AN ABSTRACT OF THE THESIS OF

Min Zhuo for the degree of Master of Science in Electrical & Computer Engineering presented on September 14, 2000. Title:

Efficient Digital Predistortion Techniques for Power Amplifier Linearization.

Redacted for Privacy

Abstract approved: _____

John T. Stonick

The importance of spectral efficiency in mobile communications often requires the use of non-constant-envelop linear digital modulation schemes. These modulation techniques carry signal information in both magnitude and phase, thus they must be linearly amplified to avoid nonlinear signal distortion which is not correctable in a typical receiver. A second difficulty in utilizing these modulation formats is that nonlinear amplification generates out-of-band power (spectral regrowth). Therefore, to achieve both high energy efficiency and spectral efficiency, some forms of linearization must be used to compensate for the nonlinearity of power amplifiers. One powerful technique that is amenable to monolithic integration is digital signal predistortion. Most predistorters try to achieve the inverse nonlinear characteristic of High Power Amplifier(HPA). In this thesis a new multi-stage digital adaptive signal predistorter is presented. The scheme is developed from the direct iterative method with low memory requirement proposed by Cavers [1] in combination with the multi-stage predistortion proposed by Stonick [2]. To make the predistorter more compact a very simple and fast method called the complementary method is proposed. The complementary method has prominent advantages over other digital predistorters in terms of stability of the algorithm, complexity of the algorithm and computational load.

Efficient Digital Predistortion Techniques for Power Amplifier Linearization

by

Min Zhuo

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Master of Science

Presented September 14, 2000
Commencement June 2001

Master of Science thesis of Min Zhuo presented on September 14, 2000

APPROVED:

Redacted for Privacy

Major Professor, representing Electrical & Computer Engineering

Redacted for Privacy

Chair of the Department of Electrical & Computer Engineering

Redacted for Privacy

Dean of the Graduate School

I understand that my thesis will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my thesis to any reader upon request.

Redacted for Privacy

Min Zhuo, Author

ACKNOWLEDGMENT

The completion of the thesis was made possible by numerous persons with their incessant guidance, encouragement and discussion. First of all, I am profoundly grateful to my major professor, Dr. John T. Stonick. He shared a great deal of his precious time and ideas during the work, and contributed to the fulfillment of my Masters Program at Oregon State University.

I am also grateful to the professors on my committee: Dr. Annette Von Jouanne and Dr. Raghu Settaluri; the graduate Council Representative, Dr. Michael E. Kassner, for their time and advice.

Special thanks are given to Ferne Simendinger and Sarah O'Leary. Their help is highly appreciated. And I sincerely thank my parents who have supported me for so many years.

I would like, finally, to thank my husband, Dr. Ji Zheng, for his encouragement and patience throughout this long process.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

LIST OF FIGURES

LIST OF FIGURES (Continued)

LIST OF FIGURES (Continued)

EFFICIENT DIGITAL PREDISTORTION TECHNIQUES FOR POWER AMPLIFIER LINEARIZATION

1. INTRODUCTION

Wireless communication is one of the fastest growing fields in the engineering world. Particularly during the past ten years, the mobile radio communication industry has grown by orders of magnitude [3]. The first U.S. cellular telephone system, AMPS, uses analog FM as modulation scheme and time division multiple access (TDMA) for its multi-access scheme. In AMPS, each channel occupies $30KHz$. In late 1991, the first U.S. digital cellular system was installed, which supports 3 users in the same $30KHz$ channel. The digital modulation scheme is $\frac{\pi}{4}$ differential quadrature phase shift keying ($\frac{\pi}{4}$DQPSK), and time division multiple access(TDMA) is used. In 1993, a digital cellular system based on code division multiple access (CDMA) was proposed by Qualcomm, Inc, which supports a variable number of users in a $1.25MHz$-wide channel using direct sequence spread spectrum. With the development of the very large-scale integration(VLSI), digital signal processing(DSP) technology and information theory, digital communication systems outperform analog communication systems in several ways [4]. Digital communication systems provide much higher capacity than the analog systems because of their improved spectrum efficiency. Also digital technology facilitates implementation of new services such as data transmission and several supplementary services that generate additional revenues. Other advantages include greater noise immunity and robustness to channel impairment, easier multiplexing of various forms of information(e.g., voice, data, and video), and great security, the detection and correction of errors.

1.1. Nonlinearity of Amplifier in Wireless Digital Communication Systems

In wireless digital communication system digitally modulated signals can be transmitted directly as baseband signals or converted to a radio frequency. In the latter case these modulated signals are often generated at low power levels and then amplified to high levels for transmission to distant destinations. Power amplifiers are usually in the last stage of the transmitter. Power amplifiers have been traditionally categorized into many classes: A, B, C, D, E, F, etc [5]. Power amplifiers can be evaluated in terms of power efficiency and linearity. Linear power amplifiers are typically less efficient than nonlinear power amplifiers. By efficiency we mean the power delivered to the load divided by the power consumed by the HPA. Most high power amplifiers(HPAs) working in radio frequency are nonlinear devices.

Digitally modulated signals can be categorized as constant-envelop signals, such as FSK signals; and variable-envelop signals, such as QPSK signals and QAM signals. Constant-envelop signals and variable-envelop signals behave differently in a nonlinear system.

The modulated RF signals which are the input signals for power amplifiers can be expressed by Equation 1.1.

$$x(t) = A(t)\cos[\omega_c t + \phi(t)] \tag{1.1}$$

where A(t) is the envelop of the signal. For constant-envelop signals A(t) does not vary with time, while for variable-envelop signals A(t) varies with time.

The simplified characteristic of a nonlinear system exhibiting a third-order memoryless nonlinearity as its input-output relationship can be represented by,

$$y(t) = x(t) + \alpha_3 x^3(t) \tag{1.2}$$

The different effects caused by the nonlinear device on constant- and variable-envelop signals is summarized in the following.

For constant-envelop signals let $A(t) = A_c$, the output is given by,

$$\begin{aligned} y(t) &= A_c \cos[\omega_c t + \phi(t)] \\ &\quad + \frac{\alpha_3 A_c^3}{4} \cos[3\omega_c t + 3\phi(t)] \\ &\quad + \frac{3\alpha_3 A_c^3}{4} \cos[\omega_c t + \phi(t)] \end{aligned} \tag{1.3}$$

The second term in Equation 1.3 represents a modulated signal around $\omega = 3\omega_c$. Since the bandwidth of the original signal, $A_c \cos[\omega_c t + \phi(t)]$, is typically much less than ω_c, we note from Carson's rule that the bandwidth occupied by $\cos[3\omega_c t + 3\phi(t)]$ is also quite small. Thus, the shape of the spectrum in the vicinity of ω_c remains unchanged. The variable-envelop signals is expressed as,

$$x(t) = x_I(t) \cos \omega_c t - x_Q(t) \sin \omega_c t \tag{1.4}$$

where $x_I(t)$ and $x_Q(t)$ are the baseband I and Q components. The output is given as,

$$\begin{aligned} y(t) &= x_I(t) \cos \omega_c t - x_Q(t) \sin \omega_c t \\ &\quad + \alpha_3 x_I^3(t) \frac{\cos 3\omega_c t + 3 \cos \omega_c t}{4} \\ &\quad - \alpha_3 x_Q^3(t) \frac{-\cos 3\omega_c t + 3 \sin \omega_c t}{4} \end{aligned} \tag{1.5}$$

The output in Equation 1.5 contains the spectra of $x_I^3(t)$ and $x_Q^3(t)$ centered around ω_c. Since these components generally occupy a broad spectrum than $x_I(t)$ and $x_Q(t)$ the spectral width grows when a variable-envelop signal passes through a nonlinear system. Since power efficiency is an important concern in today's wireless communication systems, especially in handset design, most power amplifiers used in mobile communication system are nonlinear devices. Thus, it is desirable to employ modulation schemes with constant-envelop signals which do not experience spectral regrowth when amplified by nonlinear amplifiers.

Constant envelop modulation such as FSK was widely used in the past because of its immunity to noise. In constant envelop modulation schemes the amplitude component of the signal generally does not carry information. As a result, it is not necessary to transmit the waveform with strict fidelity. Thus these modulation schemes allow power amplifiers to be operated in the nonlinear region near saturation for power efficiency. Due to the constant envelop the nonlinearity doesn't generate intermodulation products in nearby channels. However, with the increasing demand of the wireless communications, more and more services are available or under development, and the the number of wireless subscribes will soon be equal to the number of wireline customers. The continuing pressure on the limited radio spectrum available is forcing the use of more spectral efficient modulation schemes, such as 4-state quadrature phase shift keying (QPSK) and 16-state quadrature amplitude modulation (16QAM) [1] [6], which can transmit information using a narrow frequency bandwidth as compared with constant modulation schemes.

Unlike FSK signal, which has constant envelop and carries information only in frequency, QPSK and 16QAM signals with variable-envelop carry signal information in signal's amplitude and phase. Because QSPK and QAM are linear modulation schemes, any distortion in amplitude and phase caused by nonlinear devices results in an increase in bit error rate(BER), and hence degrades system performance. The above problem is more predominant in handsets for CDMA system, where the peak-to-average ratio of modulated signals can vary over a range of 3 to 12dB [6]. Furthermore power amplifiers used in handsets must operate in nonlinear range to keep long battery life. However, the nonlinear terms of a HPA transfer function give rise to spectral regrowth, which brings out-of-band interference in adjacent channels. In the mobile environment, the restrictions by FCC on out-of-band emissions are stringent. Generally in the mobile data communication systems, the adjacent channel interference (ACI) requirement ranges from -45 to -60 dBc relative to the in-band carrier level [7] [8] [9]. Therefore, highly linear power amplifiers are required for the amplification of linearly modulated signals. A simple way to achieve linear amplification is to back off the power amplifier from saturation so that it

operates in the linear region of its transfer characteristic. However, such an amplifier has a low DC-to-RF conversion efficiency as compared to an amplifier that operates in the nonlinear region near saturation. Furthermore, low DC-to-RF conversion necessitates the high current operating point, resulting in undesired thermal noise. Recall power amplifiers must have the highest possible efficiency for portable equipments, where cost and heat dissipation are prohibitive factors. A major challenge in designing a high power amplifier is to maintain linearity without compromising the power efficiency of the amplifier. Thus, the nonlinearity of amplifiers should be studied and techniques of linearization of power amplifiers in mobile communication systems need to be investigated.

1.2. Linearization Techniques

To reduce the adjacent channel interference we most somehow linearize power amplifiers used for wireless communication systems that employ linear modulation schemes. A number of linearization techniques have been studied, including: Cartesian Feedback Loop, Linear Amplification with Nonlinear Components(LINC), Feedforward Loop and Predistortion.

1.2.1. Cartesian Feedback Loop

The linearization technique of Cartesian Feedback Loop [1] [5] [7] [10] [11] forms a complete linear transmitter, as opposed to simply a linear amplifier, and its structure shown in Fig. 1.1. It takes baseband signals, in I and Q formats, and translates these signals to an RF carrier frequency at a high power level. Thus, the upconversion and power amplification processes are combined and the whole is subject to the distortion improvement of the linearization. The transmitter output is sampled just after the final RF amplifier, and synchronously demodulated, to recover quadrature cartesian components of the modulation. These signals are used to provide negative feedback, subtracting from the modulating signals to generate a loop error signal. Provided that

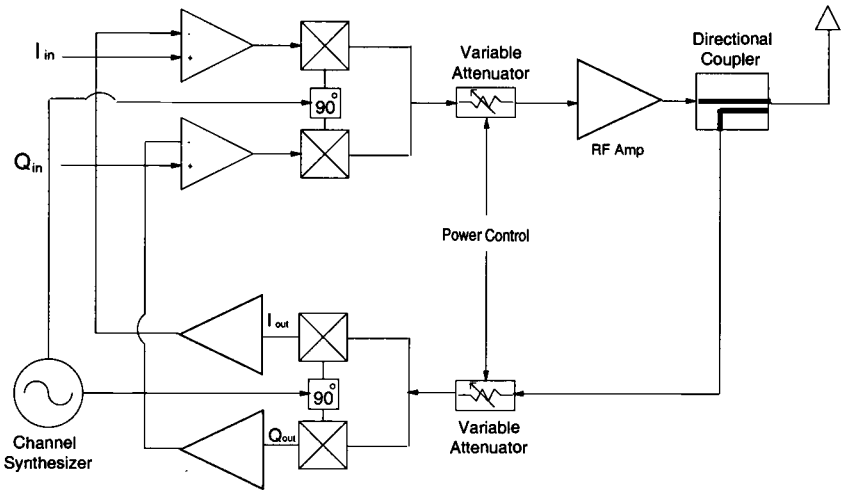

FIGURE 1.1: Structure of Cartesian Loop linearization

the loop gain is of sufficient magnitude, the feedback loop will, in theory, continuously correct for any non-linearity in the upconversion RF amplification stages. As with any feedback system, its performance is limited by the delay around the loop and the linearity improvement depends on the bandwidth over which the feedback operates. Moreover, the stability of the feedback loop is an obstacle in the design of such a system. The RF amplifiers create a significant phase shift of the feedback signals, which varies with frequency and output level. If this phase shift become excessive, oscillation will occur in the feedback loop.

1.2.2. LINC

The LINC technique [5] [7] [10] [11] is a linear transmitter technique involving RF synthesis. This means that the linear RF waveform is only created at the output of the transmitter with all of the internal processes within the transmitter being non-linear. The basic format of the LINC technique is shown in Fig.1.2. The modulating signal is generated in DSP from the input baseband information. The input signal with variable-envelop is split into two constant amplitude phase-modulated signals, which are

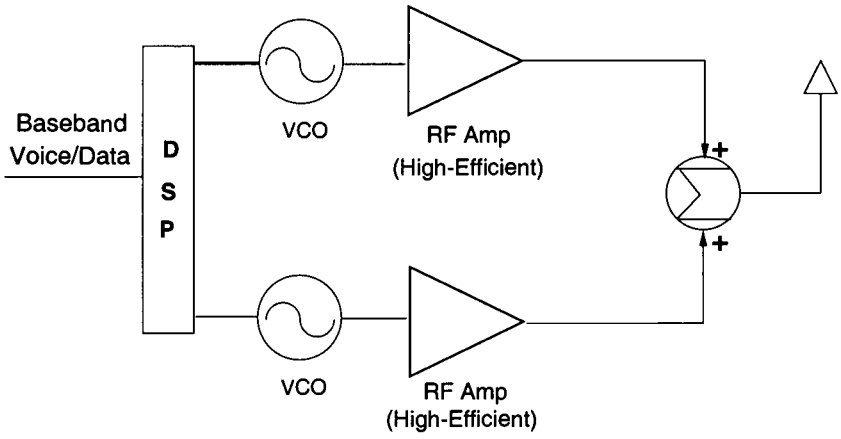

FIGURE 1.2: Structure of LINC linearization

then amplified using two well-matched nonlinear amplifiers. After RF upconversion and non-linear power amplification, the two signals will add to produce the required linear output signal with unwanted elements appearing in anti-phase and hence cancelling. The difficulties with the method are the design of two well-matched amplifier chains since gain and phase mismatch between the two signal path results in residual distortion, and a method of combining the two high-power signals from the amplifiers because the output adder might introduce significant loss since it must achieve a high isolation between two power amplifiers. Also the split of the signal gives substantial complexity since it involves the complex phase modulation.

1.2.3. Feedforward Loop

The structure of the feedforward amplifier [5] [9] [11] is shown in Fig.1.3. Its operation may be clearly seen by referring to the two-tone test spectra shown at various points throughout the diagram. The input signal is split to form two identical paths, although the ratio used in the splitting process need not to be equal. The signal in the top path is amplified by the main power amplifier and the nonlinearities in this amplifier result in intermodulation and harmonic distortions. The directional coupler, $C1$, takes

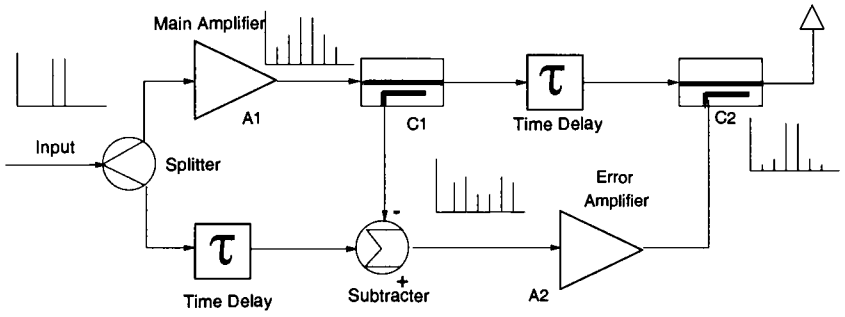

FIGURE 1.3: Structure of Feedforward Loop linearization

a sample of the main amplifier output signal and feeds it to the subtracter where a time-delayed portion of the original signal, present in the lower path, is subtracted. The result of this subtraction process is an error signal containing the distortion information from the main amplifier; ideally none of the original signal energy would remain. The error signal is then amplified linearly to the required level to cancel the distortion in the main path and fed to the output coupler. The main-path signal through coupler, $C1$, is time delayed by an amount approximately equal to the delay through the error amplifier, $A2$, and fed to the output coupler in anti-phase to the amplified error signal. The error signal will then cancel the distortion information of the main path signal leaving only an amplified version of the original input signal. The problems for Feedforward Loop are: the match of subtracter and error amplifier; the precision of estimation of time delay for analog system, and the sum of the two signal at output.

