PA efficiency enhancement using digital linearization techniques in uplink cognitive radio systems

Mouna Ben Mabrouk

N° d'ordre : – – –

Année 2015

THÈSE

présentée
pour obtenir le titre de

DOCTEUR DE L'UNIVERSITÉ DE BORDEAUX

Spécialité : Automatique, Productique, Signal et Image, Ingénierie Cognitique

par Mouna BEN MABROUK

PA Efficiency Enhancement Using Digital Linearization Techniques In Uplink Cognitive Radio Systems

Soutenue le 02 Décembre 2015 devant le jury composé de :

Mme. BAUDOIN Geneviève	Professeur à l'ESIEE, Paris	Présidente du jury
M. LOUET Yves	Professeur à Supélec, Campus de Rennes, Rennes	Rapporteur
M. BROSSIER Jean-Marc	Professeur à Grenoble INP, Grenoble	Rapporteur
Mme. DELTIMPLE Nathalie	Maître de Conférences à l'ENSEIRB-MATMECA, Bordeaux INP, Talence	Co-directrice de thèse
M. Ferré Guillaume	Maître de Conférences à l'ENSEIRB-MATMECA, Bordeaux INP, Talence	Co-directeur de thèse
M. Grivel Eric	Professeur à l'ENSEIRB-MATMECA, Bordeaux INP, Talence	Directeur de thèse

Préparée au Laboratoire IMS, UMR CNRS 5218, Université de Bordeaux, Bordeaux INP
351, avenue de la Libération - 33405 Talence Cedex

Contents

List of Figures

List of Tables

Abstract

For a battery driven terminal, the power amplifier (PA) efficiency must be optimized. Consequently, non-linearities may appear at the PA output in the transmission chain. To compensate these distortions, one solution consists in using a digital post-distorter based on a Volterra model of both the PA and the channel and a Kalman filter (KF) based algorithm to jointly estimate the Volterra kernels and the transmitted symbols. Here, we suggest addressing this issue when dealing with uplink cognitive radio (CR) system. In this case, additional constraints must be taken into account. Since the CR terminal may switch from one subband to another, the PA non-linearities may vary over time. Therefore, we propose to design a digital post-distorter based on an interacting multiple model combining various KF based estimators using different model parameter dynamics. This makes it possible to track the time variations of the Volterra kernels while keeping accurate estimates when those parameters are static. Furthermore, the single carrier case is addressed and validated by simulation results. In addition, the relevance of the proposed approach is confirmed by measurements carried on a (300-3000) MHz broadband PA.

This work is published in:

1. M. Ben Mabrouk, G. Ferré, E. Grivel and N. Deltimple, "Interacting multiple model based detector to compensate power amplifier distortions in cognitive radio," IEEE Transactions on Communications, Vol. 63, no.5, pp. 1580-1593, 2015.

2. M. Ben Mabrouk, G. Ferré, E. Grivel and N. Deltimple, "A new baseband post-distortion technique for power amplifiers in OFDM-based cognitive radio systems," IEEE International Symposium on Circuits and Systems (ISCAS), pp. 1486-1489, 2015.

3. M. Ben Mabrouk, G. Ferré, E. Grivel and N. Deltimple, "A novel digital post- distortion and detection technique for RF power amplifiers in cognitive radio systems," IEEE International Symposium Wireless Symposium (IWS), pp. 1-4, 2015.

4. M. Ben Mabrouk, E. Grivel, C. Magnant, G. Ferré and N. Deltimple, "Compensating power amplifier distortion in cognitive radio systems with adaptive interacting multiple model," Eurasip European Signal Processing Conference (EUSIPCO), 2015.

Résumé

Pour un terminal mobile alimenté sur batterie, le rendement de l'amplificateur de puissance (AP) doit être optimisé. Cette optimisation peut rendre non-linéaire la fonction d'amplification de l'AP. Pour compenser les distorsions introduites par le caractère non-linéaire de l'AP, un détecteur numérique fondé sur un modèle de Volterra peut être utilisé. Le comportement de l'AP et le canal étant modélisé par le modèle de Volterra, une approche par filtrage de Kalman (FK) permet d'estimer conjointement les noyaux de Volterra et les symboles transmis. Dans ce travail, nous proposons de traiter cette problématique dans le cadre d'une liaison montante dans un contexte radio intelligente (RI). Dans ce cas, des contraintes supplémentaires doivent être prises en compte. En effet, étant donné que la RI peut changer de bande de fréquence de fonctionnement, les non-linéarités de l'AP peuvent varier en fonction du temps. Par conséquent, nous proposons de concevoir une post-distorsion numérique fondée sur une modélisation par modèles multiples combinant plusieurs estimateurs à base de FK. Les différents FK permettant de prendre en compte les différentes dynamiques du modèle. Ainsi, les variations temporelles des noyaux de Volterra peuvent être suivies tout en gardant des estimations précises lorsque ces noyaux sont statiques. Le cas d'un signal monoporteuse est adressé et validé par des résultats de simulation. Enfin, la pertinence de l'approche proposée est confirmée par des mesures effectuées sur un AP large bande (300-3000) MHz.

Ce travail a été publié dans:

1. M. Ben Mabrouk, G. Ferré, E. Grivel and N. Deltimple, "Interacting multiple model based detector to compensate power amplifier distortions in cognitive radio," IEEE Transactions on Communications, Vol. 63, no.5, pp. 1580-1593, 2015.

2. M. Ben Mabrouk, G. Ferré, E. Grivel and N. Deltimple, "A new baseband post-distortion technique for power amplifiers in OFDM-based cognitive radio systems," IEEE International Symposium on Circuits and Systems (ISCAS), pp. 1486-1489, 2015.

3. M. Ben Mabrouk, G. Ferré, E. Grivel and N. Deltimple, "A novel digital post- distortion and detection technique for RF power amplifiers in cognitive radio systems," IEEE International Symposium Wireless Symposium (IWS), pp. 1-4, 2015.

4. M. Ben Mabrouk, E. Grivel, C. Magnant, G. Ferré and N. Deltimple, "Compensating power amplifier distortion in cognitive radio systems with adaptive interacting multiple model," Eurasip European Signal Processing Conference (EUSIPCO), 2015.

Amélioration du rendement de l'amplificateur de puissance en utilisant une technique de linéarisation numérique pour une liaison montante dans un contexte radio intelligente.

Keywords : Power amplifier, digital post-distortion, cognitive radio, Volterra model, Kalman filtering, interacting multiple model, single carrier.

Mots clés : Amplificateur de puissance, post-distortion numérique, radio intelligente, modèle de Volterra, filtrage de Kalman, modèles multiples en interaction, mono-porteuse.

Résumé

Les systèmes de communication sans-fil actuels sont conçus pour fonctionner sur des bandes de fréquences invariantes (ou allocation statique des fréquences). Ainsi, dans un contexte de téléphonie mobile, les terminaux sont conçus en tenant compte des contraintes induites par les bandes de fréquences à utiliser. Cette conception permet entre autre d'optimiser l'autonomie des terminaux en minimisant leur consommation d'énergie.

L'élément le plus consommateur d'énergie est l'amplificateur de puissance (PA) situé juste avant l'antenne à l'émission. Pour satisfaire les contraintes de haut débit des standards actuels, les PAs doivent fonctionner en mode linéaire, ce qui affecte leur rendement [1] et augmente la consommation d'énergie. L'allocation statique des bandes de fréquences à entrainer la saturation du spectre des fréquences "électroniquement" utilisables. Cependant, certaines bandes de fréquences étant sous exploitées, le déploiement des futurs systèmes de communications haut débit sans-fil passera par la réutilisation de ces dernières. On parle désormais d'allocation dynamique du spectre, c'est le concept de la radio intelligente. Afin de rendre effective cette gestion dynamique du spectre, les composants de l'étage radio fréquence, notamment le PA [2], et les traitements effectués au niveau de la couche physique des futurs terminaux intelligents doivent être en rupture avec ce qui est fait actuellement. En effet, l'accès dynamique au spectre laisse la possibilité au terminal intelligent d'utiliser une bande de fréquence dans une gamme très étendue (de quelques MHz à plusieurs GHz). De plus, en fonction du débit souhaité au niveau de la communication secondaire et de la disponibilité du spectre, on attendra du terminal intelligent qu'il :

1. émette dans des bandes de fréquence qui pourront être discontinues et possiblement très éloignées,
2. n'interfère pas avec les utilisateurs primaires,
3. minimise sa consommation d'énergie.

Le PA d'une radio intelligente doit donc être très large bande et à haut rendement. Ces deux conditions impliquent que le PA se comporte comme un système non-linéaire à mémoire. Notons que ce comportement varie en fonction du triplet (fréquence porteuse, bande passante, puissance du signal). Dans le cadre de cette

1. Le rendement du PA est défini comme le rapport entre la puissance radio fréquence de sortie et la puissance d'alimentation du PA.
2. situé juste avant l'antenne à l'émission

thèse, nous sommes amenés à traiter deux problématiques liées au traitement numérique du signal : la modélisation des distorsions induites par le PA d'une part, le développement d'un algorithme de traitement de signal permettant l'identification du modèle suivi par le PA et pouvant contrer ses non-linéarités à la réception, d'autre part. Ainsi, il s'agit de concevoir un post-distordeur numérique d'un PA pour la radio intelligente.

Lorsque le PA fonctionne proche du point de compression du gain, l'enveloppe du signal d'entrée est distordue entraînant l'apparition de nouveaux harmoniques dans la bande de communication mais aussi en dehors. Afin de modéliser ce comportement, nous avons choisi un modèle non-linéaire à mémoire, le modèle de Volterra. Nous avons choisi le modèle de Volterra car il est linéaire par rapport à ses paramètres et non linéaire par rapport au signal d'entrée ce qui rend son identification plus simple. De plus, il a l'avantage d'être stable au sens entrée bornée, sortie bornée. Les paramètres de ce modèle évoluent en fonction du temps d'autant plus que le terminal intelligent change de bande. En effet, ces paramètres sont constants lorsque le terminal intelligen envoie sur une sous-bande et variant dans le temps lorsqu'il change de bande. En conséquence de quoi, nous introduisons que le PA se comporte selon plusieurs modes ; constant et variant dans le temps.

L'idée clé du post-distordeur développé est l'identification conjointe des paramètres du modèle et du signal d'entrée en utilisant une technique d'estimation. Dans cette thèse, une approche reposant sur un filtrage optimal de type Kalman a été choisie. Pour cette technique, la représentation dans l'espace d'état qui traduit la dynamique du système joue un rôle clé et repose sur un mode donné. Pour relâcher cette contrainte relative au choix du modèle, les approches à modèles multiples peuvent être envisagées. Plusieurs algorithmes conçus pour l'identification des systèmes à modèles multiples sont définis dans la littérature, parmi lesquels on peut citer le GPB1, le GPB2 (Generalized Pseudo Bayesian) et l'IMM (Interacting Multiple Model). Comme l'IMM offre un bon compromis entre performance et coût calculatoire, nous l'avons choisi comme une structure de base du post-distordeur numérique.

Le post-distordeur a été développé pour des communications mono-porteuse. Pour valider la pertinence de l'approche de modélisation et du post-distordeur proposés, deux méthodes ont été utilisées :

— simulations au niveau système : La conception de l'IMM repose sur trois paramètres :

— le choix du type des modes,
— le choix du nombre de modes,
— le choix de la probabilité de transition entre eux.

Une étude a été menée pour valider la pertinence de cette approche en testant sa sensibilité à chacun de ces paramètres.

— des mesures sur des PAs réels : un test bench a été mis en œuvre pour représenter les différents modules de la chaine de communication. Le montage se compose d'un générateur de signaux, d'un PA commercial, d'une alimentation à courant continu, d'un oscilloscope numérique et d'un ordinateur équipé de deux logiciels ; Matlab et Labview. Les mesures ont été faites pour différentes fréquences porteuses et différentes valeurs de puissance du signal d'entrée.

La solution proposée correspond à un compromis entre le coût calculatoire et le taux d'erreur binaire enregistré. La pertinence de l'approche a été démontrée sur deux plan ; théorique et pratique. Le travail réalisé a donné lieu à plusieurs publications dans des revues et des conférences internationales.

Glossary

ACPR	Adjacent channel power ratio
ACEPR	Adjacent channel error power ratio
ADC	Analog-to-digital converter
AM-AM	Amplitude-to-amplitude modulation
AM-PM	Amplitude-to-phase modulation
BBE	Baseband equivalent
BER	Bit error rate
BPSK	Binary phase-shift keying
CDKF	Central difference Kalman filter
CMP	CR-PA model path
CP1	1-dB compression point
CR	Cognitive radio
CR-PA	Cognitive radio power amplifier
CT	Cognitive terminal
D/A	Digital-to-analog
DC	Direct current
DDR	Dynamic derivation reduction
DLA	Direct learning architecture
DO	Digital oscilloscope
DSP	Digital signal processor
DUT	Divice under test
EKF	Extended Kalman filter
ET	Envelop tracking
EVM	Error vector magnitude
FBMC	Filter bank multicarrier
FFT	Fast Fourier transform
FPGA	Field programmable gate arrays
GFDM	Generalized frequency division multiplexing
GM-EKF	Gaussian mixture extended Kalman filter
GMP	Generalized memory polynomial
GPB	Generalized pseudo Bayesian
GPIB	General purpose interface bus

GPS	Global position system
GSM	Global system for mobile communication
HW	Hammerstein-Wiener
IBO	Input back-off
ICP1	Input 1-dB compression point
IF	Intermediate frequency
IFFT	Inverse fast Fourier transform
IIP3	Input third-order intercept point
ILA	Indirect learning architecture
IMD	Intermodulation
IMM	Interacting multiple model
IP3	Third-order intercept point
KF	Kalman filter
LNA	Low noise amplifier
LPE	Low-pass equivalent
LS	Least squares
LTE	Long term evolution
LTI	Linear time-invariant
MM	Multiple model
MP	Memory polynomial
MSE	Mean square error
NARX	Non-linear autoregressive with exogenous input
NL	Non-linear
NN	Neural network
OBO	Output back-off
OFDM	Orthogonal frequency division multiplexing
OIP3	Output third-order intercept point
PA	Power amplifier
PAE	Power added efficiency
PAPR	Peak-to-average power ratio
PBO	Peak back-off
PC	Personal computer
PSD	Power spectral density
QKF	Quadrature Kalman filter
QPSK	Quadrature phase-shift keying
QoS	Quality-of-service
RADAR	Radio detection and ranging
RF	Radio frequency
SDR	Software defined radio
SFDR	Spurious-free dynamic range
SNR	Signal-to-noise ratio
SPKF	Sigma point Kalman filter

SSG	Serial signal generator
SSR	State space representation
TPM	Transition probability matrix
UKF	Unscented Kalman filter
UMTS	Universal mobile telecommunications system
VI	Virtual instrument
WH	Wiener-Hammerstein

Notations

Chapitre 1

a_τ	Impulse response values of the LTI systems before the memoryless non-linear block
b_k	Coefficients of the odd-order polynomial
B_W	Signal bandwidth
C	Capacitance
C_{par}	Parallel capacitor
c_τ	Impulse response values of the LTI systems after the memoryless non-linear block
$d()$	Distortion
DCB	Drain output capacitor
f_c	Input signal carrier frequency
$f_{DPD}()$	DPD transfert function
$f_{PA}()$	PA transfer function
G	Gain of the PA
G_{1dB}	Gain at the 1-dB compression point
G_a	Available gain
G_{Lin}	Linear gain at the 1-dB compression point
G_P	Power gain
G_t	Transducer gain
$i_C()$	Current flowing through a capacitor
$i_L()$	Current flowing through an inductor
$IP3$	Third-order intercept point
NF	PA noise figure
P_{DC}	DC power
P_{in}	Input power
$P_{in,av}$	Available power at the input
$P_{in,d}$	Power delivered by the generator
P_{out}	Output power
$P_{out,av}$	Available power at the output
$P_{out,d}$	Power delivered to the load
P_{outmax}	Peak output power
P_{sat}	Ratio of saturated output power

$S()$	PA output power spectral density
$u(t)$	Continuous PA input signal
u_n	Discrete PA input signal
$v_C()$	Voltage applied to the capacitor
VDD	DC supply voltage
v_{ds}	Voltage applied to the capacitor
$v_L()$	Voltage applied to the capacitor
W	Input signal bandwidth
W_1	Adjacent channel bandwidth
$y(t)$	Output signal of PA
η	DC-to-RF efficiency
Δt	Delay of the feedback loop

Chapitre 2

$(\sigma^2)^i$	Innovation variance
$\{\mathbf{P}^j_{k\|k}\}_{j=st,sd,ld}$	Error covariance matrices provided by each GM-EKF
$\{\hat{\mathbf{x}}^j_{k\|k}\}_{j=st,sd,ld}$	Last estimated states provided by each GM-EKF
μ^j_k	Probability that the system corresponds to the j^{th} mode m_j at time k
$\mathbf{C}_k$	Vector storing the Volterra kernels to be estimated
$\mathbf{D}_k$	Vector storing the samples to be estimated
$\mathbf{F}$	State transition matrix
$\mathbf{h}_i(k)$	Jacobian matrix
$h(\hat{\mathbf{x}}^j_{k\|k})$	Predicted observation in the m_j-GM-EKF
$h_n()$	Volterra model kernels
$\mathbf{K}^{i,st}_k$	Kalman gain for the i^{th} EKF in the m_{st}-GM-EKF
L	IFFT size
ld	Large dynamics
M	Memory depth
$\mathcal{N}(\mu, \sigma^2)$	Gaussian distribution with mean μ and variance σ^2
$\mathcal{N}(z_k; \mu, \sigma^2)$	Value of the Gaussian distribution with mean μ and variance σ^2 when the random variable is equal to z_k
n_k	Zero-mean white Gaussian vector with covariance matrix σ^2_n
P	Non-linearity order
$p(x_k\|z_{1:k})$	State vector distribution
$\mathbf{P}^{i,st}_{k\|k-1}$	*A priori* error covariance matrix
$\mathbf{S}^j_k$	Innovation covariance matrix corresponding to the m_j-GM-EKF
S_l	Symbols assumed to be uniformly distributed at the input of the IFFT in an OFDM system
u_k	PA input sample
$w_k(m_{st})$	Model noise for the m_{st}-GM-EKF
$\hat{\mathbf{x}}_{k\|k}$	*A posteriori* state estimate

$\hat{\mathbf{x}}_{k\|k}^{i,st}$	*A posteriori* state estimate of the m_{st}-GM-EKF
$\mathbf{x}_k$	State vector
y_k	PA output sample
$z_{1:k}$	Set of observations until time k
Λ_k^j	Likelihood functions corresponding to the m_j-GM-EKF
$\mathbf{\Pi}$	Transition probability matrix

Chapitre 3

$E(f)$	Discrete Fourier transform of e_n
e_n	Difference between y_n^{meas} and y_n^{mod}
f_c	Carrier frequency
P_{in}	Input power
P_{out}	Output power
$Y(f)$	Discrete Fourier transform of y_n
$\mathbf{y}_n^{meas}$	PA output signal
$\mathbf{y}_n^{mod}$	PA model output signal

Introduction

The rapid growth of wireless communications has made the problem of spectrum utilization more and more critical. On the one hand, the increasing diversity and demand for high quality-of-service (QoS) applications have resulted in an overcrowded allocated spectrum bands and a data rate decrease. On the other hand, major licensed bands, such as those allocated for television broadcasting, amateur radio, and paging, are significantly underused. This leads to spectrum leakage. One solution is to use the cognitive radio (CR) concept [Gol09]. Its basic idea is that the unlicensed users (also called the cognitive terminals (CTs)) are able to sense the spectrum and then, underlay, overlay or interweave their signals with those sent by the licensed users (also called the primary users) without impacting their transmissions. This makes it possible to enhance the spectrum utilization [Gha08] and increase the spectrum efficiency.

Consequently, CR systems require wideband spectrum sensing. For this reason, the transceiver analog part which consists of a PA, antennas, converters, filters and low noise amplifier (LNA) has to be as wideband as possible.

In this thesis, our purpose is to optimize the energy consumption of the mobile terminal by reducing the consumed power of the transmitter in an uplink CR communication. More particularly, we focus our attention on the PA as it is inherently the most power consumer components in the transmitter topology. Furthermore, the PA design usually aims at getting the lowest possible power consumption to enhance power efficiency. Nevertheless, the power efficiency is getting higher and higher when the device is driven more and more into its compression region. In return, this increases the amount of non-linearities [Vuo03]. These latter generate spectral regrowth, which leads to adjacent channel interference. Moreover, it causes signal distortions which are frequency dependent.

To avoid the PA non-linear behavior, the PA can be backed off to operate within the linear portion of its operating curve. However, amplitude modulation schemes have high peak-to-average power ratios as their envelopes have large fluctuations. This means that the PA needs to be backed off far from its saturation point, which results in very low efficiencies,

typically less than 10% [Cri06]. It also means that more than 90% of the DC power is lost and turns into heat.

To improve the PA efficiency without compromising its linearity, PA linearization is essential. As the so-called trade-off between PA efficiency and linearity has to be met, several approaches can be considered such as:

1. designing a linear PA and increasing its efficiency by using an efficiency enhancement technique,

2. designing a high-efficiency non-linear PA. Then, compensating its distortions by using a linearization technique.

The first method is generally used when linearity is required and less constraints about the PA output power are given whereas the second method is useful when the designer aims at reaching the highest possible output power or the lowest consumed power. In this PhD, we choose the second way and develop a linearization technique able at once to compensate the PA distortions and to avoid causing additional energy consumption at the transmitter.

For the recent years, digital linearization techniques have been proposed mainly thanks to the renewed possibilities offered by digital signal processing (DSP) [Vuo03]. Their key feature is to build a baseband distorter so that the transfer function of the whole system, including the PA and this distorter, is linear. There are two main families [Cri02]:

- the digital pre-distortion: in this case, the PA non-linearity is compensated before the signal transmission. This technique needs a feedback path to bring the image of the remaining non-linearities from the PA output to the pre-distorter. However, the feedback path composed of ADCs and mixers should have a large bandwidth, usually equal to 3 to 5 times the signal bandwidth, in order to pass non-linearity information (i.e. the 3^{rd} to 5^{th}-order intermodulation products) to the pre-distorter without overlapping [Zel12]. Hence, this adds extra power consumption.

- the digital post-distortion: it consists in creating a distorter element which is complementary to the distortion characteristic of the PA. In [Cri02], the authors propose to compensate the PA non-linearities at the baseband part of the receiver rather than at the radio frequency (RF) level of the transmitter.

As the spectrum usage and availability in CR may vary over time and depend on the location, the PA behavior and consequently its distortions change over time. For this reason, the linearization technique we have to propose must take into account the time-varying PA

behavior. In this PhD dissertation, we suggest using a digital post-distortion technique. This choice can be advocated by the following reasons:

- there is no need for feedback path at the transmitter,
- in an uplink communication, the detector calculations are done at the base station receiver. This decreases the amount of signal processing at the CT,
- as transmitting a non-linearly-amplified signal increases the signal power at the second and the third-order harmonics which may interfere with other users, the availability of the spectrum at these frequencies should be guaranteed. In CR systems, this can be done.

Therefore, our research efforts focus on three areas:

- modeling the PA behavior,
- developing a signal processing-based algorithm to build the post-distortion technique structure,
- validating the developed technique by measurement-based results.

This thesis is hence organized as follows:

Chapter 1 covers the general theory of RF amplifiers, including PA fundamentals and nonlinear analysis. The definitions of PA parameters and characteristics are provided. In addition, an overview of existing linearization procedures including the feedback, the feed-forward and the digital pre/post-distortion techniques is given. The theoretical concept, the main advantages and the drawbacks of each of the linearization techniques are then presented and discussed. A state of the art is also given concerning the PA behavioral model. This includes non-linear models such as the Volterra model.

In chapter 2, assuming that the PA and the channel can be modeled by a Volterra model, the digital post-distortion technique we propose consists in jointly estimating the input signal and the time-varying Volterra kernels. In a single-carrier system, we propose to study a Kalman filtering based estimator. For this reason, the state space representation (SSR) of the system is first given. Then, as the estimation issue is non-linear and the model noise in the SSR is not Gaussian but can be approximated by a sum of Gaussians, we suggest analyzing the relevance of a Gaussian-mixture extended Kalman filter (GM-EKF), initially introduced in [Ama11]. However, as the PA behavior may vary over time, the estimator must

be also able to track the time variations of the Volterra kernels when they occur. As the GM-EKF cannot meet these requirements, we propose to combine several GM-EKF based on different Volterra kernel dynamics in a multiple model (MM) structure. This leads to our first proposed algorithm called GM-EKF based IMM. It should be noted that when an IMM is used, the so-called transition probability matrix (TPM) which a design parameter managing the cooperation between the different estimators, is required. In this approach, it is *a priori* defined by using the available information that can be found in the CT database about the mean spectrum availability at a given location.
We study the performance of the algorithm using *Matlab* simulation and analyze the computational costs of the proposed algorithms.

In chapter 3, we propose an experimental setup and the associated results. Some parameters which influence the linearization performance in practical applications are studied, such as the receiver architecture and the non-linearity order of the model.

Finally, conclusions and perspectives are given in the last chapter.

Chapter 1

Power amplifiers and non-linearity

Contents

1.1 Introduction

A radio frequency (RF) transmitter is an important building block of a communications system. It converts the baseband signals containing the data to be transferred through the transport medium to the receiver. Figure 1.1 shows a simplified block diagram of a RF transmitter. The baseband signal processing is performed digitally. Thus, the baseband signals have first to be digital-to-analog (D/A) converted before being transferred to the transmitter. The conversion also requires re-construction filters to remove unwanted frequency components after the conversion. Generally, the baseband signals are in quadrature form, which allows the generation of phase, frequency and amplitude modulated signals. To combine these signals into a single analog signal for transmission, they are combined and converted into intermediate frequency (IF) or directly to RF using a quadrature modulator. This conversion can also be done using frequency and amplitude modulators, depending on the modulation method [Cri02]. If IF is used, the signal has still to be upconverted to RF. Then, the signal is amplified with a power amplifier (PA) for transmission and then fed to the antenna.

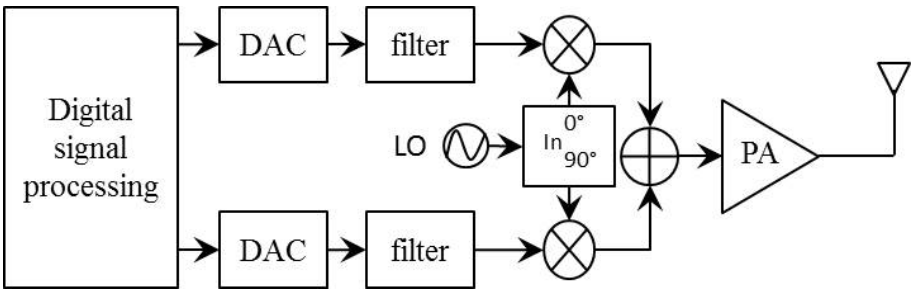

Figure 1.1: A simplified block diagram of an RF transmitter.

An amplifier is an electronic device used to increase the power of its input signal. A PA is an amplifier that is usually the final amplification stage and is designed to give a signal the necessary power for transmission to reach the receiver.

An ideal PA would have an output signal equal to the input signal multiplied by a scalar, i.e. would have a linear transfer characteristic. For this PA, the input and output powers are connected using the following relationship:

$$P_{out} = G \cdot P_{in} \tag{1.1}$$

where G is the linear power gain of the PA. An example of an ideal transfer characteristic is illustrated by the curve a in Figure 1.2. The curve b in Figure 1.2 shows the shape of a real transfer characteristic with three regions: a linear region, where the transfer characteristic

is close to the ideal, a saturation region where the output power is constant when the input power increases and an intermediate region; the compression region, where the output signal is compressed and not amplified linearly [Vuo03].

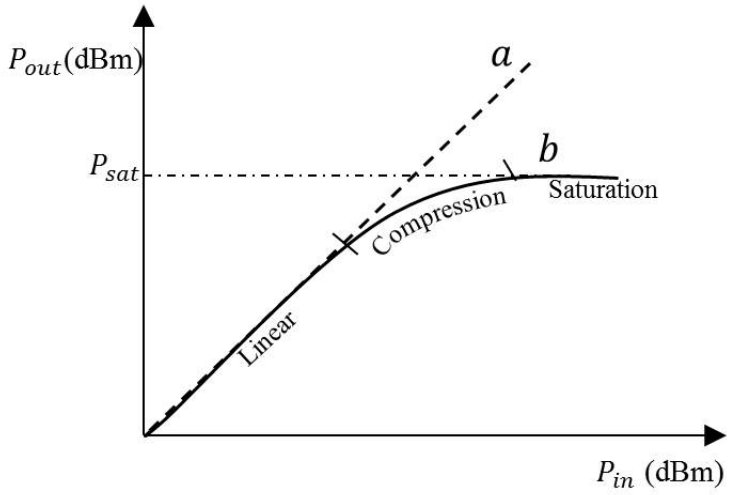

Figure 1.2: The PA (AM/AM) input-output characteristic: ideal (a) and real shape (b).

In this thesis, we are interested in the compression region as we will show that the PA get its maximum efficiency in this operation region. However, operating in this region means also producing non-linearities. In this chapter, we study the PA characteristics that illustrate its non-linearity, the effects of these non-linearities as well as the different predefined methods to deal with them. For this reason, this chapter is organized as follows:
In the second section, the PA characteristics are detailed as well as the so-called PA non-linearity effects. Further in the third section, we highlight the trade-off between efficiency and linearity. In the fourth section, we present some of the most commonly used efficiency enhancement techniques ([Cam15], [Hua15] and [Kim15]) and discuss their limits. Finally, the fifth section deals with the most famous PA linearization techniques in literature [Ken00] [Big88] [Kar89], as well as the PA modeling schemes.

1.2 PA characteristics

The main objective of this thesis is to linearize a wideband PA while enhancing its efficiency. For this reason, we are first interested in studying the PA characteristics such as the power gain, the efficiency and the linearity.

1.2.1 Gain

Gain is one of the most important characteristics of a PA. It is a measure of the amplification degree. There are three main definitions of gain: transducer gain (G_t), power gain (G_p) and available gain (G_a):

$$G_t = \frac{P_{out,d}}{P_{in,av}} \tag{1.2}$$

$$G_p = \frac{P_{out,d}}{P_{in,d}} \tag{1.3}$$

$$G_a = \frac{P_{out,av}}{P_{in,av}} \tag{1.4}$$

where $P_{in,av}$ and $P_{out,av}$ are the available power at the input and output respectively; $P_{in,d}$ and $P_{out,d}$ are the power delivered by the generator and to the load respectively (See Figure 1.3).

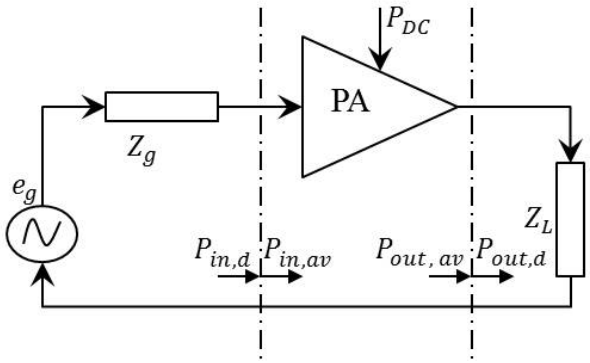

Figure 1.3: The PA power profile

The delivered power differs from the available power because of matching imperfections (See Figure 1.3). Assuming that the matching is perfectly realized. As a result,

$$P_{in,av} = P_{in,d} = P_{in}$$

$$P_{out,av} = P_{out,d} = P_{out}$$

The transducer gain is commonly used to characterize a PA. The shortened term "gain" is used with the intended meaning of "transducer gain". The gain of a PA is simply defined as the ratio of the output power to the input power:

$$G = 10\log\left(\frac{P_{out}}{P_{in}}\right) = P_{out}(dBm) - P_{in}(dBm) \tag{1.5}$$

where P_{in} and P_{out} are the input and output powers expressed in watts, whereas P_{in} (dBm) and P_{out} (dBm) are the input and output power in dBm. The gain is usually expressed in dB, whereas the power is in dBm.

1.2.2 Efficiency

The PA efficiency characterizes how much of the direct current (DC) power supply energy is converted into the RF signal. The efficiency of a PA should be as high as possible in order to make the PA consume less power. This is particularly important for battery-driven applications. There are two main definitions of efficiency [Cri06]:

- the DC-to-RF efficiency denoted by η,
- the power added efficiency (PAE).

The DC-to-RF efficiency (or drain efficiency) is the ratio of output power to the consumed power and is given as:

$$\eta = \frac{P_{out}}{P_{DC}} \tag{1.6}$$

The PAE is the ratio between the additional power provided by the PA (the difference between the output power and the input power) and the consumed power:

$$PAE = \frac{P_{out} - P_{in}}{P_{DC}} \tag{1.7}$$

If the gain is high, the PAE and drain efficiency are approximately equal. It can be shown that:

$$PAE = \eta(1 - \frac{1}{G_t}) \tag{1.8}$$

Figure 1.4 shows that the efficiency is maximized when the PA operates near the compression region. Consequently, if we aim at operating at the maximum efficiency, non-linear distortions may appear at the output. Therefore, the PA linearity is also an important criterion to be studied.

1.2.3 Linearity

In this section, the PA linearity metrics such as the 1-dB compression point (CP1) and the back-off are presented. Then, the non-linearity effects are discussed.

1.2.3.1 1-dB compression point

The CP1 is one of the major features to characterize the linearity of a PA. By definition, the 1-dB compression point of a PA is the output power level, at which the transfer characteristic of the PA deviates from the ideal linear transfer characteristic by 1 dB. In terms of gain, the CP1 is the PA power level, at which the gain decreases by 1 dB with respect to the gain in

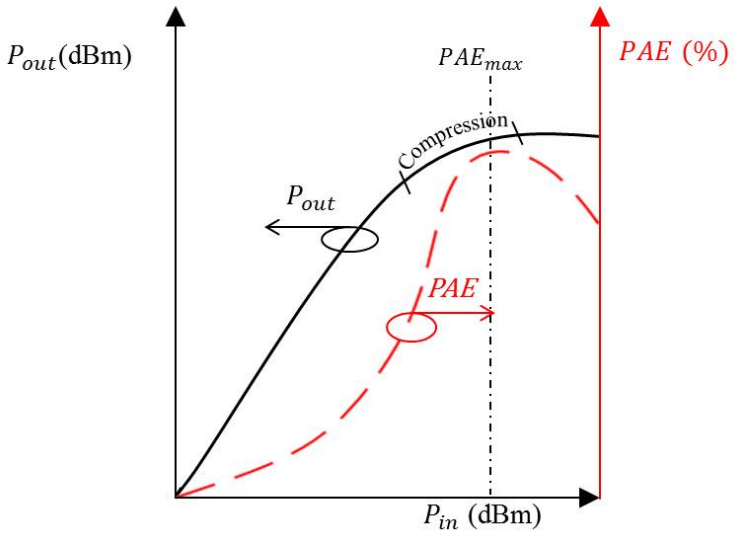

Figure 1.4: PA output power and PAE shapes.

the linear region. The above statements are illustrated in Figures 1.5 and 1.6 . The CP1 is shown on the gain and output power characteristics respectively. The higher the input 1-dB compression point (ICP1) is, the more linear the PA is. The output power and the gain at the 1-dB compression point are denoted by P_{1dB} and G_{1dB} respectively. For the gain, the relation between the real gain and the linear gain at the CP1 is as follows:

$$G_{1dB} = G_{Lin} - 1\text{dB} \tag{1.9}$$

where G_{Lin} is the PA gain in linear mode.

