Efficient baseband digital predistortion techniques for linearizing power amplifier by taking into account nonlinear memory effect

Xiaowen Feng

Thèse de Doctorat

Xiaowen FENG

Mémoire présenté en vue de l'obtention
du grade de Docteur de l'Université de Nantes
Sous le label de l'Université Nantes Angers Le Mans

École doctorale Sciences et Technologies de l'Information et Mathématiques (STIM)

Discipline : Electronique
Spécialité : Communications Numériques
Laboratoire : IETR UMR 6164

Soutenance le 2 septembre 2015

Efficient baseband digital predistortion techniques for linearizing power amplifier by taking into account nonlinear memory effect

JURY

Président	**M. Yves LOUËT**, Professeur, Centrale Supélec, Rennes
Rapporteur	**Mme Geneviève BAUDOIN**, Professeur, ESIEE, Noisy le Grand
Rapporteur	**M. Claude DUVANAUD**, Maître de Conférences/HDR, IUT Angoulême, Université de Poitiers
Examinateur	**M. Dominique DALLET**, Professeur, IMS, Bordeaux
Directeur de Thèse	**M. Yide WANG**, Professeur, Ecole polytechnique de l'université de Nantes
Encadrants	**M. Bruno FEUVRIE**, Maître de Conférences, IUT Nantes
	Mme Anne-Sophie DESCAMPS, Maître de Conférences, IUT Nantes

Acknowledgments

It is time to draw a period to my PhD endeavor. During the three years, many people have helped me and supported me. I am grateful for their help, encouragement and support.

First, I would like to thank my supervisor Prof. Yide Wang and co-supervisors Bruno Feuvrie and Anne-Sophie Descamps, for offering me the opportunity to be a PhD student under their supervision. Their professional guidance and valuable advices are essential for the completion of this thesis. I also want to thank them for their help in my daily life. They are nice and patient. I feel extremely lucky to meet them.

I would also like to thank my friends and every colleague in the lab, for their help and tolerance. I had a good time for my PhD life in France.

Furthermore, I would like to thank China Scholarship Council (CSC) for the financial support.

Finally, I would express my appreciation to my parents, my wife and my little daughter for their selfless support.

Résumé de la thèse en français

Le chapitre 1 présente le contexte et l'évolution des systèmes de communication sans fil répondant à une demande toujours plus forte des acteurs et des contenus. Naturellement, les ressources disponibles pour satisfaire cette demande sont limitées. Afin d'exploiter au mieux ces ressources et pour augmenter l'efficacité spectrale, de nouvelles techniques ou standards de communication ont été développés comme le Wide Code Division Multiple Access (WCDMA), l'Orthogonal Frequency Division Multiplexing (OFDM) et le Long Term Evolution (LTE). Malheureusement, ces techniques ont pour effet d'augmenter la dynamique des signaux entrainant une plus grande sensibilité aux non-linéarités des éléments de la chaine de transmission.

Plusieurs éléments de cette chaîne présentent des non-linéarités, en particulier des saturations qui ont pour effet de déformer le signal à transmettre. La plus forte non linéarité est apportée par l'amplificateur de puissance. L'amplification, qui doit être constante quelque soit l'amplitude du signal d'entrée, diminue puis stagne lorsque la puissance moyenne du signal est trop importante. Ces deux effets génèrent des distorsions non-linéaires en amplitude et en phase sur le signal, notamment des remontées spectrales dans les canaux adjacents et des rotations et compressions de la constellation dans le domaine temporel dégradant la qualité de la communication et perturbant les communications dans les bandes de fréquence voisines. Pour éviter de générer ces déformations, il suffirait de diminuer la puissance d'émission du signal en surdimensionnant les équipements. Mais ceci aurait deux effets négatifs. D'une part, les systèmes seraient beaucoup plus volumineux, et d'autre part le rendement serait fortement dégradé ce qui est souvent une contrainte forte sur les systèmes portatifs. Pour optimiser ce rendement, il faut s'approcher de la zone de saturation de l'amplificateur, là où le rendement est maximum. De plus, les signaux transmis sont de plus en plus large bande entrainant un effet mémoire non linéaire qui vient s'ajouter aux distorsions précédentes. Ceci entraine un spectre asymétrique et un effet de nuage sur les points de la constellation. Plusieurs critères sont introduits pour évaluer ces non-linéarités : l'ACPR (qui évalue les remontées spectrales) et l'EVM (qui mesure les distorsion de

la constellation).

Pour obtenir un bon compromis entre le rendement et la linéarité, deux sorte de méthode sont exploitées. L'une est destinée à améliorer le rendement comme les méthodes Doherty ou d'enveloppe tracking, l'autre cherche à linéariser l'amplificateur (Feedback, Feedforward, LINC, EER, Pré-distorsion). Dans cette thèse, on étudiera spécifiquement les techniques de linéarisation et plus spécifiquement la pré-distorsion numérique en bande de base. Cette pré-distorsion consiste à déformer le signal avant de l'insérer dans le PA pour que celui-ci soit globalement linéairement amplifié quelque soit son amplitude. Le pré-distorteur est donc un élément placé devant le PA. Une première méthode (LUT) qui consiste à placer en pré-distorteur une table indexée sur l'amplitude du signal d'entrée pour inverser la caractéristique de l'amplificateur est simple à mettre en œuvre mais ses performances dépendent de la taille de la LUT. Une autre méthode basée sur des modèles mathématiques est souvent exploitée. Elle donne d'excellents résultats tout en prenant en compte les effets mémoires mais au prix d'une complexité accrue. Le pré-distorteur peut aussi être judicieusement implanté par un réseau de neurone qui est performant en termes de linéarisation, mais dont la phase d'apprentissage est longue. Dans les méthodes basées sur des modèles, il existe plusieurs modèles (Hammerstein, Wiener, MP, Volterra ...) et deux architectures (directe DLA et indirecte ILA). La première architecture identifie un modèle prédéfini (MP, ...) en utilisant un algorithme du type LMS pour reproduire la caractéristique directe du PA. Ce modèle est ensuite inversé dans la pré-distorsion. Dans l'architecture ILA, c'est la caractéristique inverse du PA qui est identifiée par un modèle prédéfini. Il suffit alors de placer ce modèle devant le PA pour réaliser la pré-distorsion.

Dans cette thèse, les deux architectures sont comparées. Nous considérons principalement la technique de la prédistorsion en bande de base, basée sur l'architecture DLA et le modèle MP. Cette méthode est simplifiée par l'utilisation d'une table (Look Up Table) pour bénéficier de la simplicité de cette dernière. Pour réduire la taille de la LUT, des méthodes d'interpolation linéaire et quadratique sont proposées. Enfin, une implémentation basée sur un réseau de neurones est réalisée pour accélérer le temps de prédistorsion.

Le chapitre 2 s'attache en particulier à caractériser l'amplificateur de puissance et à décrire les principales techniques de linéarisation existantes par une bibliographie. Après avoir présenté l'évolution des systèmes de communications radio mobile de première à la quatrième génération ainsi que les architectures d'émetteur pour mettre en évidence l'importance du PA, les caractéristiques principales de l'amplificateur de puissance sont détaillées : le gain, les caractéristiques AM/AM, AM/PM, le ren-

dement. Le compromis linéarité/rendement est expliqué et justifié au moyen de la caractéristique AM/AM et de la courbe de rendement.

Les harmoniques et l'intermodulation, conséquences de la non-linéarité, sont abordés au travers d'exemples classiques avec les signaux à un ton et à deux tons. Les harmoniques peuvent être supprimées par filtrage nécessitant des circuits supplémentaires dans la chaîne de transmission mais les produits d'intermodulation ne peuvent être filtrés facilement. Pour un signal modulé, la non linéarité produit une remonté spectrale, une rotation et compression de la constellation, puisque les amplitudes les plus fortes sont diminuées par l'effet de saturation progressif. Pour caractériser quantitativement ces non-linéarités, les points à 1dB et 3dB de compression, l'IP3 l'ACPR et l'EVM sont rappelés.

Pour étudier et analyser finement les amplificateurs de puissance, il faut définir des modèles comportementaux. Le modèle de Saleh, le modèle de Rapp, le modèle polynomial, le modèle de Volterra, le modèle polynomial généralisé à mémoire, le modèle polynomial à mémoire, le modèle d'Hammerstein et de Wiener sont tous des modèles classiquement exploités. Ils sont rappelés et détaillés.

Pour lutter contre les distorsions introduites par les effets non linéaires, différentes techniques sont développées. On retrouve les techniques dites d'amélioration du rendement et celles qui tendent à linéariser le PA pour obtenir une caractéristiques AM/AM parfaitement linéaire jusqu'à la saturation. Elles sont rappelées avec leurs avantages et inconvénients (Feedback, Feedforward, LINC, EER, Pré-distorsion).

Le chapitre 3 est consacré à la pré-distorsion numérique en bande de base (DPD). Dans un premier temps, les avantages et inconvénients d'une pré-distorsion sur les signaux RF, IF ou bande de base sont synthétisés. De part l'évolution des technologies numériques, le traitement numérique des signaux en bande de base est parfaitement adapté à la pré-distorsion. La modélisation du PA en bande de base, le traitement du signal pour la pré-distorsion sont deux étapes fondamentales du système. Idéalement, la caractéristique entrée/sortie de l'ensemble (amplificateur et predisorteur) obtenue doit être parfaitement linéaire jusqu'au point de saturation. Au delà, il n'y aura pas d'amélioration possible, l'amplificateur saturera.

Trois méthodes classiques de linéarisation sont détaillées. La méthode exploitant une table indexée par l'amplitude du signal entrant traite le signal en coordonnées polaires, cartésiennes ou exploite directement le gain complexe. Le principe est explicité par la formulation suivante

$$x(n) = u(n)g(r_u(n)) \tag{1}$$

dans laquelle $u(n)$ représente le signal d'entrée, $r_u(n)$ l'amplitude du signal d'entrée, $x(n)$ le signal pré-distordu et $g(\cdot)$ le gain indexé par l'amplitude du signal d'entrée. La taille de la table et le choix de l'indexation influencent les performances du pré-distorteur (linéarisation et temps de calcum). Les entrées de la table peuvent être équidistantes par rapport à l'amplitude du signal, à la puissance du signal ou bien au logarithme de la puissance. Il existe aussi des fonctions de compression utilisant la densité de probabilité du signal d'entrée qui optimise les performances. Ces méthodes ne peuvent pas prendre en compte l'effet mémoire non-linéaire. Une deuxième méthode exploitant les réseaux de neurones est également présentée. Le réseau de neurone est connu pour sa capacité à approximer une fonction non-linéaire avec un niveau d'erreur souhaité. Plusieurs structures de réseau de neurones sont communément déployées : MultiLayer Perceptron (MLP), le Radial Basis Function artificial Neural Network (RBFNN) et le Recurrent artificial Neural Network (RNN). La structure RNN convient bien a priori pour les systèmes non linéaires à effet mémoire mais l'étape d'apprentissage est complexe. La structure classique type RBFNN lui est préférée en ajoutant comme entrée les échantillons complexes à l'instant t et ceux des instants précédents pour prendre en compte cet effet mémoire. L'étape d'apprentissage demeure un problème en termes de temps de calcul. La pré-distorsion peut aussi être réalisée à l'aide d'un modèle analytique en lieu et place du pré-distorteur. Il existe des modèles sans mémoire comme le modèle de Saleh, le modèle de Rapp, le modèle polynomial mais également des modèles plus complexes prenant en compte ces effets mémoires : modèle d'Hammerstein, de Wiener, modèle polynomial à mémoire.

Il existe deux architectures exploitant la pré-distorsion. L'architecture dite directe dans laquelle le modèle défini représente le comportement de l'amplificateur et l'architecture dite indirecte dans laquelle le modèle identifié représente la caractéristique inverse du PA. Dans ce dernier cas, il suffit de placer le modèle en amont du PA pour réaliser la pré-distorsion tandis que dans le premier cas, une méthode doit être mise en œuvre pour inverser le modèle et générer le signal pré-distordu. Pour identifier efficacement le modèle, une écriture linéaire dans les paramètres est souhaitable, c'est le cas d'un certain nombre de modèles comme celui d'Hammerstein, du modèle polynomial à mémoire. L'exemple du modèle MP illustre l'écriture classique permettant l'indentification de ses paramètres en utilisant par exemple l'algorithme LS.

Les algorithmes d'identification LMS (Least Mean Squares) et RLS (Recursive Least Squares) sont également présentés. L'algorithme LMS consiste à minimiser l'erreur quadratique moyenne entre le signal de sortie du modèle et le signal de sortie mesuré. Le modèle défini par ses paramètres $\mathbf{w}(n)$ liant la sortie et l'entrée du PA peut

s'écrire :

$$y(n) = \mathbf{x}^T(n)\mathbf{w}(n) \tag{2}$$

dans laquelle $\mathbf{x}(n)$ représente l'entrée et $\mathbf{w}(n)$ les paramètres du modèle. Les paramètres $\mathbf{w}(n)$ sont remis à jours en utilisant l'algorithme du gradient

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \mu e(n)\mathbf{x}(n) \tag{3}$$

μ contrôle la convergence et la vitesse de convergence de l'algorithme qui est détaillé dans le cas du modèle MP. L'algorithme RLS basé sur la minimisation d'une fonction de coût (erreur) introduisant un facteur d'oubli est précisé dans le cas du modèle MP.

Il est important de pouvoir comparer les deux architectures ILA et DLA. Premièrement, leur stabilité est étudiée lors de la phase d'identification utilisant l'algorithme LS en analysant le conditionnement de la matrice qui doit être inversée. Ce conditionnement est défini par l'écart entre la plus petite et la plus grande valeur propre de la matrice à inverser. Dans l'architecture ILA, la matrice à inverser est constituée par les échantillons de sortie de l'amplificateur, le modèle cherchant à inverser sa caractéristique, l'entrée du modèle devient la sortie du PA. Dans l'architecture DLA, la matrice à inverser est constituée par les échantillons du signal d'entrée du PA (modèle directe). Une comparaison utilisant un modèle de Wiener simulant le PA attaqué par un signal 16QAM avec un débit symbole de 3.84MHz est réalisée. Le nombre caractéristique du conditionnement de la matrice est tracé en fonction de la puissance moyenne du signal d'entrée. Si le signal est entaché d'erreur, le conditionnement devient moins bon. C'est le cas pour la méthode ILA dont les mesures peuvent être bruitées. Ce nombre est aussi tracé en augmentant les ordres du modèles MP. Si l'ordre est trop important, la matrice devient mal conditionnée et l'identification devient très sensible aux imperfections des mesures. Un exemple de simulation est donné en simulant l'acquisition du signal par un convertisseur analogique-numérique à 8 bits et en ajoutant un bruit ayant un rapport signal sur bruit de 35dB. Les ACPR et EVM sont tracés. Il en ressort que l'architecture DLA est plus robuste aux imperfections du système par rapport l'architecture ILA. C'est la raison pour laquelle, elle est adoptée dans ce travail de thèse.

Dans le chapitre 4, les différentes méthodes proposées sont décrites et simulées en expliquant les degrés d'optimisation apportés en vue d'une implémentation temps réel. La technique s'appuie sur l'architecture DLA et optimise la méthode originale consistant à déterminer le signal pré-distordu à partir d'une recherche de racine dans un polynôme. Cette méthode donne d'excellents résultats mais nécessite un temps de calcul trop important, qui empêche son implémentation en temps réel. Pour résoudre

ce problème, l'utilisation d'une LUT pour remplacer ce processus de recherche de racine est tout d'abord proposée. Pour s'affranchir du problème lié à la taille de la LUT, les interpolations linéaire et quadratique sont exploitées. Enfin la pré-distorsion par un réseau de neurones est réalisée.

Le modèle d'amplificateur choisi pour simuler le PA est un modèle de Wiener avec un filtre numérique possédant trois retards. Les caractéristiques AM/AM et AM/PM montrent bien les effets non-linéaires avec mémoire de l'amplificateur de puissance. Après avoir défini le modèle MP pour modéliser le PA considérant que ce dernier offre un bon compromis entre la complexité du modèle et sa capacité à suivre les effet non linéaire avec mémoire du PA, le choix de l'ordre de la non-linéarité et de la profondeur de la mémoire est étudié en observant le NMSE pour diverses valeurs. Une non linéarité de 9 et une profondeur mémoire de 3 suffisent pour obtenir un NMSE minimal. Les signaux I/Q et les caractéristiques AM/AM et AM/PM sont tracés et comparés entre le modèle de PA simulé et le modèle MP identifié par un algorithme LS. La similarité des courbes montre bien l'excellente capacité du modèle MP à suivre le comportement du PA.

Après avoir rappelé le principe de la méthode de pré-distorsion [1] utilisant une architecture DLA associée à une recherche de racine polynomiale appelée MP/root-finding, la méthode utilisant une LUT [2] est complétée et optimisée en proposant une tabulation du module et de la phase du signal (ce qui n'est pas traité dans [2]). L'algorithme appelé MP/LUT réalisant la pré-distorsion est présenté. Le phénomène de pallier introduit naturellement par une table, génère une remontée du plancher du spectre du signal. Augmenter la taille de la LUT pour le diminuer accroît l'espace mémoire nécessaire. Une interpolation linéaire, algorithme appelé MP/LILUT, associée à cette LUT permet de garder une faible taille au prix d'un faible calcul supplémentaire sur la pré-distorsion [3]. Une interpolation quadratique, algorithme appelé MP/QILUT améliore encore les performances [4] au prix d'une complexité supérieure. Les trois méthodes sont comparées dans un premier temps en mesurant le NMSE du signal pré-distordu obtenu avec la méthode proposée et l'algorithme calculant la racine d'un polynôme. Pour le même NMSE, une taille de 2000 est nécessaire pour l'algorithme MP/LUT, alors qu'une taille de 20 pour MP/LILUT et 10 pour le MP/QILUT est suffisante. De même, le NMSE entre la sortie du PA linéarisé avec la méthode MP/root-finding et les trois autres méthodes proposées est tracé. On retrouve sensiblement les mêmes résultats. Les caractéristiques AM/AM et AM/PM sont alors tracées en superposition de celles obtenues sans linéarisation. On observe une caractéristique bien linéarisée et une annulation des effets mémoire. Les ACPR sont aussi comparés pour

une puissance moyenne d'entrée de -4dBm, permettant une répartition des symboles dans les zones linéaires et non linéaires. En considérant la même taille de table pour chaque méthode, on peut conclure que la méthode de l'interpolation linéaire est un bon compromis, qu'elle améliore considérablement la méthode MP/LUT et que l'effet apporté par la méthode QILUT n'est pas considérable compte tenu de son coût de calcul supérieur. Les mêmes résultats sur l'EVM sont observés. Pour compléter l'étude, le temps de calcul du signal pré-distordu par les différentes méthodes fonction de la taille de la table est représenté sur une même courbe.

Une quatrième méthode exploitant le réseau de neurone (FFNN) est présentée. La technique propose de réaliser la séquence d'apprentissage du réseau de neurones pour modéliser le pré-distorteur avec le pré-distorteur MP/root-finding déjà réalisé. On dispose ainsi des entrées sorties d'un pré-distorteur. Pour obtenir l'effet mémoire escompté, les échantillons complexes avec une certaine profondeur mémoire sont placés à son entrée et le signal pré-distordu est placé à sa sortie. Pour différentes valeurs de profondeur mémoire, pour différentes structures du réseau de neurones, les spectres, les EVM, les ACPR sont comparés avec la méthode MP/root-finding. Il est évident que le FFNN parvient parfaitement à linéariser l'amplificateur de puissance. Le temps de calcul du signal pré-distordu est faible. La contrainte majeure de cette méthode réside dans sa phase d'apprentissage qui est relativement longue.

Dans le chapitre 5, l'ensemble des méthodes présentées dans le chapitre précédent sont testées expérimentalement.

Le système expérimental est constitué d'un PC, d'un générateur de signaux vectoriels (VSG, R&S SMU 200A), d'un oscilloscope numérique (DO, Lecroy, 4 voies, Wave master 8600, bande 6GHz, 20 GS/sec) et d'un analyseur de spectre (SA, Agilent E4440A). Le PC réalise les algorithmes pour calculer le signal pré-distordu et plus globalement pilote les différents appareils sous le logiciel Matlab (Instrument Control Toolbox). Après avoir généré les signaux bande de base, le PC démodule les signaux acquis par l'analyseur de spectre, calcule le signal pré-distordu, le renvoie au générateur pour tester les performances des méthodes. L'ensemble de ces fonctionnalités est géré par une interface graphique programmé sous Matlab.

L'amplificateur utilisé est un ZFL2500 de bande 500-2500 MHz avec un gain d'environ 31 dB. La puissance moyenne de sortie au point de compression à 1 dB est de 17.5dBm et la fréquence porteuse choisie est de 1.8GHz. Le signal initial émis modulé est un signal 16QAM de largeur de bande 3.84MHz.

Pour prendre en compte l'effet de la conversion analogique numérique sur le signal, un premier test de simulation est réalisé en quantifiant les signaux d'entrée et de sortie

en considérant un nombre de bits allant de 3 à 16. Les résultats sur l'ACPR montrent qu'un convertisseur à 10 bits suffirait. Lors de la pré-distorsion, la largeur de bande du signal pré-distordu augmente. Aussi, le générateur de fonction doit pouvoir supporter cet accroissement de bande. Pour évaluer ce phénomène, lors d'une simulation, un filtre pas bas est placé après le signal pré-distordu et la performance de la linéarisation en ACPR est calculée pour différente fréquence de coupure émulant ainsi la limitation de bande du générateur de fonction. On s'aperçoit qu'une bande cinq fois supérieure à celle du signal original est nécessaire pour obtenir une bonne performance. Enfin, une comparaison des méthodes est effectuée en analysant les spectres, l'ACPR et l'EVM des signaux linéarisés. Le temps de calcul des différents pré-distorteurs confirme les résultats de simulation. Il est à noter que seul est pris en compte le temps de calcul de la pré-distorsion et ni le temps de création des LUT ni celui de la phase d'apprentissage du réseau de neurones n'est considéré.

Le dernier chapitre synthétise les conclusions et perspectives. Après un rappel sur les communications numériques mobiles, le problème spécifique de la non-linéarité sur le PA, les caractéristiques principales du PA ainsi que les différentes méthodes destinées à linéariser le PA sont énoncées. La pré-distorsion numérique en bande de base est en particulier étudiée d'un point de vue bibliographique. Quatre méthodes originales de pré-distorsion basées sur l'architecture DLA sont proposées. Elles sont explicitées analytiquement. Des tests en simulation confortent les résultats escomptés. Enfin, une validation expérimentale confirme le bien fondé des méthodes développées.

Trois perspectives à ce travail sont énumérées : le modèle GMP un peu plus complet que le modèle MP doit pouvoir être testé sans que cela ne change les principes des méthodes proposées, une étude plus approfondie sur la constitution des LUT en termes d'espacement, et enfin l'utilisation conjointe des techniques de réduction du PAPR et de la pré-distorsion numérique en bande de base.

Contents

1

Introduction

1.1 Introduction

In modern wireless telecommunications systems, many different types of data are transmitted such as speech, email, picture, video, web page. As the demand of people is increasing for wireless applications, more and more data need to be transmitted through wireless network. However, the spectrum resources are limited. In order to improve the spectral efficiency, advanced wireless communication techniques such as Wide Code Division Multiple Access (WCDMA), Orthogonal Frequency Division Multiplexing (OFDM) and Long Term Evolution (LTE) standard have been developed. Generally, the involved signals are characterized by wide bandwidth and non-constant envelopes leading to a high Peak to Average Power Ratio (PAPR). The signals characterized by high PAPR are very sensitive to the nonlinear characteristic of transmitter.

Power Amplifier (PA) is a crucial component in transmitter. Its inherent nonlinearity directly impacts the linearity of the whole transmitter. In most cases, the power backoff can make PA operate in linear mode, but it results in a significantly low power efficiency. To have a high power efficiency, PA usually works close to its saturation region, where the nonlinear distortion is severe. The amplitudes of signals characterized by a high PAPR may exceed the saturation level of PA. It results in nonlinear distortions, such as the amplitude and phase, the adjacent channel interference and thus degradation of bit error rates. Additionally, as the bandwidth of signal continues to

increase, the memory effect is obvious and can no longer be ignored. The frequency-dependent distortion caused by the memory effect also needs to be taken into account. These distortions and interferences lead to seriously degradation of the performance of the transmission system. To correctly transmit these signals and to have a high power efficiency, one solution is to compensate the nonlinearity of PA.

The nonlinearity of PA mainly contains two aspects: static distortions and dynamic distortions. Static distortions exist in the PA inherently. It can be observed on Amplitude-to-Amplitude (AM/AM) and Amplitude-to-Phase (AM/PM) characteristics of PA. It will result in the spectral regrowth in frequency domain, the rotation and the compression of constellation in time domain. Dynamic distortions are caused by the electrical and electrothermal effects, also called the memory effects. It means that the response of PA depends not only on the current input sample, but also on the previous input samples. The memory effects will result in the spectrum asymmetry in frequency domain, and the dispersion of constellation in time domain. Generally, in the narrow band system, the memory effects can be neglected. To evaluate the nonlinearity of PA, there are several criteria such as Adjacent Channel Power Ratio (ACPR), and Error Vector Magnitude (EVM).

To make a good trade-off between the power efficiency and linearity, there are two research directions. The first one is to improve the power efficiency by advanced PA techniques, such as Doherty amplifiers [5] and envelope-tracking techniques [6]. The second one is to extend the linear range of PA by some linearization techniques. Many existing techniques are reported for the linearization of PA [7], such as Power backoff, Feedback, Feedforward, LInear amplification with Nonlinear Components (LINC), Envelope Elimination and Restoration (EER), and Predistortion.

Among these linearization techniques, predistortion is the most efficient and promising one. Depending on signal types, predistortion methods can be classified into digital predistortion and analog predistortion [8, 9]. According to the position of the predistorter in whole transmitter, predistortion methods can be classified into Radio Frequency (RF) predistortion, Intermediate Frequency (IF) predistortion and BaseBand (BB) predistortion. In this thesis, we mainly focus on the baseband digital predistortion. According to the implementation way, the baseband digital predistortion methods can be classified into three groups: LookUp Table (LUT) based methods [10, 11], model (or polynomial) based methods [12, 13] and neural network based methods [14, 15]. They have their own advantages and disadvantages. LUT based predistortion has low complexity, and is easy to implement in hardware. But it needs to overcome problems caused by the quantification error. Model based predistortion can

obtain more accuracy than LUT based predistortion, but it requires a large number of mathematical operations. Neural network based predistortion has been developed in recent years. Theoretically, artificial neural network can learn and imitate the behavior of any nonlinear system. Therefore, the neural network can be used to imitate the nonlinear behavior of predistorter. However, the implementation of neural network internal structure in hardware is difficult and the training process is complicated and time-consuming.

Digital PreDistortion (DPD) is considered as a key enabling technique for future software-defined and cognitive radio transmitters due to its simplicity and reconfigurability. In DPD, a block called predistorter is inserted at front of the PA so that the whole system has a linear behavior. The predistorter can be seen as the inverse model of PA. Generally, two steps are important for realizing DPD. One is to model the PA. The other is the identification algorithm of the predistorter.

The PA models can be classified into memoryless models and models with memory. In modern wireless communication systems, due to the increasing bandwidth of signal, PAs exhibit strong memory effects. The PA models with memory are more focused on by researchers. The well-known Volterra series model [16] is usually used to model the PA with memory effects. The number of parameters in the conventional Volterra series model increases drastically with the nonlinearity order and the memory depth. To alleviate the complexity burden, some simplified versions of Volterra series model have been proposed, such as the generalized memory polynomial model [13], Hammerstein model [17], Wiener model [18], memory polynomial model [12] and other variants[19]. Among all models of this type, the memory polynomial is mostly used because it has good modeling performance and low complexity.

There are two different approaches to determine the predistorter: Indirect Learning Architecture (ILA) and Direct Learning Architecture (DLA) [20]. In ILA, a post-distorter is assumed. The input (PA output) and output (PA input) of model which constitute the post-distorter are known. The post-distorter can be identified by using either Least Squares (LS) algorithm, Least Mean Squares (LMS) algorithm or Recursive Least Squares (RLS) algorithm [21]. Then the post-distorter is copied as the predistorter. In DLA, a PA model is first chosen and then the identification is performed to fit with PA behavior. The predistorter is obtained based on a reference error between the output of model and the output of PA. Some iterative identification algorithms have been reported, such as the Nonlinear Filtered-x LMS (NFLMS) algorithm [22] and the Nonlinear Filtered-x RLS (NFRLS) algorithm [23].