1.2.4. Predistortion

Predistortion [1] [2] [7] [9] [10] [11] is conceptually the simplest form of linearization for a power amplifier. It simply involves the creation of a distortion characteristic which is precisely complementary to the distortion characteristic of the power amplifiers and cascading the two in order to ensure that the resulting system has little or no input-output distortion. The basic form of a predistortion linearizer is shown in Fig. 1.4.

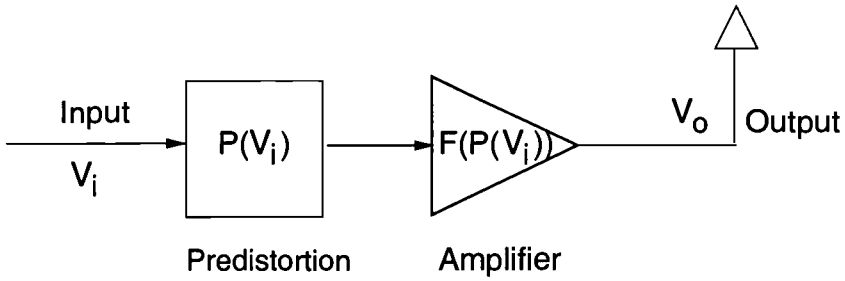

FIGURE 1.4: Scheme of Predistortion

The predistortion function, $P(\bullet)$, operates on the input signal in such a manner that its input signal is distorted precisely complementary manner to the distortion produced by the power amplifiers, $F(\bullet)$. The output signal is, ideally, an amplified but undistorted replica of the input signal. There are several kinds of predistorters as follows.

1. RF Predistortion - the non-linear predistorting element operates at the final carrier frequency.

2. IF Predistortion - the predistorting element operates at a convenient intermediate frequency, thereby possibly allowing the same design to be utilized for a number of different carriers frequencies.

3. Baseband Predistortion - prior to the advent of digital signal processing devices, this technique had few advantages over RF and IF techniques. It is now, however, a powerful tool. In this case, the predistortion characteristic is typically stored as a table within a DSP or represented by polynomials which represent the nonlinearity of the predistorter. It is possible to use feedback to provide updating information for these coefficients stored in the table or coefficients for polynomials; thus it is called adaptive baseband predistortion.

The degree of linearity improvement which can be achieved in practice depends upon a wide variety of different considerations, and in particular on the form of the transfer characteristic of the power amplifiers since the aim of the predistorter is to achieve the transfer characteristic which is inverse to those of power amplifiers.

Both RF predistortion and IF predistortion need to fabricate a circuit with a required transfer characteristic. This is not a trivial problem and a large number of different networks have been utilized over these years in an attempt to mimic types of characteristic. For such networks to achieve a high level of performance, however, they often need to be designed, or at least adjusted, for each individual amplifier.

Among these above linearization techniques, Cartesian Feedback Loop, LINC and Feedforward method, RF predistortion and IF predistortion have been utilized in complex, expensive RF and microwave systems, but they have not yet found their way into low-cost portable terminals. This is because the above methods generally complicate the design, require frequent adjustments, and become less effective as device characteristics drift. Also they suffer from limitations in bandwidth, precision or stability.

Since the baseband predistortion is implemented digitally, a greater degree of precision can be achieved when computing the predistortion coefficients, digital signal processing techniques can be employed to improve the performance and, the implementation can be programmable to fit different needs. Also, unlike analog systems, there is no concern for stability in adaptive predistortion schemes. Finally, with the availability of high speed DSP, adequate millions instruction per second (MIPS) levels are available to treat the wideband signals found in today's advanced spread spectrum systems [6].

The simplified schematic of an adaptive digital baseband predistortion system is shown in Fig. 1.5.

A fully adaptive digital predistortion system requires the addition of a predistortion circuit consisting of a digital predisorter and a look-up-table(LUT) to the transmission path in addition to a feedback path consisting of a demodulator and an adaptation circuit for updating a LUT. Most common implementations of a digital predistortion utilize standard DSPs. Such processes typical operate with a wordlength of 16 or 32 bits, which provides sufficient accuracy for most applications. In specific applications, application-specific ICs(ASIC) are designed to implement the predistortion systems, providing the flexibility in controlling wordlength and power consumption.

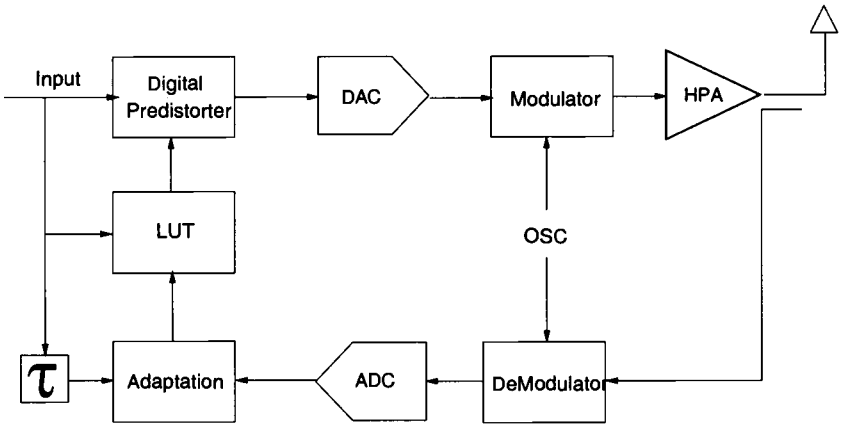

FIGURE 1.5: Structure of Digital Adaptive Predistortion

1.3. Content of Thesis

In this paper, several digital predistortion techniques are studied and simulated with the objective to increase the performance of the overall system, in terms of linearity, stability and speed of algorithm, and cost of implementation. We start from the method, called the direct iterative method proposed by Cavers, with theoretical analysis and simulation. Then we move on to a new method, called the multi-stage method, which tries to solve the problem caused by iteration and computational complexity and load. Based on the idea of the multi-stage predistorter a very simple and fast digital predistorter called the complementary method is proposed. In the following chapters comparisons on the above three methods are provided. In the end of this thesis conclusions are reached followed by several open issues for future research.

2. DIRECT ITERATIVE DIGITAL PREDISTORTION

2.1. Introduction

As stated in the last chapter the digital baseband predistorter is a robust and adaptive method for amplifier linearization in digital wireless communication systems. Most digital baseband predistorters use LUT techniques. The methods to construct the LUT differentiate the predistorters. Earlier in 1983 Saleh [12] proposed a method to adjust the transmitting constellation points on a point by point basis. The algorithm required the data to be in polar form. Later a simpler method based on Cartesian co-ordinates was proposed by Feng [12], which used only the sign of the error signal to perform the adaptation, removing the requirement for A/D converters in feedback. Nagata [1] proposed the more successful method, called as mapping predistorter, whose structure is shown in Fig. 2.1.

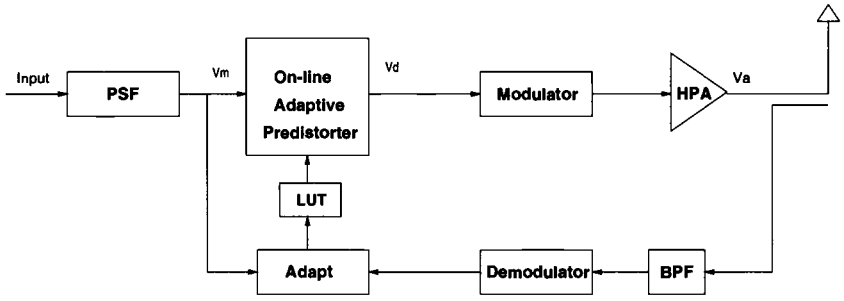

FIGURE 2.1: Structure of Mapping Predistorter

It maps the input complex signal V_m, which consists (V_{mi}, V_{mq}), to the predistorted equivalent complex value V_d consisting of (V_{di}, V_{dq}). The mapping can be implemented by two tables with the forms,

$$V_{di} = H_i(V_{mi}, V_{mq}) \tag{2.1}$$

$$V_{dq} = H_q(V_{mi}, V_{mq}) \tag{2.2}$$

The disadvantage of the construction of LUT is that the speed of convergence is slower than systems employing one-dimensional tables. If the size of LUT is kept low the necessary interpolation between points is more difficult for a two-dimensional table than a one-dimensional table. Nagata avoided interpolation by using a one-dimensional LUT which represents each possible combination of V_{mi} and V_{mq} signal. The memory requirement was large and the adaptation time was very slow.

However since distortions caused by the nonlinearity of power amplifiers depend only on the amplitude variation, it is possible to simplify the LUT by indexing the LUT with the amplitude of the complex signal, $X_m = |V_m| = \sqrt{(V_{mi}^2 + V_{mq}^2)}$. Caver proposed a new predistorter based on the complex gain in [1]. The structure of the complex gain based predistortion is shown in Fig. 2.2. What makes it different from the mapping predistortion is that the predistorter predistorts the signal only according to the amplitude of the signal.

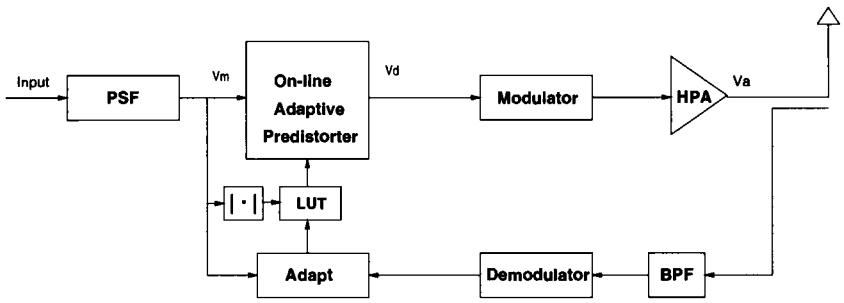

FIGURE 2.2: Structure of Complex Gain based Predistorter

2.2. Theory of Predistortion based on the Complex Gain

2.2.1. Model of a Nonlinear Power Amplifier

It is well known that high power amplifiers(HPAs) exhibit two kinds of nonlinearities: nonlinearities in the input-output power relationship; and in the input-output phase relationship. These nonlinearities are referred to as AM/AM and AM/PM conversions. Also it is reasonable for many applications to assume that the nonlinearity of HPAs is memoryless and varies extremely slowly with time. The generic expression for an HPA can be written as the following Equation 2.3. The signal designations refer to the complex baseband signals or the complex envelop of the bandpass signals.

$$V_{out} = V_{in} \cdot F(r_{in}) \tag{2.3}$$

where, $r_{in} = |V_{in}|$ is the amplitude of the input signal, V_{in} and V_{out} are complex equivalent baseband representations of the instantaneous HPA's input and output complex envelops. $F(r_{in})$ is the complex gain of the HPA. Fig. 2.3 and Fig. 2.4 show the AM/AM and AM/PM conversions for a typical class AB amplifier. The effect of compression is clearly evident at high input level. The AB class HPA with characteristics shown in fig. 2.3 and Fig. 2.4 is used as the HPA model throughout this thesis.

2.2.2. Model of Predistorter

It is obvious that the characteristic of the predistortor is the inverse of that of HPA. Specifically, the predistorter introduces amplitude distortion and phase predistortion in terms of input signal amplitude. And the distortion introduced by the predistortor cancels the distortions coming from the HPA. It is reasonable to conclude that the predistorter is also a nonlinear device with the characteristic expressed by AM/AM and AM/PM conversions, which leads us to describe the predistorter by,

$$V_d = V_m \cdot P(r_m) \tag{2.4}$$

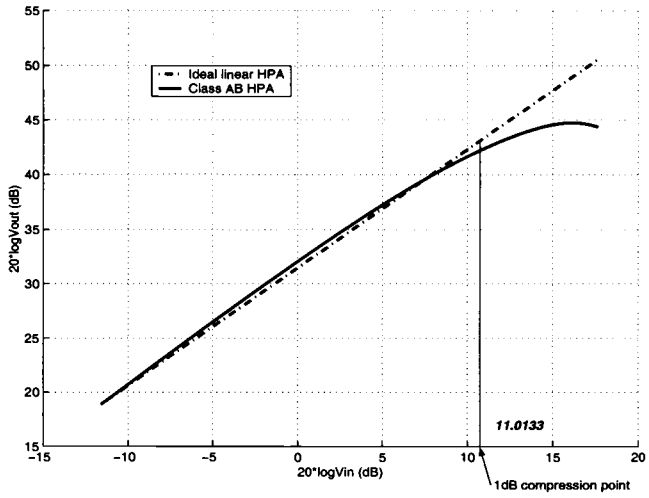

FIGURE 2.3: Nonlinear AM/AM conversion for a class AB HPA

where, $r_m = |V_m|$ is the amplitude of the input signal, V_m and V_d are complex baseband representations of the instantaneous input and output of the predistorter. $P(r_m)$ is the complex gain of predistorter. The desire is that $P(r_m)$ and $F(r_{in})$ bring distortions in a complementary manner.

2.2.3. System Analysis

Given Equation 2.3 and 2.4 the output at amplifiers is given by,

$$V_a = V_m \cdot P(r_m) \cdot F(r_m \cdot |P(r_m)|) \tag{2.5}$$

Ideally the above equation should be equal to the product of amplifier's input and ideal gain, K, hence the final equation that the predistorter should meet is given by,

$$K - P(r_m) \cdot F(r_m \cdot |P(r_m)|) = 0 \tag{2.6}$$

The problem is to find $P(\cdot)$ to make Equation 2.6 hold. Thus, the predistorter is seen to be only dependent on the signal's amplitude. Since no information for $F(r_m)$ is available, an analytical solution can't be obtained. A numerical iterative method is

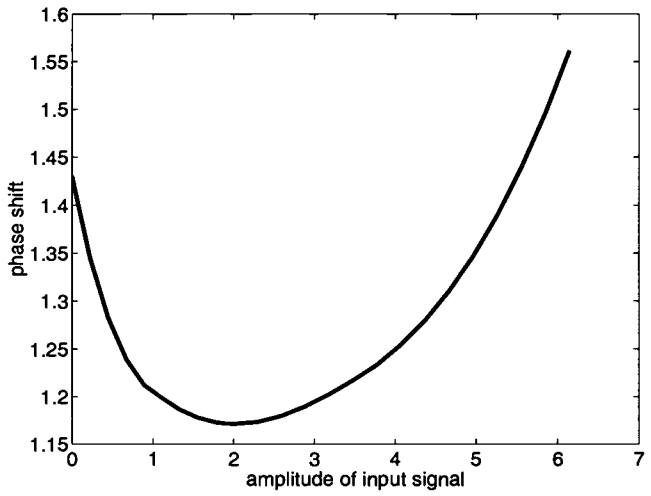

FIGURE 2.4: Nonlinear AM/PM conversion for a class AB HPA

employed to get the characteristic function of the predistorters, $P(r_m)$, by the adaptive estimator which minimizes the loop error as defined by Equation 2.7, which is the difference between the actual complex modulation envelop of the HPA and the desired modulation envelop.

$$E_g = V_a - K \cdot V_m \tag{2.7}$$

A direct link may be established between loop error and the characteristic function of HPA and predistorter, thus the loop error can be written by,

$$E_g(P) = V_m \cdot P(r_m) \cdot F(r_m \cdot |P(r_m)|) - K \cdot V_m \tag{2.8}$$

The task of the adaptive estimator is to calculate the characteristic of the predistortion such that for all values of r_m, the loop error is equal to zero. A convenient method of realizing $P(r_m)$ is to use a LUT, which is indexed by the amplitude of input signal, X_m. The table entry is a complex number $P(X_{m,i})$ representing the complex gain required to predistort the input signal with amplitude equal to $X_{m,i}$.

