thresholds of various communication standards, it is important that the output of the PA is a linearly amplified replica of the input.

This linearity provision, however, turns out to oppose the high efficiency requirement. This is because, as is shown in the gain profile of a sample measurement-based amplifier model (the ZHL model, futher details are given in Chapter 2) in Figure 1.2,

3

it is only at low input levels (and therefore, low efficiency) that power amplifiers used in communication systems act linearly. On the other hand, at high input (and high efficiency) levels which are closer to the saturation region, these amplifiers usually show relatively high levels of distortion as can be seen in the gain response of a sample PA in Figure 1.2.
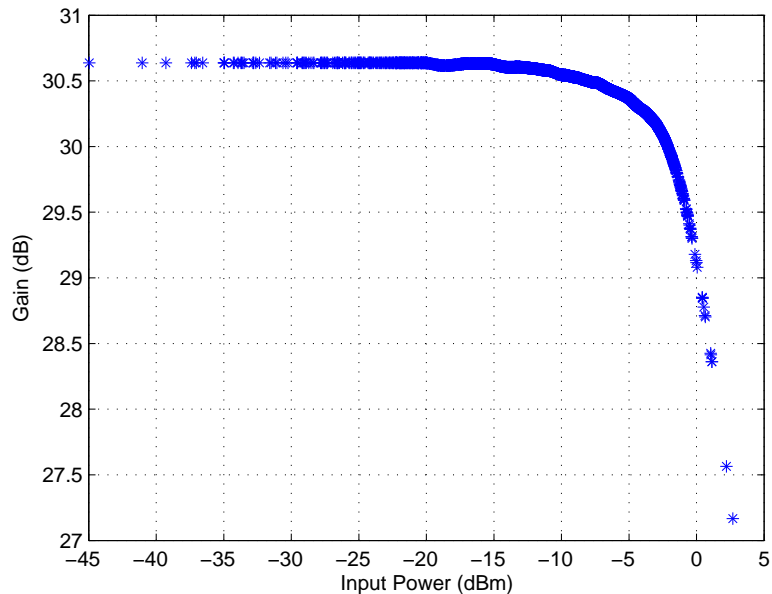
Figure 1.2: Gain response of a sample commercial amplifier

Due to the great operating cost of low efficiency, power amplifiers in wireless systems are typically operated closer to their high efficiency regions. Due to this, the high PAPR of OFDM signals induces the non-linearities of the amplifier leading to significant distortion in its output especially when the input is of high amplitude. The comparison of sample input and output signals of the ZHL amplifier model as shown in Figure 1.3 shows distortion specifically at high input levels while the portion of the signal clipped off (i.e. the distortion) is depicted in Figure 1.4. Note that in Figure 1.3 the PA input signal has been amplified by the small signal gain value (which is

30.5 dB in this example as can be observed in Figure 1.2) in order to compare it with the PA output signal.
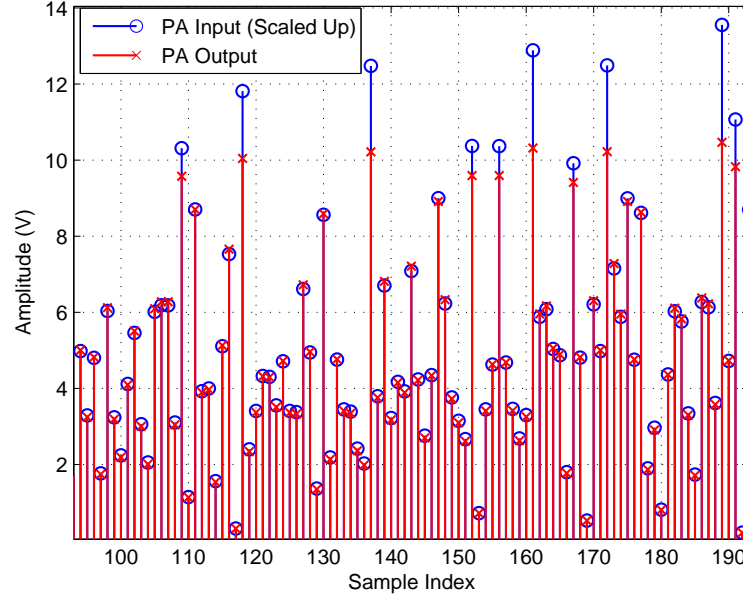
Figure 1.3: Sample input (OFDM) and ouput signals of a commercial amplifier—shows distortion at high input levels
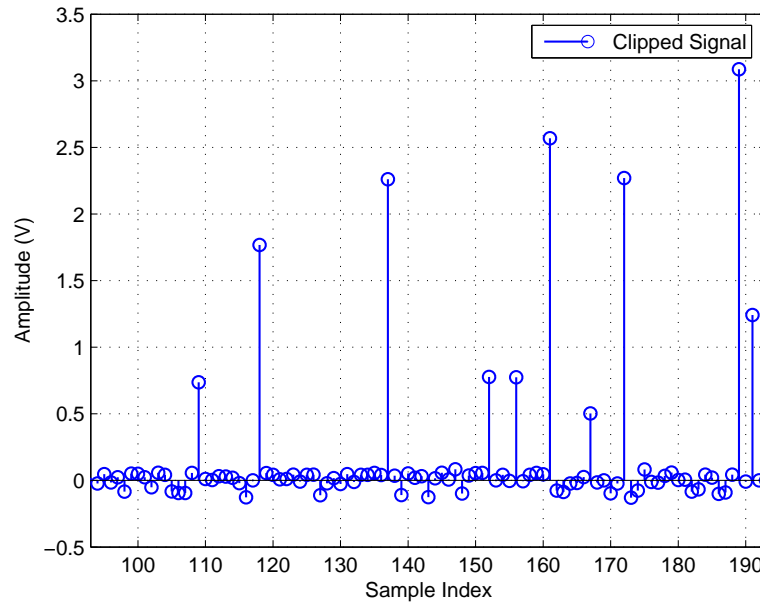
Figure 1.4: Clipped portion of signal in Figure 1.3

It is possible to avoid distortion by operating the PA at low efficiency but this is normally not done. Rather, in the industry, the common solution to this dilemma is to operate the PA at high efficiency and then, use one of several proposed techniques (as discussed in the literature review section that follows) to correct for the resulting non–linearity at the transmitter. Our goal in this work is to correct this distortion at the receiver (i.e. "post–compensation") by estimating it using compressed sensing–based approaches.

1.2   Literature Survey & Objectives

Largely due to the importance and widespread implementation of OFDM, many attempts have been made to correct the non–linear distortion induced by the power amplifier. Since government communication commissions world–wide manage the radio spectrum such that particular bands are assigned to particular services, in many of the standards that employ OFDM, the transmitted signal has to meet specific spectrum mask requirements. Most attempts at removing (or avoiding) the PA's non-linearities have therefore concentrated at the transmitter so that the resulting spectrum at the amplifier's output meets the linearity requirement and does not cause interference between adjacent channels.

A popular and preferred form of these transmitter–based techniques has been to implement upstream of the power amplifier, a "digital predistorter" (DPD) which is built to neutralize the effect of the PA i.e. in such a way that the combination of the predistorter and the (non–linear) amplifier would be a linear amplification sys-

tem [5–8]. Building the predistorter function involves characterizing the PA through a process known as "behavioral modeling" and then inverting the resulting characteristic function [5]. However, because of the extra complexity and power consumption involved, implementing a predistorter becomes an issue in applications where transmitter complexity needs to be kept at a minimum (for example in the up–link in the communication between a mobile device and a base station or in the down-link of satellite communication from a satellite station to an earth base station).

Besides predistortion, quite a number of other techniques have been proposed to mitigate the amplifier's distortion at the transmitter. These include coding schemes [9–11], companding transforms [12, 13], tone reservation [14, 15], and constellation expansion [16, 17]. (Some of these as well as others have been studied and discussed in [18, 19].) However, like predistortion, these solutions are focused on correcting the distortion before transmission and as a result, add varying degrees of complexity to the transmitter—thus making them inappropriate for use in systems where transmitter simplicity is a primary target.

Another suggested solution to avoid or reduce the amplifier's distortion is clipping [20] which involves adding a peak-reducing signal to the transmitted signal and despite having relatively low-complexity, introduces clipping noise leading to performance degradation. In the face of this, some receiver–based techniques to estimate and correct these clips have been proposed [21–26]. Although we also target compensating for distortion at the receiver in this thesis, the distortion we tackle is due to "hard-clipping" (i.e. clipping caused by the PA itself, which cannot be controlled)

unlike these previous works that dealt with distortion caused by "soft-clipping" (i.e. deliberate clipping at transmitter to avoid PA distortion and in which clipping signal can be designed to ease the recovery process at the receiver).

Although it could be argued that delaying distortion correction from the transmitter to the receiver may be inappropriate for some applications—since the spectrum at the transmitter would not meet spectrum mask regulations, there are many applications in the unlicensed industrial scientific and medical (ISM) band where some spectral leakage onto neighboring channels is still acceptable. Examples of such applications include wireless sensor networks (WSN), wireless local area networks (WLAN), radio frequency identification (RFID), Bluetooth, and micro-satellite applications. Also, these applications generally include mobile terminals which are limited in size and operate at low power—making conventional transmitter-based linearization techniques burdensome on these devices both in terms of complexity and power.

In this work, we propose a number of techniques based on the relatively new field of compressed sensing (CS) to "(post-)compensate" for the non–linear distortions caused by the transmitter's power amplifier at the receiver. Compressed sensing aims at recovering signals using much fewer measurements than as proposed by the Shannon Theorem. This is possible given that the signal to be recovered is "sparse" in some domain and the domain in which the measurements are taken is incoherent with that in which the signal is sparse [27–31]. Based on measurement-based amplifier models, we show in this thesis that the power amplifier non–linear distortions are sometimes sparse in the time-domain and we are thus, in such cases, able to employ CS to estimate

and then eliminate these distortions using a few frequency–domain free (pilot) carriers for measurement. Furthermore, to improve the estimate and preserve bandwidth, an intelligent data–aided technique which allows us to apply CS even without including free carriers in transmission is proposed.

Although in [24] and [25], compressed sensing is also applied compressed sensing at the receiver, it is used in quite a different manner. In both prior works, the transmit signal is soft-clipped before transmission and so, knowledge of the clipping signal can be and is used in estimating the clips at the receiver. In our work however, we have no such information at the receiver since we do not soft-clip the signal at the transmitter, rather, the signal is sent directly as it is and is only hard-clipped by the PA itself.

1.3   Contributions

The main points differentiating this thesis from previous work in this area are as follows:

1. To the best of our knowledge, our work presents the first ever attempt to linearize *practical amplifier prototypes* using CS. Unlike [24] and the bulk of other techniques aimed at mitigating PA distortion which implement highly simplified power amplifier models, here, we use realistic amplifier models based on laboratory measurements of power amplifier prototypes (as discussed in Section 2.1).

2. To estimate and remove the PA distortion, we implement both CS and weighted CS algorithms.

3. We enhance these more basic CS–based algorithms using a data–aided technique that gives us the ability to correct the non–linearities without the use of free carriers. We are thus able to use the bandwidth at 100% efficiency unlike common CS techniques.

4. In some situations where pre-compensation is performed at the transmitter, we can overdrive the amplifier beyond saturation point to improve the power efficiency of the system. Ordinarily, such overdrive would cause unwanted distortions at the PA output but based on the work presented in this thesis, we can overdrive the PA and then post-compensate for the resulting distortions at the receiver using the CS–based algorithms thus increasing the system efficiency. This is the final linearization approach presented in this thesis—which we refer to as "joint–compensation" since it combines both forms of compensation.

1.4 Thesis Organization

After the introduction given in this chapter, Chapter 2 goes on to describe the power amplifier models, how they were derived from laboratory measurements, and the problem we aim to solve in this thesis—the compensation of distortion caused by the PA's. Chapter 3 then presents the theory of compressed sensing and explains why and how we are able to apply it in this problem before moving to a detailed discussion of the basic and weighted CS algorithms. The motive and description for the data–aided and joint–compensation techniques come in Chapter 4. Both Chapters 3 and 4 are finalized with a presentation of simulation results of the implementation of each of the

algorithms described in the respective chapters. The thesis is concluded in Chapter 5 with a summary of the results of the research work and a proposal for future work in the area.

1.5   Notation Used

For clarity, this section briefly describes the different font-types and symbols generally used in this thesis (except where stated otherwise) to represent the different variable types. This information is summarized in the following table.

Table 1.1: Summary of notation used

Variable Type	Representation
Constants	Italicized, upper-case letters (e.g., M)
Scalars	Italicized, lower-case letters (e.g., x)
Vectors	Bold-face, lower-case letters (e.g., $\mathbf{x}$)
The ith member of the vector, $\mathbf{x}$	$\mathrm{x}(i)$
Matrices	Bold-face, upper-case letters (e.g., $\mathbf{F}$)
Variable estimates	A hat over the variable (e.g., $\hat{\mathbf{x}}$)
Variables in the frequency-domain	A check over the variable (e.g., $\check{\mathbf{x}}$)
Transpose of the matrix, $\mathbf{F}$	Superscript T (e.g., $\mathbf{F}^{\mathsf{T}}$)
Complex conjugate (Hermitian) transpose of the matrix, $\mathbf{F}$	Superscript H (e.g., $\mathbf{F}^{\mathsf{H}}$)

CHAPTER 2

POWER AMPLIFIER MODELS

& PROBLEM DESCRIPTION

Since our use of measurement based amplifier models is one of the main distinctive points of this work, this chapter serves to describe the amplifiers used, their models and how some of these models were experimentally derived from the measurements performed on the actual amplifiers. From the discussion about the amplifiers models, we will be able to give more details about the distortion-correction problem.

It should be noted however that although, in general, power amplifiers exhibit two types of non-linearities—the static (also referred to as "memoryless non-linearity") and the dynamic (also referred to as "memory effects"), we only deal with static distortions here. The difference between the two forms is that for the static distortions, the instantaneous complex gain of the PA is a function of the input sample at that instant only while in the case of memory effects, the instantaneous complex gain is a function of not just the current input sample but also a finite number of previous

input samples [5]. In this work, we assume the amplifiers have no memory effects and so, in the models we use—as detailed in Section 2.1—the PA gain is only a function of the current input sample i.e. all forms of distortion treated are static.

We should also point out that, in general, all power amplifiers have maximum input levels beyond which the amplifiers will saturate. For this reason, signals have to be scaled down to meet the maximum input requirements of the amplifier under test. Thus, here, it is the scaled–down version of the OFDM signal that is the actual PA input.

2.1   PA Models Used: Memoryless Polynomial & Look-Up Table

In this work, we have studied the Memoryless Polynomial (MP) and the Look-Up Table (LUT) models to simulate the behavior of the power amplifier. The following two subsections (2.1.1 and 2.1.2) summarily describe both models and present characteristic plots for the particular amplifiers studied in this work. Further explanation is available in [5]. Note that, in order to aid the comparison of the distortion caused by the amplifiers, the y-axis of all figures of the amplitude responses have been set to a range of 11 dB while the y-axes of the phase response figures have been set to a range of 30°

.

2.1.1 Memoryless Polynomial Model

This model is actually a particular variation of a more general model—the memory polynomial—which itself is a simplification of the more comprehensive Volterra Model used for dynamic non-linear systems. The memory polynomial is popularly used for power amplifiers exhibiting memory effects and its output is given by [32, 33]:

$$x_{mp-out}(n) = \sum_{j=0}^{M_{mp}} \sum_{i=0}^{N_{mp}} a_{ij} \cdot x_{mp-in}(n-j) \cdot |x_{mp-in}(n-j)|^{i-1} \qquad (2.1)$$

where $x_{mp-out}(n)$ is the instantaneous model output, $x_{mp-in}(n)$ is the instantaneous model input, N_{mp} and M_{mp} are the non-linearity order and the memory depth of the MP model respectively, and a_{ij} are the model coefficients.

In our implementation, $M_{mp} = 0$ (and therefore, our model is **memoryless**) while $N_{mp} = 13$. The PA characteristic curves—showing the non-linear behavior of the model—are presented in Figures 2.1 and 2.2. It is noticed from the plot of the amplitude response in Figure 2.1 the PA's non-linear gain which ranges over 3.5 dB (from 73 dB to 76.5 dB). The phase response in Figure 2.2 shows that the phase shift is also non-linear with a 16^{circ} range. Figure2.3 shows examples of the distortion with a particular set of input amplitudes compared to the corresponding output amplitudes (scaled down for comparison) while the resulting clips in this example case are illustrated in Figure 2.4.

It should be noted that the negative coefficients in the clipped portion in Figure 2.4 result when the linearly-amplified PA input is less than the actual PA output.
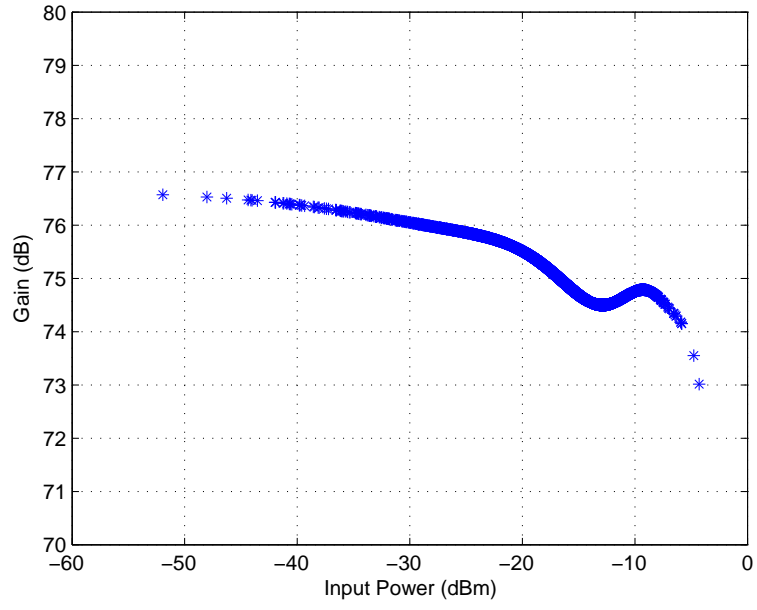
Figure 2.1: Measured gain response of MP–modeled PA
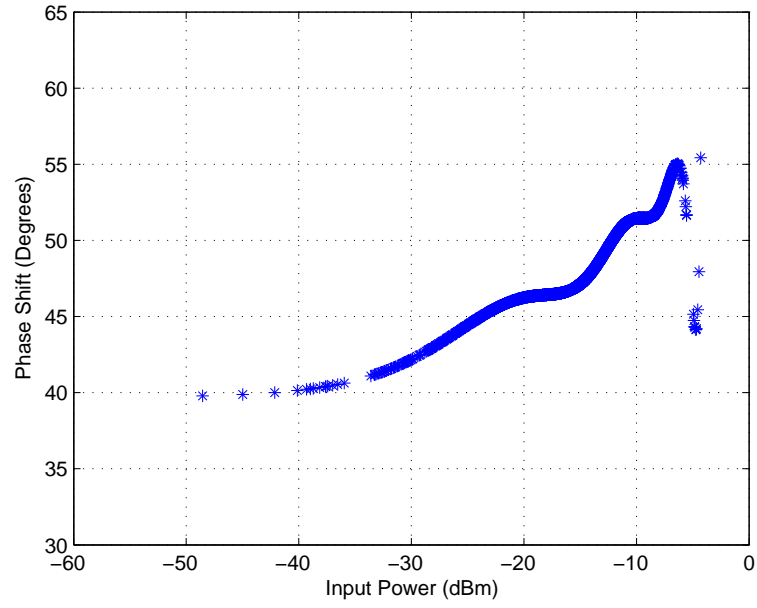
Figure 2.2: Measured phase response of MP–modeled PA

