PolyPublie
Polytechnique Montréal

POLYTECHNIQUE MONTRÉAL
UNIVERSITÉ D'INGÉNIERIE

Titre: Title:	Dynamic nonlinear behavioral modeling and adaptive predistortion for RF transmitters
Auteur: Author:	Farzaneh Taringou
Date:	2008
Type:	Mémoire ou thèse / Dissertation or Thesis
Référence: Citation:	Taringou, F. (2008). Dynamic nonlinear behavioral modeling and adaptive predistortion for RF transmitters [Mémoire de maîtrise, École Polytechnique de Montréal]. PolyPublie. https://publications.polymtl.ca/8273/

Document en libre accès dans PolyPublie
Open Access document in PolyPublie

URL de PolyPublie: PolyPublie URL:	https://publications.polymtl.ca/8273/
Directeurs de recherche: Advisors:	Fadhel M. Ghannouchi, & Roland P. Malhamé
Programme: Program:	Non spécifié

Ce fichier a été téléchargé à partir de PolyPublie, le dépôt institutionnel de Polytechnique Montréal
This file has been downloaded from PolyPublie, the institutional repository of Polytechnique Montréal

https://publications.polymtl.ca

UNIVERSITÉ DE MONTRÉAL

DYNAMIC NONLINEAR BEHAVIORAL MODELING AND ADAPTIVE PREDISTORTION FOR RF TRANSMITTERS

FARZANEH TARINGOU

DÉPARTEMENT DE GÉNIE ÉLECTRIQUE
ÉCOLE POLYTECHNIQUE DE MONTRÉAL

MÉMOIRE PRÉSENTÉ EN VUE DE L'OBTENTION
DU DIPLÔME DE MAÎTRISE ÈS SCIENCES APPLIQUÉES
(GÉNIE ÉLECTRIQUE)
MARS 2008

© FARZANEH TARINGOU, 2008.

Library and Archives Canada

Published Heritage Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et Archives Canada

Direction du Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence
ISBN: 978-0-494-41580-1
Our file Notre référence
ISBN: 978-0-494-41580-1

NOTICE:
The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:
L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

UNIVERSITÉ DE MONTRÉAL

ÉCOLE POLYTECHNIQUE DE MONTRÉAL

Ce mémoire intitulé:

MODELISATION COMPORTEMENTAL NON-LINEAIRE DYNAMIQUE ET PRÉDISTORSION ADAPTATIVE POUR LES TRANSMETTEURS RF

présenté par : FARZANEH TARINGOU

en vue de l'obtention du diplôme de : Maîtrise ès sciences appliquées

a été dûment accepté par le jury d'examen constitué de :

L. SAYDY, Ph. D., président

F.GHANNOUCHI, Ph. D., membre et directeur de recherche

R. MALHAMÉ, Ph. D., membre et co-directeur de recherche

B. SRINIVASAN, Ph. D., membre

DEDICATION

To my parents and my brother

To my best friend, Francisco

ACKNOWLEDGEMENTS

I would like to express my gratitude to my supervisors; Prof. F. Ghannouchi and Prof. R.Malhamé, for their technical and financial support and uninterrupted encouragement throughout the work involved in this thesis, pursuing my master studies. I also want to thank the member of my thesis committee, Prof. L. Saydy for his comments and suggestions on my thesis.

I also would like to thank Prof. B.Srinivasan from chemical engineering department for his very fruitful helps and technical advices in the context of system identification, Prof. C.Audet from mathematics and industrial engineering department for his kind helps concerning the optimization tools and the software NOMAD and, dear David Dousset for his helpful French skills and translations.

And finally, special thanks to my parents and my brother for being there for me and to my very missing friend, Francisco Casares for his never ending long distant supports.

RÉSUMÉ

La demande grandissante des systèmes de transmission sans fil a généré le besoin de concevoir des systèmes large bande et efficaces du point de vue énergétique. Permettant ainsi de transférer des données via des signaux modules complexes tells que OFDM, CDMA...etc. tout en minimisant les distorsions des signaux au cours de la transmission. C'est dans ce contexte que l'on doit garder à l'esprit que la principale source de distorsion est l'amplificateur de puissance. En effet, pour un certain niveau de puissance du signal d'entrée, l'amplificateur opère près de son point de compression, dans la zone dite de saturation. Il en résulte une dégradation non-linéaire en termes de gain et de déphasage relative, également connue comme les effets AM/AM et AM/PM.

Une approche communément utilisée, pour compenser les distorsions qui peuvent apparaître à la sortie de l'amplificateur, est d'inclure dans la chaîne d'amplification un élément de prédistorsion. Il consiste en un bloc fonctionnel non-linéaire qui permet de reproduire la caractéristique non-linéaire inverse de l'amplificateur de puissance. Ainsi, la combinaison des signaux issus de l'amplificateur et du prédistorseur donne la fonction identité. La procédure adaptative, qui permet la mise à jour des valeurs du prédistorseur à chaque itération, est basée sur un algorithme d'optimisation spécifique qui demande généralement un temps de calcul important. Par conséquent, avant d'utiliser l'algorithme, il est nécessaire de faire une estimation approximative des paramètres du prédistorseur, ceci servant à réduire considérablement les temps de calcul.

Un autre facteur important dont il faut tenir compte est celui du signal d'apprentissage qui utilisé dans l'algorithme adaptatif. Certains choix pourront non

seulement rendre la procédure plus complexe, mais ils sont susceptibles d'introduire des retards additionnels dans la boucle d'ajustement du fait que le calcul d'erreur dépend de la nature du signal d'entrée.

L'objectif de ce mémoire est la conception et l'implémentation d'un prédistorseur adaptatif. Ce dernier est basé sur un modèle de transmetteur spécifique et l'utilisation d'une sinusoïdale comme signal d'apprentissage, avec une fréquence qui se situe au sein de la bande passante du transmetteur. Le critère de linéarité reçoit un signal à une seule fréquence à la sortie qui a un gai linéaire et un déphasage relatif. En d'autres termes, on suppose que la non linéarité du système peut être séparée de la ***dynamique linéaire et compensée***. Lorsque le but est atteint, on a simplement un système dynamique linéaire, c'est-à-dire un système identifiable en utilisant les techniques d'identification linéaire des systèmes disponibles dans la bibliothèque de MATLAB.

L'hypothèse clé, lorsque le prédistorseur est de type statique, est que les non-linéarités statiques sont séparables de la dynamique linéaire, c'est-à-dire que le transmetteur peut être modélisé à partir de modèles à deux blocs, aussi connus sous les noms de modèles de Hammerstein ou de Wiener. Pour ces modèles, le bloc non-linéaire statique est suivi, ou bien est précédé, d'un filtre dynamique linéaire. Un modèle théorique de ce type est initialement considéré pour le transmetteur et le succès de la technique d'estimation est démontré. Le modèle est alors considéré comme une boîte noire qui remplace le système final et à laquelle on appliquera une prédistorsion adaptative. Il sera montré que du fait de l'utilisation du critère de linéarité pour une onde sinusoïdale, l'erreur de calcul est déterminée indépendamment du signal d'entrée. Cependant, des contraintes non-

linéaires doivent être considérées afin de conserver toute la fonctionnalité du système dans la zone non-linéaire.

Nous verrons également que pour une compensation de la distorsion de phase il est nécessaire d'appliquer deux ondes sinusoïdales. En effet, l'information de la phase relative dans le cadre d'une modulation en quadrature ne peut être déterminée avec uniquement un signal sinusoïdal. Finalement, les performances du prédistorseur sont évaluées en utilisant un signal WCDMA. Les courbes AM/AM et AM/PM obtenues sont étudiées ainsi que la densité du spectre de puissance du signal de sortie, où il est impératif que le phénomène non linéaire dit de recouvrement de puissance du canal adjacent soit réduit.

ABSTRACT

The increasing demand on wireless transmitter systems has generated a need for efficient wideband designs capable of transferring the data usually carried via a signal modulation scheme such as OFDM, CDMA...etc with the least possible distortion during transmission. In this context one should remember that the main source of distortion in a transmitter is the power amplifier which has a tendency to saturate when the input signal power level increases beyond a certain level, thus resulting in nonlinear degradation in terms of gain and relative phase shift respectively known as AM/AM and AM/PM effects.

One common approach to compensate for these power dependent distortions is known as *predistortion* whereby a nonlinear functional box with inverse nonlinear characteristics of that of the transmitter precedes the power amplifier where as a results, this combination gives rise to an identity operator. The adaptive procedure that updates the predistorter parameters at each iteration is based on a specific optimization algorithm which is in general computationally expensive. Therefore, it is important to produce a rough estimation of the predistorter parameters which can serve as a good initial point; this can help considerably in reducing the length of computations.

One other important element is the particular nature of the training signal that is used for operating the adaptive algorithm. Some signals will add up more complexity to the procedure as well as some feedback delay due to error calculation dependency on the input signal.

The objective of this thesis is the design and implementation of an adaptive predistorter based on a specific type of transmitter model and using single frequency waves as training signals with a frequency which falls within the bandwidth of the transmitter. The linearity criterion of the predistorter/amplifier combination is that a single frequency signal should reach the output with a linear gain and phase shift for a broad range of power levels. Once this goal is achieved, one is left with a merely linear dynamic system i.e. that is identifiable by means of linear system identification techniques available in standard MATLAB toolboxes.

Clearly, when using a static predistorter, the underlying key assumption is that the static nonlinearities are separable from linear dynamics, i.e. the transmitter can be modeled for example with 2-box models also known as Hammerstein or Wiener models where a static nonlinear function is followed by a linear dynamic filter or vice versa. We successfully illustrate the single sine wave based predistortion linearization for a theoretical Hammerstein type model.

It is also shown that for phase distortion compensation, one needs to apply two sine waves since the relative phase information in the context of quadrature modulation cannot be revealed with only one sinusoidal signal. In one last step the constructed predistorter performance is validated by launching a WCDMA signal and the AM/AM and AM/PM curves are discussed as well as the power spectrum density of the output signal where the non linear effect known as adjacent channel power regrowth must be reduced.

CONDENSÉ EN FRANÇAIS

MODELISATION COMPORTEMENTAL NON-LINÉAIRE DYNAMIQUE ET PRÉDISTORSION ADAPTATIVE POUR LES TRANSMETTEURS RF

0.1. INTRODUCTION

La conception d'un transmetteur à large bande pour les nouveaux systèmes de communications sans fil à haute vitesse, pour lesquels il existe une grande variation d'enveloppes des signaux d'entrée, est une tâche délicate et complexe. En effet, la conception nécessite des compromis entre les objectifs contradictoires de maintien d'une haute linéarité et de réalisation d'une haute efficacité énergétique. L'adoption massive des techniques de transmission multi-porteuses dans les nouveaux systèmes de communications sans fil large bande, telles que CDMA, OFDM ou encore Wimax, implique l'utilisation de signaux à enveloppe non constante et ayant des ratios de puissance maximale par rapport à la puissance moyenne (Pars) importants. Lorsqu'un signal modulé passe à travers un circuit présentant une faible non-linéarité, sa largeur de bande est fonction des non-linéarités d'ordre impair. Ce phénomène, appelé recouvrement spectral ou régénération spectrale, est une conséquence de la génération des produits des composantes fréquentielles.

Il en résulte qu'un amplificateur de puissance (PA) peut être conçu de façon à ce qu'il opère près de son point de saturation, et ce, pour obtenir une plus grande efficacité énergétique, ou bien loin en retrait de sa zone de saturation. Dans le but de faire fonctionner l'amplificateur dans sa zone de non-linéarité, tout en conservant une haute

efficacité énergétique et un comportement linéaire, il est indispensable d'ajouter un élément de distorsion non-linéaire dans le transmetteur. Les techniques de linéarisation les plus courantes actuellement sont « feedback », « feedforward » et la technique de prédistorsion. Ce dernier est l'une des plus efficaces et elle offre en outre la flexibilité de pouvoir être implémentée soit de façon numérique, soit de façon analogique. Le principe de la prédistorsion numérique, également appelé prédistorsion en bande de base, est de modifier les données à l'entrée de façon à optimiser les performances de sortie de l'amplificateur.

Nous proposons dans ce mémoire une prédistorsion adaptative basée sur la récupération d'un signal à une seule fréquence. Ce qui se distingue très nettement des autres méthodes qui utilisent, quant à elles, des signaux modulés lors de l'implémentation de la prédistorsion numérique. Concernant le critère de linéarisation, on considère simplement la réponse d'un système linéaire qui serait excité par une sinusoïde pure. Le signal de sortie de ce système linéaire doit alors être un signal sinusoïdal de même fréquence mais d'amplitude et de phase différente.

Les modèles non-linéaires à 2 blocs ont été identifiés et validés pour reproduire le comportement du système de transmission et l'élément de prédistorsion adaptative est appliqué au modèle. Une validation complémentaire de notre élément de prédistorsion a été réalisée en utilisant un signal CDMA.

0.2. MODÈLES À DEUX BLOCS, NON-LINEARITÉS STATIQUES ET EFFETS MÉMOIRE DYNAMIQUES

Dans le but de concevoir un élément de prédistorsion permettant la suppression des effets non-linéaires, il est nécessaire de disposer d'un modèle de transmetteur RF précis. On peut constater que les modèles à 2 blocs ont été extensivement décrits dans la littérature, et ce, pour modéliser le comportement non-linéaire des transmetteurs sans-fils en termes de distorsions d'amplitude AM/AM et de phase AM/PM. Cependant, nous voudrions bien faire la distinction entre les modèles classiques à 2 blocs et les modèles de Wiener/Hammerstein, qui sont intrinsèquement incapables de modéliser les variations non-linéaires de phase.

Pour les modèles de Wiener/Hammerstein, la procédure d'identification est réalisée en une seule étape. Les coefficients de la fonction statique non-linéaire, ainsi que le filtre linéaire, sont identifiés à travers un processus d'optimisation sans contraintes multidimensionnel implanté dans MATLAB.

La partie réelle du signal modulé est construite à partir des composantes en bande de base. Elle est utilisée pour l'apprentissage dans une des formes suivantes :

$$\begin{aligned} v(t) &= R(t)\cos\left[\omega_c t + \theta(t)\right] \\ v(t) &= x(t)\cos(\omega_c t) - y(t)\sin(\omega_c t) \end{aligned} \tag{1}$$

Dans le modèle de Wiener, un bloc statique non-linéaire est précédé d'un filtre linéaire, qui prend en compte la dynamique. Pour le modèle de Hammerstein, l'ordre est inversé.

La non-linéarité statique est modélisée avec un polynôme de la forme suivante:

$$f(\beta(y(t))) = \alpha_0 + \alpha_1\beta(y(t)) + \alpha_2\beta^2(y(t)) + \\ \ldots + \alpha_n\beta^n(y(t)) \tag{2}$$

où $\beta(y(t))$ est la fonction de base choisie de façon à ce que le polynôme soit en accord avec le comportement de saturation du système final. Le modèle de la fonction de transfert du bloc linéaire est défini par une simple relation entre l'entrée et la sortie, qui peut être défini comme une différence linéaire (ou une équation différentielle...) et connu sous le nom de modèle ARX: :

$$\hat{y}(t) + a_1\hat{y}(t-1) + \ldots + a_{n_a}\hat{y}(t-n_a) \\ = b_0u(t) + b_1u(t-1) + \ldots b_{n_b}u(t-n_b) + e(t) \tag{3}$$

où la condition de stabilité du BIBO (Bounded input Bounded OUTPUT) est assurée en imposant n_a = 0. Les coefficients du filtre linéaire et ceux de la fonction statique non-linéaire sont alors mis à jour de façon itérative. L'erreur est alors minimisée en utilisant le processus d'optimisation.

Puisque les modèles Wiener/Hammerstein ne sont pas en mesure de reproduire la distorsion de phase d'un système, un modèle plus général à deux blocs est utilisé. Ce dernier a une structure similaire aux modèles Wiener/Hammerstein mais la partie non-linéaire est représentée par une fonction complexe. Les performances de ces trois modèles sont évaluées avec un autre ensemble de données et la qualité du modèle considéré est calculée en utilisant la relation:

$$FIT = \left[1 - Norm(Y - \hat{Y}) / Norm(Y - Mean(Y))\right] * 100 \tag{4}$$

On constate une amélioration très sensible des performances pour le dernier modèle à deux blocs par rapport aux modèles Wiener/Hammerstein. Ce résultat a été rendu possible grâce a la prise en considération du caractère non-linéaire de la phase dans notre modèle.

0.3. MODÈLE NON-LINÉAIRE SOUS FORME DE 3-BLOCS

Notre modèle à trois blocs est en fait une combinaison du modèle de Hammerstein et du modèle non-linéaire à deux blocs, tous deux décrits dans la section 0.2, où un filtre linéaire dynamique est placé entre deux blocs non linéaires. La procédure d'identification de modèle comprend deux étapes. Les deux premiers blocs sont similaires à la structure de Wiener/Hammerstein mais utilisent un signal modulé. La sortie de cet étage est alors convertie en un signal complexe en bande de base. Le déphasage relatif est, quant à lui, calculé puis comparé au déphasage mesuré. Le déphasage résiduel est alors modélisé dans un troisième bloc non-linéaire. Le troisième modifie le déphasage du modèle dans sa globalité mais ne change en rien le gain des deux blocs qui constituent le premier étage. En d'autres termes, cela modifie les performances du filtre linéaire situé au centre des 2 blocs de telle façon que l'ensemble forme un filtre non-linéaire.

Le premier étage d'identification, qui est exactement le même que le modèle de Wiener/Hammerstein décrit dans la précédente section, permet de très bien reproduire la distorsion AM/AM.

A la sortie des deux premiers blocs, le signal est démodulé et sa composante complexe en bande de base est extraite pour permettre de retrouver les courbes AM/AM et AM/PM. Si l'on observe les relations entre le gain et la phase du système et les composantes en phase et en quadrature, on peut constater qu'il est possible, avec quelques manipulations mathématiques sur les composantes du signal, de modifier la phase pour que cette dernière soit en accord avec celle du transmetteur. Le gain, quant à lui reste inchangé :

$$Gain(t) = \frac{\sqrt{I_{out}^2(t) + Q_{out}^2(t)}}{\sqrt{I_{in}^2(t) + Q_{in}^2(t)}} \tag{5}$$

$$Phase(t) = \arctan(\frac{Q_{out}(t)}{I_{out}(t)}) - \arctan(\frac{Q_{in}(t)}{I_{in}(t)}) \tag{6}$$

En d'autres termes, le rapport entre les composantes en phase et en quadrature $\frac{Q_{out}(t)}{I_{out}(t)}$ est modifié pour donner la même réponse en phase que celle du système. Cependant, la puissance $\sqrt{I_{out}^2(t)+Q_{out}^2(t)}$ reste inchangée.

A partir de l'identification et de la validation du modèle, il a été possible d'obtenir un accord acceptable entre les mesures et la modélisation. En dépit d'une fonction non-linéaire d'ordre relativement élevé dans le modèle, un compromis entre l'efficacité et la simplicité est nécessaire ; ceci est dû à l'utilisation d'un filtre linéaire du premier ordre au lieu de deux filtres de deuxième ordre du modèle à 2 blocs.

0.3. CONSTRUCTION DE LA PRÉDISTORSION A PARTIR D'UNE ONDE SINUSOIDALE

La procédure de conception et d'élaboration de la prédistorsion numérique adaptative ou non, comme il peut être largement vu dans la littérature, utilise un signal d'apprentissage du même type que celui qui doit être transmis. Ceci mis a part, il faut garder à l'esprit que la principale cause des effets de non-linéarité est le niveau de puissance du signal à l'entrée de l'amplificateur de puissance. En effet, lorsque le niveau de puissance en entrée est trop important, l'amplificateur opère prés de sa zone de saturation, ce qui entraîne une augmentation des niveaux de puissance des produits d'intermodulation. Enfin, que se soit une simple sinusoïde ou bien un signal complexe modulé qui excite le système, aussi longtemps qu'ils ont le même niveau de puissance, ils exciteront la même zone non-linéaire d'amplificateur de puissance et le signal de sortie comportera les mêmes niveaux de distorsion.

Dans ce chapitre, un signal sinusoïdal est utilisé comme signal d'apprentissage et un processus d'adaptation directe est utilisé pour la mise à jour les coefficients du bloc de prédistorsion non-linéaire. De cette façon, la sortie du système est toujours une sinusoïde, ce qui garantit la linéarité du système considéré. Nous utilisons ce fait dans le but de déterminer le cycle de mise à jour de l'estimateur.

Le modèle à 2 blocs développé dans la précédente section est alors utilisé comme une boîte noire. Si on assigne la fonction polynomiale statique g pour la prédistorsion et la fonction non-linéaire dynamique f pour le transmetteur, alors la sortie, bien qu'elle soit distordue, est périodique avec la même période que la sinusoïde d'entrée. Une fois que

le signal de sortie $g(f(x(t)))$ est obtenu, on le développe en série de Fourier afin d'extraire la composante fondamentale :

$$g(f(x(t))) = \frac{1}{2}a_0 + \sum_{n=1}^{\infty} a_n \cos(\omega_n t) + \sum_{n=1}^{\infty} b_n \sin(\omega_n x) \tag{7}$$

$$\omega_n = n\frac{2\pi}{T} \tag{8}$$

$$f = \frac{1}{T} \tag{9}$$

où f est la fréquence fondamentale. En partant de la relation suivante pour l'extraction des coefficients :

$$a_n = \frac{2}{T}\int_{t_1}^{t_2} g(f(x(t)))\cos(\omega_n t) \tag{10}$$

$$b_n = \frac{2}{T}\int_{t_1}^{t_2} g(f(x(t)))\sin(\omega_n t) \tag{11}$$

si on sépare la composante fondamentale du reste de la série, on a alors :

$$g(f(x(t))) = \underbrace{\{a_1 \cos(\omega t) + b_1 \sin(\omega t)\}}_{Fondamentale} + \underbrace{\left\{\frac{1}{2}a_0 + \sum_{n=2}^{\infty}(a_n \cos(\omega_n t) + b_n \sin(\omega_n t))\right\}}_{Distorsion} \tag{12}$$

Il est évident que le système est linéarisé et que le signal de sortie est une sinusoïde parfaite identique à la forme d'onde de la composante fondamentale (le second terme de la série est nul). Dans le but de réduire le temps de calcul, nous avons coupé la dernière période de la forme d'onde à la sortie puis la composante fondamentale a été extraite en utilisant la relation ci-dessus. N'ayant conservé que la composante fondamentale, il reste uniquement les termes correspondant aux harmoniques non-désirées et à la distorsion du

signal de sortie. On intègre la valeur absolue de cette distorsion sur une période complète, et ce, tout en appliquant la procédure d'optimisation afin de minimiser la fonction de coût E, à chaque itération :

$$E = \int_0^T abs\left(\frac{1}{2}a_0 + \sum_{n=2}^{\infty}(a_n \cos(\omega_n t) + b_n \sin(\omega_n t))\right) \tag{13}$$

Si l'on considère que nous avons un modèle à 2 blocs pour le système et un autre bloc additionnel pour la prédistorsion, la structure complète se compose alors de 3 éléments pour lesquels l'optimiseur doit être capable d'évaluer, assez rapidement, la fonction désirée. Puisque l'évaluation de la fonction est la principale source du temps de calcul, les optimiseurs basés sur le vecteur de gradient ne sont pas favorisés. La méthode MADS (Mesh Adaptive Direct Search) que nous avons utilisée, et qui est implémentée dans NOMAD (Nonlinear Optimization for Mixed variables and Derivatives), se présente sous la forme d'un module utilisable dans MATLAB.

Dans le but de configurer l'estimateur pour éviter les valeurs qui diminuerait le signal d'entrée, une condition doit être ajoutée afin que le signal de sortie soit continuellement dans la région non-linéaire. Cela peut être réalisé en ajoutant une contrainte non-linéaire sous la forme d'une inégalité pour l'optimiseur. Ce dernier est donc défini comme étant la soustraction d'une certaine quantité, que l'on nomme K, par la valeur absolue de la forme d'onde fondamentale extraite à la sortie :

$$cx = K - \int_0^T abs(a_1 \cos(\omega t) + b_1 \sin(\omega t)) \leq 0 \tag{14}$$

Enfin cx est ajouté à l'algorithme. Après avoir ajouté cette contrainte non-linéaire, le problème d'optimisation est alors terminé.

Après avoir vérifié les problèmes de minimisation et de contraintes et choisi l'optimiseur approprié, il est possible d'appliquer le modèle au système final. Puisque le modèle a deux variables d'entrée (en phase et en quadrature), nous devons transmettre deux signaux sinusoïdaux. Ainsi, l'information de phase contenue dans un signal complexe peut être reproduite.

Les modèles mathématiques choisis pour la prédistorsion sont des polynômes avec deux différentes fonctions de base :

$$G_i^{-1}(x_i) = p_1 f_{B1}(x_i) + p_2 (f_{B1}(x_i))^3 + p_3 (f_{B1}(x_i))^5 \tag{15}$$

$$G_q^{-1}(x_{\text{int}}) = p_4 f_{B2}(x_{\text{int}}) + p_5 (f_{B2}(x_{\text{int}}))^2 + p_6 (f_{B2}(x_{\text{int}}))^{2.5} + p_7 (f_{B2}(x_{\text{int}}))^{2.8} \tag{16}$$

ou

$$f_{B1}(x_i) = \tanh^{-1}(Nx_i) \tag{17}$$

$$f_{B2}(x_{\text{int}}) = \text{normpdf}(x_{\text{int}}, \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x_{\text{int}}-\mu)^2}{2\sigma^2}} \tag{18}$$

et où N, qui est le coefficient utilisé pour la prédistorsion de gain, permet de s'assurer qu'il n'y a pas de discontinuités pour la fonction inverse de la tangente hyperbolique. Les coefficients μ, σ, qui sont des variables d'optimisation, sont initialement choisis pour simplifier le calcul numérique. Ainsi, l'algorithme ne traite pas plus de deux variables à chaque itération. Après l'obtention des coefficients par optimisation, on peut évaluer l'efficacité du linéarisateur en utilisant un signal réel du type CDMA (qui a aussi été utilisé lors de la modélisation) et observer les distorsions. Il est important de mentionner que l'onde sinusoïdale, utilisée dans la partie de l'identification, est choisie

pour que la puissance maximale soit égale ou supérieure à celle d'un signal CDMA. Enfin, les courbes AM/AM et AM/PM sont obtenues et comparées à celles obtenues sans prédistorsion. Les resultants démontrent un bon comportement linéaire.

Une fois que la distorsion non-linéaire statique est compensée, il reste la partie dynamique du modèle qui est supposément linéaire, i.e. identifiable en utilisant les technique d'identification des systèmes linéaire disponibles dans MATLAB. Dans ce cas, des structures ARX de différents ordres ont été appliquées pour modéliser le système linéarisé. La concordance entre le modèle et les mesures atteste du bon fonctionnement de l'élément de prédistorsion.

TABLE OF CONTENTS

TABLE OF FIGURES

TABLE OF TABLES

CHAPTER 1

INTRODUCTION

1.1 Motivation

WIDEBAND transmitter design for modern high-speed wireless communication systems, where highly varying envelope signals are launched in the system, is a delicate and complex issue, since it involves compromise among inconsistent requirements, such as high linearity and high power efficiency. A variety of multicarrier transmission techniques are currently employed in modern wideband wireless communication systems, among which are code division multiple access (CDMA), orthogonal frequency division multiplexing (OFDM), worldwide interoperability for microwave access (WiMAX), etc. These modulated signals with non-constant envelopes, introduce large peak-to-average power ratios (PAPRs). When a modulated signal traverses a weakly nonlinear circuit which can be characterized by a low order power series, its in bandwidth frequency content is strengthened by odd-order nonlinearities [1]. This phenomenon, known as spectral regrowth or spectral regeneration, results from the generation of intermodulation products between the individual frequency components of the spectrum whereby coding errors and adjacent channel interferences will kick in.

As a result, as things stand, a power amplifier (PA) must be either operated near its saturation area so as to achieve high system power efficiency (but this comes at the cost of distortion), or instead, at low efficiency, at a large back- off from its nonlinear saturated region, so as to avoid out-of-band emission and spectrum regrowth which

refers to the broadening of the bandwidth of the modulated signal due to the nonlinearities existing in the transmitter. In order to maintain the PA operating at nonlinear area with high power efficiency and, at the same time, to avoid any power-dependent nonlinear phenomena, a compensation approach for the nonlinear distortion that in general, could initiates from any part of the transmitter. The most largely known and developed techniques for linearization mainly fall into feedback [2]-[4], feedforward [5]-[7] and predistortion methods [8]-[10], among which predistortion has frequently come to attention, providing great amount of flexibility end efficiency in suppressing out of band frequencies. Predistortion can be implemented on either digital or analog platform [11]. The general concept of the analog predistortion technique [12] is quite simple: having characterized the nonlinear behaviour of the PAs, the predistorted device is then concatenated to the amplifier to cancel the undesirable nonlinear behaviour by exhibiting an inverse behavior of the PA, so that the overall performance sounds reasonably linear. Figure 1.1 shows a general scheme of an analog predistorter where such a device is followed by an amplifier.

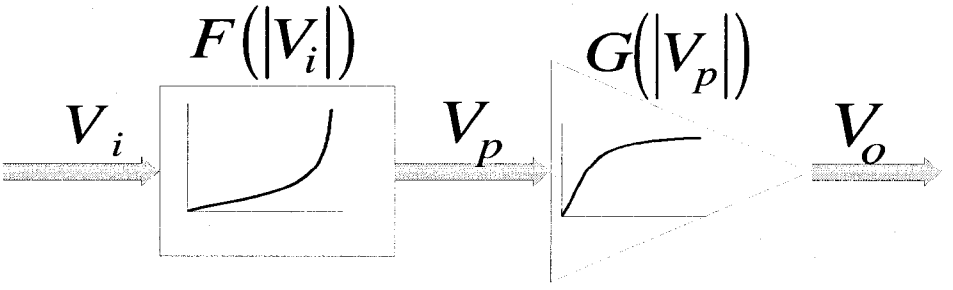


Fig. 1.1. A general scheme of an analog predistorter

In the figure, $F(|Vi|)$is the complex predistortion function and $G(|Vp|)$ is the complex gain of the PA, both of which are functions of the instantaneous magnitude of the input signal. Digital predistortion [13], also known as data predistortion, is based on the modification of the constellation of the input signal, so that the overall performance of the predistorted amplifier resembles an optimal linear amplification at the output. Digital baseband predistortion is among the most efficient methods, due to its high flexibility in digital implementation and its ability to compensate for the nonlinearity in not only the power amplifier, but also the whole transmitter that includes the power amplifier as well as the modulators/demodulators. A general scheme of a digital baseband predistortion is shown in Figure 1.2 where the baseband complex data pass through a digital to analog converter followed by a quadrature modulator which builds the passband signal before amplification. The output of the PA is demodulated by the lower quadrature modulator in the picture and having passed the analog to digital converter, the equivalent baseband signal of the output will be obtained.

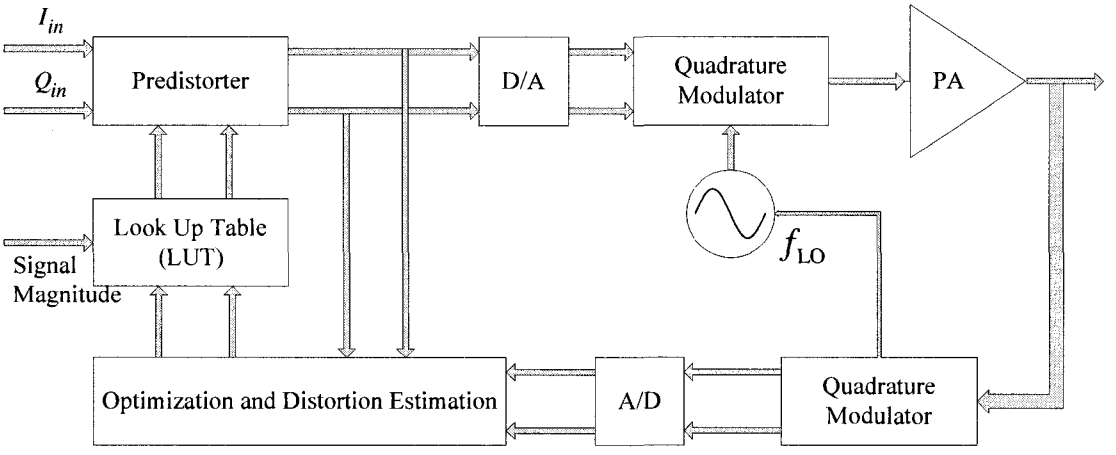


Fig. 1.2. A general scheme of a digital baseband predistortion

The relation between the baseband and passband signal can be briefly expressed as follows:

$$g(t) = x(t) + jy(t) = |g(t)|e^{j\theta(t)} \equiv R(t)e^{j\theta(t)} \tag{1.1}$$

Where $g(t)$ represents the complex baseband signal and:

$$R(t) \triangleq |g(t)| = \sqrt{x^2(t) + y^2(t)} \tag{1.2.a}$$

$$\theta(t) \triangleq \tan^{-1}\left(\frac{y(t)}{x(t)}\right) \tag{1.2.b}$$

The equivalent passband representation of the baseband signal can be expressed as:

$$v(t) = \mathrm{Re}\left\{g(t)e^{j\omega_c t}\right\} \tag{1.3}$$

Considering Nyquist sampling criteria for signal recovery, the modulated passband signal demands a considerably higher sampling rate compared to its baseband representation. Hence, employing baseband signal alleviates this burden and, at the same time, recovers the same information signal. Therefore, baseband digital predistortion is of a relatively higher interest due to a lower demand on the digital signal processing (DSP) data storage.

In this thesis, a new adaptive predistortion technique, based on launching a single frequency signal to the transmitter, is proposed. The key assumption which significantly simplifies the linearization approach is considering the nonlinear characteristics seen in the transmitters to depend exclusively on the power of the driving signal [14] and not its frequency content or bandwidth. In other words the main indicator of the signal is its power and not its frequency contents. This simplification is also referred to as

narrowband approximation [15].The training signal, when compared to various schemes of modulated signals that are largely used for identifying the digital predistortion, is a simple signal and proves to be sufficient as long as we are concerned about power dependent nonlinearities. To establish our linearization criteria, one could simply contemplate the performance of a linear system on a sinusoidal signal, which will produce a sinusoidal signal of the same frequency, incorporated with constant linear gain and linear phase shift.

Two-box nonlinear models, which have been vastly developed to mimic the behaviour of the transmission systems, have been identified and validated. The key assumption though is that a two-box model is an accurate enough representation of the PA nonlinear dynamics, at least for the conditions under which it is assumed to be utilized. The adaptive predistorter is, perforce, structured in accordance to such a model on which it will take action. Further validation of the identified predistorter is accomplished by launching a CDMA signal.

1.2 Nonlinear dynamic behaviour, two- and three-box models

In order to design a predistorter, accurate enough to eliminate the power dependent nonlinear distortions existing in the transmitter system, primarily, an accurate and robust model for the RF transmitter should be identified. The model shall remain valid for a wide enough range of input signal power that the transmitter is expected to be subjected to. For a memoryless, quasi-memoryless or quasi-linear transmitter where the output of the system can be treated for most purposes as a linear version of the input, including

weak nonlinearities [16], one widespread method is the look-up-table (LUT) [17]-[18]. The measured amplitude modulation / amplitude modulation (AM/AM) and amplitude modulation / phase modulation (AM/PM) data of the transmitter are used to build up the LUTs. Once the tables are obtained, the PA behaviour can be characterized, wherefrom, the compensator's nonlinear behaviour will be extracted. One of the main difficulties that one may come across is the requirement of massive amounts of RAM to accumulate a sufficiently accurate mapping.

As long as the signal launched through the system is narrowband, where transmitter exhibits negligible dynamic effects which sway away by the dominant static nonlinearities, an instantaneous static model and its corresponding compensator can linearize the system accurately enough and the LUT method can be efficiently applied. However, once wideband signal transmissions are considered, the transmitter dynamic behaviour can no longer be ignored, whereupon, the output of the system depends not only on the current input but also on the past inputs of the system. This type of systems is also known as system with memory and the consequent symptoms are referred to as memory effects. Therefore we fall into the dynamic nonlinear systems category where static models, such as LUTs, can no longer represent the system satisfactorily.

A variety of methods have been proposed in the past few years to account for the concurrent modeling of the nonlinearity on one hand and dynamic nature of the transmitter on the other hand. Some of these models which cover a rich research area for nonlinear dynamic system are memory polynomial models, Volterra series, two-box models and their 3-box combinations.

1.2.1 Volterra series model

The Volterra series method is a general nonlinear system approach in modeling a nonlinear system with memory [19]-[20]. If the amplifier is observed like a functional box, it can be modeled by either the Taylor or Volterra series. The Taylor series is a good tool for modeling nonlinear behaviour; however, it is only valid for memoryless nonlinearity.