DPD is one of the most fundamental building blocks in modern wireless commu-

nication systems. It is usually combined with Crest Factor Reduction (CFR) technique in practical applications [24]. CFR technique is also called PAPR reduction technique. The cascade of CFR and DPD allows to increase significantly the power efficiency of PA. For instance, a typical Class AB LDMOS PA with WCDMA waveforms may have approximately 8-15% efficiency. A DPD LogiCORE Intellectual Property (IP) has been proposed by Xilinx [25]. Its tests have be shown that the PA efficiency can grow to as much as 30-40% with CFR and DPD turned on.

In this thesis, we study DPD techniques for the linearization of PAs with memory effects. The main contributions of this thesis are summarized as follows.

1. The two different DPD learning architectures (ILA and DLA) are compared. Although several authors already reported the comparison of ILA and DLA [20, 26], to the best of our knowledge the comparison related to the condition number of matrix has not been investigated. In this thesis, we analyze the difference of ILA and DLA in terms of the condition number of the matrix to be inverted in LS identification algorithm, the modeling performance and the linearization performance of DPD.

2. A DPD method combining LUT and MP is proposed based on DLA. In original DPD method based on MP, the amplitude and phase of predistorted signal are both obtained by complicated mathematical operations. It is time-consuming. In this thesis, the processes of calculating the amplitude and phase of predistorted signal are both simplified by using LUT. The calculating time of DPD step is significantly decreased.

3. In the proposed DPD method mentioned above, although the calculating time of predistorted signal is reduced, it requires a sufficiently large table size to obtain a good linearization performance. The linear-interpolation and quadratic-interpolation techniques are introduced in LUT. They allows a trade-off between the time, table size and linearization performance.

4. We also implement DPD by using neural network. A DPD method combining a feed-forward neural network and MP is proposed. Firstly, a set of predistorted signals are obtained by using MP DPD method based on DLA. Secondly, such predistorted signals are used to train the neural network. Finally, the trained neural network is adopted as the predistorter. The proposed DPD method based on neural network has a good linearization performance. The drawback is that the training process is time-consuming.

1.2 Thesis organization

In chapter 2, the background of our study is presented briefly. We begin by the evolution of wireless communication systems. Then the transceiver, transmitter, power amplifier are introduced. After that, power amplifier nonlinearity, power amplifier behavior modeling and some criteria for evaluating the nonlinearity are described. Finally, several classical linearization techniques are presented.

Chapter 3 provides an overview of the baseband digital predistortion. It includes the principle of digital predistortion, the classification, the implementation method, the learning architecture and the identification algorithm. In addition, it emphasizes the comparison of indirect learning architecture and direct learning architecture.

Chapter 4 presents the proposed DPD methods in detail. It mainly includes two kinds of solutions. One is to combine memory polynomial and LUT. The other is to combine memory polynomial and neural network. Simulation results are given to show the performance of the proposed methods.

Chapter 5 discusses the experimental setup and results. Some parameters which influence the linearization performance in practical applications are studied, such as the number of bits of ADC and bandwidth of instruments.

Chapter 6 concludes the dissertation and provides some future possible research directions.

2

Power amplifier and linearization techniques

This chapter serves an introduction to the basic knowledge of the power amplifier and its linearization. Section 2.1 gives an evolution of wireless telecommunication systems where some advanced techniques, such as LTE, are introduced briefly. The transceiver where PA is located is discussed in section 2.2, especially, the transmitter is described. Power amplifier, key component in this thesis, is introduced in section 2.3. PA nonlinear characteristics, some criteria for measuring the nonlinearity and typical PA behavior models are presented. Finally, a survey of different linearization techniques is given in section 2.4. The techniques, such as Backoff, Feedforward, Feed-back, LINC, EER and Predistortion are presented briefly.

2.1 Wireless telecommunications systems

In modern wireless telecommunications systems, lots of data must be transmitted such as email, speech, picture, video, web page. Wireless networks connect people to each other and give them an important source of knowledge. Wireless communication technology has been continuously improved and evolved from 1G (First Generation) to 4G (Fourth Generation). Nowadays, 4G technology has been commercialized. 5G (Fifth Generation) technology will soon be realized and will be available to general

public.

MTS (Mobile Telephone Service) published in 1946 by Bell Laboratories is the first wireless mobile system which is characterized by half duplex and requires operator to establish call [27]. In 1969, IMTS (Improved Mobile Telephone Service) appeared with duplex and direct dial using DTMF (Dual Tone Multi-Frequency) keypad. MTS and IMTS systems was called 0G (Zero Generation) wireless communication system.

In the 1980s, it begins the era of 1G analog cellular network. Several typical 1G systems were made such as NMT (Nordic Mobile Telephone), AMPS (Advanced Mobile Phone System) and TACS (Total Access Communication Systems) [28]. These 1G analog radio transmission systems all based on FDMA (Frequency Division Multiple Access) techniques provided only voice services. The biggest drawback is the limitation of channel capacity restricting the data rate. Moreover, it is difficult to implement a communication between countries because different countries use different communication protocols [28].

In the late 1980s and the early 1990s, some techniques based on TDMA (Time Division Multiple Access) and CDMA (Code Division Multiple Access) standards were proposed, which were called 2G (Second Generation) wireless communication systems, such as GSM (Global System of Mobile communication), TDMA IS-54 (Interim Standard-54), IS-95, and so on. 2G system is digital cellular network, which transmits digital data stream. It provides voice and text messaging services. Compared with 1G system, 2G system can accommodate more users in the radio spectrum through either time or code multiplexing [29] and also enable to implement a communication between countries.

Due to the increasing requirement of data types and of transmission speed, the wireless communication systems are constantly evolving. GPRS (General Packet Radio Service) system has been proposed, whose data rate can be up to 171.2 kb/s [27]. Afterwards, with EDGE system, the data rate reaches 384 kb/s. Now 3G (Third Generation) system has been widely used in many wireless application products. 3G systems include WCDMA, TD-SCDMA (Time Division-Synchronous Code Division Multiple Access), and CDMA2000. In 3G system, the multimedia data, such as picture, video, web data, can be transmitted and the data rate can be up to 2 Mb/s [30].

After 3G technique, HSDPA (High Speed Downlink Packet Access) technique has been proposed. It adopts QPSK (Quadrature Phase Shift Keying) and QAM (Quadrature Amplitude Modulation) modulations, whose downlink speed can be up to 14.4 Mb/s [27]. But it does not yet belong to 4G system. In 2008, ITU-R (International Telecommunication Union-Radiocommunication) specifies a set of requirements for

4G standard. In 4G system, the transmission speed can be up to 1 Gb/s under the static environment and it can still achieve 100 Mb/s under the fast-moving environment [30]. 4G techniques such as LTE, LTE Advanced and WiMAX (Worldwide Interoperability for Microwave Access) techniques have been proposed successively, which all adopt OFDM modulation thanks to its high spectral efficiency.

With the diversification of mobile applications, 4G system is not powerful enough to meet people's requirements in wireless communication. Therefore, the emergence of advanced communication techniques comes up with 5G system. Currently, 5G system is not a term present in official document made public by telecommunication companies or standardization bodies. But in the near future, 5G system will be realized, whose transmission speed is higher than 1 Gb/s [31]. 5G system will bring almost perfect real world wireless called WWWW (World Wild Wireless Web). Thanks to these wireless communication techniques, the people can exchange all types of information via immediate way. Table 2.1 summarizes the evolution of the wireless communication techniques [27, 28, 29, 30, 31, 32].

2.2 Transceiver

In wireless communication systems, information is transmitted and received by radio waves. The block realizing the functions of the transmitting and receiving is called as transceiver (TRx) [33]. The transceiver is mainly composed of a transmitter (Tx) and a receiver (Rx). The transmitter is especially significant, because sending data without error is the premise to receive the correct data. The block diagram of the transceiver is shown in Fig. 2.1. The functions of modulation and up-conversion are realized in the transmitter. Correspondingly, the functions of demodulation and down-conversion are realized in the receiver.

Several transceiver architectures widely used are: full-duplex, half-duplex and simplex modes [34]. In full-duplex mode, the transmitter and receiver operate simultaneously in different time slots. It allows communication in both directions simultaneously. In half-duplex mode, the transmitter and receiver do not operate simultaneously. The half-duplex mode allows communication in both directions but only one direction at a time. The simplex mode only transmits information in one direction. Information flows only from the transmitter site to the receiver site. This operating model is usually used in radio broadcasting system.

In the transmitter, the baseband signals are firstly handled by DSP (Digital Signal Processing) in most cases, to perform the source and channel coding, filtering, symbol

Generation	Technique standard	Year	Description
0G	MTS	1946	First mobile telephone service, half duplex
	IMTS	1969	full duplex, up to 13 channels, 60-100 km radius, direct dial using DTMF (dual tone multi-frequency) keypad
1G	NMT	1981	Nordic mobile telephone, 12.5 kHz channel, 450 MHz, 900 MHz, FM modulation
	AMPS	1983	30 kHz channel, 825-895MHz, FM modulation
	TACS	1985	25 kHz channel, widely used up to 1990s, FM modulation
	DataTac	1990	Point-to-point wireless data communications standard, 25 kHz channels, max bandwidth 19.2 kb/s
2G	GSM	1991	TDMA, constant envelope, 200 kHz channel, GMSK modulation
	TDMA IS-54	1991	Digital AMPS, 30 kHz channel, pi/4DQPSK modulation
	PDC	1992	25 kHz channel
	IS-95	1995	spread spectrum, CDMA, 1.25 MHz channel, QPSK modulation
2.5G	GPRS	2000	200 kHz channel, max bandwidth 171.2 kb/s
2.75G	CDMA2000	2000	CDMA, double data rate, 1.25 MHz channel
	EDGE	2003	8PSK, TDMA, 200 kHz channel, max bandwidth 384 kb/s
3G	WCDMA	2004	CDMA, 3.84 MHz channel
	TD-SCDMA	2006	Uses the same band for transmit and receive, base stations and mobiles use different time slots to communicate, 1.6 MHz channel
3.5G	HSDPA	2006	download speeds up to 14.4 Mb/s
4G	LTE	2009	MIMO, OFDM, 20 MHz bandwidth, download speeds up to 100 Mb/s
	WiMAX	2010	MIMO, OFDM, download speeds up to 128 Mb/s
	LTE Advanced	2011	the support of 20 MHz bandwidths, higher cell-average and cell-edge spectral efficiencies
5G	Not published	-	WWWW, speed is higher than 1 Gb/s

Table 2.1: Evolution of wireless communication techniques

mapping, shaping and interleaving. The processed signal is then modulated by modulation techniques, such as PSK or QAM. Then the modulated signal is up-converted

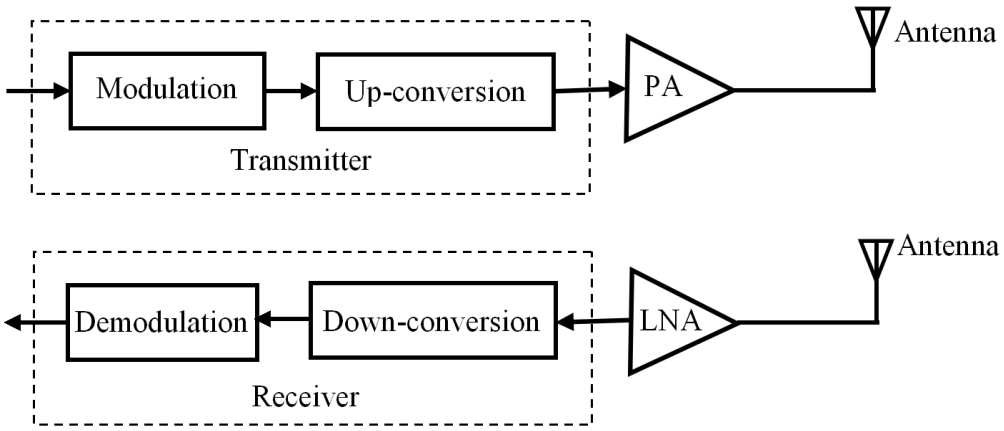


Figure 2.1: Transceiver

with a carrier frequency. At this time, the transmitted signal is still with low power. Finally, the signal is amplified by the PA and transmitted by the antenna. During this process, the nonlinearity of transmission system is unavoidably introduced into the signals. The important contribution of the nonlinearity is mostly introduced by the power amplifier.

In the following, the architectures of the transmitters are briefly introduced.

The architecture of the two-step conversion transmitter is shown in Fig 2.2. The signals are modulated at a fixed frequency by the IQ (In-phase and Quadrature) modulator. The BPF1 (Band-Pass Filter) is used to reject all undesired signals produced by the IQ modulator. The mixer shifts the spectra of signal to the final transmit frequency f_T. The BPF2 is used to reject all unwanted signals produced by the upmixing. This architecture has high-level performance and decreases the constraints of filters. But it is less suitable for integration because the complexity and physical size are increased [34].

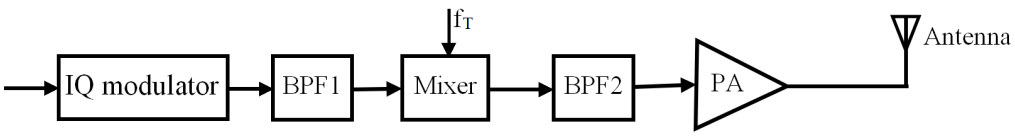


Figure 2.2: Two-step conversion transmitter

The architecture of the direct launch transmitter is shown in Fig 2.3. The baseband

IQ modulated signals are upconverted directly to the final transmitter frequency f_T by the LO (Local Oscillator). Compared with the two-step conversion transmitter, the direct launch transmitter is greatly simplified. It is better for the SoC (System On Chip) integration. The drawback is the PA and LO operate at the same frequency. This fact leads to a major source of interference and degrades the total transmitter performance [34]. The direct launch transmitter is suitable for mid-level performance, low-cost and low-power broadband applications, such as Bluetooth and WiFi.

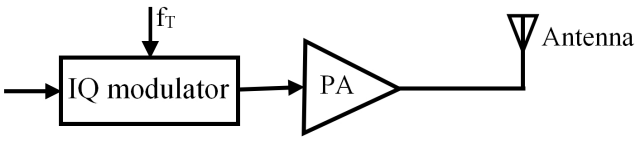


Figure 2.3: Direct launch transmitter

The architecture of FM transmitter is shown in Fig 2.4. The RF signals are directly generated by modulating the VCO (Voltage-Controlled Oscillator) frequency with the baseband signal. The main feature of FM transmitter is constant-envelope digital modulations which can be efficiently implemented. It is widely used in military and public safety applications.

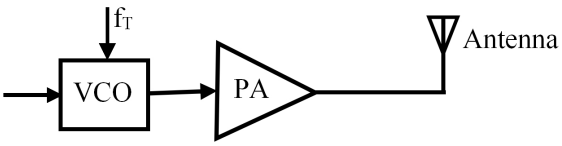


Figure 2.4: FM transmitter

In the receiver, the signal processing is the inverse operation with respect to the transmitter side. The transmitted signal is firstly received by the antenna. Then the received signal passes a LNA (Low-Noise Amplifier). Finally, the signal is down-converted and demodulated to the original signal. The common receiver architectures include superheterodyne receiver, direct conversion receiver and low IF receiver. A thorough discussion about the structure of receiver is beyond the scope of this thesis. This thesis focuses on the power amplifier linearization at the transmitter side.

2.3 Power amplifier overview

In RF transmitter, PA is a key component, which amplifies low power signal for long distance transmission. In this section, the power amplifier is introduced.

2.3.1 Power amplifier characteristics

There are two basic types of PAs: TWTA (Traveling-Wave Tube Amplifier) and SSPA (Solid-State Power Amplifier) [35]. TWTA is traditionally used at higher microwave frequencies, while SSPA is used at lower frequency bands. The former provides larger output power than the latter. But the life of TWTA is limited because of the heat generated by high voltage. SSPA is less weight, more reliable than TWTA. TWTA can be modeled by Saleh model [36]. SSPA can be modeled by Rapp model [37].

On the other hand, according to the linearity and efficiency of PA, there are many different classes of PAs, such as Class A, B, AB, C, D, and so on. Class A, B and AB PA are linear, but they are not perfectly linear. Commonly, except for class A, B and AB, other classes of PAs are not commonly used for applications requiring high linearity [33].

In the following, the parameters which describe the characteristics of PA are listed and commented, such as PA gain, AM/AM, AM/PM, power efficiency.

PA gain

Gain characteristic of PA denotes the ratio between the output power and input power. Normally, it is presented in dB unit, defined by

$$G = 10log\frac{P_{out}}{P_{in}} \tag{2.1}$$

where P_{out} and P_{in} are the output and input powers of PA, respectively. The gain is linear when the input power level is low, while it is compressed when the input power is high.

AM/AM and AM/PM characteristics

The baseband input/output signals of PA are complex (amplitude and phase). Normally, the amplitude and phase need to be analyzed separately. AM/AM characteristic denotes the relationship between the amplitude of output signal and the amplitude

of input signal. AM/AM characteristic is linear when the input power level is much smaller than the saturation point. As the input power increases, the nonlinearity of AM/AM characteristic becomes more and more obvious. Until the output power becomes saturated. In other words, the gain is compressed along with the increase of input power. AM/PM characteristic denotes the relationship between the phase shift of output signal and the amplitude of input signal. In small signal region, AM/PM characteristic is constant. As the input power increases, the shift of phase is more and more serious. An example of AM/AM and AM/PM of Saleh memoryless model is shown in Fig 2.5 and Fig 2.6. AM/AM and AM/PM characteristics are an important criteria for analyzing the linearity of power amplifier.

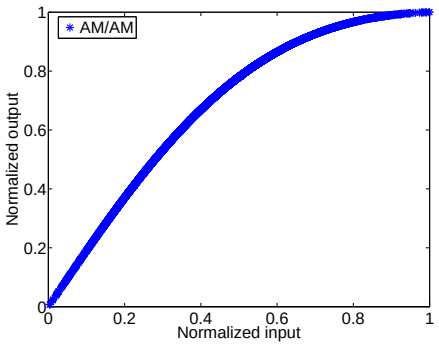


Figure 2.5: AM/AM characteristic of a memoryless PA

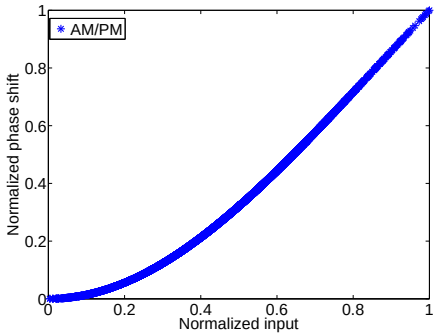


Figure 2.6: AM/PM characteristic of a memoryless PA

Power efficiency

Power efficiency of PA is most concerned by researchers in modern wireless communication system. It is because that it directly affects the power efficiency of whole transmitter. Commonly, the power efficiency of PA can be estimated by Power Added Efficiency (PAE) [38].

PAE, which involves PA input signal power, is expressed as

$$PAE = \frac{P_{out} - P_{in}}{P_{DC}} = \eta \left(1 - \frac{1}{g}\right) \tag{2.2}$$

where P_{in} and P_{out} denote the input and output powers of PA, respectively, P_{DC} denotes supplied DC (Direct Current) power, and $g = \frac{P_{out}}{P_{in}}$ is linear gain.

The primary aim of PA is to produce an amplification that has a good efficiency and low distortion. However, in practical design, there is always a trade-off between the

power efficiency and distortion. It is explained by the fact that there is low efficiency in linear region of PA, while a high efficiency is obtained in nonlinear region [32]. The relationship between the linearity and the power efficiency is shown in Fig. 2.7. The solid curve is the AM/AM characteristic, and the dashed one refers to the power efficiency. In Fig. 2.7, there are three regions: region 1 is the linear region, region 2 is the nonlinear region, and region 3 is the saturation region. In region 1, there is a good linearity for avoiding the distortion, but the power efficiency is very low. For improving the power efficiency, PA can operate in region 2, where there is a high power efficiency. But the nonlinearity is serious. The purpose of power amplifier linearization is to make the power amplifier operate at nonlinear region, and at the same time keep high power efficiency.

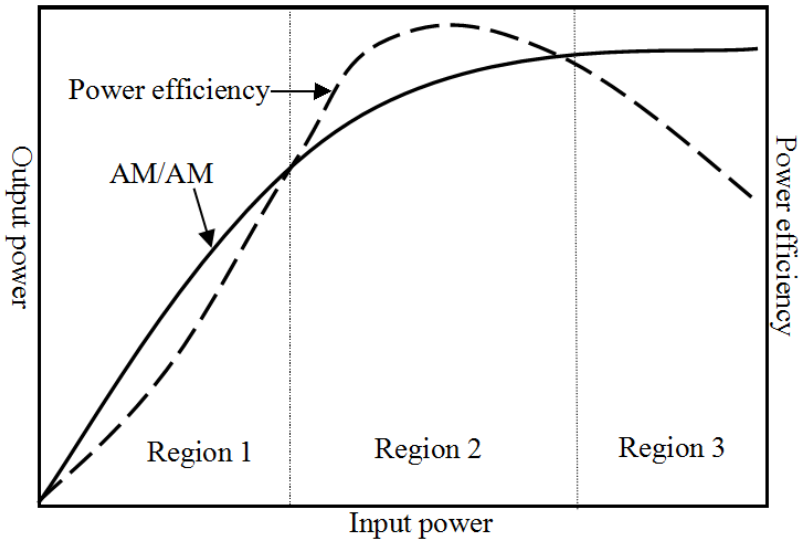


Figure 2.7: Power efficiency and AM/AM characteristic of PA

2.3.2 Power amplifier nonlinearity

In section 2.3.1, the main characteristics of PA are introduced, such as PA gain, AM/AM and AM/PM. Obviously, it can be found that the signal can not be amplified linearly when the input signal power exceeds a certain level. This is only one aspect of the nonlinearity of PA.

For an ideal linear PA, its input $r_{in}(t)$ and output $r_{out}(t)$ satisfy a linear relationship,

which can be expressed as

$$r_{out}(t) = g r_{in}(t) \tag{2.3}$$

where g is the ideal gain, $r_{out}(t)$ and $r_{in}(t)$ are the signal amplitudes of the output and input of PA, respectively. An ideal linear PA enables that all signals are increased in magnitude by the same factor; and there is a fixed phase shift between the input and output. Moreover, it does not exist memory effects, namely, the output of PA only depends on the instantaneous input sample, not any previous samples. However, perfectly linear PA dose not exist. We can only compensate the nonlinearity of PA.

The nonlinear behavior of PA includes two major components: the static nonlinearity and the memory effects [39]. The static nonlinearity, also known as the memoryless nonlinearity, corresponds to the distortion generated by PA when memory effects are not considered. They result in spectral regrowth in frequency domain as shown in Fig. 2.8, and rotation and compression of the constellation in time domain as shown in Fig. 2.9(b). The memory effects are also called the dynamic distortion, which are caused by electrical and electrothermal effects [40]. The output of PA depends not only on the current input sample, but also on the previous input samples. The memory effects result in the spectrum asymmetry, and the dispersion of constellation in time domain as shown in Fig. 2.9(c). Generally, the contribution of the static nonlinearity to the nonlinear behavior of PA is more stronger than that of the memory effects. However, both of them are equally important and need to be compensated in order to improve the linearity of the system.

In the following, the static nonlinearity is analyzed by two specific examples, where the memory effects are not considered temporarily. The first one is that a single-tone signal is taken as the input of PA. The single-tone signal is expressed as

$$V_{in}(t) = a cos(2\pi f t) \tag{2.4}$$

where a is the amplitude and f the frequency of the input signal.

For a memoryless nonlinear amplifier, the relationship of input and output signals can be expressed by Taylor polynomial as [33]

$$V_{out}(t) = g_1 V_{in}(t) + g_2 V_{in}^2(t) + g_3 V_{in}^3(t) + ... + g_n V_{in}^n(t) \tag{2.5}$$

where $g_1, g_2, ..., g_n$ are the coefficients of Taylor polynomial. When (2.4) is substituted into (2.5), we obtain PA's output responses, which contain the original signal and some harmonics. If only the first three terms of Taylor polynomial are considered, and

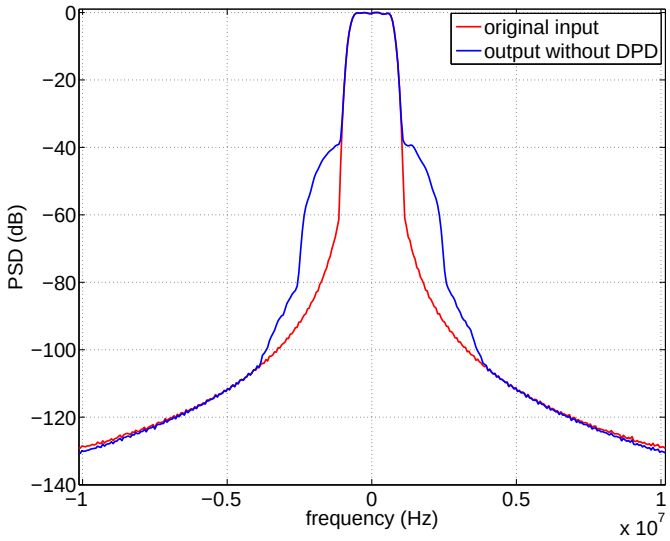


Figure 2.8: Spectral regrowth

denotes by $\theta = 2\pi ft$, (2.5) can be rewritten as

$$V_{out}(t) = \frac{g_2 a^2}{2} + \{g_1 a + \frac{3g_3 a^3}{4}\}cos(\theta) + \frac{g_2 a^2}{2}cos(2\theta) + \frac{g_3 a^3}{4}cos(3\theta) \tag{2.6}$$

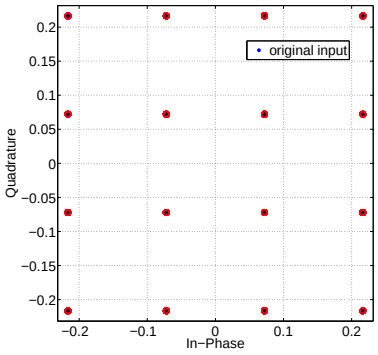
It can be seen that there are DC term, the fundamental term, the second-order harmonic term and the third-order harmonic term in the output signal. They are summarized in Table. 2.2. Actually, only the fundamental term is useful, other harmonic terms are useless. It assumes only the fundamental term $\{g_1 a + \frac{3g_3 a^3}{4}\}cos(\theta)$ is extracted. It can be seen that the amplitude of the output at the fundamental term is $g_1 a + \frac{3g_3 a^3}{4}$. As mentioned before, a is the amplitude of the input signal. When a is very small, it can assume that

$$\frac{3g_3 a^3}{4} \approx 0. \tag{2.7}$$

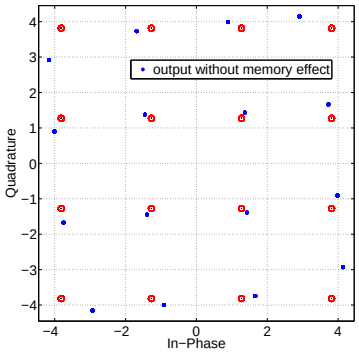
Therefore, the gain of the PA is approximated to g_1. As the value of a increases, $\frac{3g_3 a^3}{4}$ can no longer be ignored. The actual gain becomes

$$g_{actual} = g_1 + \frac{3g_3 a^2}{4g_1}. \tag{2.8}$$

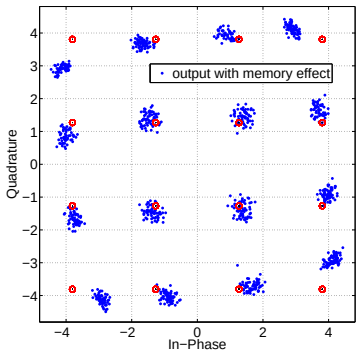
where if $\frac{3g_3 a^2}{4g_1} < 0$, the gain decreases as a increases; if $\frac{3g_3 a^2}{4g_1} > 0$, the gain increases as a increases. In most cases, the gain of PA decreases as the input power increases.



(a) Constellation of input



(b) Constellation of output without memory effect



(c) Constellation of output with memory effect

Figure 2.9: Distortions of constellation of output

This phenomenon is called gain compression.

Except for the fundamental term, the harmonic terms are redundant. They result in the spectral regrowth in frequency domain. The spectral regrowth can lead to the interference between main channel and other channels. Theoretically, these useless terms (second-order, third-order and other harmonic terms) can be filtered out by some filters.