15
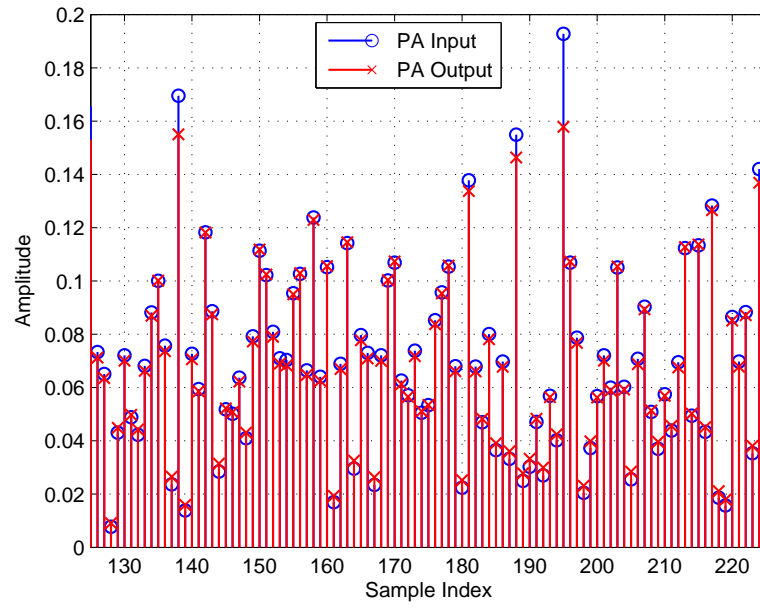
Figure 2.3: Sample input and output signals for MP–modeled PA
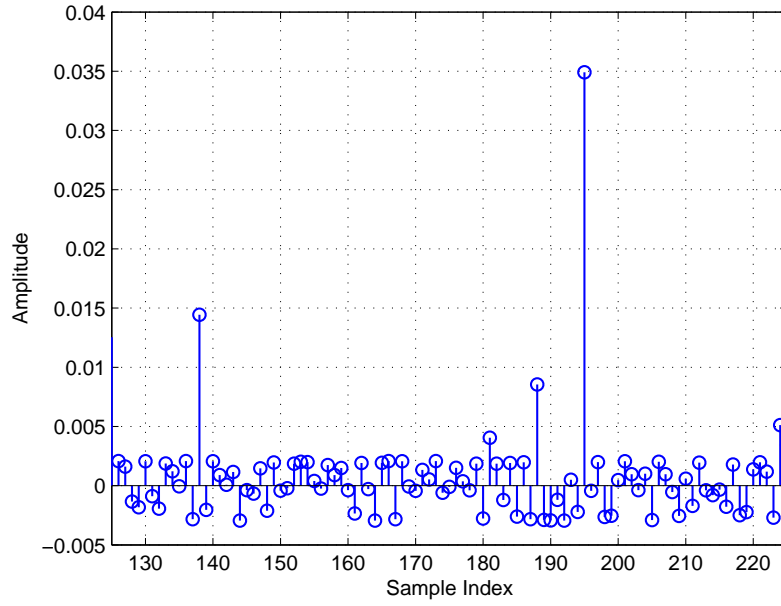
Figure 2.4: Sample clipped portion for MP–modeled PA

2.1.2 Look-Up Table Model

The look-up table model is the primary model for the static non-linearities of power amplifiers. It is a relatively simple model where a wide range of possible amplifier inputs and their corresponding (complex) gain values are saved in a look-up table so that for any given input, we simply interpolate linearly on the table to find the appropriate gain and then multiply this gain value by the input to obtain the respective output value. The output of the LUT is given by:

$$x_{lut-out}(n) = G(|x_{lut-in}(n)|) \cdot x_{lut-in}(n) \tag{2.2}$$

where $x_{lut-in}(n)$ and $x_{lut-out}(n)$ are the input and output signals respectively and $G(|x_{lut-in}(n)|)$ is the instantaneous complex gain of the PA.

The following section gives more details about specific amplifiers used in this work which were modeled using the LUT model.

2.2 Amplifiers Modeled using Look-Up Table

Due to its implementation simplicity, the look-up table model has been used for most of this work. The following are the code names for the different amplifiers modeled using the look-up table. Further description for each is given in the following sections.

1. RFMD

2. Doherty

3. ZHL

4. GaN

2.2.1 RFMD

This refers to a commercial PA from RF Micro Devices of North Carolina, USA with part number RF5198. It is designed for handset applications in the 1920MHz to 1980MHz frequency band. As can be seen in the measured gain characteristic in Figure 2.5, this amplifier is mostly linear and significant non-linear compression only comes in when the input power levels exceed 0 dB and for a range of 5 dB (i.e. from 0 dBm to 5 dBm). The phase difference, as seen in Figure 2.6 has a very short range of less than 5° (between 46.5° and 51°) and thus has very little effect on the distortion. Considering the amplitude response, it is expected that only a relatively small number of data samples would fall in the non-linear region and suffer substantial distortion. This is obvious from the sample input and output signals in Figure 2.7 as well as the corresponding clips shown in Figure 2.8.

Figure 2.5: Measured gain response of RFMD PA
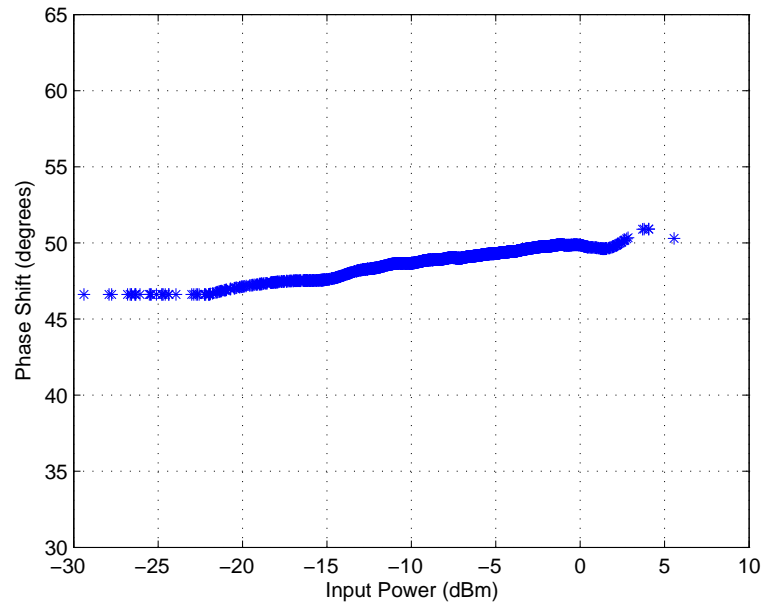
Figure 2.6: Measured phase response of RFMD PA
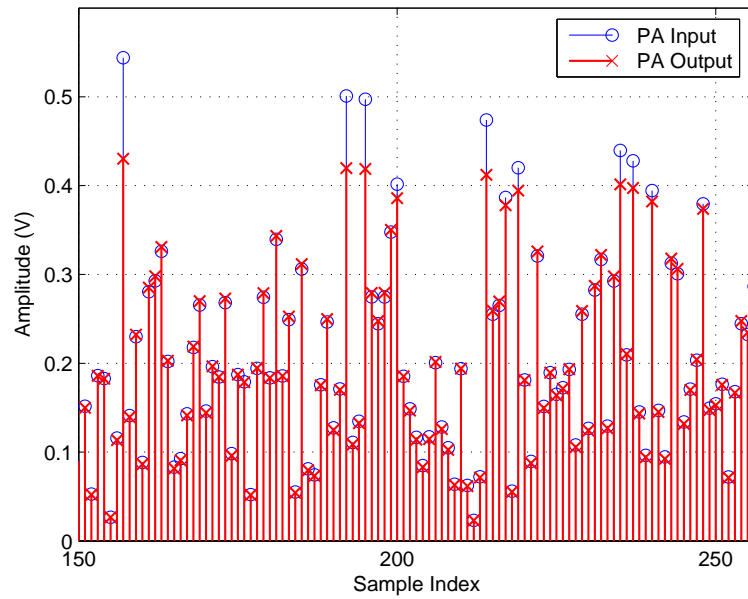
Figure 2.7: Sample input and output signals for RFMD PA
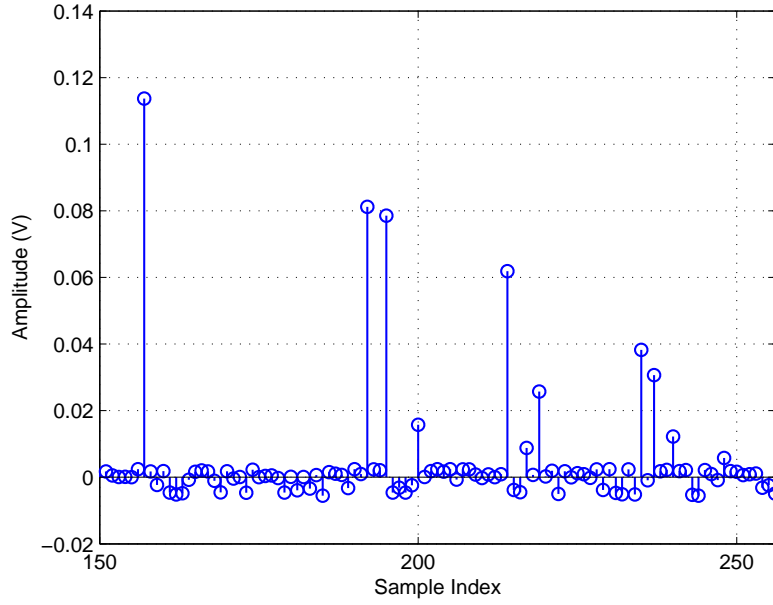
Figure 2.8: Sample clipped portion for RFMD PA

2.2.2 Doherty

The Doherty PA used here is a high-power high-efficiency laterally diffused metal oxide semiconductor (LDMOS)-based Doherty amplifier with 300 W peak power and 61 dB small-signal gain. This PA is designed for operation in the $2110 - 2170$ MHz frequency band.

Unlike the RFMD amplifier, this Doherty amplifier has a very non-linear response as illustrated in Figure 2.9. Even for low input levels, although it is relatively linear, it still has some distortions and these greatly increase at high input levels and range for about 10 dB (between -17 dBm to -7 dBm). Since much of the scaled input signal would fall in this range, the amount of distortion multiplies as compared to the RFMD amplifier. To makes things worse, the phase shift as shown in Figure 2.10 has a $30°$ range that aggravates the distortion and leads to a dense clipped signal as the

sample in Figure 2.12 shows.

Figure 2.9: Measured gain response of Doherty PA

Figure 2.10: Measured phase response of Doherty PA
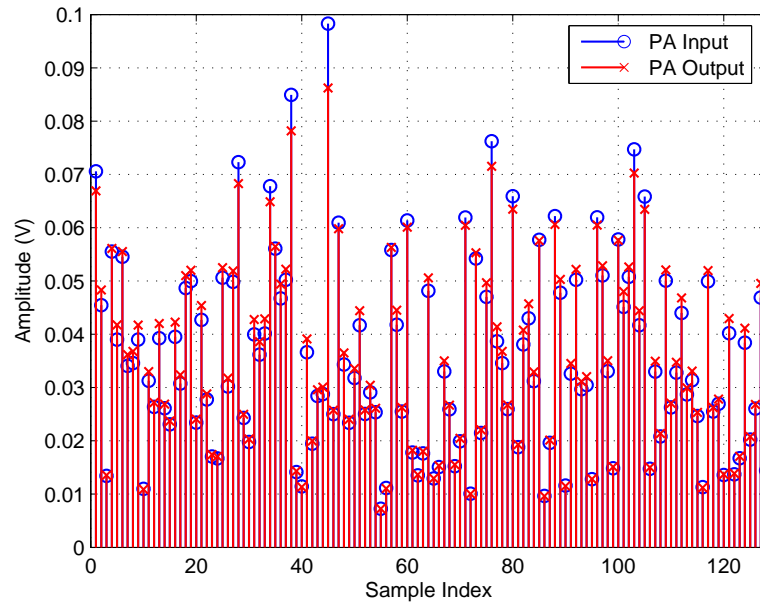
Figure 2.11: Sample input and output signals for Doherty PA
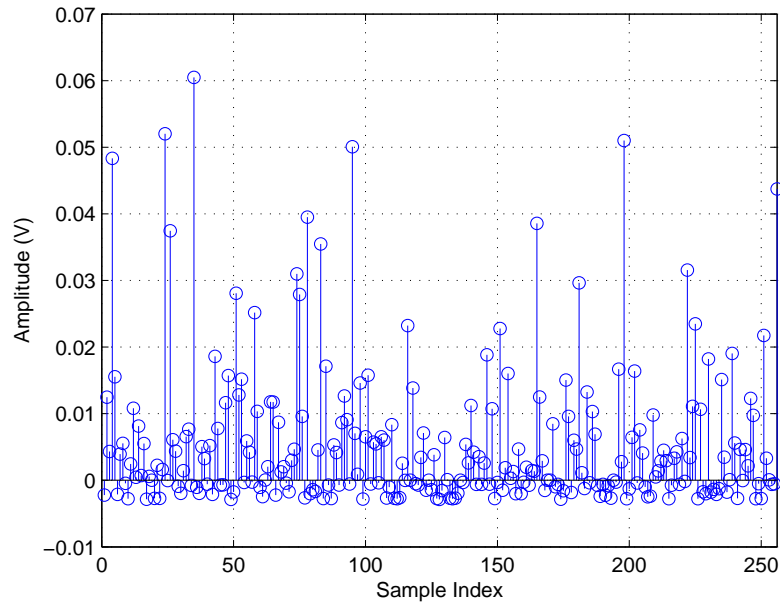
Figure 2.12: Sample clipped portion for Doherty PA

2.2.3 ZHL

ZHL refers to the "ZHL-42" which is the industry-assigned code name for a medium-high power instrumentation amplifier that operates in the frequency range of $700 - -4200$ MHz and has a small signal gain of 30 dBm. The gain response of this amplifier is somewhat similar to that of the RFMD amplifier except that it is even more linear for the initial 30 dB range of input power levels and only gets non-linear for a relatively short input range (-6 dBm to 2 dBm) as shown in Figure 2.13. Also, the phase shift of this amplifier (see Figure 2.14) is of a limited range ($-92°$ to $-89°$). As a result of this, the major clips are relatively few—much less than the Doherty amplifier but more than the RFMD since the non-linearity in the gain response lasts for a slightly longer input power range. A sample of these clips (corresponding to the input and output signals in Figure 2.15) is illustrated in Figure 2.16.
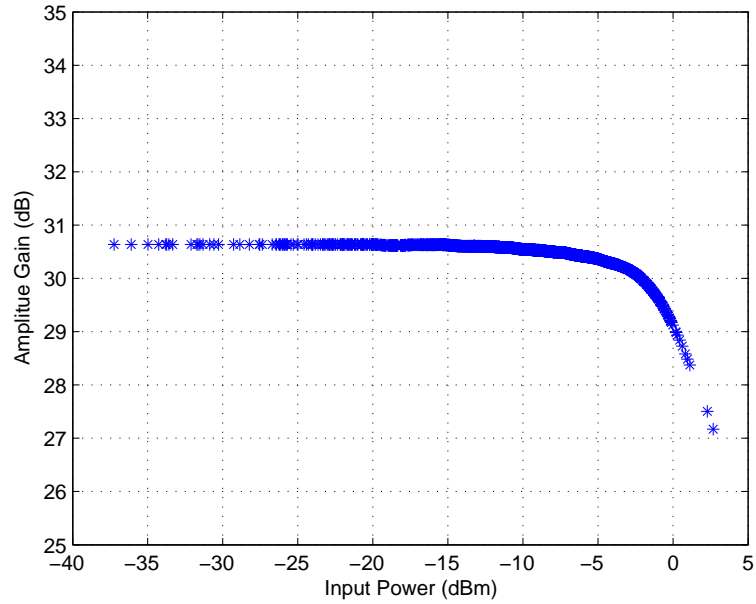
Figure 2.13: Measured gain response of ZHL PA
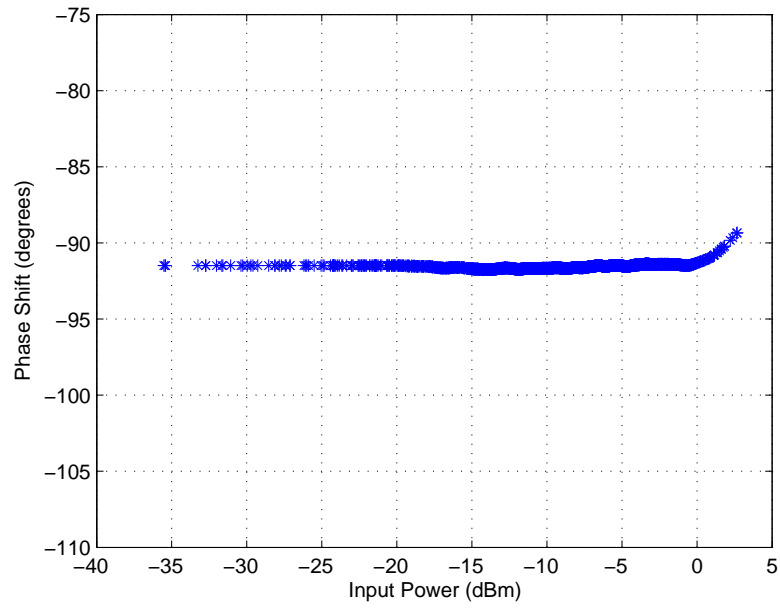
Figure 2.14: Measured phase response of ZHL PA
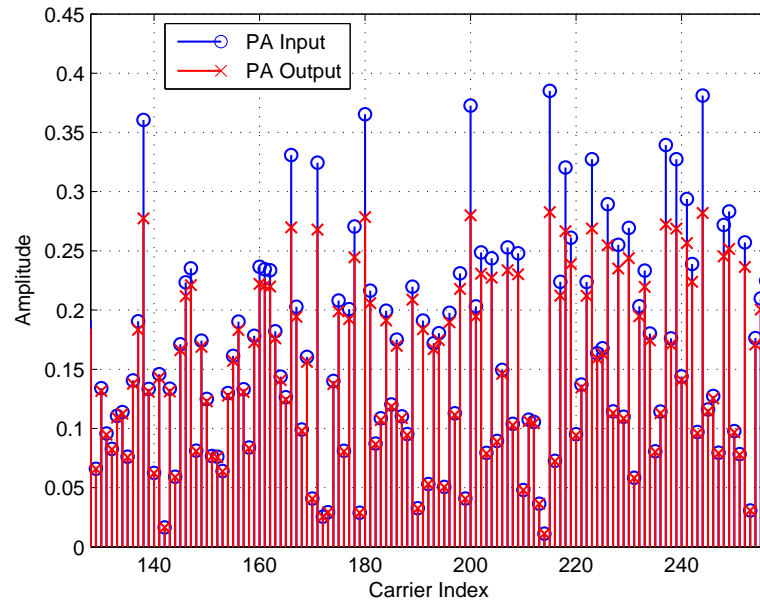
Figure 2.15: Sample input and output signals for ZHL PA
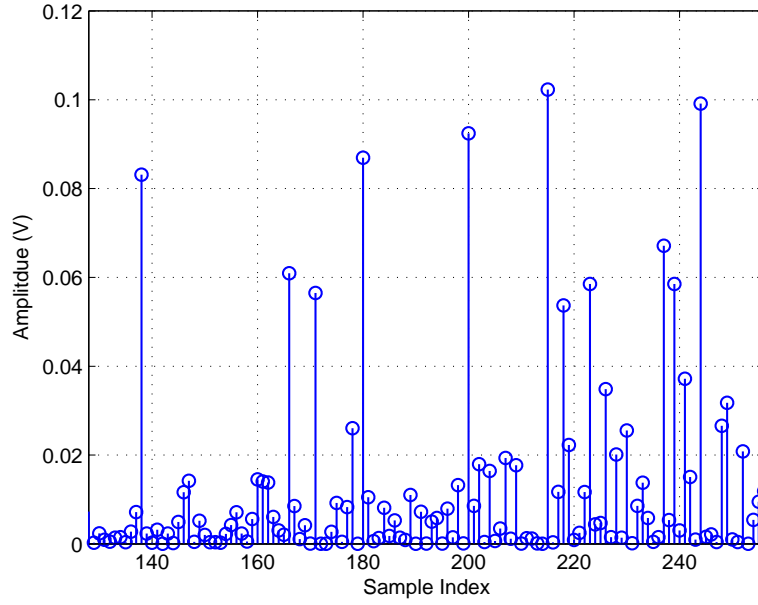
Figure 2.16: Sample clipped portion for ZHL PA