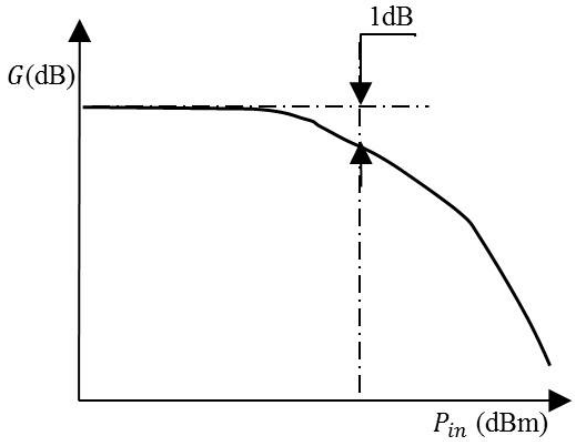

Figure 1.5: PA 1-dB compression point: gain characteristic.

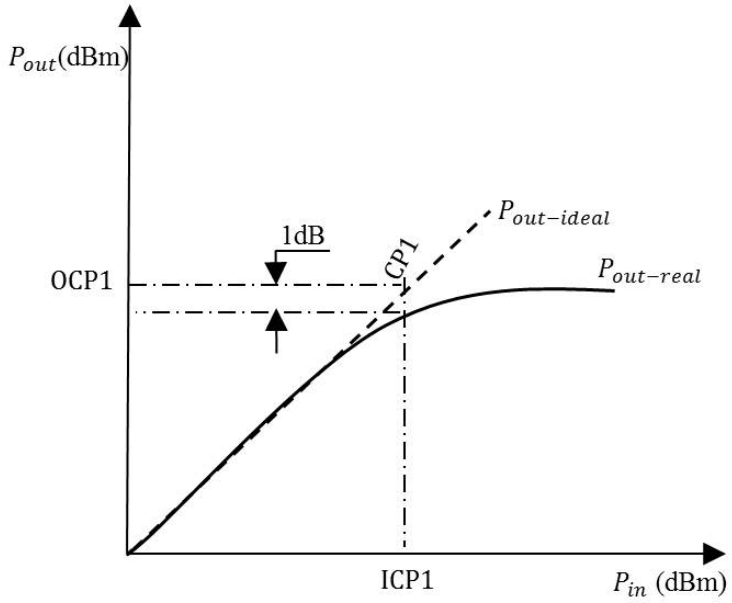

Figure 1.6: PA 1-dB compression point: P_{out} characteristic.

1.2.3.2 Back-off

For a PA, the back-off is a measure of how distant the output power is from the saturated output power P_{sat} (See Figure 1.7). There are two commonly used definitions of back-off: output back-off and peak back-off (PBO). The output back-off (OBO) is the ratio between saturated output power P_{sat} and the output power P_{out}. The OBO is usually expressed in dB, and can be written as:

$$OBO(\text{dB}) = 10log\left[\frac{P_{sat}}{P_{out}}\right] = P_{sat}(\text{dBm}) - P_{out}(\text{dBm}) \tag{1.10}$$

The PBO is the ratio between the saturated output power P_{sat} and the peak output power $P_{out_{max}}$. The PBO can be also expressed as the difference between the OBO and the peak-to-average power ratio (PAPR):

$$PBO(\text{dB}) = 10log\left[\frac{P_{sat}}{P_{out_{max}}}\right] = OBO(\text{dB}) - PAPR(\text{dB}) \tag{1.11}$$

where:

$$PAPR(\text{dB}) = 10log\left[\frac{P_{out_{max}}}{P_{out}}\right] \tag{1.12}$$

Figure 1.7 illustrates the relationships between the OBO, the PBO, the input back-off (IBO) and the PAPR curve shapes for PAs. The IBO is the ratio between the input power corresponding to $P_{out} = P_{sat}$ and the input power P_{in}.

The linearity of a conventional PA is related to its back-off: with an increase of the back-off,

the linearity of the PA increases and vice versa. Now let us focus on what happens when the PA is non-linear.

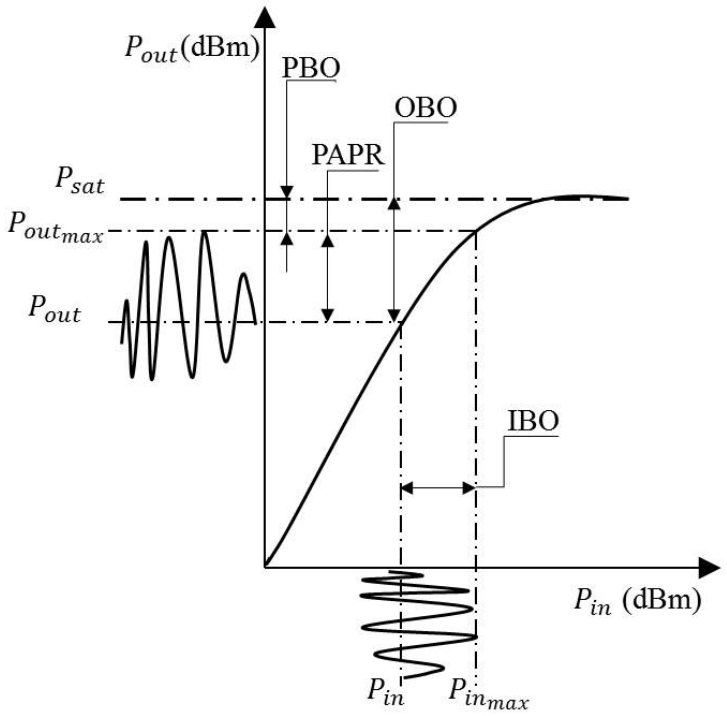

Figure 1.7: Relationships between the output back-off, the peak back-off and the peak-to-average power ratio for PAs.

1.2.3.3 Memory effects

The PAs distorted output depends not only on the instantaneous input, but on the previous inputs as well. This phenomenon is called memory effects. They depend on many conditions, such as the input signal magnitude, the carrier frequency, and the bandwidth. An interesting question can be: Why PAs exhibit memory effects?

Indeed, any power amplifier topology includes reactive elements like capacitances and inductances. Moreover, the PA active device also has parasitic capacitance in its P-N junctions. A capacitor is defined by the following differential equation:

$$i_C(t) = C\frac{dv_C(t)}{dt} \tag{1.13}$$

where, $v_C(t)$ is the voltage applied to the capacitor, $i_C(t)$ is the current flowing though the capacitor, and C is a constant called capacitance. If a capacitor is driven by a current source $i_C(t)$, the voltage between the capacitor nodes can be obtained as the integral of (1.13) from

$\tau = -\infty$ to $\tau = t$:

$$v_C(t) = \frac{1}{C}\int_{-\infty}^{t} i_C(\tau)d\tau \tag{1.14}$$

From (1.14) one can notice, that the output voltage depends not only on the instantaneous input signal, but on the entire past history.

If at some time t_0, $t_0 < t$, the value of voltage $v_C(t_0)$ is known, (1.14) can be re-written using the integral from t_0 to t:

$$v_C(t) = v_C(t_0) + \frac{1}{C}\int_{t_0}^{t} i_C(\tau)d\tau \tag{1.15}$$

As the initial condition $v_C(t_0)$ summarizes the history until the moment t_0, there is no need in specifying the entire history.

Similarly, an inductor is defined as:

$$v_L(t) = L\frac{di_L(t)}{dt} \tag{1.16}$$

Therefore, one has:

$$i_L(t) = i_L(t_0) + \frac{1}{L}\int_{t_0}^{t} v_L(\tau)d\tau \tag{1.17}$$

Equations (1.15)-(1.17) show that reactive elements are able to store energy and hence have memory. Therefore, power amplifiers exhibit memory due to the presence of reactive elements in their circuits. Another non-linearity effect, the intermodulation products, is studied in the next section.

1.2.3.4 Intermodulation products

Operating a PA under wideband signal conditions, when the amplifier operates in the compression region, causes a special kind of distortions in the output signal. They consist in additional frequencies at the output when an amplifier is excited with n-tone signal where n is greater than one. These additional frequency products are called intermodulation products.

Let a two tone sinusoidal signal to be applied to the PA input as follows:

$$u(t) = cos(2\pi f_1 t) + cos(2\pi f_2 t) \tag{1.18}$$

where $f_2 > f_1$. If a non-linear PA is considered, the output signal of the amplifier, $y(t)$, can be represented as follows:

$$y(t) = \alpha_0 + \alpha_1 u(t) + \alpha_2 u^2(t) + \alpha_3 u^3(t) + \cdots \tag{1.19}$$

The detailed expression of the PA output can be found by substituting (1.18) into (1.19). If we consider a non-linearity order equal to 3, it can be seen that when a two tone signal is

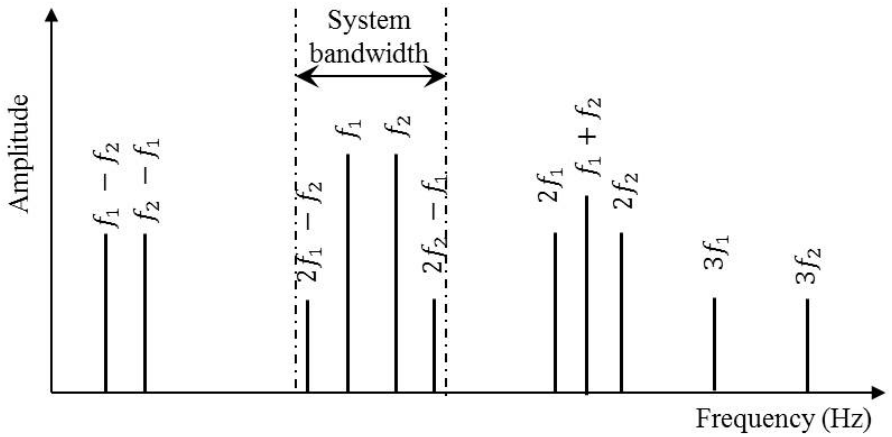

Figure 1.8: Spectrum of intermodulation products when a 2-tone signal (f_1, f_2) is applied to a PA with non-linearity order equal to 3.

applied to the power amplifier, the intermodulation products appear at $2f_1$, $2f_2$, $3f_1$, $3f_2$, $f_1 \pm f_2$, $2f_1 \pm f_2$, $f_1 \pm 2f_2$. See Figure 1.8.

$(2f_1, 2f_2)$ and $(3f_1, 3f_2)$ are the second and the third harmonics of the power amplifier respectively. They can be filtered out so the distortion caused by them can be minimized. The second order intermodulation products, $f_1 + f_2$ and $\pm(f_2 - f_1)$, can also be filtered out since they are far enough from the tones f_1 and f_2. However, the 3^{rd}-order intermodulation products, $2f_1 \pm f_2$ and $2f_2 \pm f_1$, fall within the amplifier bandwidth and cannot be filtered out. Thus, they can cause distortions at the output. These 3^{rd}-order intermodulation products are important because they are the parameters which limit the dynamic range and the bandwidth of the amplifier. A mathematical concept is defined as 3^{rd}-order intercept point in order to define the non-linearity with these parameters [Cri02].

As seen in Figure 1.9, the third-order intercept point (IP3) is the output power level (OIP3) at which the extended 3^{rd}-order harmonic slope meets that of the fundamental. At this output power, the fundamental and the 3^{rd}-order harmonic levels are equal even though operation at the IP3 is impossible since the output power usually saturates below this level. 1 dB decrease in input signal level results in 1dB decrease in fundamental tone level and 3dB decrease in all 3^{rd}-order product levels. This means if the input power is decreased by one-third of the distance in decibels from IP3 to noise floor, the third intermodulation drops to the noise level. This output power range for fundamental is called the spurious-free dynamic range (SFDR). It can be calculated by [Vuo03]:

$$SFDR(dB) = \frac{2}{3}[IP3 - P_{noise}] = \frac{2}{3}\left[IP3 + 174\text{dBm} - 10logBW(\text{dB}) - NF(\text{dB}) - G(\text{dB})\right] \tag{1.20}$$

where $IP3$ is the third-order intercept point power(dBm), BW is the PA bandwidth (Hz), NF is the PA noise figure (dB) and G is the PA gain.

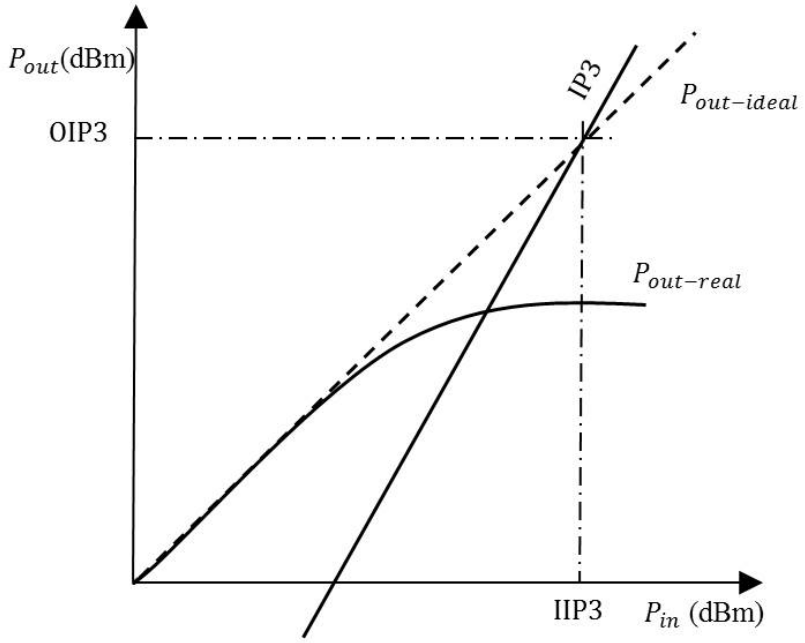

Figure 1.9: PA third-order intercept point.

This shows that the IP3 is a figure of merit for intermodulation suppression. If the IP3 power is high, then the SFDR is high which means undesired intermodulation products are suppressed. Thus, this is an important figure of merit which characterizes the linearity of a power amplifier.

1.2.3.5 Spectral regrowth and adjacent channel power ratio

The spectral regrowth is similar to the intermodulation distortion. Indeed, the intermodulation distortion is caused by the non-linearity of the amplifier when n-tone input is applied. The spectral regrowth can be observed when a modulated signal is given as input. Indeed, when a modulated signal passes through the non-linear amplifier, its bandwidth is broadened by the non-linearities (See Figure 1.10). This is caused by the mixing products between the individual frequency components of the spectrum.

The spectral regrowth leads to adjacent channel interference which is caused by the unwanted leakage of the adjacent channel. The ACPR is a commonly used figure of merit to describe linearity in modern telecommunication systems. ACPR is the ratio between the main channel power to the total adjacent channel power measured over the signal band [Vuo03]. ACPR can be calculated according to (1.21) [Vuo03]. The main and adjacent

channel powers are also shown in Figure 1.10.

$$ACPR = 10\log_{10}\left(\frac{\int_{f_c-W/2}^{f_c+W/2} S(f)df}{\int_{f_c-W_1/2+\Delta f}^{f_c+W_1/2+\Delta f} S(f)df}\right) \tag{1.21}$$

where $S(f)$ is the PA output power spectral density, f_c is the input signal carrier frequency, W is the input signal bandwidth and W_1 is the adjacent channel bandwidth. Therefore, the adjacent channel power ratio (ACPR) is kept low in order to avoid causing adjacent channel interference.

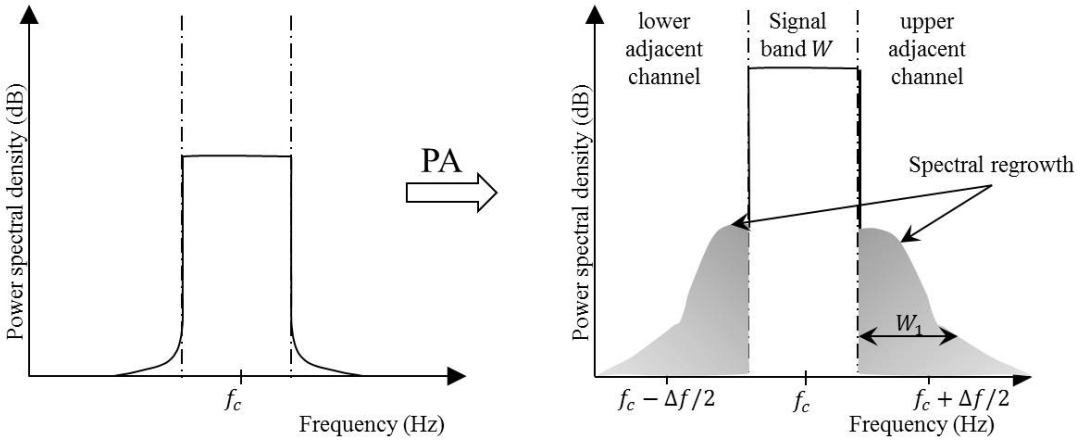

Figure 1.10: The spectral regrowth.

As it can be seen from the sections 1.2.3.3-1.2.3.5, the linearity is an important parameter that influences the performance of not only the PA but also the whole communication system. For this reason, the designer searches for building linear PAs. However, the efficiency of these PA classes is very low. Hence, another group of high-efficiency PA classes are proposed for applications where efficiency is required and linearity can be ignored. When a trade-off between them is required, the waveform engineering family is used.

1.2.4 Power amplification classes

Generally, PAs can be divided into three categories: linear amplifiers, switched amplifiers or waveform engineering amplifiers [Cri02]. With the linear classes, the signal waveform is more or less preserved however the efficiency is very low. When designing a switched PA, the theoretical efficiency may reach 100% but the signal envelop is strongly distorted. Consequently, when the efficiency increases, the linearity is deteriorated and vice versa. For this reason, the waveform engineering PAs correspond to a compromise between linearity and efficiency. More details about these classes are given in the remainder of this section.

1.2.4.1 Linear classes

The linear PA family includes four main classes: A, AB, B and C. These four types of PAs have similar circuit configuration, distinguished primarily by biasing conditions. See Figure 1.11. A class-A PA, in principle, works as a small-signal amplifier. It is probably the only "true" linear amplifier, since it amplifies over the entire input cycle such that the output is an exact scaled-up replica of the input without clipping. This "true" linearity is obtained at the expense of wasting power. To improve efficiency without sacrificing too much linearity, the concept of "reduced conduction angle" was proposed [Cri02]. The idea is to bias the active devices with low quiescent current and let the input RF signal to turn on active devices for part of the cycle. As the conduction angle shrinks, the amplifier is biased from class-AB, to class-B and class-C. Regardless of conduction angle, active devices are used as current sources. Therefore, they are often referred to as "transconductance" PAs.

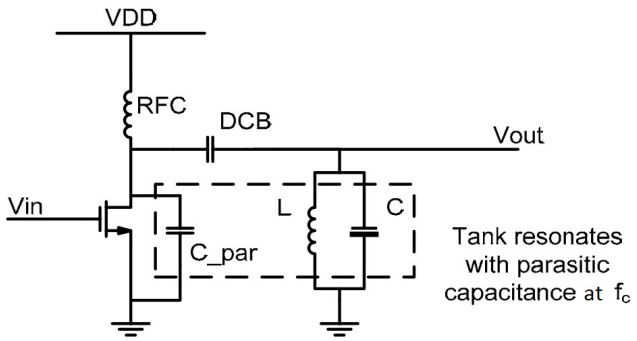

Figure 1.11: A generic topology for class-A, AB, B and C PAs.

where VDD is the DC supply, RFC is the drain output inductor, DCB is the drain output capacitor and C_par is a parallel capacitor.

1.2.4.2 Switched classes

The main classes within switched amplifiers are D and E.

A class-D amplifier is composed of a voltage controlled switch and a filtering tank. Figure 1.12 shows a voltage switching class-D amplifier. The output tuned network is tuned to the fundamental frequency. It will thus have negligible impedance at fundamental frequency and high impedance at harmonic frequencies. The analysis of such an amplifier is very straightforward due to the simple drain voltage waveform. In an ideal situation, the drain efficiency of a class-D amplifier reaches 100% as other switching type PAs.

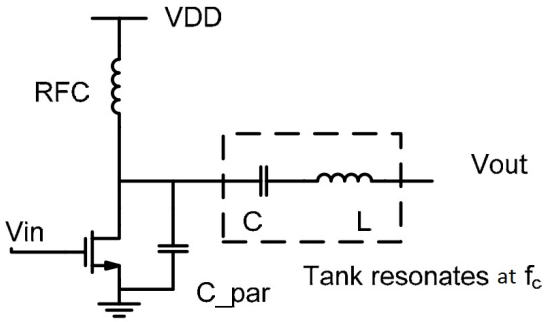

Figure 1.12: A voltage switching class-D amplifier.

Class-E PA stands out from other highly efficient switching PAs because parasitic capacitances of active devices may be absorbed into wave-shaping/matching networks. A simplest form of class-E PAs is shown in Figure 1.13. In this figure, jX is a drain output impedance. When the class-E PA is operating, the waveforms of drain current and voltage are shaped such that they do not overlap. Furthermore, the voltage will decrease gradually to zero before the active device turns on. This avoids charging/discharging capacitors at the drain, thus improve the efficiency.

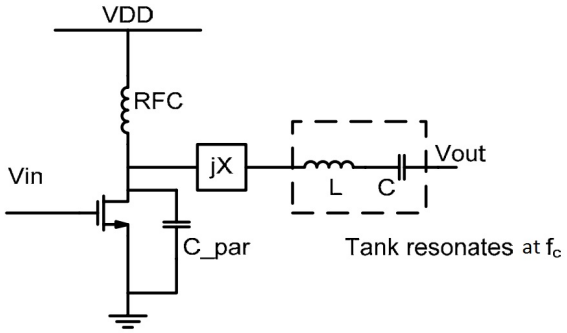

Figure 1.13: A simple class-E amplifier.

1.2.4.3 Waveform engineering classes

The main classes within waveform engineering are F and J.

Class-F is characterized by a load network that has resonance at one or more harmonic frequencies as well as at the fundamental frequency (Figure 1.14). To the author's knowledge, class-F was first proposed to improve the efficiency of overdriven transconductance amplifiers. Therefore, the active devices typically operate as a transconductor (or a current source) as those in transconductance amplifiers. However, if the input drive is large, active devices will

behave as switches just like those in switching amplifiers. In practice, with lumped elements, it is rare to see class-F amplifiers with tuned harmonics higher than the 5th harmonic, due to the complexity of a waveform shaping network.

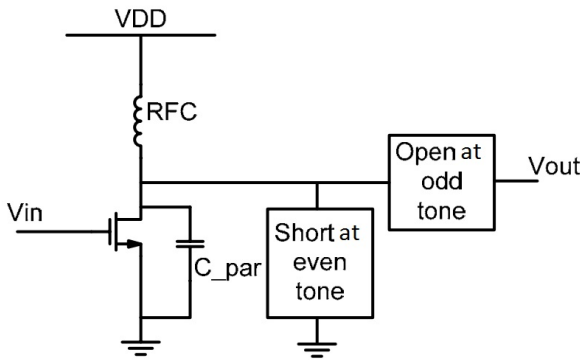

Figure 1.14: A class-F PA with tuned harmonics for waveform shaping.

The class-J amplification is characterized by a complex fundamental impedance combined with reactive harmonic terminations at the intrinsic drain of the device. It has been shown [Cri06] that such a configuration can deliver power and efficiency equivalent to the Class-B case. The consequence of adding a reactive fundamental component and reactive harmonic terminations manifests itself as an increase in the magnitude of the internal drain voltage waveform. Waveform engineering solutions also exist, which can exploit the non-linearities within the device to optimize performance [Moo10]. By only considering the 2^{nd} and 3^{rd} harmonics, the normalized intrinsic drain voltage waveform can be expressed as follows:

$$v_{ds} = 1 + acos(2\pi f_c t) + bsin(2\pi f_c t) + csin(2\pi 2 f_c t) + dsin(2\pi 3 f_c t) \tag{1.22}$$

According to the criteria outlined by Rhodes [Rho03], a family of solutions can be exploited by updating the fundamental reactive component for a set of harmonic terminations dynamically changing across frequency. In practice, difficulty arises in designing a matching network to avail of the improved performance with 3^{rd} harmonic enhancement across a wide bandwidth. Provided the normalized 2^{nd} harmonic reactance does not exceed a value of approximately 1.5 ($c = 0.5$), maximum efficiency can still be maintained and thus there would be no gain in control of the 3^{rd} harmonic. It is therefore desirable to provide minimal impedance across the 3^{rd} harmonic band otherwise it may further increase the amplitude of the drain voltage and drive the device into breakdown. The effect of presenting a low impedance 3^{rd} harmonic termination will introduce a 3^{rd} harmonic current component and distort the ideal half-sinusoidal current waveform.

Class-J theory make it possible the use the harmonic terminations to obtain high efficiency across the band of interest [Cri09]. Wright *et al.* [Wri09] demonstrated at least 60% efficiency over a 50% bandwidth while Wu *et al.* [Wu10] obtained more than 60% efficiency over a 42% bandwidth.

1.3 About the trade-off between linearity and efficiency

Modern wireless communication systems employing digital amplitude modulation schemes are experiencing continuous evolution towards high efficiency, broad bandwidth and high data rates. The PA is one of the key components in the RF front end. The PA characteristics and more particularly the efficiency directly affects the property of the whole system. High power efficiency leads to lower DC power consumption and heat dissipation, and thus has a significant impact on the overall power consumption, system stability and the requirement of DC supply and cooling system. Candidate systems used in the modern generation wireless communication systems as Long Term Evolution (LTE), employ a wide-bandwidth modulated signal, which implies that the design of broadband PA is more important in such systems.

Meanwhile, when the PA operates at its maximum efficiency, the amplification procedure is no longer linear and the signal envelop is distorted. One solution is to use the back-off technique. Once the output power is reduced from its maximum value, both the amount of amplitude conversions and distortion products is reduced. Unfortunately, the back-off reduces efficiency, making it an unattractive linearization method for amplifiers.

To summarize, when the PA efficiency is enhanced, the PA linearity is deteriorated and vice versa. Here, we address the trade-off between linearity and efficiency. To improve the trade-off linearity-efficiency, there are two ways:

- The PA can be designed to be linear at the expense of efficiency. The required efficiency level is then reached by means of an efficiency enhancement technique,
- The PA can be designed to reach higher efficiency at the expense of linearity. The linearity specification is then met by means of external linearization techniques.

In the literature, there are several efficiency enhancement and linearization techniques. For each category, there are a RF family and a digital family. In this thesis, we choose to detail and discuss the most commonly used efficiency boosting and linearization techniques

in order to emphasize our choices. The next sections, deal with the efficiency enhancement techniques:

- The RF family, such as the envelop tracking, the adaptive bias and the Doherty techniques,
- The digital family: the peak to average power ratio (PAPR) reduction techniques.

Then, we present the linearization techniques.

1.4 Enhancing the PA efficiency

In the efficiency enhancement techniques, there are two families: the RF family and the digital family. In the RF family, we choose to present the Doherty, the adaptive bias and the so-called envelop tracking techniques. These techniques have been the focus of many recent research works [Cam15], [Hua15] and [Kim15]. In the digital family, the PAPR reduction techniques are discussed.

1.4.1 RF efficiency enhancement techniques

This section describes the most commonly used efficiency enhancement techniques in the literature. These techniques aims at enhancing the efficiency at low envelope power, thus increasing the average efficiency.

1.4.1.1 Doherty amplifier

The simplest configuration of a Doherty circuit consists of two amplifiers; “main” or “carrier” amplifier and the “auxiliary” or “peaking” amplifier. The amplifiers are connected in parallel with a quarter wave transmission line (impedance inverter), as shown in Figure 1.15.

The basic concept of the Doherty amplifier [Cam15] is to allow the main amplifier to operate at the maximum efficiency (peak power) while allowing the auxiliary amplifier to deal with the modulation peaks. When the input drive is low the auxiliary amplifier is shut-down and the main amplifier operates in the linear mode. For example, if a class-B PA is used as the main amplifier and the class-C as auxiliary amplifier, the class-C is off because signal is too small. As the input power increases, the main amplifier starts to saturate and the auxiliary amplifier starts to supply current. This turn-on point of the auxiliary amplifier is called the transition point. At the transition point, the efficiency of the overall system becomes high.

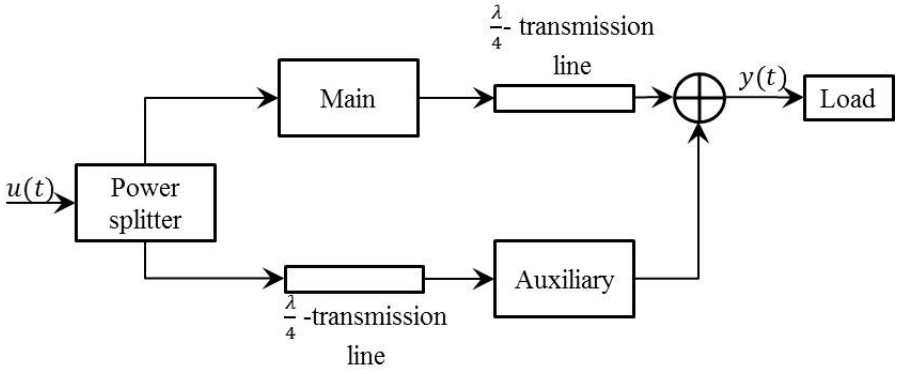

Figure 1.15: Schematic of Doherty Amplifier.

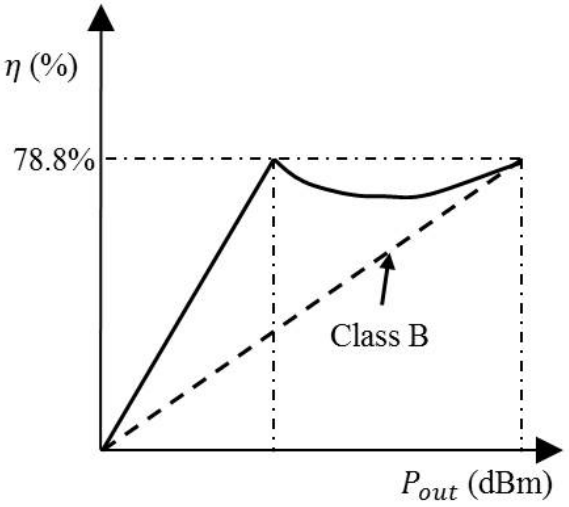

Figure 1.16: Efficiency versus the output power of Doherty amplifier (continuous line) and class-B amplifier (dashed line).

The overall efficiency characteristic of the Doherty amplifier as a function of output power is shown in Figure 1.16. If a class-B PA is used as the main amplifier and the class-C as auxiliary amplifier, the theoretical maximum efficiency is 78.8% [Cam15]. The small dip in the efficiency in between the transition point and full power is due to the lower efficiency of the auxiliary amplifier which operates in large back-off.

The performance of the Doherty topology mainly depends on the class of operation of the two amplifier blocks. Using the Doherty technique, the efficiency at back-off can be kept close to that in saturated condition.

A Doherty amplifier is narrow band because of the use of quarter-wave transmission lines and the requirements of accurate phase matching between the two paths[1] [Cri02]. Another

[1]The dimensions of the quarter-wave transmission lines are conditioned by the signal carrier frequency

drawback is the poor IMD performance mainly due to the low biasing of the auxiliary amplifier. However, other linearization scheme can be implemented to improve the linearity in the Doherty amplifier but this will add the complexity [Cam15].

1.4.1.2 Adaptive bias

The adaptive bias technique [Hua15] was primarily proposed to increase the power added efficiency of class-A amplifiers when large back-off is used. In this technique, the bias level of the amplifier is varied with the envelop level to reduce the amount of DC current drawn from the supply during back-off. When the amount of DC current is reduced, the PA consumed power decreases and its efficiency increases. The operation of the adaptive bias scheme is shown in Figure 1.17. The gate bias of the PA is varied proportionally to the signal from the envelope detector.

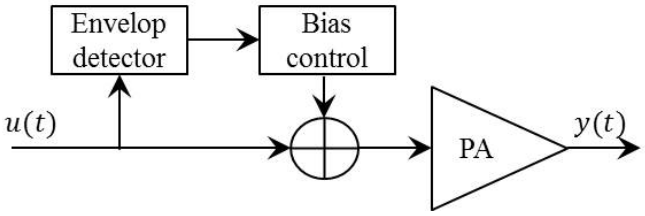

Figure 1.17: Schematic of the adaptive bias control.

1.4.1.3 Envelop tracking

In the Envelop tracking (ET) [Kim15], the RF drive contains both the amplitude and the phase information and the burden of linearity lies on the RF PA itself. This technique is also called "dynamic drain/collector biasing". As shown in Figure 1.18, the envelope detector extracts the envelope information from the RF input signal and uses it to control the collector/drain voltage of the PA through a supply modulator. The supply modulator dynamically adjusts the RF PA with just sufficient supply voltage to reduce the DC power consumption at low power levels, thus increasing the average efficiency of the whole system. Other advantage is that the linearity of the system does not depend on the bandwidth of the supply modulator.

Some drawbacks of ET techniques are:

- The theoretical average efficiency is low because of the use of a linear mode PA.

and if it changes,= we need to change the whole design.

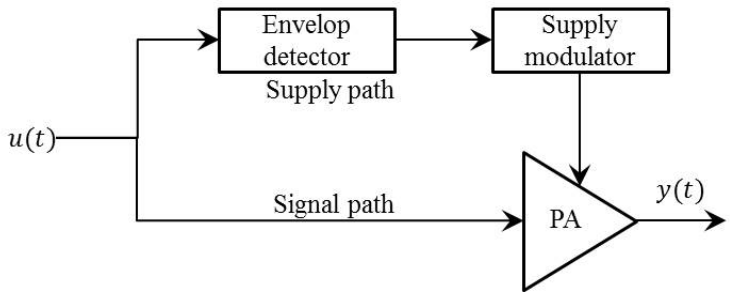

Figure 1.18: Envelope tracking architecture.

- The mismatch of the delays between the two paths (supply and signal paths) degrades the linearity,
- The power gain of the PA decreases as the supply voltage is reduced.

1.4.2 About the PAPR reduction techniques

The PAPR requires to back-off the PA and thus reduces the power efficiency of the transmitter. Waveforms that exhibit the highest spectrum efficiency such as OFDM tend to have high PAPR. A PAPR ranging from 5 to 8 dB is a goal to achieve reasonable power efficiency of the linearized RF stage [Las13], but multicarriers downlink waveforms could exhibit PAPR greater than 12 dB [Kow11], [Hus08]. One can roughly divide techniques used to reduce the PAPR into two categories:

- The signal phase modulation category,
- The signal clipping category.

The first one does not degrade the transmitted signal but requires side information. In this category, one can quote selective mapping and partial transmit sequence [Bax06]. On the other hand approaches like hardware or software clipping [Las13], and optimization of the subcarriers power [Nad11] do not require side information but may severely degrade the error vector magnitude (EVM) and the ACPR. Software clipping or subcarriers power tuning reduce the PAPR while minimizing the regrowth of both the EVM and the ACPR. It has been highlighted that software clipping of the magnitude instead of the in-phase and quadrature components allows better PAPR reduction for a given EVM budget [Las13]. In [Weg06], the author introduces another method by applying a spectrally shaped "pulse" to cancel signal peaks. The spectral shape depends on the modulation waveforms and could accommodate multicarriers and multi-standards transmitted waveforms.

1.4.3 A summary about efficiency enhancement techniques

Three RF efficiency enhancement techniques were summarized. In the Doherty technique, the signal separation and the recombination after amplification are a difficult task. In ET and adaptive bias, the burden of linearity lies on the final PA itself. So the final PA must be a linear PA and hence the average efficiency is not as high as the Doherty technique.