In terms of the Volterra series, the input and output relation of an RF power amplifier can be written as below:

$$y(t)=\int_{-\infty}^{+\infty}h_1(\tau)x(t-\tau)d\tau+\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}h_2(\tau_1,\tau_2)x(t-\tau_1)x(t-\tau_2)d\tau_1 d\tau_2+ \\ \cdots+\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}h_n(\tau_1,\cdots,\tau_n)\prod_{i=1}^{n}x(t-\tau_i)d\tau_1\cdots d\tau_n+e(n) \tag{1.4}$$

where, $y(\bullet), x(\bullet)$ are the continuous-time input and output signals of a nonlinear system respectively; e(n) is the truncation error; and, $h_n(\tau_1,.....\tau_n)$ is the n^{th}-order Volterra kernel, which is comparable to the impulse response of a linear system. A discrete time-domain Volterra series expansion with finite memory length L can be formulated as:

$$y(n)=\sum_{i=0}^{L-1}h_1(i)x(n-i)+\sum_{i=0}^{L-1}\sum_{j=0}^{L-1}h_2(i,j)x(n-i)x(n-j) \\ +\sum_{i=0}^{L-1}\sum_{j=0}^{L-1}\sum_{k=0}^{L-1}h_3(i,j,k)x(n-i)x(n-j)x(n-k)+\ldots+e(n) \tag{1.5}$$

where, $h_1(i)$, $h_2(i,j)$ and $h_3(i,j,k)$ are the linear, quadratic and cubic time-domain Volterra kernels, respectively; *e(n)* denotes the model error; and, L is the system memory length.

In the literature, several methods have been addressed for determining the kernels or the associated transfer functions [21]-[22]. The output of the Volterra model is linear, with respect to the kernels; whereby, many fast least square (LS) algorithms can be implemented to identify the kernels, following the same general procedure utilized for the identification of conventional linear digital filters, with the difference that the input vectors must be expanded. Due to the loss of the time shift in the input data vector, direct application of a linear adaptive algorithm to the Volterra series can significantly increase the complexity; therefore, a non-rectangular structure matrix, known as V-Vector, has been developed that preserves the linear time shift of a nonlinear data vector [23].

More problems arise with higher order Volterra systems, where, even with fast recursive least square (RLS) algorithms, updating the parameters turns out be considerably slow and frustrating which will severely compromise the complexity and accuracy of the model. This problem issues from the size of the matrix which increases both with the order of the Volterra kernel and the memory length of the nonlinear system. The details of this modeling approach are not considered in this thesis.

1.2.2 Memory polynomial model

Multi-branch delayed polynomial function is another technique for modeling nonlinear dynamic systems [24]-[25]. As expressed by the following equation, the input signal and its delayed versions are launched to a static polynomial function; and, the

outcomes of all the branches are added up to form the system final output. Figure 1.3 shows a typical scheme of a memory polynomial block diagram.

$$x(n) = \sum_{p=1}^{P} \sum_{q=0}^{Q} a_{pq} u(n-q) \left| u(n-q) \right|^{p-1} \tag{1.6}$$

In the above formula, for $Q = 0$, memory polynomial reduces to a memoryless model. Otherwise, Q stands for the number of tabs in the delay line; and, P is the order of the polynomial. The memory polynomial is, indeed, a simplified version of the Volterra series. As can be instantly perceived from 1.6, this model is linear with respect to its adjustable parameters and therefore is identifiable with simple linear identification techniques, like linear least squares. Further mathematical derivation on how memory polynomial can be extracted from Volterra series is beyond our goals and the reader is referred to [26]. One should however keep in mind that, occasionally, the algorithm for finding the coefficients of the polynomial functions may not converge.

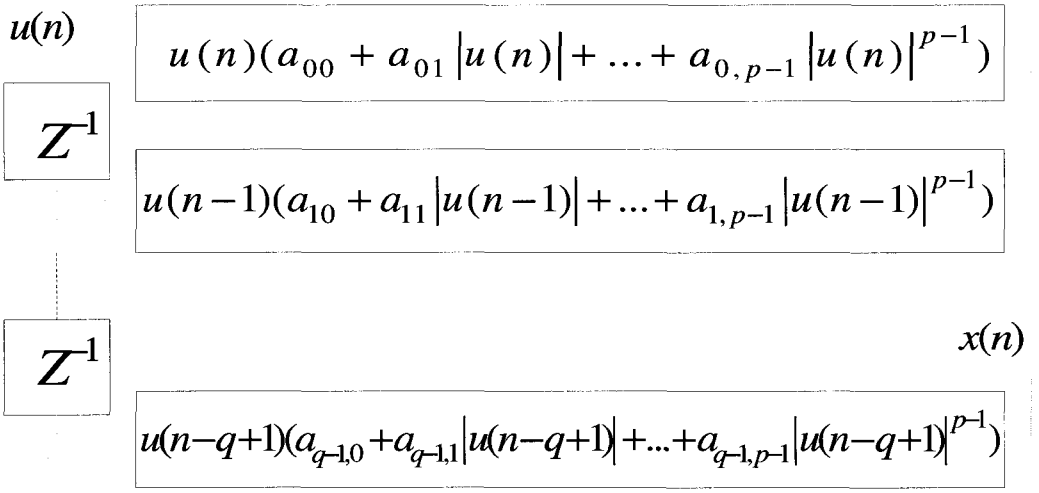


Fig. 1.3. Memory polynomial block diagram

1.2.3 Wiener/Hammerstein and general two-box models

Sometimes the nonlinearities in a system have the character of a static nonlinearity at either the input or output side, while the dynamics are linear. These types of models known as Wiener and Hammerstein are the simplest types of block-oriented models for nonlinear dynamic systems [27], where a nonlinear static block is followed by a dynamic linear filter, or vice versa. These models also prove to be a simplified version of the Volterra series, where the Volterra kernels, assigned as h , are assumed to be separable, i.e. they can be expressed as the product of the first-order kernels:

$$h_p(m_1, m_2, ..., m_p) = \prod_{j=1}^{p} h_{1,j}(m_j) \tag{1.7}$$

For a more elaborate derivation of Wiener/Hammerstein model from Volterra series, the reader is referred to [26]. A general scheme of these models is depicted in Figures 1.4.a and 1.4.b.

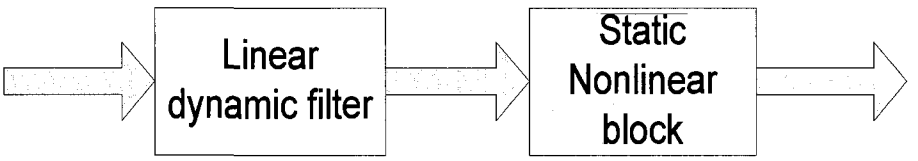


Fig. 1.4.a. Wiener model

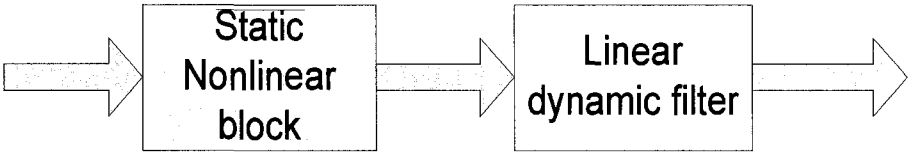


Fig. 1.4.b. Hammerstein model

For an example of Hammerstein model, considering the case in which the static nonlinearity g is known, the input will simply refine to $\bar{x}(t) = g(x(t))$ which will be treated linearly by the time invariant linear filter as the second functional box of this model. When g is unknown, it can be approximated by a polynomial of order m :

$$g(x) = \alpha_1 x + \alpha_2 x^2 + ... + \alpha_m x^m \tag{1.8}$$

Then, we let each power of x pass different numerator dynamics:

$$A(q)y(t) = B_1(q)x(t) + B_2(q)x^2(t) + ... + B_m(q)x^m(t) \tag{1.9}$$

where, $A_i(q)$ and $B_i(q)$ are polynomials in the delay operator q^{-1}. In this case, the adjustable parameter vector can be expressed as follows

$$\theta^T = \left[a_1 ... a_{na} b_1^{(1)} ... b_n^{(1)} b_1^{(2)} ... b_n^{(2)} ... b_1^{(m)} ... b_n^{(m)} \right] \tag{1.10}$$

where, n_a is the memory length. The regressor vector will be expressed as

$$\begin{aligned} \varphi^T(t) = [&-y(t-1) ... - y(t-n_a) x(t-1) ... \\ &x(t-n) x^2(t-1) ... x^2(t-n) ... x^m(t-1) ... x^m(t-n)] \end{aligned} \tag{1.11}$$

Using the above formula, the relation between the input and the output of the system can be written as below:

$$y(t) = \hat{y}(t \mid \theta) = \varphi^T(t)\theta \tag{1.12}$$

If the input/output data as well as the parameters of the filter and the polynomials are real valued, closer scrutiny of the above class of dynamical systems indicates that this type of model can indeed mimic the nonlinearity that occurs in the gain of the system. however, given that the phase shift finds its origin in the dynamical part of the model,

itself modeled as a linear filter, it will be *constant* with respect to input power for narrowband signals. Indeed, the Hammerstein and Wiener models are widely analyzed in the control theoretic literature, and, as discussed, one of their limitations is that they cannot mimic a nonlinear phase response; this is as long as the models are considered to live in the time domain. However, in the radio frequency literature [28]-[31], one can find non linear phase and gain responses, extensively modeled with Hammerstein or Wiener structures, but in contrast, the latter models are considered to live in the frequency domain (and thus act on complex signals). These models lead to good results, but their time domain interpretation is difficult. In our thesis, we shall attempt to work as far as possible with models in the time domain.

Hence, although simple Wiener and Hammerstein structures can be used to model a communication system with weakly nonlinear AM/AM distortion, they fail to seize AM/PM distortion. The two-box model approach addressed in [28]-[31], where nonlinear phase distortion is successfully captured, would follow the same implementation procedure formulated for the Wiener/Hammerstein structures; however, in this case, the model will have complex coefficients.

1.2.4 Three-box modeling approach using an RF signal

In this thesis, a three-box model is introduced and its parameters are estimated, primarily using baseband complex envelope data, which is also known as information signal. In the next step, the baseband signal is encoded into its equivalent passband representation modulated at RF frequency. The outcome is a simple real valued signal

which can be applied within a conventional Hammerstein modeling identification where the gain (AM/AM) distortion is realized by a static polynomial function; and, a finite response filter (FIR) is used to capture the dynamics.

Having identified the first two boxes, the output is decomposed back to its equivalent baseband signal. This step is necessary in order to provide us with direct access to the nonlinearity profile of the phase, using baseband domain data. While preserving the nonlinear gain profile as obtained during the first phase of identification, the third box comes into action by simply manipulating the quadrature and in-phase data vector, so that a good quality of fit for the phase variation is achieved. A scheme of such a model is shown in Figure 1.5. This model and its input/output signals will be more precisely addressed in Chapter 3.

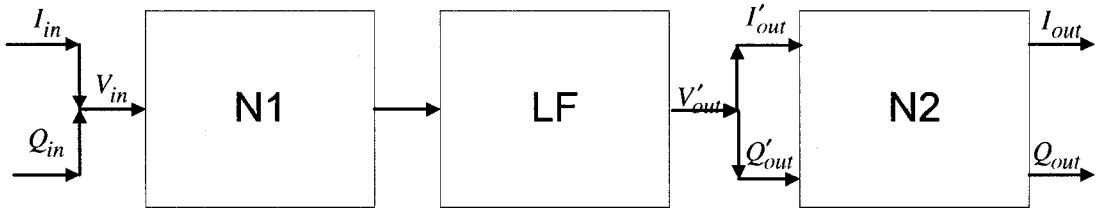


Fig. 1.5. Three-box orientation model

1.3 Objectives and outline of the thesis

1.3.1 Objectives

In this thesis, an already developed nonlinear two-box model is initially constructed for both real and complex data, and its performance is validated. Then, a new three-box model is introduced to capture the nonlinear phase variation which could not be

achieved with real valued Hammerstein/Wiener models. The resulting models' time and frequency domain responses are then discussed from the point of view of performance. Based on the structure of the complex Hammerstein model, a new adaptive predistortion identification, which employs a single frequency signal as the input excitation, is structured and implemented. The key assumption is that the complex Hammerstein model is an accurate representation of the transmitter. In other words, the static power dependent nonlinearities of the transmitter and its linear dynamic behavior can be separated. Therefore the predistorter, in a static function representation, will precedes the 2-box model and will be iteratively updated until it approaches the inverse characteristics of the static functions of the model. One should keep in mind that although a model is being used throughout the identification of the predistorter, it will be treated like a black box.

The optimization technique used to update the predistorter's adjustable parameters is mesh adaptive direct search method (MADS), which is preferred to other numerical algorithms, mainly due to the relatively large and time-consuming computational procedure during the evaluation phase of the cost function. This evaluation is a 3-stage process at each iteration including the predistorter functional box and the 2 boxes embedded in the Hammerstein model. The performance of the built-up predistorter is validated in one last step by launching a CDMA signal to the whole system and extracting the AM/AM and AM/PM curves, if the linearization procedure has worked, should indeed have all the characteristics of represent a linear system.

On one hand, the accurate performance of the built-up adaptive predistorter, which removes the need for decomposing the transmitter into two separable parts and, on the other hand, applying a simple sinusoidal signal, instead of a complicated modulated signal such as CDMA during identification, are explicated and the efficiency of the linearization in terms of suppressing out of band frequencies is compared in the power spectrum of the output of the predistorted transmitter and the transmitter alone. The power spectrum regrowth which is principally caused by the nonlinear characteristics of the transmitter and the noise-like oscillation observed in its power spectrum, which is an unmistakable signature in a nonlinear system, are largely eliminated.

1.3.2 Outline

The remainder of this thesis is organized as follows.

In Chapter 2, the performances of two box models, living respectively in the time domain and the frequency domain are contrasted. The baseband real valued models performance in terms of AM/AM and AM/PM profiles is studied and compared. Next, a two-box model or complex Hammerstein is introduced; and, the construction procedure, based on de-embedding the gain/phase static nonlinear distortions and the residual dispersion like distortion in the data, is explicated in a complex structure using complex baseband signal. Nonlinear functional boxes are modeled with polynomials, embedded with a saturation type basis function like tangent hyperbolic in the hope of achieving a better fit for. The dynamics is modeled with a simple linear filter structure known as

ARX. The complete model is validated on a set of data distinct from that used in the estimation phase.

In Chapter Three, the three-box model is introduced and a two-stage identification procedure is explained. Having constructed the modulated signal from the baseband data, the first two blocks are identified concurrently using a Hammerstein identification procedure. The output of the first two blocks is then demodulated so as to excite the last block, which shapes the profile of the nonlinear phase distortion. The complete model is validated, and the results are compared with those of the two-box model.

In Chapter Four, a new adaptive predistorter based on launching a sinusoidal wave to the system is introduced, and the linearity criteria are discussed. The adaptive predistorter construction is primarily executed on a simple hypothetical model to verify the complexity of the computational task and the optimization process. Having defined the objective function to be minimized, in order to insure that the algorithm will converge to the appropriate values, a nonlinear constraint must be contemplated within the problem where, as will be addressed later, the lack of this constraint will severely mislead the algorithm convergence direction. An adequate optimizer is chosen; and, the 2-box model of the transmitter is utilized as a black box in cascade with the static nonlinear adaptive predistorter, to which the sinusoid signal is applied. In one final step, the performance of the complete predistorted system is verified by launching a CDMA signal; and observing AM/AM and AM/PM curves before and after the predistortion, as well as the power spectrum of the output signal of the system before and after the predistortion.

In Chapter Five, the thesis is summarized, and the contributions of the present work and possible future work are outlined.

CHAPTER 2

TWO-BOX MODELS, DE-EMBEDDING NONLINEARITIES AND DYNAMIC MEMORY EFFECTS

2.1 Introduction

WIDEBAND transmitter design for modern high-speed wireless communication systems is a complicated issue, since one needs to find a compromise between inconsistent requirements. Aside from maintaining the quality of the signal at an acceptable level, the channel capacity and exploiting the best of a limited bandwidth are of essential concern. This delicate issue manifests itself more severely, in the context of wideband and highly varying envelope signals modulations schemes, such as code division multiple access (CDMA) and orthogonal frequency division multiplexing (OFDM).

The linear and nonlinear distortions that are observed throughout a transmitter, and particularly the power amplifier (PA), will compromise operating the PA at a large back-off from its nonlinear region, with a less power efficiency, or exciting the system at a higher power efficiency area, and hence challenging the emerging nonlinear distortions. In order to minimize the nonlinear distortion effects on amplitude (AM/AM) and phase (AM/PM), several techniques have been developed, such as predistortion which acts as a functional box preceding the PA with inverse characteristics of the PA. In order to design a predistorter that can effectively suppress the nonlinear effects, an accurate model of the RF system would be of great help.

Several techniques can be explored to model the dynamic nonlinear behaviour of PAs. As an example, the authors in [32] employed a two-box model and identified the behavior of the PA with a static nonlinear block followed by a dynamic weakly nonlinear filter, also referred to as an augmented Hammerstein model. The authors in [23] have used Volterra series to model PAs with memory effect, which is limited to weakly nonlinear devices (nonlinearities restricted to low order polynomials), due to the considerably expensive computational task and complicated algorithm. In [24], the Volterra series was exploited, along with a novel concept known as V-Vector algebra to ease the burden of computational complexity of the previous algorithms. The authors in [33] utilized a Wiener-Hammerstein model for power amplifiers in the frequency domain, using a polynomial to model the static nonlinearity. In [34], the Wiener-Hammerstein model and its sub-models were introduced, and a link between the generalized Volterra series and these models was studied in details.

In the basic structures of the Wiener/Hammerstein type models, there exists one static nonlinear block which is followed or preceded by a time invariant linear filter. These structures are purported to model both nonlinear amplitude and phase distortions; however the latter cannot be modeled by the real valued Hammerstein and Wiener models, or any cascade combination of these models. Therefore, they fall within a general nonlinear block-oriented model category or *complex valued* Wiener/Hammerstein models.

In this chapter, the construction and accuracy of a widely used model for power amplifiers' dynamic nonlinear behaviour is studied. Also, as mentioned in the

introduction, we would address the totally different natures of the nonlinear complex Wiener/Hammerstein models that can account for the nonlinear phase distortion as is seen in wireless communications systems, as opposed to the real-valued Wiener and Hammerstein models, which do not introduce any nonlinear phase behaviour. The first type of models is identified using complex baseband signal and consequently the identified static and dynamic functional boxes will be associated with complex valued coefficients while the real Hammerstein/Wiener models will be treated the real-valued passband signal thus leading to real parameters. Since in many applications, phase distortion might be of little concern and amplitude/amplitude (AM/AM) distortion is the main issue, real-valued Wiener or Hammerstein models can still be used to mimic system characteristics, where, the instantaneous amplitude dependence of the nonlinear behaviour is modeled with static functions separately from the system memory.

2.2 Transmitter prototype

The device under test is a three-stage laterally diffused metal oxide semiconductor (LDMOS) power amplifier with a RF linear LDMOS amplifier designed for class AB amplifier application, MHPA21010, followed by RF power field effect transistors (FETs), MRF21045 and MRF21085, as shown in Figure 2.1. The FETs are designed to be used in AB class for WCDMA base station applications up to 2170MHz. The respective biases are 28V (550mA), 28V (500mA) and 28V (1000mA); and, the RF carrier is at 2140MHz.

The drive signal is a two-channel WCDMA signal of 3 slots, and the peak-to-average power ratio (PAPR) is 11.12 for the input signal. The bandwidth (BW) is 3.84MHz/channel with 5MHz channel separation. The oversampling rate is 24, so the sampling frequency is 92.16. The actual setup, from which the input and output data are obtained, includes the up-conversion and down-conversion, modulation and demodulation blocks, and the main amplification stage

.

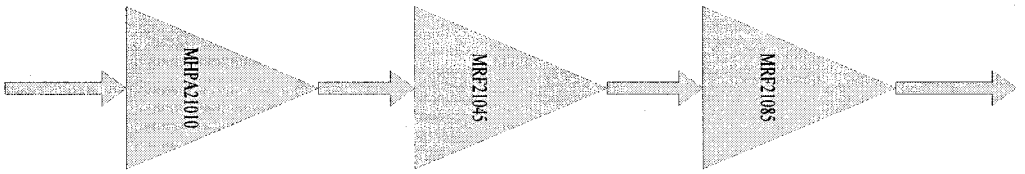


Fig. 2.1. Three-stage LDMOS power amplifier with MHPA21010 followed by MRF21045 and MRF21085

2.3 Hammerstein and Wiener model construction with passband signal

In this section, the Wiener and Hammerstein models for the transmitter prototype, as briefly described in the previous section, are identified and compared for the passband signal excitation. The identification procedure is accomplished in one stage, where the static nonlinear function parameters and the linear filter coefficients which constitute the set of adjustable variables, are identified concurrently through a MATLAB multidimensional unconstrained optimization root in a conventional global approach The minimization algorithm implements the Nelder-Mead method, simplex method or downhill simplex method, which are commonly used nonlinear optimization algorithms.

In order to model the system with a real-valued Hammerstein/Wiener model, a real signal should be used for the both input and output data strings; hence, using the measured baseband complex data, we mathematically construct the modulated signal. Considering x(t) and y(t) to be the in-phase and quadrature components of the baseband complex envelope signal, it can be expressed as follows:

$$g(t) = x(t) + jy(t) = |g(t)|e^{j\theta(t)} \equiv R(t)e^{j\theta(t)} \tag{2.1}$$

where,

$$x(t) = \mathrm{Re}\{g(t)\} \equiv R(t)\cos(\theta(t)) \tag{2.2.a}$$

$$y(t) = \mathrm{Im}\{g(t)\} \equiv R(t)\sin(\theta(t)) \tag{2.2.b}$$

And

$$R(t) \equiv |g(t)| = \sqrt{x^2(t) + y^2(t)} \tag{2.3.a}$$

$$\theta(t) \equiv \tan^{-1}\left(\frac{y(t)}{x(t)}\right) \tag{2.3.b}$$

Any physical passband waveform can be represented by:

$$v(t) = \mathrm{Re}\{g(t)e^{j\omega_c t}\} \tag{2.4}$$

where, ω_c represents the associated carrier frequency. Furthermore, the two other equivalent representations are:

$$v(t) = R(t)\cos[\omega_c t + \theta(t)] \tag{2.5.a}$$

$$v(t) = x(t)\cos(\omega_c t) - y(t)\sin(\omega_c t) \tag{2.5.b}$$

Using the experimental data x(t) and y(t) for both input and output, the equivalent

passband signals in Equation 2.5 are obtained and applied for the identification of the model's parameters. One should recall that these passband signals are artificially constructed in order to test the ability/inability of a simple real-valued Hammerstein/Wiener structure to mimic gain and phase distortions on one hand, and on the other hand, in order to compare their response to that of a complex structure. Hence, the data used throughout this section, for the identification procedure, are real values. Having estimated the parameter of the system, the equivalent baseband signal of the output of the model is extracted and compared to the corresponding measured signal.

2.3.1 Wiener model

In the Wiener model, a static nonlinear block is preceded by a linear filter, which accounts for the dynamics as previously shown in Figure 1.5a. Before any arbitrary choice for the static function to model the nonlinearity of the system, the AM/AM and AM/PM distorted curves, extracted from the measurements, are drawn to help with a better estimation of the possible mathematical function that exhibits similar behaviour. These curves are obtained using the following simple relationships between in-phase (I) and quadrature (Q) components of the input and output measured data and the gain and phase shift of such a system:

$$P_{in} = 10\log_{10}\left(\frac{(I_{in}^2 + Q_{in}^2)}{2R_{in}}\right) + 30 \qquad (2.6.a)$$

$$P_{out} = 10\log_{10}\left(\frac{(I_{out}^2 + Q_{out}^2)}{2R_{out}}\right) + 30 \qquad (2.6.b)$$

where, the power of input and output signals is expressed in dBm.

$$Gain = P_{out} - P_{in} \tag{2.7}$$

$$\Delta\varphi = \tan^{-1}\left(\frac{Q_{out}}{I_{out}}\right) - \tan^{-1}\left(\frac{Q_{in}}{I_{in}}\right) = \tan^{-1}\left(\frac{I_{in}Q_{out} - I_{out}Q_{in}}{I_{in}Q_{in} + I_{out}Q_{out}}\right) \tag{2.8}$$

Using the above equations, the AM/AM and AM/PM curves of the measurements are computed and drawn in Figure 2.2. It is worth mentioning that the measured output data is normalized to the small signal linear gain. As can be clearly seen, the AM/AM curve, expressing the gain behaviour of the transmitter, starts to deviate from a primarily constant value and decreases when the input power reaches a certain amount. As we are mostly concerned about the gain distortion, it would be helpful to plot the output power versus the input power; and, the nonlinear behaviour of the transmitter will manifest itself more evidently. This nonlinear behaviour is shown in Figure 2.3.

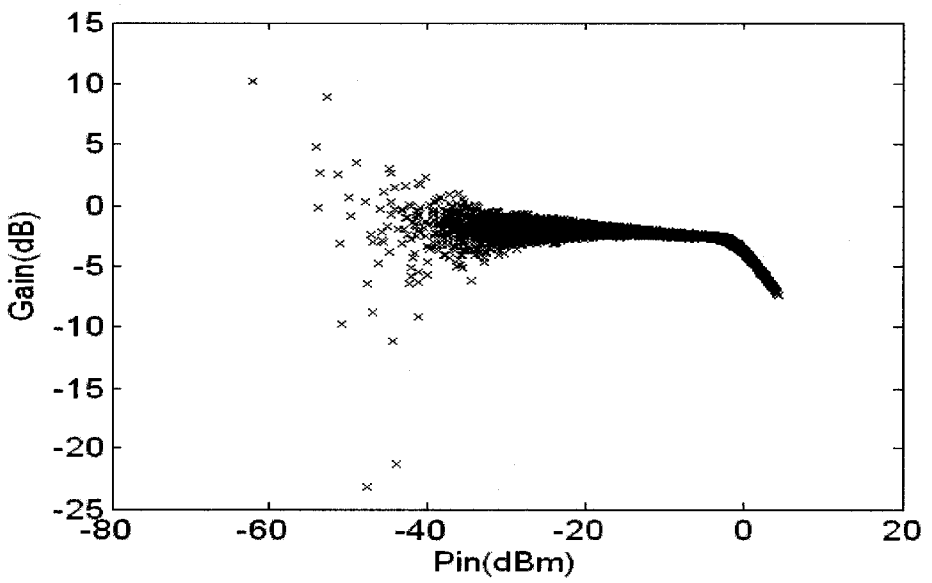


Fig. 2.2.a. The AM/AM measurement data for the transmitter prototype

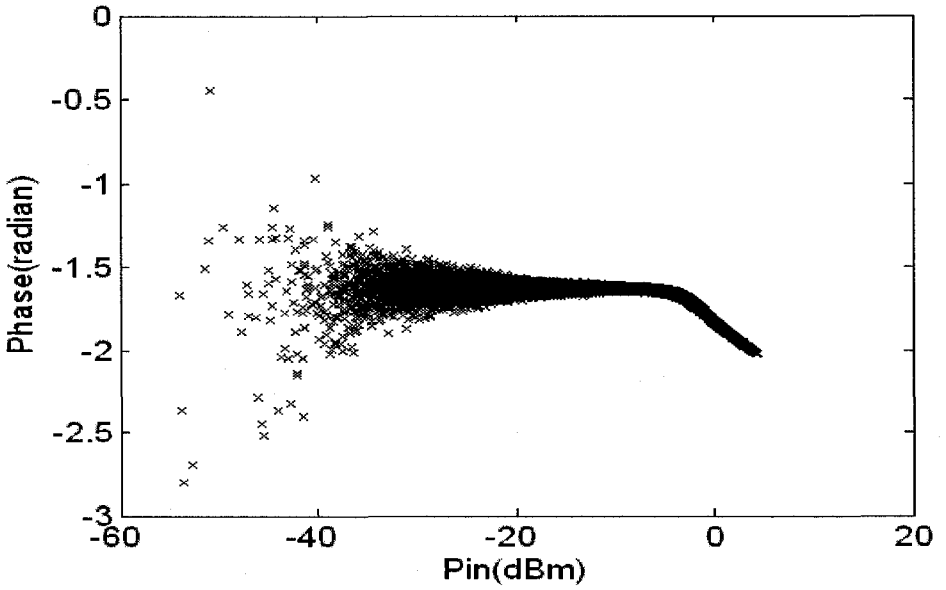


Fig. 2.2.b. The AM/PM measurement data for the transmitter prototype

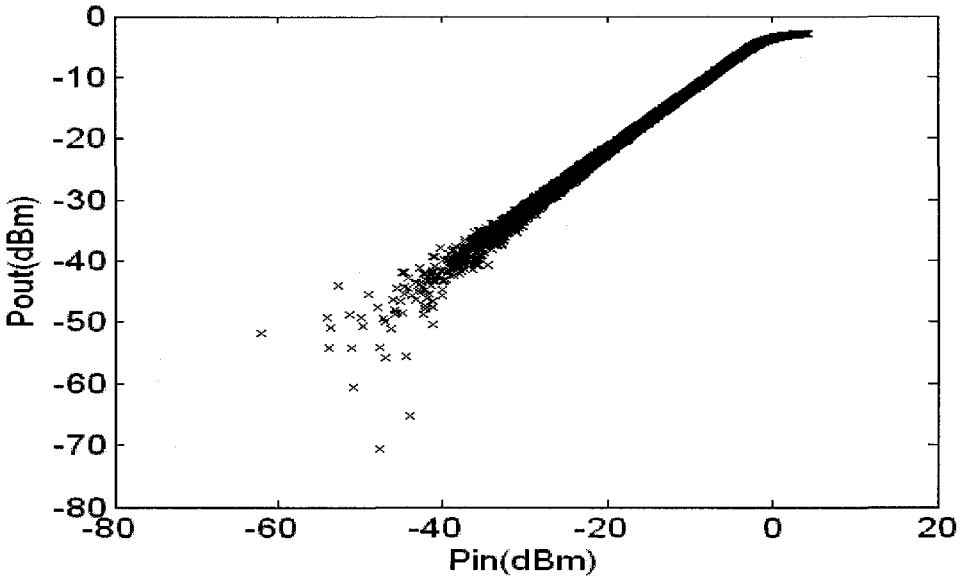


Fig. 2.3. Transmitter output power versus the input power

It can be seen that, as the input power level increases, the power amplifier reaches its saturation area, where the output power no longer exceeds a certain maximum amount and consequently the gain starts to reduce. A polynomial with odd terms, in general, is a

good model for such behaviour since the input and output powers have an odd symmetry relation; however, it may exhibit some instability and low quality of fit when the input power level increases. Hence, in order to shape the polynomial more accordingly with the behaviour observed in the actual transmitter, a mathematically known function of the similar nonlinear profile is embodied within each polynomial term as the basis function:

$$f(\beta(y(t))) = \alpha_0 + \alpha_1\beta(y(t)) + \alpha_2\beta^2(y(t)) + ... + \alpha_n\beta^n(y(t)) \tag{2.9}$$

where in the above formulation, even terms are set to zero; and, $\beta(y(t))$ is the basis function, which in our case, can be tangent hyperbolic, tangent inverse or any other function with similar behaviour. In our case, tangent hyperbolic function proves to provide relatively better results in both the Wiener and Hammerstein models. Hence, we fix $\beta(y(t)) = \tanh(y(t))$, where $y(t)$ is the real passband input of the static nonlinear block or the output of the linear filter.

The transfer function model of the linear time invariant filter, which, is assigned to take the simplest input-output relationship, described as a linear difference equation known as the ARX model:

$$\begin{aligned} &\hat{y}(t) + a_1\hat{y}(t-1) + ... + a_{n_a}\hat{y}(t-n_a) = \\ &b_0u(t) + b_1u(t-1) + ...b_{n_b}u(t-n_b) + e(t) \end{aligned} \tag{2.10}$$

where, $e(t)$ is the white noise term; $\hat{y}(t)$ is the prediction of the linear model for the output at time t; and, u is the input of the filter. Figure 2.4 shows the connection of the two blocks. The coefficients of the linear filter and those of the nonlinear static function are iteratively updated, using a global nonlinear optimization algorithm, such that the

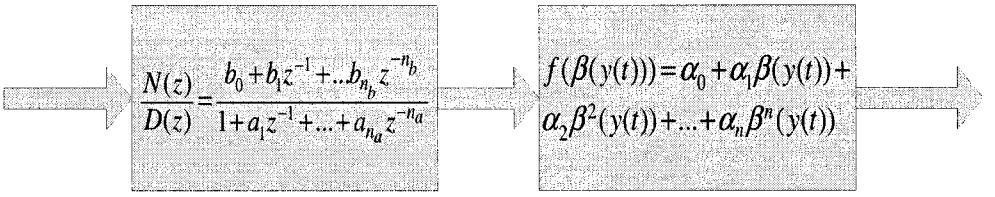


Fig. 2.4. Representation of the Wiener model's functional boxes

error is minimized. The final time plot of the estimated in-phase and quadrature components of the output of the model and the measurement are shown in Figures 2.5.a and b. Having identified these components, the estimated AM/AM and AM/PM curves are computed and plotted in Figures 2.6 and 2.7. As already mentioned, the relative phase shift prediction of the model versus the input power in figure 2.7 remains constant.