However, for multi-tone input signal, some intermodulation distortions (IMD) close to the fundamental signal are generated. It is difficult to filter these IMD products out. In the second example, it assumes that a two-tone signal is considered as the input. To assume the input signal as

$$V_{in}(t) = acos(2\pi f_1 t) + acos(2\pi f_2 t) \tag{2.9}$$

where a is the amplitude of the testing signal, f_1 and f_2 are the frequency components of the two-tone signal. It assumes $2\pi f_1 t = \theta_1$ and $2\pi f_2 t = \theta_2$, and still only the

DC term	$\frac{g_2 a^2}{2}$
Fundamental term	$\{g_1 a + \frac{3g_3 a^3}{4}\} cos(\theta)$
Second harmonic	$\frac{g_2 a^2}{2} cos(2\theta)$
Third harmonic	$\frac{g_3 a^3}{4} cos(3\theta)$

Table 2.2: Output response of single-tone signal

first three terms of the Taylor polynomial are considered. The output response can be written as

$$\begin{aligned} V_{out}(t) = {} & g_2 a^2 + (g_1 a + \frac{9g_3 a^3}{4}) cos(\theta_1) + (g_1 a + \frac{9g_3 a^3}{4}) cos(\theta_2) + ... \\ & + \frac{g_2 a^2}{2} cos(2\theta_1) + \frac{g_2 a^2}{2} cos(2\theta_2) + g_2 a^2 cos(\theta_1 + \theta_2) + g_2 a^2 cos(\theta_1 - \theta_2) + ... \\ & + \frac{g_3 a^3}{4} cos(3\theta_1) + \frac{g_3 a^3}{4} cos(3\theta_2) + \frac{3g_3 a^3}{4} cos(2\theta_1 + \theta_2) + ... \\ & + \frac{3g_3 a^3}{4} cos(2\theta_2 + \theta_1) + \frac{3g_3 a^3}{4} cos(2\theta_1 - \theta_2) + \frac{3g_3 a^3}{4} cos(2\theta_2 - \theta_1) \end{aligned} \tag{2.10}$$

The response in frequency domain is shown in Fig 2.10. Except for the fundamental tones f_1 and f_2, other harmonics and intermodulation products also are found. In practical applications, the terms $cos(2\theta_1)$, $cos(2\theta_2)$, $cos(3\theta_1)$, $cos(3\theta_2)$, $cos(\theta_1 + \theta_2)$, $cos(\theta_1 - \theta_2)$, $cos(2\theta_1 + \theta_2)$, $cos(\theta_1 + 2\theta_2)$ can be filtered out. But the terms $cos(2\theta_1 - \theta_2)$ and $cos(2\theta_2 - \theta_1)$ can not be easily filtered out, because these two terms are close to the fundamental tones. The terms $cos(2\theta_1 - \theta_2)$ and $cos(2\theta_2 - \theta_1)$ are called IMD3 products. By the same principle, there are other IMD products, such as IMD5, IMD7. The intermodulation products result in signal distortion and spectral regrowth. These characteristics constitute the static nonlinearity of PA.

The memory effects are another aspect of the nonlinearity of PA. It describes the influence of the history of the input signal on its current output value. In time domain, the memory effects refer to that the output of PA does not only depend on the current input sample, but also on the previous input samples. In frequency domain, AM/AM and AM/PM characteristics of the PA vary with the frequency of the input signal envelope. This behavior can be explained by the principle of a R-C circuit. The

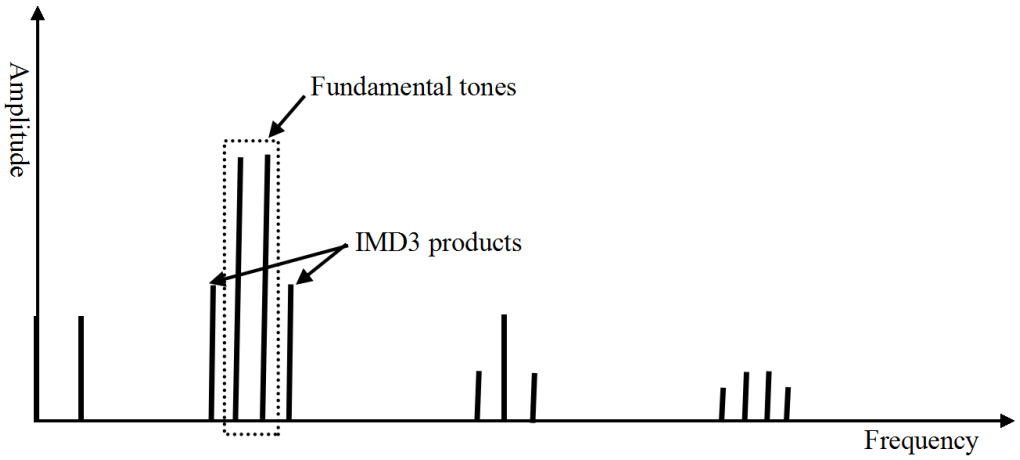


Figure 2.10: Distortion products of two-tone signal

memory effects are the result of energy storage in the system. They are a manifestation of the dynamical behavior of the system. For linear system, it can be described by time derivative or time-delay expressions. For nonlinear system, such as a PA, the output is influenced by the previous inputs.

Generally, in narrow band system, the memory effects can be ignored. But in wide band system, the time span of memory effects in PA can be comparable to the time variations of the input signal envelope [41] so that the memory effects can no longer be ignored. In AM/AM and AM/PM curves, the dispersion phenomenon indicates the presence of memory effect. The AM/AM and AM/PM characteristics of a PA with memory effects are shown in Fig. 2.11 and Fig. 2.12, respectively. In the constellation of a system with memory effects, there are not only rotation and compression distortions, but also dispersion distortions. Another characteristic is that it exists the phenomenon of spectral asymmetry.

The static nonlinearity and memory effects together constitute the nonlinearity characteristics of PA. In practical applications, they are related to each another. They jointly affect the characteristics of PA.

In the following, some criteria corresponding to the nonlinearity of PA are introduced, such as 1dB compression point, third-order Intercept Point (IP3), Adjacent Channel Power Ratio (ACPR), and Error Vector Magnitude (EVM).

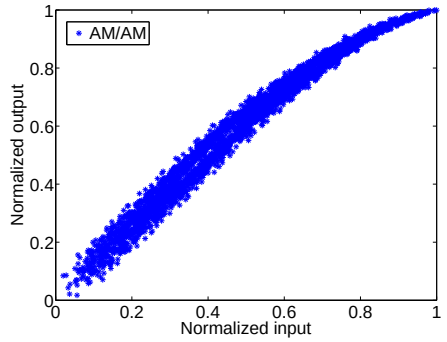


Figure 2.11: AM/AM characteristic of a PA with memory effects

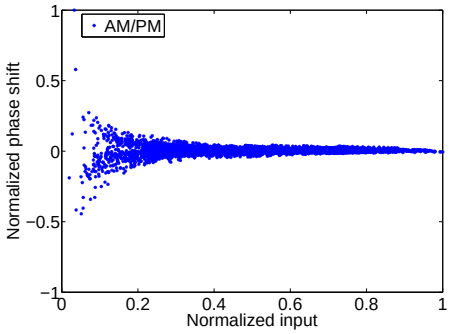


Figure 2.12: AM/PM characteristic of a PA with memory effects

1dB compression point

The gain of PA at 1dB compression point is defined as

$$G_{1dB} = G_{ideal} - 1dB \tag{2.11}$$

where G_{ideal} is the ideal gain. At 1dB compression point, the actual gain is 1dB lower than the ideal gain, as shown in Fig. 2.13. The corresponding input and output powers are called the input and output powers at 1dB compression point, respectively. The 1dB compression point is a typical parameter for evaluating PA's nonlinearity. By the same principle, another point called 3dB compression point can be defined, where the actual gain is 3dB lower than the ideal gain.

IP3

The third-order Intercept Point (IP3) relates nonlinear products caused by the third-order nonlinear terms $cos(2\theta_1 - \theta_2)$ and $cos(2\theta_2 - \theta_1)$. From (2.10), it can be seen that the amplitudes of the third-order nonlinear terms $cos(2\theta_1 - \theta_2)$ and $cos(2\theta_2 - \theta_1)$ are both proportional to a^3. When the fundamental increases by 1dB, these third-order intermodulation products increase by 3dB. IP3 point is defined as the theoretical level where the intermodulation products have equal amplitude of the fundamental tone. It is shown in Fig. 2.14.

ACPR

The spectral regrowth is generated by the nonlinearity of PA. ACPR represents the power ratio between the adjacent channel and the main channel of the signal. It

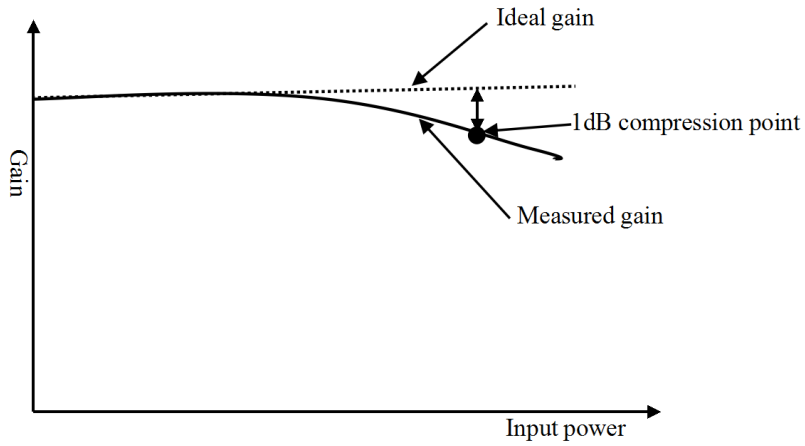


Figure 2.13: 1dB compression point

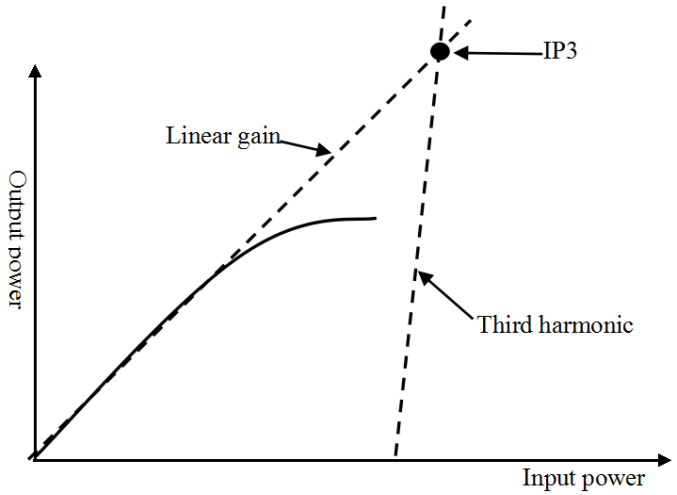


Figure 2.14: IP3

describes the degree of the spectral regrowth into the adjacent channels. ACPR is an important parameter for wireless radio standards, which is defined as follows

$$ACPR = 10log\frac{\int_{B_{adjacent}} P(f)df}{\int_{B_{main}} P(f)df} \tag{2.12}$$

where $P(f)$ represents the power spectral density of the output signal, B_{main} and $B_{adjacent}$ refer to the main and adjacent channels, respectively. An example is shown in Fig. 2.15. The offset is 5 MHz between the adjacent channel and the main channel.

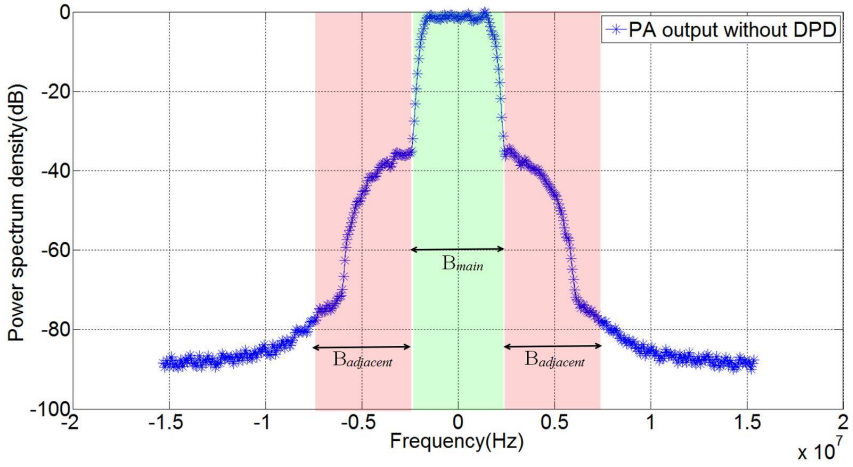


Figure 2.15: Main and adjacent channels

EVM

EVM is a measure of the shift of constellation diagram in time domain. It evaluates the error between the ideal signal and the practical measured signal in the receiver by vector format. EVM is defined as

$$EVM = \sqrt{\frac{\sum_{n=1}^{N} |S(n)_{ideal} - S(n)_{measured}|^2}{\sum_{n=1}^{N} |S(n)_{ideal}|^2}} \tag{2.13}$$

where N is the number of constellation points, S_{ideal} and $S_{measured}$ are the ideal constellation and the practical measured points, respectively. The difference between them is the error vector. It is shown in Fig. 2.16.

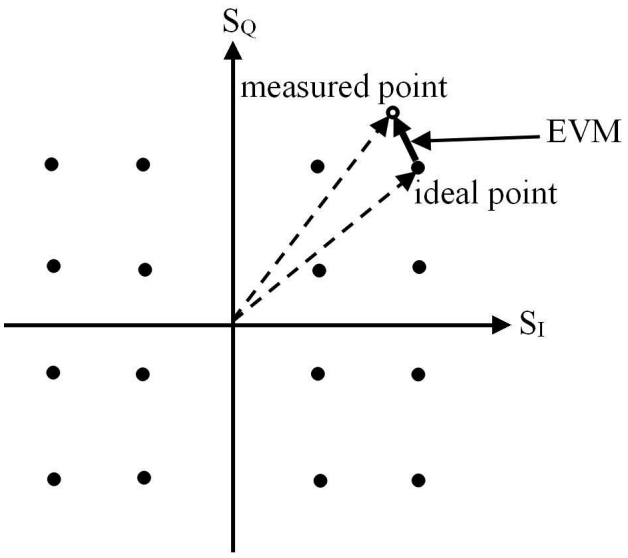


Figure 2.16: A constellation diagram for 16QAM signal

2.3.3 Power amplifier behavioral modeling

In order to compensate the nonlinearity of PA, it is necessary to model the behavior of PA accurately. In the following, some classical PA models are presented. In this thesis, all the models are considered in discrete-time domain. The PA models can be classified into memoryless models and memory models. The classical memoryless PA models include Saleh model, Rapp model and memoryless polynomial model.

Saleh model

Saleh model is usually used for modeling the behavioral of TWTA [36]. It composed of AM/AM and AM/PM functions, as follows

$$A(|x(n)|) = \frac{\alpha_a |x(n)|}{1 + \beta_a |x(n)|^2} \tag{2.14}$$

$$\Phi(|x(n)|) = \frac{\alpha_\varphi |x(n)|^2}{1 + \beta_\varphi |x(n)|^2} \tag{2.15}$$

where $|x(n)|$ is the amplitude of the input signal, $A(\cdot)$ and $\Phi(\cdot)$ are the AM/AM and AM/PM conversion functions, respectively; α_a, β_a, α_φ and β_φ are the coefficients of the model.

Rapp model

Rapp model is usually used for modeling the behavioral of SSPA [37], because the phase distortion is small for SSPA. The phase distortion can be ignored. Rapp model is represented by

$$A(|x(n)|) = \frac{|x(n)|}{[1 + (|x(n)|/V_{sat})^{2p}]^{\frac{1}{2p}}} \tag{2.16}$$

$$\Phi(|x(n)|) \approx 0 \tag{2.17}$$

where V_{sat} is the saturation voltage of PA, and p is called the smoothing factor which can affect the linear degree of the model. The higher the value of p is, the more linear the characteristic of the model is. Because the number of this model's coefficients is small, the complexity of this model is low.

Polynomial model

The polynomial model is another usual memoryless PA model. It has been used extensively for memoryless or quasi-memoryless PAs. It can be expressed by

$$y(n) = \sum_{k=0}^{K} a_k x(n)|x(n)|^k \tag{2.18}$$

where $x(n)$ and $y(n)$ are the input and output signals, respectively, $K+1$ is the maximum order of the polynomial, and a_k are the complex-valued coefficients of the model. Because the impact of the even-order polynomial terms can be ignored in the PA nonlinearity. Generally, only odd-order nonlinear terms are considered [42, 43]. The odd-order polynomial model is presented by

$$y(n) = \sum_{k=0}^{K} a_k x(n)|x(n)|^{2k} \tag{2.19}$$

The memoryless PA model is used in narrow band systems because the memory effects are not important. In wide band systems, the memoryless model results in poor performance, thus the nonlinear model with memory is adopted. The general nonlinear model with memory includes Volterra Series (VS), Generalized Memory Polynomial (GMP) model, Dynamic Deviation Reduction (DDR) model [19, 44], Memory Polynomial (MP) model, Hammerstein model, Wiener model, and other simplified VS variants.

Volterra series model

VS model is the most comprehensive model for nonlinear PAs with memory effects. It can be thought of as a Taylor series with memory [45, 46]. Its mathematical expression is described as

$$\begin{aligned} y(n) = h_0 &+ \sum_{p_1=0}^{P} h_1(p_1)x(n-p_1) \\ &+ \sum_{p_1=0}^{P}\sum_{p_2=0}^{P} h_2(p_1,p_2)x(n-p_1)x(n-p_2) \\ &+ ... \\ &+ \sum_{p_1=0}^{P}\sum_{p_2=0}^{P}...\sum_{p_Q=0}^{P} h_q(p_1,p_2...,p_Q)x(n-p_1)x(n-p_2)...x(n-p_Q) \end{aligned} \tag{2.20}$$

where $h_1(p_1)$, ... , $h_q(p_1, p_2..., p_q)$ represent the Volterra kernels, P is the memory depth, and Q is the nonlinearity order. Theoretically, it can infinitely approximate the characteristic of PA. But the number of the model's coefficients increases drastically with the nonlinearity order and memory depth. So that the complexity of the system is increased and the inverse model of PA is hard to calculate. The complexity of VS model is given [47] by

$$C(P,Q) = \sum_{q=1}^{Q} \frac{(P-1+q)!}{(q-1)!(P-1)!} \tag{2.21}$$

Generalized memory polynomial model

In order to alleviate the complexity burden, some techniques are proposed to simplify VS model. GMP model is a simplified version of VS model. It restricts the interaction between the various terms of VS model by the leading and lagging envelop terms. GMP model [13] is described by

$$
\begin{aligned}
y(n) = & \sum_{q=0}^{Q_a} \sum_{p=0}^{P_a} a_{pq} x(n-p)|x(n-p)|^q \\
& + \sum_{q=0}^{Q_b} \sum_{p=0}^{P_b} \sum_{m=1}^{M_b} b_{pqm} x(n-p)|x(n-p-m)|^q \\
& + \sum_{q=0}^{Q_c} \sum_{p=0}^{P_c} \sum_{m=1}^{M_c} c_{pqm} x(n-p)|x(n-p+m)|^q
\end{aligned} \tag{2.22}
$$

where P_a, P_b, and P_c are the memory depths of each of the branches of the model; Q_a, Q_b, and Q_c are the nonlinearity order of each of the branches of the model; M_b and M_c are the memory depths of the lagging and leading branches of the model, respectively; a_{pq}, a_{pqm}, and c_{pqm} are the coefficient of the time-aligned, lagging and leading branches of the model, respectively. Compared with VS model, GMP model requires less complexity.

Memory polynomial model

The memory polynomial model is a further-simplified version of VS model. It is widely used for the behavioral modeling and digital predistortion of PAs with memory effects [39]. The well-known MP model is given [12] by

$$
y(n) = \sum_{p=0}^{P} \sum_{q=0}^{Q} a_{pq} x(n)|x(n-p)|^q \tag{2.23}
$$

where a_{pq} represents the model's coefficients, P and $Q+1$ are the memory depth and the maximum nonlinearity order of the model, respectively. It achieves a good trade-off between the complexity and modeling accuracy of the system.

Hammerstein model

Hammerstein model is a two-box model. It consists of a memoryless nonlinear model followed by a Finite Impulse Response (FIR) filter. Its structure is shown in Fig. 2.17. The memoryless nonlinear part represents the static nonlinearity, and the FIR filter represents the memory effect. Its mathematical expression is described as

$$v(n) = \sum_{q=0}^{Q} a_q x(n)|x(n)|^q \tag{2.24}$$

$$y(n) = \sum_{p=0}^{P} h_p v(n-p) \tag{2.25}$$

where $x(n)$ represents the input of the memoryless nonlinear part, $v(n)$ represents the output of the memotyless nonlinear part (or the input of the FIR filter), and $y(n)$ represents the output of the FIR filter. a_q and h_p are the coefficients of memoryless nonlinear part and FIR filter, respectively. $Q+1$ and P represent the maximum nonlinear order and the memory depth of the model, respectively.

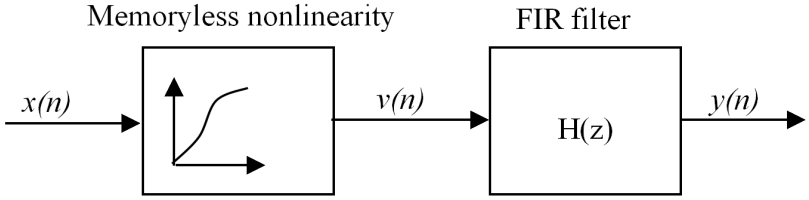


Figure 2.17: Hammerstein model

Wiener model

Wiener model has an opposite structure of Hammerstein model. It is composed of a FIR filter followed by a memoryless nonlinear part. Its structure is shown in Fig. 2.18. And the mathematical expression is described by

$$v(n) = \sum_{p=0}^{P} h_p x(n-p) \tag{2.26}$$

$$y(n) = \sum_{q=0}^{Q} a_q v(n)|v(n)|^q \tag{2.27}$$

The identification process of Wiener model is more complicated than that of Hammerstein model. Hammerstein and Wiener model are both the simplified version of

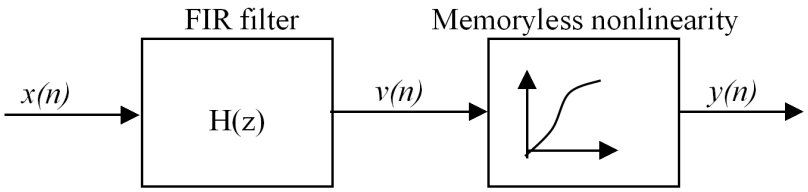


Figure 2.18: Wiener model

Volterra model. Therefore, they are less complexity than Volterra model. However, their performance is less than that of the full Volerra model.

Except for these PA models mentioned above, it still exists other models for PA modeling, such as Dynamic Deviation Reduction (DDR) model [19, 44] and so on.

2.4 Linearization techniques

PAs are inherently nonlinear. People want to realize high linearity while maintain a high power efficiency. There are two directions to achieve this goal. One is to improve the power efficiency, such as Doherty PA technique [48] and envelop tracking approaches [49, 50]. The other is to extend the linear behavior into high efficiency regions of operation. The latter is known as the linearization technique.

A large number of PA linearization techniques are proposed, including Backoff, Feedforward, Feedback, LInear amplification with Nonlinear Components (LINC), Envelope Elimination and Restoration (EER), PreDistortion, and so on. The details of these techniques are briefly presented in this section.

2.4.1 Backoff

Backoff is a general solution of PA's linearization, which makes PA work only in the linear region thanks to a power backoff [51]. In other words, the PA is operated at the region of low output power. This technique does not need additional auxiliary circuit and is easy to implement. The drawback is a loss of the power efficiency. The power efficiency is low when the PA is operated in its linear region as can be seen in Fig. 2.7. Moreover, the performance of Backoff also depends on the linearity degree

of the PA. For some highly nonlinear PA, Backoff is not suitable. Therefore, Backoff is only treated as an assistant technique rather than the main amplifier linearization technique.

2.4.2 Feedforward

The principle of Feedforward is shown in Fig. 2.19. Feedforward has two paths: the main path and the error path [52, 7].

1. The first step is to extract the distortion products. The input signal a is first separated into two parts by a coupler: signal b and signal c. Signal b is amplified by the main amplifier and the output signal d of the main power amplifier is obtained, which includes amplified original signal and unwanted distortion products. Signal d is duplicated to signal e and signal f by using a second coupler. Then signal f goes into a attenuator and generates signal g. The attenuation factor of the attenuator is calibrated such as the amplitude of original signal in g is similar to that in h. h is the delayed version of c. h is subtracted from g, then signal l only includes the distortion.

2. The second step is to amplify the distortion products. Signal l is amplified by the error power amplifier and to give signal m. The gain of the error power amplifier is calibrated such as the distortion products in signal m is equal to those in n, which is the delayed version of signal e.

3. The last step is to eliminate the distortion products. Because the amplitudes of distortion products are same in signals n and m, the distortion products are eliminated by the subtraction. As a result, the output signal y contains only the amplified original signal.

Feedforward has good linearization performance among power amplifier linearization techniques [53, 54]. It can work in both constant and nonconstant envelope signals. It has high stability and theoretically allows operation over a wide bandwidth. These advantages make Feedforward still be studied. However, the matching and drift problems make Feedforward be complicated. The gains of the main power amplifier and the error power amplifier need to match. The delay elements need to match for the signal synchronization. Especially, in wide bandwidth, these matchings are very challenging.

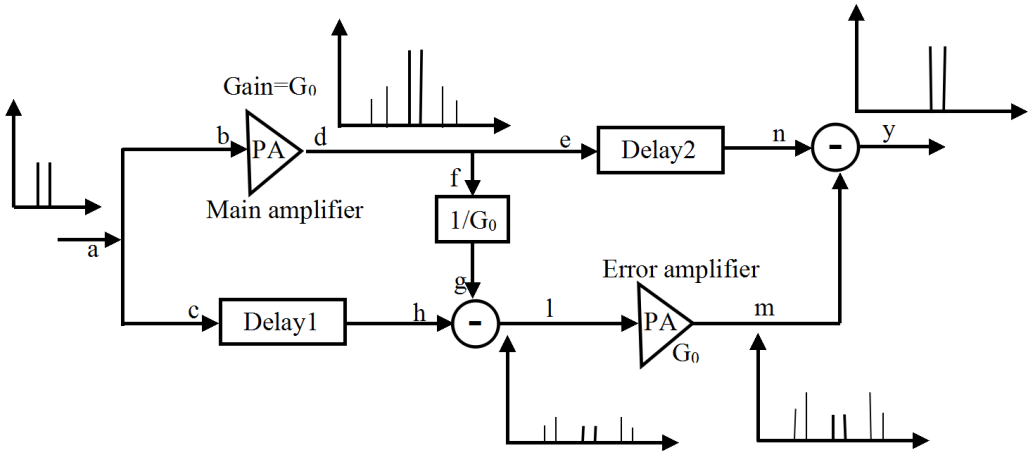


Figure 2.19: Feedforward linearization technique

2.4.3 Feedback

Feedback is a linearization technique which makes the output of the amplifier be fed back to the input of the amplifier, and compared with the input signal to reduce the distortion [33, 55]. The principle is shown in Fig. 2.20. There are two forms of feedback: positive feedback and negative feedback. Usually, for the amplifier, the feedback is the negative feedback.

Suppose the gain of the power amplifier is A, the open-loop transfer function is

$$A = \frac{V_{out}}{V_{in}} \tag{2.28}$$

where V_{in} and V_{out} are the input and output of the amplifier, respectively. The feedback signal is defined as

$$V_f = \beta V_{out} \tag{2.29}$$

where β is the feedback factor, which controls the amount of the feedback. And the relationship between the original signal V_s, the feedback signal V_f and the intput V_{in} of the amplifier is

$$V_{in} = V_s - V_f \tag{2.30}$$

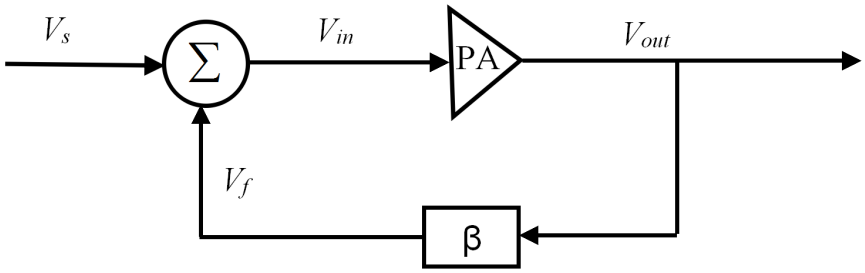


Figure 2.20: Feedback linearization technique

According (2.29) and (2.30), we get

$$V_{in} = V_s - \beta V_{out} \tag{2.31}$$

Then according (2.28) and (2.31), we can get the close-loop transfer function

$$A_f = \frac{V_{out}}{V_s} = \frac{A}{1 + \beta A} \tag{2.32}$$

In general, $\beta A >> 1$, so (2.32) becomes

$$A_f = \frac{1}{\beta} \tag{2.33}$$

From (2.33), we can see that the gain of the feedback technique is almost independent of the gain A of the amplifier and only depends on the feedback factor β. Therefore, the implementation of the feedback technique has low complexity. But feedback technique still has some disadvantages. The first is that the gain of the power amplifier A is required large enough to meet $\beta A >> 1$. The second is the delay of the feedback path, it is very difficult to make the input and output operate at the same time exactly. The third problem is the stability. The system must limit the bandwidth in order to assure the stability of the system. So that the bandwidth of the feedback technique is restricted.