2.2.4   GaN Amplifier

The Gallium Nitride (GaN) amplifier is a medium power Doherty amplifier which has been designed for use in micro-satellite applications.We see from Figure 2.17 that just like the previous Doherty amplifier, the distortion in the GaN also extends over a wide range (about 16 dB) of input power levels(-20 dBm to -4 dBm)—which is in fact a much larger range than that of the Doherty amplifier described in Section 2.2.2. On the other hand, unlike the previous Doherty PA, the maximum phase shift as shown in Figure 2.18 is only about $3°$ and so, this has only little effect on the distortion. However, due to the very non-linear amplitude response, much of the input signal would fall in non-linear region resulting in much distortion as we observe in the sample input and output signals in Figure 2.19 and the corresponding clips in Figure 2.20.
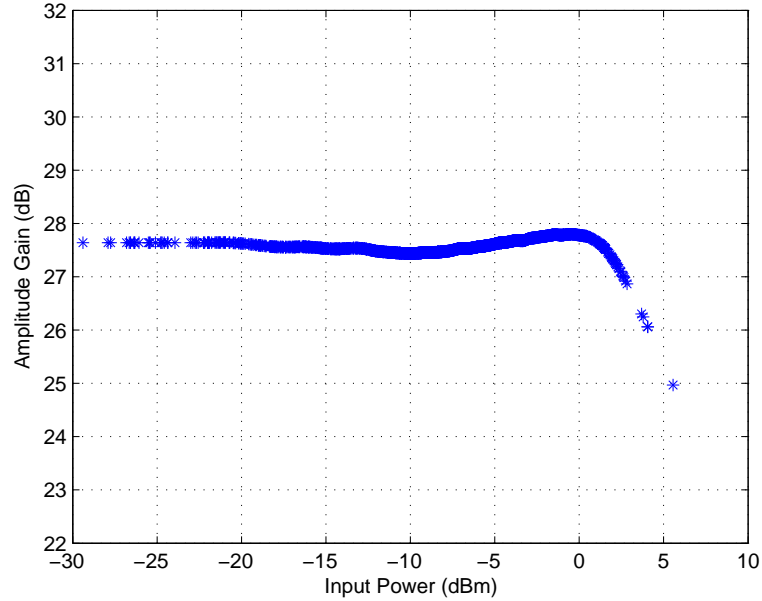
25
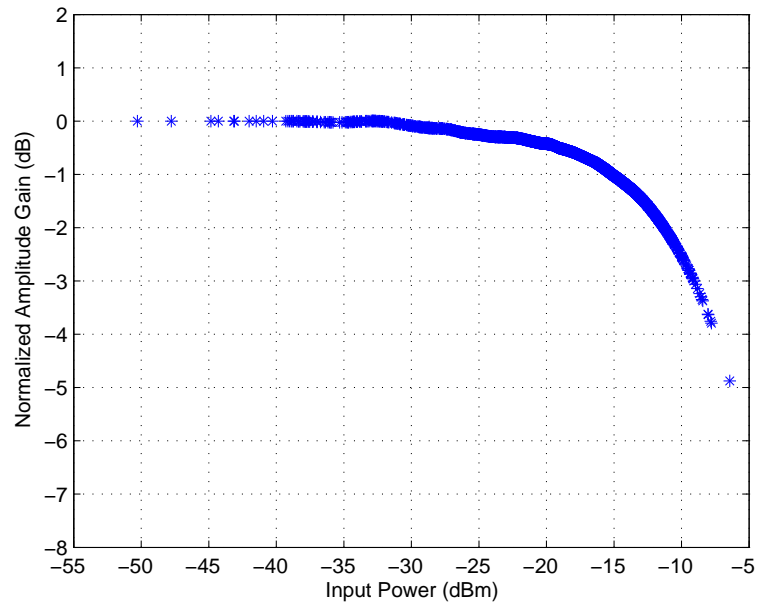
Figure 2.17: Measured gain response of GaN PA
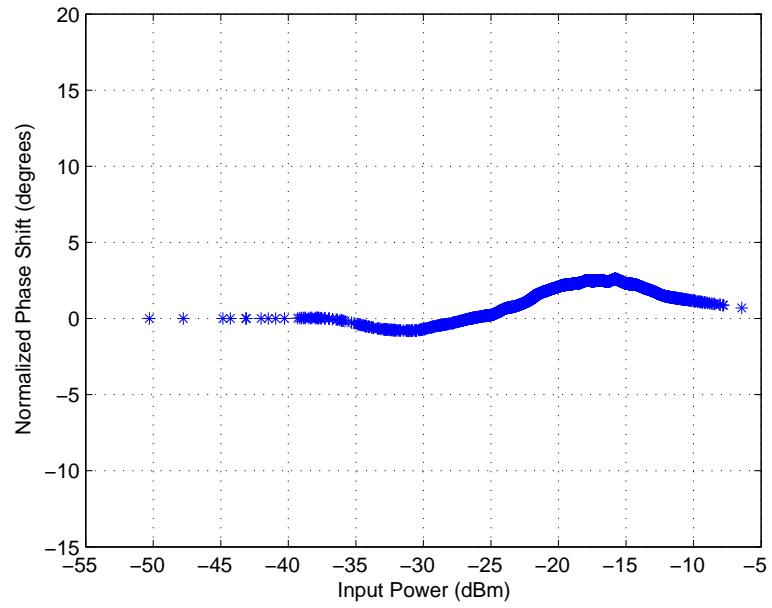
Figure 2.18: Measured phase response of GaN PA

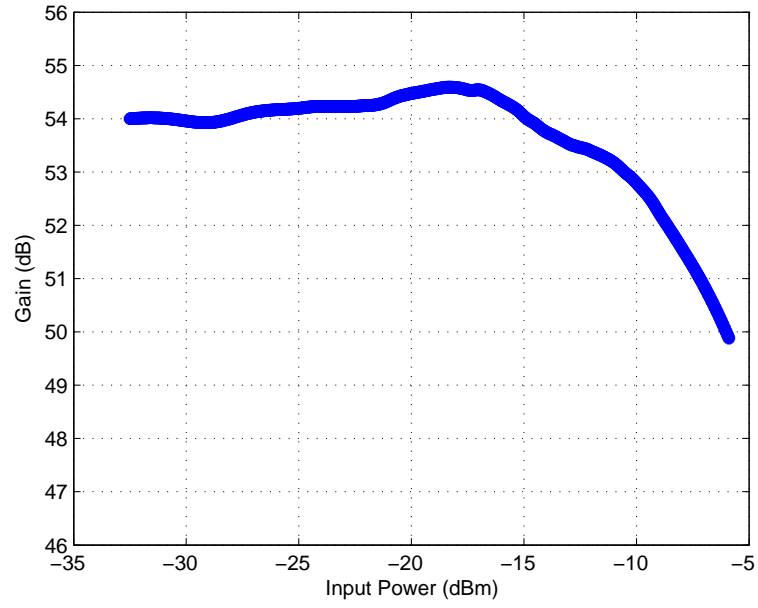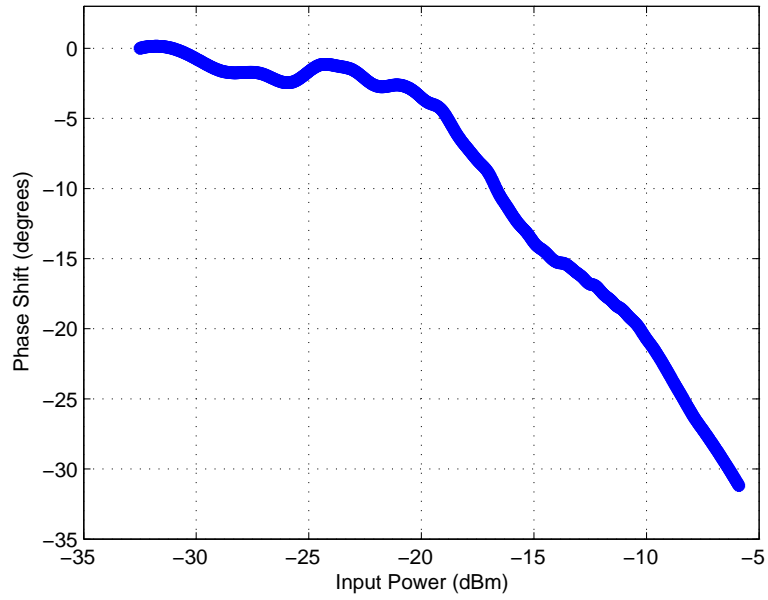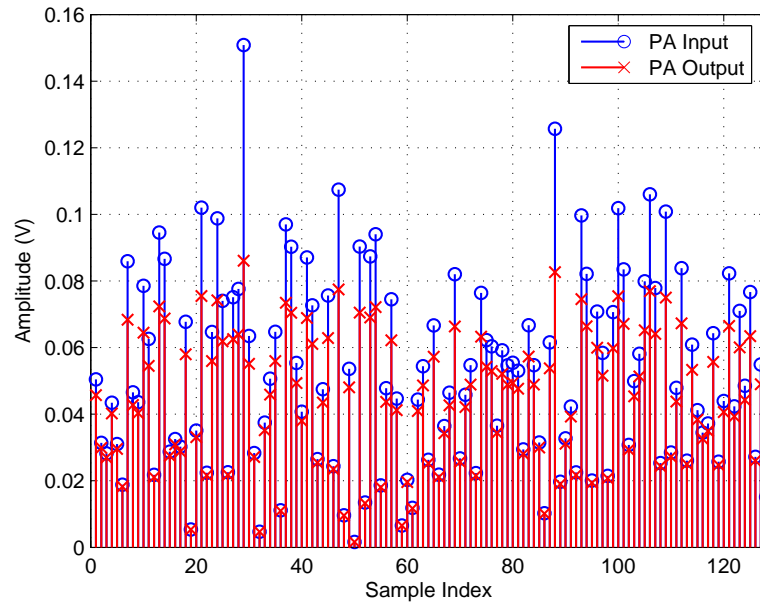**Figure 2.19:** Sample input and output signals for GaN PA
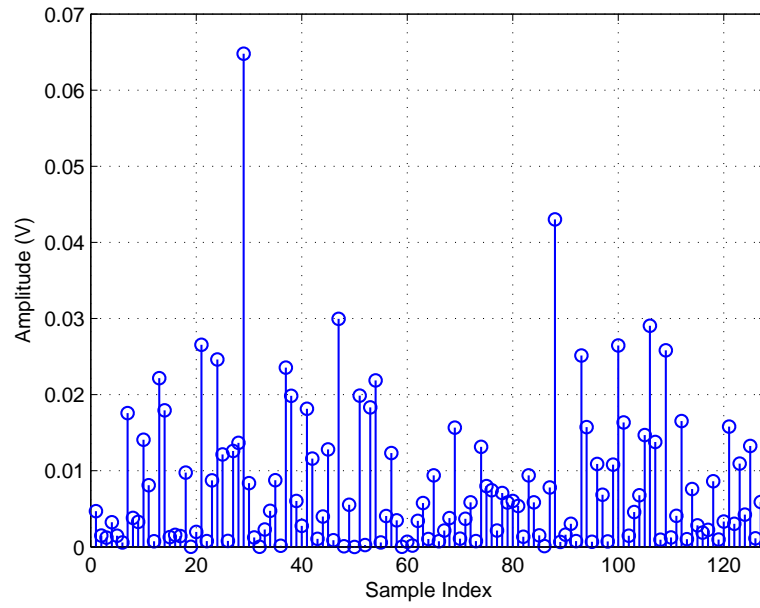
Figure 2.20: Sample clipped portion for GaN PA

2.3 Comparing the Amplifier Models Used (in terms of Gini Sparsity Index)

In the preceding sections that describe the amplifier models, we have touched on some differences between the gain and phase responses of the amplifiers and how these affect the amount of distortion caused by each amplifier. In this section, we use a sparsity measurement scale known as the Gini Sparsity Index [34] to compare and rank the four amplifiers modeled using the LUT according to the non-linearities they induce.

We are particularly interested in obtaining a measure of the sparsity because, besides giving us some information about how much distortion we are to correct, it helps us to know how effective our compressed sensing algorithm would be in estimating the distortion. This is because, as we describe in Chapter 3, an important requirement for CS to perform properly in recovering a signal is that the signal should be sparse. Thus, the more sparse a signal is, the better CS performs in estimating the signal and the more dense it becomes, the worse CS performs.

Of the 16 commonly used sparsity measures studied in [35], the Gini Index was found to be one of only two that meet all six attributes desirable for a measure of sparsity discussed in [35]. The index, denoted by $\mathcal{G}_s$, rates the sparsity of a signal from $\mathcal{G}_s = 0$ (very dense) to $\mathcal{G}_s = 1$ (very sparse, a single impulse). To calculate the Gini Index for a given vector, $\mathbf{x} = [\mathrm{x}(0)\ \mathrm{x}(1)\ \mathrm{x}(2)\ \ldots\ \mathrm{x}(N)]$ where N is the vector length, we first sort the vector from the minimum to the maximum value so that we have: $\mathrm{x}(\bar{0}) \leq \mathrm{x}(\bar{1}) \leq \mathrm{x}(\bar{2}) \leq \ldots \leq \mathrm{x}(\bar{N})$ where $\bar{0}, \bar{1}, \bar{2}, \ldots, \bar{N}$ are the new indices of the sorted

vector. The Gini Index is then found thus [34, 35]:

$$\mathcal{G}_s(\mathbf{x}) = 1 - 2 \sum_{i=1}^{N} \frac{\mathrm{x}(\bar{\mathrm{i}})}{\|\mathbf{x}\|_{l_1}} \left(\frac{N - i + \frac{1}{2}}{N} \right) \tag{2.3}$$

Since the distortion caused by every amplifier would vary depending on the input signal, to obtain a Gini Index value for the amplifiers studied, the average Gini Index for a very large number of input signals was taken. For this reason, the Gini values reported are independent of the actual value of N[1]. Table 2.1 presents a comparison of the four amplifiers modeled by the LUT model in terms of distortion due to their gain and phase responses and their Gini Sparsity Index values.

Table 2.1: Comparison of amplifier models

Amplifier Code Name	Gain Distortion Range	Phase Distortion Range	Gini Sparsity Index, $\mathcal{G}_s$
RFMD	5 dB (-5 dB $-$ 0 dB)	5°	0.718
Doherty	10 dB (-17 dB $-$ -7 dB)	30°	0.461
ZHL	8 dB (-6 dB $-$ 2 dB)	3°	0.708
GaN	16 dB (-20 dB $-$ -4 dB)	3°	0.564

As we expect, we notice in Table 2.1 that the Doherty has the lowest Gini Index value (i.e. its distortion is most dense) while the RFMD has the maximum values (its distortion is most sparse). For this reason, although all four amplifiers were studied, in this report, we present only the results of compensating for the distortion caused by the RFMD and Doherty amplifiers—thus working on both extremes—and the algorithms described can easily be extended and applied to the other amplifiers.

[1]Although, in this thesis, N was generally taken to be 256.

2.4 Experimental Set-up for deriving LUT [2]

In this section, using the RFMD amplifier as an example, we describe the experimental process of deriving the look-up tables for the amplifiers considered. The layout of the setup is shown in Figure 2.21 and the actual setup used is shown in Figure 2.22.
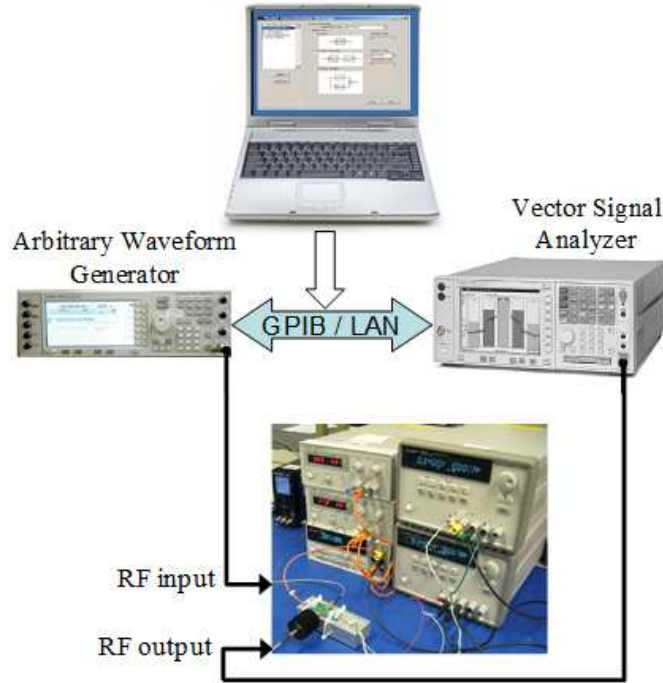
Figure 2.21: Layout of experiments for LUT measurements

We first measure the amplitude and phase responses of the RFMD amplifier by driving the amplifier with an appropriate signal (a 5 MHz LTE signal based on OFDM modulation) loaded via a signal generator. The output of the amplifier is then analyzed using a vector signal analyzer (VSA). The amplitude and phase of the input

[2]The experiment described in this section and as pictured in Figure 2.22 was performed at the iRadio Lab. a the University of Calgary, Calgary, Canada.

30
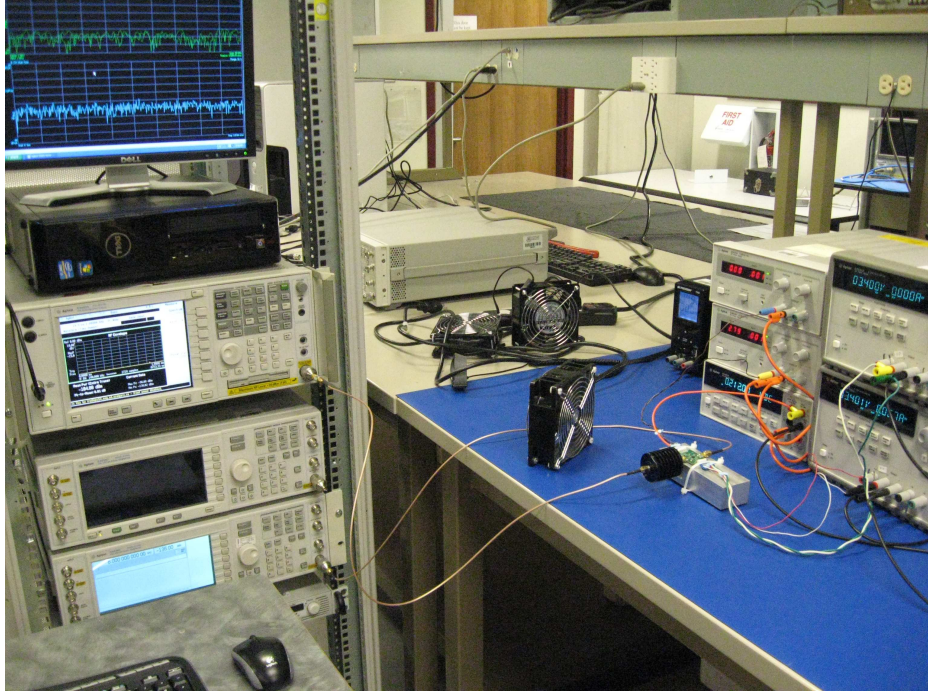
Figure 2.22: Actual experimental setup for LUT measurements

(loaded from the generator) is plotted against the corresponding amplitude and phase
of the output (as analyzed by the VSA) to give the amplitude and phase response
curves shown in blue in Figure 2.23. These curves are then fitted using the dou-
ble exponentially weighted moving average (dEWMA) method [36] and the resulting
average response in given by the red curves in Figure 2.23.
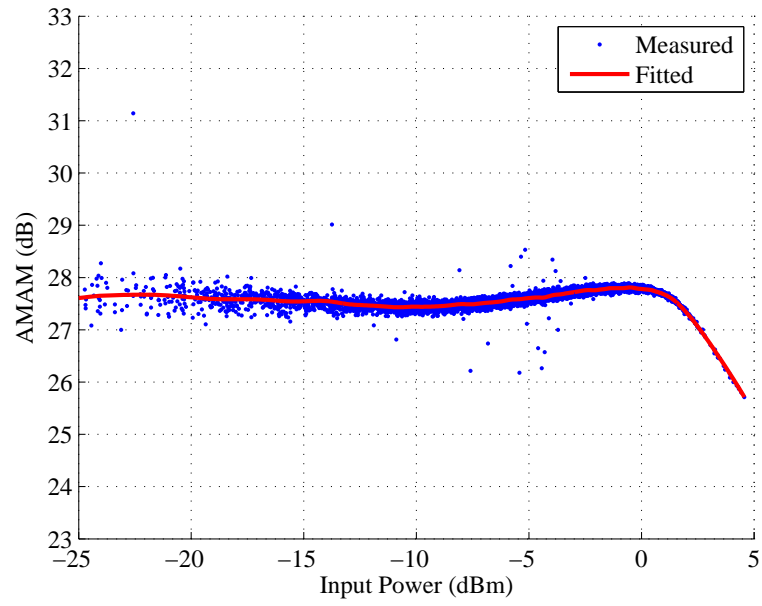
(a)
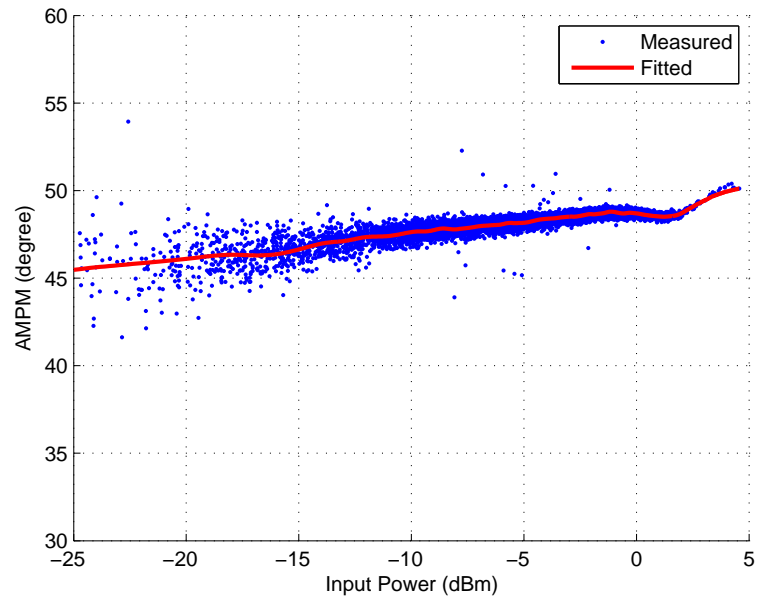
(b)

Figure 2.23: Measured and fitted characteristics of the amplifier. (a) Amplitude characteristics; (b) Phase characteristics.

