Nonlinear System Identification and Analysis with Applications to Power Amplifier Modeling and Power Amplifier Predistortion

A Thesis
Presented to
The Academic Faculty

by

Raviv Raich

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

School of Electrical and Computer Engineering
Georgia Institute of Technology
March 2004

Nonlinear System Identification and Analysis with Applications to Power Amplifier Modeling and Power Amplifier Predistortion

Approved by:

Professor G. Tong Zhou, Advisor

Professor J. Stevenson Kenney

Professor Erik I. Verriest

Professor Xiaoming Huo
(School of Industrial and Systems Engineering)

Professor Douglas B. Williams

Date Approved: 31 March 2004

To my family, my wife Genevieve,

and my daughter Yael

ACKNOWLEDGEMENTS

I would like to express my gratitude to the people without whom this dissertation would not have been completed.

First, I would like to thank my advisor, Dr. G. Tong Zhou who continuously supported my work and ideas throughout my studies at Georgia Tech. Whenever I needed, Dr. Zhou always made herself available and put a great deal of time and effort into our joint work. I am thankful for her close guidance and efforts to prepare me for an academic career.

Next, I would like to thank my thesis defense committee: Dr. Erik Verriest, Dr. Doug Williams, Dr. J. Stevenson Kenney, and Dr. Xiaoming Huo for helping to improve my dissertation with their thoughtful advice and suggestions.

I would like to thank my group members: Krishna Nagarajan, Muhammad Ikram, Yongsub Kim, Gail Rosen, Lei Ding, Hua Qian, Chunpeng Xiao, Ning Chen, Vince Emanuele, Thao Tran, and Bob Baxley for useful technical discussions and fruitful collaborations.

I would also like to thank some of the students from the center for signal and image processing (CSIP): Amer Abufadel, Volkan Cevher, Kevin Chan, Kerkil Choi, Farshid Delgosha, Erdem Ertan, Majid Fozunbal, John Glotzbach, Paul Hong, Jay Jeon, Apurva Mody, Jonathan Morris, Robert Morris, Greg Slabaugh, Martin Tobias, Nik Vasiloglou, and Rajbabu Velmurugan for their willingness to help, insightful discussions, and mostly for their friendship, which made my graduate experience much more enjoyable.

To my family, I owe a debt of gratitude for their love and support along the path of my academic pursuits.

Last but not least, I would like to express my deepest gratitude to my wife Genevieve for her continuous support and encouragement throughout my entire graduate career from the time of application to the time of graduation. Without her help, kindness, patience, and mostly her love, I would not have been able to complete this work.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

Power amplifiers (PAs) are important components of communication systems and are inherently nonlinear. When a non-constant modulus signal goes through a nonlinear PA, spectral regrowth (broadening) appears in the PA output, which in turn causes adjacent channel interference (ACI). Stringent limits on the ACI are imposed by regulatory bodies, and thus the extent of the PA nonlinearity must be controlled. PA linearization is often necessary to suppress spectral regrowth, contain adjacent channel interference, and reduce bit error rate (BER). This dissertation addresses the following aspects of power amplifier research: modeling, linearization, and spectral regrowth analysis.

We explore the passband and baseband PA input/output relationships and show that they manifest differently when the PA exhibits long-term, short-term, or no memory effects. We provide an explanation for the various memory effects in the context of AM/AM and AM/PM responses. The so-called quasi-memoryless case is especially clarified. Four particular nonlinear models with memory are further investigated. We examine bandpass nonlinearities and their ramifications in the baseband and present a baseband formula which reveals that in the quasi-memoryless case, AM/AM and AM/PM conversions are sufficient to characterize the PA. We provide experimental results to support our analysis.

We employ the indirect learning structure to identify the Hammerstein predistorter for a PA modeled by a Wiener model. We compare the performance of two Hammerstein system identification algorithms; i.e., the Narendra-Gallman and least-squares/signular value decomposition algorithms. The benefits of using the orthogonal polynomials as opposed to the conventional polynomials are explored, in the context of digital baseband PA modeling and predistorter design. A closed-form expression for the orthogonal polynomial basis is derived. We demonstrate the improvement in numerical stability associated with the use of orthogonal polynomials for predistortion.

Spectral analysis can help to evaluate the suitability of a given PA for amplifying certain signals or to assist in predistortion linearization algorithm design. With the orthogonal polynomials that we derived, spectral analysis of the nonlinear PA becomes a straightforward task. The orthogonal polynomial PA parameters directly reveal the severity of spectral regrowth, as measured by the adjacent channel power ratio. We carry out nonlinear spectral analysis with digitally modulated signal as input. We show that by taking into account the cyclostationary nature of the processes, more accurate spectral analysis results can be obtained. We demonstrate an analytical approach for evaluating the power spectra of filtered QPSK and OQPSK signals after nonlinear amplification. A salient feature of our analysis is that we do not need to assume that the PA input is Gaussian. We employed the Leonov-Shiryaev formula to obtain closed-form output PSD expressions that apply to an arbitrary-order polynomial type of nonlinearity, which may include memory effects. These results can help us make important practical decisions such as what factors contribute to spectral regrowth, and how to control or correct them in order to keep the adjacent channel interference to within limits.

Many communications devices are nonlinear and have a peak power or peak amplitude constraint. In addition to possibly amplifying the useful signal, the nonlinearity also generates distortions. A measure that takes into account both these effects is the signal-to-noise-and-distortion ratio (SNDR). We focus on SNDR optimization within the family of amplitude limited memoryless nonlinearities. We show that under the peak amplitude constraint, the nonlinearity that maximizes the SNDR is a soft limiter with gain, and the specific gain (or equivalently, the threshold of the limiter) is found according to the peak signal to noise ratio (PSNR) and the distribution of the input amplitude. We obtain a link between the capacity of amplitude-limited nonlinear channels with Gaussian noise to the SNDR. These results are also of interest in applications such as predistortion linearization and peak-to-average power ratio reduction.

CHAPTER I

INTRODUCTION

Power amplifiers (PAs) are important components of communication systems and are inherently nonlinear. For example, the so-called Class AB PAs, which are moderately nonlinear, are typically employed in wireless basestations and handsets. When digitally modulated signals go through a nonlinear PA, spectral regrowth (broadening) appears in the PA output, which in turn causes adjacent channel interference (ACI). Stringent limits on the ACI are imposed by regulatory bodies, and thus the extent of the PA nonlinearity must be controlled. PA linearization is often necessary to suppress spectral regrowth, contain adjacent channel interference, and to reduce bit error rate (BER).

PA modeling has been a popular topic of research in the last few decades. Substantial efforts have been invested in the modeling of memoryless nonlinear PAs. The power series model, or the polynomial model, is widely used in the literature to model mild nonlinear effects in the PA (see e.g., [36], [64]). More recently, there are growing interests in modeling nonlinear PAs with memory effects. The cause of memory effects can be electrical or electro-thermal [118]. High-power amplifiers (HPAs) such as those used in wireless basestations exhibit memory effects; wideband signals also tend to induce memory effects in PAs. Recently in [65], a memory polynomial model is proposed to fit nonlinear PAs with memory. The more general Volterra series (which is polynomial in nature) has also been used to model nonlinear devices with memory [72].

Prediction of spectral regrowth for a prescribed level of PA nonlinearity can be very helpful for designing communication systems. Since more linear PAs are less efficient, practitioners may wish to use the PA in a configuration that allows for maximum PA efficiency while still stay below the spectral emission limits. Such optimization strategy is feasible if we have tools for spectral regrowth analysis. In [129], it is shown that after extracting the coefficients of a polynomial PA model, it is then possible to predict spectral regrowth of

digitally modulated signals using the concept of cumulants.

If a nonlinear PA is used to transmit non-constant modulus signals, PA linearization is often necessary. Among all linearization techniques, digital baseband predistortion is one of the most cost effective. A predistorter, which (ideally) has the inverse characteristic of the PA, is used to compensate for the nonlinearity in the PA. To linearize a memoryless nonlinear PA, one can pursue lookup table (LUT) based or model based approaches. The LUT approach is easy to implement but may take a relatively long time to converge. Moreover, the piece-wise linear curve has a zig-zag appearance which may introduce additional nonlinearities that degrade the performance [68]. As for model based approaches, the polynomial model is a common choice due to its simplicity and ease of implementation [36, Sec. 3.3], [48]. Volterra series [47] and certain special cases of the Volterra series, for example, the Hammerstein model [41] and the memory polynomial model [38], have been proposed for predistorter design that includes memory effects.

1.1 Background

Power amplifiers (PAs) have been the focus of research for the past few decades. More recently, due to technological advances in the digital signal processing (DSP) area, analogue solutions are increasingly replaced by DSP type solutions. The interaction between DSP and microwave engineers expended existing research areas and introduced additional areas of research. Here, we present a literature review focusing on four areas of research: characterization and modeling of PAs, spectral regrowth analysis of PAs, PA linearization, and communication aspects of PAs.

1.2 Characterization and Modeling of Power Amplifiers

Power amplifiers are often characterized by experimental means. Measurements obtained from a PA provide information on the PA characteristics such as the AM-AM and AM-PM transfer functions, 1 dB compression point, third-order intermodulation (IMD3), and fifth-order intermodulation (IMD5). Based on these measurements, we can obtain information regarding the nonlinearity and of the PA to select an appropriate PA to match a desired

application as well as to verify that the PA meets regulatory standards. For linear time-invariant (LTI) devices, frequency response is a sufficient representation: changing the input statistics will not change the frequency response of the system. However, a nonlinear device may exhibit a different frequency response when measured with a different input signal. For example, white noise excitation may yield a different frequency response from that of single tone excitation. For this reason, an input independent representation is desired. Behavioral or systematic modeling can offer a compact representation of the PA characteristics using a relatively small set of parameters. The PA can be modeled using a circuit diagram with values of the components as the model parameters. On the other hand, it can be modeled using a parameterized "black-box". In any case, the goal in mind is to offer a parsimonious and accurate model for the PA. Discussions on spectral regrowth analysis and PA linearization will be presented later.

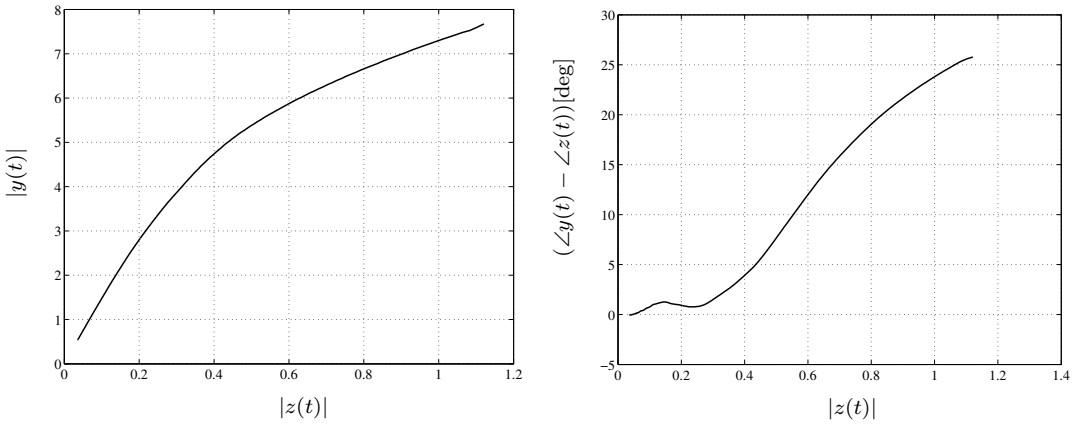
1.2.1 Memoryless Model

An instantaneous transfer function is sufficient to represent a memoryless PA. One way to model a memoryless transfer function is to use a polynomial model. The baseband polynomial model for a memoryless nonlinear PA is given by:

$$y(t) = \sum_{k=0}^{K} c_{2k+1}|z(t)|^{2k}z(t), \tag{1.1}$$

where $y(t)$ is the baseband PA output signal, $z(t)$ is the PA input signal, and c_{2k+1} is the $(2k+1)$th complex-valued polynomial coefficient. In [18], it is shown that the memoryless PA input/output baseband model given by (1.1) contains only odd-order nonlinear terms (e.g., terms associated with c_1, c_3). Bösch and Gatti pointed out in [24] that when the coefficients $\{c_{2k+1}\}$ are real-valued, the PA represented by (1.1) is strictly-memoryless. On the other hand, when the coefficients $\{c_{2k+1}\}$ are complex-valued, the PA is quasi-memoryless. However, if the PA is no longer quasi-memoryless, the baseband polynomial model in (1.1) is no longer valid and more elaborate model such as the Volterra series model should be considered.

The quasi-memoryless PA, (i.e., a PA that can be represented using (1.1)), can be characterized by two quantities. The first is the input amplitude to output amplitude



(a) AM/AM (b) AM/PM

Figure 1.1: AM/AM and AM/PM conversions measured from a real PA.

(AM/AM) conversion, i.e., the conversion from $|z(t)|$ to $|y(t)|$. The second is the input amplitude to input/output phase difference (AM/PM) conversion, i.e., the conversion from $|z(t)|$ to $\angle y(t) - \angle z(t)$. We examine both PA input $z(t)$ and PA output $y(t)$ to verify that AM/AM and AM/PM are sufficient to represent a PA given by (1.1). The amplitude of the PA output is $|y(t)| = |\sum_{k=0}^{K} c_{2k+1}|z(t)|^{2k+1}|$, which is a function only of the input amplitude $|z(t)|$. The phase of the PA output is $\angle y(t) = \angle z(t) + \angle \sum_{k=0}^{K} c_{2k+1}|z(t)|^{2k+1}$, which means that $\angle y(t) - \angle z(t)$ is a function of $|z(t)|$ only. Fig. 1.1 shows the AM/AM and AM/PM conversions taken from an actual PA. Looking at the AM/AM curve presented in Fig. 1.1(a), we observe that for a small input amplitude $|z(t)|$, the output amplitude of the PA $|y(t)|$ is a linear function of the input amplitude $|z(t)|$. Moreover, in Fig. 1.1(b), we notice that for a small input amplitude $|z(t)|$, the phase deviation $\angle y(t) - \angle z(t)$ is relatively constant as a function of the input amplitude $|z(t)|$. Many PAs share this property, which can be well characterized using the polynomial model in (1.1). The polynomial model offers a simple way to describe and analyze a memoryless nonlinear PA. However, the polynomial model has two major potential disadvantages: numerical instability and the lack of capability to model memory effects.

1.2.2 Volterra Series

In wideband applications or when high power amplifiers are employed, memory effects can no longer be ignored. In such cases, memoryless predistortion can no longer achieve the

required level of linearization performance. The cause of memory effects can be electrical or electro-thermal as suggested in [118]. The Volterra series model offers a good representation for the nonlinearity as well as the memory effects of the PA [72]. General background and theory regarding the Volterra series model can be found in [106]. In [119], the Volterra series is used to model a PA and obtain theoretical expressions for the power spectral density (PSD) at the output of the PA and for the intermodulation performance of the PA.

The baseband Volterra series model for the PA is given by:

$$y(t) = \sum_{k=0}^{K} \int \cdots \int H_{2k+1}(\tau_1, \ldots, \tau_{2k+1}) \prod_{i=1}^{k+1} z(t-\tau_i) \prod_{i=k+2}^{2k+1} z^*(t-\tau_i) d\tau_1 \cdots d\tau_{2k+1} \quad (1.2)$$

where $H_{2k+1}(\tau_1, \ldots, \tau_{2k+1})$ is the $(2k+1)$th-order kernel. When $H_{2k+1}(\tau_1, \ldots, \tau_{2k+1}) = c_{2k+1} \prod_{i=1}^{2k+1} \delta(\tau_i)$, this model simplifies to the memoryless polynomial model in (1.1). The advantage of this model is that it can approximate very accurately a very large class of nonlinear systems, and thus can capture a significant portion of the PA characteristics.

To illustrate the memory modeling capabilities of the Volterra series model, we consider the IMD3 evaluated using a two-tone test in the following example. Consider the Volterra series model given by

$$y(t) = z(t) - 0.2|z(t)|^2 z(t) + 0.1|z(t-1)|^2 z(t-1) \quad (1.3)$$

and the memoryless polynomial model

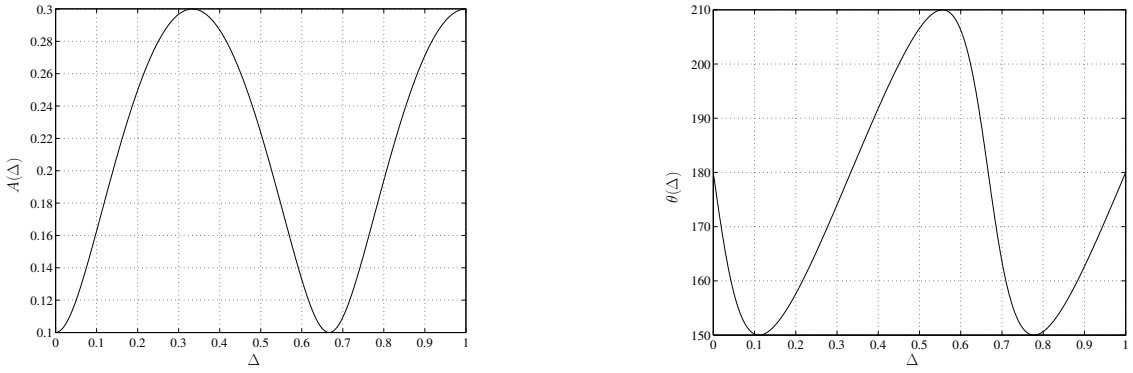
$$y(t) = z(t) - 0.2|z(t)|^2 z(t). \quad (1.4)$$

To examine the IMD3 behavior in each of the systems, we use a two-tone signal given by

$$z(t) = e^{-j2\pi \frac{\Delta}{2} t} + e^{j2\pi \frac{\Delta}{2} t} = \cos(\pi \Delta t), \quad (1.5)$$

where Δ is the frequency spacing of the two tone signal. Note that since we deal with a baseband model, the frequencies of the two-tone signal in the passband are $f_c - \frac{\Delta}{2}$ and $f_c + \frac{\Delta}{2}$, where f_c is the carrier frequency. The resulting output for the model (1.3) is given by

$$\begin{aligned} y(t) &= \cos(\pi\Delta t) - 0.2\big(\cos(\pi\Delta t) + 2\cos(2\pi\Delta t) + \cos(3\pi\Delta t)\big) \\ &\quad +0.1\big(\cos(\pi\Delta(t-1)) + 2\cos(2\pi\Delta(t-1)) + \cos(3\pi\Delta(t-1))\big), \end{aligned} \quad (1.6)$$



(a) IMD3 amplitude.

(b) IMD3 phase.

Figure 1.2: IMD3 results from a two-tone test of the Volterra series system in (1.3).

and the resulting output for the model (1.4) is given by

$$y(t) = \cos(\pi\Delta t) - 0.2\big(\cos(\pi\Delta t) + 2\cos(2\pi\Delta t) + \cos(3\pi\Delta t)\big). \tag{1.7}$$

We shall examine the third-order intermodulation (IMD3), i.e., the quantity $A(\Delta)e^{j\theta(\Delta)}$ in the component of the form $A(\Delta)\cos(3\pi\Delta t + \theta(\Delta))$ in $y(t)$. Since in (1.6), the components associated with frequency 3Δ are $-0.2\cos(3\pi\Delta t) + 0.1\cos(3\pi\Delta(t-1))$, IMD3 is given by

$$\begin{aligned} IMD3(\Delta) &= A(\Delta)e^{j\theta(\Delta)} = -0.2 + 0.1\cos(3\pi\Delta) - j0.1\sin(3\pi\Delta) && (1.8)\\ A(\Delta) &= 0.1\sqrt{(2-\cos(3\pi\Delta))^2 + \sin^2(3\pi\Delta)} = 0.1\sqrt{5 - 4\cos(3\pi\Delta)} && (1.9)\\ \theta(\Delta) &= \arctan\left(\frac{-0.1\sin(2\pi 3\frac{\Delta}{2})}{-0.2 + 0.1\cos(2\pi 3\frac{\Delta}{2})}\right). && (1.10) \end{aligned}$$

Note that $A(\Delta)$ and $\theta(\Delta)$ are the amplitude and phase of the IMD3. In Fig. 1.2, we present the amplitude and the phase of the IMD3 for the Volterra series system as given by (1.9)-(1.10). We observe that both amplitude and phase depend on the frequency spacing, which is an indication for the existence of memory effects. On the other hand, if we examine IMD3 at the output (1.7) of the memoryless system given by (1.4), we obtain

$$\begin{aligned} IMD3 &= A(\Delta)e^{j\theta(\Delta)} = -0.2 && (1.11)\\ A(\Delta) &= 0.2 && (1.12)\\ \theta(\Delta) &= -\pi. && (1.13) \end{aligned}$$

In (1.11)–(1.13), both amplitude and phase of the IMD3 are constant and therefore the IMD3 is constant w.r.t. the two-tone frequency spacing Δ. These observations illustrate

the capabilities of the Volterra series model to capture memory effects as they are observed by the IMD3 test. Similar arguments can be made for the IMD5 etc. tests.

The main disadvantage of the Volterra model is its complexity. While the $(2K+1)$th-order polynomial model is characterized by up to $2K+1$ parameters, the Volterra series is characterized by $(2K+1)$ kernels where the highest order kernel is a $(2K+1)$ dimensional function. Using the full Volterra series for real-time applications is often impractical.

1.2.3 Wiener Model

To avoid the complexity issue associated with the Volterra series representation, other models have been offered. One of these models is the Wiener model. The Wiener system is a linear time-invariant (LTI) system followed by a memoryless nonlinearity (see Fig. 1.3). The output $y(t)$ of a Wiener system is given by:

$$y(t) = f(s(t)), \qquad s(t) = \int h(\tau) z(t-\tau) d\tau, \tag{1.14}$$

where $f(\cdot)$ is the input/output transfer function of the memoryless nonlinearity and $h(\tau)$ is the impulse response of the LTI portion of the Wiener system. Using the polynomial model $f(s(t)) = \sum_{k=0}^{K} c_{2k+1}|s(t)|^{2k}s(t)$, we can offer a simple baseband representation for the memoryless nonlinearity. This model is a special case of the Volterra series model with $H_{2k+1}(\tau_1, \ldots, \tau_{2k+1}) = c_{2k+1} \prod_{i=1}^{k+1} h(\tau_i) \prod_{i=k+2}^{2k+1} h^*(\tau_i)$. In [33], Clark *et.al.* used a Wiener model to capture the nonlinear memory effects in the PA associated with wideband signals.

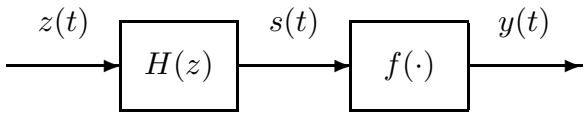


Figure 1.3: Wiener Model.

1.2.4 Hammerstein Model

The Hammerstein model is another nonlinear model with memory, which offers similar complexity to that of the Wiener model. The Hammerstein system is a memoryless nonlinearity followed by an LTI system (see Fig. 1.4). The output $y(t)$ of a Hammerstein system is given by

$$\begin{aligned} y(t) &= \int h(\tau)s(t-\tau)d\tau, \\ s(t) &= f(z(t)), \end{aligned} \tag{1.15}$$

where $f(\cdot)$ is the input/output transfer function of the memoryless nonlinearity and $h(\tau)$ is the impulse response of the LTI portion of the Hammerstein system. Using the polynomial model $f(z(t)) = \sum_{k=0}^{K} c_{2k+1}|z(t)|^{2k}z(t)$, we can offer a simple baseband representation for the memoryless nonlinearity. This model is a special case of the Volterra series model with $H_{2k+1}(\tau_1, \ldots, \tau_{2k+1}) = c_{2k+1}h(\tau_1)\prod_{i=2}^{2k+1}\delta(\tau_i - \tau_1)$. The output of a Hammerstein PA is linear w.r.t. the LTI portion parameters as opposed to the Wiener model [82]. Therefore, Hammerstein system identification can be relatively simple as compared to the Wiener system.

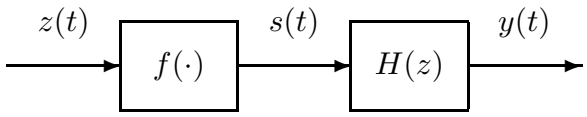


Figure 1.4: Hammerstein Model.

1.2.5 Wiener-Hammerstein Model

The Wiener-Hammerstein system is a concatenation of the Wiener system and the Hammerstein system. Since the nonlinear portion of the Wiener system is followed by the nonlinear portion of the Hammerstein system, both portions can be combined into a single block. Therefore, the Wiener-Hammerstein system is an LTI system followed by a memoryless nonlinearity followed by a second LTI system (see Fig. 1.5, which has also been referred to

as the LTI-MNL-LTI system). The PA output $y(t)$ is given by

$$\begin{aligned} y(t) &= \int h_2(\tau)s(t-\tau)d\tau, \\ s(t) &= f(w(t)), \\ w(t) &= \int h_1(\tau)z(t-\tau)d\tau, \end{aligned} \tag{1.16}$$

where $h_1(\tau)$ is the impulse response of the LTI portion of the Wiener system, $h_2(\tau)$ is the impulse response of the LTI portion of the Hammerstein system, and $f(\cdot)$ is the memoryless nonlinearity (MNL). Using the polynomial model $f(w(t)) = \sum_{k=0}^{K} c_{2k+1}|w(t)|^{2k}w(t)$, we can offer a simple baseband representation for the memoryless nonlinearity. The advantage of this model over the Hammerstein and the Wiener models is that it is more general and therefore model the PA more accurately. The added cost in complexity is reasonable. This model arises in the context of satellite communications [17]. The LTI systems represent the uplink and the downlink channels and the NL system represents the PA at the satellite.

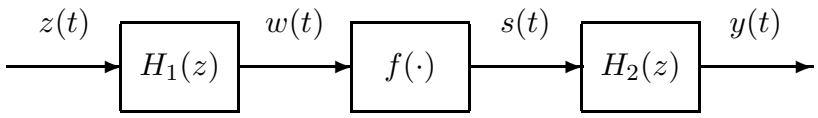


Figure 1.5: Wiener-Hammerstein Model.

1.2.6 Parallel Hammerstein Model

In Fig. 1.6, a block diagram of the parallel Hammerstein model is presented. The parallel Hammerstein model consists of multiple Hammerstein branches with a common input and an output which is the summation of the individual branch outputs. This system can be represented by

$$\begin{aligned} y(t) &= \sum_{i=1}^{L} y_i(t), \\ y_i(t) &= \int h_i(\tau)s_i(t-\tau)d\tau, \\ s_i(t) &= f_i(z(t)), \end{aligned} \tag{1.17}$$

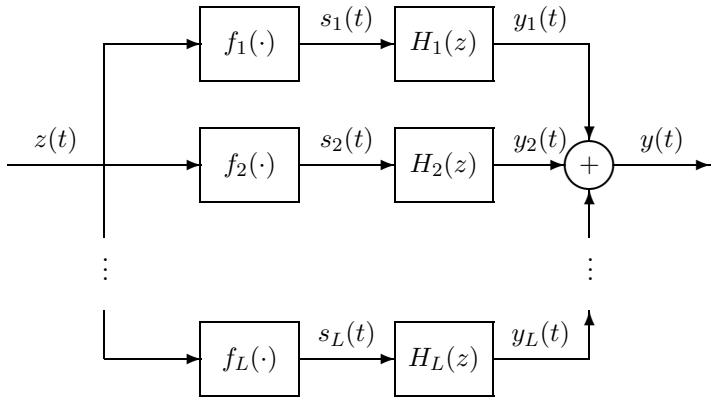


Figure 1.6: Parallel Hammerstein model.

where $h_i(\tau)$ denotes the LTI portion of the Hammerstein system and $f_i(\cdot)$ denotes the memoryless nonlinearity at the ith branch. Compared to the Wiener and Hammerstein models, this model is more general and therefore can provide a more accurate model for the PA. In the single branch case, i.e., $L = 1$, we obtain the Hammerstein model. If we set $h_i(t) = \delta(t - iT)$, then we can interchange the LTI system with the memoryless nonlinearity. The resulting system is presented in Fig. 1.7. This model was used in [65] to represent a nonlinear PA with memory effects and is given by

$$y(t) = \sum_{i=1}^{L} f_i(z(t - iT)). \tag{1.18}$$

For a simple baseband representation, each memoryless nonlinearity can be represented by a baseband polynomial model with a different set of coefficients. In such case, the model in (1.18) becomes

$$y(t) = \sum_{k=0}^{K} \sum_{i=0}^{L-1} c_{ki} |z(t - iT)|^{2k} z(t - iT), \tag{1.19}$$

which in [39] is referred to as *memory polynomial*. Note that similarly to the Hammerstein model with the polynomial nonlinearity, the memory polynomial is linear w.r.t. its coefficients $\{c_{ki}\}$. This property allows the use of linear techniques to identify a PA using the memory polynomial model.

Another useful property of the memory polynomial model is that it can be easily converted to a compact parallel Hammerstein system. This can be done by applying the

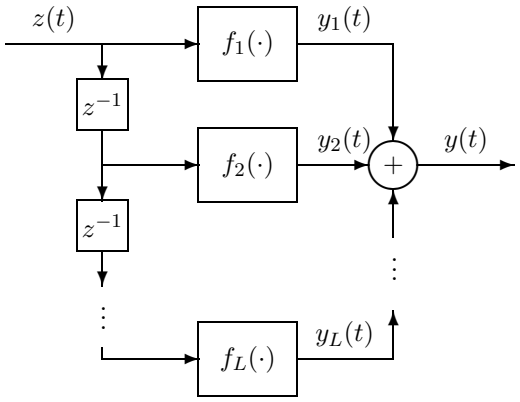


Figure 1.7: Memory Nonlinearity [65]. We use z^{-1} to denote a delay of time T. When a signal is sampled, z^{-1} represents a single sample delay of the sampled signal. With a sampling period of T our notation coincides with the conventional notation.

compact version of the singular value decomposition (SVD) to the matrix $\mathbf{C} = [c_{ki}]$. We expand $\mathbf{C}$ as $\mathbf{C} = \sum_{l=1}^{\tilde{L}} \gamma_l \mathbf{u}_l \mathbf{v}_l^H$, where γ_l, $\mathbf{u}_l$, and $\mathbf{v}_l$ are the singular values, singular column vectors, and singular row vectors of the matrix $\mathbf{C}$, respectively. Note that due to the compact representation $\tilde{L} \leq \min\{L, K+1\}$. Substituting c_{ki} with $\sum_{l=1}^{\tilde{L}} \gamma_l \mathbf{u}_{lk} \mathbf{v}_{li}^*$, we can rewrite (1.19) as

$$y(t) = \sum_{l=1}^{\tilde{L}} H_l[f_l(z(t))], \tag{1.20}$$

where

$$H_l[s_l(t)] = \sum_{i=0}^{L-1} \gamma_l \mathbf{v}_{li} s_l(t - iT), \qquad f_l(z(t)) = \sum_{k=0}^{K} \mathbf{u}_{lk} |z(t)|^{2k} z(t). \tag{1.21}$$

This is an $\tilde{L}$ branch representation as opposed to L branches in Fig. 1.7. Consider for example the case where $K = 5$ (polynomial order of $2K + 1 = 11$) and $L = 20$ (memory of 20 taps) in (1.18), i.e., a 20 branch system. This model can be easily simplified to an $\tilde{L} = 5$ branch system in (1.20).

1.3 PA Spectral Analysis

The design of PAs is constrained by spectral emission limits (e.g., Table 1.1). Nonlinear PAs create spectral regrowth, i.e., spectral broadening, which creates adjacent channel

Table 1.1: Spurious emission limits when transmitting [1]

For frequency f with $\lvert f - \text{Center Frequency}\rvert$	Greater than 780kHz	Greater than 1.98MHz
Spurious emission levels shall be less than either (a), or both (b) and (c)	(a) -42dBc/30kHz	(a) -54dBc/30kHz
	(b) -60dBc/30kHz (c) -54dBc/1.23MHz	
Spurious emission levels should be less than either (a), or both (b) and (c)	(a) -45dBc/30kHz	(a) -60dBc/30kHz
	(b) -66dBc/30kHz (c) -60dBc/1.23MHz	

interference. In general, the more nonlinear the PA is, the more spectral regrowth it will generate. It is desirable to design an efficient PA while keeping the spectral emissions of the PA below the limit. To do so, we have to consider the characteristics of the input signal to the PA as well as specific parameters of the PA. To obtain the power spectrum at the output of the PA, one can run extensive simulations to model both the input signal and the PA characteristics. To avoid that, we would like to have an analytic tool that allows us to compute the spectral regrowth for a prescribed level of PA nonlinearity. Such an analytic tool can simplify the process of optimizing the PA efficiency subject to spectral constraints.

When the PA is memoryless, the relationship between its input and output can be written by

$$y(t) = f(z(t)), \tag{1.22}$$

where $z(t)$ and $y(t)$ are the PA input and output, respectively. Given the input statistical characteristics, our goal is to obtain a closed-form expression for the output power spectrum given by

$$S_{2y}(f) = \int_{-\infty}^{\infty} c_{2y}(\tau) e^{-j2\pi\tau f} d\tau \tag{1.23}$$

where $c_{2y}(\tau) = E[y^*(t)y(t+\tau)]$ is the auto-covariance function of the output signal. In Fig. 1.8, we present a PA output power spectrum, $S_{2y}(f)$, for a nonlinear PA (solid line) and an input signal with bandwidth $\Delta = 0.1$. We observe that while the input is restricted to $[-\frac{\Delta}{2}, \frac{\Delta}{2}]$, the output spectrum (solid line) is the range $[-0.3, 0.3]$.

In the following, we present a literature review for analysis of the auto-covariance and spectrum at the output of a nonlinearity. We start by presenting basic results in the

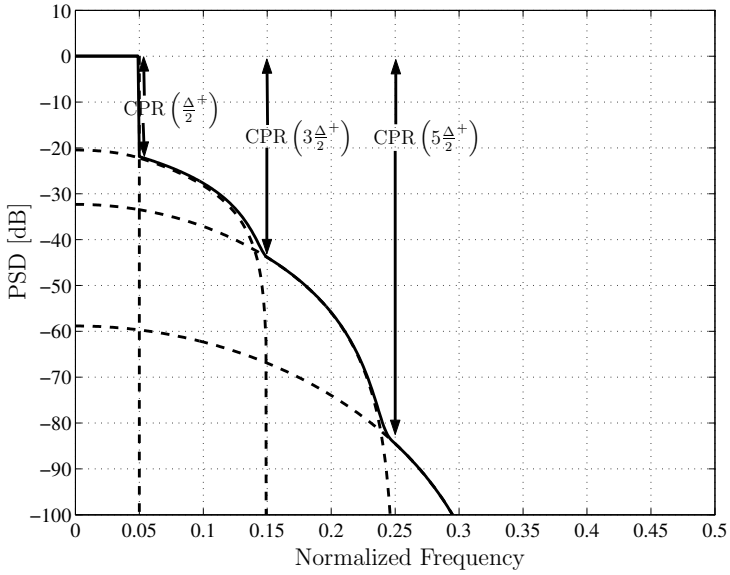


Figure 1.8: PA output power spectrum example.

analysis of the auto-covariance of a nonlinearity. In [16], the PA output auto-covariance function $c_{2y}(\tau)$ is expressed in terms of the input auto-covariance function $c_{2z}(\tau)$ in the case of real-valued Gaussian distributed input. A closed-form expression for the output auto-covariance in terms of the input auto-covariance is examined for different nonlinearities such as $f(z) = z^m$, $f(z) = z^m u(z)$, $f(z) = |z|^m$, hard- and soft-limiter. For example, with $f(z) = z^2$, the output auto-covariance is given by $c_{2y}(\tau) = c_{2z}^2(0) + 2c_{2z}^2(\tau)$. The effect of applying operators such as scaling, translation, differentiation, and integration to the input signal on the output auto-covariance are also presented. Similarly relationships are obtained in [77] for elliptically-symmetric distributions. The output auto-covariance of a μ-law nonlinearity of a real-valued Gaussian input is derived in [121].

A general closed-form expression for the output auto-covariance does not exists in the general case. Two main different approaches are taken to obtain a closed-form expression for a general nonlinearity. The first approach is presented by Barrett and Lampard, which offers the following orthogonal expansion of the nonlinearity:

$$f(z) = \sum_{k=0}^{\infty} \alpha_k \phi_k(z), \tag{1.24}$$

where $\alpha_k = E[\phi_k^*(z) f(z)]$ are the coefficients associated with the nonlinearity $f(\cdot)$ and $\phi_k(z(t))$ is orthogonal to $\phi_m(z(t+\tau))$ for all $k \neq m$ and any given t and τ, i.e., $E[\phi_k^*(z(t))\phi_m(z(t+$

$\tau))] = 0$, $\forall k \neq m$ [14]. As a result, the output auto-covariance is given by

$$\begin{aligned} c_{2y}(\tau) = E[f^*(z(t))f(z(t+\tau))] &= \sum_{k=0}^{\infty}\sum_{m=0}^{\infty} \alpha_k^*\alpha_m E[\phi_k^*(z(t))\phi_m(z(t+\tau))] \\ &= \sum_{k=0}^{\infty} |\alpha_k|^2 E[\phi_k^*(z(t))\phi_k(z(t+\tau))]. \end{aligned} \tag{1.25}$$

A closed-form expression for $\phi_k(\cdot)$ is given in three cases: real-valued Gaussian distributed input, Ricean distributed input, and an input of the form $z(t) = cos(\omega t + \theta)$ with uniformly distributed θ in $[0, 2\pi]$. An input/output cross-covariance expression for the three input distributions is obtained in [26] and [69]. The corresponding orthogonal functions $\phi_k(\cdot)$ are the Hermite, Laguerre, and Tchebycheff polynomials, respectively. For example, in the real-valued case $\phi_k(z) = H_k(z)$, the Hermite polynomials. In this case,

$$\begin{aligned} c_{2y}(\tau) &= \sum_{k=0}^{\infty} |E[f(z)H_k(z(t))]|^2 E[H_k(z(t))H_k(z(t+\tau))] \\ &= \sum_{k=0}^{\infty} |E[f(z)H_k(z(t))]|^2 \frac{(c_{2z}(\tau))^k}{k!}. \end{aligned} \tag{1.26}$$

This result is a general auto-covariance expression of any given nonlinearity $f(\cdot)$ linking the output auto-covariance and the input auto-covariance. The output power spectrum is given by

$$S_{2y}(\tau) = \sum_{k=0}^{\infty} |E[f(z)H_k(z(t))]|^2 \frac{S_{2z}(f) * S_{2z}(f) * \cdots * S_{2z}(f)}{k!}, \tag{1.27}$$

where $*$ denotes the convolution operator. Note that the output spectrum in (1.27) is a weighted sum of various kth-time convolutions of the input power spectrum. A closed-form expression for $E[\phi_k^*(z(t))\phi_k(z(t+\tau))]$ in (1.25) in terms of $c_{2z}(\tau)$ is obtained for the real-valued Gaussian [20, 27, 122] and the complex-Gaussian [21] cases. Since the closed-form auto-covariance expression is suitable for any nonlinearity, various nonlinearities have been examined. The output power spectrum expression for a bandpass nonlinearity characterized by AM-AM and AM-PM conversions is provided in [31] along with an expression for a μ-law device. In [42], a hyperbolic tangent NL with real-valued Gaussian input is analyzed.

This approach is used to analyze PA output power spectrum for various input signal distributions and PA models. In [28], for a PA with the Saleh model [104] PA and soft-limiter

are analyzed. An OFDM signal is approximated as a complex-Gaussian stationary input signal. The power spectrum of such a signal is evaluated for a traveling wave tube amplifier (TWTA) PA following the Saleh model [13, 34] and for a solid state PA (SSPA) of the form $(1+|z|^{2p})^{-1/2p}$ [34]. A general closed-form PA output power spectrum expression for a complex-Gaussian input (e.g., OFDM) is derived in [10–12] and a closed-form expression for a soft-limiter nonlinearity is obtained. In [29], spectral analysis of a PA modeled by Tchebichev polynomials is carried out.

Another approach is to model $f(\cdot)$ using the polynomial model, e.g., $f(z)=\sum_{k=0}^{K} c_k z^k(t)$ for a bandpass nonlinearity or $f(z)=\sum_{k=0}^{K} c_{2k+1}|z|^{2k}(t)z(t)$ for a baseband nonlinearity. In the real-valued Gaussian case, to obtain a closed-form expression for the output auto-covariance, $E[z^k(t)z^m(t+\tau)]$ should be evaluated. A closed-form solution is given in [101]. The Price theorem [87] and its generalization in [85] also offer a solution for this problem. In the complex-Gaussian case, to obtain a closed-form expression for the output auto-covariance, $E[|z|^{2k}(t)z^*(t)|z|^{2k}(t+\tau)z^*(t+\tau)]$ should be evaluated. This can be done using the extension of the Price theorem to the complex-case [117] or by the complex-Gaussian moments expressions in [76]. After obtaining the output auto-covariance, we can apply the Fourier transform in (1.24) to obtain the power spectrum.

Following this approach, a closed-form expression for the output power spectrum is obtained in [43] assuming a complex-Gaussian input and a 3rd-order polynomial. A 5th-order nonlinear PA is examined in [43, 113]. In [50], the analysis was carried out for a 9th-order nonlinear PA and is generalized to the nth-order in [58]. Using the same result, an IS-95 PA output spectrum is evaluated in [55], whereas for the OFDM signal adjacent channel interference (ACPR) is evaluated in [56]. When only AM-AM conversion exists, a real-valued analysis is used to obtain the output power spectrum for a 3rd-order polynomial [123–125]. Similarly, a 5th-order expression is derived in terms of the PA gain G, its 3rd-order interception point (IP3) and its 5th-order interception point (IP5) in [37], specifically for CDMA, TDMA inputs in [126] and for TDMA and 16QAM inputs in [71]. In [128], real-valued Gaussian signal is used to model the bandpass signal. As an alternative to the Gaussian moment theorem, [129] uses cumulant based approach for a 7th-order nonlinear

PA with (non-)Gaussian input. In [7], CDMA with QPSK modulation is analyzed and a closed-form output spectrum expression for a 3rd-order PA is obtained. In [127], a third-order analysis is carried out for a BPSK and QPSK input and a Bessel-series PA.

In [20], an LTI system is considered after a memoryless PA (Wiener model). For this model, the output power spectrum is multiplied with the LTI system frequency response magnitude. However, to the best of our knowledge there has been no closed-form expression for spectral regrowth for other PA models with memory effects.

1.4 Power Amplifier Linearization

In Section 1.2, we mentioned that PA modeling and characterization can help in selecting a PA to match a given application. To ensure linear amplification of a signal, a PA with a power higher than required is usually selected such that the input signal fits into the linear region of the PA. This "back-off" approach can result in a significant increase in PA power specification as well as reduced efficiency, since the PA is using a high DC power but is utilizing only a small portion of its allowed input range. The linearization approach offers a remedy to this problem. To illustrate this concept, we present an example in Table 1.2 [2]. We compare a linearized 80W PA to an unlinearized 160W PA. The 80W PA is

Table 1.2: Digital predistortion: reduced power consumption, less heat dissipation and cooling costs, and higher reliability can be achieved.

Digital Predistortion	No	Yes
PA Power Rating	160 W	80 W
Peak Power	80 W	80 W
PAR	9 dB	9 dB
Average Power	10 W	10 W
Back-off	12 dB	9 dB
Efficiency	9%	18%
Power Dissipation	101 W	45 W

linearized and can operate in the linear mode with a peak power of 80W. On the other hand, to ensure linear operation, the unlinearized 160W PA may only be utilized up to a peak power of 80W. Since only half of its power is utilized, i.e., the 160W PA uses additional 3dB back-off, its efficiency is lower than that of the 80W PA. Its power dissipation is more

than twice than that of the 80W PA.

We identify a few major advantages for the linearized 80W PA. It can be smaller and cheaper than the 160W PA since both price and size generally increase with the PA power specification. It consumes less power and therefore requires less energy to operate. Alternatively, if the PA is operating using a battery, then the battery will last longer. Since it consumes less power, less power will dissipate as heat and cooling costs can be reduced. Due to its advantages, linearization has attracted a great deal of attention in the literature. Next, we present different approaches to linearization.

1.4.1 PA Linearization Architecture

PA linearization can be implemented using different architectures. Here, we consider three of these architectures.

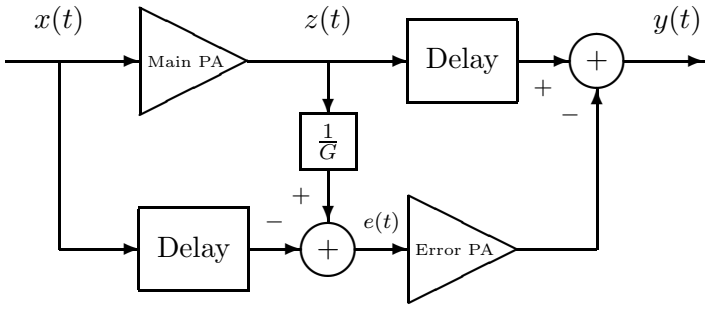


Figure 1.9: Feedforward architecture.

1.4.1.1 Feedforward linearization

The feedforward linearization technique [22] was invented by H. S. Black along with negative feedback [23]. While negative feedback suffers limited IMD suppression and instability problems, feedforward can provide in theory full IMD suppression. In Fig. 1.9, the feedforward linearization architecture is presented. The input signal, $x(t)$, is split into two branches. In the main branch, the input signal $x(t)$ is amplified by the main PA (typically a high power amplifier) yielding the PA output $z(t)$. In the secondary branch, the PA output $z(t)$ is scaled and compared with the original input $x(t)$. The resulting error

signal $e(t)$ goes through a second PA, known as the error PA. Typically, the error PA is low power and highly linear as compared to the main PA. After the error signal $e(t)$ is obtained, it is amplified and subtracted from the delayed output of the main PA. Since the error signal $e(t)$ is the nonlinear distortion, removing it from the PA output linearizes the PA. While ideally, this architecture is designed to perfectly linearize the PA, it is sensitive to changes in the parameters of the PA due to factors such as temperature, aging effects, and amplitude/phase matching, which require the gain G to continuously adapt [90].

1.4.1.2 Predistortion

While the feedforward linearization corrects for the nonlinearity after the PA, the predistortion architecture corrects for the nonlinearity before the PA. In Fig. 1.10, the input signal $x(t)$ goes first through a predistorter which ideally implements the inverse of the transfer function of the PA (up to a scaling factor). Then, the predistorted signal $z(t)$ goes through the PA, yielding $y(t)$, which is ideally a magnitude-scaled version of the input signal $y(t) = Gx(t)$. Predistortion can be preformed either in baseband or in radio frequency (RF). In analog, the expanding characteristics of the predistorter can be obtained by subtracting a compressive transfer function (of a diode for example) from a linear transfer function. With DSP, on the other hand, the predistorter can be implemented using a lookup table (typically a memory device) indexed by the input signal magnitude. A DSP is used to compute and update the values of the lookup table. When the PA exhibits memory effects, a simple memoryless mapping (such as the lookup table) can no longer be used to fully linearize the PA; thus memory effects should be included.

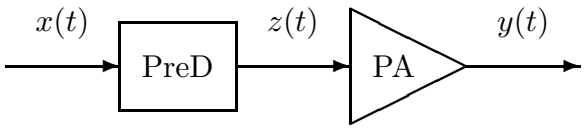


Figure 1.10: Predistortion.

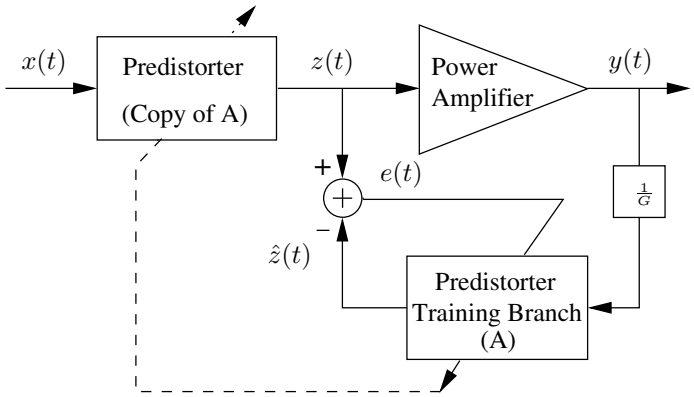


Figure 1.11: Indirect learning architecture.

1.4.1.3 Indirect learning architecture

Predistortion can be implemented using the indirect learning architecture [46] as shown in Fig. 1.11. The baseband predistorter input is denoted by $x(t)$, the baseband predistorter output/PA input is denoted by $z(t)$, and the baseband PA output is denoted by $y(t)$. The feedback path labeled "Predistorter Training Branch" (block A) has $y(t)/G$ as its input, where G is the intended gain of the PA, and $\hat{z}(t)$ is its output. The actual predistorter (copy of A) is an exact copy of the predistorter training branch. When $y(t) = Gx(t)$, the error $e(t) = z(t) - \hat{z}(t)$ is 0. To reduce the error between $y(t)$ and $Gx(t)$, we choose the predistorter parameters that minimize the error $e(t)$. The benefit of the indirect learning architecture is that, instead of assuming a model for the PA, estimating the PA parameters and then constructing its inverse, we can go directly after the predistorter[1].

1.4.2 Predistortion Models

Considering predistortion as the main architecture, we still have to select an appropriate model for the predistorted block. Issues such as bandwidth and memory effects should be taken into account.

[1]The term "indirect learning" seems counter-intuitive here, since the predistorter is learned directly; it is the PA characteristics that are learned indirectly.

1.4.2.1 *Memoryless predistorter*

When the PA is assumed memoryless, we can select a memoryless predistorter to predistort it. The input/ouput relationship of the memoryless predistorter is given by:

$$z(t) = \sum_{k=0}^{K} a_{2k+1}|x(t)|^{2k}x(t) \tag{1.28}$$

where a_{2k+1} are complex-valued coefficients, $x(t)$ is the input to the predistorter, $z(t)$ is the output of the predistorter and the input to the PA as shown in Fig. 1.10. Note that the overall response of the predistorter given by (1.28) and the PA given by (1.1) is also a polynomial of the same structure. Using the direct approach, we can first obtain the PA coefficients. Then, we would like the coefficients of the overall predistortion and PA concatenated polynomial to be zero for the nonlinear terms and constant G (PA gain) for the linear component. This cannot be done for a finite order polynomial PA. An alternative is to use the predistorter as the pth-order inverse of the polynomial model in (1.1). This means that only the first $p-1$ nonlinear terms of the polynomial resulting from the concatenation of the predistorter and the PA are set to zero. Errors in the estimates of the PA coefficients will yield error in the predistorter coefficients.

When the indirect learning architecture is considered, the coefficients of the predistorter $\{a_{2k+1}\}$ are chosen to minimize the error $e(z)$ in Fig. 1.11. One approach is to look for the coefficients a_{2k+1} that form a least squares fit to (1.28) with $x(t)$ replaced with $\frac{y(t)}{G}$. This approach does not require the estimation of the PA coefficients. Furthermore, the PA does not have to be polynomial. In [80], a third-order predistorter of the form given by (1.28) (with $K = 1$) is used to predistort a nonlinear PA in CDMA communications.

When the complex coefficient model in (1.28) uses the cartesian representation of the coefficients $\{a_{2k+1}\}$, i.e., they are separated into their real and imaginary components, it is referred to as the cartesian model [102]. An alternative to this model is the polar model [102] given by

$$z(t) = \left(\sum_{k=0}^{K} \alpha_{2k+1}|x(t)|^{2k+1}\right) e^{j\sum_{k=0}^{K}(\beta_{2k+1}|x(t)|^{2k})}. \tag{1.29}$$

where $\{\alpha_{2k+1}\}$ are the coefficients for the AM/AM transfer function and $\{\beta_{2k+1}\}$ are the

coefficients for the AM/PM transfer function. Note that this model is different from the model in (1.28). To find the coefficients $\{\alpha_{2k+1}\}$ and $\{\beta_{2k+1}\}$, we can apply the least squares fit to (1.29). While the cartesian model is linear w.r.t. its parameters, the polar model is nonlinear w.r.t. to its parameters. Therefore, a least-square solution may not be easy to obtain.

To reduce the computational complexity and adapt to time variations of the PA, a sequential estimate for the PA coefficients can be formed as opposed to a batch estimator. A sequential estimator evaluates the estimate at time $N+1$ based on the estimate at time N and the most recent data sample at time $N+1$. A batch approach takes all $N+1$ samples into account, therefore, increasing computational complexity. Adaptive approaches such as the gradient method (e.g., [19, 62, 70]) and recursive least squares (RLS) (e.g., [60]) have been offered to obtain the predistorter's coefficients (in both (1.28 and (1.29)) to overcome the complexity issue associated with the batch estimators.

1.4.2.2 Volterra Model

To predistort PAs with memory effects which were discussed earlier in the context of PA modeling, a predistorter with memory effects must be considered. A general model for such predistorter is the Volterra series model. The Volterra series model for the predistorter is similar to that used for PA modeling and is given by:

$$z(t) = \sum_{k=0}^{K} \int \cdots \int A_{2k+1}(\tau_1, \ldots, \tau_{2k+1}) \prod_{i=1}^{k+1} x(t-\tau_i) \prod_{i=k+2}^{2k+1} x^*(t-\tau_i) d\tau_1 \cdots d\tau_{2k+1} \quad (1.30)$$

where $A_{2k+1}(\tau_1, \ldots, \tau_{2k+1})$ is the $(2k+1)$th-order kernel of the Volterra predistorter. In [47] and [134], the Volterra series model is used to construct the predistorter model. While the number of basis functions representing the Volterra series model is very large, in [47] and [134] only selected basis functions are used for predistortion. For example, [134] uses terms of the form $x(n-k)x(n-l)x(n-m)$ where $k, l, m \in 0, 1, 2$, $x(n-k)x(n-l)$, where $k, l \in 0, 1, 2$, and $x(n-k)$, where $k \in 0, 1, 2$. Reference [47] uses similar terms where $k, l, m \in 0, 1, \ldots, N$. The model in [47] is a third-order Volterra series with a larger number of basis functions. In the memoryless case sometimes fifth-, seventh-, and also ninth-order models are used.

1.4.2.3 Hammerstein Model

The Hammerstein model reduces the complexity associated with the general Volterra series model. Similar to the models offered in [47] and [134] the Hammerstein model is also a special case of the Volterra series model. The Hammerstein model for the predistorter is similar to that used for PA modeling and is given by:

$$z(t) = \int \phi(\tau) s(t-\tau) d\tau, \quad s(t) = \sum_{k=0}^{K} a_{2k+1} |x(t)|^{2k} x(t), \tag{1.31}$$

where $\{a_k\}$ are the coefficients of the polynomial model of the memoryless nonlinearity and $\phi(\tau)$ is the impulse response of the LTI portion of the Wiener system. In [62], a Hammerstein model is used to predistort a high power amplifier in the context of satellite communications. The downlink channel consists of a high power amplifier followed by an LTI transfer function modeling the communication channel between the satellite and the earth.

1.4.2.4 Parallel Hammerstein Model

Similarly to PA modeling, the parallel Hammerstein predistorter can also be used for predistortion. It consists of multiple Hammerstein systems in parallel and is given by

$$z(t) = \sum_{i=1}^{L} z_i(t), \qquad z_i(t) = \int h_i(\tau) s_i(t-\tau) d\tau, \quad s_i(t) = f_i(x(t)), \tag{1.32}$$

where $h_i(\tau)$ denotes the LTI portion of the Hammerstein system at the ith branch and $f_i(\cdot)$ denotes its memoryless nonlinearity. Replacing the nonlinearity $f_i(x(t))$ with $\sum_{k=0}^{K} a_{ki} |x(t)|^{2k} x(t)$ and $h_i(\tau)$ with $\delta(\tau - iT)$, we obtain the memory-polynomial model given by

$$z(t) = \sum_{k=0}^{K} \sum_{i=0}^{L-1} a_{ki} |x(t-iT)|^{2k} x(t-iT). \tag{1.33}$$

In [39], a memory-polynomial predistorter is shown to be effective in linearizing a nonlinear PA with memory effects.