2.4.4 Linear amplification with nonlinear components

LINC technique is applied to some modulated signals where both the amplitude and phase are variable. The principle is shown in Fig. 2.21. The input signal $V_{in}(t)$ is separated into two constant-envelope signals $V_1(t)$ and $V_2(t)$. These two constant-envelope signals are amplified by two power amplifiers with same characteristic respectively. Then the two amplified signals are combined together to generate the final output signal $V_{out}(t)$.

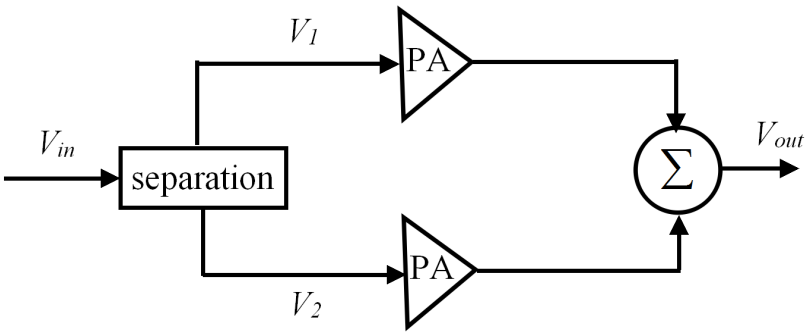


Figure 2.21: LINC linearization technique

$$\begin{aligned} V_{in}(t) &= a(t)cos[w_c t + \phi(t)] \\ &= 0.5V_0 sin[w_c t + \phi(t) + \theta(t)] - 0.5V_0 sin[w_c t + \phi(t) - \theta(t)] \\ &= V_1(t) + V_2(t) \end{aligned} \tag{2.34}$$

$$V_1(t) = 0.5V_0 sin[w_c t + \phi(t) + \theta(t)] \tag{2.35}$$

$$V_2(t) = -0.5V_0 sin[w_c t + \phi(t) - \theta(t)] \tag{2.36}$$

$$\theta(t) = sin^{-1}(a(t)/V_0) \tag{2.37}$$

where $V_{in}(t)$ is the input signal of the system, $a(t)$ is the nonconstant amplitude of input signal, and V_0 is the max amplitude of $a(t)$. The concept of LINC is very appealing,

and the actual implementation is possible. But the separation and combination of the signal are complicated and it is problematic for the perfect matching of amplifiers in two branches [56, 57].

2.4.5 Envelope elimination and restoration

EER linearization technique is firstly proposed by Khan [58]. It combines a high efficiency nonlinear PA with an envelope amplifier to implement a high efficiency linear PA. Its classical form is shown in Fig. 2.22. An envelope limiter eliminates the envelope of the RF carrier signal. The resulting constant-amplitude phase modulated signal is amplified by a high efficiency PA. Finally, the RF PA restores the envelope to the phase modulated signal by the amplitude modulation. An amplified replica of the input signal is created.

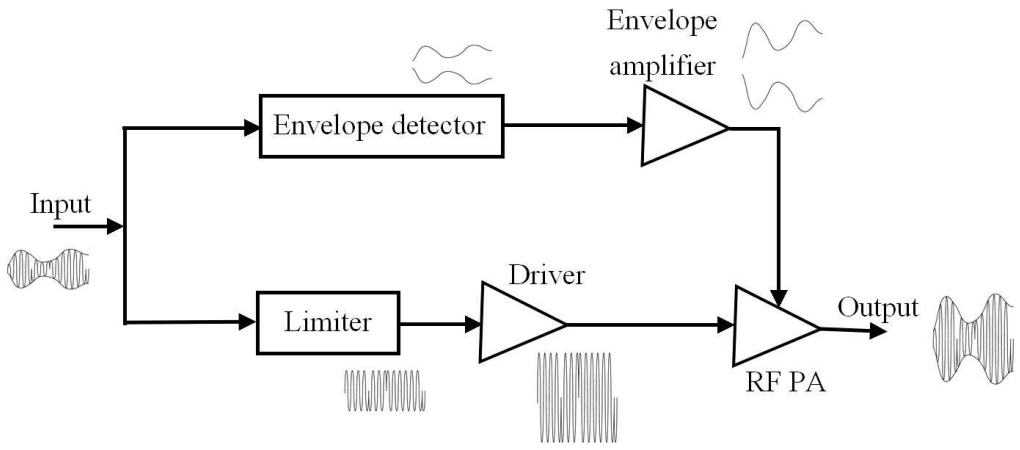


Figure 2.22: Envelope elimination and restoration technique

In contrast to linear amplifiers, EER technique allows a high efficient amplification in a wide power output range. And it produces a high average efficiency for a wide range of signals with different backoff requirements. Two of the most important factors affecting the linearity are the envelope bandwidth and the alignment between the envelope and the phase modulated signals. Generally, the envelope bandwidth must be at least twice the RF bandwidth. The misalignment between the envelope and the phase modulated signals must not exceed 10% of the inverse of the RF bandwidth [59]. This technique is widely adopted in single carrier or narrow-band applications.

2.4.6 Predistortion

Predistortion technique is one of the most effective linearization techniques. Its principle is shown in Fig. 2.23. A functional block called the PreDistorter (PD) is inserted in front of the PA, which has the inverse nonlinear characteristics of that of the PA, so that the cascaded PD-PA system has a linear behavior. Predistortion doesn't require a deep knowledge about the physic circuit of power amplifier and is easy to realize. It can be implemented entirely using digital processors (DSP and FPGA) and this offers re-configurability capabilities. So it is most efficient and promising among these linearization techniques for future communication systems.

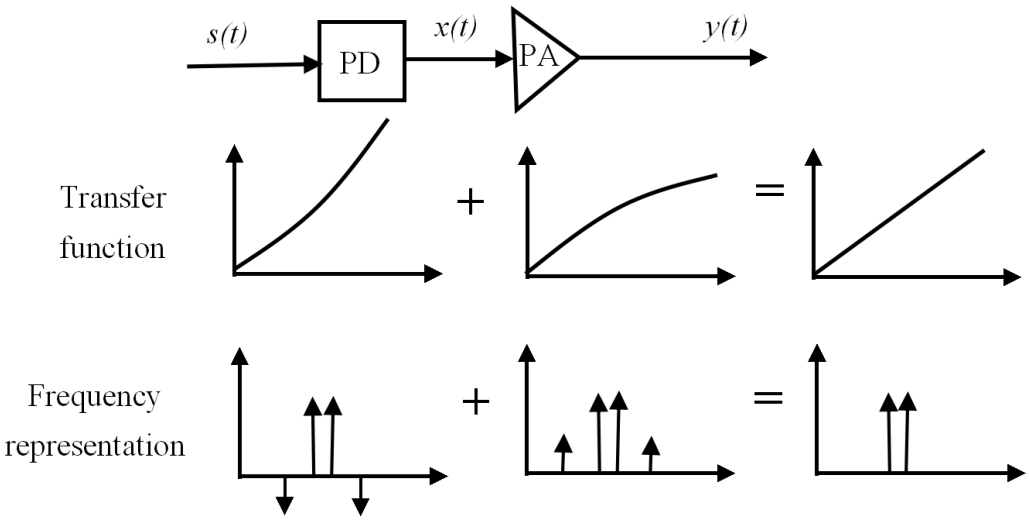


Figure 2.23: Predistortion technique

To obtain a good linearization performance, it is necessary to design a good PD which is the precise inverse model of the PA model. Therefore, the crucial issue is how to obtain an excellent PD model. More details about the predistortion is presented and discussed in the following chapters.

3

Digital predistortion

In this chapter, the DPD is described in detail. We give the purpose of DPD in section 3.1. In section 3.2, DPDs based on different implementation methods are introduced. The description about two different learning architectures of DPDs is given in section 3.3. Several identification algorithms are described in section 3.4. Finally, in section 3.5 we compare two different DPD learning architectrures (ILA and DLA) by analyzing the condition number of the matrix to be inverted in LS identification algorithm, the modeling performance and the linearization performance of DPD.

3.1 Digital predistortion overview

According to the position of predistorter, predistortion techniques can be classified into three categories: Radio Frequency (RF) predistortion, Intermediate Frequency (IF) predistortion and BaseBand (BB) predistortion. RF predistortion is realized in RF bands. Usually, the operating frequency in RF bands is high so that it is difficult to predistort the RF signal. It requires complicated analog circuit design and the improvement of the linearization performance is not good enough. Thus, RF predistortion is less used. For IF predistortion, the operating frequency is in IF bands which are lower than RF bands. One drawback of IF predistortion is that the nonlinear control is still implemented with difficulty by digital devices. Compared with RF and IF predistortion, BB predistortion has the lowest operating frequency bands. It is processed at

baseband frequency where the signal rate is relatively low, and the hardware requirement is not high. BB predistortion can be implemented by components such as FPGA or DSP. Nowadays, BB predistortion is adopted most widely, as shown in Fig. 3.1.

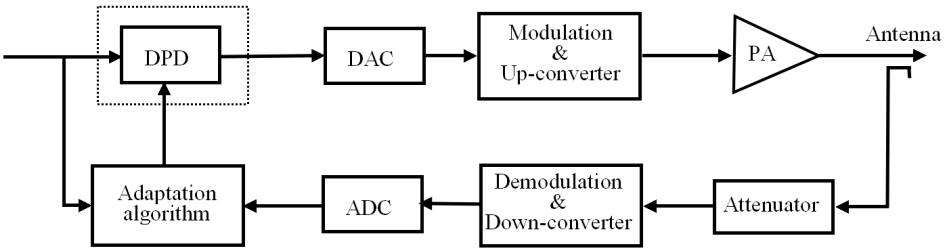


Figure 3.1: Block diagram of baseband digital predistortion

According to the signal types, predistortion techniques can be classified into analog predistortion and digital predistortion. In 1959, the concept of predistortion was firstly proposed by MacDonald [60]. It proposed to compensate the nonlinearity of the PA by using a component characterized by the PA's inverse behavior. In 1979, an analog predistorter was realized by Satoh and applied to the TWTA [61]. The linearization performance was moderate. In 1984, an analog predistortion system operated at an intermediate frequency (130MHz) was designed using diodes by Nojima et al. and adopted in the mobile telephone system [62]. After the 1980s, the predistortion techniques were developed rapidly. Some integrated chips, such as MAX2009 and MAX2010 of Maxim company, have been realized. As the digital technology developed, digital predistortion becomes possible. Compared with analog predistortion, digital predistortion is more accurate. In the past two decades, Digital PreDistortion (DPD) techniques have been more concerned with the development of the digital technology.

The aim of DPD is to make the system be linear when the PA operates in the nonlinear region. Consider the scenario depicted in Fig. 3.2. Assume r_{in} be the amplitude of the input signal, the amplitude of the PA's output signal actually is r_{out}. The relationship of r_{in} and r_{out} is nonlinear. Ideally, the desired output amplitude of the system is $r_{desired\_out}$ from the linear response. The value $r_{desired\_out}$ is used to search through the output characteristic of the PA. The proper input amplitude to the PA is determined r_{pd}. The task of the PreDistorter (PD) is to adjust the original input amplitude r_{in} to be the corrected amplitude r_{pd}. Then r_{pd} should produce the correct output amplitude to give the PD-PA system a linear response. In addition, the phase of input signal is also predistorted to obtain a constant phase difference.

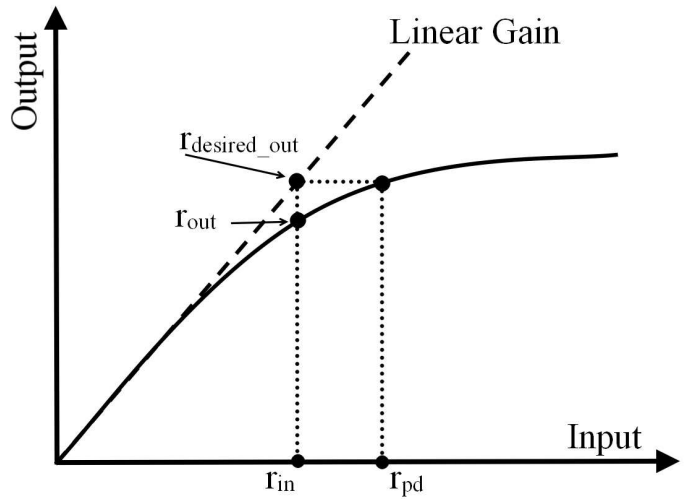


Figure 3.2: Predistortion conception

Fig. 3.3 exhibits the improvement in PA's operating region with the predistortion. In the figure, the range of input values is represented as a gray rectangle. The left rectangle is for the range of PA's operating region without predistortion. The peak power can not exist too far into the nonlinear region, otherwise the distortion is so large that it will affect the transmission. The right rectangle is for the range of operating region with predistortion. The range of input power level with the predistortion is greatly improved. But the upper level of operating point is limited by the characteristic of PA (the desired output power can not be beyond the saturation limit of PA). At the full saturation level of PA, any increase in the input power does not produce an increase in output power. Therefore, any input signal with an amplitude higher than the saturation level will not be fully corrected and distortions will increase again.

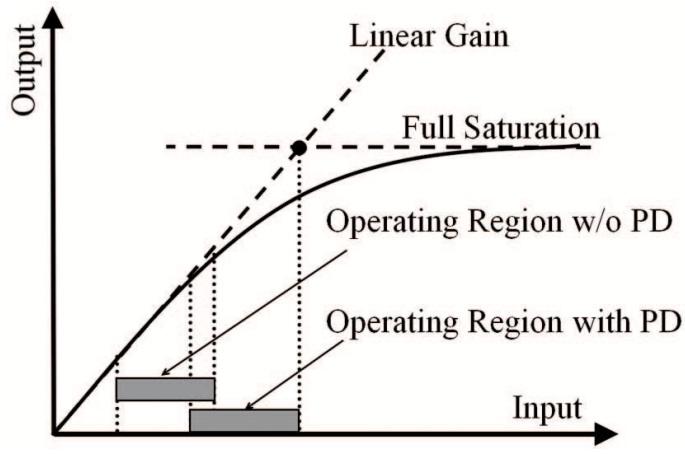


Figure 3.3: Operating region improvement

3.2 Popular solutions

A large number of DPD techniques have been published in the literature. The widely used implementation methods can be classified into three categories: Look-Up Table (LUT) based DPDs, Neural network based DPDs and model based DPDs (or called polynomial based DPDs). They are described in the following.

3.2.1 LUT based DPD

LUT based DPD is a basic method for implementing DPD technique [10, 11, 63, 64, 65]. It is a simple and efficient technique. The values of desired predistorted signals are stored in the memory units as a table. The table is indexed by the amplitude of the input signal, then the predistorted signal is extracted.

The core of LUT based DPD is the table used to generated the predistorted signal. The important task is to properly construct a LUT which represents the inverse function of PA. The contents of LUT can be implemented in Cartesian form [66, 67], polar form [68, 69] and complex-gain form [70, 71]. In LUT based DPD, compared with the Cartesian and polar forms, the complex-gain form requires less table size and less calculations. Thus, the complex-gain form is widely used. Fig. 3.4 shows a typical structure of complex-gain LUT-based predistorter. The predistorted signal is obtained by

$$x(n) = u(n)g(r_u(n)), \tag{3.1}$$

where $u(n)$ is the input signal, $r_u(n)$ the amplitude of input signal, and $g(\cdot)$ the complex gain of the predistorter.

Furthermore, how to distribute the LUT entries is also an important issue. In other words, the LUT indexing technique influences the obtained performance. Commonly, equispaced LUT is widely used because of its simplicity and low computational complexity. The LUT entries can be equispaced in linear amplitude, linear power or logarithmic power [72]. For equispaced LUT addressed using linear amplitude, it exhibits good performance by uniformly spacing the table entries along the input signal range. But there are some drawbacks, which are related to the calculation for a rectangular to polar coordinates converter. For equispaced LUT addressed using linear power, it inherently leads to dense table entries at large signal region, while few entries are assigned to small signal region. This indexing technique is suitable for class A PAs, where there is almost no gain variation at small signal region. For equispaced LUT addressed using logarithmic power, by contrast, more entries are assigned to large signal

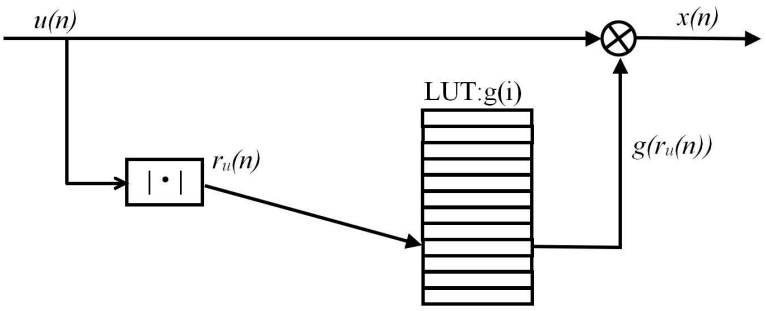


Figure 3.4: Structure of complex-gain LUT-based predistorter

region.

These LUT indexing techniques mentioned above only fit a type of specific PAs. In order to distribute the LUT entries more optimally, the nonuniform spacing method is reported by several authors [63, 65, 70, 73]. Cavers in [63] proposed a companding function to minimize the intermodulation distortion power due to LUT quantification errors. The proposed companding function depends on the Probability Density Function (PSD) of the input signal. A typical structure of complex-gain LUT-based predistorter including companding function is shown in Fig. 3.5. In LUT based DPDs, the nonuniform spacing method provides a better prospect. It is a research trend for future LUT based DPDs.

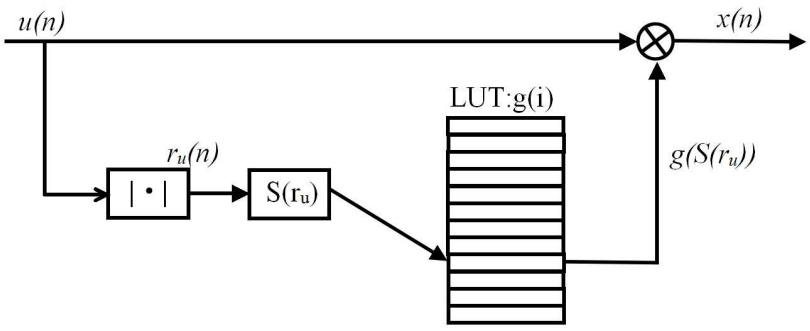


Figure 3.5: Structure of complex-gain LUT-based predistorter including companding functions

For LUT based DPD, in order to improve the linearization performance, it requires a large table size. A large number of LUT entries leads to a long convergence time. There are other LUT based DPDs, which adopt the interpolation technique. The interpolated LUT based DPDs [64, 74, 75, 3, 4] can decrease the table size while maintain good linearization performance.

The LUT based DPD has been concerned by researchers because it has simplicity and is easy to implement [11, 10, 63, 64, 65]. But it needs to overcome the distortion due to the quantification error. The LUT based DPD is applied mostly in memoryless PAs, only few is applied in PAs with memory effects.

3.2.2 Neural network based DPD

Artificial Neural Network (ANN) has an excellent capability to accurately approximate nonlinear functions [76]. In [77], it has been proven that a feed forward ANN with a single hidden layer and nonconstant, bounded and monotone-increasing continuous activation functions, can approximate any nonlinear function with any desired error. Hence, ANN can be used to model a PA and its predistorter. Several neural network based DPDs are reported in the literature [14, 78, 79, 80, 81, 82, 83].

The model of a single neuron is shown in Fig. 3.6. And the mathematical expression is given by

$$y = F(wx + b) \tag{3.2}$$

where x and y are the input and output of the neuron, respectively, w the weight value, b the bias value, and F the activation function. Usually, the activation function is sigmoid function or tanmoid function. As an example, when sigmoid function is used as the activation function, (3.2) can be written as

$$y = F(wx + b) = \frac{\lambda \cdot (1 - e^{\mu(wx+b)})}{1 + e^{\mu(wx+b)}} \tag{3.3}$$

where λ and μ are constants.

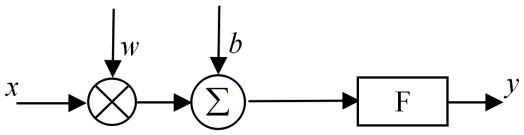


Figure 3.6: Model of a single neuron

Generally, ANN is composed of a number of neurons. Depending on neural network structure, they can be divided into three types: MultiLayer Perceptron (MLP), Radial Basis Function artificial Neural Network (RBFNN) and Recurrent artificial Neural Network (RNN) [14]. MLP and RBFNN are the two most commonly-used types of feed forward neural network. Their only fundamental difference is the way in which hidden units combine values coming from preceding layers in the network. MLP uses inner products, while RBFNN uses Euclidean distance. The structure of a simple feed forward neural network is shown in Fig. 3.7.

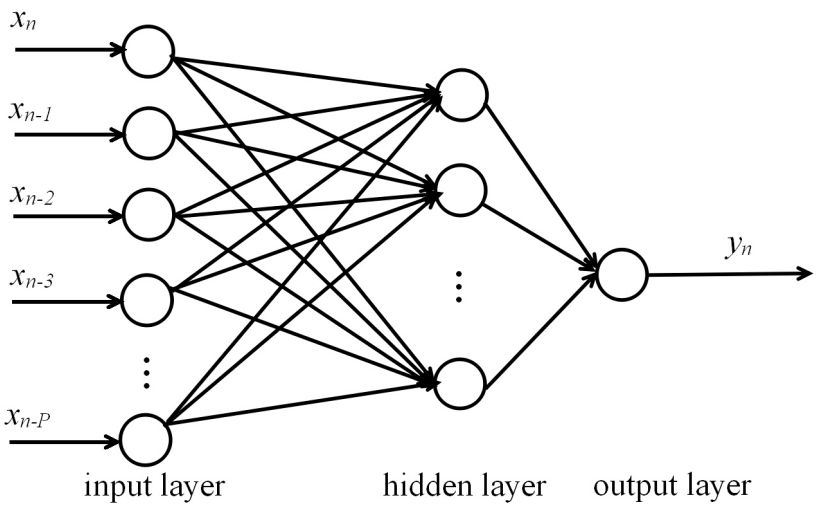


Figure 3.7: Structure of a three-layer feed forward neural network

MLP has one or more hidden layers for which the activation function is the inner product of the inputs and weights, plus a bias. The inputs are fully connected to the first hidden layer, each hidden layer is fully connected to the next, and the last hidden layer is fully connected to the outputs. It is the most popular neural network in practical applications due to its simplicity. In [79], a complex-valued MLP is used to mimic the AM/AM and AM/PM characteristics. In [80], a real-valued MLP is used to predict the I&Q outputs of a device-under-test.

RBFNN usually has only one hidden layer for which the activation function is based on the Euclidean distance between the input vector and the weight vector. The output layer performs a simple weighted sum with a linear output. In [81], a RBFNN has been used for modeling the dynamic nonlinear behavior of an RF power amplifier.

RNN is a class of ANN which includes feedback paths. The structure of a simple RNN is shown in Fig. 3.8. Indeed, RNN allows for better emulation of the interaction

between the input and output samples and the cross terms of the nonlinear system. But its training algorithm is complex, therefore, it is not popular for real-time deployment [78]. In [82, 83], RNN is used for modeling RF power amplifier and its digital predistortion.

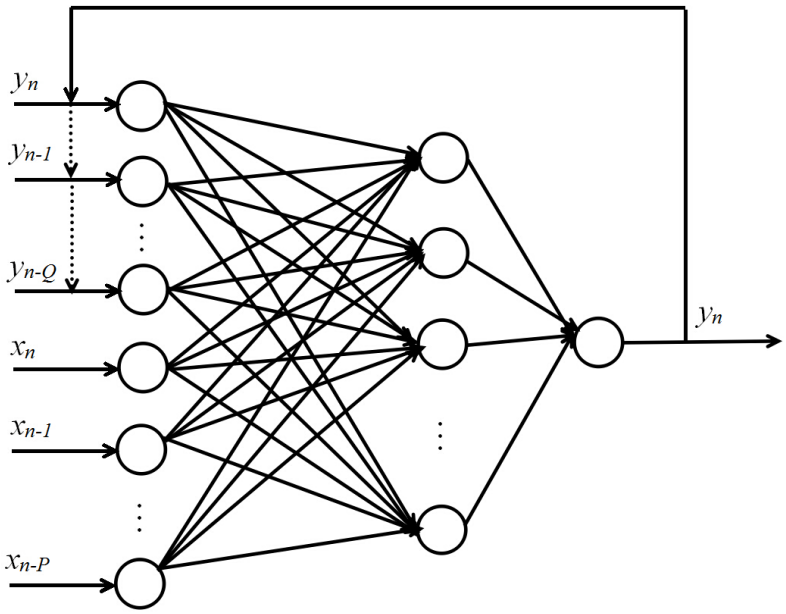


Figure 3.8: Structure of a recurrent neural network

Because neural network has a strong ability of learning and imitation, the neural network based DPDs can obtain good linearization performance. However, the training of the neural network is complex and time-consuming. The hardware implementation of the neural network is not easy in practical applications.

3.2.3 Model based DPD

Model based DPDs [12, 13, 84, 85], also called polynomial based DPDs, are mostly used for the linearization of PA. In section 2.3.2, some typical models are presented for modeling PA. Similarly, the predistorter is also a nonlinear system. Some of these models can be used to model the predistorter. For memoryless system, the polynomial model can be used to model the predistorter. For memory system, the models such as Volterra series model, generalized memory polynomial model, memory polynomial model, Hammerstein model and Wiener model all are used to model the predistorter. These models were introduced in section 2.3.2. In this section they will no longer

be repeated. It is worth noting that, the input and output of the model have different meanings between modeling the PA and modeling the predistorter.

3.3 Learning architecture

DPD architectures can be classified into two types [20, 26]: Indirect Learning Architecture (ILA) and Direct Learning Architecture (DLA). In ILA approach [12], a postdistorter model is first assumed, the measured output of PA is taken as the input of postdistorter, and the measured input of PA is taken as the output of postdistorter. Then the postdistorter is estimated by an identification algorithm such as Least Squares (LS) method. Finally, the postdistorter is placed before the PA as its predistorter directly. In DLA approach [86, 87, 1], a model of PA's behavior is first predefined and its coefficients are identified, then the predistorter is found by reversing the identified PA model. In this section, the details of ILA and DLA are presented.

3.3.1 Indirect learning architecture

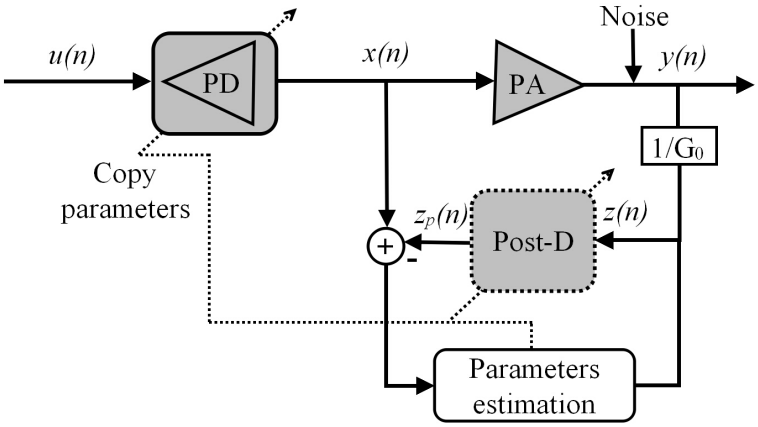


Figure 3.9: Indirect learning architecture

Fig. 3.9 shows the block diagram of ILA. Firstly, the postdistorter is identified. Usually, the postdistorter is modeled as a MP [20]. The output of the postdistorter can be expressed as

$$z_p(n) = \sum_{k=0}^{K} \sum_{l=0}^{L} a_{kl} \Phi_{kl}[z(n)] \tag{3.4}$$

where $z(n) = \frac{y(n)}{G_0}$, G_0 is the desired gain, a_{kl}, $k = 0, \cdots, K$ and $l = 0, \cdots, L$ are the complex-valued coefficients, K refers to the memory depth, $L+1$ the maximum nonlinearity order, and $\Phi_{kl}[z(n)] = z(n-k)\,|z(n-k)|^l$. Theoretically, it requires $z_p(n) = x(n)$ in a perfect system. $x(n)$ and $y(n)$ are the measured input and output of the PA, respectively. Assume that the total number of samples is N, we get

$$\mathbf{z_p} = \mathbf{Za} \tag{3.5}$$

where $\mathbf{z_p} = [z_p(1), \cdots, z_p(N)]^T$, $\mathbf{Z} = [\mathbf{u}_{00}, \cdots, \mathbf{u}_{0L}, \cdots, \mathbf{u}_{kl}, \cdots, \mathbf{u}_{K0}, \cdots, \mathbf{u}_{KL}]_{N \times KL}$, $\mathbf{u}_{kl} = [\Phi_{kl}[z(1)], \cdots, \Phi_{kl}[z(N)]]^T$, and $\mathbf{a} = [a_{00}, \cdots, a_{0L}, \cdots, a_{kl}, \cdots, a_{K0}, \cdots, a_{KL}]^T$. Classical LS method can be used to extract the coefficients

$$\hat{\mathbf{a}} = (\mathbf{Z}^H\mathbf{Z})^{-1}\mathbf{Z}^H\mathbf{z_p} \tag{3.6}$$

where $(\cdot)^H$ denotes complex conjugate transpose operation. After the identification of the postdistorter, a copy of the postdistorter is placed directly in front of the PA as its predistorter.