CHAPTER 3

POST-COMPENSATION USING

COMPRESSIVE SENSING

3.1 Introduction to Compressed Sensing

Compressed sensing (CS, also interpreted as "compressive sensing", "compressed sampling" and "compressive sampling") is a novel signal recovery technique that allows for signal reconstruction even when the sampling frequency is much lower than the Nyquist rate.

It is commonly believed that, according to Shannon's theorem, for a signal to be correctly recovered it needs to be sampled at a minimum rate of twice its maximum frequency i.e. the Nyquist frequency. However, [37–39] show that via compressed sensing, a signal can be reconstructed from much fewer measurements. This is possible due to two fundamental prerequisites in the structure of a CS problem:

1. **Sparsity:** This applies to the signal being recovered and alludes to the require-

ment that the signal, when represented in a particular domain, has most of its coefficients equal to zero and only few to be non-zero i.e. the signal is **sparse** in the given domain. An N-dimensional vector, $\mathbf{x} \in \mathbb{C}^N$ having only $s \ll N$ non-zero coefficients is said to be "S-sparse" or to have "a sparsity of S".

2. **Incoherence:** This second condition pertains the domain in which a signal is sampled. If the signal were to be sampled in the domain in which it is sparse, then, with very high probability, most of the samples would be zero. Rather, the measurements should to taken in a domain in which the signal is spread out i.e. one that is **incoherent** with the domain of sparsity such that all samples taken have some element of all the coefficients to be estimated.

As an example, to recover a signal that is sparse in the time domain, measurements should be taken in the frequency domain since these two domains are maximally incoherent with each other.[1] The incoherence of the time and frequency domains is shown in the impulse noise example in Figure 3.1 where impulse noise has been added in the time domain of an OFDM signal (which included free–carriers in the frequency domain). Despite the distortion caused by the impulse noise—as shown in Figures 3.1a and 3.1b—affecting only two signal coefficients in the time domain, it is spread out and distorts all frequency–domain coefficients (Figures 3.1c and 3.1d). This shows that a signal that is sparse in the time domain (which is the distortion in this example) is wide-spread in the frequency domain—this means that the two domains are incoherent.

[1]As a side note, it turns out that random matrices are mostly incoherent with any given basis [27, 28].
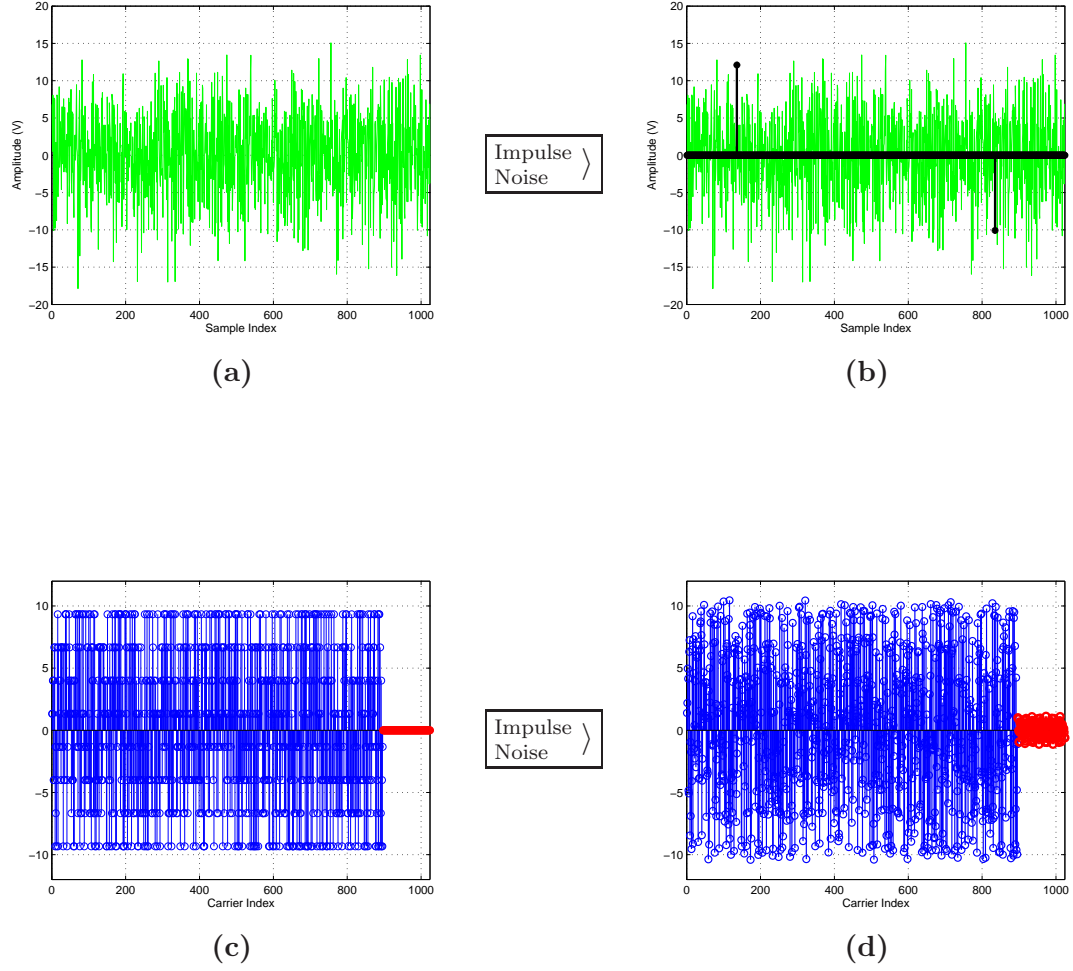
Figure 3.1: Incoherence between the time & frequency domains - impulse noise example. (a) Time domain representation of a signal before impulse noise (b) Time domain representation of the signal with impulse noise added (in black) at two locations (c) Frequency domain representation of the signal before impulse noise (d) Frequency domain representation of the signal after impulse noise shows impact (distortion) on all coefficients.

3.2 The Compressed Sensing Algorithm

Given an under-determined system such as:

$$\mathbf{b} = \mathbf{A}\mathbf{x} + \mathbf{z} \tag{3.1}$$

where $\mathbf{b}$ is an $M \times 1$ known vector, $\mathbf{A}$ is an $M \times N$ known measurement matrix, $\mathbf{z}$ is an $M \times 1$ noise vector, M is the number of measurements, N is the number of unknowns, and $M << N$. Our goal is to find the sparsest $N \times 1$ vector, $\mathbf{x}$ which solves (3.1). Thus, our first thought might be to simply use l_0-minimization:

$$\min_{\tilde{\mathbf{x}}} \|\tilde{\mathbf{x}}\|_{l_0} \quad \text{subject to} \quad \|\mathbf{A}\tilde{\mathbf{x}} - \mathbf{b}\|_{l_2} \le \epsilon \tag{3.2}$$

where ϵ restricts the amount of noise in the recovered data. To solve this problem, we need to search through the combination, $\binom{N}{S}$ possibilities (if $\mathbf{x}$ is S-sparse)—and so, when N is large, (as it usually is in OFDM), l_0-minimization becomes virtually impossible. However, as discussed in [27], given some particular conditions, the results of l_0-minimization are exactly replicated by l_1-minimization:

$$\min_{\tilde{\mathbf{x}}} \|\tilde{\mathbf{x}}\|_{l_1} \quad \text{subject to} \quad \|\mathbf{A}\tilde{\mathbf{x}} - \mathbf{b}\|_{l_2} \le \epsilon \tag{3.3}$$

where $\|\tilde{\mathbf{x}}\|_{l_1} := \sum_i |\tilde{\mathbf{x}}_i|$. This is a convex problem for which there are many tools available to solve. Besides l_1-minimization, there are other methods that can be used to solve for the sparsest $\mathbf{x}$: for example, greedy pursuit algorithms [40–42] and

36

Bayesian methods [43, 44].

To correctly estimate an S-sparse vector, $\mathbf{x}$, we need to have the number of measurements, M to be [27]:

$$M \geq S \log \left(\frac{N}{S} \right) \tag{3.4}$$

However, it has been found through experimentation [27, 45] that as a rule of thumb it is enough to use:

$$M \geq 4S \tag{3.5}$$

.

3.3 CS for PA Distortion Compensation: Why & How?

This section seeks to clarify why we are able to apply compressed sensing in compensating for the power amplifier's non-linear distortions and discusses briefly some of the main steps we take.

From the amplifier examples given in Section 2.1, we see that the signal portion clipped off by the PA ("the clipped signal") is, in general, close to being sparse since many of its coefficients are very close to zero (even though they might not be exactly zero)[2]. Figure 3.2 shows this characteristic in a sample of the distortion caused by the RFMD amplifier. We are thus able to use CS to estimate the large distortions since

---

[2]In proper terms, such a signal is referred to as a "compressible" signal and results (as detailed later on in this chapter) show that CS is very much applicable even in this case.

they are sparse relative to the rest of the clipped signal we are to estimate.
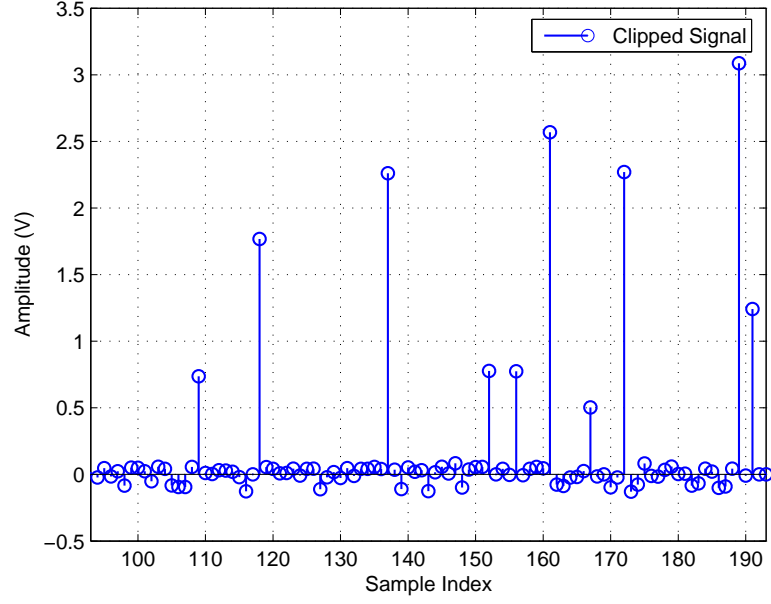
Figure 3.2: Clipped portion of RFMD PA input due to PA distortion

Regarding how compressed sensing is applied to this problem, it was previously mentioned in Section 3.1 that we need to take measurements in a domain which is incoherent with the domain where the signal is sparse. Since the distortion caused by the PA is relatively sparse in the time domain, we therefore sample in the frequency domain. We achieve this by randomly inserting free / zero / null carriers in the transmitted OFDM symbol. As explained in Section 3.1, any distortion that occurs in the time domain—such as the PA's non-linearities—would affect all frequency–domain carriers including the null ones. Thus, the frequency–domain data on the free carriers at the receiver side can serve as measurements to estimate the time–domain distortion.

3.4 Transceiver Model: Transmitter, Power Amplifier and Receiver

3.4.1 Transmitter

In OFDM, the serial stream of data, $\mathbf{d}$ to be transmitted is first divided into N parallel streams each of which is modulated (usually using some form of phase-shift keying (PSK) or quadrature amplitude modulation (QAM)) onto a sub-carrier to obtain a set of N data symbols. However, in our case, we need to reserve M free carriers for measurement for the CS algorithm, and so, we are left we $K = N - M$ data symbols, $\check{\mathbf{x}} = [\check{x}(0)\ \check{x}(1)\ \check{x}(2)\ \ldots\ \check{x}(K-1)]$. After inserting free tones and taking the inverse Discrete Fourier Transform (IDFT) of the resulting data set, our OFDM signal (which is the input of the power amplifier) becomes:

$$\mathbf{x}_i = \mathbf{F}^{\mathsf{H}} \mathbf{S}_K \check{\mathbf{x}} \tag{3.6}$$

where $\mathbf{F}^{\mathsf{H}}$ is the Hermitian conjugate of the $N \times N$ DFT matrix, $\mathbf{F}$:

$$[\mathbf{F}]_{k,l} = \frac{1}{\sqrt{N}} e^{-\frac{j2\pi kl}{N}}, \quad k, l \in 0, 1, \ldots, N-1 \tag{3.7}$$

and $\mathbf{S}_K$ is an $N \times K$ selection matrix having all entries as zero except for a single entry in each of the K columns which equals one. The multiplication $\mathbf{S}_K \check{\mathbf{x}}$ thus gives us an $N \times 1$ vector with $M = N - K$ zeros representing the free tones. Since we need these free tones to be randomly spread in the signal (due to the requirements of the

CS algorithm), the ones in the $\mathbf{S}_K$ matrix should be randomly placed.

To avoid inter-symbol interference, a cyclic prefix is appended to $\mathbf{x}_i$ before the OFDM signal advances to the power amplifier stage.

3.4.2 Power Amplifier

The models used for the power amplifier have already been described extensively in Chapter 2. Each of these amplifiers have maximum input power restrictions and so the OFDM signal has to be scaled so that its maximum power does not surpass that allowed by the particular amplifier in use. Thus the scaled input is:

$$
\begin{aligned}
\mathbf{x}_s &= c\mathbf{x}_i \\
&= c\mathbf{F}^H\mathbf{S}_K\check{\mathbf{x}}
\end{aligned}
\tag{3.8}
$$

where $0 \leq c \leq 1$ is the scale factor which varies from amplifier to amplifier. Also, in general, for any power amplifier model used, we can model the distorted output as an addition of a linearly amplified input signal and a (negative) clipped signal and thus write the power amplifier output, $\mathbf{x}_p$, as:

$$
\mathbf{x}_p = g\mathbf{x}_s + \mathbf{x}_c
\tag{3.9}
$$

where g is a constant amplification value which is an approximation of the PA's small signal gain and $\mathbf{x}_c$ is a $N \times 1$ vector representing the portion of the signal clipped by the PA (i.e. the distortion caused by the PA).

Combining equations (3.8) and (3.9), we obtain the following equation for the

transmitted signal:

$$\mathbf{x}_p = \alpha\mathbf{x}_i + \mathbf{x}_c = \alpha\mathbf{F}^\mathsf{H}\mathbf{S}_K\check{\mathbf{x}} + \mathbf{x}_c \qquad (3.10)$$

where the scale factor and the small signal gain have been combined into $\alpha = gc$. Figure 3.3 shows a sample comparison of the linearly amplified input $\alpha\mathbf{x}_i$ and the actual PA output $\mathbf{x}_p$. The corresponding clipped signal, $\mathbf{x}_c = \alpha\mathbf{x}_i - \mathbf{x}_p$ is shown in Figure 3.4.
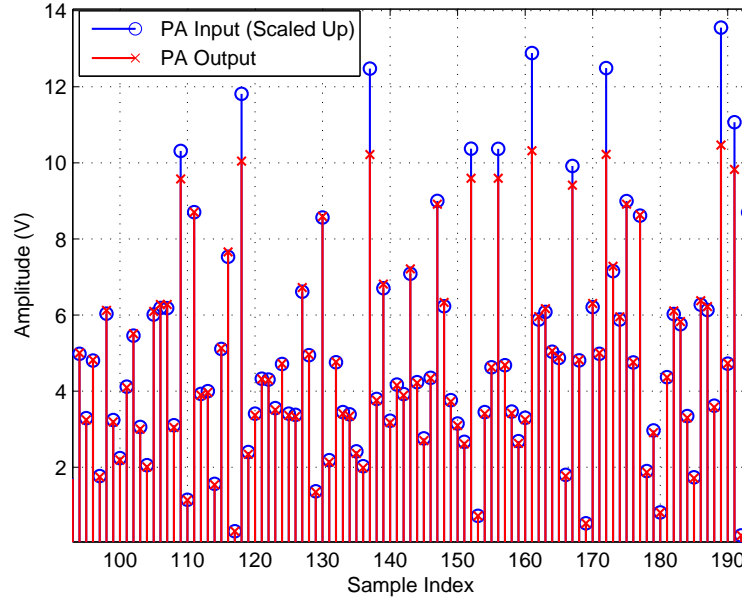
Figure 3.3: Sample input & output signals for RFMD PA

The simplified block diagram in Figure 3.5 summarizes the steps taken till this stage.
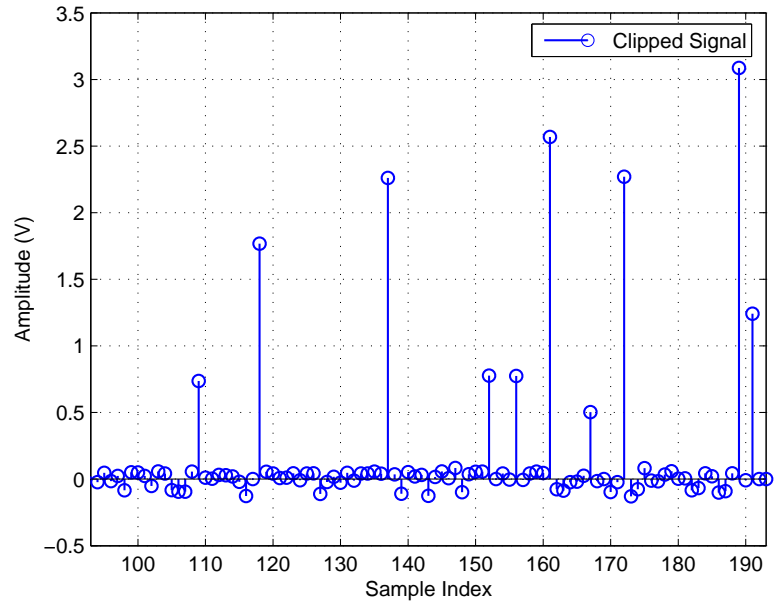
Figure 3.4: Clipped portion of RFMD PA input due to PA distortion

Figure 3.5: Simple block diagram summarizing transmitter & PA stages

3.4.3   Receiver

At the receiver, after removing cyclic prefix, we have:

$$\mathbf{y} = \mathbf{H}\mathbf{x}_p + \mathbf{z} \tag{3.11}$$

where $\mathbf{y}$ is the time–domain received signal, and $\mathbf{z} \sim \mathcal{CN}(\mathbf{0}, \sigma_{\mathbf{z}}\mathbf{I})$ both of size $N \times 1$. Due to the cyclic prefix of OFDM, $\mathbf{H}$ is a circulant matrix and in equation (3.11), represents the cyclic convolution of the channel impulse response, $\mathbf{h} = (h(0), h(1), h(2), \ldots, h(L-1))$ with $\mathbf{x}_p$ and can be decomposed as:

$$\mathbf{H} = \mathbf{F}^{\mathsf{H}}\mathbf{D}\mathbf{F}, \tag{3.12}$$

where $\mathbf{D} = diag(\check{\mathbf{h}})$ and $\check{\mathbf{h}} = \sqrt{N}\mathbf{F}\mathbf{h}$ is the DFT of the channel impulse response. (In this work, we assume the channel is known.)