Since the derivative of $E_g(P)$ is not available, Newton's method can't be applied to Equation 2.8. A reasonable alternative, which offers convergence speed intermediate

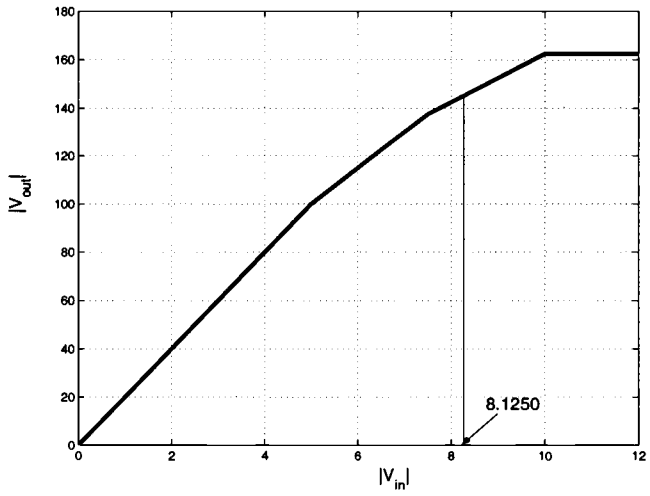

FIGURE 2.5: AM/AM conversion for a nonlinear HPA

between linear and quadratic convergence, is the secant method. The iteration for $P_{i,K+1}$ is done by Equation 2.9.

$$P_{i,K+1} = \frac{P_{i,k-1}E_g(P_{i,k}) - P_{i,k}E_g(P_{i,k-1})}{E_g(P_{i,k}) - E_g(P_{i,k-1})} \tag{2.9}$$

where, $P_{i,K}$ is the kth value of ith table entry. The update is applied when r_m equals to the ith table entry $X_{m,i}$.

2.2.4. Table Spacing in LUT

For any amplifier there is a specific range corresponding to the defined ideal gain where we try to linearize the amplifier. We will use a simple example to illustrate this point. We created a simplified HPA with the piecewise linear AM/AM characteristic shown in Fig. 2.5.

If the ideal gain is set as 20, which is the actual gain when the $r_m < 5.0$, it is obvious that the maximum linearizable input range is [0, 8.125] since the output from the HPA can't exceed 162.5, which is called the ceiling, or the saturation point for the HPA. However, if the ideal gain is set as more than 20, say 25, the maximum linearizable

input range will be [0, 6.5]. For predistortion the first step is to scale the input signal into the linearizable range if the range of the input signal exceeds this range. If the scale factor is k_s the equivalent gain K_a is the product of k_s and K, generally K_a is less then K. If the input signal occupies the range of [0, 12], in the first case: K, the ideal gain, is set as 20, the scale factor $K_s = 0.6770$, the equivalent $K_a = 13.54$; in the second case: K is set as 25, the scale factor $K_s = 0.5416$, the equivalent $K_a = 13.54$. Note K_a is constant, a function of only the ceiling of the HPA. If the ideal gain is set higher, the linearizable range will be smaller, which requires that the LUT has higher resolution. It is important to set the target ideal gain, which should depend on the nonlinearity of HPA and the dynamic range of the input signal, etc.

If the saturated output power for the HPA is P_{sat} the maximum linearizable range for amplitude of the input signal is $[0, \frac{\sqrt{P_{sat}}}{K}]$. For an actual HPA whose AM/AM characteristic is a smooth curve the predistorter tries to linearize HPA over a slightly smaller range. Generally, the ideal gain k is set as the average gain over the full input signal dynamic range. The span S is used to define the fraction of the saturated power over which we attempt linearization. The maximum output power is given by SP_{sat}. An alternative description for the span is the peak backoff (PBO) of the HPA in decibels in Equation 2.10.

$$PBO = -10 \cdot log(S) \tag{2.10}$$

Realistic values for the span S are in the range $0.95 - 0.98$. The span in turn limits the domain of the linearizer shown in Equation 2.11.

$$0 \leq {r_m}^2 \leq P_{mm} \tag{2.11}$$

where, $r_m = |V_m|$. The maximum power of V_m, P_{mm}, is given by Equation 2.12.

$$P_{mm} = \frac{SP_{sat}}{K^2} \tag{2.12}$$

P_{mm} gives the power range for the input signal over which we attempt linearization.

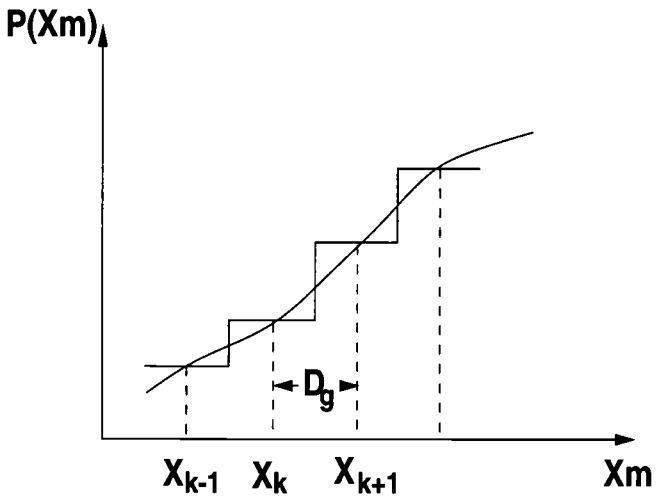

FIGURE 2.6: Quantization in the complex gain table

If the actual input signal has the dynamic range as $[0, \hat{r_m}]$. The first step is to scale the input signal into the linearizable range by multiplying the input signal by the scale factor $K_s = \frac{\sqrt{P_{mm}}}{\hat{r_m}}$. The equivalent gain equals to the product of K_s and K.

For a LUT with N table entries, the N value for X_m are equally spaced in the range of $[0, \sqrt{P_{mm}}]$, the step size is given by Equation 2.13.

$$D_g = \frac{\sqrt{P_{mm}}}{N} \tag{2.13}$$

The range and midpoint of each step, and the corresponding entries, are given for $i = 1, 2, ..., N-1$ as Equation 2.14.

$$X_{mi} = \{X_m : iD_g \leq X_m \leq (i+1)D_g\} = D_g \cdot (i + \frac{1}{2}) \tag{2.14}$$

That is, the table is optimized according to Equation. 2.6 for the midpoint of each cell, as shown in Fig. 2.6.

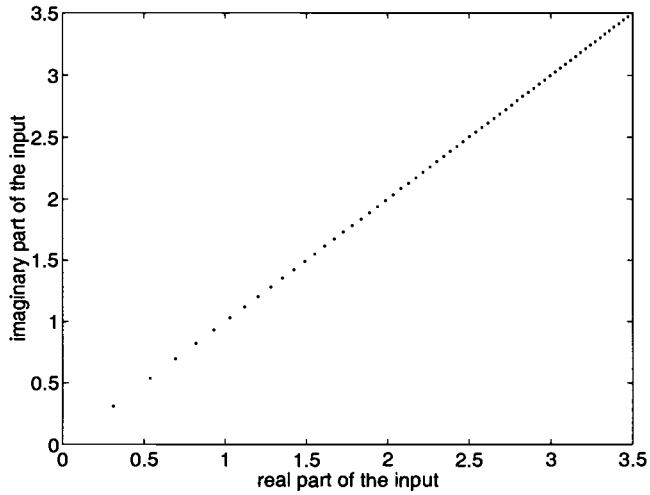

FIGURE 2.7: Constellation of the training signal

2.3. Implementation of the Direct Iterative Method

2.3.1. Simplified Simulation

In this section a simplified case is simulated. The input signal is a periodic signal with a period of N samples. In each period the sampled input signal is a monotonic function with the constant phase as $\frac{\pi}{4}$ and with the N (the size of LUT) different amplitudes equal to $x_{m,1}$, $x_{m,2}$,..., and $x_{m,N}$ contained in the LUT. The above input signal is called the training signal, and its constellation is shown in Fig. 2.7. The training signal passes the forward path through the HPA. It is then demonstrated and digitized. The difference between the recovered signal and the product of the training signal and the target ideal gain can be used to iteratively update predisorter coefficients according to Equation 2.9. The simulation is done in this way to test the direct iterative method and the effect of the secant method.

We applied this method to the class AB HPA, whose characteristic is shown in Fig. 2.3 and Fig. 2.4. The convergence of the power of the output error E_g is shown in Fig. 2.8. The curves demonstrate that convergence is slow at high output power

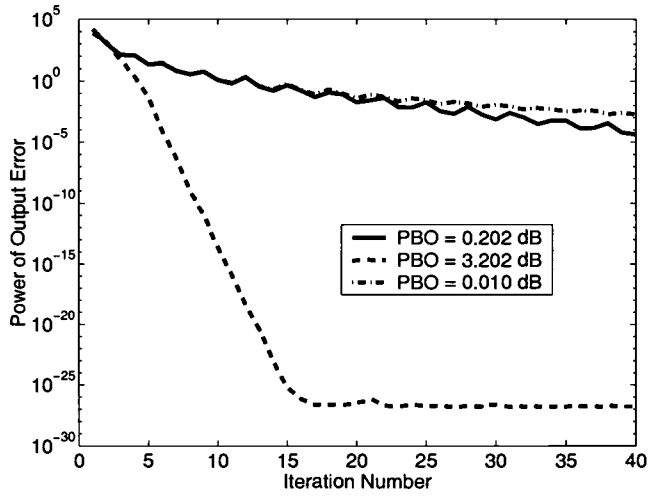

FIGURE 2.8: Convergence behavior of secant update algorithm

levels. With an additional 3dB backoff, the convergence is faster. Also it is obvious that a larger table size results in slower convergence. Here, three different cases: $PBO = 0.10dB, PBO = 0.22dB, PBO = 3.22dB$ were simulated. The larger span, the more time it takes for the predistorter to converge. To achieve the same level for error, the case with PBO = 0.10dB takes about 35 steps while the case with PBO = 3.22dB only takes 10 steps.

When the LUT is obtained the predistortion can be done by multiplication of the input signals with the correct complex coefficients of the LUT. Two kinds of baseband signals have been simulated with the results are presented in the following sections.

Case 1:OQPSK

The input signal is Offset QPSK (OQPSK), pulse-shaped by the raised-cosine-filter defined in the IS-95 standard. Using the LUT obtained by the training signal the effect of predistortion is shown in the following Figures, 2.9, 2.10 and 2.11.

The spectrum for the output signals with predistortion gives more than $30dBc$ improvement in adjacent channel interference(ACI) compared to the output signal without

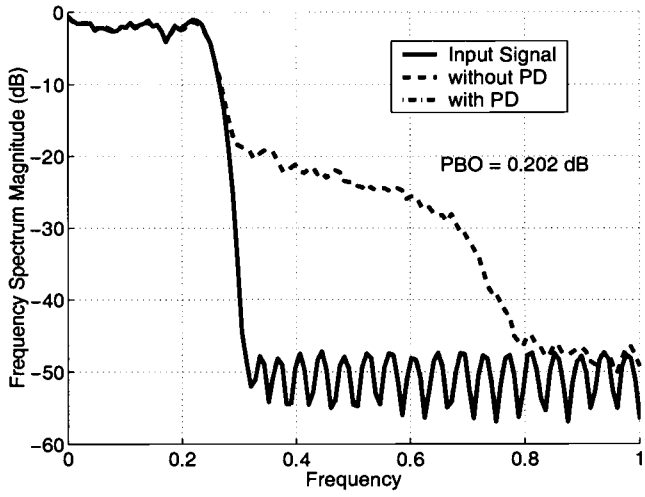

FIGURE 2.9: Spectrum for OQPSK input signal

predistortion. The input-output amplitude relationship and input-output phase relationship show that predistortion effectively corrects the AM/AM and AM/PM distortion.

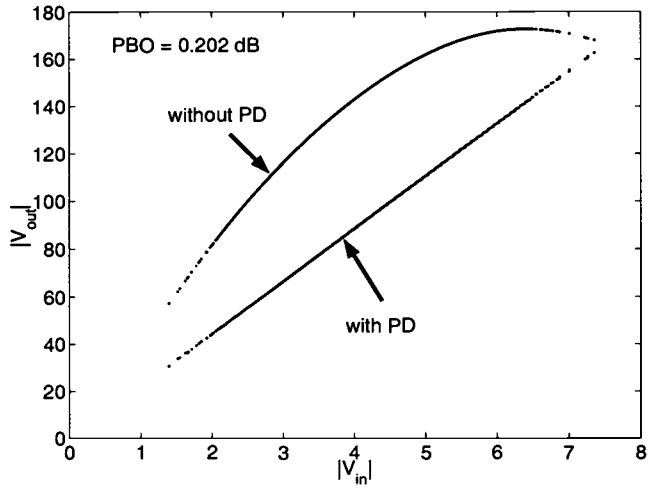

FIGURE 2.10: Input-Output Magnitude Relationship for OQPSK input signal

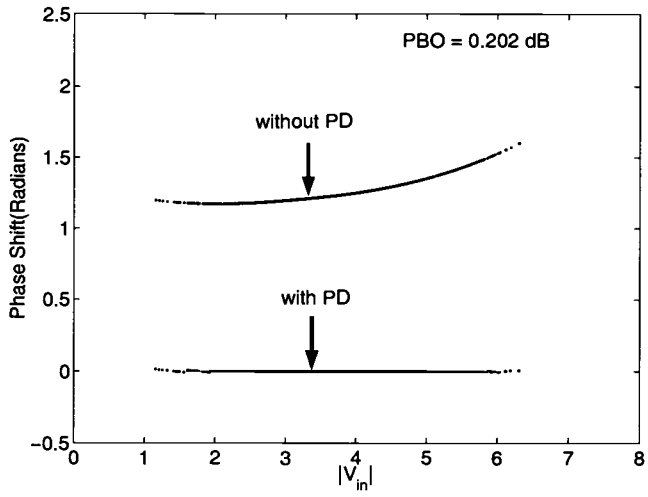

FIGURE 2.11: Phase shift for OQPSK input signal

Case 2:16QAM

The second signal is a 16QAM signal with a square root raised cosine pulse having 25% rolloff and Hamming windows to 7 symbols [1]. Using the LUT obtained by the training signal the effect of predistortion is shown in the following Figures, 2.12, 2.13 and 2.14.

The predistortion gives the system with 16QAM signal about $35dBc$ improvement in ACI over the non-predistorted signal. The input-output relationship in terms of amplitude and phase looks almost linear except for small input signal range where there is a little fluctuation. Increasing the size of the LUT can improve the performance in this small input signal range.

We will now use the constellations of the signals to explain how the predistortion works. Fig. 2.15 gives the normalized constellation of 16-QAM signal. Fig. 2.16 gives the received constellation if no predistortion is used. The nonlinear phase shift rotates the constellation, and the nonlinear compression attenuates the corner of the constellation. The constellation of the predistorted signal prior to amplification is shown in Fig. 2.17. The received constellation with predistortion and HPA is shown in Fig. 2.18. This is

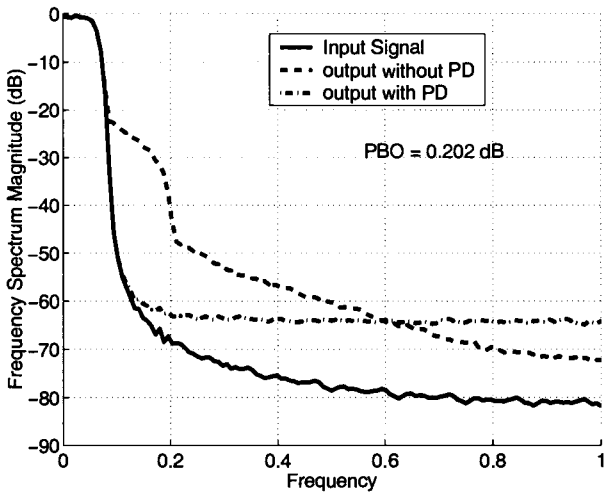

FIGURE 2.12: Spectrum for 16QAM input signal

almost identical to the input signal's constellation, which is consistent with what we observed in the spectrum.

