TAMPERE UNIVERSITY OF TECHNOLOGY
Degree Programme in Information Technology

AHMET HASIM GOKCEOGLU

PERFORMANCE OF ADAPTIVE FEEDFORWARD METHODS IN WIDEBAND POWER AMPLIFIER LINEARIZATION

Master of Science Thesis

Examiners: Professor Mikko Valkama
Professor Markku Renfors

Examiner and topic approved in the Computing and Electrical Engineering Faculty Council meeting on 4 March 2009

Abstract

TAMPERE UNIVERSITY OF TECHNOLOGY
Master's Degree Programme in Information Technology
GOKCEOGLU, AHMET: Performance of adaptive feedforward methods in wideband power amplifier linearization
Master of Science Thesis, 102 pages, 8 Appendix pages
October 2009
Major: Communication Engineering
Examiners: Professor Mikko Valkama and Professor Markku Renfors

The waveforms of the emerging communication systems aiming at high data rates and high spectral efficiencies are becoming more and more complex and thereon also sensitive to many implementation nonidealities. One good example is multicarrier waveforms having high peak-to-average power ratio (*PAPR*) and thus high sensitivity to any nonlinearity in the radio components such as transmitter power amplifier (PA). On the other hand, when emphasizing power efficiency, power amplifiers operate typically close to their saturation region and are thus heavily nonlinear. Such power efficient operation is especially important in handheld terminal equipment but also on the base-station side of, e.g., cellular systems. Thus efficient linearization of power amplifiers is seen as one essential element in obtaining proper compromise between power efficiency and spectral efficiency in the emerging communications systems.

Various approaches have been proposed and demonstrated for power amplifier linearization in the literature, among which the so-called feedforward approach is one of the most established ones. Feedforward linearizer builds on two stages, the so-called signal cancellation loop (*SCL*) and error cancellation loop (*ECL*), which aim at isolating and subtracting the distortion created by the PA from the overall linearizer output. In practice, however, the operation of feedforward linearizer is susceptible, e.g., to any parameter mismatches in the *SCL* and *ECL* components. Also the characteristics of the PA can change in time when the operating conditions or operating point are varying.

Adaptive feedforward linearizer is a promising linearization method that is able to adjust the signal cancellation loop and error cancellation loop coefficients to minimize the effects of component mismatch and to track the possible variations in the characteristics of the circuit components. In this thesis, the performance of least-mean squares (*LMS*) adaptation in terms

of steady-state *SCL* and *ECL* coefficient behavior and the corresponding suppression of inter-modulation distortion (*IMD*) are studied, covering both memoryless PA's and PA's with memory. The analysis shows that the adaptive feedforward linearizer is independent from memory effects in terms of *IMD* suppression even though convergence of *SCL* coefficient is affected by memory. An estimate for the final *IMD* suppression is driven for the memoryless amplifier model and it also holds for the model with memory under reasonable assumptions. The findings of the analysis are supported with simulation results.

Preface

This thesis is an outcome of research carried out in the projects “Advanced Techniques for RF Impairment Mitigation in Future Wireless Radio Systems” and “Understanding and Mitigation of Analog RF Impairments in Multi-antenna Transmission Systems” at the Department of Communications Engineering, Tampere University of Technology, Finland. The former is funded by the Finnish Funding Agency for Technology and Innovation (TEKES) and the latter by the Technology Industries of Finland Centennial Foundation.

I would like to thank my supervisor Professor Mikko Valkama who gave me the opportunity to join his research group and get involved in these projects. His experience and fruitful discussions definitely had a great role on the smooth progress of this work. I enjoyed every moment of being part of such a well organized yet equally friendly working environment lead by him.

A lot of credit goes to my colleagues, MSc. Ali Shahed Hagh Ghadam and MSc. Tobias Hidalgo, for sparing a great deal of time for useful discussions and sharing their experience in vast amount of fields from programming to mathematics.

I would like to thank the head of the department, Professor Markku Renfors, who puts incredible amount of effort for establishing an inspiring atmosphere and cooperative working environment within the department.

I had the opportunity to meet with many nice people from different cultures during my stay in Tampere whom I should thank for their sincere friendship. Among those, I have to mention one very special name, Emmi Juulia Keskisarja, who has the greatest influence on me. It was her spiritual support that eased the difficulties of living in a very different culture.

Finally, I owe my deepest gratitude to my parents Bayram and Nevin Gökceoglu. They have always been great guides in life yet respecting and supporting my own ideas and decisions. Whatever I succeed in my life will always be partly dedicated to these two caring, open-minded and invaluable people.

Ahmet Hasim Gökceoglu
Tampere, September 2009

Contents

List of Abbreviations

ADC	analog to digital converter
AM/AM	amplitude to amplitude mapping
AM/PM	amplitude to phase mapping
AWGN	additive white Gaussian noise
BER	bit error rate
CP	cyclic prefix
DAC	digital to analog converter
DC	direct current
DPU	digital pre-distorter unit
ECL	error cancellation loop
FFT	fast Fourier transform
GSM	global system for mobile communications
IBO	input power back off
IFFT	inverse fast Fourier transform
IIP	input intercept point
IMD	inter-modulation distortion
INM	instantaneous nonlinear mapping
ISI	inter-symbol interference
LINC	linear amplification with nonlinear components
LMS	least mean square
LTE	long term evolution
MIMO	multiple input multiple output
MMSE	minimum mean squared error

OFDM	orthogonal frequency division multiplexing
PA	power amplifier
PAPR	peak-to-average power ratio
P/S	parallel to serial converter
PSD	power spectral density
PSK	phase shift keying
QAM	quadrature amplitude modulation
RF	radio frequency
S/P	serial to parallel converter
SCL	signal cancellation loop
SER	symbol error rate
SIMDR	signal-to-inter-modulation distortion ratio
SNR	signal-to-noise ratio
SSPA	solid state power amplifier
STNR	signal-to-total-noise ratio
WLAN	wireless local area network

List of Principal Symbols

α , $\alpha[n]$	*SCL* coefficient (fixed-time varying)
$\alpha_{A,}\ \beta_A$	AM/AM parameters of *Saleh* model
$\alpha_{\phi,}\ \beta_\phi$	AM/PM parameters of *Saleh* model
α_o	complex constant in *Bussgang* model
α_{MMSE}	optimum Wiener coefficient for *SCL*
α_{MMSE}^{mem}	optimum Wiener coefficient for *SCL* (Wiener-Hammerstein memory)
α_{opt}	optimum linearizer coefficient for *SCL*
β , $\beta[n]$	*ECL* coefficient (fixed-time varying)
β_{MMSE}	optimum Wiener coefficient for *ECL*
β_{MMSE}^{mem}	optimum Wiener coefficient for *ECL* (Wiener-Hammerstein memory)
β_{opt}	optimum linearizer coefficient for *ECL*
ε_α, $\varepsilon_\alpha[n]$	mismatch of *SCL* coefficient (α_{opt} - α and α_{opt} - $\alpha\ [n]$)
$\varepsilon_\alpha^{MMSE}[n]$	mismatch between $\alpha\ [n]$ and α_{MMSE}
$\varepsilon_{\alpha,mem}^{MMSE}[n]$	mismatch between $\alpha\ [n]$ and α_{MMSE}^{mem}
ε_β, $\varepsilon_\beta[n]$	mismatch of *SCL* coefficient (β_{opt} - β and β_{opt} - $\beta\ [n]$)
$\varepsilon_\beta^{MMSE}[n]$	mismatch between $\beta\ [n]$ and β_{MMSE}
$\varepsilon_{\beta,mem}^{MMSE}$	mismatch between $\beta\ [n]$ and β_{MMSE}^{mem}
ε_τ	delay mismatch between upper and lower branches of feedforward loops
κ	linear amplitude gain in *Rapp* model
$\sigma_{\tilde{d}}^2$	variance of *IMD*
$\sigma_{\tilde{x}}^2$	variance of amplifier input
$\phi(t)$	instantaneous phase of the signal waveform before amplifier

a_n	*IMD* spectrum coefficient for the $2n+1$'th fold of amplifier input spectrum
A_m	m 'th symbol in single carrier system
A_o	clipping level in *Rapp* model
$A(t)$	instantaneous envelope of the signal waveform before amplifier
b, $\hat{b}$	sent and received bits
$b_w(t)$, $b_w[n]$	Wiener-filter before amplifier (continuous-discrete time)
$b_h(t)$, $b_h[n]$	Hammerstein-filter after amplifier (continuous-discrete time)
$b_{wh}(t)$, $b_{wh}[n]$	cascade of Wiener and Hammerstein filters (continuous-discrete time)
C_A	coupler at the output of amplifier
C_e	coupler that couples amplifier input and attenuated output into error signal
C_i	coupler at the input of amplifier
C_o	coupler at the output of linearizer
$\tilde{c}(t)$	impulse response of baseband channel
$\tilde{c}_{2p+1}$	coefficient of order $2p+1$ for baseband polynomial nonlinearity
D_k^m	*IMD* noise on k 'th subcarrier of m 'th symbol
D_{NL}	distortion terms at the output of memoryless amplifier
D_{NL}^w	distortion terms in the presence of Wiener memory model
D_{NL}^h	distortion terms in the presence of Hammerstein memory model
D_{NL}^{wh}	distortion terms in the presence of Wiener-Hammerstein memory model
$\tilde{d}(t)$, $\tilde{d}[n]$	baseband representation of *IMD* (continuous-discrete time)
$e(t)$	bandpass error signal in feedforward circuit
$\tilde{e}(t)$, $\tilde{e}[n]$	baseband equivalent error signal in feedforward circuit (continuous-discrete time)
$\tilde{e}^{mem}[n]$	baseband error signal of feedforward circuit in the presence of Wiener-Hammerstein memory
$E[.]$	expected value operator
f	frequency in Hertz
$F(.)$, $F_B(.)$	predistorter function (bandpass-baseband implementation)

$f_A(.)$	probability density function of signal envelope
f_c	center frequency
$G(.)$	instantaneous input-output mapping of amplifier
G_e	power gain of error amplifier in feedforward circuit
$g_A(.)$	AM/AM mapping
$g_\phi(.)$	AM/PM mapping
$\tilde{h}_{2p+1}$	$2p+1$'th kernel of baseband Volterra series
$H(f)$	frequency selective cancellation between upper and lower branches of feedforward circuit
IMD_{sup}	amount of *IMD* suppression from amplifier to linearizer output
J_α, J_β	cost functions to adapt α and β respectively
$\nabla_\alpha J[n]$, $\nabla_\beta J[n]$	gradients of cost functions that are used to adapt α and β respectively
k	subcarrier index
K	order of Wiener filter
L	order of Hammerstein filter
L_c	power loss of attenuator in feedforward circuit
$L_n^{(1)}(x)$	Laguerre function of the first form
m	symbol index
M	order of Wiener-Hammerstein filter
n	discrete time index
N	number of subcarriers
N_k^m	AWGN noise due to channel appearing on k'th subcarrier of m'th symbol
$\tilde{n}(t)$	complex AWGN noise
p	smoothness factor of Rapp model, nonlinearity order in polynomial and Volterra series nonlinearity model
$p(t)$	pulse-shaping signal
$p_{\tilde{i}\tilde{r}^*}[j]$	the cross-correlation of input to Wiener filter and reference signal
P_{IMDa}	power of *IMD* at the output of amplifier
P_{IMDo}	power of *IMD* at the output of linearizer

$P_{\tilde{x}}$	power of the amplifier input signal
$r(t)$	matched filter
$R_{\tilde{i}\tilde{i}^*}[j]$	the auto-correlation of input of Wiener filter
$R_{\tilde{x}\tilde{x}^*}(\tau)$, $R_{\tilde{x}\tilde{x}^*}[m]$	the auto-correlation of amplifier input (continuous-discrete lag)
$R_{\tilde{y}\tilde{x}^*}(\tau)$, $R_{\tilde{y}\tilde{x}^*}[m]$	the cross-correlation of amplifier input and output (continuous-discrete lag)
$R_{\tilde{d}}[m]$	the auto-correlation of *IMD*
$S_{\tilde{d}}(f)$	power spectral density of *IMD*
S_k^m	m 'th received sub-symbol of k 'th subcarrier
$\tilde{s}(t)$, $\tilde{s}[n]$	received baseband signal (continuous-discrete time)
t	continuous time variable
T_b	time duration between two consecutive samples of one OFDM symbol
T_s	symbol duration
$x(t)$, $x[n]$	bandpass amplifier input (continuous-discrete time)
$\tilde{x}(t)$, $\tilde{x}[n]$	baseband amplifier input (continuous-discrete time)
X_k^m	m 'th transmitted sub-symbol of k 'th subcarrier in multicarrier system
$y(t)$, $y[n]$	bandpass amplifier output (continuous-discrete time)
$\tilde{y}(t)$, $\tilde{y}[n]$	baseband amplifier output (continuous-discrete time)
$\tilde{y}^{mem}[n]$	baseband output of amplifier with Wiener-Hammerstein memory
$z(t)$, $z[n]$	bandpass amplifier output (continuous-discrete time)
$\tilde{z}(t)$, $\tilde{z}[n]$	baseband amplifier output (continuous-discrete time)
$\tilde{z}^{mem}[n]$	baseband linearizer output for amplifier with Wiener-Hammerstein memory

Chapter 1

Introduction

There is an increasing demand for having wireless access with high data rates to many different types of services. There has been great effort and development on service providers side to meet those demands and provide service to more and more users. Starting from Global System for Mobile Communications (*GSM*) [25] with radio channels of 200 kHz, the new generation 3G networks offer 5 MHz of bandwidth which is planned to be increased upto 20 MHz by long term evolution (*LTE*) [1] whereas wireless local area network (*WLAN*) supported by multiple input multiple output (*MIMO*) feature aims 40 MHz (IEEE 802.11n). This entails different choice of transmission schemes and signal waveforms such as orthogonal frequency division multiplexing (*OFDM).* One consequence of such choice is to end-up with signal waveforms that have high amplitude variations causing high peak-to-average power ratio (*PAPR).*

Amplification of the signal to be transmitted is one of the indispensable processes of wireless communications which is shown in Figure 1-1 as the last stage before antenna.

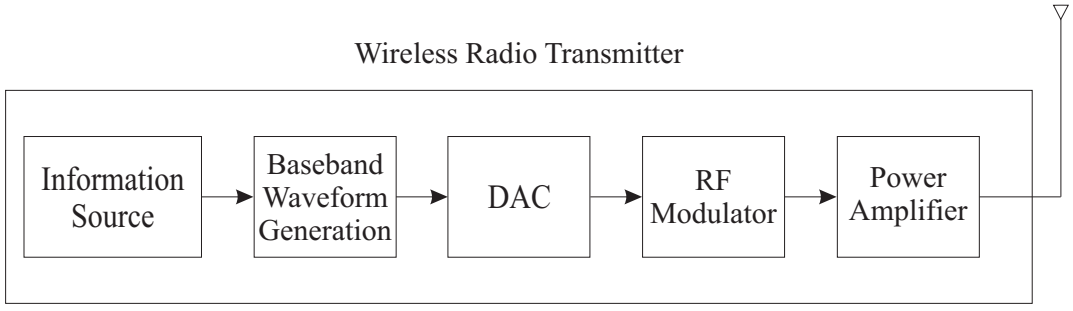


Figure 1-1. Conceptual block diagram of wireless radio transmitter.

Real world amplifiers performing this task are inherently nonlinear when their input-output relation is considered. In other words, different amplitude levels are experiencing different gains which definitely introduce distortion compared to an ideal amplification where all the

amplitude levels are scaled with the same gain. The significance of the distortion increases as *PAPR* increases since the experienced gains will vary more among different amplitude levels.

In linear time-invariant systems, the output spectrum cannot contain any frequency components that are not present in the input, which is in general not true for nonlinear systems. These additional components can be categorized under two types of distortion, referred to as harmonic and inter-modulation distortion (*IMD*) [22]. The components at integer multiples of input frequencies are called harmonic distortion, while the latter refers to all other products which are generally the linear combination of those input frequencies. Harmonic distortion can usually be eliminated rather easily with ordinary linear filtering methods, whereas the removal of *IMD* is not necessarily that trivial. There are two problematic cases stemming from *IMD* which can be categorized as spectral re-growth and in-band distortion [4], [6], [22]. Spectral re-growth refers to the expansion of the original signal band due to frequency components falling within the very close vicinity of the useful signal band. Satellite communication systems as well as the mobile communication systems listed above strictly define the allowable interference with adjacent portions of the spectrum. In other words there should be a limitation on the power of those additional frequency components originating from the nonlinear nature of the amplifier. On the other hand, in-band distortion which is another consequence of *IMD* referring to frequency components that fall within the band of signal is a source of degradation from individual link point of view such as increased symbol error rate (*SER*) or bit error rate (*BER*) [4], [15]. Backing-off the average power of amplifier input to a more linear region is one way to prevent these two shortcomings of nonlinear amplification.

The efficiency of the amplifier is another important concern which basically defines a percentage for how much of the direct current (DC) power it consumes is used for amplification. For instance in [3], a definition for efficiency is given as

$$PA_{eff} \triangleq \frac{P_{RF,out} - P_{RF,in}}{P_{DC}}$$

where $P_{RF,out}$, $P_{RF,in}$ and P_{DC} indicates the output, input and DC powers respectively. A power efficient operation is desirable for long battery life in mobile terminals, and for lower cost operation in base stations. However, the demands for efficiency conflict with the demands for linear operation since amplifiers are more efficient in the nonlinear operation region. Despite its simplicity, backing-off the input power level is not the best possible solution due to its low power efficiency. Therefore more advanced linearization methods are

proposed in literature which are offering better trade-off between power efficiency and linearity.

Feedforward linearization is one of the oldest and widely used linearization techniques especially in wideband systems where amplifiers tend to show memory effects, i.e. the output is a function of the past instances of input in addition to the instantaneous mapping. However, the basic principle of feedforward linearizer which is based on the separation of amplifier output into linearly amplified and additive distortion terms gives rise to a linearization that is independent of amplifier modeling. In the ideal operation, the identified distortion (whether it includes the memory effects or not) is then subtracted from the amplifier output leaving the linearly amplified term as the only signal to be transmitted. However, delay and coefficient mismatches and non-ideal error amplifier operation are the limiting factors for the cancellation of distortion terms in feedforward system [11], [12], [22], [34].

Adaptive feedforward linearization is offered to track aging effects of amplifier and changes in the parameters of other analog components present in the circuitry [11], [12], [28]. Thereby the accuracy of coefficients is improved yielding a better linearization performance. This thesis focuses on the performance of least mean-square (*LMS*) adaptation of *signal cancellation loop* (*SCL*) and *error cancellation loop* (*ECL*) coefficients in feedforward linearizer. As a motivation and background, Chapter 2 introduces most common power amplifier models with and without memory. The time and frequency domain modeling of the distortion introduced to waveforms due to memoryless nonlinearities presented in Chapter 2 are discussed in Chapter 3. Following that, the most popular linearization methods that have taken attention in real world implementations are discussed in Chapter 4 with emphasis on basic feedforward linearizer. The core of this thesis is presented in Chapter 5 where the optimum Wiener coefficients for *SCL* and *ECL* are derived for memoryless amplifiers as well as amplifiers with Wiener-Hammerstein memory model. In the same chapter, the achievable *IMD* suppression is analyzed in terms of the chosen step sizes. Then Chapter 6 illustrates and verifies the foundations of Chapter 5 with computer simulations. Finally conclusions are drawn in Chapter 7.

Chapter 2

Power Amplifier Modeling

In this chapter, essential mathematical models that are used to describe the behavior of power amplifiers are discussed. As a physical component, power amplifier typically has nonlinear relation between its input and output. Memoryless models, as well as the models with memory that are widely discussed in literature to describe such nonlinearities, are presented in this chapter.

2.1 Memoryless Nonlinearity Models

Power amplifiers are considered as bandpass nonlinearities, implying a nonlinear mapping between the real-valued bandpass input and output of the amplifier. When arguing about memoryless nonlinearities, this mapping is assumed to be between only the instantaneous values of the input and output of the nonlinearity and can be expressed as

$$y(t) = G(x(t)) \tag{2.1}$$

where $y(t)$ and $x(t)$ are the real-valued bandpass output and input of the nonlinearity at time t, respectively, and $G(.)$ is generally a nonlinear function. Let's assume that the input to the nonlinearity is of the general bandpass signal form

$$x(t) = A(t)\cos(2\pi f_c t + \phi(t)) \tag{2.2}$$

where $A(t)$ and $\phi(t)$ are the instantaneous amplitude (envelope) and phase of the input, respectively, and f_c is the center frequency. Then the widely accepted model for the corresponding output is of the form [24], [27], [30], and [38]

$$y(t) = g_A(A(t))\cos(2\pi f_c t + \phi(t) + g_\phi(A(t))) \tag{2.3}$$

where $g_A(.)$ and $g_\phi(.)$ are two separate functions. The baseband equivalent signal for (2.2) can be given as

$$\tilde{x}(t) = A(t)e^{j\phi(t)} \tag{2.4}$$

whereas the baseband equivalent signal for (2.3) is

$$\tilde{y}(t) = g_A(A(t))e^{j(\phi(t)+g_\phi(A(t)))} \tag{2.5}$$

then it is clear that $g_A(.)$ is a mapping from the amplitude of the baseband input to amplitude of the baseband output whereas $g_\phi(.)$ is a mapping from the amplitude of the baseband input to phase of the baseband output. Therefore $g_A(.)$, and $g_\phi(.)$ are often referred as AM/AM (amplitude-to-amplitude) and AM/PM (amplitude-to-phase) mapping and/or conversions.

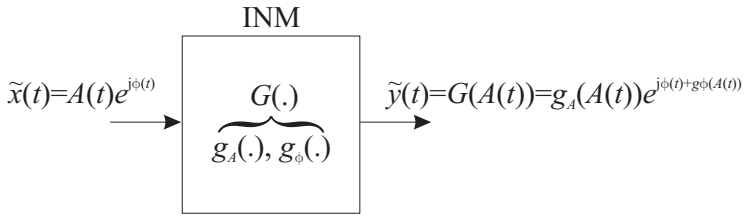


Figure 2-1. AM/AM and AM/PM characterization of a baseband equivalent memoryless nonlinearity.

In Figure 2-1, a baseband equivalent nonlinearity characterized by AM/AM and AM/PM mappings is illustrated, although there is a slight abuse of notation since g_A and g_ϕ are actually functions operating only on the amplitude of the baseband signal $\tilde{x}(t)$. The block in Figure 2-1 is referred to as instantaneous nonlinear mapping (*INM*) since the system is memoryless. This notation will be used in this thesis whenever the nonlinearity under discussion is considered to be memoryless.

The terms strictly and quasi-memoryless are also often used (e.g. [17]) to refer AM/AM and AM/PM conversions. The nonlinearity is said to be strictly memoryless when AM/PM conversion is just a constant and quasi-memoryless (nonlinearity with short term memory) when AM/PM conversion varies with $A(t)$. In order not to cause any confusion, it should be mentioned that the notion of being memoryless for a system is slightly different in signal processing and RF/Microwave literatures. A nonlinearity model characterized with AM/AM and AM/PM conversions is purely memoryless from signal processing perspective whereas it is differentiated as strictly and quasi-memoryless depending on the behavior of AM/PM conversion in RF/Microwave field.

Next some widely accepted nonlinearity models for power amplifiers that are characterized by AM/AM and AM/PM conversions will be presented, and for the simplicity of the expressions and without loss of generality we will denote $A(t) = A$.

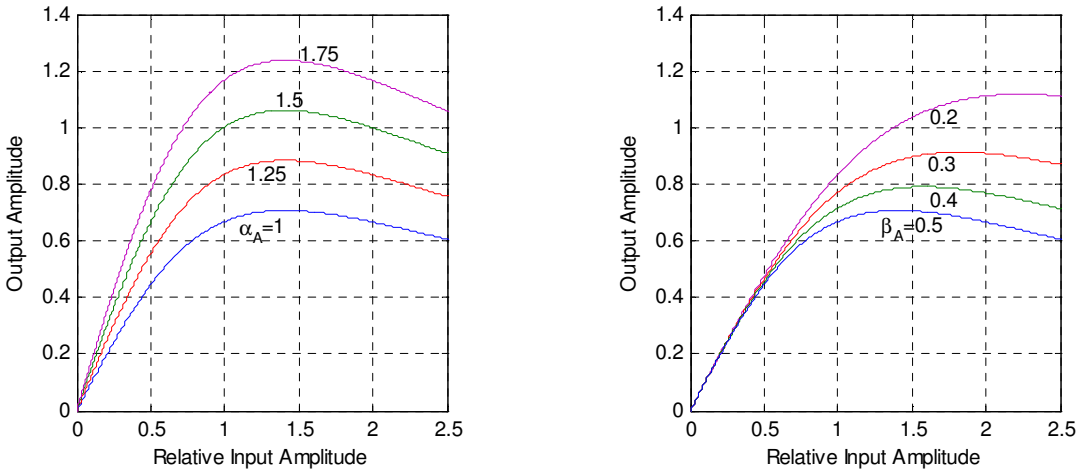
2.1.1 Saleh Model

The *Saleh* model [30] which is originally developed for traveling wave tube amplifiers (TWTA) has the following AM/AM and AM/PM characteristics

$$g_A(A) = \frac{\alpha_A A}{1 + \beta_A A^2} \tag{2.6}$$

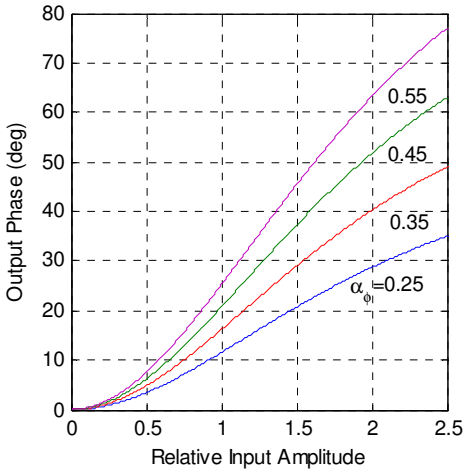
$$g_\phi(A) = \frac{\alpha_\phi A^2}{1 + \beta_\phi A^2} \tag{2.7}$$

where α_A, β_A and α_ϕ, β_ϕ are the parameter pairs characterizing the AM/AM and AM/PM conversions respectively. It is obvious that for small A, $g_A(A)$ is approximately linear in A whereas for large A, it is proportional to $1/A$ with a cofactor α_A / β_A. On the other hand, $g_\phi(A)$ is proportional to A^2 for small A with α_ϕ being the cofactor and for large A it reaches a constant level of α_ϕ / β_ϕ. AM/AM and AM/PM curves of *Saleh* model with these four parameters being varied are shown in Figure 2-2 and Figure 2-3.

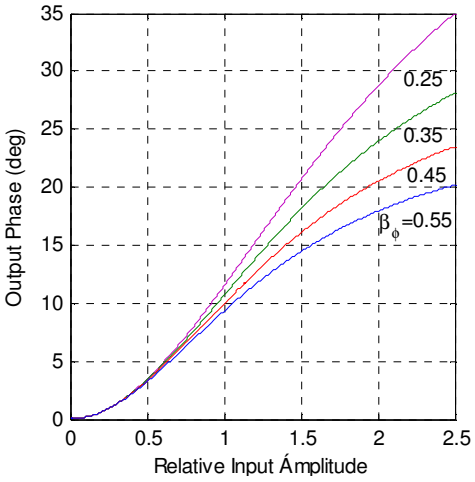


a) *b)*

Figure 2-2. AM/AM curve of Saleh model a) $\beta_A = 0.5$ *and* α_A *is varied from* 1 *to* 1.75 *b)* $\alpha_A = 1$ and β_A *varied from* 0.2 *to* 0.5.



a)



b)

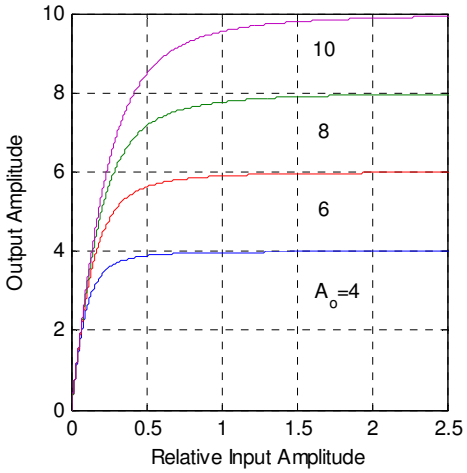
Figure 2-3. AM/PM curve of a Saleh model a) $\beta_\phi = 0.25$ *and* α_ϕ *is varied from* 0.25 *to* 0.55 *b)* $\alpha_\phi = 0.25$ *and* β_ϕ *is varied from* 0.25 *to* 0.55.

2.1.2 Rapp Model

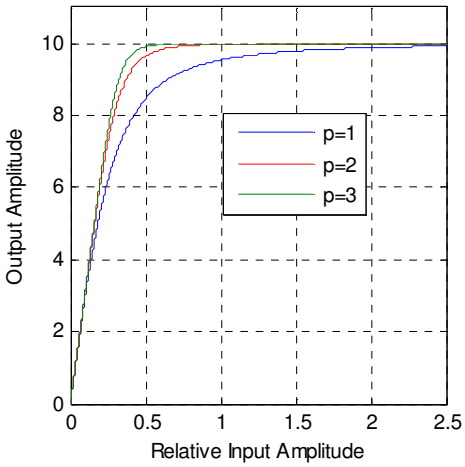
In [29], it is discussed that the behavior of the solid state power amplifiers (*SSPA*) for small input signals are more linear compared to TWTA and for large inputs the behavior is more like a clipping. Therefore *Rapp* model (also referred to as *SSPA* model) is proposed instead of *Saleh* model that has the characteristic

$$g_A(A) = \frac{\kappa A}{\left(1 + \left[\frac{\kappa A}{A_o}\right]^{2p}\right)^{1/2p}} \qquad (2.8)$$
$$g_\phi(A) = 0$$

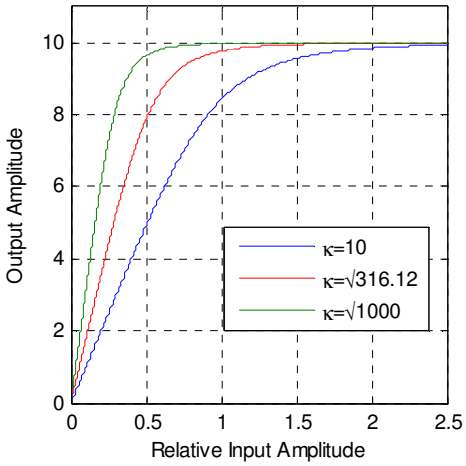
where κ is the small signal gain, A_o is the level of limiting amplitude, and p is the smoothness factor of the transition from linear region to limiting amplitude. The AM/PM conversion is assumed to be negligibly small for *SSPA*, therefore $g_\phi(A)$ is considered as zero. In Figure 2-4 a), b) and c) the dependency of the AM/AM characteristic of *Rapp* model on the parameters A_o , p , and κ is illustrated.



a)



b)



c)

Figure 2-4. SSPA model a) $p = 1$, $\kappa = \sqrt{1000}$ and A_o is varied from 4 to 10, b) $A_o = 10$, $\kappa = \sqrt{1000}$ and p is varied from 1 to 3 and, c) $A_o = 10$, $p = 1$ and κ is varied among values that correspond to 20, 25 and 30 dB power gains in the linear region.