3.3.2 Direct learning architecture

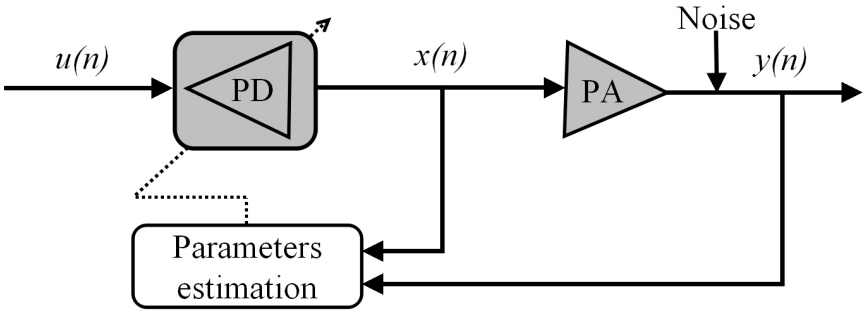


Figure 3.10: Direct learning architecture

The block diagram of DLA is illustrated in Fig. 3.10. As mentioned before, the determination of the PD needs two steps in DLA. Firstly, the model of PA's behavior needs to be predefined. Here MP is adopted as the PA model. The PA model is defined as

$$y(n) = \sum_{p=0}^{P}\sum_{q=0}^{Q} c_{pq}\Phi_{pq}[x(n)] \tag{3.7}$$

where $\Phi_{pq}[x(n)] = x(n-p)\,|x(n-p)|^q$, P and $Q+1$ are the memory depth and nonlinearity order, respectively, and c_{pq} are the model's coefficients. $x(n)$ and $y(n)$ are the input and output of the PA, respectively. The coefficients c_{pq} are extracted by LS method as mentioned before. Secondly, the identified model of the PA is reversed for the determination of the PD which is the complicated step. How to obtain the PD in DLA is presented in later section.

3.4 Identification algorithms

In DPD, the PD model needs to be identified by an adaptive algorithm. The common identification algorithms include Least Squares (LS), Least Mean Squares (LMS) and Recursive Least Squares (RLS). LS algorithm was presented in section 3.3.1. In this section, the LMS and RLS identification methods are presented, respectively.

3.4.1 Least mean squares

LMS method is an adaptive algorithm based on gradient method. Its cost function is defined by minimizing the mean square error, as follows

$$\epsilon = E\left\{|e(n)|^2\right\} \tag{3.8}$$

where $e(n)$ is the error between the desired output and actual output. The error is defined by

$$e(n) = d(n) - y(n) \tag{3.9}$$

$$y(n) = \mathbf{x}^T(n)\mathbf{w}(n) \tag{3.10}$$

where $d(n)$ is the desired output, $y(n)$ the actual output, $\mathbf{x}(n)$ the system input sequences, $\mathbf{w}(n)$ the weight function and $(\cdot)^T$ transpose operation. The update of the weight is given by

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \mu e(n)\mathbf{x}(n) \tag{3.11}$$

where μ is the step which controls the convergence speed. In LMS algorithm, the iteration is run, until a set of acceptable weight values are achieved such that the cost function ϵ is small enough.

In [88], a Nonlinear Filtered-x LMS (NFxLMS) algorithm is used to the linearization of a nonlinear system. We assume that a MP is used as the PD model. The output

of PD can be written as

$$x(n) = \sum_{p=0}^{P} \sum_{q=0}^{Q} w_{pq} \Phi_{pq}[u(n)] \tag{3.12}$$

where $\Phi_{pq}[u(n)] = u(n-p)\,|u(n-p)|^q$, P and $Q+1$ are the memory depth and maximum nonlinearity order, respectively, and w_{pq} are the PD's coefficients. $u(n)$ and $x(n)$ are the input and output of PD, respectively, as shown in Fig. 3.10. (3.12) can be rewritten in vector form as

$$x(n) = \mathbf{u}^T(n)\mathbf{w}(n) \tag{3.13}$$

where $\mathbf{u}(n) = [\Phi_{00}[u(n)], \cdots, \Phi_{0Q}[u(n)], \cdots, \Phi_{pq}[u(n)], \cdots, \Phi_{P0}[u(n)], \cdots, \Phi_{PQ}[u(n)]]$ and $\mathbf{w}(n) = [w_{00}, \cdots, w_{0Q}, \cdots, w_{pq}, \cdots, w_{P0}, \cdots, w_{PQ}]$.

The weight update equation for the PD based on gradient method is given as

$$\mathbf{w}(n+1) = \mathbf{w}(n) - \frac{1}{2}\mu \boldsymbol{\nabla}(n) \tag{3.14}$$

where μ is the step size and $\boldsymbol{\nabla}(n)$ the instantaneous estimate of the gradient of $E\left\{|e(n)|^2\right\}$ with respect to the PD coefficients $\mathbf{w}(n)$.

$$E\left\{|e(n)|^2\right\} = E\left\{|d(n) - y(n)|^2\right\} \tag{3.15}$$

where $d(n)$ is the desired output of PA, $y(n)$ the actual output of PA.

$$\boldsymbol{\nabla}(n) = -2e^*(n)\boldsymbol{\psi}(n) \tag{3.16}$$

where $\boldsymbol{\psi}(n) = \frac{\partial y(n)}{\partial \mathbf{w}(n)}$. Thus, (3.14) can be rewritten as

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \mu e^*(n)\boldsymbol{\psi}(n). \tag{3.17}$$

Assume that a MP is also used as the PA model. It is given as

$$y(n) = \sum_{k=0}^{K} \sum_{l=0}^{L} a_{kl} \Phi_{kl}[x(n)] \tag{3.18}$$

where $y(n)$ is the output of PA, $x(n)$ is the input of PA (the output of PD), a_{kl}, $k = 0, \cdots, K$ and $l = 0, \cdots, L$ are the complex-valued coefficients of PA model, $\Phi_{kl}[x(n)] = x(n-k)\,|x(n-k)|^l$, K refers to the memory depth and $L+1$ the maximum nonlinearity order.

Because the output of PA $y(n)$ is a function of $\mathbf{x}(n)$ where $\mathbf{x}(n) = [x(n), x(n-1), \cdots, x(n-K)]$ and $x(n)$ is a function of $\mathbf{w}(n)$. $\boldsymbol{\psi}(n)$ can be written as

$$\boldsymbol{\psi}(n) = \frac{\partial y(n)}{\partial \mathbf{x}(n)} \frac{\partial \mathbf{x}(n)}{\partial \mathbf{w}(n)} = \sum_{k=0}^{K} \frac{\partial y(n)}{\partial x(n-k)} \frac{\partial x(n-k)}{\partial \mathbf{w}(n)}. \tag{3.19}$$

From (3.18) we obtain [20]

$$\frac{\partial y(n)}{\partial x(n-k)} = \sum_{l=0}^{L+1} \frac{l+2}{2} a_{kl} \left| x(n-k) \right|^l. \tag{3.20}$$

According to (3.13), and assuming that $\mathbf{w}(n) \approx \mathbf{w}(n-k)$[88, 22], we obtain

$$\frac{\partial x(n-k)}{\partial \mathbf{w}(n)} = \frac{\partial \mathbf{u}^T(n-k)\mathbf{w}(n-k)}{\partial \mathbf{w}(n)} \approx \mathbf{u}(n-k). \tag{3.21}$$

By (3.21), (3.20) and (3.19), we obtain

$$\boldsymbol{\psi}(n) = \sum_{k=0}^{K} \sum_{l=0}^{L+1} \frac{l+2}{2} a_{kl} \left| x(n-k) \right|^l \mathbf{u}(n-k). \tag{3.22}$$

Finally, substituting (3.22) in (3.17), we get

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \mu e^*(n) \sum_{k=0}^{K} \sum_{l=0}^{L+1} \frac{l+2}{2} a_{kl} \left| x(n-k) \right|^l \mathbf{u}(n-k) \tag{3.23}$$

3.4.2 Recursive least squares

RLS algorithm is a least squares method based on exponentially weighted error. Its cost function is defined as

$$\xi(n) = \sum_{i=1}^{n} \lambda^{n-i} \left| e(i) \right|^2 \tag{3.24}$$

where $0 < \lambda < 1$ is the exponential forgetting factor.

In the following, a Nonlinear Filtered-x RLS (NFxRLS) algorithm is presented. In [23], NFRLS is used in a predistorter based on direct learning architecture. The minimum can be achieved by differentiating (3.24) with respect to the weight vector $\mathbf{w}(n)$ and making the result be equal to zero, namely let $\frac{\partial \xi(n)}{\partial \mathbf{w}(n)} = 0$.

$$\frac{\partial \xi(n)}{\partial \mathbf{w}(n)} = -\sum_{i=1}^{n} 2\lambda^{n-i} \boldsymbol{\psi}(i) e^*(i) \tag{3.25}$$

where $\boldsymbol{\psi}(i)$ is given in (3.22). Similarly $e(i)$ can be derived as [23]

$$e(i) \approx d(i) - \mathbf{w}(n)\boldsymbol{\psi}(i) \tag{3.26}$$

where $1 \leq i \leq n$. Combining (3.25) and (3.26),

$$\begin{aligned} \frac{\partial \xi(n)}{\partial \mathbf{w}(n)} &\approx -\sum_{i=1}^{n} 2\lambda^{n-i}\boldsymbol{\psi}(i)[d(i) - \mathbf{w}(n)\boldsymbol{\psi}(i)]^* \\ &= -\sum_{i=1}^{n} 2\lambda^{n-i}\boldsymbol{\psi}(i)d^*(i) + \sum_{i=1}^{n} 2\lambda^{n-i}\boldsymbol{\psi}(i)\mathbf{w}^*(n)\boldsymbol{\psi}^H(i) \end{aligned} \tag{3.27}$$

Making (3.27) be equal to zero, we obtain

$$\sum_{i=1}^{n} \lambda^{n-i}\boldsymbol{\psi}(i)\mathbf{w}^*(n)\boldsymbol{\psi}^H(i) = \sum_{i=1}^{n} \lambda^{n-i}\boldsymbol{\psi}(i)d^*(i) \tag{3.28}$$

which can be rewritten as

$$\mathbf{R}_{\boldsymbol{\psi}}(n)\mathbf{w}(n) = \mathbf{r}_{\boldsymbol{\psi}}(n) \tag{3.29}$$

where $\mathbf{R}_{\boldsymbol{\psi}}(n) = \sum_{i=1}^{n} \lambda^{n-i}\boldsymbol{\psi}(i)\boldsymbol{\psi}^H(i)$ and $\mathbf{r}_{\boldsymbol{\psi}}(n) = \sum_{i=1}^{n} \lambda^{n-i}d(i)\boldsymbol{\psi}^*(i)$. We can obtain the Kalman gain as in [89]

$$\mathbf{k}(n) = \frac{\mathbf{P}(n-1)\boldsymbol{\psi}(n)}{\lambda + \boldsymbol{\psi}^H(n)\mathbf{P}(n-1)\boldsymbol{\psi}(n)} \tag{3.30}$$

where $\mathbf{P}(n) = \frac{1}{\lambda}[\mathbf{P}(n-1) - \mathbf{k}(n)\boldsymbol{\psi}^H(n)\mathbf{P}(n-1)]$. The resulting weight update function is

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \mu e^*(n)\mathbf{k}(n) \tag{3.31}$$

3.5 Comparison of ILA and DLA

In this section, ILA and DLA are compared by some simulations. Firstly, their stabilities are compared in terms of the condition number of a matrix closely related to the architecture stability. Secondly, their modeling performances are analyzed by considering different number of bits of Analog-to-Digital Converter (ADC) and varying noise at the output of PA. Finally, two DPDs based ILA and DLA are compared in terms of the ACPR and EVM of linearized outputs of PA.

3.5.1 Condition number of matrix

The principles of ILA and DLA are presented in section 3.3. They both need to identify a pre-assumed model (MP model) with the measured input and output of PA. LS method is used in the model identification. From (3.6), it can be seen that the inverse of matrix $\mathbf{Z}^H\mathbf{Z}$ is required in LS algorithm. When solving the inverse problem, the condition number is usually used to measure how sensitive the solution is to changes or errors in the input. A problem with a low condition number is said to be well-conditioned, while a problem with a high condition number is said to be ill-conditioned.

In this subsection, the condition number of matrix $\mathbf{Z}^H\mathbf{Z}$ is considered to measure the stability of the two DPD architectures. In ILA, matrix $\mathbf{Z}$ is related with the output signal of PA. While in DLA, it is related with the input signal of PA. A simulation is realized to compare these two DPD architectures. In this simulation, a Wiener model is used as PA model. It is implemented as a 3-tap FIR filter with coefficients [0.7692, 0.1538, 0.0769] [90], followed by a Saleh model. Saleh model is described as

$$y(n) = \frac{\alpha_a|v(n)|}{1+\beta_a|v(n)|^2} e^{j\angle[v(n)+\frac{\alpha_\varphi|v(n)|^2}{1+\beta_\varphi|v(n)|^2}]} \tag{3.32}$$

where $y(n)$ and $v(n)$ are the output and input of Saleh model, respectively, and $\alpha_a = 20$, $\beta_a = 2.2$, $\alpha_\varphi = 2$, $\beta_\varphi = 1$ [32]. The input signal of PA is a 16QAM signal with 3.84 MHz bandwidth. The measured sequence has 1000 symbols (8000 samples). The ideal gain of this PA model is about 26 dB. The input power at 1 dB compression point is about -1 dBm.

The condition number of $\mathbf{Z}^H\mathbf{Z}$ is affected by $\mathbf{Z}$ which depends on the memory depth, nonlinearity order and value of each element (varies with the input signal power). In the first simulation, we assume that the memory depth is 3 ($K = P = 3$) and the nonlinearity order is 5 ($L = Q = 4$). The signal input power varies from -20 dBm to 5 dBm. The condition numbers of $\mathbf{Z}^H\mathbf{Z}$ of the two DPD architectures are compared. The results are shown in Fig. 3.11. It can be seen that the condition number with ILA is larger than that with DLA when the input power is close to the input saturation power. It indicates that the stability of DLA is more robust than that of ILA.

In the second simulation, we assume that the input signal power is 0 dBm. The condition number varies with the nonlinearity order and memory depth. The nonlinearity order varies from 3 to 15. The memory depth varies from 0 to 3. The comparison of two DPD architectures is shown in Fig. 3.12.

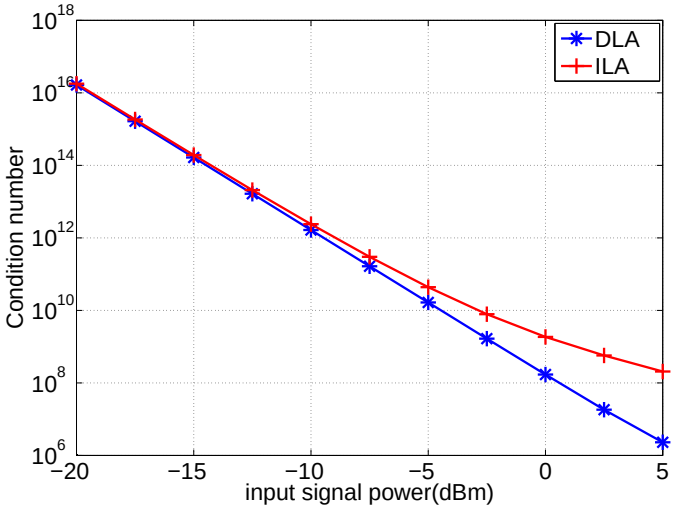


Figure 3.11: Comparison of condition number with different input signal power

From Fig. 3.11 and Fig. 3.12, it can be seen that the condition number of $\mathbf{Z}^H\mathbf{Z}$ in DLA is always lower than that in ILA. It indicates that DLA is more robust than ILA when changes or errors are present in the measurement, such as the noise and quantization error of ADC. To further validate this result, the influences of the noise and quantization error of ADC on the modeling performance are analyzed in next section.

3.5.2 Modeling performance

In practical measurement, an ADC is required to realize the signal acquisition. The number of bits of ADC will affect the accuracy of the signal acquisition and then affect the performance of the model's identification. Additionally, the measurement at the output of PA is noisy in practical application. The noise also affects the performance of the model's identification. In the third simulation, considering the number of bits of ADC in signal acquisition and the noise at the output of PA, simulations are realized to compare these two DPD architectures.

For analyzing the influence of the noise at the output of PA, it assumes that white Gaussian noise is present at the output of PA as shown in Fig. 3.9 and Fig. 3.10. An original 16QAM signal $x'(n)$ (bandwidth 3.84 MHz, power 0dBm, sequence length 8000 samples) is given as the PA input. And the output $y'(n)$ of PA is obtained by the Wiener model. The white Gaussian noise is added in the output signal $y'(n)$ to get the noisy output $y'_{noise}(n)$. The input $x'(n)$ and output $y'_{noise}(n)$ of PA are used for the

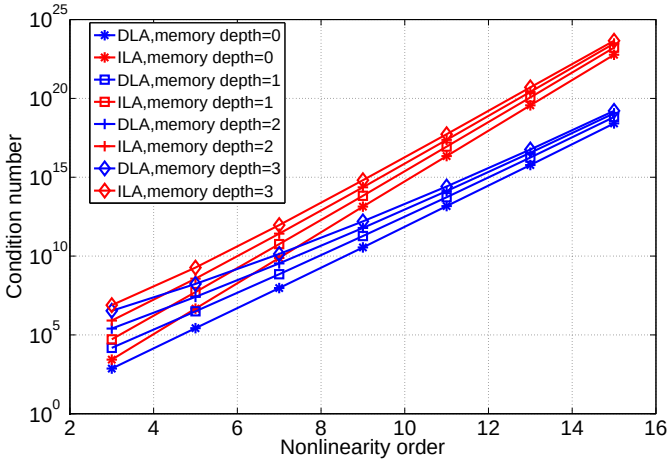


Figure 3.12: Comparison of condition number with different nonlinearity order and memory depth

identifications of (3.4) and (3.7) as shown in Fig. 3.13. In the identifications of (3.4) and (3.7), $K = P = 3$ (the memory depth = 3) and $L = Q = 4$ (the nonlinearity order = 5), respectively. The coefficients of identified MP model with DLA is different with those with ILA.

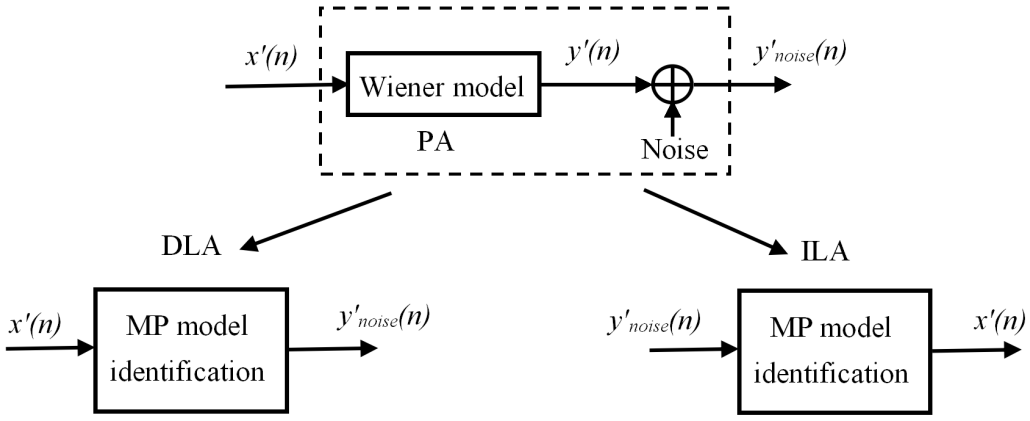


Figure 3.13: Identification with DLA and ILA

After the model identification, another 16QAM signal $x''(n)$ is used to test the identified model. The signal $x''(n)$ passes the PA as the input, then the corresponding output $y''(n)$ is obtained. For ILA, the signal $y''(n)$ is the input of the postdistorter, its output $z''_p(n)$ is obtained by the identified model (3.4). The Normalized Mean Squared

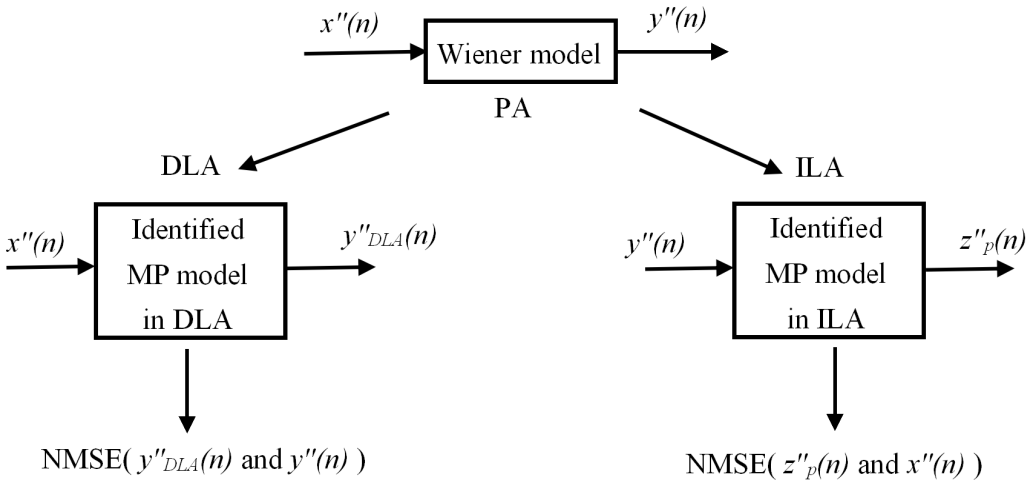


Figure 3.14: Modeling performance with DLA and ILA

Error (NMSE) between $z_p''(n)$ and $x''(n)$ is treated as the criterion of ILA modeling accuracy shown in Fig. 3.14. NMSE is defined by

$$NMSE = 10log_{10}\frac{\sum_{n=1}^{N}|x_1(n)-x_2(n)|^2}{\sum_{n=1}^{N}|x_1(n)|^2} \tag{3.33}$$

where N is the number of sample sequence, $x_1(n)$ and $x_2(n)$ denote the actual signal and modeling signal, respectively.

For DLA, the signal $x''(n)$ is the input of the identified model (3.7), the corresponding output $y''_{DLA}(n)$ is obtained. The NMSE between $y''_{DLA}(n)$ and $y''(n)$ is treated as the criterion of DLA modeling accuracy shown in Fig. 3.14.. Fig. 3.15 shows the performance of ILA and DLA in terms of NMSE when the noise is present. It can be seen that the modeling accuracy of ILA and DLA both improves as the noise power decreases. When the noise is high, the performance of DLA is better than that of ILA. When the power of noise is low (nearly ideal case), the performance of DLA and ILA is similar. In practical applications, the noise can not be ignored. DLA is more robust than ILA when the noise is present.

For the influence of the number of bits of ADC, an ADC is assumed for the signal acquisition. The signals $x'(n)$ and $y'(n)$ pass the ADC and are denoted by $x'_{ADC}(n)$ and $y'_{ADC}(n)$, respectively. The signals $x'_{ADC}(n)$ and $y'_{ADC}(n)$ are used for the identification of (3.4) and (3.7). After the model identification, the same signal $x''(n)$ is used to test the identified model. The comparison principle considering the number of bits

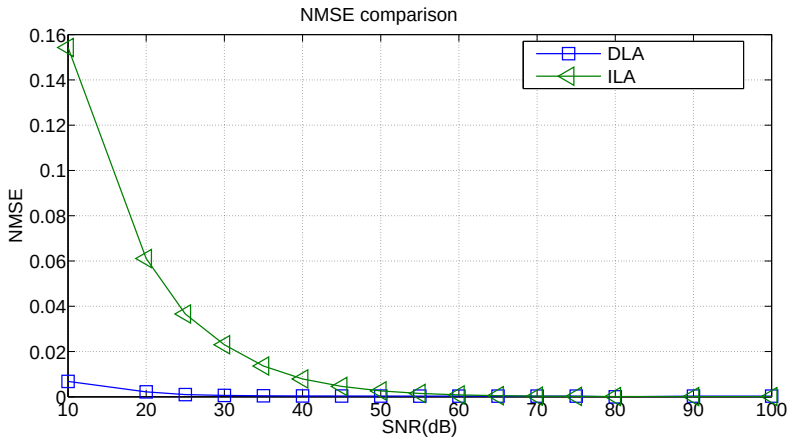


Figure 3.15: Comparison of ILA and DLA when considering noise

of ADC is the same as that considering the noise. Fig. 3.16 shows the performance of ILA and DLA in terms of NMSE with varying number of bits of ADC. It can be seen that the modeling accuracy of ILA and DLA both improves as the number of bits of ADC increases. The modeling accuracy of DLA is always better than that of ILA.

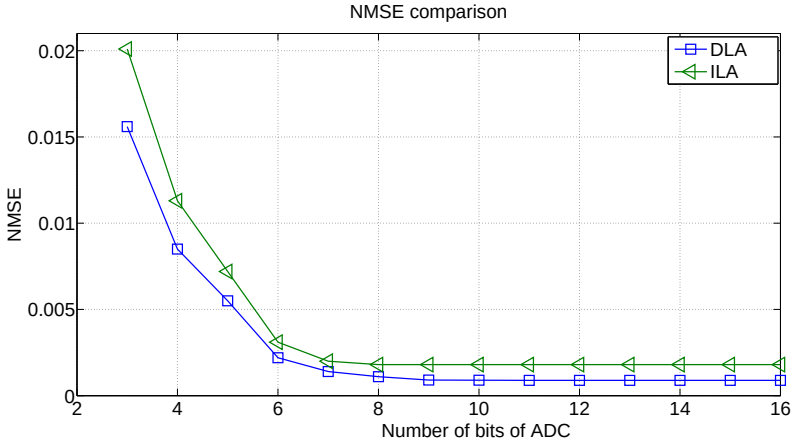


Figure 3.16: Comparison of ILA and DLA when considering number of bits of ADC

According to the above analysis, it further verifies that DLA is more robust than ILA.

3.5.3 Linearization performance

In this section, two DPDs based on different learning architectures are compared. The DPD proposed in [12] is taken as the example of ILA. And the DPD proposed in [1] is taken as the example of DLA. In this simulation, the number of bits of ADC and

the noise both are taken into account in the process of identification. We assume that the number of bits of ADC is 8, the SNR at the output of PA is 35 dB. The average power of the input signal varies from -13 dBm to -3 dBm. The input signal, which has an average power of -3 dBm, includes many samples entering the nonlinear region of PA model. Fig. 3.17 illustrates the distribution of samples of the input signal when the average power is -3 dBm. It shows that the power of about 40% of the input samples is higher than the input power (-1 dBm) at 1 dB compression point of the PA model where the linearization is needed.