Using the Fourier transform to obtain the frequency–domain form of the received signal in equation (3.11) gives us:

$$\check{\mathbf{y}} = \mathbf{F}\mathbf{y} = \mathbf{F}\mathbf{H}\mathbf{x}_p + \check{\mathbf{z}} \tag{3.13}$$

where the unitary matrix $\mathbf{F}$ leaves the DFT of $\mathbf{z}$, $\check{\mathbf{z}} \sim \mathcal{CN}(\mathbf{0}, \sigma_{\mathbf{z}}\mathbf{I})$ just like $\mathbf{z}$. Substi-

tuting for $\mathbf{x}_p$ from equation (3.10), we have:

$$\check{\mathbf{y}} = \mathbf{F} \cdot (\mathbf{F}^{\mathsf{H}}\mathbf{D}\mathbf{F} \cdot (\alpha\mathbf{F}^{\mathsf{H}}\mathbf{S}_K\mathbf{x}_i + \mathbf{x}_c)) + \check{\mathbf{z}}$$

$$\check{\mathbf{y}} = \alpha\mathbf{D}\mathbf{S}_K\mathbf{x}_i + \mathbf{D}\mathbf{F}\mathbf{x}_c + \check{\mathbf{z}} \tag{3.14}$$

3.5   Post-Compensation using Basic Compressed Sensing Algorithm

Following on from the preceding discussion, in this section, we explain specifically how we apply the CS algorithm to our problem of estimating and correcting for the PA-induced distortion of an OFDM signal at the receiver. As explained earlier in Sections 3.1 & 3.3:

1. Our ability to model the power amplifiers' non-linearities as a sparse phenomenon in the time domain is what enables us to apply CS to estimate and compensate for these distortions;

2. The measurements used in the CS algorithm for estimating these distortions are taken from the perturbations (from the zero position) of randomly–placed frequency–domain free carriers. We note that these perturbations were induced by the PA time–domain distortion itself and are signal–wide in the frequency–domain due to the incoherence of the two domains as previously discussed. A sample of these frequency–domain perturbations is shown in Figure 3.6.
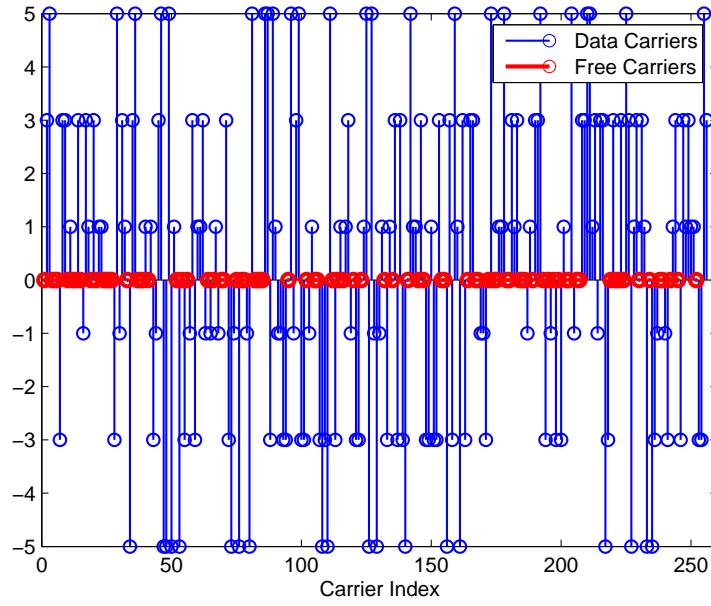
(a)
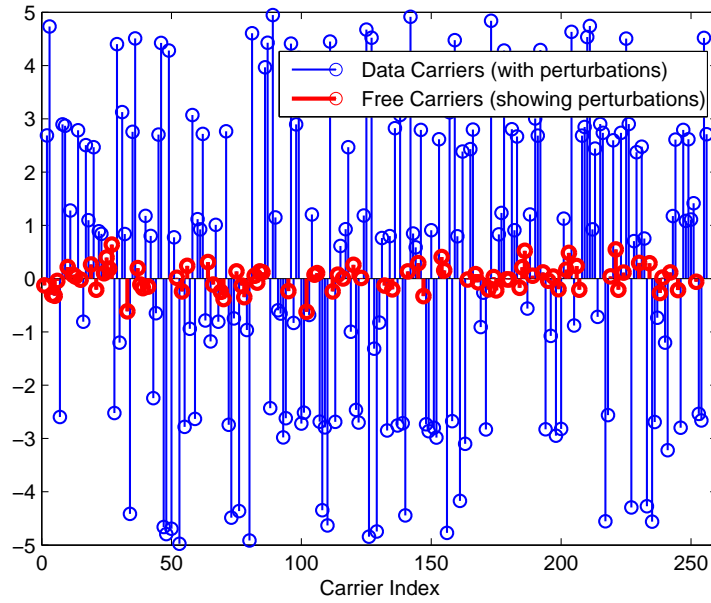
(b)

Figure 3.6: A comparison of the frequency–domain signal before and after the PA shows the PA–stimulated perturbations throughout the signal. (a) Frequency–domain signal *before* PA distortion; (b) Frequency–domain signal *after* PA distortion

Proceeding from the received signal as described in Section 3.4.3, the following details the methodology for post-compensation using CS:

We denote by $\boldsymbol{\omega}$ with a cardinality of $|\boldsymbol{\omega}| = N$, the set of all data carriers in the OFDM symbol and by $\boldsymbol{\omega}_M$ with a cardinality of $|\boldsymbol{\omega}_M| = M$, the set of free carriers used to estimate the distortion, $\mathbf{x}_c$. Given $\mathbf{S}_M$, a $N \times M$ matrix which is otherwise a zero matrix except for exactly one element equal to 1 in each column such that the columns of $\mathbf{S}_M$ form the orthogonal complement of the columns of $\mathbf{S}_K$, we derive the observation vector:

$$\check{\mathbf{y}}' = \mathbf{S}_M^{\mathsf{T}} \check{\mathbf{y}}$$

$$\check{\mathbf{y}}' = \mathbf{S}_M^{\mathsf{T}} (\alpha \mathbf{D} \mathbf{S}_K \check{\mathbf{x}} + \mathbf{D} \mathbf{F} \mathbf{x}_c + \check{\mathbf{z}})$$

$$\check{\mathbf{y}}' = \mathbf{S}_M^{\mathsf{T}} \mathbf{D} \mathbf{F} \mathbf{x}_c + \check{\mathbf{z}}' \qquad (3.15)$$

where we note that given the special structure of the $\mathbf{S}_M$ matrix as compared to the $\mathbf{S}_K$ matrix, $\mathbf{S}_M^{\mathsf{T}} \mathbf{D} \mathbf{S}_K = 0$. Re-writing equation (3.15), we have:

$$\check{\mathbf{y}}' = \boldsymbol{\Psi} \mathbf{x}_c + \check{\mathbf{z}}' \qquad (3.16)$$

where $\boldsymbol{\Psi} \triangleq \mathbf{S}_M^{\mathsf{T}} \mathbf{D} \mathbf{F}$ is the measurement matrix of size $M \times N$ and $\check{\mathbf{z}}' = \mathbf{S}_M^{\mathsf{T}} \check{\mathbf{z}}$ is an $M \times 1$ vector. Comparing equation (3.16) with (3.1), we notice that estimating the time-domain signal vector $\mathbf{x}_c$ in (3.16) is a CS problem since $\mathbf{x}_c$ which is the clipped

portion of the signal is relatively sparse. Therefore, using (3.3), we estimate $\mathbf{x}_c$ from:

$$\min_{\hat{\mathbf{x}}} \|\hat{\mathbf{x}}_c\|_{l_1} \quad \text{subject to} \quad \|\mathbf{\Psi}\hat{\mathbf{x}}_c - \check{\mathbf{y}}'\|_{l_2} \leq \epsilon \tag{3.17}$$

where $\epsilon = \sqrt{\sigma_z^2(M + \sqrt{2M})}$ [28]. As explained previously, we are not restricted to using (3.17) and any algorithm that solves (3.16) would be sufficient. As an example, Figure 3.7 shows the CS-estimate of the sample clipped signal shown in Figure 3.2.
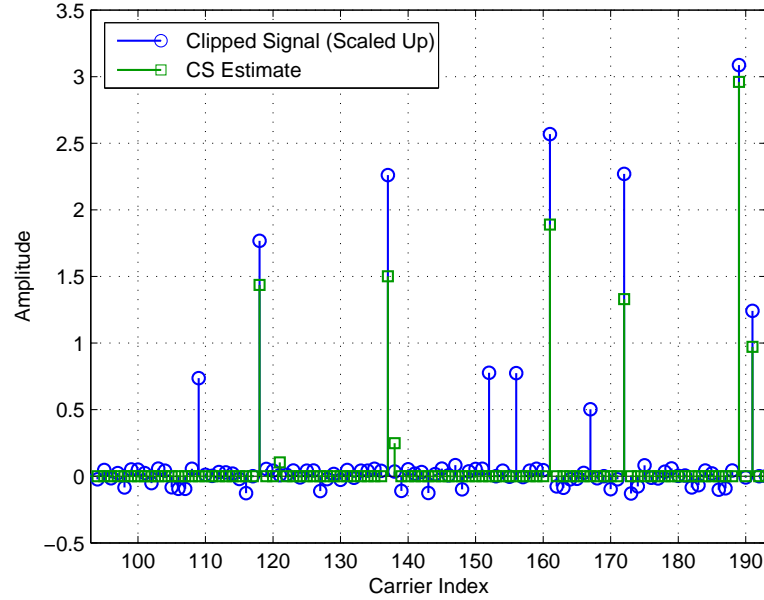
Figure 3.7: CS estimate of clipped portion of RFMD PA input

The estimate of the clipped signal, $\hat{\mathbf{x}}_c$ is then subtracted from the received signal, $\mathbf{y}$ to obtain the corrected signal,

$$\mathbf{u}_s = \mathbf{y} - \hat{\mathbf{x}}_c \tag{3.18}$$

47

This signal is then scaled and converted to the frequency–domain via the DFT:

$$\check{\mathbf{u}} = \mathbf{F}\mathbf{u}_i = \mathbf{F}\frac{\mathbf{u}_s}{\alpha} \tag{3.19}$$

$\check{\mathbf{u}}$ is demodulated and the free-carriers are removed to give an estimate of the transmitted data, $\hat{\mathbf{d}}$. The different phases of the receiver (including the post-compensation algorithm) are illustrated in Figure 3.8.
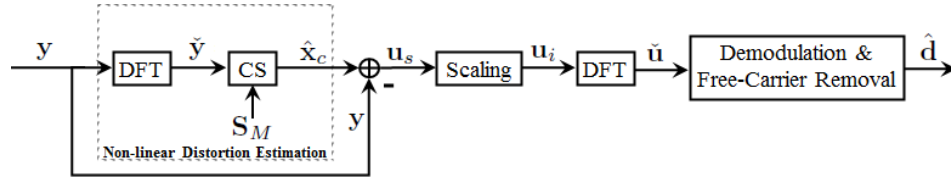
Figure 3.8: Simple block diagram summarizing the receiver stage

3.6   Improving the CS Estimate with Weighted CS

As is obvious from Figure 3.2, the power amplifier's non-linear response does not affect the high peaks exclusively, rather, the whole signal is affected—except that lower amplitudes have a much lower distortion than the peaks. However, for this reason, the CS algorithm would usually only give good estimates for the peak distortions and as for the low distortions, the estimates—as can be seen in the sample in Figure 3.7—are unusable. Also, not correcting for these minute distortions has only a negligible effect on the error rate and our ability to correctly estimate the transmitted data.[3]

[3]Although results regarding this are not presented here, this was confirmed via simulation.

For these reasons, estimates of very low amplitude can be discarded.

So, to improve our estimated clipped signal, $\mathbf{x}_c$, a first thought was to force

$$\hat{\mathbf{x}}_c(q) = 0 \quad \forall \quad q \in \{i : \hat{\mathbf{x}}_c(i) < T\} \tag{3.20}$$

where T is some threshold value. However, determining an appropriate value for T is not as straightforward as may seem and choosing a wrong value may lead to keeping some wrong estimates and discarding some correct ones. Due to this, this method of zeroing amplitudes below a particular threshold was put aside.

We rather implement an intelligent weighted CS (WCS) algorithm where we use the knowledge of the amplitude levels in the received signal as a priori information to improve the CS approximation of the distortion. Since we know that the significant clips which we need to estimate occur almost exclusively at locations with high amplitudes and that as the amplitudes reduce, the probability of clipping lessens, we need to modify the CS algorithm such that more priority is given to recovering the clips at locations where the received signal, $\mathbf{y}$ is higher magnitude (i.e. locations where clipping is more probable) and low-magnitude locations are on the other hand penalized. To achieve this, we apply a weighted CS (WCS) algorithm via reweighted l_1 minimization [46]:

$$\min_{\hat{\mathbf{x}}} \|\mathbf{w}^\mathsf{T} \hat{\mathbf{x}}_c\|_{l_1} \quad \text{subject to} \quad \|\mathbf{\Psi}\hat{\mathbf{x}}_c - \check{\mathbf{y}}'\|_{l_2} \leq \epsilon \tag{3.21}$$

where **w** is defined as the inverse of the received signal, **y**:

$$
\mathrm{w}(n) = \begin{cases} \frac{1}{\mathrm{y}(n)}, & \text{if } \mathrm{y}(n) \neq 0, \\[2mm] \infty, & \text{if } \mathrm{y}(n) = 0, \end{cases} \tag{3.22}
$$

where $n = 1, 2, \ldots, N$ [4]. Equation (3.22) assigns smaller weights to locations of larger amplitudes (where clipping is more probable) thus forcing the WCS algorithm in (3.21) to focus more on these locations and less on the other locations with less amplitude to which larger weights are assigned.

Figure 3.9 shows a comparison of the WCS and CS estimates of the sample clipped signal shown in Figure 3.2.
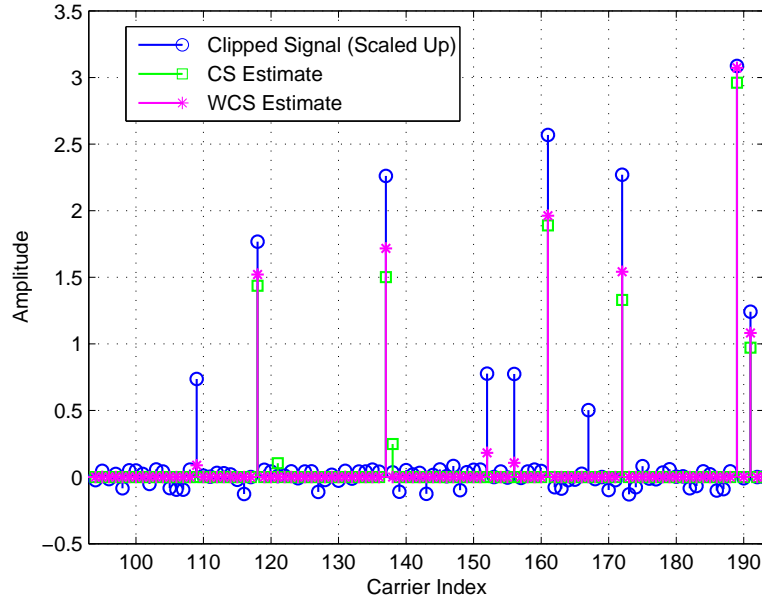
Figure 3.9: WCS & CS estimates of clipped portion of RFMD PA input

From this comparison, we see three benefits of the WCS algorithm:

[4]In simulations, ∞ in this case was taken to be 10^5

- It improves amplitude estimates at locations also detected by the CS algorithm,

- It indicates additional distortion locations not found by plain CS. and

- It removes wrong locations suggested by the plain CS algorithm.

3.7 Simulation Set-up & Results

For all simulation results presented in this thesis, the number of sub-carriers was held constant at $N = 256$ and the 64-QAM modulation scheme was used. The channel is also always assumed to be known and the SNR is varied from 15 dB to 35 dB. Although all amplifier models listed in Chapter 2 were studied, as mentioned previously, we only present the results of the RFMD and Doherty models.

Each of these performance measurements have been plotted against the SNR range. To evaluate and compare the effectiveness of the various CS-based solutions proposed in this paper, the following performance indices were used:

1. Normalized Mean Square Error (NMSE) [47]:

$$\text{NMSE} = \frac{\dfrac{1}{N}\sum_{r=1}^{N}\left|\alpha\text{x}_i(r) - \text{u}_i(r)\right|^2}{\dfrac{1}{N}\sum_{r=1}^{N}\left|\alpha\text{x}_i(r)\right|^2} \tag{3.23}$$

2. Error Vector Magnitude (EVM) [48, 49]:

$$\text{EVM}_{rms} = \sqrt{\frac{\dfrac{1}{N}\displaystyle\sum_{r=1}^{N}\left|\check{\text{x}}(r) - \check{\text{u}}(r)\right|^2}{\dfrac{1}{N}\displaystyle\sum_{r=1}^{N}\left|\check{\text{x}}(r)\right|^2}} \tag{3.24}$$

3. Bit Error Rate (BER):

$$\text{BER} = \frac{\sum bit\ errors\ (comparing\ \mathbf{d}\ \&\ \hat{\mathbf{d}})}{K} \tag{3.25}$$

We note that in all cases, we have an upper bound for the performance when no distortion-estimation technique is used at the receiver and a lower bound when the PA is assumed to be perfectly linear.

In Figures 3.10a-3.10c, we present and compare the performance (using all three performance metrics mentioned above) of both CS & WCS in recovering and correcting for the power amplifier's non-linearities. The number of free-carriers used, M, ranged from 10% to 25% of N, in steps of 5%.