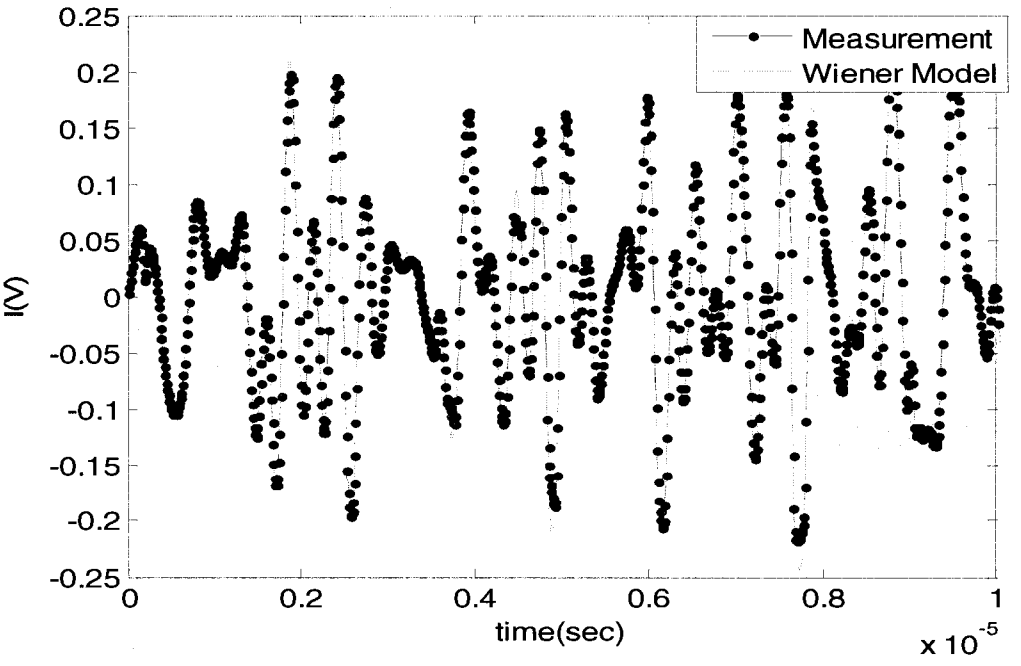


Fig. 2.5.a. In-phase components of the output of the model and the measurement

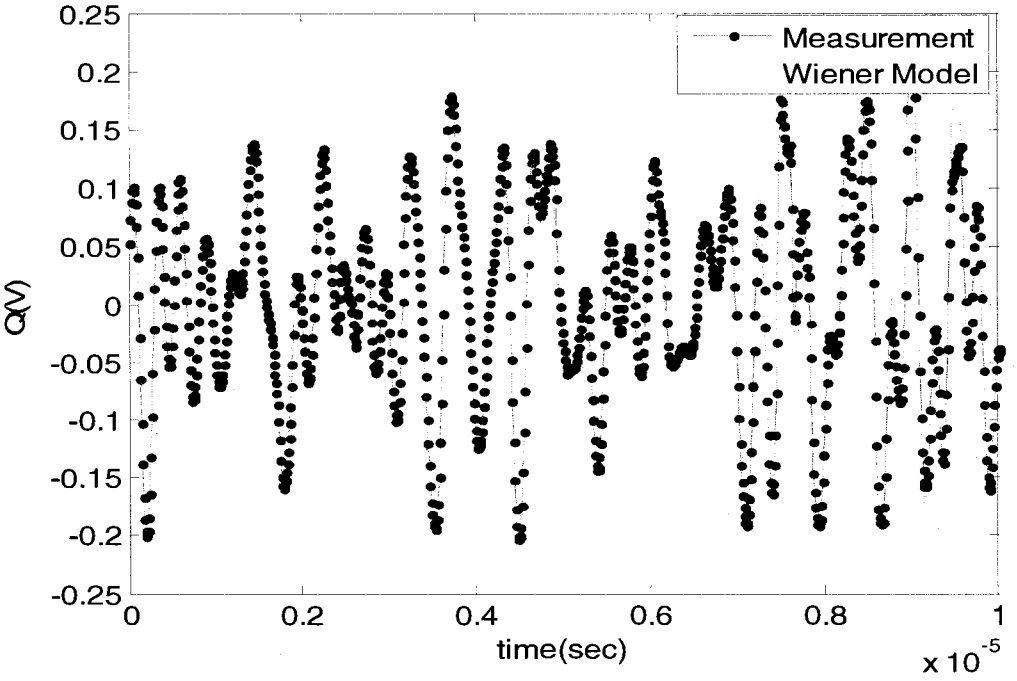


Fig. 2.5.b. Quadrature components of the output of the model and the measurement

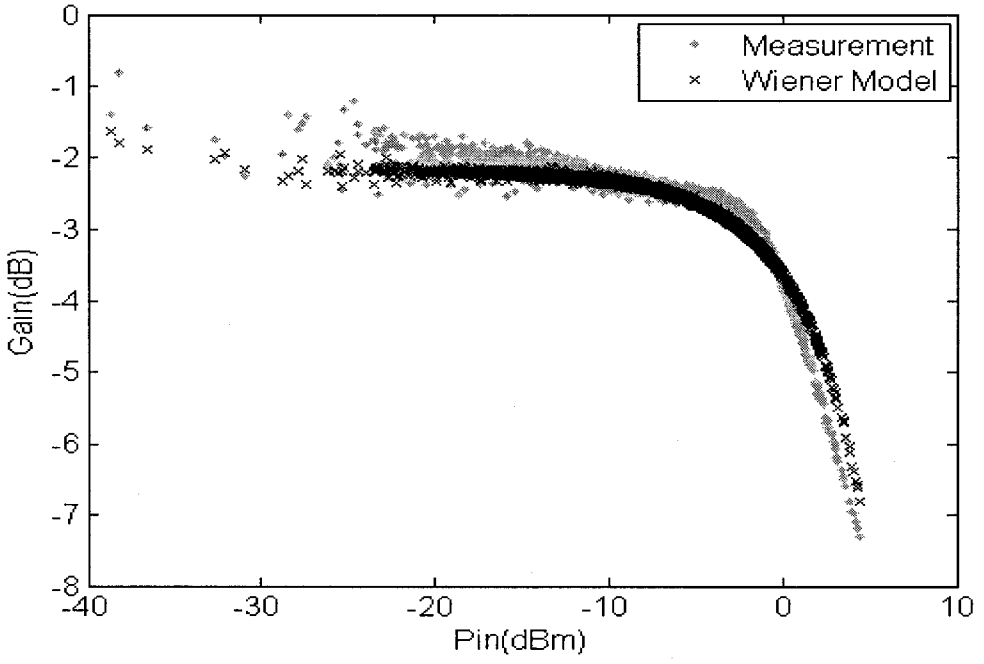


Fig. 2.6. The AM/AM estimation of the Wiener model and the measurement

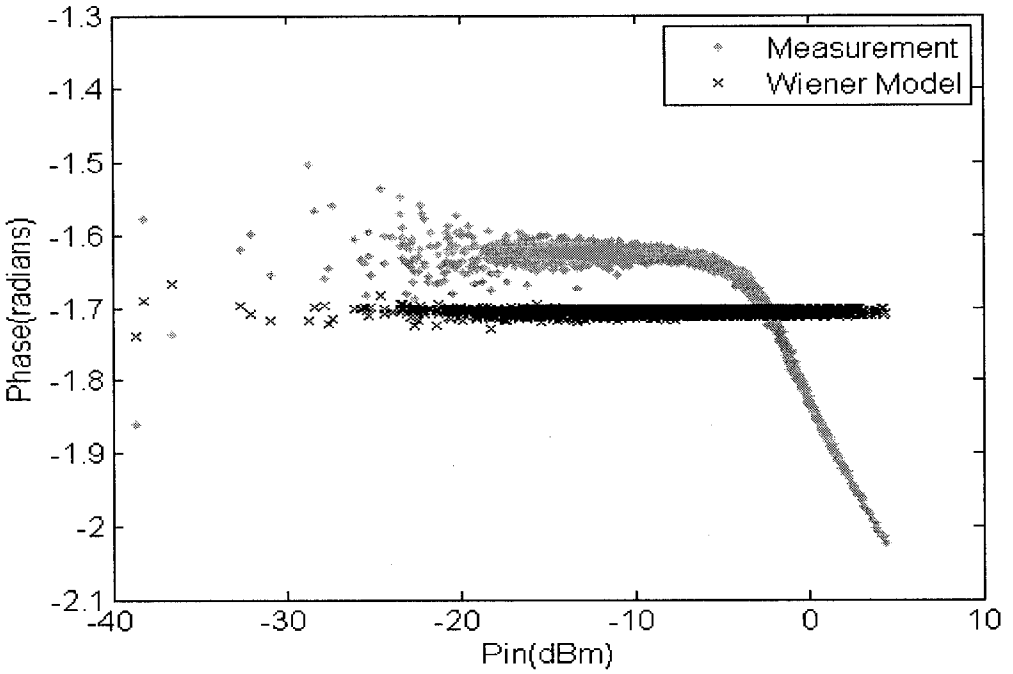


Fig. 2.7. The AM/PM estimation of Wiener model and the AM/PM measurement

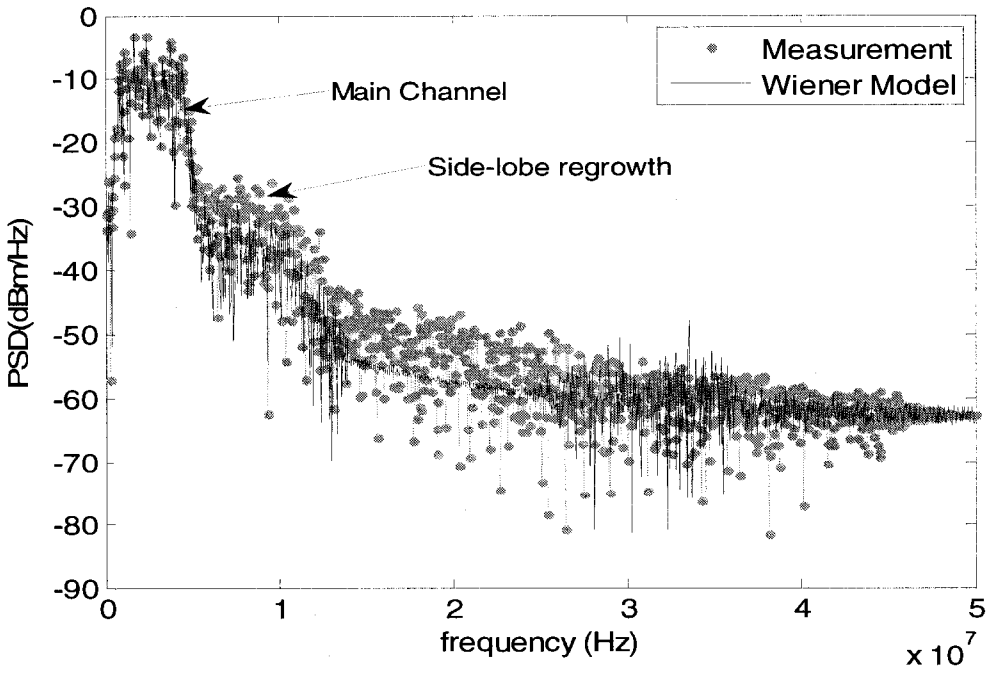


Fig. 2.8. Power spectrum density of the output signal of the Wiener model and the measurement

For a comparison in the frequency domain, the power spectrum of the real part of the baseband signal is plotted in Figure 2.8, where a high quality of fit in the main channel has been achieved and also the adjacent channel spectral broadening, due to the nonlinearity, is very well mimicked by the model. The quality of fit for the Wiener model, containing 6 adjustable parameters, can be computed using the following equation:

$$FIT = \left[1 - Norm(Y - \hat{Y}) / Norm(Y - Mean(Y))\right] * 100 \tag{2.11}$$

where, Y is the vector of the measured output, and $\hat{Y}$ is the model prediction for the output signal. The computational results are presented in Table 2.1.

Table 2.1. Quality of fit figures for the Wiener model

Component	Number of parameters	Best fit percentage
In-Phase	6	90.1495
Quadrature	6	87.6357

2.3.2 Hammerstein model

The process of identifying the Hammerstein model consists of the very same steps as the Wiener model, with the difference that the static nonlinear function will precedes the linear dynamic filter, as shown in Figure 2.9. The basis functions, the order of the static polynomial, as well as the filter order, remains the same as the Wiener model.

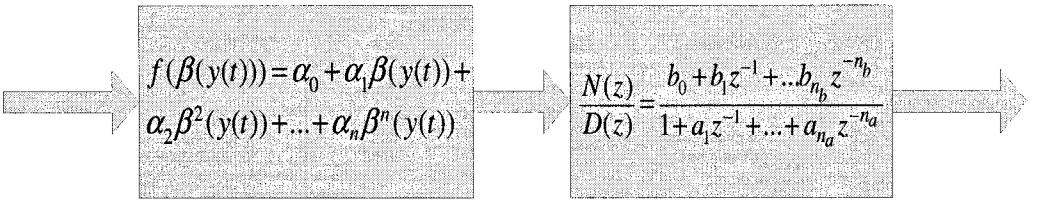


Fig. 2.9. Representation of the Hammerstein model's functional boxes

The time domain plots of the identified signals in baseband representation, in-phase and quadrature components, and their corresponding measured signals are depicted in Figures 2.10.a and 2.10.b. As a reminder, the identification is performed on the real passband signal and at the end of the identification phase; the signal is decomposed to its baseband components. From these baseband components, the AM/AM and AM/PM

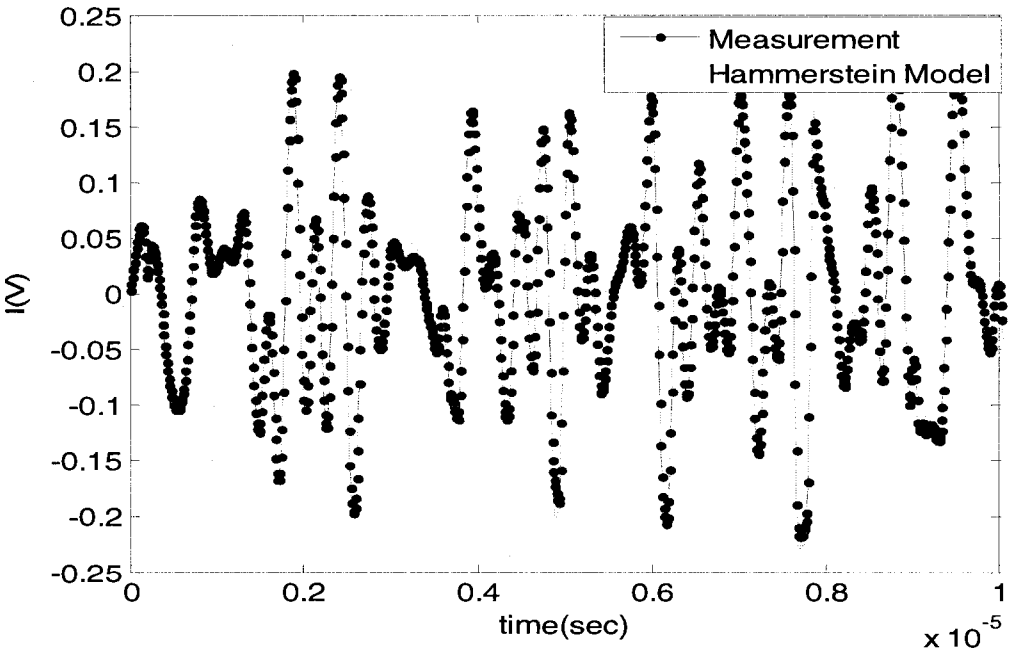


Fig. 2.10.a. In-phase component of the output of the model and the measurement

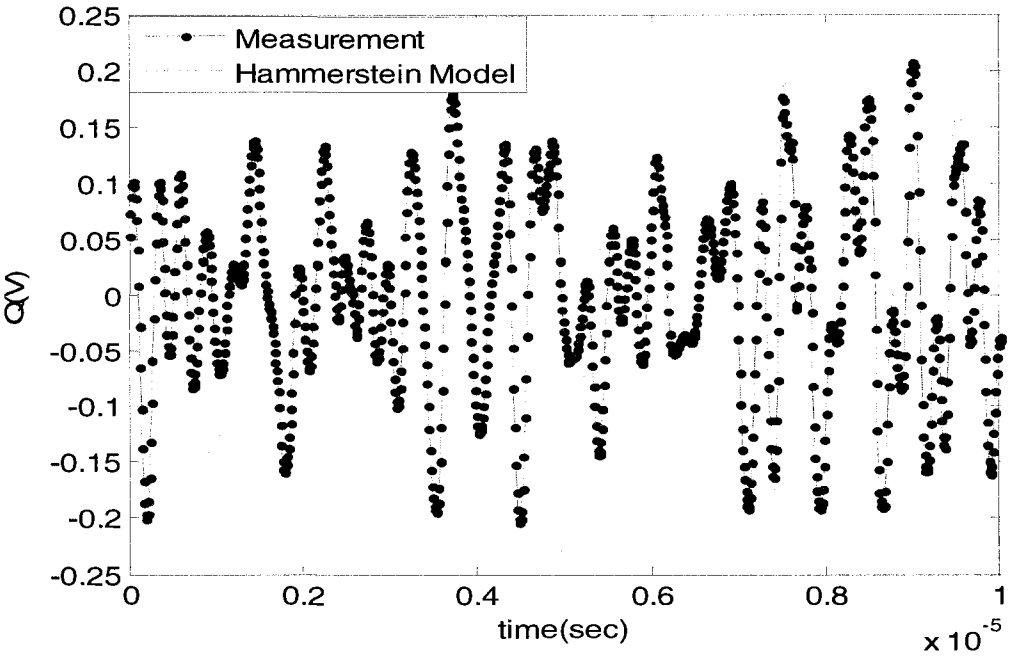


Fig. 2.10.b. Quadrature component of the output of the model and the measurement

curves are obtained and drawn in Figures 2.11 and 2.12. The quality of fit can be once again computed from equation 2.11. The final results are given in Table 2.2, which compared to the results in Table.2.1 demonstrate a relatively close performance for the Wiener and Hammerstein models. An increment of about one percentage can be seen for both the in-phase and quadrature components in Hammerstein model, which can be also perceived by observing the AM/AM curve, where the Hammerstein model follows the nonlinear behaviour more accurately as the input power increases. However, the Wiener model starts to deviate from the system nonlinearity at the same power level, resulting in a bigger error and, therefore, a lower quality of fit. Hence we can roughly conclude that for the very same order of model and the complexity of the basis functions,

Hammerstein shows a slightly better performance in terms of nonlinear behaviour modeling as the input power increases.

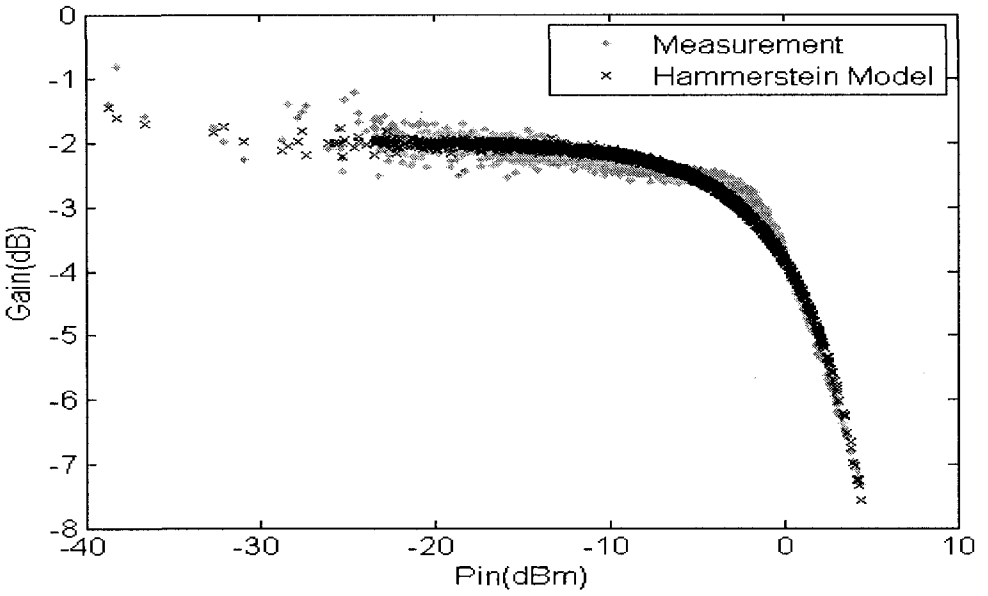


Fig. 2.11. The AM/AM estimation of the Hammerstein model and the AM/AM measurement

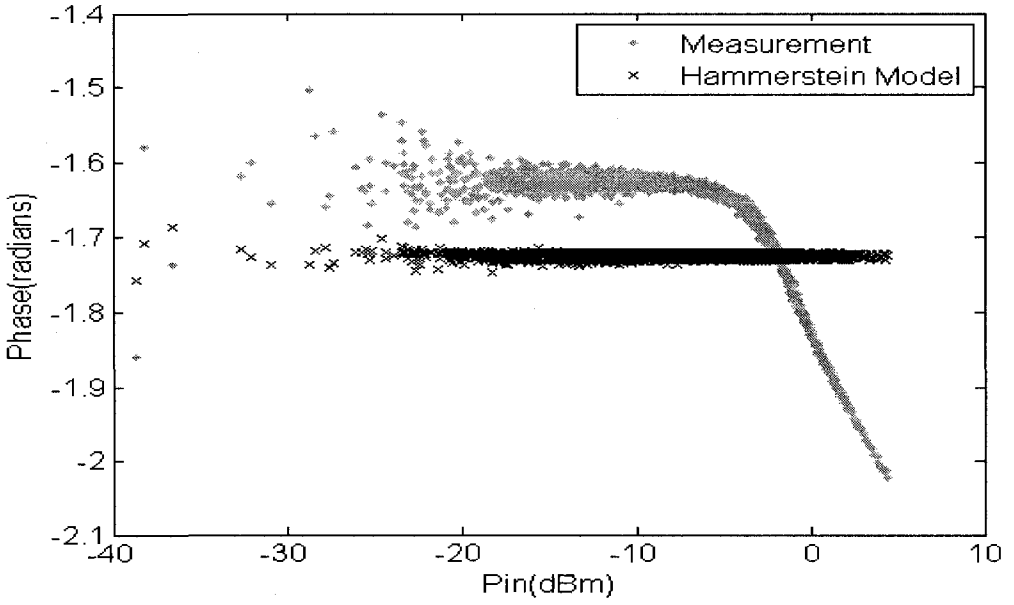


Fig. 2.12. The AM/PM estimation of the Hammerstein model and the AM/PM measurement

Table 2.2. Quality of fit figures for the Hammerstein model

Component	Number of parameters	Best fit percentage
In-Phase	6	90.5226
Quadrature	6	88.2062

A frequency-domain comparison between the power spectrum of the real part of the baseband signal of the measurement and the output of the model is drawn in Figure 2.13, where, a high quality of fit for the main channel near zero frequency, and the adjacent channel regrowth is successfully achieved.

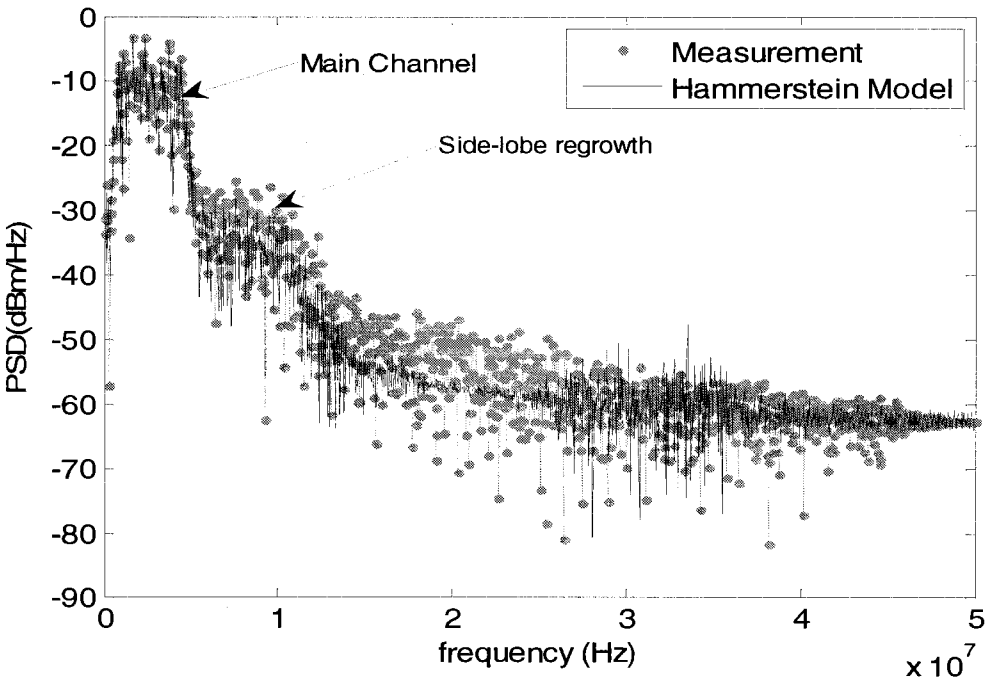


Fig. 2.13. Power spectrum density of the output signal of the Hammerstein model and the measurement

2.3.3 Two-box nonlinear model--Complex Hammerstein

A general nonlinear two-box model or a complex Hammerstein model, widely addressed in RF literature, is introduced and implemented in this section. Due to its complex structure, it can comprehend both the nonlinear profile observed in gain and phase variation of the transmitter. Phase distortion is undesirable in many telecommunication systems where the information is encoded in the phase of the signal. In this modeling (one can also refer to [35]-[36]), as previously mentioned, the identification procedure is simplified by assuming that the instantaneous power of the input signal is the main source of nonlinearities and therefore AM/AM and AM/PM curves, which are directly obtained using baseband data, are the objective of modeling instead of applying a direct identification on quadrature and in-phase components. Having identified these nonlinear profiles, with appropriate static models, the polar-rectangular relation, as described in the following formulas, is used to extract the real and imaginary part of the equivalent baseband signal that the model has provided.

$$Gain = P_{out} - P_{in} \tag{2.12}$$

$$\Delta\varphi = \tan^{-1}\left(\frac{Q_{out}}{I_{out}}\right) - \tan^{-1}\left(\frac{Q_{in}}{I_{in}}\right) = \tan^{-1}\left(\frac{I_{in}Q_{out} - I_{out}Q_{in}}{I_{in}Q_{in} + I_{out}Q_{out}}\right) \tag{2.13}$$

$$\begin{cases} Gi = Gain * cos(\Delta\varphi) \\ Gq = Gain * sin(\Delta\varphi \end{cases} \tag{2.14}$$

$$I_{out} = I_{in} * Gi - Q_{in} * Gq \tag{2.15}$$

$$Q_{out} = I_{in} * Gq + Q_{in} * Gi \tag{2.16}$$

Static polynomial model identification for both AM/AM and AM/PM, in terms of the input power, is performed as the first step. The baseband outputs are then extracted and fed to a delay line, like finite impulse filter structure. Therefore, the identification process is constituted of two steps. First, the instantaneous phase and gain nonlinear distortion are de-embedded and modeled as functions of input power. Second, using these static functions, the intermediate variables at the input of the linear filter which were primarily inaccessible, are computed. The AM/AM and AM/PM relation between the measured output and the computed intermediate variables can be observed which depicts sparse data that are concentrated around a relatively linear line, implying a linear behaviour. This can be, in one last step, modeled by a linear dynamic filter structure, which provides us with the sparsity that is seen in the measured output.

The static nonlinear modeling is performed, applying polynomial functions, embedded with certain basis functions that can shape nonlinear profiles of AM/AM and AM/PM curves more accordingly. A normal probability density function, which is available in the MATLAB function list, is selected as the basis of the polynomials. Since the AM/AM and AM/PM curves have similar type of nonlinear behaviour, the basis function is the same for both:

$$f(\beta(y(t))) = \alpha_0 + \alpha_1\beta(y(t)) + \alpha_2\beta^2(y(t)) + \ldots + \alpha_n\beta^n(y(t)) \tag{2.17}$$

where,

$$\beta(y(t)) = normpdf(y(t)) \tag{2.18}$$

$$normpdf(y(t)\,|\,\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(y(t)-\mu)^2}{2\sigma^2}} \tag{2.19}$$

The standard form of the normal probability density function is used, in which $\mu = 0$ and $\sigma = 1$.

For both gain and phase, the polynomial is of the second order; and, as can be seen in Figures 2.14 and 2.15, the estimated function is a one-to-one function and, hence, is reversible, which is necessary for building up the predistorter. However, this is not our main concern at this stage and the linearization approach will be performed independently of the information of the model contents.

Having identified the static nonlinearities of the first box, the initially inaccessible intermediate variables, at the output of this box which will act as the excitation signal for the linear filter part can be computed. The AM/AM and AM/PM curves, obtained for the computed intermediate variables and the measured output are depicted in Figures 2.16 and 2.17, where gain and phase remain relatively constant as the input power varies. A simple linear filter structure, known as auto-regressive model, of the same type mentioned in Equation 2.10, is applied to model this linear part. This linear filter in cascade with the static nonlinear functions completes the model. The overall performance of the model in the time-domain plots of the quadrature and in-phase components is shown in Figures 2.18 and 2.19.

The equivalent AM/AM and AM/PM curves were extracted and drawn, as shown in Figures 2.20 and 2.21. In a similar step as in previous sections, the quality of fit for the quadratic components of the baseband signal is computed using Equation 2.11.

It is worth pointing out that a small portion of the data, at the output of the model, as can be seen in the AM/PM curve in Figure 2.21, have fallen on the zero axes. This results

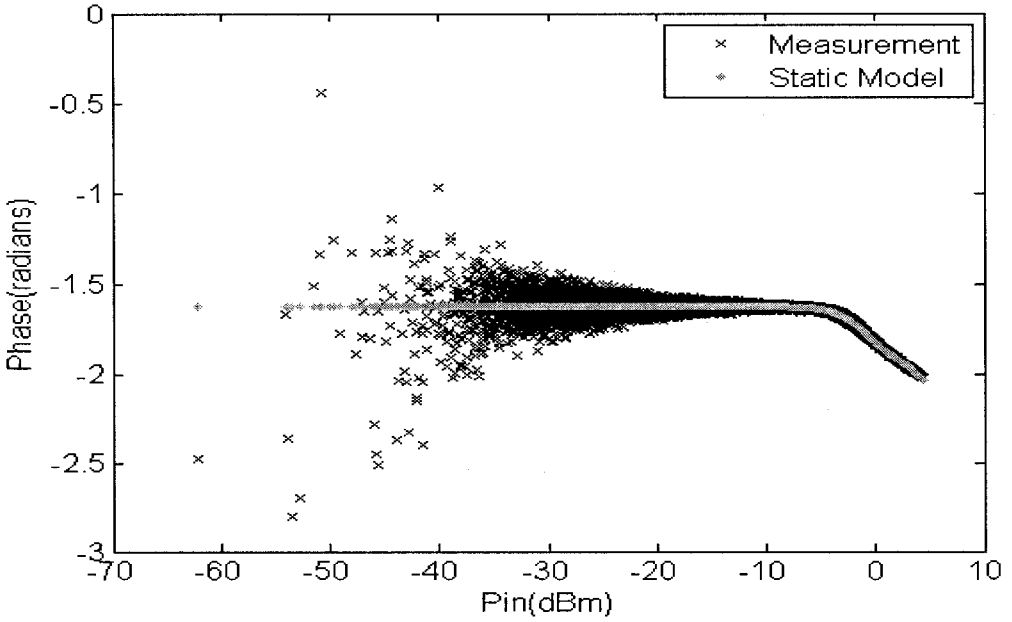


Fig. 2.14. The static AM/PM estimation of the two-box model and the AM/PM measurement

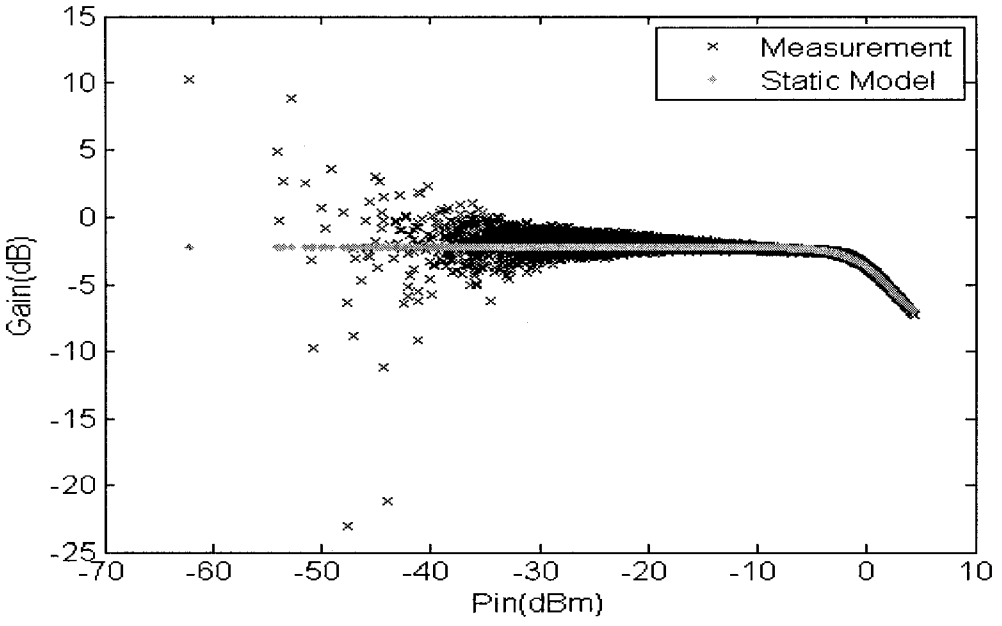


Fig. 2.15. The static AM/AM estimation of the two-box model and the AM/AM measurement

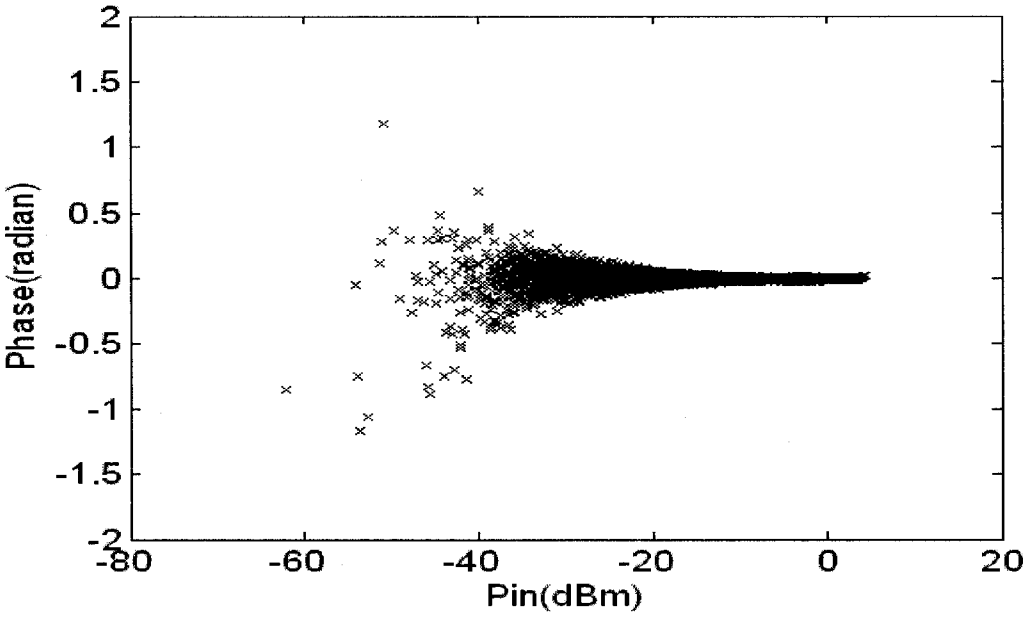


Fig. 2.16. The linear dynamic AM/PM characteristics of the transmitter extracted from the measurement

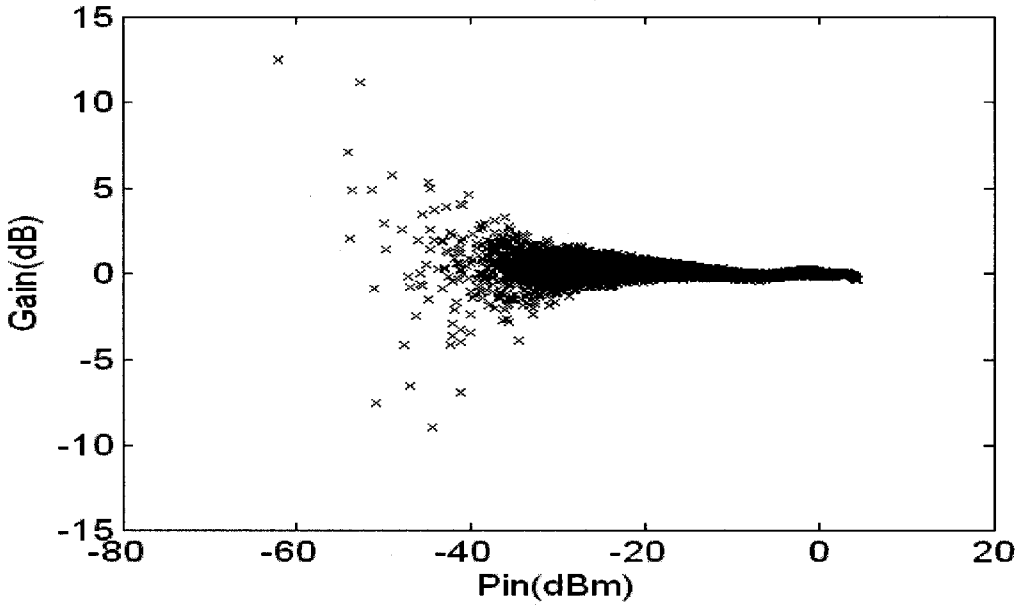


Fig. 2.17. The linear dynamic AM/AM characteristics of the transmitter extracted from the measurement