Another circuit level alternative can be investigated to enhance the PA efficiency, the class-J PAs, which provides a solution for the non-linear effects of switch-mode amplification. Therefore, they theoretically exhibit high efficiency, linearity and wideband behavior simultaneously. However, the class-J theory is based on high quality passive components (inductances and capacitances) i.e. low loss and high stability. This kind of passive components can be found in gallium (Ga)-based technologies which are high cost technologies. In the MOS technologies, the passive components have lower quality. Nevertheless, in some research works, class-J PAs are designed using LDMOS [Par14] and BiCMOS [Sar14] technologies. Unfortunately, the reached linearity and efficiency performance are considerably lower than those of the Ga-based technologies. This may lead another time to the need of using an efficiency enhancement technique.

When many constraints about the PA chip size and cost, and power consumption are required as it is the case of mobile handsets, a digital efficiency enhancement technique may be useful. A PAPR reduction technique can be used to improve the PA efficiency without increasing the PA chip size and eventually it does not depend on the used PA technology. However, these techniques cause signal distortions that have to be compensated. As a result, a linearization technique is needed.

1.5 Linearization techniques

Because of the stringent requirement on linearity and the desire to increase battery time for mobile terminals, several linearization techniques have been developed. [Ken00] has quite extensive coverage on this topic. Here, some techniques are discussed, from RF to digital techniques.

1.5.1 RF linearization techniques

In this section, we present the linearization techniques which has been the focus of the recent research works. Among the RF techniques, the feedback and feedforward techniques are detailed. These techniques are the most commonly used techniques.

1.5.1.1 Feedback

The basic structure of feedback circuit is shown in Figure 1.19.

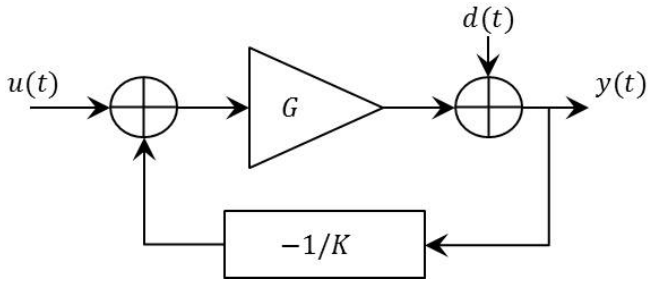

Figure 1.19: General feedback circuit for a PA.

In this structure, the input signal is $u(t)$, the gain of the PA is G, the gain of the feedback loop is $-1/K$, and the distortion is $d(t)$ which is added after the gain of the PA. The output can be obtained directly as:

$$y(t) = \frac{GK}{K+G}x(t) + \frac{K}{K+G}d(t) \tag{1.23}$$

If we assume that the amplifier gain is much greater than the feedback loop gain, i.e., $G >> K$, (1.23) can be simplified to:

$$y(t) = Kx(t) + \frac{K}{G}d(t) \tag{1.24}$$

From (1.24), we can see that the gain of the signal is lowered from G down to K, and the distortion will be significantly reduced by K/G.

1.5.1.2 Feedforward

Figure 1.20 shows a feedforward linearization scheme.

In the lower branch of the circuit, a sample of the input is subtracted from a sample of output of the main amplifier, to generate an error signal, or intermodulation products in the spectral domain. This error signal is amplified though an error amplifier, to have the same amplitude as the output error of the main amplifier. A time delay line is inserted between the two couplers in the upper branch, which make the errors from the two branches have 180 degree phase difference. The errors cancel each other in the last coupler, making the output linear again.

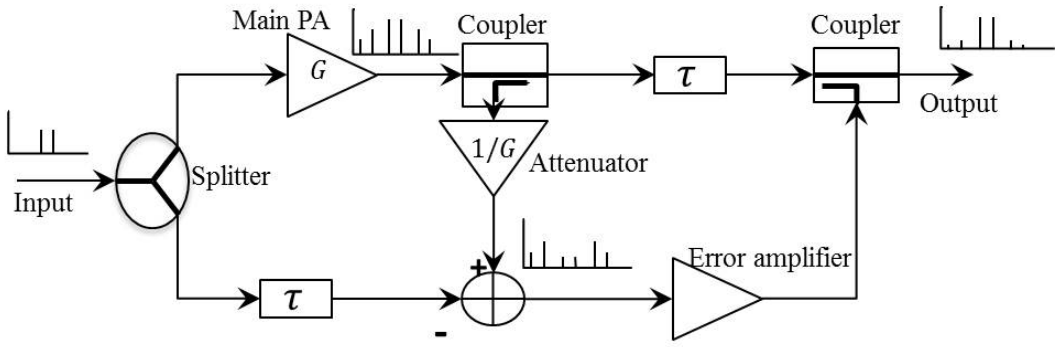

Figure 1.20: Feedforward linearization scheme.

1.5.1.3 Limitation of these techniques

In the feedback predistortion, the output signal goes back to the subtractor through the feedback loop. This will take a certain time. When considering this delay of the feedback loop, the overall equation is:

$$y(t) = Gx(t) - \frac{G}{K}y(t - \Delta t) + d(t) \tag{1.25}$$

where Δt denotes the delay of the feedback loop. It is only when $y(t)$ is equal or near to $y(t - \Delta t)$, that (1.25) can be equal to (1.23). In RF field, a small time delay can cause a great phase shift. Hence, the difference between $y(t)$ and $y(t - \Delta t)$ can be significant in a RF transmitter. On the other hand, the feedforward technique is limited in terms of power efficiency. The lower branch amplifier consumes a certain power. However, this output does not make a positive contribution, but a subtraction from the output of main amplifier. From the power point of view, the error amplifier is making additional consumption. Practically, this linearization technique has 20% of power efficiency at best. This compares poorly with other linearization schemes such as digital pre-distortion, where efficiencies greater than 50% can be achieved.

1.5.2 Digital techniques

The demand for higher flexibility and lower cost with similar performance as analog linearization schemes leads to the concept of digital linearization techniques. Signal processing techniques, which can be efficiently implemented using digital hardware such as Digital Signal Processors (DSPs) and/or Field Programmable Gate Arrays (FPGAs), are used to control an analog RF-system. The advantage is that a high degree of flexibility is maintained due to the inherent flexibility of the digital hardware which allows for changes at run-time of

the system. This is in line with the current trend to Software Defined Radio (SDR), where the ultimate goal is to define highly reconfigurable radios which can accommodate a variety of standards and transmission/receive modes, controlled entirely by software. This is only possible if the inflexible and costly analog circuitry is reduced to a minimum by replacing as much as possible by reprogrammable digital hardware.

The following sections discuss common digital linearization schemes covered in the literature. These digital techniques rely on a behavioral modeling of the PA to compensate its distortions. For this reason, we first introduce the so-called digital pre-distortion and post-distortion techniques. Then, an overview about the defined PA behavioral models is given.

1.5.2.1 Digital pre-distortion

The Digital pre-distortion of RF PAs is initiated in the early 1980s with the paper of A. M. Saleh and J. Salz [Sal83]. This and other early contributions consider data pre-distortion, i.e., the data symbols are distorted. The pulse-shaping is thus performed after the pre-distortion stage. The spectral broadening due to the non-linear amplifier cannot be avoided, but the non-linear distortion of the data is compensated. These contributions [Sal83] consider non-linear memoryless PAs.

Data pre-distortion considering also memory effects appear in the late 1980s [Big88, Kar89], using Volterra filters as models for the non-linear channel. The idea of the pre-distortion is to apply a complementary non-linearity of the PA so that the cascade of the pre-distorter and the amplifier gives a linear response.

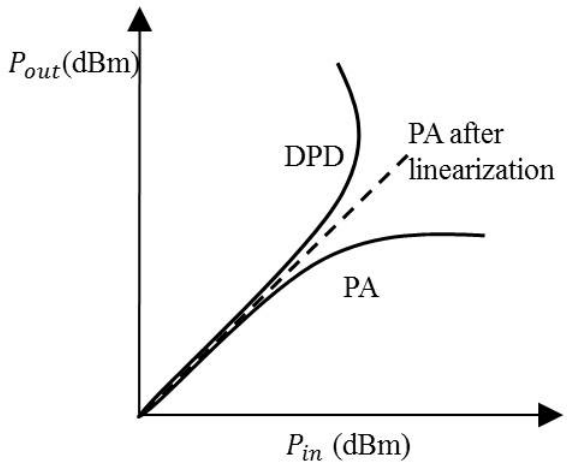

Figure 1.21: Digital pre-distortion principle.

Besides, the pre-distortion is also a behavioral modeling problem itself since it is necessary to get the reverse function of the amplifier. Figure 1.22 shows a digital pre-distortion system. Analog-to-digital converters (ADCs) are needed for converting the demodulated output signal of the transmit path, containing the non-linear PA, to digital. The scheme presented in Figure 1.22 is a signal pre-distortion scheme, where the transmit-signal, immediately before the upconversion, is pre-distorted. Since the ideal small signal gain of the

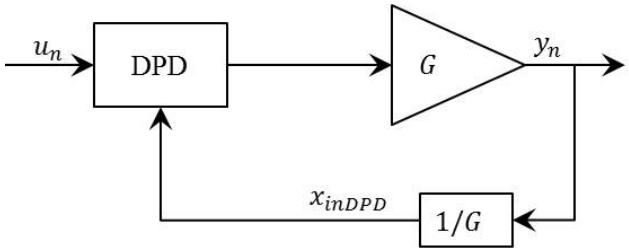

Figure 1.22: Digital pre-distortion linearization scheme.

amplifier is constant G, the cascade of the pre-distorter and the non-linear PA gives the ideal small signal gain G. Therefore, it can be stated that the input of the pre-distorter is normalized with G as follows:

$$x_{inDPD}(n) = \frac{y_n}{G} \tag{1.26}$$

where G is small signal gain, x_{out} is output of the amplifier. After the predistorter, the output is given in equation (1.26) as;

$$f_{DPD}\left(\frac{y_n}{G}\right) = u_n \tag{1.27}$$

The output of the predistorter is input to the amplifier as in equation (1.29);

$$f_{PA}(u_n) = y_n \tag{1.28}$$

$$f_{PA}\left(f_{DPD}(y_n/G)\right) = y_n \tag{1.29}$$

Here, the function of the pre-distorter f_{DPD} is equivalent to the behavioral modeling of the PA reverse function obtained by swapping the PA input and output signals with appropriate small signal gain normalization [Gha09]. After the cascade system of pre-distorter and the amplifier, the input has a linear relation with the output and the power does not saturate anymore at the level it used to (Figure 1.21).

Pre-distortion systems rely on exact inverse replication of the non-linearity of the PA, which means high sensitivity not only to memory effects but to drifting as well. Typically, some form of slow adaptation is needed for the pre-distorter. Digital pre-distorter is naturally

more flexible, but it requires a high bandwidth and dynamic range from the digital baseband, and also all baseband and intermediate frequency filters between the pre-distorter and the PA contribute to the memory effects. In addition, cascading two non-linear devices leaves a residue of high-order non-linear products that were absent in the original PA response. A pre-distorter has not only to linearize the target PA, but also it has to compensate its own distortions. This aspect of pre-distortion remains an underrated problem, and has all too often been swept under the carpet by researchers who use carefully chosen spectrum sweep ranges to display their results. There are two different approaches for the identification of the pre-inverse of the PA or pre-distortion. One of the main difficulties in extracting the behavioral model of the pre-distortion comes from the fact that the internal signal interconnecting the pre-distortion and the PA is not known *a priori*. So in order to identify the pre-distortion we are restricted by an intermediate step, in which the PA is first identified, and its model is then used for the extraction of pre-distortion coefficients. This identification approach is based on a direct learning architecture (DLA) illustrated in Figure 1.23. On the other hand, the post-inverse of the PA can be more easily identified since the input (PA output) and the output (linearly amplified version of PA input signal) signals are known. Thus, a second approach, based on an indirect learning architecture (ILA), has been proposed with the assumption that the post-inverse can be used as a pre-inverse. This learning architecture is illustrated in Figure 1.24. In this case, a post-inverse model is directly identified and used as a pre-distortion.

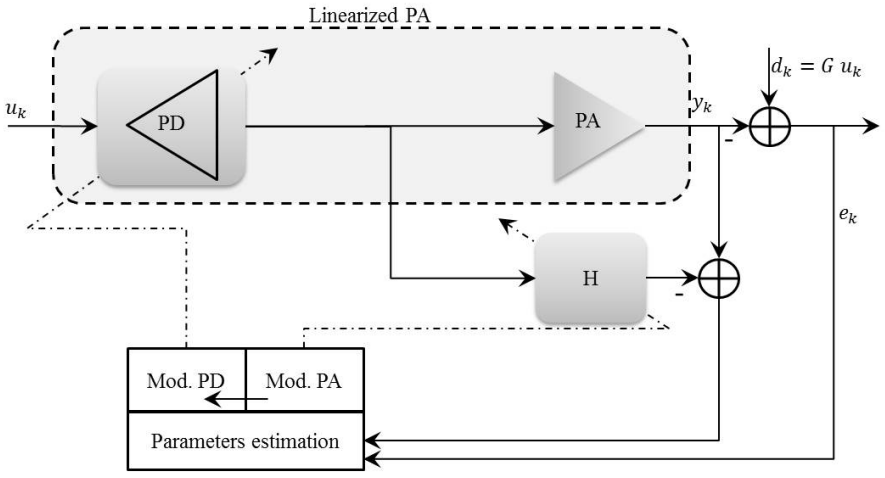

Figure 1.23: Direct Learning Architecture: DLA.

Identification algorithms differ upon the learning architecture used. In ILA, when the pre-distortion model can be written as a linear combination of a set of basis functions, the least squares (LS) technique can be used [Mor14], [Teh10]. In this case the matrix to be

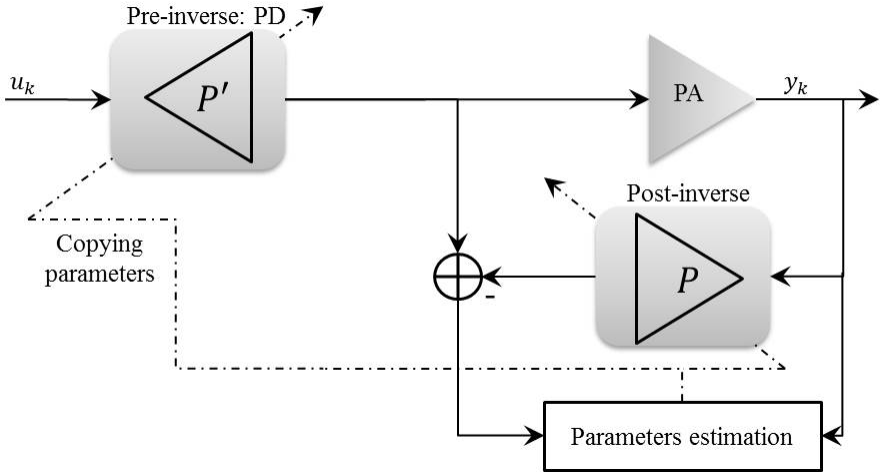

Figure 1.24: Indirect Learning Architecture: ILA.

inverted may have large condition number and the solutions over system level iterations tend to be unstable. One of the solutions to this problem is the use of orthogonal polynomials as in [Rai13]. In [Jia12], a stochastic conjugate gradient algorithm has been proposed which can smoothly estimate pre-distortion parameters over iterations and optimize the basis functions used. In [Gua12] the 1-bit ridge regression algorithm has been used to eliminate the ill-conditioning problem and a method to reduce complexity by reducing the number of samples used without affecting the condition number. For other model structures, i.e., when the output depends non-linearly on the coefficients, sophisticated algorithms must be derived (e.g., [Gha10]).

Many algorithms have been also used for DLA architecture. The main advantage of DLA is its robustness in presence of noise [Abi12]. An algorithm based on an analytical method to compute the output of the predistorter using the extracted memory polynomial model of the power amplifier has been proposed in [4]. This method is highly sensitive to noise, presenting high instability over system level iterations [Abi12]. Unlike the analytical method, non-linear filter architectures have been proposed in [Lim00], [Zho07] in which an adaptive algorithm is used to identify a model for pre-distortion from the extracted parameters of PA. The algorithms proposed are the Non-linear Filtered-x least mean squares (NFxLMS) [Lim00] and recursive least squares (NFxRLS) [Zho07] algorithms. The latter has good convergence properties but the overall computational complexity is relatively high. This limits the use of DLA over ILA since the latter offers often good compromise between complexity and performance.

1.5.2.2 Digital post-distortion

The post-distortion approach to amplifier linearization is similar to the pre-distortion, with the obvious exception that the linearizing element must be able of handling the full power capability of the PA output stage. It is therefore inherently less desirable as a linearization technique due to the restriction this places on the range of available non-linear elements which may be used in the post-distorter. In addition, the inevitable losses in this block have a significant effect on the overall efficiency of the amplifier system. Indeed, in many cases it may be more efficient to provide linearization by back-off rather than to use post-distortion.

An alternative, and rather more interesting, approach is to place the post-distorting element in the receiver rather than in the transmitter. The signal levels will then be significantly lower and the losses (other than if it is placed in the antenna path) much more reasonable. It also has the advantage in a system with infrastructure, a base station for example, that the complexity of the mobile terminal is reduced, at the expense of the base station receiver. Complexity at the base station is usually much more acceptable due to the increased size and cost which is tolerable in that part of the system. The approach proposed to allow adjustment of the post-distorters at the base station involves measuring the level of distortion present in a vacant channel and adjusting the parameters of the post distorters in the two adjacent channels to eliminate the distortion present in the non-used channel. When this vacant channel is in use, it should then, theoretically, enjoy almost interference-free reception. There are a number of inherent disadvantages with this approach, however, and these will severely limit the available performance. First, the degree of IMD reduction which can be achieved will be small, largely due to the lack of knowledge of the original signal in the transmitter which would be required for an adaptation. A linearization scheme would therefore still be required in the mobile transmitter, although its performance need not meet the full required mask. This largely removes the complexity reduction advantage in the mobile terminal.

The system also relies on vacant channels being available in order to adapt its post-distorters. This may be a problem in a heavily-loaded system and the non-regular frequency allocations of many systems may make the use of this technique difficult.

Finally, the requirement for an adaptive post-distorter for each channel in the base station will result in significantly increased complexity in that part of the system. This is only worthwhile if the cost/size/power consumption of the mobile terminal can be significantly reduced.

1.5.2.3 PA behavioral modeling

When dealing with digital pre/post-distortion techniques, it is also necessary to model the PA behavior in order to be able to build the distorter element.

In narrow band applications, the PA presents negligible memory effects and memoryless models are sufficient to compensate for the non-linearity. In this case, the static complex gain function f_{DPD} of the pre-distortion can be estimated using an equivalent baseband polynomial model. f_{DPD} can be also directly estimated for a finite number of amplitudes on the whole dynamic range of the input signal and implemented in a look-up table, see [Pre12].

When the input signal bandwidth becomes wider the time span of the PA memory becomes comparable to the time variations of the input signal envelope. Thus, the memory effects of the PA can no longer be considered as short-term. Simplified structures of the general Volterra model have been predominantly used in this case. In general, the structure of the model, either for the PA or for the pre-distortion, is defined empirically, i.e., depending on the PA used, the validity of a predefined structure is determined by an experimental evaluation of various criteria. For PA modeling evaluation many criteria have been defined in literature [Isa06], [Lan08], and they can be also used for pre-distortion modeling evaluation in addition to linearity metrics, EVM and ACPR. Two families of non-linear models with memory are considered and are based on neural networks (NNs) or non-linear autoregressive with exogenous input (NARX) models [Leo85]. In the latter family, one particular case is the polynomial NARX model. It has the advantage of being linear regarding the parameters and non-linear regarding the input and the previous output samples. When the output only depends on the input, this leads to the Volterra models [Sig12]. Among them, the memory polynomial (MP) model [Kim01] and the generalized memory polynomial (GMP) model [Mor14] have good modeling performance when compared to other model structures [Isa06], [Teh10]. Hammerstein (H), Wiener (W), Hammerstein-Wiener (HW) and Wiener-Hammerstein (WH) belong also to this category of models and have been evaluated for PA and PD modeling [Tar15], [Che14b]. Parallel structures like the parallel Hammerstein and Parallel HW models have been also used recently [Sch14]-[Gha10]. Such models are particularly interesting for the linearization of highly non-linear PAs.These models represent simplified versions of the Volterra model.

In the discrete-time domain, the full baseband Volterra series are:

$$y_n = \sum_k \frac{2k+1!}{(k+1)!k!2^{2k}} \sum_{\tau_1} \cdots \sum_{\tau_{2k+1}} h_{2k+1}(\tau_1 \cdots \tau_{2k+1}) \prod_{i=1}^{k+1} u_{n-\tau_i} \prod_{s=k+2}^{2k+1} u^*_{n-\tau_s} \tag{1.30}$$

From (1.30), it can be seen that the number of coefficients of the Volterra series increases

exponentially as the memory length and the non-linear order increase. This drawback makes the Volterra series unattractive for real-time applications. This prompts us to consider several special cases of the Volterra series. The special cases considered here include the Wiener model, the Hammerstein model, the Wiener-Hammerstein model and the memory polynomial. The Wiener model is a linear time-invariant (LTI) system followed by a memoryless non-linearity (NL) (see Figure 1.25). The both subsystems are given by:

$$u_n^L = \sum_{\tau=0}^{T} a_\tau u_{n-\tau} \tag{1.31}$$

$$y_n = \sum_{k=1}^{K} b_k u_n^L |u_n^L|^{k-1} \tag{1.32}$$

where a_τ denote the impulse response of the LTI block, b_k are the coefficients of the odd-order polynomial describing the memoryless non-linearity and K and T are respectively the non-linearity order and the memory depth. Substituting (1.31) into (1.32) gives:

$$y_n = \sum_{k=1}^{K} b_k \left[\sum_{\tau=0}^{T} a_\tau u_{n-\tau}\right] \left|\sum_{\tau=0}^{T} a_\tau u_{n-\tau}\right|^{k-1} \tag{1.33}$$

The Wiener model was used by Clark *et al.* [Cla98] to model the PA with memory effects, where improvements in modeling accuracy were observed when the Wiener model replaces the memoryless polynomial model.

The Hammerstein model is a memoryless non-linearity followed by an LTI system (see Figure 1.26). The both subsystems in this model are described by:

$$u_n^N = \sum_{k=1}^{K} f_k u_n |u_n|^{k-1} \tag{1.34}$$

$$y_n = \sum_{\tau=0}^{T} g_\tau u_{n-\tau}^N \tag{1.35}$$

where f_k are the coefficients for the memoryless non-linearity and g_τ are the impulse response values of the LTI system. Substitution of (1.34) into (1.35) leads to

$$y_n = \sum_{\tau=0}^{T} g_\tau \sum_{k=1}^{K} f_k u_{n-\tau} |u_{n-\tau}|^{k-1} \tag{1.36}$$

The Wiener-Hammerstein (W-H) model (see Figure 1.27) is an LTI system followed by a memoryless non-linearity, which in turn is followed by another LTI system. This configuration is commonly used for satellite communication channels, where the PA at the satellite transponder is driven near saturation to exploit the maximum power efficiency for the down-

link [Ben83]. The subsystems in this model are described by

$$u_n^L = \sum_{\tau=0}^{T_1} a_\tau u_{n-\tau} \tag{1.37}$$

$$u_n^{LN} = \sum_{k=1}^{K} b_k u_n^L |u_n^L|^{k-1} \tag{1.38}$$

$$y_n = \sum_{\tau=0}^{T_2} c_\tau u_{n-\tau}^{LN} \tag{1.39}$$

where a_τ and c_τ denote, respectively, the impulse response values of the LTI systems before and after the memoryless non-linear block, and b_k are the coefficients of the non-linear block. Combining (1.37), (1.38), and (1.39), we infer:

$$y_n = \sum_{\tau_1=0}^{T_1} c_{\tau_1} \sum_{k=1}^{K} b_k \left[\sum_{\tau_2=0}^{T_2} a_{\tau_2} u_{n-\tau_1-\tau_2} \right] \left| \sum_{\tau_2=0}^{T_2} a_{\tau_2} u_{n-\tau_1-\tau_2} \right|^{k-1} \tag{1.40}$$

The memory polynomial model uses the diagonal kernels of the Volterra series and can be viewed as a generalization of the Hammerstein model. In the discrete-time Volterra series (1.30), if $\tau_1 = \cdots = \tau_{2k+1} = \tau$, (1.30) becomes:

$$y_n = \sum_{k=1}^{K} \sum_{\tau=0}^{T-1} b_{\tau k} z_{n-\tau} |z_{n-\tau}|^{k-1} \tag{1.41}$$

where $b_{\tau k} = h_{2k+1}(\tau, \cdots, \tau)$ in (1.30). This model was considered for modeling PAs with memory effects in [Kim01] and for data pre-distortion of the cascade of a pulse shaping filter and a memoryless PA in [Cha01].

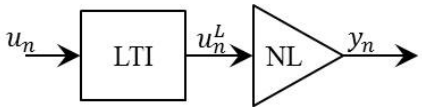

Figure 1.25: Block scheme representation of the Wiener model.

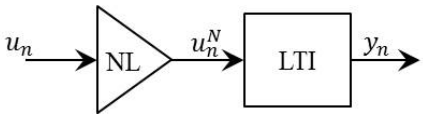

Figure 1.26: Block scheme representation of the Hammerstein model.

In order to compare their performance, the linearity parameters are used such as the ACPR which is used in the frequency domain and the EVM in the time domain. Furthermore, the complexity of the models is an important parameter to be compared since the

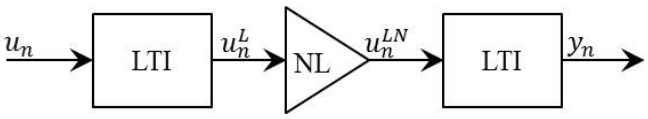

Figure 1.27: Block scheme representation of the Wiener-Hammerstein model.

pre/post-distorters are implemented in real systems. Therefore, the linearity performance and complexity of the model are used for model comparison. The performance of the digital pre/post-distortion algorithms that does not take the memory effects into account is severely degraded as the bandwidth of the input signal increases [Kim01]. Therefore, the linearity performance of the memoryless pre/post-distorters, such as look up table model pre/post-distorter, is lower than the polynomial ones that take the memory effects into account. Among the memory pre/post-distorters, the Volterra series and its simplified version memory polynomial pre/post-distorters are better in linearization. The best performance in PA behavior reproduction is given by the Volterra series, the others are originated by this model but their performances are poorer. Choosing the Volterra model to represent the PA non-linearities and memory effects can be also advocated by the following reasons ([Boy85]):

- it can represent a large class of non-linear systems,
- it is linear regarding the parameters, but non-linear regarding the input signal,
- it is stable in a bounded input-bounded output (BIBO) sense.

1.6 Conclusions

In this chapter, the fundamental parameters of power amplifiers and a non-linear analysis have been presented. PAs are characterized by their efficiency and linearity, which depend on the mode of operation. Classes-A, -B, -AB and -C preserve the input signal waveform and hence offer more linear but less efficient amplification compared to classes-D,-E ,-F and -J. Unfortunately, efficiency and linearity are trade-off parameters in PAs. Any increase in the efficiency of a PA occurs at the expense of linearity and vice versa. While efficiency is required for the power amplifier to increase battery lifetime and to minimize thermal problems, linearity is required for a spectrally efficient transmission of high date rate signals.

For application in WLAN, the requirement on the PA linearity is very stringent and, hence, it makes sense to use a linear PA (class A, AB, or B) together with an efficiency-enhancement technique. PAs exhibit high gain when operated in switched mode (class-D,

-E -F an -J). The high efficient switch mode PAs can amplify only constant envelope modulated signal (such as GSM) without distortion. Whereas, spectral efficient high data rate communication systems currently in use such as LTE needs varying envelope and phase modulated signal. Thus, modern communication systems require highly linear PAs to avoid adjacent channel interference and distortion. When dealing with a base station, a linearization technique, such as feedback, feedforward and pre-distortion can be used to reach the required linearity level. However, when dealing with a battery driven mobile terminals, where more constraints about energy consumption and circuit size are required, a digital efficiency enhancement technique such as a PAPR reduction technique can be used. In CR uplink systems, another solution can be followed, the digital post-distortion technique.

In the next chapter, we intend to develop a digital post-distortion technique for a CR-PA. Indeed, transmitting a non-linearly-amplified signal increases the signal powers at the second and the third-order harmonics (the power at higher orders can be neglected), which may interfere with communication systems operating at those frequencies. However, unlike primary systems, the digital post-distortion technique can be used in a CR context and more particularly in the uplink context. Indeed, a main feature of CR is the autonomous exploitation of locally unused frequencies to improve the spectrum utilization [Mit00]. In addition, a CR is able to adapt the system transmission and reception parameters without user intervention. Also, it can understand and follow the actions taken and the choices made by users in order to learn and become more responsive over time. Thus, it is possible to verify the spectrum state at the second and the third harmonics. For CR uplink communications, choosing the digital post-distortion technique can also be motivated by the following reasons: no feedback path is necessary at the CR terminal; in addition, the post-distortion technique calculations are performed at the cognitive base station receiver. This decreases the amount of signal processing at the CR terminal. Therefore, energy consumption can be reduced and the CR-PA efficiency can be improved.

For these reasons, we will first study the requirements of a CR environment (especially the requirement of a CR-PA behavioral modeling). Second, based on these requirements a behavioral modeling of a CR-PA will be given. Then, an estimation algorithm will be developed to compensate the CR-PA distortions at the baseband part of the receiver. The proposed approach will be presented for the single-carrier modulation. It will be evaluated by means of system level simulations.

Chapter 2

Digital post-distortion techniques based on an IMM structure combining Kalman algorithms

Contents

2.1 Introduction

In uplink cognitive radio (CR) communications, radio frequency (RF) transceivers must be efficient to save the mobile terminal battery autonomy. Therefore, when designing the CR power amplifier (CR-PA), an obvious objective is to optimize efficiency over a large bandwidth. As a consequence, as it is mentioned in the previous chapter, a trade-off between efficiency and linearity needs to be met. Indeed, the CR-PA operates in its non-linear region and then frequency-dependent distortions are generated. To solve this problem, we choose to develop a digital post-distortion and detection technique. It is based on a dynamic Volterra model to take into account the non-linear behavior of the CR-PA. The key feature of the proposed technique is the joint estimations of the model parameters and the CR-PA input samples. For this reason, the proposed digital post-distortion is based on an interacting multiple model (IMM) algorithm combining various Kalman filter (KF)-based estimators using different model parameter dynamics. This makes it possible to track the time variations of the Volterra kernels while keeping accurate estimates when those parameters are static. Furthermore, the IMM algorithm is governed by a design parameter called the transition probability matrix (TPM). This parameter can be either *a priori* defined or recursively estimated with the input samples and the Volterra kernels. In this work we consider the first case using a single-carrier signal.

This approach should be *a priori* interesting when dealing with an uplink communication as the non-linearity processing is exported to the digital part of the receiver which is the base station in this case. Another motivation is that the PA can operate freely at its maximum efficiency in addition to the reduced amount of signal processing at the baseband part of the mobile transmitter. It should be noted that the digital post-distortion is relevant in the CR context as the mobile terminal have a database about the radio environment and the already allocated resources.

2.2 CR-PA modeling requirements and system model

2.2.1 CR concept

The cognitive radio (CR) and dynamic spectrum access paradigms have emerged as a promising solution to conciliate the current spectrum demand growth and its underutilization without changes to the existing legacy wireless systems. The CR enables much higher spectrum efficiency through opportunistic spectrum access. Therefore, it is an attractive technology for future wireless communications [Gol09]. The basic idea of CR is to allow secondary/unlicensed users to access in an opportunistic and non-interfering manner some licensed bands temporarily unoccupied by primary/licensed users [Gha08].

In the next section, we study the implications of the above statements on the mobile terminal behavior and more particularly the CR-PA behavior.

2.2.2 Requirements

According to the power efficiency curve depicted in Figure 1.4 (See chapter 1), the PA must operate in the compression or the saturation region to get the maximum efficiency. However, this causes signal envelop distortion. Moreover, the compression effects, the non-linearities and the memory effects vary with the carrier frequency f_c as similarly as the PA gain [Liu04]. For traditional-communication standards such as global system for mobile communication (GSM), universal mobile telecommunications system (UMTS) or long term evolution (LTE), PAs have narrower bandwidths than the CR-PA ones. Indeed, a CT should be able to access any available sub-band B_W to satisfy its needs of data rate. As the CR-PA behavior depends on the signal bandwidth and the constellation size, it varies according to the available spectral resources.

This means that the CR-PA behavior is affected by a couple of correlated factors:

1. the input signal PAPR, which is constellation dependent,
2. the spectrum access which consists in transmitting over an available spectral resource defined by the couple of parameters (B_W, f_c).

Then, the key feature to get an accurate behavioral modeling of the CR-PA is to characterize the following phenomena:

1. the CR-PA non-linearity and the memory effects,

2. the bandwidth switching mode of a CT: when a primary user begins transmitting over a spectral resource used by a CT, the latter must switch immediately to another resource. This switching may be done in another frequency band or may not happen (for instance when no primary user utilizes the resource). See Figure 2.1.

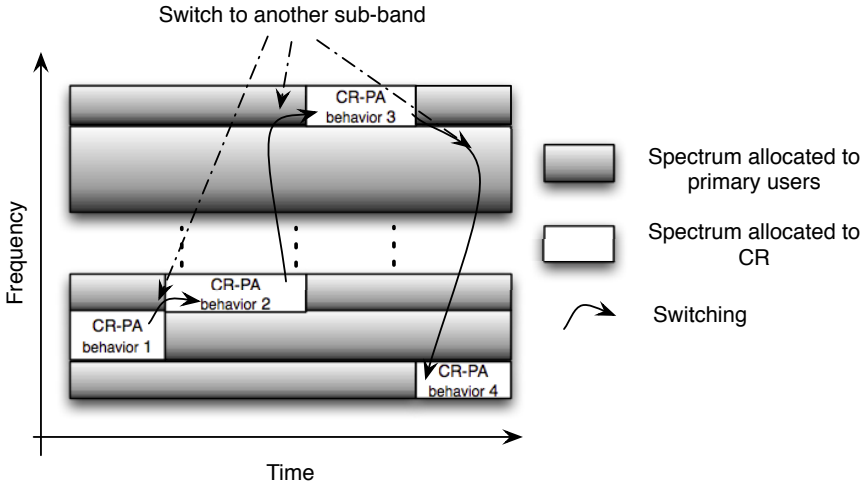

Figure 2.1: CT switching regarding available frequency resources

2.2.3 System model

The system model considered in this chapter is depicted in Figure 2.2, where u_k denotes the PA input sample at time k, y_k the output sample of the equivalent channel combining the CR-PA and the channel. n_k is an AWGN. z_k is the received sample at time k.

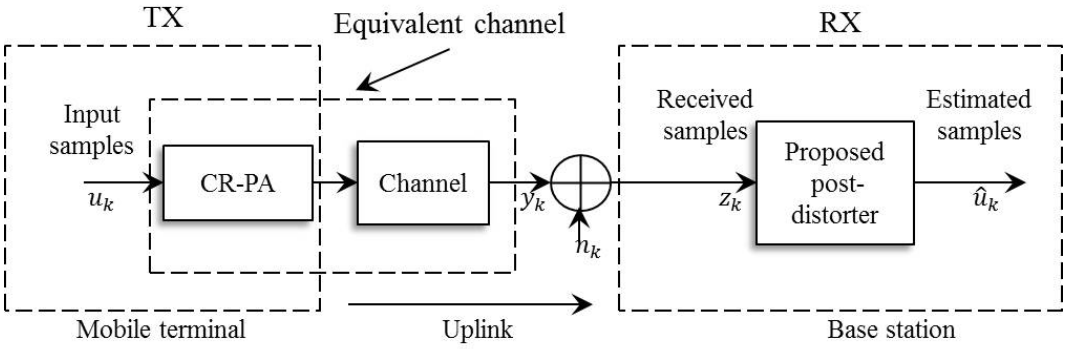

Figure 2.2: Baseband equivalent transceiver.