2.1.3 Polynomial Model

The relation between the baseband equivalent input and output of the power amplifier is often modeled with polynomials. Assuming the input signal takes values from a bounded interval (such that the output is in the desired region of polynomial model), the corresponding baseband output is given in [17] and [38] as

$$\tilde{y}(t) = \tilde{x}(t)\sum_{p=0}^{P-1} \tilde{c}_{2p+1} \left|\tilde{x}(t)\right|^{2p} \tag{2.9}$$

where $\tilde{c}_{2p+1}$ are the corresponding baseband coefficients. Combining (2.4) with (2.9) , one can obtain

$$\tilde{y}(t) = e^{j\phi}\sum_{p=0}^{P-1} \tilde{c}_{2p+1} \left[A(t) \right]^{2p+1} \tag{2.10}$$

Spurious components originating from even orders are shown in [5] to be located far away from the center frequency in the bandpass model and can be filtered away, therefore only the odd terms are present in the baseband representation. The baseband polynomial model given in (2.10) can also be related to AM/AM and AM/PM mappings via

$$g_A(A) = \left| \sum_{p=0}^{P-1} \tilde{c}_{2p+1} \left[A \right]^{2p+1} \right| \tag{2.11}$$

$$g_\phi(A) = \angle \sum_{p=0}^{P-1} \tilde{c}_{2p+1} \left[A \right]^{2p+1} \tag{2.12}$$

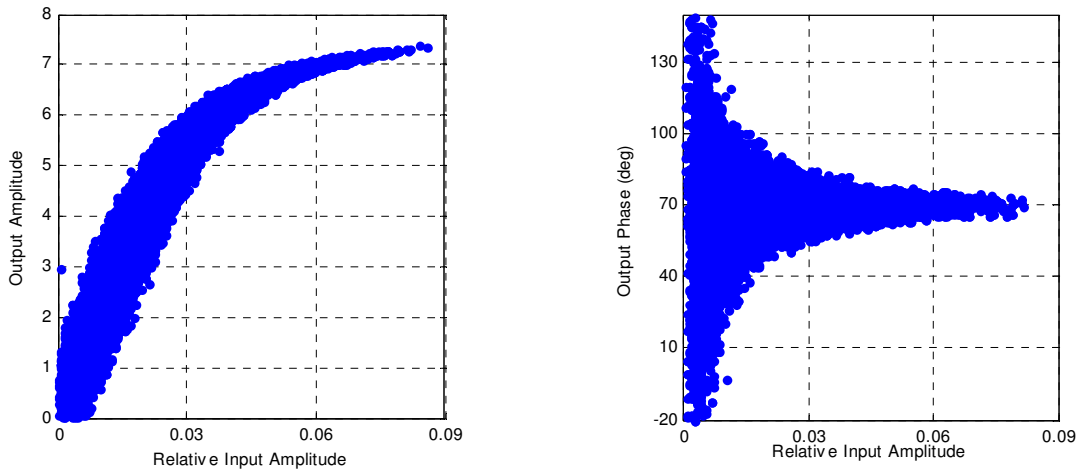
Real-valued polynomial coefficients yield a zero AM/PM mapping that corresponds to strictly memoryless case whereas complex polynomial coefficients correspond to a non-zero and non-constant AM/PM mapping implying a quasi-memoryless nonlinearity.

2.2 Nonlinearity Models with Memory

The models discussed in the previous section define a purely instantaneous mapping between the input and output of the nonlinearity. However, this might be a crude way of modeling some of real world amplifiers whose outputs depend on the past values of the input as discussed in [1], [8], [9] and [36] This phenomenon is called the memory of the amplifier which is due to both electrical and thermal effects in the underlying circuitry. The concept of memory mentioned here is different than the quasi-memory discussed in the context of AM/PM mapping. Long term memory is used to describe these electrical and thermal effects whereas AM/PM is referred as short term memory. In the presence of long term memory, it is not possible to obtain a single AM/AM and AM/PM curve to relate the input and output of the nonlinearity as shown in Figure 2-5 a) and b) respectively. Clearly Figure 2-5 a) and b) shows that for a given input amplitude value there is a set (and not a single value) of possible output

amplitude and phase values depending on the past instances of the input as a difference from the memoryless case.

Another reflection of memory effects to the amplifier characteristics is the frequency selectivity. Independent of the input power level, the response of the amplifier is not flat (although this response is different than the response of a linear system since it cannot directly be used to compute the output given the input). The most common models discussed in literature that are used to characterize nonlinearities with such long term memory effects will be presented next.



a) *b)*

Figure 2-5. The AM/AM and AM/PM relation of a nonlinearity with memory.

2.2.1 Volterra Series

One of the most powerful models that has the capability of describing a large class of nonlinearities with memory is the Volterra series with the following input and output relation

$$y(t) = c_o + \sum_{p=1}^{\infty} \int \cdots \int h_p(\tau_1, \cdots, \tau_p) \prod_{q=1}^{p} x(t-\tau_q) d\tau_1 \cdots d\tau_p \tag{2.13}$$

where c_o is a constant, and $h_p(\tau_1, \cdots \tau_p)$ is a multivariable continuous function of τ_p where $1 \le p < \infty$. These functions are referred as Volterra kernels and are used to describe the complete behavior of a nonlinear dynamic system. The baseband equivalent of (2.13) is derived in [5] and [27], [38] as

$$\tilde{y}(t) = \sum_{p=0}^{\infty} \int \cdots \int \tilde{h}_{2p+1}(\tau_1, \cdots, \tau_{2p+1}) \prod_{q=1}^{p+1} \tilde{x}(t-\tau_q) \prod_{q=p+2}^{2p+1} x^*(t-\tau_q) d\tau_1 \cdots d\tau_{2p+1} \tag{2.14}$$

where the complex baseband kernels are related to the bandpass kernels with

$$\tilde{h}_{2p+1}(t_1,\cdots,t_{2p+1}) = \left(\frac{1}{2}\right)^{2p}\binom{2p+1}{p} h_{2p+1} e^{-j2\pi f_c(\sum_{q=1}^{p+1} t_q - \sum_{q=p+2}^{2p+1} t_q)} \tag{2.15}$$

Apparently from (2.14) , only odd kernels contribute to baseband. The limits of the integrals determine the depth of the memory whereas the kernels indicate the strength. It is easy to see that, by setting $\tilde{h}_{2p+1}(t_1,\cdots,t_{2p+1}) = \tilde{c}_{2p+1}\delta(t_1,\cdots,t_{2p+1})$, i.e., there is no long term memory, we obtain the memoryless polynomial model in (2.10).

2.2.2 Wiener Model

The cascade connection of a linear system and a memoryless nonlinearity which is referred as Wiener model is one very common method for modeling nonlinearities with memory. The basic structure for such model with baseband input and output signals is shown in Figure 2-6.

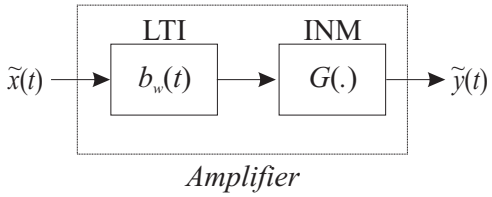


Figure 2-6 Wiener Model showing the cascade of an LTI system with memoryless nonlinearity.

The output of the system can be derived in terms of the input by first defining an intermediate signal after the linear system

$$\tilde{u}(t) = \int_{-\infty}^{\infty} b_w(\tau)\tilde{x}(t-\tau)d\tau \tag{2.16}$$

then this filtered version of the input signal is passed through memoryless nonlinear mapping

$$\tilde{y}(t) = G(\tilde{u}(t)) \tag{2.17}$$

where $G(.)$ can be one of the models presented in the previous section.

2.2.3 Hammerstein Model

Another very common model for long term memory is the Hammerstein model which is quite similar to the Wiener model with a change of order in the combination of linear system and

memoryless nonlinearity. This time the input signal is first passed through a memoryless nonlinearity and then filtered by a linear system to give the final output as seen in Figure 2-7.

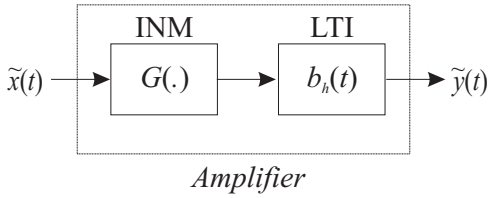


Figure 2-7. Hammerstein model showing the cascade of a memoryless nonlinearity with an LTI system.

Similar to the Wiener model we define an intermediate signal

$$\tilde{u}(t) = G(\tilde{x}(t)) \tag{2.18}$$

and then the system output can be given

$$\begin{aligned} \tilde{y}(t) &= \int_{-\infty}^{\infty} b_h(\tau)\tilde{u}(t-\tau)d\tau \\ &= \int_{-\infty}^{\infty} b_h(\tau)G(\tilde{x}(t-\tau))d\tau \end{aligned} \tag{2.19}$$

2.2.4 Wiener-Hammerstein Model

The Wiener and Hammerstein models can be combined together for a more generalized form. This model shown in Figure 2-8, first pre-filters the input signal, then the filtered signal is passed through a memoryless nonlinearity and finally a post-filtering is done to obtain the final output.

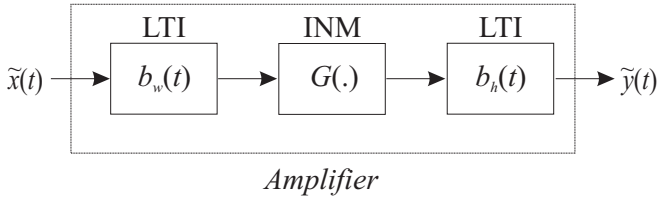


Figure 2-8. Wiener-Hammerstein model showing the cascade of two LTI systems with a memoryless nonlinearity located in between.

This time, in order to derive the input-output relation of the amplifier, we will define two intermediate signals, before and after the memoryless nonlinearity.

$$\tilde{u}_1(t) = \int_{-\infty}^{\infty} b_w(\tau)\tilde{x}(t-\tau)d\tau \tag{2.20}$$

$$\tilde{u}_2(t) = G(\tilde{u}_1(t)) \tag{2.21}$$

and the final output can be written

$$\begin{aligned}\tilde{y}(t) &= \int_{-\infty}^{\infty} b_h(\zeta)\tilde{u}_2(t-\zeta)d\zeta \\ &= \int_{-\infty}^{\infty} b_h(\zeta)G(\tilde{u}_1(t-\zeta))d\varsigma \\ &= \int_{-\infty}^{\infty} b_h(\zeta)G(\int_{-\infty}^{\infty} w(\tau)\tilde{x}(t-\zeta-\tau)d\tau)d\varsigma\end{aligned} \tag{2.22}$$

Chapter 3

Distortion Models for Signals Passing through Memoryless Nonlinearities

In the previous chapter, we introduced some widely used nonlinearity models to characterize input-output behavior of real world amplifiers. Having a good model for the amplifier might have different benefits depending on how the signal after amplifier will be processed. In some linearization methods such as predistortion where inverse of the amplifier is deployed, a good model plays a significant role. However, not all linearization methods require a very detailed modeling of the amplifier. Yet a good model can still be beneficial for the distortion analysis of given waveforms.

In this chapter, we will try to investigate the impact of power amplifier nonlinearity at signal level starting from simple tone signals and moving towards communication waveforms. In this context, the amplifier operation point, and distortion due to harmonic and *IMD* distortion will be discussed first. Then, *Bussgang theorem* will be presented and used to study the influence of nonlinearity on communication waveforms.

3.1 Harmonic and Intermodulation Distortion

Linear time-invariant systems do not produce any new frequency components that are not present in the input of the system. However, this is not true for nonlinear systems and in general the output signal occupies a wider spectrum. The following simple nonlinear system given in (3.1) will elaborate this statement.

$$y_{out} = c_1 x_{in} + c_2 x_{in}^2 + c_3 x_{in}^3 \tag{3.1}$$

Now if assuming that the input is a two tone signal, i.e. $x_{in}(t) = A\cos(2\pi f_1 t) + B\cos(2\pi f_2 t)$, the output of the system is given as

$$
\begin{aligned}
y_{out}(t) = {} & c_2(A+B)/2 + \cos(2\pi f_1 t)[c_3(3A^3/4 + 3B^2A/2) + c_1A] \\
& + \cos(2\pi f_2 t)[c_3(3B^3/4 + 3A^2B/2) + c_1B] \\
& + \cos(2\pi 2f_1 t)c_2A^2/2 + \cos(2\pi 2f_2 t)c_2B^2/2 \\
& + \cos(2\pi(f_1+f_2)t)c_2AB + \cos(2\pi(f_1-f_2)t)c_2AB \\
& + \cos(2\pi(2f_1-f_2)t)c_3 3A^2B/4 + \cos(2\pi(2f_1+f_2)t)c_3 3A^2B/4 \\
& + \cos(2\pi(2f_2-f_1)t)c_3 3AB^2/4 + \cos(2\pi(2f_2+f_1)t)c_3 3AB^2/4 \\
& + \cos(2\pi 3f_1 t)c_3A^3/4 + \cos(2\pi 3f_2 t)c_3B^3/4
\end{aligned}
\tag{3.2}
$$

In addition to the fundamental frequencies f_1 and f_2, (3.2) shows that the output spectrum has components at harmonics and different linear combinations of input frequencies. The frequencies of the form $f_1 \pm f_2$ are due to the second order intermodulation whereas third order intermodulation products are seen as $2f_1 \pm f_2$ and $2f_2 \pm f_1$. These products due to *IMD* as well as products that are originating from harmonic distortion are illustrated in Figure 3-1 (the amplitudes are not exact). It is depicted in Figure 3-1 that the additional terms due to harmonics and 2nd order *IMD* (*IMD₂*) lie far away from the fundamental frequency. However, the 3rd order *IMD* (*IMD₃*) of the form $2f_1 - f_2$ and $2f_2 - f_1$ lies in a rather close vicinity of the fundamental frequencies. It is possible to obtain more terms by increasing the order of the system and/or using an input with more tones. In [22], the other *IMD* terms originating from higher odd orders that fall near to the fundamental frequencies are shown.

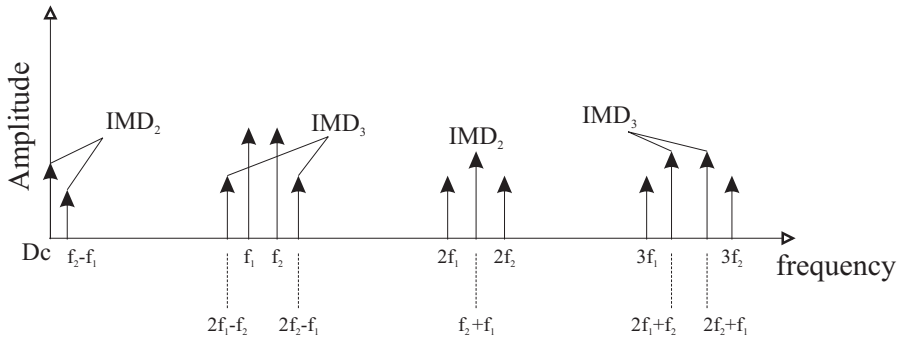


Figure 3-1. The output spectrum of a third order nonlinearity with a two tone signal.

The distortion terms that lie far away from the fundamental frequency can be eliminated by ordinary linear filtering. However, it is not feasible to think of a highly selective bandpass filter around the center frequency that only passes the fundamental frequencies and attenuates the odd order *IMD* terms. Besides there are also distortion components that fall on top of the signal such as the ones starting with c_3 in the second and third term of (3.2) which again

cannot be filtered away. This becomes more evident if the input signal and therefore the corresponding distortion terms have bandwidth.

The amplifier operation point is often mentioned to give more information about possible significance of distortion due to nonlinearity. One of the most commonly used terms in this context is 1 dB compression point that tells the input power level for which the power gain of the amplifier is 1 dB less than that of the linear region. On the other hand, input intercept point for IMD_3 (IIP_3) is another useful term defining the input power level at which the power of the fundamental frequency terms equal that of IMD_3. This is actually an artificial point that is found when the linear part of the curves are extrapolated. These two points are illustrated in Figure 3-2 in addition to the typical input-output power relation of an amplifier.

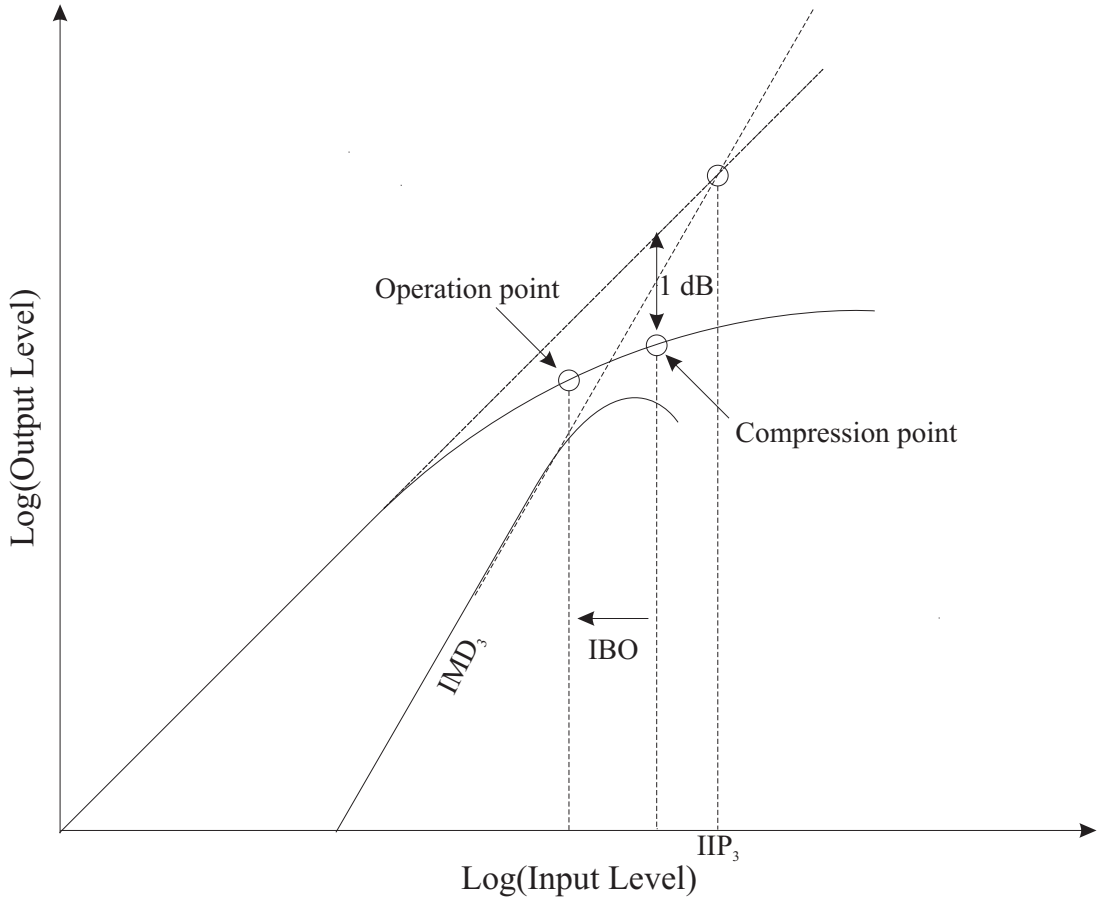


Figure 3-2. The operation , 1 dB compression, and IIP_3 points of an amplifier.

When an amplifier is said to be operating at a certain point, it refers to the average input power. The operating point is usually defined in comparison with the 1 dB compression point. When the average input power is less than that of the 1 dB compression point, the term input power back-off (*IBO*) is used to specify the difference between corresponding input power

levels. The concept of operation point and *IBO* are also illustrated in Figure 3-2. Note that when the amplifier is operating at 1 dB compression point, *IBO*=0 dB. It is also possible to define negative *IBO* values which correspond to operations closer to saturation level of the amplifier compared to 1 dB compression point.

The power of *IMD* and therefore the amount of distortion introduced is varying with the operation point. As an example, if we again consider the system given in (3.1) with an average input power of *X* dBm, then the significance of the terms $c_2 x_{in}^2 + c_3 x_{in}^3$ depend highly upon the value of *X*. However, for practical communication waveforms with non-constant envelopes and high peak-to-average power ratios (*PAPR*), the distortion can still be significant even though *X* is small. More advanced methods (other than IMD_3 curve of tone signals) are needed to measure quantitatively the distortion introduced in such cases. Rest of this chapter discusses how practical waveforms are affected both in time and frequency domains after being processed by memoryless nonlinearities.

3.2 Influence on Complex Gaussian Signals

In this section, we investigate the distortion introduced to complex Gaussian signals passing through a memoryless nonlinearity by using *Bussgang's theorem* [10], [26]. The theorem was originally stated for real Gaussian signals going through only AM/AM distortion. However the extension of theorem for complex Gaussian signals experiencing also AM/PM distortion has been developed, e.g. [15], [18]. The theorem states that the cross-correlation between output and input of nonlinearity is related to auto-correlation of the input by only a constant. Therefore in the extended version of the theorem, this relation considering a system given in Figure 2-1 can be written

$$\begin{aligned} E[\tilde{y}(t)\tilde{x}^*(t+\tau)] &= \alpha_o E[\tilde{x}(t)\tilde{x}^*(t+\tau)] \\ R_{\tilde{y}\tilde{x}^*}(\tau) &= \alpha_o R_{\tilde{x}\tilde{x}^*}(\tau) \end{aligned} \tag{3.3}$$

where α_o is a complex constant and $E[.]$ is the expectation operation.

The property indicated by expression (3.3) then enables to view the output of the nonlinearity as a sum of the scaled version of the input and a statistically uncorrelated additive distortion term, that is

$$\tilde{y}(t) = \alpha_o \tilde{x}(t) + \tilde{d}(t) \tag{3.4}$$

where $\tilde{d}(t)$ is zero-mean noise term due to the nonlinearity that is uncorrelated with the input signal and will be referred to as *IMD* from now on. The input-output relation of nonlinearity given in (3.4) is also presented in Figure 3-4 b). According to model given in Figure 2-1, $G(A(t)) = g_A(A(t))e^{j(\phi(t)+g_\phi(A(t)))}$ but considering that neither AM/AM nor AM/PM mappings operates on input phase, $\phi(t)$, then a simplified notation $G(A) = g_A(A)e^{j(g_\phi(A))}$ can be used keeping in mind that $\phi(t)$ also appears at the output. With this simplified notation, complex gain α_o is shown in [15] to be

$$\begin{aligned} \alpha_o &= E[\frac{\partial G(A)}{\partial A} + G(A)/A] \\ &= \int_0^\infty \frac{\partial G(A)}{\partial A} f_A(A)dA + \int_0^\infty (1/A)G(A)f_A(A)dA \end{aligned} \tag{3.5}$$

where $f_A(.)$ is the probability density function (pdf) of A and partial derivative with respect to A is indicated by $\partial(.)/\partial A$. It should also be remarked that for α_o to be a constant independent of t, right hand side of (3.5) should be independent of t. It is worthy to remind that A is actually standing for $A(t)$ and therefore right hand side of (3.5) is a function of t.

Denoting $\tilde{x}(t) = \tilde{x}$, $\tilde{y}(t) = \tilde{y}$ and $\tilde{d}(t) = \tilde{d}$, from (3.4) the complex amplitude gain can also be expressed as

$$\begin{aligned} \alpha_o &= R_{\tilde{y}\tilde{x}^*}(0)/R_{\tilde{x}\tilde{x}^*}(0) = E[\tilde{y}\tilde{x}^*]/E[\tilde{x}\tilde{x}^*] \\ &= (E[G(A)A]/E[A^2] \end{aligned} \tag{3.6}$$

where $(.)^*$ denotes the complex conjugation. As uncorrelatedness implies, the second term in the numerator is, $E[\tilde{d}\tilde{x}^*] = E[\tilde{d}]E[\tilde{x}^*] = 0$. The denominator is the variance of the input that can be given as

$$E[\tilde{x}\tilde{x}^*] = \int_{-\infty}^{\infty} A^2 f_A(A)dA = \sigma_{\tilde{x}}^2 \tag{3.7}$$

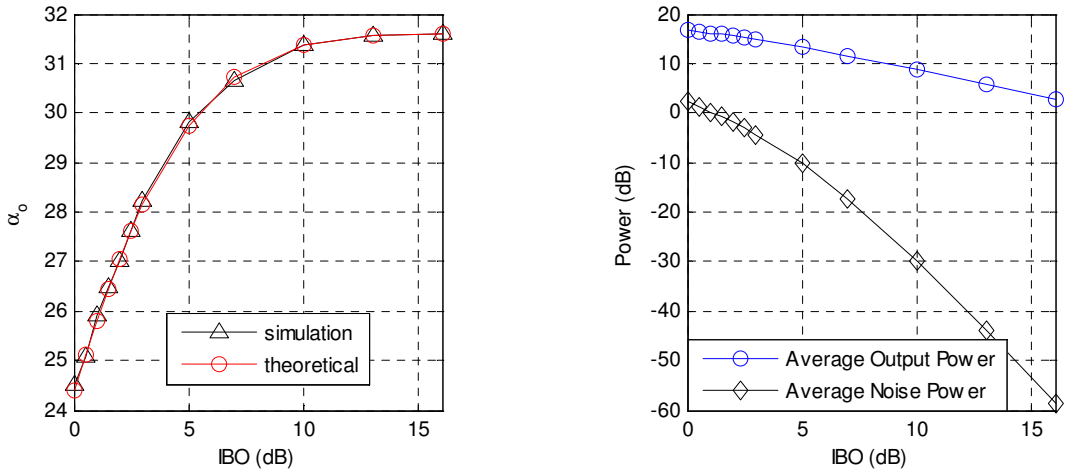
and the remaining term in numerator is

$$E[\tilde{y}\tilde{x}^*] = E[G(A)A] = \int_0^\infty A\,G(A)f_A(A)dA \tag{3.8}$$

Once the complex gain is determined, again by using the fact that noise term and the input are uncorrelated, it is possible to obtain the variance of the *IMD* term by

$$E[\tilde{d}\tilde{d}^*] = E[\tilde{y}\tilde{y}^*] - |\alpha_o|^2 E[\tilde{x}\tilde{x}^*]$$
$$= \int_0^\infty |G(A)|^2 f_A(A)dA - |\alpha_o|^2 \sigma_{\tilde{x}}^2 \tag{3.9}$$

Expressions (3.5), (3.6), and (3.9) derive an explicit expression for the complex gain and the variance of the noise in terms of the nonlinearity and input pdf. However, there might not be an analytical solution to (3.5) or (3.8) for arbitrary nonlinearities and pdf's. Nevertheless, numerical solutions are possible, and we will illustrate the dependency of the complex gain, average noise power and average output power to the operation point with some simulation results. An *SSPA* model defined in (2.8) with parameters $A_o = 10$, $p = 2$, $\kappa = \sqrt{1000}$ (corresponds to 30 dB power gain in the linear region) is used for the amplifier. The input signal is an *OFDM* signal with $N = 512$ subcarriers and 16-QAM subcarrier-modulation.



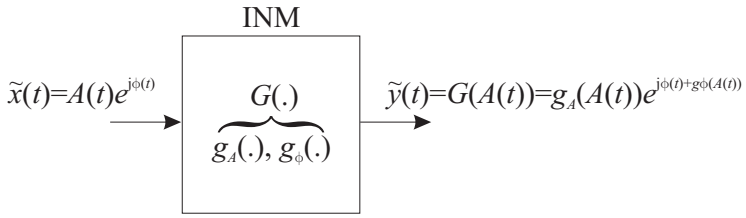
a) *b)*

Figure 3-3. a) The complex gain values of Busgang model at different IBO's from 1 *dB compression point b) Average output and IMD power at different IBO's from* 1 *dB compression point.*

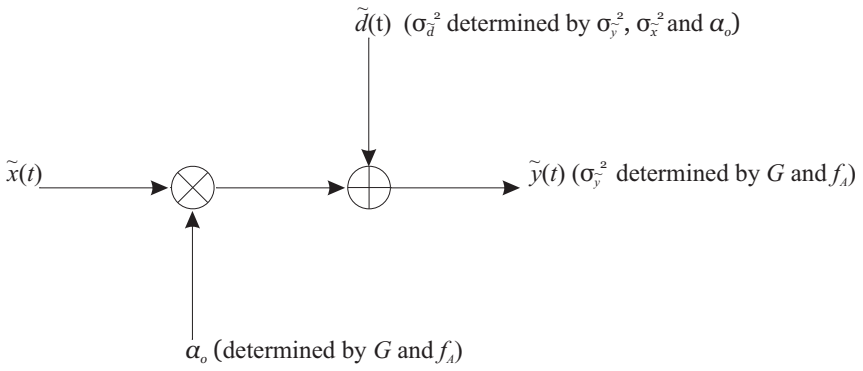
The theoretical value of α_o illustrated in Figure 3-3 a) is calculated numerically based on (3.5). On the other hand, the data at the input and output of the nonlinearity is used to approximate the expectation in (3.6) and which shows good matching with the theoretical value according to Figure 3-3. It is seen from Figure 3-3 a) that α_o increases as the *IBO* from 1 dB compression point increases. At high values of *IBO*, α_o approaches to $\kappa = \sqrt{1000}$ as expected since the amplitude values of the input signal stays in the linear region of the amplifier. The corresponding average output, and *IMD* powers are shown in Figure 3-3 b). If

we define the ratio between the useful signal and *IMD* as signal-to-*IMD* ratio (*SIMDR*), then Figure 3-3 b) points out that *SIMDR* decreases as *IBO* decreases. Although the achieved average output power is about 18 dB's for $IBO = 0$ dB, the noise appearing due to nonlinearity is about 1 dB, causing an *SIMDR* of 17 dB's. On the other extreme, when $IBO = 15$ dB, average output power is about 2 dB's whereas *IMD* power is only -60 dB, yielding an *SIMDR* of 62 dB's.

When modeling the distortion introduced to the signal due to nonlinearity, the input-output relation of the nonlinearity illustrated in Figure 2-1 and Figure 3-4 a) is replaced by Figure 3-4 b). The models are obviously connected to each other with the function $G(.)$ that characterizes the nonlinearity. In this figure, it is also illustrated which other factors does the complex-gain α_o, variance of the output $\sigma_{\tilde{y}}^2$ and variance of the *IMD* term $\sigma_{\tilde{d}}^2$ depend on.



a)



b)

Figure 3-4 Input output relation of nonlinearity given by a) AM/AM and AM/PM mapping b) Bussgang model with constant scaling α_o and distortion term $\tilde{d}(t)$.

The knowledge about the power spectral density (PSD) of $\tilde{d}(t)$ is also important for two particular reasons. The spectral components of $\tilde{d}(t)$ that fall within the bandwidth of the input signal are definitely an impairment that affects the symbol detection process on the receiver

side. On the other hand, the spectral components that cause the spectral re-growth (the components that fall to close vicinity of the input signal) may cause interference with transmissions of other neighboring channels.

In case of memoryless non-linearities, the output signal $\tilde{y}(t)$ and therefore $\tilde{d}(t)$ depend only on the magnitude of the input at the same time instant, i.e. $A(t)$. Thus for an input signal whose magnitude values at different time instants are independent of each other, i.e. with white spectrum, the corresponding *IMD* is also white. However, in general the continuous time signal at the input of the amplifier is band-limited and thus PSD of the nonlinearity noise is not white anymore. The PSD of the noise is shown in [4] and [6] to be

$$S_{\tilde{d}}(f) = \sum_{n=1}^{\infty} \frac{a_n}{(\sigma_{\tilde{x}}^2)^{2n+1}} [S_{\tilde{x}}(f) \underbrace{\otimes \cdots \otimes}_{2n+1} S_{\tilde{x}}(f)] \tag{3.10}$$

where the coefficients a_n are given as

$$a_n = \frac{4}{\sigma_{\tilde{x}}^2 (n+1)} \times \left\| \int_0^{\infty} g(A) \frac{A^2}{\sigma_{\tilde{x}}^2} e^{A^2/\sigma_{\tilde{x}}^2} L_n^{(1)}(\frac{A^2}{\sigma_{\tilde{x}}^2}) dA \right\|^2 \tag{3.11}$$

with $L_n^{(1)}(x)$ being the Laguerre function of the first form which has the general form

$$L_n^{(k)}(x) = \frac{x^{-k} e^x}{n!} (\frac{d}{dx})^n (x^{n+k} e^{-x}) \tag{3.12}$$

The PSD of the nonlinearity noise given in (3.10) can be used to compute the in-band distortion as well as the power of out-of-band components by integrating $S_{\tilde{d}}(f)$ in the desired band.