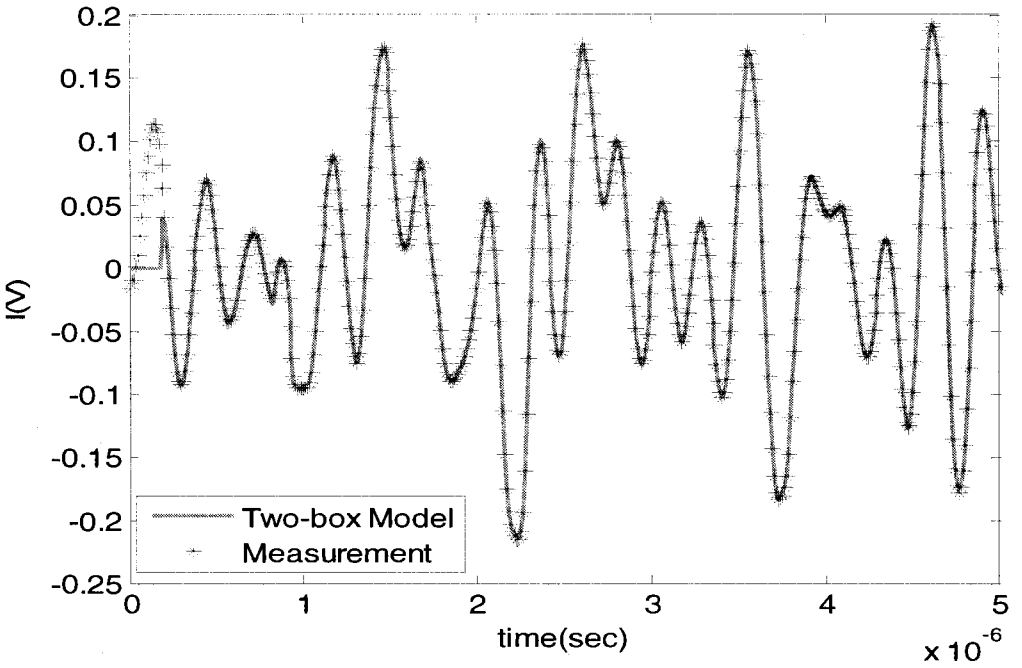


Fig. 2.18. The in-phase component of the output of the model and the measurement

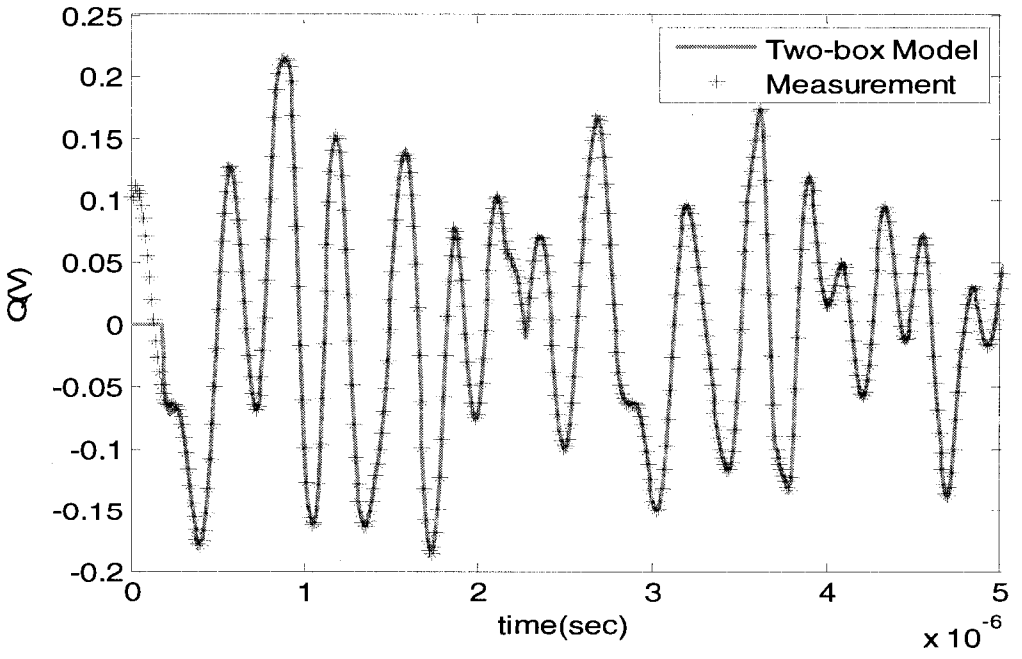


Fig. 2.19. The quadrature component of the output of the model and the measurement

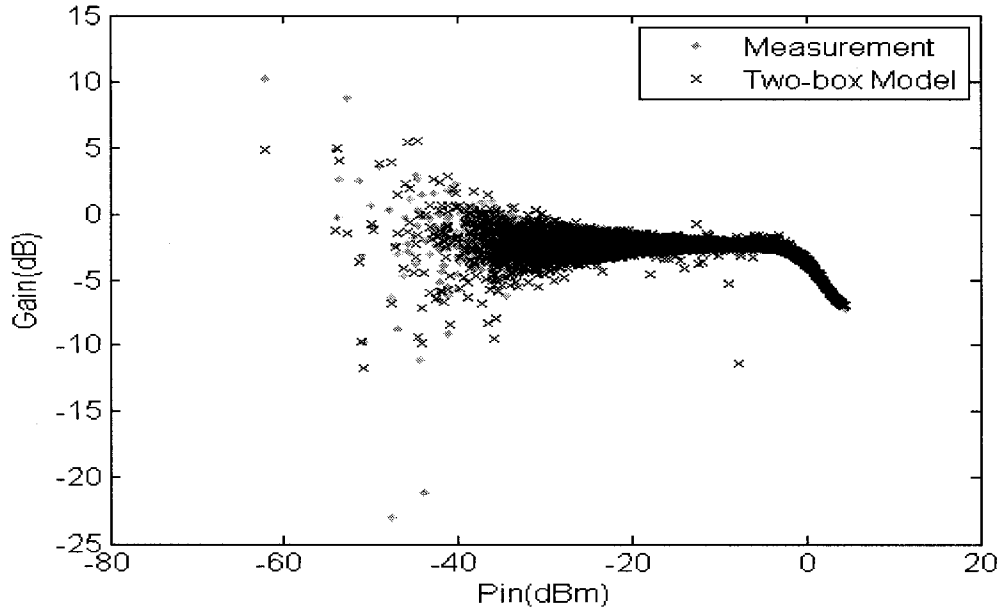


Fig. 2.20. The AM/AM estimation of the two-box model and the AM/AM measurement

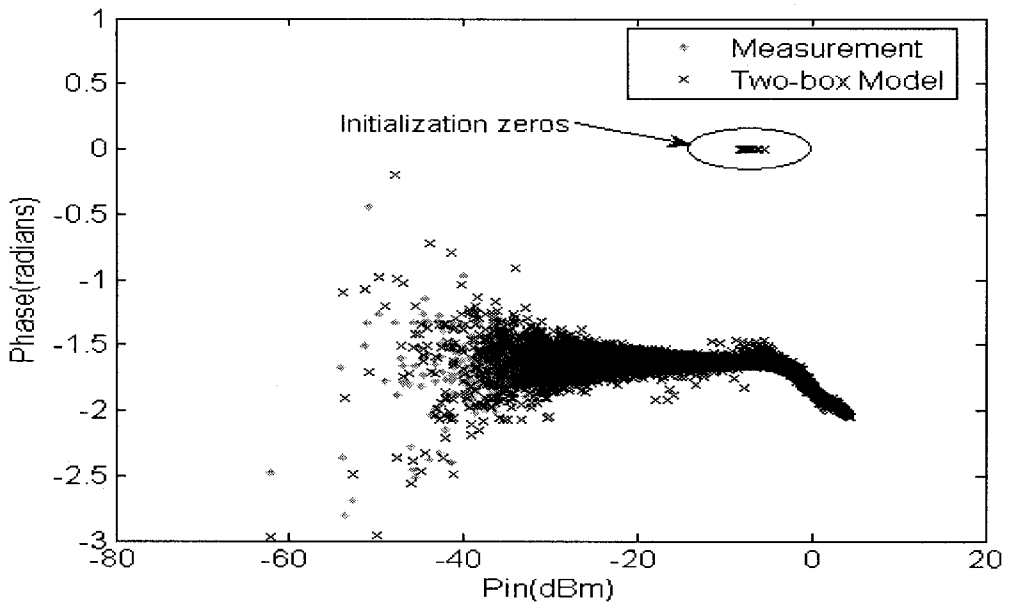


Fig. 2.21. The AM/PM estimation of the two-box model and the AM/PM measurement

from the fact that the linear filter identification, which is an iterative algorithm, assumes an initial condition for the first n-1 points of the output vector, which has been set to zero. This initial condition is not reveal in the AM/AM curve, because these points for the gain computation are mapped to minus infinity.

The number of parameters, which were identified, during nonlinear and linear identification and, with either two separate filters for each component or with one filter of complex coefficients, the final figures for the quality of fit and the order of the ARX model, are summarized in Table 2.3.

Table 2.3. Quality of fit figures for the two-box nonlinear model

	Nonlinear model		Linear ARX model			Best fit percentage	
Component	Order	Basis Fn	na	nb	nk	Complex	Separate Filters
In-phase	2	normpdf	4	16	1	98.1488	98.1498
Quadrature	2	normpdf	4	16	1	98.0723	98.0140

Although the identification results have been obtained with separate filters, the very same results have also been obtained with a complex filter, according to the percentage of the quality of the fit in Table 2.3. A more detailed study, shown in Table 2.4, reveals the fact that the parameters identified for the separate filters are very close to each other and to those of the real part of the complex filter parameters. However the imaginary part, in most cases, has a small value implying that a single linear filter with real valued

coefficients is sufficient to model both components since they are equally treated by the linear dynamics of the system.

In order to compare the performance of the model to that of the actual transmitter in the frequency domain, the power spectrum density of the real part of the baseband signal at the output of the model and the measurement are plotted in Figure 2.22. The model spectral prediction, compared to the power spectrum predicted by the Hammerstein and Wiener models demonstrate a significant enhancement, where the inconsistent oscillation in the spectrum of the real-valued Hammerstein and Wiener is no longer observed in the complex Hammerstein's spectrum, which happens to be in accordance with the betterment observed in the quality of fit calculated for the model.

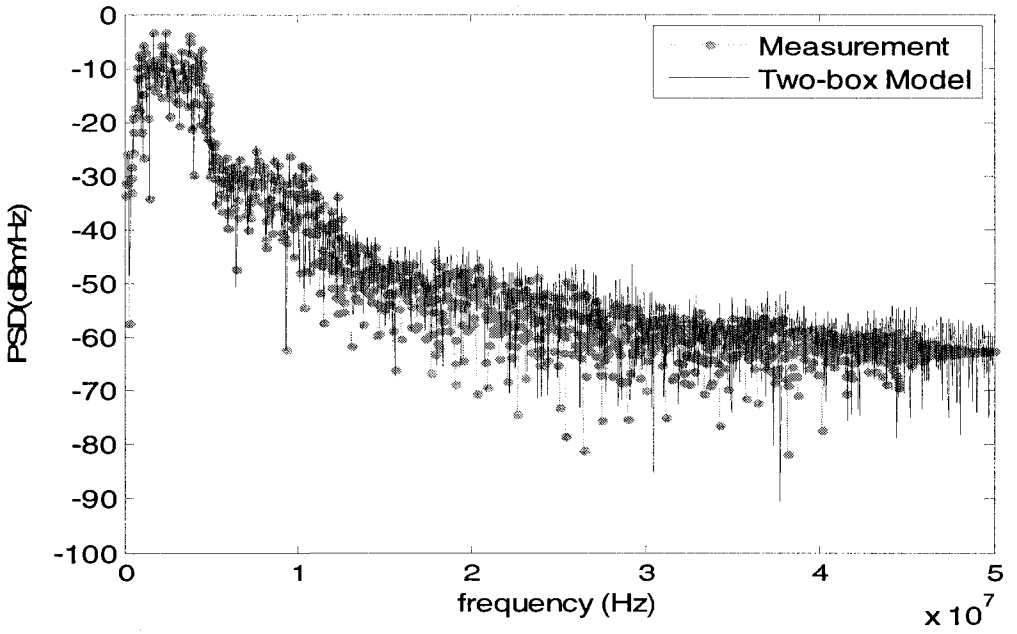


Fig. 2.22. Power spectrum density of the output signal of the two-box nonlinear model and the measurement

Table 2.4. Coefficient comparison of the separate and complex filter

Complex		na(1,4)	-0.906928	0.311763	-0.127641	0.109657
Filter			-0.033213i	0.004383i	0.022191i	-0.028843i
		nb(1,4)	2,942692	-6,264085	6,907992	-5,750543
			-0,088510i	0,278304i	-0,512000i	0,572540i
		nb(5,8)	4,310618	-2,912079	1,913753	-1,20509
			-0,530506i	0,457069i	- 0,386105i	0,309583i
		nb(9,12)	0,647345	0,296910	0,079374	0,018962
			- 0,231515i	0,168286i	- 0,144635i	0,138566i
		nb(13,16)	0,007579	0,014027	-0,013811	-0,002020
			- 0,111917i	0,074377i	- 0,037957i	0,009488i
Separate	I	na(1,4)	-0.916516	0.320543	-0.121247	0.111286
Filters		nb(1,4)	2,971008	-6,412014	7,192323	6,047791
		nb(5,8)	4,545562	-3,059638	1,991354	-1,258701
		nb(9,12)	0,739016	0,485743	0,376590	-0,310811
		nb(13,16)	0,252875	-0,134137	0,040964	0,011211
	Q	na(1,4)	-0.954957	0.375108	-0.153319	0.108779
		nb(1,4)	2,859512	-6,061107	6,642465	5,499617
		nb(5,8)	4,120139	-2,803906	1,868991	-1,185715
		nb(9,12)	0,603565	0,174767	-0,130961	0,249654
		nb(13,16)	0,177298	0,101515	-0,043503	0,002506

2.3.4 Conclusion

In this chapter, three different models have been introduced and realized. Their performance, in terms of the quality of fit for the time-domain signals, has been calculated and compared. One must also mention that the identification step was performed on a portion of measured data, while the validation results included a wider range of measurements.

It has been demonstrated that, simple real valued Wiener and Hammerstein models which are identified using passband signal, are not able to simulate a system with nonlinear phase behaviour; and, the static nonlinearity box merely serves to model the instantaneous variation of the nonlinear gain. Hence, in order to model a transmitter with phase and amplitude nonlinear profiles, a more general model, referred to as 2-box model or complex Hammerstein was introduced and identified. Although it contains the same building blocks of the simple real-valued Hammerstein, the first box is a multiple model containing two separate static nonlinear functions which accepts complex data as input and produces a complex output. The static nonlinear blocks serve to model the gain and phase variation as functions of instantaneous input power. The second box can be either a multiple linear model containing two linear ARX type filter structures, each of which serves to mimic one of the in-phase or quadrature component variations, or can be a concurrent complex structure receiving complex baseband signals. The identification of any of the two alternatives results in the very same coefficients.

The power spectral density for these models was also obtained and plotted. The real valued or the passband Wiener and Hammerstein models, quite similarly, follow the

main channel and the adjacent channel regrowth, observed in the spectrum of the measured output. However, as the power started to diminish at higher frequencies, the models failed to mimic the measurements as accurately and start to exhibit inconsistent oscillation throughout the spectrum. This was not the case of the complex two-box nonlinear model, which reveals the very same noise-like oscillation throughout the frequency band, as observed for the measured signal. This noise-like disturbance in the spectrum is, indeed, a very typical frequency domain observation of a nonlinear system.

CHAPTER 3

THREE-BOX ORIENTED NONLINEAR MODEL

3.1 Introduction

IN this chapter, a three-box model, which is capable of exhibiting nonlinear behaviour for both gain and phase of the transmitter, is proposed. Similar block-oriented structures have already been developed [37]-[39] for modeling dynamic nonlinear system, where compared to 2-box models are more general and perforce, requires a more elaborate identification technique. Figure 3.1 shows a general scheme of such a model, where two nonlinear boxes surround a linear time invariant filter.

This three-box model is, indeed, a mergence of the Hammerstein and two-box nonlinear models described in Chapter 2; however, the first 2 boxes are realized concurrently, with the same conventional procedure described for the simple Hammerstein model, identifiable with passband signals. The output of the first two boxes is then decomposed back to its equivalent baseband complex signal. The phase shift is computed and compared to that of the measurement. The residual phase variation in the measurement profile, which was not captured during the first stage of the identification of the passband signal, is then attempted to be modeled in a third single static nonlinear block. Although the third box changes the phase shift of the whole structure, it does not intervene with the performance of the first nonlinear box in modeling the static nonlinear gain. In other words, the third box manipulates the complex data at the output of the linear filter such that, while the AM/AM nonlinear profile remains unchanged, the

AM/PM is shaped according the nonlinear characteristics observed in the transmitter. The drawback of this method for modeling is that the output nonlinearity block must first extract the equivalent baseband signal of the output of the linear filter to use as an input. Hence, although using passband signal would ease the burden of performing complex identification for the first two boxes, eventually baseband signal is to be extracted to capture the residual phase distortion.

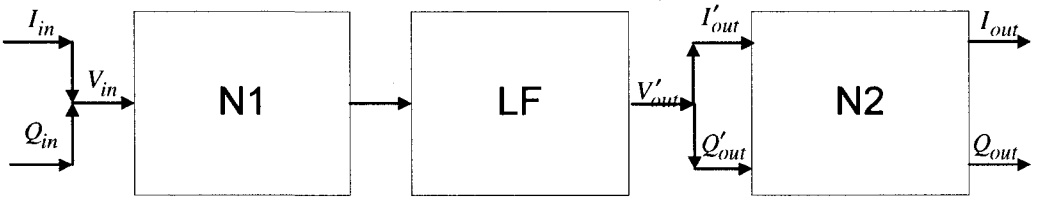


Fig. 3.1. The three-box orientation model

3.2 Three-box model's two-stage identification procedure

At microwave and millimeter wave frequencies, it is naturally more difficult to sample the signal, due to Nyquist criteria, compared to the equivalent baseband signal. Although the measurement data have been gathered by sampling the baseband signal, we mathematically re-build up the modulated passband signal as mentioned in the previous chapters for the passband Wiener and Hammerstein models. Hence, any identification with complex structures and complex parameters can be avoided. The first stage of identification is exactly the same as the Hammerstein model, described in Section 2.3.2; and, the identification ends up with the same results obtained for the as the Hammerstein model, whereupon, the AM/AM distortion profile has been successfully simulated by the

model however, this is not the case for AM/PM distortion. These results are once again plotted in Figures 3.2.a and 3.2.b.

At the output of the first two boxes, the signal is decomposed to its complex baseband signal, from which the AM/AM and AM/PM curves are computed. If we look at the relations between the gain and phase of the system and the in-phase and quadrature components, it is revealed that one can mathematically manipulate the signal components, so that the phase is shaped according to the desirable behaviour of the transmitter, while the gain remains unchanged:

$$Gain = P_{out} - P_{in} \tag{3.1}$$

$$\Delta\varphi = \tan^{-1}\left(\frac{Q_{out}}{I_{out}}\right) - \tan^{-1}\left(\frac{Q_{in}}{I_{in}}\right) = \tan^{-1}\left(\frac{I_{in}Q_{out} - I_{out}Q_{in}}{I_{in}Q_{in} + I_{out}Q_{out}}\right) \tag{3.2}$$

In other words, the ratio between quadrature and in-phase components $\frac{Q_{out}(t)}{I_{out}(t)}$ is manipulated to provide the desirable phase profile of the transmitter, while the instantaneous power $\sqrt{I_{out}^2(t) + Q_{out}^2(t)}$ does not change.

Using the AM/PM curve, the phase difference between the output of the second box and the measured output is obtained and modeled in a third box, applying the same identification procedure applied to the two-box model in Section 2.3.3. A polynomial of second order, with normal probability density function as its basis, was utilized to shape the response more accordingly. Figure 3.3 shows the final AM/PM curve obtained at the output of the third box, which indicates an acceptable agreement with that of the measurement. The AM/AM curve is not plotted, since as already explicated, it remains

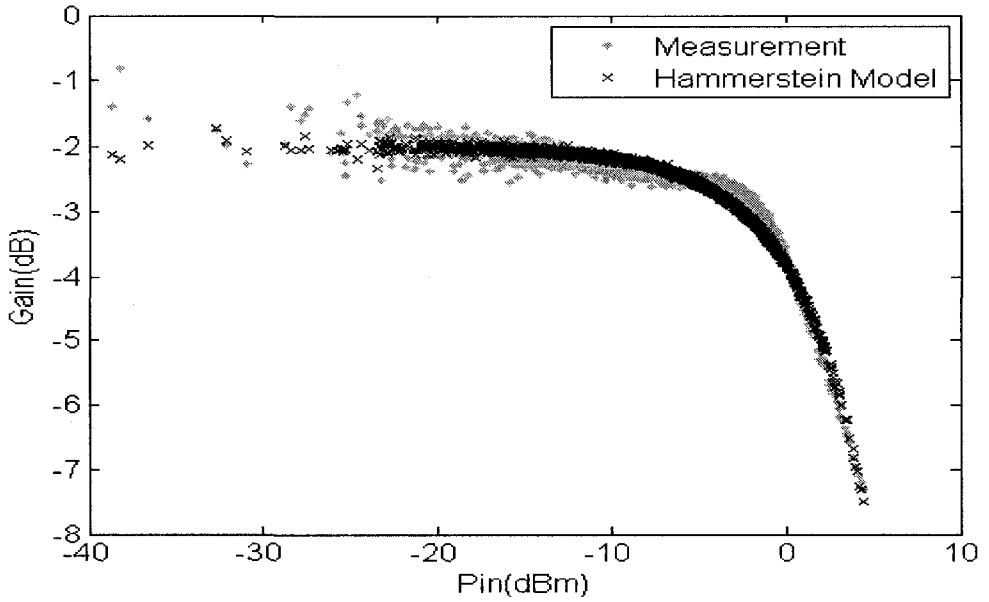


Fig. 3.2.a. The AM/AM estimation of the first 2 boxes of the three-box model and the measured AM/AM

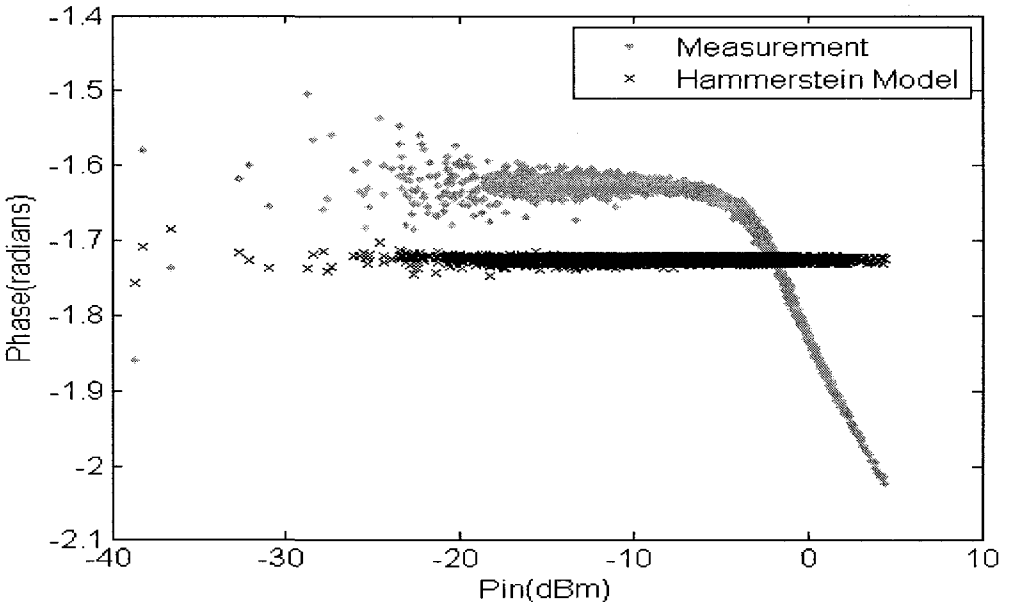


Fig. 3.2.b. The AM/PM estimation of the first 2 boxes of the three-box model and the measured AM/PM

unchanged during the third box identification. In order to observe the difference between waveforms during the first and second stage of the identification, the instantaneous power and phase signals for the second-box output, the third-box output and the measured output are drawn in Figures 3.4 and 3.5. As can be seen in Figure 3.4, no power deviation can be observed between the black curve and the blue star trace, which belong to the output of the third and second boxes, respectively. However, Figure 3.5 clearly demonstrates the performance of the third block on correcting the nonlinear phase characteristics, expected of the model; where the black and green curves, belonging to the third box output and the measured output, respectively, are quite in agreement. Once again, the blue curve proves the inability of a passband representation of Hammerstein model in introducing any nonlinear phase behavior.

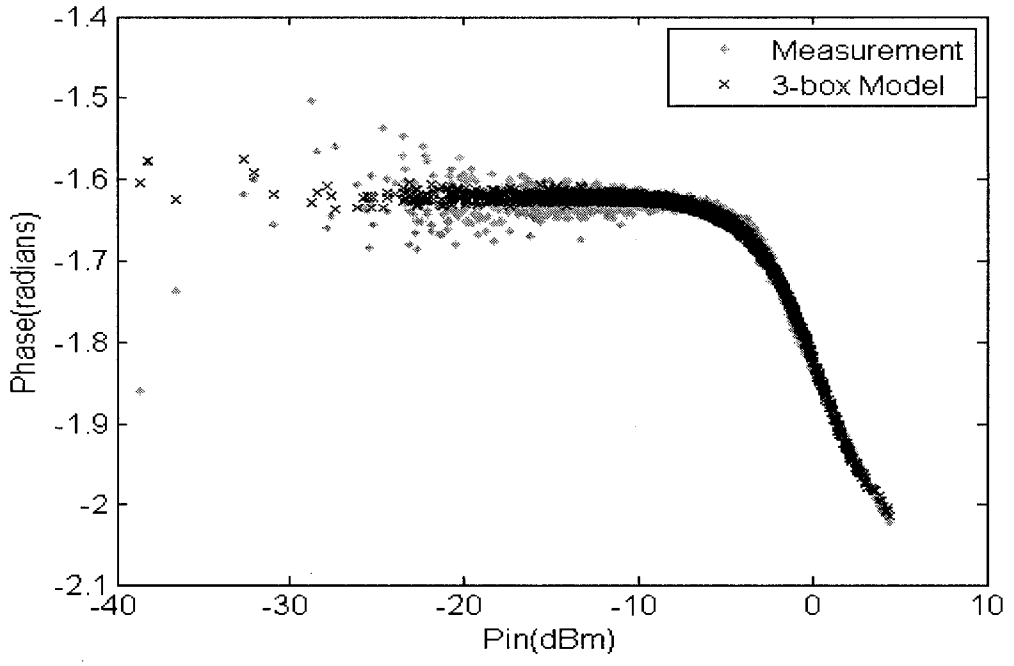


Fig. 3.3. The AM/PM estimation of the three-box model and the AM/PM measurement

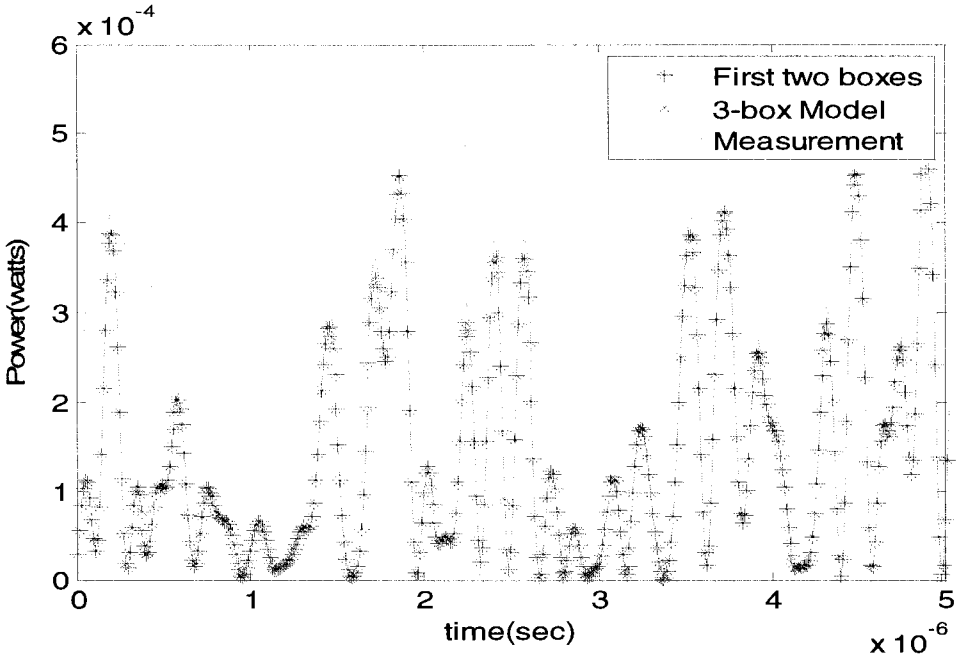


Fig. 3.4. Instantaneous power variation for the second-box output, third-box output and the measured output

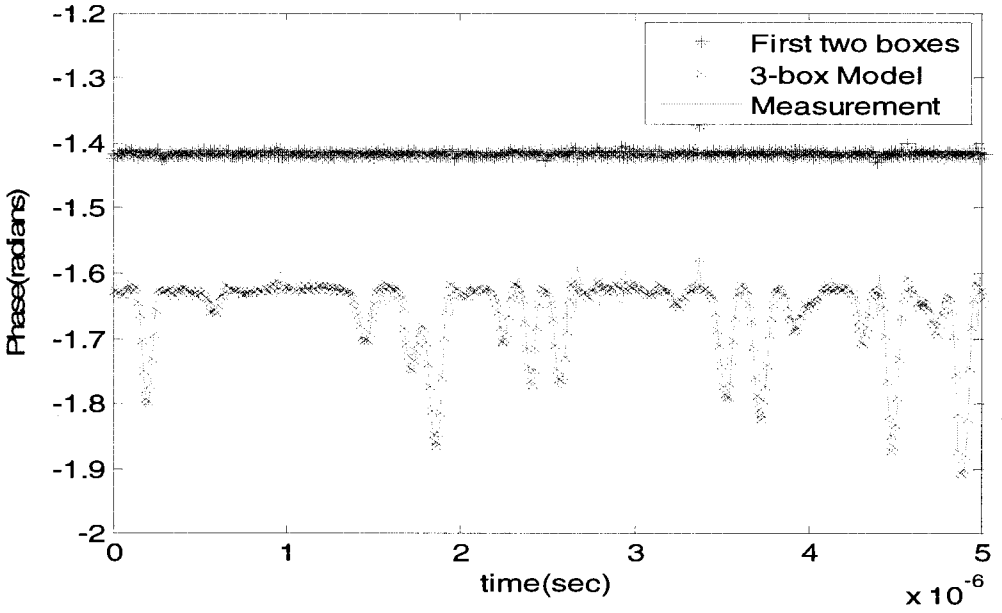


Fig. 3.5. Instantaneous phase variation for the second-box output, third-box output and the measured output

For a frequency-domain comparison, the power spectrum of the output of the three-box model and that of the measurement are plotted and compared in Figure 3.6, where a good agreement for the main channel and the adjacent channel regrowth has been achieved. The quality of fit is once again computed for the identified model, and the model is validated by applying a new set of data.

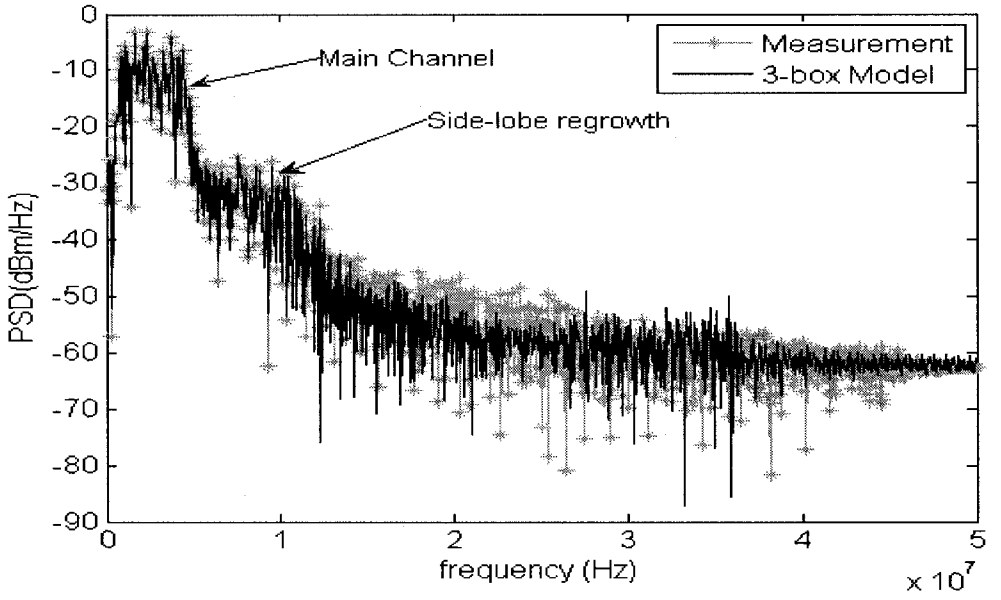


Fig. 3.6. The power spectrum of the output of the three-box model and the measurement

Table 3.1 shows the computation results of the quality of fit for both the identification and validation of the three-box and two-box nonlinear models described in Chapter 2. It can be observed that, despite a high-order nonlinear function used in the first block of the three-box model, a very simple linear filter of the first order has replaced two 16^{th}-order filters used for the two-box model. The nonlinear function for the third box is the very same as was used for the two-box model's phase function. A slightly better fitting

is obtained in both the modeling and identification stages of the three-box model. However, as long as we are concerned about the power spectrum of the model's output, both models are quite in agreement with the measurement results in terms of exhibiting similar behavior in the main channel spectrum and the adjacent channel regrwoth.

Table 3.1. 2-box and 3-box model performance summary

2-box Model	2 Parallel Nonlinear boxes		Linear ARX Filters			Best Fit Percentage for Identification	Best Fit Percentage for Validation
Component	Orders	Basis Fn	na	nb	nk	2 Separate Filters	2 Separate Filters
In-phase	2,2	normpdf	4	16	1	95.1098	96.3077
Quadrature	2,2	normpdf	4	16	1	89.6825	92.3378
3-box Model	2 Nonlinear boxes		Linear ARX Filters			Best Fit Percentage for Identification	Best Fit Percentage for Validation
Component	Orders	Basis Fn	na	nb	nk	-	-
In-phase	7,2	atanh, normpdf	1	0	0	97.7204	97.7668
Quadrature	7,2	atanh, normpdf	1	0	0	97.6559	97.6386

3.3 Conclusion

In this chapter, a three-box model was introduced and constructed. An identification process, consisting of two stages, was described. This included static nonlinear gain and phase shaping during each step, leading to our desirable nonlinear behaviour profiles. Having performed the identification and validation of the model and comparing it to the two-box model performance, a slightly better quality of fit was yielded for the same sets of data; however, the nonlinear basis functions, used in the models, are different. Despite a relatively high-order nonlinear function in the 3-box model, one could then rely on a simple single first-order linear filter, instead of two high-order filters in the two-box model. The significant shortcoming with this type of modelling is a passband to baseband decomposition in the second stage of identification, in order to provide direct access to the phase information.

CHAPTER 4

ADAPTIVE PREDISTORTION CONSTRUCTION USING SINGLE TONE SIGNAL

4.1 Introduction

DESIGN and construction procedure of digital predistortion, adaptive or non-adaptive, is normally accomplished by launching a training signal similar to the one which is to be transmitted. Highly demanded wideband transmitters that necessitate utilizing various modulated schemes, such as CDMA, OFDM or WiMAX, which take advantage of a limited bandwidth for transferring the maximum possible information over a telecommunication channel, will, perforce, encourage designers to use similar signals for the design procedure of the predistorter.

The authors in [40]-[41] used CDMA schemes to design and update the predistorter estimator; whereas, in [42]-[43], OFDM schemes were utilized. In [44], single tone, two-tone and quadrature signals were used separately as training signals and the efficiency of the predistortion was compared.

Aside from the type of the training signal and the signal that is the purpose of the transmission, one should keep in mind that the main source for nonlinear distortion in the transmitter chain is the input power level, which drives the power amplifier towards its saturation area, where consequently, inter-modulation products, harmonic generation and other nonlinear phenomena arise. Hence, whether a simple sinusoid or a complicated modulated signal is launched to the system, as long as their power level

contents fall within the same interval, they excite the very same nonlinear region, and ignoring any frequency dependent distortion for the time being, the output signal will be equally distorted by the static power dependent nonlinearity of the amplifier.

In this chapter, a single tone signal is launched as the training signal; and, a direct adaptive process is implemented to update the static nonlinear predistorter block's parameters, so that the output of the whole system remains a sinusoid signal of the same frequency, implying that the system is operating linearly. We demonstrate that changing the frequency of the input signal, as long as it falls within the pass band of the linear filter, and while its power level is being kept unchanged, will not affect the sinusoidal character of the output. Hence choosing a single frequency within the bandwidth of the transmitter suffices to eliminate the power dependent static nonlinearities and the predistorter performance remains valid for any other frequency within the bandwidth.

Since we have assumed a baseband model for the transmitter, a baseband predistortion is also required. Static polynomial functions are embodied in the predistorter box which precedes the transmitter model. The estimation error and updating law will be described in the next section. One must keep in mind that, although a model is utilized to represent the actual transmitter, the identification of a predistorter is performed independently from the model structure and information. In other words, the model is treated as a black box, which accurately resembles the actual transmitter; hence, a direct identification technique is used to update the predistorter parameters.

4.2 Hypothetical model and adaptive predistortion

In this section, a sinusoidal wave is first launched to a hypothetically built-up two-box model, which includes a saturation type static nonlinear behaviour, representing the power amplifier typical nonlinearity, followed by a low-pass simple first-order filter. The goal is to obtain a rough estimation of the capability and speed of the algorithm in updating the estimator's vector in a less expensive function evaluation and, accordingly, selecting an appropriate optimization tool before dealing with the real model which is more expensive in terms of the cost function evaluation's time .