Modeling schemes are required to predict the CR-PA AM-AM characteristic over a very wide bandwidth by adopting a dynamic model. However, no one can deny the non-linear

effect of the receiver RF stage components such as the low noise amplifier (LNA) and the analog to digital converter (ADC). Therefore, according to the state of the art about the non-linear models with memory that we hold in chapter 1, we have chosen the Volterra series. It can be written as follows [Rug81]:

$$y_k = \sum_{n=1}^{P} \frac{\begin{pmatrix} 2n+1 \\ n \end{pmatrix}}{2^{2n}} \sum_{\tau_1=0}^{M-1} \cdots \sum_{\tau_n=0}^{M-1} h_n(\tau_1, \ldots, \tau_{2n+1}, k) \prod_{s=0}^{n} u^*_{k-\tau_s} \prod_{r=n+1}^{2n+1} u_{k-\tau_r} \qquad (2.1)$$

where $h_n(\tau_1, \ldots, \tau_{2n+1}, k)$ for $n \in \{0, \cdots, P\}$ are the Volterra model kernels. Since they are time-varying, they depend on k. P is the non-linearity order and M is the memory depth ($M \leq 2$).

After choosing the model, the following step is to estimate the model parameters. In the next two sections, the joint estimations of the Volterra kernels and the transmitted samples are addressed. It is based on Kalman filtering.

2.3 Proposed algorithms for CR-PA distortions compensation in a single-carrier system

2.3.1 Introduction

In a single-carrier case, we consider that the propagation channel is non-frequency selective. When using a single-carrier signal, the channel can be frequency selective. In this case, the channel should be estimated using a training sequence and then equalized before carrying the post-distortion. The general scheme is depicted in Figure 2.2 and adjusted to the single-carrier is shown in Figure 2.3. In this case, u_k represents the transmitted symbol (PSK, QAM, etc.).

For a traditional-communication system, where the carrier frequency and the bandwidth *a priori* are fixed, the PA behavior can be described by a single model. In a KF-based estimation context, this leads to a single state space representation (SSR) of the system. However, in a CR context, the PA has to operate with various sets $\{B_w, f_c, P_{in}\}$. Switching from one sub-band to another modifies the PA non-linearity properties and hence causes sudden variations of the Volterra kernels. So, our purpose is to design an estimator which is able to provide accurate estimations of the model parameters as well as the input samples and which is able to track the model parameter modifications.

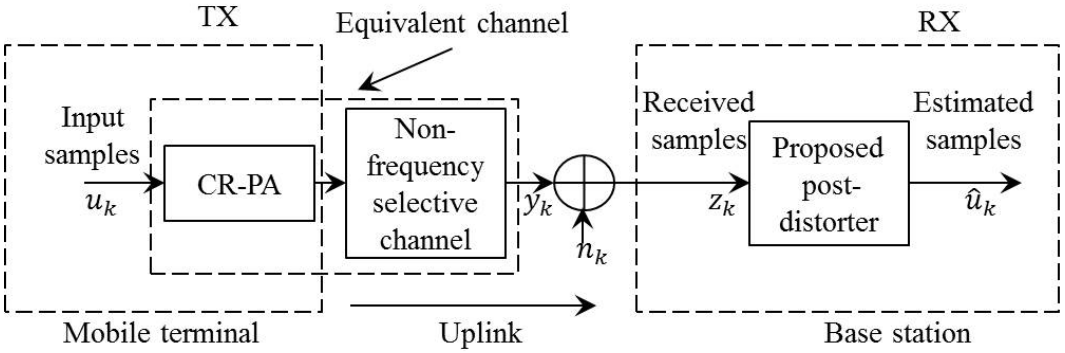

Figure 2.3: Baseband equivalent transceiver for a single-carrier transmission.

When modeling the Volterra kernels by random walks with large values of the driving process variances, the Volterra kernels can be tracked by a Kalman-algorithm based method. Nevertheless, this assumption is not well-suited during the transmission in a specific sub-band. Indeed, even if the Volterra kernels do not change, the Kalman-algorithm based estimates fluctuate much. To avoid the above problem, we propose to use a multiple-model (MM) based identification. This approach consists in taking advantage of various estimators based on different *a priori* models.
Three generations of MM methods have been proposed for the last years:

- in the first one, the estimators run in parallel and independently [Li05],
- the second one aims at introducing a cooperation strategy between the estimators[1],
- in the third one, the number of estimators varies in time [Li05]. This includes the variable-structure MM algorithms such as the likely model set algorithm [Li00].

In the following, we suggest focusing our attention on the second family by using the interacting multiple model algorithm (IMM) combining various KF based estimators. It provides a trade-off between computational cost and accuracy performance compared to generalized pseudo Bayesian (GPB) algorithms, GPB1 and GPB2 [Bar01]. Note that the IMM has been used in a wide range of applications, from global position system (GPS) navigation to target tracking in radar applications [Fau09, Jil04].

In the following, we present the joint estimations of the Volterra kernels and the input signal when the behavior of the non-linear equivalent channel is non time varying. The use

[1]As pointed by Bar-Shalom in [Bar01], using too many estimators does not necessarily improve the estimation algorithm performance. It also increases the computational cost. Hence, two or three estimators are usually used in an IMM.

of the Gaussian mixture extended Kalman filter (GM-EKF) is explained. Then, the CR-PA case, when the equivalent channel is time varying, is addressed by presenting the IMM based post-distorter.

2.3.2 SSR and GM-EKF to compensate the distortion of a non-linear equivalent channel with time invariant behavior

When the behavior of the non-linear equivalent channel does not change over time during the communication, the Volterra kernels remain constant. In order to estimate the Volterra kernels and the input signal, let us first focus our attention on the SSR of the resulting system.

Firstly, we propose to store N Volterra kernels[2] in a column vector $\mathbf{C}_k$. In the time-invariant context, it satisfies:

$$\mathbf{C}_k = \mathbf{C}_{k-1} \tag{2.2}$$

The above equation can be written in another way:

$$\mathbf{C}_k = \mathbf{C}_{k-1} + \mathbf{w}_k(m_{st}) \tag{2.3}$$

where $\mathbf{w}_k(m_{st})$ is an AWGN with zero-mean and covariance matrix $\overline{\mathbf{Q}}(m_{st})$ equal to $\mathbf{0}_{N\times N}$. The index st stands for static behavior. This type of notation will be useful in the rest of the chapter, especially when dealing with the time-varying behavior of the CR-PA.
Then, let $\mathbf{D}_k = \left[\begin{array}{ccc} u_k & \cdots & u_{k-M+1} \end{array} \right]^T$ be the vector storing the M last samples to be estimated ($M \leq 2$). It satisfies:

$$\mathbf{D}_k = \mathbf{F}'\mathbf{D}_{k-1} + \mathbf{G}'u_k \tag{2.4}$$

where $\mathbf{F}' = \left(\begin{array}{cc} \mathbf{0}_{1\times M-1} & 0 \\ \mathbf{I}_{M-1} & \mathbf{0}_{M-1\times 1} \end{array} \right)$ and $\mathbf{G}' = \left[\begin{array}{cc} 1 & \mathbf{0}_{1\times M-1} \end{array} \right]^T$. When u_k is complex the structure of $\mathbf{D}_k$ is as follows:

$$\mathbf{D}_k = \left[\begin{array}{cccccc} Re(u_k) & \cdots & Re(u_{k-M+1}) & Im(u_k) & \cdots & Im(u_{k-M+1}) \end{array} \right]^T$$

As we aim at jointly estimating the Volterra kernels and the input sample u_k, the state vector is defined by:

$$\mathbf{x}_k = \left(\begin{array}{c} \mathbf{D}_k \\ \mathbf{C}_k \end{array} \right) \tag{2.5}$$

[2] N depends on the memory depth and the non-linearity order respectively M and P.

As $\mathbf{x}_k$ is always real, we present the equations in the real form. Given (2.3)-(2.5), the resulting SSR of the system is:

$$\begin{aligned} \mathbf{x}_k &= \mathbf{F}\mathbf{x}_{k-1} + \mathbf{G}\mathbf{v}_k(m_{st}) = \mathbf{F}\mathbf{x}_{k-1} + \begin{bmatrix} \mathbf{G}' u_k \\ \mathbf{w}_k(m_{st}) \end{bmatrix} \\ z_k &= h(\mathbf{x}_k) + n_k = y_k + n_k \end{aligned} \tag{2.6}$$

where $\mathbf{F} = \begin{pmatrix} \mathbf{F}' & \mathbf{0}_{M\times N} \\ \mathbf{0}_{N\times M} & \mathbf{I}_N \end{pmatrix}$ is the one-step transition matrix, $\mathbf{G} = \begin{pmatrix} \mathbf{G}' & \mathbf{0}_{M\times N} \\ \mathbf{0}_{1\times N} & \mathbf{I}_N \end{pmatrix}$, $\mathbf{v}_k(m_{st}) = \begin{bmatrix} u_k \\ \mathbf{w}_k(m_{st}) \end{bmatrix}$, n_k is a zero-mean white Gaussian noise with variance σ_n^2 and $h(\cdot)$ is a non-linear function of the state vector, which illustrates (2.1). For complex constellations, z_k is replaced by $\mathbf{z}_k = [Re(z_k) \quad Im(z_k]^T$.

At this stage, if $z_{1:k}$ denotes the set of received samples until time k, two problems must be addressed:

1. the probability density function (pdf) of the model noise $\mathbf{G}\mathbf{v}_k(m_{st})$ should be Gaussian in order to use a Kalman filter (KF) [Ama11]. Indeed, one can be convinced by the necessity of this assumption when one derives the Kalman-filter equations by searching the relationships between the pdf of the state vector $p(\mathbf{x}_{k-1}|z_{1:k-1})$ and $p(\mathbf{x}_k|z_{1:k-1})$ on the one hand, and $p(\mathbf{x}_k|z_{1:k})$ and $p(\mathbf{x}_k|z_{1:k-1})$ on the other hand [Li00]. Here, this assumption is not satisfied because of u_k which is uniformly distributed on the finite alphabet $\gamma = \{d_i\}_{i=1,\cdots,q}$ with $q = 2^{n_b}$ and n_b the number of bits per symbol. For this reason, u_k is assumed to have a conditional pdf, denoted by $p(u_k|z_{1:k})$, which is a sum of q Gaussian distributions. Each one is centred on one value of the constellation alphabet with a very small variance λ_i^2 (see Figure 2.4). Thus, we have:

$$p(u_k) = \sum_{i=1}^{q} \frac{1}{q} \mathcal{N}(d_i, \lambda_i^2) \tag{2.7}$$

   where $\mathcal{N}(\mu, \sigma^2)$ denotes a Gaussian distribution with mean μ and variance σ^2. For the sake of simplicity, the variances $\{\lambda_i^2\}_{i=1,\cdots,q}$ can be set at the same value λ^2. As a consequence, the resulting model noise $\mathbf{G}\mathbf{v}_k(m_{st})$ is Gaussian, zero-mean and with covariance matrix $\begin{pmatrix} \lambda^2 \mathbf{G}'\mathbf{G}'^T & \mathbf{0}_{M\times N} \\ \mathbf{0}_{N\times M} & \overline{\mathbf{Q}}(m_{st}) \end{pmatrix}$.
   The drawback of this assumption is that the number of EKF is multiplied by q. Indeed, q Kalman algorithms run in parallel, each one being based on a specific assumption of the transmitted symbol.

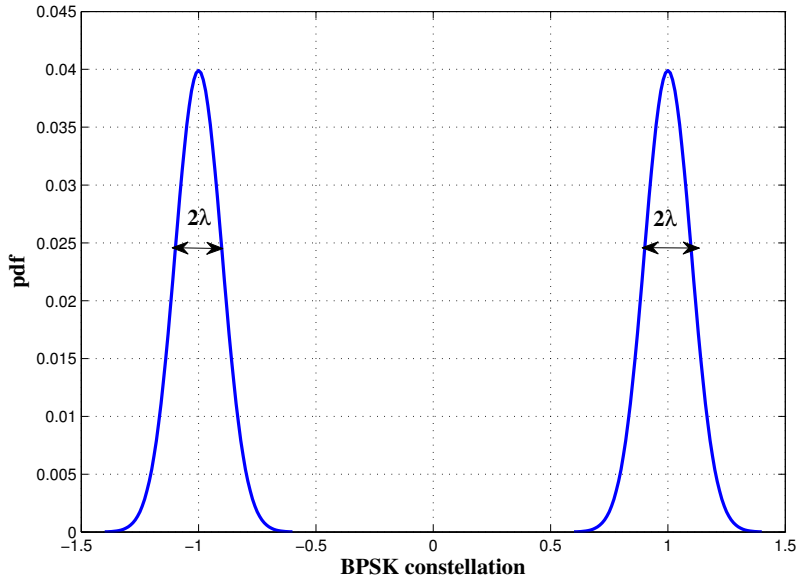

Figure 2.4: pdf of u_k for the BPSK alphabet ($q = 2$).

2. as the relationship that connects the observation z_k to the state vector $\mathbf{x}_k$ is non-linear, the standard Kalman filter cannot recursively estimate the state vector. However, there are various other ways. Particle filtering could be used, as in [Gir10]. Nevertheless, this type of method leads to a high computational cost that is not suitable to the current application. As an alternative, the sigma point Kalman filters (SPKFs) can be considered. They consist in using the so-called sigma points characterizing the state vector probability density, propagating them through the non-linear function, estimating the state vector as well as the estimation error covariance matrix by combining the resulting sigma points. They include the unscented Kalman filter (UKF), which is based on the unscented transformation [Wan02], the central difference Kalman filter (CDKF) which is based on the second-order Sterling polynomial interpolation formula [Ito00], the cubature Kalman filter [Ara09] and the quadrature Kalman filter (QKF) where the sigma points are chosen by using the Gauss-Hermite quadrature rule [Cha99]. Another solution is the EKF. It consists in using the 1^{st}-order Taylor[3] expansion of the non-linear function $h(\cdot)$. This method is popular because its computational cost is smaller than the SPKF ones and it has the advantage of retrieving the same types of equations as the standard Kalman filter. For this reason, the proposed algorithm is based on an

[3]It should be noted that the second order-EKF (SO-EKF) also exists and is based on the second-order Taylor expansion of the non-linear function [Bar01]. It is rather used if the 1st-order Taylor approximation is not sufficient to represent the non-linear function.

EKF.

The resolution of the two above problems leads a Kalman filter variant based on the EKF principle able to circumvent the non-Gaussianity issue called the GM-EKF.

2.3.2.1 A GM-EKF structure

In the following, let $\hat{\mathbf{x}}_{k|k-1}^{i,st}$ be the *a priori* estimation, provided by the i^{th} EKF $i \in \{1, \cdots, q\}$, of the state vector $\mathbf{x}_k$ given the last $(k-1)$ observations $z_{1:k-1}$. Then, let us introduce the Jacobian matrix $\mathbf{h}_i(k)$. It is the partial derivative of $h(\cdot)$ according to the state vector $\mathbf{x}_k$, which is here evaluated at the *a priori* estimation $\hat{\mathbf{x}}_{k|k-1}^{i,st}$:

$$\mathbf{h}_k^i = \frac{\partial h}{\partial \mathbf{x}_k}(\hat{\mathbf{x}}_{k|k-1}^{i,st}) \quad \text{with} \quad i = 1, \cdots, q \tag{2.8}$$

Thus, given (2.6) and (2.8), and using the Taylor expansion, the observation z_k can be expressed as follows:

$$z_k \approx h(\hat{\mathbf{x}}_{k|k-1}^{i,st}) + \mathbf{h}_k^i(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1}^{i,st}) + n_k \quad \text{with} \quad i = 1, \cdots, q \tag{2.9}$$

Thanks to (2.9), the so-called innovation of each EKF can be expressed as:

$$e_{k|k-1}^{i,st} = z_k - h(\hat{\mathbf{x}}_{k|k-1}^{i,st}) \approx \mathbf{h}_k^i \tilde{\mathbf{x}}_{k|k-1}^{i,st} + n_k \quad \text{with} \quad i = 1, \cdots, q \tag{2.10}$$

where $\tilde{\mathbf{x}}_{k|k-1}^{i,st} = \mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1}^{i,st}$ is the *a priori* error of the state estimation for the i^{th} estimator. Therefore, as done in the standard Kalman filter, the innovation can be expressed linearly from the *a priori* estimation error $\tilde{\mathbf{x}}_{k|k-1}^{i,st}$ and the measurement noise n_k.

Given the above notations, the q EKFs making up the GM-EKF operates with the two following steps:

- In the prediction step, based on (2.6) and (2.7), the *a priori* state vector estimate satisfies:

$$\begin{aligned} \hat{\mathbf{x}}_{k|k-1}^{i,st} &= E[\mathbf{x}_k | z_{1:k-1}] \underset{(2.6)}{=} \mathbf{F}\hat{\mathbf{x}}_{k-1|k-1}^{i,st} + E\left[\begin{bmatrix} \mathbf{G}' u_k \\ \mathbf{w}_k(m_{st}) \end{bmatrix}\right] \\ &\underset{(2.7)}{=} \mathbf{F}\hat{\mathbf{x}}_{k-1|k-1}^{i,st} + \begin{bmatrix} \mathbf{G}' \\ \mathbf{0}_{N\times 1} \end{bmatrix} d_i \quad \text{with } i = 1, \cdots, q \end{aligned} \tag{2.11}$$

  where $E[\cdot|z_{1:k-1}]$ denotes the conditional expectation given the set of observations $z_{1:k-1}$. The *a priori* error covariance matrix $\mathbf{P}_{k|k-1}^{i,st}$ is then deduced as follows:

$$\begin{aligned} \mathbf{P}_{k|k-1}^{i,st} &= E[\tilde{\mathbf{x}}_{k|k-1}^{i,st}(\tilde{\mathbf{x}}_{k|k-1}^{i,st})^T] \\ &= \mathbf{F}\mathbf{P}_{k-1|k-1}^{i,st}\mathbf{F}^T + \lambda^2 \begin{bmatrix} \mathbf{G}' \\ \mathbf{0}_{N\times 1} \end{bmatrix} \begin{bmatrix} \mathbf{G}' \\ \mathbf{0}_{N\times 1} \end{bmatrix}^T \quad \text{with} \quad i = 1, \cdots, q \end{aligned} \tag{2.12}$$

- In the filtering step, the innovation variance $(\sigma^2)^i$ is first computed. Then, the Kalman gain $K_k^{i,st}$ is defined and given (2.10) the estimation of the state vector is updated with the innovation weighted by the Kalman gain:

$$\begin{aligned}(\sigma^2)^i &= (\mathbf{h}_k^i)^T \mathbf{P}_{k|k-1}^{i,st} \mathbf{h}_k^i + \sigma_n^2 \quad \text{with} \quad i = 1, \cdots, q \\ K_k^{i,st} &= \mathbf{P}_{k|k-1}^{i,st} \mathbf{h}_k^i / (\sigma^2)^i \\ \hat{\mathbf{x}}_{k|k}^{i,st} &= \hat{\mathbf{x}}_{k|k-1}^{i,st} + K_k^{i,st} e_{k|k-1}^{i,st}\end{aligned} \tag{2.13}$$

  The *a posteriori* estimation error covariance matrix in the i^{th} EKF at time k can be obtained as follows:

$$\mathbf{P}_{k|k}^{i,st} = (\mathbf{I}_{M+N} - K_k^{i,st} \mathbf{h}_k^i) \mathbf{P}_{k|k-1}^{i,st} \quad \text{with} \quad i = 1, \cdots, q \tag{2.14}$$

At this stage, a hard decision is carried out between the outputs of the q EKFs to decide which estimate is the output of the GM-EKF. This decision consists in comparing the q likelihood values:

$$\beta^i = \mathcal{N}(z_k; h(\hat{\mathbf{x}}_{k|k-1}^{i,st}), (\sigma^2)^i), \quad \text{with} \quad i = 1, \cdots, q \tag{2.15}$$

where $\mathcal{N}(z_k; \mu, \sigma^2)$ denotes the value of the Gaussian distribution with mean μ and variance σ^2 when the random variable is equal to z_k.
If $i - max$ denotes the index so that $\beta^{i-max} = \underset{i=1,\cdots,q}{max}(\beta^i)$, the final estimate $\hat{\mathbf{x}}_{k|k}^{st}$ is equal to $\hat{\mathbf{x}}_{k|k}^{i-max,st}$, its *a posteriori* error covariance matrix denoted by $\mathbf{P}_{k|k}^{st}$ is $\mathbf{P}_{k|k}^{i-max}$ and the innovation variance is $S_k = (\sigma^2)^{i-max}$. In addition, $\hat{\mathbf{x}}_{k|k-1}^{i-max,st}$ is denoted by $\hat{\mathbf{x}}_{k|k-1}^{st}$. This latter quantity will be useful in the next section. Then, at the next time step, every state vector estimate $\{\hat{\mathbf{x}}_{k|k}^{i,st}\}_{i=1,\cdots,q}$ is set at the same vector, namely $\hat{\mathbf{x}}_{k|k}^{i-max,st}$. In addition, every *a posteriori* error covariance matrix $\{\mathbf{P}_{k|k}^{i,st}\}_{i=1,\cdots,q}$ is set at the same value, namely $\mathbf{P}_{k|k}^{i-max,st}$. It should be noted that the considered decision is instantaneous decision. Therefore, this approach does not correspond to a classic Viterbi approach. A summary of the m_{st}-GM-EKF steps is given in Figure 2.5.

However, the CR-PA behavior varies over time due to sub-band switching during a communication. See Figure 2.1. Therefore, a GM-EKF based on $\mathbf{w}_k(m_{st})$ is not able to track the Volterra-kernel variations. To solve this issue, we propose to consider at least two GM-EKFs combined in an IMM algorithm.

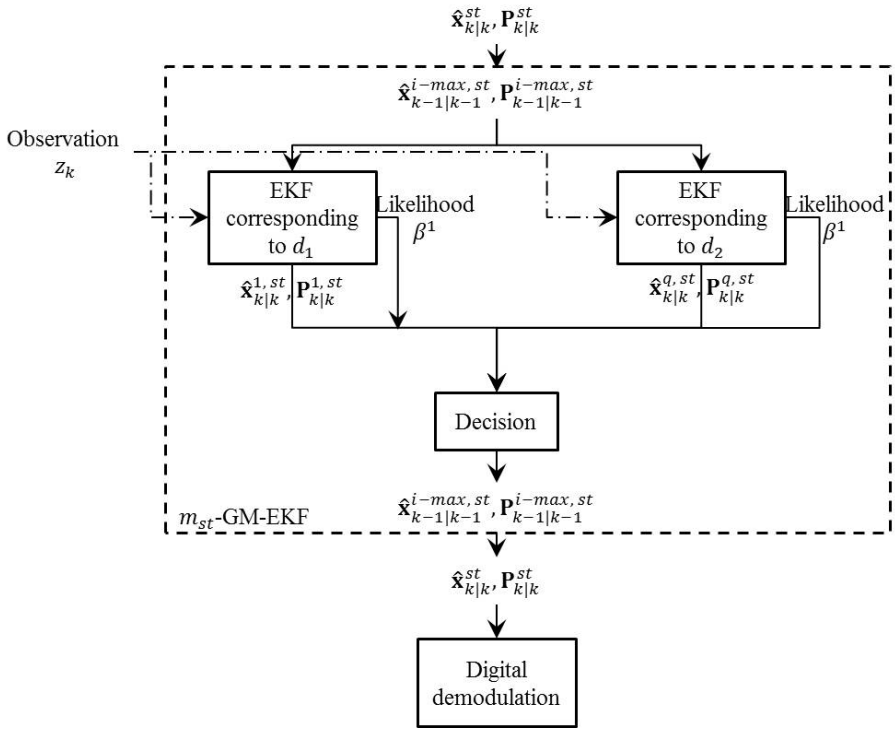

Figure 2.5: m_{st}-GM-EKF for a q-size alphabet.

2.3.3 GM-EKF-based IMM to track the time-varying equivalent channel and estimate the CR-PA input samples for single-carrier systems

2.3.3.1 From the GM-EKF to the GM-EKF based IMM

The GM-EKF, proposed in the previous section, is not able to track the Volterra-kernel variations due to the zero model noise. Furthermore, if the model noise $\mathbf{w}_k$ in (2.3) is no longer zero but a zero-mean AWGN with a non-zero variance, the GM-EKF can track these variations but it does not provide accurate estimates of static parameters. For these reasons, we propose to consider several GM-EKFs based on different *a priori* modeling of the time evolutions of the Volterra kernels. The first GM-EKF is already detailed in the previous section. For the second GM-EKF, we keep the SSR given in (2.6), but the covariance matrix of the model noise $\mathbf{w}_k$ is now defined by $\overline{\mathbf{Q}}(m_{ld})$ where the index ld stands for large dynamics. We define:

$$\overline{\mathbf{Q}}(m_{ld}) = \sigma_w^2 \mathbf{I}_N$$

The algorithm is hence able to track large variations[4]. Therefore, the static and the time-varying parameter assumptions can be considered by using the two above GM-EKFs. This leads to a 2-mode IMM combining the m_{st}-GM-EKF and the m_{ld}-GM-EKF.
We can also study what a third GM-EKF could bring into the estimation process. It can be based on a SSR where the model noise $\mathbf{w}_k$ is defined by:

$$\overline{\mathbf{Q}}(m_{sd}) = \varepsilon\sigma_w^2\mathbf{I}_N$$

where $\varepsilon << 1$. This leads to a 3-mode IMM which also includes the m_{sd}-GM-EKF. Note that the index *sd* stands for small dynamics. When introducing this GM-EKF, we aim at improving the parameter estimate accuracy and consequently the BER performance.

2.3.3.2 About the derivation of the IMM equations

Let us look at the IMM more precisely when three modes are considered i.e. the *st*, *sd* and *ld* modes. In the remainder of the chapter, $\{\hat{\mathbf{x}}^j_{k|k}\}_{j=st,sd,ld}$ denote the *a posteriori* estimates at time k provided by the m_{st}-GM-EKF, the m_{sd}-GM-EKF and the m_{ld}-GM-EKF respectively. $\{\mathbf{P}^j_{k|k}\}_{j=st,sd,ld}$ are their corresponding error covariance matrices.
The purpose of the IMM is to estimate the state vector $\mathbf{x}_k$ given the observations $z_{1:k}$ by using various SSRs. This amounts to searching for the state vector pdf $p(\mathbf{x}_k|z_{1:k})$. In the IMM algorithm, due to the various modes that can be considered, it is approximated by a Gaussian mixture:

$$p(\mathbf{x}_k|z_{1:k}) \approx \sum_{j=st,sd,ld} \mu^j_k \mathcal{N}(\mathbf{x}_k; \hat{\mathbf{x}}^j_{k|k}, \mathbf{P}^j_{k|k}) \tag{2.16}$$

where μ^j_k is the probability that the system corresponds to the j^{th} mode m_j at time k; this can be rewritten as $\mu^j_k = Pr\{M_k = m_j\}$ where $m_j = m_{st}, m_{sd}, m_{ld}$. Note that μ^j_k is also called the *a posteriori mode probability.* Given the approximation (2.16), the IMM state vector estimate $\hat{\mathbf{x}}_{k|k}$ corresponds to a weighted sum of the outputs of the GM-EKFs:

$$\hat{\mathbf{x}}_{k|k} = \sum_{j=st,sd,ld} \mu^j_k \hat{\mathbf{x}}^j_{k|k} \tag{2.17}$$

Using (2.17) and the Gaussian mixture formula (See chapter 1 in [Bar01]), the IMM *a posteriori* error covariance matrix satisfies:

$$\mathbf{P}_{k|k} = \sum_{j=st,sd,ld} \mu^j_k \{\mathbf{P}^j_{k|k} + [\hat{\mathbf{x}}^j_{k|k} - \hat{\mathbf{x}}_{k|k}][\hat{\mathbf{x}}^j_{k|k} - \hat{\mathbf{x}}_{k|k}]^T\}$$

[4]for example in the range $\pm 3\sigma_w$ with probability 99.7% at each time step

Then, the purpose of the IMM is to recursively calculate $\hat{\mathbf{x}}^j_{k|k}$, $\mathbf{P}^j_{k|k}$ and μ^j_k. The recursion equations can be deduced by expressing $p(\mathbf{x}_k|z_{1:k})$ using Bayes rules and then approximating $p(\mathbf{x}_{k-1}|z_{1:k-1}, M_k = m_j)$ with $j = st, sd, ld$ by $\mathcal{N}(\mathbf{x}_{k-1}; \hat{\mathbf{x}}^{0j}_{k-1|k-1}, \mathbf{P}^{0j}_{k-1|k-1})$, where $\{\hat{\mathbf{x}}^{0j}_{k-1|k-1}\}_{j=st,sd,ld}$ are the so-called merged means taking into account the GM-EKF estimates and the so-called mixing probabilities $\mu^{l|j}_{k-1|k-1} = Pr\{M_{k-1} = m_l|M_k = m_j\}$ where $l, j = st, sd, ld$. Introducing this merged mean can be seen as a strategy of cooperation between the different Kalman algorithms. For the proof and more details, the reader may refer to [Bar01]. A block scheme about the the whole algorithm is given in Figure 2.6.

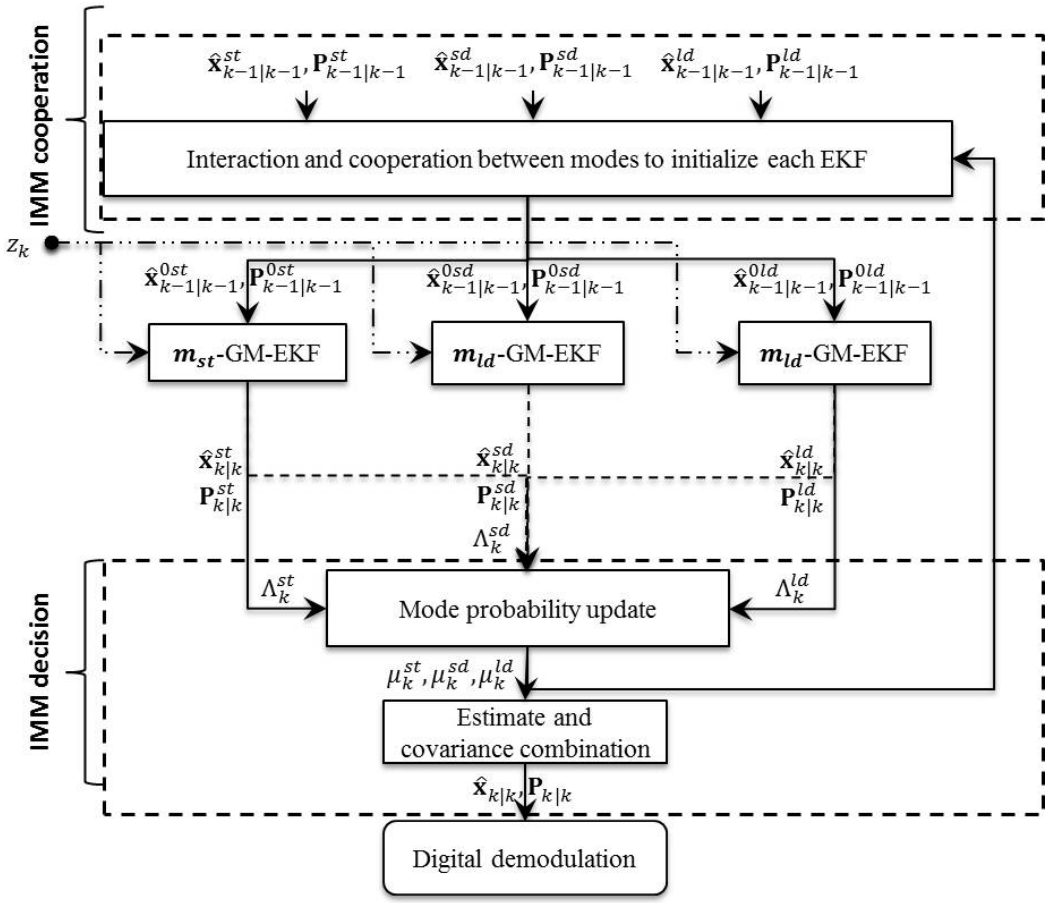

Figure 2.6: Proposed post-distorter when three GM-EKFs are used.

To update the mixing probabilities $\{\mu^{l|j}_{k-1|k-1}\}_{l,j=st,sd,ld}$ and the *a posteriori* mode probabilities $\{\mu^j_k\}_{j=st,sd,ld}$, the system is assumed to be a Markov chain (depicted in Figure 2.7) with the following transition probabilities between the modes m_l at time $(k-1)$ and m_j at

time k:

$$Pr\{M_k = m_j | M_{k-1} = m_l\} = p_{l,j}$$
$$\text{where} \quad \sum_{j=st,sd,ld} p_{l,j} = 1, \quad l = st, sd, ld \tag{2.18}$$

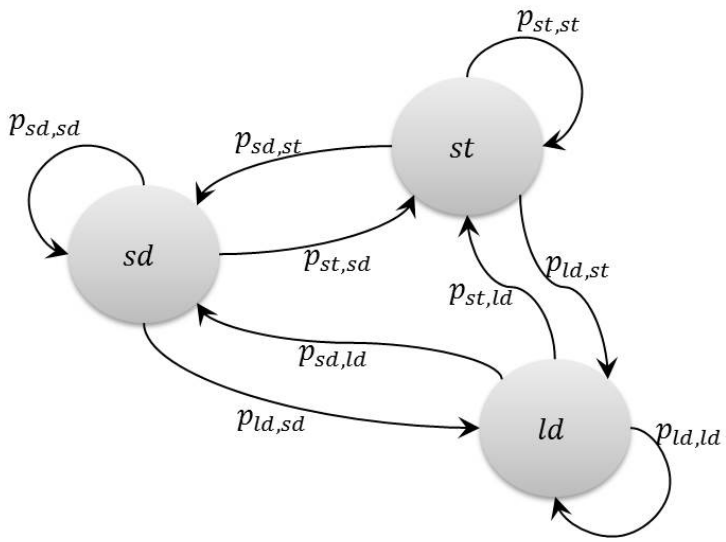

Figure 2.7: The Markov chain characterizing the switches between modes

Usually, these probabilities are stored in the so-called transition probability matrix (TPM) denoted by $\mathbf{\Pi}$:

$$\mathbf{\Pi} = [p_{l,j}], \quad j, l = st, sd, ld \tag{2.19}$$

The TPM plays a crucial role in the design of the IMM. Concerning its setting, the TPM is assumed to be known or chosen *a priori*. In this case, given the properties of the probability transition and to ensure the equiprobability between modes in the steady state, the TPM can be defined by:

$$\mathbf{\Pi} = \begin{bmatrix} p_{st,st} & p_{st,sd} & 1 - p_{st,st} - p_{st,sd} \\ 1 - p_{st,st} - p_{sd,sd} - p_{st,sd} + p_{ld,ld} & p_{sd,sd} & p_{st,st} + p_{st,sd} - p_{ld,ld} \\ p_{sd,sd} + p_{st,sd} - p_{ld,ld} & 1 - p_{sd,sd} - p_{st,sd} & p_{ld,ld} \end{bmatrix} \tag{2.20}$$

The TPM in (2.20) depends on four transition probabilities. Three of them can be deduced from the mean sojourn time in a state. Indeed, one has:

$$p_{j,j} = 1 - \frac{1}{E[t_j]} \quad \text{for} \quad j = st, sd, ld \tag{2.21}$$

where t_j is the sojourn time in mode j. If one can collect the spectral resource availability in a data base available at each CT for a given time and localization [Mit00], the mean sojourn

time in one sub-band can be estimated and $p_{st,st}$ can be deduced. For the sake of simplicity in this work, it is assumed that the mean sojourn time is the same for all sub-bands. Concerning $p_{sd,sd}$ and $p_{ld,ld}$, the mean sojourn time is assumed to be close to one sampling time to take into account the fact that the equivalent channel behavior changes instantaneously. The fourth transition probability namely $p_{st,sd}$ can be set by the practitioner.