3.3 Influence on Single-carrier Systems

In the previous section, we have introduced the *Bussgang theorem* for relating the output-to-input and input-to-input cross-correlations of a memoryless nonlinearity to which a model given in (3.4) fits. In this section we will examine some time-domain effects of such model on single-carrier communication scheme presented in Figure 3-5.

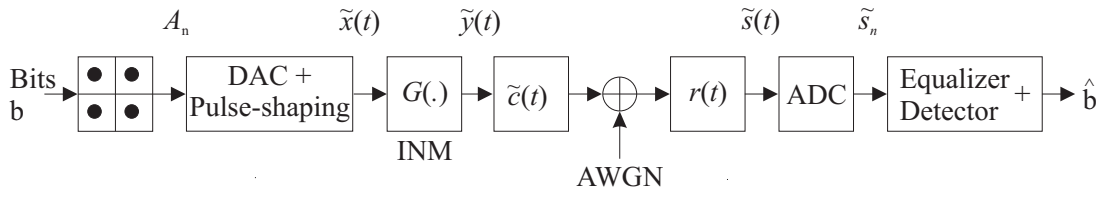


Figure 3-5 Baseband equivalent of a single carrier digital communication system.

A digital baseband equivalent of a simplified single carrier communication system [23] is illustrated in Figure 3-5 taking also the nonlinearity originating from power amplifier into account. The bit sequence to be transmitted is shown as b, which is then mapped to sequence of symbols A_n depending on the modulation scheme, e.g. 16-QAM, 8-PSK…etc. The symbol sequence is then converted to analog waveform by *digital to analog converter* (*DAC*) with pulse shape $p(t)$. The analog waveform before the amplifier can be written as

$$\tilde{x}(t) = \sum_{n=-\infty}^{\infty} A_n p(t - nT_s) \tag{3.13}$$

where T_s is the symbol period. Then this signal goes through the power amplifier whose output is denoted as $\tilde{y}(t)$. This is assumed to be the final signal to be transmitted that passes through a baseband channel with impulse response $\tilde{c}(t)$ and is also further distorted with an additive white Gaussian noise (*AWGN*). The signal is received through a receive filter *r(t)* which can be written

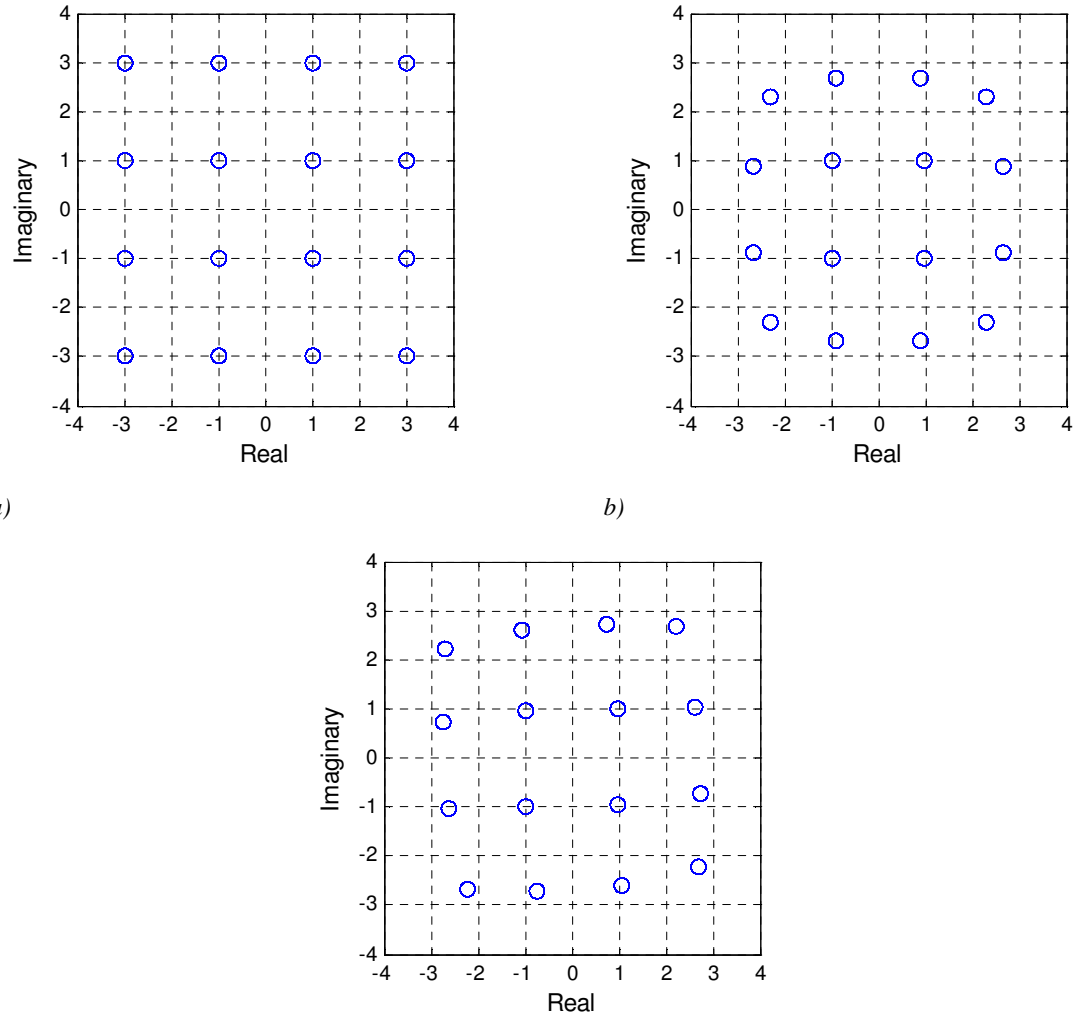
$$\tilde{s}(t) = [\tilde{y}(t) \otimes \tilde{c}(t) + \tilde{n}(t)] \otimes r(t) \tag{3.14}$$

where $\otimes$ stands for convolution operation, and $\tilde{n}(t)$ is a complex-valued white Gaussian process. Assuming that the sampling is done by *analog to digital converter* (*ADC*) after receive filter at symbol rate, then the received signal $\tilde{s}(t)$ yields the sequence $\tilde{s}_n = \tilde{s}(nTs)$ which is further processed by the equalizer and detector for the decision of received symbols and the corresponding bit sequence $\hat{b}$.

The main assumption of the *Bussgang* theorem is that the input to nonlinearity is a zero-mean Gaussian process which doesn't hold in general for single-carrier system elaborated above where the symbols are selected uniformly among the constellation points. On the other hand, single carrier systems without pulse shaping can still be shown to obey the model in (3.4). As long as $R_{yx^*}(0) = \alpha_o R_{xx^*}(0)$ is satisfied as stated in (3.6), then we still can describe the output process of the nonlinearity as the summation of the input process with a constant

scaling and an uncorrelated noise process. It is fairly easy to show that single-carrier systems without pulse shaping satisfy this condition and the complex gain for a constellation of size *M* can be given as

$$\begin{aligned}\alpha_o &= E[\tilde{y}\tilde{x}^*]/E[\tilde{x}\tilde{x}^*] \\ &= \frac{1/M\sum_{i=1}^{M} g(A_i)A_i^*}{1/M\sum_{i=1}^{M} A_iA_i^*} = \frac{\sum_{i=1}^{M} g(A_i)A_i^*}{\sum_{i=1}^{M} A_iA_i^*}\end{aligned} \tag{3.15}$$



a) *b)*

c)

Figure 3-6. The constellation of 16-QAM signal a) before the nonlinearity, b) after SSPA ($A_o = 10$, $p = 2$, $\kappa = \sqrt{1000}$), and c) after a nonlinearity with Saleh model ($\alpha_A = 1$, $\beta_A = 0.5$, $\alpha_\phi = 0.25$, $\beta_\phi = 0.25$).

In general, AM/AM is a mapping that has higher gains for small amplitudes and lower gains for larger amplitudes. This is observed as the contraction of the constellation whereas AM/PM mapping is reflected as a rotation. The constellation of a 16-QAM signal before the nonlinearity is shown in Figure 3-6 a), whereas b) and c) shows the constellation after nonlinearities with *Saleh* and *SSPA* models. The input signal is actually scaled to 1 dB compression point which changes the amplitude levels even before the amplifier, however normalized constellations are shown. Only contraction of the constellation is observable in Figure 3-6 b), since there is only AM/AM mapping in *SSPA* model whereas both contraction and rotation are visible in c) as expected.

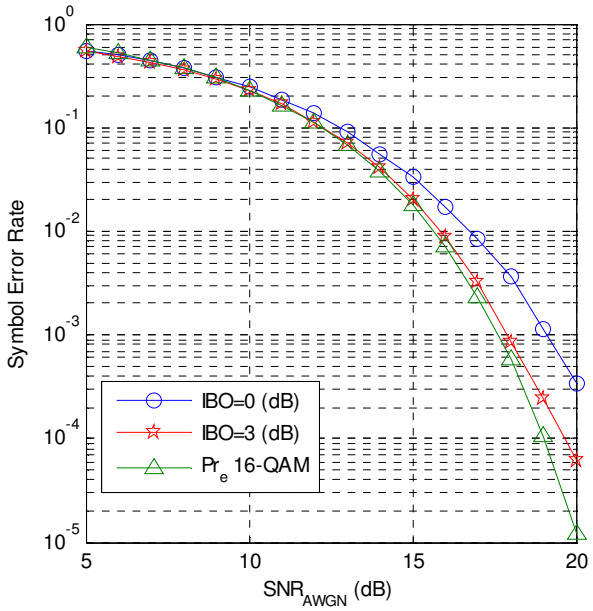


Figure 3-7. SER performance of 16-QAM for IBO= 0 *and* 3 *dB at varying* SNR_{AWGN} *values. The theoretical probability of error [23] curve is also shown considering that there is no effect of nonlinearity.*

The rotation of the constellation (due to the AM/PM mapping) can be compansated by the channel equalization process on the receiver side. The contraction of the constellation if not handled by transmitter or receiver, will increase *SER* and/or *BER* since the minimum distance is decreased. The shape of the constellation after the nonlinearity makes it rather difficult to derive an analytical form for probability of error. Since the performance analysis of single carrier systems is not the main focus of this work, we shall only illustrate the *SER* performance of 16 and 64-QAM modulations based on the simulations. In simulations, the system in Figure 3-5 is considered with pulse shape $p(t)$ being rectangular in the interval [0 T_s) and an *SSPA* model with parameters given above is used. Since we are merely paying

attention to the influence of nonlinearity, the channel response $\tilde{c}(t)$ is taken as Dirac delta $\delta(t)$ for simplicity. Finally *SER* is calculated based on the comparison of sent and received symbol sequences of length 150000.

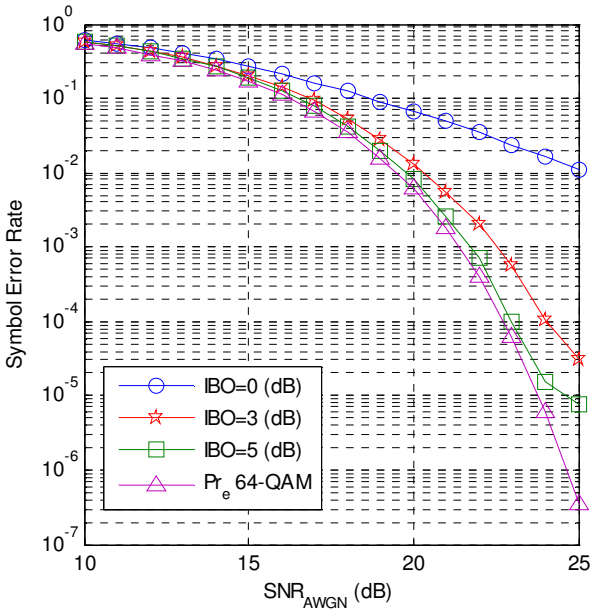


Figure 3-8. SER performance of 64-QAM for IBO=0, 3 *and* 5 *dB at varying* SNR_{AWGN} *values. The theoretical probability of error curve is also shown considering that there is no effect of nonlinearity.*

The *SER* performance degradation of 16 and 64-QAM due to the influence of nonlinearity are shown in Figure 3-7 and Figure 3-8. In both of the figures, as the power amplifier is biased towards the nonlinear region more (when *IBO* is decreased), the symbol error rate tends to increase due to *IMD* in addition to *AWGN*. It is also convenient to note that signal to noise ratio due to *AWGN* (SNR_{AWGN}) is calculated as the ratio of useful signal to *AWGN*. In other words, if we consider the signal $\tilde{z}(t) = \alpha_o \tilde{x}(t) + \tilde{d}(t) + \tilde{n}(t)$, then $SNR_{AWGN} = 10\log(|\alpha_o|^2 \sigma_{\tilde{x}}^2 / \sigma_{\tilde{n}}^2)$.

3.4 Influence on Multi-Carrier Systems

Today's high speed wireless communications are more and more getting based on systems deploying multicarrier concept. Systems with such independent carriers tend to have high *PAPR* and thus more vulnerable to nonlinear distortions. Here we will present both time and frequency domain effects of distortion introduced by memoryless nonlinearities on *OFDM* type of signal which is becoming more and more popular as a multi-carrier scheme [1], [21].

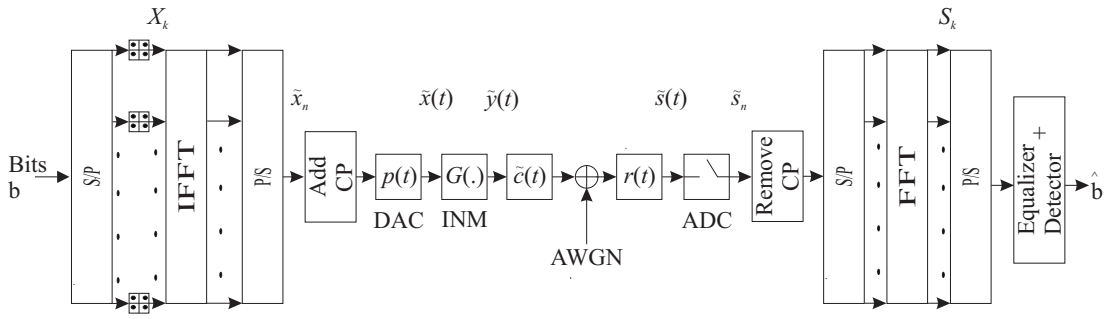


Figure 3-9. Baseband equivalent multi-carrier digital communication system employing OFDM.

A simplified diagram for a baseband *OFDM* scheme using a digital multi-carrier technique [37] is shown in Figure 3-9 with the power amplifier being modeled as a memoryless nonlinearity. The bit sequence to be transmitted is *b* which is then divided into parallel streams by a serial-to-parallel (S/P) converter. Each parallel stream of bits is then mapped into symbols X_k to modulate the $k'th$ subcarrier, where $k = 0,1\cdots N-1$ with N being the total number of sub-carriers. These N sub-carrier symbols are then processed by inverse fast Fourier transform (*IFFT*) and parallel-to-serial (P/S) conversion blocks successively, to give the $m'th$ *OFDM* symbol before the memoryless nonlinearity.

$$\tilde{x}_m[n] = 1/N \sum_{k=0}^{N-1} X_k^m e^{j2\pi kn/N} \tag{3.16}$$

The addition of cyclic prefix (*CP*) to prevent inter symbol interference (ISI) due to multi-path channel and reduce the complexity for channel equalization is not considered in (3.16). However, the following derivations are still valid since it will be removed on receiver side. Hence continuing with the signal after *DAC* with pulse shape $p(t)$

$$\tilde{x}_m(t) = \sum_{n=0}^{N-1} \tilde{x}_m[n] p(t - nT_b - mNT_b) \tag{3.17}$$

where T_b is the separation between N time domain samples of one *OFDM* symbol, i.e. one *OFDM* symbol duration $T_s = NT_b$. Finally the whole signal to be transmitted before the power amplifier can be written

$$\tilde{x}(t) = \sum_{m=-\infty}^{\infty} \tilde{x}_m(t) \tag{3.18}$$

Now again the output of the power amplifier is denoted as $\tilde{y}(t)$ which is the final signal to be transmitted. This signal goes through a channel with impulse response $\tilde{c}(t)$ and then further distorted with *AWGN*. Receiving is done by the receive filter $r(t)$, which yields the received waveform

$$\tilde{s}(t) = [\tilde{y}(t) \otimes \tilde{c}(t) + \tilde{n}(t)] \otimes r(t) \tag{3.19}$$

and assuming that $\tilde{s}(t)$ is sampled at rate T_b, we can write the samples of the *m'th OFDM* symbol as $\tilde{s}_m[n] = \tilde{s}(nT_b + mNT_b)$. After the removal of *CP* and serial to parallel (S/P) conversion block, the fast Fourier transform (*FFT*) of samples of each received *OFDM* block is taken which gives

$$S_k^m = \sum_{n=0}^{N-1} s_m[n] e^{-j2\pi nk/N} \tag{3.20}$$

The sequence after *FFT* block is then serialized by the P/S converter, and sent to the equalizer and detector for the final decision to be done for the received bits $\hat{b}$.

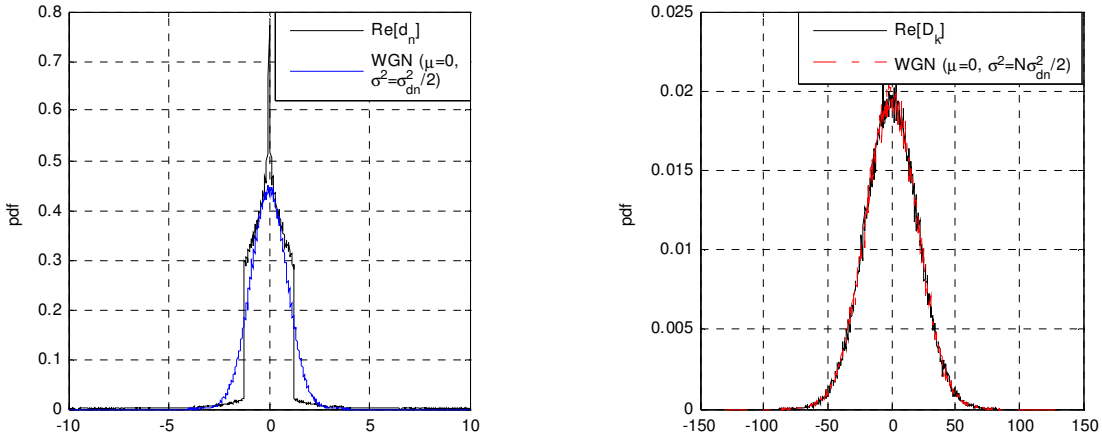
The time domain OFDM sequence $\tilde{x}[n]$ described by the expression (3.16) can be considered as a complex Gaussian process for large number of sub-carriers (by central limit theorem). Then the *Bussgang* theorem introduced in Section 3.2 can directly be used to analyze the distortion introduced by the nonlinearity. It is also shown in [15] that the model in (3.4) is still valid with α_o being just a complex constant for two special cases of pulse-shaping applied to time domain *OFDM* sequence. Either a rectangular filter or a filter band-limited to [*-1/(2 T_b), 1/(2 T_b)*] satisfies the conditions for α_o to be time independent.

We can consider that the baseband *OFDM* signal, $\tilde{x}(t)$ has real and imaginary components being i.i.d Gausian, then the magnitude is *Rayleigh* distributed, i.e. $f_A(A) = \frac{A}{\sigma_{\tilde{x}}^2} \exp(-\frac{A}{2\sigma_{\tilde{x}}^2})$. The closed form solutions to both (3.8) and (3.9) and thus to α_o is given in [31] for this *Rayleigh* distributed amplitude going through a polynomial nonlinearity.

The variance of *IMD*, $\tilde{d}$, can be obtained either analytically or numerically, but the distribution is not straightforward. In general, the output of the nonlinearity is non-Gaussian, therefore the difference between the output and input is also expected to be non-Gaussian. However, more information is available for the distribution of *IMD* falling on each subcarrier. It is worth mentioning once again that spectrum of $\tilde{d}$ is white if and only if spectrum of $\tilde{x}$ is white. In that case with an additional condition that the channel response is $c(t) = \delta(t)$, the signal after *FFT* block on receiver side which is given by (3.20) can be re-written

$$\begin{aligned} S_k^m &= \sum_{n=0}^{N-1} s_m[n] e^{-j2\pi nk/N} \\ &= \sum_{n=0}^{N-1} (\alpha_o \tilde{x}_m[n] + \tilde{d}_m[n] + \tilde{n}_m[n]) e^{-j2\pi nk/N} \\ &= \alpha_o X_k^m + D_k^m + N_k^m \end{aligned} \tag{3.21}$$

where D_k^m and N_k^m are the components of noise terms on subcarrier k coming from nonlinearity and channel respectively. It can be seen that D_k^m is superposition of N independent time domain noise samples of $\tilde{d}$, and by central limit theorem it can be claimed that it has Gaussian distribution for large N. For an *OFDM* signal with $N = 512$, passing through an *SSPA* model with the same parameters as in Section 3.2, the distribution of the real part of *IMD* is shown in Figure 3-10 a). It is obviously far from being Gaussian when compared to the distribution of *WGN* with the same variance. On the other hand, Figure 3-10 b) indicates that the noise appearing on subcarrier k due to $\tilde{d}$, or in other words D_k has Gaussian distribution.

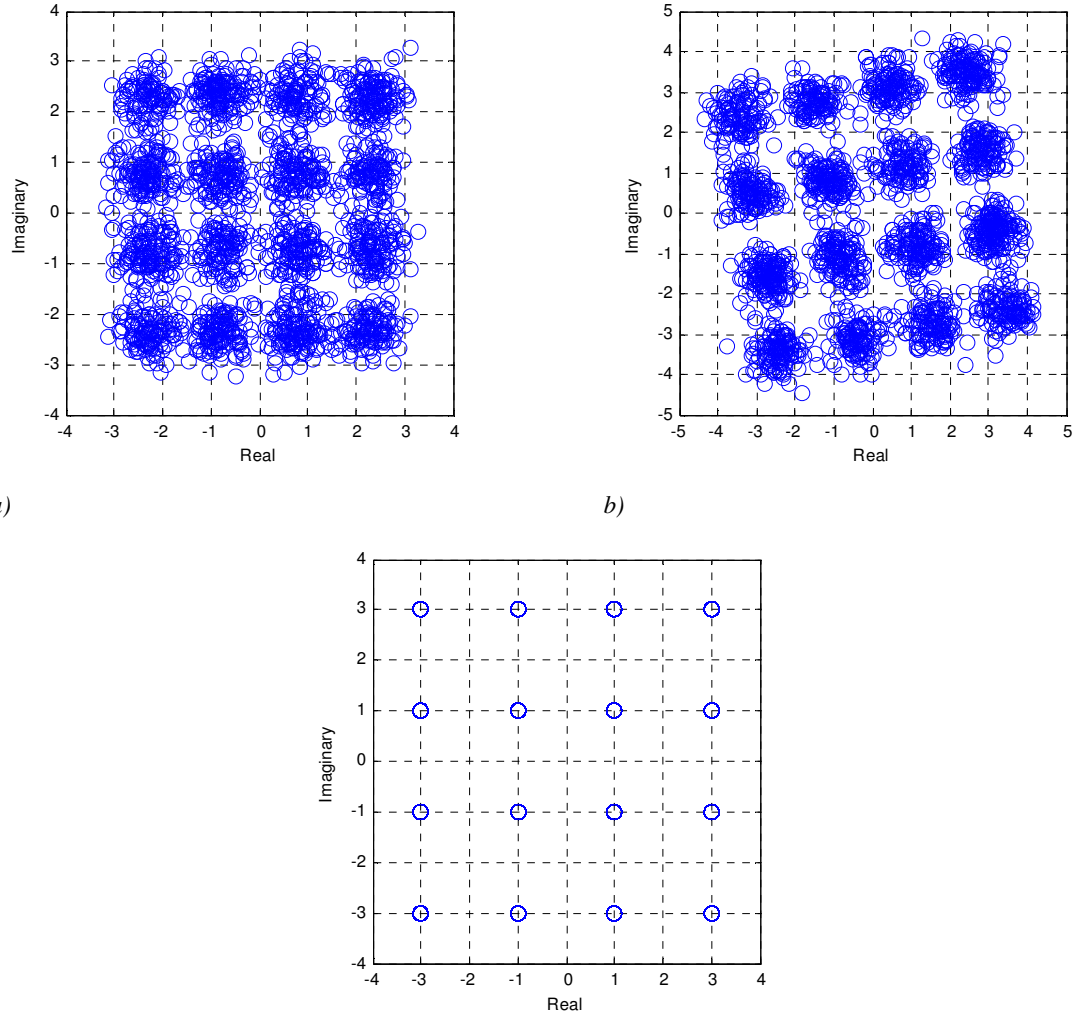


a) *b)*

Figure 3-10.Distribution of zero mean white Gausian noise a) and real part of IMD b) and real part of IMD that falls on subcarrier k.

At this point, we can state that there are two noise terms D_k^m and N_k^m on top of the desired subcarrier symbol X_k^m that are independent from each other and having Gaussian distribution. In order to obtain a normalized constellation, we can divide (3.21) by the amplitude gain of the amplifier in linear region. This complex constant appearing as a cofactor will cause a rotation and contraction of the original constellation that can be observed when

Figure 3-11 a) and c) are considered. However the amount of contraction is same for all constellation points as a difference from single-carrier case. In addition to that, we also expect to see a Gaussian cloud around the constellation points due to D_k^m and N_k^m as well which is illustrated in
Figure 3-11 a) and b).



a) *b)*

c)

Figure 3-11. Constellation of subcarrier $k = 10$ *for a) SSPA model (* $A_o = 10$, $p = 2$, $\kappa = \sqrt{1000}$ *), b) Saleh model (* $\alpha_A = 1$, $\beta_A = 0.5$, $\alpha_\phi = 0.25$, $\beta_\phi = 0.25$ *) c) ideal symbol locations.*

The two noise terms being independently Gaussian also allows us to use the usual probability of error derivations for the basic constellations, e.g. for 16-QAM ([23], pp. 327).

The main idea is to take the advantage of the noise terms being independent and superpose their powers to obtain a signal-to-total-noise-ratio ($STNR_k = |\alpha_o|^2 |X_k|^2 / (|D_k|^2 + |N_k|^2)$) and use this quantity for deriving the probability of error. For clarity, we shall again point out to different noise terms that are given as $SIMDR$ ($SIMDR_k = |\alpha_o|^2 |X_k|^2 / |D_k|^2$) and SNR_{AWGN} ($SNR_{AWGNk} = |\alpha_o|^2 |X_k|^2 / |N_k|^2$). The probability of error for 16-QAM modulation and the obtained *SER* are shown in Figure 3-12.

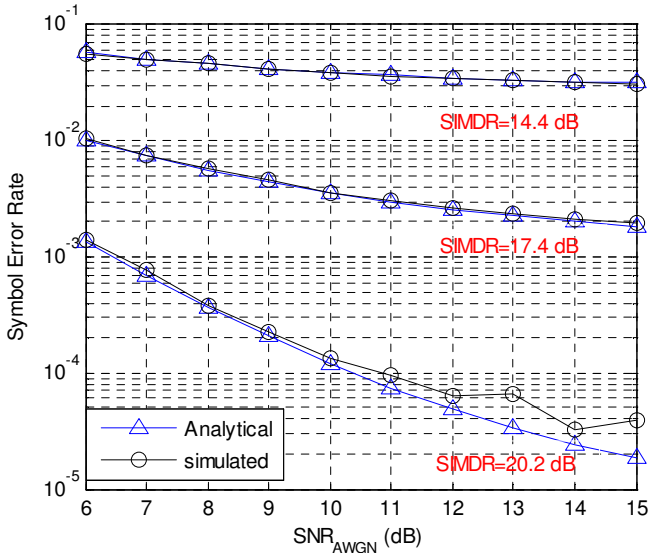


Figure 3-12. Analytical and simulated SER for an OFDM signal with 16-QAM subcarrier-modulation when SNR_{AWGN} is varied from 6 to 15 dB for three different values of $SIMDR_k$.

Chapter 4

Linearization Methods

In the previous chapter, the impacts of memoryless nonlinearities on signals that have importance from communication theory point of view are studied. Spectral re-growth and in-band distortion are shown to be the two significant impairments. The former can cause unacceptable interference with transmissions of neighboring channels whereas the latter degrades the system performance. With the increasing concerns for power efficiency, then power amplifier linearization becomes more essential in the compensation of these impairments. In this chapter some very widely used linearization methods will be presented.

4.1 Feedback Linearization

Feedback linearization which is illustrated in Figure 4-1 has been widely implemented for the linearization of audio amplifiers. The basic principle is to compare the amplifier output with the signal to be amplified and feed the error signal which is any distortion other than linear amplification. Here a *Bussgang* model (eventhough, interestingly, the name *Bussgang* is never mentioned explicitly in feedback linearization techniques to the author's knowledge) is assumed for the output of amplifier which consists of linearly amplified version of the input and additive distortion term.

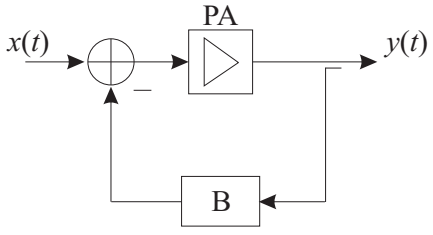


Figure 4-1. Simplified structure for feedback linearizer.

The distortion terms are shown to be reduced in [22] by a factor of B/A_o where B is the attenuation value on the feedback path and A_o is the linear amplitude gain of the amplifier with the assumption that $B << A_o$. On the other hand, there is a sacrifice from the gain and the signal is amplified only by a factor of B.

The feedback linearization of RF signals is differentiated from the linearization of audio signals. One of the main reasons is that the bandwidth and location of the signal in the spectrum which might impose a feedback loop design up-to several GHz. The design specifications can be extremely difficult to achieve especially considering the delay in the feedback path that might be several cycles of the centre frequency. Cartesian feedback method [16] illustrated in Figure 4-2 is one way to deal with this problem. The amplifier output is down-converted to baseband and the comparison is done with the inphase and quadrature signals. Thus the delay in each loop has less significance due to slowly varying baseband signals. This baseband processing can be enhanced by the use of DSP to compensate for the nonidealities introduced in the down-conversion process due to mixers.