4.2.1 Linearity criteria and cost function definition

A linear system's response to a sinusoid gives rise to a sinusoid at the output with some gain and phase shift, but with the same frequency. This is a basic characteristic of a linear system which will be the utilized as a figure of linearity in the rest of this chapter. We use this fact to find the update law for the predistorter estimation. The complete scheme of the adaptive system, including the predistorter, is shown in Figure 4.1, in the most simplified scheme possible where, f represents the predistorter static function; and, the power amplifier or transmitter in general, is replaced with a two-box model consisting of the static nonlinear functional box and the cascaded linear filter. In an ideal case, f is the inverse of the static nonlinear part of the model. Since the predistorter parameters' estimator is initiated randomly, the initial output is a distorted signal. The optimization algorithm for updating the estimator's coefficients takes advantage of the Fourier series concept. Before going into the details of the algorithm,

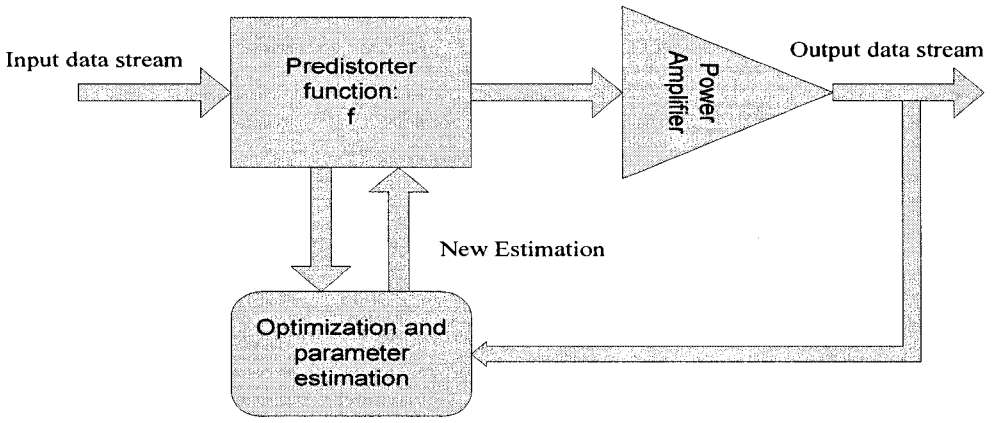


Fig. 4.1. Adaptive predistorter structure

the components of each block are introduced. The first block, which comes as the predistorter, is a polynomial of the sixth order which will be iteratively updated:

$$f(x) = \alpha_0 + \alpha_1 x + \alpha_2 x^2 + \alpha_3 x^3 + \alpha_4 x^4 + \alpha_5 x^5 + \alpha_6 x^6 \tag{4.1}$$

In the above formulation x is the excitation signal chosen to be:

$$x(t) = A\sin(\omega t + \varphi) \tag{4.2}$$

The second and third blocks, which constitute the hypothetical Hammerstein model, have been chosen as follows:

$$\begin{cases} \bar{g}(x(t)) = \tan^{-1}(x(t)) \rightarrow G(s) = L[\bar{g}(.)](s) \\ W(s) = \dfrac{1}{s+1} \qquad\qquad \rightarrow \tilde{g}(t) = L^{-1}[W(s)(G(s))](t) \end{cases} \tag{4.3}$$

where, $\bar{g}$ represents the static nonlinear function which is selected to be tangent inverse and G stands for the its Laplace transform. $W(s)$ is the linear filter which follows the

static block, $\tilde{g}$ represents the output waveform of the linear filter and eventually L and L^{-1} denote the Laplace transform and its inverse.

The following steps are taken over a sufficiently long period of time, so that the transient dynamics is not an issue. Since the excitation signal is a periodic waveform, the output waveform, although distorted, would be also periodic. Hence Fourier series concept can be applied to extract the fundamental frequency term from the rest of frequency contents:

$$g(t) = L^{-1}[W(s)L[\bar{g}(f(t))](s)] = \frac{1}{2}a_0 + \sum_{n=1}^{\infty} a_n \cos(\omega_n t) + \sum_{n=1}^{\infty} b_n \sin(\omega_n t) \tag{4.4}$$

$$\omega_n = n\frac{2\pi}{T} \tag{4.5}$$

In the above equations, ω_n represents nth frequency component. The following well known relation is used for the computation of the Fourier series coefficients:

$$a_n = \frac{2}{T}\int_{t_1}^{t_2} g(t)\cos(\omega_n t) \tag{4.6}$$

$$b_n = \frac{2}{T}\int_{t_1}^{t_2} g(t)\sin(\omega_n t) \tag{4.7}$$

If we separate the fundamental term from the rest of the series, we obtain:

$$g(t) = \underbrace{\{a_1 \cos(\omega t) + b_1 \sin(\omega t)\}}_{Fundamental} + \underbrace{\left\{\frac{1}{2}a_0 + \sum_{n=2}^{\infty}(a_n \cos(\omega_n t) + b_n \sin(\omega_n t))\right\}}_{Distortion} \tag{4.8}$$

where, the first term on the right side is the fundamental frequency term; and, the second term accounts for the rest of the harmonic products in the distorted wave form.

It is obvious that, if the system is linearized and the output is a perfect sinusoid, then it will be equal to its fundamental waveform. In other words, the second term will tend to zero. Since the output waveform is periodic, in order to reduce the computational task, we simply cut the last period of the output waveform and extract its fundamental frequency term using 4.6 and 4.7. Having subtracted the fundamental term, the residual undesirable harmonics will be obtained that is to be minimized. The integration of the absolute value of this distortion over the length of one period of the signal is assigned as the cost function, and the optimization procedure will iteratively update the predistorter parameters so as, having met certain termination conditions, a minimum value for this function be obtained. Figures 4.2.a and 4.2.b show the steps taken to extract the total distortion at the output.

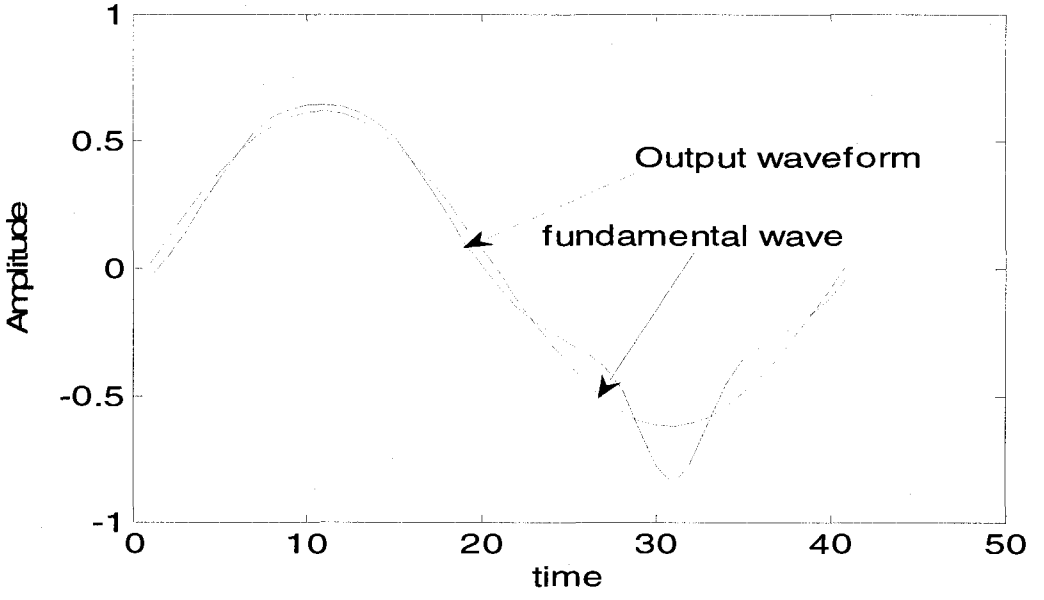


Fig. 4.2.a. The distorted signal and its extracted fundamental frequency component

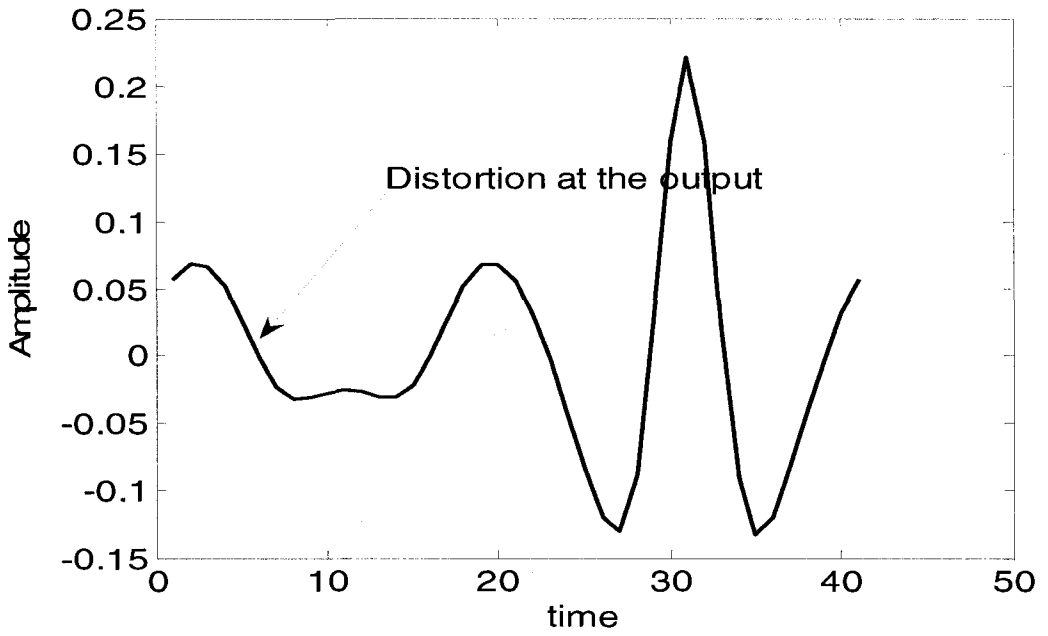


Fig. 4.2.b. The distortion obtained after subtracting the fundamental frequency component

4.2.2 Nonlinear optimization and constraints definition

As seen in the previous section, the cost function that is to be minimized, in our case, is a nonlinear function; and, a numerical optimization method should be selected to calculate the objective function iteratively, using the last update of the estimator's parameters. Considering that fact that we have chosen a two-box model to simulate the actual transmitter behaviour; and adding a third box to take over the role of the predistorter, the complete structure consists of three boxes, for which the optimizer must be capable of performing the evaluation sufficiently fast. Since the function evaluation is the main time-consuming step, any optimization technique which proceeds by means of gradient vector evaluation, can significantly slow down the convergence time.

The numerical optimization method selected in this thesis, is mesh adaptive direct search method (MADS). MADS is implemented using NOMAD (nonlinear optimization for mixed variables and derivatives), which is a suite of MATLAB functions [45]-[46]. The underlying optimizer used in the code of NOMAD is an implementation of a class of the MADS filter algorithm. This type of method uses a conceptual discretization of the space of decision variables into a stencil or pattern of points. Mesh is an iteration dependent global discretization of the decision variable space. The meshes in the structural algorithm must fulfill certain technical conditions, in order to satisfy the convergence theory.

Another point that makes NOMAD interesting is the use of surrogate functions that serve for computationally expensive functions: a common choice for the search step is to initially invest a moderate number of function evaluations to form inexpensive surrogates for the objective and constraints. In other words, a surrogate is a function that can be used as a stand-in during each optimization phase, when we are dealing with expensive functions in term of computational complexity.

Aside from surrogate function capabilities of NOMAD, it also applies a surrogate optimizer for the search strategies. For example, FMINCON in the MATLAB optimization toolbox can be set as a surrogate optimizer. This optimizer uses the Nelder-Mead or simplex method in its algorithm and can find a local optimum if the objective function varies smoothly. However, like all general-purpose multidimensional optimization algorithms, Nelder-Mead occasionally gets stuck and need to restart the

algorithm with a new simplex starting at the current best value, while NOMAD has an adaptive capability of rotating its coordinates and proceeding with it.

Having chosen the proper optimizer and using the previous section to define the objective function, we can proceed to estimate the best possible parameters to be assigned to the predistorter. The cost function has the following form as previously described:

$$E = \int_0^T abs\left(\frac{1}{2}a_0 + \sum_{n=2}^{\infty}(a_n \cos(\omega_n t) + b_n \sin(\omega_n t))\right) \tag{4.9}$$

The algorithm tries to find α_i in (4.1), such that E reaches a minimum. In an ideal case, E would be zero. The question then arises concerning the sufficient information that the algorithm should be given, in order to not only converge but also converge to proper values. In a first try, the optimization is performed on a simple cost function as defined in 4.9; it is observed that the algorithm converges quite fast, in a relatively small number of iterations. However, the results, as shown in Figure 4.3, are not in accordance with expected results for a well predistorted or a linearized system.

In order to explicate the reason on how and why the algorithm terminates without converging to the expected parameters, one needs to more precisely study the nonlinear model itself and the exciting signal, before applying any predistorter. They have been plotted in the same graph in Figure 4.4, along with the identity function and the sinusoidal excitation. It can be seen that, in order to operate in the nonlinear region, the sinusoidal amplitude must be large enough to excite the saturation area of the function f, which is assigned as the behavioral model of the transmitter. However, if the excitation

amplitude is small, the transmitter operates at its linear region; and, the output will inherently be a sinusoid wave. As can be seen in Figure 4.4, the excitation signal is large enough to drive the system towards its nonlinear region. Now a more careful study of the Figure.4.3 after performing the optimization will reveal that the output of the system with predistortion, at the last iteration, is considerably weakened compared to that of the first iteration in Figure.4.2.a. The fact that the optimizer converges and the algorithm terminates is due to the small integration error between the output and its extracted fundamental term. Since both waveforms are extremely weakened, the integration error is small enough for the optimizer to terminate; however, the two waveforms are by no means alike. In other words, the predistorter does not converge by approaching to an inverse functional behaviour, but by degrading the amplitude of the input excitation and consequently that of its extracted fundamental term.

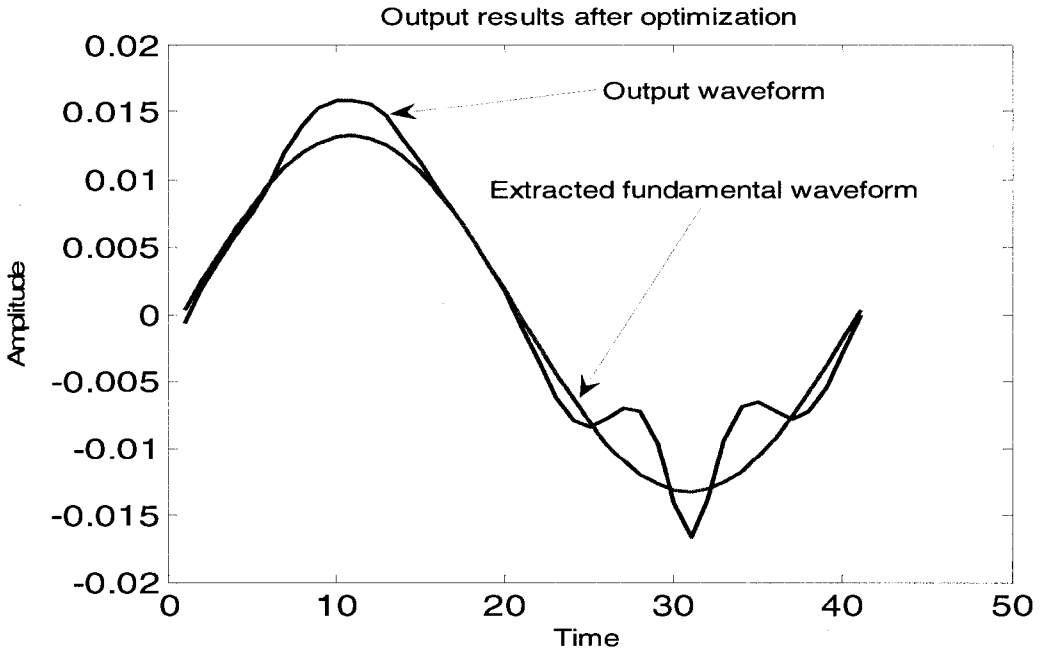


Fig. 4.3. The predistorted signal and its fundamental frequency component after optimization

The other possible manner, by which the optimizer could deviate from the proper convergence directions, can happen by initializing the algorithm such that it operates at a linear region where no nonlinear distortion would be detected and perforce, the minimization error will be again small enough for the procedure to terminate. It is obvious that, in order to prevent the estimator from accepting values that will degrade the input signal amplitude , an additional condition must be imposed on the optimizer, so that the output signal level does not decrease below a certain value to ensure that we are operating in a nonlinear region. This can be performed by adding a nonlinear constraint to the optimization process, which is defined as the integration of the absolute value of the extracted fundamental waveform at the output which is set to takes values more than a given amount.

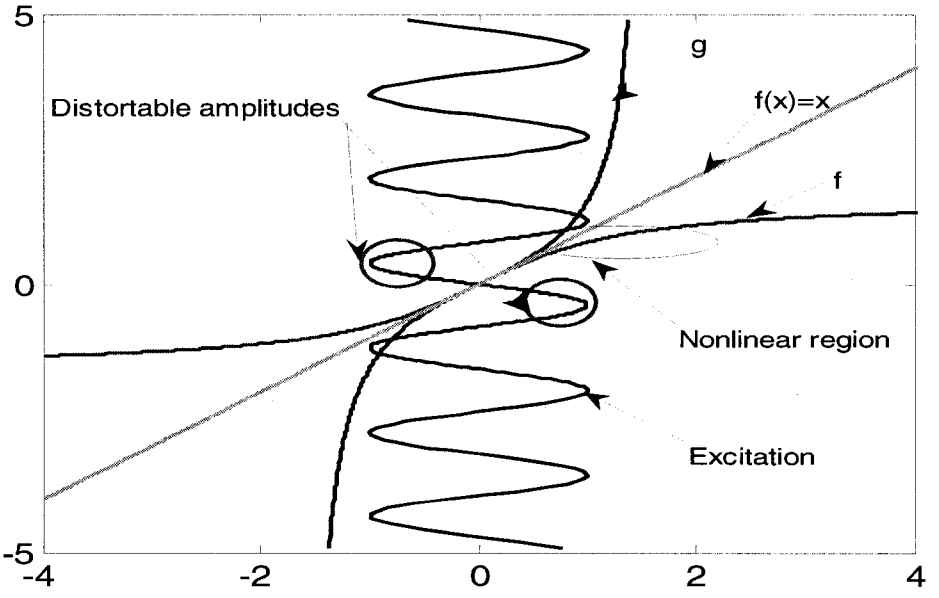


Fig. 4.4. Static nonlinear function, its inverse, identity function and the sinusoid excitation comparison

Herein, we refer to this given value by K. The constraint must be evaluated during the main objective function call at each iteration. The software NOMAD enables us to implement nonlinear constraints optimization problems which would evaluate the constraints during the cost function evaluation.

$$cx = K - \int_0^T abs(a_1 \cos(\omega t) + b_1 \sin(\omega t)) \leq 0 \tag{4.10}$$

Hence, cx is added to the algorithm as an additional nonlinear constraint. Having considered this nonlinear constraint, the optimization problem is complete. First, we launch the sinusoidal waveform to the system without a predistorter, so that we can later compare with the result obtained by the adaptive predistorter. Figure 4.5 shows the distorted output in the absence of a predistorter. Then, a polynomial function; of sixth order and only with odd terms, precedes the model to act as the predistorter. As can be seen in Figure 4.6, a very good fitting between the output waveform and its fundamental term is obtained; and, in comparison to Figure 4.5, the saturation affect that starts near 0.7 is very well compensated, as shown in Figure 4.6.

It is worth mentioning that the polynomial function used as the predistorter is not embodied with any type of basis function like tan(.). In the case that a basis function like tan(.) is used, further limitation must be taken into consideration. Since this function is periodic, for example, with period π, the input signal's amplitude cannot increase beyond $\frac{\pi}{2}$ starting from positive values; otherwise, the algorithm will come across the discontinuity of the function, and it may never converge. This issue will be addressed in the next sections.

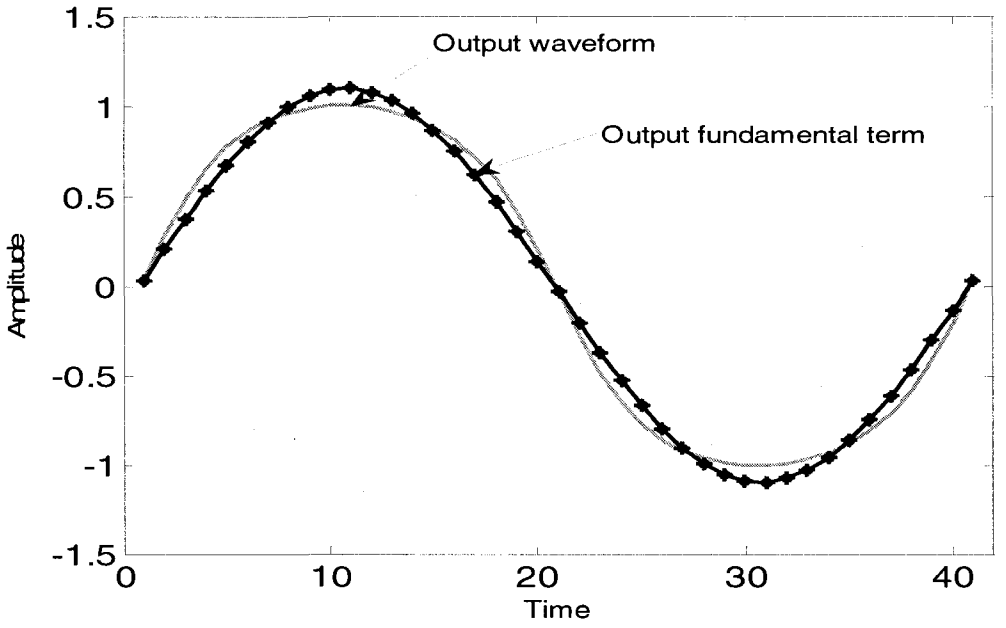


Fig. 4.5. The distorted output of the hypothetical model and its extracted fundamental component before using predistortion

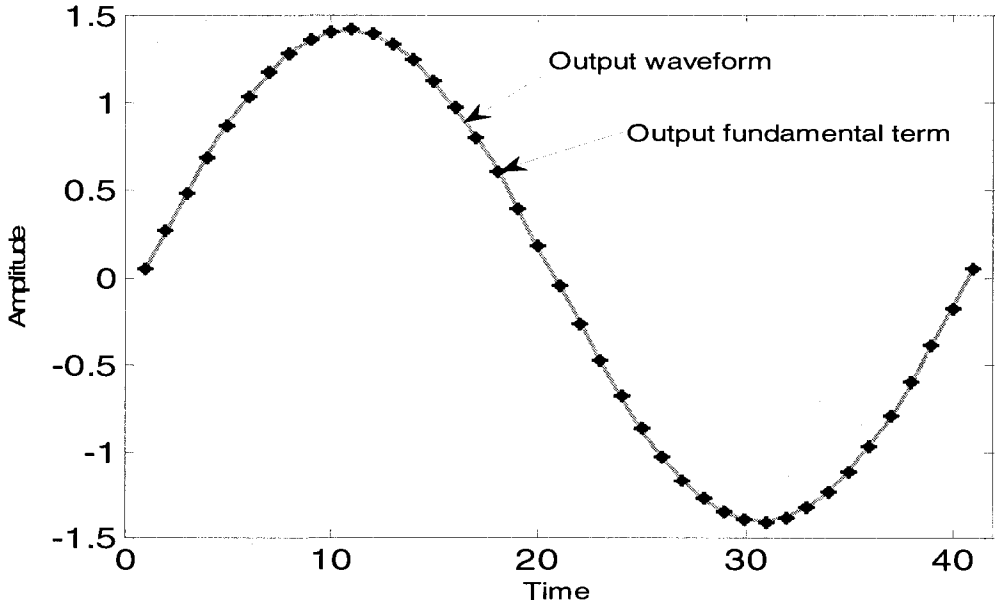


Fig. 4.6. The output of the hypothetical model and its extracted fundamental component after using predistortion

At the end of optimization, the performance history of the algorithm, the convergence criteria, the objective function last value and the parameters' vector can be observed in the graphical user interface (GUI) of NOMAD. Figure 4.7 shows the performance history for the results obtained.

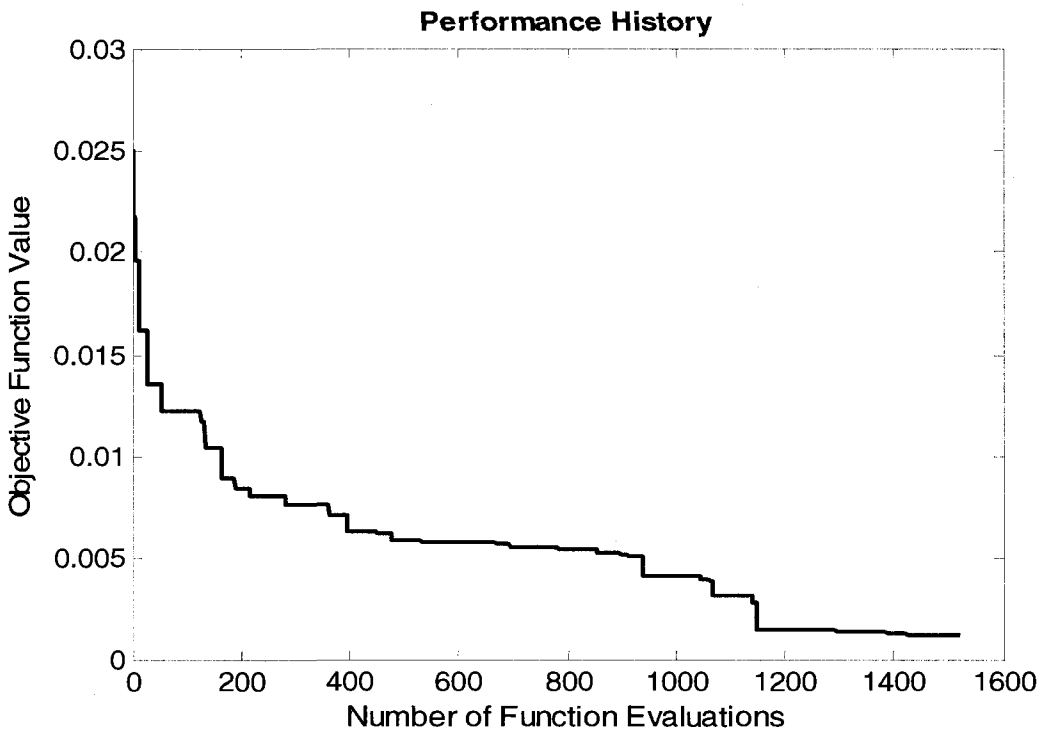


Fig. 4.7. NOMAD algorithm performance history

Table 4.1 contains the objective function values for the feasible and least infeasible solution obtained by NOMAD, as well as other information concerning the number of function calls, mesh size, etc. All this information can be observed using the graphical user interface (GUI) of NOMADm which can also provide us with the meshing parameters and termination criterion as well as the filter violation intervals.

Table 4.1. NOMAD statistics and solutions obtained

Best feasible solution		**Least infeasible solution**	
Objective Function Value	0.00121312	Objective Function Value	0.008642
Constraints Violation measure	0	Constraints Violation measure	3.59e-008
Statistics			
Final Mesh Size		6.103515625e-005	
MADS Iterations		110	
Poll Steps Executed		110	
Consecutive Poll Failures		4	
Functions Evaluations		1520	
Gradient Evaluations		0	
CPU time		2314.9688	
Cache Hits		199	
Interrupted by User		No	

In one last step we will show that the single frequency based predistorter will remain efficient and valid for other input signal frequencies as long as they fall within the passband of the linear filter embedded in the transmitter model. In other words for all the frequencies that the linear filter will approximately treat them equally, the power-

dependent static nonlinearity will be efficiently eliminated by the static predistorter which is obtained by one single frequency signal.

Based on the assumption that the linear filter amplitude and phase response remain relatively constant for the frequencies ω_1 and ω_2 , the power-dependent static nonlinearity of the model will distort the both signals in the same way as depicted for example in Figure.4.8 where since the linear filter performs equally, both waveforms have similarly sustained a static power-dependent distortion at the same amplitude.

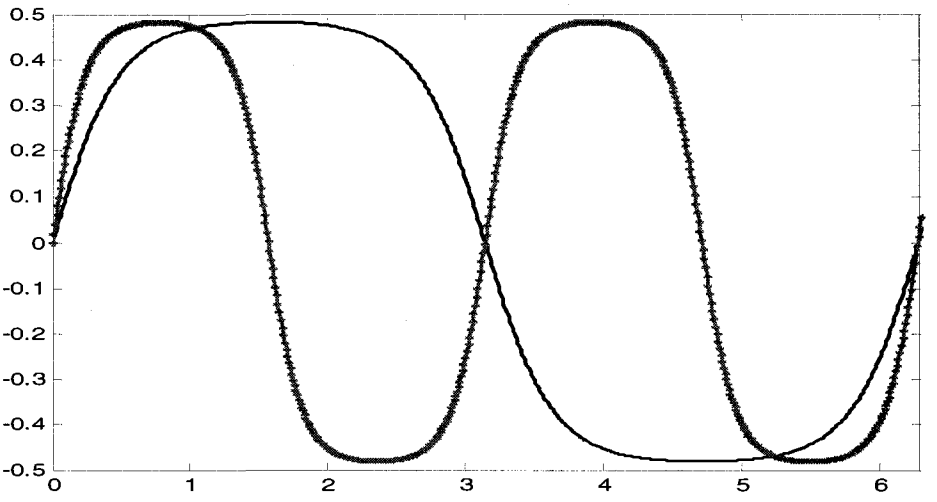


Fig. 4.8 Distorted output signal of the two-box model for two sinusoidal excitation of equal amplitude and different frequencies

Now assuming that the predistortion was constructed by launching the signal ω_1, its extracted fundamental frequency component and the distortion can be once again represented by the expansion of its Fourier series as follows:

$$g_1(t) = \underbrace{\{a_1 \cos(\omega_1 t) + b_1 \sin(\omega_1 t)\}}_{Fundamental} + \underbrace{\left\{\frac{1}{2}a_0 + \sum_{n=2}^{\infty}(a_n \cos(\omega_n t) + b_n \sin(\omega_n t))\right\}}_{Distortion} \quad (4.11)$$

Since the optimization proceeds with respect to the coefficients of the fundamental term as well as those of the distortion term, and since the final values of the predistorter will inherently depend on these parameters, any other signal with different set of Fourier coefficient will give rise to a different cost function and the constructed predistorter will not be valid anymore. Now expanding the distorted version of the signal ω_2 as follows:

$$g_2(t) = \underbrace{\{a_2' \cos(\omega_2 t) + b_2' \sin(\omega_2 t)\}}_{Fundamental} +$$

$$\underbrace{\left\{\frac{1}{2}a_0' + a_1' \cos(\omega_1 t) + b_1' \sin(\omega_1 t) + \sum_{n=3}^{\infty}(a_n' \cos(\omega_n t) + b_n' \sin(\omega_n t))\right\}}_{Distortion} \quad (4.12)$$

We prove that based on the assumption that the linear filter treats both frequency relatively equally, their corresponding coefficients will be equal and therefore the cost function defined for the parameters of the ω_1 will remain valid for ω_2. Using the following relations for obtaining the coefficients of the Fourier series, we have:

$$a_n - jb_n = \frac{1}{T_1}\int_0^{T_1} g_1(t) e^{-jn\omega_1 t} dt \quad (4.13)$$

Assigning the frequencies' ratio as α:

$$\alpha = \frac{\omega_2}{\omega_1} = \frac{T_1}{T_2} \quad (4.14)$$

where T denotes the period of the corresponding frequencies. The last period of distorted signal for both frequencies under the given assumptions, can be expressed by:

$$g_2(t) = g_1(\alpha t) \tag{4.15}$$

We emphasize once again that the linear filter with a good approximation treats both the signal equally. Therefore for signal at ω_2 :

$$a_n' - jb_n' = \frac{1}{T_2}\int_0^{T_2} g_2(t)e^{-jn\omega_2 t}dt = \frac{\alpha}{T_1}\int_0^{\frac{T_1}{\alpha}} g_1(\alpha t)e^{-jn\alpha\omega_1 t}dt \tag{4.16}$$

Changing the integration variable, we obtain:

$$t' = \alpha t \tag{4.17}$$

$$a_n' - jb_n' = \frac{\alpha}{T_1}\int_0^{T_1} g_1(t')e^{-jn\omega_1 t'}\frac{dt'}{\alpha} = \frac{1}{T_1}\int_0^{T_1} g_1(t')e^{-jn\omega_1 t'}dt' = a_n - jb_n \tag{4.18}$$

which proves that the corresponding fundamental and distortion coefficients for both frequencies is the same and therefore performing the minimization on one signal ω_1, will insure that the corresponding coefficient for other frequencies in the passband of the transmitter are minimized. However, if the phase and amplitude response of the linear filter for the range of the frequencies of the input excitation does not remain constant, therefore different frequency contents will sustain different amplitude and phase variation [47], this type of distortion will introduce some bias in the Fourier series coefficient for each input frequency and therefore will results in an error in the cost function evaluation. This residual distortion which is referred to as linear distortion [47] can not be captured by the static nonlinear predistorter and will exhibit itself at the output in a dispersion-like sparsity. This type of distortion can be treated in a post-distortion procedure using equalization techniques [47].

4.3 Construction of the complete predistorted system with the two-box model

4.3.1 Theoretical and practical criteria and limitations

For constructing the predistorter, one should remember that, theoretically, for any one-to-one function, an inverse function can be assigned so that their combination results in an identity function. However, according to the order of functions (i.e. which one precedes and is excited by the input first and which one is cascaded), we are limited by the domain of the first function. This is described in Figure 4.9, where a nonlinear function of similar behaviour as our power amplifier and its theoretical inverse are plotted. Although the nonlinear function, like the functions in the building blocks of our model, can take any value from zero to infinity, the real amplifier input power level is limited. This limit is shown as Pin(max) in Figure 4.9, until which we wish to linearize the amplifier.

As can be seen for the normalized data; employed during 2-box model identification, the nonlinear model of the power amplifier shows a saturation tendency and falls beneath the identity curve, indicating that the maximum power delivered Pout(max) is, at no point, bigger than the input power. The predistorter is expected to have the behaviour of the red curve so that the combination will result in the identity function. However, a closer look reveals the fact that, since the predistorter precedes in the transmitter line as the first functional block, we are limited by the domain of input values for which, the predistorter's function is defined.

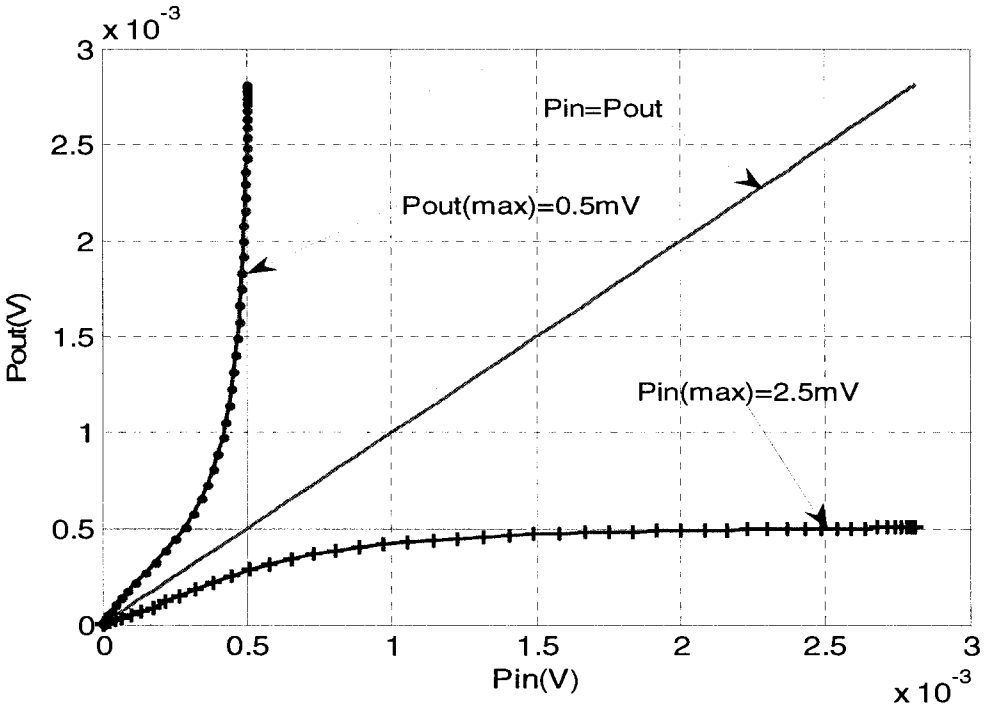


Fig. 4.9. Nonlinear function behavior (-+), its inverse (-.) and identity function (-) and their range and domain

The definition of the domain of any function is the set of values for which the function is defined; furthermore, the set of values to which the domain is mapped by the function is called the range. These basic definitions are better described for an exemplary function in Figure 4.10.