Given the last estimated states $\{\hat{\mathbf{x}}^j_{k-1|k-1}\}_{j=st,sd,ld}$ provided by each GM-EKF and its associated error covariance matrices $\{\mathbf{P}^j_{k-1|k-1}\}_{j=st,sd,ld}$, the IMM estimator structure consists of the steps described in Algorithm 1.

Algorithm 1 An IMM combining GM-EKFs for the joint estimations of the Volterra kernels and the input samples for single-carrier systems

- Computing the mixing probabilities, for $j = st, sd, ld$:

$$\mu_{k-1|k-1}^{l|j} = \frac{1}{c^j} p_{l,j} \mu_{k-1}^{l} \quad j, l = st, sd, ld \quad \text{where} \quad c^j = \sum_{l=st,sd,ld} p_{l,j} \mu_{k-1}^{l}$$

  Deducing the merged means and covariance matrices from the GM-EKF estimates:

$$\hat{\mathbf{x}}_{k-1|k-1}^{0j} = \sum_{l=st,sd,ld} \hat{\mathbf{x}}_{k-1|k-1}^{l} \mu_{k-1|k-1}^{l|j} \quad j = st, sd, ld$$

$$\mathbf{P}_{k-1|k-1}^{0j} = \sum_{l=st,sd,ld} \mu_{k-1|k-1}^{l|j} \{\mathbf{P}_{k-1|k-1}^{l} + [\hat{\mathbf{x}}_{k-1|k-1}^{l} - \hat{\mathbf{x}}_{k-1|k-1}^{0j}][\hat{\mathbf{x}}_{k-1|k-1}^{l} - \hat{\mathbf{x}}_{k-1|k-1}^{0j}]^T\}$$

- State prediction and update:

  Whereas the input of a GM-EKF used alone is the previous *a posteriori* estimate, namely $\hat{\mathbf{x}}_{k-1|k-1}^{j}$ with $j = st, sd$ or ld, the input of the GM-EKF when it is used in an IMM structure is $\hat{\mathbf{x}}_{k-1|k-1}^{0j}$. Then, each GM-EKF provides $\hat{\mathbf{x}}_{k|k-1}^{j}$, $\hat{\mathbf{x}}_{k|k}^{j}$ and $\mathbf{P}_{k|k}^{j}$. In addition, the likelihood functions corresponding to each GM-EKF are obtained as follows:

$$\Lambda_k^j = \mathcal{N}(z_k; h(\hat{\mathbf{x}}_{k|k-1}^{j}), S_k^j) \quad j = st, sd, ld$$

  where $h(\hat{\mathbf{x}}_{k|k-1}^{j})$ is the predicted observation in the m_j-GM-EKF by using $\hat{\mathbf{x}}_{k-1|k-1}^{0j}$ and the innovation covariance matrix S_k^j.

- Updating the mode probabilities:

$$\mu_k^j = \frac{1}{c} \Lambda_k^j c^j, \quad j = st, sd, ld \quad \text{where} \quad c = \sum_{l=st,sd,ld} \Lambda_k^j c^j$$

- Mixing the GM-EKFs estimates to get a final state estimate:

$$\hat{\mathbf{x}}_{k|k} = \sum_{j=st,sd,ld} \mu_k^j \hat{\mathbf{x}}_{k|k}^{j}$$

  It can be seen as a soft decision between the outputs of the GM-EKFs, i.e. a weighted sum of the outputs of the GM-EKFs.

2.3.4 Simulations and results in the single-carrier case

2.3.4.1 Simulation protocols

We consider the system model described in Figure 2.3 with a BPSK digital modulation. The non-linearity order P and the memory depth M of the Volterra model are respectively

equal to 3 and 2. In addition, an AWGN[5] channel n_k is considered. Its variance σ_n^2 is set to have a given signal-to-noise ratio (SNR) equal to $10\log_{10}\frac{P_y}{P_n}$ where $P_y = \frac{1}{N_s}\sum_{k=1}^{N_s}|y_k|^2$ and $P_n = \frac{1}{N_s}\sum_{k=1}^{N_s}|n_k|^2$ respectively denote the power of the received signal and the noise, with N_s the number of samples under study. Moreover, the variance λ^2 of the Gaussians in the GM-EKF is set at 10^{-5}. Concerning the parameters defining the random walks, $\sigma_w^2 = 1$ and $\varepsilon = 10^{-2}$.

Then, let us define the CR-PA model path (CMP). A CMP is a set of Volterra-parameter values recorded during a communication. In our simulations, we use three scenarios:

- the first scenario is a toy example, where the CMP is generated by a Markov chain with three states. The parameter sequences are random walks with covariance $\overline{\mathbf{Q}}(m_{st})$, $\overline{\mathbf{Q}}(m_{sd})$ or $\overline{\mathbf{Q}}(m_{ld})$. The transition matrix between the states *st*, *sd* and *ld* is the following:

$$\begin{bmatrix} 0.99 & 0.006 & 0.004 \\ 0.004 & 0.4 & 0.596 \\ 0.006 & 0.594 & 0.4 \end{bmatrix} \tag{2.22}$$

  Remark: the matrix defined in (2.22) has the same structure as the TPM introduced in (2.20) where $p_{st,st}$, $p_{sd,sd}$, $p_{st,sd}$ and $p_{ld,ld}$ are respectively equal to 0.99, 0.4, 0.006 and 0.4. As a consequence, the model parameters can be successively static, slightly time-varying and strongly time-varying. Every switch between the states is possible. The properties of the generated signal exactly match the model used at the receiver. This is a good way to evaluate the performance of our approach in a perfect case. We can also analyze its limits. The simulations are based on one hundred realizations of this Markov chain. One of them is depicted in Figure 2.8 (a). For the sake of the figure clarity, we only present three kernels. For each realization of the model parameters, the channel equivalent output samples y_k are generated and then disturbed by the additive noise n_k for a given SNR.

  Note that in the ideal case, the TPM in the post-distorter should be set to the matrix (2.22). However, this piece of information cannot be known *a priori*. For this reason, given the observations generated with this transition matrix, we have then studied the performance of our algorithm when the TPM is set to another value. This is of interest to study the sensitivity of the approach to the choice of the TPM.

[5]If we consider a Rayleigh or Rice channel model, the consequence is an increase in the Volterra model memory depth M.

- in the second scenario, the Volterra parameters remain constant during a given period of time and suddenly switch to another value. Unlike the previous scenario, there is no possible switch between the *sd* mode and the *ld* one. It should be noted that this CMP could be approximated by a Markov chain where the TPM is defined by:

$$\begin{bmatrix} 1-\rho & \rho/2 & \rho/2 \\ 1/(1+\delta) & \delta/(1+\delta) & 0 \\ 1/(1+\delta) & 0 & \delta/(1+\delta) \end{bmatrix} \tag{2.23}$$

  where $(\rho, \delta) << 1$. See Figure 2.8 (b). In this case, the mean sojourn time to have static Volterra parameters would be $1/\rho$ whereas the mean sojourn times to have small and large variations in time of the Volterra parameters would be equal to $1+\delta$ [Ben15].

- in the third scenario, the CMP is generated by a five-state TPM defined by:

$$\begin{bmatrix} 0.99 & 0.004 & 0.003 & 0.002 & 0.001 \\ 0.004 & 0.496 & 0.04 & 0.26 & 0.2 \\ 0.003 & 0.23 & 0.497 & 0.23 & 0.04 \\ 0.002 & 0.04 & 0.2 & 0.498 & 0.26 \\ 0.001 & 0.23 & 0.26 & 0.01 & 0.499 \end{bmatrix} \tag{2.24}$$

  The variances of the driving process in the random walks are 0, 0.001, 0.05, 0.5 and 1. Then, at the receiver, the transmitted samples and the Volterra kernels are estimated using the 3-mode IMM algorithm governed by the TPM defined in (2.22). This scenario makes it possible to generate the Volterra parameters whose pdf at each time step does not exactly correspond to the one that can be obtained at the receiver by using an IMM with 2 or 3 estimators. This tends to be more representative of what could happen in real cases.

In the next section, the sensitivity of the TPM is addressed based on the first scenario. Then, we study the algorithm sensitivity to the choice of the number of mode for the three scenarios.

2.3.4.2 Results

2.3.4.2.1 About the algorithm sensitivity to the TPM choice In Figures 2.9-2.11, the sensitivity of our approach to some elements of the TPM is studied. For this purpose, the BER is computed for various SNRs and different values of the transition probabilities. According to Figure 2.9, the lowest BER is obtained for $p_{st,st} = 0.99$. However, the BER

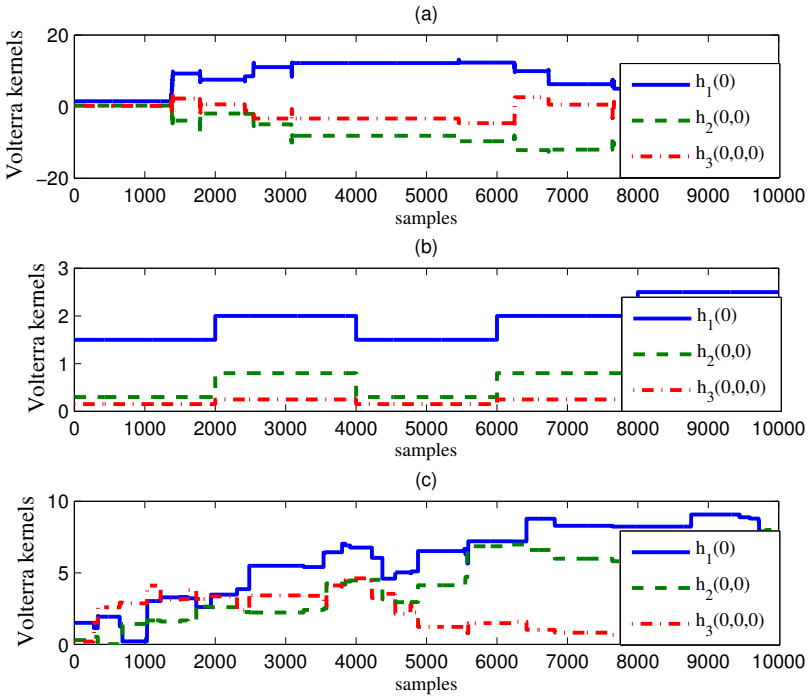

Figure 2.8: CMPs generated for the three scenarios.

does not grossly increase until $p_{st,st} = 0.97$. It means that there is a tolerable error margin to set $p_{st,st}$. In addition, this margin is not sensitive to the SNR according to the simulation we did. Another margin appears also in Figure 2.10 showing the existence of a tolerable error margin about the value of $p_{sd,sd}$. The same observation is depicted by Figure 2.11. As a consequence, in practice, *a priori* setting the TPM would not be so problematic.

2.3.4.2.2 About the sensitivity of the proposed algorithm to the number of modes

- Let us first look at the first scenario

We test the proposed algorithm sensitivity by comparing the BER performance between:

· a 2-mode IMM which combines m_{st}-GM-EKF and m_{ld}-GM-EKF,

· a 3-mode IMM which combines m_{st}-GM-EKF, m_{sd}-GM-EKF and m_{ld}-GM-EKF.

The comparison of BERs between the 2-mode IMM and the 3-mode IMM illustrates a trade-off between computational cost and required QoS. Indeed, Figure 2.12 shows a gain of 2dB in SNR for a BER= 10^{-4} with the 3-mode IMM when compared to the 2-mode IMM.

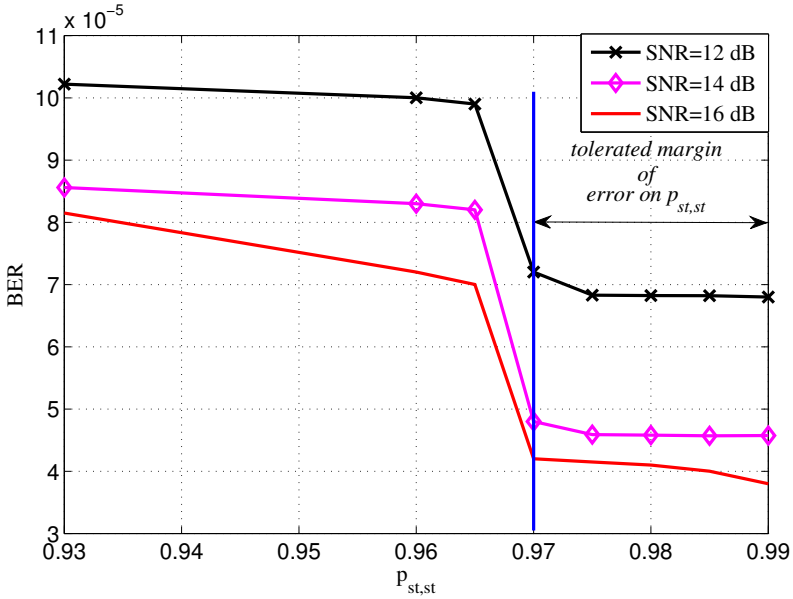

Figure 2.9: The proposed-algorithm sensitivity to errors on $p_{st,st}$.

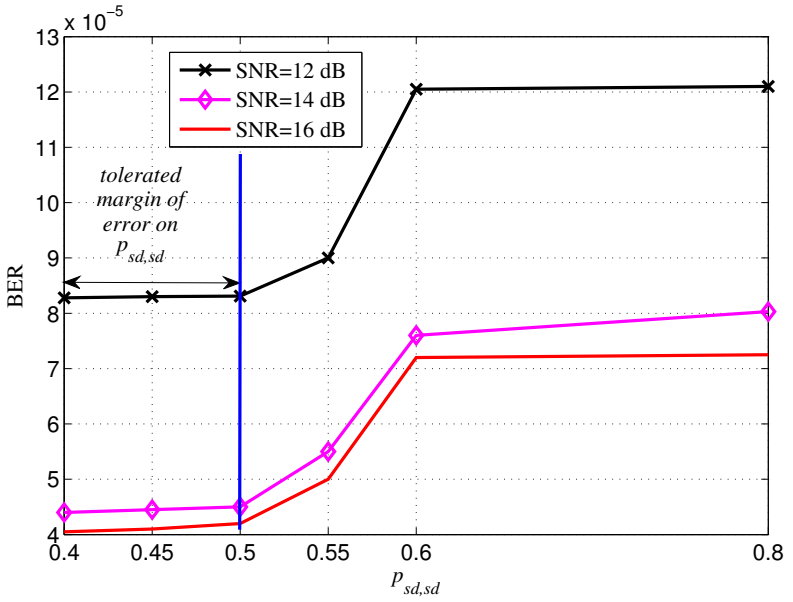

Figure 2.10: The proposed-algorithm sensitivity to errors on $p_{sd,sd}$.

This can be explained by the impact of the parameter estimate fluctuations on the signal estimate. Indeed, to compute the estimated state vector we make a soft decision between the output estimates of each GM-EKF. The contribution of each GM-EKF on the final estimate is weighted by the mode probability. The contribution of the m_{ld}-GM-EKF is 1/2 at the steady state step in the 2-mode IMM, but it is equal to 1/3 at the steady state step for the 3-mode IMM. This can reduce the amount of fluctuations on the parameter estimation

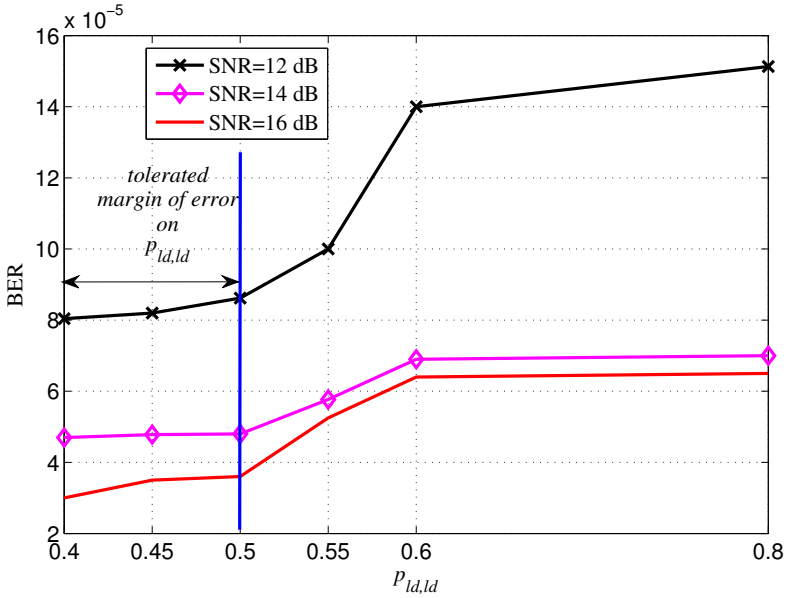

Figure 2.11: The proposed-algorithm sensitivity to errors on $p_{ld,ld}$.

which has a direct impact on the BER since we deal with the joint estimations of the CR-PA input samples and the Volterra parameters.

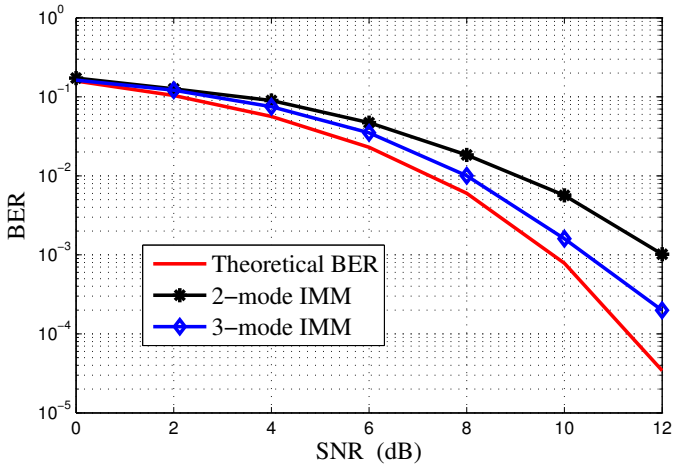

Figure 2.12: BER comparison between the 2-mode IMM and the 3-mode IMM applied to a toy example. The TPM used for the 3-mode IMM is (2.22).

- Let us now study the second scenario

Here, we suggest comparing a m_{st}-GM-EKF alone, a m_{sd}-GM-EKF alone, a m_{ld}-GM-EKF alone, a 2-mode IMM and a 3-mode IMM. The Volterra parameter tracking performance are presented in Figure 2.13 for one realization. For the sake of clarity, we present only one

parameter in the figure.

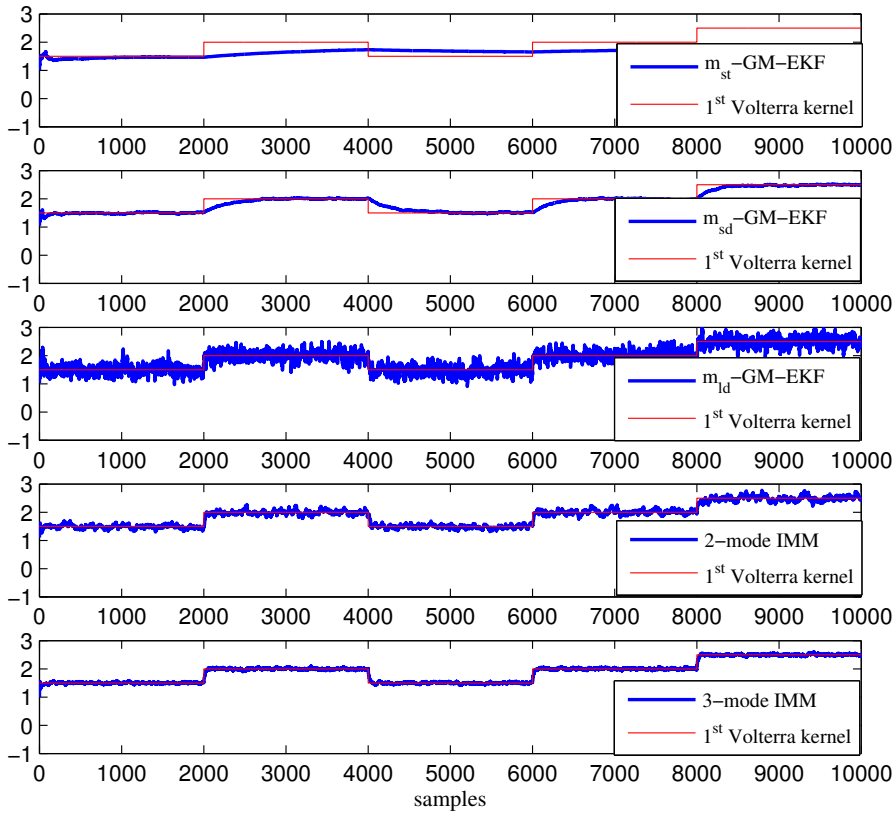

Figure 2.13: Volterra parameter tracking performance of the m_{st}-GM-EKF, the m_{sd}-GM-EKF, the m_{ld}-GM-EKF, the 2-mode IMM and the 3-mode IMM.

When only using a m_{st}-GM-EKF, the Volterra parameters can be estimated at the beginning but we cannot track their variations. Using small dynamics with a m_{sd}-GM-EKF, the tracking can be done but it requires several observations. Using a m_{ld}-GM-EKF, the parameter tracking is faster but the estimate fluctuates much. The time spent to reach the new Volterra value and the uncertainties on the estimates increase the BER. These uncertainties are considerably reduced with a 2-mode IMM. When using a 3-mode IMM, we obtain both the convergence rapidity and the parameter estimate accuracy. This can also be seen in Figure 2.14 and Figure 2.15. Indeed, the 3-mode IMM provides the best parameter estimates in terms of mean square error (MSE). This impacts the BER performance (Figure 2.16) calculated at the switching period (or the parameter estimation convergence period), assumed to be equal to 50 samples in our simulations.

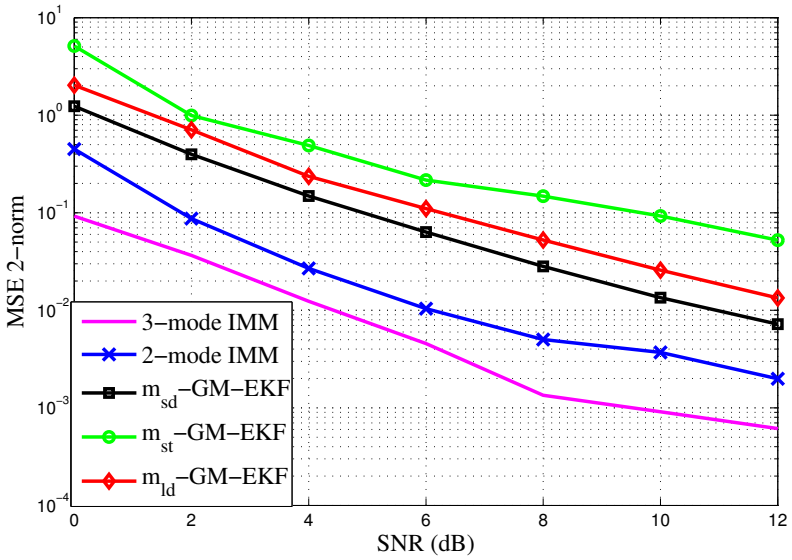

Figure 2.14: Mean square error comparison of the parameters estimation. The used norm is the 2-norm.

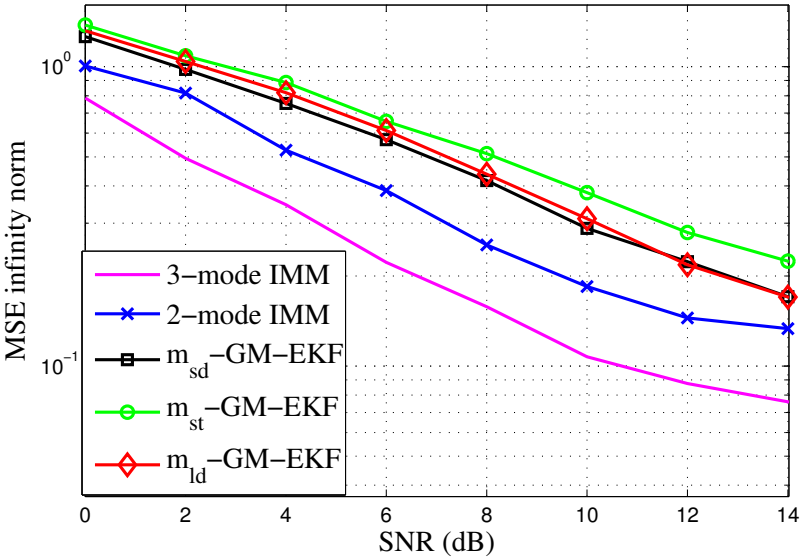

Figure 2.15: Mean square error comparison of the parameters estimation. The used norm is the infinity norm.

The BER performance of m_{ld}-GM-EKF is better than m_{st}-GM-EKF but worse than m_{sd}-GM-EKF. However, the m_{ld}-GM-EKF has the highest model noise covariance and the m_{st}-GM-EKF has the lowest one. This means that increasing the covariance of a single GM-EKF improves the BER but to a certain limit. By inserting a second GM-EKF such as the case of the 2-mode IMM, the SNR decreases by 1dB when the BER is equal to 10^{-3}. It

decreases by 2dB when the three GM-EKF are combined together in the 3-mode IMM.

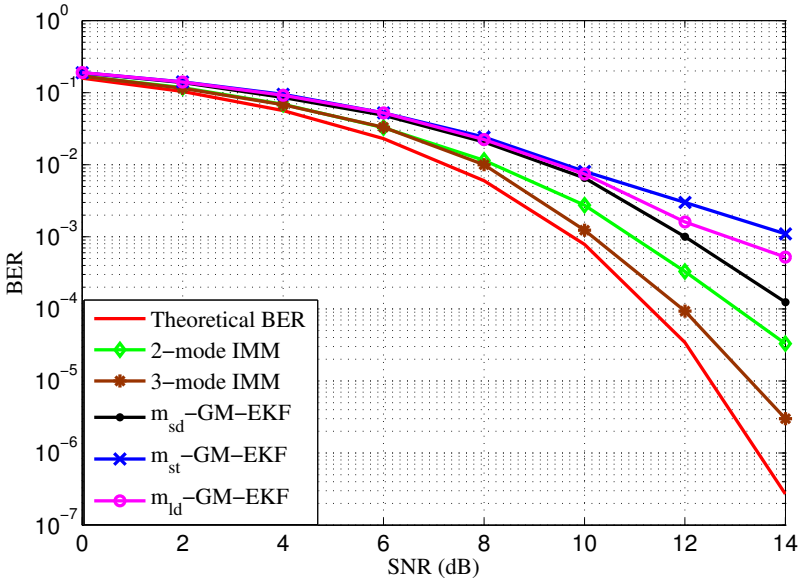

Figure 2.16: BER comparison applied to the second scenario.

Figure 2.17 illustrates the BER evolution within sampling time for a SNR equal to 12dB. Here, the 3-mode IMM provides a very small BER in the outer of the switching period which is not the case of the 2-mode IMM and the m_{sd}-GM-EKF. This can be explained by the weight of the m_{ld}-GM-EKF in each IMM algorithm (see Figure 2.18).
Indeed, the m_{sd}-GM-EKF cooperates with the m_{ld}-GM-EKF at the switching period and with the m_{st}-GM-EKF outer. As shown in Figure 2.18 the contribution of the m_{ld}-GM-EKF output in the final estimate is less or equal to 1/3 for the 3-mode IMM, while it is, approximately, 1/2 for the 2-mode IMM. Decreasing the m_{ld}-GM-EKF weight in the final estimate reduces the instantaneous fluctuations of the parameters estimate depicted in Figure 2.13.

- Let us look now at the third scenario

When what is designed at the receiver does not exactly correspond to what is done at the transmitter, the MSE is slightly higher. See Figure 2.19. Nevertheless, the IMM is still able to track the Volterra parameters and to estimate the input signal. See Figure 2.20. The BER is slightly deteriorated but still close to the theoretical curve.

In this section, we study the limits of the proposed algorithm when the TPM is predefined and a single-carrier modulation is used. First, the sensitivity of the algorithm to errors made when calculating the TPM is studied. Then, the sensitivity to the number of the modes is

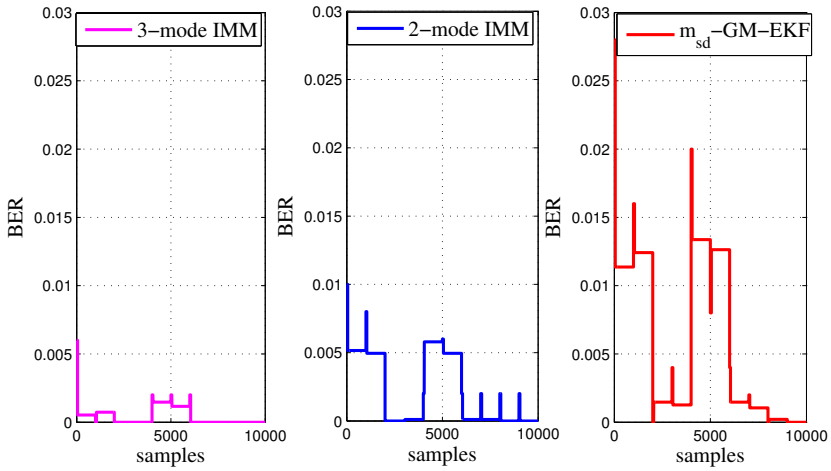

Figure 2.17: Comparison of the BER evolution within sampling time $SNR = 12dB$.

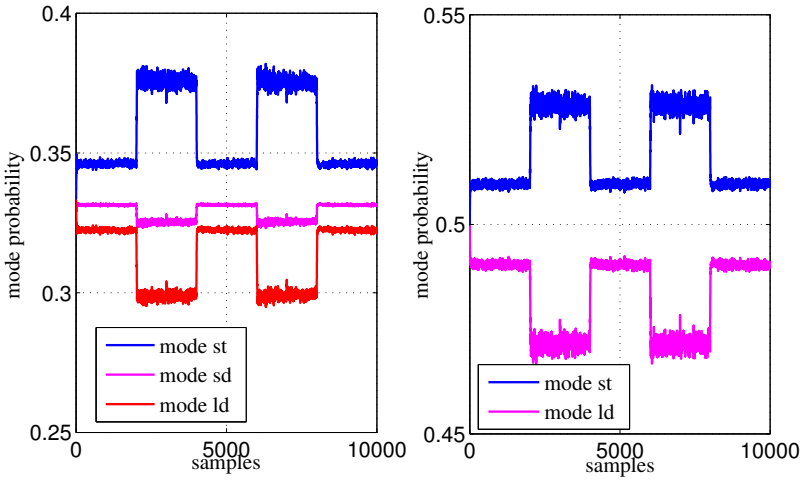

Figure 2.18: Mode probability evolution within sampling time for 3-mode IMM (left) and 2-mode IMM (right) $SNR = 12dB$.

evaluated. And finally, the case when the predefined Volterra kernels dynamics number at the receiver is different from the ones at the transmitter.

2.4 Conclusions

In this chapter, we propose a digital post-distortion algorithm based on an IMM structure for a CR-PA in an uplink communication. We started with a single-carrier system. The post-distorter algorithm is an IMM structure combining a bank of GM-EKFs. The simulation results concerning this case deal with the sensitivity of the proposed algorithm to the fixing

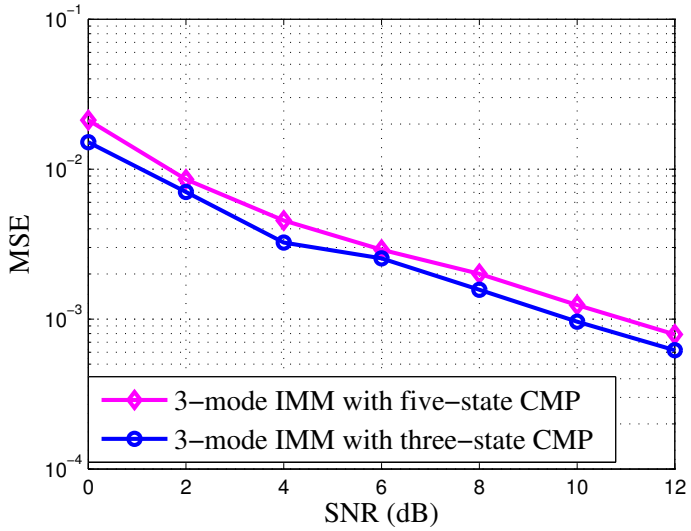

Figure 2.19: MSE-based analysis of the proposed algorithm sensitivity to the number of the CMP dynamics.

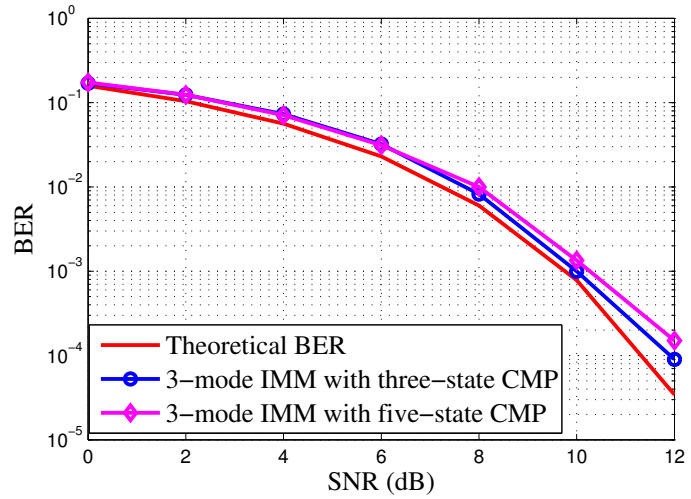

Figure 2.20: BER-based analysis of the proposed algorithm sensitivity to the number of the CMP dynamics.

errors of the TPM when considering a three-state Markov chain. In addition, we analyze, in terms of BER, the sensitivity of the proposed digital post-distorter to the number of GM-EKFs chosen for the IMM algorithm.

The originality of our approaches is to transmit a non-linearity amplified signal without any filtering and then to design a corresponding post-distorter at the receiver. This can be done in a CR system. One of our motivations is to reduce the consumed power at the mobile terminal. Indeed, compared to other linearization techniques usually implemented at the transmitter (i.e. digital pre-distortion, cartesian feedback, polar loop, etc.) the whole

linearization processing in our case is performed at the receiver. As a consequence, this is a strong advantage as in uplink the transmitter is battery driven unlike the receiver. Concerning the estimation algorithm itself, its complexity depends on the number of the EKFs. For a single-carrier modulation scheme, a i-mode IMM combines $i \times q$ EKFs.

In this chapter, the theoretical concept of the proposed post-distorter was validated by system level simulations. The next chapter deals with the measurement-based validation. The measurements are carried on a broadband commercial PA.