Note that due to the great non-linearity of the Doherty PA, neither CS nor WCS are able to recover the distortions it causes.[5] For this reason, all curves for the Doherty PA (with and without recovery) almost exactly overlap each other and so, only a single curve (representing the case when WCS is used with the PA) is shown in the results presented in Figures 3.10a and 3.10c. (In the EVM plot of Figure 3.10b, the curve for

---

[5]A distortion-recovery technique suitable for highly non-linear amplifiers like the Doherty is the joint compensation which is discussed in Chapter 4.

the Doherty PA has been left out as its minimum value is 32% which is high above the RFMD EVM range displayed.) Also note that in Figure 3.10a, only the SNR range from 20 dB to 35 dB has been displayed because before 20 dB, all recovery techniques perform essentially equally in terms of NMSE.
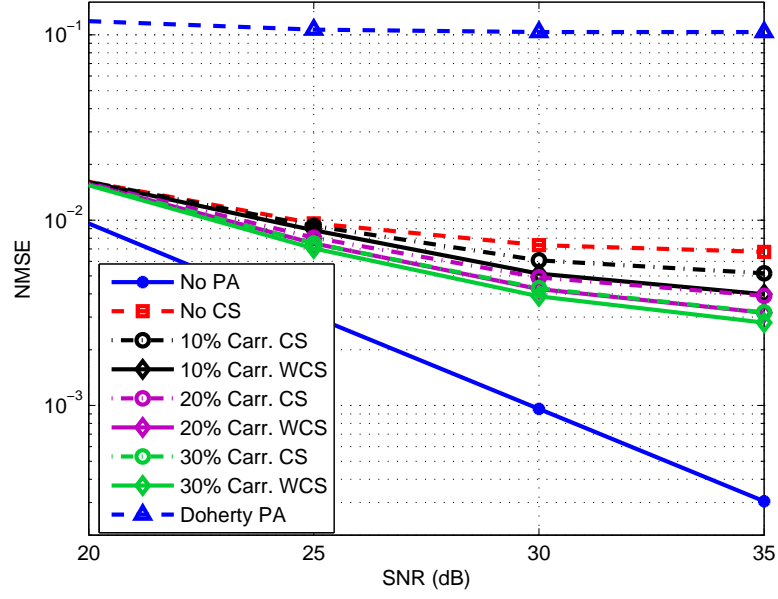
(a)
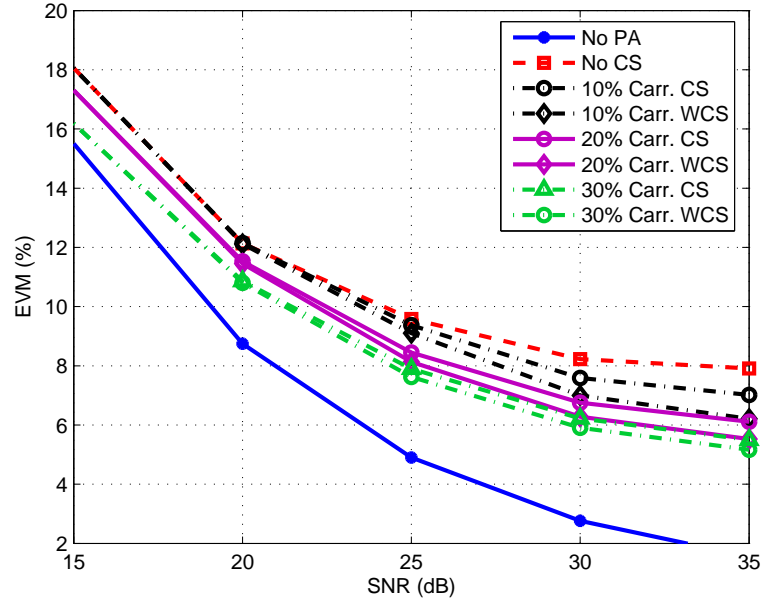
(b)
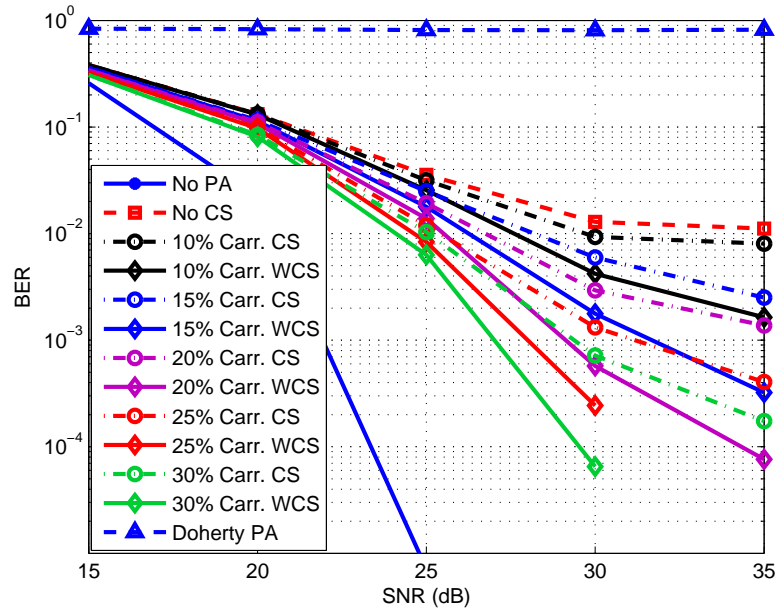
(c)

Figure 3.10: Comparison of performance of WCS and CS for RFMD amplifier using (a) NMSE (b) EVM (c) BER measurements

From these results we note:

1. Both CS and WCS are able to estimate some of the PA distortions and thus correct the received signal to an extent. Also, we notice that the more free-carriers we use, the better the distortion estimate is.

2. WCS gives good improvement over CS. Especially in terms of NMSE & EVM, it seems to be equivalent to using 10% extra carriers for CS—as we see the 10% & 20% WCS curves almost exactly overlapping the 20% & 30% CS curves receptively. In terms of BER, WCS shows even better improvement over CS.

3.8   Chapter Conclusion

After describing the basic CS algorithm, its sparsity and incoherence requirements, and how we apply it to our distortion-estimation problem, we have also explained in this chapter, the weighted CS technique. The results presented for the RFMD PA show that although basic CS is able to estimate and correct some distortion, WCS significantly improves upon these estimates.

CHAPTER 4

DATA-AIDED CS AND JOINT COMPENSATION

In this chapter, we discuss two advanced methods by which the results of the CS algorithm can further be improved. These are:

1. Data-Aided Compressed Sensing ("Data-Aided CS" or "DACS")

2. Joint Compensation (i.e. Pre-Compensation along with Post-Compensation)

4.1 DATA-AIDED CS ALGORITHMS

4.1.1 Motivation

In both the basic and weighted CS algorithms described in Chapter 3, we have always included free-carriers in the transmitted frequency–domain signal in order to obtain, at the receiver, the measurements which CS uses to estimate the PA distortion. However,

if we could find some way to obtain measurements from the data-carriers also, we would be able to improve our estimation algorithm from two important angles:

- Further estimation of left-over distortion even after implementing CS.

- Ability to implement CS without any free-carriers thus using the bandwidth at full efficiency.

We notice that the free-carrier measurements used in CS—as is clear from Figure 3.6—are simply the perturbations (i.e. amounts of deviation) of these free-carriers from their original zero position. Thus, similarly for the data-carriers, if at the receiver, we have knowledge of their original locations in the constellation, we would be able to calculate their respective perturbations from these locations and thus obtain extra measurements to use in the CS algorithm. The question that now presents itself is: how do we find the original constellation locations of at least some of these data-carriers?

One option would be to assume that all constellation points of the received signal, $\check{\mathbf{y}}$, within a given short threshold distance, R of an original constellation point, $\check{x}(i)$ have deviated from that original constellation point and thus consider as reliable all perturbations, $\check{p}(l)$ such that:

$$\check{p}(l) = |\check{x}(i) - \check{y}(j)| \leq R, \quad i, j \in 0, 1, \ldots, k-1 \tag{4.1}$$

where k is the number of frequency-domain data-carriers. This is equivalent to drawing circles of radii, R around all original constellation points and taking all deviations

within the circles as reliable perturbations $\check{p}(l)$ to be used as measurements in the CS algorithm.

A problem with this however is in choosing a suitable value for the threshold R. On the one hand, selecting a minuscule value for R would make sure that each receiver point is correctly compared with its respective original point, but this would give us only few measurements of little amplitude which would lead to only negligible improvement in the estimation problem. On the other hand, if we chose R to be too large, we stand the risk of comparing wrong points and thus obtaining erratic measurements.

To avoid this problem, we use a more tenable approach to obtain reliable perturbations. It is obvious that each constellation point of the received signal would—with very high probability—have deviated from one of the four original constellation locations surrounding it and not other constellation locations. In Figure 4.1 for example, $\check{u}(2)$ would most probably have deviated from one of either $\mathcal{X}, \mathcal{X}_a, \mathcal{X}_b,$ or $\mathcal{X}_c$. Thus, we can simply select the closest of these neighboring points (in this case, $\mathcal{X}$) as most probably being the original constellation point of $\check{u}(2)$ and then take the distance between the two points to be a possible reliable perturbation.

Since this algorithm makes use of the received data to obtain measurements to be used in CS, we refer to it as the Data-Aided CS (DACS) algorithm. The preceding is a simplified discussion of this algorithm and the following section gives a more detailed mathematical description.
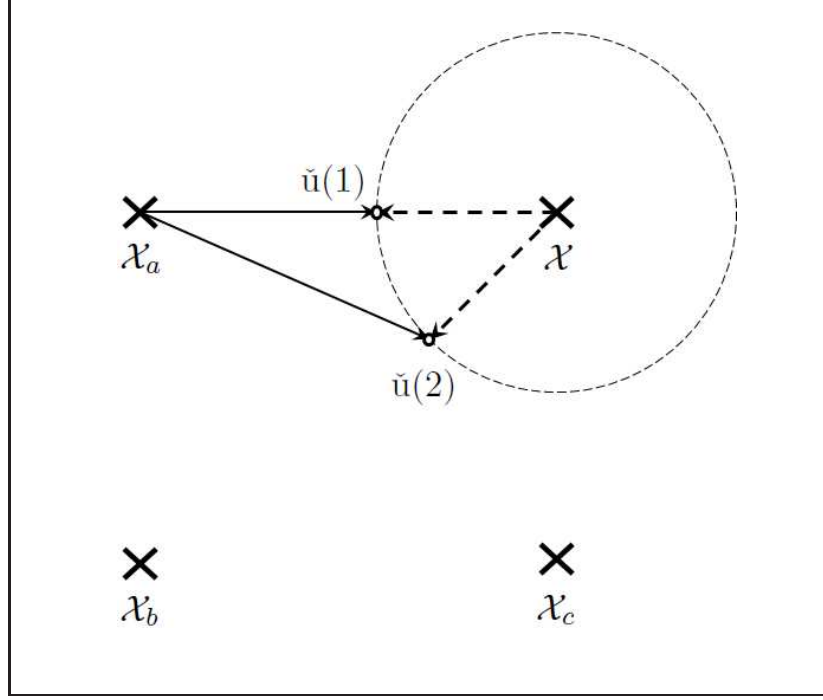
Figure 4.1: Description of the data-aided algorithm

4.1.2 Description

From the above discussion, we understand that we are able to implement the DACS algorithm based on our assumption that a number of data carriers, $\check{\mathbf{x}}' \subset \check{\mathbf{x}}$ even after the power amplifier's distortion, are still within their respective decision regions. We denote the set of these carriers by $\boldsymbol{\omega}_{M'}$ with cardinality $|\boldsymbol{\omega}_{M'}| = M'$ such that $M' < N$. If $\mathcal{X}$ consists of the M-ary constellation points, then at these locations of reliable perturbations, it is highly probable that:

$$\check{\mathbf{x}}' \approx \mathcal{X}_{M'} \tag{4.2}$$

Just as we had $\mathbf{S}_M$ which selected the free-carriers in the case of basic CS, we let $\mathbf{S}_{M'}$ denote the $N \times M'$ selection matrix—a zero matrix except for a single element in

each column which equals 1—that selects the "reliable" carriers. Given $\mathbf{u}$ and $\mathbf{\check{u}}$ to be time and frequency-domain versions of the partially corrected signal (see Equations 3.18 and 3.19) respectively, we know that $\mathbf{u}$ is a sum of: (i) an estimate of the input signal, $\mathbf{\hat{x}}_i$, (ii) left-over distortion, $\mathbf{\tilde{x}}_c = \mathbf{x}_c - \mathbf{\hat{x}}_c$ (after removing some non-linearities), and (iii) noise, $\mathbf{z}$, i.e.:

$$\mathbf{u} = \mathbf{\hat{x}}_i + \mathbf{\tilde{x}}_c + \mathbf{z} \tag{4.3}$$

And in the frequency-domain:

$$\mathbf{\check{u}} = \mathbf{\check{\hat{x}}} + \mathbf{DF}\mathbf{\tilde{x}}_c + \mathbf{\check{z}} \tag{4.4}$$

To obtain the data-carrier measurements for DACS, we multiply (4.4) by $\mathbf{S}_{M'}^{\mathsf{T}}$ on both sides:

$$\mathbf{S}_{M'}^{\mathsf{T}}\mathbf{\check{u}} = \mathbf{S}_{M'}^{\mathsf{T}}\mathbf{\check{\hat{x}}} + \mathbf{S}_{M'}^{\mathsf{T}}\mathbf{DF}\mathbf{\tilde{x}}_c + \mathbf{S}_{M'}^{\mathsf{T}}\mathbf{\check{z}}$$

$$\mathbf{S}_{M'}^{\mathsf{T}}\mathbf{\check{u}} = \mathbf{\check{\hat{x}}}' + \mathbf{S}_{M'}^{\mathsf{T}}\mathbf{DF}\mathbf{\tilde{x}}_c + \mathbf{S}_{M'}^{\mathsf{T}}\mathbf{\check{z}} \tag{4.5}$$

Subtracting the corresponding correct constellation locations from (4.5) yields:

$$\mathbf{S}_{M'}^{\mathsf{T}}\mathbf{\check{u}} - \mathcal{X}_{M'} = \mathbf{\check{\hat{x}}}' - \mathcal{X}_{M'} + \mathbf{S}_{M'}^{\mathsf{T}}\mathbf{DF}\mathbf{\tilde{x}}_c + \mathbf{S}_{M'}^{\mathsf{T}}\mathbf{\check{z}}$$

But since (4.2) gives $\mathbf{\check{x}}' - \mathcal{X}_{M'} = \mathbf{0}$, we have

$$\mathbf{S}_{M'}^{\mathsf{T}}\mathbf{\check{u}} - \mathcal{X}_{M'} = \mathbf{S}_{M'}^{\mathsf{T}}\mathbf{DF}\mathbf{\tilde{x}}_c + \mathbf{S}_{M'}^{\mathsf{T}}\mathbf{\check{z}} \tag{4.6}$$

which we write as

$$\check{\mathbf{u}}' = \mathbf{\Psi}'\tilde{\mathbf{x}}_c + \check{\mathbf{z}}'' \tag{4.7}$$

where $\check{\mathbf{u}}' \triangleq \mathbf{S}_{M'}^{\mathsf{T}}\check{\mathbf{u}} - \mathcal{X}_{M'}$ and $\mathbf{\Psi}' \triangleq \mathbf{S}_{M'}^{\mathsf{T}}\mathbf{DF}$. We point out here that in estimating the residual distortion, $\tilde{\mathbf{x}}_c$, we don't have to use all M' reliable carriers; rather we can limit the number of carriers used to Q such that $\boldsymbol{\omega}_Q \subset \boldsymbol{\omega}_{M'} \subset \boldsymbol{\omega}$—which would be sufficient. So, we re-define our variables of (4.7) thus:

$$\check{\mathbf{u}}' \triangleq \mathbf{S}_R^{\mathsf{T}}\check{\mathbf{u}} - \mathcal{X}_R \text{ and } \mathbf{\Psi}' \triangleq \mathbf{S}_R^{\mathsf{T}}\mathbf{DF}$$

Comparing (4.7) with (3.1), we see we can use CS to estimate $\tilde{\mathbf{x}}$ thus:

$$\min_{\hat{\tilde{\mathbf{x}}}} \|\hat{\tilde{\mathbf{x}}}_c\|_{l_1} \quad \text{subject to} \quad \|\mathbf{\Psi}'\hat{\tilde{\mathbf{x}}}_c - \check{\mathbf{u}}'\|_{l_2} \leq \epsilon \tag{4.8}$$

For further refinement of the estimates, we can iterate the DACS algorithm as given above a number of j times. The whole procedure is illustrated in the functional block digram of Figure 4.2
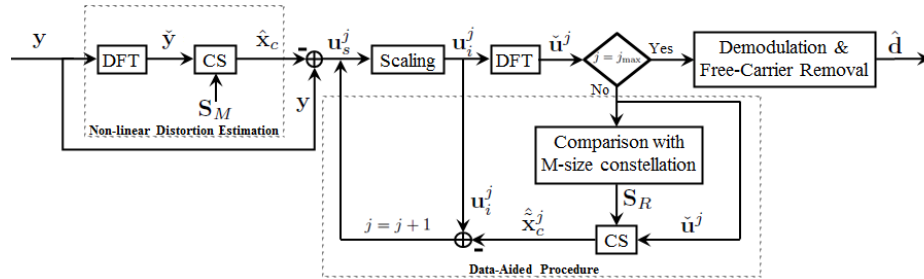
Figure 4.2: Block diagram illustrating the DACS algorithm

61

4.1.3 Choosing Data-Carriers: Constructing $\boldsymbol{\omega}_R$

One of the most crucial steps in the DACS algorithm as described above is how we select the set of R data-carriers, $\boldsymbol{\omega}_R$ (out of a possible $N - M$) that would be used to obtain measurements used in runs of DACS. Considering all possible ways, we would have a total of $\binom{N-M}{R}$ options to build $\boldsymbol{\omega}_R$.

To satisfy the requirements of the CS algorithm, the R carriers need to be randomly distributed in the signal. However, constructing $\boldsymbol{\omega}_R$ by a simple random selection would mean that less-reliable carriers are selected. And since the improvement offered by DACS greatly depends on the reliability of the carriers, such a random selection increases our risk of obtaining a wrong estimate for the remnant distortion.

Rather, we employ a smart technique in building $\boldsymbol{\omega}_R$ such that it consists of the most reliable carriers by using the posterior probability:

$$\mathcal{R}(n) = \log \frac{\Pr(\langle \breve{u}(n) \rangle = \mathcal{X} | \breve{u}(n))}{\Pr(\langle \breve{u}(n) \rangle = \mathcal{X}^* | \breve{u}(n))} \tag{4.9}$$

where $n = 1, 2, 3, \ldots, (N - M)$, $\langle . \rangle$ represents the respective decoded symbol, $\mathcal{X}$ is the constellation point closest to $\breve{u}(n)$ while, $\mathcal{X}^*$ represents the nearest neighboring point. (4.9) thus gives us a function that helps us measure the reliablity of decoding $\breve{u}(n)$ to $\mathcal{X}$ relative to $\mathcal{X}^*$.

To explain this technique, we consider the sample case as shown in Figure 4.1. As in the figure, every point, $\breve{u}(n)$ would have four closest neighbors surrounding it from the constellation. We assume $\mathcal{X}$ to be the correct symbol, and the three other

neighbors are denoted by $\mathcal{X}_a$, $\mathcal{X}_b$ and $\mathcal{X}_c$. We observe in this example that:

$$|\breve{u}(1) - \mathcal{X}| = |\breve{u}(2) - \mathcal{X}| \tag{4.10}$$

However,

$$\frac{|\breve{u}(1) - \mathcal{X}|}{|\breve{u}(1) - \mathcal{X}_a|} > \frac{|\breve{u}(2) - \mathcal{X}|}{|\breve{u}(2) - \mathcal{X}_a|} \tag{4.11}$$

i.e. $$\mathcal{R}(1) > \mathcal{R}(2) \tag{4.12}$$

Meaning that $\breve{u}(1)$ is a more reliable carrier than $\breve{u}(2)$.

After determining the reliability of all $N - M$ data-carriers, we choose the R carriers of highest reliability values, i.e.:

$$\boldsymbol{\omega}_R \triangleq \arg\{\mathcal{R}_s\}_{s=1}^R \tag{4.13}$$

It turns out that the set of carriers, $\boldsymbol{\omega}_R$, is almost always a randomly dispersed set from $(\boldsymbol{\omega} - \boldsymbol{\omega}_M)$. It thus satisfies the CS requirement that measurements are taken randomly, and so, from this angle, the performance of CS in estimating the distortion is not inhibited by our intelligent method of selecting $\boldsymbol{\omega}_R$.

4.1.4   Simulation Results

To evaluate the performance of the DACS algorithm, we again considered the RFMD amplifier with an OFDM signal of the same parameters given in Chapter 3 (i.e. $N =$

256, 64-QAM, SNR range: 15 dB to 35 dB in 5 dB steps) and the same performance measurements (NMSE, EVM and BER). Again, just as in Chapter 3, the results are upper-bounded by the case when no distortion estimation method is used ("No CS") and lower-bounder when the power amplifier is by-passed ("No PA") i.e. we have only noise.