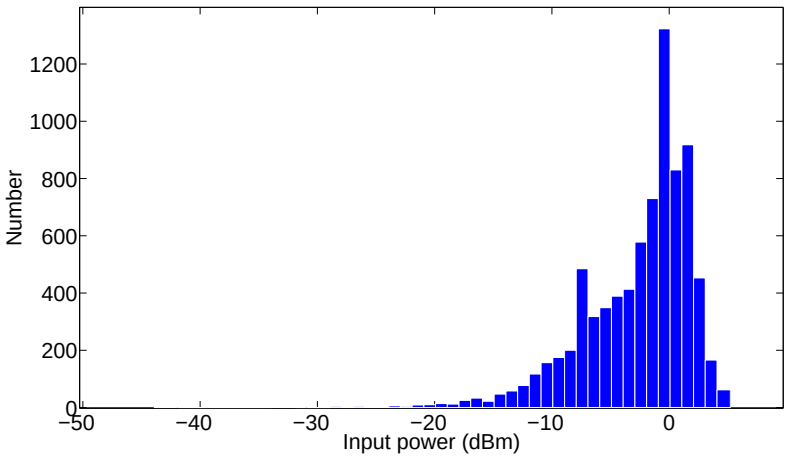


Figure 3.17: Histogram distribution of the input signal when the average power is -3 dBm

The original input $x'(n)$ and output $y'(n)$ of PA firstly pass the ADC, then the noise is added on the output signal of PA. The processed input and output signals are used for the identification of (3.4) and (3.7). For ILA, the identified postdistorter model is directly placed in front of PA as the predistorter [12]. For DLA, the predistorter is determined by inverting the identified PA model [1].

The linearized outputs of two different learning architectures are obtained by the DPD solutions mentioned above. The ACPR and EVM of linearized outputs of two architectures are calculated, respectively. Fig. 3.18 shows the evolution of ACPR of linearized outputs with varying input power from -13 dBm to -3 dBm. And Fig. 3.19 shows the evolution of EVM of linearized outputs with varying input power. It can be seen that the linearization performance of DLA is more robust than that of ILA when considering the number of bits of ADC and the noise at the output of PA. In this thesis, we mainly focus on the DPD based on DLA.

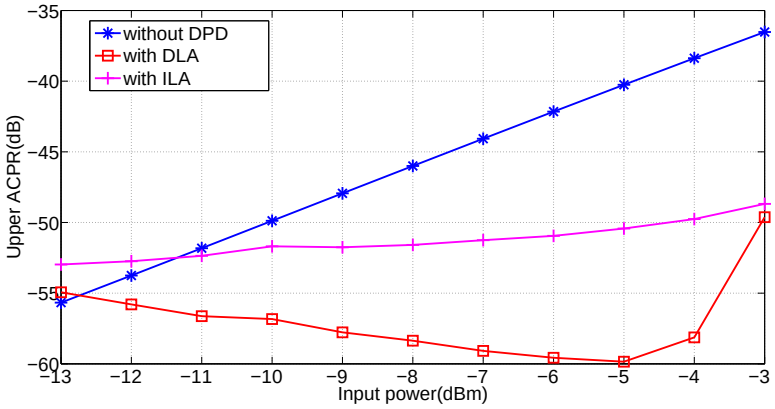


Figure 3.18: ACPR comparison of DPDs based different learning architectures

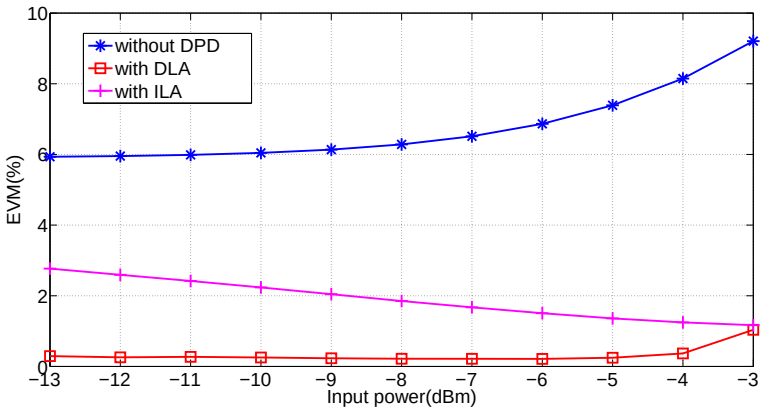


Figure 3.19: EVM comparison of DPDs based different learning architectures

4

Proposed digital predistortion methods

In this chapter, our proposed DPD methods are described in detail. In section 4.1, the behavior modeling of PA is made using MP model. The nonlinearity order, memory depth and coefficients of MP model are determined. In section 4.2, a DPD with root-finding procedure in [1] is presented. This DPD method has a good linearization performance, but it is very time-consuming. Our proposed DPD methods are based on this root-finding-based DPD. In section 4.3, the proposed MP/LUT-based DPD, MP/LILUT-based DPD and MP/QILUT-based DPD are depicted respectively. And their simulation results are summarized. In section 4.4, the proposed NN-based DPD is described and its simulation results are given.

4.1 Power amplifier modeling

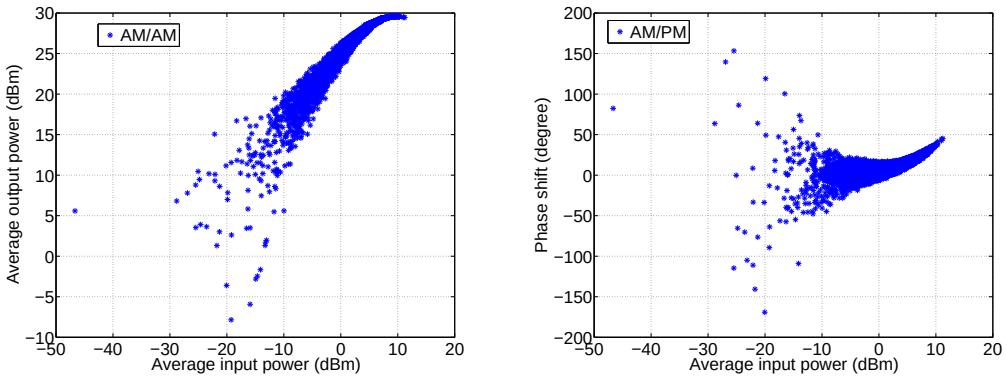
In this chapter, we consider that the Wiener model mentioned in section 3.5.1 is used to simulate the PA. To remind this model, it is repeated as follows

$$v(n) = h_0 x(n) + h_1 x(n-1) + h_2 x(n-2) \tag{4.1}$$

$$y(n) = \frac{\alpha_a |v(n)|}{1 + \beta_a |v(n)|^2} e^{j\angle[v(n) + \frac{\alpha_\varphi |v(n)|^2}{1+\beta_\varphi |v(n)|^2}]} \tag{4.2}$$

where $x(n)$ and $y(n)$ are the input and output of the simulated PA, respectively, $v(n)$ the input of Saleh model (the output of the filter), and $h_0 = 0.7692$, $h_1 = 0.1538$, $h_2 = 0.0769$, $\alpha_a = 20$, $\beta_a = 2.2$, $\alpha_\varphi = 2$, $\beta_\varphi = 1$. A 16QAM signal with 3.84 MHz bandwidth is used to test and analyze the characteristics of this simulated PA. The test sequence has 1000 symbols (8 samples/symbol).

Fig. 4.1 shows AM/AM and AM/PM characteristics based on the instantaneous samples of input and output of the simulated PA. It can be seen that the simulated PA suffers from the nonlinearity and memory effects. Fig. 4.2 presents the gain performance of the simulated PA with varying average input power. The average input power varies from -30 dBm to 5 dBm. The gain in linear region is about 26 dB. The average input powers at 1 dB compression point and at 3 dB compression point are about -1 dBm and 4.3 dBm, respectively.



(a) AM/AM characteristic of the instantaneous samples of simulated PA

(b) AM/PM characteristic of the instantaneous samples of simulated PA

Figure 4.1: AM/AM and AM/PM characteristics of the instantaneous samples of simulated PA

As mentioned before, the first step for realizing DPD with DLA is to obtain a mathematical model which can accurately describe the nonlinear behavior of PA. Several PA models have been described in section 2.3.3. Because the memory effects can no longer be ignored in modern wide-band wireless communication system, it is necessary to adopt a model with memory effects. In this thesis, MP model is adopted to model the PA. MP model has lower complexity and can closely mimic the nonlinear behavior of PA with memory effects. It offers a good trade-off between computational complexity and modeling accuracy. In addition, the even orders of polynomial have

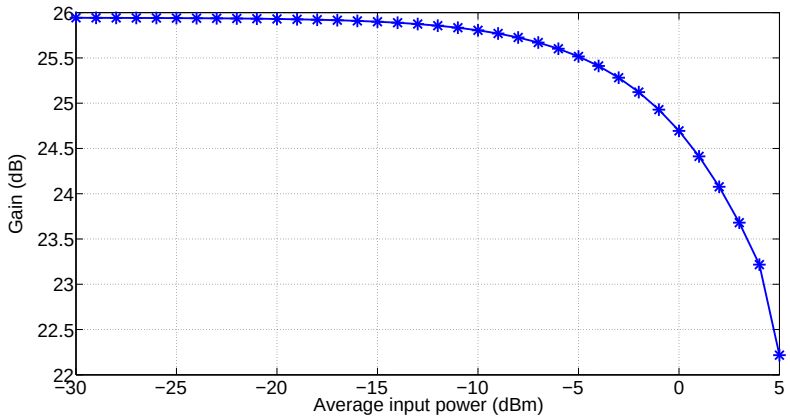


Figure 4.2: Gain characteristic of simulated PA

very little effect on the final amplified and filtered signal, the odd-order MP model is widely used [91, 32]. The odd-order MP model is described as

$$y(n) = \sum_{p=0}^{P} \sum_{q=0}^{Q} c_{pq} \Phi_{pq}[x(n)] \tag{4.3}$$

where $\Phi_{pq}[x(n)] = x(n-p)\,|x(n-p)|^{2q}$, P and $2Q+1$ are the memory depth and highest nonlinearity order (only odd orders are considered), respectively, and c_{pq} are the model's coefficients. $x(n)$ and $y(n)$ are the input and output of the PA, respectively. In the following, we will observe the modeling performance of MP model.

In MP model, three parameters need to be determined: P, Q and c_{pq}. Generally, a set of estimated P and Q is given, then c_{pq} are extracted by LS method as described in section 3.3.1. In order to estimate the modeling performance of MP model, we compare the outputs of the identified MP model and the simulated PA. Fig. 4.3 shows NMSE between the output of the identified MP model and that of the simulated PA for different memory depths and nonlinearity orders. The memory depth varies from 1 to 10. The nonlinearity order varies from 1 to 15 (only odd order). The smallest NMSE (-50.17 dB) is achieved with the memory depth 3 and nonlinearity order 9. To observe the results clearly, the curves of nonlinearity order 7, 9 11 and 13 are shown in Fig. 4.4. In the later section, the memory depth is determined to be 3 (P=3) and the nonlinearity order is determined to be 9 (Q=4) for MP model.

To further assess the accuracy of MP model, some tests are conducted. The in-phase and quadrature components of the output signal of MP model (P=3, Q=4) are compared with that of the simulated PA. Fig. 4.5 presents the first 100 samples of the

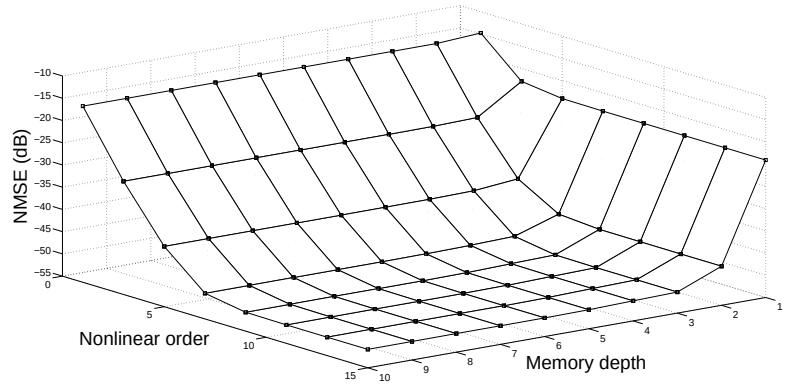


Figure 4.3: NMSE between the output of the identified MP model and that of the simulated PA

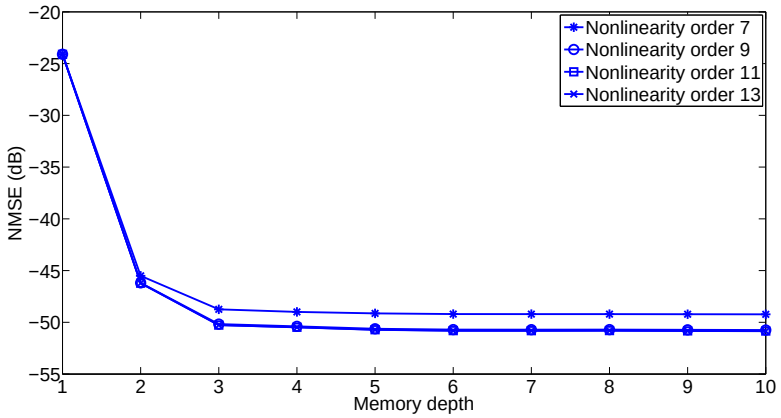


Figure 4.4: NMSE with nonlinearity order 7, 9, 11 and 13

output of MP model and that of the simulated PA. It is observed that MP model has a good fitting of the simulated PA. In addition, Fig. 4.6 compares AM/AM and AM/PM characteristics of MP model and the simulated PA. It is clear that MP model has a good ability to mimic the nonlinear behavior of the PA with memory effects.

4.2 MP based DPD with root-finding method

In this section, the MP DPD with root-finding procedure [1] is presented. The ideal cascaded PD-PA system can be expressed as

$$y(n) = \sum_{p=0}^{P} \sum_{q=0}^{Q} c_{pq} \Phi_{pq}[x(n)] = G_0 u(n) \tag{4.4}$$

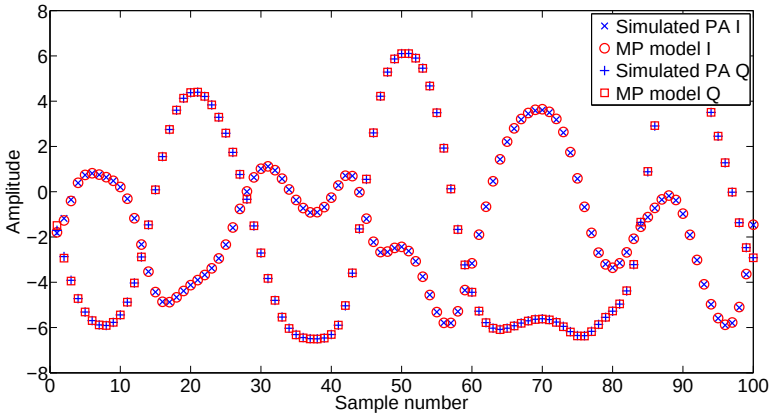
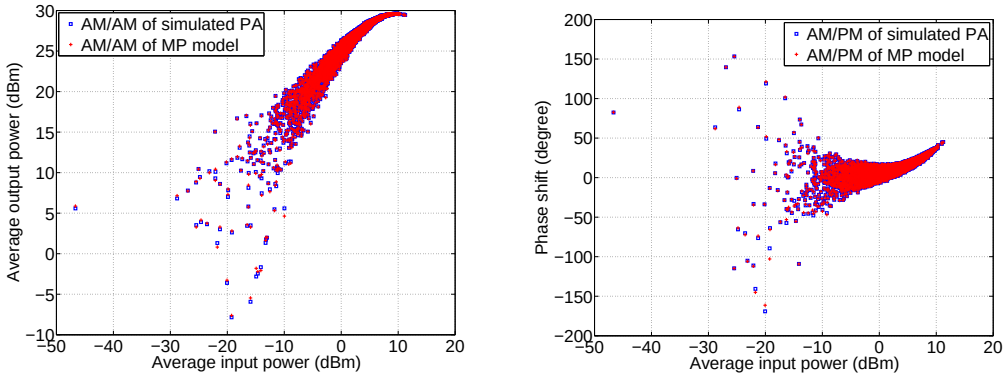


Figure 4.5: In-phase and quadrature components of the output of MP model and simulated PA



(a) Comparison of AM/AM characteristics of MP model and simulated PA

(b) Comparison of AM/PM characteristics of MP model and simulated PA

Figure 4.6: Comparison of AM/AM and AM/PM characteristics of MP model and simulated PA

where $\Phi_{pq}[x(n)] = x(n-p)\,|x(n-p)|^{2q}$, P and $2Q+1$ the memory depth and highest nonlinearity order, respectively, c_{pq} are the coefficients of PA model and G_0 is the ideal gain. $u(n)$ denotes the input of PD-PA system, $x(n)$ the predistorted signal (namely the input of PA) and $y(n)$ the output of PD-PA system, as shown in Fig. 3.10. According to the memory behavior of PA, the output $y(n)$ of PA can be represented as the sum of two parts: the static part $s(n)$ depending only on the current input sample (p=0), and

the dynamic part $d(n)$ which depends on the previous input samples (p is from 1 to P).

$$\begin{cases} y(n) = s(n) + d(n) \\ s(n) = \sum_{q=0}^{Q} c_{0q} \Phi_{0q}[x(n)] \\ d(n) = \sum_{p=1}^{P} \sum_{q=0}^{Q} c_{pq} \Phi_{pq}[x(n)] \end{cases} \tag{4.5}$$

From (4.4) and (4.5), $s(n)$ can be rewritten as

$$s(n) = e^{j\angle x(n)} \sum_{q=0}^{Q} c_{0q} |x(n)|^{2q+1} = G_0 u(n) - d(n) \tag{4.6}$$

where $\angle x(n)$ and $|x(n)|$ are the phase and amplitude of the predistorted signal $x(n)$, respectively. By definition, $d(n)$ depends only on the previous samples, the right-hand side of (4.6) and coefficients c_{0q} are known at instant n. $|x(n)|$ can be calculated by taking the absolute value.

$$\left| \sum_{q=0}^{Q} c_{0q} |x(n)|^{2q+1} \right| = |G_0 u(n) - d(n)| \tag{4.7}$$

(4.7) is a high-order nonlinear equation, the amplitude $|x(n)|$ of the predistorted signal can be found by a root-finding procedure [1]. The corresponding phase $\angle x(n)$ is then calculated by

$$\angle x(n) = arg \left\{ \frac{G_0 u(n) - d(n)}{\sum_{q=0}^{Q} c_{0q} |x(n)|^{2q+1}} \right\}. \tag{4.8}$$

Finally, the predistorted signal $x(n)$ is given by

$$x(n) = |x(n)| e^{j\angle x(n)}. \tag{4.9}$$

The DPD mentioned above can achieve good linearization performance, but the calculation of the predistorted signal is time-consuming. The amplitude and phase of the predistorted signal are calculated by (4.7) and (4.8), respectively. The calculations of these two equations are both complicated and time-consuming, especially the root-finding procedure in (4.7). For convenience, the MP DPD with root-finding procedure [1] is denoted as MP/root-finding-based DPD in the later section.

4.3 Proposed DPDs with combination of MP and LUT

In this section, three proposed baseband DPDs with DLA are presented in detail. In the first DPD solution, (4.7) and (4.8) both are replaced by using the conventional non-interpolated LUT, denoted as MP/LUT-based DPD in the later section. In the second DPD solution, the linear interpolation is introduced in LUT based on MP/LUT-based DPD, denoted as MP/LILUT-based DPD. In the third DPD solution, the quadratic interpolation is introduced in LUT based on MP/LUT-based DPD, denoted as MP/QILUT-based DPD.

4.3.1 MP and LUT based DPD

In [2], the amplitude calculation of $|x(n)|$ in (4.7) is replaced by a conventional non-interpolated LUT, but the phase calculation still depends on (4.8). For our proposed MP/LUT-based DPD, (4.7) and (4.8) both are considered and replaced by using LUT.

In this subsection, MP/LUT-based DPD is proposed. (4.8) can be rewritten as

$$\angle x(n) = arg\{G_0 u(n) - d(n)\} - arg\{\sum_{q=0}^{Q} c_{0q}|x(n)|^{2q+1}\}. \tag{4.10}$$

Before performing the DPD, a LUT including the information of the amplitude and phase of the predistorted signal should be constructed. Firstly, the dynamic range of the amplitude $|x(n)|$ of the predistorted signal needs to be estimated. Depending on the analysis of Fig. 3.2 and Fig. 3.3, the maximum value of $|x(n)|$ is set to be the amplitude r_{in_sat} which represents the input amplitude of the beginning point of the saturation region. Suppose that a sample is near the saturation region in which the amplitude of the predistorted signal should be $r_{pd} = r_{in\_sat}$, we find the corresponding point r_{in} according to the characteristic of the PA. Then define that $\frac{r_{pd}}{r_{in}} = \alpha$. The maximum dynamic range of $|x(n)|$ is defined as

$$|x(n)| \in [0, min(\alpha \cdot max(|u(n)|), r_{in\_sat})] \tag{4.11}$$

where $u(n)$ is the input of PD-PA system. α depends on the characteristic of PA. In this simulated PA, the value of α is around 2. Secondly, the dynamic range of $|x(n)|$ is decomposed into M intervals with equal length (the length of each interval is equal to Δx), denoted by $R(m)$. The LUT is shown in Table 4.1. In this table,

$$R(m) = m\Delta x \tag{4.12}$$

$$E(m) = \left| \sum_{q=0}^{Q} c_{0q} R(m)^{2q+1} \right| \tag{4.13}$$

$$\theta(m) = arg\{\sum_{q=0}^{Q} c_{0q} R(m)^{2q+1}\} \tag{4.14}$$

where $E(m)$ is the LUT input and $\theta(m)$ the phase of the term $\sum_{q=0}^{Q} c_{0q} R(m)^{2q+1}$.

LUT input	LUT output	
	Amplitude	Phase
$E(0)$	$R(0)$	$\theta(0)$
...	...	...
$E(m)$	$R(m)$	$\theta(m)$
...	...	...
$E(M-1)$	$R(M-1)$	$\theta(M-1)$
$E(M)$	$R(M)$	$\theta(M)$

Table 4.1: Values of LUT in MP/LUT-based DPD

After the LUT is constructed, we use the proposed MP/LUT-based DPD algorithm described in Algorithm 1. The proposed MP/LUT-based DPD can significantly improve the time efficiency of DPD process than the root-finding method, but the quantification error is large when the table size is small. Its linearization performance is proportional to the table size. In order to achieve a good linearization performance, the table size must be sufficiently large. Therefore, some interpolation techniques [92, 93] are introduced in LUT for reducing the table size.

4.3.2 MP and linear-interpolated LUT based DPD

The linear interpolation is widely used due to its simplicity. In [74, 75], linear-interpolated LUTs have been used to improve the performance of digital predistorter. The principle of linear interpolation is shown in Fig. 4.7. Suppose that a point (a_o, b_o) is between two points (a_i, b_i) and (a_{i+1}, b_{i+1}). a_o is known, while b_o needs to be calcu-

Algorithm 1 Proposed MP/LUT-based DPD

1. Initialize $n = 0$, $d(0) = 0$.
2. Begin loop
   {
3. Calculate the value of $G_0 u(n) - d(n)$, denoted by $s(n)$.
4. Find a value $E(m)$ from Table 4.1, which is the closest to $|s(n)|$. m is determined.
5. Find the corresponding value of $|x(n)|$ by
   $$|x(n)| = LUT_{amp}\{|s(n)|\} = R(m)$$
   where LUT_{amp} refers to the amplitude output of LUT.
6. Calculate the corresponding phase $\angle x(n)$ by
   $$\angle x(n) = arg\{s(n)\} - LUT_{pha}\{|s(n)|\} = arg\{s(n)\} - \theta(m)$$
   where LUT_{pha} refers to the phase output of LUT.
7. Calculate $x(n) = |x(n)|e^{j\angle x(n)}$.
8. Calculate $n = n + 1$.
9. Calculate
   $$d(n) = \sum_{p=1}^{P} \sum_{q=0}^{Q} c_{pq} x(n-p) \, |x(n-p)|^{2q}.$$
10. } Goto loop

lated. If the curve between the points (a_i, b_i) and (a_{i+1}, b_{i+1}) is approximated as linear, there is:

$$\frac{b_o - b_i}{a_o - a_i} = \frac{b_{i+1} - b_i}{a_{i+1} - a_i} \tag{4.15}$$

Then b_o can be calculated as:

$$b_o = b_i + (a_o - a_i)s_o \tag{4.16}$$

where $s_o = \frac{b_{i+1} - b_i}{a_{i+1} - a_i}$ is the slope of the curve. In case of small table size (b_{i+1} and b_i are relatively far), linear-interpolated LUT provides a more appropriate corresponding value than the conventional non-interpolated LUT, but at the cost of some small additional computations due to linear interpolation. To further reduce the mathematical computations, the slope of each interval is precomputed and stored in LUT.

The Linear-Interpolated LUT (LILUT) is shown in Table 4.2. In the table,

$$R(m) = m\Delta x \tag{4.17}$$

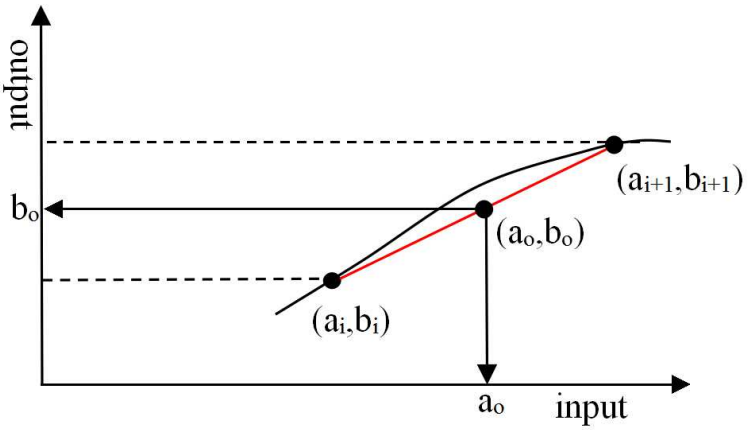


Figure 4.7: Principle of linear interpolation

$$E(m) = \left| \sum_{q=0}^{Q} c_{0q} R(m)^{2q+1} \right| \tag{4.18}$$

$$K_R(m) = \frac{R(m+1) - R(m)}{E(m+1) - E(m)} \quad (0 \leq m \leq M-1) \tag{4.19}$$

$$\theta(m) = arg\{\sum_{q=0}^{Q} c_{0q} R(m)^{2q+1}\} \tag{4.20}$$

$$K_\theta(m) = \frac{\theta(m+1) - \theta(m)}{E(m+1) - E(m)} \quad (0 \leq m \leq M-1) \tag{4.21}$$

where $K_R(m)$ is the slope between the points $(E(m), R(m))$ and $(E(m+1), R(m+1))$, and $K_\theta(m)$ the slope between the points $(E(m), \theta(m))$ and $(E(m+1), \theta(m+1))$. The proposed MP/LILUT-based DPD algorithm is described in Algorithm 2.

4.3.3 MP and quadratic-interpolated LUT based DPD

The quadratic interpolation is another interpolation technique, where the interpolated point is calculated depending on the nearest three points. Compared with the linear interpolation, the quadratic interpolation is more accurate. Namely, there is still room for improving the interpolation accuracy. The principle of quadratic interpolation is described as follows. Suppose that there are three adjacent discrete points

LUT input	LUT output			
	Amplitude	Slope for amplitude	Phase	Slope for phase
$E(0)$	$R(0)$	$K_R(0)$	$\theta(0)$	$K_\theta(0)$
...	...	...	...	...
$E(m)$	$R(m)$	$K_R(m)$	$\theta(m)$	$K_\theta(m)$
...	...	...	...	...
$E(M-1)$	$R(M-1)$	$K_R(M-1)$	$\theta(M-1)$	$K_\theta(M-1)$
$E(M)$	$R(M)$	-	$\theta(M)$	-

Table 4.2: Values of LUT in MP/LILUT-based DPD

Algorithm 2 Proposed MP/LILUT-based DPD

1. Initialize $n = 0$, $d(0) = 0$.
2. Begin loop
   {
3. Calculate the value of $G_0 u(n) - d(n)$, denoted by $s(n)$.
4. Find two adjacent values $E(m)$ and $E(m+1)$ from Table 4.2, which are the two closest to $|s(n)|$. m is determined.
5. Calculate the corresponding value of $|x(n)|$ by
   $$|x(n)| = LILUT_{amp}\{|s(n)|\} = R(m) + [|s(n)| - E(m)]K_R(m)$$
   where $LILUT_{amp}$ refers to the amplitude output of LILUT.
6. Calculate the corresponding phase $\angle x(n)$ by
   $$\angle x(n) = LILUT_{pha}\{|s(n)|\} = arg\{s(n)\} - \{\theta(m) + [|s(n)| - E(m)]K_\theta(m)\}$$
   where $LILUT_{pha}$ refers to the phase output of LILUT.
7. Calculate $x(n) = |x(n)|e^{j\angle x(n)}$.
8. Calculate $n = n + 1$.
9. Calculate
   $$d(n) = \sum_{p=1}^{P} \sum_{q=0}^{Q} c_{pq} x(n-p) \, |x(n-p)|^{2q}.$$
10. } Goto loop

(a_{i-1}, b_{i-1}), (a_i, b_i) and (a_{i+1}, b_{i+1}). The quadratic interpolation function is

$$b' = f_{i-1}(a')b_{i-1} + f_i(a')b_i + f_{i+1}(a')b_{i+1} \tag{4.22}$$

where (a', b') is the interpolated point, a' is known and b' needs to be calculated. $f_{i-1}(a')$, $f_i(a')$ and $f_{i+1}(a')$ are the weighting functions [64].

$$f_{i-1}(a') = \frac{(a' - a_i)(a' - a_{i+1})}{(a_{i-1} - a_i)(a_{i-1} - a_{i+1})} \tag{4.23}$$

$$f_i(a') = \frac{(a' - a_{i-1})(a' - a_{i+1})}{(a_i - a_{i-1})(a_i - a_{i+1})} \tag{4.24}$$

$$f_{i+1}(a') = \frac{(a' - a_{i-1})(a' - a_i)}{(a_{i+1} - a_{i-1})(a_{i+1} - a_i)} \tag{4.25}$$

An example for the comparison of linear interpolation and quadratic interpolation is shown in Fig. 4.8. Because there is more referenced information in quadratic interpolation than in linear interpolation, the curve with quadratic interpolation is smoother than that with linear interpolation. But the quadratic interpolation is more complicated and requires more computations than the linear interpolation.