Chapter 3

About the measurement-based validation of the proposed digital post-distortion

Contents

3.1 Introduction

In this chapter, we validate the theoretical performance of the proposed post-distortion by means of experimental results. Then, we answer some remaining questions related to its relevance:

- how to fix the Volterra non-linearity order and memory depth?
- how to obtain the baseband observation z_k which carry the footsteps of the PA output signal spectrum components?
- does the algorithm have the same BER performance when applied on measured signals as it is in system level simulations?
- is the proposed approach really able to track the CR-PA behavior variation over time?

For this reason, this chapter presents experimental setup and results for verifying the performance of the digital post-distorter. The experimental investigations are carried out for a wideband power amplifier (model AAMCS-AMP) from MACOM. The PA is driven into saturation while the performance without and with the proposed post-distortion are investigated. Different scenarios of the test signals are considered. The measured results verify the feasibility of the proposed post-distortion and demonstrate close performance to the simulation-based results. Therefore, this chapter is organized as follows:
The second section deals with the PA characterization while in the third section the test bench used to validate the proposed technique is detailed. Then, the measurement results are discussed.

3.2 PA characterization

3.2.1 Test bench

To characterize the DUT, we use the test bench depicted in Figure 3.1. It consists of:

- a signal generator : Agilent E4428C ESG Analog Signal Generator,
- a wideband attenuator, gain=-30dB,
- a power meter: Power meter HP 435A,
- a DC power supply: Agilent E3646A.

- the DUT: Power amplifier AAMCS-AMP-300M-3000M,

  As it can be seen from the previous chapter, the proposed algorithm is validated by system level simulations. To confirm these results by measurement, we use a wideband PA to be able to vary the carrier frequency of the PA input signal. This makes it possible to study the PA behavior variation and evaluate the proposed approach performance. Hence, the used DUT is able to cover frequencies from 300MHz to 3GHz and delivers an output power of 26 dBm at 3 GHz. A gain of 26 ± 2dB is achieved over the entire band and the PA consumes 3.78W from a 15V DC supply. The saturation power is 30dBm and the 1-dB compression point (CP1) is 26dBm.

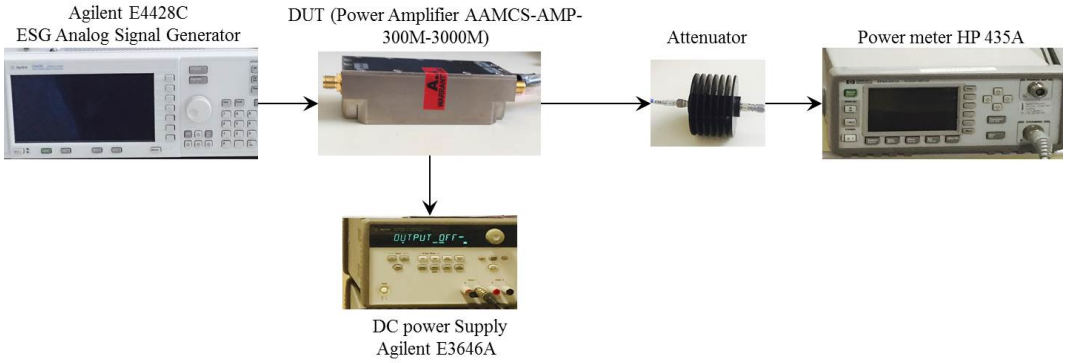

Figure 3.1: PA characterization test bench.

At the signal generator level, the PA input signal characteristics are defined:

- the power P_{in},
- the carrier frequency f_c and bandwidth B_w,
- the shaping filter: rectangular filter.

3.2.2 PA Characterization

3.2.2.1 AM-AM and AM-PM characterization

The AM-AM and AM-PM characteristics of the PA are shown in Figure 3.2 and Figure 3.3. The 1 dB compression appears at the output power of 27 dBm and input power of 0 dBm at $f_c = 2.8GHz$. At the same point, the AM-PM modulation is 0.25 degrees. In addition, at this point, the PAE efficiency is equal to 14%. Although the DUT has high gain (Figure 3.4) and P_{out} (Figure 3.2) that enable it to respond to the high level requirements of the majority of the communications standards, its PAE is very low. This is due to the consumed power.

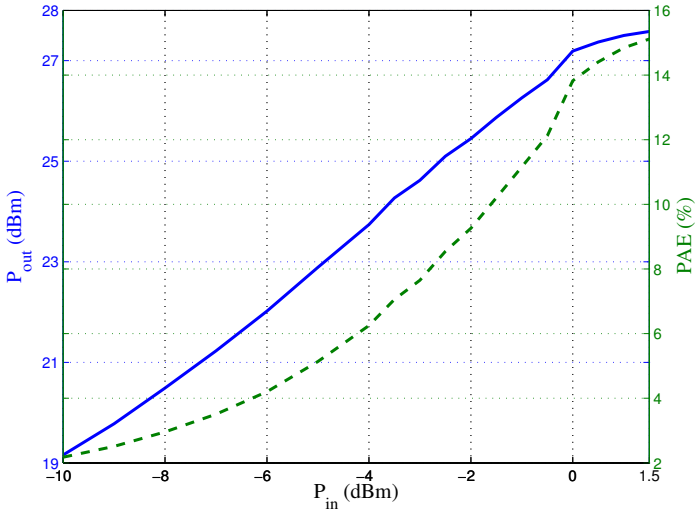

Figure 3.2: Measured AM-AM characteristic for $f_c = 2.8GHz$.

Indeed, the needed DC voltage V_{DC} and current I_{DC} to operate the PA are respectively 15V and 0.252A. In other words, the DC power supply of the PA is 3.78W which is quite high for an amplifier implemented in a mobile terminal. However, as our purpose is to validate a theoretical assumption about a wideband PA behavior, this DUT is valid to confirm the proposed modeling and post-distortion technique relevance.

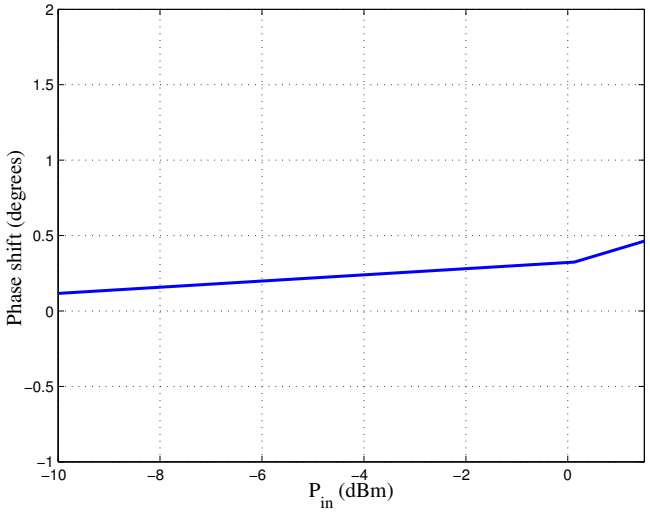

Figure 3.3: Measured AM-PM characteristic for $f_c = 2.8GHz$.

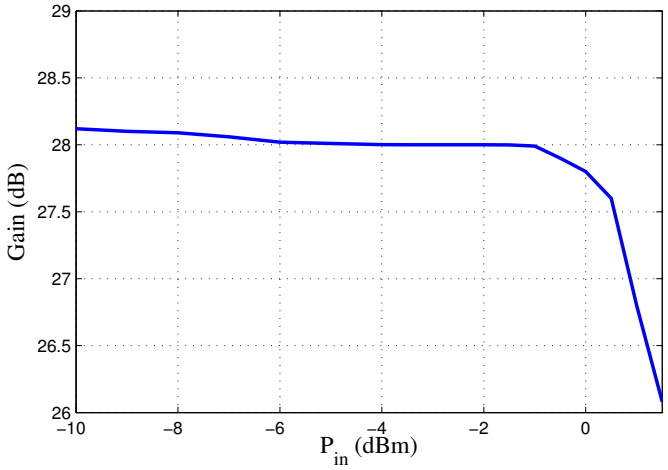

Figure 3.4: Measured Power gain for $f_c = 2.8GHz$.

3.2.2.2 Frequency and input power dependency

Using the same test bench, our purpose is to characterize the PA behavior dependency on the carrier frequency. For this reason, we show a three dimension representation of the P_{out}, PAE and gain in Figures 3.5, 3.6 and 3.7.

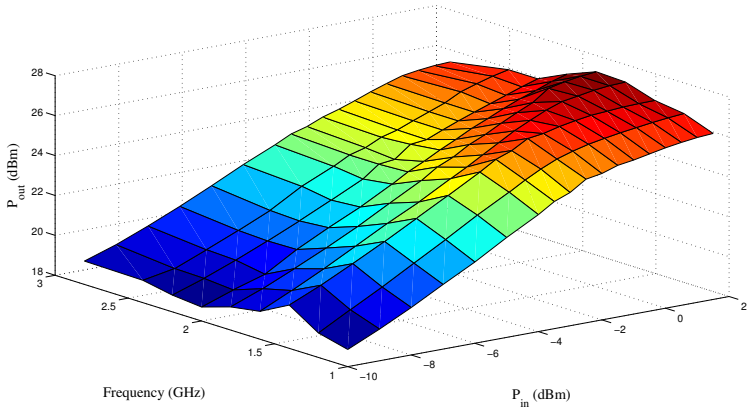

Figure 3.5: AM-AM characteristic evolution with carrier frequency and P_{in} variations.

On the one hand, the PA performance in terms of P_{out}, PAE and gain is maximal over narrower wideband, between 1.6GHz and 1.8GHz. On the other hand, in Figure 3.5, one can notice that the input compression point at 1dB, OCP1 varies with frequency. This means that depending on the carrier frequency the PA non-linear amplification starts more or less early in terms of input power. When we take a look at the values variation of P_{out}, PAE and gain, we verify that the PA behavior varies when the carrier frequency is changed which is coherent with our assumption about the PA behavior variation with carrier frequency.

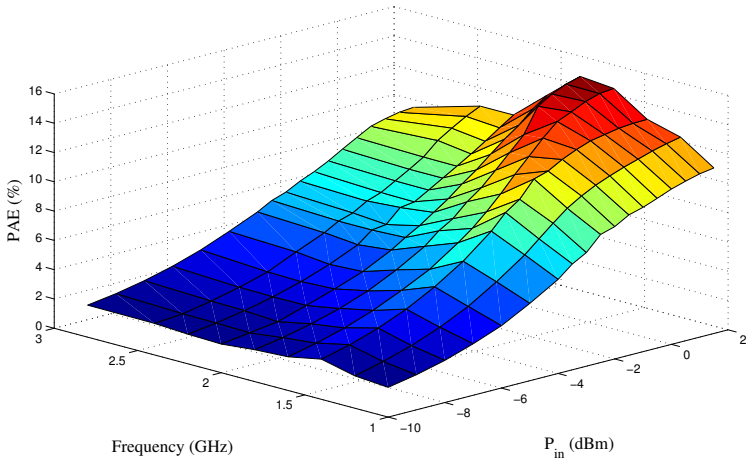

Figure 3.6: PAE characteristic evolution with carrier frequency and P_{in} variations.

This result has been heavily studied and verified in literature [Liu04], [Lan08], [Tar15] and [Gha10]. Indeed, the PA behavior change not only as a function of the carrier frequency but also as a function of the input signal bandwidth.

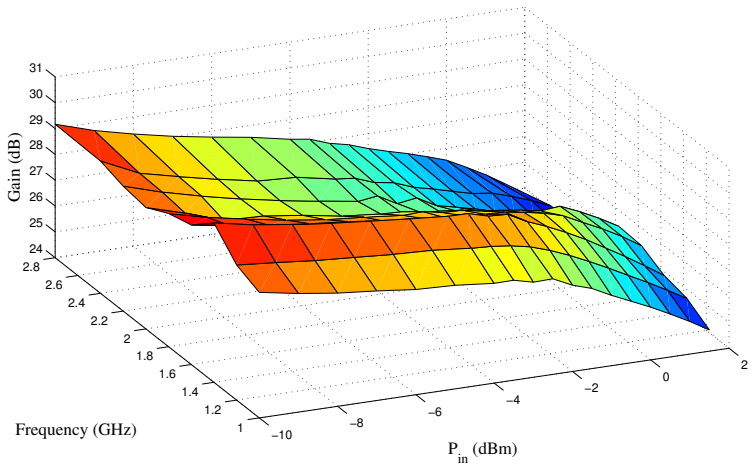

Figure 3.7: Power gain characteristic evolution with carrier frequency and P_{in} variations.

3.2.2.3 Non-linear characterization

In order to characterize its non-linear behavior, for this section and the upcoming ones, the RF signal power level should be high enough to drive the DUT to work in its non-linear zone at the CP1 which means that the P_{in}=1.5dBm.

In this section, the classical two-tone analysis is used to quantify the memory effects in power amplifiers. For this reason, a two-tone input signal with 2 MHz of frequency spacing

is injected to the PA. This frequency spacing is the highest value that can be reached by the signal generator. As shown by the table 3.1 and Figure 3.8, the first tone frequency f_1 is 2.799 GHz whereas the second tone frequency f_2 is 2.801 GHz. One can notice that in addition to the amplified versions of f_1 and f_2, spurious frequencies appear at other frequencies. To identify these frequencies, we calculate the intermodulation products. As mentioned in table 3.1, some of these frequencies correspond to the 3^{rd}, 5^{th}, 7^{th}, 9^{th} and 13^{th} order intermodulation products.

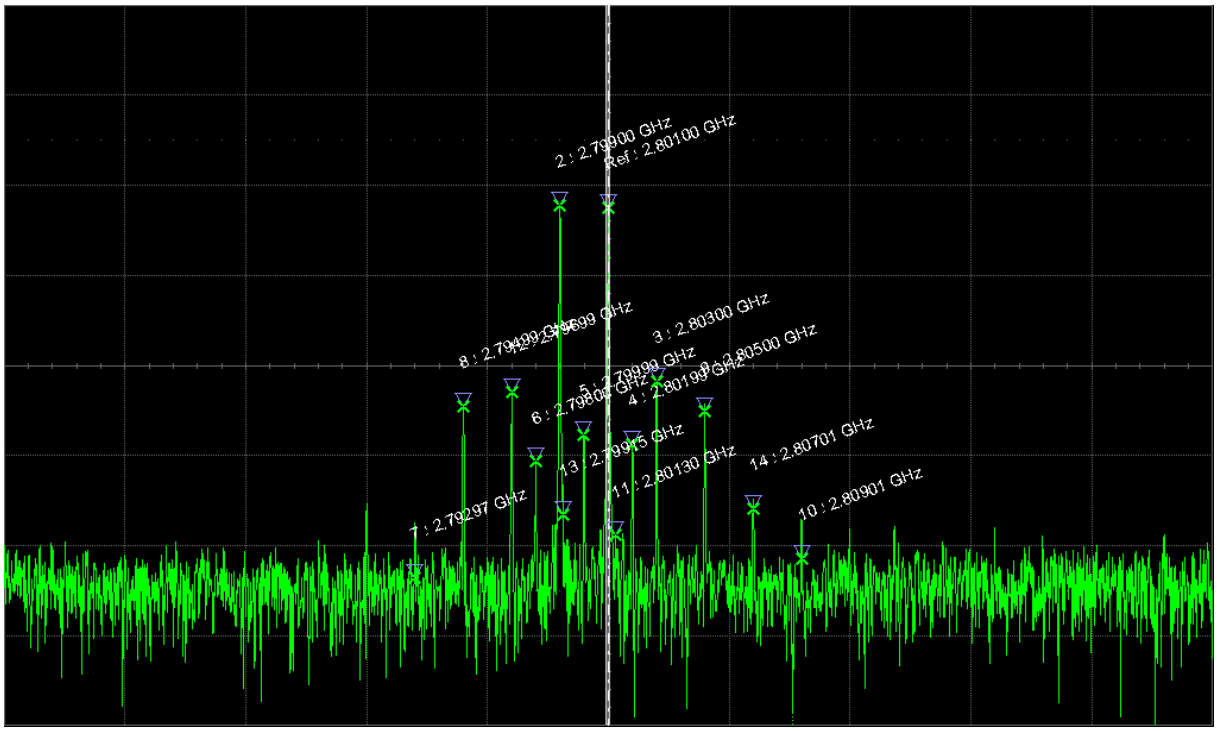

Figure 3.8: PA output spectrum for a two tone input signal.

The higher the intermodulation order is, the lower the power level is. For this reason, it would be wise [Afs12] to take this into account when trying to model the PA. Indeed, at the set of frequencies, marked as $\{3, 6, 7, 10, 11, 13, 14\}$, the ACPR is lower than -40 dBc. For instance, this is in concordance with the LTE mask [Hon13]. Therefore, we consider these products insignificant and we do not take them into account when modeling the DUT. This result can be exploited when defining the memory depth M and the non-linearity order P of the Volterra model. These assumptions are based on a modeling metric called the adjacent channel error power ratio (ACEPR) defined in [Isa06]. This metric is computed to illustrate the adopted model reproduction of the PA ACPR. It is expressed as follows:

$$ACEPR = \frac{\int_{adj} |E(f)|^2 df}{\int_{ch} |Y(f)| df} \tag{3.1}$$

Table 3.1: PA output spectrum for a two tone input signal

Markers	Intermodulation product	Frequency (GHz)	Amplitude (dBm)
Ref	f_2	2.80100	16.7
2	f_1	2.79900	17.1
3	$2f_2 - f_1$	2.80300	-21.8
4	frequency offset	2.80199	-35.8
5	frequency offset	2.79999	-33.8
6	$5f_1 - 4f_2$	2.79800	-40.7
7	$4f_1 - 3f_2$	2.79297	-65.5
8	$3f_1 - 2f_2$	2.79499	-27.4
9	$3f_2 - 2f_1$	2.80500	-28.6
10	$5f_2 - 4f_1$	2.80901	-65
11	$7f_2 - 6f_1$	2.8130	-60.2
12	$2f_1 - f_2$	2.79699	-25.2
13	$7f_1 - 6f_2$	2.79905	-63.6
14	$4f_2 - 3f_1$	2.80701	-62.2

where $E(f)$ is the Fourier transform of e_n which expresses the difference between the PA output signal y_n^{meas} and the PA model output signal y_n^{mod}:

$$e_n = y_n^{meas} - y_n^{mod} \tag{3.2}$$

The more similar a calculated output signal of a model is to an experimental one, the lower the ACEPR value is, i.e., the model is closer to the PA behavior. By fixing a threshold equal to 3 dB for this metric, it becomes possible to neglect some intermodulation products without them the ACEPR still below the threshold. This two tone analysis is done around several frequencies as it is shown by the Figure 3.9. Based on the results of these tests and the fixed ACEPR threshold, the non-linearity order is defined for each frequency. The highest order for all the frequencies is 5. For this reason, the non-linearity order P is fixed at 5. Then, the memory depth is defined to be equal to 2.

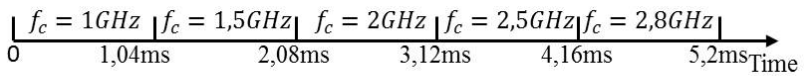

Figure 3.9: Frequency change over time.

Afterwards, the Volterra kernels are identified using a Kalman filter as its is shown by Figure 3.10. The estimated coefficients are only used to be able to compare system level simulation based results to the measurement based results. Note that, when applying the

digital post-distortion on the measured PA amplifier output signal, the obtained *a priori* information about the Volterra kernels is not used. The Volterra kernels are assumed to be unknown and they are jointly estimated with the PA input samples as it has been proposed in chapter 2.

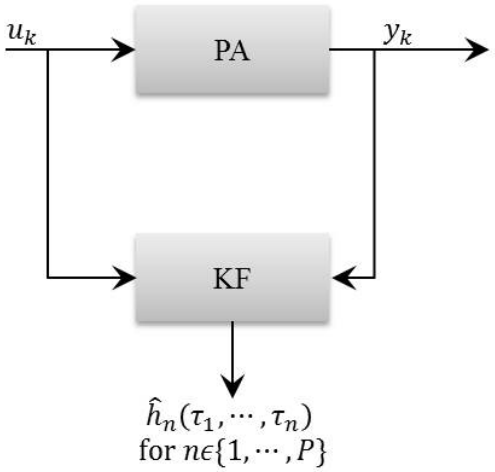

Figure 3.10: The PA model-parameters extraction procedure.

The extracted Volterra kernels are presented in Figure 3.11. For the sake of the figure clarity, we only present some of them. As long as the carrier frequency of the PA input signal varies, the values of the Volterra kernels vary. The test bench used to vary the input signal and retrieve the PA input and output signal are explained in the upcoming section 3.3.1.

Let us now check to what extent the obtained Volterra model is able to reproduce the PA power spectral density (PSD). The PSD of the PA output and the Volterra model output are then calculated to verify the relevance of the above procedure. For this reason, a single-carrier signal is injected to the PA, the characteristics of which are:

- a QPSK constellation,
- 5MHz of bandwidth W,
- and a carrier frequency f_c fixed at 2 GHz.

3.3 Performance analysis and evaluation

3.3.1 Test bench

The test bench in Figure 3.12 is built to experimentally validate the proposed digital post-distortion. The test bench consists of the DUT broadband PA (model AAMCS-AMP from

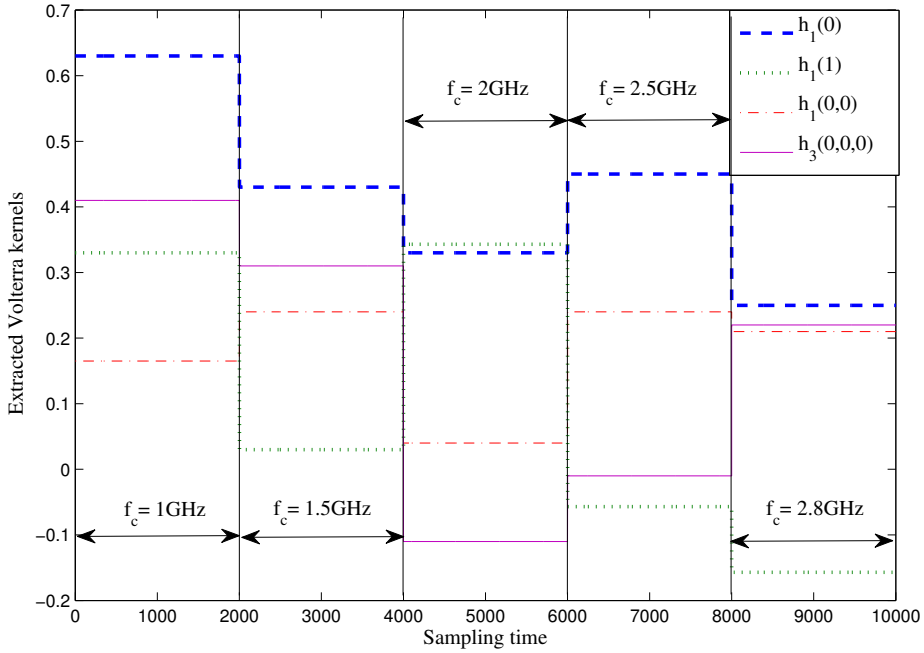

Figure 3.11: Some of the extracted Volterra kernels.

MACOM) defined in 3.2.1, a serial signal generator (SSG), a digital oscilloscope (DO), a DC supply and a personal computer (PC) on which the Labview software is installed. In this section, a detailed review about these devices and software characteristics is given and then the experimental procedure is explained.

3.3.1.1 Serial signal generator

The SSG used in the test bench is HP ESG-D Series E4433B signal generator. It has random waveform generator option. This option provides to load baseband signal to the generator and it carries the baseband signal to the desired carrier frequency. Therefore, in the test bench the SSG is used:

1. to receive the baseband signal,
2. to convert it to analog like digital to analog converter,
3. to modulate it around the desired carrier frequency.

The baseband input data is generated using *Matlab*, these I/Q signals are loaded to the SSG. In the arbitrary waveform generator option, the SSG has sampling clock and reconstruction

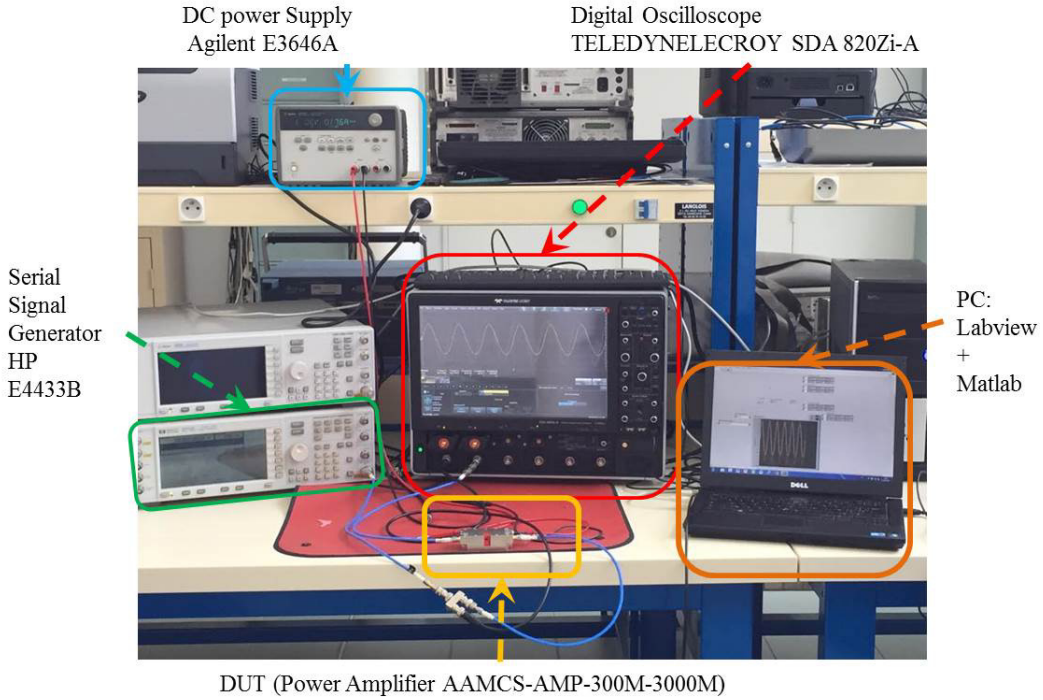

Figure 3.12: Test bench.

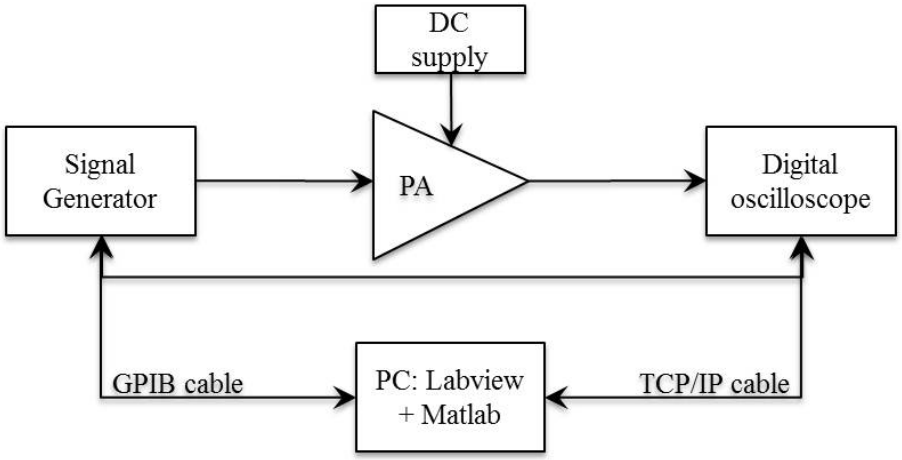

Figure 3.13: Bloc diagram of the used test bench.

filter. Their settings are fixed to be the same as those used to generate the baseband signal using *Matlab*.

3.3.1.2 Digital oscilloscope

A digital oscilloscope (DO) is an oscilloscope which can store and analyze a digital signal. It typically provides the advanced trigger, storage, display and measurement features. The analog input signal is sampled and then converted into a digital record of the signal amplitude

at each sample time. In addition, the user can define the set of elements such as the RF filter, the mixer, the I/Q demodulator and the ADCs, that make it possible to obtain the baseband version of the digitized signal.

3.3.1.3 Labview

The Labview (Laboratory Virtual instrument Engineering Workbench) software is a graphical programmatic software used to control and to simulate the measurement instruments. Using Labview, it becomes possible to set up a test bench controlled by a PC. The Labview file extension is VI (like Virtual Instrument). It consists of two windows:

1. to display the measurement results and curves,

2. to detail the program different items and functions represented by a block diagram.

Therefore, we use the Labview software in our test bench to automatically manage and change the measurement settings. In addition, it makes it possible to retrieve the digitized signals by the DO and save them in Excel format. This makes it possible to post process the samples using *Matlab*. Finally, it allows us to recapture the displayed curves on the DO screen.

3.3.1.4 Experimental procedure

The test bench described by (Figure 3.12 and Figure 3.13) is designed to be fully automatic by using Labview software. See Figure 3.14. The complex envelop of the transmitted signal is generated then modulated by the carrier frequency at the SSG. Then, the signal is transferred via a GPIB cable to the DO. The SSG is also connected to the PA which is linked to the DO. The latter is connected to the PC via an ethernet cable. The PA input and output signals are digitized at the DO. These signals are retrieved at the PC and saved in Excel format using the Labview software. The digital post-distortion, written in Matlab, is carried on the PA output signal saved in these files.

The PA input signal corresponds to the one defined in 3.2.2.3.

Let us now describe the way according to which the test bench is controlled by Labview. As it is mentioned above, the communication between the DO and the PC is an ethernet communication. When initializing this communication, it becomes possible to set the DO parameters. Indeed, the DO has several channels and each channel is associated to a given entry. Hence, the first channel is connected to the SSG and the second channel is linked to the PA output. This has to be specified from the PC.

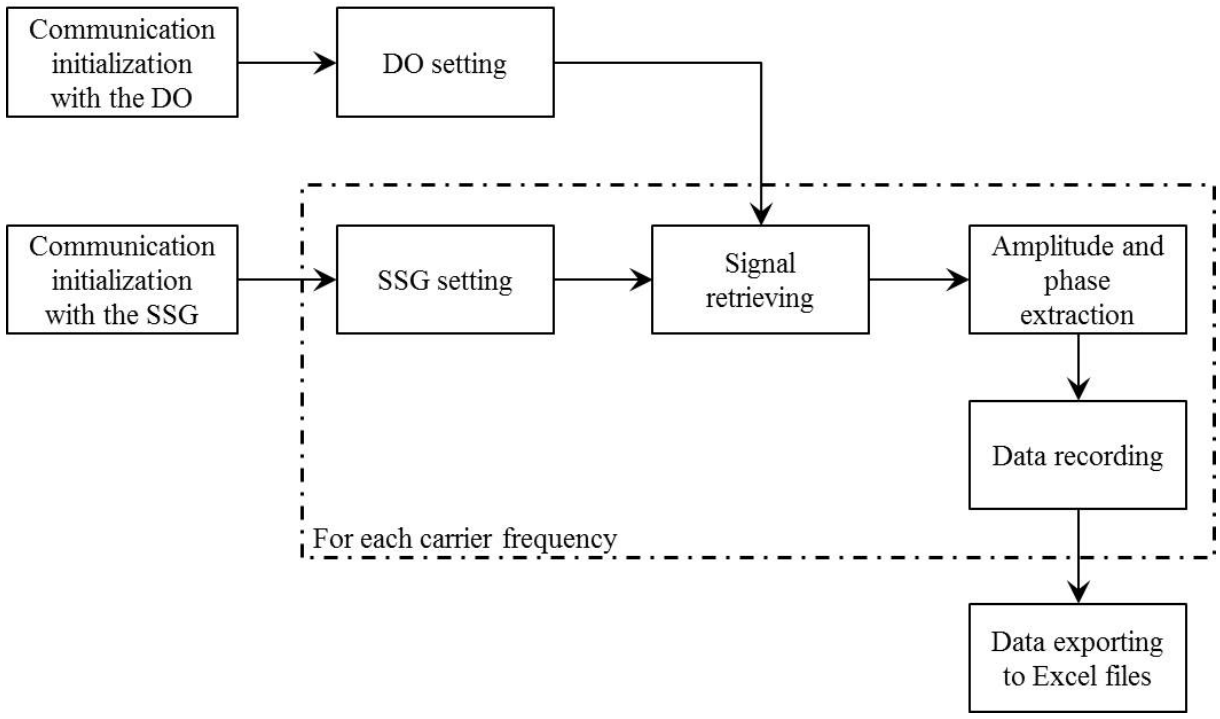

Figure 3.14: Bloc scheme of the Labview VI used to monitor the measurement instruments.

In parallel, the communication between the PC and the SSG is done using USB-GPIB connector. When initializing this communication it becomes possible to send a set of commands to the SSG in order to generate signals with given amplitude and frequency:

1. the first command indicates the generated signal frequency,

2. the second one indicates the generated signal amplitude,

After this, the signals are recaptured. These signals are the input signals of each DO channel. The extracted data by the DO are processed and stored thanks to several commands. These commands make it possible to manipulate data and tables. Using these tables, the PC rebuild the curves figures of the DO. Then, it extract the signal amplitudes and phases. These data are then stored in tables and exported to Excel files.

Summarizing, we present the programming algorithm used for the performed graphical program (Algorithm 2).

3.3.1.5 A summary about the used test bench to evaluate the algorithm performance

The test bench can be summarized, as it is depicted in Figure 3.15, as follows:

Algorithm 2 Graphic programming algorithm

Main program{

- Initialization of the communication with the DO,
- Initialization of the communication with the SSG,
- For frequencies from 300 MHz to 3 GHz by a step of 200 MHz:{
  - retrieve the shape of the input signal,
  - calculate the input signal amplitude,
  - extract the input signal phase,
  - retrieve the shape of the output signal,
  - calculate the output signal amplitude,
  - extract the output signal phase,
  - store the data in Excel tables,

  }
- close the communication with the DO,
- close the communication with the SSG,
- display the Excel tables,

}

- the PC generates the baseband version of the PA input signal, apply the proposed post-distortion technique and retrieve the bitmapped version of the PA output signal,
- the SSG up converts the baseband signal and modulate it by a carrier frequency,
- the DUT is the PA,
- the DO digitizes the PA output signal.