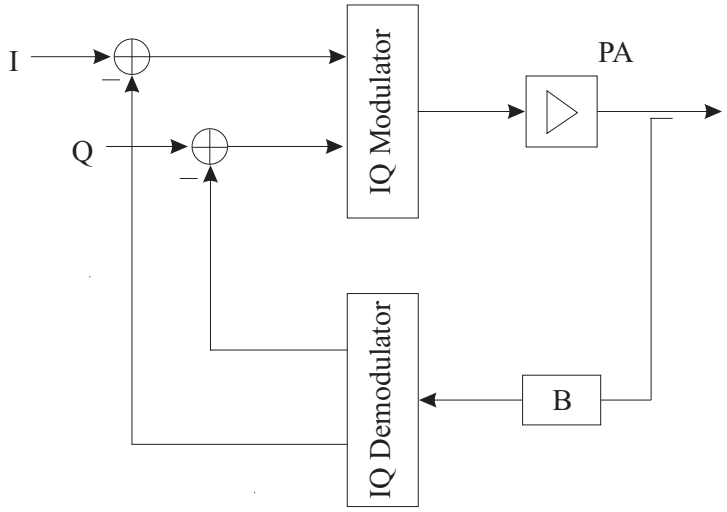


Figure 4-2. Simplified structure of Cartesian feedback.

4.2 *LINC*

Linear amplification with nonlinear components (*LINC*) [14] is another widely studied linearization method. The method is based on separating the signal to be amplified that might have both amplitude and phase variations into two constant envelope phase modulated components. Then the two constant envelope signals are amplified by two nonlinear

amplifiers and the output of the amplifiers are combined to yield a linear amplification. The basic structure of a *LINC* linearizer that deploys this idea is shown in Figure 4-3.

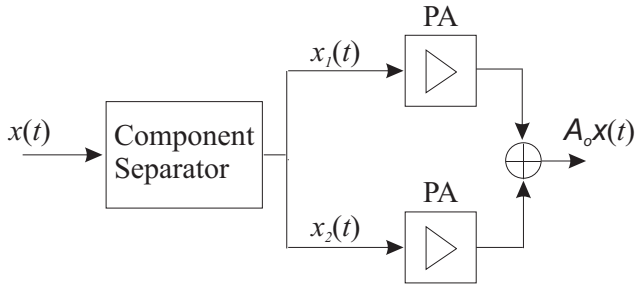


Figure 4-3.Simplified structure of LINC linearizer.

If the general bandpass model is assumed for the signal to be amplified, say $x(t) = A(t)\cos(2\pi f_c t + \theta(t))$, then the constant envelope signals after component separator are $x_{1,2} = \pm A_m / 2 \sin(2\pi f_c t + \theta(t) \pm \phi(t))$. It is shown in [14] and [16] that A_m being maximum of $A(t)$ and $\phi(t) = \sin^{-1}(A(t)/A_m)$ is a proper choice for re-combining amplifier outputs to produce a linearly amplified version of $x(t)$.

The major drawbacks of the *LINC* method are the difficulties in the realization of component separator, synchronization of two branches and isolating the power amplifiers.

4.3 Pre-distortion

One very intuitive approach to linearization problem is to use a pre-distorter that reverses the nonlinearity introduced by the amplifier. In other words, given a good model for the nonlinearity of amplifier, the idea is to either pre-distort (before amplifier) or post-distort (after amplifier) the signal or both so that the cascade of the nonlinearities yields a linearized response. However, due to drawbacks related to practical implementation ([22], [16]) pre-distortion is more commonly used linearization method compared to post-distortion.

Pre-distorter can be implemented at different transmitter stages such as at RF or IF stages or at baseband ([22], [16], [35]). The simplified illustrations for RF and baseband implementations are given in Figure 4-4 and Figure 4-5.

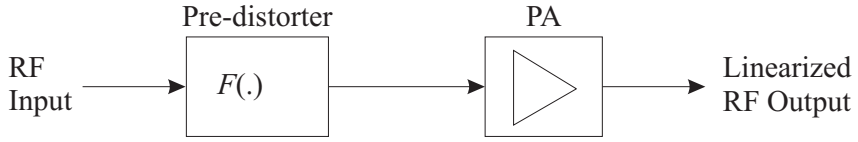


Figure 4-4. Pre-distorter implemented at RF.

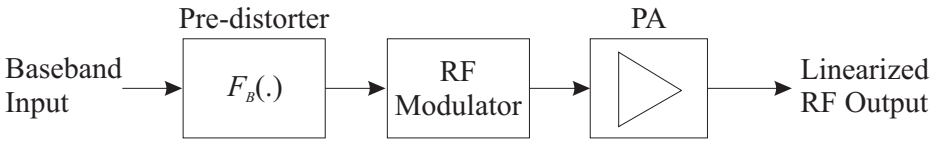


Figure 4-5. Pre-distorter implemented in baseband.

The functions $F(.)$ and $F_B(.)$ try to mimic the inverse of amplifier nonlinearity $G(.)$ at RF and baseband respectively. This directly implies that a good model for $G(.)$ such as the ones introduced in Chapter 2 is essential. Therefore poor modeling of amplifier nonlinearity will directly cause poor linearization performance [3].

4.3.1 Adaptive Digital Pre-distortion

Another possible source of poor performance is the use of fixed pre-distorter and neglecting the aging effects of amplifier. Then the pre-distorter will no longer match the inverse of the amplifier. Adaptive digital pre-distortion methods operating at baseband or low *IF* are proposed ([17], [3], [16]) to track the changes in the amplifier model and update the pre-distorter accordingly which is illustrated in Figure 4-6.

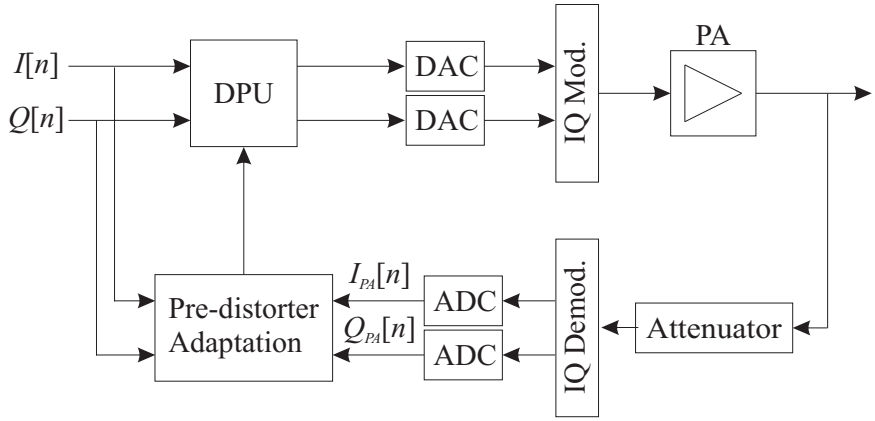


Figure 4-6. Adaptive pre-distorter implemented at baseband.

Similar to the Cartesian feedback linearizer in Figure 4-2, the feedback loop is closed at baseband but as a further operation *ADC's* are used to have discrete observations of the down-converted power amplifier output. These signals denoted as $I_{PA}[n]$ and $Q_{PA}[n]$ as well as the original baseband signals $I[n]$ and $Q[n]$ are then used to adapt the pre-distorter. The adaptation of pre-distorter may include some or all of the following processes:

1. Choosing a model for amplifier
2. Estimating the parameters of the chosen amplifier model

3. Designing a proper pre-distorter unit to inverse the stated amplifier model

Once the parameters of pre-distorter are identified, then they can be passed to the digital pre-distorter unit (DPU) that distorts the baseband signals $I[n]$ and $Q[n]$ to have a linearized output after power amplifier.

Adaptive pre-distortion technique might still have poor linearization performance if there is a model mismatch at the first place. This actually emphasizes the importance of the first step in pre-distorter adaptation which is mentioned above. Another complication of adaptive digital pre-distortion is the required digital signal processing (DSP) unit for the implementation of necessary identification and adaptation processes. Depending on the complexity of the model, the required DSP might shadow the efficiency of the power amplifier.

4.4 Feedforward Linearization

Feedforward linearization was first introduced in late 1920's by Black [7] and after 1960's, it started to take more attraction due to increase in the operating central frequencies and bandwidths of the communication systems. Since then it has found a vast area of application in military, satellite and cellular communications [22].

4.4.1 Basic Circuitry and Operation

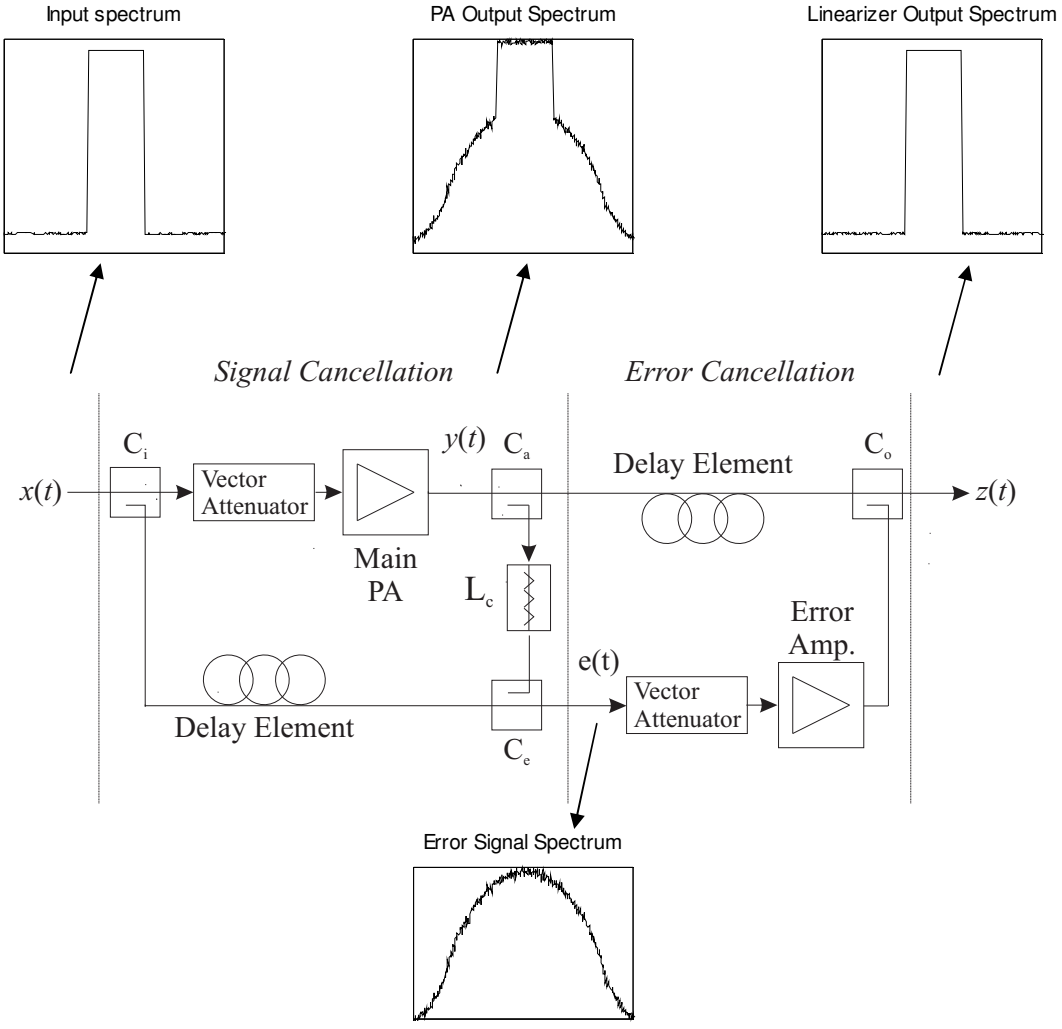


Figure 4-7. The basic structure and components of feedforward linearizer.

The basic structure and circuitry of feedforward linearizer with *RF* signals are shown in Figure 4-7. The input is denoted as $x(t)$ whereas amplifier and linearizer outputs are shown as $y(t)$ and $z(t)$ respectively. The two loops, *signal cancellation* and *error cancellation* are

the basic processing units of feedforward linearizer being responsible for the identification and cancellation of the main amplifier's distortion.

4.4.2 Signal Cancellation Loop

The first stage of feedforward linearization is to identify the distortion introduced by the main amplifier. For this purpose, the input signal $x(t)$ is separated into a lower and upper branch of *signal cancellation loop* (*SCL*) by the input coupler C_i. The upper branch first goes through a vector attenuator (a scaling and phase shift) and then passes through the main amplifier. The output of the amplifier, $y(t)$, is again split into two by the coupler C_A, where the branch going down is passed through an attenuator of power loss denoted as L_c. The role of the attenuator is to bring the signal level approximately to the level of lower branch of *SCL* that merely passes through a delay line. The two signals coming from the upper and lower branches are coupled by C_e to yield the error signal $e(t)$ which should (supposedly) contain the attenuated version of the distortion introduced by the main amplifier.

4.4.3 Error Cancellation Loop

The second stage of feedforward linearization is to eliminate the distortion of main amplifier that has already been identified by *SCL*. The distortion of the amplifier that is identified as $e(t)$ has a different scaling compared to the distortion in the upper branch due to the attenuator L_c. Therefore an error amplifier in Figure 4-7 is used to bring the level of distortion approximately back to the level of upper branch. The scaling and phase is fine tuned by a second vector attenuator placed in the lower branch before the error amplifier. The distortion that is identified and scaled properly is then cancelled from the amplifier output by the coupler C_o, to yield a distortion free, linearized output $z(t)$.

4.4.4 Simplified Baseband Equivalent Model

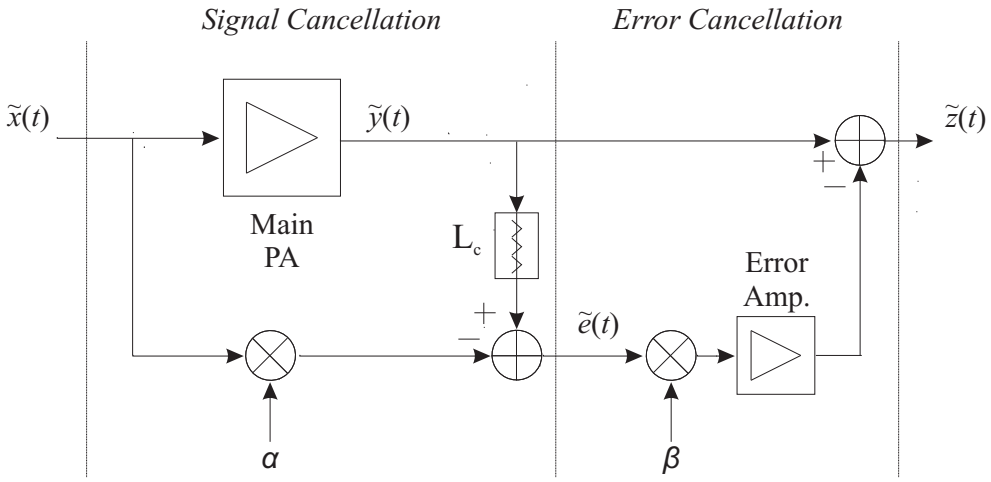


Figure 4-8. The simplified baseband equivalent model for feedforward linearizer.

We will next analyze the feedforward linearizer with a simplified baseband equivalent model shown in Figure 4-8. The vector modulators are replaced by two complex constants α and β, and ideal splitting is assumed for C_i and C_a whereas C_e and C_o are modeled as subtractors. We will also assume that the error amplifier is working in a linear manner with amplitude gain $\sqrt{G_e}$. The main amplifier is first considered to be a memoryless nonlinearity whose input signal is one of the types discussed in Chapter 3, enabling a characterization given by (3.4), i.e. $\tilde{y}(t) = \alpha_o \tilde{x}(t) + \tilde{d}(t)$. Then the error signal can be written as

$$\begin{aligned} \tilde{e}(t) &= \sqrt{L_c}\tilde{y}(t) - \alpha\tilde{x}(t) \\ &= \sqrt{L_c}(\alpha_o\tilde{x}(t) + \tilde{d}(t)) - \alpha\tilde{x}(t) \\ &= (\alpha_o\sqrt{L_c} - \alpha)\tilde{x}(t) + \sqrt{L_c}\tilde{d}(t) \end{aligned} \tag{4.1}$$

The expression (4.1) shows that when $\alpha = \alpha_o\sqrt{L_c}$, the only content of the error signal $\tilde{e}(t)$ will be the attenuated version of the *IMD* terms, i.e. $\sqrt{L_c}\tilde{d}(t)$. Once the *IMD* terms are identified by *SCL*, the error signal is processed by *ECL* to yield

$$\begin{aligned} \tilde{z}(t) &= \tilde{y}(t) - \beta\sqrt{G_e}\tilde{e}(t) \\ &= \alpha_o\tilde{x}(t) - (1 - \beta\sqrt{G_eL_c})\tilde{d}(t) \end{aligned} \tag{4.2}$$

which basically tells that for $\beta = 1/\sqrt{G_eL_c}$, the *IMD* terms are removed from the output of the amplifier yielding only linear amplification. Here we assumed that the error amplifier does not introduce a phase term and simply scales the signal, implying a real valued coefficient β.

However, the analysis can be extended to complex case in a fairly easy way by just placing an additional multiplicative term $e^{j\theta_e}$ to the second term of first line of (4.2) with θ_e being the constant phase shift introduced by the error amplifier. In that case β should also compensate for this phase shift as well. However for the rest of this work an error amplifier that introduces only amplitude scaling is considered.

Based on the above, from linearization (*IMD* cancellation) point of view, the optimum values for *SCL* and *ECL* coefficients are

$$\begin{aligned} \alpha_{opt} &= \alpha_o \sqrt{L_c} \\ \beta_{opt} &= 1/\sqrt{G_e L_c} \end{aligned} \tag{4.3}$$

and the corresponding linearizer output with these coefficients is

$$\tilde{z}(t)\,|_{\alpha_{opt},\beta_{opt}} = \alpha_o \tilde{x}(t) \tag{4.4}$$

The expression (4.4) points out that an *IMD*-free, perfectly linearized signal can be obtained independent of the type of input signal (bandwidth, *PAPR* …etc) and the exact knowledge for the behaviour of memoryless nonlinearity. However, there are certain practical considerations in real world implementations causing imperfect linearization. These can mainly be categorized under error amplifier demands, delay and coefficient mismatches in the loops that are addressed in more details below.

4.4.5 Error Amplifier Demands

The error amplifier in *ECL* is assumed to be working linearly when deriving the optimum coefficients in (4.3) and the linearizer output in (4.4). The output power of the error amplifier should only be as strong as the power of the *IMD* terms which is much smaller than the power levels of the main power amplifier. Therefore the assumption of linear error amplifier is usually valid.

Another important issue discussed in [28] is the coupling loss of C_o. In order not to lower the overall power gain, small coupling ratio is chosen for the upper branch of *ECL*. Thereby the main amplifier goes through a small attenuation whereas the error signal experiences much more loss. This loss should be compensated by the error amplifier, which increases the demands on it.

Multiple feedforward loops are discussed in [22] to overcome any further distortion introduced by the error amplifier if linear operation assumption no longer holds.

4.4.6 Timing Synchronization

The timing synchronization between the upper and lower branches of *SCL* and *ECL* is another issue that must be handled. In other words, the delays shown in Figure 4-7 should be adjusted such that the signals with correct timing instants are subtracted. This delay between the upper and lower branches of the loops are tried to be matched at the operating center frequency with the delay line. However, the imperfections of the components and the signal bandwidth becomes a limiting factor for exact matching.

The impacts of timing mismatch on the cancellation of corresponding signals are not very easy to analyze. The signal characteristics such as symbol rate and applied pulse-shaping are also playing role on the degree of cancellation in the presence of timing mismatch. If we consider the difference of two signals with equal amplitudes and phases but with a delay mismatch of ε_τ

$$w(t) = v(t - \varepsilon_\tau / 2) - v(t + \varepsilon_\tau / 2) \tag{4.5}$$

Then the corresponding spectrum for the difference signal is

$$\begin{aligned} W(f) &= A\,V(f)(e^{-j2\pi f\varepsilon_\tau/2} - e^{j2\pi f\varepsilon_\tau/2}) \\ &= j2V(f)\sin(2\pi f\varepsilon_\tau/2) \end{aligned} \tag{4.6}$$

considering that $V(f)$ is the original signal to be cancelled, then we have a frequency selective cancellation that can be given as

$$\begin{aligned} H(f) &= j2\sin(\pi f\varepsilon_\tau) \\ &\approx j2\pi f\varepsilon_\tau \end{aligned} \tag{4.7}$$

where the approximation is valid for $|f\varepsilon_\tau| << 1$. Similar expression is also derived in [12]. We see from (4.7) that the signal bandwidth and center frequency are at least the two signal characteristics affecting the degree of cancellation in the presence of delay mismatch.

4.4.7 Coefficient Sensitivity

In this thesis, we will mainly focus on the sensitivity of *IMD* suppression to the *SCL* and *ECL* coefficients α and β. We can extend the linearizer output given in (4.2) to the case that α is arbitrary and does not necessarily match the optimum value

$$\begin{aligned}\tilde{z}(t) &= \tilde{y}(t) - \beta\sqrt{G_e}\tilde{e}(t) \\ &= [\alpha_o\tilde{x}(t) + \tilde{d}(t)] - [\beta\sqrt{G_e}((\alpha_o\sqrt{L_c} - \alpha)\tilde{x}(t) + \sqrt{L_c}\tilde{d}(t))] \\ &= [\alpha_o - \beta\sqrt{G_e}(\alpha_o\sqrt{L_c} - \alpha)]\tilde{x}(t) + (1 - \beta\sqrt{G_eL_c})\tilde{d}(t)\end{aligned} \tag{4.8}$$

then defining the mismatches between the optimum and actual values of coefficients as

$$\varepsilon_\alpha = \alpha_{opt} - \alpha = \alpha_o\sqrt{L_c} - \alpha \tag{4.9}$$

$$\varepsilon_\beta = \beta_{opt} - \beta = 1/\sqrt{L_cG_e} - \beta \tag{4.10}$$

the expression (4.8) can be re-written as

$$\tilde{z}(t) = (\alpha_o - \beta\sqrt{G_e}\varepsilon_\alpha)\tilde{x}(t) + \sqrt{G_eL_c}\varepsilon_\beta\tilde{d}(t) \tag{4.11}$$

The cofactor of the first term in (4.11) is the overall amplitude gain and the second term is the *IMD* term at the output of the linearizer which we shall denote as

$$IMD_o = \sqrt{G_eL_c}\varepsilon_\beta\tilde{d}(t) \tag{4.12}$$

Its power can easily be related to the power of *IMD*

$$P_{IMDo} = G_eL_c\left|\varepsilon_\beta\right|^2 P_{IMDa} \tag{4.13}$$

and from (4.13) it is easy to see that the power of the IMD_o is directly proportional to the absolute square of the mismatch in β. We can define the suppression of *IMD* based on (4.13) as

$$IMD_{\sup} = 10 * \log(\frac{P_{IMDa}}{P_{IMDo}}) = 10 * \log(1/\left|\varepsilon_\beta\right|^2) \tag{4.14}$$

If we denote the normalized mismatch as $\varepsilon_\beta^n = (\beta_{opt} - \beta)/\beta_{opt}$, then for $\varepsilon_\beta^n = \pm 10^{-m}$ with m being a positive integer, an increase in m by one corresponds to a 20 dB improvement in IMD_{sup}. This is illustrated in Figure 4-9 where feedforward linearization is applied to an *SSPA* amplifier (parameters same as in Section 3.2) with α being fixed to α_{opt} and m is varied from 1 to 3. The corresponding spectrum for linearizer output is shown in Figure 4-10.

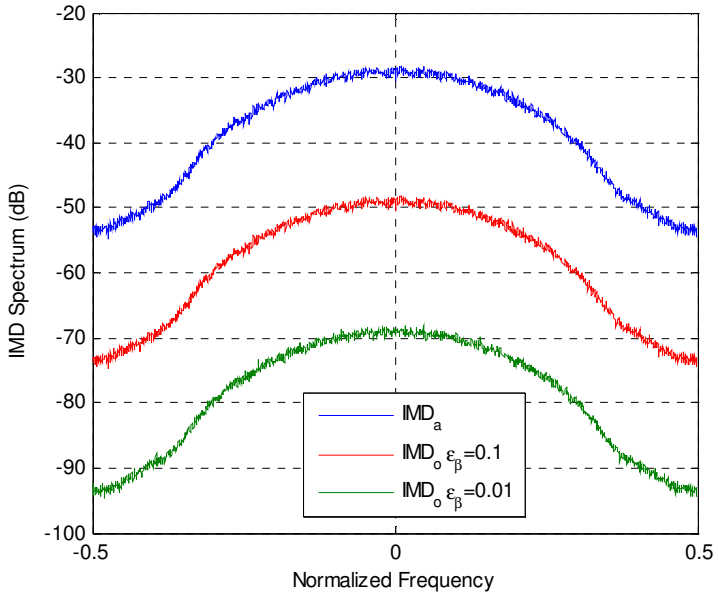


Figure 4-9. The IMD spectrum for $\alpha = \alpha_{opt}$ *and* ε_{β}^{n} *set to* 0.1, 0.01 *respectively.*

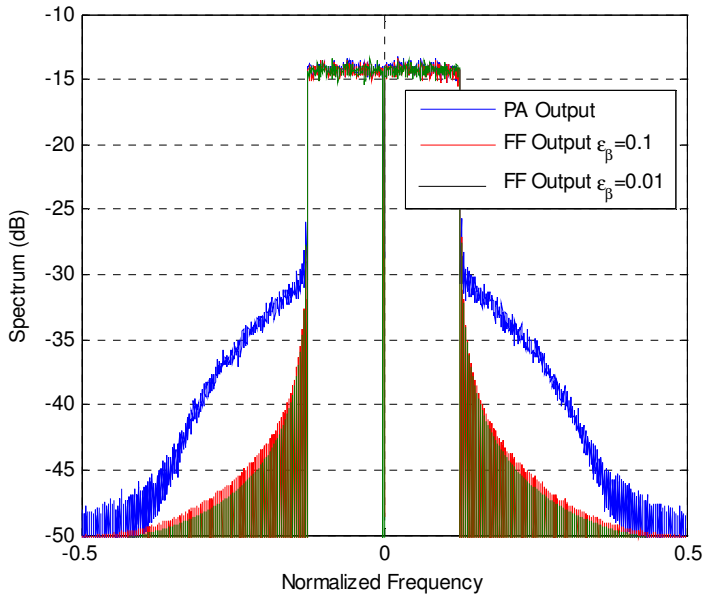


Figure 4-10. The power spectral densities of amplifier and linearizer outputs for $\alpha = \alpha_{opt}$, *and* ε_{β}^{n} *set to* 0.1, 0.01 *respectively.*

We have only discussed the mismatch in β which is apparently the only factor affecting the amount of *IMD* suppression according to (4.13). However, ε_{α} is also critical for at least two important reasons. According to (4.13) it is easy to observe the effect of ε_{α} on the overall amplitude gain of the useful signal at the linearizer output. Positive values of ε_{α} will decrease

the overall amplitude gain, whereas more gain can be achieved with a negative ε_α. On the other hand, it is not feasible to set β to β_{opt} and choose arbitrarily large negative mismatch in α, to obtain a linearized output with high gain. This is due to the second and rather implicit impact that ε_α has on the feedforward circuit. The greater the value of ε_α (either positive or negative), the greater the power of the signal going through the error amplifier, and the error amplifier will no longer be operating linearly. In summary, ε_α should be small enough in order not to cause significant changes on the gain of the useful signal and keep the signal levels low enough for distortionless error amplification process.

4.4.8 Feedforward Linearization of Amplifiers with Memory

The concept of separating the linear amplification and rest of mapping done by the amplifier is one very powerful feature of feedforward linearization. All the above analysis presented for a memoryless nonlinearity can be extended to a nonlinearity with memory with only few modifications. The major difference is that there will be additional terms other than *IMD* but still they will be identified by the *SCL* and eventually subtracted from the amplifier output by *ECL*. This can be analytically shown for some of the memory models and for generality of presentation we will consider Wiener-Hammerstein model below.

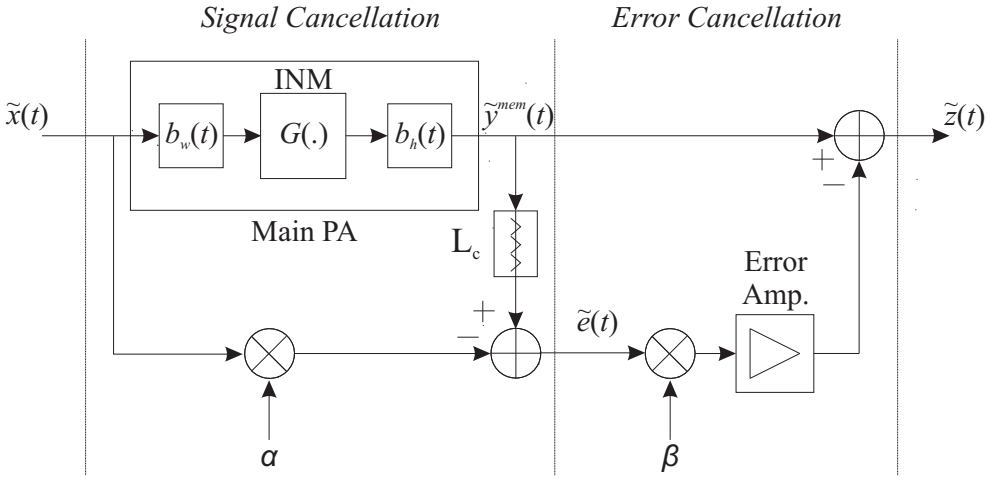


Figure 4-11. Feedforward Linearization applied to a nonlinearity with Wiener-Hammerstein memory with b_w *and* b_h *being the Wiener and Hammerstein filters respectively.*

The modified structure with Wiener-Hammerstein memory model is shown in Figure 4-11. We first define the output right after the memoryless nonlinearity, which can be seen as a Wiener model

$$\begin{aligned}\tilde{y}_w(t) &= \alpha_o^w \int_{-\infty}^{\infty} b_w(\tau)\tilde{x}(t-\tau)d\tau + \tilde{d}_w(t) \\ &= \tilde{u}(t) + \tilde{d}_w(t)\end{aligned} \tag{4.15}$$

Note that we have denoted the complex gain and *IMD* term as α_o^w and $\tilde{d}_w$. The obvious reason is that they both depend on the distribution of the signal before the memoryless nonlinearity which is in this case the filtered process $\tilde{x}_w(t) = \int_{-\infty}^{\infty} b_w(\tau)\tilde{x}(t-\tau)d\tau$. However, in order to be able to compare with the memoryless case, we will consider a mild memory for the Wiener filter wherever appropriate throughout this work such that $\alpha_o^w \approx \alpha_o$ and $\sigma^2_{\tilde{d}_w} \approx \sigma^2_{\tilde{d}}$ are maintained. Having said that, the final output of the amplifier with memory is

$$\begin{aligned}\tilde{y}^{mem}(t) &= \int_{-\infty}^{\infty} b_h(\varsigma)\tilde{u}(t-\varsigma)d\varsigma + \int_{-\infty}^{\infty} b_h(\varsigma)\tilde{d}(t-\varsigma)d\varsigma \\ &= \alpha_o \int_{-\infty}^{\infty} b_h(\varsigma) \int_{-\infty}^{\infty} b_w(\tau)\tilde{x}(t-\tau-\varsigma)d\tau d\varsigma + \int_{-\infty}^{\infty} b_h(\varsigma)\tilde{d}(t-\varsigma)d\varsigma\end{aligned} \tag{4.16}$$

The expression (4.16) can be further simplified if we define the cascade of the Wiener and Hammerstein filters as

$$\begin{aligned}b_{wh}(t) &= b_w(t) \otimes b_h(t) \\ &= \int_{-\infty}^{\infty} b_w(\tau)b_h(t-\tau)d\tau\end{aligned} \tag{4.17}$$

then we can re-write (4.16) as

$$\tilde{y}^{mem}(t) = \alpha_o \int_{-\infty}^{\infty} b_{wh}(\tau)\tilde{x}(t-\tau)d\tau + \int_{-\infty}^{\infty} b_h(\tau)\tilde{d}(t-\tau)d\tau \tag{4.18}$$

Note that the derivations can be very easily adapted to plain Wiener or Hammerstein models as well by the following substitutions into (4.18)

$$\begin{aligned}b_{wh}(t) &= \begin{cases} b_h(t) & Hammerstein \\ b_w(t) & Wiener \end{cases} \\ b_h(t) &= \begin{cases} b_h(t) & Hammerstein \\ \delta(t) & Wiener \end{cases}\end{aligned} \tag{4.19}$$

As (4.19) implies, the input is either filtered with b_h or b_w depending on the model used whereas the filtering of *IMD* is done only for the Hammerstein model.