Let us define the domain and range of the main function and its inverse by the following notations:

For the main function: $\begin{cases} D_1 = \text{Function domain} \\ R_1 = \text{Function range} \end{cases}$

For the inverse function: $\begin{cases} D_2 = \text{Function domain} \\ R_2 = \text{Function range} \end{cases}$

Due to the definition of inverse function we have:

$$D_1 = R_2 \tag{4.19}$$

$$D_2 = R_1 \tag{4.20}$$

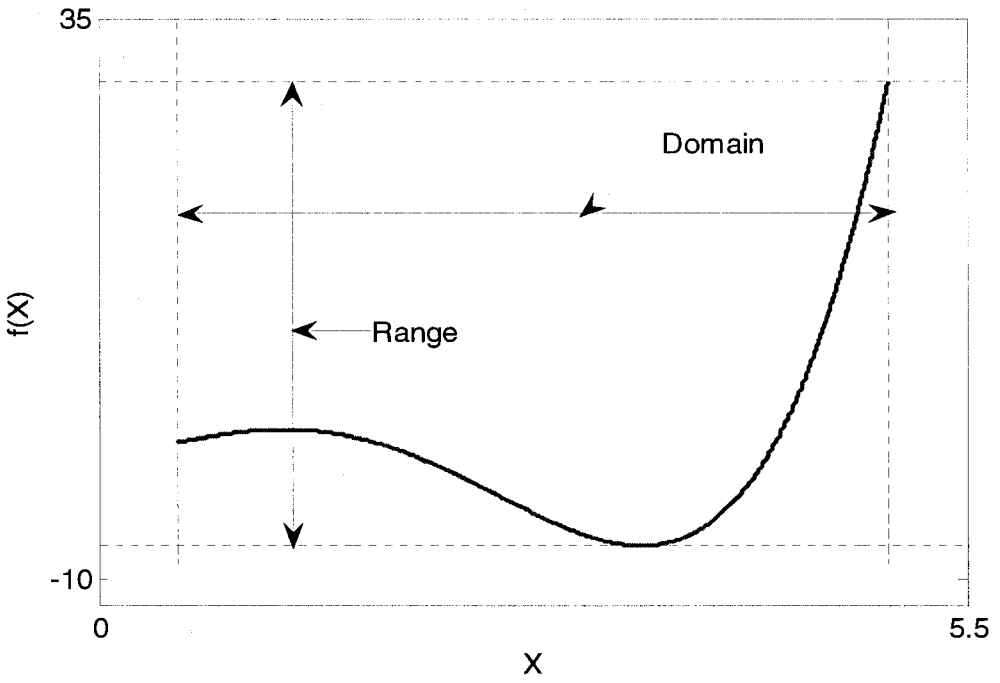


Fig. 4.10.Graphical definition of the domain and range of a function

In our case, $R_1 \subset D_1 \Rightarrow D_2 \subset D_1$, as can be seen in Figure 4.9; and, perforce, we are limited by a smaller domain D_2. That is to say that the input signal power level cannot be more than the maximum value in this domain.

On the other hand, depending on the functional blocks used in the inverse model (even though we try to extend domain D_2 to reach the maximum value in domain D_1), the output may tend to infinity. Such a high power level output cannot be launched to the actual transmitter; and, if the inverse function model contains periodic basis functions, such as tangent or tangent hyperbolic inverse, then the discontinuous characteristics of this type of functions which can severely compromise the predistorter performance must be taken into consideration.

These effects, which are essentially due to the lack of inverses of periodic functions, are shown in Figure 4.11. Hence, in the latter case, the period of the inverse model must be large enough to cover the entire input signal domain, so that the discontinuities are not hit. In fact, we take advantage of basis functions like $\tan(.)$, due to their similar behavioural profile to what is expected of an inverse function for the transmitter; therefore, a polynomial with such basis functions of a low order can replace a high-order polynomial with identity basis functions. It would be all at the cost of paying attention to the period of the function not to be crossed over by the drive signal domain. One must remember that the period of the basis function can be either treated as one of the optimization variable or fixed at a certain value at which the previously considerations would be fulfilled; however, the optimizer performs with one less degree of freedom. In the latter case, where the period of the function is taken as a variable, one needs to apply another nonlinear constraint on the optimizer so that we are assured that no possible discontinuous area will be reached where it can mislead the algorithm and the proper convergence may never be achieved.

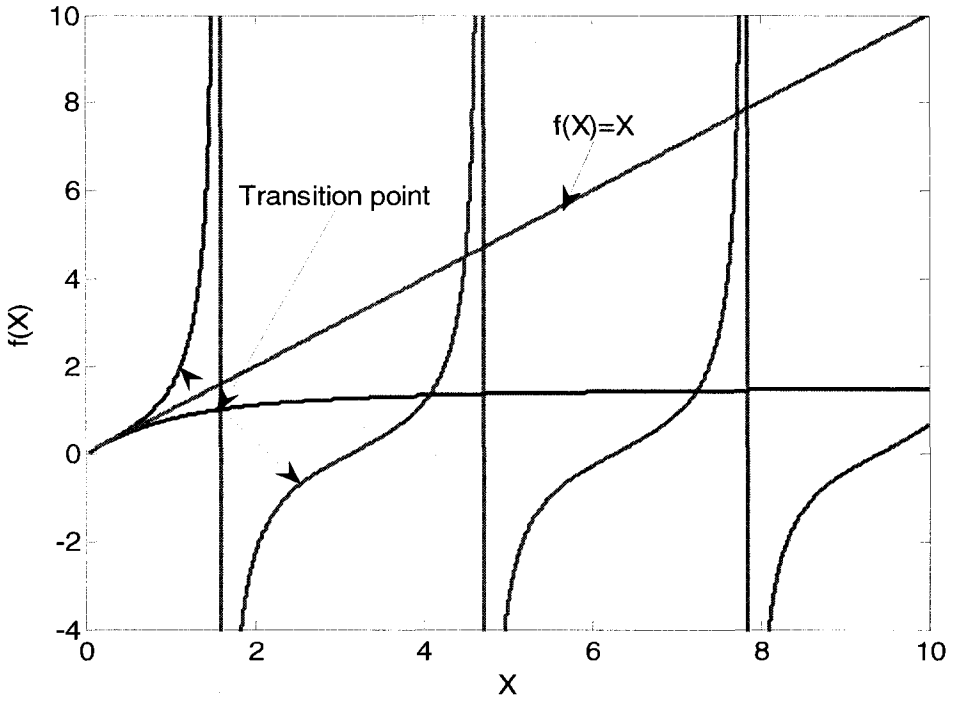


Fig. 4.11. Predistorting a nonlinear system using periodic basis functions and the problem of lack of inverses

To avoid all above constraints, we construct the predistorter for linearly normalized data at a new level, as shown in Figure 4.12. In other words, the transmitter model is reconstructed for a new level of power. As can be seen, the domain and range of the functions are equal; hence, no problems will be encountered for launching a higher level of input signal power since it will fall within the domain of the predistorter. One must remember that in practice, working with power amplifier; due to their amplification, the output power range always covers the input power interval. This normalization is just for the predistorter construction and will be later undone.

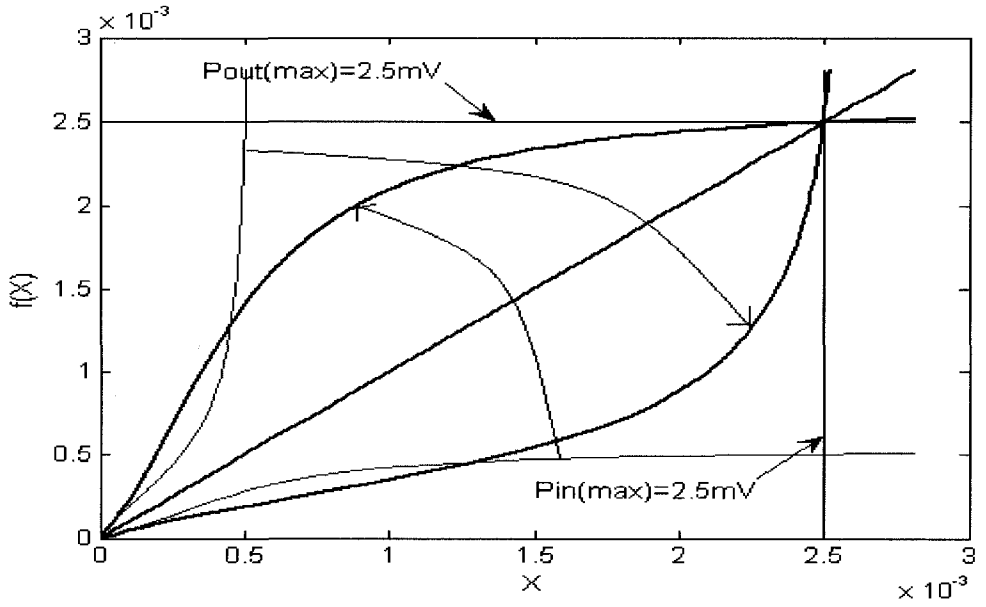


Fig. 4.12. Normalizing the signal power level to unit gain at the maximum power to be transmitted (light blue curve to dark blue curve), hence changing the inverse function domain (light red to dark red curve)

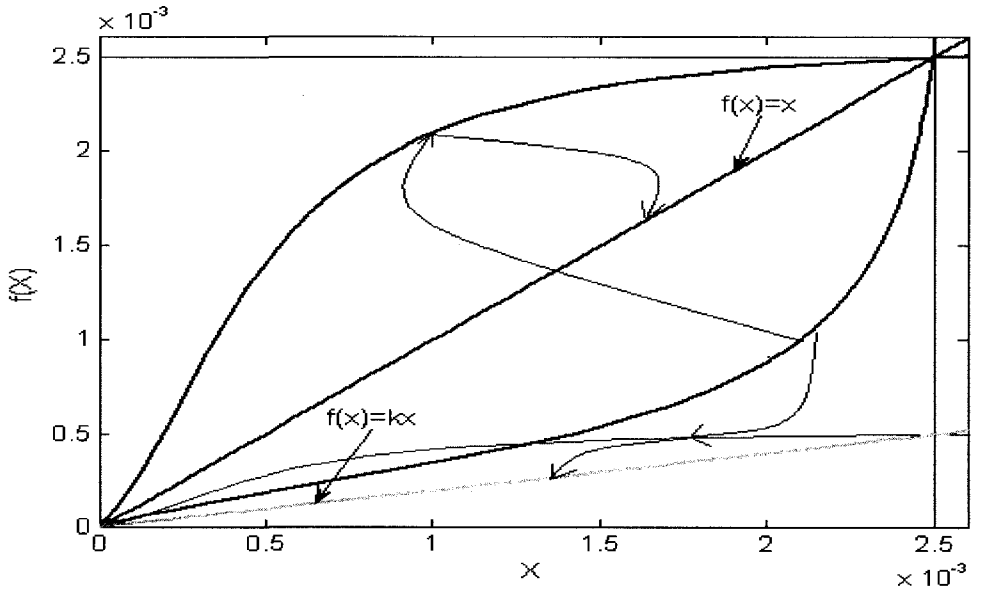


Fig. 4.13. Different linearization levels obtained by applying the predistortion function (red curve) on the two power level system models (blue curves)

The predistorter built at this new level of power, can still be utilized in order to linearize the system at the initial power level before normalization, but it will be at the cost of degradation in the small signal gain. This is shown in Figure 4.13, where the inverse function of the red curve is applied on both blue curves; and, despite the fact that in both cases the same maximum power is transferred linearly, for the lower power profile, we have a decrease of gain by factor k, which is, indeed, the inverse of the factor of normalization. This can be proven easily.

For the lower level power system model and its inverse profile, the corresponding functions are assigned as f and f^{-1}, respectively; and, those of the normalized system model and its corresponding inverse are assigned as g and g^{-1}, where $g = kf$. Hence:

$$\begin{cases} f\left(f^{-1}(x)\right) = x \\ g\left(g^{-1}(x)\right) = x \Rightarrow kf\left(g^{-1}(x)\right) = x \Rightarrow \\ kf = g \end{cases} \tag{4.21}$$

$$f\left(g^{-1}(x)\right) = \frac{x}{k} \tag{4.22}$$

This proves that the second configuration in Figure 4.14 also gives rise to a linear system. However, if the order of boxes is changed, then this linear transformation is no longer valid, as can be more explicated as follows:

$$\begin{cases} g^{-1}\left(g(x)\right) = x \Rightarrow g^{-1}\left(kf(x)\right) = x \\ g^{-1}\left(kf(x)\right) \neq kg^{-1}\left(f(x)\right) \neq x \end{cases} \tag{4.23}$$

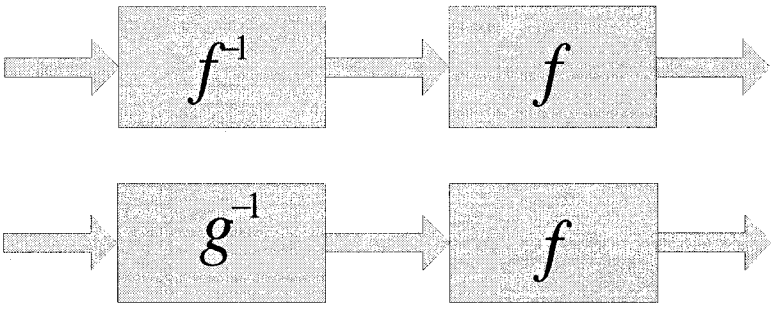


Fig. 4.14. Function inversibility at different amplitude levels

Therefore, in order to compensate for this gain loss, a linear post-distortion has to be performed using a linear amplifier. In any case, linearization, regardless of the type of model and its nonlinear profile, will always introduce some reduction in gain, compared to that of the small signal gain.

This fact, also mentioned in [47] in a modeling approach, is shown in Figure 4.15, where the best linear approximation of a system, with the help of predistorter and the small signal gain or linear gain, are compared and one can notice the bias introduced by the linearization, which results in a decrease in the amount of linear gain. It can be seen that this bias depends on the maximum power that we are willing to transmit linearly; the lower the maximum power, the smaller the bias. Hence although we avoid harmonic generation, intermodulation effects and other nonlinear phenomena, smaller gain will be achieved for small signal applications.

Having explained the above problems, we have utilized a model for the system, for which, the domain and the range are the same and equal to the maximum power intended to be transmitted.

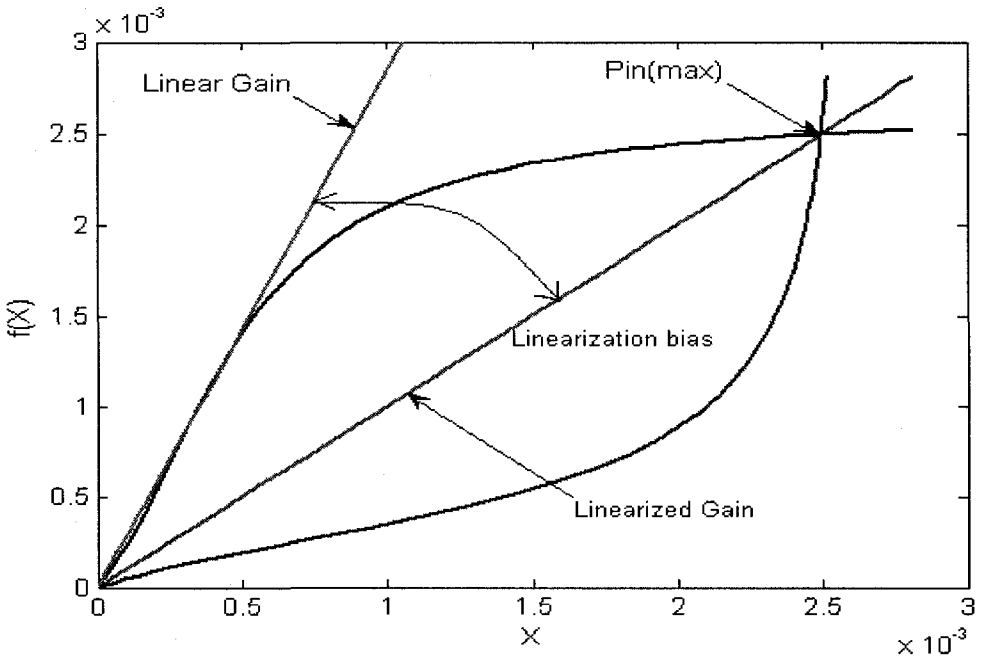


Fig. 4.15. Gain decrease introduced due to the linearization bias

4.3.2 Predistorter construction and linearity criteria for a two-box model

After verifying the minimization problem and its constraints in the previous section and having chosen the proper optimizer, we can now perform the very same procedure, but this time on the actual model of the transmitter. We use the two-box model realized in Chapter 2, which provides us with a relatively good quality of fit. However, one should remember that this model has a complex structure and the input and output signals are inherently complex numbers, which in here they consist of the baseband complex input and output signals. Since a simple sinusoidal signal is being utilized as the excitation, the other input is simply set to zero; or, the total power due to be fed to the system can be divided into two sinusoids, each of which serves as one of the predistorter's inputs.

We found that the latter method was easier for the optimizer to resolve; therefore, two single frequency waves of the same amplitude and frequency drive the predistorter. First, we excite the two-box model with sinusoids without applying any predistortion. Figure 4.16 shows the output waveforms obtained, along with their equivalent input waveforms; and, it can be seen that both components are distorted. As already mentioned, the total power of inputs is high enough to drive the system towards its nonlinear region.One must pay attention, that the input signal power level is the same for both components; however, due to the nature of the model at the output, the amplitude information is carried by one of the components, which here is the quadrature (Q) component which starts to be clipped as the power increases; and then, phase distortion information is embedded on the in-phase (I) component of the output. Furthermore, this information distribution is such that, the maximum of output power, which is the sum of the maximum power of I and Q components, is the same as the actual signal which is to be transmitted. Here, we would like to mention once again that the model is implemented in base-band domain which dramatically simplifies the analysis. Consequently the predistorter is also structured to operate in base-band domain. Indeed applying the modulated signal, to simulate each cycle of the RF carrier, needs up to thousands more operations. In order to see the performance of the model on a very low power input, we launched two sinusoids of small amplitudes so that the distortion effects do not appear. The results are presented in Figure 4.17 where, although more power is carried by the quadrature component of the output, both waveforms are sinusoids – a requirement in order to reduce the phase distortion to zero. Using this fact, the linearity

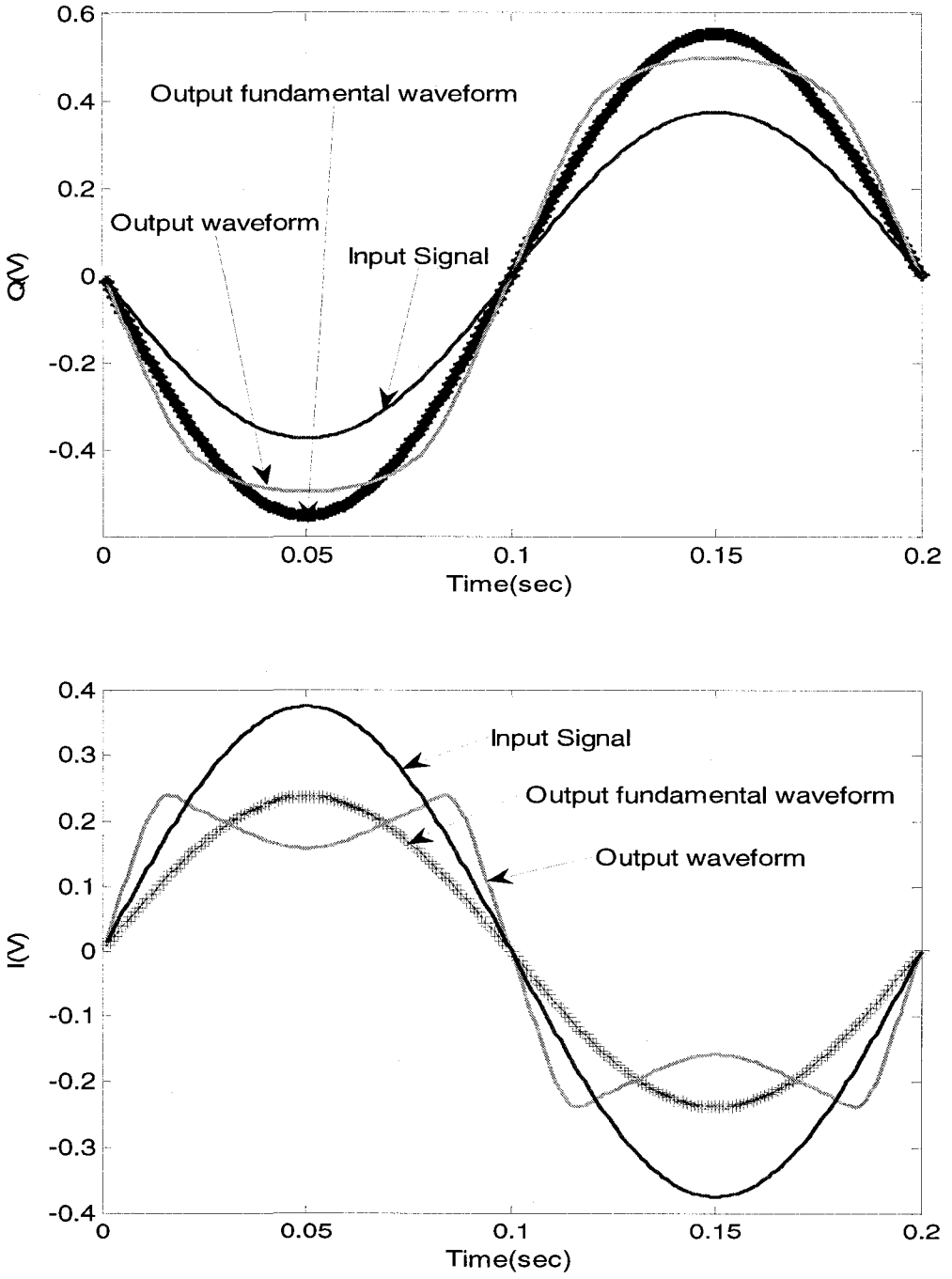


Fig. 4.16. The in-phase (lower picture) and quadrature (upper picture) components of the input, output, output fundamental frequency component, before running the optimization

criteria can be defined as to obtain two sinusoidal waves at the output port, but as expressed in a very similar manner in Section 4.2, this is to be achieved while the maximum output power does not decrease below the value we are willing to transfer linearly. In other words, regardless of the power distribution between the final two sinusoids by the optimization algorithm, the total power should not be less than a certain amount, which is to be transmitted linearly to satisfy the power efficiency condition.

Since nearly all optimization algorithms are sensitive to the initial point, it must be carefully chosen, so that, iteratively, it leads up to the proper convergence point. The modeling procedure can be helpful, so that we can roughly guess the expected profile of the inverse functions; and, aside from choosing the proper polynomial order, we can roughly select a basis function that suits the inverse model behaviour. In this context, it is preferable to set the initial parameter vector, with a uniform order for all the adjustable variables. For further explanation, suppose an example of the initial parameter vector chosen to be as follows:

$$P0 = [a0\ a1\ a2\ a3\ a4] = [1\ 100\ 1000\ 0.001\ 1e\text{-}5] \tag{4.24}$$

i.e. with component variables spread over a large interval; and, since the optimizer would normally work with uniform meshing in all dimensions, then a fine mesh may detect the influence of small variables on the variation of the cost function. However, the effect of large variables may remain undetected; and, this may mislead the algorithm. Hence, it is of critical importance to put the variable in a uniform order, which would

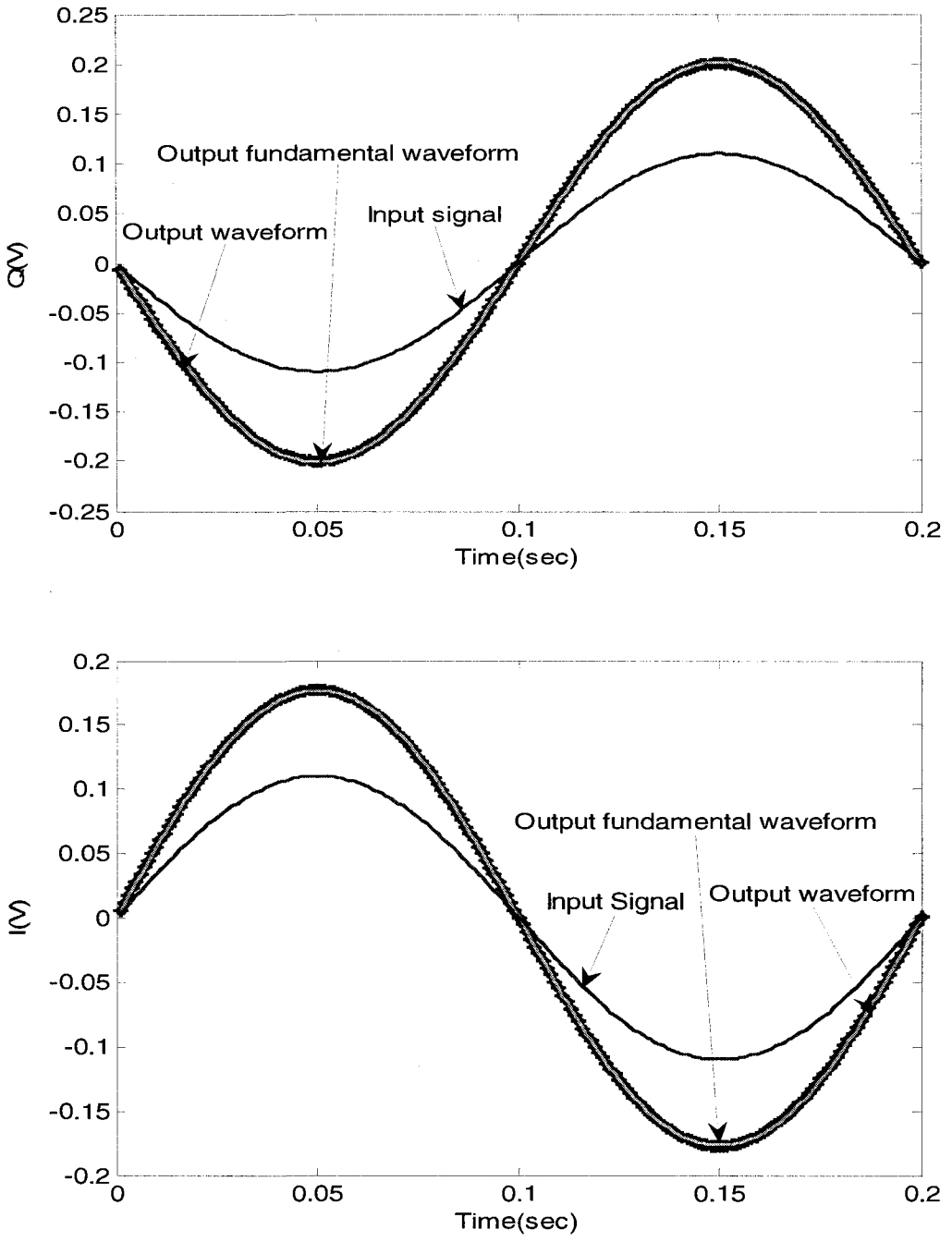


Fig. 4.17. The in-phase (lower picture) and quadrature (upper picture) components of the input, output, output fundamental frequency component, before running the optimization for a low power excitation

preferably work with a fine mesh. For the above example, the uniform format can be expressed as follows:

P0= [a0 a1 a2 a3 a4....] = [1 100 1000 0.001 1e-5...] $\Rightarrow$ (4.25)

$$P0=[0.1 \quad 0.1 \quad 0.1 \quad 0.1 \quad 0.1]*\begin{bmatrix} 10 & 0 & 0 & 0 & 0 \\ 0 & 1e3 & 0 & 0 & 0 \\ 0 & 0 & 1e4 & 0 & 0 \\ 0 & 0 & 0 & 1e-2 & 0 \\ 0 & 0 & 0 & 0 & 1e-4 \end{bmatrix} \tag{4.26}$$

The matrix acts like a scaling factor, and the new variable vector will be:

$$P0=[0.1 \quad 0.1 \quad 0.1 \quad 0.1 \quad 0.1] \tag{4.27}$$

Therefore, uniform meshing results in uniform changes in the variables; and, multiplication by the scaling matrix above, readjusts the variables in the right order. It is worth to mention that the software NOMAD can also provide us with scaling capabilities available as one of its setting options.

The practical problems that were described in detail in Section 4.3 all concerned the gain nonlinear distortion. However, when it comes to phase distortion, there will not be range and domain problems for the input and output power, mainly because the predistorter for phase nonlinearity should operate on the same power level that is driving the real system. In other words, the output power of the gain predistorter block will be launched

into the transmitter, but we take one branch from it to construct the phase distorter. It is described in Figure 4.18.

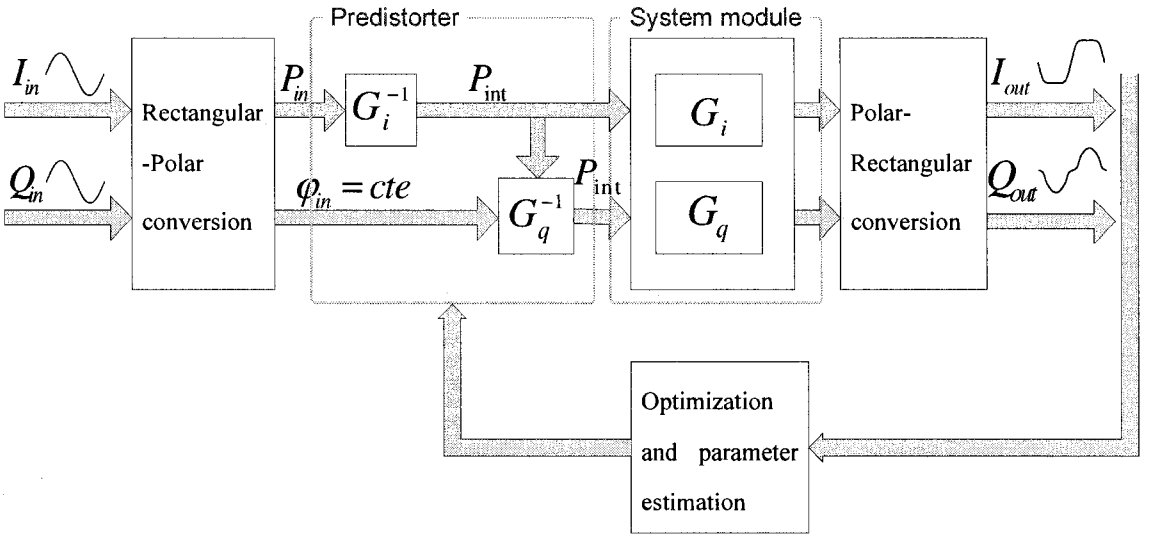


Fig. 4.18. Adaptive predistortion system block diagram

The polar-rectangular relations are the same as expressed in Section 2.3.3, so they are not repeated here. G_i and G_q represent the system nonlinear distortion for gain and phase, respectively, and G_i^{-1} and G_q^{-1} are their corresponding inverses that hoped to be obtained through numerical optimization of the predistorter variables.

The functional blocks chosen for the predistorter are polynomials embodied with two different basis functions:

$$G_i^{-1}(x_i) = p_1 f_{B1}(x_i) + p_2 (f_{B1}(x_i))^3 + p_3 (f_{B1}(x_i))^5 \tag{4.28}$$

$$G_q^{-1}(x_{\text{int}}) = p_4 f_{B2}(x_{\text{int}}) + p_5 (f_{B2}(x_{\text{int}}))^2 + p_6 (f_{B2}(x_{\text{int}}))^{2.5} + p_7 (f_{B2}(x_{\text{int}}))^{2.8} \tag{4.29}$$

where:

$$f_{B1}(x_i) = \tanh^{-1}(Nx_i) \tag{4.30}$$

$$f_{B2}(x_{\text{int}}) = normpdf(x_{\text{int}}, \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x_{\text{int}}-\mu)^2}{2\sigma^2}} \tag{4.31}$$

The multiplication factor N seen in the basis function of the gain is to assure that we do not reach a discontinuous area of the tangent hyperbolic inverse function. Although the coefficient μ, σ can also be taken as optimization variables, they are initially chosen at proper values, so that the algorithm does not need to treat two more variables at each iteration, alleviating the computational process. As previously described, the two input data are sinusoidal waves, and their maximum total power is assigned such that they can successfully drive the transmitter towards its nonlinear region, until the point at which the maximum power will be transmitted. Hence we proceed by expressing the complex input data as:

$$I_{in} = A\sin(\omega t + \varphi) \tag{4.32}$$

$$Q_{in} = A\sin(\omega t) \tag{4.33}$$

where, φ is a constant that can be set to zero. The output data are initially distorted and can be expressed in terms of its individual frequency contents using Fourier series expansion:

$$I_{out}(t) = \underbrace{\{a_{1i}\cos(\omega t) + b_{1i}\sin(\omega t)\}}_{Fundamental} + \underbrace{\left\{\frac{1}{2}a_{0i} + \sum_{n=2}^{\infty}(a_{ni}\cos(\omega_n t) + b_{ni}\sin(\omega_n t))\right\}}_{Distortion} \tag{4.34}$$

$$Q_{out}(t) = \underbrace{\{a_{1q}\cos(\omega t) + b_{1q}\sin(\omega t)\}}_{Fundamental} + \underbrace{\left\{\frac{1}{2}a_{0q} + \sum_{n=2}^{\infty}(a_{nq}\cos(\omega_n t) + b_{nq}\sin(\omega_n t))\right\}}_{Distortion} \tag{4.35}$$

In order to approach a linear behavior, the summation of all the terms in the series, except for the fundamental frequency term, should tend to zero, which will be formulated as the objective function of the minimization in an integration form:

$$E_i = \int_0^T abs\left(\frac{1}{2}a_{0i} + \sum_{n=2}^{\infty}(a_{ni}\cos(\omega_n t) + b_{ni}\sin(\omega_n t))\right) \tag{4.36}$$

$$E_q = \int_0^T abs\left(\frac{1}{2}a_{0q} + \sum_{n=2}^{\infty}(a_{nq}\cos(\omega_n t) + b_{nq}\sin(\omega_n t))\right) \tag{4.37}$$

It follows that:

$$E_{tot} = \int_0^T abs\left(\frac{1}{2}(a_{0i}+a_{0q}) + \sum_{n=2}^{\infty}((a_{nq}+a_{ni})\cos(\omega_n t) + (b_{ni}+b_{nq})\sin(\omega_n t))\right) \tag{4.38}$$

As can be instantly perceived, the error is a function of distortion terms. Similar arguments concerning nonlinear constraint and the convergence direction of the algorithm discussed in section 4.2.2 for the hypothetical model, applies to the 2-box model as well. It remains an unmistakable requirement in order to guarantee that the output power does not decrease below the value we are willing to transmit linearly – Pout(max). This imposes a nonlinear constraint on the optimization, similar to what was shown in Section 4.2.2, with the difference that now the output is the sum of two waveforms. In order to ensure that the algorithm does not converge to unrealistic values that are too big to feed the actual system, an upper bound should also be imposed on the optimizer, resulting in a second nonlinear constraint. It is of interest to point out that the first constraint is applied on the output of the model, while the second constraint is applied on the intermediate stage, which is after predistortion and before feeding the

model. In other words, the input of the system should not be increased over a certain amount, known as Pin (max); and, the output of the system should not be less than the desired maximum power. This can be described in closed form expressions as:

$$\max(P_{\text{int}}) = \max(I_{\text{int}}^2 + Q_{\text{int}}^2) \leq N_1 \max(P_{\text{in}}) = K_1 \tag{4.39}$$

$$\max(P_{\text{out}}) = \max(I_{\text{out}}^2 + Q_{\text{out}}^2) \geq N_2 \max(P_{\text{out-des}}) = K_2 \tag{4.40}$$

In our case, we have normalized the data such that $\max(P_{in}) = \max(P_{out-des})$. The quantities N_1 and N_2 are chosen near to one, in order give some tolerance to the optimizer. In fact, the two upper inequalities can be written in the form of equality constraints, but since equality constraints are difficult to handle, due to numerical consideration and also because of the basic properties of the MADS algorithm, we avoid using equality constraints.

The optimization results for these two sinusoids are plotted in Figure 4.19, while the system response before using a predistorter has already been shown in Figure 4.16. It can be observed that the amplitude of the output waveforms in both components has reached that of the input signal which is expected since the linearized gain of the system is set to one. Furthermore, both outputs exhibit sinusoidal wave behaviour; and, they coincide very well with their fundamental frequency component.