1.5 Communications Aspects of Power Amplifiers

Power amplifiers (PAs) are peak power limited in addition to being nonlinear. As such, ideal linearization via predistortion is not possible. Linearized PA can only be linearized to

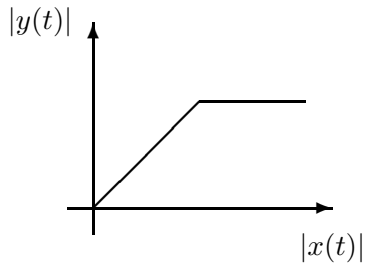


Figure 1.12: Soft-limiter.

the soft-limiter in Fig. 1.12. Because a peak power limited PA cannot be truly linearized, it introduces some distortion. This distortion can distort the constellation of a transmitted communication signal. In Fig. 1.13, we show the constellation of a 16-QAM signal ($\circ$) and its distorted version ($\times$). The nonlinear distortion changes the constellation and therefore makes detection of the symbols more sensitive to noise. The nonlinear distortion is not

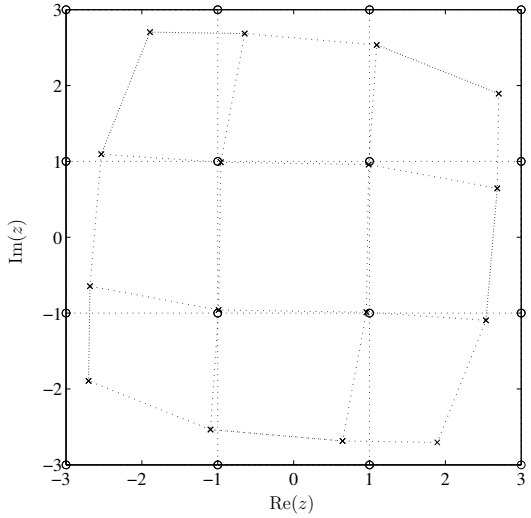


Figure 1.13: The constellation of a 16QAM signal.

exclusive to the PAs and can also be created by other devices such as mixers [18, 64, 72], magnetic recording channels [135], or when companding [59, 120] or clipping [83, 84, 103, 108, 115] are used for the purpose of peak-to-average power ratio (PAPR) reduction. We would like to investigate what undesirable effects are caused by the nonlinearity. There has been a lot of research devoted to this issue. Nonlinearity causes increase in symbol-error-rate (SER) [64,108], spectral regrowth [108,129], and reduction in channel capacity [107,109,115].

1.6 Organization of this Dissertation

In the rest of this document, we will present our results in areas of PA modeling, PA linearization, PA spectral analysis, and nonlinearity considerations in the context of communications applications. In the next two chapters, we describe results on PA modeling. In Chapter 2 ([95]), we present nonlinear PA modeling of memoryless, quasi-memoryless, and memory PAs. In Chapter 3 ([130]), we discuss the issue of baseband modeling of PAs and verify these results with measured data. In the following three chapters, we present results in PA linearization. In Chapter 4 ([41]), we examine memory predistortion of PAs using a Hammerstein predistorter. In Chapters 5 ([91,93]) and 6 ([96,98]), we introduce the application of orthogonal polynomials to predistortion when the input signal is uniformly and Gaussian distributed, respectively. In Chapter 6 ([96,98]), we discuss the application of such orthogonal polynomials to spectral regrowth analysis. In the following three chapters we present additional spectral regrowth analysis results. In Chapter 7 ([97,99]), we discuss a spectral analysis for a PA with a nonstationary input signal. In Chapter 8 ([100]), we examine a spectral analysis for QPSK and OQPSK signals. In Chapter 9 ([131, 132]), we present spectral analysis results for a memory polynomial PA. In Chapter 10 ([89, 92, 94]), we evaluate amplitude-limited nonlinearities and their effects on communications. Finally, in Chapter 11, we summarize this work and suggest topics for future research. For the reader's convenience, we have made an effort to keep every chapter as self contained as possible.

CHAPTER II

ON THE MODELING OF MEMORY NONLINEAR EFFECTS OF POWER AMPLIFIERS FOR COMMUNICATION APPLICATIONS

Understanding power amplifier (PA) nonlinearity is a first step towards linearization efforts. We first explore the passband and baseband PA input/output relationships and show that they manifest differently when the PA exhibits long-term, short-term, or no memory effects. We then explain the various memory effects in the context of AM/AM and AM/PM responses. The so-called quasi-memoryless case is especially clarified. Four particular nonlinear models with memory are further investigated.

2.1 Introduction

Power amplifier (PA) is an indispensable component of a communication system and is inherently nonlinear. PAs can be classified according to the degree of nonlinearity that they exhibit, which in turn dictates efficiency. High linearity implies low efficiency, which means reduced power that is delivered to the load. High nonlinearity, on the other hand, causes spectral regrowth and increases bit error rate. Therefore, PA linearization is often necessary with the goal of improving linearity while maintaining good efficiency.

Predistortion is a popular approach to linearize a PA for which PA modeling often serves as an important first step. Before choosing a linearization method, one must decide whether the PA exhibits memory effects. The cause of memory effects can be electrical or electrothermal [118]. High-power amplifiers (HPAs) such as those used in wireless basestations exhibit memory effects; wideband signals also tend to induce memory effects in the PA. In such cases, memoryless predistortion can be ineffective. Thus, accurate representation of the memory effects in nonlinear PAs is crucial to linearization efforts.

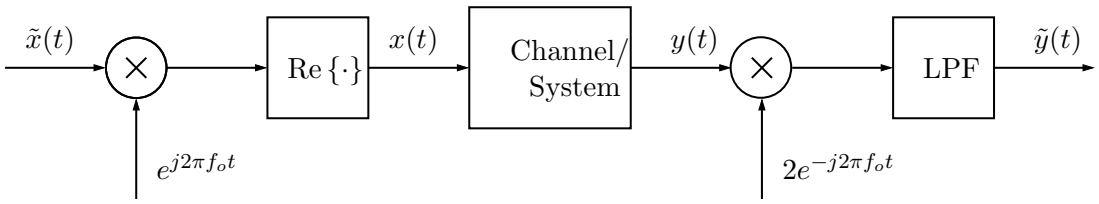


Figure 2.1: Block diagram of a bandpass system.

Denote by $\tilde{x}(t)$ the baseband input to the PA and by $\tilde{y}(t)$ the baseband output of the PA. The so-called AM/AM conversion is defined as the function mapping from the baseband input signal amplitude, $|\tilde{x}(t)|$, to the baseband output signal amplitude, $|\tilde{y}(t)|$; AM/PM conversion is defined as the function mapping from the baseband input signal amplitude, $|\tilde{x}(t)|$, to the baseband output phase deviation, $\angle\tilde{y}(t) - \angle\tilde{x}(t)$. Linearization of a nonlinear device which can be completely characterized by AM/AM and AM/PM conversions is rather straightforward. For example, the predistorter can be implemented by a lookup table (LUT). The LUT creates nonlinear function mappings (AM/AM and AM/PM) for the predistorter that are complementary to those of the PA.

Let us consider as an example, a baseband $(2K+1)$th-order nonlinear PA model [18],

$$\tilde{y}(t) = \sum_{k=0}^{K} a_{2k+1}|\tilde{x}(t)|^{2k}\tilde{x}(t), \tag{2.1}$$

where a_{2k+1} is the coefficient of the $(2k+1)$th-order nonlinear term.

From a signal processing point of view, the system in (2.1) is considered memoryless, since the output $\tilde{y}(t)$ depends only on the input $\tilde{x}(t)$ at the same time instant. In the RF/Microwave literature however, if $\{a_{2k+1}\}$ are real-valued, the model in (2.1) is referred to as (strictly-) memoryless since the AM/PM conversion is constant. Otherwise, the system described by (2.1) is considered quasi-memoryless, which implies that short-term memory effects exist in the system (e.g., [24]). The symptom of a quasi-memoryless system is that AM/PM conversion varies with $|\tilde{x}(t)|$ [18]. Therefore, confusion often arises as to the exact nature of memory effects in nonlinear devices used in communications applications.

When long-term memory effects are present, AM/AM and AM/PM conversions are insufficient to characterize the PA, and more elaborate models such as the Volterra series can be used [24].

In this chapter, we will examine what memory effects in the passband mean, and what memory effects in the baseband mean. Our hope is to resolve the confusing notions and conflicting terminologies of memory nonlinear effects used in different literatures.

In Section 2.2, the nonlinear baseband model and its relationship to the nonlinear bandpass model is introduced. In addition, a discussion on the various degrees of memory effects is presented. In Section 2.3, we investigate four special systems with memory. Finally, Section 2.4 summarizes this chapter.

2.2 Baseband Representation

In this section, we derive the input/output relationship of the baseband signal from the input/output relationship of the corresponding bandpass signal. We then consider two cases, quasi-memoryless and (strictly-) memoryless.

In Fig. 2.1, a block diagram depicts the upconversion of the complex-valued baseband signal to the passband, the transmission of the real-valued bandpass signal through the channel, and the downconversion of the bandpass output into a complex-valued baseband signal. The relationship between the real-valued bandpass input signal, $x(t)$, and the complex-valued baseband input signal, $\tilde{x}(t)$, is given by:

$$\begin{aligned} x(t) = \mathrm{Re}\left\{\tilde{x}(t)e^{j2\pi f_o t}\right\} \qquad &\Leftrightarrow \\ \tilde{x}(t) = 2\mathrm{LPF}\left[x(t)e^{-j2\pi f_o t}\right], \end{aligned} \tag{2.2}$$

where f_o is the carrier frequency, $\mathrm{LPF}[\cdot]$ denotes the lowpass filtering operation, and $\mathrm{Re}\{\cdot\}$ denotes the real-value part. Similarly, the relationship between the real-valued bandpass output signal, $y(t)$, and the complex-valued baseband output signal, $\tilde{y}(t)$, is given by:

$$\begin{aligned} y(t) = \mathrm{Re}\left\{\tilde{y}(t)e^{j2\pi f_o t}\right\} \qquad &\Leftrightarrow \\ \tilde{y}(t) = 2\mathrm{LPF}\left[y(t)e^{-j2\pi f_o t}\right]. \end{aligned} \tag{2.3}$$

The "channel"[1], in general, can be linear or nonlinear (NL). Subject to certain requirements ([106]), the Volterra series can be used to represent a time-invariant (TI) NL system as

[1]Channel is used here with quotation marks, since it refers to the PA, although traditionally it refers to a communication channel.

follows:

$$y(t) = \sum_k \int \cdots \int h_k(\boldsymbol{\tau}_k) \prod_{i=1}^{k} x(t-\tau_i) d\boldsymbol{\tau}_k, \tag{2.4}$$

where $\boldsymbol{\tau}_k = [\tau_1, \ldots, \tau_k]^T$, $h_k(\cdot)$ is the kth-order Volterra kernel, and $d\boldsymbol{\tau}_k = d\tau_1 d\tau_2 \cdots d\tau_k$. Assuming that $x(t)$ is band-limited with bandwidth $B_x \ll f_o$, and substituting (2.2) and (2.3) into (2.4), the complex-valued baseband output, $\tilde{y}(t)$, simplifies to [17]:

$$\tilde{y}(t) = \sum_k \int \cdots \int \tilde{h}_{2k+1}(\boldsymbol{\tau}_{2k+1}) \prod_{i=1}^{k+1} \tilde{x}(t-\tau_i) \prod_{i=k+2}^{2k+1} \tilde{x}^*(t-\tau_i) d\boldsymbol{\tau}_{2k+1}, \tag{2.5}$$

where

$$\tilde{h}_{2k+1}(\boldsymbol{\tau}_{2k+1}) = \frac{1}{2^{2k}} \binom{2k+1}{k} \times h_{2k+1}(\boldsymbol{\tau}_{2k+1}) e^{-j2\pi f_o \left(\sum_{i=1}^{k+1} \tau_i - \sum_{i=k+2}^{2k+1} \tau_i \right)}, \tag{2.6}$$

and $(\cdot)^*$ denotes complex conjugation. Define the kth-dimensional Fourier transform of $h_k(\boldsymbol{\tau}_k)$:

$$H_k(\mathbf{f}_k) = \int \cdots \int h_k(\boldsymbol{\tau}_k) e^{-j2\pi \mathbf{f}_k^T \boldsymbol{\tau}_k} d\boldsymbol{\tau}_k, \tag{2.7}$$

where $\mathbf{f}_k = [f_1, f_2, \ldots, f_k]^T$. It follows that the Fourier transform of (2.6) is

$$\tilde{H}_{2k+1}(\mathbf{f}_{2k+1}) = \frac{1}{2^{2k}} \binom{2k+1}{k} \times H_{2k+1} \left(\mathbf{f}_{2k+1} + f_o [\mathbf{1}_{k+1}^T, -\mathbf{1}_k^T]^T \right), \tag{2.8}$$

where $\mathbf{1}_k = [1, \ldots, 1]^T$ is a k-dimensional column vector of ones. For example, with $k = 1$, Eq. (2.8) yields

$$\tilde{H}_3(f_1, f_2, f_3) = \frac{3}{4} H_3 \left(f_1 + f_o, f_2 + f_o, f_3 - f_o \right). \tag{2.9}$$

We would like to examine ramifications of the baseband input/output relationship (2.5) under two scenarios.

2.2.1 The Quasi-memoryless Case

If the signal $x(t)$ is narrowband such that $\tilde{x}(t-\tau_i) \approx \tilde{x}(t)$ over the support of each kernel, $h_k(\boldsymbol{\tau}_k)$, we can replace $\tilde{x}(t-\tau_i)$ by $\tilde{x}(t)$ in (2.5) to obtain

$$\begin{aligned}\tilde{y}(t) &= \sum_k \int \cdots \int \tilde{h}_{2k+1}(\boldsymbol{\tau}_{2k+1}) d\boldsymbol{\tau}_{2k+1} \; |\tilde{x}(t)|^{2k} \tilde{x}(t) \\ &= \sum_k \tilde{H}_{2k+1}(\mathbf{0}_{2k+1}) |\tilde{x}(t)|^{2k} \tilde{x}(t) \qquad (2.10) \\ &= \sum_k \frac{1}{2^{2k}} \binom{2k+1}{k} H_{2k+1}\left(f_o[\mathbf{1}_{k+1}^T, -\mathbf{1}_k^T]^T\right) \\ &\qquad\qquad |\tilde{x}(t)|^{2k} \tilde{x}(t), \qquad (2.11)\end{aligned}$$

where $\mathbf{0}_k = [0, \ldots, 0]^T$ is a column vector of k zeros. Since the right-hand side (RHS) of (2.11) depends on $\tilde{x}(t)$ only, it is regarded as memoryless from a signal processing point of view. Comparing with (2.1), we see that

$$\begin{aligned}a_{2k+1} &= \tilde{H}_{2k+1}(\mathbf{0}_{2k+1}) \\ &= \frac{1}{2^{2k}} \binom{2k+1}{k} H_{2k+1}\left(f_o[\mathbf{1}_{k+1}^T, -\mathbf{1}_k^T]^T\right). \qquad (2.12)\end{aligned}$$

Therefore, either $\tilde{H}_{2k+1}(\mathbf{0}_{2k+1})$ or $H_{2k+1}\left(f_o[\mathbf{1}_{k+1}^T, -\mathbf{1}_k^T]^T\right)$ can determine the polynomial baseband relationship between $\tilde{x}(t)$ and $\tilde{y}(t)$. The baseband relationship (2.10) is memoryless but the physical device as described by $h_k(\boldsymbol{\tau}_k)$ has memory. The fact that the signal is narrowband makes the kernel's Fourier-transform, $H_k(\mathbf{f}_k)$, wideband in comparison. Since $h_k(\boldsymbol{\tau}_k)$ occupies a short time span relative to the time variation of $\tilde{x}(t)$, the memory effect is regarded as short-term. Moreover, $\tilde{H}_{2k+1}(\mathbf{0}_{2k+1}) \propto H_{2k+1}(f_o[\mathbf{1}_{k+1}^T, -\mathbf{1}_k^T \;]^T)$ is complex-valued in general, and hence $\angle\tilde{y}(t) - \angle\tilde{x}(t)$ is not constant (AM/PM conversion is present). Such a system is referred to as quasi-memoryless in the RF/microwave literature.

2.2.2 The (strictly-) Memoryless Case

If the physical device is strictly-memoryless, we will have

$$h_k(\boldsymbol{\tau}_k) = c_k \prod_{i=1}^{k} \delta(\tau_i), \qquad (2.13)$$

where c_k is real-valued. Substituting (2.13) in (2.6), we obtain

$$\tilde{h}_{2k+1}(\boldsymbol{\tau}_{2k+1}) = \tilde{c}_{2k+1} \prod_{i=1}^{2k+1} \delta(\tau_i), \qquad (2.14)$$

where $\tilde{c}_{2k+1} = \frac{1}{2^{2k}}\binom{2k+1}{k} c_{2k+1}$. Since c_{2k+1} is real-valued, both $\tilde{c}_{2k+1}$ and $\tilde{h}_{2k+1}(\boldsymbol{\tau}_{2k+1})$ are real-valued as well. In other words, not only the passband system is strictly-memoryless, but also the baseband system. As a result, the complex-valued baseband output is:

$$\begin{aligned}\tilde{y}(t) &= \sum_k \tilde{H}_{2k+1}(\mathbf{0}_{2k+1})\ |\tilde{x}(t)|^{2k}\tilde{x}(t) \\ &= \sum_k \tilde{c}_{2k+1}\ |\tilde{x}(t)|^{2k}\tilde{x}(t). \end{aligned} \tag{2.15}$$

If a PA's baseband input/output relationship obeys (2.15), its AM/AM characteristic is

$$|\tilde{y}(t)| = |\tilde{x}(t)|\ |G(|\tilde{x}(t)|)|, \tag{2.16}$$

where the complex gain is

$$G(r) = \sum_k \tilde{c}_{2k+1}\ r^{2k}, \quad r = |\tilde{x}(t)|, \tag{2.17}$$

and the AM/PM relationship is described by

$$\angle\tilde{y}(t) - \angle\tilde{x}(t) = \angle G(|\tilde{x}(t)|). \tag{2.18}$$

Note that since the coefficients, $\tilde{c}_{2k+1}$, are real-valued, the phase of $G(r)$ is either 0 or π. However, in order for the phase to change, the gain, $G(r)$, must be zero at that point of phase change, thus making the PA output at the point zero. Since PAs do not attenuate the signal to zero, we conclude that the phase change, $\angle\tilde{y}(t) - \angle\tilde{x}(t)$, remains constant. Therefore, a strictly-memoryless NL PA has no AM/PM conversion.

2.3 Some Special Cases

In Section 2.2, we have seen that a strictly-memoryless NL system has AM/AM conversion but no AM/PM conversion (i.e., the AM/PM conversion is constant). In contrast, a quasi-memoryless NL system exhibits both AM/AM and AM/PM conversions. In this section, we would like to investigate the opposite; i.e., whether a constant AM/PM conversion implies that the nonlinear system is strictly-memoryless. We will see that two special cases of the Volterra model, i.e., the Hammerstein and Wiener systems, are not memoryless but can still be free of the AM/PM conversion when a narrowband input signal is applied.

2.3.1 Diagonal Kernel System

Let us consider a diagonal kernel system, whose kth-order kernel is given by:

$$h_k(\boldsymbol{\tau}_k) = \phi_k(\tau_1) \prod_{i=1}^{k} \delta(\tau_i - \tau_1). \tag{2.19}$$

Note that the kernel values are nonzero only along the diagonal $\tau_1 = \tau_2 = \ldots = \tau_k$. Substituting (2.19) into (2.4), the bandpass input/output relationship is given by

$$y(t) = \sum_k \int \phi_k(\tau_1)\ x^k(t - \tau_1) d\tau_1. \tag{2.20}$$

The RHS of (2.20) is also known as a memory polynomial of $x(t)$. If the input $\tilde{x}(t)$ to such a system is narrowband, the baseband input/output relationship is given by (2.1) with

$$a_{2k+1} = \frac{1}{2^{2k}} \binom{2k+1}{k} \Phi_{2k+1}(f_o), \tag{2.21}$$

where $\Phi_k(f)$ is the Fourier transform of $\phi_k(\tau)$. Note that even though $\phi_k(\tau)$ is real-valued, $\Phi_k(f_o)$ is not real-valued in general and hence the AM/PM conversion is non-constant. However, special cases of the diagonal kernel system exist where the AM/PM conversion is constant.

2.3.2 Hammerstein System

If a memoryless nonlinearity is followed by a linear time-invariant (LTI) system, the overall system is called a Hammerstein system. The Hammerstein system is a special case of the diagonal kernel system, where $\phi_k(\tau) = c_k \phi(\tau)$. Substituting this kernel relationship into (2.20), we obtain

$$\begin{aligned} y(t) &= \int \phi(\tau_1) \sum_k c_k x^k(t - \tau_1)\ d\tau_1 \\ &= \phi(t) \star u(t), \quad u(t) = \sum_k c_k x^k(t), \end{aligned} \tag{2.22}$$

where $\star$ denotes convolution. The above equation describes that $x(t)$ first goes through a memoryless NL and then an LTI system to yield $y(t)$, and thus the overall system from $x(t)$ to $y(t)$ is Hammerstein. If the input $x(t)$ to the Hammerstein system is narrowband, the baseband input/output relationship is given by (2.1) with

$$a_{2k+1} = \frac{1}{2^{2k}} \binom{2k+1}{k} c_{2k+1} \Phi(f_o), \tag{2.23}$$

where $\Phi(f)$ is the Fourier transform of $\phi(\tau)$. Since for the Hammerstein system, $h_k(\boldsymbol{\tau}_k) = c_k\phi(\tau_1)\prod_{i=1}^{k}\delta(\tau_i-\tau_1)$ and $\{c_k\}$ are real-valued, a_{2k+1} in (2.23) has $\angle a_{2k+1} = \angle\Phi(f_o)$, which does not depend on k. As a result, the AM/PM conversion is constant.

2.3.3 Separable Kernel System

Consider the case where the kth-order Volterra kernel can be written as a product of k one-dimensional functions,

$$h_k(\boldsymbol{\tau}_k) = \prod_{i=1}^{k}\psi_k(\tau_i), \tag{2.24}$$

where each $\psi_k(\cdot)$ is real-valued and one-dimensional. This system is referred to as a separable kernel system. Substituting (2.24) in (2.4) and simplifying, the bandpass input/output relationship is

$$y(t) = \sum_k \left(\int \psi_k(\tau)x(t-\tau)d\tau\right)^k. \tag{2.25}$$

If the input $x(t)$ to such a system is narrowband, the baseband input/output relationship of the separable kernel system is given by (2.1) with

$$a_{2k+1} = \frac{1}{2^{2k}}\binom{2k+1}{k}\left|\Psi_{2k+1}\left(f_o\right)\right|^{2k}\Psi_{2k+1}\left(f_o\right), \tag{2.26}$$

where $\Psi_k\left(f\right)$ is the Fourier transform of $\psi_k(\tau)$. Note that although $\psi_k(\tau)$ is real-valued, $\Psi_k(f_o)$ is not real-valued in general, and hence the AM/PM conversion is non-constant. However, special cases of the separable kernel system exist where the AM/PM conversion is constant.

2.3.4 Wiener System

If an LTI system is followed by a memoryless nonlinearity, the overall system is called a Wiener system. The Wiener system is a special case of the separable kernel system, where $\psi_k(\tau) = \sqrt[k]{c_k}\psi(\tau)$. As such,

$$y(t) = \sum_k c_k\left(\int \psi(\tau)x(t-\tau)d\tau\right)^k. \tag{2.27}$$

Since $u(t) = \psi(t) \star x(t)$, $y(t) = \sum_k c_k u^k(t)$, the system that transforms $x(t)$ into $y(t)$ is Wiener. The kth-order Volterra kernel of the Wiener system is given by:

$$h_k(\boldsymbol{\tau}_k) = c_k \prod_{i=1}^{k} \psi(\tau_i), \tag{2.28}$$

where c_k and $\psi(\cdot)$ are real-valued. If the input $x(t)$ to the Wiener system is narrowband, the baseband input/output relationship is given by (2.1) with

$$a_{2k+1} = \frac{1}{2^{2k}} \binom{2k+1}{k} c_{2k+1} \left|\Psi(f_o)\right|^{2k} \Psi(f_o), \tag{2.29}$$

where $\Psi(f)$ is the Fourier transform of $\psi(\tau)$. Since $\angle a_{2k+1} = \angle \Psi(f_o)$ is independent of k, the resulting AM/PM conversion is constant.

2.4 Conclusion

Volterra series is a general nonlinear model with memory. In this chapter, we first describe the *baseband* input/output relationship of a Volterra nonlinear PA whose input is a bandpass communication signal. When a *physical* PA exhibits short-term memory effects, it is said to be quasi-memoryless, because the corresponding *baseband* input/output relationship is memoryless. This is a potential point of confusion upon examining different literatures. In general, the so-called AM/AM and AM/PM conversions cannot fully describe the PA, but they can if the PA is quasi-memoryless. On the other hand, if the physical PA is (strictly-) memoryless, the output is simply a polynomial function of the input, and the resulting AM/PM conversion is constant. Thus, AM/AM conversion alone can fully characterize a (strictly-) memoryless PA. This however, is not the only case where AM/PM conversion is absent. We show that if a narrowband signal goes through a Wiener or a Hammerstein nonlinear system, the AM/PM conversion is also constant. It is important to distinguish the long-memory, short-memory, and memoryless nonlinear cases, since they affect our choice of the linearization strategy.

CHAPTER III

ON BASEBAND REPRESENTATION OF PASSBAND NONLINEARITIES [1]

Modeling, analysis and compensation of nonlinearities in the transmitter (power amplifier in particular) has attracted a lot of attention recently. Given the same passband polynomial or Volterra representation of the nonlinear system, two different baseband formulations have appeared in the literature. The purpose of this chapter is to address the discrepancy between the two and to affirm that proper conjugation must be applied in the baseband representation of passband nonlinearities. Experimental validation is also provided.

3.1 Introduction

The topic of RF/microwave device nonlinearity [72], especially power amplifier (PA) nonlinearity [35], has attracted a lot of attention recently. This is because the modern, more spectrally efficient transmission formats (such as OFDM and CDMA) tend to have large peak-to-average power ratios but analog devices are linear only over a limited input amplitude range. All PAs exhibit varying degrees of nonlinearity, and the mildly nonlinear class AB PAs are often used in digital communication systems. In order for the PA to deliver a reasonable amount of average power, either the signal has to experience some nonlinearity thus creating in-band distortion as well as adjacent channel interference, or the PA has to be linearized. Digital baseband predistortion is a promising approach to PA linearization.

For quasi-memoryless nonlinearities, polynomials have been used extensively to model the PA [35, 50, 95, 123, 128, 129], or to construct the predistorter [39, 81, 112, 113]. Denote by $x(t)$ the baseband equivalent input to the device and by $z(t)$ the baseband equivalent output of the device. Two different versions of the baseband polynomial model exist.

[1] This chapter was published in [130] and is a result of joint work with Hua Qian, Lei Ding, and G. Tong Zhou.

Version 1 (for example, eq. (2.225) of [18], eq. (5) of [50], eq. (5) of [81], eq. (3) of [112], eq. (1) of [113]):

$$z(t) = x(t) \sum_{k=0}^{K} a_{2k+1} |x(t)|^{2k} \tag{3.1}$$

$$= \sum_{k=0}^{K} a_{2k+1} \, [x(t)]^{k+1} [x^*(t)]^k. \tag{3.2}$$

Version 2 (for example, eq. (7) of [50], eq. (2.7) of [123], eq. (6) of [128]):

$$z(t) = \sum_{k=0}^{K} a_{2k+1} \, [x(t)]^{2k+1}. \tag{3.3}$$

Since the baseband quantities are complex valued in general, equations (3.1) and (3.3) are fundamentally different. Using these different models to characterize or linearize the devices, or to predict spectral regrowth [50, 112, 113, 128, 129], conflicting results may be produced.

Similar discrepancy also arises in the treatment of (baseband) Volterra series, [18,47,134], which is a general nonlinear model with memory. The purpose of this chapter is to provide both theoretical and experimental justifications for (3.1)-(3.2) and to emphasize the use of proper conjugations in baseband formulations.

3.2 Passband and baseband representations

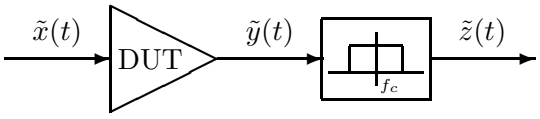


Figure 3.1: Passband configuration.

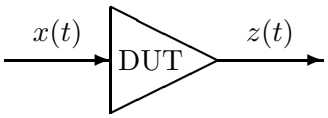


Figure 3.2: Baseband configuration.

In Fig. 3.1, we denote by $\tilde{x}(t)$, the passband input to the device under test (DUT), by $\tilde{y}(t)$, the passband output of the DUT before the zonal filter, and by $\tilde{z}(t)$, the passband output of the DUT after the zonal filter. The zonal filter [18, p. 68] only allows

frequency contents centered around f_c to pass un-altered, whereas frequency contents appearing around the dc, $2f_c$, $3f_c$ etc, are eliminated.

We shall first present the Volterra series, and then specialize to the memoryless nonlinear cases.

In [72, Ch. 4], the Volterra series is proposed to model nonlinear RF/microwave devices. This means for Fig. 3.1, we can relate $\tilde{x}(t)$ and $\tilde{y}(t)$ by [18, 72],

$$\tilde{y}(t) = \sum_{\ell=1}^{L} \int \cdots \int \tilde{h}_\ell(\boldsymbol{\tau}_\ell) \prod_{i=1}^{\ell} \tilde{x}(t-\tau_i) d\boldsymbol{\tau}_\ell, \tag{3.4}$$

where $\boldsymbol{\tau}_\ell = [\tau_1, \ldots, \tau_\ell]$, $\tilde{h}_\ell(\cdot)$ is the ℓth-order Volterra kernel, and $d\boldsymbol{\tau}_\ell = d\tau_1 d\tau_2 \cdots d\tau_\ell$. Note that both even and odd ℓ's are allowed in (3.4), and all quantities in (3.4) are real-valued. The strictly memoryless case corresponds to $\tilde{h}_\ell(\boldsymbol{\tau}_\ell) = \tilde{a}_\ell \ \delta(\boldsymbol{\tau}_\ell)$, in which case (3.4) reduces to the familiar polynomial (or power series) model [35, p. 181], [72, p. 156]

$$\tilde{y}(t) = \sum_{\ell=1}^{L} \tilde{a}_\ell \ [\tilde{x}(t)]^\ell. \tag{3.5}$$

Note that no conjugation appears in (3.5) since all quantities involved are real-valued.

The baseband relationship that describes Fig. 3.2 is [18]:

$$\begin{aligned} z(t) = \sum_{k=0}^{K} \int \cdots \int h_{2k+1}(\boldsymbol{\tau}_{2k+1}) \times \\ \prod_{i=1}^{k+1} x(t-\tau_i) \prod_{i=k+2}^{2k+1} x^*(t-\tau_i) d\boldsymbol{\tau}_{2k+1}, \end{aligned} \tag{3.6}$$

where $*$ denotes conjugation, and

$$\begin{aligned} h_{2k+1}(\boldsymbol{\tau}_{2k+1}) = \frac{1}{2^{2k}} \binom{2k+1}{k} \tilde{h}_{2k+1}(\boldsymbol{\tau}_{2k+1}) \times \\ e^{-j2\pi f_c \left(\sum_{i=1}^{k+1} \tau_i - \sum_{i=k+2}^{2k+1} \tau_i \right)}. \end{aligned} \tag{3.7}$$

The relationship between the L in (3.5) and the K in (3.6) is given by

$$2K+1 = \begin{cases} L, & \text{if } L \text{ is odd,} \\ L-1, & \text{if } L \text{ is even.} \end{cases}$$

We emphasize that the baseband expression (3.6) is between $x(t)$ and $z(t)$ (c.f. Fig. 3.2) but the passband relationship (3.4) is between $\tilde{x}(t)$ and $\tilde{y}(t)$ (c.f. Fig. 3.1). The zonal

filter represents the frequency selectivity of the DUT and the antenna and is absorbed into the DUT block in Fig. 3.2. Moreover, notice that there are $k+1$ un-conjugated copies of x but k conjugated copies of x in (3.6). Volterra predistorters are proposed in [47, 134]; unfortunately, the above conjugation pattern is not always expressed correctly in the baseband Volterra models.

If the system is strictly memoryless, the corresponding baseband expression is (3.1) with

$$a_{2k+1} = \frac{1}{2^{2k}} \binom{2k+1}{k} \tilde{a}_{2k+1}, \tag{3.8}$$

obtained by substituting $\tilde{h}_\ell(\boldsymbol{\tau}_\ell) = \tilde{a}_\ell \, \delta(\boldsymbol{\tau}_\ell)$ into (3.6)-(3.7). Since $\tilde{a}_{2k+1}$ is real-valued, so is a_{2k+1} by virtue of (3.8).

When $\tilde{x}(t)$ is narrowband, it is shown in [95] that the baseband counterpart of (3.4) is again (3.1) but with

$$a_{2k+1} = \frac{1}{2^{2k}} \binom{2k+1}{k} \tilde{H}_{2k+1}(\underbrace{f_c, \ldots, f_c}_{k+1}, \underbrace{-f_c, \ldots, -f_c}_{k}), \tag{3.9}$$

where $\tilde{H}_{2k+1}(\cdot)$ on the right hand side of (3.9) is the $(2k+1)$-dimensional Fourier transform of the Volterra kernel $\tilde{h}_{2k+1}(\boldsymbol{\tau}_{2k+1})$ evaluated at the above frequency tuple. From (3.9), it is apparent that a_{2k+1} is a function of the carrier frequency f_c. If $x(t)$ is narrow-band, a_{2k+1} is approximately constant over the signal bandwidth and the model in (3.1) is sufficient. However, if $x(t)$ is wideband, the Volterra model (3.6) is more appropriate [18, 47, 134].

It is well known that (qusi-)memoryless PAs are characterized by their amplitude-to-amplitude ($|x(t)|$ to $|z(t)|$; i.e., AM/AM) and amplitude-to-phase ($|x(t)|$ to $\angle z(t) - \angle x(t)$; i.e., AM/PM) conversions. Indeed, from (3.1), we infer that

$$|z(t)| = |x(t)| \left| \sum_{k=0}^{K} a_{2k+1} |x(t)|^{2k} \right|, \tag{3.10}$$

$$\angle z(t) - \angle x(t) = \angle \sum_{k=0}^{K} a_{2k+1} |x(t)|^{2k}, \tag{3.11}$$

which depend on $|x(t)|$ but not on $\angle x(t)$. Table 3.1 summarizes the nonlinearities discussed so far (N.A. stands for "not applicable").

Table 3.1: Nonlinear systems and their representations

	Strictly Memoryless	**Quasi-Memoryless**	**With Memory**
Passband	eq. (3.5)	eq. (3.4)	eq. (3.4)
Baseband	eq. (3.1)	eq. (3.1)	eq. (3.6)
	a_{2k+1} real	a_{2k+1} complex	
AM/AM	non-constant	non-constant	N.A.
AM/PM	constant	non-constant	N.A.

On the other hand, the baseband model (3.3) without conjugation is problematic. It describes a (quasi-)memoryless system, but $|z(t)|$ and $\angle z(t) - \angle x(t)$ depend on both $|x(t)|$ and $\angle x(t)$, thus contradicting the notion of AM/AM and AM/PM conversions.

3.3 Experimental Results

We have assembled a digital baseband predistortion linearization testbed that consists of a high speed digital I/O system, a digital to analog converter, an upconversion chain, a DUT, a downconversion chain, and an analog-to-digital converter. In this experiment, the DUT is a Siemens CGY0819 handset PA operating at the cellular band (824-849 MHz). The input is a 1.35 MHz bandwidth signal centered at 836 MHz. We constructed two predistorters, one according to (3.1) and the other one according to (3.3). Fig. 3.3 shows the PA output power spectral density (PSD) measured by a spectrum analyzer. The PA is operated near the 1dB compression point, so a significant amount of nonlinearity is present. Line (a) is the PA output PSD without predistortion and shows spectral regrowth (broadening). Line (b) is the PA output PSD when a predistorter (3.1) with $K = 4$ is applied. In this case, approximately 20 dB of spectral regrowth suppression was achieved. In contrast, line (c) is the PA output PSD when a predistorter (3.3) with $K = 4$ is applied: instead of reducing the nonlinear effects, the predistorter aggravated the spectral regrowth.

3.4 Conclusions

In this chapter, we examined bandpass nonlinearities and their ramifications in the baseband. We clarified that in the baseband representation, proper conjugation must be applied. Such conjugation designation ensures that in the quasi-memoryless case, AM/AM

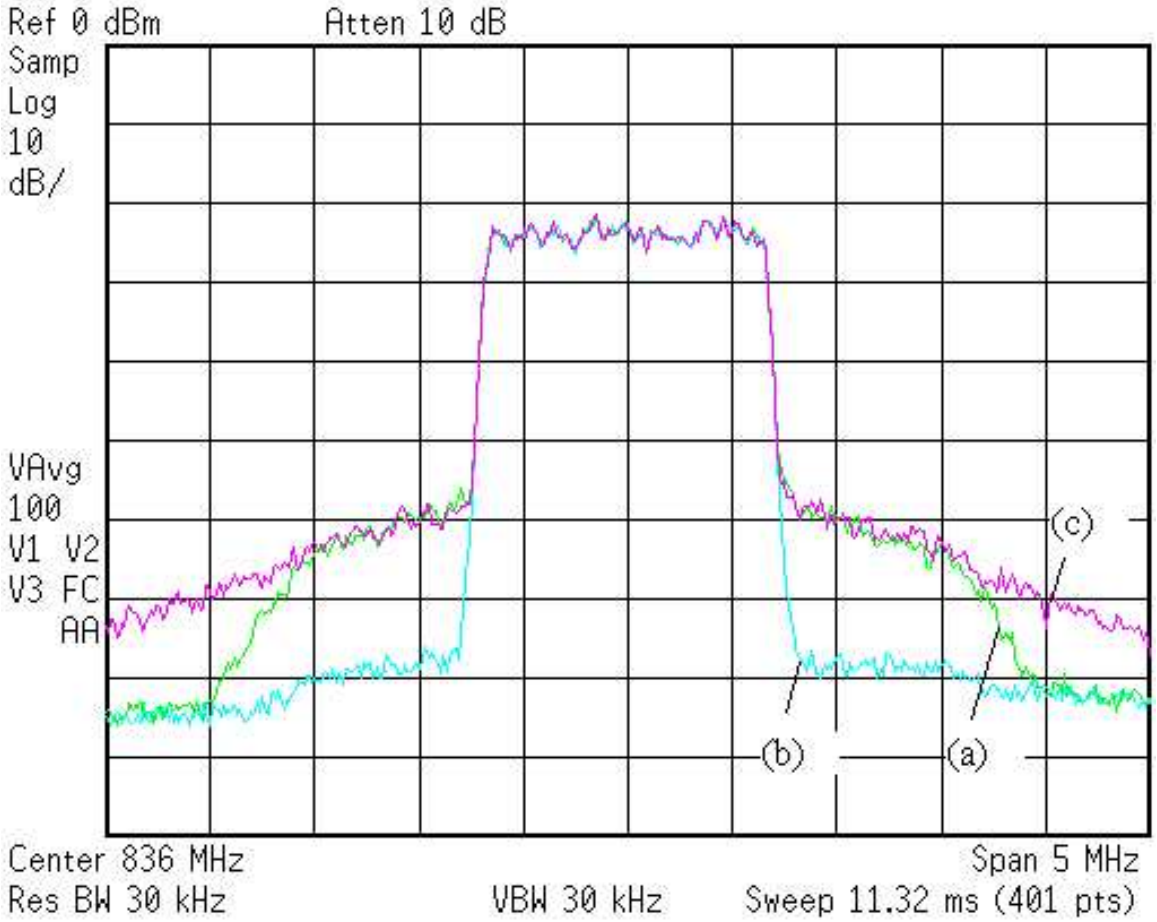


Figure 3.3: Measured PA output PSD: (a) without predistortion; (b) with predistorter (3.1); (c) with predistorter (3.3).

and AM/PM conversions are sufficient to characterize the device. Experimental results support our analysis.

CHAPTER IV

A HAMMERSTEIN PREDISTORTION LINEARIZATION DESIGN BASED ON THE INDIRECT LEARNING ARCHITECTURE[1]

Power amplifiers (PAs) are inherently nonlinear devices and are used in virtually all communications systems. Digital baseband predistortion is a highly cost effective way to linearize the PAs, but most existing architectures assume that the PA has a memoryless nonlinearity. For wider bandwidth applications such as WCDMA, PA memory effects can no longer be ignored, and memoryless predistortion has limited effectiveness. In this chapter, we model the PA as a Wiener system and construct a Hammerstein predistorter, obtained using an indirect learning architecture. Linearization performance is demonstrated on a 3-carrier UMTS signal.

4.1 Introduction

Power amplifiers (PAs) are indispensable components in a communication system and are inherently nonlinear. It is well known that there is an approximate inverse relationship between the PA efficiency and its linearity. Hence, nonlinear PAs are desirable from an efficiency point of view. The price paid for higher efficiency is that nonlinearity causes spectral regrowth (broadening) which leads to adjacent channel interference. It also causes in-band distortion which degrades the bit error rate (BER) performance. Newer transmission formats such as CDMA and OFDM are especially vulnerable to PA nonlinearities, due to their high peak to average power ratio; i.e. large fluctuations in their signal envelopes. In order to comply with spectral masks imposed by regulatory bodies and to reduce BER, PA linearization is necessary.

[1]This chapter was published in [41] and is a result of joint work with Lei Ding and G. Tong Zhou.

Of all linearization techniques, digital baseband predistortion is among the most cost effective. A predistorter is a functional block that precedes the PA. It generally creates an expanding nonlinearity since the PA has a compressing characteristic. Ideally, we would like the PA output to be a scalar multiple of the input to the predistorter-PA chain. For a memoryless PA, (i.e.; the current output depends only on the current input), memoryless predistortion is sufficient. There has been intensive research on memoryless predistortion during the past decade [35].

For wider bandwidth applications such as WCDMA, PA memory effects can no longer be ignored. Moreover, higher power amplifiers such as those used in wireless basestations exhibit memory effects. The cause of memory effects can be electrical or electro-thermal as suggested in [118]. Memoryless predistortion for a PA with memory often results in poor linearization performance. Although Volterra series is a general nonlinear model with memory, its predistortion is complex and its real-time implementation difficult. In [33], Clark *et.al.* used a Wiener model; i.e., a linear time-invariant (LTI) system followed by a memoryless nonlinearity, to capture the nonlinear memory effects in the PA associated with wideband signals. In this chapter, we also adopt the Wiener PA model, which has the advantage that its predistortion can be easily carried out. A Hammerstein system is a memoryless nonlinearity followed by a LTI system, and can therefore linearize a Wiener PA model.

In the current literature, predistorters with memory mainly fall into the data predistorter category [46, 61], in the sense that predistortion is applied before the pulse shaping filter. The main drawback of data predistortion is its dependence on the signal constellation and the pulse shaping filter. Both Volterra model based [46] and Hammerstein model based [61] data predistorters have been proposed. In [46], a Volterra data predistorter is constructed using the indirect learning architecture. In [61], the Hammerstein data predistorter is obtained using a stochastic gradient method.

As opposed to data predistortion, we shall pursue signal predistortion in this chapter; i.e., predistortion occurs after the pulse shaping filter. To construct a Hammerstein predistorter, one approach is to first identify the Wiener PA and then find the Hammerstein

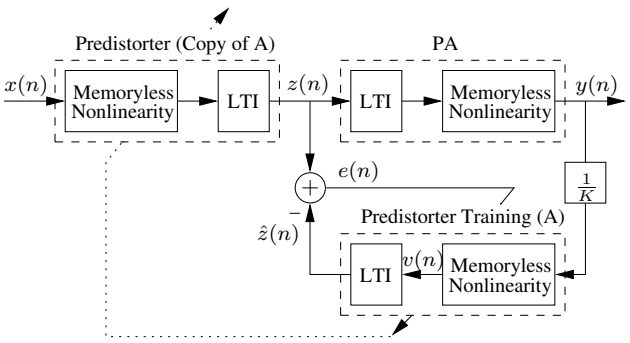


Figure 4.1: The indirect learning architecture for the Hammerstein predistorter.

predistorter as its inverse. Since Wiener system identification is generally more difficult to carry out than Hammerstein system identification, we pursue an alternative approach which generates the Hammerstein predistorter without first identifying the Wiener PA. Unlike [61], our Hammerstein predistorter will be constructed using an indirect learning architecture similar to the one used in [46]. In this setup, finding the predistorter is essentially equivalent to identifying a Hammerstein system.

4.2 Indirect Learning Architecture

Fig. 4.1 shows the indirect learning structure that is used for Hammerstein predistorter identification. The PA has a Wiener structure (LTI followed by memoryless nonlinearity). The feedback path labeled "Predistorter Training" (block A) has a Hammerstein structure if we view $y(n)/K$ as its input and $\hat{z}(n)$ as its output. The actual predistorter is an exact copy of the feedback path (copy of A); it has $x(n)$ as its input and $z(n)$ as its output. Ideally, we would like $y(n) = Kx(n)$, which renders $z(n) = \hat{z}(n)$ and the error term $e(n) = 0$. Given $y(n)$ and $z(n)$, our task is to find the parameters of block A, which yields the predistorter. The algorithm converges when the error energy $||e(n)||^2$ is minimized.

Here we consider that the PA characteristics do not change rapidly with time; changes in PA characteristics are often due to temperature drift, aging etc which have long time constants. After gathering a block of $y(n)$ and $z(n)$ data samples, the training branch (block A) can process the data off-line, which lowers the requirement of the processing power of the predistortion system. Once the predistorter identification algorithm has converged, the

new set of parameters are plugged into the high speed predistorter, which can be readily implemented by application-specific integrated circuits (ASIC) or field programmable gate arrays (FPGA). When the predistorter coefficients have been found and it is believed that the PA characteristics are hardly changing, the setup in Fig. 4.1 can be run in open loop; i.e., we temporarily shut down the training branch, until changes in PA characteristics require a new predistorter.

4.3 Identification of The Hammerstein Predistorter

The predistorter training branch can be described by:

$$v(n) = \sum_{k=0}^{(K-1)/2} c_{2k+1}\, y(n)|y(n)|^{2k}, \tag{4.1}$$

$$z(n) = \sum_{p=1}^{P} a_p z(n-p) + \sum_{q=0}^{Q} b_q v(n), \tag{4.2}$$

which implies that for the predistorter, we model the memoryless nonlinearity as an odd-order polynomial and the LTI system as a general pole/zero system. Combining the two equations above, we obtain

$$z(n) = \sum_{p=1}^{P} a_p z(n-p) + \sum_{q=0}^{Q} b_q \left(\sum_{k=0}^{(K-1)/2} c_{2k+1} y(n-q)|y(n-q)|^{2k} \right). \tag{4.3}$$

Given $y(n)$ and $z(n)$, our objective is to estimate the a_p, b_q and c_{2k+1} coefficients. Parameter estimation of this model is a classical Hammerstein system identification problem. If no additional assumptions are made on the system's input signal $y(n)$, iterative Newton and Narendra-Gallman algorithms are the two most popular iterative estimation methods [45]. The two algorithms exhibit similar performance as shown in [45]. The main drawback of these algorithms is that they are sensitive to the initial guesses and may converge to a local minimum. A recent method proposed by Bai [9] uses an optimal two stage identification algorithm, which can lead to a global optimum. The model structure introduced in [9] is a Hammerstein system followed by a memoryless nonlinearity. However, we can easily modify the results of [9] to suit our model. Note that for a given set of $\{y(n), z(n)\}$ values,

the b_q's and the c_{2k+1}'s are not unique (i.e.; multiplying b_q with a constant and dividing c_{2k+1} by the same constant yields the same model). To avoid this problem, we assume that $\sum_{q=0}^{Q} |b_q|^2 = 1$ and the real part of b_0 is positive as suggested in [9].

Next, we will review the Narendra-Gallman (NG) and the optimal two stage identification (LS/SVD) algorithms.

4.3.1 Narendra-Gallman algorithm

The NG algorithm starts with initial guesses for the a_p and b_q coefficients, denoted by $a_p^{(0)}$ and $b_q^{(0)}$, respectively. At the ith iteration (4.3) can be rewritten as

$$z(n) - \sum_{p=1}^{P} a_p^{(i)} z(n-p) = \sum_{k=0}^{(K-1)/2} c_{2k+1} u_{2k+1}(n) \tag{4.4}$$

$$u_{2k+1}(n) = \sum_{q=0}^{Q} b_q^{(i)} y(n-q)|y(n-q)|^{2k}.$$

At this stage, our objective is to solve for c_{2k+1}. Using matrix notation we can reformulate (4.4) as

$$\mathbf{z}_0 - \mathbf{Z}\mathbf{a}^{(i)} = \mathbf{U}\mathbf{c}, \tag{4.5}$$

where $\mathbf{Z} = [\mathbf{z}_1, \ldots, \mathbf{z}_P]$, $\mathbf{z}_l = [\mathbf{0}_l^T, z(1), \ldots, z(N-l)]^T$, where $\mathbf{0}_l$ is a $l \times 1$ all-zero vector, $\mathbf{a}^{(i)} = [a_1^{(i)}, \cdots, a_P^{(i)}]^T$, $\mathbf{U} = [\mathbf{u}_1, \cdots, \mathbf{u}_K]$, $\mathbf{u}_{2k+1} = [u_{2k+1}(1), \cdots, u_{2k+1}(N)]^T$, and $\mathbf{c} = [c_1, \cdots, c_K]^T$. The least-squares solution for (4.5) is

$$\hat{\mathbf{c}}^{(i+1)} = (\mathbf{U}^H\mathbf{U})^{-1}\mathbf{U}^H\left(\mathbf{z}_0 - \mathbf{Z}\mathbf{a}^{(i)}\right), \tag{4.6}$$

where ^{H} denotes Hermitian transpose. In the second step, based on the $c_{2k+1}^{(i+1)}$'s obtained, we rewrite (4.3) as,

$$\mathbf{z}_0 = \mathbf{Z}\mathbf{a} + \mathbf{V}\mathbf{b} = [\mathbf{Z}\ \mathbf{V}]\begin{bmatrix}\mathbf{a}\\ \mathbf{b}\end{bmatrix}, \tag{4.7}$$

where $\mathbf{V} = [\mathbf{v}_0\mathbf{v}_1, \cdots, \mathbf{v}_Q]$, $\mathbf{v}_l = [\mathbf{0}_l^T, v(1), \cdots, v(N-l)]^T$, $\mathbf{b} = [b_0, \cdots, b_Q]^T$, and $v(n)$ is given in (4.1). The least-squares solution for (4.7) is,

$$\begin{bmatrix}\hat{\mathbf{a}}^{(i+1)}\\ \hat{\mathbf{b}}^{(i+1)}\end{bmatrix} = \left([\mathbf{Z}\ \mathbf{V}]^H[\mathbf{Z}\ \mathbf{V}]\right)^{-1}[\mathbf{Z}\ \mathbf{V}]^H\mathbf{z}_0, \tag{4.8}$$

With the new $\hat{\mathbf{a}}^{(i+1)}$ and $\hat{\mathbf{b}}^{(i+1)}$ estimates, we can go back to the first step and continue until the algorithm converges.

4.3.2 Optimal two stage identification algorithm

Since the difficulty in estimating the b_q's and c_{2k+1}'s is that they appear together as the coefficient on the r.h.s. of (4.3), if we define

$$d_{q,2k+1} = b_q c_{2k+1}, \tag{4.9}$$

we can first estimate $d_{q,2k+1}$ using least-squares and then find b_q and c_{2k+1} from $d_{q,2k+1}$. Substituting (4.9) into (4.3), we obtain

$$\begin{aligned} z(n) = &\sum_{p=1}^{P} a_p z(n-p) \\ &+ \sum_{q=0}^{Q} \sum_{k=0}^{(K-1)/2} d_{q,2k+1}\ g_{q,2k+1}(n), \end{aligned} \tag{4.10}$$

where $g_{q,2k+1}(n) = y(n-q)|y(n-q)|^{2k}$. Rewriting in a matrix form, we obtain

$$\mathbf{z}_0 = \mathbf{Z}\mathbf{a} + \mathbf{G}\mathbf{d} = [\mathbf{Z}\ \mathbf{G}] \begin{bmatrix} \mathbf{a} \\ \mathbf{d} \end{bmatrix}, \tag{4.11}$$

where $\mathbf{G} = [\mathbf{g}_{01}, \cdots, \mathbf{g}_{0K}, \cdots, \mathbf{g}_{Q1}, \cdots, \mathbf{g}_{QK}]$, $\mathbf{g}_{q,2k+1} = [g_{q,2k+1}(1), \cdots, g_{q,2k+1}(N)]^T$, and $\mathbf{d} = [d_{01}, \cdots, d_{0K}, \cdots, d_{Q1}, \cdots, d_{QK}]^T$. The least-squares solution for (4.11) is

$$\begin{bmatrix} \hat{\mathbf{a}} \\ \hat{\mathbf{d}} \end{bmatrix} = \left([\mathbf{Z}\ \mathbf{G}]^H [\mathbf{Z}\ \mathbf{G}]\right)^{-1} [\mathbf{Z}\ \mathbf{G}]^H \mathbf{z}_0, \tag{4.12}$$

Equation (4.9) can be alternatively expressed as

$$\mathbf{D} = \begin{bmatrix} d_{01} & d_{03} & \cdots & d_{0K} \\ d_{11} & d_{13} & \cdots & d_{1K} \\ \vdots & \vdots & & \vdots \\ d_{Q1} & d_{Q3} & \cdots & d_{QK} \end{bmatrix} = \mathbf{b}\mathbf{c}^T,$$

where $\mathbf{b} = [b_0, \ldots, b_Q]^T$, $\mathbf{c} = [c_1, \ldots, c_K]^T$. Since the matrix $\mathbf{D}$ has rank one, a natural way to estimate $\hat{\mathbf{b}}$ and $\hat{\mathbf{c}}$ from $\hat{\mathbf{D}}$ is to perform a singular value decomposition (SVD) on $\hat{\mathbf{D}}$ and

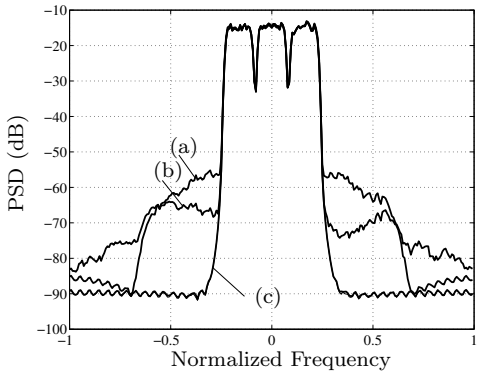


Figure 4.2: Comparison of the PSDs for pole/zero Wiener PA and pole/zero Hammerstein predistorter. (a) Output without predistortion; (b) Output with memoryless predistortion; (c) Output with Hammerstein predistortion, NG and LS/SVD algorithms (similar performance).

then find the eigenvectors corresponding to the largest singular value. Let the SVD of $\hat{\mathbf{D}}$ be given by,

$$\hat{\mathbf{D}} = \sum_{i=1}^{\min[(Q+1),(K+1)/2]} \sigma_i \boldsymbol{\mu}_i \boldsymbol{\nu}_i^H, \tag{4.13}$$

where $\boldsymbol{\mu}_i$ and $\boldsymbol{\nu}_i$ are $Q+1$ and $(K+1)/2$ dimensional orthonormal vectors, respectively. Then $\hat{\mathbf{b}}$ and $\hat{\mathbf{c}}$ can be estimated as

$$\hat{\mathbf{b}} = s_\mu \boldsymbol{\mu}_1, \quad \hat{\mathbf{c}} = s_\mu \sigma_1 \boldsymbol{\nu}_1^*, \tag{4.14}$$

where * denotes complex conjugate and s_μ is the first non-zero element of $\boldsymbol{\mu}_1$. These estimates can be shown to be the closest $\hat{\mathbf{b}}$ and $\hat{\mathbf{c}}$ to $\hat{\mathbf{D}}$ in the least-squares sense [9].

In summary, the NG algorithm is a simple and robust algorithm. Although it may have convergence problems, it can perform well in many cases as will be shown in the next section. The LS/SVD algorithm avoids the potential local minimum problem of the NG algorithm. However, using SVD to find the b_q's and c_{2k+1}'s may not result in the best b_q's and c_{2k+1}'s that minimize the squared error criterion. Our examples in the next section will show that both work well for identifying the Hammerstein predistorter although one may outperform the other in a particular scenario.