In the following, we will be using the discrete time equivalents of the signals $\tilde{x}(t)$, $\tilde{y}(t)$, $\tilde{d}(t)$, $\tilde{e}(t)$ and $\tilde{z}(t)$ in order to analytically show that the performance of feedforward linearization is independent of the memory effects. We will also assume that the Wiener and Hammerstein filters are *FIR* with orders *K* and *L* respectively. Then (4.18) will be modified as

$$\begin{aligned}\tilde{y}^{mem}[n] &= \alpha_o \sum_{m=0}^{K+L} b_{wh}[m]\tilde{x}[n-m] + \sum_{l=0}^{L} b_h[l]\tilde{d}[n-l] \\ &= \alpha_o b_{wh}[0]\tilde{x}[n] + \alpha_o \sum_{m=1}^{K+L} b_{wh}[m]\tilde{x}[n-m] + \sum_{l=0}^{L} b_h[l]\tilde{d}[n-l]\end{aligned} \tag{4.20}$$

and the error signal going to the lower branch of *ECL* is then given as

$$\begin{aligned}\tilde{e}^{mem}[n] &= (\alpha_o b_{wh}[0]\sqrt{L_c} - \alpha)\tilde{x}[n] \\ &\quad +\alpha_o \sqrt{L_c} \sum_{m=1}^{K+L} b_{wh}[m]\tilde{x}[n-m] + \sqrt{L_c}\sum_{l=0}^{L} b_h[l]\tilde{d}[n-l]\end{aligned} \tag{4.21}$$

Note that for $b_{wh}[0] = 1$, the cofactor of $\tilde{x}[n]$, appearing as the first term of (4.21), is actually ε_α given in (4.9). This will be the assumption for the rest of this thesis. Comparing (4.1) and (4.21), the second and third terms of (4.21) contain past instances of input and *IMD* which is a different feature compared to memoryless case with only instantaneous *IMD*. These terms will also be considered as distortion due to nonlinearity and altogether be denoted as D_{NL}. The exact expressions of the distortion terms for Wiener, Hammerstein and Wiener-Hammerstein models as well as the memoryless case are given below

$$\begin{aligned}D_{NL} &= \tilde{d}[n] \quad (memoryless) \\ D_{NL}^{w} &= \alpha_o \sum_{m=1}^{K} b_w[k]\tilde{x}[n-k] + \tilde{d}[n] \\ D_{NL}^{h} &= \alpha_o \sum_{l=1}^{L} b_h[l]\tilde{x}[n-l] + \sum_{l=0}^{L} b_h[l]\tilde{d}[n-l] \\ D_{NL}^{wh} &= \alpha_o \sum_{m=1}^{K+L} b_{wh}[m]\tilde{x}[n-l] + \sum_{l=0}^{L} b_h[l]\tilde{d}[n-l]\end{aligned} \tag{4.22}$$

The actual output of the linearizer will reveal more about the independence of feedforward method from memory effects.

$$\begin{aligned}\tilde{z}^{mem}[n] &= (\alpha_o\sqrt{L_c} - \beta\sqrt{G_e}\varepsilon_\alpha)\tilde{x}[n] \\ &+\alpha_o(1-\beta\sqrt{G_eL_c})\sum_{m=1}^{K+L} b_{wh}[m]\tilde{x}[n-m] \\ &+(1-\beta\sqrt{G_eL_c})\sum_{l=0}^{L} b_h[l]\tilde{d}[n-l]\end{aligned} \tag{4.23}$$

Now combining (4.22) with (4.23) we can write

$$\begin{aligned}\tilde{z}^{mem}[n] &= (\alpha_o - \beta\sqrt{G_e}\varepsilon_\alpha)\tilde{x}[n] + (1-\beta\sqrt{G_eL_c})D_{NL}^{wh} \\ &= (\alpha_o - \beta\sqrt{G_e}\varepsilon_\alpha)\tilde{x}[n] + \sqrt{G_eL_c}\varepsilon_\beta D_{NL}^{wh}\end{aligned} \tag{4.24}$$

The only difference between (4.11) and (4.24) is the terms that are considered as distortion. Therefore given the same *SCL* and *ECL* coefficients, the amount of suppression of distortion terms is same independent of memory. Similar analysis is carried out in [34] for only Wiener type memory and in general agrees the derivations here. Thus the optimum coefficients derived for memoryless case in (4.3) are still the optimum values since when plugged-in to (4.24) yield perfectly linearized output $\tilde{z}_{opt}^{mem}[n] = \alpha_o\tilde{x}[n]$. Therefore we can state that

$$\alpha_{opt}^{mem} = \alpha_{opt} = \alpha_o\sqrt{L_c}, \quad \beta_{opt}^{mem} = \beta_{mem} = \frac{1}{\sqrt{G_eL_c}} \tag{4.25}$$

In the rest of this thesis, only α_{opt} and β_{opt} will be used when referring to the optimum coefficients from linearization point of view since those values are independent of memory.

Chapter 5

Performance Analysis of LMS Adaptation

The feedforward linearizer shown in Figure 4-7 and/or Figure 4-8 with fixed *SCL* and *ECL* coefficients is incapable of tracking possible changes occurring to the parameters of the analog components. For instance, the aging effects will change the value of α_o for the main amplifier, and possibly causing an increase in the mismatch between upper and lower branches of *SCL*. Similarly, the power loss of attenuator and/or the linear gain of the error amplifier might change through time. The mismatch between the upper and lower branches of *ECL* will most probably increase if the coefficient is fixed which eventually will cause a degradation in the performance of linearization. Therefore an adaptive mechanism is essential to maintain the desired performance through the life time of operation.

The gradient based methods implemented both digitally or in analog form are proposed in literature [11]-[13] to adapt *SCL* and *ECL* coefficients. As a difference from the existing works in literature, we will derive the optimal Wiener coefficients for both loops, examine the convergence of both coefficients and derive the achievable *IMD* reduction under *LMS* adaptation. The other novel side of the work is the extension of analysis for nonlinearities with Wiener-Hammerstein memory and comparison with the memoryless case.

5.1 Optimum Wiener Coefficient for *SCL*

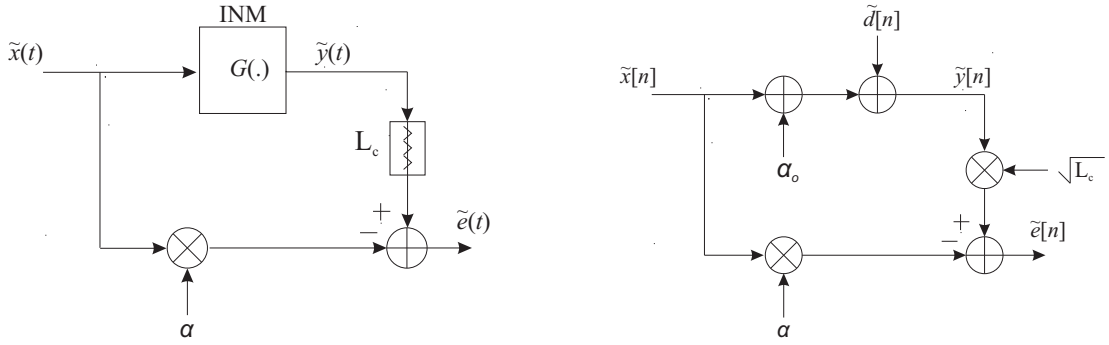
The optimum linearizer coefficient for *SCL*, α_{opt}, is derived in (4.3) and furthermore in (4.25) it is shown to be the optimal coefficient even when the memory effects are included. The multiplication of the lower branch of *SCL* with a complex coefficient as in Figure 4-8 can be viewed as a trivial filtering operation causing a scaling and a phase rotation only. Although this is true for both the continuous and discrete type of input signals, we will concentrate on the discrete case to be able to derive the so-called optimum Wiener filter coefficient as

formulated in [20]. We will first do the derivation for a memoryless power amplifier in the upper branch of *SCL* and then extend this derivation to Wiener-Hammerstein memory model.

5.1.1 Memoryless Model

We start the derivation by first replacing the original *SCL* shown in both Figure 4-8 and Figure 5-1 a) with an equivalent discrete baseband model as in Figure 5-1 b). The main amplifier is considered to be a memoryless nonlinearity and the input signal is assumed to obey the conditions for the model in (3.4) to hold.

.



a). *b)*

Figure 5-1. SCL with a) continuous baseband signals b) discrete baseband signals and main amplifier replaced by the equivalent Bussgang model.

The upper branch of the model presented in Figure 5-1 b) forms the reference signal and we try to find the value of the coefficient α to minimize the mean-squared error between the filter output and reference signal. Thus the cost function to be minimized is

$$J_\alpha = E[|\tilde{e}[n]|^2] \tag{5.1}$$

where the discrete equivalent of the error signal in (4.1) is given by

$$\tilde{e}[n] = (\alpha_o\sqrt{L_c} - \alpha)\tilde{x}[n] - \sqrt{L_c}\tilde{d}[n] \tag{5.2}$$

The optimum Wiener coefficient that minimizes the cost function in (5.1) can now be calculated by setting the gradient of the cost function to zero

$$\nabla_\alpha J[n] = \frac{\partial J_\alpha}{\partial \alpha} = 0 \tag{5.3}$$

Minimizing the gradient is easily shown to be equivalent to

$$E[\tilde{e}^*[n]\tilde{x}[n]] = 0 \tag{5.4}$$

Then we plug-in (5.2) into (5.4) to arrive at the minimum mean-squared error (*MMSE*) solution as

$$\alpha_{MMSE} = \alpha_o\sqrt{L_c} + \sqrt{L_c}E[\tilde{d}^*[n]x[n]] / E[|x[n]|^2] \tag{5.5}$$

Using then the original assumption that $\tilde{d}$ and $\tilde{x}$ are uncorrelated, the second term on the right hand side of (5.5) vanishes and we obtain

$$\alpha_{MMSE} = \alpha_o\sqrt{L_c} \tag{5.6}$$

The same result can also be obtained with the well known solution of Wiener-Hopf equations [20]. If we denote the input sequence to Wiener filter under adaptation as $\tilde{i}[n]$ and the reference signal as $\tilde{r}[n]$, then the auto-correlation of input in this case becomes $R_{\tilde{i}\tilde{i}^*}[j] = R_{\tilde{x}\tilde{x}^*}[j] = E[\tilde{x}[n+j]\tilde{x}^*[n]]$ whereas the cross correlation between the filter input and reference signal is $p_{\tilde{i}\tilde{r}^*}[j] = E[\tilde{x}[n+j]\tilde{y}^*[n]\sqrt{L_c}]$. Then the optimum Wiener filter in this special case of one tap becomes

$$\begin{aligned}\alpha_{MMSE} &= (R_{\tilde{x}\tilde{x}^*}{}^{-1}[0]p_{\tilde{i}\tilde{r}^*}[0])^* \\ &= (\frac{\alpha_o^*\sqrt{L_c}E[|\tilde{x}[n]|^2] + \sqrt{L_c}E[\tilde{x}[n]\tilde{d}^*[n]]}{E[|\tilde{x}[n]|^2]})^* \\ &= \alpha_o\sqrt{L_c}\end{aligned} \tag{5.7}$$

Note that the conjugation of right hand side is done since the filter output is calculated directly by α instead of α^* as a difference to the notation in [20].

An important remark is that the *MMSE* solution is identical to the *SCL* coefficient, α_{opt}, derived in (4.3) which was shown to be optimum from linearization point of view.

5.1.2 Wiener-Hammerstein Memory Model

Now having derived the optimum Wiener coefficient for memoryless nonlinearities, we will now extend the derivation to a generalized Wiener-Hammerstein memory model for which the amplifier output is given by (4.20). Then in that case, the cross-correlation between the input to adapt *SCL* coefficient and reference signal becomes

$$
\begin{aligned}
p_{\tilde{i}\tilde{r}^*}^{mem}[j] &= E[\tilde{x}[n+j](\tilde{y}^{mem}[n])^*\sqrt{L_c}] \\
&= \alpha_o^*\sqrt{L_c}\sum_{m=0}^{M} b_{wh}[m]E[\tilde{x}[n+j]\tilde{x}^*[n-m]] + \sqrt{L_c}\underbrace{\sum_{l=0}^{L} b_h[l]E[\tilde{x}[n+j]\tilde{d}^*[n-l]]}_{0} \\
&= \alpha_o^*\sqrt{L_c}\sum_{m=0}^{M} b_{wh}[m]R_{\tilde{x}\tilde{x}^*}[j-m]
\end{aligned}
\tag{5.8}
$$

and again using the identity in (5.7) we obtain

$$
\begin{aligned}
\alpha_{MMSE}^{mem} &= (R_{\tilde{u}\tilde{u}^*}{}^{-1}[0]p_{\tilde{i}\tilde{r}^*}^{mem}[0])^* \\
&= \frac{\alpha_o\sqrt{L_c}}{P_{\tilde{x}}}\sum_{m=0}^{M} b_{wh}[m]R_{\tilde{x}\tilde{x}^*}[-m]
\end{aligned}
\tag{5.9}
$$

where $P_{\tilde{x}} = R_{\tilde{x}\tilde{x}^*}[0]$.

The expression in (5.9) points out that there is a certain offset between the optimum Wiener and linearizer coefficients which can be written as

$$
\begin{aligned}
\varepsilon_{\alpha}^{off} &= \alpha_{opt} - \alpha_{MMSE}^{mem} \\
&= \alpha_o\sqrt{L_c} - \frac{\alpha_o\sqrt{L_c}}{P_{\tilde{x}}}\sum_{m=0}^{M} b_{wh}[m]R_{\tilde{x}\tilde{x}^*}[-m] \\
&= \alpha_o\sqrt{L_c}\left(1 - \frac{\sum_{m=0}^{M} b_{wh}[m]R_{\tilde{x}\tilde{x}^*}[-m]}{P_{\tilde{x}}}\right) \\
&= \frac{-\alpha_o\sqrt{L_c}}{P_{\tilde{x}}}\sum_{m=1}^{M} b_{wh}[m]R_{\tilde{x}\tilde{x}^*}[-m]
\end{aligned}
\tag{5.10}
$$

The offset is zero if and only if the input is white, i.e. $R_{\tilde{x}\tilde{x}^*}(m)$ is zero for all integers *m* except zero. In general the deviation of *MMSE* solution from the optimum linearizer coefficient, α_{opt}, is determined by the combined effect of the strength of Wiener-Hammerstein filter taps (amount of memory) and auto-correlation of the input signal.

5.2 *LMS* adaptation of *SCL* coefficient

The optimum Wiener coefficient of *SCL* for memoryless case is derived in (5.6) and (5.7). Now we would like to define the procedure for actually adapting this coefficient and investigate its behavior in terms of convergence to the optimum linearizer and Wiener coefficients.

5.2.1 Basic Recursion

The basic recursion of *LMS* for updating the *SCL* coefficient is

$$\begin{aligned} \alpha^*[n+1] &= \alpha^*[n] - \hat{\mu}\hat{\nabla}_\alpha J[n] \\ &= \alpha^*[n] + 2\hat{\mu}\tilde{e}^*[n]x[n] \end{aligned} \tag{5.11}$$

or equivalently

$$\alpha[n+1] = \alpha[n] + 2\hat{\mu}\tilde{e}[n]x^*[n] \tag{5.12}$$

where $\hat{\mu}$ is the step size and $\hat{\nabla}_\alpha J[n]$ is the instantaneous gradient estimate obtained by removing the expectation operation, $E[.]$, from the left hand side of (5.3) yielding an instantaneous product of the error and input signals. Now also denoting $\mu = 2\hat{\mu}$, the well known *LMS* adaptation can be written as

$$\alpha[n+1] = \alpha[n] + \mu\tilde{e}[n]x^*[n] \tag{5.13}$$

5.2.2 Convergence Behaviour of Memoryless Model

The *LMS* adaptation of *SCL* coefficient applying the recursion given in (5.13) for a memoryless amplifier model is shown in Figure 5-2.

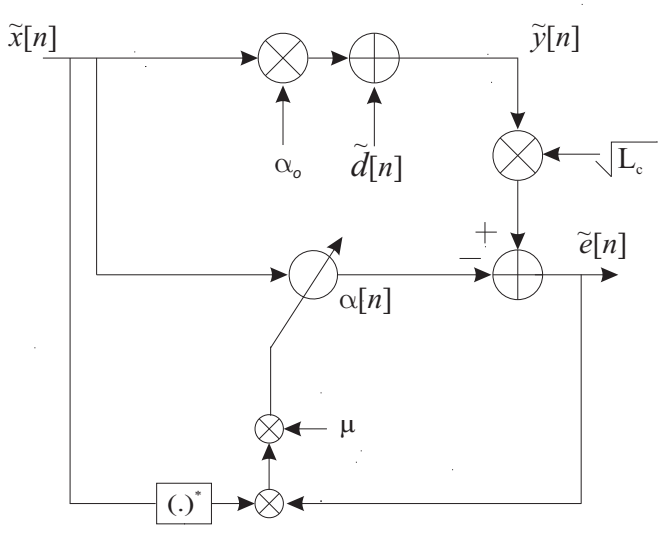


Figure 5-2. LMS adaptation of SCL coefficient for a memoryless amplifier model.

It is shown in Appendix A in (A.6) that $\lim_{n\to\infty} E[\varepsilon_\alpha^{MMSE}[n]] = 0$ for $0 < \mu_\alpha < 2/P_{\tilde{x}}$. This states that $\alpha[n]$ is convergent to α_{MMSE} in the mean sense if the input $\tilde{x}$ is white. Under

the same conditions, also the mean square of the mismatch is given in Appendix (A.10) as $\lim_{n\to\infty} E[\left|\varepsilon_{\alpha}^{MMSE}[n]\right|^2] = \frac{\mu_\alpha L_c P_{IMDa}}{2-\mu_\alpha P_{\tilde{x}}}$.

5.2.3 Convergence Behaviour of Wiener-Hammerstein Memory Model

In Section 5.1.2, the optimum Wiener coefficient assuming a Wiener-Hammerstein type of memory was derived in (5.9) and shown in general to have an offset from the optimum linearizer coefficient, α_{opt}. When the structure of adaptation is kept same with only one coefficient to be adapted, as illustrated in Figure 5-3, then $\alpha[n]$ is shown to be converging to α_{MMSE}^{mem} in the mean sense in Appendix A.

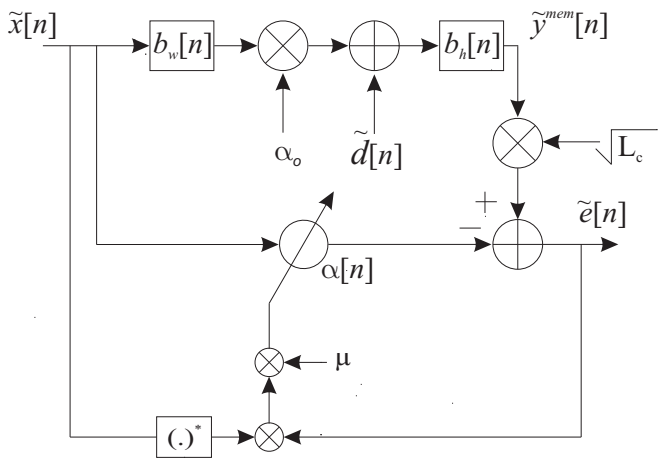


Figure 5-3. LMS adaptation in the presence of Wiener-Hammerstein memory.

The main assumption is once again that the input is white (or we have a small-step size). However, it is more convenient to study the convergence in terms of optimum linearizer coefficient α_{opt} for which we can write the mismatch between the true and optimum value as

$$\begin{aligned}\varepsilon_{\alpha,mem}[n] &= \alpha_{opt} - \alpha^{mem}[n] \\ &= \alpha_{opt} - \alpha_{MMSE}^{mem} + \alpha_{MMSE}^{mem} - \alpha^{mem}[n] \\ &= \varepsilon_{\alpha}^{off} - \varepsilon_{\alpha,mem}^{MMSE}[n]\end{aligned} \tag{5.14}$$

where $\varepsilon_{\alpha,mem}^{MMSE}$ is defined in Appendix A. Then we can write

$$\lim_{n\to\infty} E[\varepsilon_{\alpha,mem}[n]] = \varepsilon_{\alpha}^{off} \tag{5.15}$$

since the second term is shown to be zero in (A.13) and $\varepsilon_{\alpha}^{off}$ has a deterministic value. Expression (5.15) states that the mean convergence of *SCL* coefficient under *LMS* adaptation has constant $\varepsilon_{\alpha}^{off}$ deviation from optimum linearizer coefficient α_{opt}.

We can also study the mean square of this mismatch

$$\begin{aligned}\lim_{n\to\infty} E[|\varepsilon_{\alpha,mem}[n]|^2] &= \lim_{n\to\infty} E[|\varepsilon_{\alpha}^{off} - \varepsilon_{\alpha,mem}^{MMSE}[n]|^2] \\ &= |\varepsilon_{\alpha}^{off}|^2 + E[|\varepsilon_{\alpha,mem}^{MMSE}[n]|^2] - 2\,\mathrm{Re}\{\varepsilon_{\alpha}^{off} E[(\varepsilon_{\alpha,mem}^{MMSE}[n])^*]\} \\ &= |\varepsilon_{\alpha}^{off}|^2 + E[|\varepsilon_{\alpha,mem}^{MMSE}[n]|^2] \\ &\approx |\varepsilon_{\alpha}^{off}|^2 + \frac{\mu_\alpha L_c P_{IMDa}}{2-\mu_\alpha P_{\tilde{x}}} = |\varepsilon_{\alpha}^{off}|^2 + E[|\varepsilon_{\alpha,mem}[n]|^2]\end{aligned} \tag{5.16}$$

where the last term in the second line is dropped due to (A.13). The approximation holds for low auto-correlation values of input and weak memory as explained in Appendix A. Expression (5.16) shows that the mean square mismatch in case of Wiener-Hammerstein memory model additionally contains the term $|\varepsilon_{\alpha}^{off}|^2$.

5.3 Optimum Wiener Coefficient for *ECL*

The optimum linearizer coefficient for *ECL* is derived in (4.3) and furthermore in (4.25) it is shown to be the optimal coefficient even when the memory effects are included. Similar to discussion in Section 5.1, we can view the multiplication with β as a trivial filtering. In order to be able to derive the optimum Wiener coefficient and then proceed towards the *LMS* adaptation, we will again consider *ECL* operating with discrete baseband signals as shown in Figure 5-4. There the error amplifier is considered to be performing only linear amplification.

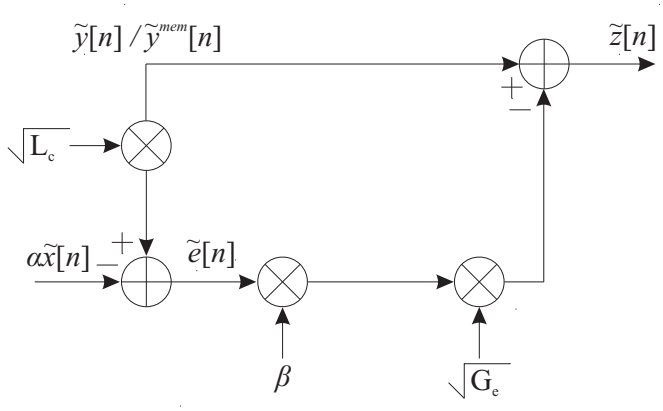


Figure 5-4. ECL with baseband signals and error amplifier replaced by the equivalent amplitude multiplier.

The model presented in Figure 5-4 gives the intuition that processing of *ECL* is conditioned on the value of *SCL* coefficient α. In other words, we would expect the optimum Wiener coefficient for β to depend on α. We will check the validity of this expectation by

first deriving the optimum Wiener coefficient of *ECL* for memoryless amplifier model and then extending the derivation for Wiener-Hammerstein memory model.

5.3.1 Memoryless Model

Now let's first consider that the upper-branch of *ECL* is the output of a memoryless nonlinearity and serves as a reference signal. In that case, the cost function is the power of the linearizer output

$$J_\beta = E[|\tilde{z}[n]|^2] \tag{5.17}$$

where the discrete equivalent of the linearizer output in (4.11) can be written as

$$\tilde{z}[n] = (\alpha_o - \beta[n]\sqrt{G_e}\varepsilon_\alpha[n])\tilde{x}[n] + \sqrt{G_e L_c}\varepsilon_\beta[n]\tilde{d}[n] \tag{5.18}$$

where both α and β are considered to be varying in time. In addition, the time dependent mismatch is defined as $\varepsilon_\alpha[n] = \alpha_{opt} - \alpha[n]$. Then we can obtain the optimum Wiener coefficient for β by again using the well known equality

$$\beta_{MMSE} = R_{\tilde{u}\tilde{u}^*}[0]^{-1} p_{\tilde{u}\tilde{r}^*}[0] \tag{5.19}$$

where the auto-correlation of filter input with lag zero can be written as

$$\begin{aligned} R_{\tilde{u}\tilde{u}^*}[0] &= R_{\tilde{e}\tilde{e}^*}[0] = G_e E[|\tilde{e}[n]|^2] \\ &= G_e(E[|\varepsilon_\alpha[n]|^2 |\tilde{x}[n]|^2] + L_c E[|\tilde{d}[n]|^2]) \end{aligned} \tag{5.20}$$

Note that we have actually integrated the scaling introduced by the error amplifier to (5.20). This is theoretically equivalent to exchanging the order of error amplifier and β so that the signal $\tilde{e}$ is amplified first before the calculation of gradient signal and being fed to β.

In the expression (5.20), the first term in parenthesis is separable only if the amplifier input $\tilde{x}$ is white so that $\varepsilon_\alpha[n]$ depends only on the values of $\tilde{x}$ earlier than *n* and independent of $\tilde{x}[n]$. Under these assumptions (5.20) can be simplified to

$$\begin{aligned} R_{\tilde{e}\tilde{e}^*}[0] &= G_e(E[|\varepsilon_\alpha[n]|^2]E[|\tilde{x}[n]|^2] + L_c E[|\tilde{d}[n]|^2]) \\ &= G_e(E[|\varepsilon_\alpha[n]|^2]P_{\tilde{x}} + L_c P_{IMDa}) \end{aligned} \tag{5.21}$$

Next we write the cross-correlation between the filter input and reference signal as

$$\begin{aligned} p_{\tilde{i}\tilde{r}^*}[0] &= E[\sqrt{G_e}\tilde{e}[n]\tilde{y}^*[n]] \\ &= \sqrt{G_e}E[(\varepsilon_\alpha[n]\tilde{x}[n] + \sqrt{L_c}\tilde{d}[n])(\alpha_o^*\tilde{x}^*[n] + \tilde{d}^*[n])] \\ &= \sqrt{G_e}(\alpha_o^* E[\varepsilon_\alpha[n]\tilde{x}[n]\tilde{x}^*[n]] + \sqrt{L_c}E[\tilde{d}[n]\tilde{d}^*[n]]) \end{aligned} \tag{5.22}$$

where we again implicitly dropped the cross terms between $\tilde{x}$ and $\tilde{d}$. Considering the last line of (5.22), the first expectation in parenthesis is again separable if and only if $\tilde{x}$ is white. Then we can simplify (5.22) to the form

$$\begin{aligned} p_{\tilde{i}\tilde{r}^*}[0] &= \sqrt{G_e}(\alpha_o^* E[\varepsilon_\alpha[n]\tilde{x}[n]\tilde{x}^*[n]] + \sqrt{L_c}E[\tilde{d}[n]\tilde{d}^*[n]]) \\ &= \sqrt{G_e}(\alpha_o^* E[\varepsilon_\alpha[n]]E[\tilde{x}[n]\tilde{x}^*[n]] + \sqrt{L_c}E[\tilde{d}[n]\tilde{d}^*[n]]) \\ &= \sqrt{G_e}(\alpha_o^* E[\varepsilon_\alpha[n]]P_{\tilde{x}} + \sqrt{L_c}P_{IMDa}) \end{aligned} \tag{5.23}$$

Finally the optimum Wiener coefficient for *ECL* given in (5.19) becomes

$$\beta_{MMSE} = \frac{\alpha_o E[\varepsilon_\alpha^*[n]]P_{\tilde{x}} + \sqrt{L_c}P_{IMDa}}{\sqrt{G_e}(E[|\varepsilon_\alpha[n]|^2]P_{\tilde{x}} + L_c P_{IMDa})} \tag{5.24}$$

Our first intuition after the model in Figure 5-4 was that β_{MMSE} is likely to depend on the value of α. This is now shown to be the case by (5.24) where β depends both on the mean and mean square of the mismatch in α. Furthermore, if we consider that α is also a coefficient that is being adapted by *LMS* (but has reached to steady state), the term $E[\varepsilon_\alpha^*[n]]$ is zero due to (A.6) and $E[|\varepsilon_\alpha[n]|^2]$ is equivalent to $E[|\varepsilon_\alpha^{MMSE}[n]|^2]$. The simplified form of (5.24) in that case is

$$\beta_{MMSE} = \frac{\sqrt{L_c}P_{IMDa}}{\sqrt{G_e}(E[|\varepsilon_\alpha^{MMSE}[n]|^2]P_{\tilde{x}} + L_c P_{IMDa})} \tag{5.25}$$

We can also consider β_{MMSE} under a deterministic mismatch in α. This being the case, we have

$$\beta_{MMSE}\,|_{\varepsilon_\alpha} = \frac{\alpha_o\varepsilon_\alpha^* P_{\tilde{x}} + \sqrt{L_c}P_{IMDa}}{\sqrt{G_e}(|\varepsilon_\alpha|^2 P_{\tilde{x}} + L_c P_{IMDa})} \tag{5.26}$$

Note that when $E[\varepsilon_\alpha^*[n]]$, $E[|\varepsilon_\alpha[n]|^2]$ in (5.24) and/or ε_α^*, $|\varepsilon_\alpha|^2$ in (5.26) are zero, then $\beta_{MMSE} = \beta_{opt}$ and/or $\beta_{MMSE}\,|\,\varepsilon_\alpha = \beta_{opt}$. For the case of deterministic mismatch in α, the condition that ε_α^*, $|\varepsilon_\alpha|^2$ are being zero implies an injection of α that perfectly matches α_{opt}. In the more interesting case where α is also being adapted, then $E[\varepsilon_\alpha^*[n]]$ is zero under the conditions given in Appendix A. However, $E[|\varepsilon_\alpha[n]|^2]$ can only be approximately zero if the

step size used to adapt α is very small or a theoretical steepest-descent adaptation with true gradient is applied. In order to summarize all these, we write

$$\begin{aligned} &\alpha \text{ deterministic: } \beta_{MMSE} \mid \varepsilon_\alpha = \beta_{opt}, \quad if\ \alpha[n] = \alpha_{opt}\ (\varepsilon_\alpha = 0) \\ &\alpha \text{ adapted:} \begin{cases} \beta_{MMSE} = \beta_{opt}, & if\ \nabla_\alpha J[n]\ is\ used \\ \beta_{MMSE} \approx \beta_{opt}, & if\ \mu_\alpha \approx 0 \end{cases} \end{aligned} \tag{5.27}$$

It is also beneficial to define an offset term for β as we have done for α in (5.10). However, this offset is mostly due to the mismatch in α (as can be seen from the conditions given above) instead of the memory and can be obtained by subtracting (5.25) from (4.3)

$$\begin{aligned} \beta_{off} &= \beta_{opt} - \beta_{MMSE} \\ &= \frac{1}{\sqrt{G_e L_c}} - \frac{\sqrt{L_c} P_{IMDa}}{\sqrt{G_e}(E[\left|\varepsilon_\alpha^{MMSE}[n]\right|^2] P_{\tilde{x}} + L_c P_{IMDa})} \\ &= \frac{P_{\tilde{x}}(E[\left|\varepsilon_\alpha^{MMSE}[n]\right|^2]}{\sqrt{G_e L_c}(E[\left|\varepsilon_\alpha^{MMSE}[n]\right|^2] P_{\tilde{x}} + L_c P_{IMDa})} \end{aligned} \tag{5.28}$$

We see from (5.28) that the converged value of β has an offset from the optimum linearizer coefficient and this offset is a function of mean square of the steady-state noise on top of the converged value of α.