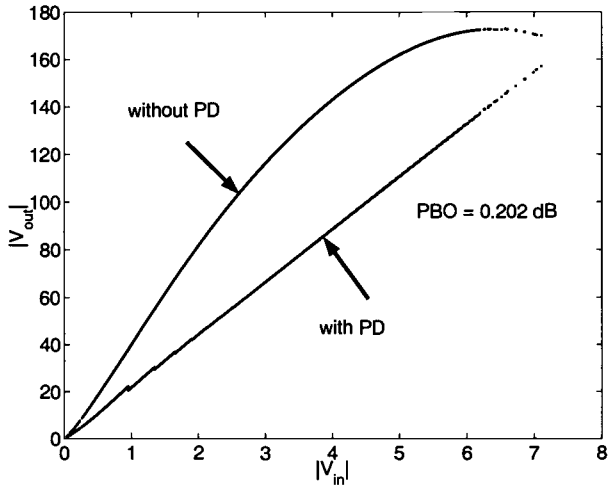

FIGURE 2.13: Input-Output Magnitude Relationship for 16QAM input signal

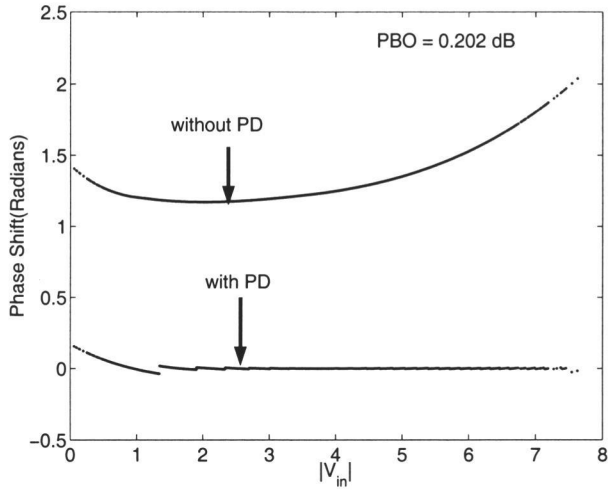

FIGURE 2.14: Phase shift for 16QAM input signal

2.3.2. Phase Shift and Magnitude Adjustment

In the last subsection, we demonstrated that the iteration method and the associated secant method were effective for our simple examples. However, when the above method put into practice the simplified situation doesn't exist. Since the predistortion

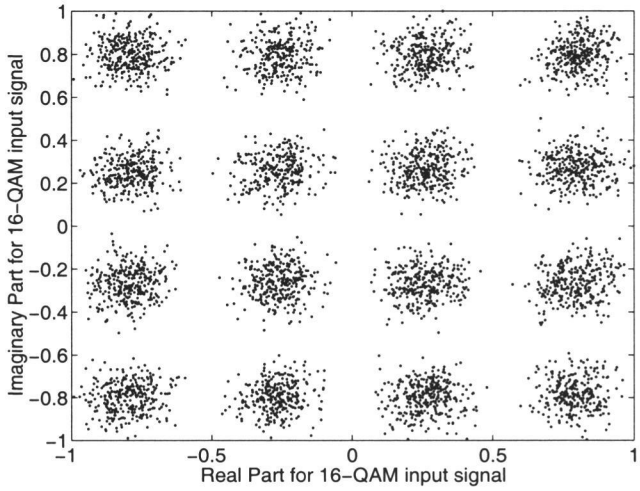

FIGURE 2.15: Normalized constellation for the input signal

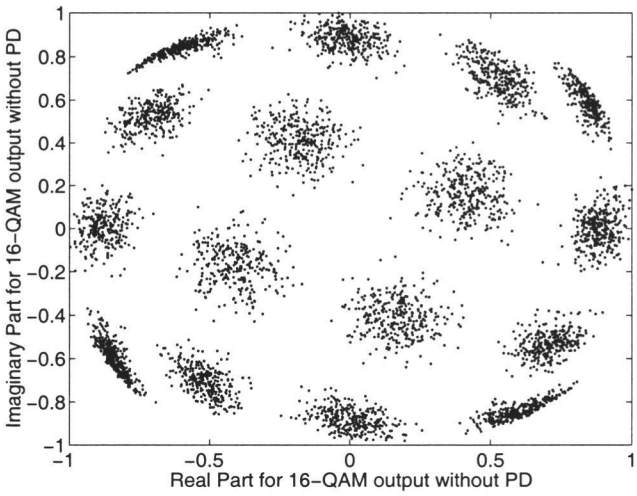

FIGURE 2.16: Normalized constellation for the output signal without predistortion

is running online no training signal can be fed to update P_m, instead, actual QAM or QPSK signals must be used. Under this situation, the amplitude of the signal can't be exactly equal to the center value of the table entry. Furthermore, a signal with the same amplitude may not have the same phase. The situation means that if the secant method is used to get predistortion coefficients, each point in the signal constellation should be

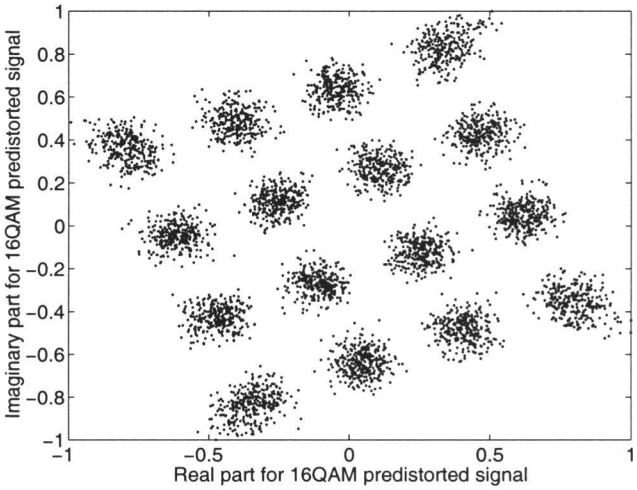

FIGURE 2.17: Normalized constellation for the predistorted signal

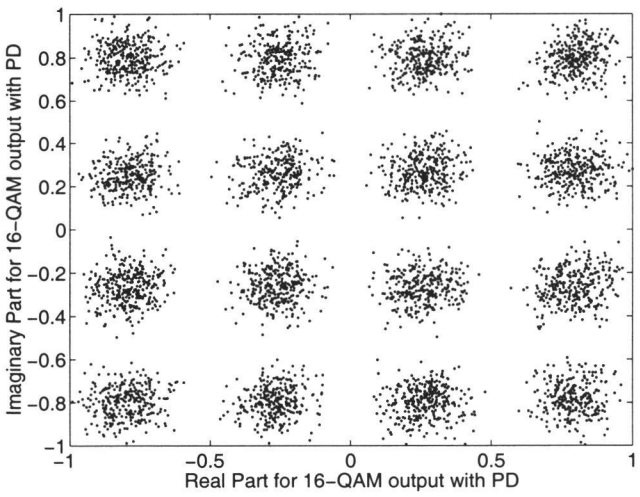

FIGURE 2.18: Normalized constellation for the output signal with predistortion

used as a index for the LUT. This results in a very large size table as in the mapping method. In this section the phase shift and magnitude adjustment are used with the secant method to keep the small size table, but without any loss in precision.

From Equation 2.9, to iteratively update P_i in every iterative step the values of $V_{m,i}$ must have the same amplitude as $X_{m,i}$ and the phase of $V_{m,i}$ should be kept constant. However this is impractical. However we are fortunate because both the nonlinearity of the predistortion and the HPA only depend on the signal's amplitude. Thus, the output V_a shifted by a phase $-\angle V_m$ is equal to the output that would have been obtained for an input which has the same amplitude as V_m, but, zero phase. Equation 2.15 gives the method for phase shift.

$$E_g(P) = V_a \cdot e^{-j\angle V_m} - KV_m \cdot e^{-j\angle V_m} \tag{2.15}$$

In addition to this phase shift, a magnitude adjustment is needed to map an equivalent error to an input signal value with exact table amplitude, X_i, when an input signal with amplitude other than X_i is transmitted. First, relationship between the error and the amplitude input signal should be found. Here, the error in each iteration step is investigated via simulation: the training signal is the input signal, the LUT has 64

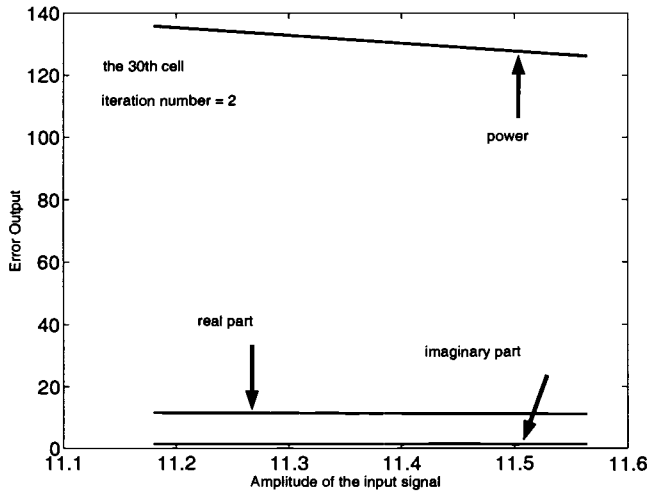

FIGURE 2.19: Error Output for 30th cell and the iteration step = 2

values and $PBO = 0.202dB$. From Fig. 2.8 only in 7 steps the iteration converge for the above case. In each step the predistortion is done according to the updated coefficients. There are 64 bins that covers the entire range for the input signal. For each bin, the input signal and its error are studied, and the results for the 30th bins at steps = 2,4,7 are shown in Fig. 2.19, 2.20 and 2.21.

From the results it is easy to see: 1) The power of the error decrease over time. 2) In each iteration, both the real and imaginary parts of the error are linear with respect to amplitude differences between the input signal and the center value of the bin, X_i. 3) When the algorithm converges the error for the midpoint of the bin is almost equal to zero.

The linear relationship between the real(imaginary) part of error and the amplitude difference from the center value in each bin gives us a way to calculate the complex error for an input signal with an amplitude not equal to X_i. Before we can update any predistortion coefficient, two input signals, $V_{m,i1}$ and $V_{m,i2}$, whose amplitude falls into the same bin, and their corresponding errors, E_{i1} and E_{i2}, must be collected. With the above information the real part and imaginary of the error corresponding to the midpoint of the bin can be linearly interpolated using the $r_{m,i1}$, $r_{m,i2}$ and the real and imaginary

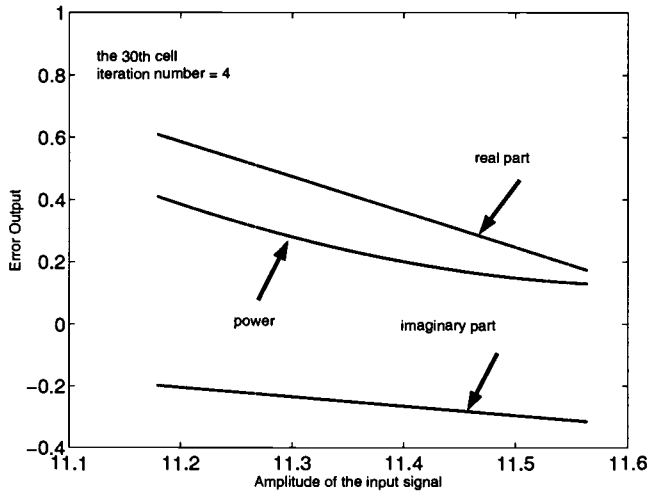

FIGURE 2.20: Error Output for 30th cell and the iteration step = 4

parts of E_{i1} and E_{i2}. The linear calculation to find the real and imaginary parts of equivalent error is illustrated in Fig. 2.22 and given by Equations 2.16 and 2.17.

$$Re(E_{g,i}) = S_R \cdot r_{i,1} + Re(Ei, 2) - S_R \cdot r_{i,2} \tag{2.16}$$

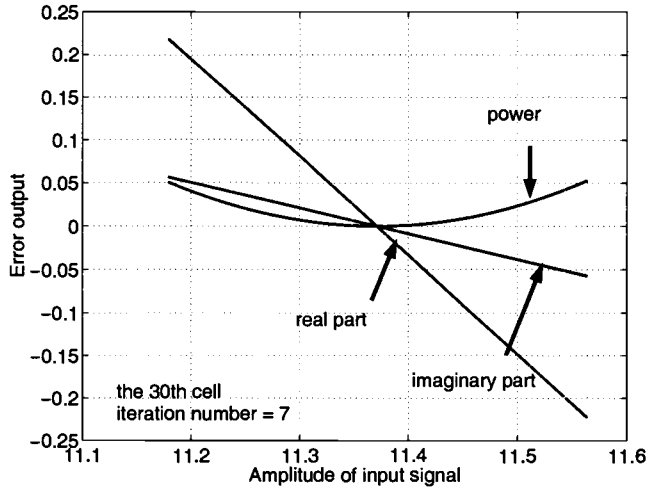

FIGURE 2.21: Error Output for 30th cell and the iteration step = 7

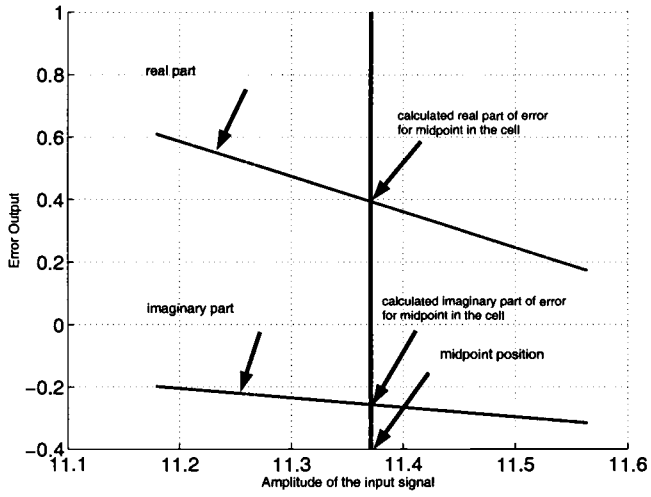

FIGURE 2.22: Magnitude adjustment method based on linear interpolation

$$Im(E_{g,i}) = S_I \cdot r_{i,1} + Im(Ei,2) - S_I \cdot r_{i,2} \tag{2.17}$$

where, $S_R = \frac{Re(E_{i,2}) - Re(E_{i,1})}{r_{i,2} - r_{i,1}}$, $S_I = \frac{Im(E_{i,2}) - Im(E_{i,1})}{r_{i,2} - r_{i,1}}$, $Re(\cdot)$ is the real part of the variable, and $Im(\cdot)$ is the imaginary part of the variable. r_i is the amplitude of the midpoint of the ith bin, $r_{i,1}$ and $r_{i,2}$ are the amplitudes of $|V_{i,1}|$ and $|V_{i,2}|$.

The structure correcting both the phase and magnitude is shown in Fig. 2.23.

Recall the process of linear interpolation to find the equivalent error for the midpoint in each bin. Thus, a second LUT is used to store the two different values of magnitude of signal and the corresponding complex value for error. After $V_{i,1}$, $E_{i,1}$, and $V_{i,2}$, $E_{i,2}$ are collected, the corresponding E_g can be estimated by linear interpolation for real part and imaginary part.

2.4. Simulation Results

Simulation results are shown in Fig. 2.24 for the OQPSK signal, and in Fig. 2.25 for the 16QAM signal. Comparing to these results obtained by using the training signal the results with OQPSK and 16QAM signals give nearly identical performance. As

discussed previously, the first step of the predistortion is to scale the input signals to the linearizable range. By scaling the signals to the linearizable range the out-of-band power will decrease because in this range HPA looks more linear than over the original range. Thus to be fair in Fig. 2.24 and Fig. 2.25 the spectra for output signals are shown by dotted lines, in these cases, the input signals are only scaled down to the linearizable range before going to the HPAs. The predistorter is seen to reduce ACI by anther 20dB for 16QAM signal. However even though the performance is impressive the computational load required to construct the LUT for the phase shift and magnitude adjustment, the sorting of the signal, the filling the table and the linear interpolation consume a lot of time and power. In the next chapter another method will be proposed, which doesn't need the phase shift and magnitude adjustment.

In the direct iterative method the predistorter gives AM/AM and AM/PM conversions. The nonlinear characteristics of the predistorter in terms of AM/AM and AM/PM are shown in Fig. 2.26 and Fig. 2.27. With a 32-size LUT and a 128-size LUT the characteristics of predistorters are almost the same, however, a 128-size LUT gives more coefficients in the range for small input signal, and thus makes the predistorter perform better in this region.