4.4 Simulations

In this section, we illustrate through computer simulations the performance of the Hammerstein predistorter identified using the indirect learning architecture. In the first example, the LTI portion of the Wiener PA model has a pole/zero form, whose system function is given by

$$H(z) = \frac{1 + 0.3z^{-2}}{1 - 0.2z^{-1}}. \tag{4.15}$$

For the memoryless nonlinear portion of the Wiener PA model, we use a 5th order nonlinearity with coefficients,

$$\begin{aligned} c_1 &= 14.9740 + 0.0519j, \quad c_3 = -23.0954 + 4.9680j, \\ c_5 &= 21.3936 + 0.4305j, \end{aligned} \tag{4.16}$$

which were extracted from an actual Class AB PA.

The baseband input signal is a 3-carrier Universal Mobile Telecommunications System (UMTS) signal. Hammerstein predistorter identification is carried out based on 8000 data samples. The predistorter parameters usually converge after a few iterations. Next, we compare the spectra of the input and output signals to asses the effectiveness of the predistorter in reducing spectral regrowth. In this example, we assume that the LTI portion of the Hammerstein predistorter is a pole/zero system with two poles and one zero (correct model orders for the inverse of the $H(z)$ of (4.15)). In addition, we make the assumption that the nonlinearity of the predistorter is 5th order.

Performance of predistorter identified with the LS/SVD and NG algorithms is demonstrated in Fig. 4.2. Both algorithms fully suppress the spectral regrowth exhibited by the PA output when no predistortion is applied. In contrast, we observe in Fig. 4.2 that 5th order memoryless predistortion does not fully suppress the spectral regrowth.

In the second example, the LTI portion of the Wiener PA is $H(z) = 1 + 0.3z^{-2}$ (FIR), and the LTI portion of the Hammerstein predistorter is assumed to be FIR as well. Our objective here is to see whether the algorithm can correctly identify an FIR filter that approximates the inverse of the FIR system in the PA. We assume that the FIR system in the predistorter has 15 taps. The results are shown in Fig. 4.3. The two algorithms

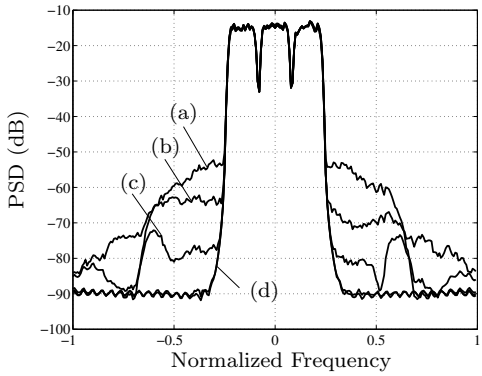


Figure 4.3: Comparison of the PSDs for FIR Wiener PA and 15-tap FIR Hammerstein predistorter. (a) Output without predistortion; (b) Output with memoryless predistortion; (c) Output with Hammerstein predistortion (NG); (d) Output with Hammerstein predistortion (LS/SVD).

exhibit different behaviors this time: the NG algorithm performs worse than the LS/SVD algorithm. When examining the concatenated response of the two LTI blocks (one from the Wiener PA and the other from the Hammerstein predistorter), we observe that the predistorter's LTI system identified by the NG algorithm can only compensate for the PA's LTI system within the signal bandwidth. However, the LS/SVD algorithm is able to find a good FIR system for the predistorter, both within and outside of the signal bandwidth.

In the third example, we perturbed the Wiener PA model coefficients so it is a full Volterra model (not Wiener any more). Our objective is to see whether the Hammerstein predistorter has any robustness. The result is shown in Fig. 4.4. We still observe significant reduction of spectral regrowth with the Hammerstein predistorter.

In all cases, memoryless predistortion is not very effective in suppressing spectral regrowth, which underscores the notion that PA memory effects must be taken into account when designing the predistorter.

4.5 Conclusions

We employed the indirect learning structure to identify the Hammerstein predistorter for a PA modeled by a Wiener model. We compared the performance of two Hammerstein system identification algorithms; i.e., the NG and LS/SVD algorithms, in this context.

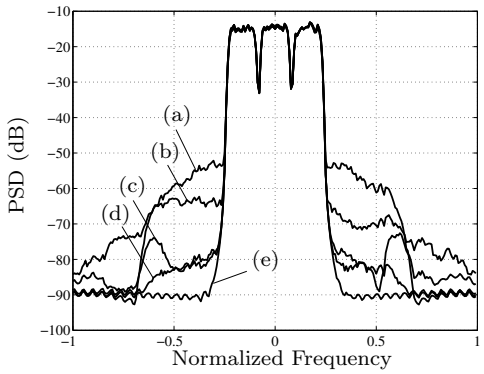


Figure 4.4: Comparison of the PSDs for full Volterra PA and 15-tap FIR Hammerstein predistorter. (a) Output without predistortion; (b) Output with memoryless predistortion; (c) Output with Hammerstein predistortion (NG); (d) Output with Hammerstein predistortion (LS/SVD); (e) Input signal.

For a Wiener model with a simple pole/zero LTI structure, both algorithms show similar performance. However, when the LTI portion of the Wiener PA as well as that of the Hammerstein predistorter are FIR, the LS/SVD algorithm outperforms the NG algorithm. Simulation results illustrate the effectiveness of the proposed predistorter design.

Acknowledgment: This work was supported in part by National Science Foundation grant MIP 9703312 and by the State of Georgia's Yamacraw Initiative. The authors would also like to thank Dr. Zhengxiang Ma and Dr. Dennis R. Morgan of Bell Laboratories, Lucent Technologies for inspiring this work.

CHAPTER V

ORTHOGONAL POLYNOMIALS FOR POWER AMPLIFIER MODELING AND PREDISTORTER DESIGN[1]

The polynomial model is commonly used in power amplifier (PA) modeling and predistorter design. However, the conventional polynomial model exhibits numerical instabilities when higher order terms are included. In this chapter, we introduce a novel set of orthogonal polynomials, which can be used for PA as well as predistorter modeling. Theoretically, the conventional and the orthogonal polynomial models are "equivalent" and thus should behave similarly. In practice, however, the two approaches can perform quite differently in the presence of quantization noise and with finite precision processing. Simulation results show that the orthogonal polynomials can alleviate the numerical instability problem associated with the conventional polynomials and generally yield better PA modeling accuracy as well as predistortion linearization performance.

5.1 Introduction

Power amplifier (PA) is a major source of nonlinearity in a communication system. To increase efficiency, PAs are often driven into their nonlinear region, thus causing spectral regrowth (broadening) as well as in-band distortion. PA linearization is often necessary to suppress spectral regrowth, contain adjacent channel interference, and to reduce bit error rate (BER).

The power series model, or the polynomial model, is widely used in the literature to describe nonlinear effects in the PA (see e.g., [36], [64]). In [129], it is shown that after extracting the polynomial coefficients of the PA, it is then possible to predict spectral

[1]This chapter was published in [91, 93] and is a result of joint work with Hua Qian and G. Tong Zhou.

regrowth of digitally modulated signals using the concept of cumulant. Recently in [65], a memory polynomial model is proposed to fit nonlinear PAs with memory. The more general Volterra series (which is polynomial in nature) has also been used to model nonlinear devices with memory [72].

If the nonlinear PA is used to transmit non-constant modulus signals, PA linearization is often necessary. Among all linearization techniques, digital baseband predistortion is one of the most cost effective. A predistorter, which (ideally) has the inverse characteristic of the PA, is used to compensate for the nonlinearity in the PA. To linearize a memoryless nonlinear PA, one can pursue lookup table (LUT) based or model based approaches. The LUT approach is easy to implement but may take a relatively long time to converge. Moreover, the piece-wise linear curve has a zig-zag appearance which may introduce additional nonlinearities that degrade the performance [68]. As for model based approaches, the polynomial model is a common choice due to its simplicity and ease of implementation [36, Sec. 3.3], [48]. Volterra series [47] and certain special cases of the Volterra series, for example, the Hammerstein model [41] and the memory polynomial model [38], have been proposed for predistorter design that includes memory effects.

Higher-order polynomials present a challenge for both PA modeling and predistorter design. As we show in Section 2, in the process of solving for the model coefficients, a matrix inversion is needed which can cause a numerical instability problem if higher-order polynomial terms are included. The objective of this chapter is to derive a set of orthogonal polynomial basis and to model the PA or the predistorter using such basis functions. The resulting orthogonal polynomial model coefficients can be extracted with much improved numerical stability.

To the best of our knowledge, [78, 79] are the only published results on orthogonal polynomials for predistorter design. Our approach is different and has the following advantages: (i) Our orthogonal polynomial basis functions are expressed in closed form (non-iterative), and their coefficients $\{U_{lk}\}$ are free of round-off errors. (ii) Our basis functions are predetermined and can be implemented with little demand on the computation resources. In [78, 79], the basis functions are calculated online and iteratively, thus requiring much

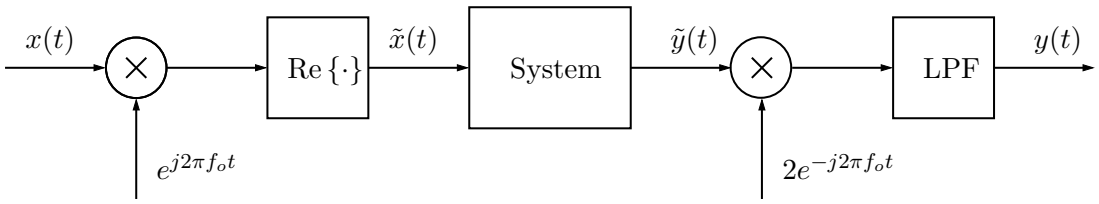


Figure 5.1: Block diagram of a bandpass communication system.

more computational power. (iii) Our basis set consists of both even and odd-order terms whereas that of [78, 79] allows odd-powered series only. Moreover, our basis function expressions are for generally complex-valued baseband data; their application to nonlinear systems with memory is also prescribed.

In Section 5.2, we first introduce the conventional polynomial model and point out its deficiencies. Next, we derive novel orthogonal polynomial basis functions and illustrate their benefit in PA modeling. In Section 5.3, we formulate a predistortion linearization algorithm with orthogonal polynomials. Numerical examples are presented alongside theoretical analysis. Finally, conclusions are drawn in Section 5.4.

5.2 The polynomial model

5.2.1 The conventional polynomial model

In Fig. 5.1, we denote by $\tilde{x}(t)$ the passband input to a nonlinear system (e.g., a power amplifier or a predistorter), and by $\tilde{y}(t)$ the corresponding passband output. If the nonlinear system obeys the polynomial model,

$$\tilde{y}(t) = \sum_{k=1}^{K} \tilde{b}_k \; \tilde{x}^k(t) \tag{5.1}$$

then it can be shown that the corresponding baseband input $x(t)$ and the baseband output $y(t)$ are related by [18, p. 69]

$$y(t) = \sum_{\substack{k=1 \\ k \text{ odd}}}^{K} b_k \; |x(t)|^{k-1} x(t), \tag{5.2}$$

where

$$b_k = 2^{1-k} \binom{k}{\frac{k-1}{2}} \tilde{b}_k. \tag{5.3}$$

Since (5.1) is a physical (passband) model, $\tilde{b}_k$ is real-valued; so is b_k according to (5.3). In reality, however, model (5.1) is not exact. As shown in [40], by including even-order nonlinear terms in (5.2); i.e.,

$$y(t) = \sum_{k=1}^{K} b_k \ |x(t)|^{k-1} x(t), \tag{5.4}$$

modeling accuracy can be improved. Moreover, as shown in [95], when the nonlinear system exhibits short-term memory effects (i.e., the PA is quasi-memoryless), the baseband model (5.4) is still applicable but b_k is now generally complex-valued.

Let us define $\phi_k(x) = |x|^{k-1}x$. Eq. (5.4) then becomes

$$y(t) = \sum_{k=1}^{K} b_k \ \phi_k(x(t)). \tag{5.5}$$

Given the PA input $x(t)$ and PA output $y(t)$ measurements, we would like to extract the PA parameters $\{b_k\}$. Define the N-by-1 input data vector

$$\mathbf{x} = [x(t_1), \ldots, x(t_N)]^T,$$

the N-by-1 output data vector

$$\mathbf{y} = [y(t_1), \ldots, y(t_N)]^T,$$

and the K-by-1 parameter vector

$$\mathbf{b} = [b_1, \ b_2, \ldots, b_K]^T.$$

Next define

$$\boldsymbol{\phi}_k(\mathbf{x}) = [\phi_k(x(t_1)), \ldots, \phi_k(x(t_N))]^T,$$

and the N-by-K matrix

$$\boldsymbol{\Phi} = [\boldsymbol{\phi}_1(\mathbf{x}) \ \ \boldsymbol{\phi}_2(\mathbf{x}) \ \ \ldots \ \ \boldsymbol{\phi}_K(\mathbf{x})].$$

We can now represent (5.5) as

$$\mathbf{y} = \boldsymbol{\Phi}\mathbf{b}. \tag{5.6}$$

The least squares (LS) solution for $\mathbf{b}$ is

$$\mathbf{b}_{LS} = (\boldsymbol{\Phi}^H \boldsymbol{\Phi})^{-1} \boldsymbol{\Phi}^H \mathbf{y}. \tag{5.7}$$

The inversion of the K-by-K matrix $\mathbf{\Phi}^H\mathbf{\Phi}$ in (5.7) can experience a numerical instability problem. To understand the problem, let us first examine its expected value $E[\mathbf{\Phi}^H\mathbf{\Phi}]$.

First, notice that $\phi_k^*(x)\phi_l(x) = |x|^{k+l} = r^{k+l}$, where $r = |x|$. For a stationary random process $x(t)$, we infer that $E[\boldsymbol{\phi}_k^H(\mathbf{x})\boldsymbol{\phi}_l(\mathbf{x})] = NE[r^{k+l}]$. Therefore, $E[\mathbf{\Phi}^H\mathbf{\Phi}] = N\mathbf{P}$, where the (k,l)th element of $\mathbf{P}$ is $E[r^{k+l}]$.

Consider as an example, r uniformly distributed in $[0,1]$, which gives rise to $E[r^{k+l}] = 1/(k+l+1)$. The resulting $\mathbf{P}$ matrix is known as a segment of the generalized Hilbert matrix with $p = 2$ [110], which is ill-conditioned.

The condition number of a matrix is defined as $\rho = |\lambda_{\max}/\lambda_{\min}|$, where $\lambda_{\max}$ and $\lambda_{\min}$ are its maximum and minimum eigenvalue, respectively. It can be used to predict the numerical stability associated with matrix inversion. In general, when the condition number is much larger than 1, the numerical error involved in inverting the matrix can be significant.

Let us consider two additional probability density functions (PDFs) for which closed form expressions of the moments are available. If r is exponentially distributed with parameter λ; i.e., the PDF

$$f_r(r) = 1/\lambda\ e^{-r/\lambda}, \qquad r \geq 0, \tag{5.8}$$

it can be shown that $E[r^k] = k!\lambda^k$. As another example, consider the Rayleigh PDF

$$f_r(r) = \frac{r}{\sigma^2}\ e^{-r^2/(2\sigma^2)}, \qquad r \geq 0. \tag{5.9}$$

Its kth-order moment is $E[r^k] = (\sqrt{2}\sigma)^k\ \Gamma(k/2+1)$, where the Gamma function is defined as $\Gamma(x) = \int_0^\infty t^{x-1}\ e^{-t}\ dt$. When k is even, $E[r^k] = (\sqrt{2}\sigma)^k(k/2)!$; when k is odd, $E[r^k] = \sqrt{\pi/2}\ \sigma^k\ 1\cdot 3\cdot 5\cdots k$.

In Figure 5.2, we show the condition number ρ of the K-dimensional matrix $\mathbf{P}$ whose (k,l)th element is $E[r^{k+l}]$. We consider three PDFs for r: (i) r is uniformly distributed in $[0,1]$; (ii) r is exponentially distributed with parameter $\lambda = 0.5$; (iii) r is Rayleigh distributed with parameter $\sigma^2 = 1.5$. We observe that the condition number increases exponentially as a function of K and becomes very large even for a moderate K. This implies that in practice, the inversion of $\mathbf{\Phi}^H\mathbf{\Phi}$ in (5.7) can be difficult.

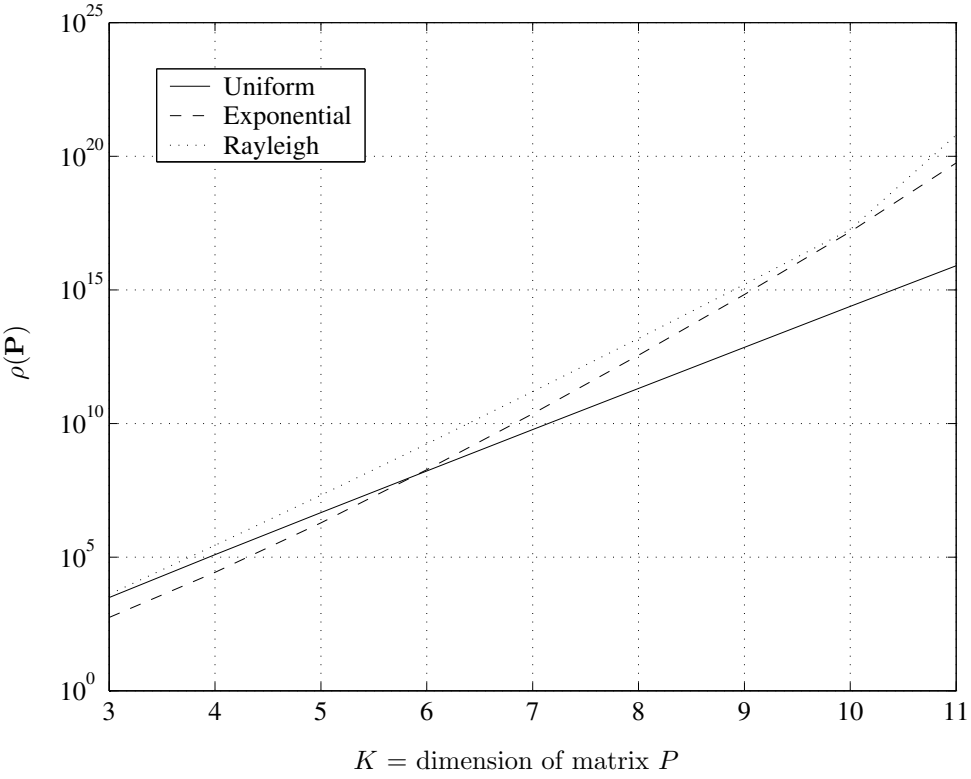


Figure 5.2: Condition number of the size $K \times K$ matrix $\mathbf{P}$, whose (k, l)th element is $E[r^{k+l}]$. Case (i): r is uniformly distributed in [0, 1]; Case (ii): r is exponentially distributed with $\lambda = 0.5$; Case (iii): r is Rayleigh distributed with $\sigma^2 = 1.5$.

5.2.2 Orthogonal polynomial bases

To alleviate the numerical instability problem associated with the basis set $\mathbf{\Phi}$ in (5.5), we consider orthogonal polynomials. Instead of (5.6), we write

$$\mathbf{y} = \mathbf{\Psi}\boldsymbol{\beta}, \tag{5.10}$$

where the new set of basis,

$$\mathbf{\Psi} = [\boldsymbol{\psi}_1(\mathbf{x}) \;\; \boldsymbol{\psi}_2(\mathbf{x}) \;\; \ldots \;\; \boldsymbol{\psi}_K(\mathbf{x})],$$

spans the same space as $\mathbf{\Phi}$. The least squares solution to (5.10) is

$$\boldsymbol{\beta}_{LS} = (\mathbf{\Psi}^H\mathbf{\Psi})^{-1}\mathbf{\Psi}^H\mathbf{y}. \tag{5.11}$$

We consider the following requirements for $\mathbf{\Psi}$:

1. Orthogonality: Any two different basis functions, $\psi_k(x)$ and $\psi_l(x)$, are orthogonal; i.e.,

$$E[\psi_k^*(x)\psi_l(x)] = 0, \qquad \forall\ k \neq l. \tag{5.12}$$

2. Form of the basis: We consider polynomial basis

$$\psi_k(x) = \sum_{l=1}^{k} \phi_l(x)\ U_{lk} = \sum_{l=1}^{k} U_{lk}\ |x|^{l-1}x, \tag{5.13}$$

where U_{lk} is generally complex-valued and $U_{lk} = 0$ for $l > k$. Note that $\psi_k(x)$ has order k.

Therefore, we seek an upper triangular matrix $\mathbf{U}$ whose (l,k)th element is U_{lk}, to construct the orthogonal polynomial basis $\mathbf{\Psi} = \mathbf{\Phi}\ \mathbf{U}$ such that

$$E[\mathbf{\Psi}^H\mathbf{\Psi}] = \mathbf{U}^H E[\mathbf{\Phi}^H\mathbf{\Phi}]\mathbf{U} = N\mathbf{U}^H\mathbf{P}\mathbf{U}$$

is diagonal.

As we show in Section 5.2.1, the (k,l)th element of $\mathbf{P}$ is $E[r^{k+l}]$. Given a PDF for r, the orthogonal polynomial basis construction problem becomes finding the upper triangular

matrix $\mathbf{U}$ such that $\mathbf{U}^H\ \mathbf{P}\ \mathbf{U} = \text{diag}(d_1, \ldots, d_K)$. Therefore in theory, orthogonal polynomials are tied to the PDF of the signal amplitude. Since $\mathbf{P}$ always tends to be ill-conditioned (see the examples in Section 5.2.1), solving for $\mathbf{U}$ is often a numerically challenging task.

Let $r = |x|$ be uniformly distributed in [0, 1], and require that the squared norm of the basis be preserved; i.e.,

$$d_k = E[|\psi_k(x)|^2] = E[|\phi_k(x)|^2] = \int_0^1 r^{2k}\ dr = \frac{1}{2k+1}, \qquad 1 \le k \le K. \tag{5.14}$$

We show in the Appendix that the matrix $\mathbf{U}$ that solves this problem has an elegant expression:

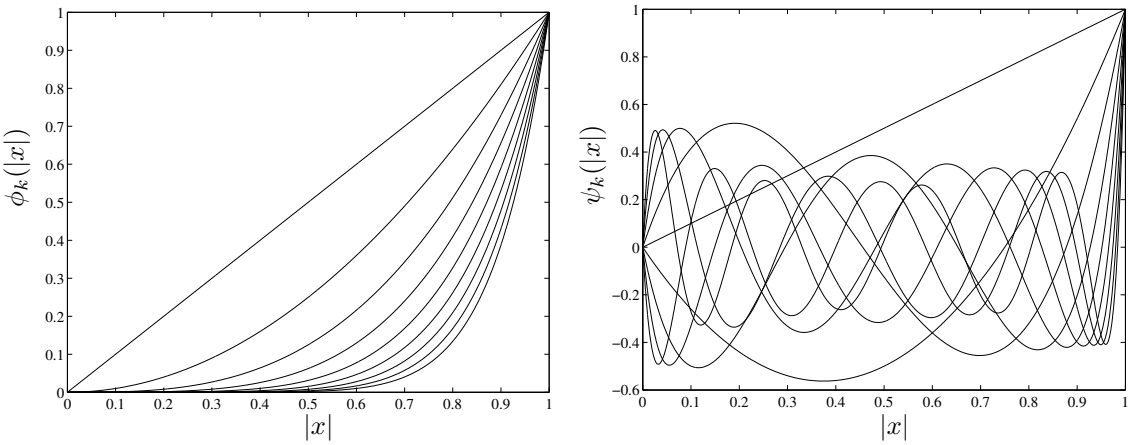
$$U_{lk} = \begin{cases} (-1)^{l+k} \frac{(k+l)!}{(l-1)!(l+1)!(k-l)!}, & l \le k, \\ 0, & l > k. \end{cases} \tag{5.15}$$

Therefore, the kth-order orthogonal polynomial basis function for the uniformly distributed $|x|$ is

$$\psi_k(x) = \sum_{l=1}^{k} (-1)^{l+k} \frac{(k+l)!}{(l-1)!(l+1)!(k-l)!} |x|^{l-1} x. \tag{5.16}$$

Table 5.1 shows the first 7 such orthogonal polynomials. Notice that if we replace the complex valued basis functions $|x|^{k-1}x, k = 1, 2..., K$, with real-valued basis functions, $|x|^k, k = 1, 2..., K$, we obtain the real-valued orthogonal polynomials defined in the region $[0, 1]$, which are known as the shifted Legendre polynomials [52], except that the $|x|^0$ polynomial is not included. Although the construction of orthogonal basis is often an iterative procedure, we were able to obtain novel, *closed form* expression (5.16) for complex-valued $x(t)$.

Fig. 5.3(a) shows $\phi_k(|x|) = |x|^k$ vs. $|x|$ and Fig. 5.3(b) shows $\psi_k(|x|)$ vs. $|x|$. When $x \to 0$, the conventional polynomial basis function $\phi_k(x)$ is on the order of $o(x^k)$ which goes to 0 a lot faster than x when $k > 1$. Implementing $\phi_k(x)$ with a lookup table can therefore introduce more error when a low input value is forced to zero due to quantization. The orthogonal polynomial basis function, $\psi_k(x)$, is free of such problem, since it is on the order $o(x)$ for any k. Moreover, $\psi_k(x)$ has k different roots whereas $\phi_k(x)$ has k repeated roots at $x = 0$. This implies that $\psi_k(x)$ has a richer "frequency content" and hence has better interpolating properties.



(a) Conventional polynomial basis (b) Orthogonal polynomial basis

Figure 5.3: Comparison between $\phi_k(|x|)$ and $\psi_k(|x|)$.

Table 5.1: Orthogonal polynomial basis functions $\psi_k(x)$ for $1 \leq k \leq 7$.

$\psi_1(x)$	$=$	x
$\psi_2(x)$	$=$	$4\|x\|x - 3x$
$\psi_3(x)$	$=$	$15\|x\|^2x - 20\|x\|x + 6x$
$\psi_4(x)$	$=$	$56\|x\|^3x - 105\|x\|^2x + 60\|x\|x - 10x$
$\psi_5(x)$	$=$	$210\|x\|^4x - 504\|x\|^3x + 420\|x\|^2x - 140\|x\|x + 15x$
$\psi_6(x)$	$=$	$792\|x\|^5x - 2310\|x\|^4x + 2520\|x\|^3x - 1260\|x\|^2x + 280\|x\|x - 21x$
$\psi_7(x)$	$=$	$3003\|x\|^6x - 10296\|x\|^5x + 13860\|x\|^4x - 9240\|x\|^3x + 3150\|x\|x^2 - 504\|x\|x + 28x$

If r is uniformly distributed in $[0, 1]$, the condition number of $E[\mathbf{\Psi}^H\mathbf{\Psi}] = N\mathbf{U}^H\mathbf{P}\mathbf{U}$ is $d_1/d_K = (2K + 1)/3$, which is orders of magnitudes smaller than that of $E[\mathbf{\Phi}^H\mathbf{\Phi}] = N\mathbf{P}$; c.f. Figure. 5.2.

5.2.3 Discussions on the orthogonal polynomial basis

In theory, we can calculate the orthogonal polynomial basis for any given PDF $f_r(r)$; i.e., we can find the $\mathbf{U}$ matrix that makes $\mathbf{U}^H\mathbf{P}\mathbf{U}$ diagonal, for $\mathbf{P}$ generated from the given distribution. However, we were only able to obtain closed form expression of U_{lk} for the uniform distribution case (closed form solution may not exist for other distributions). Interestingly, the basis functions (5.16) (see also Table 5.1) consist of integer-valued coefficients which are free of round-off errors. For almost all other PDFs, U_{lk} will be non-integer valued.

Although $\{\psi_k(x)\}$ of (5.16) are derived assuming that the input amplitude, $r = |x|$, is uniformly distributed, we would like to show that such $\psi_k(x)$ can be used even if r is not uniformly distributed.

For a given PDF $f_r(r)$, we can form matrix $\mathbf{P}$ whose (k,l)th element is $E[r^{k+l}]$. Hence $E[\mathbf{\Psi}^H\mathbf{\Psi}] = N\mathbf{U}^H\mathbf{P}\mathbf{U}$ is known, where $\mathbf{U}$ is given by (5.15). When $f_r(r)$ is not uniform in $[0,1]$, $\mathbf{U}^H\mathbf{P}\mathbf{U}$ is probably not diagonal, but our hope is that its condition number does not become huge. Keep in mind that our primary concern is to obtain $(\mathbf{\Psi}^H\mathbf{\Psi})^{-1}$ accurately; the exact orthogonality of $\mathbf{\Psi}$ for any given data $\mathbf{x}$ is of secondary importance.

Let us consider the "truncated exponential" distribution.

$$f_r(r) = \begin{cases} \frac{1}{1-e^{-\frac{1}{\lambda}}}\frac{1}{\lambda}e^{-\frac{r}{\lambda}}, & 0 \le r \le 1, \\ 0, & \text{otherwise}, \end{cases} \tag{5.17}$$

and the "truncated Rayleigh" distribution,

$$f_r(r) = \begin{cases} \frac{1}{1-e^{-\frac{1}{2\sigma^2}}}\frac{r}{\sigma^2}e^{-\frac{r^2}{2\sigma^2}}, & 0 \le r \le 1, \\ 0, & \text{otherwise}. \end{cases} \tag{5.18}$$

Suppose that v is a uniformly distributed r.v. in [0, 1]. Transformations

$$r = -\lambda \ln(1 - v(1 - e^{-1/\lambda}))$$

and

$$r = \sqrt{-2\sigma^2 \ln(1 - v(1 - e^{-1/(2\sigma^2)}))}$$

generate r.v. r with the PDF in (5.17) and (5.18), respectively.

Four specific cases are investigated:

(i) r is truncated Rayleigh distributed (c.f. (5.18)) with $\sigma^2 = 0.1086$;

(ii) r is truncated Rayleigh distributed (c.f. (5.18)) with $\sigma^2 = 0.5$;

(iii) r is truncated exponentially distributed (c.f. (5.17)) with $\lambda = 0.2127$;

(iv) r is truncated exponentially distributed (c.f. (5.17)) with $\lambda = 1$.

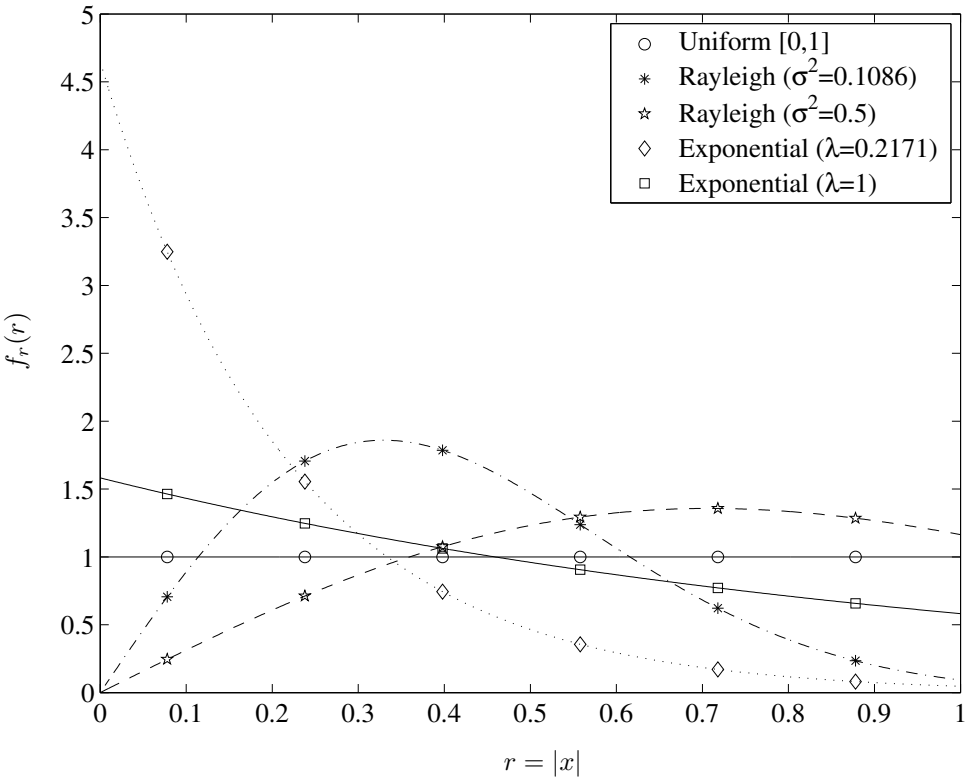
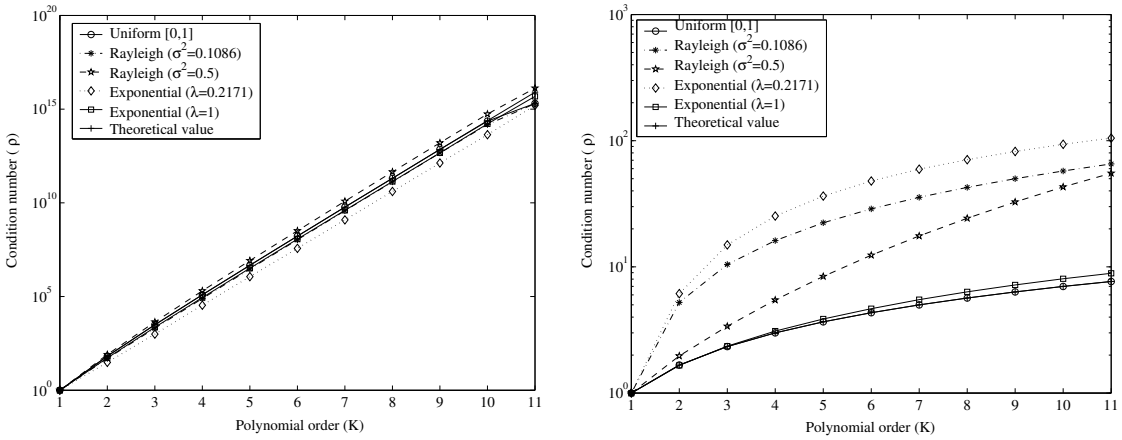


Figure 5.4: PDF for (i) truncated Rayleigh distribution with $\sigma^2 = 0.1086$; (ii) truncated Rayleigh distribution with $\sigma^2 = 0.5$; (iii) truncated exponential distribution with $\lambda = 0.2171$; (iv) truncated exponential distribution with $\lambda = 1$; along with the uniform distribution in $[0, 1]$.

The above four PDFs along with the uniform distribution are shown in Fig. 5.4.

Fig. 5.5(a) shows the condition number for the matrix $\mathbf{\Phi}^H\mathbf{\Phi}$ for each of the PDFs, calculated from 1,000,000 samples. We note that, for each PDF, the condition number grows exponentially with the polynomial order K. Fig. 5.5(b) shows the condition number for the matrix $\mathbf{\Psi}^H\mathbf{\Psi}$ for each of the PDFs, which increases at a much slower rate as K increases and are within 100 for the cases tested. The low condition number will ensure better numerical stability when finite precision computation of the model coefficients is carried out. Notice from Fig. 5.4 and Fig. 5.5(b) that the closer a PDF resembles the uniform distribution (e.g., case (iv) as opposed to case (iii)), the lower the condition number when the orthogonal polynomial basis functions (5.16) are used.

In practice, we do not require $r = |x|$ to be exactly in $[0, 1]$ in order for the orthogonal polynomial basis $\psi_k(x)$ to be used. There are two scenarios however, where a simple scaling



(a) Conventional polynomial case (b) Orthogonal polynomial case

Figure 5.5: Comparison of the condition numbers of (a) $\mathbf{\Phi}^H\mathbf{\Phi}$ and of (b) $\mathbf{\Psi}^H\mathbf{\Psi}$ for the PDFs given in Fig. 5.4. 1,000,000 samples were drawn from the following PDFs: (i) truncated Rayleigh distribution with $\sigma^2 = 0.1086$; (ii) truncated Rayleigh distribution with $\sigma^2 = 0.5$; (iii) truncated exponential distribution with $\lambda = 0.2171$; (iv) truncated exponential distribution with $\lambda = 1$; along with the uniform distribution in $[0, 1]$. For (a), the theoretical value is the condition number of matrix $\mathbf{P}$, whose (k, l)th entry is $\frac{1}{k+l+1}$, $1 \leq k, l \leq K$. For (b), the theoretical value of the condition number is $\frac{2K+1}{3}$.

operation is needed. One situation is if $\Pr(r > 1) \gg 0$; i.e., a significant number of r values exceed 1. Another situation is if $\Pr(0 \leq r \leq \delta) \approx 1$ for some $\delta \ll 1$; i.e., the r values concentrate around zero. For both cases, we first scale x to obtain $\check{x} = \mu x$ and then apply the orthogonal polynomials to $\check{x}$. When $\Pr(r > 1) \gg 0$, μ should be chosen such that the maximum value of $|\check{x}|$ is around 1. For the second situation, μ should be selected such that the "center of gravity" of $|\check{x}|$ is shifted closer to 1. A step-by-step algorithm is outlined below.

Step 1. Determine $\mu > 0$ such that the majority of μr values lie in $[0, 1]$, or that μr spreads over much of the $[0, 1]$ interval. One possible scaling factor is $\mu = 1/\max_t |x(t)|$, if $|x(t)|$ is bounded. Let $\check{x} = \mu x$.

Step 2. Form $\psi_k(\check{x})$ according to (5.16); i.e.,

$$\psi_k(\check{x}(t)) = \sum_{l=1}^{k} (-1)^{l+k} \frac{(k+l)!}{(l-1)!(l+1)!(k-l)!} |\check{x}(t)|^{l-1} \check{x}(t).$$

Then form

$$\boldsymbol{\psi}_k(\check{\mathbf{x}}) = [\psi_k(\check{x}(t_1)), \ldots, \psi_k(\check{x}(t_N))]^T,$$

$$\boldsymbol{\Psi} = [\boldsymbol{\psi}_1(\check{\mathbf{x}}) \;\; \boldsymbol{\psi}_2(\check{\mathbf{x}}) \;\; \ldots \;\; \boldsymbol{\psi}_K(\check{\mathbf{x}})].$$

Step 3. Solve for $\boldsymbol{\beta}$ using (5.11); i.e.,

$$\boldsymbol{\beta}_{LS} = (\boldsymbol{\Psi}^H \boldsymbol{\Psi})^{-1} \boldsymbol{\Psi}^H \mathbf{y}.$$

Step 4. The model is then

$$y(t) = \sum_{k=1}^{K} \beta_k \; \psi_k(\mu x(t)).$$

Next, we would like to point out major differences between our orthogonal polynomial basis (5.16) and that of [78, 79].

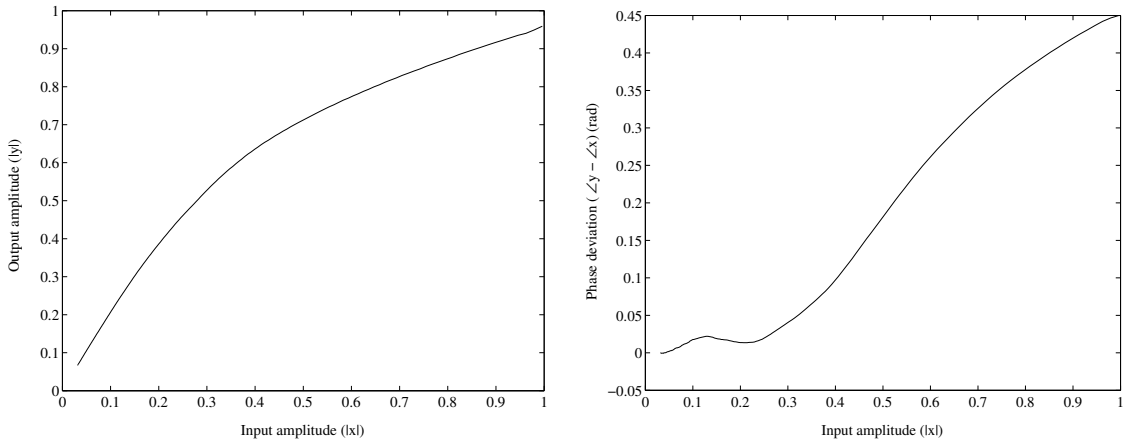
(i) Our basis set $\{\psi_k(x)\}$ includes both even and odd k values, whereas in [78, 79], only odd k values are allowed. Even-order terms are beneficial; see [40].

(ii) In [78, 79], an "orthogonal polynomial basis function calculator" constructs the basis functions using the available data $\mathbf{x}$, whereas our $\psi_k(\cdot)$ functions are already constructed and are available in closed form. For real data applications, speed is an important concern and the optimization of the basis functions for each data vector is not necessary. Since our $\psi_k(x)$ functions are pre-determined, a lookup table can be built to further accelerate the speed. In [78, 79], matrix inversion is avoided because of the exact orthogonality of the basis functions for the data present, but the computation load is shifted to constructing the basis functions in the first place. Our standpoint is that, we do not need $\boldsymbol{\Psi}^H \boldsymbol{\Psi}$ to be exactly diagonal for every realization of $\mathbf{x}$. As long as the condition number of $\boldsymbol{\Psi}^H \boldsymbol{\Psi}$ is low, we can proceed with its inversion and obtain the model coefficients via linear least squares as in (5.11).

(iii) In solving for the U_{lk} coefficients for the construction of $\boldsymbol{\Psi} = \boldsymbol{\Phi}\mathbf{U}$, we desire the U_{lk} coefficients to be very accurate. Most orthogonal bases (including that of [78, 79]) are constructed iteratively, often leading to less reliable U_{lk} values for larger k's. Our non-iterative, closed form, integer valued solution (5.15) ensures that the $\psi_k(x)$ basis functions are free of round-off errors.

5.2.4 PA modeling example

If the baseband input to a nonlinear PA is $x(t)$ and the corresponding baseband output is $y(t)$, then $|y(t)|$ vs. $|x(t)|$ is the so-called AM/AM conversion, and $\angle y(t) - \angle x(t)$ vs. $|x(t)|$ is the so-called AM/PM conversion. We first measured the AM/AM and AM/PM characteristics of a 2-stage GaAsFET RFIC PA (Class AB) as shown in Fig. 5.6(a) and Fig. 5.6(b), respectively.



(a) AM/AM transfer characteristic of the PA (b) AM/PM transfer characteristic of the PA

Figure 5.6: AM/AM and AM/PM transfer characteristics of an actual class AB PA.

Next, a polynomial PA model is constructed to fit the measured data. When the conventional polynomial basis is used, the resulting PA model output is

$$\hat{y}(t) = \sum_{k=1}^{K} b_k \ \phi_k(x(t)). \tag{5.19}$$

When the orthogonal polynomial basis is used,

$$\hat{y}(t) = \sum_{k=1}^{K} \beta_k \ \psi_k(x(t)). \tag{5.20}$$

Theoretically, when additional higher order polynomial terms are included in the PA model, modeling error becomes progressively smaller. To give a quantitative measure of the approximation accuracy, we define a normalized mean square error,

$$\text{NMSE (dB)} = 10 \log_{10} \left[\frac{\sum_{n=1}^{N} |y(t_n) - \hat{y}(t_n)|^2}{\sum_{n=1}^{N} |y(t_n)|^2} \right], \tag{5.21}$$

where $y(t_n)$ is the measured PA output and $\hat{y}(t_n)$ is the PA model output.

In Fig. 5.7, we show the modeling errors (NMSEs) for the conventional (solid line) and the orthogonal polynomial (dash dotted line) PA models. The $\{b_k\}$ coefficients in (5.19) and the $\{\beta_k\}$ coefficients in (5.20) were computed from (5.7) and (5.11), respectively, and the computation environment was C with 32-bit floating point precision. We observe that for the orthogonal polynomials, when the maximum polynomial order increases, modeling error constantly decreases. However, this is not the case for the conventional polynomials. In fact, for the PA under study, the orthogonal polynomial PA model shows superiority starting at $K = 7$, which is not a very high polynomial PA model order for realistic PAs.

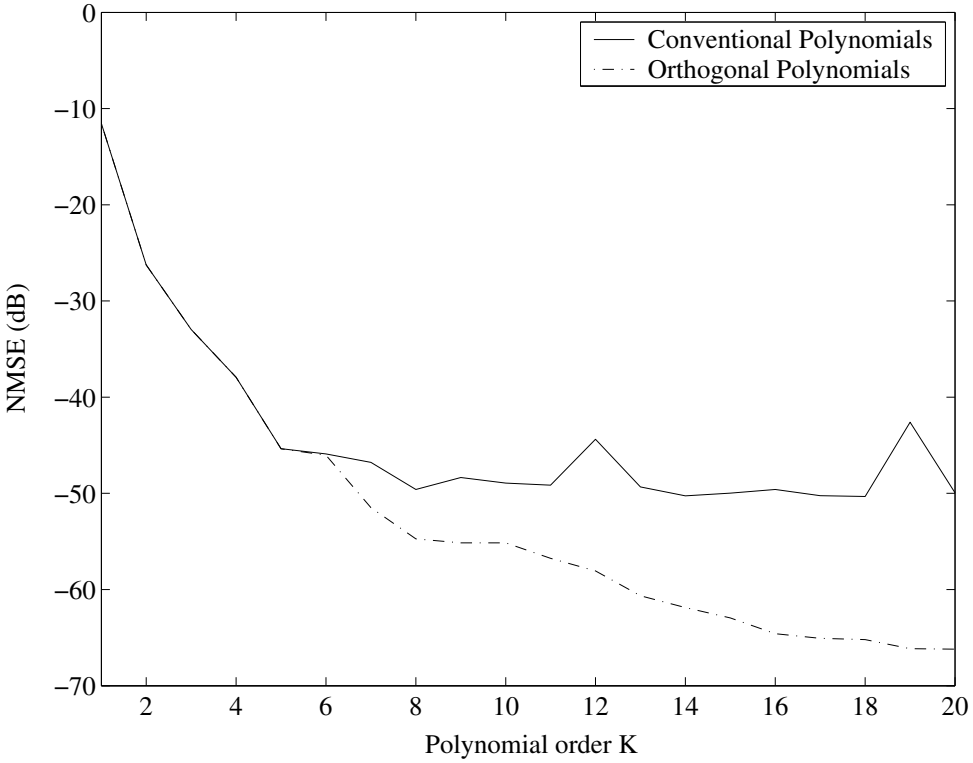


Figure 5.7: PA modeling errors for the conventional polynomial model (solid line) and the orthogonal polynomial model (dot-dashed line) when K polynomial basis functions are used.

5.2.5 Extension to the memory polynomial case

For wideband applications (e.g., $x(t)$ is wideband CDMA) and/or with high power amplifiers (e.g., basestation PAs), memory effects show up in the PA [65], [118], [33]. In [65], the memory polynomial model is shown to be a good model for characterizing nonlinear PAs

with memory effects.

Here, we assume uniform sampling with sampling period T. The memory polynomial PA model is given by

$$y[n] = \sum_{k=1}^{K} \sum_{q=0}^{Q} b_{kq}\ x[n-q]|x[n-q]|^{k-1}, \tag{5.22}$$

where $x[n] = x(t_n) = x(nT)$ and $y[n] = y(t_n) = y(nT)$, K is the highest polynomial order, and Q is the maximum delay.

Let us define the N-by-1 shifted input vector

$$\mathbf{x}_q = [\mathbf{0}_{1\times q},\ x[1],\ \ldots,\ x[N-q]]^T,$$

the N-by-1 vector

$$\boldsymbol{\phi}_k(\mathbf{x}_q) = [\mathbf{0}_{1\times q},\ \phi_k(x[1]),\ \ldots,\ \phi_k(x[N-q])]^T,$$

the N-by-K matrix

$$\boldsymbol{\Phi}_q = [\boldsymbol{\phi}_1(\mathbf{x}_q)\ \boldsymbol{\phi}_2(\mathbf{x}_q)\ \ldots\ \boldsymbol{\phi}_K(\mathbf{x}_q)],$$

and the N-by-$K(Q+1)$ matrix

$$\boldsymbol{\Phi} = [\boldsymbol{\Phi}_0\ \boldsymbol{\Phi}_1\ \ldots\ \boldsymbol{\Phi}_Q].$$

The $K(Q+1)$ parameter vector is

$$\mathbf{b} = [b_{10},\ b_{20},\ \ldots,\ b_{K0},\ \ldots, b_{1Q},\ b_{2Q}, \ldots,\ b_{KQ}]^T.$$

We can then rewrite (5.22) as $\mathbf{y} = \boldsymbol{\Phi}\mathbf{b}$ and solve for $\mathbf{b}$ using linear least squares as in (5.7).

As one might expect, the "conventional memory polynomial" model (5.22) may suffer the same numerical instability problem described earlier. To alleviate such problem, we suggest to replace $\phi_k(x)$ by $\psi_k(x)$. The result is the "orthogonal memory polynomial" model,

$$y[n] = \sum_{k=1}^{K} \sum_{q=0}^{Q} \beta_{kq}\ \psi_k(x[n-q]). \tag{5.23}$$

By replacing ϕ_k by ψ_k, and b_{kq} by β_{kq}, we can write $\mathbf{y} = \boldsymbol{\Psi}\boldsymbol{\beta}$, and solve for $\boldsymbol{\beta}$ using linear least squares similar to (5.11).

Although orthogonality holds for $\mathbf{\Phi}_q$ for each q; i.e., $E[\mathbf{\Phi}_q^H \mathbf{\Phi}_q]$ is diagonal, orthogonality does not hold for the different delayed elements; i.e., $E[\mathbf{\Phi}^H \mathbf{\Phi}]$ is not exactly diagonal. However, we still expect the orthogonal memory polynomial model (5.23) to be numerically more robust than the conventional memory polynomial model (5.22).

5.3 Orthogonal polynomial predistortion

5.3.1 Predistortion via the indirect learning architecture

The baseband model (5.4) can be used for PA modeling as well as predistorter design. For the latter, we advocate the use of the indirect learning architecture [46] as shown in Fig. 5.8.

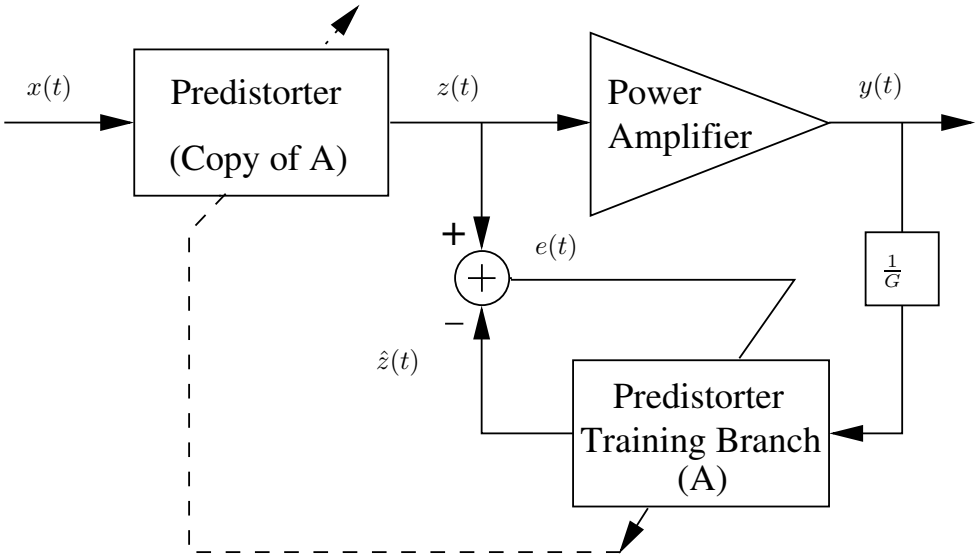


Figure 5.8: Indirect learning architecture.

The baseband PA input is denoted by $x(t)$, the baseband PA output / predistorter input is denoted by $z(t)$, and the baseband PA output is denoted by $y(t)$. The feedback path labeled "Predistorter Training Branch" (block A) has $y(t)/G$ as its input, where G is the intended gain of the PA, and $\hat{z}(t)$ is its output. The actual predistorter (copy of A) is an exact copy of the predistorter training branch. Since when $y(t) = Gx(t)$, the error $e(t) = z(t) - \hat{z}(t)$ is 0, the predistorter parameters can be found by minimizing $\|e(t)\|^2$. The benefit of the indirect learning architecture is that, instead of assuming a model for the PA, estimating the PA parameters and then constructing its inverse, we can go directly after

the predistorter[1].

For the predistorter training branch, if the conventional polynomial model is adopted, we have,

$$\hat{z}(t) = \sum_{k=1}^{K} a_k \left| \frac{y(t)}{G} \right|^{k-1} \frac{y(t)}{G}, \tag{5.24}$$

where G is the intended linear gain of the PA. Based on a set of PA input $\{z(t_i)\}_{i=1}^{N}$ and output $\{y(t_i)\}_{i=1}^{N}$ measurements, the least-squares solution can be obtained for the predistorter coefficients, $\mathbf{a} = [a_1, \ldots, a_K]^T$, similar to (5.7), but with $y(t)/G$ now playing the role of $x(t)$, and $z(t)$ playing the role of $y(t)$. Once the coefficients $\{a_k\}$ are found, they are plugged into the predistorter:

$$z(t) = \sum_{k=1}^{K} a_k \ |x(t)|^{k-1} x(t) = \sum_{k=1}^{K} a_k \ \phi_k(x(t)). \tag{5.25}$$

This procedure can be repeated, one benefit of the iterations being that more diverse values of $z(t)$ and $y(t)$ are possible and are helpful for obtaining more accurate model parameter estimates. To initialize, $\mathbf{a} = [1\ 0\ \ldots\ 0]^T$ can be used. Such recursive procedure enables the predistorter to linearize even a (slowly) time-varying PA.

As in PA modeling, in solving for the $\{a_k\}$ coefficients for the conventional polynomial predistorter (5.25), the numerical instability problem may show up. Instead of (5.25), we prefer to use the orthogonal polynomial predistorter,

$$z(t) = \sum_{k=1}^{K} \alpha_k \ \psi_k(x(t)), \tag{5.26}$$

where $\psi_k(x)$ is given by (5.16) (see also Table 5.1).

Since $\psi_k(x)$ is a linear combination of $\{\phi_l(x)\}_{l=1}^{k}$, models (5.26) and (5.25) are equivalent, in theory. However, in practice, sampling the input and output of a PA using a finite precision analog-to-digital converter (ADC) may introduce error to the samples. Furthermore, since obtaining the LS estimates of the predistorter coefficients requires a matrix inversion, the digital signal processor (DSP) precision may impact the accuracy of the resulting matrix inverse. With the orthogonal polynomial predistorter model (5.26), numerical problems due to quantization and finite precision calculations in the DSP can be significantly alleviated.

[1]The term "indirect learning" seems counter-intuitive here, since the predistorter is learned directly; it is the PA characteristics that are learned indirectly.

5.3.2 Simulations – predistortion linearization

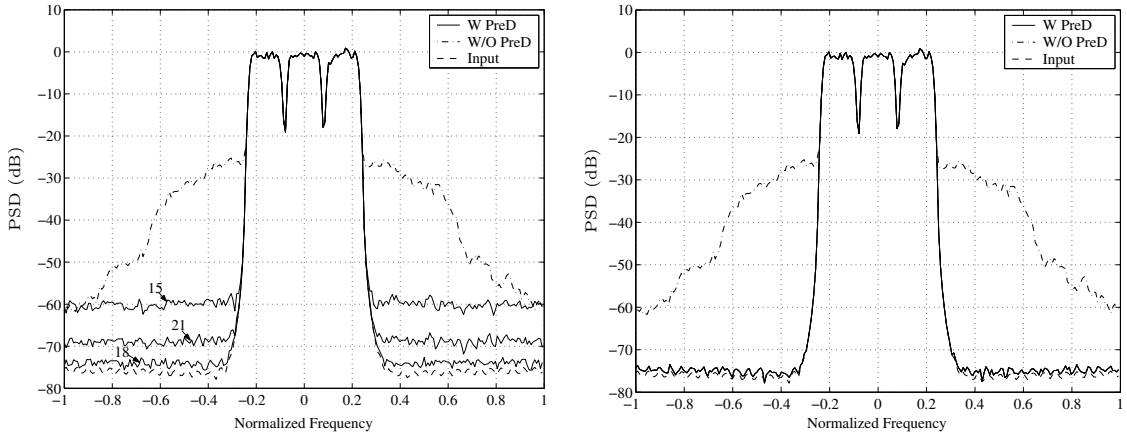
The numerical problems associated with estimating the predistorter coefficients will affect the performance of the predistorter. We will examine the problem from the view point of spectral regrowth suppression. We utilize the indirect learning architecture shown in Fig. 5.8 to carry out predistortion linearization. The predistorter's input, $x(t)$, is a three carrier UMTS signal [105]. The simulation environment is C, with 32-bit floating point accuracy. The plots are done in MATLAB®.