For a better understanding of what happens to the input waves after being launched into the system, we may look more closely at the output of each stage. To observe the individual characteristics of the predistorter and the transmitter model, the input signal's total power and the predistorter's total output power are compared in Figure 4.20, where

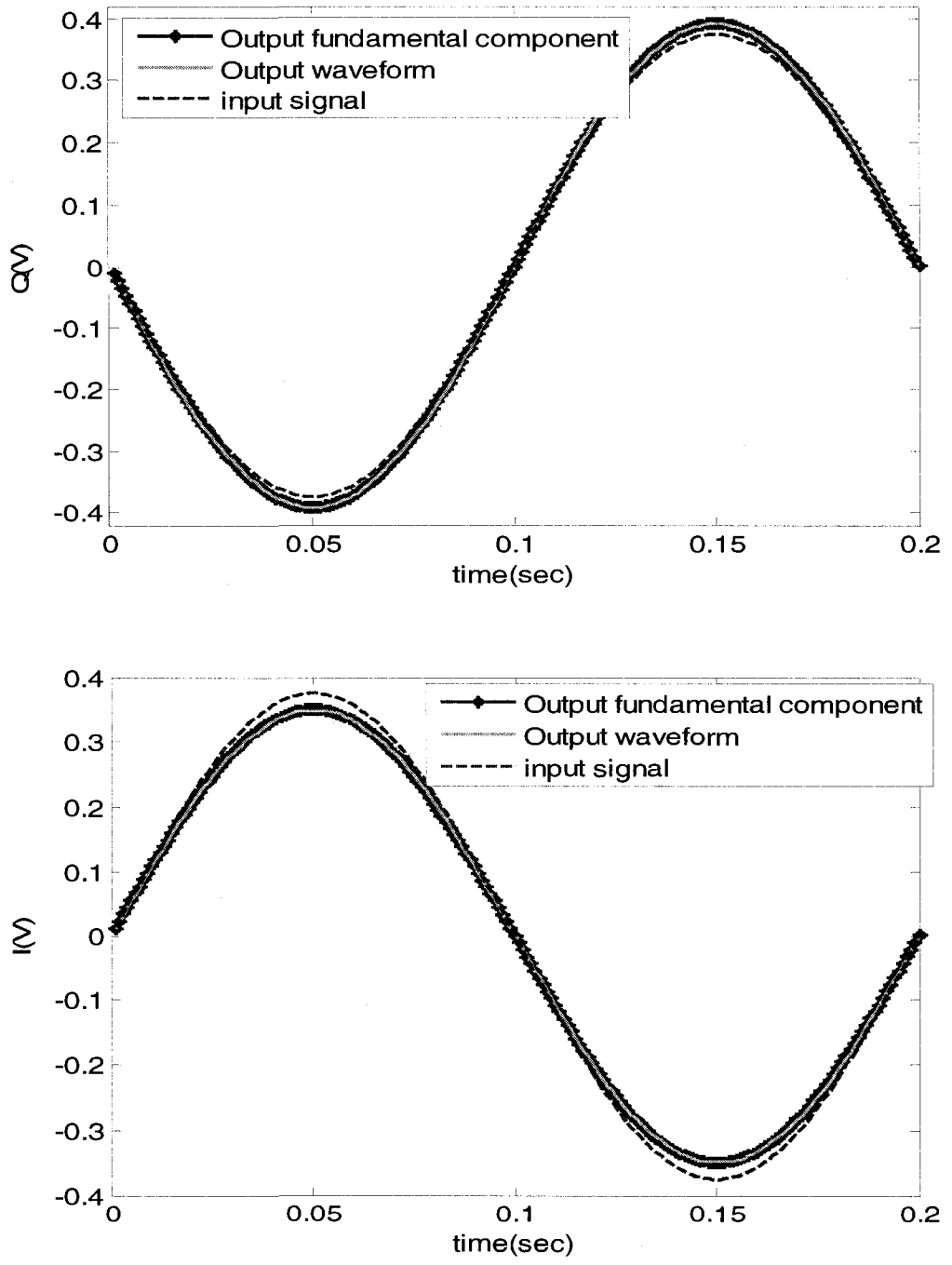


Fig.4.19. The in-phase (lower picture) and quadrature (upper picture) components of the input, output and output fundamental frequency component after running the optimization

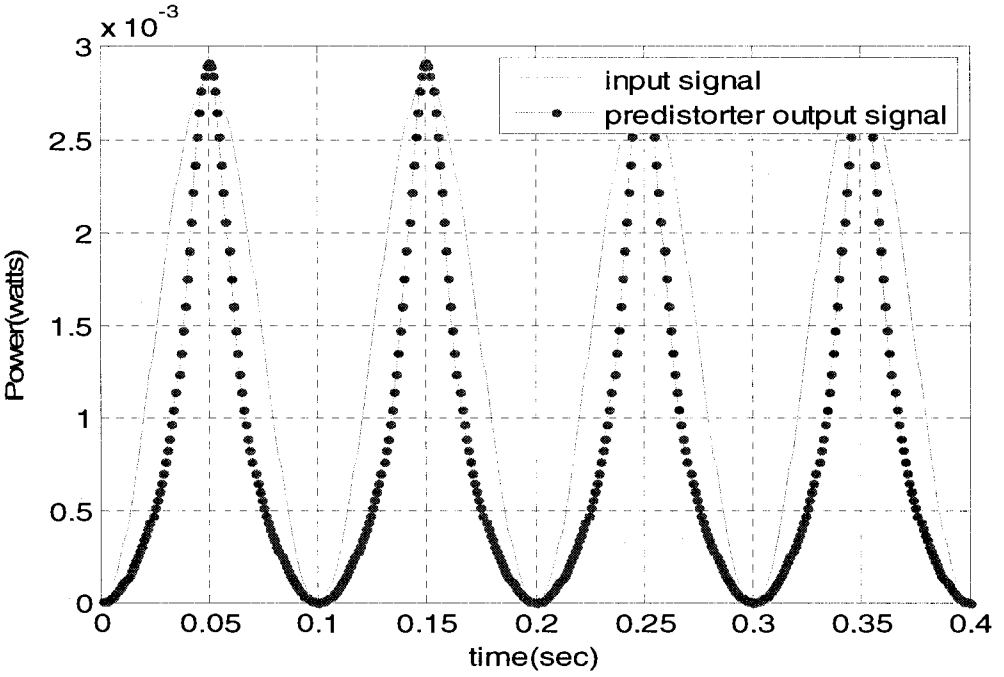


Fig. 4.20. The input signal power and the predistorter output signal power

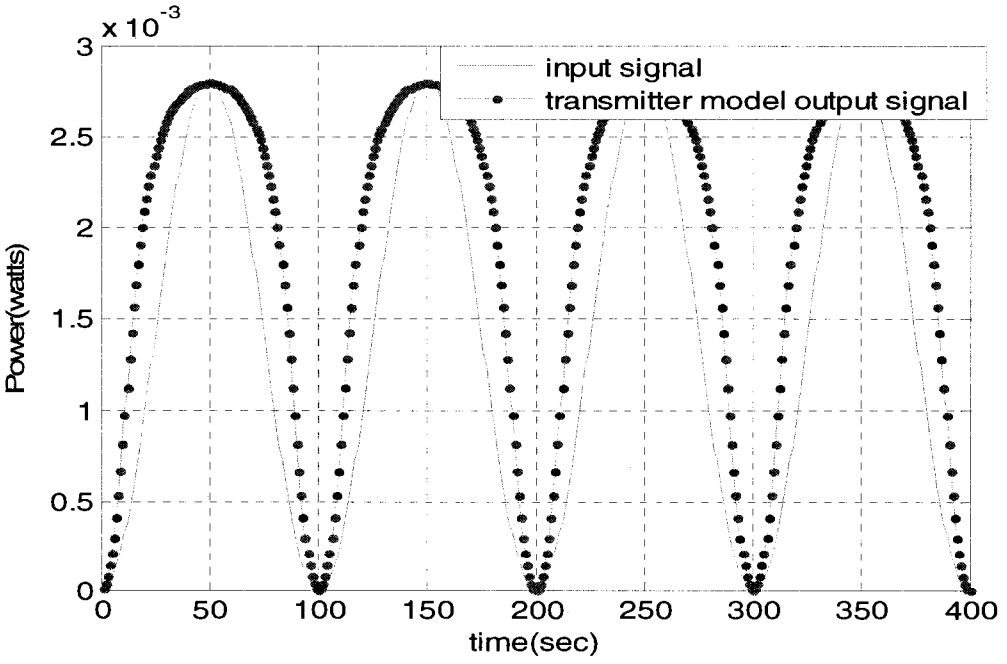


Fig. 4.21. The input signal power and the transmitter model output signal power without predistortion

the output power profile is in accordance with what is expected from an inverse function. The transmitter's model performance alone, with the same input, is plotted in Figure 4.21, where the clipping effects start to kick in when the input power level increases. As can be seen in the graphs, the maximum power at the output of the predistorter increases slightly beyond that of the input, which is due to the tolerance provided for the constraints of the optimization algorithm.

It is also interesting to look separately at each baseband component at the output of the predistorter. They are plotted in Figure 4.22, along with their equivalent input components and the output of the transmitter model, acting alone without predistorter. Although not much information can be extracted from these curves, their combination leads to the nearly sinusoidal signals in Figure 4.19.

To see how linearly the complete predistorted system is operating, the output power is plotted versus input power in Figure 4.23. By paying attention to the points where the plot intersects the grid, one can deduce that the system is pursuing a linear manner. More precisely, the input-output power ratio is plotted in Figure 4.24, where one can observe that the ratio varies in the interval of 0.995 to 1.0321, which is quite agreeable.

As for the phase distortion, the individual performance of the predistortion block and the 2-box model in the absence of predistorter, in exhibiting nonlinear phase profiles, are plotted in Figure 4.25, which were obtained for the same excitation signal. However, the distortion of the system is initiated at zero where that of the predistorter starts, just to ease the comparison. For a more precise evaluation of the phase distortion cancellation, the final phase shift of the predistorted system is also depicted in Figure 4.26, as can be

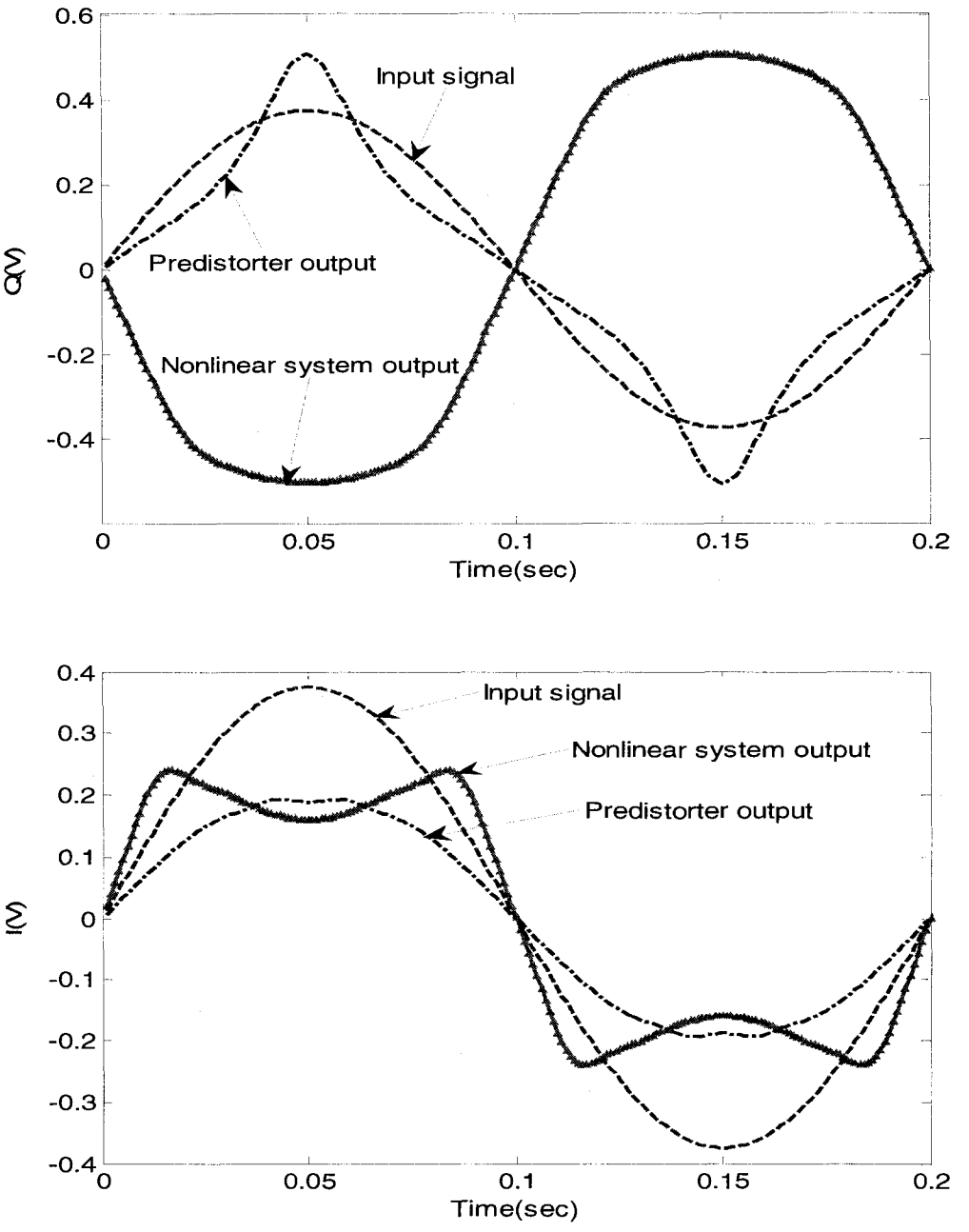


Fig. 4.22. Predistorter output signal, nonlinear system output signal without predistortion, and the input signal comparison for quadrature (upper picture) and In-phase components (lower picture)

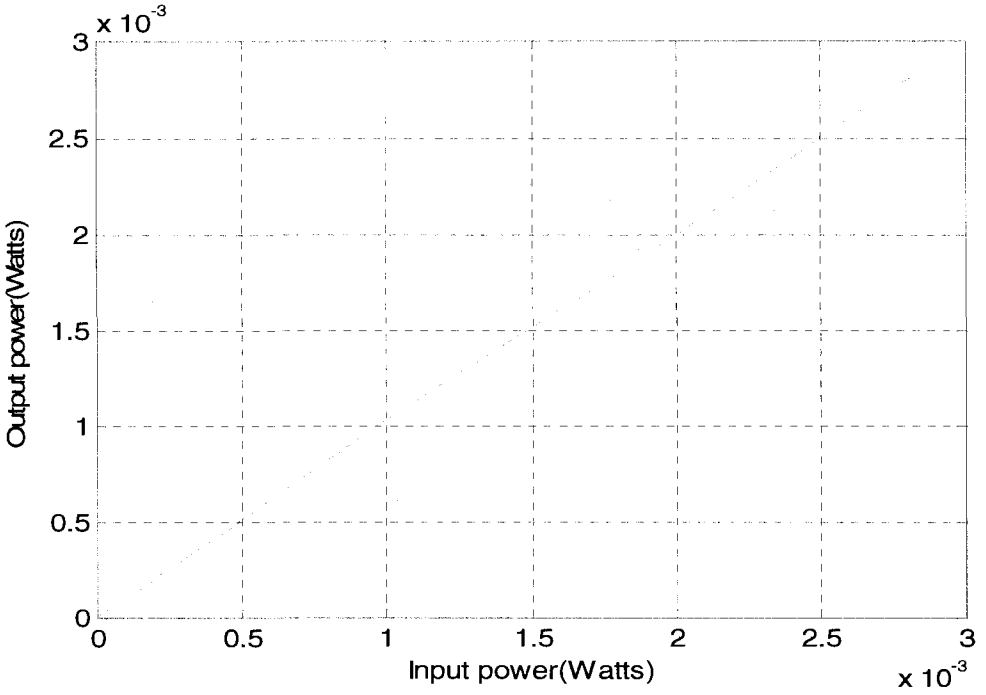


Fig. 4.23. The predistorted output signal power versus input signal power

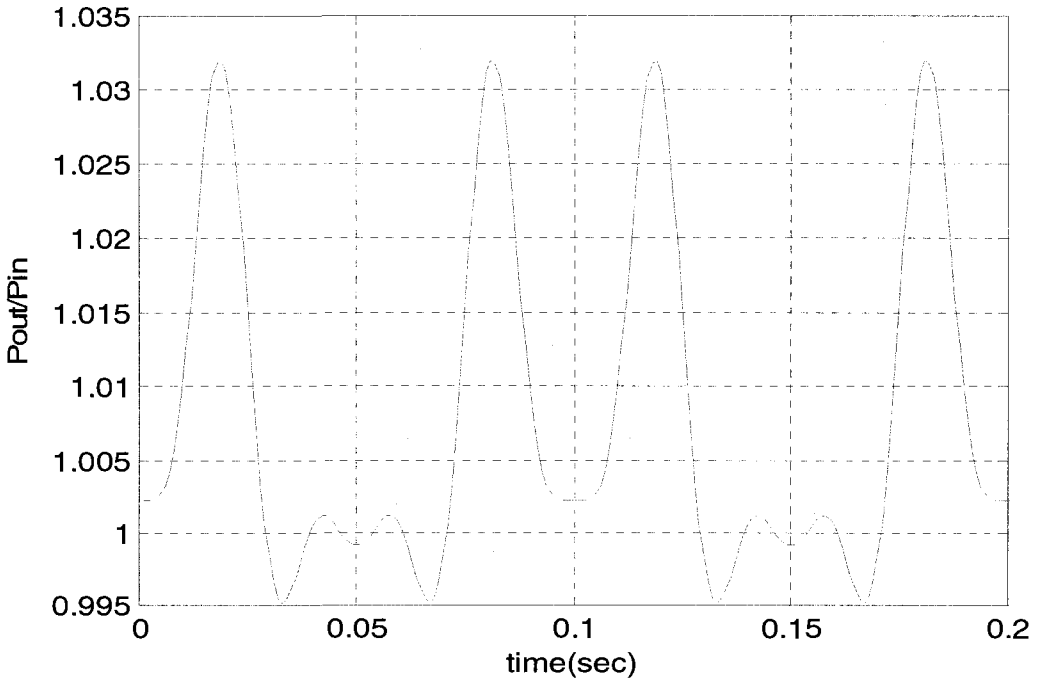


Fig. 4.24. The output-input power ratio for the linearized system in time plot

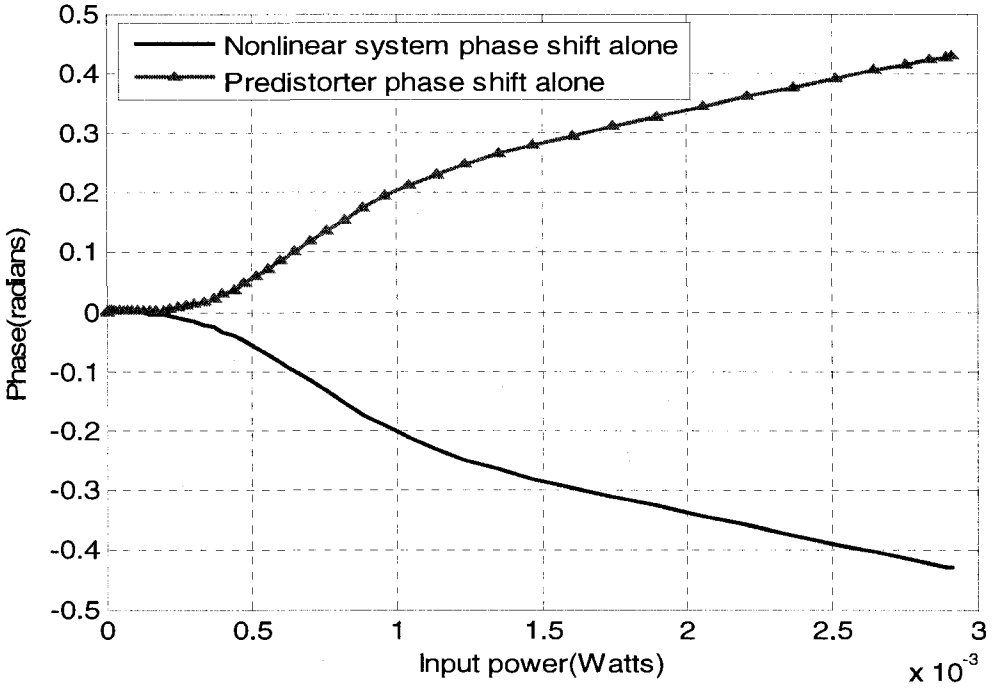


Fig. 4.25. Relative phase variation of the predistorter alone and the nonlinear model alone

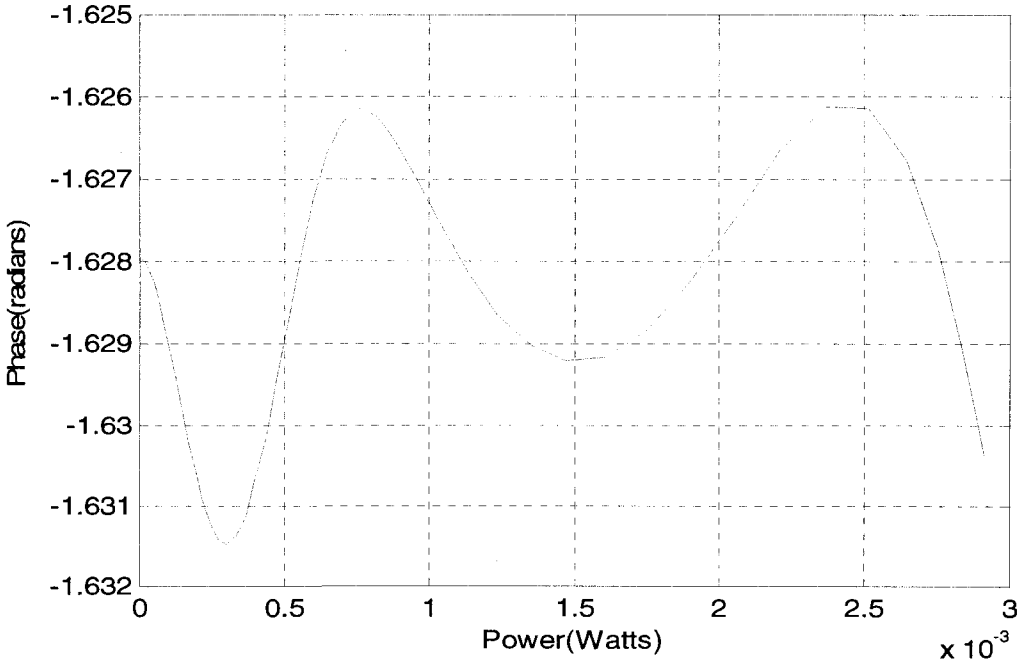


Fig. 4.26. Phase variation for the linearized system

observed, the phase variation, for the given input power range, falls in the interval of -1.6261 to -1.6315, which is acceptable as a figure of linearity compared to the strong phase variation introduced by the transmitter model alone.

In one last step, we calculate the quality of fit for the output quadrature and in-phase components and their equivalent fundamental frequency extracted terms, using the following relation, which was also utilized in Chapters 2 and 3:

$$FIT = \left[1 - Norm(Y - \hat{Y}) / Norm(Y - Mean(Y))\right] * 100 \qquad (4.41)$$

where, Y is set to be the output component, and $\hat{Y}$ is its equivalent extracted fundamental. The results are shown in Table 4.2. These results have been obtained after a couple of optimizations with surrogate functions.

Table 4.2. Quality of fit figures for the in-phase and quadrature components at the output of the linearized model

Component	Number of parameters	Best fit percentage
In-Phase	7	99.5105
Quadrature	7	99.3848

4.4 Complete predistorted system and linearization validation with CDMA signal

Having constructed the predistortion with parameters obtained at the last iteration of the optimization process, we can now validate the efficiency of the linearizer by launching another signal and observing the performance of the predistorter in cancelling

nonlinear distortions. The signal used for the validation phase is the same CDMA signal that was utilized in Chapters 2 and 3 for modeling. It is worth mentioning that the sinusoidal wave used in the previous section has been chosen such that its maximum power equates or is more than that of the CDMA signal. The predistorter output power is compared with the input power in Figure 4.27 As can be observed, for higher input power, the corresponding output power is higher. In other words, higher gain is applied on a higher input level, implying the nonlinear characteristics we have expected.

Figure 4.28 reveals this inverse behaviour more clearly. Looking at the power relation between the output of the system and the input power in Figure 4.29, it can be seen that the system is linearized to a good degree. This power ratio, which should be around one for the normalized model, is plotted in Figure 4.30. It shows a variation in the interval of

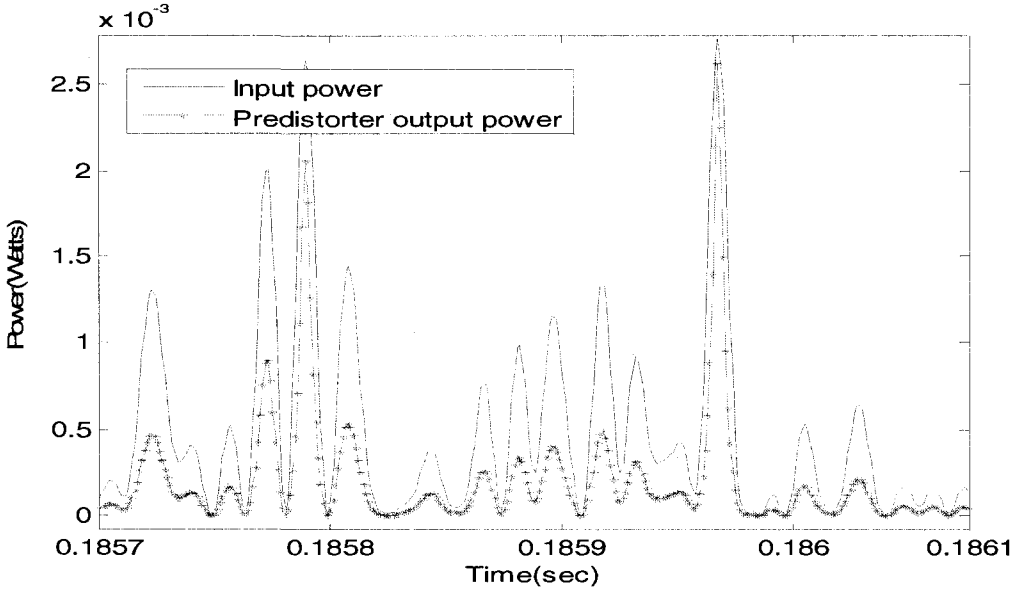


Fig. 4.27. The input power signal and the predistorter output power signal plots

0.995 to 1.0322. Although this is shown for a small portion of data, the result for the complete data frame is very similar. As for the phase distortion, the individual performance of the predistorter and that of the transmitter model have been plotted, as shown in Figure 4.31. Their overall phase distortion result, which is supposed to have been suppressed, is shown in Figure 4.32, where the phase variation is in the interval of -1.6216 to -1.6315 radians. It can be seen that the system is linearized, in terms of phase and gain distortion, where the linearized gain is one, and the phase distortion is concentrated around the small signal linear phase shift, which is about -1.6 radians. The final output power signal is plotted in Figure 4.33, along with the input power signal. For the adjusted linearized gain to be one, the two curves must coincide with good precision, which can be also compared with the one plotted in Figure 4.27.

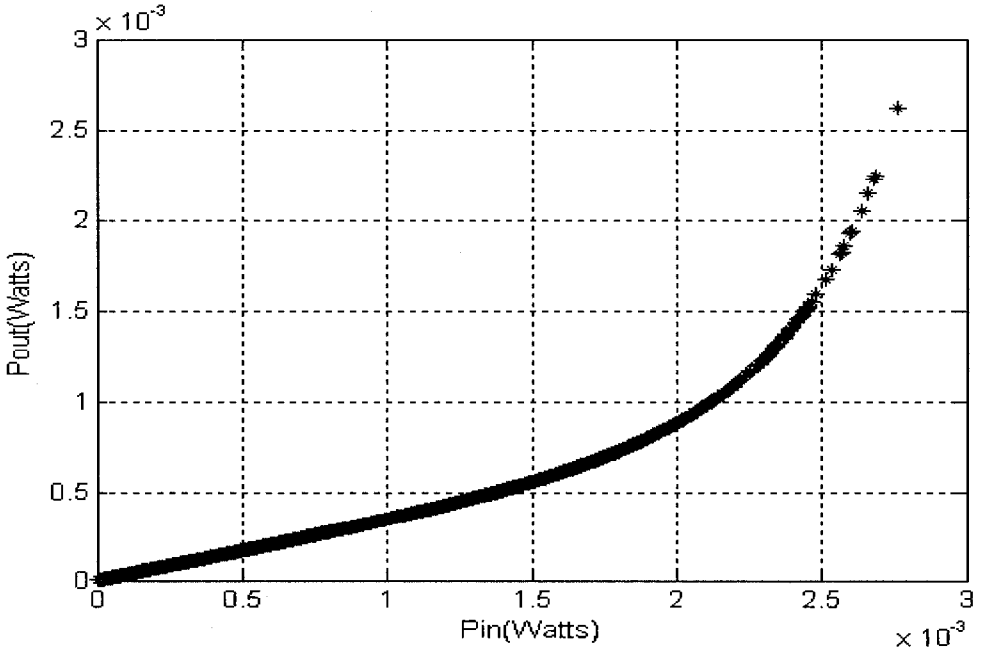


Fig. 4.28. The predistorter output power versus input power

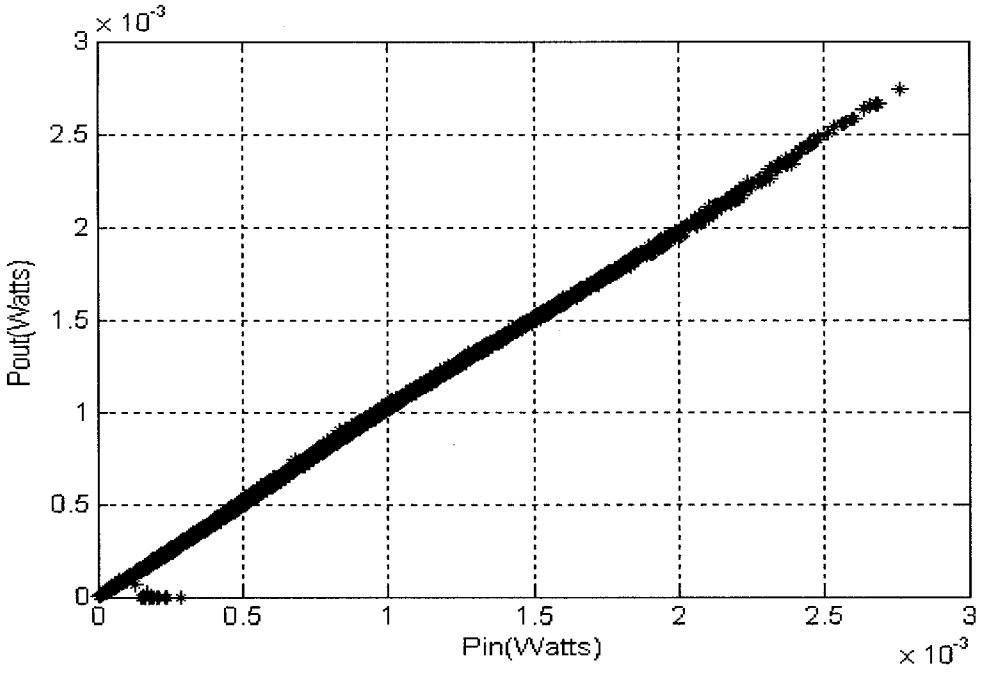


Fig. 4.29. The output power versus input power for the linearized system

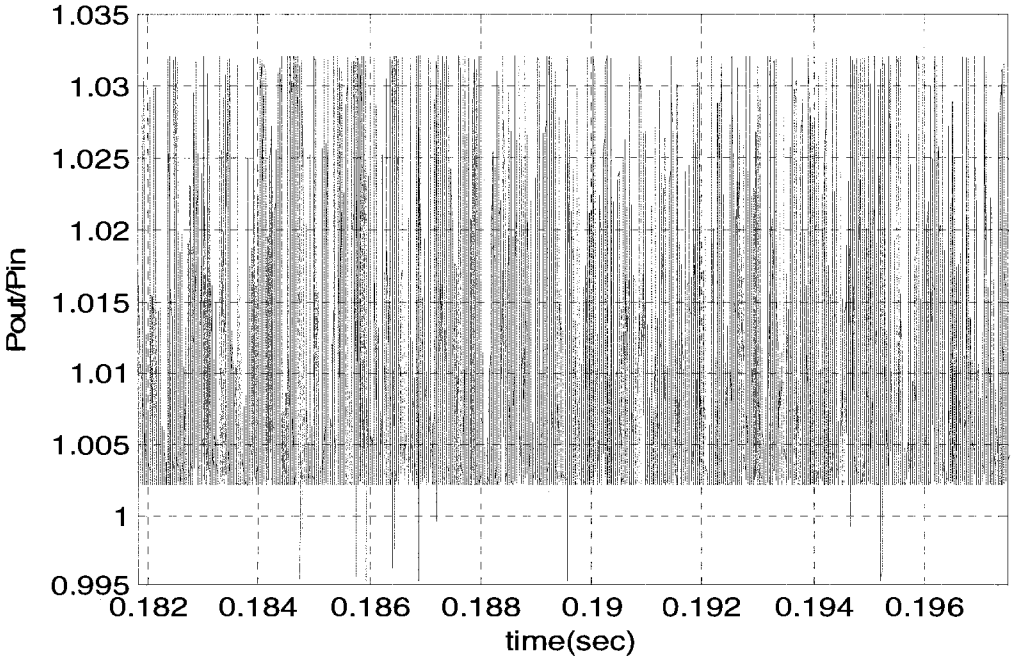


Fig. 4.30. Output-input power ratio for the linearized system

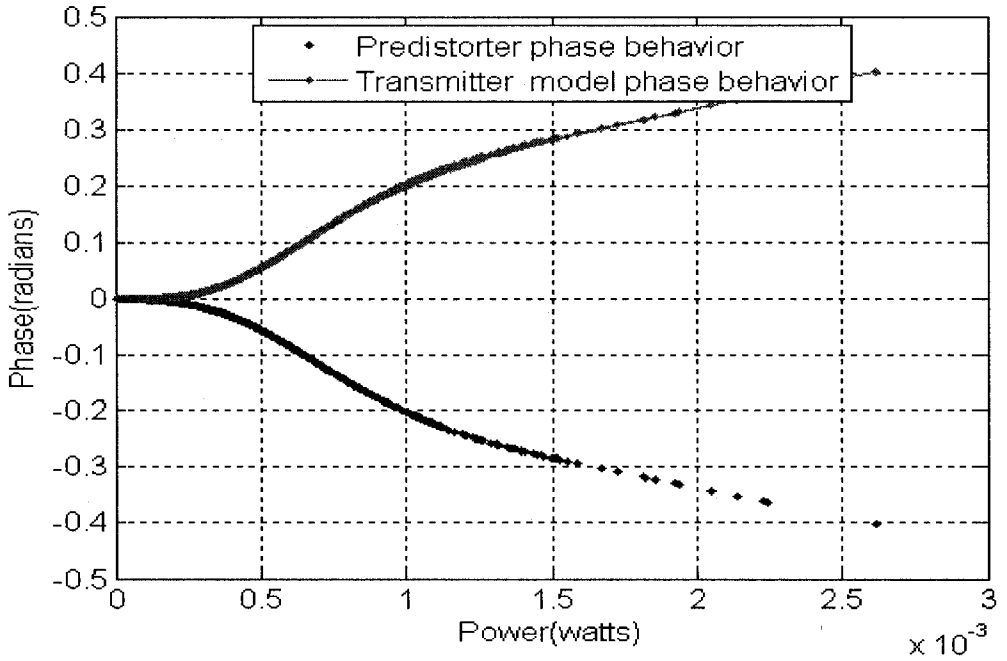


Fig. 4.31. Relative phase variation of the predistorter alone and the transmitter model alone

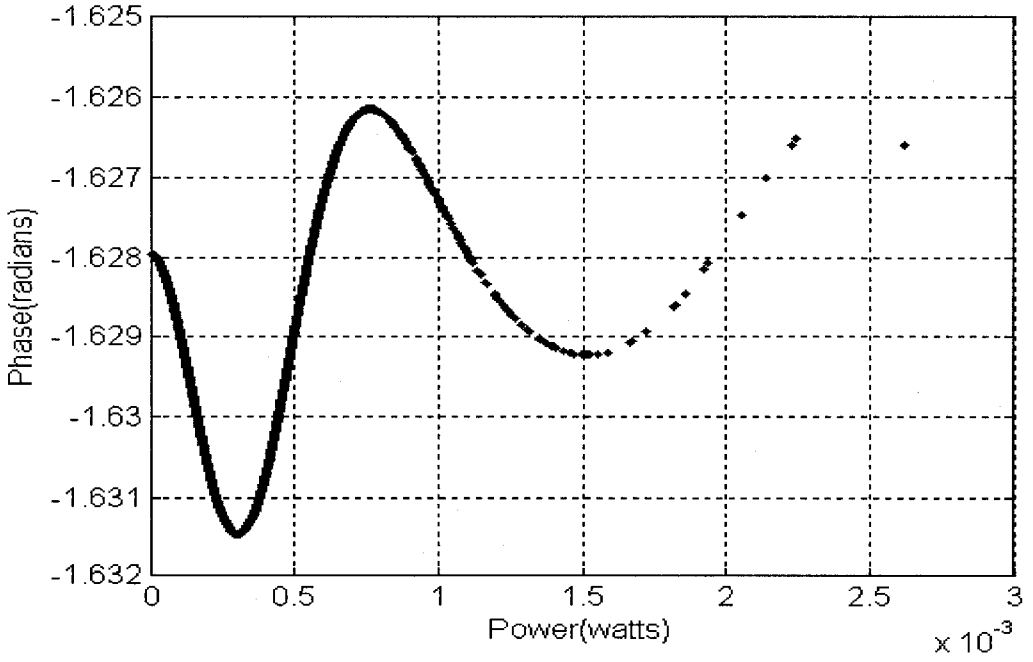


Fig. 4.32. Phase variation of the linearized system

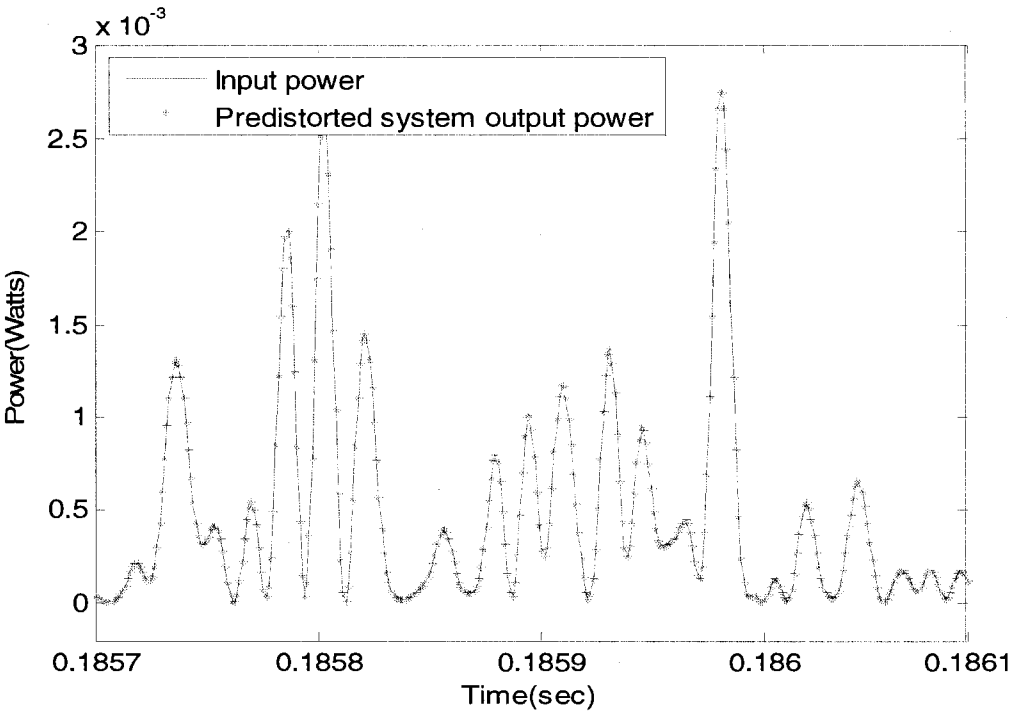


Fig. 4.33. The output signal power of the predistorted system and the input signal power

The power spectrum of output signal can also be observed for the predistorted system, when compared with the response of the system before using the predistorter. The results are plotted in Figures 4.34.a, 4.34.b and 4.34.c, where good results in terms of adjacent channel spectrum regrowth suppression have been obtained. Meanwhile, the main channel power level for both curves remains at the same level. The adjacent channel spectrum regrowth suppression is roughly about 15-20dBm. This spectrum re-growth suppression which is one of the main purposes of linearization technique is indeed an unmistakable signature of the nonlinearity.