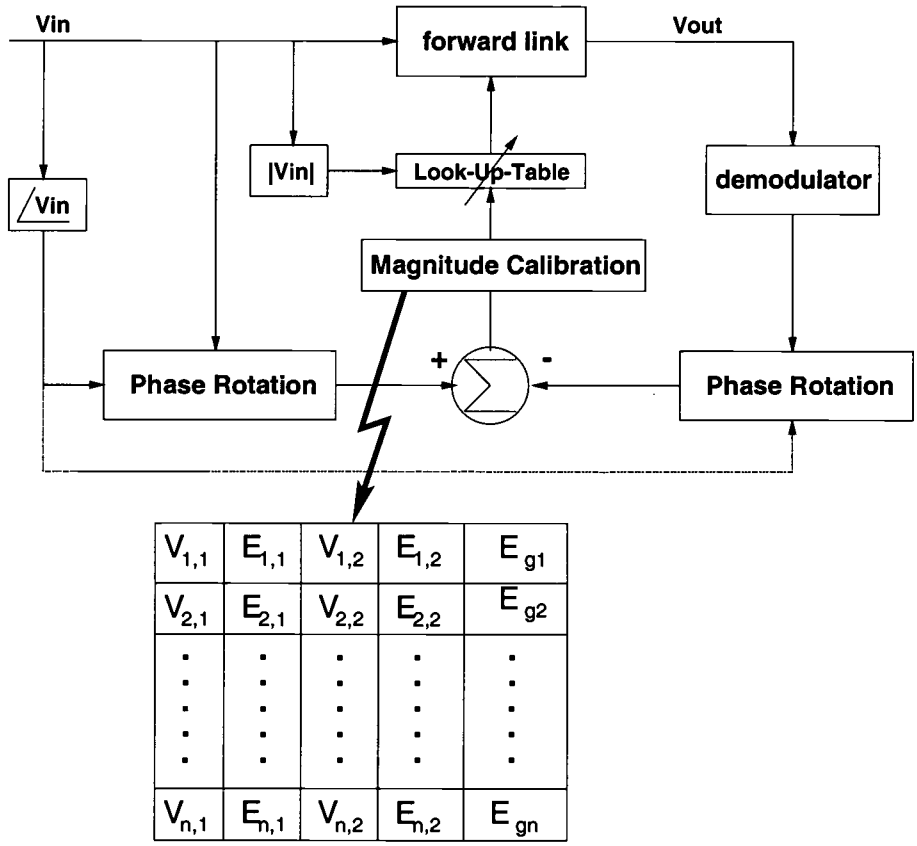

FIGURE 2.23: Structure of Direct Iterative Predistorter

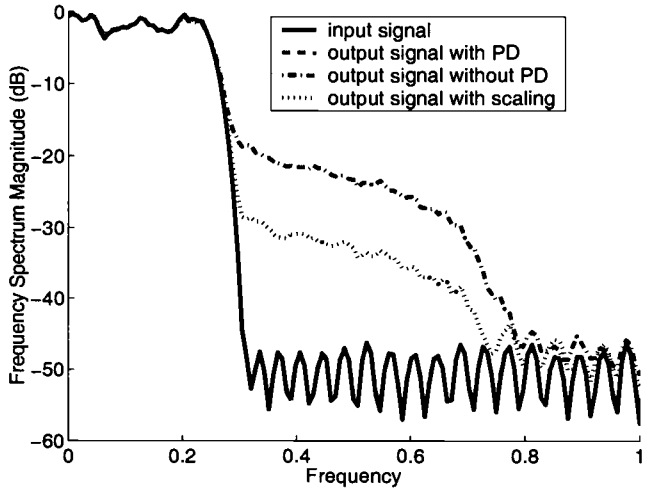

FIGURE 2.24: Spectrum of the output signal for OQPSK input signal

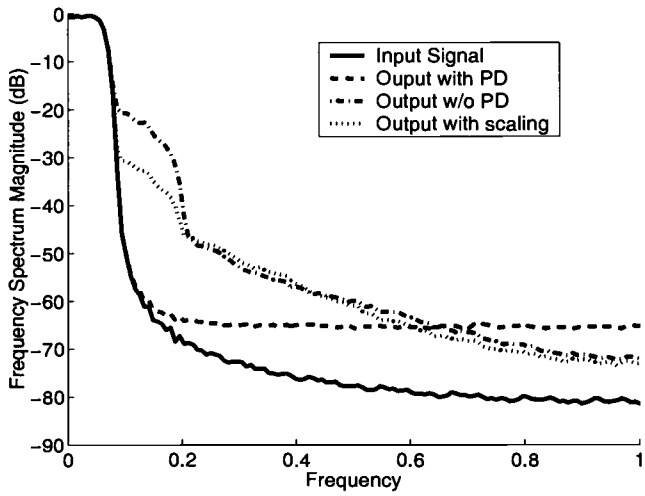

FIGURE 2.25: Spectrum of the output signal for 16-QAM signal

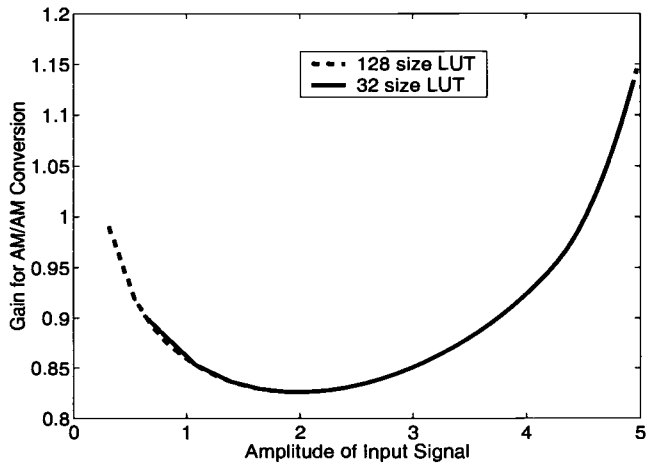

FIGURE 2.26: AM/AM conversion for the predistorter

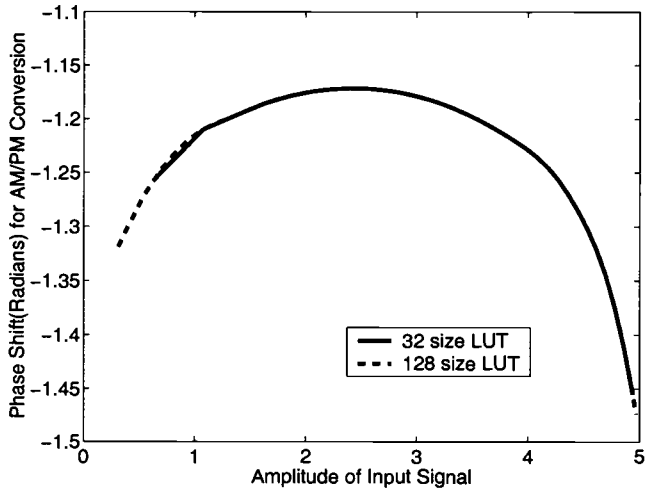

FIGURE 2.27: AM/PM conversion for the predistorter

3. MULTI-STAGE DIGITAL PREDISTORTION

3.1. Problem of the Direct Iterative Method

In the last chapter we presented the direct iterative method for the predistortion. A drawback is that for each step of updating the LUT the algorithm has to wait until the next time the HPA output data associated with the updated predistorter coefficient is used by LUT to see the effect of the previous update of the complex coefficient. This slows convergence, furthermore, we found that the convergence is sensitive to the starting points chosen for the iteration. When the secant method is used to update the complex coefficients, the phase shift and the magnitude adjustments must be done to find the equivalent error, which introduces more computational load and further slows the speed. More memory must be used to keep another table for the magnitude adjustment. The process of filling the tables consumes a lot of time and power. Also in the iteration any noise or measurement error unavoidably influences the convergence process. The problem is that this method uses feedback to directly solve the inverse problem, which makes the direct iterative method unstable in some cases.

From the Equation 2.6 it is obvious that if the characteristic of the HPA, which is expressed as F, can be estimated then a perfect solution for the predistortion of this model can be found. To avoid delays caused by feedback in the previous approaches, we first find an explicit expression for F. In [2] two polynomials are used, one is an odd order polynomial to describe the AM/AM conversion, the other is an even polynomial to describe the AM/PM conversion, both of which have the real coefficients. Once a model for the HPA's AM/AM and AM/PM is constructed, the next step is to find predistortion polynomials for the AM/AM and the AM/PM, which are inverse to the models for the HPA's AM/AM and AM/PM conversions. These inverse models are used to predistort the input signal's magnitude and phase. This two-step approach is called the multi-stage predistortion.

In our approach a multi-stage predistorter is also adopted because of its reported

advantage [2]. However, instead of two real polynomials we use one complex polynomial. The predistortion is then done by simultaneously predistorting the magnitude and phase.

First, we construct a Complex Gain Polynomial(CGP) model (feedforward model) for the HPA as opposed to the real polynomials used in [2]. Since this is an identification problem we do not pass our modeled signal through the HPA and thus eliminate the delay associated with the direct iterative method. Second, we construct the LUT based on the complex gain polynomial as in [2], which is done closed loop with respect to our model but open loop with respect to the actual HPA. Predistortion entails multiply by the input signal by the complex coefficient according to LUT as in the direct iterative method.

3.2. Theory and Implementation of the Multi-Stage Predistorter

A descriptive block diagram for our new multi-stage digital adaptive predistorter is shown in Fig. 3.1. The output V_a is decided by Equation 2.5 and the predistorter tries to achieve the function described by Equation 2.6. To determine $P(r_m)$ the first step in the multi-stage method is to find a CGP to describe F. From Fig. 3.1 it can be seen that the CGP is obtained by collecting a block of data (V_d, V_f) and using a curve-fitting method. Since there is a delay in the transmission path the data V_d must be delayed by the same amount to align with V_f.

After the CGP is obtained, a numerical iterative method is employed to find the complex numbers as the roots of Equation 2.6 for all table entries. Although iteration is still used to solve the inverse problem, the difference is that the iteration now use synthetic data instead of measured data, and the iteration loops through our model and not through the actual system.

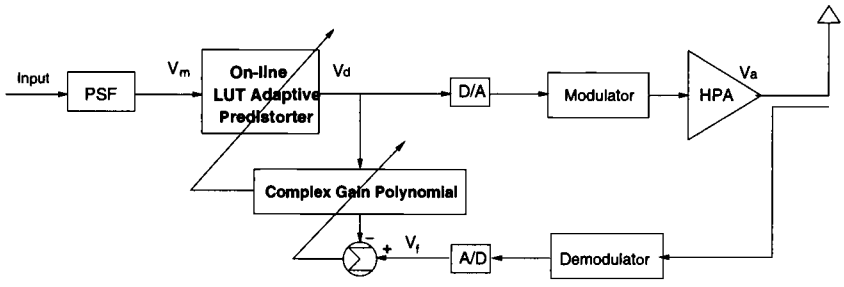

FIGURE 3.1: Structure of the multi-stage predistortion

3.2.1. Complex Gain Polynomial

As is well known a Taylor series is a valid representation for a memoryless nonlinear function. Generally, a Taylor model can be written as Equation. 3.1.

$$P(r_m) = \sum_{i=0}^{\infty} a_i \cdot r_m^i \tag{3.1}$$

Our CGP is basically a truncated Taylor's representation. What order truncation is good depends on the characteristic of the HPA and system requirement for linearity etc. Traditionally the nonlinearity of a HPA is described by IIP_3. However, experience and analysis indicate that using only the IIP_3 is not enough to describe the nonlinearity. Using a high order CGP obviously gives more precision, however, it also brings more computational load. As an example we collected 128 data samples of V_d and V_f and generated different orders of CGP to model the HPA. We used $P_i(r_m)$, i = 3,5,7,11,13. The calculated output according to CGP can be described by the Equation 3.2.

$$\hat{V}_a = V_d \cdot P_i(r_m) \tag{3.2}$$

The error in amplitude between the actual output V_a from the HPA and the calculated output $\hat{V}_a$ is calculated as shown in Equation. 3.3. The error in phase shift between the actual output V_a and calculated output $\hat{V}_a$ is calculated as shown in Equation. 3.4.

$$E_{am} = V_a - \hat{V}_a \tag{3.3}$$

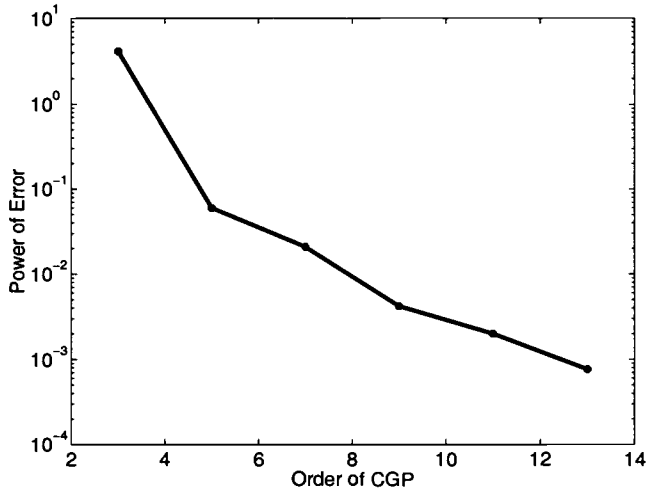

FIGURE 3.2: Effect of different orders of the CGP on error in amplitude

$$E_\theta = \angle V_a - \angle \hat{V}_a \tag{3.4}$$

Fig. 3.2 shows the power of the amplitude error between the actual output and calculated output for the different orders of CGP. Fig. 3.3 shows the power of the phase error between the actual output and calculated output for different orders of CGP. The above results are for the class AB HPA depicted in the 1st chapter .

From the above results the 5th order CGP is enough to describe the nonlinearity of the class AB HPA; the 3th order CGP gives too much error. Higher orders for the CGP give less error, however, they also create more computational load. For any specific HPA the proper order of the CGP depends on the specific nonlinear characteristic of the HPA. Fig. 3.4 gives the spectrum for output signal using predistortion with different orders of the CGP.

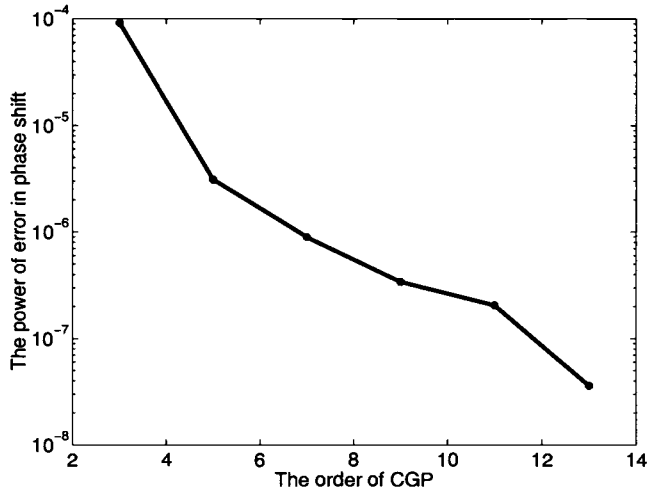

FIGURE 3.3: Effect of different orders of the CGP on error in phase

3.2.2. Construction of LUT

If the HPA has a saturation power of P_{sat}, and the LUT has N table entries, then the step size is given by Equation 3.5.

$$\triangle = \frac{\sqrt{P_{sat}}}{K} \tag{3.5}$$

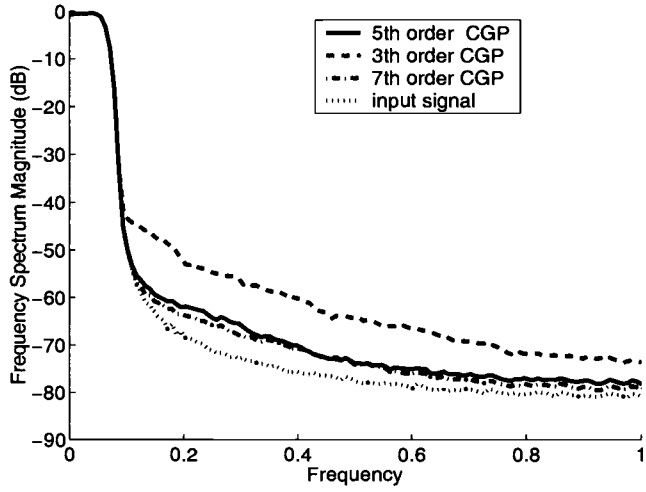

FIGURE 3.4: Spectrum of the output signal with different orders of the CGP

And the corresponding table entries are gives for $i = 0, 1, ..., N-1$ in Equation 3.6.

$$r_{m,i} = \triangle \cdot (i + \frac{1}{2}) \tag{3.6}$$

The initial value for $P(r_{m,i})$ must be set in the root-finding problem when an iterative method is employed. Since the nonlinearity of the HPA doesn't change dramatically from its neighbor, the initial value for the next table entry is set to be the root of the previous table entry. For the first table entry the initial value is set to be 1. After the root for the first table entry is calculated, the initial value for the root of the second table entry is set equal to the value of the first root, and the third to the second, etc.

When solving for the root for each table index, a threshold, ξ, must be specified to stop the root-finding process. When the value of Equation. 2.6 is less than ξ, the associated $P(r_m)$ is taken as the predistorter coefficient for the corresponding table entry. When the amplitude of the input signal approaches the saturation point of the HPA it takes more time to reach the threshold. As discussed earlier for any specific HPA there exists a limited range over which the output can be made linear to the input. In the direct iterative method the PBO is specified before the iteration process, which means that users must have some prior knowledge about the HPA to set the reasonable PBO. If the PBO is set unreasonably the predistorter gives worse performance. The drawback is that the algorithm can't set proper value for PBO automatically. In the multi-stage method the algorithm detects the linearizable range during the second stage. The algorithm tries to find roots for all table indexes. When it finds there is no root for a table index, say, $r_{m,imax}$, it means that the $[r_{m,1}, r_{m,imax-1}]$ is the linearizable range. Through this method the largest possible linearizable range can be obtained. Also there is no need to find the roots for table entries whose index is larger than $imax$. And no prior knowledge about the HPA is needed for the algorithm automatically to detect the linearizable range.