5.3.2.1 Memoryless predistortion

The PA input $z(t)$ and output $y(t)$ are assumed to obey the arctan model:

$$y(t) = \left(\gamma_1 \tan^{-1}(\zeta_1|z(t)|) + \gamma_2 \tan^{-1}(\zeta_2|z(t)|)\right) e^{j\angle z(t)}, \tag{5.27}$$

where $\gamma_1 = 8.00335 - j4.61157$, $\gamma_2 = -3.77167 + j12.03758$, $\zeta_1 = 2.26895$, and $\zeta_2 = 0.8234$. This PA model fits well measured data from an actual Class AB PA discussed in Section 5.2.4. The intended linear gain is set to $G = 7$.



(a) Conventional polynomial predistorter (b) Orthogonal polynomial predistorter

Figure 5.9: Memoryless predistortion example. Dash dotted lines show the PA output PSD without predistortion; solid lines show the PA output PSD with predistortion (results are shown for iteration numbers 15, 18 and 21); dashed lines show the PA input PSD. For easy visual comparison, output PSDs are normalized with respect to the input PSD. (a) Conventional polynomial predistorter with $K = 7$. The predistorter did not converge, revealing a numerical instability problem. (b) Orthogonal polynomial predistorter with $K = 7$. The predistorter converged and could fully suppress spectral regrowth.

Fig. 5.9(a) shows the PSD at the output of the PA for the conventional polynomial

predistorter with a polynomial order $K = 7$. The PSD is presented for iterations 15, 18, and 21, respectively and shows no sign of convergence. In contrast, Fig. 5.9(b) shows the PSD at the output of the PA for the orthogonal polynomial predistorter with the same order $K = 7$. In this case, the predistorter shows stability and effectiveness.

5.3.2.2 Predistortion with memory

When a nonlinear PA exhibits memory effects, memoryless predistortion is often ineffective. Indeed, the exact inverse of a nonlinear system with memory should be another nonlinear system with memory.

In this example, we assume that the nonlinear system to be compensated for obeys the Wiener-Hammerstein (W-H) model; i.e., a linear time invariant (LTI) system followed by a memoryless nonlinearity, which in turn is followed by another LTI system. The LTI blocks before and after the memoryless nonlinearity, which are denoted by $H(z)$ and $G(z)$, respectively, are

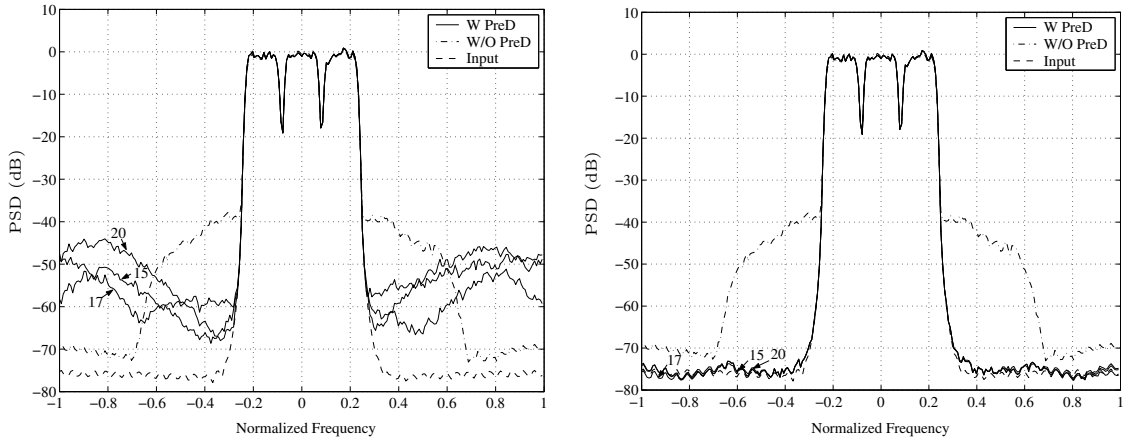
$$H(z) = \frac{1}{1.5}\frac{1+0.25z^{-2}}{1+0.4z^{-1}}, \qquad G(z) = \frac{1}{0.52}\frac{1-0.1z^{-1}}{1-0.2z^{-1}}. \tag{5.28}$$

For the memoryless nonlinear portion of the W-H model, we choose the arctan model defined in (5.27) with the same parameters.

Using the indirect learning architecture, conventional and orthogonal memory polynomial predistorters can be constructed following the ideas of Section 5.2.5. Here we show the conventional and the orthogonal memory polynomial predistorters with 5 delay taps ($Q = 5$) and 7th order polynomials ($K = 7$).

Performance of the conventional memory polynomial predistorter is shown in Fig. 5.10(a). Its performance is not satisfactory; especially noticeable is the worsening of the alternate channel spectral regrowth. Moreover, the conventional memory polynomial predistorter did not converge even after 20 iterations; the instability was caused by ill conditioning of the $\mathbf{\Phi}^H\mathbf{\Phi}$ matrix. Increasing Q or K did not alleviate the problem either.

In Fig. 5.10(b), we show performance of the the orthogonal memory polynomial predistorter for the same PA. We observe that spectral regrowth was well suppressed and the



(a) Conventional memory polynomial predistorter (b) Orthogonal memory polynomial predistorter

Figure 5.10: Predistortion linearization performance of a memory polynomial predistorter for a Wiener-Hammerstein system. Dash-dotted lines show the PA output PSD without predistortion; solid lines show the PA output PSD with memory polynomial predistortion (results are shown for iteration numbers 15, 17, 20); dashed lines show the PA input PSD. For easy visual comparison, output PSDs are normalized with respect to the input PSD. (a) Conventional memory polynomial predistorter with $K = 7$ and $Q = 5$. The predistorter did not converge, revealing a numerical instability problem. (b) Orthogonal memory polynomial predistorter with $K = 7$ and $Q = 5$. The predistorter converged and could suppress most of the spectral regrowth.

predistorter converged. In general, linearizing a nonlinear system with memory is a difficult task, but our proposed orthogonal memory polynomial predistorter is promising.

5.4 Conclusions

In this chapter, the benefits of using the orthogonal polynomials as opposed to the conventional polynomials are explored, in the context of digital baseband PA modeling and predistorter design. Closed-form expression for the orthogonal polynomial basis is derived. We demonstrated using simulation examples, the numerical instability problem associated with the conventional polynomials, when the polynomial order is high (e.g., $K \geq 7$). Extension to an orthogonal polynomial predistorter with memory is also considered.

5.A Orthogonality Proof

We prove here the orthogonality of the polynomials given in (5.16) when $r = |x|$ is uniformly distributed in $[0, 1]$, and find the norm of such orthogonal polynomials.

The inner-product of two functions $\psi_k(x)$ and $\psi_l(x)$ is defined as

$$\langle \psi_k(x), \psi_l(x) \rangle = E[\psi_k^*(x)\psi_l(x)].$$

When $\langle \psi_k(x), \psi_l(x) \rangle = 0$, we say that $\psi_k(x)$ and $\psi_l(x)$ are orthogonal. The norm of $\psi(x)$ is given by

$$\|\psi(x)\| = \sqrt{\langle \psi(x), \psi(x) \rangle}. \tag{5.29}$$

Recall that the Gamma function is given by:

$$\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt, \qquad x \geq 1. \tag{5.30}$$

We henceforth assume that $r = |x|$ is uniformly distributed in [0, 1].

Theorem 1. *Let* $\psi_n^{(p)}(x) = \sum_{i=1}^n U_{in}^{(p)} |x|^{i+\frac{p}{2}-1} e^{j\angle x}$, *where* $j = \sqrt{-1}$ *and*

$$U_{in}^{(p)} = (-1)^{n+i} \frac{\Gamma(n+p+i-1)}{\Gamma(i)\Gamma(i+p)\Gamma(n-i+1)}. \tag{5.31}$$

For any $n \in \mathbb{Z} \geq 1$, *and* $p \in \mathbb{R} \geq 0$, *the basis* $\left\{\psi_1^{(p)}(x), \ldots, \psi_N^{(p)}(x)\right\}$ *is an orthogonal basis with the squared-norm* $\left\|\psi_n^{(p)}(x)\right\|^2 = \frac{1}{2n+p-1}$.

Proof. We start by introducing the generalized segmented $n \times n$ Hilbert matrix, $\mathbf{H}_n$, given by

$$(\mathbf{H}_n)_{il} = \frac{1}{i+l+p-1} \qquad \forall 1 \leq l, i \leq n. \tag{5.32}$$

Its inverse, $\mathbf{H}_n^{-1}$, is given by [67]

$$\left(\mathbf{H}_n^{-1}\right)_{il} = \frac{(-1)^{i+l}}{p+i+l-1} \frac{\Gamma(n+p+i)}{\Gamma(i)\Gamma(i+p)\Gamma(n-i+1)} \frac{\Gamma(n+p+l)}{\Gamma(l)\Gamma(l+p)\Gamma(n-l+1)}. \tag{5.33}$$

Using (5.31) and invoking the property $\Gamma(x) = (x-1)\Gamma(x-1)$, we can rewrite (5.33) as:

$$\left(\mathbf{H}_n^{-1}\right)_{il} = \frac{(n+p+i-1)(n+p+l-1)}{p+i+l-1} U_{in}^{(p)} U_{ln}^{(p)}. \tag{5.34}$$

Since $\mathbf{H}_n^{-1}$ is the inverse of $\mathbf{H}_n$, it follows that for $1 \leq i, k \leq n$,

$$\sum_{l=1}^n \left(\mathbf{H}_n^{-1}\right)_{il} (\mathbf{H}_n)_{lk} = \delta_{ik}. \tag{5.35}$$

Substitution of (5.34) and (5.32) into (5.35) yields:

$$\sum_{l=1}^{n} \frac{(n+p+i-1)(n+p+l-1)}{p+i+l-1} U_{in}^{(p)} U_{ln}^{(p)} \frac{1}{l+k+p-1} = \delta_{ik}. \tag{5.36}$$

Replacing $k = n$ in (5.36) and simplifying, we obtain

$$\sum_{l=1}^{n} U_{ln}^{(p)} \frac{1}{p+i+l-1} = \frac{\delta_{in}}{(n+p+i-1)U_{in}^{(p)}} = \frac{\delta_{in}}{(2n+p-1)U_{nn}^{(p)}}. \tag{5.37}$$

The inner-product of $\psi_n^{(p)}(x)$ and $|x|^{i+\frac{p}{2}-1}e^{j\angle x}$ yields

$$\begin{aligned}
\left\langle \psi_n^{(p)}(x), |x|^{i+\frac{p}{2}-1}e^{j\angle x} \right\rangle &= \sum_{l=1}^{n} U_{ln}^{(p)} \int_0^1 |x|^{l+\frac{p}{2}-1} |x|^{i+\frac{p}{2}-1} dr \\
&= \sum_{l=1}^{n} U_{ln}^{(p)} \frac{1}{i+l+p-1} \\
&= \frac{\delta_{in}}{(2n+p-1)U_{nn}^{(p)}}.
\end{aligned} \tag{5.38}$$

Eq. (5.38) indicates that $\psi_n^{(p)}(x)$ is orthogonal to $|x|^{i+\frac{p}{2}-1}e^{j\angle x}$ for any $1 \le i \le n-1$. Thus $\psi_n^{(p)}(x)$ is orthogonal to any linear combination of $\{|x|^{i+\frac{p}{2}-1}e^{j\angle x}\}_{i=1}^{n-1}$. Therefore, $\psi_n^{(p)}(x)$ is orthogonal to any $\psi_l^{(p)}(x)$ with $1 \le l \le n-1$. This property holds for $n = 2, \ldots, N$, and therefore, the basis $\left\{\psi_1^{(p)}(x), \ldots, \psi_N^{(p)}(x)\right\}$ is orthogonal (i.e., each function in the basis is orthogonal to the others). The squared norm of $\psi_n^{(p)}(x)$ can be calculated as follows:

$$\begin{aligned}
\left\|\psi_n^{(p)}(x)\right\|^2 &= \left\langle \psi_n^{(p)}(x), \psi_n^{(p)}(x) \right\rangle \\
&= \sum_{i=1}^{n} U_{in}^{(p)} \left\langle \psi_n^{(p)}(x), |x|^{i+\frac{p}{2}-1}e^{j\angle x} \right\rangle \\
&= \sum_{i=1}^{n} U_{in}^{(p)} \frac{\delta_{in}}{(2n+p-1)U_{nn}^{(p)}} \\
&= \frac{1}{2n+p-1}.
\end{aligned} \tag{5.39}$$

□

Note that the first orthogonal function (i.e., for $n = 1$) is $\psi_1^{(p)}(x) = U_{11}^{(p)}|x|^{\frac{p}{2}}e^{j\angle x} = |x|^{\frac{p}{2}}e^{j\angle x}$. In this chapter, we use the special case where $p = 2$, such that $U_{ln} = U_{ln}^{(2)}$ and $\psi_n(x) = \psi_n^{(2)}(x)$; and such that the first polynomial, $\psi_1(x) = x$. When the input to the PA is small, the PA is approximately linear. Therefore, having the orthogonal basis with the

first polynomial being x, is appropriate. If the PA is not approximately linear for a small input signal, a different choice of p (other than $p = 2$) may be appropriate. The norm for the orthogonal polynomials used in this chapter, $\psi_n(x)$, (when $p = 2$) is $\frac{1}{2n+1}$.

CHAPTER VI

ORTHOGONAL POLYNOMIALS FOR COMPLEX GAUSSIAN PROCESSES

Power amplifiers are the major source of nonlinearity in communications systems. Such nonlinearity causes spectral regrowth as well as in-band distortion, which lead to adjacent channel interference and increased bit error rate. Polynomials are often used to model the nonlinear power amplifier or its predistortion linearizer. In this chapter, we present a novel set of orthogonal polynomials for baseband Gaussian input to replace the conventional polynomials, and show how they alleviate the numerical instability problem associated with the conventional polynomials. The orthogonal polynomials also provide an intuitive means of spectral regrowth analysis.

6.1 Introduction

Power amplifiers (PAs) are the major source of nonlinearity in communications systems. To achieve high efficiency from a given PA, the PA is often driven into its nonlinear region. When a non-constant envelope signal goes through a nonlinear PA, spectral regrowth (broadening) appears in the PA output, which in turn causes adjacent channel interference (ACI). Due to stringent limits on the ACI imposed by regulatory bodies, PA nonlinearity must be limited. PA linearization is often necessary to suppress spectral regrowth, contain adjacent channel interference, and reduce bit error rate (BER).

When modeling a nonlinear PA or the predistorter linearizing such a PA, the power series model, or the polynomial model, is often used (see e.g., [15, 36, 64, 113]). After extracting the polynomial coefficients of the PA, it is then possible to predict spectral regrowth present at the PA output [21, 50, 129, 131]. The polynomial model is widely used since it is relatively simple to construct and to analyze. However, it suffers a major draw back: computations associated with parameter estimation for the polynomial model tend to

be numerically unstable [91, 116]. Orthogonal polynomials offer a remedy to this problem; they also provide an intuitive means of spectral regrowth analysis, as we will see later in this chapter.

In [106], the use of orthogonal basis functions for Volterra nonlinear systems was discussed, under the assumption of real-valued Gaussian input. In [116], orthogonal polynomials were introduced to reduce numerical errors in the direct inversion of memoryless nonlinearities. Moreover, the spectrum at the output of a nonlinearity was expressed in terms of the orthogonal polynomial coefficients, but only the real-valued case was considered. In [91], a set of orthogonal polynomial basis functions was derived when the magnitude of the input signal is uniformly distributed. Orthogonal frequency division multiplexing (OFDM) signals, which are widely used in modern communication systems, are approximately complex-Gaussian distributed. In this chapter, we will derive a closed-form expression for orthogonal polynomials when the input is complex Gaussian distributed. We will replace the conventional polynomial model by the orthogonal polynomial model in order to improve numerical stability in parameter extraction for baseband models, as well as to simplify spectral regrowth analysis. Although the theoretical results are derived assuming Gaussian distribution of the input, as we will show in the chapter, the orthogonal polynomial model ensures better numerical stability even if the input distribution is not exactly Gaussian.

This chapter is organized as follows. The problems associated with PA modeling and predistortion are described in Section 6.2. Derivation of the orthogonal polynomials is presented and some of their properties are outlined in Section 6.3. In Section 6.4, we apply the orthogonal polynomials to carry out spectral regrowth analysis. In Section 6.5, we show the benefits of the orthogonal polynomials for improving numerical stability in predistortion linearization. Finally, conclusions are drawn in Section 6.6.

6.2 Problem Formulation

We say that $f(z)$ and $g(z)$ are orthogonal functions if and only if

$$E[f^*(z)g(z)] = 0, \tag{6.1}$$

where * stands for complex conjugation and $E[\cdot]$ denotes statistical expectation. We are prompted to consider orthogonal polynomials due to some numerical stability problems that we encountered in PA modeling and PA linearization.

6.2.1 The Numerical Instability Problem in PA Modeling

We model the baseband input/output relationship of a (quasi-)memoryless PA by [17]

$$\begin{aligned} y(t) &= \sum_{k=0}^{K} a_{2k+1}|z(t)|^{2k}z(t) \\ &= \sum_{k=0}^{K} a_{2k+1}\phi_{2k+1}(z(t)), \end{aligned} \tag{6.2}$$

where

$$\phi_{2k+1}(z(t)) = |z(t)|^{2k}z(t) \tag{6.3}$$

is the conventional polynomial basis function. For justification of the odd-order model and the conjugation pattern $[z(t)]^{k+1}[z^*(t)]^k = |z(t)|^{2k}z(t)$, see [17]. Next, define the following vector notations:

$$\begin{aligned} \mathbf{y} &= [y(t_1), y(t_2), \ldots, y(t_N)]^T, \\ \boldsymbol{\phi}_{2k+1}(\mathbf{z}) &= [|z(t_1)|^{2k}z(t_1), |z(t_2)|^{2k}z(t_2), \ldots, |z(t_N)|^{2k}z(t_N)]^T, \\ \boldsymbol{\Phi}(\mathbf{z}) &= [\boldsymbol{\phi}_1(\mathbf{z}), \boldsymbol{\phi}_3(\mathbf{z}), \ldots, \boldsymbol{\phi}_{2K+1}(\mathbf{z})], \\ \mathbf{a} &= [a_1, a_3, \ldots, a_{2K+1}]^T. \end{aligned} \tag{6.4}$$

Using this vector notation, eq. (6.2) can be written as

$$\mathbf{y} = \boldsymbol{\Phi}(\mathbf{z})\mathbf{a}. \tag{6.5}$$

Therefore, the least-squares (LS) estimate of $\mathbf{a}$ based on the PA input measurements $\mathbf{z}$ and the PA output measurements $\mathbf{y}$ is

$$\widehat{\mathbf{a}}_{LS} = \boldsymbol{\Phi}(\mathbf{z})^{\dagger}\mathbf{y}, \tag{6.6}$$

where $\boldsymbol{\Phi}(\mathbf{z})^{\dagger} \triangleq (\boldsymbol{\Phi}(\mathbf{z})^H\boldsymbol{\Phi}(\mathbf{z}))^{-1}\boldsymbol{\Phi}(\mathbf{z})^H$ is the pseudo-inverse of $\boldsymbol{\Phi}(\mathbf{z})$. The matrix $\boldsymbol{\Phi}(\mathbf{z})^H\boldsymbol{\Phi}(\mathbf{z})$ is often ill-conditioned, therefore the inversion of such matrix will incur numerical errors.

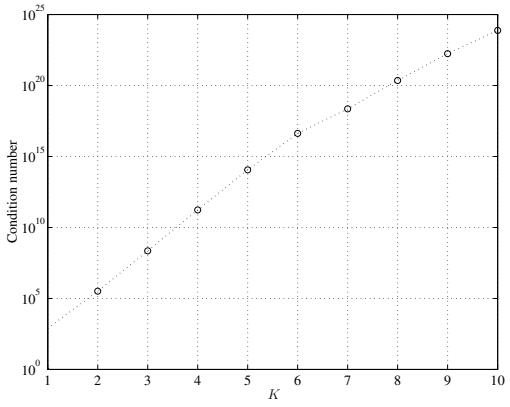


Figure 6.1: Condition number of $\mathbf{\Phi}(\mathbf{z})^H\mathbf{\Phi}(\mathbf{z})$ (averaged over 500 independent Monte-Carlo runs) as a function of K.

Consider as an example, a simple nonlinear model $y(t) = a_1 z(t) + a_3|z(t)|^2 z(t)$, with $a_1 = 15.0008 + 0.0908j$ and $a_3 = -23.0826 + 3.3133j$; i.e., the true parameter vector is $\mathbf{a} = [a_1, a_3, 0, \ldots, 0]^T$. The input $z(t)$ was an i.i.d. complex Gaussian distributed with zero-mean and standard deviation 0.16. The number of samples N was 1000. Define the condition number of a matrix as $\rho = \left|\frac{\lambda_{\max}}{\lambda_{\min}}\right|$ where $\lambda_{\max}$ and $\lambda_{\min}$ are respectively, the maximum and minimum eigenvalues of the matrix. The condition number of the resulting matrix $\mathbf{\Phi}(\mathbf{z})^H\mathbf{\Phi}(\mathbf{z})$ (averaged over 500 independent Monte-Carlo runs) is shown in Fig. 6.1. We observe that the condition number increases exponentially with K. Therefore, the higher the nonlinear model order, the more susceptible the solution (6.6) is to numerical errors. We would like to see a reduction of the condition number (and hence improvement in numerical stability) when the orthogonal polynomial basis, $\{\psi_1(z(t)), \psi_3(z(t)), \ldots, \psi_{2K+1}(z(t))\}$, is used to replace the conventional polynomial basis, $\{\phi_1(z(t)), \phi_3(z(t)), \ldots, \phi_{2K+1}(z(t))\}$.

We will show that with the orthogonal polynomials, we will be inverting a $\mathbf{\Psi}(\mathbf{z})^H\mathbf{\Psi}(\mathbf{z})$ matrix with the corresponding $E[\mathbf{\Psi}(\mathbf{z})^H\mathbf{\Psi}(\mathbf{z})] = \mathbf{I}$, if $z(t)$ is complex Gaussian distributed. Since an identity matrix has a condition number equal to 1, we expect a significant reduction in condition number as compared to Fig. 6.1.

6.2.2 The Numerical Instability Problem in Predistortion Linearization

We advocate the indirect learning architecture [46] as shown in Fig. 6.2. The baseband

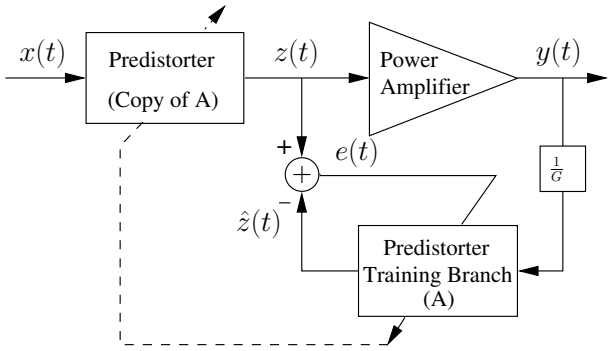


Figure 6.2: The indirect learning architecture. The predistorter parameters are obtained from the $z(t)$ and $y(t)$ samples, and then used to generate future $z(t)$ values from $x(t)$.

predistorter input is denoted by $x(t)$, the baseband predistorter output/PA input is denoted by $z(t)$, and the baseband PA output is denoted by $y(t)$. The feedback path labeled "Predistorter Training Branch" (block A) has $y(t)/G$ as its input, where G is the intended gain of the PA, and $\hat{z}(t)$ is its output. The actual predistorter (copy of A) is an exact copy of the predistorter training branch. When $y(t) = Gx(t)$, the error $e(t) = z(t) - \hat{z}(t)$ is 0. To reduce the error between $y(t)$ and $Gx(t)$, we choose the predistorter parameters that minimizes the error energy in $e(t)$. The benefit of the indirect learning architecture is that, instead of assuming a model for the PA, estimating the PA parameters and then constructing its inverse, we can go directly after the predistorter.

We use an iterative approach to obtain the predistorter coefficients. In Fig. 6.2, let us denote the predistorter input signal, the predistorter output/PA input signal, and the PA output signal, at the ith iteration by $x^{(i)}(t)$, $z^{(i)}(t)$, and $y^{(i)}(t)$, respectively. To initialize, the predistorter passes its input to the PA as is; i.e., $z^{(0)}(t) = x^{(0)}(t)$. Based on measured PA input $z^{(0)}(t)$ and output $y^{(0)}(t)$ values, a predistorter model $f^{(0)}(\cdot)$ is obtained so that $z^{(0)}(t) \approx f^{(0)}(y^{(0)}(t)/G)$. In the next iteration, the estimated predistorter is copied to the main branch to generate $z^{(1)}(t) = f^{(0)}(x^{(1)}(t))$. We then supply $z^{(1)}(t)$ as input to the PA to obtain output $y^{(1)}(t)$. Next, we update the predistorter based on the new PA input and output measurements with the objective of achieving $z^{(1)}(t) \approx f^{(1)}(y^{(1)}(t)/G)$. At the ith iteration, we implement the predistorter $f^{(i)}(\cdot)$ as

$$z^{(i+1)}(t) = f^{(i)}(x^{(i+1)}(t)),$$

and obtain an improved predistorter estimate $f^{(i+1)}(\cdot)$ by solving

$$z^{(i+1)}(t) \approx f^{(i+1)}(y^{(i+1)}(t)/G).$$

This procedure is repeated until the predistorter estimate has converged according to a pre-selected criterion. For notational simplicity however, we will omit $^{(i)}$ from now on.

We point out that before convergence, even if $x(t)$ is Gaussian, $y(t)$ is non-Gaussian. However, as we will show in Section 6.5.1, the orthogonal polynomials derived in this chapter can alleviate the condition number of the regressor matrix even if the signal distribution deviates from complex Gaussian.

The conventional polynomial predistorter is [113], [15]

$$z(t) = \sum_{k=0}^{K} b_{2k+1}\phi_{2k+1}\left(x(t)\right), \tag{6.7}$$

where $\phi_{2k+1}(\cdot)$ is defined as in (6.3). We estimate the coefficients $\{b_{2k+1}\}$ by applying a least-squares (LS) fit to the measured data $y(t)$ and $z(t)$ based on:

$$z(t) = \sum_{k=0}^{K} b_{2k+1}\phi_{2k+1}\left(\frac{y(t)}{G}\right), \tag{6.8}$$

so at convergence the concatenation of the predistorter and the PA will approximate a memoryless linear system with gain G. Using the $\mathbf{\Phi}(\cdot)$ notation introduced in (6.4) and define $\mathbf{b} = [b_1, b_3, \ldots, b_{2K+1}]^T$, the vector format of (6.8) is given by

$$\mathbf{z} = \mathbf{\Phi}(G^{-1}\mathbf{y})\mathbf{b}. \tag{6.9}$$

The LS estimator of $\mathbf{b}$ given by

$$\widehat{\mathbf{b}}_{LS} = \mathbf{\Phi}(G^{-1}\mathbf{y})^{\dagger}\mathbf{z} \tag{6.10}$$

requires the inversion of $\mathbf{\Phi}(G^{-1}\mathbf{y})^H\mathbf{\Phi}(G^{-1}\mathbf{y})$, which is often ill-conditioned. Similar to the PA modeling problem, we would like to consider the use of orthogonal polynomials to reduce the numerical instability problem in implementing (6.10). The PA modeling and predistorter construction problems are very similar by virtue of the indirect learning architecture. The difference is that in PA modeling, a model is assumed for the PA, whereas in predistortion, a model is assumed for the inverse of the PA.

6.3 Orthogonal Polynomials

Here, we would like to find an orthogonal basis which spans the same space as spanned by $\{\phi_1(z(t)), \ldots, \phi_{2K+1}(z(t))\}$. This is applicable to both PA modeling and predistorter design. We assume that $z(t)$ is complex Gaussian distributed. Obtaining closed-form expressions for orthogonal polynomials for an arbitrary distribution is generally a difficult problem and the derivations are not easily generalized.

We seek a set of orthogonal polynomials $\{\psi_{2k+1}(z)\}$ of the form:

$$\psi_{2k+1}(z) = \sum_{l=0}^{k} \mathbf{\Delta}_{lk}\ \phi_{2l+1}(z), \tag{6.11}$$

where the conventional polynomial basis function $\phi_{2l+1}(\cdot)$ is defined as in (6.3) and $\mathbf{\Delta} = [\mathbf{\Delta}_{lk}]$ is an upper triangular matrix. To find such orthogonal polynomials, we start with the following lemma for a zero mean, unit variance complex Gaussian process $\tilde{z}(t)$.

Lemma 1. *Define*

$$\tilde{\psi}_{2m+1}(\tilde{z}(t)) = \sum_{k=0}^{m} (-1)^{m-k}\ \frac{\sqrt{m+1}}{(k+1)!} \binom{m}{k}\ \phi_{2k+1}(\tilde{z}(t)), \tag{6.12}$$

where $\phi_{2k+1}(\cdot)$ is given by (6.3). If $\tilde{z}(t)$ is complex Gaussian distributed with zero mean and unit variance, then for any $l < m$ and for any τ, $\tilde{\psi}_{2m+1}(\tilde{z}(t))$ and $\phi_{2l+1}(\tilde{z}(t+\tau))$ are orthogonal; i.e.,

$$E[\tilde{\psi}^*_{2m+1}(\tilde{z}(t))\phi_{2l+1}(\tilde{z}(t+\tau))] = 0. \tag{6.13}$$

Proof. See Appendix 6.A. □

The above Lemma paves the way for proving that (6.12) forms an orthonormal basis.

Theorem 1. *Let $\tilde{z}(t)$ be a zero mean, unit variance, complex Gaussian process with auto-covariance function $c_{2\tilde{z}}(\tau)$. Then $\{\tilde{\psi}_1(\tilde{z}(t)), \tilde{\psi}_3(\tilde{z}(t)), \ldots, \tilde{\psi}_{2K+1}(\tilde{z}(t))\}$ defined in (6.12) forms an orthonormal basis. The following two properties hold:*

1. *For any $n \neq m$ and for any τ, the random processes $\tilde{\psi}_{2m+1}(\tilde{z}(t))$ and $\tilde{\psi}_{2n+1}(\tilde{z}(t+\tau))$ are orthogonal to each other.*

2. *The auto-correlation (auto-covariance) function of the random process* $\tilde{\psi}_{2m+1}(\tilde{z}(t))$ *is* $|c_{2\tilde{z}}(\tau)|^{2m}c_{2\tilde{z}}(\tau)$. *Since* $c_{2\tilde{z}}(0)=1$, $E[|\tilde{\psi}_{2m+1}(\tilde{z}(t))|^2]=1$.

Proof. See Appendix 6.B. □

Table 6.1: The first five Orthonormal polynomials for unit variance $\tilde{z}$.

$\tilde{\psi}_1(\tilde{z})$	$=$	$\tilde{z}$
$\tilde{\psi}_3(\tilde{z})$	$=$	$\sqrt{2}\left(-1+\frac{1}{2}\lvert\tilde{z}\rvert^2\right)\tilde{z}$
$\tilde{\psi}_5(\tilde{z})$	$=$	$\sqrt{3}\left(1-\lvert\tilde{z}\rvert^2+\frac{1}{6}\lvert\tilde{z}\rvert^4\right)\tilde{z}$
$\tilde{\psi}_7(\tilde{z})$	$=$	$\sqrt{4}\left(-1+\frac{3}{2}\lvert\tilde{z}\rvert^2-\frac{1}{2}\lvert\tilde{z}\rvert^4+\frac{1}{24}\lvert\tilde{z}\rvert^6\right)\tilde{z}$
$\tilde{\psi}_9(\tilde{z})$	$=$	$\sqrt{5}\left(1-2\lvert\tilde{z}\rvert^2+\lvert\tilde{z}\rvert^4-\frac{1}{6}\lvert\tilde{z}\rvert^6+\frac{1}{120}\lvert\tilde{z}\rvert^8\right)\tilde{z}$

In Table 6.1, the first five orthonormal polynomials obtained from (6.12) are shown.

We remark that $\tilde{\psi}_{2m+1}(\tilde{z})$ is related to the associated Laguerre polynomials by

$$\tilde{\psi}_{2m+1}(\tilde{z})=\sqrt{m+1}\;L_m^1(|\tilde{z}|^2)\tilde{z}, \tag{6.14}$$

where $L_n^k(t)\triangleq\sum_{m=0}^{n}\frac{(-1)^m}{m!}\binom{n+k}{m+k}t^m$ are the associated Laguerre polynomials [5, p. 775].

For a zero-mean complex Gaussian process $z(t)$ with variance $\sigma_z^2\neq 1$, orthogonal polynomials can be obtained using the following corollary.

Corollary 1. *If* $z(t)$ *is a zero-mean complex Gaussian process with variance* σ_z^2, *then*

$$\psi_{2m+1}(z(t))=\tilde{\psi}_{2m+1}(z(t)/\sigma_z)=\sum_{k=0}^{m}(-1)^{m-k}\frac{\sqrt{m+1}}{(k+1)!}\binom{m}{k}\left|\frac{z(t)}{\sigma_z}\right|^{2k}\frac{z(t)}{\sigma_z}, \tag{6.15}$$

$$=\sum_{k=0}^{m}\frac{(-1)^{m-k}}{\sigma_z^{2k+1}}\;\frac{\sqrt{m+1}}{(k+1)!}\binom{m}{k}\;\phi_{2k+1}(z(t)), \tag{6.16}$$

is orthogonal to $\psi_{2n+1}(z(t+\tau))$, $\forall m\neq n$, $\forall\tau$.

We remark also that the orthogonality in (6.17) is stronger than the orthogonality in (6.1), since it holds even at different time-delayed terms (i.e., $\tau\neq 0$):

$$E\left[\psi_{2k+1}^*(z(t))\psi_{2l+1}(z(t+\tau))\right]=0,\quad \forall k\neq l,\forall\tau. \tag{6.17}$$

Next, we comment on our expression and other known orthogonal polynomials.

6.3.1 Wiener G-Functionals

Similar to our orthogonal polynomials, Wiener's G-functionals also form an orthogonal basis [106], [75]. However, there are major differences between the two. First, the Wiener G-functionals are derived for real-valued Gaussian input whereas our orthogonal polynomials are derived for complex-valued Gaussian input. Second, our orthogonal polynomials are for the odd-order baseband model whereas the Wiener G-functionals include all-order Volterra kernels. Most importantly, the Wiener G-functionals are derived iteratively using the Gram-Schmidt procedure (e.g., [74]) whereas our expression (6.15) is in closed-form.

6.3.2 Hermite Polynomials

Table 6.2: The first six Hermite-related orthonormal polynomials

$p_0(w)$	$=$	1
$p_1(w)$	$=$	w
$p_2(w)$	$=$	$\frac{1}{\sqrt{2}}(w^2-1)$
$p_3(w)$	$=$	$\frac{1}{\sqrt{6}}(w^3-3w)$
$p_4(w)$	$=$	$\frac{1}{\sqrt{24}}(w^4-6w^2+3)$
$p_5(w)$	$=$	$\frac{1}{\sqrt{120}}(w^5-10w^3+15w)$

Assume that w is a zero mean real-valued Gaussian random variable with unit variance. For memoryless polynomial nonlinearities, the Wiener G-functionals are simplified to the following polynomials:

$$p_n(w) = \frac{(-1)^n}{\sqrt{2^n n!}} H_n\left(\frac{w}{\sqrt{2}}\right), \tag{6.18}$$

where $H_n(w)$ are the Hermite polynomials given by $H_n(w) = (-1)^n e^{w^2} \frac{\partial^n}{\partial w^n} e^{-w^2}$ [5, Ch. 22, pp. 771-802]. Hermite polynomials have the following property:

$$\int_{-\infty}^{\infty} H_n(w) H_m(w) e^{-w^2} dw = \begin{cases} \sqrt{\pi} n! 2^n, & n = m, \\ 0, & n \neq m. \end{cases} \tag{6.19}$$

Based on (6.18)-(6.19), we infer that

$$\int_{-\infty}^{\infty} p_n(w) p_m(w) \frac{1}{\sqrt{2\pi}} e^{-\frac{w^2}{2}} dw = \begin{cases} 1, & n = m, \\ 0, & n \neq m. \end{cases} \tag{6.20}$$

Therefore, for w Gaussian distributed with zero mean and unit variance, the set of polynomials $\{p_n(w)\}$ forms an orthonormal basis. The first six $\{p_n(w)\}$ functions are presented in Table 6.2. Note that this set of orthogonal polynomials includes even and odd order terms and assumes a real-valued standard Gaussian distribution.

6.4 Spectral Analysis

Next, let us examine the PA's input/output relationship represented using the orthogonal polynomial model,

$$\begin{aligned} y(t) &= \sum_{k=0}^{K} \alpha_{2k+1} \psi_{2k+1}(z(t)) \\ &= \sum_{k=0}^{K} \alpha_{2k+1} \tilde{\psi}_{2k+1} \left(\frac{z(t)}{\sigma_z} \right) \\ &= \sum_{k=0}^{K} y_{2k+1}(t). \end{aligned} \tag{6.21}$$

The auto-covariance function $c_{2y}(\tau)$ is given by

$$c_{2y}(\tau) = \sum_{k=0}^{K} \sum_{l=0}^{K} \mathrm{cum}\{y_{2k+1}^{*}(t), y_{2l+1}(t+\tau)\}. \tag{6.22}$$

From the first property of Theorem 1, we infer that $\mathrm{cum}\{y_{2k+1}^{*}(t), y_{2l+1}(t+\tau)\} = 0$ for $k \neq l$, and hence the cross-terms in (6.22) vanish. The auto-covariance function of $y_{2k+1}(t)$ can be computed using the second property of Theorem 1, resulting in

$$c_{2y}(\tau) = \sum_{k=0}^{K} |\alpha_{2k+1}|^2 \left| \frac{c_{2z}(\tau)}{\sigma_z^2} \right|^{2k} \frac{c_{2z}(\tau)}{\sigma_z^2}. \tag{6.23}$$

Applying the Fourier transform to both sides of (6.23), we obtain the PSD at the output of the PA:

$$S_{2y}(f) = \sum_{k=0}^{K} |\alpha_{2k+1}|^2 \underbrace{\frac{S_{2z}(f)}{\sigma_z^2} * \cdots * \frac{S_{2z}(f)}{\sigma_z^2}}_{k+1} * \underbrace{\frac{S_{2z}(-f)}{\sigma_z^2} * \cdots * \frac{S_{2z}(-f)}{\sigma_z^2}}_{k}. \tag{6.24}$$

Therefore, the orthogonal polynomial PA model coefficient α_{2k+1} is directly linked to the amount of spectral regrowth generated by the $(2k+1)$-convolution term in (6.24). Since $\psi_{2k+1}(\cdot)$ and $\phi_{2k+1}(\cdot)$ are related through (6.15), it can be shown that the conventional and

the orthogonal polynomial PA coefficients are related through

$$\alpha_{2k+1} = \sum_{l=k}^{K} \frac{(l+1)!}{\sqrt{k+1}} \binom{l}{k} \sigma_z^{2l+1} a_{2l+1}. \tag{6.25}$$

A relatively linear PA will have a relatively large $|\alpha_1|$ and relatively small $|\alpha_3|$, $|\alpha_5|$, and so on. Therefore, the coefficients in the orthogonal polynomial PA model (6.21) have clear meanings in the context of spectral regrowth. By inspecting $|\alpha_1|$, $|\alpha_3|$, $|\alpha_5|$, ..., we immediately have a sense of the severity of spectral regrowth. In contrast, the a_1, a_3, a_5 etc. coefficients from a conventional polynomial PA model do not provide such direct link to spectral regrowth.

The closed-form expression for the PA output PSD (6.24) can be used to compute the channel power ratio (CPR), which is defined as $\mathrm{CPR}(f) = \frac{\int_{f_1}^{f_2} S_{2y}(f)df}{\int_{f_3}^{f_4} S_{2y}(f)df}$ [50], where $[f_1, f_2]$ is in the main channel and $[f_3, f_4]$ is in another channel. Here, we take $f_4 - f_3 = f_2 - f_1$ and let $[f_1, f_2]$ be an infinitely small interval surrounding the zero-frequency, and $[f_3, f_4]$ surrounding frequency f. Then

$$\mathrm{CPR}(f) = \frac{S_{2y}(f)}{S_{2y}(0)}. \tag{6.26}$$

As an example, consider a rectangular shaped PA input PSD, $S_{2z}(f) = u(f+\frac{\Delta}{2}) - u(f-\frac{\Delta}{2})$, where $u(\cdot)$ is the step function. From (6.24), the PA output spectrum can be shown to be

$$S_{2y}(f) = \sum_{k=0}^{K} |\alpha_{2k+1}|^2 \sum_{l=0}^{2k+1} (-1)^k \binom{2k+1}{l} \times \frac{u\left(\frac{f}{\Delta} - l + k + \frac{1}{2}\right)\left(\frac{f}{\Delta} - l + k + \frac{1}{2}\right)^{2k}}{2k!}. \tag{6.27}$$

Assuming that $|\alpha_1| \gg |\alpha_3| \gg \cdots \gg |\alpha_{2K+1}|$ (a reasonable assumption for mildly nonlinear PAs),

$$\mathrm{CPR}\left(\frac{\Delta}{2}(2n-1)^+\right) \approx \frac{|\alpha_1|^2 (2n)!}{|\alpha_{2n+1}|^2}. \tag{6.28}$$

for $n \geq 1$. Note that $\frac{\Delta}{2}^+$ is just outside the main channel, $\frac{3\Delta}{2}^+$ is just outside the adjacent channel, $\frac{5\Delta}{2}^+$ is just outside the alternate channel, so $\mathrm{CPR}\left(\frac{\Delta}{2}^+\right)$ is a measure for the adjacent channel power ratio; $\mathrm{CPR}\left(3\frac{\Delta}{2}^+\right)$ can be used to gauge the alternate channel

power ratio, and so on. In dB scale, (6.28) turns into

$$\begin{aligned}&\mathrm{CPR}\left(\tfrac{\Delta}{2}(2n-1)^{+}\right)_{[\mathrm{dB}]} \approx \\ &\qquad |\alpha_1|^2_{[\mathrm{dB}]} - |\alpha_{2n+1}|^2_{[\mathrm{dB}]} + 10\log_{10}(2n)!. \end{aligned} \tag{6.29}$$

Figure 6.3 shows the PSD of a PA obeying model (6.21) with coefficients satisfying

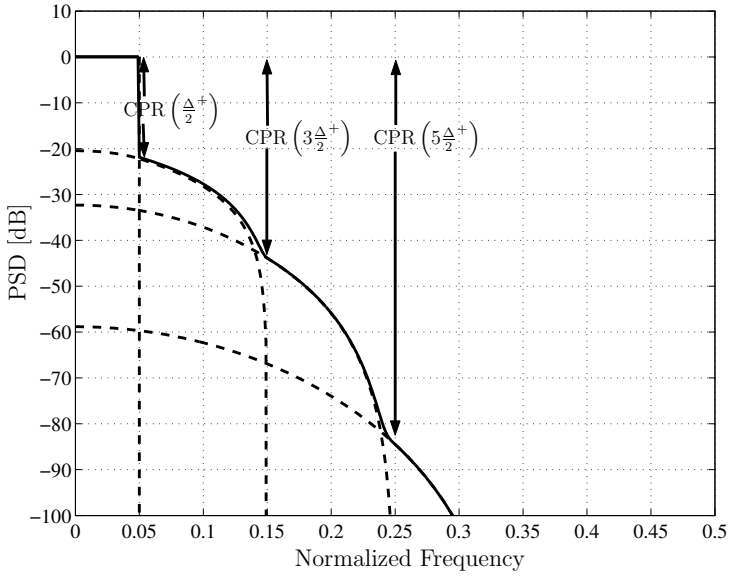


Figure 6.3: From top to bottom, the dashed lines depict the PSDs generated by $y_1(t)$, $y_3(t)$, $y_5(t)$, and $y_7(t)$, respectively. The PSD of $y(t)$ is shown as the solid line. Various CPR measurements are indicated.

$|\alpha_1|^2 = 0$dB, $|\alpha_3|^2 = -19.2$dB, $|\alpha_5|^2 = -30.1$dB, and $|\alpha_7|^2 = -55.9$dB. To verify the CPR expression in (6.29), take for example, $\mathrm{CPR}(5\frac{\Delta}{2}^{+})_{[\mathrm{dB}]}$. Here with $\Delta = 0.1$, we obtain $\mathrm{CPR}\,(0.25^{+}) \approx 0 - (-55.9) + 28.6 = 84.5$dB. We conclude that by mere observation of the magnitudes of the orthogonal polynomial coefficients, we can have a good picture of the spectral regrowth generated by the nonlinear PA.

Next, we examine the so-called (conventional) memory polynomial PA model whose input/output relationship is given by

$$\begin{aligned} y(t) &= \sum_{k=0}^{K} a_{2k+1}(t) * \Big[|z(t)|^{2k} z(t)\Big] \\ &= \sum_{k=0}^{K} a_{2k+1}(t) * \phi_{2k+1}(z(t)), \end{aligned} \tag{6.30}$$

where $*$ is the convolution operator. This model has been shown to be effective for PA modeling [65] and predistorter design [39] when the nonlinear high power or wideband PA exhibits memory effects.

Expressing the same $y(t)$ in terms of the orthogonal memory polynomial basis functions, we write

$$\begin{aligned} y(t) &= \sum_{k=0}^{K} \alpha_{2k+1}(t) * \psi_{2k+1}(z(t)) \\ &= \sum_{k=0}^{K} \alpha_{2k+1}(t) * \tilde{\psi}_{2k+1}\left(\frac{z(t)}{\sigma_z}\right). \end{aligned} \tag{6.31}$$

Using Theorem 1, the auto-covariance $c_{2y}(\tau)$ of the PA output, $y(t)$, can be shown to relate to that of the PA input through

$$\begin{aligned} c_{2y}(\tau) = \sum_{k=0}^{K} & \alpha_{2k+1}(\tau) * \alpha^*_{2k+1}(-\tau) \\ & * \left[\left| \frac{c_{2z}(\tau)}{\sigma_z^2} \right|^{2k} \frac{c_{2z}(\tau)}{\sigma_z^2} \right]. \end{aligned} \tag{6.32}$$

Applying the Fourier transform to both sides of (6.32) yields

$$S_{2y}(f) = \sum_{k=0}^{K} |\mathcal{A}_{2k+1}(f)|^2 \underbrace{\frac{S_{2z}(f)}{\sigma_z^2} * \cdots * \frac{S_{2z}(f)}{\sigma_z^2}}_{k+1} * \underbrace{\frac{S_{2z}(-f)}{\sigma_z^2} * \cdots * \frac{S_{2z}(-f)}{\sigma_z^2}}_{k}, \tag{6.33}$$

where $\mathcal{A}_{2k+1}(f)$ is the Fourier transform of $\alpha_{2k+1}(t)$. In the memoryless case, $\alpha_{2k+1}(t) = \alpha_{2k+1}\delta(t)$, and (6.33) becomes (6.24).

6.5 Numerical Stability Improvement

Let us define the $\mathbf{\Psi}(\mathbf{z})$ matrix similar to the way the $\boldsymbol{\phi}(\mathbf{z})$ matrix is defined in (6.5) (i.e., replace ϕ by ψ), and define $\boldsymbol{\alpha} = [\alpha_1, \alpha_3, \ldots, \alpha_{2K+1}]^T$. Eq. (6.21) can be rewritten as

$$\mathbf{y} = \mathbf{\Psi}(\mathbf{z})\ \boldsymbol{\alpha}. \tag{6.34}$$

The LS solution for the orthogonal polynomial coefficients $\boldsymbol{\alpha}$ is

$$\widehat{\boldsymbol{\alpha}}_{LS} = \mathbf{\Psi}(\mathbf{z})^{\dagger}\ \mathbf{y}. \tag{6.35}$$

Although (6.6) and (6.34) are theoretically equivalent; i.e., $\mathbf{y} = \mathbf{\Phi}(\mathbf{z})\mathbf{a} = \mathbf{\Psi}(\mathbf{z})\boldsymbol{\alpha}$, because the condition number of $E[\mathbf{\Psi}(\mathbf{z})^H\mathbf{\Psi}(\mathbf{z})]$ is much smaller than that of $E[\mathbf{\Phi}(\mathbf{z})^H\mathbf{\Phi}(\mathbf{z})]$, the solution in (6.35) will have better numerical properties than (6.6).

Orthogonal polynomials are also advantageous when it comes to PA predistortion. To overcome the numerical instability problem, we propose to replace the conventional polynomial predistorter model (6.7) by the following model based on the orthogonal polynomials:

$$z(t) = \sum_{k=0}^{K} \beta_{2k+1}\psi_{2k+1}\left(x(t)\right). \tag{6.36}$$

Recall that the argument of $\tilde{\psi}_{2k+1}(\cdot)$ in (6.12) is $\tilde{z}(t)$, which is assumed to have unit variance. In (6.15), $\psi_{2k+1}(\cdot)$ is defined through $\tilde{\psi}_{2k+1}(\cdot)$, whose argument $z(t)/\sigma_z$ is standardized. As a result,

$$\psi_{2k+1}(y(t)/G) = \tilde{\psi}_{2k+1}\Big(\frac{y(t)/G}{\sigma_y/G}\Big) = \tilde{\psi}_{2k+1}\Big(\frac{y(t)}{\sigma_y}\Big) = \psi_{2k+1}(y(t)).$$

This means that $\psi_{2k+1}(\cdot)$ is "insensitive" to the multiplicative constant and thus the intended gain G has to be realized via a two-step procedure. First, given measured data $y(t)$ and $z(t)$, we obtain the LS estimates

$$\widehat{\boldsymbol{\gamma}}_{LS} = \mathbf{\Psi}(\mathbf{y})^{\dagger}\mathbf{z}, \tag{6.37}$$

of $\boldsymbol{\gamma} = [\gamma_1, \gamma_3, \ldots, \gamma_{2K+1}]^T$ for $z(t) = \sum_{k=0}^{K} \gamma_{2k+1}\psi_{2k+1}\left(y(t)\right)$. If we plug the $\{\gamma_{2k+1}\}$ coefficients directly into the predistorter as in $z(t) = \sum_{k=0}^{K} \gamma_{2k+1}\psi_{2k+1}(x(t))$, we will not be able to realize the intended gain G. Instead, we map $\{\gamma_{2k+1}\}$ to $\{\beta_{2k+1}\}$ via

$$\boldsymbol{\beta} = \mathbf{M}\boldsymbol{\gamma}, \tag{6.38}$$

where $\mathbf{M}$ is a $(K+1)$-by-$(K+1)$ matrix with entries

$$\mathbf{M}_{kl} = \begin{cases} \sqrt{\frac{k+1}{l+1}}\binom{l+1}{k+1}\epsilon^{2k+1}(\epsilon^2-1)^{l-k}, & l \geq k, \\ 0, & l < k, \end{cases} \tag{6.39}$$

and $\epsilon = \frac{G\sigma_x}{\sigma_y}$, and then implement the predistorter as

$$z(t) = \sum_{k=0}^{K} \beta_{2k+1}\psi_{2k+1}(x(t)), \tag{6.40}$$

or $\mathbf{z} = \mathbf{\Psi}(\mathbf{x})\boldsymbol{\beta}$ in matrix form. With this predistorter, the resulting linearized PA gain will be close to G. Again in theory, the predistorters in (6.7) and (6.40) are equivalent; i.e., $\mathbf{z} = \mathbf{\Phi}(\mathbf{x})\mathbf{b} = \mathbf{\Psi}(\mathbf{x})\boldsymbol{\beta}$, but in practice, the solution in (6.37)-(6.38); i.e,

$$\widehat{\boldsymbol{\beta}}_{LS} = \mathbf{M}\widehat{\boldsymbol{\gamma}}_{LS}, \qquad \widehat{\boldsymbol{\gamma}}_{LS} = \mathbf{\Psi}(\mathbf{y})^{\dagger}\mathbf{z}, \tag{6.41}$$

is numerically superior to the solution in (6.10). Theoretical justification of this two-step procedure is provided in Appendix 6.C.

6.5.1 Deviation from the Gaussian Assumption

Since our orthogonal polynomials are derived with the complex Gaussian distribution in mind, one may wonder what happens if the distribution of $z(t)$ deviates from the Gaussian assumption?

Recall that if $z(t)$ is complex Gaussian distributed, its amplitude $r(t) = |z(t)|$ is Rayleigh distributed, its phase is uniformly distributed in $[-\pi, \pi)$, and the amplitude and the phase are mutually independent. To examine the robustness of the orthogonal polynomials in reducing the condition number when $z(t)$ is not complex Gaussian distributed, we consider $r(t) = |z(t)|$ central Chi-square distributed with varying degrees of freedom. The phase of $z(t)$ is still uniformly distributed in $[-\pi, \pi)$ and is independent of the amplitude. Since $r(t)$ is not Rayleigh distributed, $z(t)$ is not complex Gaussian any more. Figure 6.4(b) shows the Rayleigh and the $\chi^2(p)$ PDFs with $p = 4, 8, 12$, respectively. The $\chi^2(p)$ PDF is given by

$$f_r(r) = \frac{\left(\frac{r}{\delta}\right)^{\frac{p}{2}-1} e^{-\frac{r}{\delta}}}{\delta\Gamma(\frac{p}{2})}, \qquad \delta^2 = \frac{\Gamma(\frac{p}{2})}{\Gamma(\frac{p}{2}+2)}. \tag{6.42}$$

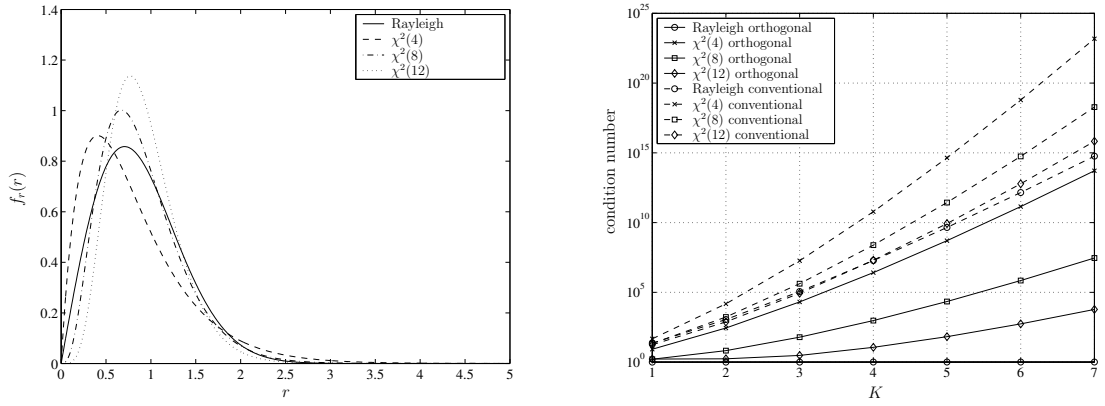
It can be shown that the kth-order moment of r is

$$E[r^k] = \frac{\Gamma(\frac{p}{2}+k)}{\Gamma(\frac{p}{2})\left((\frac{p}{2}+1)\frac{p}{2}\right)^{\frac{k}{2}}}. \tag{6.43}$$

Specifically, $E[r^2] = 1$, $\forall p$. This ensures that the corresponding $z(t)$ has unit variance, so that we conduct a fair comparison with a zero mean, unit variance complex Gaussian process.

Our objective is to show that the condition number of $E[\mathbf{\Psi}(\mathbf{z})^H\mathbf{\Psi}(\mathbf{z})]$ is much smaller than that of $E[\mathbf{\Phi}(\mathbf{z})^H\mathbf{\Phi}(\mathbf{z})]$. Since $\phi_{2k+1}(z) = |z|^{2k}z$, it follows that the (l,m)th entry of the $E[\mathbf{\Phi}(\mathbf{z})^H\mathbf{\Phi}(\mathbf{z})]$ matrix is $E[r^{2(l+m-1)}]$, which can be calculated using (6.43). Thus the $E[\mathbf{\Phi}(\mathbf{z})^H\mathbf{\Phi}(\mathbf{z})]$ matrix can be obtained. since $\mathbf{\Psi}(\mathbf{z}) = \mathbf{\Phi}(\mathbf{z})\mathbf{\Delta}$ where $\mathbf{\Delta}$ is an upper-triangular matrix (c.f. (6.11)), we can calculate $E[\mathbf{\Psi}(\mathbf{z})^H\mathbf{\Psi}(\mathbf{z})] = \mathbf{\Delta}^H E[\mathbf{\Phi}(\mathbf{z})^H\mathbf{\Phi}(\mathbf{z})]\mathbf{\Delta}$ as well. In Fig. 6.4(a), we plot the condition number of $E[\mathbf{\Phi}(\mathbf{z})^H\mathbf{\Phi}(\mathbf{z})]$ (dashed lines) and that of $E[\mathbf{\Psi}(\mathbf{z})^H\mathbf{\Psi}(\mathbf{z})]$ (solid lines) as a function of K (the polynomial order is $2K+1$). When $z(t)$ is Gaussian distributed, the condition number of $E[\mathbf{\Psi}(\mathbf{z})^H\mathbf{\Psi}(\mathbf{z})]$ is 1. Although when $z(t)$ is non-Gaussian distributed, the condition number of $E[\mathbf{\Psi}(\mathbf{z})^H\mathbf{\Psi}(\mathbf{z})]$ is larger than 1, in all cases, use of the orthogonal polynomials resulted in orders of magnitude of reduction in the condition number.