At the PC level, the digitized PA output signal is brought back to the baseband frequency using a *Matlab* program the main idea of which is based on the RF front end and receiver architecture depicted in Figure 3.16. This latter is made up by a classical RF front end for the fundamental, the 2^{nd} and the 3^{rd} harmonic on three parallel paths. For the sake of simplicity, this architecture is given for $P = 3$ in Figure 3.16. For higher orders, we need supplementary paths. This makes it possible to avoid aliasing after the analog-to-digital conversion. Then, the digital version of the signal at these frequencies is summed up and

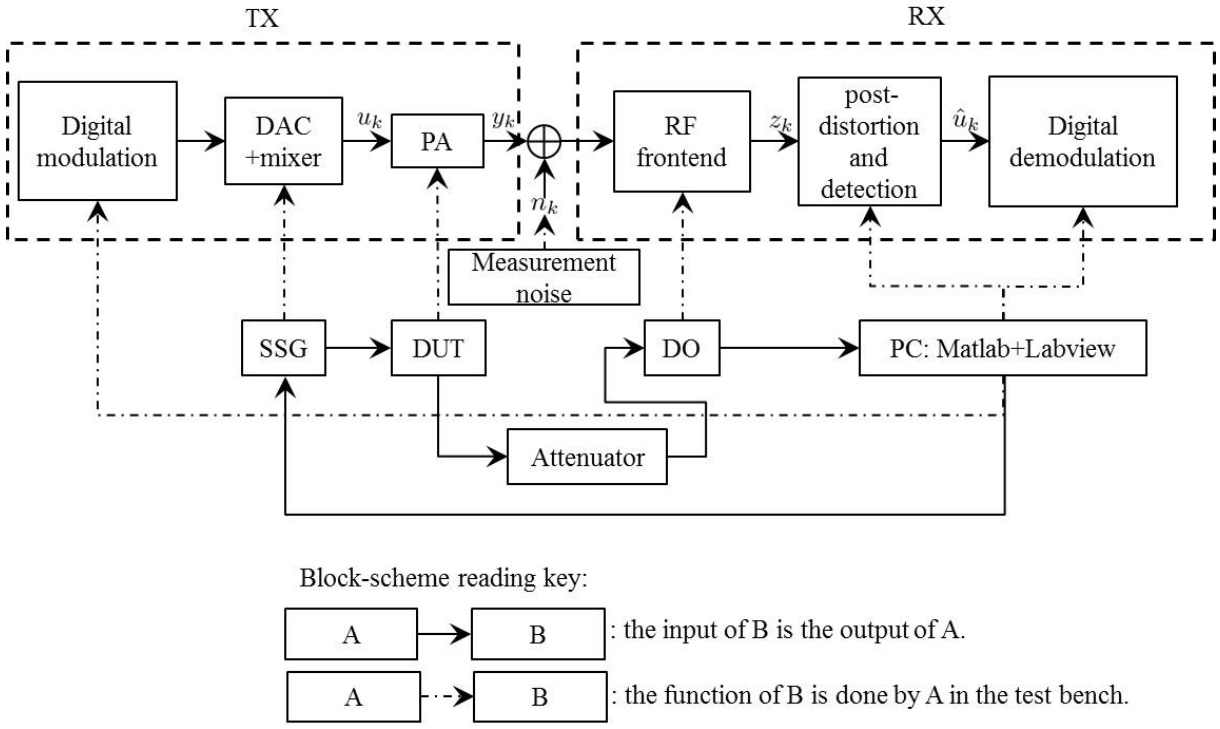

Figure 3.15: Correspondence between the system model and the test bench.

forwarded to the matched filter. Finally, the matched filter output signal is sampled at the transmitted symbol rate.

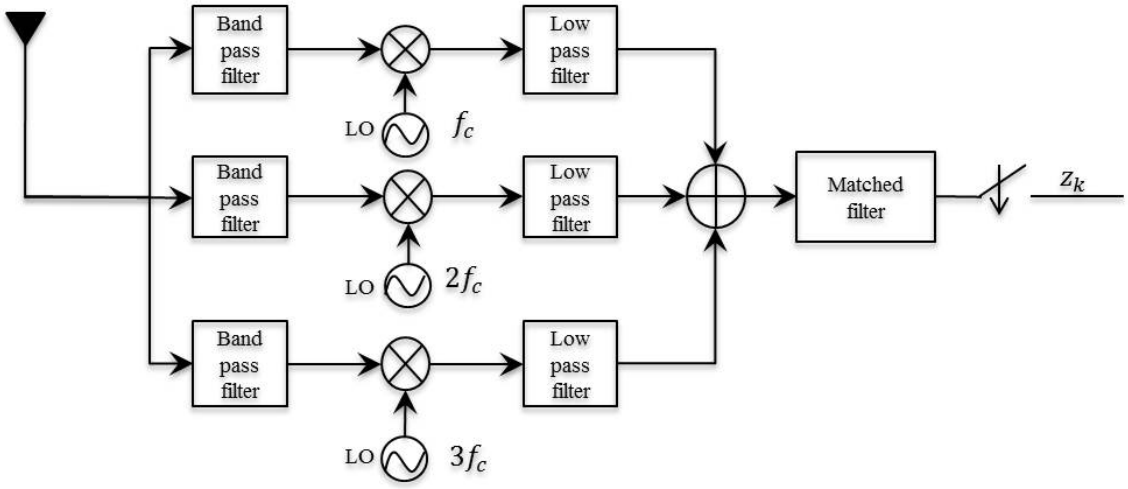

Figure 3.16: The used RF front end to restore the baseband PA non-linearities until the 3^{rd} order.

3.3.2 Measurement results

As in the previous chapter, we show that the 3-mode IMM has better performance in terms of BER and Volterra kernels tracking, it is chosen as the main structure of the post-distorter evaluated in this section. Also, we choose to *a priori* define the transition probability matrix (TPM) as we can calculate its components. In addition the generated signal is the one specified in 3.2.2.3, a single-carrier signal its characteristics are:

- a QPSK constellation,
- 5MHz of bandwidth W,
- and a carrier frequency f_c fixed at 2 GHz.

Afterwards, in the measurement result part, we first propose and analyze a sophisticated receiver architecture to obtain the baseband observation z_k. Then, in the second part, we study the algorithm performance regarding the signal constellation retrieving in terms of EVM. Finally, in the third part, the algorithm tracking of the PA behavior variation is studied.

3.3.2.1 Receiver architecture validation

Using the signal specified in 3.2.2.3, we compare the BER between several receiver architectures, one of them is depicted in Figure 3.16. Indeed, the fixed non-linearity order for the Volterra model is 5. Hence, the receiver architecture need 5 paths to be able to retrieve the spectrum around the 2^{nd}, the 3^{rd}, the 4^{th} and 5^{th} harmonic. It should be noted that the calculated BER has no sense without indicating the SNR. However, as in our case it is not possible to calculate it. Instead, some measurement parameters may help the reader to carry out similar measurements such as the sampling frequency $f_s = 1.92MHz$ and $P_{in} = ICP1$ =1.5dBm.

As this architecture (Figure 3.16) is complicated to be implemented for a high order non-linearity, we analyze the performance of the post-distortion technique when using five different receiver architectures:

1. architecture 1: When only the spectrum around the fundamental is retrieved,
2. architecture2: When, in addition to the fundamental, the spectrum version around the 2^{nd} harmonic is brought back to the baseband frequencies,

3. architecture 3: When, in addition to the fundamental, the spectrum versions around the 2^{nd} and the 3^{rd} harmonics are brought back to the baseband frequencies,

4. architecture 4 When, in addition to the fundamental, the spectrum versions around the 2^{nd}, the 3^{rd} and the 4^{th} harmonics are brought back to the baseband frequencies,

5. architecture 5: When, in addition to the fundamental, the spectrum versions around the 2^{nd}, the 3^{rd}, the 4^{th} and the 5^{th} harmonics are brought back to the baseband frequencies.

Using the first architecture the BER is quite high but it get lower when the spectrum around the 2^{nd} order harmonic is also brought back to the baseband frequencies and taken into account in the observation z_k. It also get lower when the 3^{rd} architecture is used to obtain z_k. However, the BER stay constant when the spectrum around the 4^{th} and the 5^{th} order harmonics are added. This can be explained by two facts:

1. the power around these frequencies is insignificant,

2. the variance of the estimation error which depends on the interferences caused by the spectra around the harmonics with the fundamental one is higher than the noise power.

According to the above results (illustrated in table 3.2), it might be interesting to consider the 3^{rd} architecture as it provides less complexity and better BER than the other four architectures. It should be noted that these results are related to the DUT and might not be the same for other PAs.

Table 3.2: Recorded BER for different receiver architectures

Scenario	BER
Architecture 1 (f_c)	10^{-3}
Architecture 2 (f_c, $2f_c$)	1.3×10^{-4}
Architecture 3 (f_c, $2f_c$, $3f_c$)	10^{-5}
Architecture 4 (f_c, $2f_c$, $3f_c$, $4f_c$)	10^{-5}
Architecture 5 (f_c, $2f_c$, $3f_c$, $4f_c$, $5f_c$)	10^{-5}

3.3.2.2 Comparison between simulations and measurements: About the signal constellation

The evaluation is performed by studying the constellation of the PA output signal with and without post-distortion. Then, the results are compared to the system level simulation based

results.

In Figure 3.17(a), the rotations and the dispersions of the constellation are due to the Volterra model, extracted in 3.2.2.3. In Figure 3.18(a), the rotations and the dispersions of the constellation are due to the PA nonlinearities and memory effects. As it can be seen, these distortions, of the PA and the Volterra model, are compensated by the proposed digital post-distortion (Figure 3.17(b) and Figure 3.18(b)).

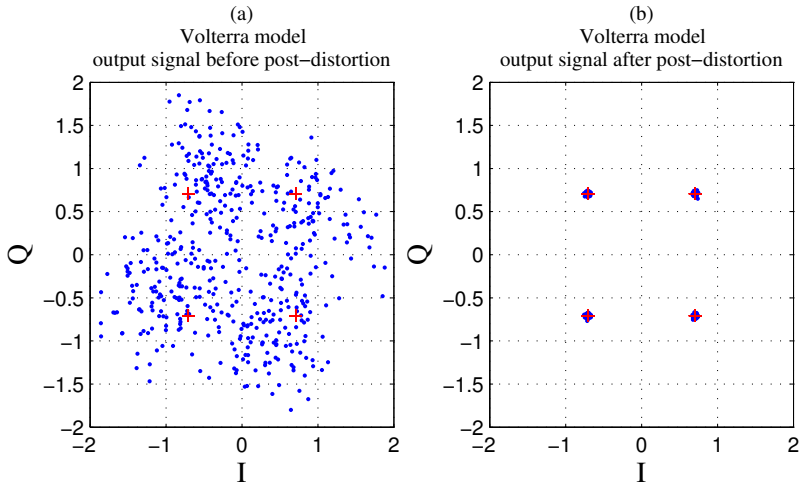

Figure 3.17: DUT output signal constellation comparison before and after post-distortion.

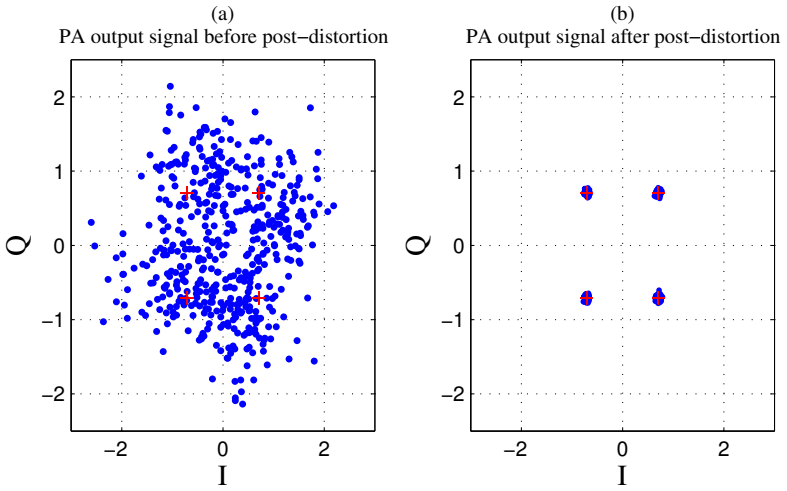

Figure 3.18: Volterra model output signal constellation comparison before and after post-distortion.

In Figure 3.19, the measured EVM is compared to the simulated one. One can notice that the measured EVM is 1% higher at $P_{in} = 1.5dBm$ than the simulated EVM, when applying the proposed approach on measured signals. This can be explained by the measurement

impairments and noise that is not taken into account by the algorithm assumptions. However, the measured EVM is still close to the simulated EVM.

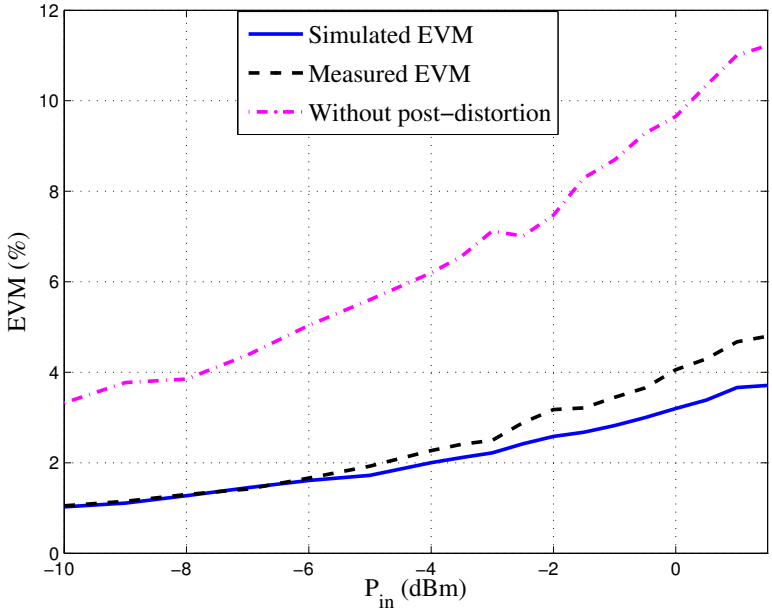

Figure 3.19: Error vector magnitude comparison between simulations and measurements.

Another common effect of the PA distortion on the constellation is shown via the eye diagrams, in Figure 3.20 and Figure 3.21, calculated at the compression point using a single-carrier signal.

The PA distortions, non-linearities and memory effects, appear as closure of the eye pattern in Figure 3.21. The eye pattern is re-expanded after the post-distortion compensation (Figure 3.22). This observation is consistent with the previous symbol-constellation results.

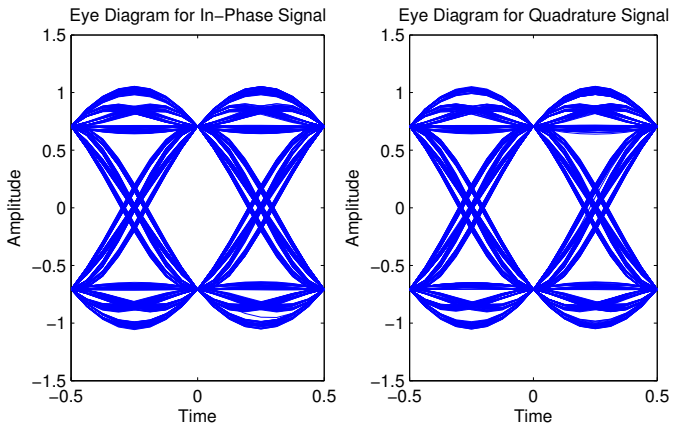

Figure 3.20: Eye diagram of the PA input-signal complex envelop.

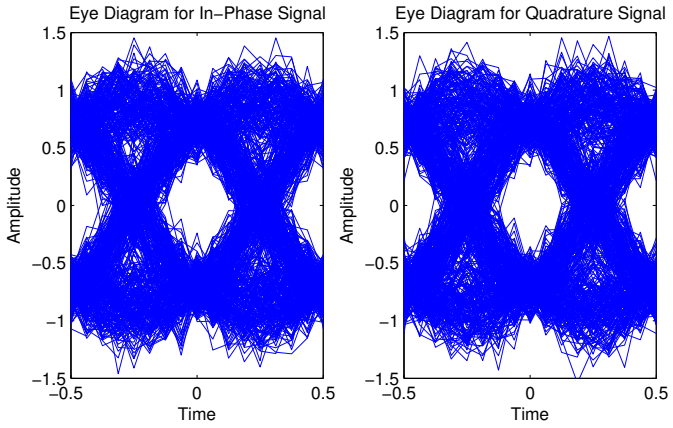

Figure 3.21: Eye diagram of the PA output-signal complex envelop.

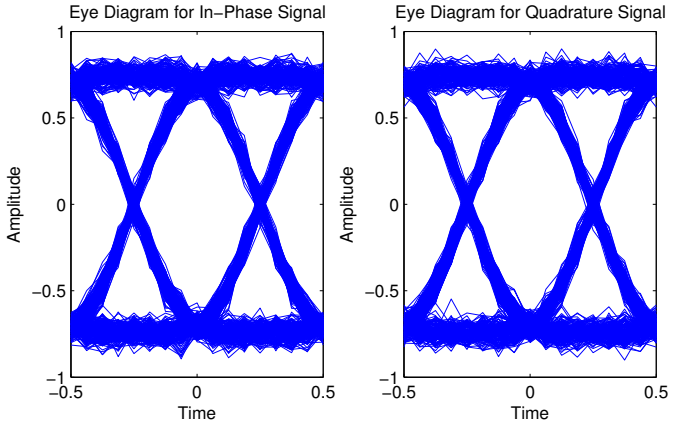

Figure 3.22: Eye diagram of the received signal after post-distortion.

3.3.2.3 Comparison between simulations and measurements: About the algorithm tracking of the PA behavior variation

Now let us evaluate the algorithm tracking of the PA-behavior change when the carrier frequency evolves. First, the carrier frequency of the PA input signal is modified several times. See Figure 3.9. Then, using the test bench defined in 3.3.1, the corresponding PA output signals are brought back to the baseband frequency using the third architecture defined in 3.3.2.1. Then, they are concatenated. Finally, the post-distortion is carried on the resulting signal. From Figure 3.23, one can notice that the PA effect on the constellation is frequency dependent as it is verified previously in section 3.2.2.2. However, the proposed method is able to track the PA behavior variation to give a better estimate of the transmitted signal. See Figure 3.24 which corresponds to the symbol constellation from 0 to $T_0 = 5.2$ms. This confirms the algorithm capability to track the PA behavior variation when the carrier frequency evolve over time.

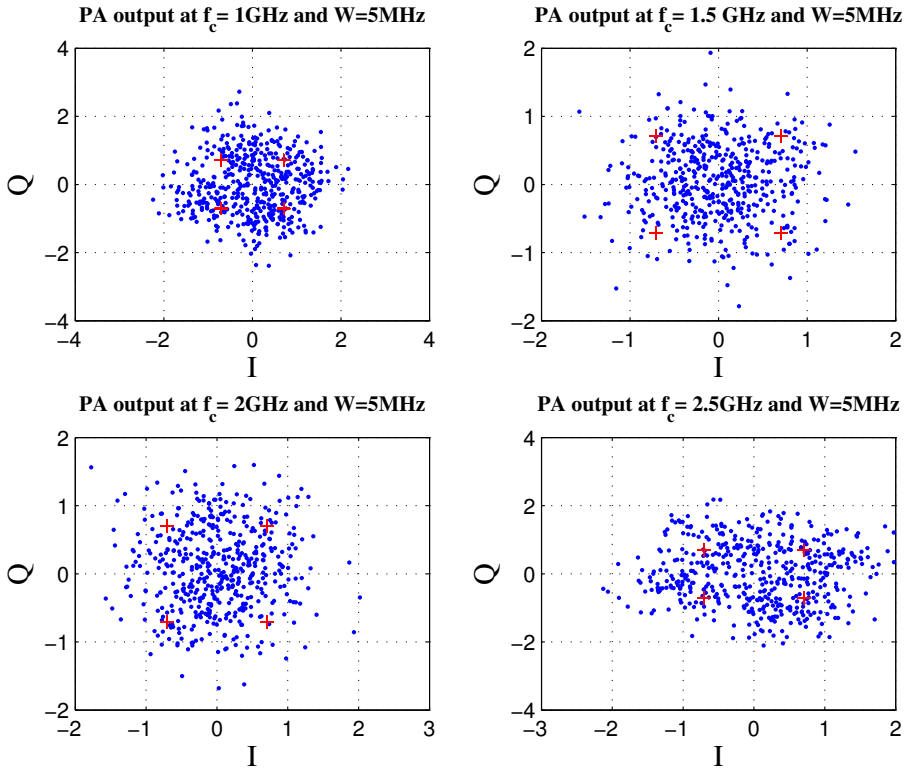

Figure 3.23: Frequency impact on the PA distortion of the signal constellation at P_{in} = ICP1.

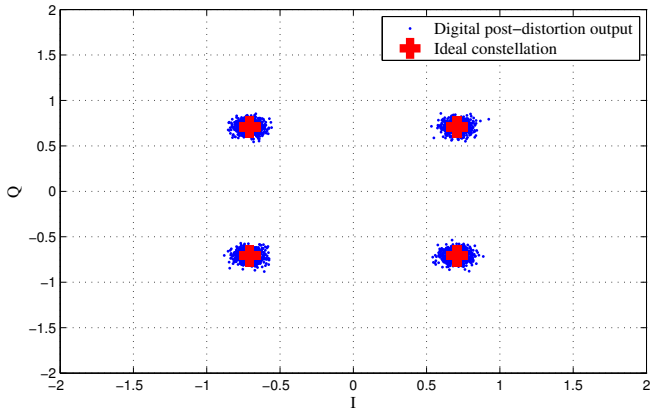

Figure 3.24: Received symbol constellation after the proposed digital post-distortion at P_{in} = ICP1.

In Figure 3.25, the estimated Volterra kernels by the post-distorter are presented as well as the extracted Volterra kernels by the procedure explained in 3.2.2.3. As it is shown, the post-distorter is able to track the CR-PA behavior variation over time due to the frequency variation. This is in line with simulation results presented in the previous chapter.

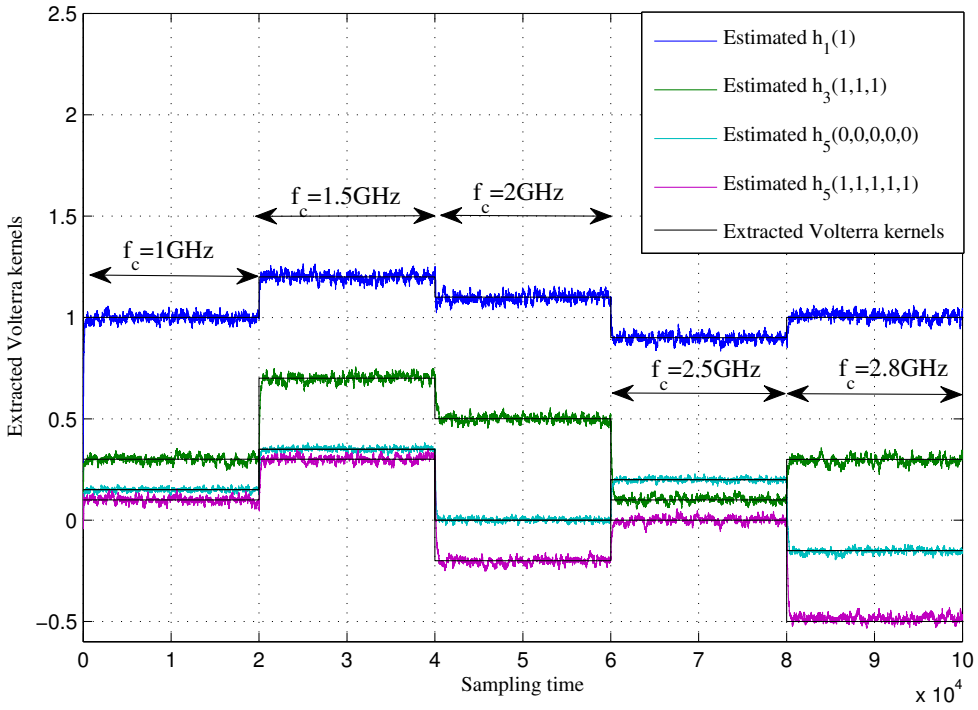

Figure 3.25: Estimated Volterra kernels.

3.4 Results comparison with literature

As the digital post-distortion for PAs has never been developed for the current communication systems due to the so-called spectral regrowth and ACPR, there is no existent approach in literature to be compared with. For this reason, the proposed post-distortion is compared to the digital pre-distortion. Therefore, we select four recently research works based on Volterra series modeling for wideband PAs. These techniques are briefly described and compared with our approach in the remainder of this section.

1. In [Feh14a], a digital pre-distortion consisting of derivations of the low-pass equivalent (LPE) Volterra series is studied.

2. Starting with the passband Volterra-series formulation, a baseband equivalent (BBE) Volterra-series expression suitable for linearizing PAs driven with intra-band carrier aggregated signals is derived in [Feh14a]. The BBE Volterra series is inherently compact and reduces the model complexity to its basic essentials without the need for pruning, which could potentially impact its linearization capacity when dealing with a wideband PA.

3. [Che14a] aims at an effective digital pre-distortion that can jointly mitigate the frequency-dependent I/Q imbalance and PA nonlinearities. As the computational complexity grows up exponentially to describe all the impairments via a Volterra-series pre-distorter, a modified dynamic-derivation-reduction based (DDR) Volterra series pre-distorter is proposed to reduce the complexity.

4. The authors in [Feh14b] propose a new transmitter architecture suitable for effective transmission of interband carrier aggregated signals. For the purpose of illustrating the concept, only two component carriers will be used. This architecture involves dual-band crest factor reduction and digital pre-distortion modules which have been specifically devised to jointly enhance the linearity and efficiency performance of the transmitter when driven with carrier aggregated signals.

In Table 3.3, a comparison between our work and the aforementioned research works is proposed. Firstly, let us look at the column of the EVM. As it can be seen, the first, the second and the fourth technique, provide better EVM than our work. This can be explained by the high considered non-linearity order $P = 5$. However, this increases the complexity on these digital pre-distortions, which makes it difficult to implement them in mobile terminals. Nevertheless, the reached EVM in our work still acceptable by the current communication systems [Che14a].

Now, let us take a look at the ACPR column. This ACPR corresponds to the ACPR1 calculated at 5MHz of the main carrier frequency. The proposed post-distortion approach provide the lowest ACPR= -54dBc. Therefore, the proposed approach is a good compromise between EVM and ACPR. In addition, the proposed approach is based on fewer Volterra kernels which makes the estimation procedure simpler.

Table 3.3: EVM and ACPR-based comparison between the proposed post-distortion technique and recent digital pre-distortion techniques existent in literature

	Technique	Parameter	Number of coefficients	EVM	ACPR1 at 5MHz
1	Classical low-pass equivalent Volterra series [Feh14a]	$P = 7$, $M = 2$	231	2%	-48dBc
2	Baseband equivalent Volterra series [Feh14a]	$P = 7$, $M = 2$	91	2.1%	-48dBc
3	DPD proposed in [Che14a]	$P = 5$, $M = 2$	44	4.2%	-40dBc
4	DPD proposed in [Feh14b]	$P = 5$, $M = 2$	44	3.5%	-50dBc
	This work	$P = 5$, $M = 2$	44	4%	-54dBc

3.5 Conclusion

In this chapter, the proposed digital post-distortion technique has been applied on a wide-band PA. First of all, the used DUT has been characterized by means AM/AM and AM/PM characteristics. Then, the DUT behavior dependency on carrier frequency is demonstrated validating our assumptions about the CR-PA behavior variation as a function of the carrier frequency. A non-linear characterization is carried on the PA in order to find a relevant way to define the Volterra model non-linearity order and memory depth. Besides, a receiver architecture is proposed and analyzed in order to retrieve the so-called received baseband observation z_k. Based on these interpretations, the algorithm is then applied on the DUT operating near the compression point to get the maximum efficiency. The distortions generated by the PA are successfully compensated at the receiver. This can be seen from the reached EVM which is equal to 4% and ACPR which is equal to -54dBc at 5MHz from the carrier frequency.

The proposed approach feasibility and performance is verified. This technique makes it possible for the PA implemented in the mobile terminal to work at the compression region where the efficiency is the highest possible given that the linearity is met at the digital part of the receiver (the base station). This may decrease the power consumption of the mobile terminal transmitter. However, this aspect is not quantified in this thesis but it can be the subject of further works.

Conclusions and perspectives

This PhD dissertation deals with enhancing the PA efficiency using a digital post-distortion for uplink CR systems.

After presenting generalities about CR and giving the state of the art of the different strategies to meet the trade-off between linearity and efficiency, we suggested designing a technique based on a dynamic Volterra model of the equivalent channel consisting of the PA and the channel. The main advantage of the model is its linearity regarding its parameters, which makes their estimations easier. These parameters are assumed to be time-varying and to exhibit sudden changes when the CT switches from one sub-band to another. However, the choice of the model order and its memory plays a key role. Selecting too small values could not be suited to model the non-linearities, but choosing higher values exponentially increases the number of the parameters to be estimated and consequently the algorithm computational cost.

Therefore, according to various tests we did, we have proposed to set the non-linearity order to $P = 3$ and the memory depth to $M = 2$. This corresponds to a "good" compromise between modeling accuracy, data storage and computational cost. Then, three methods have been proposed.

As the input samples can be estimated if the Volterra kernels are known, we have proposed to jointly estimate the Volterra kernels and the input samples. This hence leads to a non-linear estimation issue that can be solved by using an EKF. This corresponds to a good compromise in terms of accuracy and computational cost compared to the other non-linear Kalman-based estimators. However, this type of approach requires Gaussian assumptions for the model noise and the measurement noise in the state space representation. For this reason, a Gaussian-Mixture EKF is used in the single-carrier case. As the PA behavior can remain unchanged or be time-varying, a single model cannot be representative of the way the Volterra kernels evolve over time. To circumvent this problem, we have proposed to take into account different assumptions about the time dynamics of the Volterra kernels. Based on these assumptions, different estimators are combined in a global structure called the IMM

algorithm to jointly estimate the input samples and track the Volterra parameters.

Using an IMM structure seems to be an attractive solution. Nevertheless, some aspects must be addressed:

1. The number of estimators in the IMM structure: the IMM algorithms we propose are based either on two or three estimators. One is based on the assumptions that the Volterra kernels are constant where the others assume that the Volterra kernels are random walks. They differ from each other by the variance of the driving process. Using two Kalman filters has the advantage of reducing the computational cost, but using three Kalman filters can provide better estimates and lower BER. Considering more than three estimators combined in an IMM structure would not improve the estimation accuracy and would increase the whole computational cost.

2. The setting of the TPM: this matrix plays a key role in the IMM structure. In this PhD, we have proposed two ways to obtain this matrix. On the one hand, we propose to *a priori* set the transition probability by taking into account the mean sojourn time in one model assumption. Nevertheless, this requires some additional information about the spectrum availability and location that can be found in the CT database. On the other hand, the TPM is jointly estimated with the input samples and the Volterra kernels by using [Li05]. In this latter case, the resulting adaptive IMM has the advantage of being "smart", but at the price of a higher computational cost. It should be noted that these recursive algorithms require at least 50 samples at the receiver to converge to the true values of the Volterra kernels.

The statements mentioned above are studied and analyzed using *Matlab* simulations. Then, the relevance of the proposed approaches is confirmed using real signals. The GM-EKF-based IMM reaches 4% of EVM and -54dBc of ACPR at 5MHz. Given this ACPR, the proposed algorithm can be also used as the basic structure of a digital pre-distortion technique in a downlink communication.

These works can be extended in a number of directions, including:

- The proposed algorithm main structure is based on an IMM structure. This structure combines two or three Kalman-based estimators each one is based on an assumption on the Volterra kernel dynamics in order to track the CR-PA behavior. The main drawback of the proposed approach is its computational cost. To solve this problem, an adaptive Kalman-based [Hui15] estimator can be used instead.

- In this thesis, we are interested in single carrier-modulations. However, recently, post-OFDM multicarrier modulation schemes seem to be more suitable for the 5G for example. Therefore, it would be interesting to update the proposed algorithm according to the specifications of these modulation schemes.

- The proposed post-distortion is developed using a behavioral model. Behavioral models are black box models. They do not take into account the inner phenomena of the modeled system. It might be interesting to build digital linearization techniques based on physical models of PAs.

- The main purpose of this thesis is to reduce the power consumption of the mobile terminal. However, we could not evaluate this power reduction. Also, in literature, no values about the digital linearization power consumption are given to compare with. For this reason, we could develop a platform that provides realistic environment including the different components of the mobile terminal transceiver. This may facilitate the evaluation of the power consumption of the different techniques suggested in literature including this work. Then, it could be of interest to make a comparative analysis between these techniques based on power consumption.

- In these works, the measurements are based on a commercial wideband PA (300MHz-3GHz) with low efficiency. However, one of the main purposes of the CRs is to reduce the energy consumption of the mobile terminal by enhancing the efficiency of the PAs. According to the recently published literature, the class-J PA seems to be a promising architecture in terms of efficiency and bandwidth. This topology performance relies on high quality passive components. This can be achieved by using the recent silicon on insulator (SOI) technologies. Therefore, it becomes possible to design a fully integrated wideband class-J PA.

Appendix A

Kalman filtering

KF, firstly presented in [Kal60] by Rudolf E. Kalman, is based on a state space representation of the system, described by two equations. When the state space equations are linear and the noises are additive white zero-mean Gaussian processes, they satisfy:
State equation:

$$\mathbf{x}_n = \mathbf{F}_n \mathbf{x}_{n-1} + \mathbf{G}\mathbf{w}_n \tag{A.1}$$

Measurement equation:

$$\mathbf{z}_n = \mathbf{H}_n \mathbf{x}_n + \mathbf{n}_n \tag{A.2}$$

where $\mathbf{x}_n$ is the state vector of size U at time n and $\mathbf{z}_n$ is the measurement column vector of size K at time n. The model noise w and the observation noise n are uncorrelated white zero-mean Gaussian vectors with covariance matrices $\mathbf{Q}$ and $\mathbf{R}$, respectively. In addition, $\mathbf{F}_n$ is the transition matrix of size $U \times U$ from time $n-1$ to n, G is the input gain matrix of size $U \times U$ and $\mathbf{H}_n$ is the measurement matrix $K \times U$ at time n.

The KF operates in two steps: the prediction step and the filtering step. In the prediction step, the KF uses the estimated state at the previous instant to deduce the current state, without taking into account the current observation. This is the so-called *a priori* estimation of the state vector defined as follows:

$$\hat{\mathbf{x}}_{n|n-1} = \mathbb{E}[\mathbf{x}_n | \mathbf{z}_0, \mathbf{z}_1, \cdots, \mathbf{z}_{n-1}] \tag{A.3}$$

Given (A.1) and (A.3) and as w is a zero-mean AWGN, the *a priori* estimation of the state vector can be expressed as follows:

$$\hat{\mathbf{x}}_{n|n-1} = \mathbf{F}_n \hat{\mathbf{x}}_{n-1|n-1} \tag{A.4}$$

where $\hat{\mathbf{x}}_{n|n-1}$ is the estimation of the state vector at time $n-1$ given the set of observations $\{\mathbf{z}_0, \mathbf{z}_1, \cdots, \mathbf{z}_{n-1}\}$.