3.2.3. Method of Interpolation

After all the roots are specified the LUT is complete. The predistortion process is quite simple, and the complex input signal is multiplied by the complex predistortion coefficient which is specified by the amplitude of the input signal. In general the input signal doesn't have an amplitude that is exactly one of the table entries. Under this situation interpolation is used to estimate the corresponding complex coefficient from the table entries. In the last chapter the nearest neighbor interpolation was used, which is by far the simplest way, however, for points other than the midpoint in each cell the resulting coefficient can be far from the best value. In the direct iterative method no interpolation methods other than the nearest neighbor interpolation can be used because the linear relationship between the real(imaginary) part of the error and the amplitude difference is only valid when nearest neighbor interpolation is used. The above linear relationship must be kept true to make the magnitude adjustment valid. Fortunately in the multi-stage method a more precise interpolation method can be used since the construction of LUT is only dependent on the CGP. Here, we use three interpolation methods for comparison: nearest neighbor, linear and spline interpolation. Fig. 3.5 shows the spectrum for the output signal with predistortion for the different interpolation methods. For each case a LUT with 64 table entries and a 5th order CGP were used.

As expected the nearest neighbor interpolation gives the poorest results, and spline interpolation gives the best result. However, the difference between the linear interpolation and the spline interpolation is not significant. Furthermore, the linear interpolation is faster and requires far less computation than spline interpolation.

3.2.4. Size of Table

A LUT is used because the computation required to calculated the predistortion real time is prohibitive. Obviously, the larger the table size, the more precise the predistortion. However, it is a tradeoff between cost and performance. For a 16QAM signal and the class AB HPA, if linear interpolation is used, we found that a LUT with 16

table entries was needed. Fig. 3.6 gives the spectrum for a 8-LUT, 16-LUT, 32-LUT and 64-LUT.

3.3. Simulation Results

From the above discussion for the class AB HPA with the 5th order CGP, a LUT with 32 table entries and linear interpolation method is seen to be the best tradeoff between performance and cost. In Fig. 3.7 and 3.8 the spectra for the output signals are shown when OQPSK and 16QAM signal are used as input signals.

The 5th-order CGP used to represent the class AB HPA is given in Equation 3.7.

$$\begin{aligned} P(r_m) &= (0.0107 + 0.0047i) \cdot r_m^5 + (-0.2317 - 0.1316i) \cdot r_m^4 \\ &\quad + (1.9296 + 1.2995i) \cdot r_m^3 + (-8.3443 - 5.9910i) \cdot r_m^2 \\ &\quad + (16.2592 + 11.1304i) \cdot r_m + (4.6449 + 30.9432i) \end{aligned} \tag{3.7}$$

The LUT predistorts the input signal according to the 32 complex coefficients stored in the LUT. The predistorter provides nonlinear AM/AM and AM/PM conver-

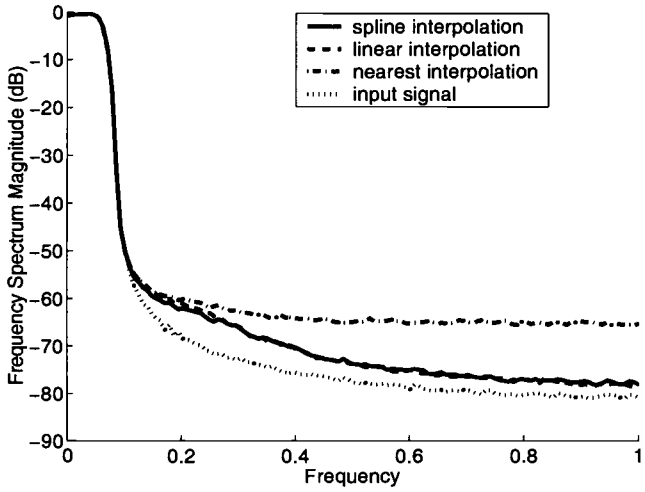

FIGURE 3.5: Spectrum of the output signal with different interpolation methods

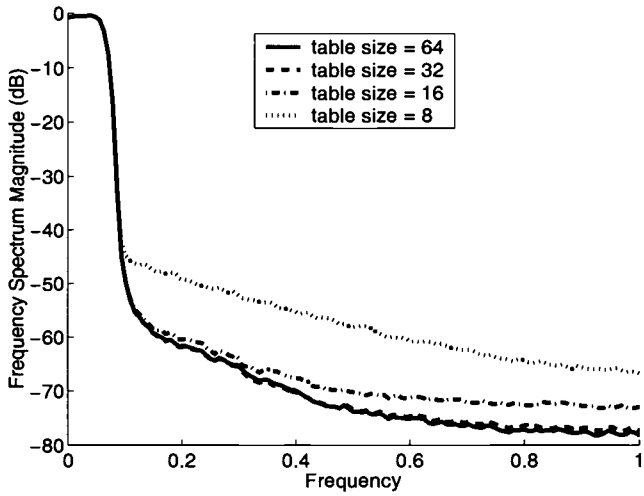

FIGURE 3.6: Spectrum of the output signal with different size LUTs

sions as shown in Fig. 3.9 and Fig. 3.10. Also, for 16QAM signal the input-output relationships in amplitude and phase are shown in Fig. 3.11 and Fig. 3.12.

Better improvement for predistortion can be achieved by using a higher order of CGP(to get a better representation for HPA's nonlinearity), however, the larger the LUT, the more complex the computation. The results in Fig. 3.13 show the simulation

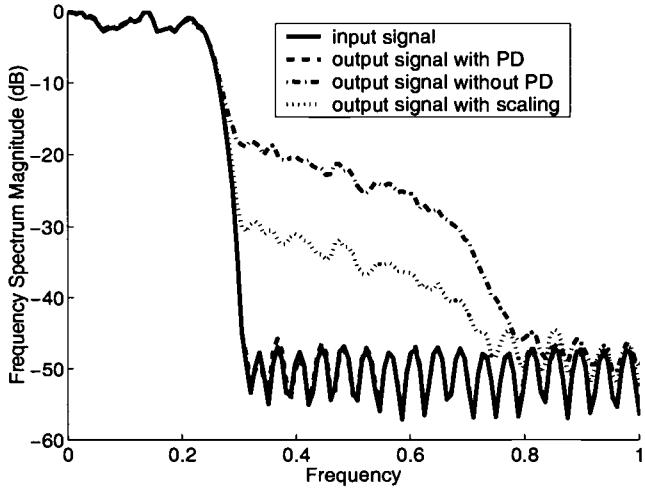

FIGURE 3.7: Spectrum of the output signal with OQPSK input signal

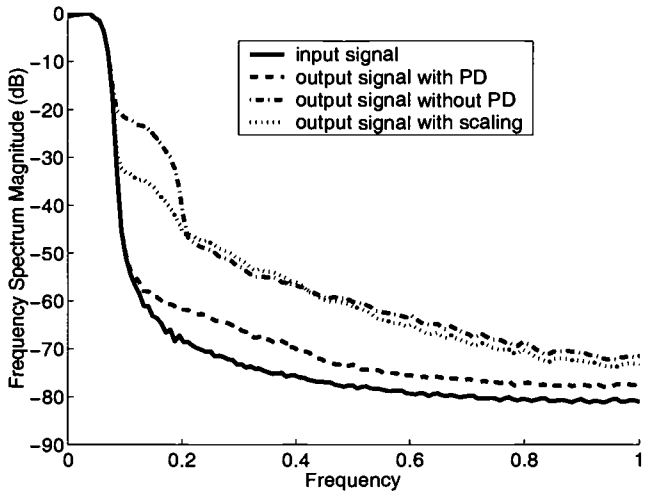

FIGURE 3.8: Spectrum of the output signal with 16-QAM input signal

results when a 9th-order CGP, a 128 size LUT and the linear interpolation are applied to the class AB HPA. The reduction of out-of-band power emission is almost 40dB.

In summary, the predistorter has nonlinear characteristics represented by AM/AM and AM/PM conversions shown in Fig. 3.9 and Fig. 3.10. Compared to the AM/AM and AM/PM conversions for the case with the 5th-order CGP and the 32-size table these

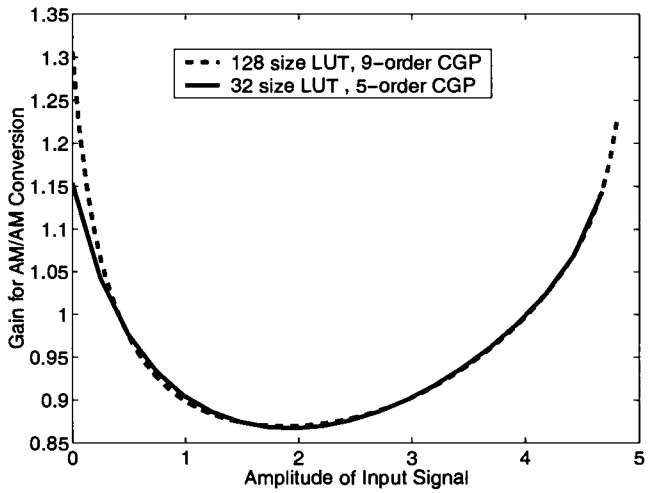

FIGURE 3.9: AM/AM conversion for the predistorter

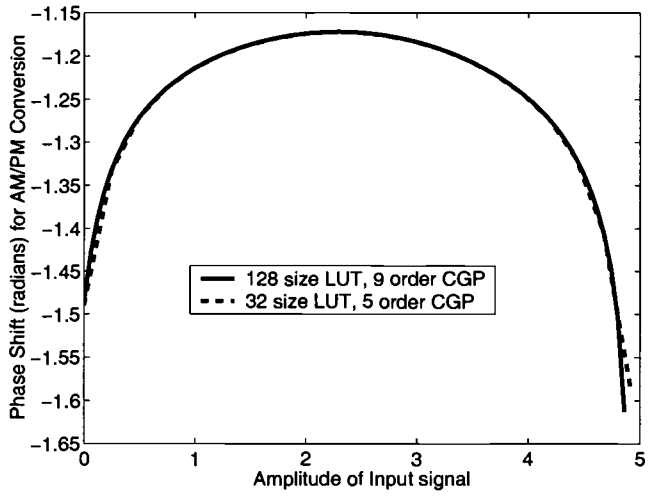

FIGURE 3.10: AM/PM conversion for the predistorter

conversions with the 9th-order CGP and the 128 size LUT are similar except in this case the predistorter operates over a larger range of the input signal.

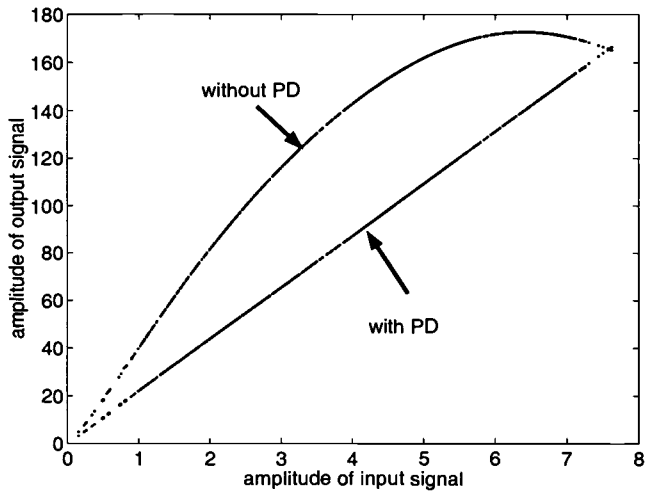

FIGURE 3.11: Input-output relationship on amplitude for 16QAM input signal

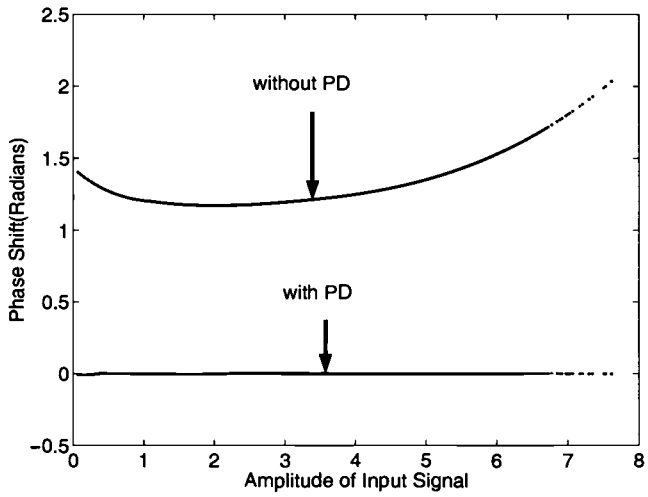

FIGURE 3.12: Phase shift for 16QAM input signal

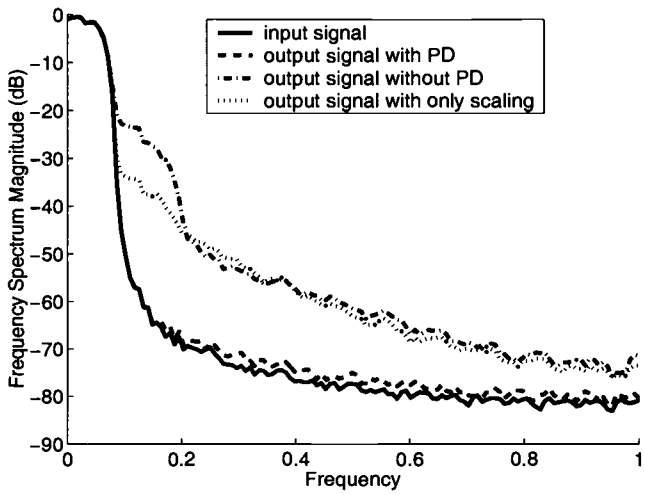

FIGURE 3.13: Spectrum of the output signal for 16-QAM signal

4. COMPLEMENTARY DIGITAL PREDISTORTION

4.1. Problem of the Multi-Stage Method

The multi-stage method doesn't require the feedback and online iteration. However, the problem of this method is its computational load. The explicit expression of function $P(r_m)$ can only be obtained through significant computation [2]. In the multi-stage method we presented in the last chapter, although we don't find the explicit expression of function $P(r_m)$, instead, a numerical method is employed to find the root of the objective function. The process of finding-root is burdensome. The computational load of this process limits the update speed. The above problem will limit the use of the multi-stage predistortion in applications where the power is limited and the algorithm must be kept as simple as possible. In this chapter a simpler method is introduced to find the predistorter's AM/AM and AM/PM conversions.

4.2. Introduction

Theoretically the predistorter distorts the signal in such a manner that its output signal is distorted in a precisely complementary manner to the distortion produced by the HPA. The output signal is therefore, ideally, an amplified, but undistorted replica of the input signal as shown in Fig. 4.1.

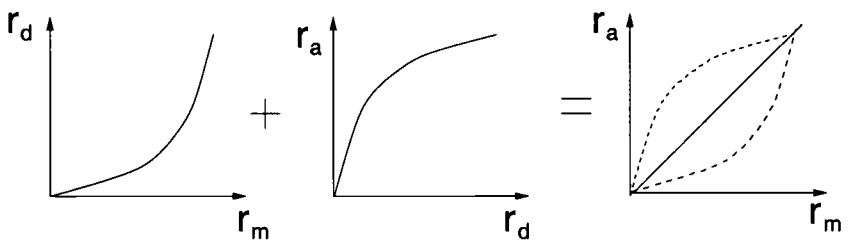

FIGURE 4.1: Theory of Complementary System

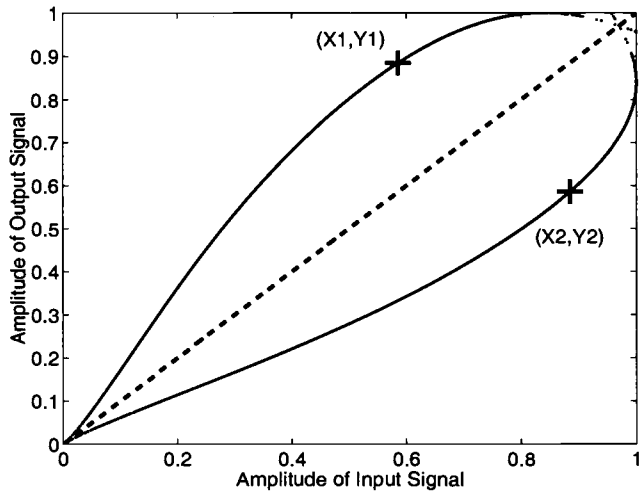

FIGURE 4.2: AM/AM conversions for nonlinear system A and B

However, the inverse characteristic is difficult to determine, especially for AM/AM conversion. In [2] the AM/AM and AM/PM conversions for the predistorter were obtained by polynomial fitting. As in [2] the forward model is identified first from the measured data, and then the inverse to the forward model is computed by iteration method, which is a major limitation due to computational load. We propose a simpler method to find predistortion polynomials based on the complementary nature exhibited by the nonlinearity of the predistorter and the HPA, which gives the method the name as the complementary method.