Keep in mind that our goal is to avoid numerical instability in least-squares parameter estimation and utilizing orthogonal polynomials is only a means to that end. Maintaining exact orthogonality in $\{\psi_{2k+1}(z)\}$ for every $z(t)$ distribution is not our goal; keeping the condition number of $\mathbf{\Psi}(\mathbf{z})^H\mathbf{\Psi}(\mathbf{z})$ generally low is. Our predistortion example next further illustrates the effectiveness of our orthogonal polynomials in maintaining numerical stability even when $z(t)$ is not Gaussian distributed.



(a) (b)

Figure 6.4: (a) PDF of $r(t) = |z(t)|$. When $r(t)$ has the Rayleigh distribution, the corresponding $z(t)$ is complex Gaussian. Central χ^2 distributions with degrees of freedom $p = 4, 8, 12$ are also shown: for these $r(t)$ distributions, the corresponding $z(t)$ is not Gaussian any more. (b) The condition number of $E[\mathbf{\Psi}(\mathbf{z})^H\mathbf{\Psi}(\mathbf{z})]$ is shown as solid lines; the condition number of $E[\mathbf{\Phi}(\mathbf{z})^H\mathbf{\Phi}(\mathbf{z})]$ is shown as dashed lines. Orthogonal polynomials lowered the condition number for all these distributions.

6.5.2 Numerical Examples

The simulation environment in this section is C, floating point data with 32-bit accuracy. This precision is used in high-end digital signal processors (DSPs). Other DSPs may use a fixed-point operation which is known to be less accurate than the floating-point format. An example is given to demonstrate how the numerical problems associated with estimating the predistorter coefficients affect the performance of the predistorter in terms of spectral regrowth suppression. We utilize the indirect learning architecture to perform predistortion linearization. The predistorter's input, $x(t)$, is a three carrier Universal Mobile Telephone Service (UMTS) signal [30]. Both conventional polynomials and orthogonal polynomials are considered for the construction of the predistorter of order 9 (i.e., $K = 4$).

The PA is assumed to have the following input/output relationship:

$$y(t) = \left(\eta_1 \tan^{-1}(\xi_1|z(t)|) + \eta_2 \tan^{-1}(\xi_2|z(t)|)\right) e^{j\angle z(t)}, \tag{6.44}$$

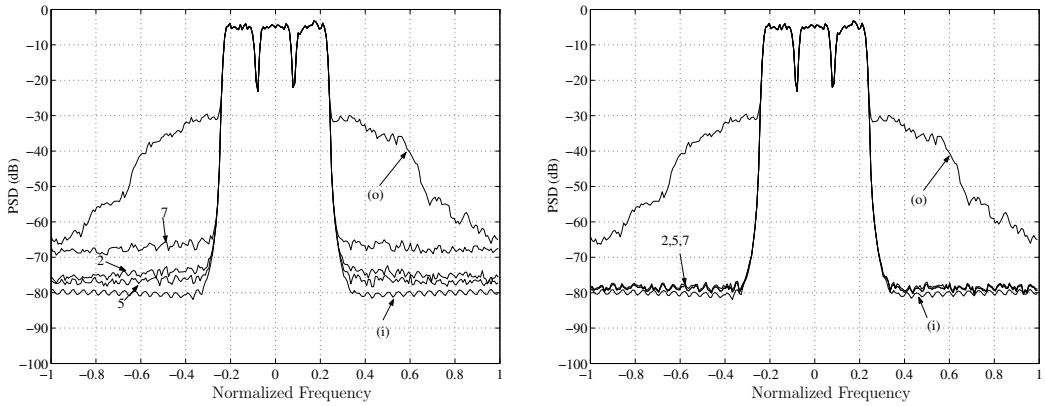
where $\eta_1 = 8.0034 - j4.6116$, $\eta_2 = -3.7717 + j12.0376$, $\xi_1 = 2.2690$, and $\xi_2 = 0.8234$. This PA model fits well measured data from an actual class AB PA[1]. The intended linearized PA gain is set to $G = 11$.

Fig. 6.5(a) shows the PSD at the output of the PA for the conventional polynomial predistorter (6.7) of order 9. The PSD is presented for iterations 2, 5, and 7, respectively and shows no sign of convergence. In contrast, Fig. 6.5(b) shows the PSD at the output of the PA for the orthogonal polynomial predistorter (6.36) with the same nonlinearity order 9. In this case, the predistorter shows stability and fully suppresses the spectral regrowth. We remark that although the UMTS signal is not exactly Gaussian distributed, the orthogonal polynomials (6.15) derived for Gaussian processes are still beneficial.

6.6 Conclusions

PA is a main source of nonlinearity in a communications system, and generates in-band distortion as well as adjacent channel interference for non-constant envelope signals. The

[1]Many different PA models have been proposed in the literature. Using the indirect learning architecture, the polynomial predistorter can be used to compensate for a variety of nonlinear PA models.



(a) Conventional polynomial predistorter (b) Orthogonal polynomial predistorter

Figure 6.5: Ninth-order ($K = 4$) polynomial predistorters are used to linearize the PA model in (6.44). Line (i) corresponds to the PSD of the input signal; line (o) corresponds to the PSD of the output signal without predistortion. Lines 2, 5, 7 indicate the PSDs of the linearized PA output at the 2nd, 5th, and 7th iterations. The output PSDs have been lowered by $20 \log_{10} |G|$ dB to be overlaid on top of the input PSD to facilitate spectral regrowth comparisons. The conventional polynomial predistorter in (a) did not converge due to numerical instability problems. In contrast, the orthogonal polynomial predistorter in (b) converged and demonstrated superior spectral regrowth suppression capability.

polynomial model has been widely used in PA modeling as well as predistortion linearization design. For either task, obtaining a least squares solution for the model parameters can be numerically challenging due to ill-conditioning of the regressor matrix to be inverted. For bandpass communications signals, we have obtained novel closed-form expression of baseband orthogonal polynomials for complex Gaussian input. We demonstrated that the orthogonal polynomials are effective in alleviating the numerical instability problem in least squares parameter estimation, even if the input signal is not exactly Gaussian distributed. Furthermore, with orthogonal polynomials, spectral analysis of the nonlinear PA becomes a straightforward task. The orthogonal polynomial PA parameters directly reveal the severity of spectral regrowth, as measured for example, by the adjacent channel power ratio.

Acknowledgment

The authors wish to thank Hua Qian for his help with some of the simulations.

6.A Proof of Lemma 1

First, we substitute (6.12) into the left-hand side of (6.13) to obtain,

$$E[\tilde{\psi}^*_{2m+1}(\tilde{z}(t))\phi_{2l+1}(\tilde{z}(t+\tau))] = \sqrt{m+1}\sum_{k=0}^{m}(-1)^{m-k}\ \frac{1}{(k+1)!}\binom{m}{k}\times \mathrm{cum}\{\phi^*_{2k+1}(\tilde{z}(t)), \phi_{2l+1}(\tilde{z}(t+\tau))\}. \tag{6.45}$$

In [131], we proved that if $z(t)$ is complex Gaussian distributed, then

$$\mathrm{cum}\{\phi^*_{2k+1}(z(t)), \phi_{2l+1}(z(t+\tau))\} = \sum_{s=0}^{\min(k,l)}\ \frac{1}{s+1}\binom{k}{s}\binom{l}{s}(k+1)!(l+1)! |c_{2z}(\tau)|^{2s}c_{2z}(\tau)[c_{2z}(0)]^{k+l-2s}. \tag{6.46}$$

Note that $\tilde{z}(t)$ has variance $c_{2\tilde{z}}(0) = 1$ here. Substituting (6.46) into (6.45), we rewrite (6.45) as

$$(l+1)!\sqrt{m+1}\sum_{k=0}^{m}(-1)^{m-k}\ \binom{m}{k}\times \sum_{s=0}^{\min(k,l)}\ \frac{1}{s+1}\binom{k}{s}\binom{l}{s}|c_{2\tilde{z}}(\tau)|^{2s}c_{2\tilde{z}}(\tau). \tag{6.47}$$

Since $l < m$, $\sum_{k=0}^{m}\sum_{s=0}^{\min(k,l)} = \sum_{s=0}^{l}\sum_{k=s}^{m}$. We further rewrite (6.47) as

$$(l+1)!\sqrt{m+1}\sum_{s=0}^{l}\ \frac{1}{s+1}\ \binom{l}{s}\ |c_{2\tilde{z}}(\tau)|^{2s}c_{2\tilde{z}}(\tau)\times \sum_{k=s}^{m}(-1)^{m-k}\binom{m}{k}\binom{k}{s}. \tag{6.48}$$

Let us focus on the k-dependent terms in (6.48), namely, $\sum_{k=s}^{m}(-1)^{m-k}\binom{m}{k}\binom{k}{s}$. Since $s \le l < m$, we have $m > s$. It is straightforward to show that

$$\begin{aligned}
&\sum_{k=s}^{m}(-1)^{m-k}\binom{m}{k}\binom{k}{s}\\
&= (-1)^{m-s}\binom{m}{s}\sum_{k=s}^{m}\binom{m-s}{k-s}(-1)^{s-k}\\
&= (-1)^{m-s}\binom{m}{s}\sum_{k'=0}^{m-s}\binom{m-s}{k'}(-1)^{k'}\\
&= (-1)^{m-s}\binom{m}{s}(1+(-1))^{m-s} = 0, \quad \forall m > s.
\end{aligned} \tag{6.49}$$

Substituting (6.49) into (6.48), we conclude that for $l < m$, the polynomials $\tilde{\psi}_{2m+1}(\tilde{z}(t))$ and $\tilde{\phi}_{2l+1}(\tilde{z}(t+\tau))$ are orthogonal thus Proving Lemma 1.

6.B Proof of Theorem 1

Define an upper-triangular matrix $\mathbf{U}$ as:

$$\mathbf{U}_{km} = \begin{cases} (-1)^{m-k}\frac{\sqrt{m+1}}{(k+1)!}\binom{m}{k}, & k \leq m, \\ 0, & k > m. \end{cases} \tag{6.50}$$

Property 1): Since $\tilde{\psi}_{2l+1}(\tilde{z}(t)) = \sum_{k=0}^{l} \mathbf{U}_{kl}\ \phi_{2k+1}(\tilde{z}(t))$, we can write

$$E[\tilde{\psi}^*_{2m+1}(\tilde{z}(t))\tilde{\psi}_{2l+1}(\tilde{z}(t+\tau))] = \sum_{k=0}^{l} \mathbf{U}_{kl}\ E[\tilde{\psi}^*_{2m+1}(\tilde{z}(t))\phi_{2k+1}(\tilde{z}(t+\tau))]. \tag{6.51}$$

When $l < m$, we have $k < m$ for the summands on the right-hand side of (6.51). Since (6.51) is zero according to Lemma 1, $\tilde{\psi}_{2m+1}(\tilde{z}(t))$ is orthogonal to $\tilde{\psi}_{2l+1}(\tilde{z}(t+\tau))$ for any $l < m$. Interchanging t with $t+\tau$, we prove the orthogonality for $l > m$ as well. Thus, $\tilde{\psi}_{2m+1}(\tilde{z}(t))$ is orthogonal to $\tilde{\psi}_{2l+1}(\tilde{z}(t+\tau))$ for any $l \neq m$, $\forall\tau$.

Property 2): The auto-correlation function of $\tilde{\psi}_{2m+1}(\tilde{z}(t))$ is:

$$\begin{aligned} &E[\tilde{\psi}^*_{2m+1}(\tilde{z}(t))\tilde{\psi}_{2m+1}(\tilde{z}(t+\tau))] \\ &= \sum_{l=0}^{m} \mathbf{U}_{lm}\ E[\tilde{\psi}^*_{2m+1}(\tilde{z}(t))\phi_{2l+1}(\tilde{z}(t+\tau))]. \end{aligned} \tag{6.52}$$

Since $\tilde{\psi}_{2m+1}(\tilde{z}(t))$ is orthogonal to $\tilde{\phi}_{2l+1}(\tilde{z}(t+\tau))$ for $l < m$, (6.52) simplifies to

$$\mathbf{U}_{mm}\ E[\tilde{\psi}^*_{2m+1}(\tilde{z}(t))\phi_{2m+1}(\tilde{z}(t+\tau))]. \tag{6.53}$$

From (6.49), we learned that $\sum_{k=s}^{m}(-1)^{m-k}\binom{m}{k}\binom{k}{s} = 0$ if $m > s$; it equals to 1 if $m = s$. Replacing $l = m$ in (6.45)-(6.48), it is straightforward to show that only the $s = k = l = m$ term survives the double summation in (6.48) and hence $E[\tilde{\psi}^*_{2m+1}(\tilde{z}(t))\phi_{2m+1}(\tilde{z}(t+\tau))]$ simplifies to $\frac{(m+1)!}{\sqrt{m+1}}\ |c_{2\tilde{z}}(\tau)|^{2m}c_{2\tilde{z}}(\tau)$. Moreover, $U_{mm} = \sqrt{m+1}/(m+1)!$ from (6.50). Combining these results, (6.53) is simplified to

$$\begin{aligned} &\frac{\sqrt{m+1}}{(m+1)!}\frac{(m+1)!}{\sqrt{m+1}}|c_{2\tilde{z}}(\tau)|^{2m}c_{2\tilde{z}}(\tau) \\ &\qquad = |c_{2\tilde{z}}(\tau)|^{2m}c_{2\tilde{z}}(\tau). \end{aligned} \tag{6.54}$$

When $\tau = 0$, we obtain $E[|\tilde{\psi}_{2m+1}(\tilde{z}(t))|^2] = [c_{2\tilde{z}}(0)]^{2m+1} = 1$ for the unit variance $\tilde{z}(t)$. Therefore, the basis $\{\tilde{\psi}_{2m+1}(\tilde{z}(t))\}$ is orthonormal.

6.C Derivation of the Orthogonal Polynomial Predistorter

Here, we would like to show that the orthogonal polynomial predistorter given by:

$$\mathbf{z}^{(i+1)} = \mathbf{\Psi}(\mathbf{x})\widehat{\boldsymbol{\beta}}_{LS}, \qquad \widehat{\boldsymbol{\beta}}_{LS} = \mathbf{M}\widehat{\boldsymbol{\gamma}}_{LS}, \qquad \widehat{\boldsymbol{\gamma}}_{LS} = \mathbf{\Psi}(\mathbf{y})^{\dagger}\mathbf{z}^{(i)}, \tag{6.55}$$

is theoretically equivalent to the conventional polynomial predistorter:

$$\mathbf{z}^{(i+1)} = \mathbf{\Phi}(\mathbf{x})\widehat{\mathbf{b}}_{LS}, \qquad \widehat{\mathbf{b}}_{LS} = \mathbf{\Phi}(\mathbf{y}/G)^{\dagger}\mathbf{z}^{(i)}. \tag{6.56}$$

Note that we have omitted $^{(i)}$ in $\mathbf{x}^{(i)}$ and $\mathbf{y}^{(i)}$ for notational simplicity.

Let $\mathbf{D}(s)$ be a diagonal matrix whose diagonal elements are $[s, s^3, \ldots, s^{2K+1}]$. It follows easily that

$$\mathbf{\Phi}(\mathbf{x}) = \mathbf{\Phi}(\mathbf{x}/s)\ \mathbf{D}(s). \tag{6.57}$$

It also holds that

$$\mathbf{D}(\sigma s) = \mathbf{D}(\sigma)\ \mathbf{D}(s). \tag{6.58}$$

Writing the orthogonal polynomials (6.15) for $\mathbf{x}$ and $\mathbf{y}$ in vector format, we obtain

$$\mathbf{\Psi}(\mathbf{x}) = \mathbf{\Phi}(\mathbf{x})\mathbf{D}(1/\sigma_x)\mathbf{U}, \tag{6.59}$$

$$\mathbf{\Psi}(\mathbf{y}) = \mathbf{\Phi}(\mathbf{y})\mathbf{D}(1/\sigma_y)\mathbf{U}, \tag{6.60}$$

where $\mathbf{U}$ is defined in (6.50). For the $\mathbf{M}$ matrix given in (6.39), it can be shown that

$$\mathbf{UM} = \mathbf{D}(G\sigma_x/\sigma_y)\mathbf{U}. \tag{6.61}$$

Specifically, by substituting the $\mathbf{U}$ and $\mathbf{M}$ expressions, the (k,l)th element on either side of (6.61) can be shown to be

$$\frac{1}{\sqrt{l+1}}\frac{1}{k!}\binom{l+1}{k+1}\epsilon^{2k+1}(-1)^{l-k},$$

where $\epsilon = G\sigma_x/\sigma_y$. Similar tools as used in Appendix 6.A are helpful here.

From (6.61), we can write

$$\mathbf{M} = \mathbf{U}^{-1}\mathbf{D}(G\sigma_x/\sigma_y)\mathbf{U}. \tag{6.62}$$

Starting with (6.55) and using (6.56), (6.57), (6.58), (6.60), and (6.62), we prove that $\widehat{\boldsymbol{\beta}}_{LS} = \mathbf{U}^{-1}\mathbf{D}(\sigma_x)\widehat{\mathbf{b}}_{LS}$ as follows:

$$\begin{aligned}
\widehat{\boldsymbol{\beta}}_{LS} &= \mathbf{M}\boldsymbol{\Psi}(\mathbf{y})^{\dagger}\mathbf{z}^{(i)} \\
&= \mathbf{M}(\boldsymbol{\Phi}(\mathbf{y})\mathbf{D}(1/\sigma_y)\mathbf{U})^{\dagger}\mathbf{z}^{(i)} \\
&= \mathbf{M}(\boldsymbol{\Phi}(\mathbf{y}/G)\mathbf{D}(G/\sigma_y)\mathbf{U})^{\dagger}\mathbf{z}^{(i)} \\
&= \mathbf{M}\mathbf{U}^{-1}\mathbf{D}(\sigma_y/G)\boldsymbol{\Phi}(\mathbf{y}/G)^{\dagger}\mathbf{z}^{(i)} \\
&= \mathbf{U}^{-1}\mathbf{D}(G\sigma_x/\sigma_y)\mathbf{U}\mathbf{U}^{-1}\mathbf{D}(\sigma_y/G)\widehat{\mathbf{b}}_{LS} \\
&= \mathbf{U}^{-1}\mathbf{D}(\sigma_x)\widehat{\mathbf{b}}_{LS}.
\end{aligned} \tag{6.63}$$

Next, by invoking (6.55), (6.58), (6.59), and (6.63), we show that the orthogonal polynomial predistorter $\mathbf{z}^{(i+1)} = \boldsymbol{\Psi}(\mathbf{x})\widehat{\boldsymbol{\beta}}_{LS}$ is equivalent to the conventional polynomial predistorter $\mathbf{z}^{(i+1)} = \boldsymbol{\Phi}(\mathbf{x})\widehat{\mathbf{b}}_{LS}$:

$$\begin{aligned}
\mathbf{z}^{(i+1)} &= \boldsymbol{\Psi}(\mathbf{x})\widehat{\boldsymbol{\beta}}_{LS} \\
&= \boldsymbol{\Phi}(\mathbf{x})\mathbf{D}(1/\sigma_x)\mathbf{U}\widehat{\boldsymbol{\beta}}_{LS} \\
&= \boldsymbol{\Phi}(\mathbf{x})\mathbf{D}(1/\sigma_x)\mathbf{U}\mathbf{U}^{-1}\mathbf{D}(\sigma_x)\widehat{\mathbf{b}}_{LS} \\
&= \boldsymbol{\Phi}(\mathbf{x})\widehat{\mathbf{b}}_{LS}.
\end{aligned}$$

CHAPTER VII

STATISTICAL ANALYSIS OF A BANDPASS NONLINEARITY WITH NONSTATIONARY INPUT

Power amplifier is an important component of a communication system and is inherently nonlinear. When a non-constant envelope signal goes through a nonlinear power amplifier, spectral regrowth (broadening) appears at the output of the power amplifier. To satisfy regulatory requirements on out of band emissions, spectral regrowth must be contained. In this chapter, we first provide some general statistical analysis results for nonlinear systems with (non)stationary Gaussian input. We then derive a novel closed-form expression for the output power spectral density when the power amplifier is quasi-memoryless and cyclostationarity of the digitally modulated input is taken into account. We compare our results with the conventional analysis where stationary input is assumed. We emphasize the importance of paying attention to the cyclostationary nature of the input when excess bandwidth is present.

7.1 Introduction

Power amplifiers (PAs) are the major source of nonlinearity in communications systems. To achieve high efficiency from a given PA, the PA is sometimes driven into its nonlinear region. When a non-constant envelope signal goes through a nonlinear PA, spectral regrowth (broadening) appears at the PA output, which in turn causes adjacent channel interference (ACI). Due to stringent limits on the ACI imposed by regulatory bodies, PA nonlinearity must be contained.

It would be very helpful if we can predict spectral regrowth for a prescribed level of PA nonlinearity. Since more linear PAs are less efficient, practitioners may wish to use the PA in a configuration that allows for maximum PA efficiency while satisfying the spectral emission limits. Such an optimization strategy is feasible if we have tools for spectral analysis for

the nonlinear device.

A digital communication signal is cyclostationary (and thus nonstationary). The PA is nonlinear. Therefore, we are faced with a challenging task of carrying out spectral analysis on a nonlinear system with nonstationary input. Many authors have investigated spectral regrowth of the nonlinear PA [7, 44, 49, 113, 133], but to the best of our knowledge, there have been no published results that take into account the nonstationary or cyclostationary nature of the input. In the current literature, the input is assumed to be stationary. Indeed, nonlinear spectral analysis alone is already a challenging task. Interestingly, we will show in this chapter that under curtain conditions (no excess bandwidth), the stationary assumption is valid.

If the input is Gaussian and stationary, the PA output power spectral density (PSD) has been derived for a fifth-order polynomial nonlinear PA model in [113] and [44]. Using a moment-based approach in [49] and a cumulant-based approach in [133], spectral analysis has been extended to a polynomial model of any order. In addition, [133] provides spectral analysis results for a polynomial PA model with memory.

When the PA input is non-Gaussian, theoretical analysis becomes more complicated; however results are available in [129] for a 7th-order nonlinear PA with stationary (non-)Gaussian inputs. In [7], spectral analysis for a CDMA signal is presented assuming stationarity. In [8], uniform phase randomization is used to "stationarize" a cyclostationary input.

In this chapter, we first prove in Section 7.2 fundamental results on covariance analysis of a general nonlinearity with (non-)stationary Gaussian input. We then specialize to the polynomial model studied in [7,44,49,113,133], and present a closed-form expression for the output covariance function when the input is (non-)stationary Gaussian and the nonlinearity order is arbitrary. In Section 7.3, we examine the (cyclo)stationarity of digitally modulated signals, and present spectral analysis results that take into account (cyclo)stationarity of the input signal. We offer a comparison between the estimated PSD for the output of the PA, as well as analytical expressions for the PSD with and without the stationarity assumption. We show that when cyclostationarity of the input signal is taken into account, the PSD

predicted by our formula matches well the PSD calculated from the data. We summarize our findings in Section 7.4. The rather technical proofs of this chapter are deferred to the appendices.

7.2 Basic Results

The following baseband PA model has been frequently used in the literature [7, 8, 17, 18, 44, 49, 113, 129, 133]

$$y(t) = \sum_{k=0}^{K} a_{2k+1}|x(t)|^{2k}x(t), \tag{7.1}$$

where $x(t)$ is the baseband PA input signal, $y(t)$ is the baseband PA output signal, and $\{a_{2k+1}\}$ are complex-valued coefficients that can be extracted from standard characterizations such as amplitude-to-amplitude (AM/AM) and amplitude-to-phase (AM/PM) conversions of the PA. The highest nonlinearity order is $2K + 1$. The fact that only odd-order nonlinear terms appear in (7.1) is attributed to the bandpass nonlinear nature of the PA [17, 18].

We see from (7.1) that the PA complex gain is $G(x(t)) = y(t)/x(t) = \sum_{k=0}^{K} a_{2k+1}\,|x(t)|^{2k}$, which is a function of input amplitude $r = |x(t)|$ only. Writing the complex gain as $G(r) = A(r)\; e^{j\Phi(r)}$, we refer to $A(r)$ as the AM/AM conversion, and $\Phi(r)$ as the AM/PM conversion. A linear PA would have constant $A(r)$ and $\Phi(r)$ characteristics. If $A(r)$ is non-constant but $\Phi(r)$ is, the corresponding PA is called strictly memoryless. If both $A(r)$ and $\Phi(r)$ are non-constant, the resulting PA is called quasi-memoryless. Equation (7.1) can be used to describe both types of memoryless nonlinearity, and hence we do not distinguish the two in subsequent analysis.

Similar to [44, 49, 113, 133], we assume that the input $x(t)$ is complex Gaussian distributed, which is well-motivated for applications such as OFDM (orthogonal frequency division multiplexing). We would like to examine the auto-covariance function of $y(t)$ and then its PSD. The problem is non-trivial and requires a series of steps. We start with a general theorem on the correlation of (nonlinear) transformations of Gaussian random variables.

Theorem 2. *Let u and v be jointly complex Gaussian random variables and denote their cross-covariance by*

$$\sigma_{uv} = E[(u - E[u])(v - E[v])^*]. \tag{7.2}$$

The cross-correlation between two functions of u and v, $f(u)$ and $g(v)$, can be expressed as

$$E[f(u)g^*(v)] = \sum_{n=0}^{\infty} \sum_{k=0}^{n} \frac{1}{n!} \binom{n}{k} \sigma_{uv}^k (\sigma_{uv}^*)^{n-k} \\ E\left[\frac{\partial^n f(u)}{\partial u^k \partial (u^*)^{n-k}}\right] E\left[\frac{\partial^n g(v)}{\partial v^k \partial (v^*)^{n-k}}\right]^*, \tag{7.3}$$

where $\binom{n}{k} = \frac{n!}{k!(n-k)!}$.

Proof. See Appendix 7.A. □

We remark that in (7.3), $\frac{\partial f(u)}{\partial u}$ denotes the partial derivative of $f(u)$ w.r.t. u holding u^* constant, and $\frac{\partial f(u)}{\partial u^*}$ denotes the partial derivative of $f(u)$ w.r.t. u^* holding u constant [63, p. 518]. For example, we have $\frac{\partial u}{\partial u} = 1$, $\frac{\partial u}{\partial u^*} = 0$. The notations $\frac{\partial g(v)}{\partial v}$, $\frac{\partial g(v)}{\partial v^*}$ are similarly defined.

Theorem 2 is a very general result and does not require u and v to be zero-mean. Next, we verify (7.3) by way of some simple examples.

(i) If u and v are independent, we have $\sigma_{uv} = 0$. Therefore, only the $k = n = 0$ term contributes to the right-hand side (RHS) of (7.3), giving rise to

$$E[f(u)g^*(v)] = E[f(u)]E[g^*(v)].$$

(ii) If $f(u) = u$, $g(v) = v$, then we only need to consider the terms with $n = k = 0$ and $n = k = 1$ on the RHS of (7.3). This leads to $E[uv^*] = E[u]E[v^*] + \sigma_{uv}$, which is the same as (7.2).

(iii) Let us consider the case with $u = v$, $f(u) = g(u) = u^p$. Thus, $\sigma_{uv} = \sigma_u^2$, and the left-hand side (LHS) of (7.3) becomes $E[|u|^{2p}]$. Since $\frac{\partial^m u^p}{\partial (u^*)^m} = 0$, $\forall m \neq 0$, only the $n = k$ terms survive on the RHS of (7.3). Moreover, if u is zero-mean complex Gaussian distributed, $E[u^m] = 0$ for $m \neq 0$. Since

$$\frac{\partial^n (u^p)}{\partial u^n} = \frac{p!}{(p-n)!} u^{p-n}, \quad p \geq n, \tag{7.4}$$

we infer that $E\left[\frac{\partial^n(u^p)}{\partial u^n}\right] \neq 0$, only if $p = n$. Therefore, for u zero-mean complex Gaussian distributed, only the $n = k = p$ term survives on the RHS of (7.3), and we obtain

$$E[|u|^{2p}] = \frac{1}{p!}\sigma_u^{2p}(p!)(p!) = p!\sigma_u^{2p}. \tag{7.5}$$

Equation (7.5) is a well-known result obtained by Reed in [101].

(iv) If $f(u) = u$, only the $n = k = 0$ and $n = k = 1$ terms contribute to the RHS of (7.3). Hence,

$$E[ug^*(v)] = E[u]E[g^*(v)] + \sigma_{uv}E[g'(v)]^*,$$

where $g'(v) = \frac{\partial g(v)}{\partial v}$. Equivalently,

$$\text{cov}\{u, g(v)\} = \sigma_{uv}E[g'(v)]^* \propto \sigma_{uv}. \tag{7.6}$$

If a (nonlinear) system has $x(t)$ as input, $y(t) = g(x(t))$ as output, and $x(t)$ is stationary, then by replacing $u = x(t+\tau)$ and $v = x(t)$ in (7.6), and recognizing that $\sigma_{uv} = c_{2x}(\tau)$, $\text{cov}(u, g^*(v)) = c_{xy}(\tau)$, we obtain

$$c_{xy}(\tau) = c_{2x}(\tau)\ E[g'(x(t))]^* \propto c_{2x}(\tau), \tag{7.7}$$

which is the celebrated Bussgang Theorem [86, p. 307].

The above examples illustrate the generality of our Theorem 2. Equation (7.3) says that we can turn the joint expectation $E[f(u)g^*(v)]$ into a linear combination of products of individual expectations. Next, we specialize to the case where the gain function, $f(u)/u$, depends on the input power $|u|^2$ only; i.e., if

$$f(u) = u\xi(uu^*). \tag{7.8}$$

The following lemma will assist in simplifying (7.3) for this special case.

Lemma 2. *Suppose that $f(u) = u\xi(uu^*)$ and up to $m + l$ derivatives of $\xi(\cdot)$ exist. If u is zero-mean complex Gaussian distributed then*

$$E\left[\frac{\partial^{m+l}f(u)}{\partial u^m \partial (u^*)^l}\right] = 0, \quad \text{if } m \neq l + 1. \tag{7.9}$$

Proof. See Appendix 7.B. □

Recall that if the device is quasi-memoryless, it is characterized by the AM-AM, AM-PM conversions, which can be represented by (7.8).

Using Lemma 2, we can reduce the double summation on the RHS of (7.3) to a single summation. We present the following corollary for nonlinear transformations of the type in (7.8).

Corollary 2. *Let u and v be zero-mean complex Gaussian distributed. Assume that $f(u)/u$ depends only on $|u|^2$, $g(v)/v$ depends only on $|v|^2$, and all derivatives of $f(u)$ and $g(v)$ exist. Then*

$$E[f(u)g^*(v)] = \sum_{s=0}^{\infty} \frac{|\sigma_{uv}|^{2s}\sigma_{uv}}{s!(s+1)!} E\left[\frac{\partial^{2s+1} f(u)}{\partial u^{s+1}\partial (u^*)^s}\right] \times E\left[\frac{\partial^{2s+1} g(v)}{\partial v^{s+1}\partial (v^*)^s}\right]^*. \tag{7.10}$$

Proof. According to Lemma 2, we only need to consider the terms on the RHS of (7.3) that satisfy $k = n - k + 1$; i.e., $n = 2k - 1$. Replacing $n = 2s + 1$, $k = s + 1$, $n - k = s$ in (7.3) and realizing that $\frac{1}{(2s+1)!}\binom{2s+1}{s+1} = \frac{1}{s!(s+1)!}$, we obtain (7.10). □

Now, let us consider the polynomial model

$$\begin{aligned} f(u) &= \sum_{k=0}^{K} a_{2k+1} u^{k+1} (u^*)^k \\ &= u \sum_{k=0}^{K} a_{2k+1} |u|^{2k}. \end{aligned} \tag{7.11}$$

If we let $u = x(t + \tau)$, $v = x(t)$, $f(u) = y(t + \tau)$, and $g(v) = f(v) = y(t)$, then we have $\sigma_{uv} = c_{2x}(t; \tau)$. We can apply Theorem 2 to relate the auto-covariance function of $y(t)$, $c_{2y}(t; \tau) = E[f(u)g^*(v)]$, to that of the input, $c_{2x}(t; \tau) = E[uv^*]$. Note that for u complex Gaussian distributed, $E[f(u)] = E[g(v)] = 0$.

Theorem 3. *Suppose that $x(t)$ is a zero-mean, complex Gaussian random process and*

$$\begin{aligned} y(t) &= \sum_{k=0}^{K} a_{2k+1}[x(t)]^{k+1}[x^*(t)]^k \\ &= x(t)\sum_{k=0}^{K} a_{2k+1}|x(t)|^{2k}. \end{aligned} \tag{7.12}$$

Then the auto-covariance of $y(t)$, $c_{2y}(t;\tau)$, is related to that of $x(t)$, $c_{2x}(t;\tau)$, via

$$\begin{aligned} c_{2y}(t;\tau) = &\sum_{s=0}^{K} \frac{|c_{2x}(t;\tau)|^{2s} c_{2x}(t;\tau)}{(s+1)} \\ &\left(\sum_{l=s}^{K} a_{2l+1}\binom{l}{s}(l+1)!(c_{2x}(t;0))^{l-s}\right) \\ &\left(\sum_{k=s}^{K} a_{2k+1}\binom{k}{s}(k+1)!(c_{2x}(t+\tau;0))^{k-s}\right)^*. \end{aligned} \tag{7.13}$$

Proof. From (7.11) and (7.4), we infer that

$$\frac{\partial^{2s+1} f(u)}{\partial u^{s+1}\partial(u^*)^s} = \sum_{k=s}^{K} a_{2k+1}\frac{(k+1)!}{(k-s)!}\frac{k!}{(k-s)!}|u|^{2(k-s)}. \tag{7.14}$$

Next, we utilize (7.5) to obtain

$$E\left[\frac{\partial^{2s+1} f(u)}{\partial u^{s+1}\partial(u^*)^s}\right] = \sum_{k=s}^{K} a_{2k+1}\frac{(k+1)!}{(k-s)!}k!(\sigma_u^2)^{k-s}. \tag{7.15}$$

Similarly,

$$E\left[\frac{\partial^{2s+1} g(v)}{\partial v^{s+1}\partial(v^*)^s}\right] = \sum_{l=s}^{K} a_{2l+1}\frac{(l+1)!}{(l-s)!}l!(\sigma_v^2)^{l-s}. \tag{7.16}$$

Substituting (7.15) and (7.16) into (7.10) and replacing with $\sigma_u^2 = c_{2x}(t+\tau;0)$, $\sigma_v^2 = c_{2x}(t;0)$, $\sigma_{uv} = c_{2x}(t;\tau)$, and $E[f(u)g^*(v)] = c_{2y}(t;\tau)$, we obtain (7.13). □

Theorem 3 describes the auto-covariance of the output of a baseband polynomial nonlinearity when the input is (non-)stationary Gaussian. It is a general result. The next corollary examines the case when $x(t)$ is stationary, which makes $c_{2x}(t;\tau) = c_{2x}(\tau)$, $\forall t$.

Corollary 3. *Suppose that $x(t)$ is a zero-mean, stationary complex Gaussian random process and $y(t)$ is related to $x(t)$ through (7.12). Then the auto-covariance function of $y(t)$ is*

given by

$$c_{2y}(\tau) = \sum_{s=0}^{K} \frac{|c_{2x}(\tau)|^{2s} c_{2x}(\tau)}{(s+1)} \left| \sum_{l=s}^{K} a_{2l+1} \binom{l}{s} (l+1)! (c_{2x}(0))^{l-s} \right|^2 . \tag{7.17}$$

Proof. For $x(t)$ stationary, replace $c_{2x}(t;0) = c_{2x}(t+\tau;0) = c_{2x}(0)$, $c_{2x}(t;\tau) = c_{2x}(\tau)$ in (7.13). □

Corollary 3 can be shown to agree with results in [21, 49, 133] for stationary Gaussian $x(t)$.

7.3 Digitally-Modulated Signals and Spectral Analysis for Polynomial Nonlinearity

In this section, we would like to specialize to digitally modulated $x(t)$ which is cyclostationary in general. We first examine the (time-varying) covariance function of $x(t)$, and then that of $y(t)$, and then the PSD of $y(t)$.

7.3.1 Digitally-Modulated Signals

Consider the following baseband representation of a digitally-modulated signal:

$$x(t) = \sum_{k=-\infty}^{\infty} s_k h(t - kT), \tag{7.18}$$

where s_k is the kth symbol, $h(t)$ is impulse response of the pulse shaping filter, and T is the symbol period. Applying the continuous-time Fourier transform (CTFT) to both sides of (7.18), we obtain

$$X(f) = H(f) S(e^{j2\pi T f}), \tag{7.19}$$

where the CTFT of $x(t)$ is defined as

$$X(f) = \mathcal{F}\{x(t)\} = \int x(t) e^{-j2\pi t f} dt, \tag{7.20}$$

the CTFT of $h(t)$ is

$$H(f) = \mathcal{F}\{h(t)\} = \int h(t) e^{-j2\pi t f} dt, \tag{7.21}$$

and the discrete-time Fourier transform (DTFT) of s_k is defined as

$$S(e^{j2\pi Tf}) = \sum_{k=-\infty}^{\infty} s_k e^{-j2\pi kTf}. \tag{7.22}$$

Note that $S(e^{j2\pi Tf})$ is periodic in f with period $\frac{1}{T}$, i.e., information contained in $f \in [-\frac{1}{2T}, \frac{1}{2T}]$ is repeated every $\frac{1}{T}$. To preserve the information in $S(e^{j2\pi Tf})$, the pulse shaping filter, $H(f)$, must have a bandwidth greater than or equal to $\frac{1}{T}$.

Assume that $\{s_k\}$ is zero-mean, i.i.d with variance $\gamma_{2s} = E[|s_k|^2]$. The mean and covariance function of $x(t)$ are respectively,

$$E[x(t)] = 0, \tag{7.23}$$

$$\begin{aligned} c_{2x}(t;\tau) &= \text{cum}\,\{x^*(t), x(t+\tau)\} \\ &= \gamma_{2s} \sum_{k=-\infty}^{\infty} h^*(t-kT)h(t+\tau-kT). \end{aligned} \tag{7.24}$$

Note that $x(t)$ is not wide-sense stationary (WSS) in general since (7.24) may depend on t.

In Appendix 7.C, we show that $c_{2x}(t;\tau)$ can be separated into t-dependent terms and τ-dependent terms as follows:

$$c_{2x}(t;\tau) = \frac{\gamma_{2s}}{T} \sum_{m=-\infty}^{\infty} \rho_{\frac{m}{T}}(\tau) e^{-j\frac{2\pi}{T}mt}, \tag{7.25}$$

where

$$\rho_u(\tau) = \int H^*(f+u)H(f)e^{j2\pi f\tau}df. \tag{7.26}$$

Inverse CTFT is defined as

$$x(t) = \mathcal{F}^{-1}\{X(f)\} = \int X(f)e^{j2\pi ft}df.$$

From (7.26), we see that $\rho_u(\tau)$ and $H^*(f+u)H(f)$ form a CTFT pair. The time average of (7.25) is

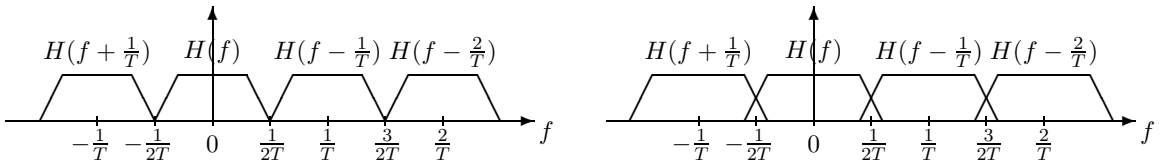
$$c_{2x}(\tau) \triangleq \overline{c_{2x}(t;\tau)} = \frac{\gamma_{2s}}{T}\rho_0(\tau) = \frac{\gamma_{2s}}{T}\int |H(f)|^2 e^{j2\pi f\tau}df, \tag{7.27}$$

where

$$\overline{f(t)} \triangleq \lim_{\Delta\to\infty} \frac{1}{2\Delta}\int_{-\Delta}^{\Delta} f(t)dt$$

represents the time averaging operation and $\rho_0(\tau) = \mathcal{F}^{-1}\left\{|H(f)|^2\right\}$.

With respect to $h(t)$, we consider the following two cases:



(a) $H(f) = 0$ for $|f| > \frac{1}{2T}$.

(b) $H(f) \neq 0$ for some $|f| > \frac{1}{2T}$.

Figure 7.1: Relationship among $H(f + \frac{m}{T})$ for different m's.

(i) **No excess bandwidth case.** When $H(f)$ is band-limited to bandwidth $\frac{1}{T}$, i.e., $H(f) = 0$ for $|f| > \frac{1}{2T}$, the only m for which $H(f + \frac{m}{T})$ overlaps with $H(f)$, is $m = 0$ (see Fig. 7.1(a)). In this case, $\rho_{\frac{m}{T}}(\tau) = 0$ except for $m = 0$. Therefore, the time-dependent term $e^{-j\frac{2\pi}{T}mt}$ in (7.25) is immaterial and

$$c_{2x}(t;\tau) = \frac{\gamma_{2s}}{T}\rho_0(\tau). \tag{7.28}$$

Since $c_{2x}(t;\tau)$ does not depend on t in (7.28), $c_{2x}(t;\tau) = c_{2x}(\tau)$ and $x(t)$ is WSS. If $x(t)$ is Gaussian (either real or complex) then it is also strict sense stationary (SSS). In this case, $c_{2y}(\tau)$ is given by (7.17) and the PSD of $y(t)$ has been investigated in [133] and [49] for an arbitrary nonlinear order.

(ii) **Excess bandwidth case.** When the bandwidth of $H(f)$ exceeds $\frac{1}{T}$ but does not exceed $\frac{2}{T}$, i.e., $H(f) \neq 0$ for some $|f| > \frac{1}{2T}$, but $H(f) = 0$ for $|f| > \frac{1}{T}$, the only m values for which $H(f + \frac{m}{T})$ overlaps with $H(f)$ are $m = 0$, $m = -1$, and $m = 1$ (see Fig. 7.1(b)). In this case, only $\rho_0(\tau)$, $\rho_{\frac{1}{T}}(\tau)$, and $\rho_{-\frac{1}{T}}(\tau)$ are non-zero, and hence from (7.25),

$$c_{2x}(t;\tau) = \frac{\gamma_{2s}}{T}\left(\rho_0(\tau) + \rho_{-\frac{1}{T}}(\tau)e^{j\frac{2\pi}{T}t} + \rho_{\frac{1}{T}}(\tau)e^{-j\frac{2\pi}{T}t}\right). \tag{7.29}$$

In this case $c_{2x}(t;\tau)$ is a function of both t and τ, meaning that $x(t)$ is not WSS. Since $c_{2x}(t;\tau)$ is a periodic function of t, $c_{2y}(t;\tau)$ is also a periodic function of t by virtue of (7.13). Spectral analysis of $y(t)$ is still feasible, as we will see in the next section.

7.3.2 Cyclostationary Spectral Analysis

In the nonstationary (excess bandwidth) case, the spectrum of the PA output is given by:

$$S_{2y}(f) = \mathcal{F}\left\{c_{2y}(\tau)\right\}, \tag{7.30}$$

where $c_{2y}(\tau)$ is the time-averaged version of $c_{2y}(t;\tau) = \operatorname{cum}\{y^*(t), y(t+\tau)\}$. The time average of (7.13) is

$$c_{2y}(\tau) = \overline{c_{2y}(t;\tau)} = \tag{7.31}$$
$$\sum_{s=0}^{K}\sum_{l=s}^{K}\sum_{k=s}^{K}\frac{1}{(s+1)}\binom{l}{s}(l+1)!\binom{k}{s}(k+1)!a_{2l+1}a_{2k+1}^*$$
$$\overline{|c_{2x}(t;\tau)|^{2s}c_{2x}(t;\tau)(c_{2x}(t;0))^{l-s}(c_{2x}^*(t+\tau;0))^{k-s}}.$$

Unfortunately, time-average of a product is not the same as the product of individual time-averages (e.g., $\overline{c_{2x}(t;\tau)c_{2x}(t,0)} \neq c_{2x}(\tau)c_{2x}(0)$) so (7.31) is not easily simplified.

For the digitally-modulated $x(t)$ of (7.18), we substitute (7.26) and (7.29) into (7.31) to obtain a closed-form expression for $c_{2y}(\tau)$ in terms of $H(f)$, γ_{2s}, and T. For simplicity, we describe the result for a PA given by (7.1) with $K = 1$, i.e., including only the linear and cubic nonlinear terms. In this case, (7.31) becomes

$$c_{2y}(\tau) = |a_1|^2\underbrace{\overline{c_{2x}(t;\tau)}}_{①} + 2a_1a_3^*\underbrace{\overline{c_{2x}(t;\tau)c_{2x}^*(t+\tau;0)}}_{②}$$
$$+2a_1^*a_3\underbrace{\overline{c_{2x}(t;\tau)c_{2x}(t;0)}}_{③} + 4|a_3|^2\underbrace{\overline{c_{2x}(t;\tau)c_{2x}(t;0)c_{2x}^*(t+\tau;0)}}_{④}$$
$$+2|a_3|^2\underbrace{\overline{|c_{2x}(t;\tau)|^2c_{2x}(t;\tau)}}_{⑤}. \tag{7.32}$$

Substituting (7.29) into (7.32) and taking the CTFT on both sides of (7.32), we show in Appendix 7.D that the PA output PSD is

$$S_{2y}(f) = \frac{\gamma_{2s}}{T}\Big|a_1H(f) + a_3\frac{\gamma_{2s}}{T}\Big(\rho_0(0)H(f) \tag{7.33}$$
$$+\rho_{\frac{1}{T}}(0)H(f+\tfrac{1}{T}) + \rho_{-\frac{1}{T}}(0)H(f-\tfrac{1}{T})\Big)\Big|^2$$
$$+2|a_3|^2(\frac{\gamma_{2s}}{T})^3\Big(|H(f)|^2 \star |H(f)|^2 \star |H(-f)|^2$$
$$+2[H^*(f-\tfrac{1}{T})H(f)] \star [H(-f-\tfrac{1}{T})H^*(-f)] \star |H(f)|^2$$
$$+2[H^*(f+\tfrac{1}{T})H(f)] \star [H(-f+\tfrac{1}{T})H^*(-f)] \star |H(f)|^2$$
$$+2[H^*(f+\tfrac{1}{T})H(f)] \star [H^*(f-\tfrac{1}{T})H(f)] \star |H(-f)|^2\Big),$$

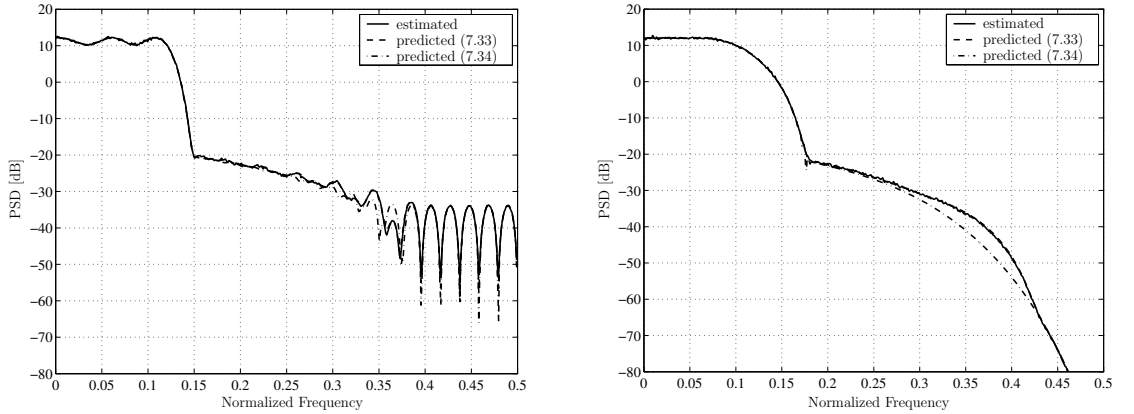
where $\star$ denotes convolution.

In the zero excess bandwidth (WSS) case, $H(f)$ and $H(f\pm\frac{1}{T})$ do not overlap, $H(f)H(f\pm\frac{1}{T})=0$, $\rho_{\frac{1}{T}}(0)=\rho_{-\frac{1}{T}}(0)=0$, and thus (7.33) simplifies to

$$S_{2y}(f)=\frac{\gamma_{2s}}{T}\left|a_1+a_3\frac{\gamma_{2s}}{T}\rho_0(0)\right|^2|H(f)|^2 \\ +2\left(\frac{\gamma_{2s}}{T}\right)^3|a_3|^2|H(f)|^2\star|H(f)|^2\star|H(-f)|^2. \tag{7.34}$$

Next, we verify (7.33) and compare it with (7.34) using computer simulations.

7.4 Numerical Examples



(a) with IS-95 pulse shaping filter.

(b) with root raised cosine filter (roll-off factor is 0.5).

Figure 7.2: PA output PSD for a third-order nonlinear PA. The solid line is the estimated PSD based on output samples; the dashed line corresponds to (7.33), and the dash-dotted line is generated using equation (7.34).

Consider the PA given in (7.1) with $K=1$ and $a_1=15.0008+0.0908j$ and $a_3=-23.0826+3.3133j$. Here, we explore the PA output PSD $S_{2y}(f)$ when the PA input $x(t)$ is given by (7.18) with the following pulse shaping filter (i) IS-95 pulse shaping filter [1], or (ii) root raised cosine filter given by:

$$H(f)=\begin{cases}1 & |f|<(1-\beta)f_c \\ \frac{1}{2}+\frac{1}{2}\cos\left(\pi\frac{f-(1-\beta)f_c}{2\beta f_c}\right) & (1-\beta)f_c\le|f|\le(1+\beta)f_c \\ 0 & |f|>(1+\beta)f_c\end{cases} \tag{7.35}$$

with cut-off frequency $f_c=\frac{1}{2T}$, and roll-off factor $\beta=0.5$ (50% excess bandwidth). For both (i) and (ii), sampling rate is 4 samples per symbol. γ_{2s} is selected such that the variance of $x(t)$ is 0.017, and $|x(t)|$ enters into the compression region of the PA.

Figures 7.2(a) and 7.2(b) show the PA output PSD corresponding to filters (i) and (ii), respectively. The solid line is the PA output PSD obtained from 2^{17} samples of $y(t)$. The dashed line is the PA output PSD calculated based on (7.33). The dash-dotted line is the PA output PSD calculated based on (7.34) (i.e., assuming a stationary input data model). From both figures, we observe that the dashed line and the solid line coincide, thus verifying the theoretical expression in (7.33). The small gap (in the adjacent channel) between the solid line and the dashed-dotted line in Fig. 7.2(b) indicates that (7.34) cannot be used to accurately predict the PA output PSD when the input has excess bandwidth. Therefore, treating digitally-modulated signals with excess bandwidth as stationary underestimates out-of-band emission by as much as 6 dB for the example shown. For a different PA or a different input drive level, the discrepancy between stationary and nonstationary spectral analysis can be more or less than what we see here. The discrepancy is negligible in Fig. 7.2(a) because the filter has basically no excess bandwidth, except that small ripples are present outside the passband $[-\frac{1}{2T}, \frac{1}{2T}]$.

7.5 Conclusions

Power amplifiers are used in most communication systems and are inherently nonlinear. Spectral analysis can help to evaluate the suitability of a given PA for amplifying certain signals or to assist in predistortion linearization algorithm design. In this chapter, we investigated bandpass nonlinearities with Gaussian inputs. We first presented general results on covariance analysis of (nonlinearly) transformed Gaussian random variables. We then specialized to the case of digitally modulated signals. We showed that when the pulse shaping filter has no excess bandwidth, the input signal is wide sense stationary and previous nonlinear spectral analysis results apply. When the pulse shaping filter has excess bandwidth, the input is cyclostationary. We then derived a closed-form expression for the PSD at the output of the PA. We showed that by taking into account cyclostationary nature of the processes, more accurate spectral analysis results can be obtained.

7.A Proof of Theorem 1

We follow the approach of [63, Appendix 15B], denote the Fourier transform (FT) of $f(u)$ by $F(\omega_u)$, and express $f(u)$ in terms of $F(\omega_u)$ as

$$f(u) = \int F(\omega_u) e^{j\mathrm{Re}\{\omega_u^* u\}} d\omega_u. \tag{7.36}$$

Similarly we express $g(v)$ in terms of its FT $G(\omega_v)$ as

$$g(v) = \int G(\omega_v) e^{j\mathrm{Re}\{\omega_v^* v\}} d\omega_v. \tag{7.37}$$

Using (7.36) and (7.37), the correlation between $f(u)$ and $g(v)$ is

$$E[f(u)g^*(v)] = \iint F(\omega_u) G^*(\omega_v) E\left[e^{j\mathrm{Re}\{\omega_u^* u - \omega_v^* v\}}\right] d\omega_u d\omega_v. \tag{7.38}$$

Recall that $\Phi_{uv}(\omega_u, \omega_v) = E\left[e^{j\mathrm{Re}\{\omega_u^* u + \omega_v^* v\}}\right]$ is the joint characteristic function of u, v. For u, v jointly complex Gaussian distributed

$$\begin{aligned}\Phi_{uv}(\omega_u, \omega_v) = e^{-\frac{1}{4}(|\omega_u|^2\sigma_u^2 + |\omega_v|^2\sigma_v^2 + 2\mathrm{Re}\{w_u^*\sigma_{uv}w_v\})} \\ \times e^{j\mathrm{Re}\{\omega_u^*\mu_u + \omega_v^*\mu_v\}},\end{aligned} \tag{7.39}$$

where $\mu_u = E[u]$, $\mu_v = E[v]$, σ_u^2 is the variance of u, and σ_v^2 is the variance of v [63, p. 559]. Since

$$\begin{aligned}\Phi_u(\omega_u) &= E\left[e^{j\mathrm{Re}\{\omega_u^* u\}}\right] \\ &= \Phi_{uv}(\omega_u, 0) = e^{-\frac{1}{4}|\omega_u|^2\sigma_u^2 + j\mathrm{Re}\{\omega_u^*\mu_u\}}\end{aligned} \tag{7.40}$$

and

$$\begin{aligned}\Phi_v(\omega_v) &= E\left[e^{j\mathrm{Re}\{\omega_v^* v\}}\right] \\ &= \Phi_{uv}(0, \omega_v) = e^{-\frac{1}{4}|\omega_v|^2\sigma_v^2 + j\mathrm{Re}\{\omega_v^*\mu_v\}},\end{aligned} \tag{7.41}$$

we can write

$$\Phi_{uv}(\omega_u, \omega_v) = \Phi_u(\omega_u)\Phi_v(\omega_v) e^{-\frac{1}{2}\mathrm{Re}\{w_u^*\sigma_{uv}w_v\}}. \tag{7.42}$$

Substituting (7.42) into (7.38), we obtain

$$\begin{aligned}E[f(u)g^*(v)] = \iint F(\omega_u) G^*(\omega_v) \Phi_u(\omega_u)\Phi_v(-\omega_v) \\ e^{\frac{1}{2}\mathrm{Re}\{w_u^*\sigma_{uv}w_v\}} d\omega_u d\omega_v.\end{aligned} \tag{7.43}$$

Using the Taylor series expansion of $e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$ and the fact that $\Phi_v(-\omega_v) = \Phi_v^*(\omega_v)$, we obtain

$$\begin{aligned} E[f(u)g^*(v)] &= \iint F(\omega_u)G^*(\omega_v)\Phi_u(\omega_u)\Phi_v^*(\omega_v) \\ &\qquad \sum_{n=0}^{\infty} \frac{(\frac{1}{2}\mathrm{Re}\left\{w_u^*\sigma_{uv}w_v\right\})^n}{n!} d\omega_u d\omega_v \\ &= \sum_{n=0}^{\infty} \frac{1}{4^n n!} \iint F(\omega_u)G^*(\omega_v)\Phi_u(\omega_u)\Phi_v^*(\omega_v) \\ &\qquad (w_u^*\sigma_{uv}w_v + w_v^*\sigma_{uv}^*w_u)^n d\omega_u d\omega_v. \end{aligned} \tag{7.44}$$

Recall the binomial expression $(w_u^*\sigma_{uv}w_v + w_v^*\sigma_{uv}^*w_u)^n = \sum_{k=0}^{n} \binom{n}{k} (w_u^*\sigma_{uv}w_v)^k (w_v^*\sigma_{uv}^*w_u)^{n-k}$. We then rewrite (7.44) as

$$\begin{aligned} E[f(u)g^*(v)] &= \sum_{n=0}^{\infty}\sum_{k=0}^{n} \frac{1}{n!}\binom{n}{k}\sigma_{uv}^k(\sigma_{uv}^*)^{n-k} \\ &\qquad \int \left(\frac{w_u^*}{2}\right)^k \left(\frac{w_u}{2}\right)^{n-k} F(\omega_u)\Phi_u(\omega_u)d\omega_u \\ &\qquad \int \left(\frac{w_v}{2}\right)^k \left(\frac{w_v^*}{2}\right)^{n-k} G^*(\omega_v)\Phi_v^*(\omega_v)d\omega_v. \end{aligned} \tag{7.45}$$

Based on (7.36) and

$$\begin{aligned} \frac{\partial^k \left[e^{j\mathrm{Re}\{\omega_u^* u\}}\right]}{\partial u^k} &= \left(\frac{j\omega_u^*}{2}\right)^k e^{j\mathrm{Re}\{\omega_u^* u\}}, \\ \frac{\partial^{n-k} \left[e^{j\mathrm{Re}\{\omega_u^* u\}}\right]}{\partial (u^*)^{n-k}} &= \left(\frac{j\omega_u}{2}\right)^{n-k} e^{j\mathrm{Re}\{\omega_u u\}}, \end{aligned}$$

we infer that

$$\begin{aligned} &\int \left(\frac{w_u^*}{2}\right)^k \left(\frac{w_u}{2}\right)^{n-k} F(\omega_u)\Phi_u(\omega_u)d\omega_u \\ &= E\left[\int \left(\frac{w_u^*}{2}\right)^k \left(\frac{w_u}{2}\right)^{n-k} F(\omega_u) e^{j\mathrm{Re}\{\omega_u^* u\}} d\omega_u\right] \\ &= \frac{1}{j^n} E\left[\frac{\partial^k}{\partial u^k}\frac{\partial^{n-k}}{\partial (u^*)^{n-k}} \int F(\omega_u) e^{j\mathrm{Re}\{\omega_u^* u\}} d\omega_u\right] \\ &= \frac{1}{j^n} E\left[\frac{\partial^n f(u)}{\partial u^k \partial (u^*)^{n-k}}\right]. \end{aligned} \tag{7.46}$$

Similarly,

$$\begin{aligned} &\int \left(\frac{w_v}{2}\right)^k \left(\frac{w_v^*}{2}\right)^{n-k} G^*(\omega_v)\Phi_v^*(\omega_v)d\omega_v \\ &= \left(\frac{1}{-j}\right)^n E\left[\frac{\partial^n g(v)}{\partial v^k \partial (v^*)^{n-k}}\right]^*. \end{aligned} \tag{7.47}$$

Putting (7.45), (7.46), and (7.47) together, we thus prove (7.3).