5.3.2 Wiener-Hammerstein Memory Model

We can easily extend the derivations of previous section to the case where the output of the amplifier is given as (4.20). Similar to (5.17) we will write the cost function as

$$J_\beta^{mem} = E[\left| z^{mem}[n] \right|^2] \tag{5.29}$$

where the modified linearizer output given in (4.23) with coefficients varying with time is given by

$$\begin{aligned} \tilde{z}^{mem}[n] = &(\alpha_o \sqrt{L_c} - \beta[n]\sqrt{G_e}\varepsilon_\alpha[n])\tilde{x}[n] \\ &+\alpha_o(1 - \beta[n]\sqrt{G_e L_c}) \sum_{m=1}^{K+L} b_{wh}[m]\tilde{x}[n-m] \\ &+(1 - \beta[n]\sqrt{G_e L_c}) \sum_{l=0}^{L} b_h[l]\tilde{d}[n-l] \end{aligned} \tag{5.30}$$

We will also modify (4.21) considering that α is varying with time, and write

$$\begin{aligned}\tilde{e}^{mem}[n] &= (\alpha_o\sqrt{L_c} - \alpha[n])\tilde{x}[n] \\ &\quad +\alpha_o\sqrt{L_c}\sum_{m=1}^{K+L} b_{wh}[m]\tilde{x}[n-m] + \sqrt{L_c}\sum_{l=0}^{L} b_h[l]\tilde{d}[n-l] \\ &= \varepsilon_\alpha[n]\tilde{x}[n] + \alpha_o\sqrt{L_c}\sum_{m=1}^{K+L} b_{wh}[m]\tilde{x}[n-m] + \sqrt{L_c}\sum_{l=0}^{L} b_h[l]\tilde{d}[n-l]\end{aligned} \tag{5.31}$$

Then the auto-correlation of the filter input given originally in (5.20) will be modified as

$$\begin{aligned}R_{\tilde{e}\tilde{e}^*}[0] &= E[\tilde{e}^{mem}[n](\tilde{e}^{mem}[n])^*] \\ &= E[\left|\varepsilon_{\alpha,mem}[n]\right|^2 |\tilde{x}[n]|^2] + 2\,\mathrm{Re}\{\alpha_o\sqrt{L_c}\sum_{m=1}^{M} b_{wh}[m]E[\varepsilon^*_{\alpha,mem}[n]\tilde{x}[n-m]\tilde{x}^*[n]]\} \\ &\quad +|\alpha_o|^2 L_c\sum_{m=1}^{M}\sum_{r=1}^{M} b_{wh}[m]b^*_{wh}[r]E[\tilde{x}[n-m]\tilde{x}^*[n-r]] \\ &\quad +L_c\sum_{l=0}^{L-1}\sum_{k=0}^{L-1} b_h[l]b^*_h[k]E[\tilde{d}[n-l]\tilde{d}^*[n-k]]\end{aligned} \tag{5.32}$$

If we are again considering that α is also being adapted by *LMS* adaptation as described in Section 5.2, i.e. $\varepsilon_\alpha[n]$ depends on values of $\tilde{x}$ prior to instant *n*, then the expectation in the first term of (5.32) is separable if and only if $\tilde{x}$ is white. However, even white $\tilde{x}$ will not be sufficient to separate the product in the second term. In this second term, there are products of the form $\varepsilon^*_{\alpha,mem}[n]\tilde{x}[n-m]$ and since $\varepsilon_{\alpha,mem}[n]$ depends on values of $\tilde{x}$ prior to instant *n,* we cannot blindly separate the expectation. At this point, we can only assume that the step-size used in the adaptation of α, μ_α, is small so that the dependency of $\varepsilon_{\alpha,mem}[n]$ to $\tilde{x}$ at such instants is not high. Then we can make an approximation and separate those products to simplify (5.32) as

$$\begin{aligned}R_{\tilde{e}\tilde{e}^*}[0] &= E[\left|\varepsilon_{\alpha,mem}[n]\right|^2]P_{\tilde{x}} + 2\,\mathrm{Re}\{\alpha_o\sqrt{L_c}E[\varepsilon^*_{\alpha,mem}[n]]\sum_{m=1}^{M} b_{wh}[m]R_{\tilde{x}\tilde{x}^*}[-m] \\ &\quad +|\alpha_o|^2 L_c\sum_{m=1}^{M}\sum_{r=1}^{M} b_{wh}[m]b^*_{wh}[r]R_{\tilde{x}\tilde{x}^*}[r-m] \\ &\quad +L_c\sum_{l=0}^{L-1}\sum_{k=0}^{L-1} b_h[l]b^*_h[k]R_{\tilde{d}}[k-l]\end{aligned} \tag{5.33}$$

The cross-correlation between the filter input and reference signal is given by

$$
\begin{aligned}
p_{\tilde{i}\tilde{r}^*}[0] &= E[\sqrt{G_e}e^{mem}[n](y^{mem}[n])^*] \\
&= \sqrt{G_e}(\alpha_o^* E[\varepsilon_{\alpha,mem}[n]\tilde{x}[n]\tilde{x}^*[n]] + \alpha_o^* \sum_{m=1}^{M} b_{wh}^*[m]E[\varepsilon_{\alpha,mem}[n]\tilde{x}[n]\tilde{x}^*[n-m]] \\
&+|\alpha_o|^2 \sqrt{L_c}\sum_{m=1}^{M} b_{wh}[m]E[\tilde{x}^*[n]\tilde{x}[n-m]] \\
&+|\alpha_o|^2 \sqrt{L_c}\sum_{m=1}^{M}\sum_{r=1}^{M} b_{wh}[m]b_{wh}^*[r]E[\tilde{x}[n-m]\tilde{x}^*[n-r]] \\
&+\sqrt{L_c}\sum_{l=0}^{L}\sum_{k=0}^{K} b_h[l]b_h^*[k]E[\tilde{d}[n-l]\tilde{d}^*[n-k]])
\end{aligned}
\tag{5.34}
$$

We will again assume the same conditions as before to make the expectations in first and second term of (5.34) separable over products of expectations yielding

$$
\begin{aligned}
p_{\tilde{i}\tilde{r}^*}[0] = \sqrt{G_e}(&\alpha_o^* E[\varepsilon_{\alpha,mem}[n]]P_{\tilde{x}} + \alpha_o^* E[\varepsilon_{\alpha,mem}[n]]\sum_{m=1}^{M} b_{wh}^*[m]R_{\tilde{x}}[m] \\
&+|\alpha_o|^2 \sqrt{L_c}\sum_{m=1}^{M} b_{wh}[m]R_{\tilde{x}}[-m] \\
&+|\alpha_o|^2 \sqrt{L_c}\sum_{m=1}^{M}\sum_{r=1}^{M} b_{wh}[m]b_{wh}^*[r]R_{\tilde{x}}[r-m] \\
&+\sqrt{L_c}\sum_{l=0}^{L}\sum_{k=0}^{K} b_h[l]b_h^*[k]R_{\tilde{d}}[k-l])
\end{aligned}
\tag{5.35}
$$

By denoting then

$$
\begin{aligned}
T_1 &= \sum_{m=1}^{M} b_{wh}[m]R_{\tilde{x}\tilde{x}^*}^*[m] \\
T_2 &= \sum_{m=1}^{M}\sum_{r=1}^{M} b_{wh}^*[m]b_{wh}[r]R_{\tilde{x}\tilde{x}^*}^*[r-m] \\
T_3 &= \sum_{l=0}^{L}\sum_{k=0}^{L} b_h^*[l]b_h[k]R_{\tilde{d}}^*[k-l]
\end{aligned}
\tag{5.36}
$$

we are finally ready to write the optimum Wiener coefficient by plugging-in (5.33), (5.35) into (5.19) and using (5.36) as

$$
\beta_{MMSE}^{mem} = \frac{(\alpha_o E[\varepsilon_{\alpha,mem}^*[n]]P_{\tilde{x}} + \alpha_o E[\varepsilon_{\alpha,mem}^*[n]]T_1 + |\alpha_o|^2\sqrt{L_c}(T_1^* + T_2) + \sqrt{L_c}T_3}{\sqrt{G_e}(E[|\varepsilon_{\alpha,mem}|^2]P_{\tilde{x}} + 2\,\mathrm{Re}\{\alpha_o\sqrt{L_c}E[\varepsilon_{\alpha,mem}^*[n]]T_1\} + |\alpha_o|^2 L_c T_2 + L_c T_3)} \tag{5.37}
$$

The optimum coefficient can again be conditioned on the value of $\alpha[n]$ and therefore $\varepsilon_\alpha[n]$

$$\beta_{MMSE}^{mem}\ |_{\varepsilon_\alpha} = \frac{(\alpha_o \varepsilon_{\alpha,mem}^{*} P_{\tilde{x}} + \alpha_o \varepsilon_{\alpha,mem}^{*} T_1\ + |\alpha_o|^2 \sqrt{L_c}(T_1^{*} + T_2) + \sqrt{L_c} T_3}{\sqrt{G_e}(\left|\varepsilon_{\alpha,mem}\right|^2 P_{\tilde{x}} + 2\,\mathrm{Re}\{\alpha_o \sqrt{L_c} \varepsilon_{\alpha,mem}^{*} T_1\} + |\alpha_o|^2 L_c T_2 + L_c T_3)} \tag{5.38}$$

Note that the conditions, white input and/or small μ_α, which were assumed to hold when deriving (5.33) and (5.35) are not necessary for this conditional case given by (5.38).

It was shown earlier for the memoryless case that under the conditions stated in (5.27), $\beta_{MMSE} = \beta_{opt}$. We also would like to see whether this property still holds when the corresponding conditions for the case with memory are met. When there is memory, the mean and mean square of the mismatch $\varepsilon_{\alpha,mem}$ are given in (5.15) and (5.16) respectively. If we then consider pure steepest-descent adaptation method, the last term in (5.16) is zero, yielding $E[\left|\varepsilon_{\alpha,mem}\right|^2] = \left|\varepsilon_\alpha^{off}\right|^2$. Then inserting (5.15) and this simplified form of (5.16) into (5.37) we obtain

$$\begin{aligned}
&(E[\left|\varepsilon_{\alpha,mem}\right|^2] = \left|\varepsilon_\alpha^{off}\right|^2,\ E[\varepsilon_{\alpha,mem}[n]] = \varepsilon_\alpha^{off}) \\
&\beta_{MMSE}^{mem} = \frac{(-|\alpha_o|^2 \sqrt{L_c} T_1^{*} - |\alpha_o|^2 |T_1|^2 \sqrt{L_c} / P_{\tilde{x}} + |\alpha_o|^2 \sqrt{L_c}(T_1^{*} + T_2) + \sqrt{L_c} T_3}{\sqrt{G_e}(-|\alpha_o|^2 |T_1|^2 L_c / P_{\tilde{x}} + |\alpha_o|^2 L_c T_2 + L_c T_3)} \\
&\quad = \frac{(-|\alpha_o|^2 |T_1|^2 \sqrt{L_c} / P_{\tilde{x}} + |\alpha_o|^2 \sqrt{L_c} T_2 + \sqrt{L_c} T_3)}{\sqrt{G_e L_c}(-|\alpha_o|^2 |T_1|^2 \sqrt{L_c} / P_{\tilde{x}} + |\alpha_o|^2 \sqrt{L_c} T_2 + \sqrt{L_c} T_3)} \\
&\quad = \frac{1}{\sqrt{G_e L_c}} = \beta_{opt}
\end{aligned} \tag{5.39}$$

This shows that in case of power amplifier with memory, and when α is adapted with steepest-descent method (or *LMS* with very small step size) such that $E[\left|\varepsilon_{\alpha,mem}^{MMSE}[n]\right|^2] = 0$ and $E[\left|\varepsilon_{\alpha,mem}[n]\right|^2] = \left|\varepsilon_\alpha^{off}\right|^2$, then $\beta_{MMSE}^{mem} = \beta_{opt}$.

The above analysis can further be extended to the general case of *LMS* adaptation where both $E[\left|\varepsilon_\alpha^{MMSE}[n]\right|^2]$ and $E[\left|\varepsilon_{\alpha,mem}^{MMSE}[n]\right|^2]$ are non-zero. If we consider the numerator and denominator of β_{MMSE} with steepest-descent ($E[\left|\varepsilon_\alpha^{MMSE}[n]\right|^2] = E[\left|\varepsilon_{\alpha,mem}^{MMSE}[n]\right|^2] = 0$)

$$\beta_{MMSE} = \frac{n_\beta}{d_\beta} = \beta_{opt} \tag{5.40}$$

and β_{MMSE}^{mem} in the same way

$$\beta_{MMSE}^{mem} = \frac{\tilde{n}_\beta}{\tilde{d}_\beta} = \beta_{opt} \tag{5.41}$$

where

$$
\begin{aligned}
n_\beta &= \sqrt{G_e} P_{IMDa}, \quad d_\beta = \sqrt{G_e} L_c P_{IMDa} \\
\tilde{n}_\beta &= \alpha_o (\varepsilon_\alpha^{off})^* P_{\tilde{x}} + \alpha_o (\varepsilon_\alpha^{off})^* T_1 + |\alpha_o|^2 \sqrt{L_c} (T_1^* + T_2) + \sqrt{L_c} T_3 \\
\tilde{d}_\beta &= \sqrt{G_e} \left(\left| \varepsilon_\alpha^{off} \right|^2 P_{\tilde{x}} + 2\,\mathrm{Re}\{\alpha_o \sqrt{L_c} (\varepsilon_\alpha^{off})^* T_1\} + |\alpha_o|^2 L_c T_2 + L_c T_3 \right)
\end{aligned}
\tag{5.42}
$$

Then with *LMS* adaptation, non-zero $E[\left|\varepsilon_\alpha^{MMSE}[n]\right|^2]$ and $E[\left|\varepsilon_{\alpha,mem}^{MMSE}[n]\right|^2]$ reflects to the denominator as

$$
\beta_{MMSE} = \frac{n_\beta}{d_\beta + \sqrt{G_e} P_{\tilde{x}} E[\left|\varepsilon_\alpha^{MMSE}[n]\right|^2]}, \quad \beta_{MMSE}^{mem} = \frac{\tilde{n}_\beta}{\tilde{d}_\beta + \sqrt{G_e} P_{\tilde{x}} E[\left|\varepsilon_{\alpha,mem}^{MMSE}[n]\right|^2]} \tag{5.43}
$$

Therefore if we desire $\beta_{MMSE}^{mem} = \beta_{MMSE}$, then the relation between $E[\left|\varepsilon_\alpha^{MMSE}[n]\right|^2]$ and $E[\left|\varepsilon_{\alpha,mem}^{MMSE}[n]\right|^2]$ should be

$$
\begin{aligned}
E[\left|\varepsilon_{\alpha,mem}^{MMSE}[n]\right|^2] &= \frac{(\tilde{n}_\beta d_\beta - n_\beta \tilde{d}_\beta) + \tilde{n}_\beta \sqrt{G_e} P_{\tilde{x}} E[\left|\varepsilon_\alpha^{MMSE}[n]\right|^2]}{n_\beta \sqrt{G_e} P_{\tilde{x}}} \\
&= \frac{\tilde{n}_\beta}{n_\beta} E[\left|\varepsilon_\alpha^{MMSE}[n]\right|^2]
\end{aligned}
\tag{5.44}
$$

where we have used (5.40) to cancel the term in parenthesis in the first line of (5.44). Expression (5.44) defines the relation between the mean square of the steady-state noise on top of the converged value of α in memoryless and memory cases such that the mean convergence of β is identical in both cases.

An offset can also be defined for the case when there is memory however, we have shown that $\beta_{MMSE}^{mem} = \beta_{MMSE}$ when $E[\left|\varepsilon_{\alpha,mem}^{MMSE}[n]\right|^2] = \frac{\tilde{n}_\beta}{n_\beta} E[\left|\varepsilon_\alpha^{MMSE}[n]\right|^2]$. This implies that $\beta_{off}^{mem} = \beta_{off}$ and (5.28) can also be used to give the mismatch in β where there is memory.

5.4 *LMS* Adaptation of *ECL* Coefficient

We have derived the optimum Wiener coefficient for *ECL* in (5.24) and (5.37) assuming a memoryless amplifier and Wiener-Hammerstein memory respectively. Now we would like to define a practical adaptation mechanism for the *ECL* coefficient β using the *LMS* algorithm, as we previously did for the *SCL* coefficient α.

5.4.1 Basic Recursion

The gradient of the cost function given in (5.17) is first given by

$$\nabla_\beta J[n] = \frac{\partial J_\beta}{\partial \beta} = -2\sqrt{G_e}E[\tilde{z}^*[n]\tilde{e}[n]] \tag{5.45}$$

then a practical instantaneous gradient estimate $\hat{\nabla}_\beta J[n]$ obtained by removing the expectation operator from (5.45). Thus a basic update equation of the form

$$\beta[n+1] = \beta[n] + 2\hat{\mu}_\beta\sqrt{G_e}\tilde{z}[n]\tilde{e}^*[n] \tag{5.46}$$

is obtained which can further be simplified by denoting $\mu_\beta = 2\hat{\mu}_\beta$

$$\beta[n+1] = \beta[n] + \mu_\beta\sqrt{G_e}\tilde{z}[n]\tilde{e}^*[n] \tag{5.47}$$

The adaptation of the *ECL* coefficient defined by the recursion in (5.47) is illustrated in Figure 5-5.

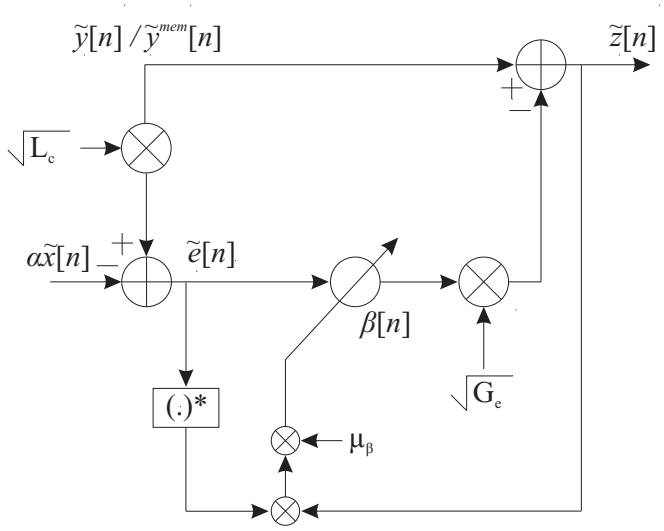


Figure 5-5. LMS adaptation of ECL coefficient.

5.4.2 Convergence Behaviour

It is shown in Appendix B that, for the memoryless case, convergence of β to β_{MMSE} given in (5.24) is achieved in the mean sense if the amplifier input $\tilde{x}$ is white and the mismatch ε_α is rather small. In short this can be stated as $\lim_{n\to\infty} E[\varepsilon_\beta^{MMSE}[n+1]] = 0$ for

$0 < \mu_\beta < 2/G_e P_{\tilde{e}}$ which is also given in (B.5). It is also discussed in Appendix B that under same assumptions with memoryless case, a Wiener-Hammerstein memory that does not cause a significant frequency selectivity to $\tilde{d}$ is also convergent in the mean sense.

The mean square of the mismatch for the memoryless model in steady state is derived in (B.11) as $E[|\varepsilon_\beta^{MMSE}|^2] \approx \mu_\beta |\alpha_o|^2 P_{\tilde{x}}/2$. Furthermore it is discussed that this expression can also be used to calculate the mean square of mismatch for Wiener-Hammerstein memory model as well if the step sizes μ_α and μ_β are small enough. On the other hand, this expression depicts the variance of the noise around the converged point of β, i.e. around β_{MMSE} which is not necessarily β_{opt}. We are interested in the mean square of the mismatch between β and β_{opt} since that is the actual quantity being responsible for the amount of *IMD* suppression as shown in (4.13). We can easily express this term by

$$\begin{aligned} E[|\varepsilon_\beta[n]|^2] &= E[|\beta_{opt} - \beta[n]|^2] = E[|\beta_{opt} - \beta_{MMSE} + \beta_{MMSE} - \beta[n]|^2 \\ &= E[|\beta_{off} - \varepsilon_\beta^{MMSE}|^2] = |\beta_{off}|^2 - 2\,\mathrm{Re}\{\beta_{off}^* E[\varepsilon_\beta^{MMSE}]\} + E[|\varepsilon_\beta^{MMSE}|^2] \\ &= |\beta_{off}|^2 + E[|\varepsilon_\beta^{MMSE}|^2] \end{aligned} \quad (5.48)$$

where β_{off} and $E[|\varepsilon_\beta^{MMSE}|^2]$ are given in (5.28) and (B.11) respectively. Note that while simplifying from second line to third, we have used the fact that $E[\varepsilon_\beta^{MMSE}] = 0$ as pointed out in (B.5). In addition since we have similar values for β_{off} and $E[|\varepsilon_\beta^{MMSE}|^2]$ in the case where there is memory, the mean square of the mismatch and therefore the *IMD* suppression is similar. While deriving $E[|\varepsilon_\beta^{MMSE}|^2]$ given in (B.11), it is already assumed that α has a small steady state noise, thus β_{off} is rather small and we can approximate (5.48) by

$$E[|\varepsilon_\beta[n]|^2] \approx E[|\varepsilon_\beta^{MMSE}|^2] \quad (5.49)$$

5.5 *IMD* Suppression Analysis

The amount of suppression of *IMD* terms for fixed β is given in (4.13) . It is also discussed in the same chapter that the amount of suppression for the case when there is memory is exactly same if the definition of *IMD* is modified accordingly, i.e. the distortion terms presented in (4.22). We can state an equivalent formulation for *IMD* suppression with β being a random quantity

$$IMD_{\sup} = 10 * \log(1 / E[|\varepsilon_\beta|^2]) \quad (5.50)$$

where $E[|\varepsilon_\beta|^2]$ is given in (5.49).

The *IMD* suppression depends only on the mean-square mismatch in β which depends on only μ_α, μ_β, α_o and $P_{\tilde{x}}$ considering (5.49) and (B.11). Assuming that $P_{\tilde{x}}$ is a known quantity, and α_o can be approximately calculated by the observations obtained from input and output of the amplifier, then the choice of μ_α and μ_β tell the achievable amount of *IMD* suppression. Here we are assuming that the choice of μ_α and μ_β are free, however it is small to satisfy the assumptions made during the analysis.

Chapter 6

Simulation Examples

6.1 Convergence of *SCL* and *ECL* Coefficients

The simulations in this part present the convergence behavior of *SCL* and *ECL* coefficients under LMS adaptation. An *OFDM* signal with $N = 1024$ subcarriers having spacing of 15 kHz and 16-QAM subcarrier-modulation is used as an input to amplifier. The amplifier model for the memoryless case is an *SSPA* with $\kappa = \sqrt{1000}$ (30 dB power gain), $A_o = 10$ and $p = 1$. The input signal is scaled to 1 dB compression point and *IBO* is set to 0 dB. The value of attenuator loss and error amplifier gain is set to $L_c = -30$ and $G_e = 30$ dB respectively. The simulations are averaged over 150 independent realizations.

6.1.1 Mean Convergence of SCL Coefficient

In addition to the common parameters stated above, the number of active sub-carriers is set to $N_a = 256$. For the simulations with memory, two different Wiener and Hammerstein filter combinations are used before and after *SSPA*. The coefficients of Wiener filter are fixed to $b_{w1} = [1.000\ 0.005]$, and the coefficients of Hammerstein filter are varied between two different simulations as $b_{h1} = [1.000 - 0.03]$ and $b_{h2} = [1.000 - 0.03]$ respectively. The step size $\mu_\alpha = 0.05$ is used for adaptation (both for amplifier with and without memory).

The optimum value for the complex constant is calculated according to $\alpha_o = (\underline{\tilde{x}}^H \underline{\tilde{y}})/(\underline{\tilde{x}}^H \underline{\tilde{x}})$ where $\underline{\tilde{x}}$ and $\underline{\tilde{y}}$ are the data vectors showing the input and output of the memoryless amplifier. The vectors are chosen large enough to approximate the expectation in (3.6) ($N_s = 120 * 1024$ corresponding to $L = 120$ *OFDM* symbols).

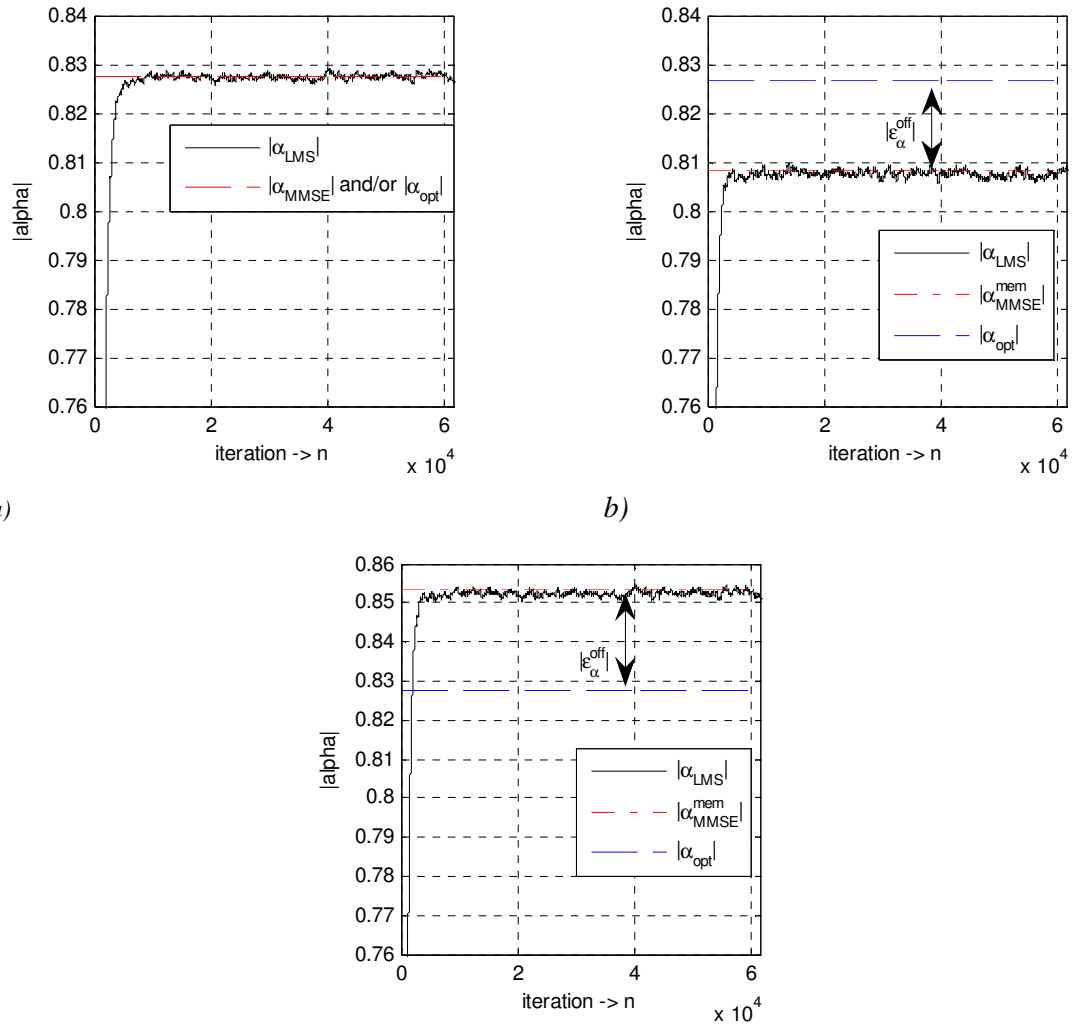


Figure 6-1. The mean convergence of absolute value of α *under LMS adaptation with step size* $\mu_\alpha = 0.05$ *and when* 120 *OFDM symbols (*$N = 1024$, $N_a = 256$ *and* $M = 16$*) are used as an input to a amplifier a) memoryless (SSPA with* $\kappa = \sqrt{1000}$, $A_o = 10$ $p = 1$*) b) Wiener-Hammerstein memory with* b_{w1} - b_{h1} *c) Wiener-Hammerstein memory with* b_{w1} - b_{h2}.

SSPA which is the common nonlinearity block for both amplifier with and without memory yields a real α_o, nevertheless the magnitudes of the related quantities are plotted in Figure 6-1 a), b) and c).

The *SCL* coefficient α is shown to be converging to optimum Wiener coefficient α_{MMSE} (which is also equal to optimum linearizer coefficient α_{opt}) in the mean sense by

Figure 6-1 a). This was the expected scenario when a memoryless amplifier model and a white input signal $\tilde{x}$ are used. Although N_a is smaller than N and thus yielding a non-white signal $\tilde{x}$, the step size $\mu_\alpha = 0.05$ is small enough to maintain convergence in the mean sense.

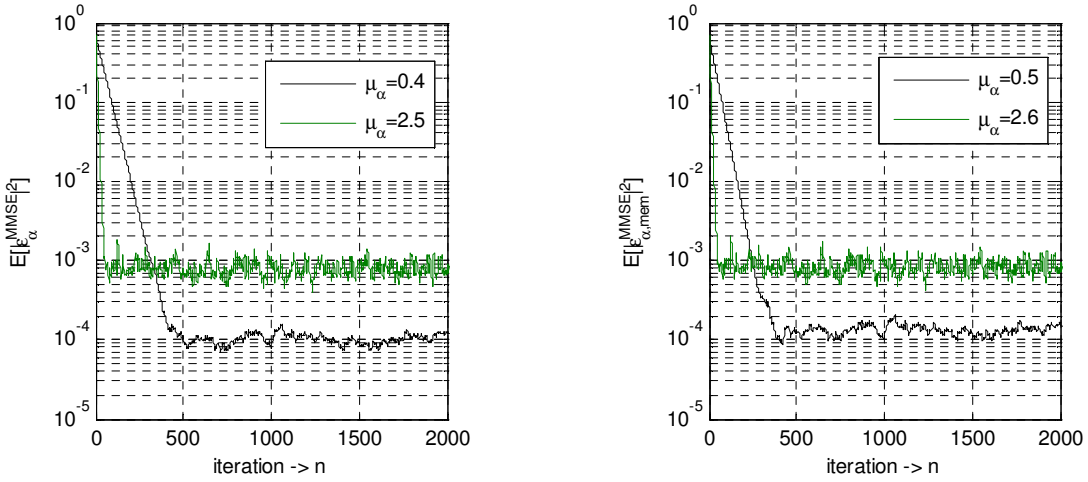
The effect of memory on the convergence of α is depicted in Figure 6-1 b) and c) where the expected ε_α^{off} given in (5.10) is clearly observable. Actually N_a is deliberately chosen smaller than N to introduce correlation within the samples of $\tilde{x}$, so that we obtain a non-zero value for ε_α^{off} which is described in (5.10). It is true that ε_α^{off} is non-zero even if $\tilde{x}$ is white, as long as filter denoted by b_w introduces some correlation. However; as discussed in Section 4.4, in order to have similar value for α_o with the memoryless case, this filter should be chosen with small memory taps which in turn introduces only small amount of correlation.