4.3. Theory and Implementation of the Complementary Predistorter

4.3.1. Symmetric Systems

Consider a nonlinear system A and a nonlinear system B, both of their nonlinear characteristics can be represented by AM/AM and AM/PM conversions shown in Fig. 4.2 and 4.3.

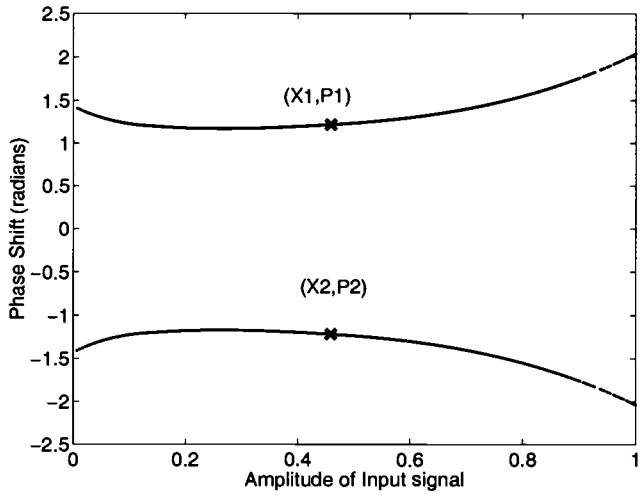

FIGURE 4.3: AM/PM conversions for nonlinear system A and B

The two systems are connected in cascade. For the overall cascaded system the input-output relationship will be perfect linear. This point can be easily seen since: (1): The AM/AM conversion curve for system A is symmetric to that for system B with respect to the line $Y = X$, (2): The AM/PM conversion curve for system A is symmetric to that for system B with respect to the line $Y = 0$.

If the HPA and the predistorter can be made to have the same properties, respectively , as system A and system B, to be detail, the predistorter looks like the system B, and the HPA looks like the system A, the overall system becomes an ideal linear amplification system. This is the idea for the complementary predistortion.

4.3.2. Complementary AM/AM and AM/PM Conversions

From the previous chapters and [2] we know that it is possible to find a CGP to represent the HPA. From the complex gain, it is easy to get AM/AM conversion(the magnitude of complex gain) and AM/PM conversion(the phase of complex gain).

The question is if it is possible from the HPA's AM/AM and AM/PM conversions to find the complementary characteristics for the AM/AM and the AM/PM conversions

for the predistorter? For AM/PM conversion it is trivial [2]. However for AM/AM it is not easy.

If the HPA can be normalized like system A, the answer is straightforward. We notice it is not difficult to normalize the HPA's input signal and the HPA's output signal as in the following Equations.

$$V_{in}^{n} = \frac{V_{in}}{|V_{in}^{max}|} \tag{4.1}$$

$$V_{out}^{n} = \frac{V_{out}}{|V_{out}^{max}|} \tag{4.2}$$

where, $|V_{in}^{max}|$ is the maximum amplitude of the input signal, and $|V_{out}^{max}|$ is the maximum amplitude of the output signal.

As discussed earlier, the normalized AM/AM conversion for the predistorter should be symmetric to the normalized AM/AM conversion for the HPA with respect to the line $Y = X$. Thus, we get the normalized AM/AM conversion curves for both the HPA and the predistorter as shown in Fig. 4.2. The normalized AM/AM curve for the predistorter gives the range for the input signal and the output signal for the predistorter as $(0, 1)$. However, since the output signal from the predistorter is the input signal for HPA, and the HPA operates in the range as $(0, |V_{in}^{max}|)$, the output signal from the predistorter obtained by mapping the normalized AM/AM curve should be re-scaled to fit the HPA's input signal range by multiplying the normalized output signal by $|V_{in}^{max}|$.

The AM/PM conversion for the predistortion can be obtained in a simple way similar to [2].

4.3.3. Implementation

By collecting a block of data (v_d, v_f), the AM/AM conversion for the HPA can be represented by a block of data, $(|v_d|, |v_f|)$, and AM/PM conversion for the HPA can be represented by a block of data, $(|v_d|, \theta)$, where θ is the phase shift between v_d and v_f.

The normalized AM/AM conversion for the HPA is represented by normalizing the amplitude of v_d and v_f, $(|v_d^n|, |v_f^n|)$. The normalized AM/AM conversion for the predistorter can be easily obtained as $(|v_f^n|, |v_d^n|)$. The following step is to scale the range of the normalized output signal to fit the range for the HPA's input signal by multiplying $|v_d^n|$ by a factor, $|V_{in}^{max}|$. Thus, the input-output relationship in terms of amplitude of signal for the predistorter is represented by a block of data, $(|v_f^n|, |v_d^n| \cdot |V_{in}^{max}|)$. We can use a real polynomial to represent $(|v_f^n|, |v_d^n| \cdot |V_{in}^{max}|)$ for AM/AM conversion for the predistorter, which is called the normalized amplitude polynomial(NAP). The amplitude correction can be done according to NAP.

The AM/PM curve for predistorter is easily represented by a block of data, $(|v_d|, -\theta)$, where θ is the phase shift between v_d and v_f. Also a real polynomial can be used to represent the AM/PM conversion for the predistorter, which is called the phase shift polynomial(PSP). The phase correction can be done according to PSP. The predistortion on phase follows the predistortion on amplitude. The block of the complementary predistorter is given in Fig. 4.4.

4.4. Simulation Results

Simulations were performed with the class AB HPA and a 16QAM signal. It was found that that a 7th-order NAP and 6th-order PSP give good results.

The 7th-order NAP used to represent the normalized AM/AM conversion for the predistorter is given in Equation 4.3.

$$\begin{aligned} r_d &= 51.5936 \cdot r_m^7 - 168.8278 \cdot r_m^6 + 218.7429 \cdot r_m^5 \\ &\quad -142.1784 \cdot r_m^4 + 48.8585 \cdot r_m^3 \\ &\quad -8.4926 \cdot r_m^2 + 1.2802 \cdot r_m - 0.004 \end{aligned} \tag{4.3}$$

The 5th-order PSP to represent the AM/PM conversion for the predistorter is given in Equation 4.4.

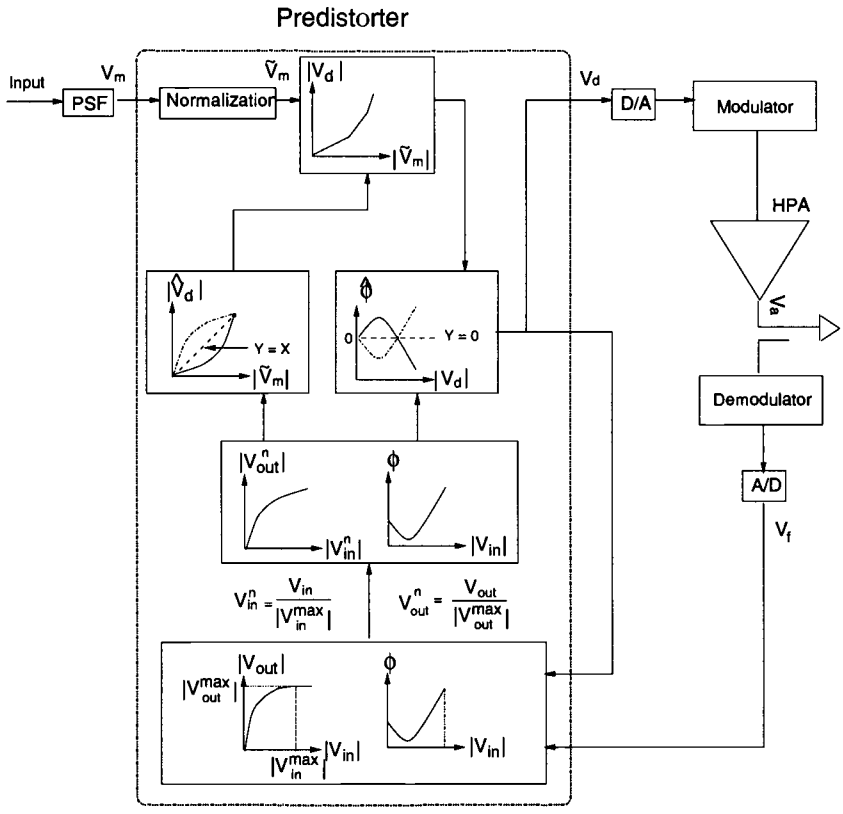

FIGURE 4.4: Structure of the Complementary Predistorter

$$\begin{aligned} \theta &= -0.0006 \cdot r_d^5 + 0.0077 \cdot r_d^4 - 0.0503 \cdot r_d^3 \\ &\quad +0.1841 \cdot r_d^2 - 0.3341 \cdot r_d + 1.4004 \end{aligned} \tag{4.4}$$

The improvement in ACI is shown in Fig. 4.5, and the nonlinearity of the AM/AM and AM/PM conversions are shown in Fig. 4.6 and Fig. 4.7. As for the direct iterative method and the multi-stage method the AM/AM conversion and AM/PM conversion can also be represented as in Fig. 4.8 and Fig. 4.9.

Ideally, all of these predistortion methods under the best situation introduce the equivalent pre-distortion for the input signal. Fig. 4.10 and Fig. 4.11 give the input-output relationships in terms of amplitude and phase shift for the predistorter based on the complementary method and the predistorter based on the multi-stage method(a

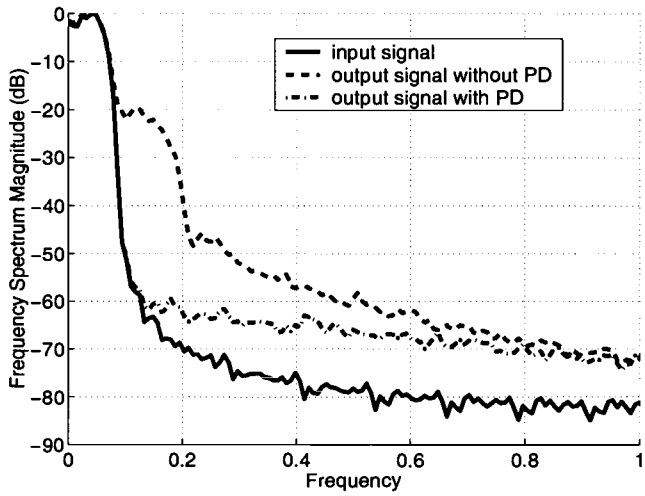

FIGURE 4.5: Spectrum for output signal

9th-order CGP and a 128-size LUT). As mentioned earlier, when a 9th-order CGP and a 128-size LUT are used in the multi-stage method for the AB class HPA and 16-QAM input signal the spectral regrowth is negligible. Compared to the AM/AM conversion and AM/PM conversion for the predistorter given by the multi-stage method in this best case the complementary method has almost the same AM/AM and AM/PM conversions

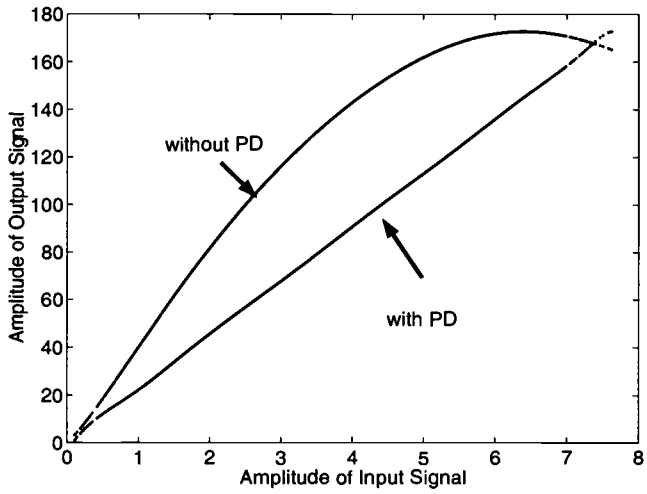

FIGURE 4.6: Input-output relationship on amplitude for 16QAM input signal

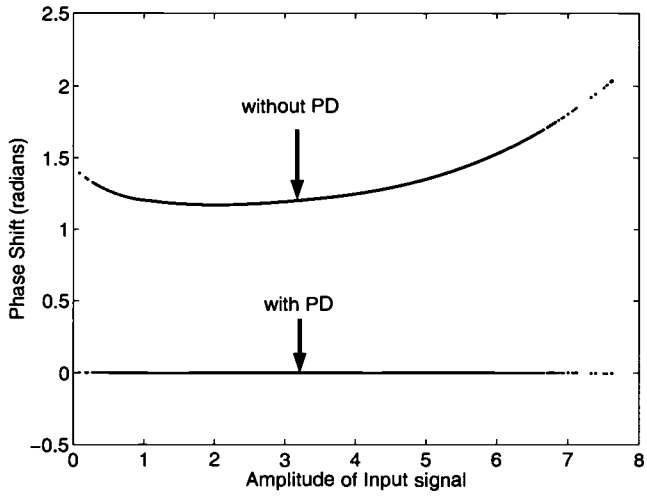

FIGURE 4.7: Phase shift for 16QAM input signal

for the predistorter. This illustrates that even the complementary method gives the significant reduction in computation the performance is fairly good.

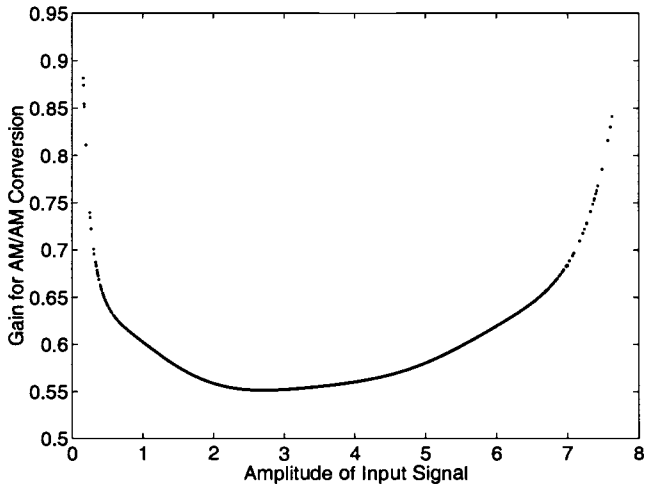

FIGURE 4.8: Gain for AM/AM conversion by NAP

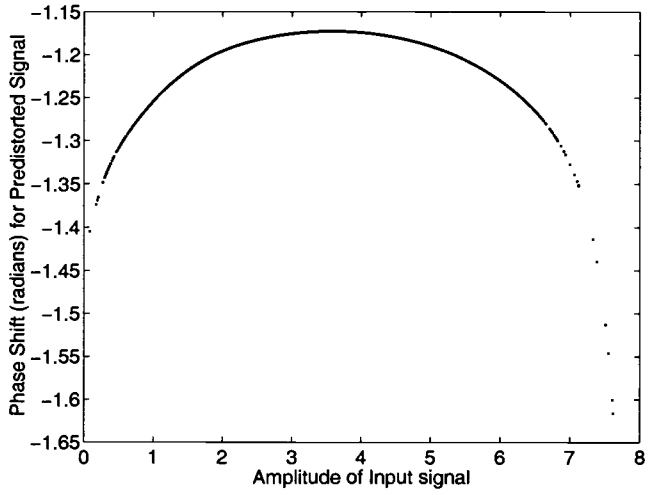

FIGURE 4.9: Phase Shift for AM/PM conversion by PSP

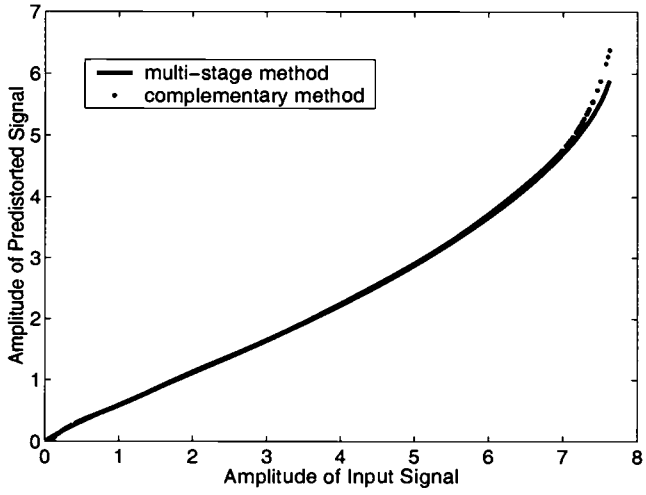

FIGURE 4.10: Input-output relationship on amplitude for the predistorter

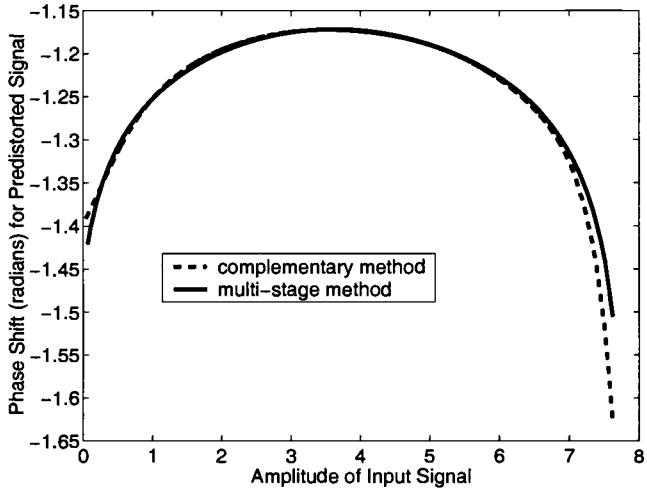

FIGURE 4.11: Phase shift introduced by the predistorter

5. COMPARISONS OF PREDISTORTION METHODS

In the previous chapters three methods are discussed: direct iterative method, multi-stage method and complementary method. In this chapter comparisons in terms of the complexity, stability and speed for the approaches are made. We also study the adaptability of the predistorter for the drift in HPA's characteristics and the sensitivity to the noise.