7.B Proof of Lemma 1

We first find

$$\frac{\partial^l f(u)}{\partial (u^*)^l} = \frac{\partial^l [u\xi(uu^*)]}{\partial (u^*)^l} = u^{l+1}\xi^{(l)}(uu^*), \tag{7.48}$$

where $\xi^{(l)}(p) = \frac{\partial^l \xi(p)}{\partial p^l}$. Next, we take the mth derivative of (7.48) w.r.t. u, apply the Leibnitz's Theorem for the derivative of a product [5, p. 12], and obtain

$$\frac{\partial^{m+l} f(u)}{\partial u^m \partial (u^*)^l} = \sum_{k=0}^{m} \binom{m}{k} \frac{\partial^k \left(u^{l+1}\right)}{\partial u^k} \frac{\partial^{m-k}[\xi^{(l)}(uu^*)]}{\partial u^{m-k}} \tag{7.49}$$

$$= \sum_{k=0}^{m} \binom{m}{k} \frac{(l+1)!\, u^{l+1-k}}{(l+1-k)!} (u^*)^{m-k} \xi^{(l+m-k)}(uu^*). \tag{7.50}$$

We realize that

$$u^{l+1-k}(u^*)^{m-k} = |u|^{l+1+m-2k} e^{j(l+1-m)\angle u}.$$

Therefore, (7.50) can be written as

$$\frac{\partial^{m+l} f(u)}{\partial u^m \partial (u^*)^l} = e^{j(l+1-m)\angle u} g(|u|). \tag{7.51}$$

Recall that for u zero-mean complex Gaussian distributed, the phase $\angle u$ and the magnitude $|u|$ are mutually independent and $\angle u$ is uniformly distributed in $[0, 2\pi)$. Hence $E[e^{j(l+1-m)\angle u}] = 0$ if $m \neq l+1$. As a result, we obtain (7.9).

7.C Derivation of (7.25)

The inverse CTFT of $H(f)$ is

$$h(t) = \int H(f) e^{j2\pi f t} df. \tag{7.52}$$

Substituting (7.52) into (7.24), we obtain

$$\begin{aligned} c_{2x}(t;\tau) &= \gamma_{2s} \sum_{k=-\infty}^{\infty} \int H^*(f_1) e^{-j2\pi f_1 (t-kT)} df_1 \\ &\qquad \int H(f_2) e^{j2\pi f_2 (t+\tau-kT)} df_2 \\ &= \gamma_{2s} \iint H^*(f_1) e^{-j2\pi f_1 t} H(f_2) e^{j2\pi f_2 (t+\tau)} \\ &\qquad \sum_{k=-\infty}^{\infty} e^{j2\pi (f_1 - f_2) kT} df_1 df_2. \end{aligned} \tag{7.53}$$

Using the fact that $\sum_{k=-\infty}^{\infty} e^{j2\pi fkT}$ is the Fourier series expansion of $\sum_{m=-\infty}^{\infty} \frac{1}{T}\delta(f-\frac{m}{T})$, we rewrite (7.54) as

$$\begin{aligned} c_{2x}(t;\tau) &= \gamma_{2s} \iint H^*(f_1)e^{-j2\pi f_1 t}H(f_2)e^{j2\pi f_2(t+\tau)} \\ &\qquad \sum_{m=-\infty}^{\infty} \frac{1}{T}\delta(f_1-f_2-\frac{m}{T})df_1 df_2 \\ &= \frac{\gamma_{2s}}{T}\sum_{m=-\infty}^{\infty} \int H^*(f_2+\frac{m}{T})e^{-j2\pi f_2 t}e^{-j\frac{2\pi}{T}mt} \\ &\qquad H(f_2)e^{j2\pi f_2(t+\tau)}df_2 \\ &= \frac{\gamma_{2s}}{T}\sum_{m=-\infty}^{\infty} e^{-j\frac{2\pi}{T}mt}\int H^*(f+\frac{m}{T})H(f)e^{j2\pi f\tau}df, \end{aligned} \tag{7.54}$$

which yields (7.25) and (7.26).

7.D Derivation of (7.33)

We substitute (7.29) into (7.32) and write out each of the time-averages. Recall that the time average of $e^{j\alpha t}$ is zero except when α is 0 modulo 2π.

$$①= \frac{\gamma_{2s}}{T}\rho_0(\tau) \tag{7.55}$$

$$②= \left(\frac{\gamma_{2s}}{T}\right)^2 \Big(\rho_0(\tau)\rho_0^*(0)+\rho_{\frac{1}{T}}(\tau)\rho_{\frac{1}{T}}^*(0)e^{j\frac{2\pi}{T}\tau} +\rho_{-\frac{1}{T}}(\tau)\rho_{-\frac{1}{T}}^*(0)e^{-j\frac{2\pi}{T}\tau}\Big) \tag{7.56}$$

$$③= (\frac{\gamma_{2s}}{T})^2\left(\rho_0(\tau)\rho_0(0)+\rho_{\frac{1}{T}}(\tau)\rho_{-\frac{1}{T}}(0)+\rho_{-\frac{1}{T}}(\tau)\rho_{\frac{1}{T}}(0)\right)$$

$$\begin{aligned} ④= (\frac{\gamma_{2s}}{T})^3\Big(&\rho_0(\tau)|\rho_0(0)|^2+\rho_{\frac{1}{T}}(\tau)\rho_0(0)\rho_{\frac{1}{T}}^*(0)e^{j\frac{2\pi}{T}\tau} \\ &+\rho_{\frac{1}{T}}(\tau)\rho_{-\frac{1}{T}}(0)\rho_0^*(0)+\rho_0(\tau)|\rho_{\frac{1}{T}}(0)|^2e^{j\frac{2\pi}{T}\tau} \\ &+\rho_0(\tau)|\rho_{-\frac{1}{T}}(0)|^2e^{-j\frac{2\pi}{T}\tau}+\rho_{-\frac{1}{T}}(\tau)\rho_{\frac{1}{T}}(0)\rho_0^*(0) \\ &+\rho_{-\frac{1}{T}}(\tau)\rho_0(0)\rho_{-\frac{1}{T}}^*(0)e^{-j\frac{2\pi}{T}\tau}\Big) \end{aligned} \tag{7.57}$$

$$\begin{aligned} ⑤= (\frac{\gamma_{2s}}{T})^3\Big(&|\rho_0(\tau)|^2\rho_0(\tau)+2|\rho_{-\frac{1}{T}}(\tau)|^2\rho_0(\tau) \\ &+2|\rho_{\frac{1}{T}}(\tau)|^2\rho_0(\tau)+2\rho_{\frac{1}{T}}(\tau)\rho_{-\frac{1}{T}}(\tau)\rho_0^*(\tau)\Big). \end{aligned} \tag{7.58}$$

To obtain $S_{2y}(f)$, we take the CTFT of (7.32), i.e.,

$$\begin{aligned} S_{2y}(f) = \mathcal{F}\{c_{2y}(\tau)\} &= |a_1|^2\mathcal{F}\{①\}+4\mathrm{Re}\left(a_1^*a_3\mathcal{F}\{③\}\right) \\ &\quad +4|a_3|^2\mathcal{F}\{④\}+2|a_3|^2\mathcal{F}\{⑤\}. \end{aligned} \tag{7.59}$$

From (7.26), we see that the CTFT of $\rho_u(\tau)$ is $H^*(f+u)H(f)$. Therefore, the CTFT of (7.55)-(7.58) is respectively,

$$\mathcal{F}\{①\} = \frac{\gamma_{2s}}{T}|H(f)|^2 \tag{7.60}$$

$$\begin{aligned}\mathcal{F}\{②\} = (\frac{\gamma_{2s}}{T})^2\Big(&|H(f)|^2\rho_0^*(0) + \rho_{\frac{1}{T}}^*(0)H^*(f)H(f-\tfrac{1}{T}) \\ &+\rho_{-\frac{1}{T}}^*(0)\rho_{-\frac{1}{T}}(\tau)H^*(f)H(f+\tfrac{1}{T})\Big)\end{aligned} \tag{7.61}$$

$$\begin{aligned}\mathcal{F}\{③\} = (\frac{\gamma_{2s}}{T})^2\Big(&\rho_0(0)|H(f)|^2 + \rho_{-\frac{1}{T}}(0)H^*(f+\tfrac{1}{T})H(f) \\ &+\rho_{\frac{1}{T}}(0)H^*(f-\tfrac{1}{T})H(f)\Big)\end{aligned} \tag{7.62}$$

$$\begin{aligned}\mathcal{F}\{④\} = \\ &(\frac{\gamma_{2s}}{T})^3\Big(|\rho_0(0)|^2|H(f)|^2 + \rho_0(0)\rho_{\frac{1}{T}}^*(0)H(f-\tfrac{1}{T})H^*(f) \\ &+\rho_{-\frac{1}{T}}(0)\rho_0^*(0)H^*(f+\tfrac{1}{T})H(f) + |\rho_{\frac{1}{T}}(0)|^2|H(f-\tfrac{1}{T})|^2 \\ &+|\rho_{-\frac{1}{T}}(0)|^2|H(f+\tfrac{1}{T})|^2 + \rho_{\frac{1}{T}}(0)\rho_0^*(0)H^*(f-\tfrac{1}{T})H(f) \\ &+\rho_0(0)\rho_{-\frac{1}{T}}^*(0)H(f+\tfrac{1}{T})H^*(f)\Big)\end{aligned} \tag{7.63}$$

$$\begin{aligned}\mathcal{F}\{⑤\} = (\frac{\gamma_{2s}}{T})^3\Big(&|H(f)|^2 \star |H(f)|^2 \star |H(-f)|^2 \\ &+2[H^*(f-\tfrac{1}{T})H(f)] \star [H(-f-\tfrac{1}{T})H^*(-f)] \star |H(f)|^2 \\ &+2[H^*(f+\tfrac{1}{T})H(f)] \star [H(-f+\tfrac{1}{T})H^*(-f)] \star |H(f)|^2 \\ &+2[H^*(f+\tfrac{1}{T})H(f)] \star [H^*(f-\tfrac{1}{T})H(f)] \star |H(-f)|^2\Big).\end{aligned} \tag{7.64}$$

Using the fact that $\rho_{\frac{1}{T}}(0) = \rho_{-\frac{1}{T}}^*(0)$, we simplify (7.62) as

$$\begin{aligned}\mathcal{F}\{③\} = (\frac{\gamma_{2s}}{T})^2H(f)\Big(&\rho_0(0)H(f) + \rho_{\frac{1}{T}}(0)H(f+\tfrac{1}{T}) \\ &+\rho_{-\frac{1}{T}}(0)H(f-\tfrac{1}{T})\Big)^*\end{aligned} \tag{7.65}$$

and realize that

$$\mathcal{F}\{②\} = \mathcal{F}\{③\}^*. \tag{7.66}$$

Moreover, we use the fact that $H^*(f-\frac{1}{T})H(f+\frac{1}{T}) = 0$ to simplify (7.63) as

$$\begin{aligned}\mathcal{F}\{④\} = (\frac{\gamma_{2s}}{T})^3\Big|&\rho_0(0)H(f) + \rho_{-\frac{1}{T}}(0)H(f-\tfrac{1}{T}) \\ &+\rho_{\frac{1}{T}}(0)H(f+\tfrac{1}{T})\Big|^2.\end{aligned} \tag{7.67}$$

Substituting (7.60), (7.65)-(7.67) and (7.64) back into (7.59), we obtain (7.33).

CHAPTER VIII

ANALYZING SPECTRAL REGROWTH OF QPSK AND OQPSK SIGNALS

In this chapter, a comparison is made between the spectral regrowth of quadrature phase shift keyed (QPSK) and offset QPSK (OQPSK) signals as they go through non-linear amplifications. Contrary to existing approaches that assume the power amplifier input is Gaussian, our analysis is carried out without the Gaussian assumption, by using higher-order statistics. We show that it is possible to assess quantitatively, whether and how much OQPSK is beneficial in reducing spectral regrowth. Simple closed form formulas are obtained when the pulse shape filter is time-limited. A particular measure of spectral broadening is also provided.

8.1 Introduction

QPSK is a popular modulation format that is used in many applications (e.g., IS-95 CDMA). Let us denote a QPSK symbol by s_m where $s_m = [\pm 1 \pm j]$ with probability 0.25 each. A significant drawback of QPSK is the ± 180-degree phase change at the $1 + j \leftrightarrow -1 - j$ and the $1 - j \leftrightarrow -1 + j$ transitions. Such transitions are undesirable if the waveform is to be filtered and subsequently processed by a nonlinear power amplifier (PA).

Nonlinear PAs are used in communication systems for improved efficiency because generally, there is an inverse relationship between linearity and efficiency [35]. Higher efficiency means that a larger percentage of the dc (e.g., battery) power is delivered to the load, thus increasing battery life and minimizing heat dissipation.

Figure 8.1 shows in solid line, the AM/AM (amplitude to amplitude) conversion of a nonlinear PA. The dashed line shows in comparison, a linear AM/AM response. Although $|s_m| = \sqrt{2}$ is constant modulus, the envelope of a filtered QPSK signal could fluctuate, thus leading to nonlinear distortions. In Figure 8.1, this means that a filtered QPSK signal could

traverse the A-C region of the PA response. In addition to the PA compression at the larger amplitudes, the filtered QPSK signal also experiences cut-off when the input amplitude is close to zero.

A remedy is to employ offset QPSK (OQPSK) modulation. In OQPSK, the I- and Q-symbol streams are offset in time by half the symbol period, thus avoiding the ±180-degree phase change. For illustration purpose, we can imagine that in Figure 8.1, instead of the A-C region, the filtered OQPSK signal traverses through the B-C region of the PA, thus avoiding cut-off.

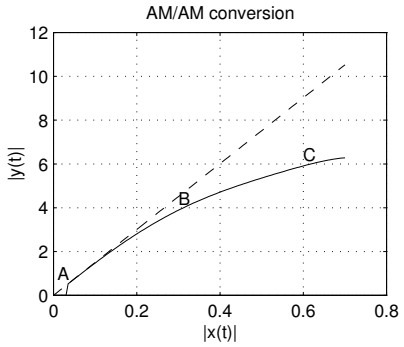


Figure 8.1: AM/AM characteristic of a nonlinear PA (solid line).

Despite of this merit of OQPSK, some concerns were raised about the overall effectiveness of employing OQPSK. First of all, it does not lend itself to differential encoding [6]. Furthermore, argument has been made that although the QPSK signal experiences the cut-off effect, it spends a very small percentage of time in the cut-off region. On the other hand, the region that the OQPSK signal spends its time with (e.g., the B-C region in Figure 1) is more compressed and hence more nonlinear than the A-B region that the QPSK signal frequently visits.

This chapter attempts to offer a means of quantitatively analyzing spectral regrowth of a communication signal passing through a nonlinear device. Specifically, we compare the power spectra of filtered QPSK and OQPSK signals after their nonlinear amplification. Although spectral analysis is routinely carried out for communications signals, the nonlinearity present in the PA complicates the problem. In [50], the authors analyzed the spectral regrowth pattern when the input signal is Gaussian, but their results have limited

applicability since many communication signals are non-Gaussian. In [57], a Volterra system approach was adopted. We address here spectral regrowth of a memoryless nonlinear device.

8.2 Problem Formulation

A linearly modulated signal is expressed in the baseband as

$$x(t) = \frac{A}{\sqrt{2}} \sum_{m=-\infty}^{\infty} s_m \; h\left(t - mT\right), \tag{8.1}$$

where $s_m = a_m + jb_m$ is the mth symbol transmitted, $h(t)$ is the baseband pulse shape filter, A is a real-valued input scale factor, and T is the symbol period. We assume that a_m and b_m are *i.i.d.* and are mutually independent. We refer to the resulting s_m as circular complex symmetric.

When s_m is QPSK, we have $a_m \in \{1, -1\}$ with equal probability 0.5, and similarly for b_m.

A filtered OQPSK signal may be written as:

$$\begin{aligned} x(t) = \frac{A}{\sqrt{2}} \sum_{m=-\infty}^{\infty} a_m h\left(t - mT\right) \\ + j \; \frac{A}{\sqrt{2}} \sum_{m=-\infty}^{\infty} b_m h\left(t - mT - \frac{T}{2}\right), \end{aligned} \tag{8.2}$$

where a_m, b_m, $h(t)$, A, and T are the same as in the filtered QPSK case. The only difference is that the imaginary part of (8.2) has a $T/2$ delay relative to that of (8.1).

Next, $x(t)$ is input to a PA to yield output $y(t)$. Ideally, we would like $y(t) = \alpha \; x(t)$, where α is a constant with $|\alpha| > 1$. But in reality, all PAs are inherently nonlinear. In the case of a memoryless nonlinear PA, we can approximate its baseband input/output relationship by [18, p. 735]:

$$y(t) = x(t) \sum_{n=0}^{\infty} a_{2n+1} \; |x(t)|^{2n}, \tag{8.3}$$

from which we infer that the complex gain is

$$\frac{y(t)}{x(t)} = \sum_{n=0}^{\infty} a_{2n+1} \; |x(t)|^{2n}. \tag{8.4}$$

It is seen that the complex gain is a function of the input amplitude $|x(t)|$ only. This is consistent with the fact that a memoryless nonlinear PA is often characterized by its AM/AM (i.e., $|y(t)|$ vs. $|x(t)|$) and AM/PM (i.e., $\angle y(t) - \angle x(t)$ vs. $|x(t)|$) characteristics. If $|x(t)|$ is constant such as the case of (8.1) with a rectangular shaped $h(t)$ (see also Section 3), then $x(t)$ will not experience any nonlinear distortion since the gain in (8.4) is constant. Our objective here is to analyze the power spectral density (PSD) of $y(t)$ and its dependence on the PA parameters $\{a_{2n+1}\}$, the baseband filter $h(t)$, and the input scale factor A.

8.3 Analysis

Although our analysis on spectral regrowth can be generalized to accommodate higher-order nonlinearities, for simplicity, we illustrate our approach using a 3rd-order nonlinear model:

$$\begin{aligned} y(t) &= a_1 x(t) + a_3 \ |x(t)|^2 x(t) \\ &= a_1 x(t) + a_3 \ x^2(t) x^*(t). \end{aligned} \tag{8.5}$$

Since s_m has a symmetric distribution, $y(t)$ has zero-mean. Therefore, the auto-correlation and auto-covariance functions of $y(t)$ coincide. We define the auto-correlation function of $y(t)$ at time t and lag τ as follows:

$$c_{2y}(t; \tau^*) = E[y(t) \ y^*(t+\tau)]. \tag{8.6}$$

In (8.6), τ^* indicates that conjugation is applied to the lagged copy, $y(t+\tau)$. Note that τ itself is always a real number.

Since $y(t)$ is cyclostationary, its time-averaged auto-correlation function is

$$\bar{c}_{2y}(\tau) = \frac{1}{T} \int_0^T c_{2y}(t; \tau^*) \ dt. \tag{8.7}$$

The power spectrum of $y(t)$ is the Fourier transform of $\bar{c}_{2y}(\tau)$:

$$\begin{aligned} S_{2y}(f) &= \mathcal{F}_{\tau \to f} \left\{ \bar{c}_{2y}(\tau) \right\} \\ &= \int_{-\infty}^{\infty} \bar{c}_{2y}(\tau) \ e^{-j2\pi\tau f} \ d\tau. \end{aligned} \tag{8.8}$$

We would like to examine $S_{2y}(f)$ for the PA model in (8.5) and the input as in (8.1) or (8.2). Substituting (8.5) into (8.6), we obtain

$$\begin{aligned} c_{2y}(t;\tau) &= E\left[y(t)y^*(t+\tau)\right] \\ &= |a_1|^2 E\left[x(t)x^*(t+\tau)\right] \\ &\quad + a_1 a_3^* E\left[x(t)|x(t+\tau)|^2 x^*(t+\tau)\right] \\ &\quad + a_3 a_1^* E\left[x^*(t+\tau)|x(t)|^2 x(t)\right] \\ &\quad + |a_3|^2 E\left[|x(t)|^2 x(t)|x(t+\tau)|^2 x^*(t+\tau)\right]. \end{aligned} \tag{8.9}$$

Alternatively, we write

$$\begin{aligned} c_{2y}(t;\tau) = |a_1|^2\ \phi_{11}(t;\tau) + a_1 a_3^*\ \phi_{13}(t;\tau) \\ + a_3 a_1^*\ \phi_{31}(t;\tau) + |a_3|^2\ \phi_{33}(t;\tau) \end{aligned} \tag{8.10}$$

where

$$\begin{aligned} \phi_{11}(t;\tau) &= \mathrm{cov}\left\{x(t), x^*(t+\tau)\right\} \\ \phi_{13}(t;\tau) &= \mathrm{cov}\left\{x(t)|x(t+\tau)|^2, x^*(t+\tau)\right\} \\ \phi_{31}(t;\tau) &= \mathrm{cov}\left\{x^*(t+\tau), |x(t)|^2 x(t)\right\} \\ \phi_{33}(t;\tau) &= \mathrm{cov}\left\{|x(t)|^2 x(t), |x(t+\tau)|^2 x^*(t+\tau)\right\}. \end{aligned} \tag{8.11}$$

Our next step is to expand the above covariance terms using the Leonov-Shiryaev formula [25]. Under the circular symmetry assumption of s_m, we infer that $x(t)$ of (8.1) is

circular symmetric as well. Therefore, we find for filtered QPSK,

$$\begin{aligned}
\phi_{11}(t;\tau) &= c_{2x}(t;\tau^*) \\
\phi_{13}(t;\tau) &= c_{4x}(t;\tau^*,\tau,\tau^*) + 2c_{2x}(t;\tau^*)c_{2x}(t+\tau^*;0) \\
\phi_{31}(t;\tau) &= c_{4x}(t;0^*,0,\tau^*) + 2c_{2x}(t;\tau^*)c_{2x}(t;0^*) \\
\phi_{33}(t;\tau) &= c_{6x}(t;0^*,0,\tau^*,\tau,\tau^*) \\
&\quad +4c_{4x}(t;0^*,\tau,\tau^*)c_{2x}(t;\tau^*) \\
&\quad +2c_{4x}(t;\tau^*,\tau,\tau^*)c_{2x}(t;0^*) \\
&\quad +2c_{4x}(t;0^*,0,\tau^*)c_{2x}(t+\tau;0^*) \\
&\quad +c_{4x}(t;0,\tau^*,\tau^*)c_{2x}(t^*;\tau) \\
&\quad +4c_{2x}(t;\tau^*)c_{2x}(t;0^*)c_{2x}(t+\tau;0^*) \\
&\quad +2c_{2x}(t;\tau^*)c_{2x}(t;\tau^*)c_{2x}(t^*;\tau).
\end{aligned}$$

Note that the OQPSK signal (8.2) is no longer circular symmetric and hence the corresponding ϕ_{13}, ϕ_{31}, and ϕ_{33} expressions contain additional terms.

The kth-order cumulant of $x(t)$ at time t and lags $(\tau_1, \ldots, \tau_{k-1})$ is defined as

$$\begin{aligned}
&c_{kx}(t;\tau_1,\ldots,\tau_{\ell-1},\tau_\ell^*,\ldots,\tau_{k-1}^*) \\
&\triangleq \mathrm{cum}\{x(t), x(t+\tau_1),\ldots,x(t+\tau_{\ell-1}), \\
&\qquad x^*(t+\tau_\ell),\ldots,x^*(t+\tau_{k-1})\}.
\end{aligned}$$

Note that a conjugated lag in the argument of $c_{kx}(\cdot)$; e.g., τ_ℓ^*, implies that the corresponding term in the cumulant; e.g., $x^*(t+\tau_\ell)$, has conjugation. For the $x(t)$ in (8.1), we have

$$\begin{aligned}
&c_{kx}(t;\tau_1,\ldots,\tau_\ell^*,\ldots,\tau_{k-1}^*) \\
&= \gamma_{ks}\ \left(\frac{A}{\sqrt{2}}\right)^k \sum_m h(t-mT)h(t-mT+\tau_1)\ldots
\end{aligned}$$

$$h(t-mT+\tau_{\ell-1})h^*(t-mT+\tau_\ell)\ldots h^*(t-mT+\tau_{k-1}),$$

and

$$\gamma_{ks} = \mathrm{cum}\,\{s(t), s(t+\tau_1),\ldots,s(t+\tau_{\ell-1}),$$

$$s^*(t+\tau_\ell), \ldots, s^*(t+\tau_{k-1})\} .$$

Next, let us analyze $S_{k\ell}(f)$, which is the Fourier transform of $\bar{\phi}_{k\ell}(\tau)$, the time-average of $\phi_{k\ell}(t;\tau)$.

Interestingly, when $h(t) = 0$, $\forall\ |t| > T/2$, the $\phi_{k\ell}(t;\tau)$ expressions can be simplified considerably. As a result, we obtain

$$S_{11}(f) = \frac{A^2}{T} |H(-f)|^2 \tag{8.12}$$

$$S_{13}(f) = \frac{A^4}{T} H(-f) \left[H(f) \circledast H^*(-f) \circledast H^*(-f)\right] \tag{8.13}$$

$$S_{31}(f) = \frac{A^4}{T} H^*(-f) \left[H^*(f) \circledast H(-f) \circledast H(-f)\right] \tag{8.14}$$

$$S_{33}(f) = \frac{A^6}{T} |H^*(f) \circledast H(-f) \circledast H(-f)|^2 , \tag{8.15}$$

where $H(f)$ is the Fourier transform of $h(t)$, and $\circledast$ is the convolution operator.

When $h(t)$ is real valued and symmetric, we obtain a surprisingly simple expression for the PSD of $y(t)$:

$$S_{2y}(f) = |a_1|^2 A^2 \frac{1}{T} \left| H(f) + \frac{a_3}{a_1} A^2 H_3(f) \right|^2 \tag{8.16}$$

where $H_3(f) = H(f) \circledast H(f) \circledast H(f)$.

We make the following remarks regarding (8.16):

Remark 1: Potential spectral regrowth is indicated by the $H_3(f)$ term which generally expands the bandwidth of $H(f)$ through the triple convolution.

Remark 2: The severity of spectral regrowth is determined by the coefficient $(a_3/a_1)A^2$ in (8.16). If the PA is inherently very nonlinear; i.e., the a_3/a_1 ratio is high, then one needs to reduce the input amplitude factor A in order to minimize spectral regrowth – this is referred to as input back-off. In general, spectral regrowth becomes more severe as A increases.

Now let us consider two baseband filters often studied in the literature [88]:

$$h(t) = \begin{cases} 1 & |t| \le \frac{T}{2} \\ \\ 0 & |t| > \frac{T}{2} \end{cases} \tag{8.17}$$

and

$$h(t) = \begin{cases} \sqrt{2}\cos(\pi\frac{t}{T}) & |t| \leq \frac{T}{2} \\ \\ 0 & |t| > \frac{T}{2}. \end{cases} \tag{8.18}$$

For the rectangular pulse (8.17), we obtain

$$H(f) = \frac{1}{f_o}\frac{\sin(\frac{f\pi}{f_o})}{(\frac{f\pi}{f_o})}, \tag{8.19}$$

and

$$H_3(f) = \frac{1}{f_o}\frac{\sin(\frac{f\pi}{f_o})}{(\frac{f\pi}{f_o})} = H(f). \tag{8.20}$$

Substituting (8.19)-(8.20) into (8.16), we infer that there is no spectral regrowth when the rectangular pulse is used for the $x(t)$ in (8.1). This is expected since in this case, the resulting $|x(t)| = A$ has constant envelope.

For the sinusoidal pulse (8.18), we have

$$H(f) = \frac{\sqrt{2}}{2\pi f_o}\frac{\cos(\frac{f\pi}{f_o})}{\frac{1}{4} - \left(\frac{f}{f_o}\right)^2}, \tag{8.21}$$

and

$$H_3(f) = \frac{3\sqrt{2}}{2\pi f_o}\frac{\cos(\frac{f\pi}{f_o})}{\left(\frac{1}{4} - \left(\frac{f}{f_o}\right)^2\right)\left(\frac{9}{4} - \left(\frac{f}{f_o}\right)^2\right)}. \tag{8.22}$$

This $H_3(f)$ can be shown to have a wider mainlobe than $H(f)$.

One way to quantify spectral regrowth is to use a notion of bandwidth $\sqrt{\langle f^2\rangle}$ where

$$\langle f^2\rangle \triangleq \frac{\int_{-\infty}^{\infty}(f - \langle f\rangle)^2 S(f)df}{\int_{-\infty}^{\infty} S(f)df} \tag{8.23}$$

and

$$\langle f\rangle \triangleq \frac{\int_{-\infty}^{\infty} fS(f)df}{\int_{-\infty}^{\infty} S(f)df}. \tag{8.24}$$

Note that for a symmetric spectrum $S(f)$, the corresponding $\langle f\rangle = 0$.

Substitution of (8.16), (8.21), and (8.22) into (8.23) yields the following bandwidth formula for the cosine pulse (8.18):

$$\sqrt{\langle f^2 \rangle} = \frac{f_o}{2} \sqrt{\frac{1 - 3\ \mathrm{Re}(\beta) + 4.5\ |\beta|^2}{1 - 3\ \mathrm{Re}(\beta) + 2.5\ |\beta|^2}},$$
$$\beta = -\frac{a_3}{a_1}\ A^2. \tag{8.25}$$

When the PA is linear, we have $a_3 = 0$ and hence $\beta = 0$. The bandwidth formula (8.25) yields $0.5 f_o$ as the bandwidth of a linear system. Therefore the ratio, $\sqrt{\langle f^2 \rangle}/(0.5 f_0)$, can be used as a measure of bandwidth expansion and from (8.25), it is obvious that this ratio is > 1 for any $\beta \neq 0$.

When $x(t)$ is OQPSK, the analysis is generally more involved. But with either (8.17) or (8.18), the OQPSK signal in (8.2) has $|x(t)| = A$ and hence the corresponding PA output $y(t) = (a_1 + a_3\ A^2)x(t)$ does not experience any spectral regrowth.

8.4 Simulations

In this section, we present a numerical example to verify the accuracy of the expressions (8.12)-(8.16). $1,000$ symbols s_m were generated and a filtered QPSK signal $x(t)$ was obtained from equation (8.1) with the cosine pulse (8.18). The scale factor was $A = 1$ and the sampling period was $\frac{1}{40}T$ seconds. The resulting $x(t)$ went through nonlinear amplification as described by (8.5) with $a_1 = 1$ and $a_3 = -0.3 \exp(j\frac{\pi}{4})$.

Figure 8.2 shows the theoretical $S_{11}(f)$ (c.f. (8.12)) in solid line and its estimate in dashed line. The estimate is nothing but the PSD estimate of $x(t)$. Close agreement between the two is observed.

Figure 8.3 shows the theoretical $S_{33}(f)$ (c.f. (8.15)) in solid line and its estimate – the PSD estimate of $x(t)^2 x^*(t)$ in dashed line. Similar agreement is observed. Comparing with Figure 8.2, we see that the bandwidth of $S_{33}(f)$ has increased from that of $S_{11}(f)$. Indeed, $\sqrt{\langle f^2 \rangle}$ of $S_{33}(f)$ is $\sqrt{1.8}$ times or 34% larger than that of $S_{11}(f)$ (c.f. (8.23)). Moreover, evaluation of the (8.12)-(8.15) terms reveals that $S_{33}(f)$ is the major contributor to spectral regrowth in $S_{2y}(f)$.

In Figure 8.4, we show a comparison between the output spectrum $S_{2y}(f)$ when the

input is QPSK (solid line) or OQPSK (dashed line). The bandwidth of the QPSK signal is indeed larger than that of the OQPSK signal. In fact, equation (8.25) with $\beta = 0.3 \exp(j\frac{\pi}{4})$ tells that the bandwidth increase was 14%.

8.5 Conclusions

We have described in this chapter, an analytical approach for evaluating the power spectra of filtered QPSK and OQPSK signals after nonlinear amplification. A salient feature of our analysis is that we do not need to assume that the PA input is Gaussian. In the QPSK case, we were able to obtain a simple closed form expression for the output PSD when the PA is cubic nonlinear and the baseband filter is time-limited. We treated the cosine pulse filtered QPSK/OQPSK signals in detail and provided a measure of bandwidth expansion. We are currently working on applying our analysis to more general scenarios.

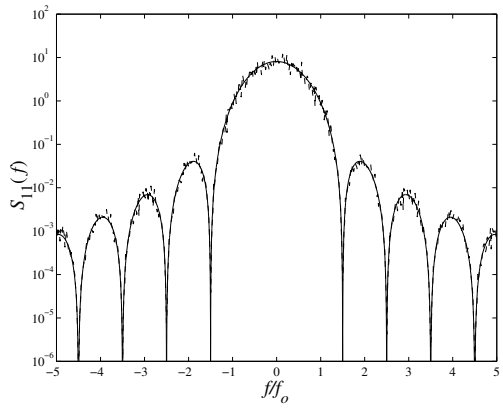


Figure 8.2: The theoretical $S_{11}(f)$ (solid line) and its estimate (dashed line).

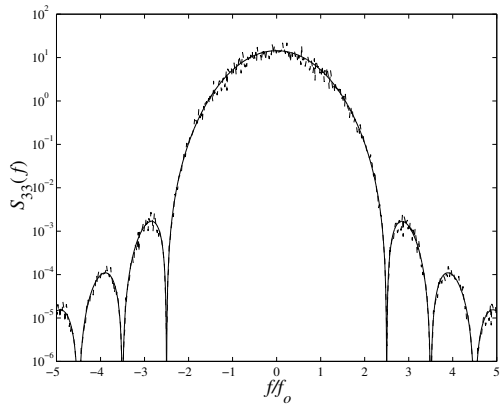


Figure 8.3: The theoretical $S_{33}(f)$ (solid line) and its estimate (dashed line).

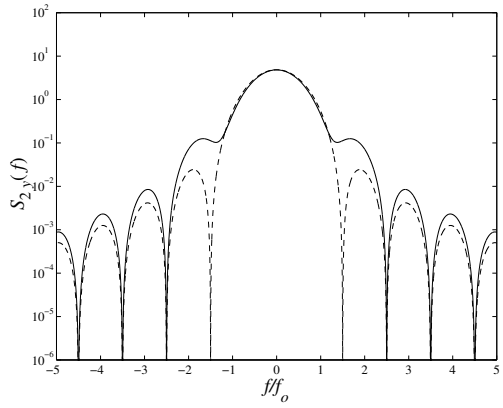


Figure 8.4: The PA output PSD $S_{2y}(f)$ when the input is QPSK (solid line) or OQPSK (dashed line).

CHAPTER IX

SPECTRAL ANALYSIS OF POLYNOMIAL NONLINEARITY WITH APPLICATIONS TO RF POWER AMPLIFIERS

The majority of the nonlinearity in a communication system is attributed to the power amplifier (PA) present at the final stage of the transmitter chain. In this chapter, we consider Gaussian distributed input signals (such as OFDM), and PAs that can be modeled by memoryless or memory polynomials. We derive closed form expressions of the PA output power spectral density, for an arbitrary nonlinear order, based on the so-called Leonov-Shiryaev formula. We then apply these results to answer practical questions such as the contribution of AM/PM conversion to spectral regrowth, and the relationship between memory effects and spectral asymmetry.

9.1 Introduction

Power amplifiers (PAs) are important components of communications systems and are inherently nonlinear. For example, the so-called Class AB PAs, which are moderately nonlinear, are typically employed in wireless basestations and handsets. When a non-constant modulus signal goes through a nonlinear PA, spectral regrowth (broadening) appears in the output, which in turn causes adjacent channel interference (ACI). Stringent limits on ACI are imposed by the standard bodies and thus the extent of the PA nonlinearity must be controlled.

We are interested in predicting the amount of spectral regrowth for a given level of PA nonlinearity. Since more linear PAs are less efficient, one may want to maximize nonlinearity (and hence optimize efficiency) subject to the spectral mask constraint. Such optimization

strategy is feasible if we have tools for spectral regrowth analysis of the nonlinear output.

If the PA input is Gaussian, the PA output power spectral density (PSD) has been derived for a 5th-order nonlinear PA in [113], [44]. In [50], the analysis was carried out for a 9th-order nonlinear PA. The results in [21] are fairly general but developed for bandpass signals, whereas references [44], [50], [113] and this chapter adopt a baseband nonlinear formulation. In [49], a general expression is given without proof. When the PA input is non-Gaussian, theoretical analysis becomes more complicated, but results are available in [129] for a 7th-order nonlinear PA with (non-)Gaussian inputs.

The objective of this chapter is to derive closed-form expressions for the PA output PSD (or output auto-covariance function) for an arbitrary nonlinear order, for both the memoryless and memory baseband polynomial PA models. The PA input is assumed to be Gaussian distributed, which is a reasonable assumption for OFDM signals [44], forward link CDMA signals with a large number of Walsh-coded channels at the same frequency [7], or signals at the satellite-borne relay [21]. The Gaussian assumption significantly reduces the complexity of the analysis. Equipped with these formulas, we can then answer practical questions such as how important or necessary it is to correct for the AM/PM distortion in the PA, and possible mechanisms for spectral asymmetry in the PA output spectrum.

We would like to emphasize that the PA models considered in this chapter belong to the polynomial family [75], [72]; i.e., polynomials or Taylor series for the (quasi-) memoryless case, and Volterra series for the case with memory. Polynomials and Volterra series are frequently used in PA modeling; see e.g., [18], [21], [35], [44], [50], [72], [113], [129].

The organization of the chapter is as follows. In Section 9.2, we outline the approach of spectral analysis for a baseband nonlinear system with cyclostationary input, suitable for digital communication signals. We will investigate the well known (quasi-) memoryless PA model in Section 9.3, and then study the relatively recent memory polynomial model in Section 9.4. Conclusions are drawn in Section 9.5. In order not to interrupt the flow of the chapter, we defer the rather technical proofs of our theorems to Section 9.6.

9.2 Cyclostationary input and spectral analysis

A digital communication signal $x(t)$ is represented by

$$x(t) = \sum_k s_k \; h(t - kT), \tag{9.1}$$

where s_k is the kth symbol, $h(t)$ is the pulse shaping filter, and T is the symbol period. Thus, $x(t)$ is strict-sense cyclostationary in general [51, Ch. 12], [54].

Let us denote by $\text{cum}\{\cdot\}$, the cumulant operator. The first-order cumulant is the mean; the second-order cumulant is the covariance. General definitions and properties of cumulants can be found in [25]. The auto-covariance function of the PA input signal $x(t)$ at time t and lag τ is defined as

$$c_{2x}(t;\tau) = \text{cum}\{x^*(t), x(t+\tau)\}. \tag{9.2}$$

Closed-form spectral analysis for a nonlinear system with nonstationary (or cyclostationary) input is in general extremely difficult (if at all possible), even under the Gaussian $x(t)$ assumption. Therefore, we focus our attention on the case where the bandwidth of the pulse shaping filter is limited to $1/T$ (i.e., $h(t)$ has no excess bandwidth). Denote by $H(f)$ the Fourier transform (FT) of $h(t)$; i.e.,

$$H(f) = \int h(t) \; e^{-j2\pi ft} \; dt; \tag{9.3}$$

this assumption implies that $H(f) = 0$, $\forall |f| > 1/(2T)$.

If s_k is zero-mean, i.i.d. with variance σ_s^2, we show next that $x(t)$ in (9.1) is wide-sense stationary; i.e., $c_{2x}(t;\tau) = c_{2x}(\tau)$, $\forall t$.

First, it is straightforward to show that

$$c_{2x}(t;\tau) = \sigma_s^2 \sum_k h^*(t - kT) h(t + \tau - kT) \tag{9.4}$$

for the $x(t)$ in (9.1). Next, recall the inverse FT relationship

$$h(t) = \int H(f) \; e^{j2\pi ft} \; df. \tag{9.5}$$

Substituting (9.5) into (9.4) and using the fact that

$$\sum_m \frac{1}{T} \; \delta(f - \frac{m}{T}) = \sum_k e^{j2\pi fkT}, \tag{9.6}$$

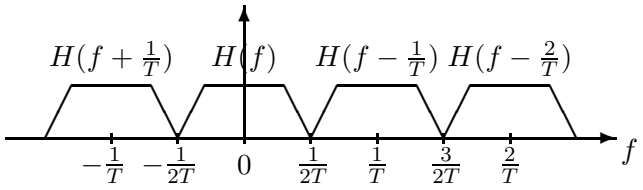


Figure 9.1: When $H(f)$ has no excess bandwidth, $H^*(f+m/T)H(f)=0$, $\forall m \neq 0$.

we obtain

$$c_{2x}(t;\tau) = \frac{\sigma_s^2}{T} \sum_m e^{-j2\pi mt/T} \int H^*(f+m/T)H(f)\ e^{j2\pi f\tau}\ df. \tag{9.7}$$

From (9.7), it is clear that the t-dependence in $c_{2x}(t;\tau)$ comes from the $e^{-j2\pi mt/T}$ term, if $m \neq 0$. Equation (9.7) can also be viewed as a synthesis equation for the time-varying correlation function in terms of cyclic correlation with cycles $-2\pi m/T$. The bandwidth of $H(f)$ affects the number of cycles present in $c_{2x}(t;\tau)$ [111], [32].

Since the bandwidth of $H(f)$ is limited to $1/T$, $H(f+m/T)$ and $H(f)$ do not overlap if $m \neq 0$ (see Fig. 9.1), and hence the product $H^*(f+m/T)H(f)=0$, $\forall m \neq 0$. As a result, only the $m=0$ term survives in the summation in (9.7) and

$$c_{2x}(t;\tau) = \frac{\sigma_s^2}{T} \int |H(f)|^2\ e^{j2\pi f\tau}\ df, \tag{9.8}$$

which is not a function of t. Therefore, under the no excess bandwidth assumption, $c_{2x}(t;\tau) = c_{2x}(\tau)$, $\forall t$, meaning that $x(t)$ is wide-sense stationary.

Since all cumulants of order ≥ 3 vanish for Gaussian processes, a wide-sense stationarity Gaussian $x(t)$ is also strict-sense stationarity. From now on, we will drop the t-dependence and express the auto-covariance function of $x(t)$ as $c_{2x}(\tau)$.

We point out that (wide-sense) stationarity of $x(t)$ is assumed in [21], [44], [50], [113], [129], often without justification.

The power spectral density (PSD) of $x(t)$ is defined as the Fourier Transform (FT) of $c_{2x}(\tau)$:

$$S_{2x}(f) = \int c_{2x}(\tau)\ e^{-j2\pi f\tau}\ d\tau. \tag{9.9}$$

Next, we will relate the PSD of the baseband PA output $y(t)$ to that of the baseband PA input $x(t)$, when $x(t)$ and $y(t)$ obey polynomial nonlinear relationships.

9.3 Quasi-memoryless PA model

The following model is commonly used to describe memoryless PAs in the baseband; see e.g., [18, p. 69],

$$y(t) = \sum_{k=0}^{K} a_{2k+1}\ [x(t)]^{k+1}\ [x^*(t)]^k \tag{9.10}$$

$$= x(t)\ \sum_{k=0}^{K} a_{2k+1}\ |x(t)|^{2k}, \tag{9.11}$$

where $\{a_{2k+1}\}$ are the (complex-valued) coefficients for the PA. We see from (9.11) that the complex gain is $G(x(t)) = y(t)/x(t) = \sum_{k=0}^{K} a_{2k+1}\ |x(t)|^{2k}$, which is a function of $r = |x(t)|$ only.

Writing the complex gain as $G(r) = A(r)\ e^{j\Phi(r)}$, we refer to $A(r)$ as the AM/AM conversion, and $\Phi(r)$ as the AM/PM conversion. A linear PA would have constant $A(r)$ and $\Phi(r)$ characteristics. If $A(r)$ is non-constant but $\Phi(r)$ is, the corresponding PA is called strictly memoryless. If both $A(r)$ and $\Phi(r)$ are non-constant, the resulting PA is called quasi-memoryless. Eq. (9.10) can be used to describe both types of memoryless nonlinearity, and hence we do not distinguish the two in subsequent analysis.

9.3.1 Closed form expression for spectral regrowth

We assume that $x(t)$ is circular complex in the sense that

$$\text{cum}\{x(t), x(t+\tau)\} = 0, \qquad \forall \tau. \tag{9.12}$$

Let us write $x(t) = x_R(t) + jx_I(t)$, where $x_R(t)$ and $x_I(t)$ are the real and imaginary parts of $x(t)$, respectively. It can be shown that eq. (9.12) is equivalent to

$$\text{cum}\{x_R(t), x_R(t+\tau)\} = \text{cum}\{x_I(t), x_I(t+\tau)\},$$

$$\text{cum}\{x_R(t), x_I(t+\tau)\} = -\text{cum}\{x_I(t), x_R(t+\tau)\}.$$

Processes satisfying (9.12) have also been referred to as complex video processes [101]. This assumption is commonly used; see [21], [44], [50], [113], [129].

We now present the first theorem which relates the output PSD $S_{2y}(f)$ to the input PSD $S_{2x}(f)$ and (quasi-) memoryless PA parameters $\{a_{2k+1}\}$.

Theorem 2. *Assume that $x(t)$ is stationary, zero-mean, complex Gaussian distributed and satisfies (9.12). If the output $y(t)$ is related to the input $x(t)$ through (9.10), then the autocorrelation function of $y(t)$ is*

$$c_{2y}(\tau) = \sum_{m=0}^{K} \alpha_{2m+1} \; |c_{2x}(\tau)|^{2m} \; c_{2x}(\tau), \tag{9.13}$$

where the constant coefficient

$$\alpha_{2m+1} = \frac{1}{m+1} \left| \sum_{k=m}^{K} a_{2k+1} \; \binom{k}{m} \; (k+1)! \; [c_{2x}(0)]^{k-m} \right|^2, \tag{9.14}$$

and

$$\binom{k}{m} = \frac{k!}{m!(k-m)!}.$$

The PSD of $y(t)$ is related to that of $x(t)$ through

$$S_{2y}(f) = \sum_{m=0}^{K} \alpha_{2m+1} \; \underbrace{S_{2x}(f) \star \cdots \star S_{2x}(f)}_{m+1} \; \star \; \underbrace{S_{2x}(-f) \star \cdots \star S_{2x}(-f)}_{m}, \tag{9.15}$$

where $\star$ denotes convolution.

Proof: See Section 9.6.1.

Some remarks are now in order:

(R1) From (9.15), we infer that if $S_{2x}(f)$ has bandwidth B_x, $y(t)$ has bandwidth $B_y = (2K+1)B_x$, due to the spectral expansion caused by the convolution.

(R2) If $S_{2x}(f)$ is symmetric; i.e., $S_{2x}(f) = S_{2x}(-f)$, then $S_{2y}(f)$ is symmetric as well. This means that a (quasi-) memoryless PA will not lead to spectral asymmetry in the PA output.

(R3) If $S_{2x}(f)$ is asymmetric, the $2m$ times spectral convolution on the RHS of (9.15) will yield a more symmetric spectrum for larger m.

Next, we would like to provide detailed expressions for the 9th-order nonlinear PA; i.e., $K = 4$ in (9.10). Equation (9.15) yields for $K = 4$,

$$\alpha_1 = |a_1 + 2a_3c_{2x}(0) + 6a_5c_{2x}^2(0) + 24a_7c_{2x}^3(0) + 120a_9c_{2x}^4(0)|^2, \tag{9.16}$$

$$\alpha_3 = 2|a_3 + 6a_5c_{2x}(0) + 36a_7c_{2x}^2(0) + 240a_9c_{2x}^3(0)|^2, \tag{9.17}$$

$$\alpha_5 = 12\ |a_5 + 12a_7c_{2x}(0) + 120a_9c_{2x}^2(0)|^2, \tag{9.18}$$

$$\alpha_7 = 144\ |a_7 + 20a_9c_{2x}(0)|^2, \tag{9.19}$$

$$\alpha_9 = 2880\ |a_9|^2. \tag{9.20}$$

It is important to cross-verify (9.16)-(9.20) with previously published results to validate our closed form expression. We shall compare with three references below.

- In [113], $c_{2x}(\tau)$ was defined as $0.5\text{cum}\{x^*(t), x(t+\tau)\}$ (equation (27) of [113]). Once we have taken care of this scaling difference, equations (9.16)-(9.20) can be shown to agree with equation (38)[1] of [113], which holds for up to 5th-order nonlinearities.

- In [129], $x(t)$ was assumed to be circular complex symmetric which renders $c_{2x}(\tau)$ real-valued. Except for the $[c_{2x}(\tau)]^{2m+1}$ vs. $|c_{2x}(\tau)|^{2m}c_{2x}(\tau)$ difference, equations (9.16)-(9.20) agree with the expressions presented in Section III.B of [129], where a 7th-order nonlinear model was considered.

- In [50], the output PSD expression was obtained for a 9th-order nonlinear PA model[2]. Our equations (9.16)-(9.20) agree with the expressions[3] found on p. 1068 of [50].

In conclusion, previously published results in [113], [50], [129] can be regarded as special cases of our closed form expression (9.15).

9.3.2 Case study: The effect of AM/PM conversion on spectral regrowth

Although by reducing the input power level to the PA (i.e., with input back-off), one can reduce the amount of spectral regrowth, the efficiency of the PA is also diminished.

[1] Reference [113] has a typo in equation (38): $48\mathcal{R}\{\eta_1\eta_3^*\}$ should be $48\mathcal{R}\{\eta_1\eta_5^*\}$.

[2] Although the baseband input-output relationship is incorrectly expressed in equation (7) of [50], the correct baseband model was used in equation (A.5) of [50].

[3] Reference [50] has a typo on p. 1068: $15\tilde{a}_9R_{zo}$ should be $20\tilde{a}_9R_{zo}$.

Some form of PA linearization is often sought in order to achieve both good linearity and efficiency. In order to adopt an effective linearization strategy, it is important to understand the nonlinear effects present and their manifestation in terms of spectral regrowth[4]. For a given (quasi-) memoryless PA, it is useful to assess the relative contributions from the AM/AM and AM/PM conversions to spectral regrowth. We can do so using Theorem 1.

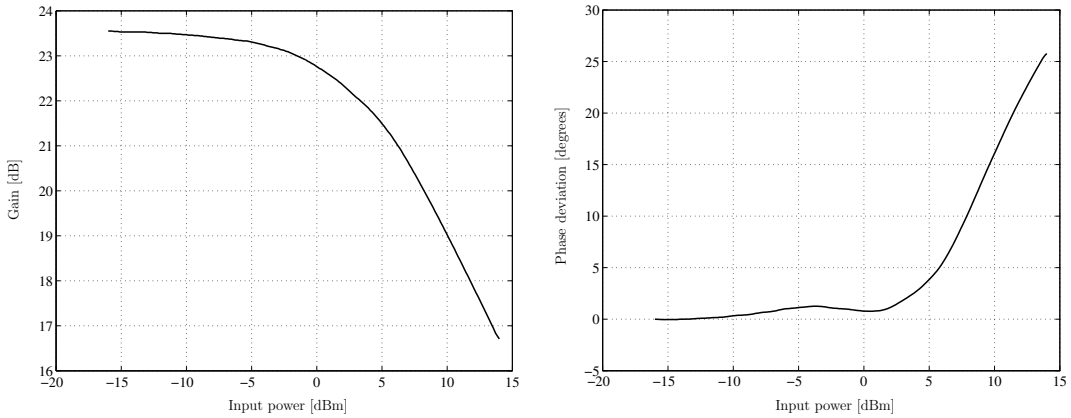
Given measured PA AM/AM characteristic $A(r)$ and AM/PM characteristic $\Phi(r)$, we can then calculate the complex gain $G(r) = A(r)\, e^{j\Phi(r)}$. Note that although the PA output $y(t)$ is a nonlinear function of the PA input $x(t)$, $y(t)$ is linear in the model coefficients $\{a_{2k+1}\}$. Therefore, regressing $rG(r)$ w.r.t. the basis $\{r, r^3, \ldots, r^{2K+1}\}$, we can estimate the model parameters $\{a_{2k+1}\}$ via linear least squares. Afterwards, we apply Theorem 1 to calculate the output PSD $S_{2y}(f)$.

To assess the individual contribution from the AM/AM conversion to $S_{2y}(f)$, we set[5], $\Phi(r) = 0$ and find the $\{a_{2k+1}\}$ coefficients corresponding to $G(r) = A(r)$. On the other hand, to evaluate the individual contribution of the AM/PM effect to spectral regrowth, we set $A(r) = A$ (the intended linear gain of the PA), and find the $\{a_{2k+1}\}$ coefficients corresponding to $G(r) = A\, e^{j\Phi(r)}$ as described in the previous paragraph.

Example 1. Fig. 9.2 shows the AM/AM and AM/PM characteristics of an actual Class AB PA. Table 9.1 lists the extracted PA model parameters for three scenarios: (i) when both AM/AM and AM/PM conversions are present; (ii) when only the AM/AM conversion is present ($\Phi(r) = 0$); and (iii) when only the AM/PM conversion is present ($A(r) = 11.75$ was used so that the corresponding output power $c_{2y}(0)$ remains the same as in case (i) and case(ii)).