In the results presented in the plots of Figure 4.3, we have compared the performance of two iterations of DACS using 30% and 40% of the data-carriers after using WCS with 10% free-carriers. We also include the performance of two iterations of DACS using 40% of the data-carriers *without* a previous run of WCS.
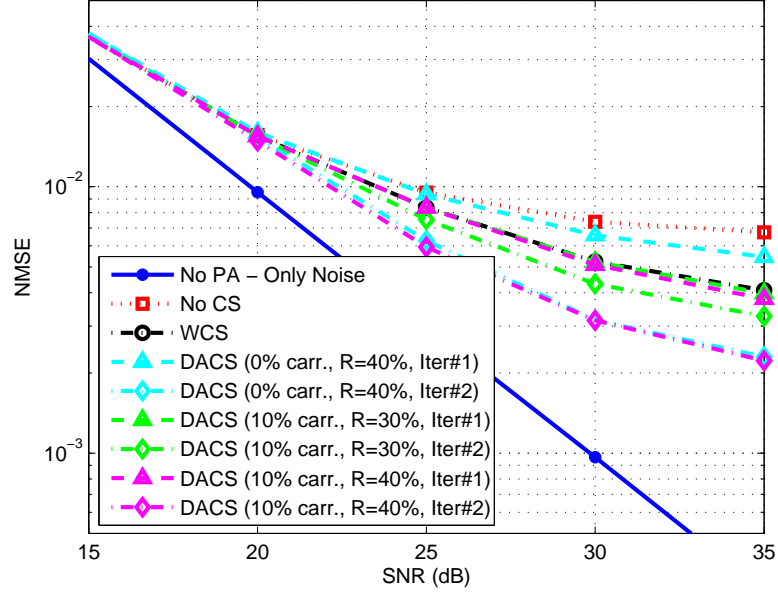
(a)
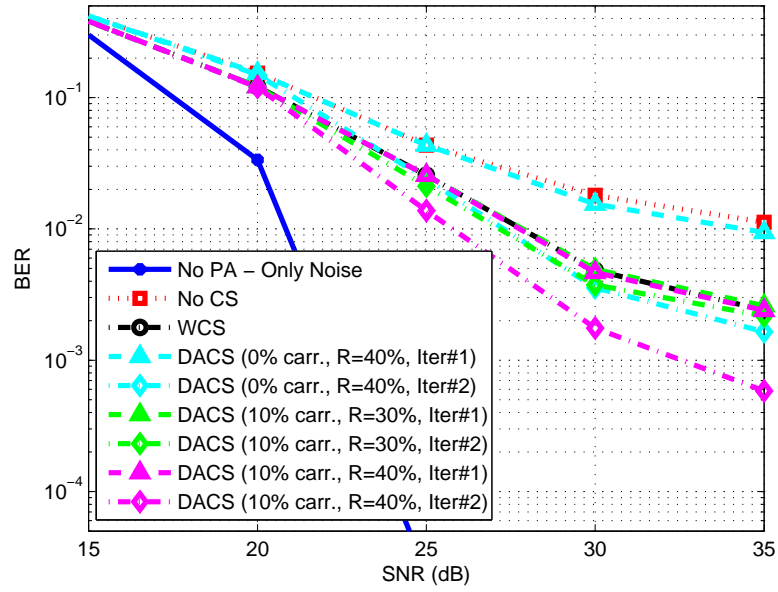
(b)
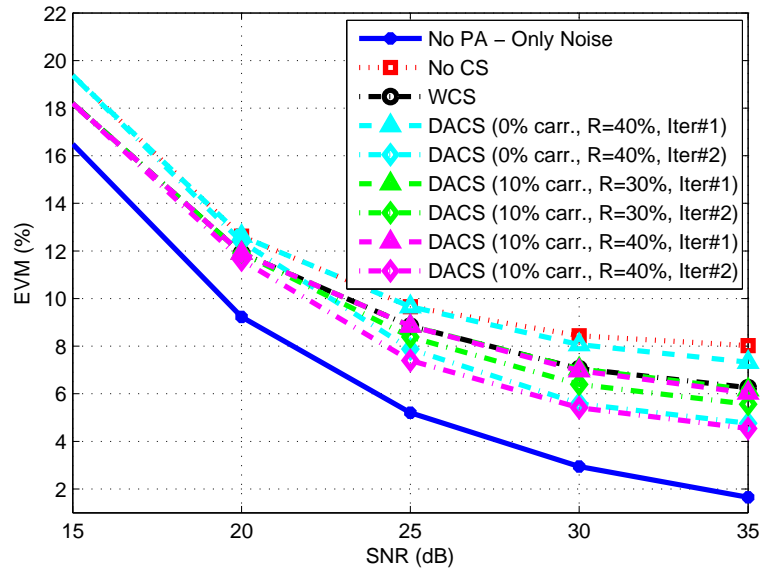
(c)

Figure 4.3: Comparison of performance of different runs of DACS for the RFMD amplifier using (a) NMSE (b) EVM (c) BER

The results are quite encouraging as the two plots which use DACS after WCS show the enhancement DACS gives to the WCS estimates—especially the second iteration of each of these which improves over the first iterations' estimates to a large degree. We also note the expected effect of using a larger number of data-carriers for DACS as we see that using 40% gives much better results than 30% carriers—particularly in the BER curves.

4.1.4.1 Blind Distortion Estimation with DACS

Since the Data-Aided CS algorithm uses only data-carriers in estimating the PA distortions, we can improve bandwidth efficiency by doing away with free-carriers (used for basic CS and WCS) and rely only on DACS for distortion estimation. However, since we would implement neither CS nor WCS, we expect distortion estimates would be worse (than if we were to use CS or WCS before DACS). The results of this case of blind estimation in which we include zero free–carriers in the transmitted signal ($M = 0$) are presented in Figures 4.3a–4.3c as the curves with 0% carriers.

Comparing the results for blind estimation, we see that we are able to get good performance in estimating the distortion especially after the second iteration when we use more data-carriers (40% in our case). In all curves, this blind estimation turns out better to be than using 10% free-carriers and 30% data-carriers and in the EVM and NMSE curves, it gets as good as using 10% free-carriers and 40% data-carriers. Thus by using DACS, we can exclude free-carriers from our transmitted signal thus improving bandwidth efficiency while at the same time, still having a relatively good estimate of the distortion.

4.2 JOINT COMPENSATION

We discuss in this section how transmitter-side pre-compensation methods can be combined with our receiver-side, CS-based post-compensation algorithms. This combination is termed, "joint compensation".

4.2.1 Motivation

As was noted in Chapter 2, not all power amplifiers lend themselves readily to distortion estimation using CS. This is especially because such amplifiers are very non-linear and so, the distortion they cause is quite condensed i.e. very non-sparse. In such cases, CS fails to estimate the distortion correctly as one of its fundamental requirements (i.e. that the signal being estimated is sparse) is not met. An example of such a highly non-linear amplifier is the Doherty PA (previously described in Sub-section 2.2.2) which has a relatively very low Gini Sparsity Index of about 0.461. It should be noted that, as we mentioned in that sub-section, the distortions of the Doherty PA are even made worse by the large range of the phase response. A sample of the distortion due to this amplifier is shown in Figure 4.4. In situations such as this, we are forced to use a transmitter-side pre-compensation method such as pre-distortion to correct for the PA distortion.

Also, even for some amplifiers where the distortion is sparse and receiver-based CS techniques can be used to estimate and remove the distortion, we may be forced to use pre-compensation especially if the transmitted signal has to meet stringent spectrum mask requirements as is the case with many of today's communication standards.
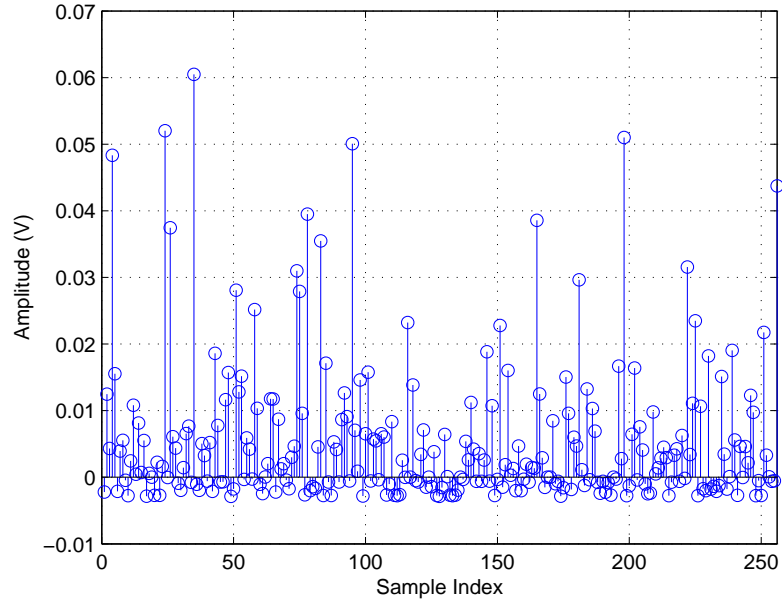
Figure 4.4: Sample of distortion at the output of the Doherty PA

In these two cases where pre-compensation techniques are required, digital pre-distortion (described in Section 1.2) is a popular choice. However, in particular cases, a mismatch between the pre-distorter function and the power amplifier may inhibit the otherwise excellent performance of pre-distortion. Some reasons why such a mismatch might occur include:

1. Aging, long term drifts of the biasing, heat, a change in the characteristics of the transmitted signal (including changes in the carrier frequency or signal average power) or a combination of these may lead to a change in the amplifier's behavior.

2. Changes in the surroundings of the mobile terminal may cause fluctuations in the antenna impedance which would lead to a load mismatch thus changing the amplifier's behavior.

In situations such as these, the pre-distorter would normally be able to correct for much of the amplifier distortion, however, due to the mismatch between it and the PA, there would be some remnant distortion on the transmitted signal. If this remanant distortion is relatively sparse, we can therefore use the CS-based post-compensation algorithms presented in this thesis to estimate the left-over distortion at the receiver and then correct the received signal.

This implementation of post-compensation after using pre-compensation is what we refer to as "joint compensation" and we describe it in the following section using the example case when the transmitted signal has a higher maximum power than is normally allowed by the PA. While the block diagram of the receiver remains the same as Figure 4.2 (since we stil use DACS for post-compensation), the only change on the transmitter side is our inclusion of the pre-distorter upstream of the power amplifier as illustrated in Figure 4.5.
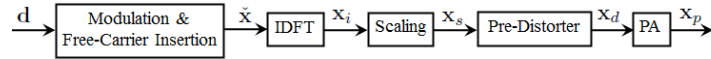
Figure 4.5: Block diagram of transmitter in joint compensation

4.2.2   Description

Under normal conditions, even after implementing a pre-distorter function before the amplifier, the combination is still only operated below a certain maximum input power level which can reach but must not exceed the saturation power. This is because operating the PA in the saturation region would lead to non-linear distortions.

However, since we are able to correct these distortions using the CS-based algorithms presented in this thesis, it becomes possible to operate the power amplifier at a higher average power level than it would ordinarily operate at i.e. overdrive the amplifier so that it operates closer to saturation and thus at higher efficiency. Figure 4.6 compares the gain characteristic of the Doherty amplifier before pre-distortion, after pre-distortion and after pre-distortion again but with the amplifier overdriven (by up to 6 dB). As we see in the figure, the cascade of the digital pre-distorter (DPD) with the Doherty PA gives a perfectly linear amplifier. But when the overdrive is applied, the amplifier is driven into the saturation region resulting in distortion as shown in the sample clips in Figure 4.7 when the overdrive was 2 dB.
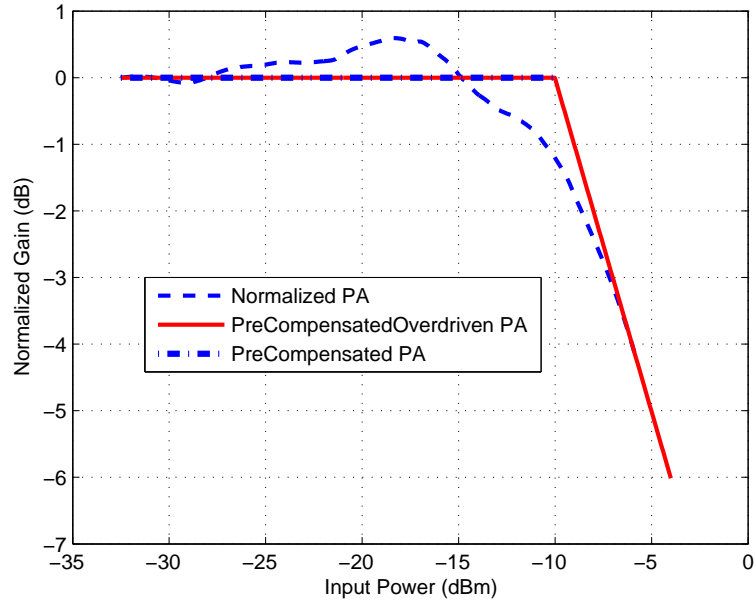
Figure 4.6: Comparsion of gain characteristics of different versions of the Doherty PA

Although overdriving the amplifier might cause spectrum regrowth, this is not a problem in the ISM band. To study the joint compensation algorithm, we implement it for the overdriven Doherty amplifier ranging the overdrive level from 0 dB to 6 dB.
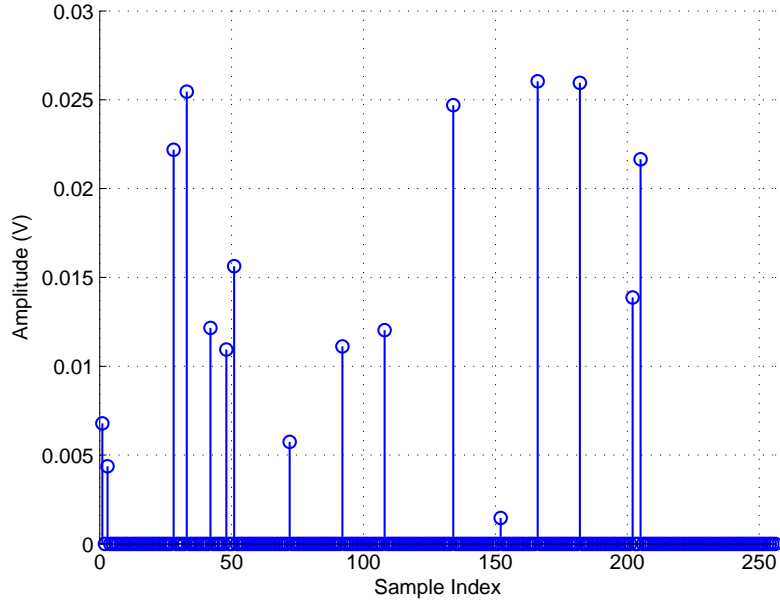
Figure 4.7: Clips due to the overdriven precompensated Doherty PA (overdrive = 2 dB)

4.2.2.1 Overdrive and Sparsity

We know that without overdrive (i.e. overdrive = 0 dB), the pre-distorted amplifier has no distortions. It is also obvious that as the overdrive increases from 0 dB to 6 dB, we go further and further into the saturation region of the PA and thus, the amount of distortion increases causing a decrease in the sparsity of the clipped portion. Figure 4.8 compares the sparsity levels of the clipped signal as the overdrive increases. We use the Gini Index [34, 35] which, as mentioned previously, rates the sparsity of a signal from $\mathcal{G}_s = 0$ (very dense) to $\mathcal{G}_s = 1$ (very sparse, a single impulse).

In general, we expect that when the overdrive is too low, the distortion is relatively negligible and of too little amplitude to recover using CS. At medium values of overdrive, DACS should perform well in estimating the clips till a certain range of overdrive levels above which the distortion should get too sparse for CS to correctly
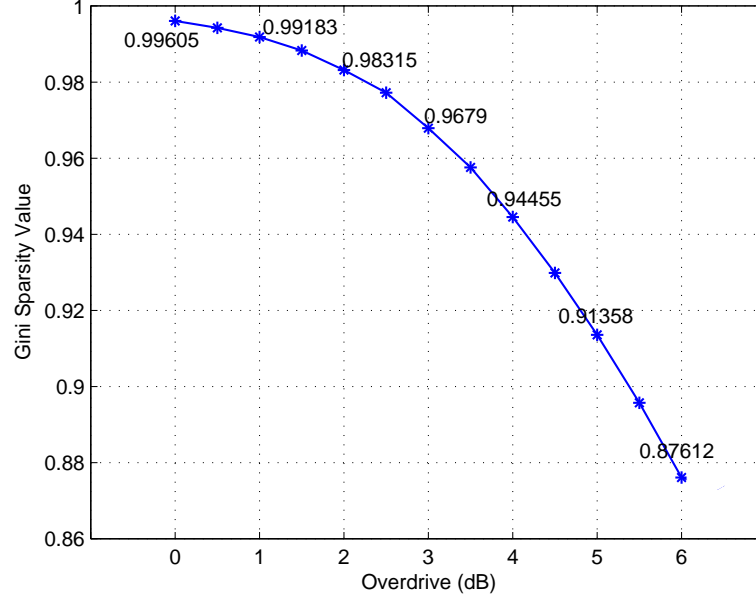
Figure 4.8: Gini sparsity index of overdriven Doherty PA

estimate.

4.2.3   Simulation Results (using Doherty Amplifier)

Using parameters similar to those used in DACS in simulating the joint compensation, we compare the performance of the amplifier without the DPD and with the DPD (both with and without overdrive). For each of the overdrive values used (2 dB, 3 dB and 4 dB), we compare the results of the second iteration of DACS ("DACS^2") versus the case where no distortion estimation is done. For all runs of DACS (except one), WCS is first used with 10% free-carriers to estimate some distortion. In the single exception to this (at an overdrive of 3 dB "OD=3dB"), DACS is again run twice but no free-carriers are included in the transmitted signal i.e. no previous distortion estimation was used. This is labeled as "0-DACS^2".

It should be noted that here, performance is lower bounded by the perfectly linear case of the non-overdriven amplifier cascaded with the digital pre-distorter and upper bounded by the results when the DPD is not used (whether or not CS is applied in this case does not matter since the huge distortion due to the PA makes any improvement offered by CS negligible).

(a)

(b)

(c)
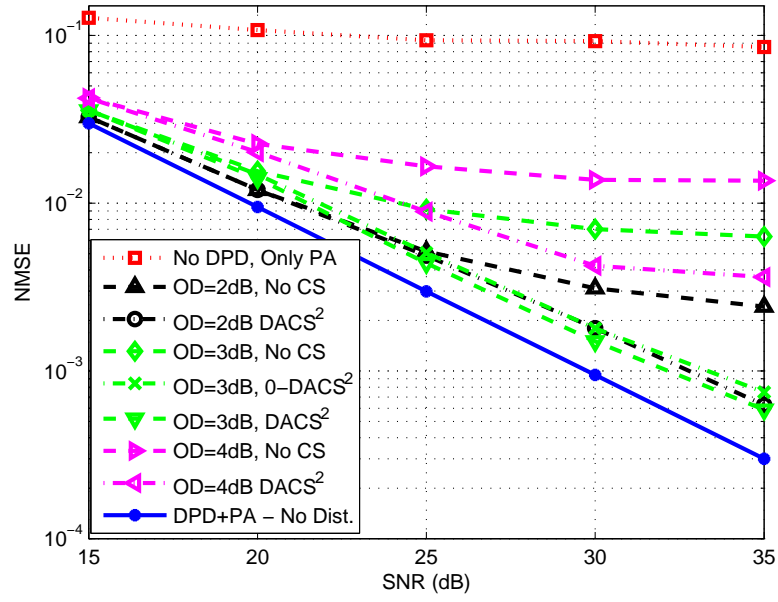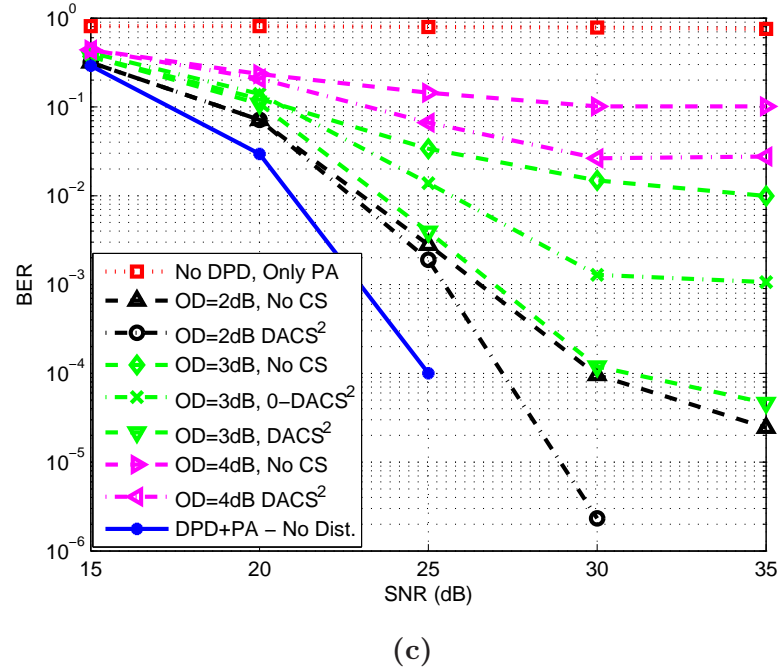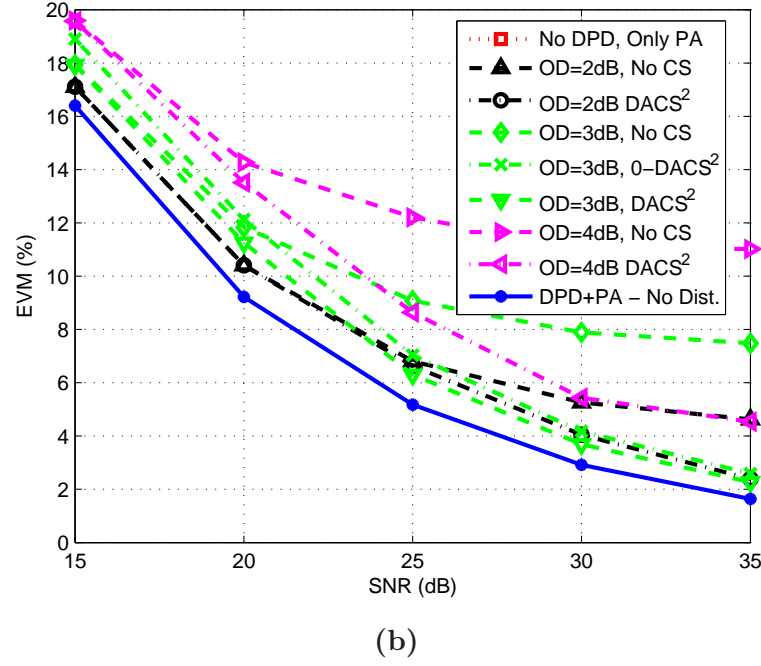
Figure 4.9: Comparison of performance of joint compensation for the Doherty PA via (a) NMSE (b) EVM (c) BER (Note: the 'No CS' case is not shown in the EVM plot as even it's minimum value (at 35 dB) is too high (approx. 29%))

From the Figures 4.9a-4.9c, we note that:

1. A combination of any form of the DPD with the PA corrects for much of the PA distortion. However, the more we overdrive the PA, the more non-linearities occur and the less effective the DPD is in correcting for this distortion.[1]

2. Using DACS for post-compensation after pre-distortion (joint compensation) allows us to overdrive the amplifier by about 3 dB and still maintain relatively good performance—as we see that at an SNR of 35 dB, the 3 dB overdriven amplifier gives an EVM value of about 2.5% which is very close to the EVM of the perfectly linear amplifier (approx. 1.8%).