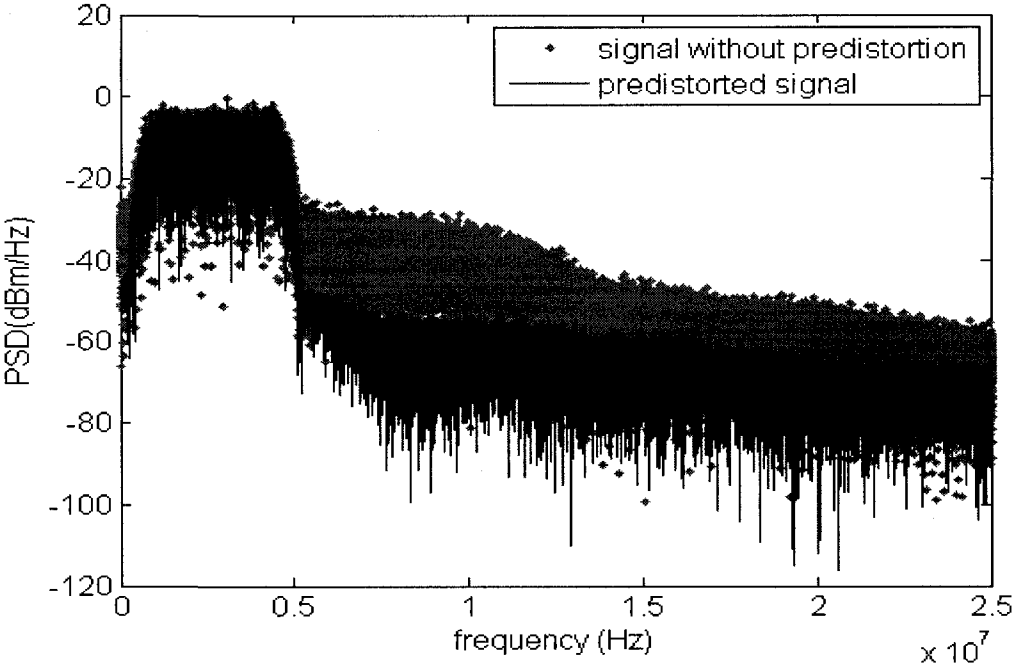


Fig. 4.34.a. Power spectrum density of the predistorted signal and the signal without predistortion

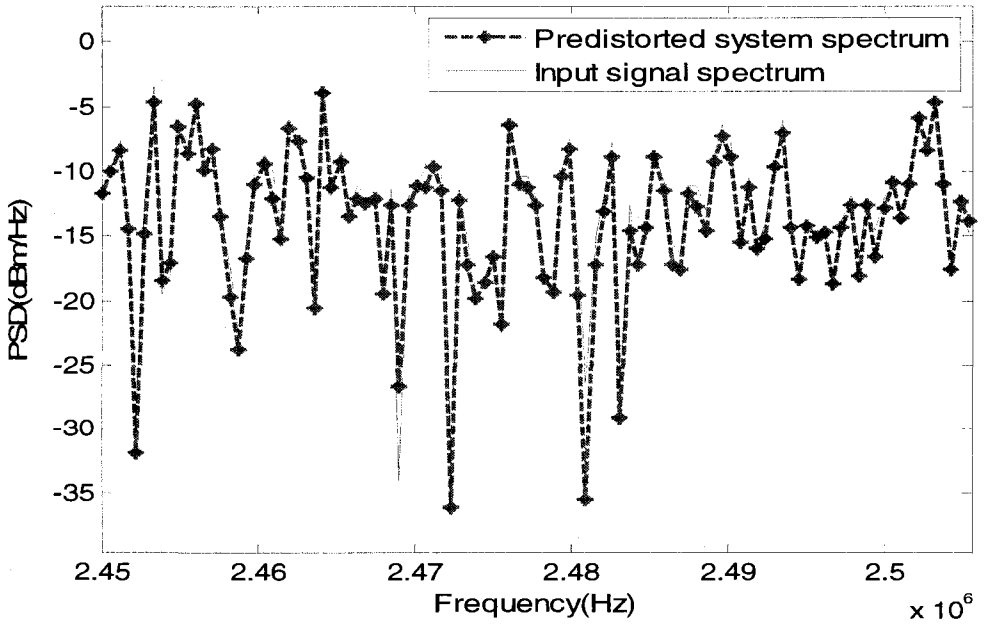


Fig. 4.34.b. Power spectrum density in a closer view for the main channel

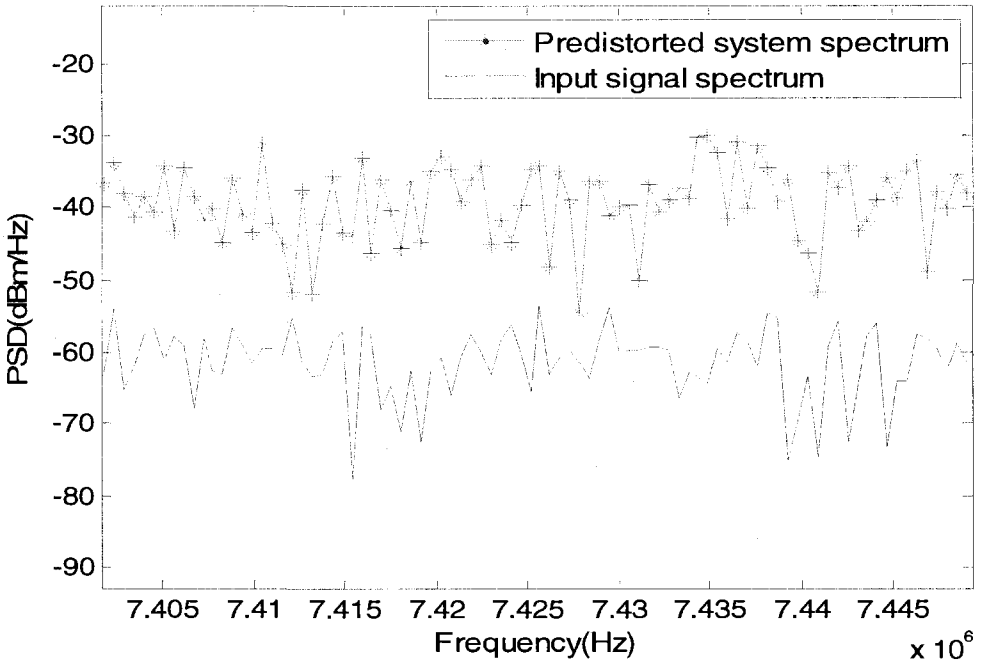


Fig. 4.34.c. Power spectrum density in a closer view for the adjacent channel

A linear gain adjustment must now be applied, so that the linearized gain is set back to the actual linearized gain before normalizing the transmitter model profile, which is, indeed, less than the unit gain we have obtained. The AM/AM and AM/PM curves are obtained and compared to those of the system without predistortion. These curves are plotted in Figures 4.35.a and 4.35.b, representing a relatively good linear behaviour. The slight deviation seen in the AM/AM curve for a higher input power level can be further suppressed using a higher order model for gain predistorter, consequently resulting in a more difficult optimization process. Having linearized the system, the linear output-input relation can now be identified by any linear identification procedure, where a variety of linear methods available in standard MATLAB toolboxes can be applied for this purpose or for performing any equalization.

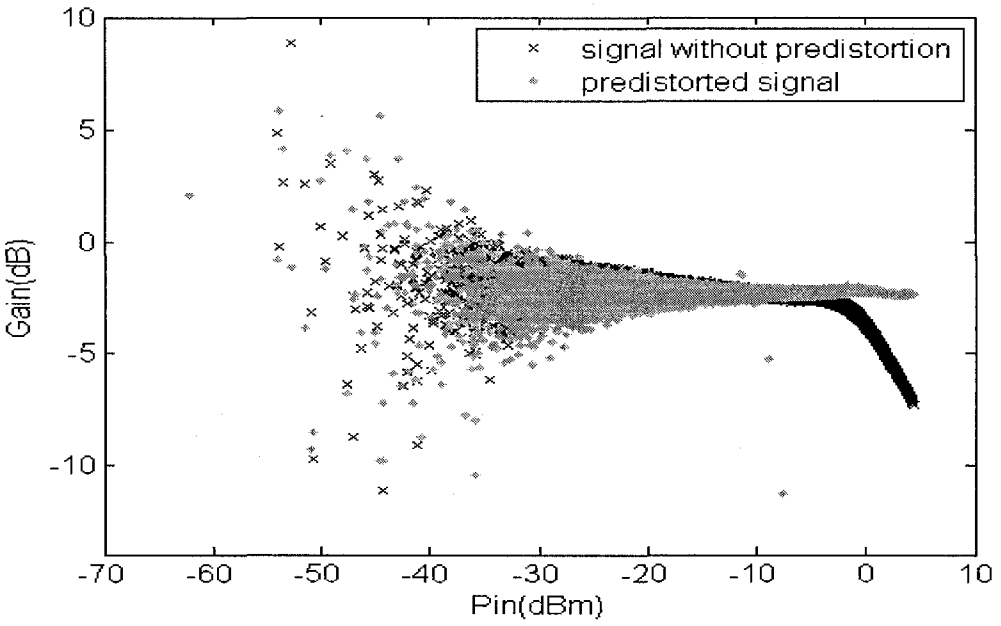


Fig. 4.35.a. The AM/AM curves for the predistorted CDMA signal and the signal without predistortion

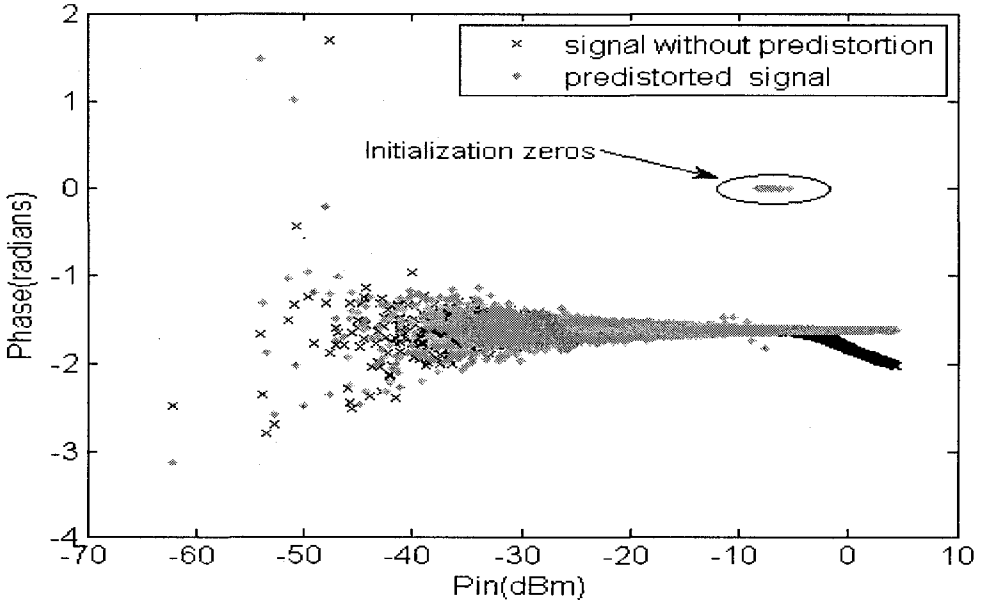


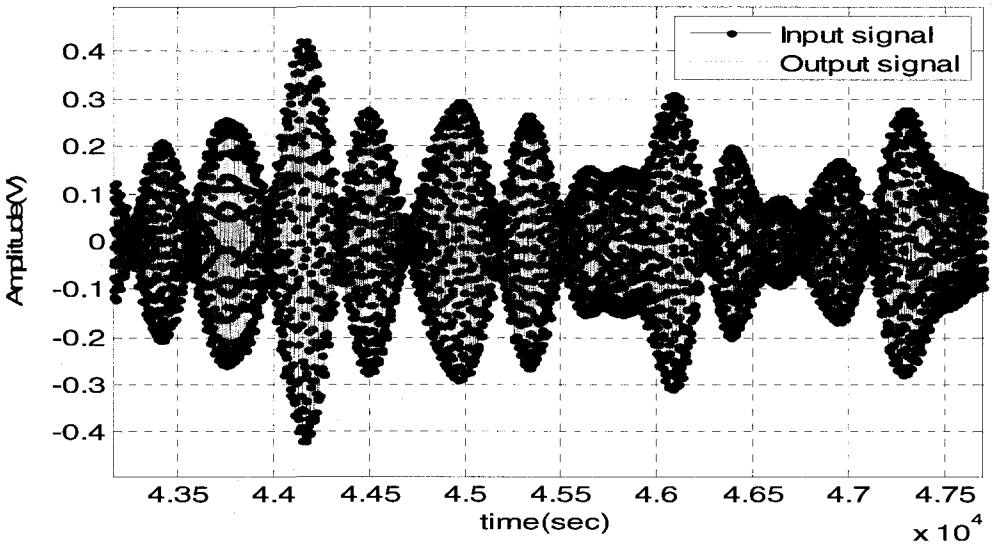
Fig. 4.35.b. The AM/PM curves for the predistorted CDMA signal and the signal without predistortion

Now we can build up the single real valued passband signal from the two quadratic components and identify it with a real valued linear filter. Later, and if necessary, according to the application, a simple frequency transfer will convert the band-pass model to its equivalent low pass. Therefore, the training signals for the input and output are expressed in the following forms:

$$v_{in}(t) = I_{in}(t)\cos(\omega_c t) - Q_{in}(t)\sin(\omega_c t) \tag{4.42}$$

$$v_{out}(t) = I_{out}(t)\cos(\omega_c t) - Q_{out}(t)\sin(\omega_c t) \tag{4.43}$$

Both band-pass signals are plotted in Figure 4.36.a. Figure 4.36.b provides a closer view. The linear identification is performed on the unit gain output before the gain adjustment. ARX structures with various orders have been attempted using linear system identification techniques available in standard MATLAB toolboxes to model the linear part. The fitting results for the models are given in Table 4.3, where it can be seen that



4.36.a. The input and output modulated signals for the linearized system

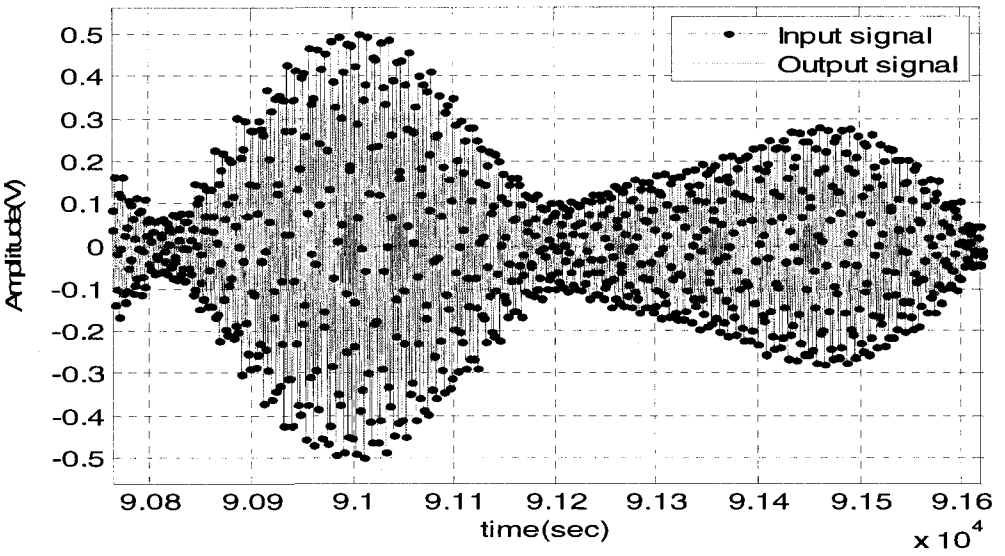


Fig. 4.36.b. The input and output modulated signals for the linearized system in a smaller interval of time

the best percentage have been achieved by the FIR filter, proving that we have, to a good degree, linearized the transmitter model. Further frequency shaping and filtering techniques can be now applied according to the application.

Table 4.3 Quality of fit of the linear model for a different number of tabs

ARX order [n_a n_b n_k]	Fitting percentage
[4 16 1]	94.21
[4 4 1]	94.23
[0 16 1]	94.67
[0 32 1]	96.39
[0 64 1]	96.8
[0 128 1]	97.47

CHAPTER 5

CONCLUSIONS

5.1 Thesis Summary

POWER amplifier nonlinear behaviour is a delicate task to tackle when it comes to wideband transmitter design for modern high-speed wireless communication systems, where highly variable envelope signals are launched into the system. The nonlinear behaviour that shows up at high powers gives rise to signal distortion, in terms of gain and phase shift. Hence, the operating power as the main source of nonlinearity, must be either at a large enough back-off from the nonlinear region to maintain the input-output linearity or, on the other hand, in the nonlinear region transmitting the maximum power and dealing with nonlinear effects and harmonics generation.

Since we are interested in transmitting the maximum power available by a transmitter in many applications, a variety of alternatives have been developed to linearize the system up to its maximum power operating region. One of the most common methods is the utilization of a predistorter, which is characterized in such way as to introduce the inverse behaviour of the transmitter; and, linear behaviour will be observed when the predistorter is cascaded with. In order to characterize a predistorter well, one must initially have a good identification of the transmitter itself. This would be of great help when it comes to optimizing the predistorter, where a rough estimate of the inverse model can ease the convergence by initializing the algorithm at a proper point, aside from the possibility of eliminating a couple of optimization variables.

In the introductory chapter, a summary of a variety of modeling and identification approaches were presented and compared. In Chapter 2, three types of models were constructed; and, their abilities to reproduce the actual transmitter behaviour, in terms of gain and phase, were compared. Simple passband Hammerstein and Wiener models were introduced and, despite the fact that they succeeded very well in simulating the nonlinear gain distortion seen in the transmitter, they failed to pursue the nonlinear phase variation of the system. This is due to their inherent structure, which makes them incapable of producing any nonlinear phase, since the filter embedded in them is a linear real valued type and the static nonlinearity serves to produce nonlinear gain only.

Due to this significant weakness in exhibiting nonlinear phase behavior, passband Wiener/Hammerstein models are not efficient to characterize the most important features of the nonlinearities of the power amplifier. As a consequence, a general two-box nonlinear model, which is widely used in RF transmitter modeling, was introduced and constructed. This model, as opposed to the previously mentioned Hammerstein and Wiener models, is capable of reproducing the phase distortion observed in the transmitter; however it has a completely complex structure via which both gain and phase nonlinear variation or their equivalent rectangular representation known as in-phase and quadrature components are respectively identified. The quality of fit for this type of model exhibits a considerable improvement when compared to that of the passband Wiener/Hammerstein models.

In Chapter 3, a three-box model was introduced and constructed using a modulated signal. This model, which included two nonlinear blocks surrounding a linear filter, used

a modulated signal and was implemented in a two-stage identification process, where the first two blocks were identified concurrently using the RF signals as the training data. The first nonlinear block modeled the nonlinearity observed in the gain variation of the transmitter and was built up in the very same manner as the Hammerstein model. The output of the linear filter was then transformed to its equivalent baseband data in order to have direct access to the phase variation. A comparison between the phase, provided by the first two boxes and the actual phase distortion of the transmitter, led to a second nonlinear box for phase behaviour. The last two blocks, however, can be considered as one nonlinear phase filter, due to the fact that the third block, indeed, changed the constant phase provided by the linear filter. The quality of fit calculated for the same set of data for the three-box model and the two-box model developed in Chapter 2 was seen to be slightly better, both for the identification and validation steps, which were performed on different sets of data, however this type of model has the disadvantage of operating on both passband and complex data, in that, the third block needs to perform a transformation on its passband input before shaping the phase variation.

In Chapter 4, a new predistortion method was proposed, which utilized a simple single tone sinusoidal wave for training. The key assumptions throughout this chapter are as follows: the complex 2-box model is an accurate representation of the transmitter; the nonlinear distortions observed in the transmitter are exclusively power (and not frequency) driven whereby, the main feature of the training signal will be its power level and not its frequency contents or its bandwidth.

The idea came from the fact that, for a well linearized system, a sinusoidal input will remain a sinusoid of the same frequency, but with different amplitude and some phase shift. We are concerned about both gain distortion and phase distortion and the actual system receives two input data, known as quadrature and in-phase data components, for wideband transmission schemes. Therefore, one needs to verify the linearity criteria for such a system, where a single sinusoidal cannot reveal the phase distortion. Aside from practical points to be considered for the matter of transmitting the maximum power, a series of critical considerations on the optimizer needed to be taken so that the chance of convergence increased. Through the formulation of the problem of the response of a linear system with two sine inputs, it was deduced that the output must be also two sinusoids of the same frequency, with amplitudes so that the maximum output power was transmitted and, at the same time, phase linearity was held. The rough estimation of the model inverse also served as a good choice of an initial point. In addition, some scaling factor needed to be considered for the matter of mesh refinement or coarsening that took place at each iteration.

Having constructed the predistorter with the sinusoidal wave, we could now validate its performance by launching another signal, which is, indeed, the goal of transmission, in our case a CDMA signal, to the same predistorted system. The results obtained in Chapter 4 exhibit a relatively good linear behavior, where the AM/AM and AM/PM curves had a constant value tendency. Moreover, the whole system can be again identified by a linear ARX model. The results show that up to 94% of the data fit the linear model, which confirms the effectiveness of the linearizer. It is worth mentioning

that, although we have used a model instead of a real setup, it was treated like a black box; and, the information of the model was not used somewhere in the algorithm, aside from an initial point estimation.

5.2 Future work

The 3-box modeling approach can lead up to a novel pre-post distortion method to compensate for the nonlinear distortion in the transmitter where part of the nonlinearity can be eliminated before the transmitter and part of in a post distortion procedure. Other possible oriented box configuration for modeling the power amplifier could be further investigated using the modulated RF signal and a nonlinear filter which accounts for the phase distortion as well as the gain distortion.

The sine wave adaptive predistortion approach proposed could be further evaluated for other types of transmission scheme such as QAM, OFDM, etc and the possibility of using a single sine instead of two could be studied by separating the quadrature and in-phase components of the fundamental term at the output itself.

REFERENCES

[1] S.A. Mass, "Volterra Analysis of Spectral Regrowth," *IEEE Microwave& Guided Wave Lett.,* Vol. 7, No. 7, July 1997.

[2] B. Shi, L. Sundstrom, "Linearization of RF power amplifier using power feedback," VTC 2003-Spring., Vol. 2, pp. 1345-1349, April 2003.

[3] T.A. Johansen, K.J. Hunt, "A computational approach to approximate input/state feedback linearization," *Decision and Control, 2000. Proc.*, Vol. 5, pp. 4467-4472, June 2000.

[4] I. Meier, J.B. De Swardt, "Error-feedback for amplifier linearization," *COMSIG, 1998 Proc*, pp. 381-386, June 1998.

[5] F. Tabatabai, H.S. Al-Raweshidy, "Feedforward Linearization Technique for Reducing Nonlinearity in Semiconductor Optical Amplifier," *Lightwave Tech.*, Vol. 25, pp. 2667-2674, September 2007.

[6] M. Faulkner, "Amplifier linearization using RF feedback and feedforward techniques," *IEEE Trans. Vehicular Tech.*, Vol. 74, pp. 209-215, February 1998.

[7] A.J. Zozaya, E. Bertran, "On the performance of Cartesian feedback and feedforward linearization structures operating at 28 GHz," *2004 IEEE Trans. Broadcast.*, Vol. 50 , pp. 382-389, December 2004.

[8] A. Birafane, A.B. Kouki, "Phase-only predistortion for LINC amplifiers with chireix-outphasing combiners," *IEEE Trans. Microwave Theory & Tech.*,Vol. 53, pp. 2240-2250, June 2005.

[9] S. McBeath, D. Pinckley, "Digital memory-based predistortion," *2005 IEEE MTT-S Int, Microwave Symp-Dig.*, June 2005.

[10] R.N. Braithwaite, "Memory Correction of a Doherty Power Amplifier with a WCDMA Input using Digital Predistortion," *Microwave Symp Dig, 2006 IEEE MTT-S Int.*, Vol. 1, pp. 1526-1529, June 2006.

[11] H. Alasadi, R. Boutros, M. Ibnkahla, "Comparison between digital and analog predistortion for satellite communications," *Elec & Comp Eng, 2003. Canadian Conference.*, Vol. 1, pp. 183-186, May 2003.

[12] F.C. Lin, D.M. Holburn, V.A. Lalithambika, R.J. Mears, C.H. Hum, S.D. Walker, "A CMOS analogue predistortion circuit for wideband optical fiber links," *Elec & Comp Eng, 2004. Canadian Conference.*, Vol. 1, pp. 245-248, May 2004.

[13] S. Andreoli, H.G. McClure, P. Banelli, S. Cacopardi, "Digital linearizer for RF amplifiers," *IEEE Trans. Broadcast.*, Vol. 43, pp. 12-19, March 1997.

[14] L. Hanzo, S. X. NG, T. Keller, W. Webb, Quadrature Amplitude Modulation—From basics to adaptive Trellis-Coded, Turbo-Equalized and Space-Time Coded OFDM, CDMA and MC-CDMA systems, The Atrium, Southern Gate, Chichester, England: John Wiley& Sons Ltd, 2004.

[15] J. Vuolevi, "Analysis, Measurement and Cancellation of the Bandwidth and Amplitude Dependence of Intermodulation Distortion in RF Power Amplifiers," *Academic Dissertation to be presented with the assent of the Faculty of Technology, University of Oulu.*, Nov 2001.

[16] S. Maas, *Nonlinear Microwave Circuits.* Norwood, MA: Artech House, 1988.

[17] O. Hammi, F.M. Ghannouchi, S. Boumaiza, B. Vassilakis, "A Data-Based Nested LUT Model for RF Power Amplifiers Exhibiting Memory Effects," *Microwave& Wireless Comp Lett, IEEE.*, Vol. 17, pp. 712-714, October 2007.

[18] H.H. Chen, C.H, Lin, P.C. Huang, J.T. Chen, "Joint Polynomial and Look-Up-Table Predistortion Power Amplifier Linearization," *IEEE Trans. Circuits & Sys.*, Vol. 53, pp. 612-616, August 2006.

[19] A. Zhu, T.J. Brazil, "Behavioral modeling of RF power amplifiers based on pruned volterra series," *Microwave& Wireless Comp Lett, IEEE.*, Vol. 14, pp. 563-565, December 2004.

[20] N. Norholm, C. Iversen, T. Larsen, "GaAs MESFET large-signal modelling for multiport Volterra series analysis," *Circuits, Devices & Sys, IEE Proc.*, Vol. 144, pp. 40-44, February 1997.

[21] A. Zhu, J.C. Pedro, T.R. Cunha, "Pruning the Volterra Series for Behavioral Modeling of Power Amplifiers Using Physical Knowledge," *IEEE Trans. Microwave Theory & Tech.*, Vol. 55, pp. 813-821, May 2007.

[22] Z. Qingsheng, E. Zafiriou, "Control-relevant identification of Volterra series models," *American Control Conference, 1994.*, Vol. 2, pp. 2050-2054, June 1994.

[23] I. Vassiliou, S. Sangiovanni-Vincentelli, "A frequency-domain, Volterra series-based behavioral simulation tool for RF systems," *Custom Integ Cir, Proc IEEE.*, pp. 21-24, May 1999.

[24] Z. Anding, T.J. Brazil, "An adaptive Volterra predistorter for the linearization of RF high power amplifiers," *Microwave Symp Dig, 2002 IEEE MTT-S Int.*, Vol. 1, pp. 461-464, June 2002.

[25] A. Ahmed, S.M. Endalkachew, G. Kompa, "Power amplifier linearization using memory polynomial predistorter with non-uniform delay taps," *Microwave Symp Dig, 2004 IEEE MTT-S Int.*, Vol. 3, pp. 1871-1874, June 2004.

[26] D.R. Morgan, Z. Ma, J. Kim, M.G. Zierdt, J. Pastalan, "A Generalized Memory Polynomial Model for Digital Predistortion of RF Power Amplifiers," *IEEE Trans. Signal Processing.*, Vol. 56, pp. 3852-3860, October 2006.

[27] L. Ljung, System Identification—Theory for User, Englewood Cliffs, NJ: Prentice Hall, 1987.

[28] Y. Wang, X. Xiang, Y. Ke-Chu, "Baseband Wiener Predistorter for Linearizing Power Amplifier," *Signal Processing, The 8th Int.*, Vol. 3.

[29] P. Gilabert, G. Montoro, E. Bertran, "On the Wiener and Hammerstein models for power amplifier predistortion," *APMC Proc, 2005.* , Vol. 2, December 2005.

[30] P. Crama, Y. Rolain, "Broadband measurement and identification of a Wiener-Hammerstein model for an RF amplifier," *ARFTG Dig.*, pp. 49-57, December 2002.

[31] L. Taijun, S. Boumaiza, F.M. Ghannouchi, "Deembedding static nonlinearities and accurately identifying and modeling memory effects in wide-band RF

transmitters," *IEEE Trans. Microwave Theory & Tech.*, Vol. 53, pp. 3578-3587, November 2005.

[32] L. Taijun, S. Boumaiza, F.M. Ghannouchi, "Augmented hammerstein predistorter for linearization of broad-band wireless transmitters," *IEEE Trans. Microwave Theory & Tech.*, Vol. 54, pp. 1340 -1349, June 2006.

[33] M. Sano, S. Lianming, "Identification of Hammerstein-Wiener system with application to compensation for nonlinear distortion," *SICE 2002. Proc.*, Vol. 3, pp. 1521-1526, August 2004.

[34] D.R. Morgan, Z. Ma, J. Kim, M.G. Zierdt, J. Pastalan, "A Generalized Memory Polynomial Model for Digital Predistortion of RF Power Amplifiers," *IEEE Trans. Signal Processing.*, Vol. 56 , pp. 3852-3860, October 2006.

[35] P. Jantunen, G. Gamez, T. Laakso, "Measurements and modelling of nonlinear power ampliers," *Signal Processing Symp, Proc 2004.*, Vol. 1, pp. 328-331, 2004.

[36] P. Gilabert, G. Montoro, E. Bertran, "On the Wiener and Hammerstein models for power amplifier predistortion," *Asia-Pacific Conference Proc.*, Vol. 2, pp. 4-7, December 2005.

[37] B. Er-Wei, "An optimal two stage identification algorithm for Hammerstein-Wiener nonlinear systems," *American Control Conference, Proc.*, Vol. 5, pp. 2756-2760, June 1998.

[38] D.R. Morgan, Z. Ma, J. Kim, M.G. Zierdt, J. Pastalan, "A Generalized Memory Polynomial Model for Digital Predistortion of RF Power Amplifiers," *IEEE Trans. Signal Processing.*, Vol. 54, pp. 3852-3860, October 2006.

[39] M. Sano, S. Lianming, "Identification of Hammerstein-Wiener system with application to compensation for nonlinear distortion," *SICE Proc.*, Vol. 3, pp. 1521-1526, August 2002.

[40] P. Varahman, Z. Atlasbaf, N.V. Heydarian, "Adaptive digital predistortion for power amplifiers used in CDMA applications," *Applied Electromag, Asia-Pacific.*, 20-21 December 2005.

[41] C. Potter, "System analysis of a W-CDMA base-station PA employing adaptive digital predistortion," *Microwave Symp Dig, IEEE MTT-S Int.*, Vol. 1, pp. 21-24, June 2002.

[42] H. Zhi-yong, G. Jian-hua G. Shu-jian , W. Gang "An improved look-up table predistortion technique for HPA with memory effects in OFDM systems," *IEEE Trans., Broadcast.*, Vol.52, pp. 87-91, March 2006.

[43] A.N. D'Andrea, V. Lottici, R. Reggiannini, "Nonlinear predistortion of OFDM signals over frequency-selective fading channels," *IEEE Trans., Comm.*, Vol. 49, pp. 837-843, May 2001.

[44] Qiu, J.X.; Abe, D.K.; Antonsen, T.M., Jr.; Danly, B.G.; Levush, B.; Myers, R.E, "Linearizability of TWTAs using predistortion techniques," *IEEE Trans., Electron Dev.*, vol.52, pp. 718 - 727 May 2006

[45] C. Audet, J.E. Dennis Jr., "Mesh Adaptive Direct Search Algorithms for constrained optimization," *SIAM Journal on Optimization,.*, Vol. 17, No.1, pp. 188-217, 2006.

[46] M. Abramson, NOMADM optimization software. http://www.afit.edu/en/ENC/Faculty/ MAbramson/NOMADm.html

[47] K. S. Shanmugam, Digital and analog communication systems, John Wiley & Sons, 1985

[48] T.P. Dobrowiecki, J. Schoukens, "Practical choices in the FRF measurement in presence of nonlinear distortions," *IEEE Trans., Instrumentation & measurement.*, Vol. 50, pp. 2-7, February 2001.