First, we would like to verify that the closed form expression (9.15) is accurate. We generated 65,536 samples of the PA input $x(t)$, by passing a zero-mean, i.i.d., circular complex Gaussian process through a 48-tap lowpass filter; the variance of $x(t)$ was set to $\sigma_x^2 = c_{2x}(0) = 0.32^2$. The PA output $y(t)$ was formed according to $y(t) = x(t)A(|x(t)|)\, e^{j\Phi(|x(t)|)}$.

[4]The error vector magnitude should also be reduced, which is not the subject of this chapter.
[5]If we set $\Phi(r) = c$, the PSD $S_{2y}(f)$ can be shown to be independent of the constant c.



(a) AM/AM (b) AM/PM

Figure 9.2: Measured AM/AM and AM/PM characteristics of a Class AB PA.

The sample and the theoretical $S_{2x}(f)$ and $S_{2y}(f)$ are shown in Fig. 9.3. They show good agreement and verify that our PSD formula (9.15) is accurate.

Next, we apply (9.15) to predict spectral regrowth for the above three scenarios. From Fig. 9.4, we see that for the particular PA given in Fig. 9.2 and for the Gaussian input described above, both AM/AM and AM/PM conversions contribute significantly to spectral regrowth. If one does not apply any linearization technique to the PA, the output PSD will be at the level indicated by the solid line in Fig. 9.4. If with a linearization method, we can completely correct for the AM/AM distortion, the resulting $S_{2y}(f)$ would be given by the dash-dotted line, which is attributed solely to the AM/PM conversion. The remaining spectral regrowth is still high and additional linearization, aimed at reducing the AM/PM distortion, may be necessary.

In [66], a predistortion linearization algorithm was implemented for a handset which only corrects the AM/AM distortion of the PA. Example 1 however, shows that one should be careful not to under-estimate the effects of AM/PM distortion. Of course, one has to evaluate the particular $A(r)$ and $\Phi(r)$ characteristics to draw pertinent conclusions.

Scenarios	(i) AM/AM + AM/PM	(ii) AM/AM only	(iii) AM/PM only
a_1	$14.8526 - j0.1337$	14.8469	$11.7443 - j0.1562$
a_3	$-23.1899 + j6.9785$	-23.3505	$0.4681 + j5.9639$
a_5	$30.5226 - j1.9699$	33.8272	$-4.7569 + j6.9758$
a_7	$-21.5517 - j4.7097$	-25.4177	$4.8612 - j13.7023$
a_9	$6.0311 + j2.7527$	7.3773	$-1.5655 + j5.6319$

Table 9.1: Estimated polynomial PA model coefficients for three scenarios: (i) when both AM/AM and AM/PM conversions are present; (ii) when only the AM/AM conversion is present ($\Phi(r) = 0$); and (iii) when only the AM/PM conversion is present ($A(r) = 11.75$ was used).

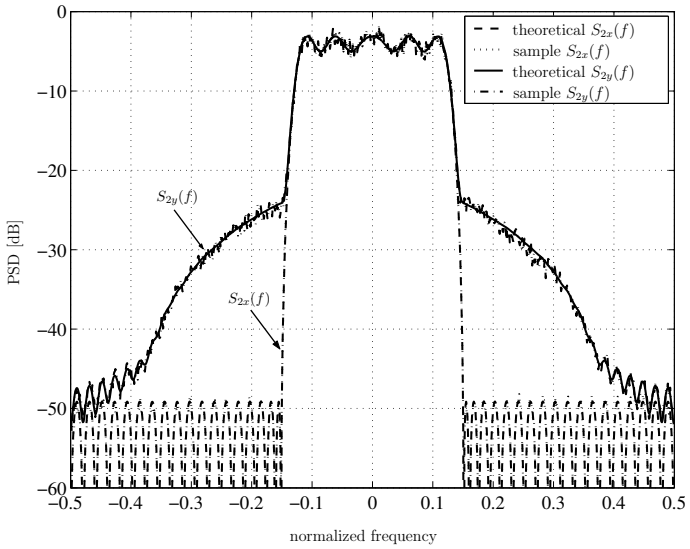


Figure 9.3: The theoretical $S_{2x}(f)$ is shown as the dashed line, the sample $S_{2x}(f)$ is shown as the dotted line; the theoretical $S_{2y}(f)$ is shown as the solid line, and the sample $S_{2y}(f)$ is shown as the dash-dotted line. The sample and the theoretical PSDs are very close (the dashed line and the dotted line almost coincide; the solid line and the dash-dotted line almost coincide), indicating that formula (9.15) is accurate. Note that we have lowered $S_{2y}(f)$ by 21.4dB to facilitate easier visual comparison between $S_{2x}(f)$ and $S_{2y}(f)$.

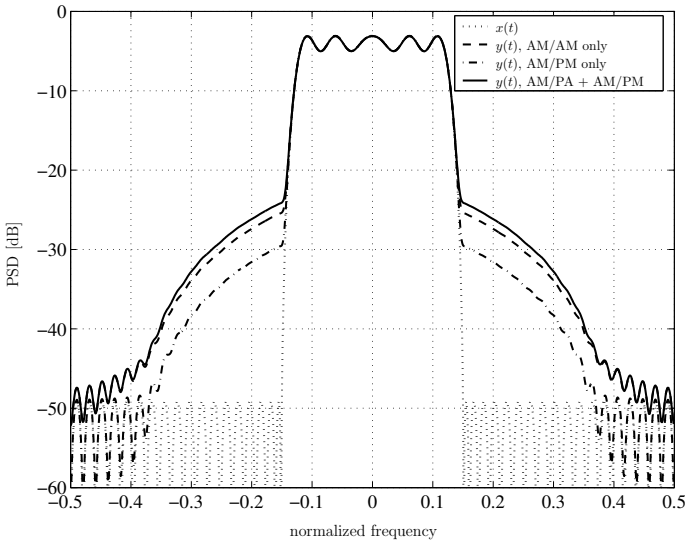


Figure 9.4: The theoretical $S_{2x}(f)$ is shown as the dotted line, the theoretical $S_{2y}(f)$ is shown as the solid line for scenario (i), as the dashed line for scenario (ii), and as the dash-dotted line for scenario (iii). In this example, we observe that both AM/AM and AM/PM conversions contribute to spectral regrowth in $S_{2y}(f)$, hence correcting only one of the distortions does not go far enough to suppress spectral regrowth.

9.4 Memory polynomial PA model

For low power amplifiers and/or narrowband input, the PA can be regarded as (quasi-) memoryless. However, high-power amplifiers (HPAs) such as those used in wireless basestations exhibit memory effects; wideband signals (such as WCDMA) also tend to induce memory effects in the PA. In general, the cause of memory effects can be electrical or electro-thermal [118]. When long-term memory effects are present, AM/AM and AM/PM conversions are insufficient to characterize the PA, and more elaborate models such as the Volterra series can be used; e.g., [72], [24].

Although the Volterra series is a general nonlinear model with memory [75], its application to practical systems is limited due to the drastic increase in computational complexity when higher order nonlinearities are included. Recently in [65], [39], it has been shown that the so-called memory polynomial model is a good framework for studying nonlinear PAs with memory effects; it is also a good model for predistorters. When only odd-order

nonlinear terms are considered, the PA output is related to the input as follows:

$$y(t) = \sum_{k=0}^{K} \int h_{2k+1}(\tau)\ |x(t-\tau)|^{2k} x(t-\tau)\ d\tau \tag{9.21}$$

$$= \sum_{k=0}^{K} \int h_{2k+1}(\tau)\ [x(t-\tau)]^{k+1} [x^*(t-\tau)]^k\ d\tau \tag{9.22}$$

$$= \sum_{k=0}^{K} \underbrace{h_{2k+1}(t) \star \phi_{2k+1}(x(t))}_{y_{2k+1}(t)} \tag{9.23}$$

where $\phi_{2k+1}(x(t)) = [x(t)]^{k+1}[x^*(t)]^k$.

To the best of our knowledge, there has been no published results on spectral regrowth analysis for nonlinear PAs with memory.

9.4.1 Closed form expression

We present here, a simple closed-form expression for the output PSD of the memory polynomial model (9.21).

Theorem 3. *Assume that $x(t)$ is stationary, zero-mean, complex Gaussian distributed and satisfies (9.12). If the output $y(t)$ is related to the input $x(t)$ through (9.21), then the PSD of $y(t)$ is related to that of $x(t)$ through*

$$S_{2y}(f) = \sum_{m=0}^{K} \alpha_{2m+1}(f)\ \underbrace{S_{2x}(f) \star \cdots \star S_{2x}(f)}_{m+1}\ \star\ \underbrace{S_{2x}(-f) \star \cdots \star S_{2x}(-f)}_{m}, \tag{9.24}$$

where

$$\alpha_{2m+1}(f) = \frac{1}{m+1} \left| \sum_{k=m}^{K} H_{2k+1}(f)\ \binom{k}{m}\ (k+1)!\ [c_{2x}(0)]^{k-m} \right|^2. \tag{9.25}$$

and

$$H_{2k+1}(f) = \int h_{2k+1}(t)\ e^{-j2\pi ft}\ dt, \tag{9.26}$$

is the FT of the $(2k+1)$st-order kernel $h_{2k+1}(t)$.

Proof: See Section 9.6.2.

Remarks:

(R4) The (quasi-) memoryless model (9.10) can be regarded as a special case of the memory polynomial model (9.21) with $h_{2k+1}(t) = a_{2k+1}\ \delta(t)$. Therefore, Theorem 1 can be regarded as a special case of Theorem 2 with $H_{2k+1}(f) = a_{2k+1}$.

(R5) Since the baseband kernel $h_{2k+1}(t)$ is generally complex valued, its FT is not guaranteed to be conjugate symmetric. Therefore, even if $S_{2x}(f)$ is symmetric, $S_{2y}(f)$ may not be symmetric.

9.4.2 Case study: Asymmetric spectral regrowth and memory effects

It is commonly known that asymmetry in the PSD of $y(t)$ is indicative of memory effects in the PA (e.g., [35]). Since the memory polynomial model has been shown to be a good model for nonlinear PAs with memory, next, we shall carry out quantitative analysis on spectral asymmetry of a PA with memory, by applying Theorem 2. We use the adjacent channel power ratio (ACPR) defined as [50]

$$\text{ACPR} = \frac{\int_{f_3}^{f_4} S_{2y}(f)\ df}{\int_{f_1}^{f_2} S_{2y}(f)\ df}, \tag{9.27}$$

as the performance metric, where f_1 and f_2 are the frequency limits of the main channel, and f_3 and f_4 are the frequency limits of the adjacent channel. The two bandwidths $(f_2 - f_1)$ and $(f_4 - f_3)$ need not be the same and indeed are not for many current standards [64, p. 39]. For $\text{ACPR}_{\text{LOWER}}$, we use f_3, f_4 as limits for the lower adjacent channel. Similarly, for $\text{ACPR}_{\text{UPPER}}$, we use f_3, f_4 as limits for the upper adjacent channel.

Example 2. In Table 9.2, we show the memory polynomial kernel coefficients extracted from a PA which is known to exhibit memory effects. The sampling rate was $f_s = 150$MHz. To calculate the ACPR, we used $[-0.15, 0.15]$ as the normalized frequency limits for the main channel, $[-0.45, -0.15]$ as the normalized frequency limits for the lower adjacent channel, and $[0.15, 0.45]$ as the normalized frequency limits for the upper adjacent channel. In Fig. 9.5, we plot $\text{ACPR}_{\text{LOWER}}$ as the solid line, and $\text{ACPR}_{\text{UPPER}}$ as the dashed-dotted line, as a function of the input signal power $\sigma_x^2 = c_{2x}(0)$. The two curves do not coincide, implying spectral asymmetry in $S_{2y}(f)$. At low input power levels, the ACPR curves are approximately constant – this is because the PA is approximately linear when it is largely

	$q=0$	$q=1$	$q=2$
$h_1[q]$	$1.1330+j0.0696$	$-0.2027+j0.0338$	$0.0854-j0.0341$
$h_3[q]$	$-0.2348-j0.0876$	$0.1809+j0.2447$	$-0.0439-j0.0640$
$h_5[q]$	$0.2675-j0.4113$	$-0.1376-j0.1862$	$0.0888+j0.0197$
$h_7[q]$	$-0.2686+j0.2694$	$0.0273+j0.0504$	$-0.0457+j0.0093$

Table 9.2: Memory polynomial PA coefficients extracted for a real PA with maximum nonlinearity order $2K+1=7$ and maximum lag $Q=2$.

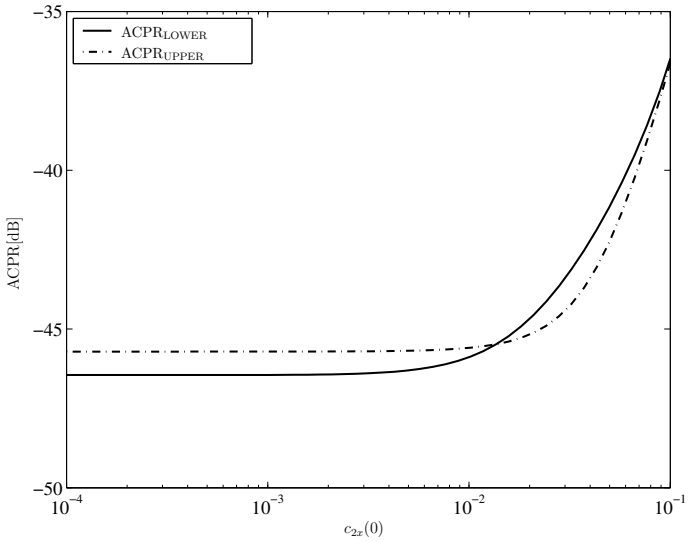


Figure 9.5: $\text{ACPR}_{\text{LOWER}}$ (solid line) and $\text{ACPR}_{\text{UPPER}}$ (dash-dotted line) as a function of the input power $c_{2x}(0)$, for a PA with memory.

backed-off, and spectral regrowth was almost absent. As the PA is driven into compression, adjacent channel power increases sharply. Plots similar to Fig. 9.5 can be used to select the input power level to ensure that spectral emission requirements are met.

9.5 Conclusions

The focus of this chapter was on polynomial type of PA nonlinearities and Gaussian inputs. The objective was to obtain analytical expressions for the PA output power spectral density. We employed the little known Leonov-Shiryaev formula (see Proofs section) to obtain closed form output PSD expressions that apply to an arbitrary-order nonlinearity, and showed that they embody as special cases, previously reported results for memoryless nonlinear PAs of specific orders. Our spectral regrowth analysis on the PA model with memory is the first of its kind. These results can help us make important practical decisions such as what factors

contribute to spectral regrowth, and how to control or correct them in order to keep the adjacent channel interference to within limits.

9.6 Proofs of Theorems

9.6.1 Proof of Theorem 1

Define $\phi_{2k+1}(x(t)) = [x(t)]^{k+1}[x^*(t)]^k$. We can rewrite (9.10) as

$$y(t) = \sum_{k=0}^{K} a_{2k+1}\ \phi_{2k+1}(x(t)). \tag{9.28}$$

Since $x(t)$ is assumed to be zero-mean, Gaussian distributed, only the second-order statistics of $x(t)$ are non-zero. Moreover, all odd-order moments of $x(t)$ are zero [101]. Therefore, $E[\phi_{2k+1}(x(t))] = 0$, and $E[y(t)] = 0$.

The auto-correlation (auto-covariance) function of $y(t)$ is

$$c_{2y}(\tau) = \text{cum}\{y^*(t), y(t+\tau)\} \tag{9.29}$$

$$= \sum_{k=0}^{K}\sum_{l=0}^{K} a^*_{2k+1}\ a_{2l+1}\ \text{cum}\{\phi^*_{2k+1}(x(t)), \phi_{2l+1}(x(t+\tau))\}. \tag{9.30}$$

First, we would like to express $\text{cum}\{\phi^*_{2k+1}(x(t)), \phi_{2l+1}(x(t+\tau))\}$ in terms of $c_{2x}(\tau)$.

Since $\phi_{2k+1}(x(t))$ is zero-mean,

$$\text{cum}\{\phi^*_{2k+1}(x(t)), \phi_{2l+1}(x(t+\tau))\} = E\{[x^*(t)]^{k+1}[x(t)]^k[x(t+\tau)]^{l+1}[x^*(t+\tau)]^l\}. \tag{9.31}$$

It is possible to use the moment theorem for complex Gaussian processes [101] to simplify (9.31), but as the authors of [50] found out, it "requires overwhelmingly complex manual expansion of the moment expressions." We adopt another approach here, which employs the so-called Leonov-Shiryaev formula [25, p. 89].

To utilize the Leonov-Shiryaev formula, we start with a two-way table. We list the individual elements that form the product $\phi^*_{2k+1}(x(t)) = [x^*(t)]^{k+1}x^k(t)$ in the first row, and display the individual elements that form the product $\phi_{2l+1}(x(t+\tau)) = [x(t+\tau)]^{l+1}[x^*(t+$

$\tau)]^l$ in the second row.

$$\begin{array}{cc} \underbrace{x^*(t)\ \cdots\ x^*(t)}_{k+1} & \underbrace{x(t)\ \cdots\ x(t)}_{k} \\ \underbrace{x(t+\tau)\ \cdots\ x(t+\tau)}_{l+1} & \underbrace{x^*(t+\tau)\ \cdots\ x^*(t+\tau)}_{l} \end{array} \tag{9.32}$$

Next, we partition the above $(2k+2l+2)$ elements into subsets, according to the following criteria:

(i) The joint cumulant of the elements in any subset is non-zero.

(ii) For each partition, there must be at least one subset that contains elements from both rows of (9.32). We shall refer to such subset as a "hooking" subset.

When both conditions (i) and (ii) are satisfied, the corresponding partition is called a "valid" partition. We must find all valid partitions of the two-way table in order to simplify (9.31).

Since $x(t)$ is zero-mean, Gaussian, and satisfies (9.12), the only non-zero cumulants of $x(t)$ are

$$c_{2x}(\tau) = \text{cum}\{x^*(t), x(t+\tau)\}$$

and its variants

$$c_{2x}(0) = \text{cum}\{x^*(t), x(t)\},$$

$$c^*_{2x}(\tau) = \text{cum}\{x(t), x^*(t+\tau)\}.$$

Therefore, to meet requirement (i), we only need to consider two element subsets, and the two elements within the subset must have different conjugation.

To illustrate the above concept, let us consider the following two-way table which would be needed if we are interested in evaluating $\text{cum}\{\phi^*_5(x(t)), \phi_3(x(t+\tau))\}$:

$$\begin{array}{ccccc} x^*(t) & x^*(t) & x^*(t) & x(t) & x(t) \\ x(t+\tau) & x(t+\tau) & x^*(t+\tau) & & \end{array}$$

One valid partition of the above 8 elements is:

$$\{x^*(t), x(t+\tau)\},\ \{x^*(t), x(t)\},\ \{x^*(t), x(t)\},\ \{x(t+\tau), x^*(t+\tau)\},$$

and there are 12 such possibilities (consider each element unique). In this partition, there is only one hooking subset $\{x^*(t), x(t+\tau)\}$.

Another valid partition is:

$$\{x^*(t), x(t+\tau)\},\ \{x^*(t), x(t+\tau)\},\ \{x(t), x^*(t+\tau)\},\ \{x^*(t), x(t)\},$$

and the multiplicity also happens to be 12. In this partition, the first three subsets are hooking subsets.

These are the only valid partitions for the above 8 element example.

Once we have found all valid partitions, we take the cumulant of the elements in each subset, multiply the resulting cumulants from all subsets of a given partition, and then sum over all valid partitions. For the above 8 element example, we have

$$\begin{aligned}
&\mathrm{cum}\{\phi_5^*(x(t)), \phi_3(x(t+\tau))\} \\
&= 12 c_{2x}(\tau) c_{2x}(0) c_{2x}(0) c_{2x}(0) \\
&\quad + 12 c_{2x}(\tau) c_{2x}(\tau) c_{2x}^*(\tau) c_{2x}(0) \\
&= 12 c_{2x}(\tau) c_{2x}^3(0) + 12 |c_{2x}(\tau)|^2 c_{2x}(\tau) c_{2x}(0).
\end{aligned}$$

Now for the general two-way table in (9.32), we realize the following. For each partition to be valid, there need to be $(2m+1)$ hooking subsets: $(m+1)$ subsets are of the form $\{x^*(t), x(t+\tau)\}$, m subsets are of the form $\{x(t), x^*(t+\tau)\}$, and $0 \le m \le \min(k, l)$. To come up with these $(2m+1)$ hooking subsets, there are

$$\frac{(k+1)k\cdots(k+1-m)(l+1)l\cdots(l+1-m)}{(m+1)!}$$

$$\times \frac{k(k-1)\cdots(k-m+1)l(l-1)\cdots(l-m+1)}{m!} \tag{9.33}$$

different possibilities.

Apart from the $(2m+1)$ hooking subsets, the remaining elements must be grouped into $(k-m)$ subsets of the form $\{x^*(t), x(t)\}$, and $(l-m)$ subsets of the form $\{x(t+\tau), x^*(t+\tau)\}$. The multiplicity number for this stage is

$$(k-m)!(l-m)! \tag{9.34}$$

Multiplying (9.33) and (9.34), we find that the multiplicity number for a partition that involves exactly $(m+1)$ subsets of $\{x^*(t), x(t+\tau)\}$, m subsets of $\{x(t), x^*(t+\tau)\}$, $(k-m)$ subsets of $\{x^*(t), x(t)\}$, and $(l-m)$ subsets of $\{x(t+\tau), x^*(t+\tau)\}$, is

$$\frac{1}{m+1}\binom{k}{m}\binom{l}{m}(k+1)!(l+1)! \tag{9.35}$$

Now take the cumulant of each subset and multiply the resulting cumulants. We infer that the contribution from any one partition described above to (9.31) is

$$[c_{2x}(\tau)]^{m+1}[c_{2x}^*(\tau)]^m[c_{2x}(0)]^{k-m}[c_{2x}(0)]^{l-m}.$$

Summing over all valid partitions, we obtain,

$$\begin{aligned}&\text{cum}\{\phi_{2k+1}^*(x(t)), \phi_{2l+1}(x(t+\tau))\}\\ &= \sum_{m=0}^{\min(k,l)} \frac{1}{m+1}\binom{k}{m}\binom{l}{m}(k+1)!(l+1)!|c_{2x}(\tau)|^{2m}c_{2x}(\tau)[c_{2x}(0)]^{k+l-2m}.\end{aligned} \tag{9.36}$$

Substituting (9.36) into (9.30), we obtain,

$$\begin{aligned}c_{2y}(\tau) = \sum_{k=0}^{K}\sum_{l=0}^{K} a_{2k+1}^*\ a_{2l+1} \sum_{m=0}^{\min(k,l)} \frac{1}{m+1}\binom{k}{m}\binom{l}{m}\\ \times(k+1)!(l+1)!|c_{2x}(\tau)|^{2m}c_{2x}(\tau)[c_{2x}(0)]^{k+l-2m}.\end{aligned} \tag{9.37}$$

The above equation can be simplified once we realize the following:

- $\sum_{k=0}^{K}\sum_{l=0}^{K}\sum_{m=0}^{\min(k,l)}$ is equivalent to $\sum_{m=0}^{K}\sum_{k=m}^{K}\sum_{l=m}^{K}$.
- Since $c_{2x}(0) = E[|x(t)|^2]$ is real-valued,

$$\sum_{k=m}^{K} a_{2k+1}^*\binom{k}{m}(k+1)!\ [c_{2x}(0)]^{k-m} = \left[\sum_{l=m}^{K} a_{2l+1}\binom{l}{m}(l+1)!\ [c_{2x}(0)]^{l-m}\right]^*.$$

Therefore,

$$c_{2y}(\tau) = \sum_{m=0}^{K} \alpha_{2m+1}\ |c_{2x}(\tau)|^{2m}\ c_{2x}(\tau), \tag{9.38}$$

where

$$\alpha_{2m+1} = \frac{1}{m+1}\left|\sum_{k=m}^{K} a_{2k+1}\binom{k}{m}(k+1)!\ [c_{2x}(0)]^{k-m}\right|^2. \tag{9.39}$$

Since the FT of $c_{2x}(\tau)$ is $S_{2x}(f)$, the FT of $c_{2x}^*(\tau)$ is $S_{2x}(-f)$. Thus, the input-output PSD relationship is given by

$$S_{2y}(f) = \sum_{m=0}^{K} \alpha_{2m+1}\ \underbrace{S_{2x}(f)\star\cdots\star S_{2x}(f)}_{m+1}\ \star\ \underbrace{S_{2x}(-f)\star\cdots\star S_{2x}(-f)}_{m}. \tag{9.40}$$

9.6.2 Proof of Theorem 2

Define

$$f_{kl}(\tau) = \int h_k^*(t)\ h_l(t+\tau)\ dt \tag{9.41}$$

as the (deterministic) cross-correlation function between the kernels $h_k(t)$ and $h_l(t)$.

Define

$$g_{kl}(\tau) = \text{cum}\{\phi_k^*(x(t)), \phi_l(x(t+\tau))\} \tag{9.42}$$

as the (statistical) cross-correlation function between $\phi_k(x(t))$ and $\phi_l(x(t))$. The expression for $g_{(2k+1)(2l+1)}(\tau)$ was found previously as (9.36).

From the linear systems theory, it is well known that if $y_k(t) = h_k(t) \star u_k(t)$, $y_l(t) = h_l(t) \star u_l(t)$, then $\text{cum}\{y_k^*(t), y_l(t+\tau)\} = f_{kl}(\tau) \star \text{cum}\{u_k^*(t), u_l(t+\tau)\}$, where $f_{kl}(\tau)$ is given in (9.41).

Since in the memory polynomial model (9.23), $y_{2k+1}(t) = h_{2k+1}(t) \star \phi_{2k+1}(x(t))$, we use our linear systems knowledge to infer

$$c_{2y}(\tau) = \sum_{k=0}^{K} \sum_{l=0}^{K} f_{(2k+1)(2l+1)}(\tau) \star g_{(2k+1)(2l+1)}(\tau). \tag{9.43}$$

Recall that the FT of $f_{kl}(\tau)$ is $H_k^*(f)H_l(f)$. Thus, the FT (9.43) yields

$$S_{2y}(f) = \sum_{k=0}^{K} \sum_{l=0}^{K} H_{2k+1}^*(f)\ H_{2l+1}(f)\ G_{(2k+1)(2l+1)}(f), \tag{9.44}$$

where $G_{(2k+1)(2l+1)}(f)$ is the FT of $g_{(2k+1)(2l+1)}(\tau)$ given by (9.36).

Following the similar procedure as in Section 9.6.1, we can simplify $S_{2y}(f)$ to (9.24)-(9.25).

Acknowledgment: The authors would like to thank Ning Chen for many insightful discussions on this chapter. Appreciation also goes to Dr. J. S. Kenney for providing the PA measurements used in Fig. 9.2.

CHAPTER X

OPTIMIZATION OF SNDR FOR AMPLITUDE LIMITED NONLINEARITIES[1]

Many communications components are nonlinear and have a peak power or peak amplitude constraint. Nonlinearity generates distortions and thus an appropriate performance measure is the signal-to-noise-and-distortion ratio (SNDR). In this chapter, we are interested in finding the nonlinear mapping that maximizes the SNDR subject to the peak amplitude constraint. The answer is a soft limiter with gain calculated based on the noise power and the probability density function of the input amplitude. We also investigate a bounding relationship between the SNDR and capacity of the nonlinear channel. The results of this chapter can be applied for efficient transmission of high peak-to-average power ratio signals such as OFDM or for optimal linearization of nonlinear devices.

10.1 Introduction

Many components in a communication system have a peak power (or peak amplitude) constraint. For example, power amplifiers (PAs) are peak power limited in addition to being nonlinear. Denote by x a zero-mean complex baseband signal with variance σ_x^2 and by v a zero-mean additive noise process with variance σ_v^2. Let us consider the received signal modeled by

$$y = h(x) + v, \tag{10.1}$$

where $h(\cdot)$ is a memoryless nonlinear mapping with peak amplitude constraint

$$|h(x)| \leq A. \tag{10.2}$$

Model (10.1) is of interest, for example, in transmission systems involving nonlinear components such as PAs or mixers [18, 64, 72], for nonlinear magnetic recording channels

[1]This chapter was published in [89,92,94] and is a result of joint work with Hua Qian and G. Tong Zhou.

[135], or when companding [59, 120] or clipping [83, 84, 103, 108, 115] is involved for the purpose of peak-to-average power ratio (PAPR) reduction.

Two questions can be asked: (i) What undesirable effects are caused by the nonlinearity? (ii) How can we best choose or modify (if possible) the nonlinearity $h(\cdot)$ so that the undesirable nonlinear effects are minimal?

There has been a lot of research devoted to the first question. Nonlinearity causes increase in symbol-error-rate (SER) [64, 108], spectral regrowth [108, 129], and reduction in channel capacity [107, 109, 115]. However, to the best of our knowledge, optimization of nonlinearity under the peak amplitude constraint has not been studied extensively. We use the signal-to-noise-and-distortion ratio (SNDR) as the optimization criterion in this chapter. We will show that the ideal linearizer (which is overall nonlinear) with gain properly selected according to the distribution of $|x|$ and the channel noise power, maximizes the SNDR. We will also point out a bounding relationship between the SNDR and channel capacity, further motivating the SNDR consideration.

The organization of this chapter is as follows. In Section 10.2, we derive the SNDR expression and find the optimal nonlinear mapping according to the SNDR criterion. In Section 10.3, we relate the SNDR to the capacity of the nonlinear channel in (10.1). Conclusions are drawn in Section 10.4. The rather technical proofs of this chapter are deferred to the appendices.

10.2 The SNDR Criterion and Its Optimization

The SNDR criterion has been used in [83, 84, 103]. We start by reviewing its definition and then present our novel results on its optimization.

10.2.1 SNDR Definition

The nonlinear mapping in (10.1) can be decomposed as

$$h(x) = \alpha x + d, \tag{10.3}$$

where d is the distortion created by $h(\cdot)$ and α is chosen such that d is uncorrelated with x; i.e., $E[x^* d] = 0$.

From (10.3), we obtain

$$E[x^*h(x)] = \alpha E[|x|^2] + E[x^*d] = \alpha E[|x|^2]. \tag{10.4}$$

Thus,

$$\alpha = \frac{E[x^*h(x)]}{E[|x|^2]} = \frac{E[x^*h(x)]}{\sigma_x^2}. \tag{10.5}$$

The distortion power is given by

$$\varepsilon_d = E[|d|^2] = E[|h(x)|^2] - |\alpha|^2\sigma_x^2. \tag{10.6}$$

The SNDR is defined as

$$\text{SNDR} = \frac{|\alpha|^2\sigma_x^2}{\varepsilon_d + \sigma_v^2} = \frac{|E[x^*h(x)]|^2/\sigma_x^2}{E[|h(x)|^2] - |E[x^*h(x)]|^2/\sigma_x^2 + \sigma_v^2}. \tag{10.7}$$

We see from (10.7) that the SNDR depends on the distribution of x, the nonlinear mapping $h(\cdot)$, and the noise power σ_v^2.

10.2.2 Optimization of the SNDR

First, let us examine the angle of $h(x)$, $\angle h(x)$, which enables maximization of the SNDR. Let us write $h(x) = |h(x)|e^{j\angle h(x)}$, and

$$|E[x^*h(x)]| = |E[|x||h(x)|e^{j(\angle h(x) - \angle x)}]|. \tag{10.8}$$

Since the right hand side (RHS) of (10.8) is $\leq E[|x||h(x)|]$ with equality holds if and only if

$$\angle h(x) - \angle x = \text{constant}, \tag{10.9}$$

we infer that as far as $\angle h(x)$ is concerned, the numerator of the SNDR expression in (10.7) is maximized, the denominator of the SNDR in (10.7) is minimized, and thus the SNDR is maximized if we have (10.9). Condition (10.9) implies that the so-called AM-PM conversion is absent. Without loss of generality, we can set the constant in (10.9) to be zero, and work with $\angle h(x) = \angle x$.

A type of nonlinearity that is of particular interest in communication systems exhibits the so-called amplitude-to-amplitude (AM-AM) conversion, meaning that $|h(x)|$ only depends on $|x|$. A system with AM-AM, but no AM-PM conversion is called a strictly memoryless system [18, Chap. 2]. Therefore, we consider $h(\cdot)$ of the form

$$h(x) = Ag(\frac{|x|}{\sigma_x})e^{j\angle x}, \tag{10.10}$$

where $0 \leq g(\cdot) \leq 1$. This ensures that $|h(x)| \leq A$ and $\angle h(x) = \angle x$. The above standardization in both the argument and the amplitude limit of $g(\cdot)$ will make the notations less cumbersome.

Let $\gamma = |x|/\sigma_x$ and substitute (10.10) into (10.7) to obtain

$$\text{SNDR} = \frac{|E[\gamma g(\gamma)]|^2}{E[g^2(\gamma)] - |E[\gamma g(\gamma)]|^2 + \frac{\sigma_n^2}{A^2}}. \tag{10.11}$$

We see from (10.11) that because of (10.9), the distribution of the phase of x does not affect the SNDR; the distribution of $|x|$, or equivalently, the probability density function (PDF) of γ does.

The SNDR optimization problem can be stated as follows: For a given distribution of $\gamma = |x|/\sigma_x$, find the nonlinear mapping $g(\cdot)$ with $0 \leq g(\cdot) \leq 1$, such that the SNDR expression in (10.11) is maximum.

Fig. 10.1 (a)-(f) show some example $g(\cdot)$ functions that satisfy the constraint $0 \leq g(\cdot) \leq 1$. It is expected that they will exhibit different SNDR behavior. The results to be presented next will shed light on which $g(\cdot)$ is the most desirable.

Lemma 3. *Among all $g(\cdot)$ functions satisfying $0 \leq g(\cdot) \leq 1$, the $g(\cdot)$ that maximizes the SNDR expression in (10.11) must be of the form*

$$g(\gamma) = \begin{cases} \frac{\gamma}{\eta}, & \gamma \in S, \\ 1, & \gamma \notin S, \end{cases} \tag{10.12}$$

where S is a subset of $[0, \infty)$, and η is determined from

$$\eta = \frac{C_1}{C_o + \frac{\sigma_n^2}{A^2}}, \tag{10.13}$$

with $C_o = E[I(\gamma \notin S)]$, $C_1 = E[\gamma I(\gamma \notin S)]$, and $I(\cdot)$ is the indicator function.

Proof. See Appendix I. □

This result rules out the $g(\cdot)$'s such as those shown in Fig. 10.1 (a), (d), (e) and (f) as candidates for the optimal nonlinear mapping. Functions depicted in Fig. 10.1 (b) and (c) meet the requirements in Lemma 1. Our next result further elucidates on the optimal solution for $g(\cdot)$.

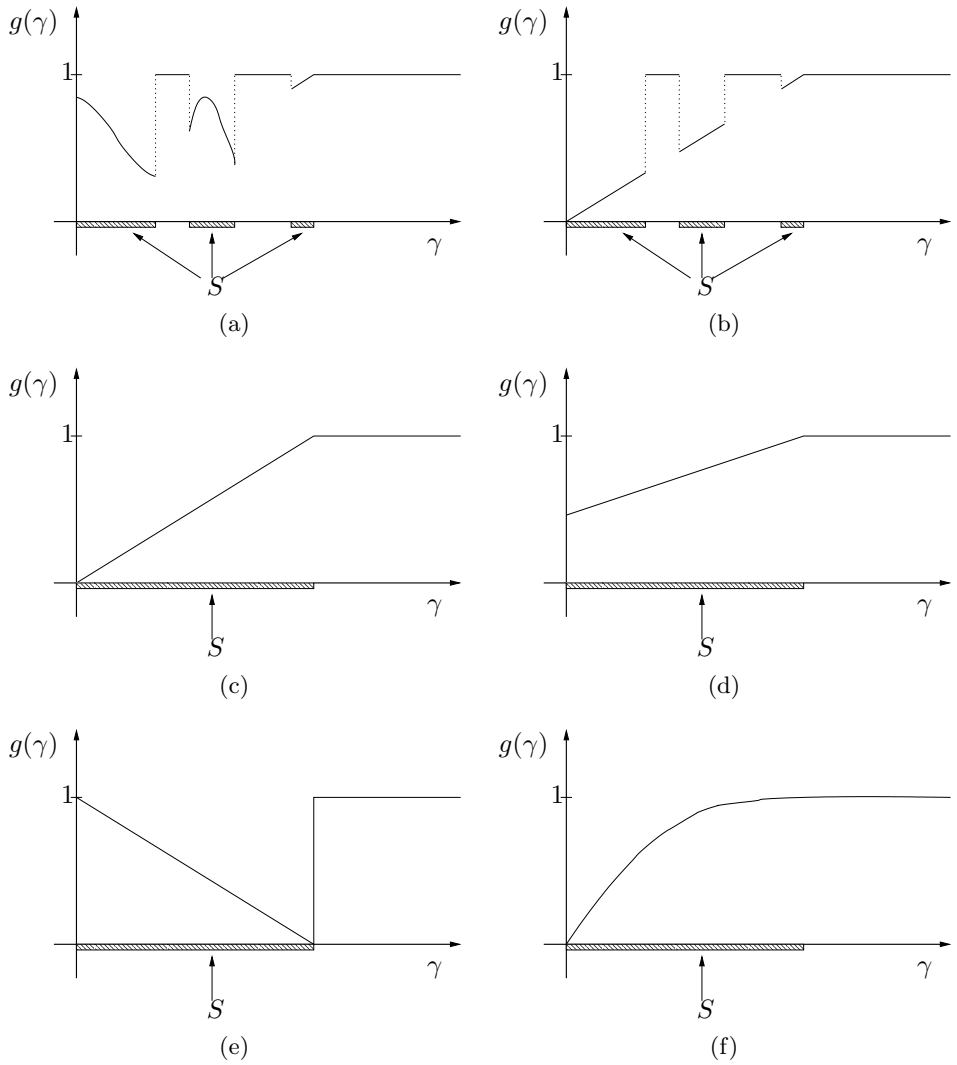


Figure 10.1: Nonlinear mappings $g(\cdot)$ that satisfy the $0 \leq g(\cdot) \leq 1$ constraint.

Theorem 4. *Within the class of $g(\cdot)$ satisfying $0 \leq g(\cdot) \leq 1$, the following $g(\cdot)$ maximizes the SNDR expression in (10.11):*

$$g(\gamma) = \begin{cases} \frac{\gamma}{\eta^\star}, & 0 \leq \gamma < \eta^\star, \\ 1, & \gamma \geq \eta^\star, \end{cases} \tag{10.14}$$

where the threshold $\eta^\star$ is found from $\eta^\star = T^{-1}(\frac{A^2}{\sigma_v^2})$, with

$$T(\eta) = \frac{\eta}{C_1(\eta) - \eta C_o(\eta)}, \tag{10.15}$$

$$C_o(\eta) = \int_\eta^\infty p(\gamma) d\gamma, \tag{10.16}$$

$$C_1(\eta) = \int_{\eta}^{\infty} \gamma p(\gamma) d\gamma, \tag{10.17}$$

and $p(\gamma)$ *is the PDF of* γ*. The optimal SNDR is found as*

$$\text{SNDR}^{\star} = \frac{1}{\frac{1}{R(\eta^{\star})} - 1}, \tag{10.18}$$

where

$$R(\eta^{\star}) = \frac{C_1^2(\eta^{\star})}{C_o(\eta^{\star}) + \frac{\sigma_\eta^2}{A^2}} + \bar{C}_2(\eta^{\star}), \tag{10.19}$$

and

$$\bar{C}_2(\eta) = \int_{0}^{\eta} \gamma^2 p(\gamma) d\gamma. \tag{10.20}$$

Proof. See Appendix II. □

From (10.16) and (10.17), it is straightforward to show that the derivative of $[C_1(\eta) - \eta C_o(\eta)]$ is $-C_o(\eta)$, which is negative. This means that the denominator of $T(\eta)$ in (10.15) is a monotonically decreasing function of η. Therefore, $T(\eta)$ is a monotonically increasing function of η and its inverse $T^{-1}(\cdot)$ exists.

Theorem 1 establishes that the nonlinearity in the shape of Fig. 10.1 (c) is optimal. Intuitively, the zig-zagged and non-smooth functions of Fig. 10.1 (a), (b), (e) must be generating nonlinear distortions that lower the SNDR. Our examples in Section 10.3.3 will illustrate the subtle difference in SNDR among the $g(\cdot)$ functions shown in Fig. 10.1 (c), (d), (f).

Clipping is a popular approach to reduce the PAPR of certain signals such as OFDM [84, 103, 108, 115]. A soft clipper (limiter) is given by

$$h(x) = \begin{cases} x, & |x| < A, \\ Ae^{j\angle x}, & |x| \geq A. \end{cases} \tag{10.21}$$

The lower the threshold A, the smaller the PAPR, but the larger the clipping probability. If we are to choose the A value so as to maximize the SNDR, then we should set $A^{\star} = \sigma_x \, \eta^{\star}$, where $\eta^{\star}$ is given by Theorem 1.

If a given system nonlinearity $u(\cdot)$ is undesirable and it is possible to apply a predistortion mapping $f(\cdot)$, then according to Theorem 1, it is best to make $u(f(\cdot))$ equal to the $g(\cdot)$

function given in (10.14) (assume that $u(\cdot)$ is normalized to have a maximum amplitude of 1). This is the well-known linearization strategy [64]. However, what was little understood before, was the selection of the threshold η, or equivalently, the gain factor $1/\eta$. Theorem 1 says that the optimal (in terms of SNDR) $\eta^\star$ depends on the PDF of γ and the peak signal-to-noise ratio PSNR $= A^2/\sigma_v^2$. Examples are given in the next subsection to illustrate the calculation of $\eta^\star$ and its utility.

10.2.3 Examples – Optimal Threshold Selection and Application

Example 1. When x is zero-mean complex Gaussian distributed, $\gamma = \frac{|x|}{\sigma_x}$ is Rayleigh distributed with PDF

$$p(\gamma) = 2\gamma e^{-\gamma^2}, \qquad \gamma \geq 0. \tag{10.22}$$

Note that (10.22) ensures that $E[\gamma^2] = 1$. Substituting (10.22) into (10.16) and (10.17), we obtain

$$C_o(\eta) = e^{-\eta^2}, \tag{10.23}$$

$$C_1(\eta) = \eta e^{-\eta^2} + \sqrt{\pi} Q(\sqrt{2}\eta), \tag{10.24}$$

where

$$Q(\eta) = \int_\eta^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}\, dt. \tag{10.25}$$

It follows then that

$$T(\eta) = \frac{\eta}{\sqrt{\pi} Q(\sqrt{2}\eta)}. \tag{10.26}$$

The optimal $\eta^\star = T^{-1}(A^2/\sigma_v^2)$ does not have a closed-form expression but can be easily calculated numerically.

Example 2. When $|x|$ is uniformly distributed, $\gamma = \frac{|x|}{\sigma_x}$ is uniformly distributed as well. Let γ be uniformly distributed between 0 and $\sqrt{3}$ so that $E[\gamma^2] = 1$. It follows that

$$C_o(\eta) = 1 - \frac{\eta}{\sqrt{3}}, \quad 0 \leq \eta \leq \sqrt{3}, \tag{10.27}$$

$$C_1(\eta) = \frac{3 - \eta^2}{2\sqrt{3}}, \quad 0 \leq \eta \leq \sqrt{3}, \tag{10.28}$$

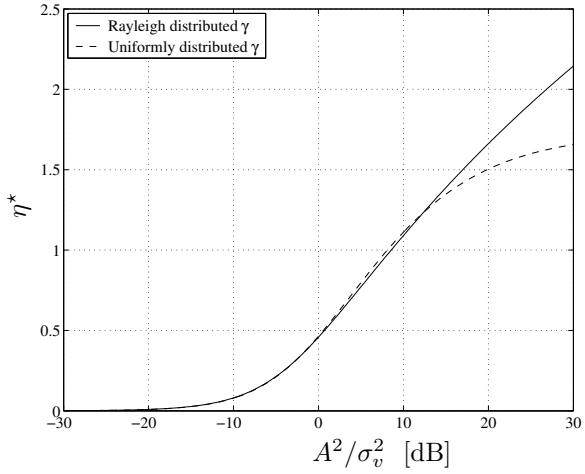


Figure 10.2: Optimal threshold $\eta^\star$ as a function of PSNR $= A^2/\sigma_v^2$ for two distributions of γ: Rayleigh (solid line) and uniform (dashed line).

and hence

$$T(\eta) = \frac{\eta}{\frac{\eta^2}{2\sqrt{3}-\eta} + \frac{\sqrt{3}}{2}}. \tag{10.29}$$

Setting the above $T(\eta)$ equal to A^2/σ_v^2, we obtain a closed-form solution for the optimal $\eta^\star$:

$$\eta^\star = \sqrt{3}\left(\left(1+\frac{\sigma_n^2}{A^2}\right) - \sqrt{\frac{\sigma_n^2}{A^2}\left(\frac{\sigma_n^2}{A^2}+2\right)}\right), \tag{10.30}$$

which is seen to depend only on the PSNR $= A^2/\sigma_v^2$.

Fig. 10.2 shows the optimal $\eta^\star$ as a function of PSNR for the above two examples. An interesting observation is made from Fig. 10.2: At high noise levels (i.e., low A^2/σ_v^2), strong clipping is suggested. This is because with a low $\eta^\star$, the linear gain $1/\eta^\star$ is larger, giving more energy $|\alpha|^2\sigma_x^2$ to the signal to counter-act the noise.

We point out that when it comes to linearization, there are other considerations in the selection of η, such as spectral spreading [108, 129] and SER.

Example 3. Application of optimal clipping with gain for OFDM transmission.

OFDM is well known for its robustness against frequency selective fading channels and for its high spectral efficiency. It has shown tremendous potential for high speed digital communication systems. It has been accepted as standards in many applications such as digital subscriber line [3] and digital audio/video broadcasting [4].

Denote by $\{X(k)\}_{k=0}^{N-1}$ the frequency domain OFDM signal drawn from a known constellation $\mathcal{C}$, and by N the number of sub-carriers. Nyquist-rate sampled time domain OFDM signal is given by

$$x(n) = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} X(k)\ e^{j\frac{2\pi kn}{N}}, \quad 0 \le n \le N-1. \tag{10.31}$$

It is well-known that $|x(n)|$ exhibits high peaks, especially for N large [114].

Theoretically, the worst case peak power of $x(n)$ is $N\sigma_x^2$, but that happens with near zero probability. Let us consider as an example, a QPSK sequence $\{X(k)\}_{k=0}^{N-1}$ with $N = 512$ and $X(k) \in \mathcal{C} = \{\sigma_x e^{\pm j\frac{\pi}{4}}\}$. Let us choose $A = 2.6283$ so that $Pr\{|x(n)|/\sigma_x > A\} = 0.001$ and call this the "unclipped" case, i.e., treat $h(x) \approx x$. In the unclipped case (assuming AWGN channel), the received signal is $y(n) = x(n) + v(n)$. We then follow standard procedures to estimate $X(k)$ from $y(n)$.

Next, we describe an alternative method of OFDM transmission called "optimal clipping with gain". Instead of transmitting x, we transmit an optimally designed $h(x)$. The received signal (assuming AWGN channel) is $y(n) = h(x(n)) + v(n) = \alpha x(n) + d(n) + v(n)$. The receiver decodes $X(k)$ as usual (treats $d(n) + v(n)$ as noise) so there is no additional complexity at the receiver. Modification at the transmitter is described below.

Step 1. Calculate $\eta^\star$ for given A and σ_v^2 values and distribution of $|x|$ according to equations (10.15)-(10.17), and then obtain the optimum $g(\cdot)$ in (10.14).

Step 2. Map x to $h(x)$ using (10.10). Transmit $h(x)$.

Fig. 10.3 illustrates the PDF of $|x|$ and that of $|h(x)|$. Although non-negligible clipping is present in $h(x)$, the idea is to have increased signal power in $h(x)$ which dominates the increase in the distortion power. As we show in Fig. 10.4 (some results are highlighted in Table 10.1), this simple operation at the transmitter allows significant SER improvement.

10.3 Relationship between SNDR and Capacity

Capacity for memoryless nonlinear channels has been studied in [107, 109, 115]. In this section, we are not interested in finding the input distribution that achieves the capacity.

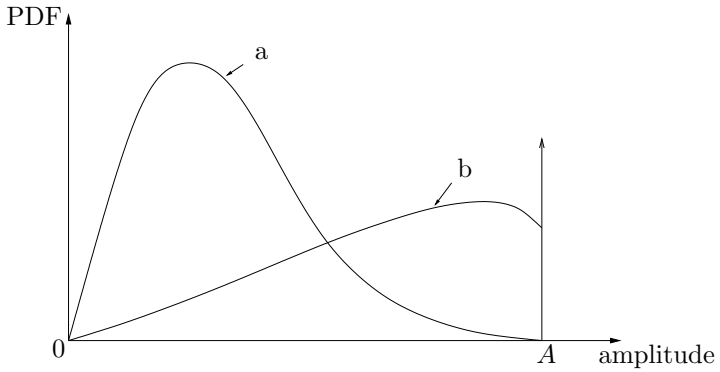


Figure 10.3: PDF of $|x|$ (line a) and PDF of $|h(x)|$ (line b).

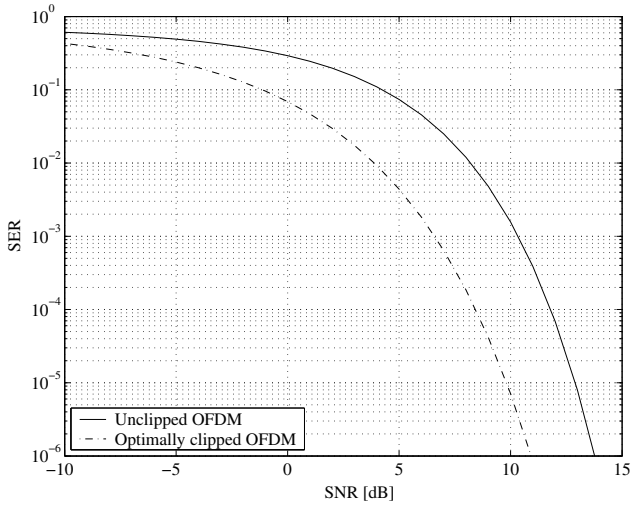


Figure 10.4: SER vs. SNR $= \sigma_x^2/\sigma_v^2$ curves for unclipped OFDM and optimally clipped OFDM with threshold $\eta^\star$.

Table 10.1: SER vs. SNR for unclipped OFDM and OFDM with optimum clipping using threshold $\eta^\star$.

SNR	**−5 dB**	**0 dB**	**5 dB**	**10 dB**
$\boldsymbol{\eta^\star}$	0.6678	0.9896	1.2956	1.5768
SER for unclipped OFDM	4.915×10^{-1}	2.921×10^{-1}	7.394×10^{-2}	1.565×10^{-3}
SER with optimum clipping	2.391×10^{-1}	6.900×10^{-2}	4.408×10^{-3}	7.089×10^{-6}

Instead, we describe an interesting relationship between SNDR and capacity for a given input distribution, thus underscoring the importance of looking into SNDR.

10.3.1 Lower Bound on Capacity

The capacity is given by:

$$C = \max_{p_x} I(y; x), \tag{10.32}$$

where $I(y;x) = H(x) - H(x|y) = H(y) - H(y|x)$ is the mutual information between y and x, $H(x)$ is the entropy of x given by $E[-\log p_x(x)]$, and $p_x(x)$ is the PDF of x (base 2 logarithm is used throughout this chapter). Note that since the channel statistic $p_{y|x}(y|x)$ is available, using the Bayes rule, we can relate p_x to p_y and perform the maximization in (10.32) w.r.t. p_y. First, we find a lower bound on the mutual information as follows:

$$\begin{aligned} I(y;x) &= H(x) - H(x|y) \\ &\overset{①}{=} H(x) - H(x+\beta y|y) \\ &\overset{②}{\geq} H(x) - H(x+\beta y) \\ &\overset{③}{\geq} H(x) - \log(\pi e \sigma^2_{x+\beta y}) \\ &\overset{④}{=} H(x) - \log(\pi e(\sigma_x^2 + 2\mathrm{Re}\{\beta\sigma_{xy}\} + |\beta|^2\sigma_y^2)), \end{aligned} \tag{10.33}$$

where ① is because $H(x|y) = H(x+\beta y|y)$; ② is because the conditional entropy is smaller than or equal to the entropy; ③ is due to the entropy of a complex Gaussian r.v. always greater than or equal to the entropy of a complex r.v. with the same variance, and that the former is $\log(\pi e \sigma^2)$ [73]; ④ is obtained by expanding $\sigma^2_{x+\beta y}$. To obtain a tighter bound, we maximize the right hand side (RHS) of (10.33) with respect to β; the maximum occurs

with $\beta = -\frac{\sigma_{xy}^*}{\sigma_y^2}$, yielding

$$I(y;x) \geq H(x) - \log\left(\pi e\left(\sigma_x^2 - \frac{|\sigma_{xy}|^2}{\sigma_y^2}\right)\right). \tag{10.34}$$

Note that this bound applies for any input distribution p_x. By adding and subtracting $\log(\pi e\sigma_x^2)$ on the RHS of (10.34), we obtain

$$\begin{aligned} I(y;x) &\geq H(x) - \log(\pi e\sigma_x^2) + \log\left(\frac{\sigma_x^2}{\sigma_x^2 - \frac{|\sigma_{xy}|^2}{\sigma_y^2}}\right) \\ &= H(x) - \log(\pi e\sigma_x^2) + \log\left(\frac{\sigma_y^2}{\sigma_y^2 - \frac{|\sigma_{xy}|^2}{\sigma_x^2}}\right). \end{aligned} \tag{10.35}$$

Assuming that x and v are uncorrelated, we substitute $\sigma_{xy}/\sigma_x = AE[\gamma g(\gamma)]$, and $\sigma_y^2 = A^2E[|g(\gamma)|^2] + \sigma_v^2$ into (10.35) to obtain

$$\begin{aligned} I(y;x) &\geq H(x) - \log(\pi e\sigma_x^2) + \log\left(\frac{A^2E[|g(\gamma)^2] + \sigma_v^2}{A^2E[|g(\gamma)|^2] + \sigma_v^2 - A^2|E[\gamma g(\gamma)]|^2}\right) \\ &= H(x) - \log(\pi e\sigma_x^2) + \log(1 + \text{SNDR}), \end{aligned} \tag{10.36}$$

by referring to (10.11). Since $C \geq I(y;x)$ for any input distribution p_x, by setting p_x to be the PDF of a zero-mean complex Gaussian r.v., we obtain

$$C \geq \log(1 + \text{SNDR}), \tag{10.37}$$

with the SNDR evaluated for a complex Gaussian x.

10.3.2 Upper Bound on Capacity

Next, we present an upper bound for the capacity. Using the PDF of y that maximizes the capacity; i.e.,

$$p_y^* = \arg\max_{p_y}[H(y) - H(y|x)], \tag{10.38}$$

we can write the capacity as

$$\begin{aligned} C = I(y;x)|_{p_y^*} &= H(y)|_{p_y^*} - H(y|x) \\ &= H(y)|_{p_y^*} - H(v) \\ &= H(y)|_{p_y^*} - \log(\pi e\sigma_v^2). \end{aligned} \tag{10.39}$$

Next, we bound the entropy $H(y)$ with the entropy of a complex Gaussian y, yielding

$$\begin{aligned} C &\leq \log(\pi e \sigma_y^2) - \log(\pi e \sigma_v^2) \\ &= \log\left(1 + \frac{A^2\alpha^2 + \sigma_d^2}{\sigma_v^2}\right) \\ &= \log\left(1 + \frac{A^2 E[|g(\gamma)|^2]}{\sigma_v^2}\right) \\ &\leq \log\left(1 + \frac{A^2}{\sigma_v^2}\right). \end{aligned} \tag{10.40}$$

Since $|\alpha^2|\sigma_x^2 \leq |h(x)|^2 \leq A^2$ and $\varepsilon_d \geq 0$, we must have

$$\text{SNDR} = \frac{|\alpha^2|\sigma_x^2}{\varepsilon_d + \sigma_v^2} \leq \frac{A^2}{\sigma_v^2}.$$

This relationship can also be inferred by comparing (10.37) with (10.40). When the SNDR is maximized w.r.t. $g(\cdot)$ and the result is close to $\frac{A^2}{\sigma_v^2}$, the lower bound (10.37) and the upper bound (10.40) are close to each other; and we will have a good idea about the capacity.