6.1.2 Mean Convergence of *ECL* Coefficient

The first aim of these simulations is to illustrate the impact of mean square of mismatch in α on β_{MMSE} and mean convergence of β for amplifier models with and without memory. In other words the derivations (5.25) and (5.37) are going to be justified. For this purpose, β_{MMSE} and β_{LMS} is studied by simulations for two different values of μ_α (also slightly different for amplifier with and without memory, see Figure 6-2 a) and b)). In addition, the number of active subcarriers is chosen to be $N_a = 1024$. By that the spectrum of amplifier input is guaranteed to be white (for the memoryless case) which enables *IMD* and thus signal before β to be white (with assumption of small ε_α). As discussed in Appendix A, this is the necessary condition for the mean convergence of β to β_{MMSE}. Also in these simulations it is shown that $\beta_{MMSE} = \beta_{MMSE}^{mem}$ for similar values of $E[\left|\varepsilon_\alpha^{MMSE}\right|^2]$ and $E[\left|\varepsilon_{\alpha,mem}^{MMSE}\right|^2]$. Remember that the exact relation to be satisfied is given in (5.44). For the simulations where amplifier is modeled with Wiener and Hammerstein memory, the chosen coefficients for filter are $b_{w2} = [1.00 - 0.05\ 0.01]$ and $b_{h3} = [1.00\ 0.03]$ which gives more frequency selective response compared to filters used in the previous part. One final comment related to the simulations is that, the adaptation of β is started after the convergence of α is obtained. Regarding all the other simulations in this chapter, β is started to be adapted after 10000 iterations of α.

The expected value of mean-square of mismatch between α and its *MMSE* solution in the memoryless case is shown in Figure 6-2 a) for two different values of step size. As expected small step size provides a smaller steady state level with the cost of longer convergence time (which still requires only half of one *OFDM* symbol). Then the values of β_{MMSE}

corresponding to these steady state levels of $E[|\varepsilon_\alpha^{MMSE}|^2]$ are shown in Figure 6-3 b) to which the β under LMS adaptation converges in the mean sense as expected.



a) *b)*

Figure 6-2. The expected value of mean-squared mismatch between the MMSE solution of α and value obtained from LMS adaptation for a) memoryless amplifier (SSPA with $\kappa = \sqrt{1000}$, $A_o = 10$, $p = 1$) when $\mu_\alpha = 0.4$ and μ_α =2.5) b) Wiener-Hammerstein memory (b_{w2} - b_{h3}) when $\mu_\alpha = 0.5$ and $\mu_\alpha = 2.6$.

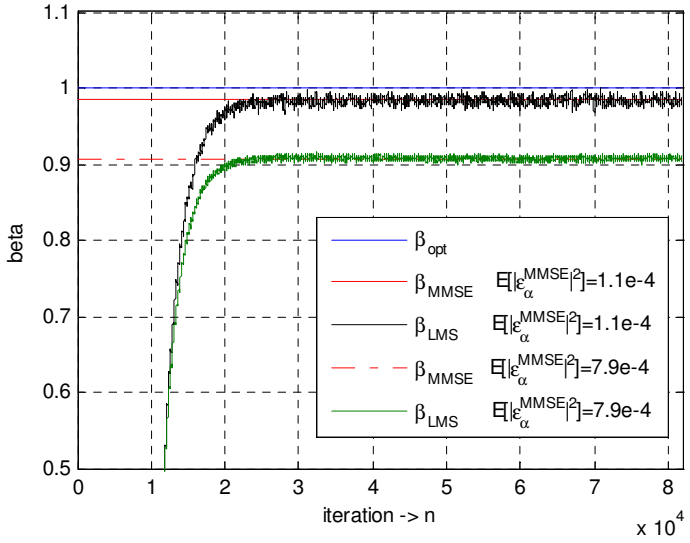


Figure 6-3. Optimum linearizer coefficient β_{opt}, mmse solution β_{MMSE} and the mean convergence of β under LMS adaptation corresponding to two different values of $E[|\varepsilon_\alpha^{MMSE}|^2]$ that are shown in Figure 6-2 a) for memoryless amplifier.

Similar results are also obtained for the amplifier with memory and illustrated in Figure 6-2 b) and Figure 6-3 b). Note that slightly different step sizes are used for the simulations related to memory so that the obtained $E[|\varepsilon_{\alpha,mem}^{MMSE}|^2]$ approximately satisfies (5.44) and similar *MMSE* solution and mean convergence are achieved.

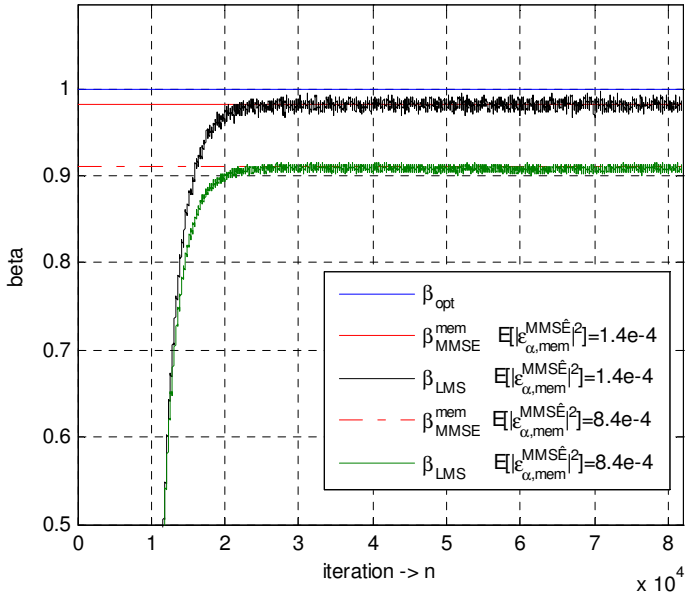


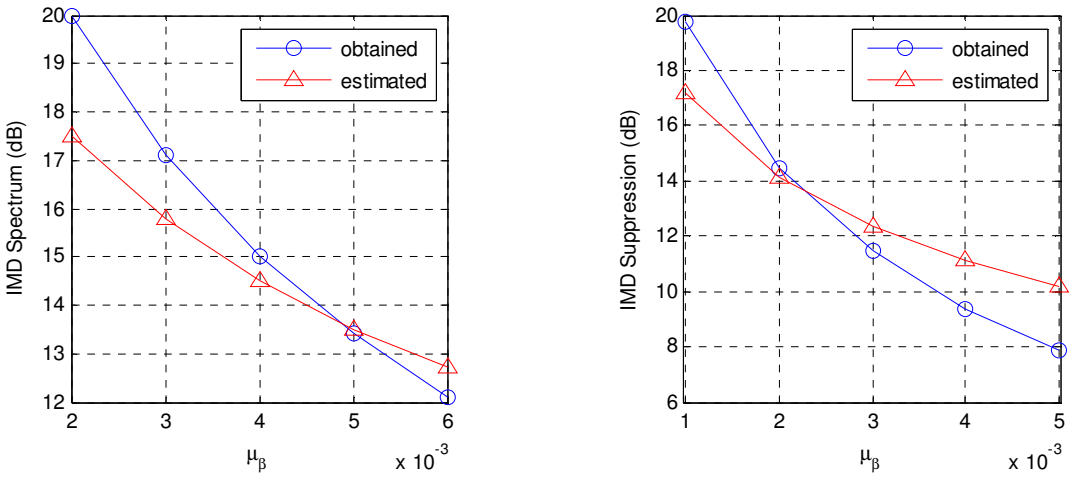
Figure 6-4. Optimum linearizer coefficient β_{opt}, mmse solution β_{MMSE} and the mean convergence of β under LMS adaptation corresponding to two different values of $E[|\varepsilon_{\alpha,mem}^{MMSE}|^2]$ that are shown in Figure 6-2 b) for amplifier with Wiener-Hammerstein memory .

6.2 Estimated and Achieved *IMD* Suppressions

The amount of *IMD* suppression from amplifier output to linearizer output is given by (5.50). We have an estimate for $E[|\varepsilon_\beta|^2]$ via (5.48) under the assumption of amplifier input $\tilde{x}$ being white and the mismatch in α being small enough as stated in Appendix B. Due to these we again will choose $N_a = N = 1024$ and $\mu_\alpha = 0.05$. Other parameters related to feedforward circuit and amplifier are same with the previous part except the operation point is now chosen to be 3 dB compression point. The observed *PAPR* of the amplifier input signal is about 10.5 dB.

The estimated and obtained values for *IMD* suppression for memoryless case are observable in Figure 6-5 a). The estimated values have still some deviation from the obtained values even though the simulation parameters are chosen in order to satisfy the necessary assumption for estimation. The reason is due to the rough approximation of (B.9), however

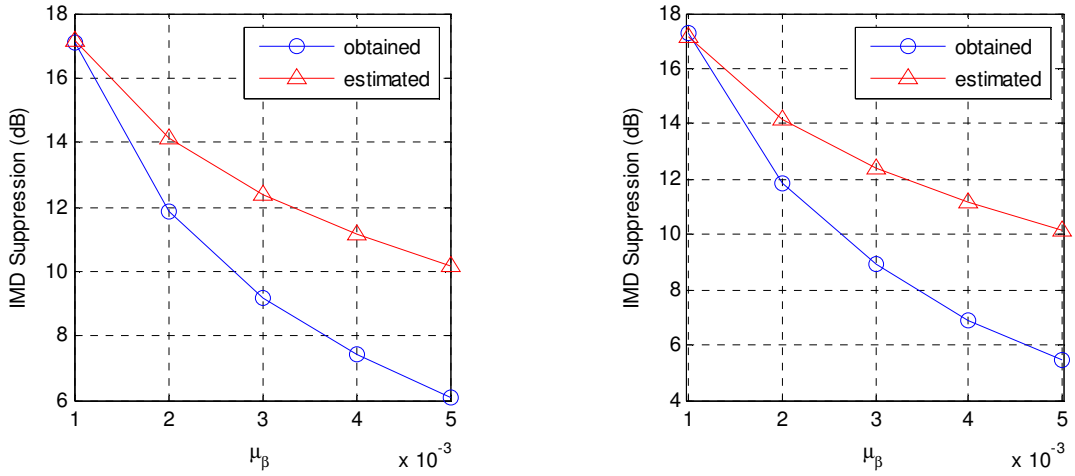
the estimated values stay within 1.5 dB range from actual values in the worst case. Same estimated values for *IMD* suppression are used for the case where there is memory and presented with the obtained values in Figure 6-5 b). It can be seen that performance in terms of *IMD* suppression is very close to the memoryless case as expected.



a) *b)*

Figure 6-5. The estimated and obtained IMD suppression when OFDM input signal with $N_a = 1024$ *is used as an input to a a) memoryless amplifier (SSPA with* $\kappa = \sqrt{1000}$, $A_o = 10$, $p = 1$*) b) amplifier with Wiener-Hammerstein memory (* b_{w2} - b_{h3} *) where in both cases the operation point is 3 dB compression point.*

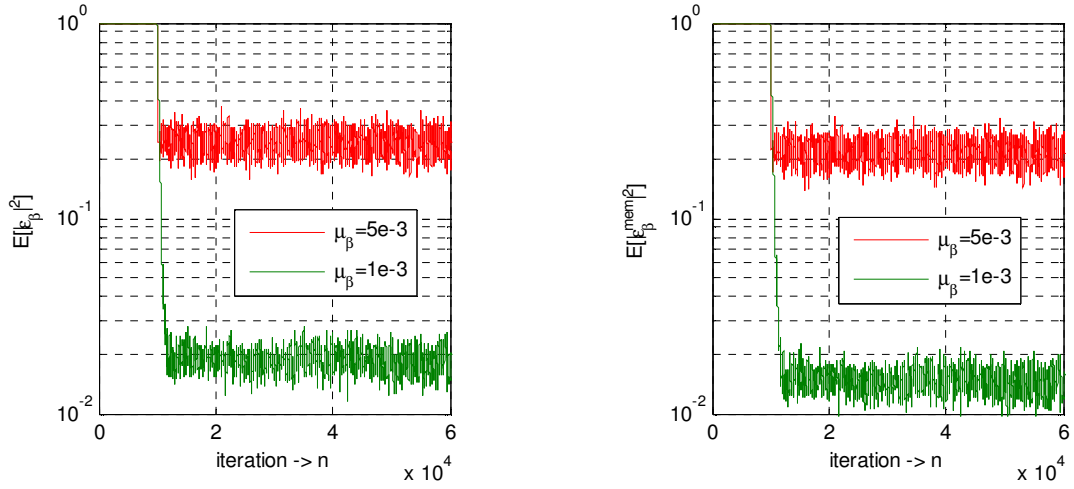
Another set of simulations are carried out by relaxing the condition on $\tilde{x}$ for being white by setting $N_a = 512$. We can see by Figure 6-6 a) and b) that the accuracy for the estimation of *IMD* suppression is degraded slightly. However, when μ_β is small then good accuracy is still achieved which is also expected. Mainly ε_β^{MMSE} has smaller dependency to $\tilde{d}$ at any time instant therefore it is still not a poor approximation when passing from (B.6) to (B.7).



a) *b)*

Figure 6-6. The estimated and obtained IMD suppression when OFDM input signal with $N_a = 512$ is used as an input to a) a memoryless amplifier (SSPA with $\kappa = \sqrt{1000}$, $A_o = 10$, $p = 1$) b) amplifier with Wiener-Hammerstein memory (b_{w2} - b_{h3}) where in both cases the operation point is 3 dB compression point..

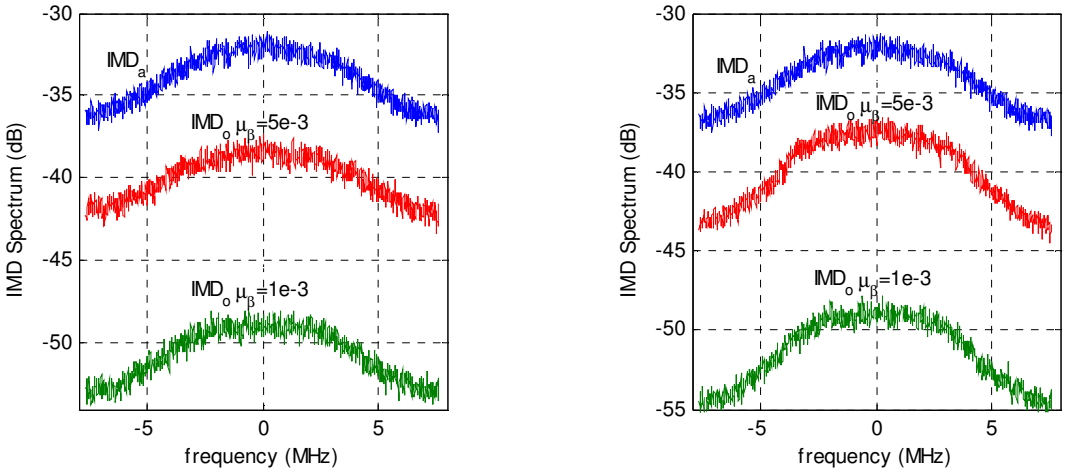
The suppression of *IMD* depends on the value of mean-square of the mismatch in β that can be controlled by μ_β. Different steady state levels of this mean-square of mismatch are shown in Figure 6-7 a) and b) corresponding to different step sizes.



a) *b)*

Figure 6-7. The expected value of mean-squared mismatch between the optimum value of β and the value obtained from LMS adaptation for two different step sizes ($\mu_\beta = 0.005$ and $\mu_\beta = 0.001$) when OFDM signal with $N_a = 512$ is used as an input to a a) memoryless amplifier (SSPA with $\kappa = \sqrt{1000}$, $A_o = 10$, $p = 1$) and b) amplifier with Wiener-Hammerstein memory (b_{w2} - b_{h3}).

The *IMD* spectra for amplifier with and without memory corresponding to those mismatch values are plotted in Figure 6-8 a) and b) which are also showing correspondence to the levels indicated by the blue curve in Figure 6-6 a) and b). For instance when $\mu_\beta = 5e-3$, an *IMD* suppression of 6 dB is obtained whereas 17 dB of attenuation is achieved for $\mu_\beta = 1e-3$ for both memoryless and memory cases.



a) *b)*

Figure 6-8. IMD spectrum at the output of amplifier and linearizer that corresponds to $\mu_\beta = 0.005$ *and* $\mu_\beta = 0.001$ *when OFDM signal with* $N_a = 512$ *is used as an input to a a) memoryless amplifier (SSPA with* $\kappa = \sqrt{1000}$, $A_o = 10$, $p = 1$*) and b) amplifier with Wiener-Hammerstein memory (* b_{w2} - b_{h3} *).*

Now we are also able to observe the spectral re-growth at the output of power amplifier and suppression of *IMD* corresponding to different choice of μ_β from the spectrum plots in Figure 6-9 and Figure 6-10. We see that when there is memory, the performance remains almost same with the memoryless case in terms of *IMD* suppression as expected.

Finally we repeat the simulations with $N_a = 256$. The plots showing the estimated and obtained *IMD* suppression in Figure 6-11 points out that the accuracy of estimation degraded compared to $N_a = 512, 1024$ since $\tilde{x}$ further deviates from being white.

Similar to the simulations with $N_a = 512$, different steady state values for the mean square mismatch in β are obtained depending on the step size as shown in Figure 6-12 a) and b). These values of mean square mismatch then determine the amount of *IMD* suppression as shown in Figure 6-13 a) and b), Figure 6-14 and Figure 6-15.

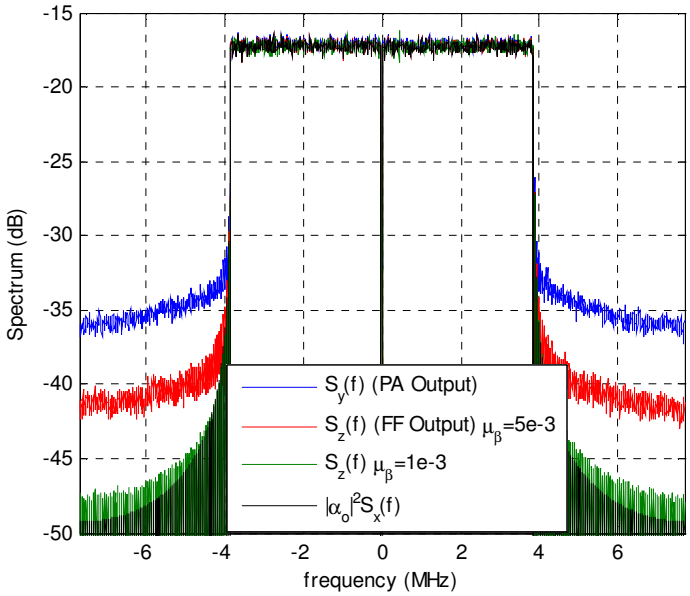


Figure 6-9. The spectra of linearly amplified input, power amplifier output, and linearizer output (for $\mu_\beta = 0.005$ *and* $\mu_\beta = 0.001$*) when memoryless amplifier model (SSPA with* $\kappa = \sqrt{1000}$, $A_o = 10$, $p = 1$*) and OFDM input signal with* $N_a = 512$ *is used.*

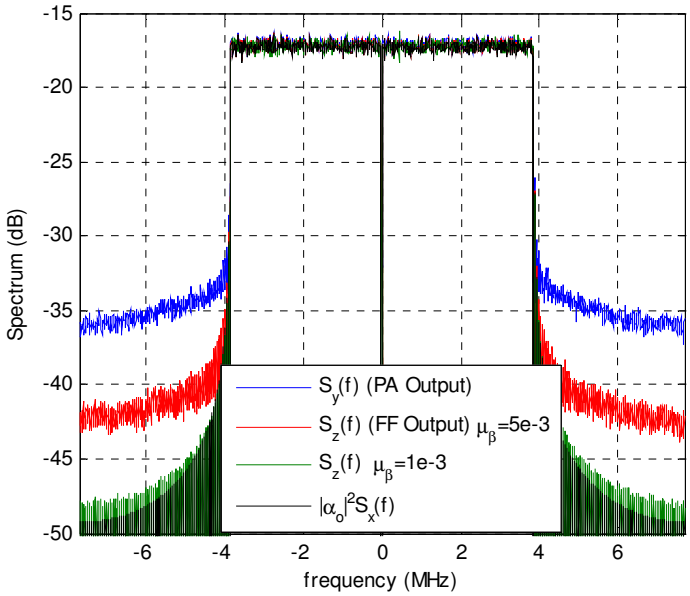
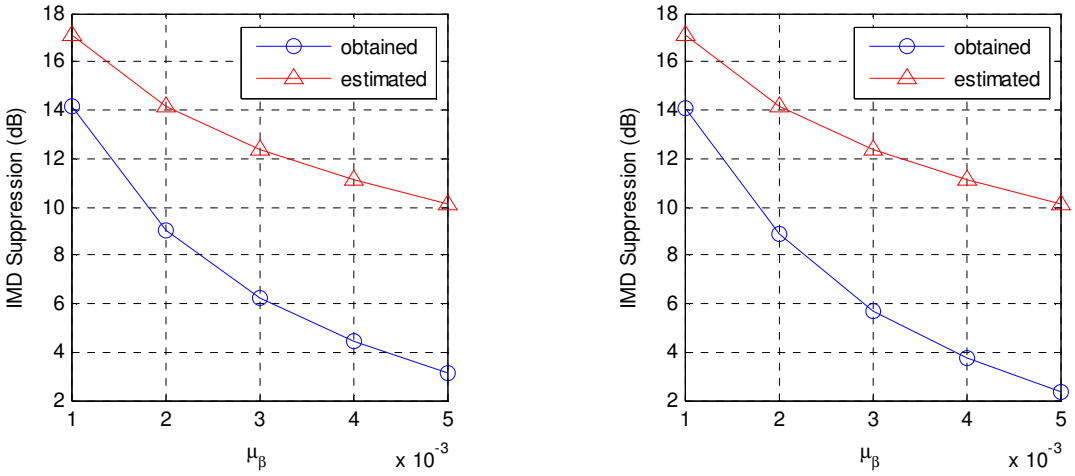
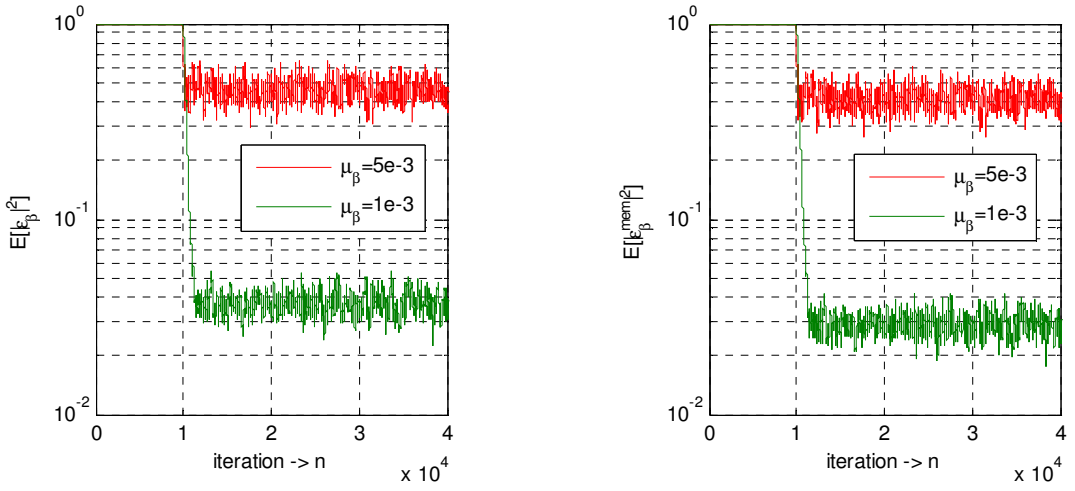


Figure 6-10. The spectra of linearly amplified input, power amplifier output, and linearizer output (for $\mu_\beta = 0.005$ *and* $\mu_\beta = 0.001$*) when amplifier with Wiener-Hammerstein memory (* b_{w2} - b_{h3} *) and OFDM input signal with* $N_a = 512$ *is used.*



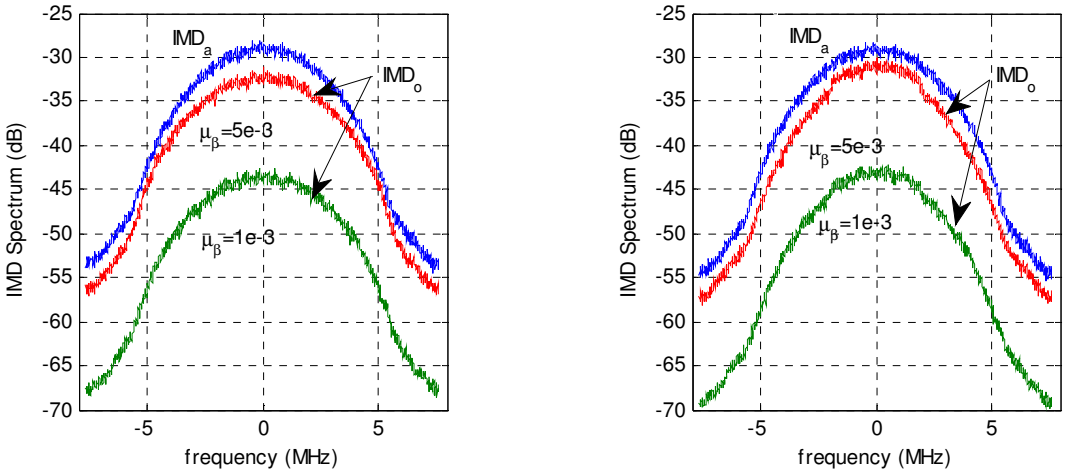
a) *b)*

Figure 6-11. The estimated and obtained IMD suppression when OFDM input signal with $N_a = 256$ is used as an input to a) a memoryless amplifier (SSPA with $\kappa = \sqrt{1000}$, $A_o = 10$, $p = 1$) b) amplifier with Wiener-Hammerstein memory (b_{w2} - b_{h3}) where in both cases the operation point is 3 dB compression point..



a) *b)*

Figure 6-12. The expected value of mean-squared mismatch between the optimum value of β and the value obtained from LMS adaptation for two different step sizes ($\mu_\beta = 0.005$ and $\mu_\beta = 0.001$) when OFDM signal with $N_a = 256$ is used as an input to a a) memoryless amplifier (SSPA with $\kappa = \sqrt{1000}$, $A_o = 10$, $p = 1$) and b) amplifier with Wiener-Hammerstein memory (b_{w2} - b_{h3}).



a) *b)*

Figure 6-13. IMD spectrum at the output of amplifier and linearizer that corresponds to $\mu_\beta = 0.005$ *and* $\mu_\beta = 0.001$ *when OFDM signal with* $N_a = 256$ *is used as an input to a a) memoryless amplifier (SSPA with* $\kappa = \sqrt{1000}$, $A_o = 10$, $p = 1$*) and b) amplifier with Wiener-Hammerstein memory (* b_{w2} - b_{h3} *).*

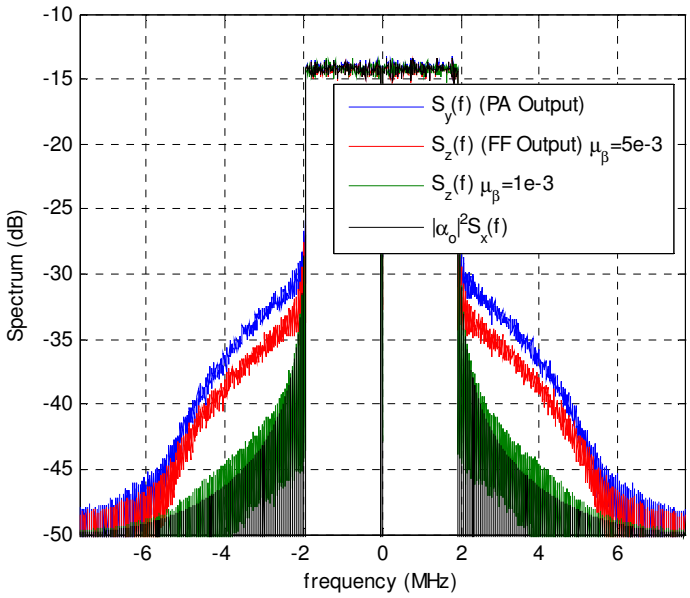


Figure 6-14. The spectra of linearly amplified input, power amplifier output, and linearizer output (for $\mu_\beta = 0.005$ *and* $\mu_\beta = 0.001$*) when memoryless amplifier model (SSPA with* $\kappa = \sqrt{1000}$, $A_o = 10$, $p = 1$*) and OFDM input signal with* $N_a = 256$ *is used.*

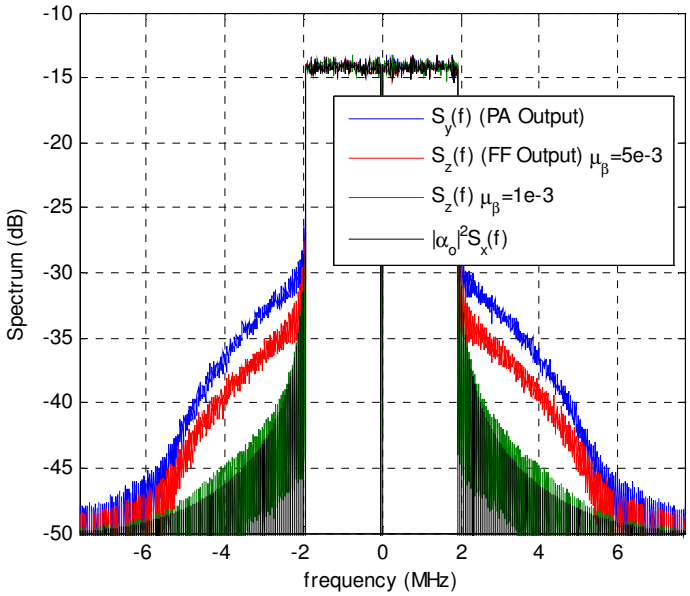


Figure 6-15. The spectra of linearly amplified input, power amplifier output, and linearizer output (for $\mu_\beta = 0.005$ *and* $\mu_\beta = 0.001$*) when amplifier with Wiener-Hammerstein memory (* b_{w2} - b_{h3} *) and OFDM input signal with* $N_a = 256$ *is used.*

Chapter 7

Summary and Conclusion

This thesis studied the performance of *LMS* adaptation of feedforward coefficients in terms of convergence and achievable *IMD* suppression. To the author's knowledge such analysis has not presented in literature so far. The extension of the analysis for the case where the main amplifier is modeled with Wiener-Hammerstein memory is another novel aspect.

The analysis in Sections 5.3 and 5.4 showed that the choice of step size to adapt *SCL* and *ECL* has two important impacts concerning the amount of *IMD* suppression. The former determines how close *ECL* coefficient converges to the optimum linearizer coefficient in the mean sense while the latter determines the variance of the fluctuation after convergence. An estimate for the amount of *IMD* suppression is then derived as the combined effect of these two phenomena under the assumption of white amplifier input and small step size for *SCL* coefficient. The average power of amplifier input and the linear gain in *Bussgang* model is also assumed to be known when deriving this estimate.