Then, let us introduce the *a priori* estimation error:

$$\tilde{\mathbf{x}}_{n|n-1} = \mathbf{x}_n - \hat{\mathbf{x}}_{n|n-1} \tag{A.5}$$

and its corresponding covariance matrix $\mathbf{P}_{n|n-1}$ defined by:

$$\mathbf{P}_{n|n-1} = \mathbb{E}[\tilde{\mathbf{x}}_{n|n-1}\tilde{\mathbf{x}}_{n|n-1}^H] \tag{A.6}$$

Given (A.4) and (A.5), the *a priori* estimation error of the state vector can be rewritten as:

$$\begin{aligned} \tilde{\mathbf{x}}_{n|n-1} &= \mathbf{F}_n\mathbf{x}_{n-1} + \mathbf{G}\mathbf{w}_n - \mathbf{F}_n\hat{\mathbf{x}}_{n-1|n-1} && \text{(A.7)} \\ &= \mathbf{F}_n\tilde{\mathbf{x}}_{n|n-1} + \mathbf{G}\mathbf{w}_n && \text{(A.8)} \end{aligned}$$

Then, the covariance matrix of $\tilde{\mathbf{x}}_{n|n-1}$ satisfies:

$$\begin{aligned} \mathbf{P}_{n|n-1} &= \mathbf{F}_n\mathbb{E}[\tilde{\mathbf{x}}_{n-1|n-1}\tilde{\mathbf{x}}_{n-1|n-1}^H]\mathbf{F}_n^H + \mathbf{G}\mathbb{E}[\mathbf{w}_n\mathbf{w}_n^H]\mathbf{G}^H && \text{(A.9)} \\ &= \mathbf{F}_n\mathbf{P}_{n-1|n-1}\mathbf{F}_n^H + \mathbf{G}\mathbf{Q}\mathbf{G}^H && \text{(A.10)} \end{aligned}$$

At this stage, the filtering step uses the current observation to correct the a priori estimated state vector, in order to obtain the *a posteriori* estimation. The current state can be estimated by following the kind of update equation obtained in the RLS algorithm:

$$\hat{\mathbf{x}}_{n|n} = \hat{\mathbf{x}}_{n|n-1} + \mathbf{K}_n\tilde{\mathbf{z}}_n \tag{A.11}$$

where $\mathbf{K}_n$ is the Kalman gain and $\tilde{\mathbf{z}_n}$ is the innovation defined as follows:

$$\begin{aligned} \tilde{\mathbf{z}}_n &= \mathbf{z}_n - \hat{\mathbf{z}}_{n|n-1} && \text{(A.12)} \\ &= \mathbf{z}_n - \mathbf{H}_n\hat{\mathbf{x}}_{n|n-1} && \text{(A.13)} \\ &= \mathbf{H}_n\tilde{\mathbf{x}}_{n|n-1} + \mathbf{n}_n && \text{(A.14)} \end{aligned}$$

with $\hat{\mathbf{z}}_{n|n-1}$ the prediction of $\mathbf{z}_n$ based on $\hat{\mathbf{x}}_{n|n-1}$.
Now, let us introduce the *a posteriori* estimation error at time n:

$$\begin{aligned} \tilde{\mathbf{x}}_{n|n} &= \mathbf{z}_n - \hat{\mathbf{x}}_{n|n} && \text{(A.15)} \\ &= [\mathbf{I}_U - \mathbf{K}_n\mathbf{H}_n]\tilde{\mathbf{x}}_{n|n-1} - \mathbf{K}_n\mathbf{n}_n && \text{(A.16)} \end{aligned}$$

Given (A.16), its corresponding covariance matrix $\mathbf{P}_{n|n}$ satisfies:

$$\mathbf{P}_{n|n} = [\mathbf{I}_U - \mathbf{K}_n\mathbf{H}_n]\mathbf{P}_{n|n-1}[\mathbf{I}_U - \mathbf{K}_n\mathbf{H}_n]^H + \mathbf{K}_n\mathbf{R}\mathbf{K}_n^H \tag{A.17}$$

The Kalman gain is defined to obtain the MMSE estimation of the state vector. This means that:

$$\frac{\partial(\mathbb{E}[\hat{\mathbf{x}}_{n|n}^H \hat{\mathbf{x}}_{n|n}])}{\partial \mathbf{K}_n} = \frac{\partial tr(\mathbf{P}_{n|n})}{\partial \mathbf{K}_n} = \mathbf{0} \tag{A.18}$$

or equivalently by using (A.17),

$$\frac{\partial tr([\mathbf{I}_U - \mathbf{K}_n\mathbf{H}_n]\mathbf{P}_{n|n-1}[\mathbf{I}_U - \mathbf{K}_n\mathbf{H}_n]^H + \mathbf{K}_n\mathbf{R}\mathbf{K}_n^H)}{\partial \mathbf{K}_n} = \mathbf{0} \tag{A.19}$$

This leads to:

$$\mathbf{K}_n = (\mathbf{P}_n^{\mathbf{xz}})(\mathbf{P}_n^{\mathbf{zz}})^{-1} \tag{A.20}$$

where $\mathbf{P}_n^{\mathbf{zz}}$ is the innovation covariance matrix defined as follows:

$$\mathbf{P}_n^{\mathbf{zz}} = \mathbf{H}_n\mathbf{P}_{n|n-1}\mathbf{H}_n^H + \mathbf{R} \tag{A.21}$$

and $\mathbf{P}_n^{\mathbf{xz}}$ is defined as:

$$\mathbf{P}_n^{\mathbf{xz}} = \mathbf{P}_{n|n-1}\mathbf{H}_n^H \tag{A.22}$$

Finally by combining (A.11) and (A.20), it can be easily shown that the covariance matrix of the a posteriori estimation error can be updated as follows:

$$\mathbf{P}_{n|n} = \mathbf{P}_{n|n-1} - \mathbf{K}_n\mathbf{H}_n\mathbf{P}_{n|n-1} \tag{A.23}$$

$$= [\mathbf{I}_U - \mathbf{K}_n\mathbf{H}_n]\mathbf{P}_{n|n-1} \tag{A.24}$$

This is another expression of (A.17).

At this stage, Kalman filter can be summarized in Algorithm 3 as follows:

Algorithm 3 Kalman filter

- Initialize the values $\hat{\mathbf{x}}_{0|0}$ and $\mathbf{P}_{0|0}$,

The prediction step:

- Update the state vector:

$$\hat{\mathbf{x}}_{n|n-1} = \mathbf{F}_n \hat{\mathbf{x}}_{n-1|n-1}$$

- Update the error covariance matrix:

$$\mathbf{P}_{n|n-1} = \mathbf{F}_n \mathbf{P}_{n-1|n-1} \mathbf{F}_n^H + \mathbf{G}\mathbf{Q}\mathbf{G}^H$$

The filtering step:

- Update the Kalman gain

$$\mathbf{K}_n = \mathbf{P}_{n|n-1} \mathbf{H}_n^H (\mathbf{H}_n \mathbf{P}_{n|n-1} \mathbf{H}_n^H + \mathbf{R})^{-1}$$

- Update the state vector:

$$\hat{\mathbf{x}}_{n|n} = \hat{\mathbf{x}}_{n|n-1} + \mathbf{K}_n (z_n - \mathbf{H}_n \hat{\mathbf{x}}_{n|n-1})$$

- Update the error covariance matrix:

$$\mathbf{P}_{n|n} = (\mathbf{I} - \mathbf{K}_n \mathbf{H}_n) \mathbf{P}_{n|n-1}$$

Appendix B

Extended Kalman filtering

When considering a non-linear system, the EKF consists in analytically propagating the estimation through the system dynamics, by means of a first-order linearization using Taylor expansion.

The non-linear state space equations that describe the system satisfy:

State equation:

$$\mathbf{x}_n = \mathbf{f}_n(\mathbf{x}_{n-1}) + \mathbf{G}\mathbf{w}_n \tag{B.1}$$

Measurement equation:

$$\mathbf{z}_n = \mathbf{h}_n(\mathbf{x}_n) + \mathbf{n}_n \tag{B.2}$$

where $\mathbf{f}_n$ and $\mathbf{h}_n$ are non-linear functions. As proposed in Appendix A, $\mathbf{x}_n$ is the state vector of size U at time n and $\mathbf{z}_n$ is the measurement vector of size K at time n. The noises $\mathbf{w}_n$ and $\mathbf{n}_n$ are uncorrelated additive white zero-mean Gaussian processes with covariance matrices $\mathbf{Q}$ and $\mathbf{R}$, respectively. In addition, $\mathbf{G}$ is the input gain matrix of size $U \times U$.

The first-order Taylor expansion of (B.1) around $\hat{\mathbf{x}}_{n-1|n-1}$ is:

$$\begin{aligned} \mathbf{f}_n(\mathbf{x}_{n-1}) &\approx \mathbf{f}_n(\hat{\mathbf{x}}_{n-1|n-1})\nabla_{\mathbf{x}}\mathbf{f}_n|_{\hat{\mathbf{x}}_{n-1|n-1}}(\mathbf{x}_{n-1} - \hat{\mathbf{x}}_{n-1|n-1}) \\ &\approx \mathbf{f}_n(\hat{\mathbf{x}}_{n-1|n-1}) + \nabla_{\mathbf{x}}\mathbf{f}_n|_{\hat{\mathbf{x}}_{n-1|n-1}}\mathbf{x}_{n-1} - \nabla_{\mathbf{x}}\mathbf{f}_n|_{\hat{\mathbf{x}}_{n-1|n-1}}\hat{\mathbf{x}}_{n-1|n-1} \end{aligned} \tag{B.3}$$

where $\nabla_{\mathbf{x}}\mathbf{f}_n|_{\hat{\mathbf{x}}_{n-1|n-1}}$ denotes the Jacobian matrix of $\mathbf{f}_n$ composed of the partial derivatives of $\mathbf{f}_n$ regarding $\mathbf{x}$ and evaluated for $\hat{\mathbf{x}}_{n-1|n-1}$.

Given (B.3), (B.1) can be rewritten as follows:

$$\mathbf{x}_n = \mathbf{f}_n(\hat{\mathbf{x}}_{n-1|n-1}) + \nabla_{\mathbf{x}}\mathbf{f}_n|_{\hat{\mathbf{x}}_{n-1|n-1}}\mathbf{x}_{n-1} - \nabla_{\mathbf{x}}\mathbf{f}_n|_{\hat{\mathbf{x}}_{n-1|n-1}}\hat{\mathbf{x}}_{n-1|n-1} + \mathbf{G}\mathbf{w}_n \tag{B.4}$$

Using () and (B.4), the *a priori* estimation of the state is defined as follows:

$$\begin{aligned} \hat{\mathbf{x}}_{n|n-1} &= \mathbf{f}_n(\hat{\mathbf{x}}_{n-1|n-1}) + \nabla_{\mathbf{x}}\mathbf{f}_n|_{\hat{\mathbf{x}}_{n-1|n-1}}\mathbb{E}[\mathbf{x}_{n-1}|\mathbf{z}_0, \cdots, \mathbf{z}_{n-1}] - \nabla_{\mathbf{x}}\mathbf{f}_n|_{\hat{\mathbf{x}}_{n-1|n-1}}\hat{\mathbf{x}}_{n-1|n-1} \\ &+ \mathbf{G}\mathbb{E}[\mathbf{w}_n|\mathbf{z}_0, \cdots, \mathbf{z}_{n-1}] \end{aligned} \tag{B.5}$$

As $\mathbf{w}_n$ is a white zero-mean Gaussian process, (B.5) can be rewritten as:

$$\begin{aligned}\hat{\mathbf{x}}_{n|n-1} &= \mathbf{f}_n(\hat{\mathbf{x}}_{n-1|n-1}) + \nabla_{\mathbf{x}}\mathbf{f}_n|_{\hat{\mathbf{x}}_{n-1|n-1}}\hat{\mathbf{x}}_{n-1|n-1} - \nabla_{\mathbf{x}}\mathbf{f}_n|_{\hat{\mathbf{x}}_{n-1|n-1}}\hat{\mathbf{x}}_{n-1|n-1} \\ &= \mathbf{f}_n(\hat{\mathbf{x}}_{n-1|n-1}) \end{aligned} \tag{B.6}$$

Thus, given (B.4) and (B.6), the *a priori* error estimation can be defined as:

$$\begin{aligned}\tilde{\mathbf{x}}_{n|n-1} &= \mathbf{x}_n - \hat{\mathbf{x}}_{n|n-1} \\ &\approx \mathbf{f}_n(\hat{\mathbf{x}}_{n-1|n-1}) + \nabla_{\mathbf{x}}\mathbf{f}_n|_{\hat{\mathbf{x}}_{n-1|n-1}}\mathbf{x}_{n-1} - \nabla_{\mathbf{x}}\mathbf{f}_n|_{\hat{\mathbf{x}}_{n-1|n-1}}\hat{\mathbf{x}}_{n-1|n-1} \\ &+ \mathbf{G}\mathbf{w}_n - \mathbf{f}_n(\hat{\mathbf{x}}_{n-1|n-1}) \\ &\approx \nabla_{\mathbf{x}}\mathbf{f}_n|_{\hat{\mathbf{x}}_{n-1|n-1}} + \mathbf{G}\mathbf{w}_n \end{aligned} \tag{B.7}$$

and its corresponding covariance matrix $\mathbf{P}_{n|n-1}$ is defined by:

$$\mathbf{P}_{n|n-1} = \nabla_x\mathbf{f}_n|_{\hat{\mathbf{x}}_{n-1|n-1}}\mathbf{P}_{n-1|n-1}\nabla_x\mathbf{f}_n^H|_{\hat{\mathbf{x}}_{n-1|n-1}} + \mathbf{G}\mathbf{Q}\mathbf{G}^H \tag{B.8}$$

It should be noted that the *a priori* estimation of the state vector (B.6) does not depend on the Jacobian matrix $\nabla_x\mathbf{f}_n|_{\hat{\mathbf{x}}_{n-1|n-1}}$. However, $\nabla_x\mathbf{f}_n|_{\hat{\mathbf{x}}_{n-1|n-1}}$ is required to calculate the a priori error estimation covariance matrix $\mathbf{P}_{n|n-1}$.

Now, let us focus our attention on the *a posteriori* estimation of the state vector. For this purpose, it is necessary to calculate the first-order Taylor expansion of (B.2) around $\hat{\mathbf{x}}_{n|n-1}$ as follows:

$$\mathbf{h}_n(\mathbf{x}_n) \approx \mathbf{h}_n(\hat{\mathbf{x}}_{n|n-1})\nabla_{\mathbf{x}}\mathbf{h}_n|_{\hat{\mathbf{x}}_{n|n-1}}(\mathbf{x}_n - \hat{\mathbf{x}}_{n|n-1}) \tag{B.9}$$

where $\nabla_{\mathbf{x}}\mathbf{h}_n|_{\hat{\mathbf{x}}_{n|n-1}}$ denotes the Jacobian matrix of $\mathbf{h}_n$ composed of the partial derivatives of $\mathbf{h}_n$ regarding $\mathbf{x}and evaluated for_{n|n-1}$.

Then, given (B.2) and (B.9), the innovation can be expressed as:

$$\begin{aligned}\mathbf{z}_n - \mathbf{h}_n(\hat{\mathbf{x}}_{n|n-1}) &\approx \mathbf{h}_n(\mathbf{x}_n) + \mathbf{n}_n - \mathbf{h}_n(\mathbf{x}_n) + \nabla_{\mathbf{x}}\mathbf{h}_n|_{\hat{\mathbf{x}}_{n|n-1}}(\mathbf{x}_n - \hat{\mathbf{x}}_{n|n-1}) \\ &\approx \nabla_{\mathbf{x}}\mathbf{h}_n|_{\hat{\mathbf{x}}_{n|n-1}}(\mathbf{x}_n - \hat{\mathbf{x}}_{n|n-1}) + \mathbf{n}_n \end{aligned} \tag{B.10}$$

When looking at the linear case, the innovation is defined as follows:

$$\mathbf{z}_n - \mathbf{h}_n\hat{\mathbf{x}}_{n|n-1} = \mathbf{h}_n(\mathbf{x}_n - \hat{\mathbf{x}}_{n|n-1}) + \mathbf{n}_n \tag{B.11}$$

A similarity can be noticed between (B.10) and (B.11). The EKF can be hence easily derived. Indeed, the way the state vector estimation can be updated and the definition of the Kalman gain can be obtained similarly as in appendix A, by replacing $\mathbf{h}(n)$ by $\nabla_{\mathbf{x}}\mathbf{h}_n|_{\hat{\mathbf{x}}_{n|n-1}}$.

Thus, to obtain the a posteriori estimation of the state vector, one has:

$$\hat{\mathbf{x}}_{n|n} = \hat{\mathbf{x}}_{n|n-1} + \mathbf{K}_n(z_n - \mathbf{h}_n(\hat{\mathbf{x}}_{n|n-1})) \tag{B.12}$$

where $\mathbf{K}_n$ is the Kalman gain defined by:

$$\mathbf{K}_n = (\mathbf{P}_n^{\mathbf{xz}})(\mathbf{P}_n^{\mathbf{zz}})^{-1} \tag{B.13}$$

where $\mathbf{P}_n^{\mathbf{zz}}$ is the innovation covariance matrix defined as follows:

$$\mathbf{P}_n^{\mathbf{zz}} = \nabla_{\mathbf{x}}\mathbf{h}_n|_{\hat{\mathbf{x}}_{n|n-1}}\mathbf{P}_{n|n-1}\nabla_{\mathbf{x}}\mathbf{h}_n|^H_{\hat{\mathbf{x}}_{n|n-1}} + \mathbf{R} \tag{B.14}$$

and $\mathbf{P}_n^{\mathbf{xz}}$ is defined as:

$$\mathbf{P}_n^{\mathbf{xz}} = \mathbf{P}_{n|n-1}\nabla_{\mathbf{x}}\mathbf{h}_n|^H_{\hat{\mathbf{x}}_{n|n-1}} \tag{B.15}$$

Finally, by combining (B.12) and (B.13), it can be easily shown that the estimation error covariance matrix is updated as follows:

$$\begin{aligned} \mathbf{P}_{n|n} &= \mathbf{P}_{n|n-1} - \mathbf{K}_n\nabla_{\mathbf{x}}\mathbf{h}_n|_{\hat{\mathbf{x}}_{n|n-1}}\mathbf{P}_{n|n-1} \\ &= [\mathbf{I}_U - \mathbf{K}_n\nabla_{\mathbf{x}}\mathbf{h}_n|_{\hat{\mathbf{x}}_{n|n-1}}]\mathbf{P}_{n|n-1} \end{aligned} \tag{B.16}$$

EKF is summarized in Algorithm 4:

Algorithm 4 Extended Kalman filter

- Initialize the values $\hat{\mathbf{x}}_{0|0}$ and $\mathbf{P}_{0|0}$,
- Calculate the Jacobian matrix $\nabla_x \mathbf{f}_n|_{\hat{\mathbf{x}}_{n-1|n-1}}$,

The prediction step:

- Update the state vector:

$$\hat{\mathbf{x}}_{n|n-1} = \mathbf{f}_n(\hat{\mathbf{x}}_{n-1|n-1})$$

- Update the error covariance matrix:

$$\mathbf{P}_{n|n-1} = \nabla_x \mathbf{f}_n|_{\hat{\mathbf{x}}_{n-1|n-1}} \mathbf{P}_{n-1|n-1} \nabla_x \mathbf{f}_n^H|_{\hat{\mathbf{x}}_{n-1|n-1}} + \mathbf{G}\mathbf{Q}\mathbf{G}^H$$

- Calculate the Jacobian matrix $\nabla_x \mathbf{h}_n|_{\hat{\mathbf{x}}_{n|n-1}}$,

The filtering step:

- Update the Kalman gain $\mathbf{K}_n$

$$\mathbf{K}_n = \mathbf{P}_{n|n-1} \nabla_x \mathbf{h}_n^H|_{\hat{\mathbf{x}}_{n|n-1}} (\nabla_x \mathbf{h}_n|_{\hat{\mathbf{x}}_{n|n-1}} \mathbf{P}_{n|n-1} \nabla_x \mathbf{h}_n^H|_{\hat{\mathbf{x}}_{n|n-1}} + \mathbf{R})^{-1}$$

- Update the state vector:

$$\hat{\mathbf{x}}_{n|n} = \hat{\mathbf{x}}_{n|n-1} + \mathbf{K}_n(z_n - \mathbf{h}_n(\hat{\mathbf{x}}_{n|n-1}))$$

- Update the error covariance matrix:

$$\mathbf{P}_{n|n} = (\mathbf{I} - \mathbf{K}_n \nabla_x \mathbf{h}_n|_{\hat{\mathbf{x}}_{n|n-1}}) \mathbf{P}_{n|n-1}$$

Appendix C

The numerical integration algorithm

Algorithm 5 The numerical integration algorithm

- Initially, setting the weights $p_0^{(s)} = \frac{1}{g}$
- Updating the weights:

$$p_k^{(s)} = \frac{\mu_{k-1}\mathbf{\Pi}^{(s)}\Lambda_k}{\mu_{k-1}\hat{\mathbf{\Pi}}_{k-1}\Lambda_k} p_{k-1}^{(s)} \quad \text{with} \quad s = 1, \cdots, q$$

- Estimating $\mathbf{\Pi}$: $\quad \hat{\mathbf{\Pi}}_k = \sum_{s=1}^{q} \mathbf{\Pi}^{(s)} p_k^{(s)}$

Bibliography

[Abi12] M. Abi Hussein, V. A. Bohara and O. Venard, "On the system level convergence of ILA and DLA for digital predistortion," International Symposium in Wireless Communication Systems (ISWCS), pp. 870–874, 2012.

[Afs12] S. Afsardoost, T. Eriksson and C. Fager, "Digital predistortion using a vector-switched model," IEEE Transactions on Microwave Theory and Techniques, Vol. 60, n°4, pp. 1166–1174, 2012.

[Ama11] R. Amara and S. Marcos, "A blind network of extended Kalman filters for nonstationary channel equalization," IEEE International Conference on Acoustics, Speech and Signal Processing, (ICASSP), pp. 2117–2120, 2011.

[Ara09] I. Arasaratnam and S. Haykin, "Cubature Kalman filters," IEEE Transactions on Automatic Control, Vol. 54, n°6, pp. 1254–1269, 2009.

[Bar01] Y. Bar-Shalom, X. Rong Li and T. Kirubarajan, Estimation with applications to tracking and navigation, John Wiley & Sons, 2001.

[Bax06] R. Baxley and G. Zhou, "Comparison of selected mapping and partial transmit sequence for crest factor reduction in OFDM," IEEE Military Communications Conference (MILCOM), pp. 1–4, 2006.

[Ben15] M. Ben Mabrouk, G. Ferré, E. Grivel and N. Deltimple, "A novel digital post-distortion and detection technique for RF power amplifiers in cognitive radio systems," IEEE International Symposium Wireless Symposium (IWS), pp. 1–4, 2015.

[Ben83] S. Benedetto and E. Biglieri, "Nonlinear equalization of digital satellite channels," IEEE Journal on Selected Areas in Communications, Vol. 1, n°1, pp. 57–62, 1983.

[Big88] E. Biglieri, S. Barberis and M. Catena, "Analysis and compensation of nonlinearities in digital transmission systems," IEEE Journal on Selected Areas in Communications, Vol. 6, n°1, pp. 42–51, 1988.

[Boy85] S. Boyd and L. O. Chua, "Nonlinear system modeling and identification using Volterra-PARAFAC models," IEEE Transactions on Circuits and Systems, Vol. 32, n°11, pp. 1150–1161, 1985.

[Cam15] V. Camarchia, M. Pirola, R. Quaglia, S. Jee, Y. Cho and B. Kim, "The Doherty power amplifier: review of recent solutions and trends," IEEE Transactions on Microwave Theory and Techniques, Vol. 63, n°2, pp. 559–571, 2015.

[Cha01] S. Chang and E. J. Powers, "A simplified predistorter for compensation of nonlinear distortion in OFDM systems," IEEE Global Communications Conference (GLOBECOM), Vol. 5, n°23, pp. 3080–3084, 2001.

[Cha99] S. Challa, Y. Bar-Shalom and V. Krishnamurthy, "Nonlinear filtering using Gauss-Hermite quadrature and generalised edgeworth series," American Control Conference (ACC), Vol. 5, n°6, pp. 3397–3401, 1999.

[Che14a] C. F. Cheang, K. F. Un, W. H. Yu, P. I. Mak and R. P. Martins, "A combinatorial impairment-compensation digital predistorter for a sub-GHz IEEE 802.11af-WLAN CMOS transmitter covering a 10x-wide RF bandwidth," IEEE Transactions on Circuits and Systems-I, Vol. 62, n°4, pp. 1025–1032, 2014.

[Che14b] S. Chen, "An efficient predistorter design for compensating nonlinear memory high power amplifiers," IEEE Transactions on Broadcasting, Vol. 57, n°4, pp. 856–865, 2014.

[Cla98] C. J. Clark, G. Chrisikos, M. S. Muha and A. A. Moulthrop, "Time-domain envelope measurement technique with application to wideband power amplifier modeling," IEEE Transactions on Microwave Theory and Techniques, Vol. 46, n°12, pp. 2531–2540, 1998.

[Cri02] S. C. Cripps, Advanced techniques in RF power amplifier design, Artech House, 2002.

[Cri06] S. C. Cripps, RF power amplifiers for wireless communications, 2nd edition, Artech House, 2006.

[Cri09] S. C. Cripps, P. J. Tasker, A. L. Clarke, J. Lees and J. Benedikt, "On the continuity of high efficiency modes in linear RF power amplifiers," IEEE Microwave and Wireless Components Letters, Vol. 19, n°10, pp. 665–667, 2009.

[Fau09] F. Faurie, A. Giremus and E. Grivel, "Fault detection combining interacting multiple model and multiple solution separation for aviation satellite navigation system," IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Vol. 58, n°8, pp. 3273–3276, 2009.

[Feh14a] B. Fehri and S. Boumaiza, "Baseband equivalent Volterra series for behavioral modeling and digital predistortion of power amplifiers driven with wideband carrier aggregated signals," IEEE Transactions on Microwave Theory and Techniques, Vol. 62, n°11, pp. 2594–2603, 2014.

[Feh14b] B. Fehri and S. Boumaiza, "Joint dual-band crest factor reduction and digital predistortion of power amplifiers driven by inter-band carrier aggregated signals," Microwave Symposium (IMS), 2014 IEEE MTT-S International, pp. 1–4, 2014.

[Gha08] A. Ghasemi and E. S. Sousa, "Spectrum sensing in cognitive radio networks: requirements, challenges and design trade-offs," IEEE Communications Magazine, Vol. 46, n°4, pp. 32–39, 2008.

[Gha09] F. M. Ghannouchi and O. Hammi, "Behavioral modeling and predistortion," IEEE Microwave Magazine, Vol. 10, n°7, pp. 52–64, 2009.

[Gha10] F. Ghannouchi, F. Taringou and O. Hammi, "A dual branch Hammerstein-Wiener architecture for behavior modeling of wideband RF transmitters," IEEE MTT-S International in Microwave Symposium Digest (MTT), pp. 1692–1695, 2010.

[Gir10] A. Giremus, E. Grivel, J. Grolleau and M. Najim, "A Rao Blackwellized particle filter for joint channel/symbol estimation in MC-DS-CDMA," IEEE Transactions on Communications, Vol. 58, n°8, pp. 2292–2304, 2010.

[Gol09] A. Goldsmith, S. A. Jafar, I. Maric and S. Srinivasa, "Breaking spectrum gridlock with cognitive radios: an information theoretic perspective," Proceedings of the IEEE, Vol. 97, n°5, pp. 894–914, 2009.

[Gua12] L. Guan and A. Zhu, "Optimized low-complexity implementation of least squares based model extraction for digital predistortion of RF power amplifiers," IEEE

Transactions on Microwave Theory and Techniques, Vol. 60, n°3, pp. 594–603, 2012.

[Hon13] T. Hone, A. Aref, J. Guan and R. Negra, "Achieving linearity for outphasing amplifiers targeting LTE applications and beyond," Asia-Pacific Microwave Conference Proceedings (APMC), pp. 444–446, 2013.

[Hua15] T. Y. Huang, Y. H. Lin and H. Wang, "A K-Band adaptive-bias power amplifier with enhanced linearizer using 0.18um CMOS process," IEEE International Microwave Symposium (IMS), pp. 1–3, 2015.

[Hui15] L. Huihui and A. J. Zhang, "Improved adaptive Kalman filtering algorithm for vehicular positioning," 2015 34th Chinese Control Conference (CCC), pp. 5125–5129, 2015.

[Hus08] S. Hussain and Y. Louet, "Peak to average power ratio analysis of multi-carrier and multi-standard signals in software radio context," 3rd International Conference on Information and Communication Technologies: From Theory to Applications (ICTTA), pp. 1–5, 2008.

[Isa06] M. Isaksson, D. Wisell and D. Ronnow, "A comparative analysis of behavioral models for RF power amplifiers," IEEE Transactions on Microwave Theory and Techniques, Vol. 54, n°1, pp. 348–359, 2006.

[Ito00] K. Ito and X. Kaiqi, "Gaussian filters for nonlinear filtering problems," IEEE Transactions on Automatic Control, Vol. 45, n°5, pp. 910–927, 2000.

[Jia12] H. Jiang, X. Yu and P. Wilford, "Digital predistortion using stochastic conjugate gradient method," IEEE Transactions on Broadcasting, Vol. 58, n°1, pp. 114–124, 2012.

[Jil04] V. P. Jilkov and X. R. Li, "Online Bayesian estimation of transition probabilities for Markovian jump systems," IEEE Transactions on Signal Processing, Vol. 52, n°6, pp. 1620–1630, 2004.

[Kar89] G. Karam and H. Sari, "Improved data predistortion using intersymbol interpolation," IEEE Journal on Selected Areas in Communications, Vol. 1, n°6, pp. 286–291, 1989.

[Ken00] P. B. Kenington, High linearity RF amplifier design, Artech House, 2000.

[Kim01] J. Kim and J. Konstantinou, "Digital predistortion of wideband signals based on power amplifier model with memory," Electronics Letters, Vol. 37, n°23, pp. 1417–1418, 2001.

[Kim15] J. Kim, D. Kim, Y. Cho, D. Kang, B. Park, K. Moon, S. Koo and B. Kim, "Highly efficient RF transmitter over broad average power range using multilevel envelope-tracking power amplifier," IEEE Transactions on Circuits and Systems I, Vol. 62, n°6, pp. 1648–1657, 2015.

[Kow11] S. Kowlgi and C. Berland, "Linearity considerations for multi-standard cellular base station transmitters," European Microwave Conference (EuMC, pp. 226 – 229, 2011.

[Lan08] P. Landin, M. Isaksson and P. Handel, "Comparison of evaluation criteria for power amplifier behavioral modeling," IEEE MTT-S International in Microwave Symposium Digest (MTT), pp. 1441–1444, 2008.

[Las13] N. Lashkarian, H. Tarn and C. Dick, "Crest factor reduction in multi-carrier WCDMA transmitters," IEEE 16th International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), pp. 321–325, 2013.

[Leo85] I. J. Leontaritis and S. A. Billings, "Input-output parametric models for nonlinear systems part I: deterministic non-Linear systems," International Journal of Control, Vol. 41, n°2, pp. 303–328, 1985.

[Li00] X. R. Li and Y. Zhang, "Multiple-model estimation with variable structure part V: likely-model set algorithm," IEEE Transactions on Aerospace and Electronic Systems, Vol. 36, n°4, pp. 448–466, 2000.

[Li05] X. R. Li and V. P. Jilkov, "Survey of maneuvering target tracking. Part V. Multiple-model methods," IEEE Transactions on Aerospace and Electronic Systems, Vol. 41, n°10, pp. 1255–1321, 2005.

[Lim00] Y. H. Lim, Y. S. Cho, I. W. Cha and D. H. Youn, "An adaptive nonlinear prefilter for compensation of distortion in nonlinear systems," IEEE Transactions on Signal Processing, Vol. 46, n°6, pp. 1726–1730, 2000.

[Liu04] J. Liu, H. Arslan, L. P. Dunleavy and J. Paviol, "Impact of carrier frequency dependent power amplifier behavior on 802.11a WLAN system," European Conference on Wireless Technology (ECWT), pp. 289–292, 2004.

[Mit00] J. Mitola III. Cognitive radio: an integrated agent architecture for software defined radio. PhD thesis, Royal Intitute of Technology (KTH), Sweden, 2000.

[Moo10] J. Moon, J. Kim and B. Kim, "Investigation of a class-J power amplifier with a nonlinear Cout for optimized operation," IEEE Transactions on Microwave Theory and Techniques, Vol. 58, n°11, pp. 2800–2811, 2010.

[Mor14] D. Morgan, J. Kim, M. Zierdt and J. Pastalan, "Parametric identification of parallel Hammerstein systems," IEEE Transactions on Signal Processing, Vol. 54, n°10, pp. 3852–3860, 2014.

[Nad11] C. Nader, P. Landin, W. Van Moer and N. Bjorsell, "Performance evaluation of peak-to-average power ratio reduction and digital pre-distortion for ofdm based systems," IEEE Transactions on Microwave Theory and Techniques, Vol. 59, n°12, pp. 3504–3511, 2011.

[Par14] D. R. Parveg, P. Singerl, A. Wiesbauer, H. M. Nemati and C. Fager, "A broadband, efficient, overdriven class-J RF power amplifier for burst mode operation," European Microwave Conference (EuMC), pp. 1666–1669, 2014.

[Pre12] C. Presti, D. Kimball and P. Asbeck, "Closed-loop digital predistortion system with fast real-time adaptation applied to a handset WCDMA PA module," IEEE Transactions on Microwave Theory and Techniques, Vol. 60, n°3, pp. 604–618, 2012.

[Rai13] R. Raich, H. Qian and G. Zhou, "Orthogonal polynomials for power amplifier modeling and predistorter design," IEEE Transactions on Vehicular Technology, Vol. 53, n°5, pp. 1468–1479, 2013.

[Rho03] J. D. Rhodes, "Output universality in maximum efficiency linear power amplifiers," International Journal of Circuit Theory and Applications, Vol. 31, n°4, pp. 385–405, 2003.

[Rug81] W. J. Rugh, Nonlinear system theory: the Volterra/ Wiener approach, The Johns Hopkins University Press, 1981.

[Sal83] A. A. M. Saleh and J. Salz, "Adaptive linearization of power amplifiers in digital radio systems," The Bell System Technical Journal, Vol. 62, n°4, pp. 1019–1033, 1983.

[Sar14] A. Sarkar and B. Floyd, "A 28-GHz class-J power amplifier with 18-dBm output power and 35% peak PAE in 120-nm SiGe BiCMOS," IEEE 14th Topical Meeting on Silicon Monolithic Integrated Circuits in RF Systems (SiRF), pp. 71–73, 2014.

[Sch14] M. Schoukens, R. Pintelon and Y. Rolain, "Parametric identification of parallel Hammerstein systems," IEEE Transactions on Instrumentation and Measurement, Vol. 60, n°12, pp. 3931–3938, 2014.

[Sig12] Z. Sigrist, E. Grivel and B. Alcoverro, "Estimating second-order Volterra system parameters from noisy measurements based on an LMS variant or an errors-in-variables method," Signal Processing, Vol. 92, pp. 1010–1020, 2012.

[Tar15] F. Taringou, O. Hammi and F. Ghannouchi, "Behaviour modelling of wideband RF transmitters using Hammerstein-Wiener models," IEEE Transactions on Microwave Theory and Techniques, Vol. 4, n°4, pp. 282–290, 2015.

[Teh10] A. Tehrani, H. Cao, M. Isaksson and C. Fager, "A comparative analysis of the complexity/accuracy tradeoff in power amplifier behavioral models," IEEE Transactions on Microwave Theory and Techniques, Vol. 58, n°6, pp. 1510–1520, 2010.

[Vuo03] J. Vuolevi and T. Rahkonen, Distortion in RF power amplifiers, Artech House, 2003.

[Wan02] E. Wan and R. Merwe, Kalman filtering and neural networks, John Wiley & Sons, 2002.

[Weg06] A. Wegener, "High-performance crest factor reduction processor for w-CDMA and OFDM applications," IEEE Symposium on Radio Frequency Integrated Circuits (RFIC), pp. 1–4, 2006.

[Wri09] P. Wright, J. Lees, P. J. Tasker and S. C. Cripps, "A Methodology for realizing high efficiency class-J in a linear and broadband PA," IEEE Transactions on Microwave Theory and Techniques, Vol. 57, n°12, pp. 3196–3204, 2009.

[Wu10] D. Y. T. Wu, F. Mkadem and S. Boumaiza, "Design of a broadband and highly efficient 45W GaN power amplifier via simplified real frequency technique," IEEE MTT-S International Microwave Symposium Digest (MTT), pp. 1090–1093, 2010.

[Zel12] J. Zeleny, C. Dehos, P. Rossona and A. Kaiser, "Receiver-aided predistortion of power amplifier non-linearities in cellular networks," Institution of Engineering and Technology Science, Measurement & Technology, Vol. 6, n°7, pp. 168–175, 2012.

[Zho07] D. Zhou and V. E. De Brunner, "Novel adaptive nonlinear predistorters based on the direct learning algorithm," IEEE Transactions on Signal Processing, Vol. 55, n°1, pp. 120–133, 2007.