5.1. Computational Complexity

The main part computation for the direct iterative method is the iteration by the secant method. Since no derivative of E_g can be known the secant method is applied with the convergence speed between linear convergence and quadratic convergence. Each update involves 4 complex subtractions and 3 complex multiplications. Also the phase shift and the magnitude adjustment contain 1 complex multiplication, 10 real subtractions and 4 real multiplications in each update. After the complex coefficient for each table index is obtained, the predistortion is done by multiplication of the input signal by the complex coefficient.

The multi-stage method first finds the CGP for the HPA. It collects hundreds of sampled input data and output data and uses a curve-fitting method to find the polynomial complex coefficient in a Least Square Sense. The next step is to find the inverse to the HPA in the form of complex coefficients. Since the derivative of the objective function is available a Newton method can be used, which gives the quadratic convergence. Each update involves one complex multiplication, which is much simpler than the secant method in the direct iterative method. The implementation of the predistortion is similar to that in the direct iterative method. However, different interpolation methods can be used to achieve better tradeoff between cost and performance.

Compared to the direct iterative method and the multi-stage method the com-

putational complexity for the complementary method is the simplest one. It collects hundreds of sampled input data and output data to find the input-output relationship in terms of amplitude and phase shift for the predistorter. NAP and PSP are obtained by a curve-fitting method. The predistortion is implemented by operation of real polynomial evaluation. Simple operations, like normalization, scale, and swap, are used to get the block of data to represent AM/AM and AM/PM conversions for the predistortion from (v_d, v_f). Furthermore, there is no LUT necessary here to store the predistortion.

5.2. Computation Speed

From the analysis of computational complexity the complementary method is simplest, and the direct secant method is the most complex one. From simulation it is found that the secant method takes the longest time. For 16QAM signal with 5000 sampled data, the simulation running on SUN Ultra Sparc 10 workstations takes about 4-6 seconds to converge for the direct iterative method, and about 0.4-0.6 seconds for the multi-stage method, about 0.12-0.20 seconds for the complementary method.

5.3. Algorithm Stability

Among the three methods only the direct iterative method has the potential problem of instability because of feedback.

5.4. Sensitivity to Noise

Among the three methods, the direct iterative method is most sensitive to noise because the method attempts to directly solve the inverse problem. The convergence behaviour as a function of the iteration number is shown in Fig. 5.1.

We know that when there is no noise after the 7th step iteration step, the resulting

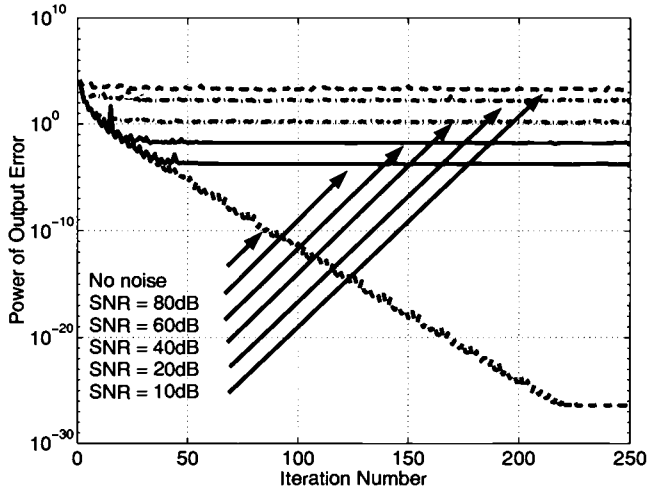

FIGURE 5.1: Convergence behavior with the AWGN

predistortion coefficient is close enough to make the output equal to the ideal linear value. The ratio of the power of the error to the power of the signal is about $40dB$. When noise is present, the corresponding spectra for the output signals are shown in Fig. 5.2.

When SNR for the input signal is less than $40dB$ the predistorter gives improvement in ACI, but not as much as when SNR for the input signal is greater than $40dB$.

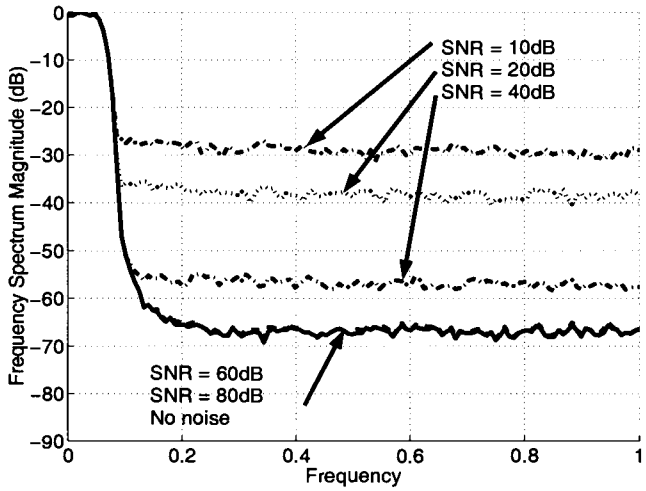

FIGURE 5.2: Spectrum of output signal with the AWGN

The multi-stage method is involved in the forward problem to characterize the HPA nonlinearity by the CGP in the first stage. The second stage has nothing to do with the measured data, so it has nothing to do with noise. Thus, if in the first stage there are ways to make the CGP insensitive to noise the predistorter will gain immunity to noise. The multi-stage method uses curve-fitting method to find the CGP based on a least square sense from the sampled data (v_d, v_f). Block processing can be employed to reduce the influence from noise. There are several ways to improve the algorithm when noise is present. Since the nonlinearity is memoryless the data pairs (v_d, v_f) can be re-ordered based on the amplitude of v_d. The re-ordered data pair can be treated as a very low frequency process and can be filtered to reject the out-of-band noise [2]. Another way is that the corrected output can be obtained by average in time domain from the re-ordered output.

5.5. Adaptability of Algorithms

For the direct iterative method the predistortion coefficients contained in LUT are updated online. When the iteration converges the complex coefficients will be kept unchanged according to the Equation 2.9. If the characteristic of the HPA changes the loop error becomes an indicator for the iteration and, the complex coefficient will be updated. The direct iterative method is inherent an adaptive algorithm since the feedback loop always keeps track of the drift in the HPA.

The multi-stage method achieves adaptability by updating the CGP. Periodic sampling of input data and output data, comparing the actual output and the estimated output, which calculated from the input data and the CGP which is stored in memory, and the resulting error between the actual output and estimated output is an indicator for the necessary of re-generation of the CGP. Other way to get the indicator for re-generation of the CGP can be measurement of the adjacent channel power emission. The complementary method also needs to judge whether it is necessary to re-sample the data to estimate the input-output relationship for the HPA to replace the obsolete one

like the multi-stage method. For some applications periodic update might be an efficient solution.

5.6. Conclusion

The direct iterative method is the most computationally complex. It does keep track of the drift in the characteristic for the HPA, however the problem of potential instability exists for the feedback structure. The multi-stage method is better than the direct iterative method when it comes to the computation load and computational complexity. The improvement on linearity is better than the direct iterative method, especially for small input signals, and no stability problems exist. The complementary method has the simplest structure, and the lightest computational load.

6. CONCLUSIONS AND FUTURE RESEARCH

6.1. Conclusions

The increasing demand for wireless communications, and the pressure on the limited radio spectrum available are forcing the development of more spectral efficient modulation schemes. The result is that linearization of the power amplifiers associated with linear modulation schemes become a critical design issue. In this thesis, several predistortion methods have been studied: direct iterative method, multi-stage method and complementary method. By analysis and simulation the direct iterative method requires the most computation and has potential instability problems. Also, noise severely hurts the system performance. The multi-stage methods don't use feedback, and instead use a two-stage structures. The multi-stage methods heavily rely on the precision of characterization of nonlinearity of the HPA represented by the CGP. The improvement on linearity of the system by multi-stage methods is significant. Furthermore, block processing of the sampled data can improve the performance when noise is present. Although the multi-stage methods decrease the computational load, and there is no stability problem, we still want to simplify the method to meet the requirement for low cost and fast speed. The complementary method is a simplified multi-stage method that uses symmetry in the normalized AM/AM for the predistorter and the HPA to get the normalized magnitude polynomial, and symmetry in the AM/PM for the predistorter and the HPA to get the phase shift polynomial as in [2]. No LUT is needed to store the predistorter coefficients. Instead two polynomials for magnitude and phase predistortion are generated. The algorithm is the fastest and simplest. Also like the multi-stage method block processing of sampled data can be used to increase SNR for AM/AM and AM/PM conversions for the HPA.

6.2. Future Research

There are many topics for future research in the linearization of power amplifiers. In this paper several possible digital adaptive predistortion methods were studied. All the methods are based on characterization of the HPA in the digital domain, and they all depend on demodulation. The demodulator increases the complexity in the predistortion which makes it might not the best solution for some applications. To overcome the above problem a method based on the adjacent channel power emission is being investigated. Theoretically the adjacent channel power emission is related with HPA's nonlinearity. By monitoring the out-of-band power the distortion introduced by HPA can be estimated [8]. Adaptation is accomplished by iterative adjustment of the predistortion coefficients to minimize the out-out-band power. The approach on how to best estimate the out-of-band power is still not solved. Also the numerical method used to iteratively adjust the predistortion coefficients based on out-of-band power need to be studied to achieve efficiency and stability.

In the direct iterative method and multi-stage method uniform spacing for LUT was used. There is no doubt that if more table values are used in the area where the HPA's characteristic varies sharply that the predistorter will give better performance [13]. The method to carry out the optimum table space, the tradeoff between uniform spacing and optimum spacing are the possible topics for future research.

For the direct iteration method the phase shift and the magnitude adjustment are each a large computation burden. To speed the algorithm new methods must be found to avoid the phase shift and the magnitude adjustment. For the multi-stage method and the complementary methods polynomials are used to represent the complex gain and the predistorter's AM/AM and AM/PM conversions. System performance is heavily dependent on the accuracy of these polynomials. Different HPAs require different orders of polynomials. To design a baseband digital predistorter that can be used with different systems a function to automatically choose proper order of polynomial must be developed.

For all methods there is time delay between the input data and the output data. Exact estimation of time delay is important. In analog predistortion system adjustment of the delay line will be necessary, which even makes the adaptability more difficult. For digital predistortion methods the time delay still exists. However DSP techniques can be used here to make the estimation of the time delay accurate and efficient as in [2].

BIBLIOGRAPHY

1. James K. Cavers. Amplifier linearization using a digital predistorter with fast adaptation and low memory requirement. *IEEE Trans. Vehicular Tech.*, 39(4):374–382, 1990.

2. J.T. Stonick, V.L. Stonick, and J.M.F. Moura. Mulit-stage adaptive predistortion of hpa staturation effects for digital television transmission. *ICASSP99*, 1999.

3. Theodore S. Rappaport. *Wireless Communications: Principles and Practice.* Prentice Hall, 1995.

4. Bernard Sklar. *Wireless Communications Fundamentals and Application.* Prentice Hall, 1988.

5. Behzad Razavi. *RF Microelectronics.* Prentice Hall, 1997.

6. Frank Zavosh, Mike Thomas, Chris Thron, and Tracy Hall. Digital predistortion techniques for rf power amplifiers with cdma applications. *IEEE Microwave Journal*, 42(10):22–50, 1999.

7. Anit Lohtia, Paul A. Goud, and Colin G. Englefield. Power amplifier linearization using cubic spline interpolation. *Proceedings of the 43th IEEE Vehicular Tech. Conf*, pages 679–479, 1993.

8. S.Stapleton and F. Costescu. An adaptive predistorter for a power amplifier based on adjacent channel emissions. *IEEE Trans. Vehicular Tech.*, 41(1):49–56, 1992.

9. Toshio Nojima, , and Tohru Konno. Cuber predistortion linearizer for relay equipment in 800 mhz abnd land mobile telephone system. *IEEE Trans. Vehicular Tech.*, 34(4):169–177, 1985.

10. A. Bateman, D.M. Haines, and R. J. Wilkinso. Linear transceiver architectures. *Proceedings of the 38th IEEE Vehicular Tech. Conf.*, pages 478–484, 1988.

11. Peter Kenington. Linearised rf amplifier and transmitter technique. *http://www.wsil.com.*

12. Michael Faulkner and Mats Johansson. Adaptive linearization using predistortion - experimental results. *IEEE Trans. Vehicular Tech.*, 43(2):323–332, 1994.

13. James K. Cavers. Optimum table spacing in predistorting amplifier linearizers. *IEEE Trans. Vehicular Tech.*, 48(5):1699–1705, 1999.

14. Andrew S. Wright and Willem G. Durtler. Experimental performance of an adaptive digital linearized power amplifier. *IEEE Trans. Vehicular Tech.*, 41(4), 1992.

15. Derek S. Hilborn, Shawn P. Stapleton, and James K. Caver. An adaptive direct conversion transmitter. *IEEE Trans. Vehicular Tech.*, 43(2):223–233, 1994.

16. James K. Cavers. Adaptation behavior of a feedforward amplifier linearier. *IEEE Trans. Vehicular Tech.*, 44(1):31–40, 1995.

17. James K. Cavers. The effect of quadurature modulator and demodulator errors on adaptive digital predistorters for amplifier linearization. *IEEE Trans. Vehicular Tech.*, 46(2):456–466, 1997.

18. James K. Cavers. New methods for adaptation of quadrature modulators and demodulators in amplifier linearization circuits. *IEEE Trans. Vehicular Tech.*, 46(3):707–716, 1997.

19. Qiang Wu, Heng Xiao, and Fu Li. Linear rf power amplifier design for cdma signals: A spectrum analysis approach. *Microwave Journal*, 35(10), 1999.

20. J.T. Stonick, V.L. Stonick, J.M.F. Moura, and R.S.Zborowski. Memoryless polynomial adaptive predistortion. *ICASSP95*, pages 981–984, 1995.

21. J.T. Stonick, V.L. Stonick, and J.M.F. Moura. Linearization of am/am distortion using the predistortion lms algorithm in a multi-stage adaptive predistorter. *IEEE Topical Workshop on Power Amplifiers for Wireless Communications*, 1999.

22. Jia Sun, Bin Li, and Michael Yan Wah Chia. Linearised and highly efficient cdma power amplifier. *IEE Electronics Letters*, 42(10):22–25, 1997.

23. Hyun Woo Kang, Yong Soo Cho, and Dae Hee Youn. On compensating nonlinear distortions of an ofdm system using an efficient adaptive predistort. *IEEE Trans. Vehicular Tech.*, 41(4), 1992.

24. Kevin Morris and Peter Kenington. Power amplifier linearisation using predistortion techniques. *IEE Colloquium on RF and Microwave Components for Communications System*, 1997.

25. Peter Kenington. Achieving high efficiency using linearised power amplifiers. *IEEE Topical Workshop on Power Amplifiers for Wireless Communications*, 1999.

26. Shigeru Ono, Noriaki Kondoh, and Yoshihito Shimazaki. Digital cellular system with linear modulation. *Proceedings of the 39th IEEE Vehicular Tech. Conf*, pages 44–49, 1989.

27. Couch II. Leon W. *Digital and Analog Communication Systems.* Prentice Hall, 1995.