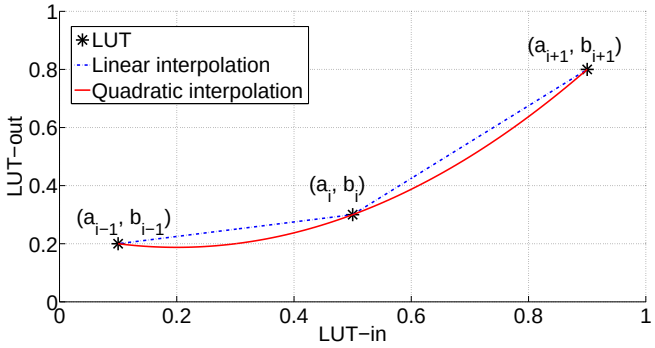


Figure 4.8: Comparison of linear interpolation and quadratic interpolation (LUT consists of 3 entries)

The contents of LUT in MP/QILUT-based DPD is the same as that in MP/LUT-based DPD, as shown in Table 4.1. The proposed MP/QILUT-based DPD algorithm is described in Algorithm 3.

4.3.4 Simulation results

In this subsection, the simulation results of MP/root-finding-based DPD, MP/LUT-based DPD, MP/LILUT-based DPD and MP/QILUT-based DPD are presented and compared. In the simulation, the Wiener model mentioned in section 3.5.1 is used as the simulated PA. The odd-order MP model (4.3) is used to model the simulated PA. For MP model, the memory depth is 3 (P=3) and the highest nonlinearity order is 9

Algorithm 3 Proposed MP/QILUT-based DPD

1. Initialize $n = 0$, $d(0) = 0$.
2. Begin loop
   {
3. Calculate the value of $G_0u(n) - d(n)$, denoted by $s(n)$.
4. Find three adjacent values $E(m-1)$, $E(m)$ and $E(m+1)$ from Table 4.1, which are the three closest to $|s(n)|$. m is determined.
5. Calculate the weighting functions $f_{m-1}(|s(n)|)$, $f_m(|s(n)|)$ and $f_{m+1}(|s(n)|)$ by
$$f_{m-1}(|s(n)|) = \frac{(|s(n)|-E(m))(|s(n)|-E(m+1))}{(E(m-1)-E(m))(E(m-1)-E(m+1))}$$
$$f_m(|s(n)|) = \frac{(|s(n)|-E(m-1))(|s(n)|-E(m+1))}{(E(m)-E(m-1))(E(m)-E(m+1))}$$
$$f_{m+1}(|s(n)|) = \frac{(|s(n)|-E(m-1))(|s(n)|-E(m))}{(E(m+1)-E(m-1))(E(m+1)-E(m))}$$
6. Calculate the corresponding value of $|x(n)|$ by
$$|x(n)| = QILUT_{amp}\{|s(n)|\} = f_{m-1}(|s(n)|)R(m-1) + f_m(|s(n)|)R(m) + f_{m+1}(|s(n)|)R(m+1)$$
where $QILUT_{amp}$ refers to the amplitude output of QILUT.
7. Calculate the corresponding phase $\angle x(n)$ by
$$\angle x(n) = QILUT_{pha}\{|s(n)|\} = arg\{s(n)\} - \{f_{m-1}(|s(n)|)\theta(m-1) + f_m(|s(n)|)\theta(m) + f_{m+1}(|s(n)|)\theta(m+1)\}$$
where $QILUT_{pha}$ refers to the phase output of QILUT.
8. Calculate $x(n) = |x(n)|e^{j\angle x(n)}$.
9. Calculate $n = n + 1$.
10. Calculate
$$d(n) = \sum_{p=1}^{P}\sum_{q=0}^{Q} c_{pq}x(n-p)\,|x(n-p)|^{2q}.$$
11. } Goto loop

(Q=4). The tested input signal $u(n)$ is a 16QAM modulated signal with 3.84 MHz bandwidth, which is filtered by a raised cosine filter. The roll-off factor of filter is 0.22. The input sequence has 1000 symbols (8 samples/symbol). The average power of input signal is set to be -4 dBm.

The proposed DPD solutions are compared with MP/root-finding-based DPD in terms of NMSE between the predistorted signal by the proposed DPDs and that by MP/root-finding-based DPD, NMSE between the desired linearized PA output and actual PA output, ACPR of PA output, EVM of PA output, and the consumed time of

DPD process.

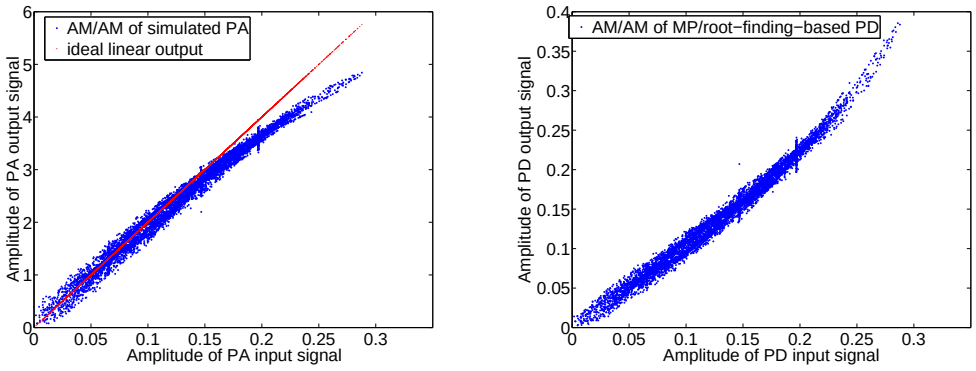
NMSE between the predistorted signal by proposed DPDs and that by MP/root-finding-based DPD

NMSE between the predistorted signal by the proposed DPD and that by MP/root-finding-based DPD is defined as

$$NMSE_{PD} = 10log_{10}\frac{\sum_{n=1}^{N}|x(n)_{root-finding} - x(n)_{proposed}|^2}{\sum_{n=1}^{N}|x(n)_{root-finding}|^2} \tag{4.26}$$

where N is the number of samples, $x(n)_{root-finding}$ the predistorted signal by MP/root-finding-based DPD, and $x(n)_{proposed}$ the predistorted signal by the proposed DPD.

In this subsection, NMSE between the predistorted signal by the proposed DPDs and that by MP/root-finding-based DPD can be used to measure the accuracy of the proposed DPDs. The proposed MP/LUT-based PD, MP/LILUT-based PD and MP/QILUT-based PD are based on MP/root-finding-based PD. AM/AM and AM/PM characteristics of MP/root-finding-based PD is shown in Fig. 4.9 and Fig. 4.10. It shows that AM/AM and AM/PM characteristics of MP/root-finding-based PD is the inverse of that of simulated PA. For the proposed MP/LUT-based PD, MP/LILUT-based PD and MP/QILUT-based PD, the predistortion accuracy is related to the table size. If the table size is large enough, they should have the same predistortion ability as MP/root-finding-based PD.



(a) AM/AM characteristic of simulated PA

(b) AM/AM characteristic of MP/root-finding-based PD

Figure 4.9: AM/AM characteristics of simulated PA and MP/root-finding-based PD

Fig. 4.11(a) shows the evolution of NMSE between the predistorted signal by MP/LUT-based DPD and that by MP/root-finding-based DPD versus the table size.

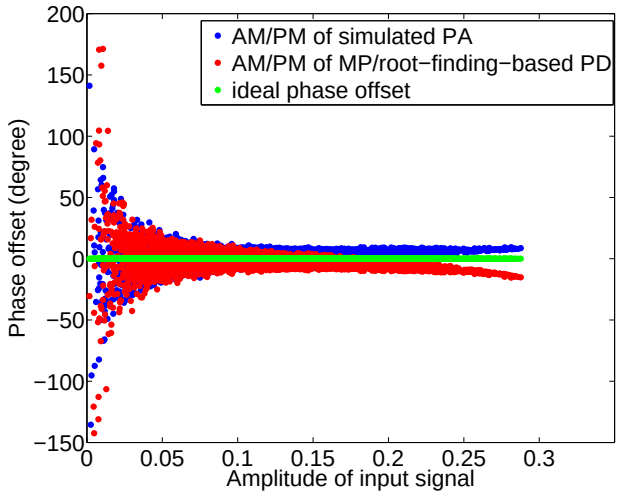


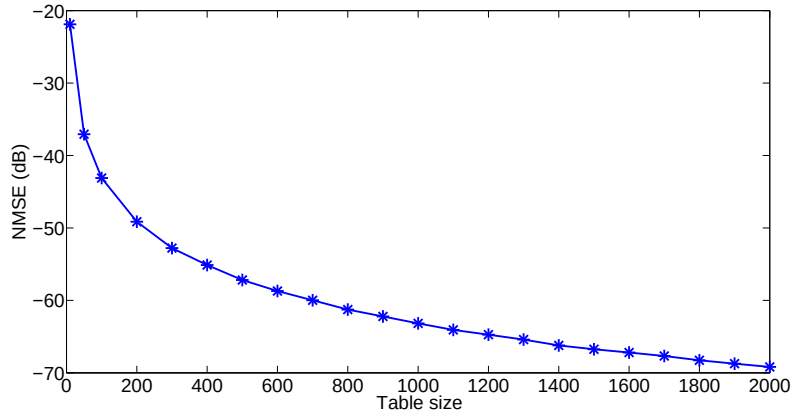
Figure 4.10: AM/PM characteristics of simulated PA and MP/root-finding-based PD

It shows that NMSE value improves with the increasing table size. NMSE value can achieve -69.2 dB when the table size is up to 2000.

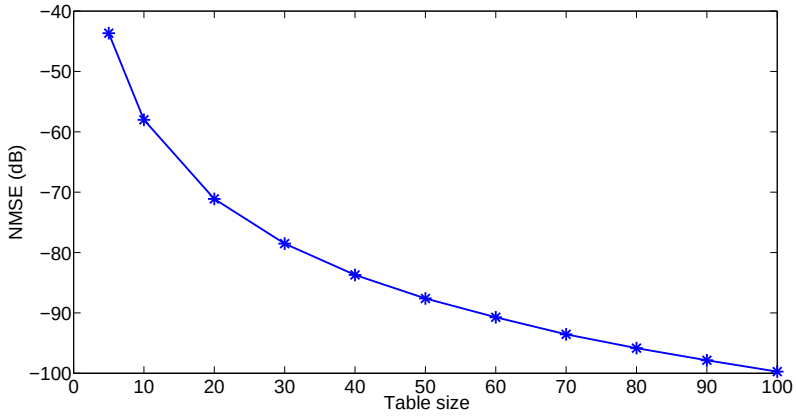
Fig. 4.11(b) shows the evolution of NMSE between the predistorted signal by MP/LILUT-based DPD and that by MP/root-finding-based DPD versus the table size. It shows that MP/LILUT-based DPD requires less table size than MP/LUT-based DPD for the same NMSE accuracy. NMSE value can achieve -71.1 dB with table size of only 20 (far smaller than 2000).

Fig. 4.11(c) shows the evolution of NMSE between the predistorted signal by MP/QILUT-based DPD and that by MP/root-finding-based DPD versus the table size. It shows that MP/QILUT-based DPD further decreases the table size than MP/LILUT-based DPD for the same NMSE accuracy. NMSE value can achieve -71.8 dB with the table size of only 10.

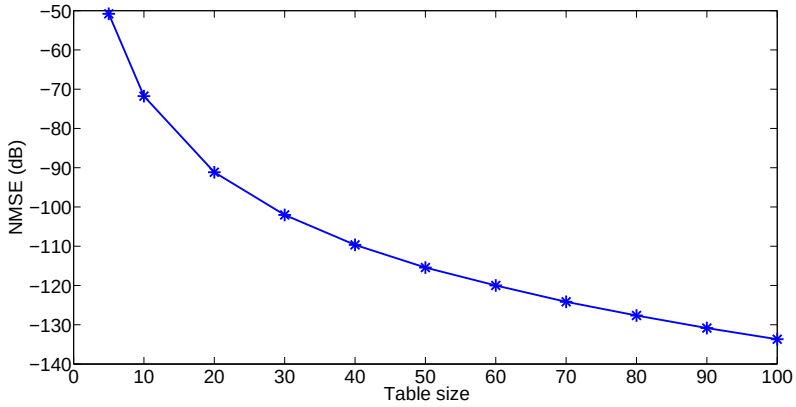
From Fig. 4.11, it shows that the predistorted signal by MP/QILUT-based DPD is the most closest to that by MP/root-finding-based DPD with same table size. For instance, with same table size 100, NMSE values by MP/LUT-based DPD, MP/LILUT-based DPD and MP/QILUT-based DPD are -43.1 dB, -99.7 dB and -133.7 dB, respectively. In other words, when obtaining the same NMSE accuracy, MP/QILUT-based DPD requires the least table size. As mentioned above, when NMSE value is about -70 dB, MP/LUT-based DPD, MP/LILUT-based DPD and MP/QILUT-based DPD require a table size of 2000, 20 and 10, respectively.



(a) MP/LUT-based DPD



(b) MP/LILUT-based DPD



(c) MP/QILUT-based DPD

Figure 4.11: Evolution of NMSE between the predistorted signal by proposed DPD and that by MP/root-finding-based DPD versus table size

NMSE between the desired PA output and actual linearized PA output

NMSE between the desired PA output and actual linearized PA output is used to measure the linearity metrics of linearized PA. It is defined as

$$NMSE = 10log_{10}\frac{\sum_{n=1}^{N}|G_0u(n) - y(n)|^2}{\sum_{n=1}^{N}|G_0u(n)|^2} \tag{4.27}$$

where G_0 is the ideal gain, $u(n)$ and $y(n)$ are the input and output of PD-PA system, respectively.

Fig. 4.12(a) shows the evolution of NMSE between desired PA output and actual PA output with MP/LUT-based DPD versus table size. It indicates that the linearity of output improves with the increase of table size. NMSE between desired PA output and actual PA output with MP/root-finding-based DPD achieves -60.7 dB. For MP/LUT-based DPD, NMSE achieves -60.3 dB when the table size is up to 2000. MP/LUT-based DPD requires table size of 2000 to obtain similar linearity performance as MP/root-finding-based DPD.

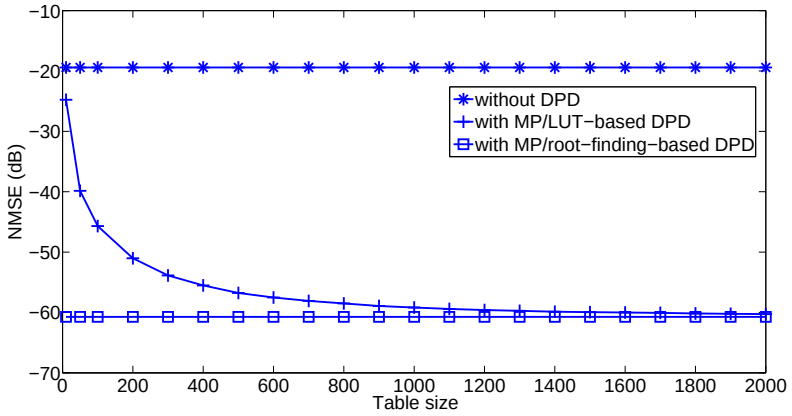
Fig. 4.12(b) shows the evolution of NMSE between desired PA output and actual PA output with MP/LILUT-based DPD versus table size. It can be seen that MP/LILUT-based DPD requires table size of 20 to obtain similar linearity performance (NMSE -60.3 dB) as MP/root-finding-based DPD.

Fig. 4.12(c) shows the evolution of NMSE between desired PA output and actual PA output with MP/QILUT-based DPD versus table size. It can be seen that MP/QILUT-based DPD only requires table size of 10 to obtain similar linearity performance (NMSE -60.2 dB) as MP/root-finding-based DPD. From Fig. 4.12(b), with table size of 10, MP/LILUT-based DPD obtains NMSE performance -56.2 dB (greater about 4 dB than MP/QILUT-based DPD).

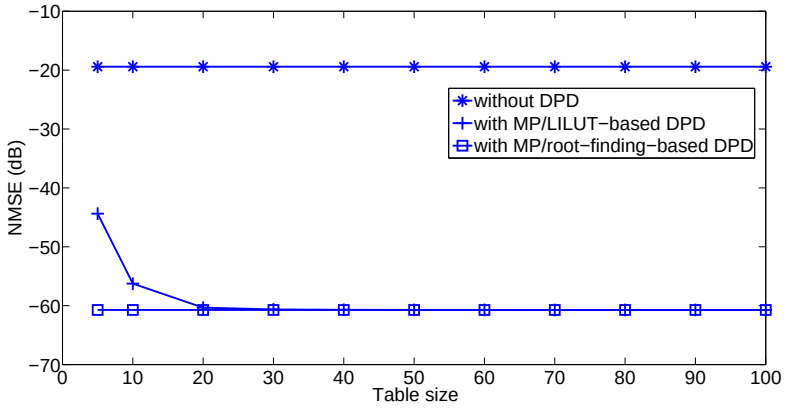
To see the linearity of PA output with DPD intuitively, Fig. 4.13 shows AM/AM characteristics of linearized PA. It further validates the results of Fig. 4.12.

ACPR of PA output

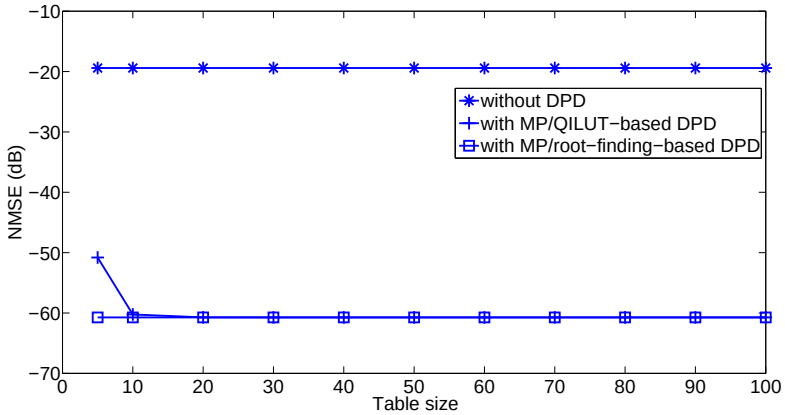
ACPR is a metric used to measure the linearity of PA output, which is defined in section 2.3.2. In the simulation, the offset of adjacent channel is set to 5 MHz. Fig. 4.14(a) shows the evolution of ACPR of PA output with MP/LUT-based DPD versus table size. This ACPR refers to the upper ACPR value. ACPR performance of MP/root-finding-based DPD is -61.2 dB. MP/LUT-based DPD requires a table size of 1000 to achieve the similar ACPR performance (-61.1 dB) as MP/root-finding-based



(a) MP/LUT-based DPD

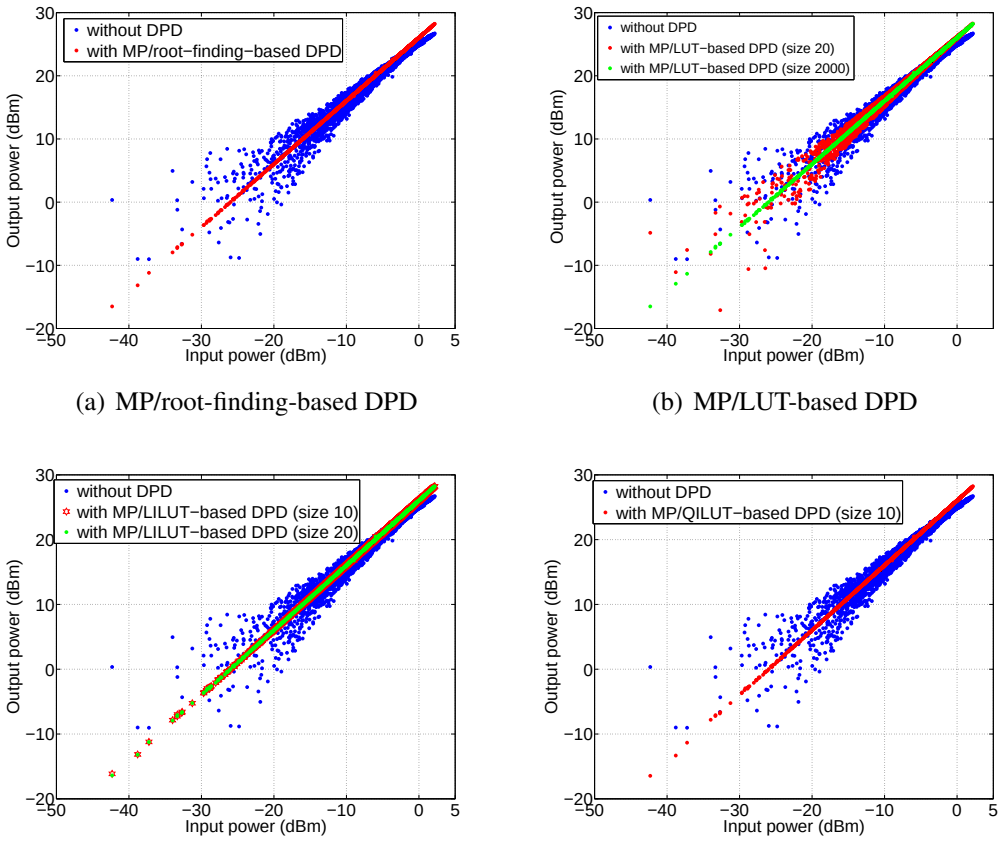


(b) MP/LILUT-based DPD



(c) MP/QILUT-based DPD

Figure 4.12: Evolution of NMSE between desired PA output and actual PA output with proposed DPD versus table size



(a) MP/root-finding-based DPD

(b) MP/LUT-based DPD

(c) MP/LILUT-based DPD

(d) MP/QILUT-based DPD

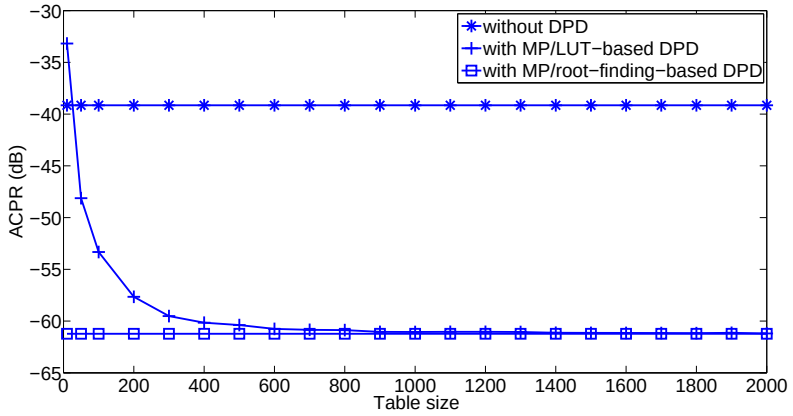
Figure 4.13: AM/AM characteristics of linearized PA

DPD.

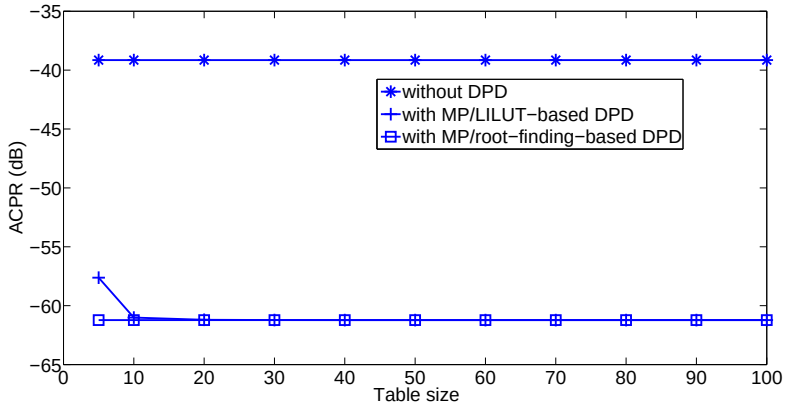
Fig. 4.14(b) shows the evolution of ACPR of PA output with MP/LILUT-based DPD versus table size. It shows that MP/LILUT-based DPD with table size of 10 can achieve similar ACPR performance (-61.0 dB) as MP/root-finding-based DPD. When the table size is up to 20, it can achieve the same ACPR performance (-61.2 dB) as MP/root-finding-based DPD.

Fig. 4.14(c) shows the evolution of ACPR of PA output with MP/QILUT-based DPD versus table size. It shows that MP/QILUT-based DPD requires the table size of 10 to achieve the same ACPR performance (-61.2 dB) as MP/root-finding-based DPD.

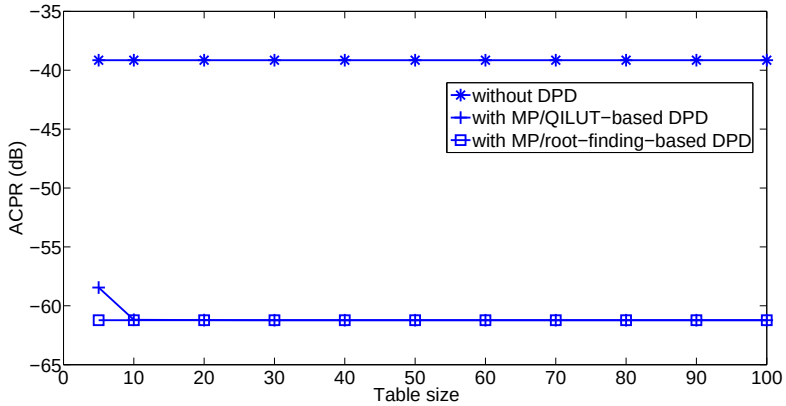
ACPR results shown above are obtained at the average input power -4 dBm (the average output power is 21.4 dBm). In the following, the proposed MP/LUT-based DPD, MP/LILUT-based DPD and MP/QILUT-based DPD are compared in terms of different output powers. In order to obtain the best performance, the table sizes of



(a) MP/LUT-based DPD



(b) MP/LILUT-based DPD



(c) MP/QILUT-based DPD

Figure 4.14: Evolution of ACPR of PA output with proposed DPD versus table size

MP/LUT-based DPD, MP/LILUT-based DPD and MP/QILUT-based DPD are set to 1000, 20 and 10, respectively. Fig. 4.15 summaries the evolution of ACPR of PA

output with different output power for proposed DPDs. In Fig. 4.15(a), it shows that MP/LUT-based DPD with a table size of 1000 can achieve the similar performance as MP/root-finding-based DPD and MP/LUT-based DPD with table size of 20 has very bad performance. In Fig. 4.15(b) and Fig. 4.15(c), they give the performance of MP/LILUT-based DPD and MP/QILUT-based DPD with different table sizes. It can be seen that MP/LILUT-based DPD with table size of 10 has a similar performance as MP/QILUT-based DPD with table size of 10. When the table size is 10, the difference of ACPR performance is very small between MP/LILUT-based DPD and MP/QILUT-based DPD. To see the difference of MP/LILUT-based DPD and MP/QILUT-based DPD, the case of table size of 5 is given in Fig. 4.15(d). From Fig. 4.15(d), it shows MP/QILUT-based DPD with table size of 5 is better than MP/LILUT-based DPD with table size of 5. It indicates that, when the table size is small, the ACPR performance of MP/QILUT-based DPD is better than that of MP/LILUT-based DPD. In Fig. 4.16, the spectrums of outputs of linearized PA with proposed DPDs are given at the average output power 18.6 dBm.