The estimate is shown to be matching good with the *IMD* suppression values obtained from simulations as presented in Section 6.2. Furthermore when the condition on white amplifier input is relaxed, the accuracy of estimation is still reasonable according to those simulations where small step size is used to adapt *ECL* coefficient.

The independency of feedforward linearization from memory effects has been discussed in literature for fixed coefficients. However, the analysis of possible impacts of memory on the adaptive system has been missing. This analysis is carried out in this thesis for Wiener-Hammerstein modeled memory effects. We have shown that memory effects introduces an offset to the mean convergence of *SCL* coefficient. This offset is shown in Section 5.1.2 to be a function of auto-correlation of the amplifier input and the strength of memory. Despite this offset in the convergence of *SCL* coefficient, the convergence of *ECL* coefficient in the presence of Wiener-Hammerstein memory is shown in Section 5.3 to be same/similar to memoryless case. Moreover when same/similar step sizes as in memoryless case are used for

the adaptation of both coefficients, the performance in terms of *IMD* reduction is shown to be same/similar in memoryless case. These foundations are also supported by the simulations in Section 6.2.

Adaptive feedforward linearization is a powerful alternative among the power amplifier linearization methods. In the ideal implementation scenario, it is independent of input signal characteristics such as bandwidth and *PAPR* as well as the nonlinearity experienced by this signal. However, there are practical considerations that bring limitations to those parameters and the overall linearizer performance. The demands on linearly operating error amplifier and synchronization of upper and lower branches of *SCL* and *ECL* are important limiting factors of basic feedforward operation. In the adaptive case, additional circuitry is needed for down-conversion and sampling of the necessary signals. Those will bring further restrictions on the bandwidth of the amplifier input. Besides, the impairments introduced by the additional circuitry to up and down-convert and sample the signals in the feedforward structure is another performance limiting factor which is not well studied in literature and is left as a future work.

Bibliography

[1] 3GPP Technical Specification Group Radio Access Network, "Requirements for Evolved UTRA and UTRAN",TR 25913, Jan. 2008

[2] P. Aaen, J. Pla, J. Wood, *Modeling and Characterization of RF and Microwave Power FET's*, Cambridge University Press, 2007.

[3] E. Aschbacher, *Digital Pre-distortion of Microwave Power Amplifiers*, Vienna University of Technology, PhD. Thesis, September 2005.

[4] P. Banelli, "Theoretical analysis and performance of *OFDM* signals in nonlinear *AWGN* channels", *IEEE Trans. on Communications,* vol. 48, pp. 430-441, March 2000.

[5] S. Benedetto, E. Biglieri, and R. Dafarra, "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.

[6] N. M. Blachman, "The output signals and noise from a nonlinearity with amplitude-dependent phase shift", *IEEE Trans. on Information Theory*, vol. IT-25, pp. 77-79, Jan. 1979.

[7] H. S. Black., *Translating System,* U.S. patent 1,686,792, Oct. 1928.

[8] W. Bosch and G. Gatti, "Measurement and simulation of memory effects in predistortion linearizers," *IEEE Trans. Microwave Theory Tech,* vol. 37, pp. 1885-1890, Dec. 1989.

[9] S. Boumaiza, F. M. Ghannouchi, "Thermal memory effects modeling and compensation in RF Power Amplifiers and Predistortion Linearizers", *IEEE Trans on Microwave Theory and Tech.,* vol. 51, pp. 2427-2433, Dec. 2003.

[10] J. J. Bussgang, "Crosscorrelation functions of amplitude-distorted Gaussian signals," Research Lab. Electron, M.I.T., Cambridge, MA, Tech. Rep. 216, Mar. 1952.

[11] J. K. Cavers, "Adaptation behavior of a feedforward amplifier linearizer", *IEEE Trans. on Vehicular Technol.,* vol. 44, pp. 31-39*,* Feb. 1995.

[12] J. K. Cavers, "Convergence behavior of an adaptive feedforward linearizer", in *Proc. Vehicular Technology Conference,* June 1994.

[13] J. Chen, H. Tsai, Y. Chen, "The optimal RLS parameter tracking algorithm for a power amplifier feedforward linearizer", *IEEE Trans. on Circuits and Systems-II: Analog and Digital Signal Signal Processing*, vol. 46, pp. 464-68, Apr. 1999

[14] D. C. Cox, "Linear amplification with nonlinear components", *IEEE Trans on Communications*, vol. COM-22, pp. 1942-1945, Dec. 1994.

[15] D. Dardari, V. Tralli, A. Vaccari, "A theoretical characterization of nonlinear distortion effects in *OFDM* systems", *IEEE Trans. on Communications*, vol. 48, pp. 1755-1764, Oct. 2000.

[16] J. L. Dawson, *Feedback Linearization of RF Amplifiers*, Stanford University, PhD. Thesis, Aug. 2003.

[17] L. Dei, *Digital RF Power Amplifier Linearizers*, Georgia Institute of Technology, PhD. thesis, March 2004.

[18] N. Y. Ermolova, Sven-Gustav Haggman, "An extension of Bussang's theory to complex- valued signals", in *Proc.* of the 6th Nordic Signal Processing Symposium-NORSIG 2004, June 9-11, 2004.

[19] S. J. Grant, J. K. Cavers, "A DSP controlled adaptive feedforward amplifier linearizer," in *Proc.* 5th International Conference on Universal Personal Communications*,* vol. 2, 1996, pp. 788-92.

[20] S. Haykin, *Adaptive Filter Theory*, 3rd ed. Prentice Hall, 1996.

[21] IEEE Computer Society and IEEE Microwave Theory and Techniques Society, "IEEE standard for local and metropolitan area networks, part 16: air interface for fixed broadband wireless access systems", June 2004

[22] P. E. Kenington, *High-linearity RF Amplifier Design*, Artech House, 2000.

[23] E. A. Lee, D. G. Messershmitt, *Digital Communication*, 2nd ed. Kluwer Academic Publishers, 2000.

[24] J. Minkoff, "The role of AM-to-PM conversion in memoryless nonlinear systems", *IEEE Transactions on Communications,* vol. com-33, no. 2, Feb. 1985

[25] Mobile Radio Communications, John Wiley and Sons, 2nd edition, 1999.

[26] A. Papoulis, Probability, *Random Variables, and Stochastic Processes*, 4th ed. McGraw-Hill, 2002.

[27] 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, Oct. 2002, pp. 7-10

[28] R. G. Randall, *A Broadband DSP Based Feedforward Amplifier Linearizer*, University of Calgary, MSc. Thesis, 2001.

[29] C.Rapp, "Effects of the HPA-nonlinearity on a 4-DPSK/*OFDM* signal for a digital sound broadcasting system" in *Proc.* Conf. Rec. ECSC'91, Luettich, Oct. 1991.

[30] A. A. M. *Saleh*, "Frequency-independent and frequency-dependent nonlinear models of TWT amplifiers," *IEEE Transactions on Communications*, vol. COM-29, no. 11, pp. 1715-1720, Nov. 1981.

[31] T. Schenk, RF Impairments in Multiple Antenna *OFDM*, Eindhoven University of Technology, PhD thesis, 2006.

[32] H. Seidel, H.R. Beurrier, and A. N. Friedman, "Error controlled high power linear amplifiers at VHF," *Bell Systems Technical Journal,* May/June 1968.

[33] H. Seidel, "A microwave feedforward experiment," *Bell Systems Technical Journal,* Nov. 1971

[34] A. Shahed, A. Hasim Gokceoglu, M. Valkama, "Coefficient sensitivity analysis for feedforward amplifier linearizer with memory", in *Proc.* The 11th International Symposium on Wireless Personal Multimedia Communications*,* 2008.

[35] I. Teikari, *Digital Predistortion Linearization Methods for RF Power Amplifiers*, Helsinki University of Technology, PhD. Thesis, Sep. 2008

[36] J. H. K. Vuolevi, T. Rahkonen, and J. P. A. Manninen, "Measurement and simulation of memory effects in predistortion linearizers," *IEEE Trans. Microwave Theory Tech,* vol. 49, pp. 1383-1389, Aug. 2001

[37] S. B. Weinstein and P. M. Ebert, "Data transmissiom in frequency division multiplexing using the discrete Fourier transform," *IEEE Trans. Commun.*, vol. COM-19, Oct. 1971.

[38] G. T. Zhou, H. Qian, L. Ding, R. Raich, "On the baseband representation of a bandpass nonlinearity," *IEEE Trans. Signal Process*., vol. 53, no. 8, pp. 2953-2957, Aug. 2005.

Appendix A

Mean and Mean-Squared Convergence of *SCL* Coefficient

We will follow similar steps and assumptions as in [20] to carry out the analysis for mean and mean squared convergence of *SCL* coefficient (1 tap filter) under *LMS* adaptation. We will start with the model where main amplifier is a memoryless nonlinearity and then extend the analysis for the case where Wiener-Hammerstein model is used for long term memory effects.

A.1 Memoryless Model

Starting from (5.13) we first subtract α_{MMSE} from both sides to yield

$$\alpha[n+1] - \alpha_{MMSE} = \alpha[n] - \alpha_{MMSE} + \mu_\alpha \tilde{x}^*[n]\tilde{e}[n] \tag{A.1}$$

Now by defining $\varepsilon_\alpha^{MMSE}[n] = \alpha[n] - \alpha_{MMSE}$ ($\varepsilon_\alpha^{MMSE}[n] = -\varepsilon_\alpha$ for memoryless case), we get

$$\begin{aligned} \varepsilon_\alpha^{MMSE}[n+1] &= \varepsilon_\alpha^{MMSE}[n] + \mu_\alpha \tilde{x}^*[n]e[n] \\ &= \varepsilon_\alpha^{MMSE}[n] + \mu_\alpha \tilde{x}^*[n](\sqrt{L_c}\tilde{y}[n] - \alpha[n]\tilde{x}[n]) \end{aligned} \tag{A.2}$$

and by modifying the second term of (A.2) in parenthesis as $\sqrt{L_c}\tilde{y}[n] - \alpha_{MMSE}\tilde{x}[n] + \alpha_{MMSE}\tilde{x}[n] - \alpha[n]\tilde{x}[n]$ we obtain

$$\varepsilon_\alpha^{MMSE}[n+1] = \varepsilon_\alpha^{MMSE}[n] - \mu_\alpha \varepsilon_\alpha^{MMSE}[n]\tilde{x}[n]\tilde{x}^*[n] + \mu_\alpha \tilde{x}^*[n]e_o[n] \tag{A.3}$$

where $e_o[n] = \sqrt{L_c}\tilde{y}[n] - \alpha_{MMSE}\tilde{x}[n]$ denotes the error signal when the filter is computing its output with optimum Wiener coefficients. Note that $e_o[n]$ is actually $\tilde{d}[n]$ since it is shown that the optimum Wiener coefficient is identical to optimum linearizer coefficient for *SCL* given in (4.3) . Now applying the expectation operation to (A.3) we obtain

$$E[\varepsilon_{\alpha}^{MMSE}[n+1]] = E[\varepsilon_{\alpha}^{MMSE}[n]] - \mu_{\alpha}E[\varepsilon_{\alpha}^{MMSE}[n]\tilde{x}[n]\tilde{x}^{*}[n]] + \mu_{\alpha}\underbrace{E[\tilde{x}^{*}[n]e_{o}[n]]}_{0} \quad \text{(A.4)}$$

The last term vanishes since by definition the error is orthogonal to input for the optimum Wiener coefficients. The second term can be separated as $E[\varepsilon_{\alpha}^{MMSE}[n]]E[\tilde{x}[n]\tilde{x}^{*}[n]]$ under *independence assumption* ([20], pp. 392) which implies that the input sequence $\tilde{x}$ to be white. In the case where second term is separable, we can simplify (A.4) as

$$\begin{aligned} E[\varepsilon_{\alpha}^{MMSE}[n+1]] &= E[\varepsilon_{\alpha}^{MMSE}[n]](1-\mu_{\alpha}E[\tilde{x}[n]\tilde{x}^{*}[n]]) \\ &= E[\varepsilon_{\alpha}^{MMSE}[n]](1-\mu_{\alpha}P_{\tilde{x}}) \\ &= E[\varepsilon_{\alpha}^{MMSE}[0]](1-\mu_{\alpha}P_{\tilde{x}})^{n+1} \end{aligned} \quad \text{(A.5)}$$

where we have assumed that the power of the input can be given $P_{\tilde{x}} = E[\tilde{x}[n]\tilde{x}^{*}[n]]$ under the stationary input assumption. The last line of (A.5) tells that for a proper choice of the step-size, the mismatch between the adapted and optimum Wiener coefficient vanishes asymptotically

$$\lim_{n\to\infty} E[\varepsilon_{\alpha}^{MMSE}[n]] = 0 \quad 0 < \mu_{\alpha} < 2/P_{\tilde{x}} \quad \text{(A.6)}$$

The mean squared error of mismatch can also be obtained under the previous white input assumption. The noisy gradient is modeled as a summation of true gradient and a zero mean noise $\hat{\nabla}_{\alpha}J[n] = \nabla_{\alpha}J[n] + 2N_{\alpha}[n]$ in [20] which is then plugged-into (A.1) to yield

$$\begin{aligned} \varepsilon_{\alpha}^{MMSE}[n+1] &= \varepsilon_{\alpha}^{MMSE}[n] - \mu_{\alpha}(P_{\tilde{x}}\varepsilon_{\alpha}[n] + N_{\alpha}[n]) \\ &= \varepsilon_{\alpha}^{MMSE}[n](1-\mu_{\alpha}P_{\tilde{x}}) - \mu_{\alpha}N_{\alpha}[n] \end{aligned} \quad \text{(A.7)}$$

where the equality $\tilde{x}^{*}[n]e[n] = -\frac{1}{2}\hat{\nabla}_{\alpha}J[n] = -(P_{\tilde{x}}\alpha[n] - P_{\tilde{x}}\alpha_{MMSE}) - N_{\alpha}[n]$ is utilized. Then taking the absolute square and applying expectation to (A.7) yields

$$\begin{aligned} E[\left|\varepsilon_{\alpha}^{MMSE}[n+1]\right|^{2}] = E[\left|\varepsilon_{\alpha}^{MMSE}[n]\right|^{2}](1-\mu_{\alpha}P_{\tilde{x}})^{2} - 2\mu_{\alpha}(1-\mu_{\alpha}P_{\tilde{x}})\,\mathrm{Re}\{E[\varepsilon_{\alpha}^{MMSE}[n]N_{\alpha}^{*}[n]]\} \\ +\mu_{\alpha}^{2}E[\left|N_{\alpha}[n]\right|^{2}] \end{aligned} \quad \text{(A.8)}$$

which can be further simplified by assuming that ε_{α} and N_{α} are uncorrelated and therefore dropping the second term out of the equation. In addition if the steady-state is reached where the true gradient is zero and the adaptation is merely noise, i.e, $\tilde{x}^{*}[n]e_{o}[n] \approx -N_{\alpha}[n]$, the last term in (A.8) becomes $E[|N_{\alpha}|^{2}] \approx E[|e_{o}[n]|^{2}]E[|\tilde{x}[n]|^{2}] = L_{c}P_{IMDa}P_{\tilde{x}}$. However; it is worth mentioning that this is actually a crude approximation. It is true that $e_{o}[n] = \tilde{d}[n]$ is

uncorrelated with $\tilde{x}[n]$ but this does not guarantee $|\tilde{e}_o|^2$ to be uncorrelated with $|\tilde{x}|^2$. In fact the variance of $\tilde{d}[n]$ directly depends on the variance of $\tilde{x}[n]$ as illustrated in (3.9).

$$\begin{aligned} E[\left|\varepsilon_{\alpha}^{MMSE}[n+1]\right|^2] &= E[\left|\varepsilon_{\alpha}^{MMSE}[n]\right|^2](1-\mu_{\alpha}P_{\tilde{x}})^2 + \mu_{\alpha}^2 L_c P_{IMDa} P_{\tilde{x}} \\ &= E[\left|\varepsilon_{\alpha}^{MMSE}[0]\right|^2](1-\mu_{\alpha}P_{\tilde{x}})^{2(n+1)} + \mu_{\alpha}^2 L_c P_{IMDa} P_{\tilde{x}} \sum_{i=0}^{n} (1-\mu_{\alpha}P_{\tilde{x}})^{2i} \end{aligned} \quad \text{(A.9)}$$

We again consider that the step size is selected properly to make the first term in (A.9) to vanish asymptotically. Then the mean square of the mismatch is

$$\begin{aligned} \lim_{n\to\infty} E[\left|\varepsilon_{\alpha}^{MMSE}[n]\right|^2] &= \frac{\mu_{\alpha}^2 L_c P_{IMDa} P_{\tilde{x}}}{1-(1-\mu_{\alpha}P_{\tilde{x}})^2} = \frac{\mu_{\alpha}^2 L_c P_{IMDa} P_{\tilde{x}}}{\mu_{\alpha}P_{\tilde{x}}(2-\mu_{\alpha}P_{\tilde{x}})} \\ &= \frac{\mu_{\alpha} L_c P_{IMDa}}{2-\mu_{\alpha}P_{\tilde{x}}} \end{aligned} \quad \text{(A.10)}$$

which is well in-line with the general result [20] $E[|\varepsilon|^2] = \mu J_o / (2-\mu P_{in})$ where J_o is the value of the cost function with optimum Wiener coefficient and P_{in} is the input power of Wiener filter.

A.2 Wiener-Hammerstein Memory Model

When the Wiener-Hammerstein memory model is used for the main amplifier, and the adaptation structure is not changed, i.e. the structure in Figure 5-3, the above analysis needs only a few modifications. Starting with the mean convergence, (A.3) still holds but now e_o is not only the *IMD* term $\tilde{d}$ but also includes the terms from input and past values of *IMD,* i.e.

$$e_o^{mem}[n] = \alpha_o \sqrt{L_c} \sum_{m=0}^{M} b_{wh}[m]\tilde{x}[n-m] - \alpha_{MMSE}^{mem}\tilde{x}[n] + \sqrt{L_c}\sum_{l=1}^{L} b_h[l]\tilde{d}[n-l] \quad \text{(A.11)}$$

where $\varepsilon_{\alpha}^{off}$ is defined in (5.10). However, this does not change anything from mean convergence point of view since $e_o^{mem}[n]$ is still orthogonal to $\tilde{x}[n]$ by definition. We can re-write (A.4) as

$$E[\varepsilon_{\alpha,mem}^{MMSE}[n+1]] = E[\varepsilon_{\alpha,mem}^{MMSE}[n]] - \mu_{\alpha}E[\varepsilon_{\alpha,mem}^{MMSE}[n]\tilde{x}[n]\tilde{x}^*[n]] + \mu_{\alpha}\underbrace{E[\tilde{x}^*[n]e_o^{mem}[n]]}_{0} \quad \text{(A.12)}$$

where $\varepsilon_{\alpha,mem}^{MMSE}[n] = \alpha^{mem}[n] - \alpha_{MMSE}^{mem}$. Furthermore the white input (or the small step-size) assumption will allow the second term of (A.12) to have exactly the same form as (A.5). Then

under those conditions we can again state that the *SCL* coefficient α this is asymptotically convergent in the mean sense, that is

$$\lim_{n\to\infty} E[\varepsilon_{\alpha,mem}^{MMSE}[n]] = 0 \quad 0 < \mu_\alpha < 2/P_{\tilde{x}} \tag{A.13}$$

There is a slight change, however, for the mean square analysis due to the term $E[|N_\alpha^{mem}|^2] = E[|e_o^{mem}[n]|^2]E[|\tilde{x}[n]|^2]$ where

$$\begin{aligned} E[|e_o^{mem}[n]|^2] &= |\alpha_o|^2 L_c \sum_{m=0}^{M}\sum_{p=0}^{M} b_{wh}[m]b_{wh}^*[p]R_{\tilde{x}}[p-m] + R_{\tilde{x}}[0]|\alpha_{MMSE}^{mem}|^2 \\ &\quad -2\,\mathrm{Re}\{(\alpha_{MMSE}^{mem})^*\alpha_o\sqrt{L_c}\sum_{m=0}^{M} b_{wh}[m]R_{\tilde{x}}[-m]\} \\ &\quad +L_c\sum_{l=0}^{L}\sum_{k=0}^{L} b_h[l]b_h^*[k]R_{\tilde{d}}[k-l] \\ &= |\alpha_o|^2 L_c \sum_{m=0}^{M}\sum_{p=0}^{M} b_{wh}[m]b_{wh}^*[p]R_{\tilde{x}}[p-m] - R_{\tilde{x}}[0]|\alpha_{MMSE}^{mem}|^2 \\ &\quad +L_c\sum_{l=0}^{L}\sum_{k=0}^{L} b_h[l]b_h^*[k]R_{\tilde{d}}[k-l] \end{aligned} \tag{A.14}$$

instead of $E[|e_o[n]|^2] = L_c P_{IMDa}$ of the memoryless case. Therefore the mean square can be written as

$$\lim_{n\to\infty} E[|\varepsilon_{\alpha,mem}^{MMSE}[n]|^2] = \frac{\mu_\alpha E[|e_o^{mem}[n]|^2]}{2-\mu_\alpha P_{\tilde{x}}} \tag{A.15}$$

However; we can approximate $E[|e_o^{mem}[n]|^2]$ by $L_c P_{IMDa}$, if the memory is rather weak. Then a second order statistic of the mismatch as given in (A.10) is also valid for Wiener-Hammerstein memory case as well.

Appendix B

Mean and Mean-Squared Convergence of *ECL* Coefficient

We will follow a similar analysis as in Appendix A while deriving the mean and mean-squared convergence of *ECL* coefficient both for amplifier with and without memory.

B.1 Memoryless Model

Starting with the memoryless case, we can re-write the basic update equation given in (5.47)

$$\beta[n+1] - \beta_{MMSE} = \beta[n] - \beta_{MMSE} + \mu_\beta \sqrt{G_e}\tilde{z}[n]\tilde{e}^*[n] \tag{B.1}$$

and then denoting $\varepsilon_\beta^w[n] = \beta[n] - \beta_{MMSE}$

$$\begin{aligned} \varepsilon_\beta^{MMSE}[n+1] &= \varepsilon_\beta^{MMSE}[n] + \mu_\beta \tilde{e}^*[n]\tilde{z}[n] \\ &= \varepsilon_\beta^{MMSE}[n] + \mu_\beta \sqrt{G_e}\tilde{e}^*[n](\tilde{y}[n] - \sqrt{G_e}\beta[n]\tilde{e}[n]) \end{aligned} \tag{B.2}$$

Writing the second term in parenthesis as $\tilde{y}[n] + \sqrt{G_e}(-\beta_{MMSE}\tilde{e}[n] + \beta_{MMSE}\tilde{e}[n] - \beta[n]\tilde{e}[n])$, (B.2) can be re-structured

$$\begin{aligned} \varepsilon_\beta^{MMSE}[n+1] &= \varepsilon_\beta^{MMSE}[n] + \mu_\beta \sqrt{G_e}\tilde{e}^*[n]\tilde{z}_o[n] - G_e\mu_\beta\varepsilon_\beta^{MMSE}[n]\tilde{e}[n]\tilde{e}^*[n] \\ &= \varepsilon_\beta^{MMSE}[n](1 - G_e\mu_\beta\tilde{e}[n]\tilde{e}^*[n]) + \mu_\beta\sqrt{G_e}\tilde{e}^*[n]\tilde{z}_o[n] \end{aligned} \tag{B.3}$$

where $\tilde{z}_o[n]$ is the optimal error signal in Wiener filter structure (or in this case linearizer output calculated with optimal Wiener coefficient). Now applying expectation to both sides of (B.3)

$$\begin{aligned} E[\varepsilon_\beta^{MMSE}[n+1]] &= E[\varepsilon_\beta^{MMSE}[n](1 - G_e\mu_\beta\tilde{e}[n]\tilde{e}^*[n])] + \mu_\beta\sqrt{G_e}E[\tilde{e}^*[n]\tilde{z}_o[n]] \\ &= E[\varepsilon_\beta^{MMSE}(1 - G_e\mu_\beta\tilde{e}[n]\tilde{e}^*[n])] \end{aligned} \tag{B.4}$$

since $\tilde{z}_o[n]$ is orthogonal to $\tilde{e}[n]$ by definition. The right hand side of (B.4) is separable under *independence assumption,* which requires the filter input $\tilde{e}$ to be white. For the memoryless case, $\tilde{e}$ that is given in (5.2), whiteness can be achieved by having white $\tilde{d}$ with ε_α being small enough to make $\tilde{d}$ the dominant signal. As discussed in Section 3.2, $\tilde{x}$ should be white for $\tilde{d}$ to be white. Since mathematical tractability of β requires α to be adjusted first (if ε_α is a random process) to the level that enough suppression of $\tilde{x}$ is achieved, we will use this assumption throughout this work wherever appropriate. Once the expectation in (B.4) is separated under these assumptions, then it is easy to see that

$$\begin{aligned}\lim_{n\to\infty} E[\varepsilon_\beta^{MMSE}[n+1]] &= E[\varepsilon_\beta^{MMSE}[n]](1 - G_e\mu_\beta E[\tilde{e}[n]\tilde{e}^*[n]]) \\ &= E[\varepsilon_\beta^{MMSE}[0]](1-\mu_\beta G_e P_{\tilde{e}})^{n+1} = 0 \qquad 0 < \mu_\beta < \frac{2}{G_e P_{\tilde{e}}}\end{aligned} \tag{B.5}$$

where $P_{\tilde{e}} \approx L_c P_{IMDa}$. Therefore convergence to the optimum Wiener coefficient given in (5.24) is achieved in the mean sense under the assumptions stated above and with the proper choice of step size indicated in (B.5).

We will also follow the same procedure as we did in Appendix A for the second order characterization of ε_β^{MMSE}. We start our derivation by re-writing (B.1) where the product $\sqrt{G_e}\tilde{z}[n]\tilde{e}^*[n]$ is replaced with the noisy gradient that is the superposition of true gradient and zero-mean gradient noise, $\frac{-1}{2}\hat{\nabla}_\beta J = \frac{-1}{2}(\nabla_\beta J + 2\tilde{N}_\beta) = -(G_e P_{\tilde{e}}\varepsilon_\beta^{MMSE}[n] + \tilde{N}_\beta[n])$

$$\varepsilon_\beta^{MMSE}[n+1] = \varepsilon_\beta^{MMSE}[n](1-\mu_\beta P_{\tilde{e}} G_e) - \mu_\beta \tilde{N}_\beta[n] \tag{B.6}$$

Then multiplying (B.6) by its complex conjugate and applying expectation, we obtain

$$\begin{aligned}E[\left|\varepsilon_\beta^{MMSE}[n+1]\right|^2] = E[\left|\varepsilon_\beta^{MMSE}[n]\right|^2](1-\mu_\beta P_{\tilde{e}} G_e)^2 \\ -2\mu_\beta(1-\mu_\beta P_{\tilde{e}} G_e)\,\mathrm{Re}\{E[\varepsilon_\beta^{MMSE}[n]\tilde{N}_\beta^*[n]]\} \\ +\mu_\beta^2 E[\left|\tilde{N}_\beta\right|^2]\end{aligned} \tag{B.7}$$

The second term on right hand side of (B.7) will be dropped by assuming that ε_β^w and $\tilde{N}_\beta$ are uncorrelated. After this, when the iteration is propagated till *n=0*

$$\begin{aligned}E[\left|\varepsilon_\beta^{MMSE}[n+1]\right|^2] = E[\left|\varepsilon_\beta^{MMSE}[0]\right|^2](1-\mu_\beta P_{\tilde{e}} G_e)^{2(n+1)} \\ +\mu_\beta^2 E[\left|\tilde{N}_\beta\right|^2]\sum_{i=0}^{n}(1-\mu_\beta P_{\tilde{e}} G_e)^{2i}\end{aligned} \tag{B.8}$$

We will also assume that steady-state is reached and the adaptation is merely noise, i.e. $\sqrt{G_e}\tilde{z}[n]\tilde{e}^*[n] \approx \tilde{N}_\beta$. In that case the mean square of gradient noise can be approximated by $E[\left|\tilde{N}_\beta\right|^2] \approx G_e E[|\tilde{z}_o[n]|^2]E[|\tilde{e}[n]|^2] \approx |\alpha_o|^2 P_{\tilde{x}} P_{\tilde{e}} G_e$. As discussed in Appendix A, the approximation is crude due to the dependency of the variance of *IMD* to variance of amplifier input. Finally regarding (B.8) with a properly chosen step size, the mean squared error converges asymptotically to

$$\begin{aligned} \lim_{n\to\infty} E[\left|\varepsilon_\beta^{MMSE}[n+1]\right|^2] &= \frac{\mu_\beta^2 |\alpha_o|^2 P_{\tilde{x}} P_{\tilde{e}}}{1-(1-\mu_\beta P_{\tilde{e}} G_e)^2} \\ &= \frac{\mu_\beta |\alpha_o|^2 P_{\tilde{x}}}{(2-\mu_\beta P_{\tilde{e}} G_e)} \end{aligned} \tag{B.9}$$

which is again of the well known form $E[|\varepsilon|^2] = \mu J_o / (2 - \mu P_{in})$ as also mentioned in Appendix A. Remembering that for the memoryless amplifier model, $\tilde{e}[n]$ is given in (5.2), under the assumption of white $\tilde{x}$, we can alternatively write $P_{\tilde{e}}$ and thus (B.9) as

$$E[\left|\varepsilon_\beta^{MMSE}\right|^2] = \frac{\mu_\beta |\alpha_o|^2 P_{\tilde{x}}}{(2-\mu_\beta(E[\left|\varepsilon_\alpha^{MMSE}\right|^2]P_{\tilde{x}} + L_c P_{IMDa})G_e)} \tag{B.10}$$

where $E[\left|\varepsilon_\alpha^{MMSE}\right|^2]$ is given in (A.10). In addition even if the estimate of $E[\left|\varepsilon_\alpha^{MMSE}\right|^2]$ is rather crude due to the reason explained in Appendix A, the dominating term in the denominator of (B.10) is 2 with the assumption that both error amplifier input and μ_β are small. Then for the mean-square error we can write

$$E[\left|\varepsilon_\beta^{MMSE}\right|^2] \approx \frac{\mu_\beta |\alpha_o|^2 P_{\tilde{x}}}{2} \tag{B.11}$$

B.2 Wiener-Hammerstein Memory Model

Now let's do the derivations starting from mean convergence for the case where there is Wiener-Hammerstein type memory. Expression (B.5) stated that β is convergent to β_{MMSE} in the mean sense when $\tilde{e}$ is white which requires $\tilde{x}$ to be white and ε_α to be small. However, when there is memory such that $\tilde{e}$ is given in (5.31), small ε_α and white $\tilde{x}$ does not suffice for $\tilde{e}$ to be white since the memory part of $\tilde{x}$ still remains. In addition, even if $\tilde{d}$ is white, the last term of (5.31) which is a filtered version of $\tilde{d}$ is not whit. However, we can assume that the memory is not strong to cause a high frequency selectivity and the memory weights of $\tilde{x}$ at times prior to *n* are rather small. In that case, the dominant term will be the

slightly filtered $\tilde{d}$ which can still be approximated as white. Under these assumptions in addition to the assumption in the memoryless case, we can state that β is convergent in the mean sense to the optimum Wiener coefficient given in (5.37).

The exact expression for mean square mismatch is rather difficult due to the expression of $E[|z_o[n]|^2]$ when there is memory. However, if $E[\left|\varepsilon_{\beta,mem}^{MMSE}[n]\right|^2]$ is small, i.e. μ_α and μ_β are small, so that the dominating term of (5.30) is the first one, then we can use (B.11) for the mean square error in Wiener-Hammerstein memory model as well.