10.3.3 Example on Capacity and Bounds

Assume that x is complex Gaussian distributed and the corresponding γ has the PDF in (10.22). Let us consider four nonlinear mappings:

$$g_1(\gamma) = \begin{cases} \frac{\gamma}{\eta^\star}, & \gamma \leq \eta^\star, \\ 1, & \gamma > \eta^\star, \end{cases} \tag{10.41}$$

where $\eta^\star$ is calculated as explained in Example 1 in Section 10.2.3,

$$g_2(\gamma) = \begin{cases} \frac{\gamma}{\eta}, & \gamma \leq \eta, \\ 1, & \gamma > \eta, \end{cases} \tag{10.42}$$

$$g_3(\gamma) = \begin{cases} \beta + \frac{1-\beta}{\eta}\gamma, & \gamma \leq \eta, \\ 1, & \gamma > \eta, \end{cases} \tag{10.43}$$

$$g_4(\gamma) = \begin{cases} -(\frac{\gamma}{\eta})^2 + \frac{2\gamma}{\eta}, & \gamma \leq \eta, \\ 1, & \gamma > \eta. \end{cases} \tag{10.44}$$

Fig. 10.1 (c), (d) and (f) illustrate $g_2(\gamma)$, $g_3(\gamma)$ and $g_4(\gamma)$ respectively.

For the nonlinearity $g_3(\gamma)$ in (10.43), it can be shown that

$$E[\gamma g_3(\gamma)] = \frac{1}{\eta}\left((1-\beta)(1-e^{-\eta^2}+\sqrt{\pi}\eta Q(\sqrt{2}\eta)) + \frac{1}{2}\beta\sqrt{\pi}\eta\right), \tag{10.45}$$

$$E[g_3(\gamma)^2] = \frac{1}{\eta^2}\left((1-\beta)^2(1-e^{-\eta^2}) + 2\beta(1-\beta)\sqrt{\pi}\eta(\frac{1}{2}-Q(\sqrt{2}\eta)) + \beta^2\eta^2\right), \tag{10.46}$$

and the corresponding SNDR can be calculated according to (10.11). Note that the SNDR expression for the nonlinearity $g_2(\gamma)$ in (10.42) can be obtained as a special case of the above with $\beta = 0$. For the nonlinearity $g_4(\gamma)$ in (10.44), it can be shown that

$$E[\gamma g_4(\gamma)] = \frac{1}{\eta^2}\left(-\frac{1}{2}\eta e^{-\eta^2} + 2\eta + (\frac{1}{2}\eta^2+\frac{3}{2})\sqrt{\pi}Q(\sqrt{2}\eta) - \frac{3}{4}\sqrt{\pi}\right), \tag{10.47}$$

$$E[g_4(\gamma)^2] = \frac{1}{\eta^4}\left(4\eta^2 - 6\sqrt{\pi}\eta(\frac{1}{2}-Q(\sqrt{2}\eta)) - 2e^{-\eta^2}+2\right), \tag{10.48}$$

and the SNDR can be calculated using (10.11).

Next, we set $\beta = 0.5$ for (10.43) and $\eta = 2.6283$ for (10.42)-(10.44), which ensures that $Pr\{|x(n)|/\sigma_x > \eta\} = 0.001$. We show in Fig. 10.5, $\log(1+A^2/\sigma_v^2)$ (line a), $\log(1+\text{SNDR}_1)$ (line b), $\log(1+\text{SNDR}_2)$ (line c), $\log(1+\text{SNDR}_3)$ (line d), $\log(1+\text{SNDR}_4)$ (line e), as a function of A^2/σ_v^2, where SNDR_i corresponds to $g_i(\gamma)$, $i = 1,\ 2,\ 3,\ 4$. Notice that line a is higher than line b, line b is higher than lines c-e. Capacity lies in between lines a and b. This example also illustrates that the nonlinearity $g_1(\gamma)$ in (10.41) yields a higher SNDR as compared to the other nonlinearities, as predicted by Theorem 1. Note that $g_1(\gamma)$ and $g_2(\gamma)$ are different, since $\eta^\star$ in (10.41) is obtained as a function of A^2/σ_v^2, whereas η in (10.42) remains constant.

10.4 Conclusions

Many communications devices are nonlinear and have a peak power or peak amplitude constraint. In addition to possibly amplifying the useful signal, nonlinearity also generates distortions. A measure that takes into account both these effects is the signal-to-noise-and-distortion ratio (SNDR). The focus of this chapter is on SNDR optimization within the family of amplitude limited memoryless nonlinearities. We showed that under the peak

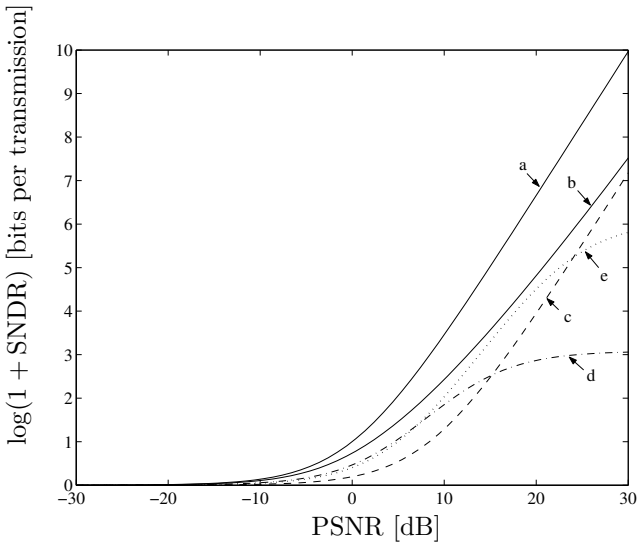


Figure 10.5: Bounds on capacity: line a corresponds to the upper bound $\log(1 + A^2/\sigma_v^2)$; line b corresponds to the lower bound $\log(1+\text{SNDR})$ obtained using the optimum $g_1(\cdot)$ with optimum threshold $\eta^\star$ and complex Gaussian x; lines c-e are looser lower bounds obtained with $g_2(\cdot)$, $g_3(\cdot)$, and $g_4(\cdot)$.

amplitude constraint, the nonlinearity that maximizes the SNDR is a soft limiter with gain, and the specific gain (or equivalently, the threshold of the limiter) is found according to the peak signal to noise ratio (PSNR) and the distribution of the input amplitude. When the input is complex Gaussian distributed, the optimum $\log(1 + \text{SNDR})$ also provides a tight lower bound on the nonlinear channel capacity. We provided a clipped OFDM transmission example to illustrate how simple optimum clipping can improve the system performance. The results of this chapter are also of interest in applications such as predistortion linearization and peak-to-average power ratio reduction.

10.A Proof of Lemma 1

Since we will be optimizing w.r.t. to a function, we introduce here the notion of functional derivative [53]

$$\frac{\delta F[g(\gamma)]}{\delta g(\gamma_o)} \triangleq \lim_{\epsilon \to 0} \frac{F[g(\gamma) + \epsilon\delta(\gamma - \gamma_o)] - F[g(\gamma)]}{\epsilon}, \tag{10.49}$$

where $\delta(\cdot)$ denotes the Dirac delta function. From (10.49), we infer that

$$\frac{\delta g(\gamma)}{\delta g(\gamma_o)} = \delta(\gamma - \gamma_o), \tag{10.50}$$

$$\frac{\delta g^2(\gamma)}{\delta g(\gamma_o)} = 2g(\gamma)\delta(\gamma - \gamma_o). \tag{10.51}$$

Due to the constraint $g(\cdot) \leq 1$, we assume the following form for $g(\cdot)$:

$$\begin{aligned} g(\gamma) &< 1, \gamma \in S, \\ g(\gamma) &= 1, \gamma \notin S, \end{aligned} \tag{10.52}$$

where S is a subset of $[0, \infty)$. To maximize the SNDR w.r.t. $g(\cdot)$, we need

$$\frac{\delta \text{SNDR}}{\delta g(\gamma_o)} = 0, \qquad \forall\ \gamma_o \in S. \tag{10.53}$$

Denote by $I(\cdot)$ the indicator function. Since $I(\gamma \notin S) + I(\gamma \in S) = 1$, we infer that for the $g(\cdot)$ in (10.52),

$$\begin{aligned} E[g^2(\gamma)] &= E[I(\gamma \notin S)g^2(\gamma)] + E[I(\gamma \in S)g^2(\gamma)] \\ &= E[I(\gamma \notin S)] + E[I(\gamma \in S)g^2(\gamma)] \\ &= C_o + E[I(\gamma \in S)g^2(\gamma)], \end{aligned} \tag{10.54}$$

where

$$C_o = E[I(\gamma \notin S)]. \tag{10.55}$$

Similarly,

$$E[\gamma g(\gamma)] = C_1 + E[I(\gamma \in S)\gamma g(\gamma)], \tag{10.56}$$

where

$$C_1 = E[\gamma I(\gamma \notin S)]. \tag{10.57}$$

It follows easily that $C_o \geq 0$, $C_1 \geq 0$. Substituting (10.54) and (10.56) into (10.11), we obtain

$$\text{SNDR} = \frac{\left(C_1 + E[I(\gamma \in S)\gamma g(\gamma)]\right)^2}{C_o + E[I(\gamma \in S)g^2(\gamma)] - \left(C_1 + E[I(\gamma \in S)\gamma g(\gamma)]\right)^2 + \frac{\sigma_v^2}{A^2}}. \tag{10.58}$$

Note that C_o and C_1 are functions of the set S and the probability density function (PDF) of γ but are not functions of $g(\cdot)$. For notational simplicity, we omit their S-dependence in this section.

For the SNDR expression in (10.58), let us denote

$$Q[g(\gamma)] = C_1 + E[I(\gamma \in S)\gamma g(\gamma)], \tag{10.59}$$

$$N[g(\gamma)] = (Q[g(\gamma)])^2, \tag{10.60}$$

$$D[g(\gamma)] = C_o + E[I(\gamma \in S)g^2(\gamma)] - (Q[g(\gamma)])^2 + \frac{\sigma_v^2}{A^2}. \tag{10.61}$$

Then SNDR $= \frac{N[g(\gamma)]}{D[g(\gamma)]}$. Condition (10.53) requires

$$\frac{\delta N[g(\gamma)]}{\delta g(\gamma_o)} D[g(\gamma)] = \frac{\delta D[g(\gamma)]}{\delta g(\gamma_o)} N[g(\gamma)]. \tag{10.62}$$

Denote by $p(\gamma)$ the PDF of the random variable γ. Then

$$E[I(\gamma \in S)\gamma g(\gamma)] = \int I(\gamma \in S)\gamma g(\gamma) p(\gamma) d\gamma. \tag{10.63}$$

Taking the functional derivative of (10.63) w.r.t. $g(\gamma_o)$ and using (10.50), we obtain

$$\begin{aligned} \frac{\delta E[I(\gamma \in S)\gamma g(\gamma)]}{\delta g(\gamma_o)} &= \int I(\gamma \in S)\gamma \delta(\gamma - \gamma_o) p(\gamma) d\gamma \\ &= \gamma_o p(\gamma_o), \end{aligned} \tag{10.64}$$

for $\gamma_o \in S$. Similarly, using (10.51) we obtain

$$\frac{\delta E[I(\gamma \in S)g^2(\gamma)]}{\delta g(\gamma_o)} = 2g(\gamma_o)p(\gamma_o). \tag{10.65}$$

Therefore,

$$\frac{\delta N[g(\gamma)]}{\delta g(\gamma_o)} = 2Q[g(\gamma)]\gamma_o p(\gamma_o) \tag{10.66}$$

$$\frac{\delta D[g(\gamma)]}{\delta g(\gamma_o)} = 2g(\gamma_o)p(\gamma_o) - 2Q[g(\gamma)]\gamma_o p(\gamma_o). \tag{10.67}$$

Substituting (10.59)-(10.61), (10.66)-(10.67) into (10.62) and simplifying, we obtain

$$g(\gamma_o) = \frac{\gamma_o}{\eta}, \quad \eta = \frac{C_1 + E[I(\gamma \in S)\gamma g(\gamma)]}{C_o + E[I(\gamma \in S)g^2(\gamma)] + \frac{\sigma_v^2}{A^2}}, \tag{10.68}$$

as the solution for (10.53). Since (10.68) holds $\forall\ \gamma_o \in S$, we must have

$$g(\gamma) = \begin{cases} \frac{\gamma}{\eta}, \gamma \in S, \\ 1, \gamma \notin S. \end{cases} \tag{10.69}$$

Substituting (10.69) into (10.68), we obtain

$$\eta = \frac{C_1 + \frac{1}{\eta}\bar{C}_2}{C_o + \frac{1}{\eta^2}\bar{C}_2 + \frac{\sigma_v^2}{A^2}}, \tag{10.70}$$

where

$$\bar{C}_2 = E[I(\gamma \in S)\gamma^2]. \tag{10.71}$$

Similar to the notations of (10.55) and (10.57), we may define $C_2 = E[I(\gamma \notin S)\gamma^2]$ and infer that $\bar{C}_2 = E[\gamma^2] - C_2 = 1 - C_2$. Since both sides of (10.70) contain η, we cross multiply and further simplify it to

$$\eta = \frac{C_1}{C_o + \frac{\sigma_v^2}{A^2}}. \tag{10.72}$$

In summary, under the peak amplitude constraint, the optimal $g(\cdot)$ that maximizes the SNDR is given by (10.69) where η is given by (10.72).

10.B Proof of Theorem 1

Comparing (10.52) with (10.69), we infer that $\gamma < \eta$ on S. Therefore, the set S must be a subset of $S^\star \triangleq [0, \eta)$; i.e., $S \subseteq S^\star$. The objective here is to determine the optimal S.

Since $g(\gamma) = \frac{\gamma}{\eta}$ for $\gamma \in S$, we infer that

$$E[I(\gamma \in S)\gamma g(\gamma)] = \frac{1}{\eta}E[I(\gamma \in S)\gamma^2] = \frac{\bar{C}_2}{\eta},$$

$$E[I(\gamma \in S)g^2(\gamma)] = \frac{1}{\eta^2}E[I(\gamma \in S)\gamma^2] = \frac{\bar{C}_2}{\eta^2}.$$

Hence, we can rewrite (10.58) as

$$\text{SNDR} = \frac{\left(C_1 + \frac{1}{\eta}\bar{C}_2\right)^2}{C_o + \frac{1}{\eta^2}\bar{C}_2 - \left(C_1 + \frac{1}{\eta}\bar{C}_2\right)^2 + \frac{\sigma_v^2}{A^2}}. \tag{10.73}$$

From (10.72), we infer that $C_o + \frac{\sigma_v^2}{A^2} = \frac{C_1}{\eta}$. Thus, (10.73) can be further simplified to

$$\text{SNDR} = \frac{\left(C_1 + \frac{1}{\eta}\bar{C}_2\right)^2}{\frac{1}{\eta}C_1 + \frac{1}{\eta^2}\bar{C}_2 - \left(C_1 + \frac{1}{\eta}\bar{C}_2\right)^2} = \frac{1}{\frac{1}{\eta C_1 + \bar{C}_2} - 1}. \tag{10.74}$$

Recall that C_o, C_1, $\bar{C}_2$ and η are all functions of S. According to (10.74), maximizing SNDR w.r.t. S is equivalent to maximizing (c.f. (10.72))

$$R(S) = \eta(S)C_1(S) + \bar{C}_2(S) = \frac{C_1^2(S)}{C_o(S) + \frac{\sigma_v^2}{A^2}} + \bar{C}_2(S) \tag{10.75}$$

w.r.t. S (we now emphasize the S-dependence).

Let us denote

$$\bar{S} = S^\star \backslash S, \tag{10.76}$$

$$\Delta_i = C_i(S) - C_i(S^\star) = E[I(\gamma \notin S)\gamma^i] - E[I(\gamma \notin S^\star)\gamma^i] = E[I(\gamma \in \bar{S})\gamma^i] \geq 0, \quad i = 0, 1, 2 \tag{10.77}$$

In the sequel, we will show that $R(S^\star) - R(S) \geq 0$, $\forall S \subseteq S^\star$, thus establishing that $S^\star$ maximizes the SNDR.

Denote $G = C_o(S^\star) + \frac{\sigma_v^2}{A^2}$, and thus $C_o(S) + \frac{\sigma_v^2}{A^2} = G + \Delta_0$. Based on (10.75), we write

$$R(S^\star) - R(S) = \frac{C_1^2(S^\star)}{C_o(S^\star) + \frac{\sigma_v^2}{A^2}} + \bar{C}_2(S^\star) - \frac{C_1^2(S)}{C_o(S) + \frac{\sigma_v^2}{A^2}} - \bar{C}_2(S) \tag{10.78}$$

$$= \frac{C_1^2(S^\star)}{G} - \frac{C_1^2(S)}{G + \Delta_0} + C_2(S) - C_2(S^\star). \tag{10.79}$$

Recall that $C_1(S) = C_1(S^\star) + \Delta_1$, $C_2(S) = C_2(S^\star) + \Delta_2$. We infer that

$$R(S^\star) - R(S) = \frac{(G + \Delta_0)C_1^2(S^\star) - G(C_1(S^\star) + \Delta_1)^2 + \Delta_2 G(G + \Delta_0)}{G(G + \Delta_0)}. \tag{10.80}$$

The denominator $G(G + \Delta_0)$ is always positive. The numerator can be shown to be

$$(C_1(S^\star)\sqrt{\Delta_0} - G\sqrt{\Delta_2})^2 + 2GC_1(S^\star)(\sqrt{\Delta_2}\sqrt{\Delta_0} - \Delta_1) + G(\Delta_2\Delta_0 - \Delta_1^2).$$

Recall the Cauchy-Schwartz inequality $(E[\theta\phi])^2 \leq E[\theta^2]E[\phi^2]$. Letting $\theta = \gamma I(\gamma \in \bar{S})$ and $\phi = I(\gamma \in \bar{S})$, we infer that $\Delta_2\Delta_0 \geq \Delta_1^2$. Therefore, the above numerator is ≥ 0.

We have thus proved that $R(S^\star) - R(S) \geq 0$; i.e., $R(S^\star) \geq R(S)$, $\forall\ S \subseteq S^\star$. This implies that the optimal S, in the sense of maximizing the SNDR, is $S^\star = [0, \eta)$.

Now $(\gamma \notin S^\star)$ means $(\gamma \geq \eta)$. We thus express

$$C_o(S) = C_o(\eta) = E[I(\gamma \geq \eta)] = Pr(\gamma \geq \eta), \tag{10.81}$$

$$C_1(S) = C_1(\eta) = E[\gamma I(\gamma \geq \eta)], \tag{10.82}$$

$$\bar{C}_2(S) = \bar{C}_2(\eta) = E[I(\gamma < \eta)\gamma^2]. \tag{10.83}$$

Next, we infer from (10.72) that

$$T(\eta) \triangleq \frac{\eta}{C_1(\eta) - \eta C_o(\eta)} = \frac{A^2}{\sigma_v^2}. \tag{10.84}$$

The above $T(\eta)$ is a nonlinear function of η and can be shown to be monotonically increasing in η. Therefore, (10.84) can be used to solve for the optimal $\eta^\star = T^{-1}(A^2/\sigma_v^2)$ for a given PDF of γ. The optimal SNDR is

$$\text{SNDR}^\star = \frac{1}{\frac{1}{R(\eta^\star)} - 1}, \tag{10.85}$$

where

$$R(\eta^\star) = \frac{C_1^2(\eta^\star)}{C_o(\eta^\star) + \frac{\sigma_v^2}{A^2}} + \bar{C}_2(\eta^\star). \tag{10.86}$$

Lastly, we point out that $S^\star$ may not be the only maximizer of the SNDR. These cases should make (10.80) zero. One such case is when $p(\gamma) = 0$, $\forall \gamma \in \bar{S}$, which yields $\Delta_i = 0$ for $i = 0, 1, 2$. This means that S can be a partial set of $S^\star$, which does not include values of γ for which $p(\gamma) = 0$. This is intuitive, since the probability of $\gamma \in \bar{S}$ is 0, the value $g(\gamma)$ for $\gamma \in \bar{S}$ does not affect the SNDR.

CHAPTER XI

CONCLUSIONS

In this dissertation, we considered the application of signal processing to the research on nonlinear power amplifiers in the following areas: characterization and modeling, spectral regrowth analysis, linearization, and communication aspects.

11.1 Contributions

In this thesis, we presented results in following areas:

- In the area of PA modeling, we derived the passband and baseband PA input/output relationships. We examined the differences in formulation when the PAs exhibits long-term, short-term, or no memory effects. We clarified the appropriate formulation of the baseband representation and verified against experimental measurements of an actual PA.

- In the area of PA linearization, we presented a Hammerstein model predistorter for a Wiener model PA. We offered an algorithm to obtain the predistorter coefficients and analyzed and compared its performance to existing algorithms. We presented closed-form expressions for an orthogonal polynomial predistorter. We demonstrated the improvement in numerical stability associated with the use of orthogonal polynomials for predistortion.

- In the area of PA spectral regrowth analysis, we presented spectral analysis for a PA modeled using orthogonal polynomials. We specialized the spectral analysis results to the case of digitally modulated signals. We showed that by taking into account the cyclostationary nature of the processes, more accurate spectral analysis results can be obtained. Using the Leonov-Shiryaev formula, we analyzed spectral regrowth at the output of a PA with QPSK and OQPSK modulated input. We obtained closed-form

output PSD expressions that apply to an arbitrary-order nonlinearity, which may include memory effects.

- In the area of applications of PAs to communications, we performed an SNDR optimization within the family of amplitude limited memoryless nonlinearities. We showed that under the peak amplitude constraint, the nonlinearity that maximizes the SNDR is a soft limiter with gain, and the specific gain (or equivalently, the threshold of the limiter) was found. We obtained a link between the capacity of amplitude-limited nonlinear channels with Gaussian noise to the SNDR.

Following is the list of publications resulted from the work presented in this thesis:

Journal papers

J1. R. Raich and G. T. Zhou, "Orthogonal polynomials for complex Gaussian processes," *IEEE Trans. on Signal Processing*, accepted Oct. 2003.

J2. G. T. Zhou and R. Raich, "Spectral analysis of polynomial nonlinearity with applications to RF power amplifiers," *EURASIP Journal on Applied Signal Processing, Special Issue on Nonlinear Signal and Image Processing*, accepted Dec. 2003.

J3. R. Raich, H. Qian, and G. T. Zhou, "Orthogonal polynomials for power amplifier modeling and predistorter design," *IEEE Trans. on Vehicular Technology*, submitted Jan. 2003.

J4. G. T. Zhou, H. Qian, L. Ding, and R. Raich, "On baseband representation of passband nonlinearities," *IEEE Communications Letters*, submitted Nov. 2003.

J5. R. Raich and G. T. Zhou, "Statistical analysis of a bandpass nonlinearity with nonstationary input," *IEEE Trans. on Signal Processing*, submitted Nov. 2003.

J6. R. Raich, H. Qian, and G. T. Zhou, "Optimization of SNDR for amplitude limited nonlinearities," *IEEE Trans. on Communications*, submitted Dec. 2003.

Conference papers

C1. R. Raich and T. Zhou, "Analyzing spectral regrowth of QPSK and OQPSK signals," in *Proc. IEEE Intl. Conf. Acoust., Speech, Signal Processing*, vol. 4, (Salt Lake City, Utah), pp. 2673–2676, May 2001.

C2. J. S. Kenney, W. Woo, L. Ding, R. Raich, H. Ku, and G. T. Zhou, "The impact of memory effects on predistortion linearization of RF power amplifiers," in *Proc. 8th Intl. Symposium on Microwave and Optical Technology (ISMOT'2001)*, (Montreal, Quebec, Canada), pp. 189–193, June 2001.

C3. L. Ding, R. Raich, and G. T. Zhou, "A Hammerstein predistortion linearization design based on the indirect learning architecture," in *Proc. IEEE Intl. Conf. Acoust., Speech, Signal Processing*, vol. 3, (Orlando, Florida), pp. 2689–2692, May 2002.

C4. Y. C. Park, W. Woo, R. Raich, J. S. Kenney, and G. T. Zhou, "Adaptive predistortion linearization of RF power amplifiers using lookup tables generated from subsampled data," in *Proc. IEEE Radio and Wireless Conference*, pp. 233–236, Aug. 2002.

C5. R. Raich and G. T. Zhou, "on the modeling of memory nonlinear effects of power amplifiers for communication applications," in *Proc. 10th IEEE DSP Workshop*, (Pine Mountain, GA), pp. 7–10, Oct. 2002.

C6. R. Raich, H. Qian, and G. T. Zhou, "Digital baseband predistortion of nonlinear power amplifiers using orthogonal polynomials," in *Proc. IEEE Intl. Conf. Acoust., Speech, Signal Processing*, vol. 6, pp. 689–692, Apr. 2003.

C7. G. T. Zhou and R. Raich, "Closed-form expressions of output power spectrum for nonlinear power amplifiers with or without memory," in *Proc. IEEE - EURASIP Workshop on Nonlinear Signal and Image Processing*, (Grado, Italy), June 2003.

C8. R. Raich and G. T. Zhou, "Theory and applications of orthogonal polynomials for gaussian input," in *Proc. IEEE Workshop on Statistical Signal Processing*, (St. Louis, Missouri), pp. 97–100, Sep. 2003.

C9. R. Raich and G. T. Zhou, "Spectral analysis for bandpass nonlinearity with cyclostationary input," in *Proc. IEEE Intl. Conf. Acoust., Speech, Signal Processing*, May 2004, to appear.

C10. H. Qian, R. Raich, and G. T. Zhou, "On the benefits of deliberately introduced baseband nonlinearities in communication systems," in *Proc. IEEE Intl. Conf. Acoust., Speech, Signal Processing*, May 2004, to appear.

C11. R. Raich, H. Qian, and G. T. Zhou, "Signal to noise and distortion ratio considerations for nonlinear communication channels," in *Proc. IEEE 6th CAS Symposium on Emerging Technologies: Frontiers of Mobile and Wireless Communication*, submitted Jan. 2004.

Patents

P1. R. Raich, Q. Hua, and G. Zhou, "Orthogonal polynomials for power amplifier modeling and predistorter design." US provisional patent filed October 15, 2003. GTRC ID 2772.

11.2 Future Work

The work presented in this thesis can be extended to the following areas:

- Analysis of amplitude-limited nonlinearities in fading channels. In Chapter 10, we presented an SNDR analysis of amplitude-limited nonlinearities when the channel is additive white Gaussian noise (AWGN). This work can be extended to include fading channels and symbol-error-rate (SER) analysis.

- PA modeling and predistortion of a temperature-varying nonlinearity model. The memoryless model $y(t) = f(z(t))$ can be extended such that $f(\cdot)$ will depend on the PA output power, which is related to the PA temperature. The output of such PA is given by

$$y(t) = f(z(t), p(t))$$
$$p(t) = \int\int z(t-u)z^*(t-v)P(u,v)dudv,$$

where $p(t)$ denotes the PA output power and $P(u,v) = P^*(v,u)$. Using this formulation, the PA output signal $y(t)$ depends on the PA input signal $x(t)$ and the instantaneous PA output power $p(t)$. Given this model, PA parameters such as the functions $f(\cdot,\cdot)$ and $P(\cdot,\cdot)$ can be estimated to identify the PA. Similar approach can be applied with the PA predistortion since the PA described can be predistorted by a similar model:

$$z(t) = g(x(t), q(t))$$
$$q(t) = \int\int x(t-u)x^*(t-v)Q(u,v)dudv.$$

Note that the predistorter characteristics will vary based on its input power. Both PA modeling and predistortion for the temperature based model can be further investigated.

REFERENCES

[1] TIA/EIA-95. *Mobile Station - Base Station Compatibility Standard for Dual-Mode Wideband Spread Spectrum Cellular Systems*, TIA, 1997.

[2] "Operation and Performance of the ISL5239 Pre-distortion Linearizer", Application Note AN1022, July 2002.

[3] ANSI, "Network and customer installation interfaces-asymmetric digital subscriber line (ADSL) metallic interface," 1995.

[4] ETSI, "Radio broadcasting systems; digital audio broadcasting (DAB) to mobile, protable and fixed receivers," Feb. 1995.

[5] Abramowitz, M. and Stegun, I. A., eds., *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. New York: Dover Publications, Inc., 1972.

[6] Akaiwa, Y. and Nagata, Y., "Highly efficient digital mobile communication with a linear modulation method," *IEEE Journal on Selected Areas in Communications*, vol. 5, pp. 890–895, June 1987.

[7] Aparin, V., "Analysis of CDMA signal spectral regrowth and waveform quality," *IEEE Trans. on Microwave Theory and Techniques*, vol. 49, pp. 2306–2314, Dec. 2001.

[8] Aparin, V. and Larson, L. E., "Analysis and reduction of cross-modulation distortion in CDMA receivers," *IEEE Trans. on Microwave Theory and Techniques*, vol. 51, pp. 1591–1602, May 2003.

[9] Bai, E. W., "An optimal two stage identification algorithm for Hammerstein-Wiener nonlinear systems," in *Proc. American Contr. Conf.*, pp. 2756–2760, 1998.

[10] Banelli, P., "Theoretical analysis and performance of OFDM signals in nonlinear fading channels," *IEEE Trans. on Wireless Communications*, vol. 2, pp. 284–293, Mar. 2003.

[11] Banelli, P., Baruffa, G., and Cacopardi, S., "Effects of HPA nonlinearity on frequency multiplexed OFDM signals," *IEEE Trans. on Broadcasting*, vol. 47, pp. 123–136, June 2001.

[12] Banelli, P. and Cacopardi, S., "Theoretical analysis and performance of OFDM signals in nonlinear AWGN channels," *IEEE Trans. on Communications*, vol. 48, pp. 430–441, Mar. 2000.

[13] Barbieri, A., Cariolaro, G., and Vangelista, L., "Nonlinear models of TWT revisited for OFDM systems," in *proc. 38th Midwest Symposium on Circuits and Systems*, vol. 1, pp. 522–525, Aug. 1995.

[14] Barrett, J. and Lampard, D., "An expansion for some second-order probability distributions and its application to noise problems," *IEEE Trans. on Inform. Theory*, vol. 1, pp. 10–15, Mar. 1955.

[15] Baudoin, G. and Jardin, P., "Adaptive polynomial pre-distortion for linearization of power amplifiers in wireless communications and WLAN," in *Proc. Intl. Conf. on Trends in Communications (EUROCON'2001)*, vol. 1, pp. 157–160, July 2001.

[16] Baum, R., "The correlation function of Gaussian noise passed through nonlinear devices," *IEEE Trans. on Inform. Theory*, vol. 15, pp. 448–456, July 1969.

[17] Benedetto, S., Biglieri, E., and Daffara, R., "Modeling and performance evaluation of nonlinear satellite links — a Volterra series approach," *IEEE Trans. on Aerospace and Electronic Systems*, vol. 15, pp. 494–507, July 1979.

[18] Benedetto, S. and Biglieri, E., *Principles of Digital Transmission with Wireless Applications*. Plenum Series in Telecommunications, Kluwer Academic Publishers, July 1999.

[19] Besbes, H., Le-Ngoc, T., and Lin, H., "A fast adaptive polynomial predistorter for power amplifiers," in *Global Telecommunications Conference, 2001. GLOBECOM '01. IEEE*, vol. 1, pp. 659–663, 2001.

[20] Blachman, N., "The uncorrelated output components of a nonlinearity," *IEEE Trans. on Inform. Theory*, vol. 14, pp. 250–255, Mar. 1968.

[21] Blachman, N., "The output signals and noise from a nonlinearity with amplitude-dependent phase shift," *IEEE Trans. on Inform. Theory*, vol. 25, pp. 77–79, Jan. 1979.

[22] Black, H. S., "Translating system." United States Patent No. 1,686,792, Oct. 9 1929.

[23] Black, H. S., "Inventing the negative feedback amplifier," *IEEE Spectrum*, vol. 14, pp. 55–60, Dec. 1977.

[24] Bösch, W. and Gatti, G., "Measurements and simulation of memory effects in predistortion linearizers," *IEEE Trans. on Microwave Theory and Techniques*, vol. 37, pp. 1885–1890, December 1989.

[25] Brillinger, D. R., *Time Series: Data Analysis and Theory*. San Francisco: Holden-day Inc., 1981.

[26] Brown, J., "On a cross-correlation property for stationary random processes," *IEEE Trans. on Inform. Theory*, vol. 3, pp. 28–31, Mar. 1957.

[27] Cambanis, S. and Liu, B., "On the expansion of a bivariate distribution and its relationship to the output of a nolinearity," *IEEE Trans. on Inform. Theory*, vol. 17, pp. 17–25, Jan. 1971.

[28] Cariolaro, G., Michieletto, G., Stivanello, G., and Vangelista, L., "Spectral analysis at the output of a TWT driven by an OFDM signal," in *proc. of the 1994 Singapore International Conference on Communication Systems (ICCS)*, vol. 2, pp. 653–657, Nov. 1994.

[29] Chang, H.-C., Chen, S.-W., Moe, D., and Hackman, S., "CDMA ACPR prediction using two-tone intermodulation spectrum by modeling device non-linearity with tchebyshev polynomials," in *Proc. Wireless Communications Conference*, pp. 117–120, Aug. 1997.

[30] Chia, S., "The Universal Mobile Telecommunication System," *Communications Magazine*, vol. 30, pp. 54–62, Dec. 1992.

[31] Chie, C., "A modified barrett-lampard expansion and its application to bandpass nonlinearities with both AM-AM and AM-PM conversion," *IEEE Trans. on Communications*, vol. 28, pp. 1859–1866, Nov. 1980.

[32] Ciblat, P., Loubaton, P., Serpedin, E., and Giannakis, G. B., "Asymptotic analysis of blind cyclic correlation-based symbol-rate estimators," *IEEE Trans. on Inform. Theory*, vol. 48, pp. 1922–1934, July 2002.

[33] Clark, C. J., Chrisikos, G., Muha, M. S., Moulthrop, A. A., and Silva, C. P., "Time-domain envelope measurement technique with application to wideband power amplifier modeling," *IEEE Trans. on Microwave Theory and Techniques*, vol. 46, pp. 2531–2540, Dec. 1998.

[34] Costa, E., Midrio, M., and Pupolin, S., "Impact of amplifier nonlinearities on OFDM transmission system performance," *IEEE Communications Letters*, vol. 3, pp. 37–39, Feb. 1999.

[35] Cripps, S. C., *RF Power Amplifiers for Wireless Communications*. Norwood, MA: Artech House, 1999.

[36] Cripps, S. C., *Advanced Techniques in RF Power Amplifier Design*. Norwood, MA: Artech House, 2002.

[37] Dardari, D., "Spectral properties and signal-to-distortion ratio of arbitrary spectrum Gaussian signals in the presence of nonlinearities," in *Proc. Global Telecommunications Conference (GLOBECOM)*, vol. 2, pp. 1147–1151, Nov. 2001.

[38] Ding, L., Zhou, G. T., Morgan, D. R., Ma, Z., Kenney, J. S., Kim, J., and Giardina, C. R., "Memory polynomial predistorter based on the indirect learning architecture," in *Proc. IEEE Global Telecommunications Conference*, (Taipei, Taiwan), Nov 2002.

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

[40] Ding, L. and Zhou, G., "Effects of even-order nonlinear terms on predistortion," in *Proc. 10th IEEE DSP Workshop (DSP'2002)*, (Pine Mountain, GA), Oct 2002.

[41] Ding, L., Raich, R., and Zhou, G. T., "A Hammerstein predistortion linearization design based on the indirect learning architecture," in *Proc. IEEE Intl. Conf. Acoust., Speech, Signal Processing*, vol. 3, (Orlando, Florida), pp. 2689–2692, May 2002.

[42] Ekstrom, J. L., "The correlation function of tanh limited noise," in *proc. 35th Midwest Symposium on Circuits and Systems*, vol. 2, pp. 970–971, Aug. 1992.

[43] Ermolova, N. Y., "Analysis of nonlinear effects in OFDM communications systems," in *IEEE proc. 53rd Vehicular Technology Conference (VTC)*, vol. 1, pp. 737–740, May 2001.

[44] Ermolova, N. Y., "Spectral analysis of nonlinear amplifier based on the complex gain Taylor series expansion," *IEEE Communications Letters*, vol. 5, pp. 465–467, Dec. 2001.

[45] Eskinat, E., Johnson, S. H., and Luyben, W. L., "Use of Hammerstein models in identification of nonlinear systems," *AIChE J.*, vol. 37, pp. 255–267, Feb. 1991.

[46] Eun, C. and Powers, E. J., "A predistorter design for a memory-less nonlinearity preceded by a dynamic linear system," in *Proc. Global Telecommunications Conference*, vol. 1, (Singapore), pp. 152–156, November 1995.

[47] Eun, C. and Powers, E. J., "A new Volterra predistorter based on the indirect learning architecture," *IEEE Trans. Signal Processing*, vol. 45, pp. 223–227, Jan. 1997.

[48] Faulkner, M. and Mattsson, T., "Spectral sensitivity of power amplifiers to quadrature modulator misalignment," *IEEE Trans. on Vehicular Technology*, vol. 41, pp. 516–525, November 1992.

[49] Gard, K., Steer, M. B., and Larson, L. E., "Generalized autocorrelation analysis of spectral regrowth from bandpass nonlinear circuits," in *Proc. IEEE MTT-S International Microwave Symposium Digest*, vol. 1, pp. 9–12, May 2001.

[50] Gard, K. G., Gutierrez, H. M., and Steer, M. B., "Characterization of spectral regrowth in microwave amplifiers based on the nonlinear transformation of a complex Gaussian process," *IEEE Trans. on Microwave Theory and Techniques*, vol. 47, pp. 1059–1069, July 1999.

[51] Gardner, W. A., *Introduction to Random Processes with Applications to Signals and Systems*. New York: McGraw-Hill, 2nd ed. ed., 1990.

[52] Gautschi, W., "On the preceding paper "A Legendre polynomial integral" by James L. Blue," *Mathematics of Computation*, vol. 33, pp. 742–743, April 1979.

[53] Gelfand, I. M. and Fomin, S. V., *Calculus of Variations*. Prentice-Hall, Inc., 1963.

[54] Giannakis, G. B., "Cyclostationary signal analysis." in *The Digital Signal Processing Handbook,* (V. K. Madisetti and D. B. Williams, eds.), Chapter 17, Boca Raton, FL: CRC, 1998.

[55] Goh, T. S. L. and Pollard, R. D., "ACPR prediction of CDMA systems through statistical behavioural modelling of power amplifiers with memory," in *proc. 8th International Conference on Communication Systems (ICCS)*, vol. 2, pp. 1189–1193, Nov. 2002.

[56] Goh, T. S. L., Pollard, R. D., and Boussakta, S., "ACPR prediction of multicarrier systems through behavioural modelling of power amplifiers using measured two-tone transfer characteristics and statistical techniques," in *proc. 7th IEEE High Frequency Postgraduate Student Colloquium*, vol. 2, p. 10, Sept. 2002.

[57] Gusmão, A., Gonçalves, V., and Esteves, N., "A novel approach to modeling of OQPSK-type digital transmission over nonlinear radio channels," *IEEE Journal on Selected Areas in Communications*, vol. 15, pp. 647–655, May 1997.

[58] Gutierrez, H., Gard, K., and Steer, M. B., "Nonlinear gain compression in microwave amplifiers using generalized power-series analysis and transformation of input statistics," *IEEE Trans. on Microwave Theory and Techniques*, vol. 48, pp. 1774–1777, Oct. 2000.

[59] Huang, X., Lu, J., Chuang, J., and Zheng, J., "Companding transform for the reduction of peak-to-average power ratio of OFDM signals," in *Proc. IEEE Vehicular Technology Conference*, vol. 2, pp. 835–839, May 2001.

[60] Jin, M., Shin, S. K., Oh, D., and Kim, J., "Reduced order RLS polynomial predistortion," in *proc. 2003 International Symposium on Circuits and Systems (ISCAS)*, vol. 4, pp. 333–336, May 2003.

[61] Kang, H. W., Cho, Y. S., and Youn, D. H., "On compensating nonlinear distortions of an OFDM system using an efficient adaptive predistorter," *IEEE Trans. on Communications*, vol. 47, pp. 522–526, Apr. 1999.

[62] Kang, H. W., Cho, Y. S., and hee Youn, D., "An efficient adaptive predistorter for nonlinear high power amplifier in satellite communication," in *Circuits and Systems, 1997. ISCAS '97., Proceedings of 1997 IEEE International Symposium on*, vol. 4, pp. 2288–2291, June 1997.

[63] Kay, S. M., *Fundamentals of Statistical Signal Processing : Estimation Theory*. Norwood, MA: Englewood Cliffs, New Jersey, 2000.

[64] Kenington, P. B., *High-Linearity RF Amplifer Design*. Norwood, MA: Artech House, 2000.

[65] Kim, J. and Konstantinou, K., "Digital predistortion of wideband signals based on power amplifier model with memory," *IEE Electronics Letters*, vol. 37, pp. 1417–1418, Nov. 2001.

[66] Kusunoki, S., Yamamoto, K., Hatsugai, T., Tagami, K., Nagaoka, H., Tominaga, N., Osawa, K., Tanabe, K., Sakurai, S., and Iida, T., "Power amplifier module with digital adaptive predistortion for cellular phone," in *Proc. IEEE MTT-S International Microwave Symposium Digest*, vol. 2, pp. 765–768, June 2002.

[67] Lavoie, J. L., "On inverses of finite segments of the generalized Hilbert matrix," *Mathematics of Computation*, vol. 18, pp. 141–143, Jan. 1964.

[68] Lee, K. and Gardner, P., "Comparison of different adaptation algorithms for adaptive digital predistortion based on edge standard," in *Microwave Symposium Digest, IEEE MTT-S International*, vol. 2, pp. 1353 –1356, 2001.

[69] Leipnik, R., "The effect of instantaneous nonlinear devices on cross-correlation," *IEEE Trans. on Inform. Theory*, vol. 4, pp. 73–76, June 1958.

[70] Levy, Y., Karam, G., and Sari, H., "Adaptation of a digital predistortion technique based on intersymbol interpolation," in *Global Telecommunications Conference, 1995. GLOBECOM '95., IEEE*, vol. 1, pp. 145–150, Nov. 1995.

[71] Liu, C., Xiao, H., Wu, Q., and Li, F., "Spectrum analysis of nonlinear distortion of RF power amplifiers for wireless signals," in *proc. International Conference on Communication Technology Proceedings (ICCT)*, vol. 2, pp. 1468–1471, Apr. 2003.

[72] Maas, S. A., *Nonlinear Microwave Circuits*. Piscataway, NJ: IEEE Press, 1997.

[73] Massey, J. L. and Neeser, F. D., "Proper complex random processes with applications to information theory," *IEEE Trans. on Inform. Theory*, vol. 39, pp. 1293–1302, July 1993.

[74] Mathews, V. J., "Orthogonalization of correlated Gaussian signals for Volterra system identification," *IEEE Signal Processing Lett.*, vol. 2, pp. 188–190, Oct. 1995.

[75] Mathews, V. J. and Sicuranza, G. L., *Polynomial Signal Processing*. John Wiley & Sons, 2000.

[76] McGee, W., "Circularly complex Gaussian noise–A price theorem and a Mehler expansion (corresp.)," *IEEE Trans. on Inform. Theory*, vol. 15, pp. 317–319, Mar. 1969.

[77] McGraw, D. and Wagner, J., "Elliptically symmetric distributions," *IEEE Trans. on Inform. Theory*, vol. 14, pp. 110–120, Jan. 1968.

[78] Midya, P., "Polynomial predistortion linearizing device, method, phone and base station." United States Patent No. 6,236,837, July 30, 1998.

[79] Midya, P. and Grosspietsch, J., "Scalar cost function based predistortion linearizing device, method, phone and basestation." United States Patent No. 6,240,278, July 30, 1998.

[80] Morris, K. and Kenington, P., "Power amplifier linearisation using predistortion techniques," in *RF and Microwave Components for Communication Systems(Digest No: 1997/126), IEE Colloquium on*, Apr. 1997.

[81] Namiki, J., "An automatically controlled predistorter for multilevel quadrature amplitude modulation," *IEEE Trans. on Communications*, vol. 31, pp. 707–712, May 1983.

[82] Nelles, O. and Isermann, R., "Identification of nonlinear dynamic systems classical methods versus radial basis function networks," in *Proceedings of the American Control Conference, 1995.*, vol. 5, pp. 3786–3790, 1995.

[83] Ochiai, H. and Imai, H., "Performance analysis of deliberately clipped OFDM signals," *IEEE Journal on Selected Areas in Communications*, vol. 18, pp. 2270–2277, Nov. 2000.

[84] Ochiai, H. and Imai, H., "Performance of the deliberate clipping with adaptive symbol selection for strictly band-limited OFDM systems," *IEEE Trans. Communications*, vol. 50, pp. 89–101, Jan. 2002.

[85] Papoulis, A., "Comments on 'An extension of Price's theorem' by McMahon, E. L.," *IEEE Trans. on Inform. Theory*, vol. 11, pp. 154–154, Jan. 1965.

[86] Papoulis, A., *Probability, Random Variables, and Stochastic Processes*. McGraw-Hill, 1991.

[87] Price, R., "A useful theorem for nonlinear devices having Gaussian inputs," *IEEE Trans. on Inform. Theory*, vol. 4, pp. 69–72, June 1958.

[88] Proakis, J. G., *Digital Communications*. McGraw-Hill, 3rd ed., 1995.

[89] Qian, H., Raich, R., and Zhou, G. T., "On the benefits of deliberately introduced baseband nonlinearities in communication systems." accepted to *Proc. IEEE Intl. Conf. Acoust., Speech, Signal Processing*, May 2004.

[90] Raab, F. H., Asbeck, P., Cripps, S., Kenington, P. B., Popovic, Z. B., Pothecary, N., Sevic, J. F., and Sokal, N. O., "Power amplifiers and transmitters for RF and microwave," *IEEE Trans. on Microwave Theory and Techniques*, vol. 50, pp. 814–826, Mar. 2002.

[91] Raich, R., Qian, H., and Zhou, G. T., "Digital baseband predistortion of nonlinear power amplifiers using orthogonal polynomials," in *Proc. IEEE Intl. Conf. Acoust., Speech, Signal Processing*, vol. 6, pp. 689–692, Apr. 2003.

[92] Raich, R., Qian, H., and Zhou, G. T., "Optimization of SNDR for amplitude limited nonlinearities." submitted to *IEEE Trans. Communications*, Dec. 2003.

[93] Raich, R., Qian, H., and Zhou, G. T., "Orthogonal polynomials for power amplifier modeling and predistorter design." submitted to *IEEE Trans. on Vehicular Technology*, Jan. 2003.

[94] Raich, R., Qian, H., and Zhou, G. T., "Signal to noise and distortion ratio considerations for nonlinear communication channels." submitted to *IEEE 6th CAS Symposium on Emerging Technologies: Frontiers of Mobile and Wireless Communication*, Jan. 2004.

[95] Raich, R. and Zhou, G. T., "On the modeling of memory nonlinear effects of power amplifiers for communication applications," in *Proc. 10th IEEE DSP Workshop*, (Pine Mountain, GA), pp. 7–10, Oct. 2002.

[96] Raich, R. and Zhou, G. T., "Orthogonal polynomials for complex Gaussian processes." accepted to *IEEE Trans. Signal Processing*, Oct. 2003.

[97] Raich, R. and Zhou, G. T., "Statistical analysis of a bandpass nonlinearity with nonstationary input." submitted to *IEEE Trans. Signal Processing*, Nov. 2003.

[98] Raich, R. and Zhou, G. T., "Theory and applications of orthogonal polynomials for Gaussian input," in *Proc. IEEE Workshop on Statistical Signal Processing*, (St. Louis, Missouri), pp. 97–100, Sept. 2003.

[99] Raich, R. and Zhou, G. T., "Spectral analysis for bandpass nonlinearity with cyclostationary input." accepted to *Proc. IEEE Intl. Conf. Acoust., Speech, Signal Processing*, May 2004.

[100] Raich, R. and Zhou, T., "Analyzing spectral regrowth of QPSK and OQPSK signals," in *Proc. IEEE Intl. Conf. Acoust., Speech, Signal Processing*, vol. 4, (Salt Lake City, Utah), pp. 2673–2676, May 2001.

[101] Reed, I. S., "On a moment theorem for complex Gaussian processes," *IRE Trans. on Information Theory*, vol. 8, pp. 194–195, Apr. 1962.

[102] Rey, C., "Adaptive polar work-function predistortion," *IEEE Trans. on Microwave Theory and Techniques*, vol. 47, pp. 722–726, Jun. 1999.

[103] Rinne, J. and Renfors, M., "The behavior of orthogonal frequency division multiplexing signals in an amplitude limiting channel," in *Proc. IEEE International Conference on Communications*, vol. 1, pp. 381–385, May 1994.

[104] Saleh, A. M., "Frequency-independent and frequency-dependent nonlinear models of TWT amplifiers," *IEEE Trans. on Communications*, vol. COM-29, pp. 1715–1720, November 1981.

[105] Samukic, A., "Umts universal mobile telecommunications system: development of standards for the third generation," in *Proc. of Global Telecommunications Conference (GLOBECOM 98)*, vol. 4, pp. 1976 –1983, 1998.

[106] Schetzen, M., *The Volterra and Wiener Theories of Nonlinear Systems*. Malabar, Florida: Krieger Publishing Company, Inc., January 1980.

[107] Shamai, S. and Bar-David, I., "The capacity of average and peak-power-limited quadrature Gaussian channels," *IEEE Trans. on Inform. Theory*, vol. 41, pp. 1060–1071, July 1995.

[108] Shi, Q., "OFDM in bandpass nonlinearity," *IEEE Trans. Consumer Electronics*, vol. 42, pp. 253–258, Aug. 1996.

[109] Smith, J. G., "The information capacity of amplitude- and varianced-constrained scalar Gaussian channel," *Information and Control*, vol. 18, pp. 203–219, Apr. 1971.

[110] Smith, R. B., "Two theorems on inverse of finite segments of the generalized Hilbert matrix," *Mathematical Tables and Other Aids to Computation*, vol. 13, pp. 41–43, January 1959.

[111] Spooner, C. M. and Gardner, W. A., "The cumulant theory of cyclostationary time-series. ii. development and applications," *IEEE Trans. Signal Processing*, vol. 42, pp. 3409–3429, Dec. 1994.

[112] Stapleton, S. P. and Costescu, F. C., "An adaptive predistorter for a power amplifier based on adjacent channel emissions," *IEEE Trans. on Vehicular Technology*, vol. 41, pp. 49–56, Feb. 1992.

[113] Stapleton, S. P., Kandola, G. S., and Cavers, J. K., "Simulation and analysis of an adaptive predistorter utilizing a complex spectral convolution," *IEEE Trans. on Vehicular Technology*, vol. 41, pp. 387–394, Nov. 1992.

[114] Tellado, J., *Multicarrier Modulation with Low PAR - Applications to DSL and Wireless*. New York, MA: Kluwer Academic Publishers, 2002.

[115] Tellado, J., Hoo, L., and Cioffi, J., "Maximum-likelihood detection of nonlinearly distorted multicarrier symbols by iterative decoding," *IEEE Trans. on Communications*, vol. 41, pp. 1060–1071, July 1995.

[116] Tsimbinos, J. and Lever, K. V., "Nonlinear system compensation based on orthogonal polynomial inverses," *IEEE Trans. on Circuits and Systems I: Fundamental Theory and Application*, vol. 48, pp. 406–417, Apr. 2001.

[117] van den Bos, A., "Price's theorem for complex variates," *IEEE Trans. on Inform. Theory*, vol. 42, pp. 286–287, Jan. 1996.

[118] Vuolevi, J. H. K., Rahkonen, T., and Manninen, J. P. A., "Measurement technique for characterizing memory effects in RF power amplifiers," *IEEE Trans. on Microwave Theory and Techniques*, vol. 49, pp. 1383–1388, Aug. 2001.

[119] Wang, T. and Brazil, T. J., "Using volterra mapping based behavioural models to evaluate ACI and cross modulation in CDMA communication systems," in *High Frequency Postgraduate Student Colloquium, 2000*, pp. 102–108, 2000.

[120] Wang, X., Tjhung, T., and Wu, Y., "On the SER and spectral analyses of A-law companded multicarrier modulation," *IEEE Trans. Vehicular Technology*, vol. 52, pp. 1408–1412, Sept. 2003.

[121] Weiner, D., Spina, J., and Fitch, A., "On the correlation function of signal plus Gaussian noise passed through nonlinear devices," *IEEE Trans. on Inform. Theory*, vol. 17, pp. 613–614, Sept. 1971.

[122] Wise, G., Traganitis, A., and Thomas, J., "The effect of a memoryless nonlinearity on the spectrum of a random process," *IEEE Trans. on Inform. Theory*, vol. 23, pp. 84–89, Jan. 1977.

[123] Wu, Q., Testa, M., and Larkin, R., "Linear RF power amplifier design for CDMA signals," in *proc. IEEE MTT-S International Microwave Symposium Digest*, vol. 2, pp. 851–854, June 1996.

[124] Wu, Q., Testa, M., and Larkin, R., "On design of linear RF power amplifier for CDMA signals," *International Journal of RF and Microwave Computer-Aided Engineering*, vol. 8, pp. 283–292, Dec. 1998.

[125] Wu, Q., Xiao, H., and Li, F., "Linear RF power amplifier design for CDMA signals: A spectrum analysis approach," *Microwave Journal*, Dec. 1998.

[126] Xiao, H., Wu, Q., and Li, F., "Spectrum approach to the design of RF power amplifier for CDMA signals," in *Proc. 9th IEEE Signal Processing Workshop on Statistical Signal and Array Processing*, pp. 132–135, Sept. 1998.

[127] Yang, T., Reed, J. H., and Hsia, T. C., "Spectral correlation of BPSK and QPSK signals in a nonlinear channel with AM/AM and AM/PM conversions," in *proc. IEEE International Conference on Communications*, vol. 2, pp. 627–632, June 1992.

[128] Yi, S.-J., Nam, S., Oh, S.-H., and Han, J.-H., "Prediction of a CDMA output spectrum based on intermodulation products of two-tone test," *IEEE Trans. on Microwave Theory and Techniques*, vol. 49, pp. 938–946, May 2001.

[129] Zhou, G. T. and Kenney, J. S., "Predicting spectral regrowth of nonlinear power amplifiers," *IEEE Trans. on Communications*, vol. 50, pp. 718–722, May 2002.

[130] Zhou, G. T., Qian, H., Ding, L., and Raich, R., "On baseband representation of passband nonlinearities." submitted to *Communications Letters*, Nov. 2003.

[131] Zhou, G. T. and Raich, R., "Closed-form expressions of output power spectrum for nonlinear power amplifiers with or without memory," in *Proc. IEEE - EURASIP Workshop on Nonlinear Signal and Image Processing*, (Grado, Italy), June 2003.

[132] Zhou, G. T. and Raich, R., "Spectral analysis of polynomial nonlinearity with applications to RF power amplifiers." accepted to *EURASIP Journal on Applied Signal Processing, special issue on nonlinear signal and image processing*, Dec. 2003.

[133] Zhou, G. T. and Raich, R., "Spectral analysis of polynomial nonlinearity with applications to RF power amplifiers." accepted to *EURASIP Journal on Applied Signal Processing, special issue on nonlinear signal and image processing*, Dec. 2003. See also *Proc. IEEE - EURASIP Workshop on Nonlinear Signal and Image Processing*, (Grado, Italy), June 2003.

[134] Zhu, A. and Brazil, T. J., "An adaptive Volterra predistorter for the linearization of RF high power amplifiers," in *Microwave Symposium Digest, 2002 IEEE MTT-S International*, vol. 1, pp. 461 – 464, 2002.

[135] Ziperovich, P., "Performance degradation of PRML channels due to nonlinear distortions," *IEEE Trans. Magnetics*, vol. 27, pp. 4825–4827, Nov. 1991.

VITA

Raviv Raich was born in Israel. He received both the B.Sc. and M.Sc. degrees in electrical engineering from Tel-Aviv University, Tel-Aviv, Israel, in 1994 and 1998, respectively. Since 2000, he has been pursuing the Ph.D. degree in electrical engineering at Georgia Institute of Technology, Atlanta, Georgia, USA.

From 1994 to 1997, he served as an electronic engineer in the Israeli Defense Force. During 1998 he was with the Department of Electrical Engineering—Systems, Tel-Aviv University. During the same year, he was a consultant for Tadiran Electronic Systems, Ltd., Holon, Israel. During 1999 and 2000, he worked as a researcher with the communications team, Industrial Research Ltd., Wellington, New Zealand. His main research interests are predistortion linearization of nonlinear power amplifiers for wireless applications, statistical signal processing for communications, estimation and detection theory.