3. The curves denoting the case when no free-carriers are used for WCS before implementing the DACS algorithm ("0-DACS[2]") show that for a slight reduction in performance, we can increase bandwidth efficiency if we exclude free-carriers at the transmitter.

---

[1]We notice that according to Figures 4.8 and 4.9, CS seems to only work when the Gini Index $\mathcal{G} \geq 0.950$ but from results presented in Chapter 3, it worked for the RFMD amplifier which had a Gini Index of just $\mathcal{G} = 0.718$. The reason for this discrepancy in the threshold in both cases is discussed in the Appendix.

CHAPTER 5

CONCLUSION & FUTURE WORK

5.1 Conclusion

In an attempt to correct the non-linear distortion caused by power amplifiers in OFDM systems, we have been able to show that these non-linearities can sometimes be modeled as sparse phenomenon and so, methods based on the compressed sensing can be used to correct these distortions by estimating and then canceling them at the receiver side. In particular, four distinct algorithms have been presented in this thesis:

1. **"Basic" Compressed Sensing (CS):** This refers to the simplest of algorithms where CS is applied directly using measurements from free-carriers inserted in the frequency-domain transmitted signal.

2. **Weighted CS (WCS):** This improves upon the CS algorithm by giving larger weights to less probable clipping locations and smaller weights to the more prob-

able ones. Doing this forces CS to focus on the more probable locations and thus improves the distortion estimate. Results show that this is equivalent to using 10% (or more) extra free-carriers in basic CS.

3. **Data-Aided CS (DACS):** As a further enhancement of the CS algorithm, we were able to extract measurements from the data-carriers for more runs of CS. Besides allowing us to improve upon the WCS results, this enables us to estimate for the PA distortion even without including any free-carriers on the frequency-domain transmitted signal thus saving bandwidth and enhancing system throughput. The results of DACS are encouraging especially after the second iteration.

4. **Joint Compensation:** The above three methods only perform well when the distortion caused by the amplifier is sparse. In cases where a highly non-linear power amplifier such as the Doherty is used, we implement joint compensation by pre-compensation (via a pre-distorter function at the transmitter) along with post-compensation (via DACS, for example, at the receiver). Using joint compensation allows us to increase the operating average power of the PA by about 3 dB since our post-compensation algorithm would correct for the resulting extra distortions. Even in cases where the PA is not overdriven, this algorithm helps remove any residual distortion that may occur due to a mismatch between the DPD and the PA.

5.2 Future Work

Based on the work presented in this thesis, the following are possible areas which could be considered for further work:

1. In the actual implementation of OFDM in communication standards, free-carriers are usually available at the band edges. As opposed to randomly placed carriers within the signal bandwidth (as adopted in this work), carriers at the edges are readily available. It would be interesting to see how our algorithm can be modified to take advantage of these carriers. One benefit of doing this is that the resulting matrix would be of higher structure and so the structure could possibly be used to reduce the complexity of the algorithm [43].

2. In solving the CS problem, we focused in this thesis on convex relaxation algorithms. Another option would be to consider Bayesian-based recovery for Gaussian and non-Gaussian priors [43, 50]. The advantage of Bayesian techniques is that they are of lower complexity and are also able to incorporate a prior, statistical information about the clipped signal.

3. In our work, we assumed the channel to be known. Future work could explore applying the algorithms presented to a more realistic situation with the entire communication link and then compensating for distortion at the the receiver when channel impairments are present.

4. It is also possible to evaluate the feasibility of implementing (and perhaps, modifying) the techniques proposed in this thesis for power amplifiers which exhibit

memory effects.

APPENDIX: Reason for Variation in Sparsity Recovery Threshold based on the Gini Index

From results presented in Figures 3.10a-3.10c in Chapter 3, we notice that the CS recovery performs well for the RFMD applifier which had a Gini Sparsity Index of $\mathcal{G} = 0.718$. However, in the results of the overdriven Doherty PA presented in Figures 4.3a-4.3b, the performance of even two iterations of DACS hardly gives good results when the overdrive is 4 dB or above which, according to Figure 4.8, translates to a Gini range of $\mathcal{G} \leq 0.945$. The obvious question is: how come CS-recovery breaks for a system with a high sparsity index but still works very well when the sparsity is relatively much lower.

To answer this question, we note that one major difference between the distortion due to the RFMD amplifier (or any general amplifier for that matter) and the distortion due to the overdriven pre-distorted amplifier system is that the later case excludes the many minor clips (i.e. coefficients which are very close to zero but no exactly zero) present in the former case. This is noticed whenthe distortions of both amplifiers as presented in Figures 3.2 and 4.7 are compared.

Apparently, it seems the cause of the relatively low Gini index for RFMD and ZHL might be the very many close-to-zero but non-zero amplitudes which are absent in the overdriven pre-distorted Doherty amplifier. To test if these minor clips are the cause of the discrepancy in the threshold value, the Gini Index for the normal distortion caused by the RFMD amplifier was compared against the index value for the same distortion however, after zeroing-out the lowest 200 amplitudes (i.e. leaving only the

56 highest amplitudes). For the former (normal distortion), the normal Gini Index value for the RFMD amplifier was obtained ($\mathcal{G} = 0.711$) while for the later, it was much higher ($\mathcal{G} = 0.908$).
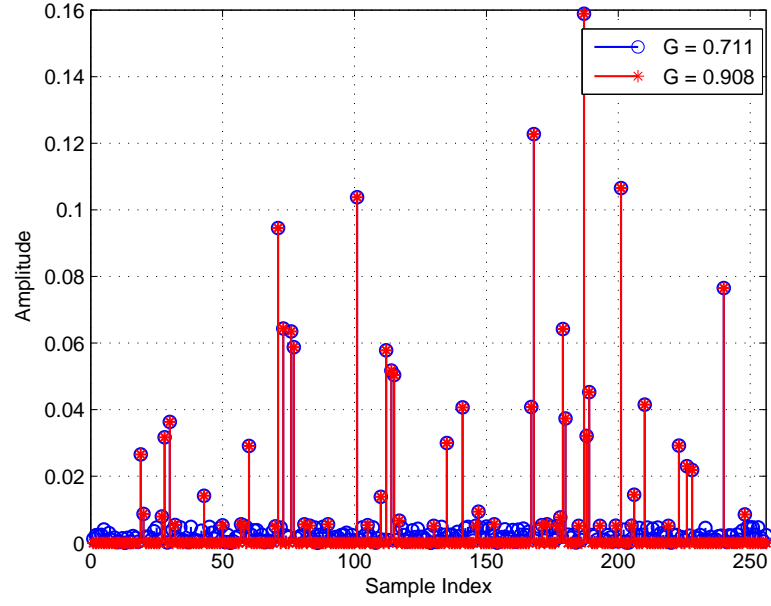
Figure 5.1: Gini index comparison

REFERENCES

[1] IEEE, "Ieee standard for local and metropolitan area networks." IEEE Std 802.16-2009, 2009.

[2] E. T. S. I. E. 3rd Generation Partnership Project (3GPP), "Lte; evolved universal terrestrial radio access (e-utra); user equipment (ue) radio transmission and reception 1." 3GPP TS 36.101 version 10.6.0 Release 10, 2012.

[3] E. T. S. I. (ETSI), "Radio broadcasting systems; digital audio broadcasting (dab) to mobile, portable and fixed receivers." EN 300 401 version 1.3.3, 2001.

[4] G. Fettweis and E. Zimmermann, "Ict energy consumption - trends and challenges," in *Proceedings of the 11th International Symposium on Wireless Personal Multimedia Communications*, September 2008.

[5] F. Ghannouchi and O. Hammi, "Behavioral modeling and predistortion," *IEEE Microwave Magazine*, vol. 10, pp. 52–64, December 2009.

[6] H.-H. Chen, C.-H. Lin, P.-C. Huang, and J.-T. Chen, "Joint polynomial and look-up-table predistortion power amplifier linearization," *Circuits and Systems II: Express Briefs, IEEE Transactions on*, vol. 53, pp. 612–616, August 2006.

[7] A. Ghadam, S. Burglechner, A. Gokceoglu, M. Valkama, and A. Springer, "Implementation and performance of dsp-oriented feedforward power amplifier linearizer," *Circuits and Systems I: Regular Papers, IEEE Transactions on*, vol. 59, pp. 409–425, February 2012.

[8] K. Rawat, M. Rawat, and F. Ghannouchi, "Compensating i-q imperfections in hybrid rf/digital predistortion with an adapted lookup table implemented in an fpga," *Circuits and Systems II: Express Briefs, IEEE Transactions on*, vol. 57, pp. 389–393, May 2010.

[9] T. Jiang and G. Zhu, "Complement block coding for reduction in peak-to-average power ratio of ofdm signals," *Communications Magazine, IEEE*, vol. 43, pp. S17–S22, September 2005.

[10] P. Fan and X.-G. Xia, "Block coded modulation for the reduction of the peak to average power ratio in ofdm systems," *Consumer Electronics, IEEE Transactions on*, vol. 45, pp. 1025–1029, November 1999.

[11] K. Sathananthan and C. Tellambura, "Coding to reduce both par and picr of an ofdm signal," *Communications Letters, IEEE*, vol. 6, pp. 316–318, August 2002.

[12] T. Jiang, W. Xiang, P. Richardson, D. Qu, and G. Zhu, "On the nonlinear companding transform for reduction in papr of mcm signals," *Wireless Communications, IEEE Transactions on*, vol. 6, pp. 2017–2021, June 2007.

[13] T. Jiang, Y. Yang, and Y.-H. Song, "Exponential companding technique for papr reduction in ofdm systems," *Broadcasting, IEEE Transactions on*, vol. 51, pp. 244–248, June 2005.

[14] B. Krongold and D. Jones, "An active-set approach for ofdm par reduction via tone reservation," *Signal Processing, IEEE Transactions on*, vol. 52, pp. 495–509, February 2004.

[15] N. Andgart, P. Ödling, A. Johansson, and P. O. Börjesson, "Designing tone reservation par reduction," *EURASIP J. Appl. Signal Process.*, vol. 2006, pp. 82–82, January 2006.

[16] B. Krongold and D. Jones, "Par reduction in ofdm via active constellation extension," *Broadcasting, IEEE Transactions on*, vol. 49, pp. 258–268, September 2003.

[17] Y. Kou, W.-S. Lu, and A. Antoniou, "A new peak-to-average power-ratio reduction algorithm for ofdm systems via constellation extension," *Wireless Communications, IEEE Transactions on*, vol. 6, pp. 1823–1832, May 2007.

[18] S. H. Han and J. H. Lee, "An overview of peak-to-average power ratio reduction techniques for multicarrier transmission," *Wireless Communications, IEEE*, vol. 12, pp. 56–65, April 2005.

[19] T. Jiang and Y. Wu, "An overview: Peak-to-average power ratio reduction techniques for ofdm signals," *Broadcasting, IEEE Transactions on*, vol. 54, pp. 257–268, June 2008.

[20] H. Ochiai and H. Imai, "Performance of the deliberate clipping with adaptive symbol selection for strictly band-limited ofdm systems," *Selected Areas in Communications, IEEE Journal on*, vol. 18, pp. 2270–2277, November 2000.

[21] D. Kim and G. Stuber, "Clipping noise mitigation for ofdm by decision-aided reconstruction," *Communications Letters, IEEE*, vol. 3, pp. 4–6, January 1999.

[22] H. Chen and A. Haimovich, "Iterative estimation and cancellation of clipping noise for ofdm signals," *Communications Letters, IEEE*, vol. 7, pp. 305–307, July 2003.

[23] H. Saeedi, M. Sharif, and F. Marvasti, "Clipping noise cancellation in ofdm systems using oversampled signal reconstruction," *Communications Letters, IEEE*, vol. 6, pp. 73–75, February 2002.

[24] E. Al-Safadi and T. Al-Naffouri, "On reducing the complexity of tone-reservation based papr reduction schemes by compressive sensing," in *Global Telecommunications Conference, 2009. GLOBECOM 2009. IEEE*, pp. 1–6, December 2009.

[25] E. B. Al-Safadi, "Receiver based papr reduction in ofdm by compressed estimation of sparse signals," Master's thesis, King Fahd University of Petroleum & Minerals, 2010.

[26] L. Xia, Z. Li, T. Youxi, and L. Shaoqian, "Analysis of the performance of iterative estimation and cancellation of clipping non-linear distortion in ofdm," in *Future Generation Communication and Networking (FGCN 2007)*, vol. 2, pp. 193–197, December 2007.

[27] E. Candés and M. Wakin, "An introduction to compressive sampling," *IEEE Signal Processing Magazine*, vol. 25, pp. 21–30, March 2008.

[28] E. J. Candés, J. K. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Communications on Pure and Applied Mathematics*, vol. 59, pp. 1207–1223, August 2006.

[29] E. J. Candés and T. Tao, "The dantzig selector: Statistical estimation when p is much larger than n," *Annals of Statistics*, vol. 35, no. 6, pp. 2313–2351, 2007.

[30] E. Candés and P. Randall, "Highly robust error correction by convex programming," *Information Theory, IEEE Transactions on*, vol. 54, pp. 2829–2840, July 2008.

[31] J. Tropp, "Just relax: convex programming methods for identifying sparse signals in noise," *Information Theory, IEEE Transactions on*, vol. 52, pp. 1030–1051, March 2006.

[32] L. Ding, G. Zhou, D. Morgan, Z. Ma, J. Kenney, J. Kim, and C. Giardina, "A robust digital baseband predistorter constructed using memory polynomials," *Communications, IEEE Transactions on*, vol. 52, pp. 159–165, January 2004.

[33] D. Morgan, Z. Ma, J. Kim, M. Zierdt, and J. Pastalan, "A generalized memory polynomial model for digital predistortion of rf power amplifiers," *Signal Processing, IEEE Transactions on*, vol. 54, pp. 3852–3860, October 2006.

[34] C. Gini, "Measurement of inequality of incomes," *The Economic Journal*, vol. 31, pp. 124–126, March 1921.

[35] N. Hurley and S. Rickard, "Comparing measures of sparsity," *Information Theory, IEEE Transactions on*, vol. 55, pp. 4723–4741, October 2009.

[36] S. Butler and J. Stefani, "Supervisory run-to-run control of polysilicon gate etch using in situ ellipsometry," *Semiconductor Manufacturing, IEEE Transactions on*, vol. 7, pp. 193–201, May 1994.

[37] E. J. Candés, J. K. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *Information Theory, IEEE Transactions on*, vol. 52, pp. 489–509, February 2006.

[38] E. J. Candés and T. Tao, "Near optimal signal recovery from random projections: Universal encoding strategies," *IEEE Transactions on Information Theory*, vol. 52, pp. 5406–5425, December 2006.

[39] D. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, pp. 1289–1306, April 2006.

[40] S. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *Signal Processing, IEEE Transactions on*, vol. 41, pp. 3397–3415, December 1993.

[41] J. Tropp and A. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *Information Theory, IEEE Transactions on*, vol. 53, pp. 4655–4666, December 2007.

[42] G. Rath and C. Guillemot, "Sparse approximation with an orthogonal complementary matching pursuit algorithm," in *Acoustics, Speech and Signal Process-*

ing, 2009. ICASSP 2009. IEEE International Conference on, pp. 3325–3328, april 2009.

[43] A. Quadeer, S. Ahmed, and T. Al-Naffouri, "Structure based bayesian sparse reconstruction using non-gaussian prior," in *Communication, Control, and Computing (Allerton), 2011 49th Annual Allerton Conference on*, pp. 277–283, September 2011.

[44] S. Ji, Y. Xue, and L. Carin, "Bayesian compressive sensing," *Signal Processing, IEEE Transactions on*, vol. 56, pp. 2346–2356, June 2008.

[45] E. Candés and J. Romberg, "Sparsity and incoherence in compressive sampling," *Inverse Problems*, vol. 23, pp. 969–985, July 2007.

[46] E. Candés, M. Wakin, and S. Boyd, "Enhancing sparsity by reweighted l_1 minimization," *Journal of Fourier Analysis and Applications*, vol. 14, pp. 877–905, 2008. 10.1007/s00041-008-9045-x.

[47] D. Wisell and M. Isaksson, "Derivation of a behavioral rf power amplifier model with low normalized mean-square error," in *Microwave Integrated Circuit Conference, 2007. EuMIC 2007. European*, pp. 485–488, Oct. 2007.

[48] M. D. McKinley, K. A. Remley, M. Myslinski, J. S. Kenney, D. Schreurs, and B. Nauwelaers, "EVM Calculation for Broadband Modulated Signals," 2004.

[49] S. Forestier, P. Bouysse, R. Quere, A. Mallet, J.-M. Nebus, and L. Lapierre, "Joint optimization of the power-added efficiency and the error-vector measurement of

20-ghz phemt amplifier through a new dynamic bias-control method," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 52, pp. 1132–1141, april 2004.

[50] P. Schniter, L. Potter, and J. Ziniel, "Fast bayesian matching pursuit," in *Information Theory and Applications Workshop, 2008*, pp. 326–333, February 2008.

[51] M. Honkanen and S.-G. Haggman, "New aspects on nonlinear power amplifier modeling in radio communication system simulations," in *Proc. IEEE Int. Symp. on Personal, Indoor, and Mobile Radio Comm. PIMRC '97*, vol. 3, pp. 844–848, September 1997.

[52] O. Hammi, J. Sirois, S. Boumaiza, and F. Ghannouchi, "Design and performance analysis of mismatched doherty amplifiers using an accurate load-pull-based model," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 54, pp. 3246–3254, August 2006.

[53] S. S. Das, *Techniques to enhance spectral efficiency of OFDM wireless systems.* PhD thesis, CTIF, Aalborg University, 2007.

[54] E. B. Al-Safadi and T. Y. Al-Naffouri, "Peak reduction and clipping mitigation by compressive sensing," *CoRR*, vol. abs/1101.4335, 2011.

Vitae

- Damilola Sadiq Owodunni.

- Born in Lagos, Nigeria on April 26th, 1987.

- Received Bachelor of Science (B.Sc.) degree in Electrical Engineering (with emphasis in Computer Engineering) from University of Mississippi, Mississippi, USA in May 2009.

- Joined King Fahd University of Petroleum & Minerals in September 2009.

- Email: dksadiq@kfupm.edu.sa

- Permanent Address: Lagos, Nigeria