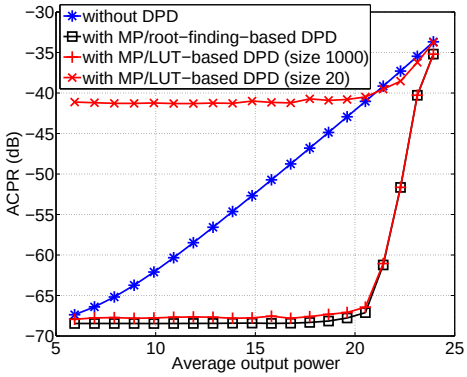
EVM of PA output

EVM is also a metric to measure the linearity of PA output, which is defined in sections 2.3.2. Fig. 4.17(a) shows the evolution of EVM of PA output with MP/LUT-based DPD versus table size. MP/root-finding-based DPD achieves EVM of 0.18%. MP/LUT-based DPD requires a table size of 300 to achieve an EVM of 0.25%.

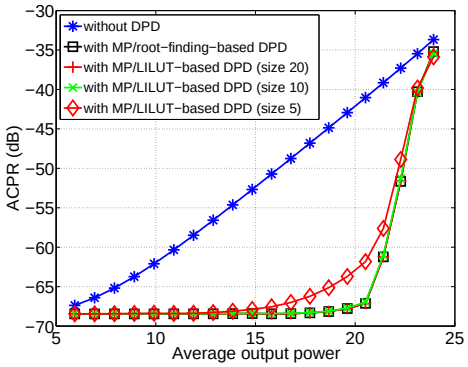
Fig. 4.17(b) shows the evolution of EVM of PA output with MP/LILUT-based DPD versus table size. MP/LILUT-based DPD requires a table size of 10 to achieve an EVM of 0.19%.

Fig. 4.17(c) shows the evolution of EVM of PA output with MP/QILUT-based DPD versus table size. MP/QILUT-based DPD requires a table size of 10 to achieve an EVM of 0.18%.

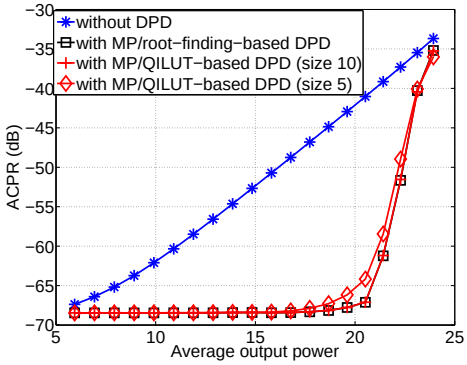
From Fig. 4.17, it shows that MP/LUT-based DPD requires more table size than MP/LILUT-based DPD and MP/QILUT-based DPD for the same performance. MP/LILUT-based DPD and MP/QILUT-based DPD both can achieve a good EVM performance with table size of 10. The difference between MP/LILUT-based DPD and MP/QILUT-based DPD is very small. In the following, EVMs of MP/LUT-based DPD, MP/LILUT-based DPD and MP/QILUT-based DPD are compared in terms of different output powers. In order to obtain the best performance, the table sizes of MP/LUT-based DPD, MP/LILUT-based DPD and MP/QILUT-based DPD are set to 1000, 10



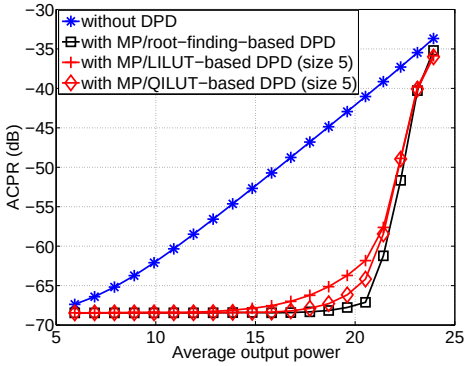
(a) MP/LUT-based DPD



(b) MP/LILUT-based DPD



(c) MP/QILUT-based DPD



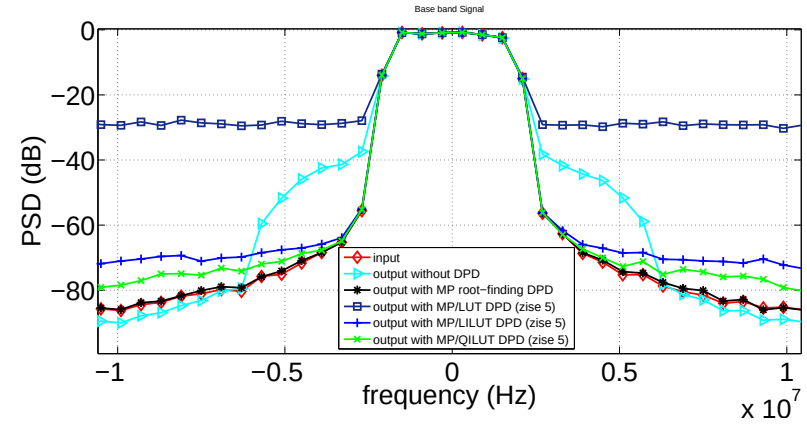
(d) Comparison of MP/LILUT-based DPD and MP/QILUT-based DPD

Figure 4.15: ACPR of linearized PA output with proposed DPDs at different output power

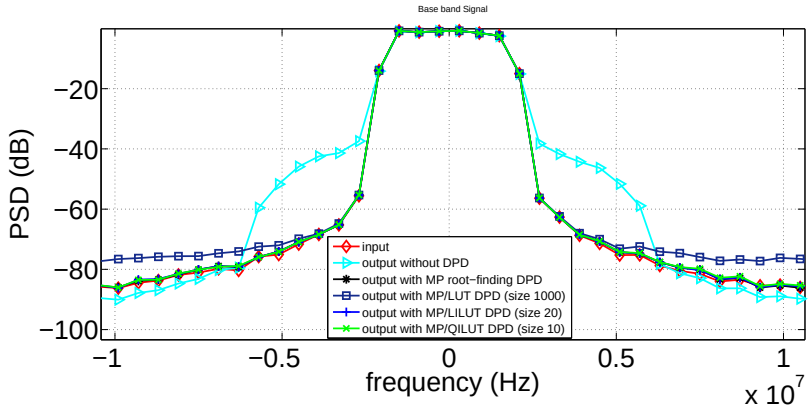
and 10, respectively. The results are summarized in Table 4.3 and Fig. 4.18. It further verifies that MP/LUT-based DPD requires bigger table size than MP/LILUT-based DPD and MP/QILUT-based DPD for the same performance. In EVM performance, the difference between MP/LILUT-based DPD and MP/QILUT-based DPD is small.

Consumed time of DPD process

In this subsection, the consumed time of proposed DPDs is analyzed. The consumed time of MP/LUT-based DPD, MP/LILUT-based DPD and MP/QILUT-based DPD is shown in Fig. 4.19. The consumed time refers to the time for processing one predistorted sample in DPD algorithm (not include the time of constructing LUT). It shows that MP/root-finding-based DPD requires about 180 us. It is the most consuming among these DPD methods. The consumed time of MP/LUT-based DPD,



(a)



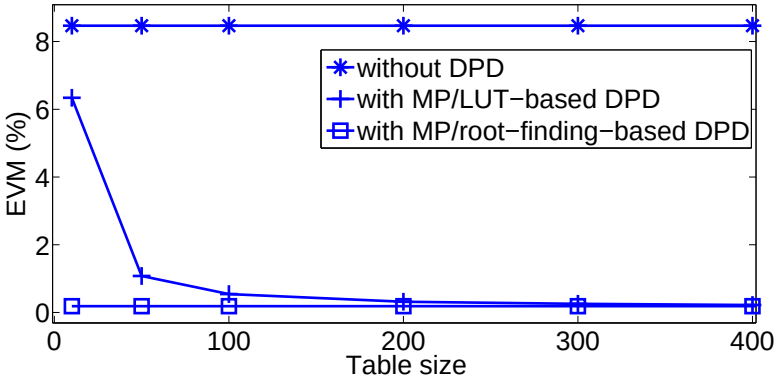
(b)

Figure 4.16: Spectrums of output of linearized PA with proposed DPDs at average output power 18.6 dBm

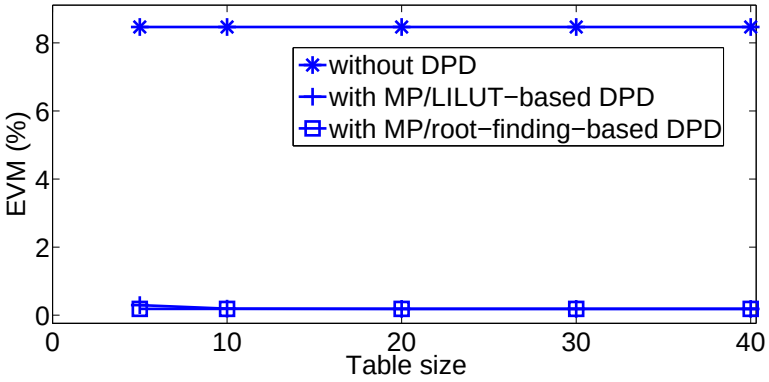
MP/LILUT-based DPD and MP/QILUT-based DPD all increases slightly with the table size. MP/LUT-based DPD requires the least time (only 30 us with table size 2000). According the above analysis of NMSE, ACPR and EVM, MP/LILUT-based DPD and MP/QILUT-based DPD require less table size. It is enough with table size of 20. MP/LILUT-based DPD with table size of 20 requires about 45 us. MP/QILUT-based DPD with table size of 20 requires about 151 us.

4.4 Proposed DPD with combination of MP and NN

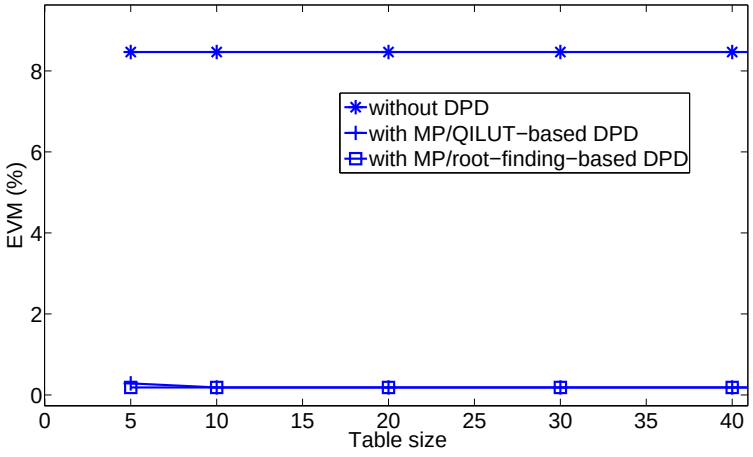
In this section, a new DPD method with combination of neural network and MP is proposed.



(a) MP/LUT-based DPD



(b) MP/LILUT-based DPD



(c) MP/QILUT-based DPD

Figure 4.17: Evolution of EVM of PA output with proposed DPD versus table size

output power (dBm)	without DPD	MP/LUT (size 1000)	MP/LILUT (size 10)	MP/QILUT (size 10)
5.93	5.732%	0.044%	0.004%	0.004%
6.92	5.734%	0.044%	0.005%	0.005%
7.92	5.738%	0.045%	0.007%	0.007%
8.92	5.744%	0.044%	0.008%	0.008%
9.91	5.751%	0.045%	0.011%	0.011%
10.90	5.763%	0.046%	0.012%	0.012%
11.89	5.780%	0.047%	0.014%	0.014%
12.88	5.805%	0.048%	0.019%	0.019%
13.86	5.843%	0.048%	0.024%	0.024%
14.84	5.900%	0.054%	0.031%	0.029%
15.81	5.987%	0.059%	0.039%	0.037%
16.77	6.116%	0.066%	0.050%	0.047%
17.73	6.310%	0.076%	0.063%	0.061%
18.64	6.598%	0.091%	0.082%	0.078%
19.60	7.019%	0.111%	0.107%	0.101%
20.52	7.622%	0.139%	0.144%	0.132%
21.41	8.466%	0.192%	0.192%	0.184%
22.28	9.617%	0.253%	0.273%	0.247%
23.12	11.14%	0.597%	0.596%	0.582%
23.93	13.12%	1.634%	1.592%	1.614%

Table 4.3: EVM with different output power

4.4.1 MP and neural network based DPD

Neural network has an ability of learning the dynamic behavior of nonlinear system. We would like to use the neural network to learn the dynamic nonlinear behavior of the predistorter. Feed Forward Neural Network (FFNN) shown in Fig .3.7 is used to model the predistorter.

It exists some neural network based DPDs in the literature. In most cases, the

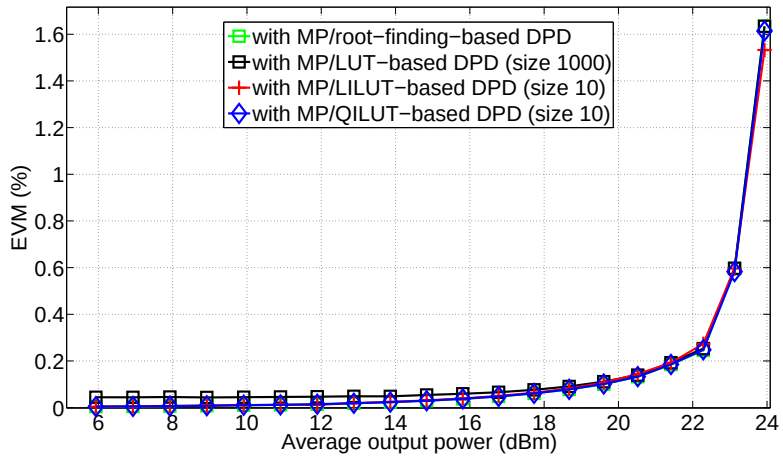


Figure 4.18: EVM of linearize PA output with proposed DPDs at different output power

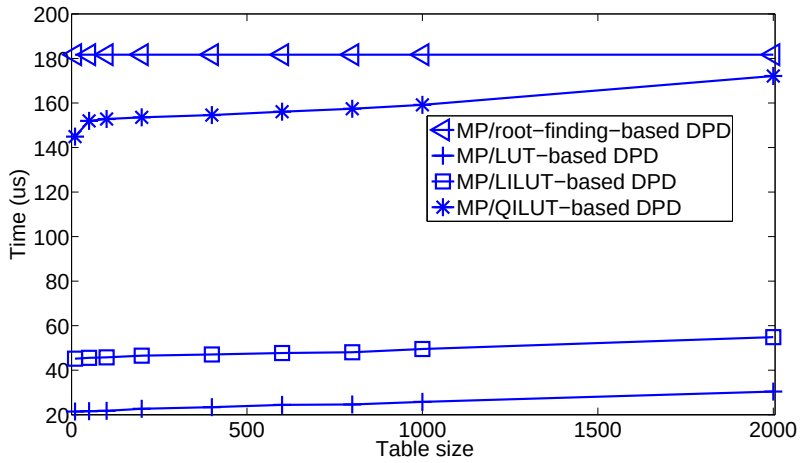


Figure 4.19: Consumed time with different table size

training samples of neural network in the predistorter are the measured input and output of the real PA. The output of PA is taken as the input sample of neural network. The input of PA is taken as the output sample of neural network. Such architecture based on indirect learning algorithm has a problem of modeling accuracy, especially when the measured signal is noisy [20]. In this thesis, the training samples are derived from MP/root-finding-based DPD. Firstly, the predistorted signal is calculated based on MP model with root-finding method. Secondly, such predistorted signal obtained by MP model is used to train FFNN. Finally, the trained FFNN model is used to calculate the predistorted signal as a predistorter, so that the cascade of FFNN and PA achieves a linear behavior.

The structure of FFNN is shown in Fig. 4.20. FFNN is trained by the system input signal $u(nT)$ and the predistorted signal $x(nT)$ obtained by MP/root-finding-based DPD. This FFNN is based on real-valued input and output. Since $u(nT)$ and $x(nT)$ are both complex signals, they are considered in FFNN model as two parts: the real part and the imaginary part. $Real\{\cdot\}$ and $Imag\{\cdot\}$ represent the real and imaginary parts of the complex signal, respectively. The input of FFNN includes the current and previous samples. The FFNN is implemented by MATLAB ANN toolbox. The FFNN generation function is *newff* and the training function is *train*.

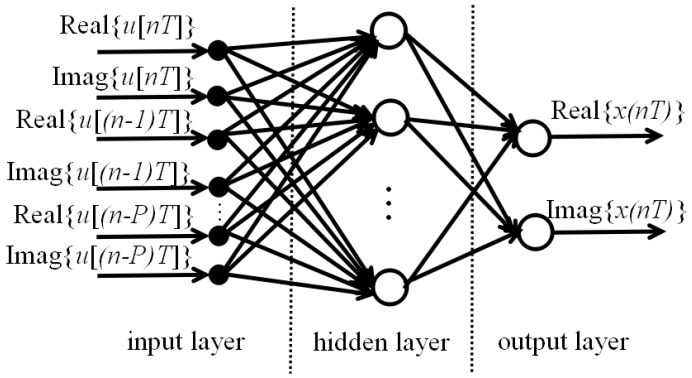


Figure 4.20: Structure of three-layer feed-forward neural network

4.4.2 Neural network training

In order to obtain a good modeling performance of FFNN, some important parameters of FFNN need to be determined. They are number of hidden layers, number of neurons, transfer function, Back Propagation (BP) network training function, BP weight/bias learning function, performance function. These parameters setting mainly depends on empirical analysis. The training sample sequence is 1000.

Number of hidden layers

The universal approximation theorem [76, 77, 94] has proven that a FFNN with a single hidden layer can approximate any nonlinear function with any desired error. Therefore, FFNN with a single hidden layer is adopted to realize the predistorter.

Number of neurons in each layer

The number of neurons of input layer depends on the memory depth. The number of neurons of hidden layer depends on the empirical analysis. The number of neurons of output layer depends on the number of FFNN output, so it is set to be 2. To determine the number of neurons of input layer, we compare NMSE (NMSE between the sample output and the output of trained neural network) performance of NN with different memory depth. Because the number of neurons of hidden layer is also not known, the number of neurons of hidden layer is assumed to be 10. The maximum number of iterations is 300. Table 4.4 shows NMSE performance of NN with different memory depths. It shows that NN achieves the best performance with memory depth $P = 3$. Therefore, the number of neurons of input layer is set to be 8. In the following, the number of neurons of hidden layer needs to be determined. Table 4.5 shows NMSE performance of NN with different numbers of neurons in hidden layer. It shows NN achieves the best performance when the number of neurons of hidden layer is 25. Henceforward, the number of neurons of hidden layer is set to be 25.

Memory depth	NMSE (dB)
P=1	-57.06
P=2	-58.25
P=3	-61.03
P=4	-59.27
P=5	-58.53

Table 4.4: NMSE performance versus memory depth

Transfer function

A popular choice of transfer function in output layer is the linear function *purelin*. For the transfer function of hidden layer, *tansig* (tan sigmoid function) and *logsig* (log sigmoid function) are used mostly. Table 4.6 shows NMSE performance of NN with different transfer functions in hidden layer. It shows *tansig* has better performance than *logsig*. Thus, *tansig* function is used as the transfer function of hidden layer.

Number of neurons	NMSE (dB)
5	-36.69
10	-61.08
15	-67.30
20	-71.42
25	-71.99
30	-69.02

Table 4.5: NMSE performance versus number of neurons in hidden layer

Transfer function	NMSE (dB)
tansig	-72.35
logsig	-68.63

Table 4.6: NMSE performance versus transfer function in hidden layer

Other parameters

In FFNN, there are still other parameters to be set. They are set as follows. BP network training function is *trainlm*. BP weight/bias learning function is *learngdm*. The performance function is *mse*. Table 4.7 summarizes the topology of the considered FFNN and its training setup. To obtain a better modeling performance, the number of iterations of NN is up to 1000 or MSE performance function achieves 10^{-10}.

4.4.3 Simulation results

The trained FFNN model is inserted in the front of PA as the predistorter. Some simulation results are obtained. NN-based DPD is compared with MP/root-finding-based DPD in terms of NMSE between the desired linearized PA output and actual PA output, ACPR of PA output, EVM of PA output, and the consumed time of DPD process.

Fig. 4.21 shows the AM/AM characteristics of linearized PA with NN-based DPD at average output power 18.6 dBm. NMSE between the desired linearized PA output and actual PA output with NN-based DPD achieves -64.47 dB. NMSE between the desired linearized PA output and actual PA output with MP/root-finding-based DPD is

Characteristic	Setting
Neural network	three-layer
Number of neurons	8-25-2
Transfer function	tansig-purelin
BP network training function	trainlm
BP weight/bias learning function	learngdm
Performance function	mse
Number of training samples	1000

Table 4.7: Neural network training configuration

-68.61 dB. It shows that NN-based DPD has a good linearization performance.

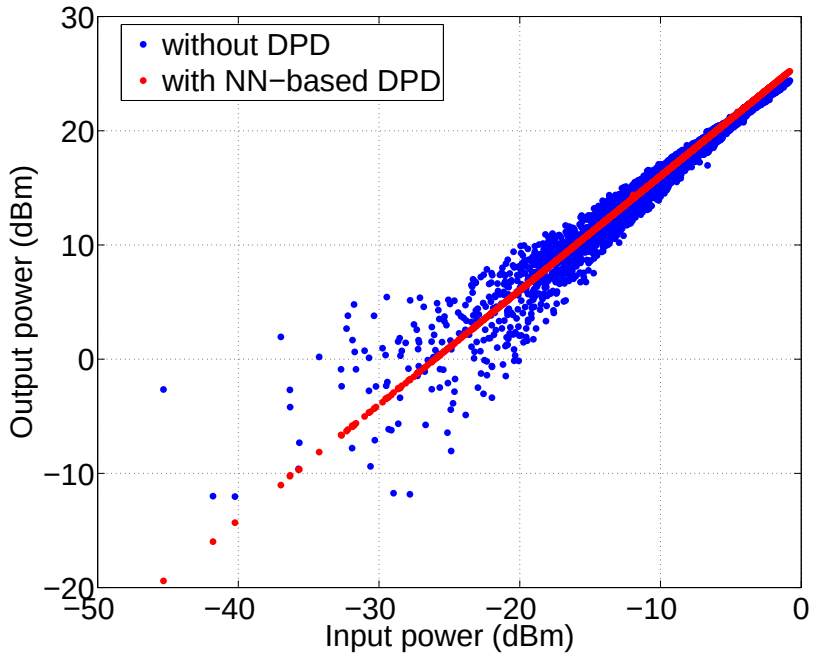


Figure 4.21: AM/AM characteristics of linearized PA with NN-based DPD

Fig. 4.22 shows the spectrum of output of linearized PA with NN-based DPD and MP/root-finding-based DPD at average output power 18.6 dBm. ACPR with MP/root-finding-based DPD is -68.15 dB and ACPR with NN-based DPD -68.00 dB. It shows that the linearization performance of NN-based DPD is very close to that of MP/root-finding-based DPD.

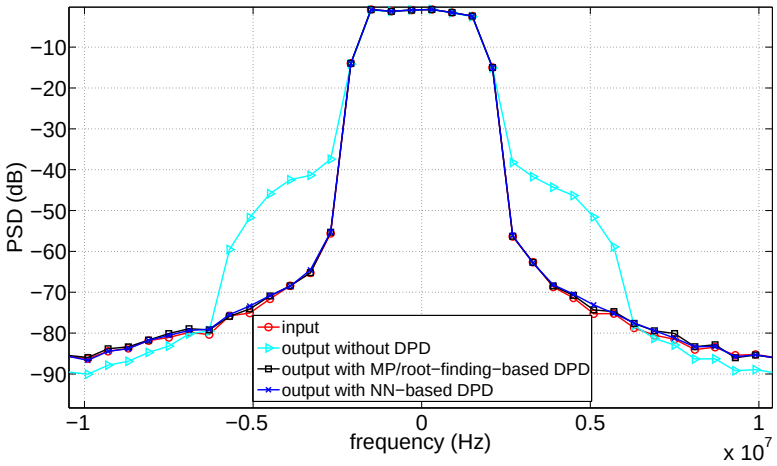
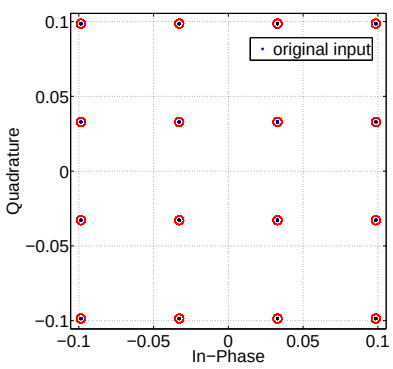


Figure 4.22: Spectrum of output of linearized PA with NN-based DPD

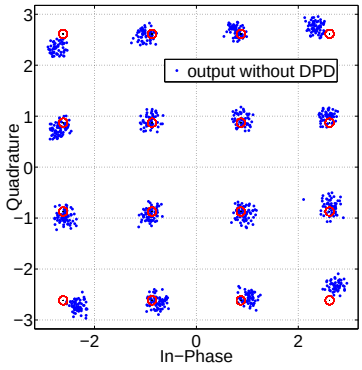
Fig. 4.23 shows the constellations of PA input, PA output without DPD, PA output with MP/root-finding-based DPD and PA output with NN-based DPD at average output power 18.6 dBm. The EVM of PA output without DPD is 8.46%. The EVMs of PA output with MP/root-finding-based DPD and with NN-based DPD are 0.07% and 0.06%, respectively. According to EVM results, it further indicates NN-based DPD can achieve similar linearization performance as MP/root-finding-based DPD.

The simulation results shown above are based on an average output power of 18.6 dBm. In the following, the simulation results with different output power are given. Fig. 4.24 summaries the evolution of ACPR of PA output with different output power for MP/root-finding-based DPD and NN-based DPD. Fig. 4.25 summaries the evolution of EVM of PA output with different output power for MP/root-finding-based DPD and NN-based DPD. They show the performance of NN-based DPD is almost same as that of MP/root-finding-based DPD.

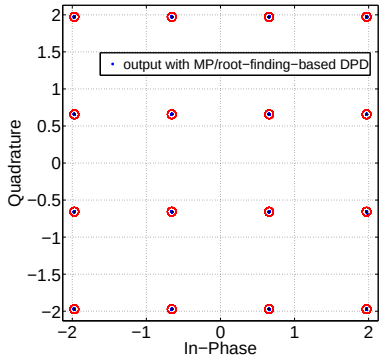
It is known that the training of NN is very time-consuming. Especially when a good modeling performance is desired, the number of iterations is large. However, the trained NN is less time-consuming. The consumed time with MP/root-finding-based DPD and NN-based DPD is presented in Table 4.8. The time refers to the consumed time for processing one predistorted sample.



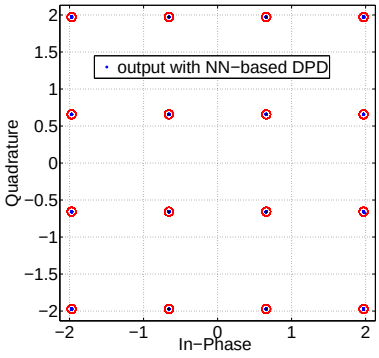
(a) Input



(b) Output without DPD (8.46%)



(c) Output with MP/root-finding-based DPD (0.07%)



(d) Output with NN-based DPD (0.06%)

Figure 4.23: Constellation of linearize PA output with NN-based DPD

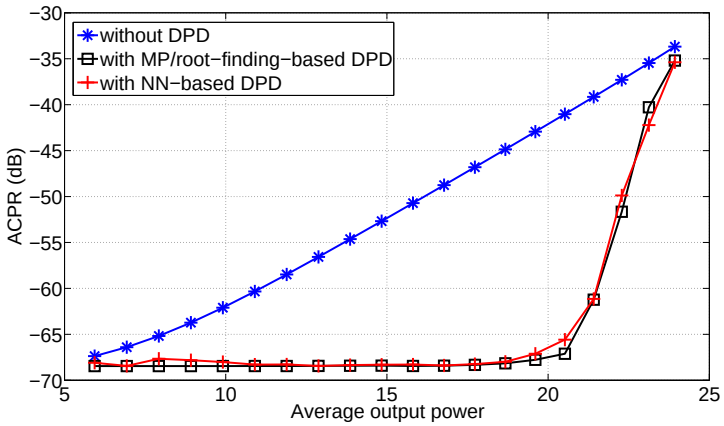


Figure 4.24: ACPR of linearize PA output with NN-based DPD at different output power

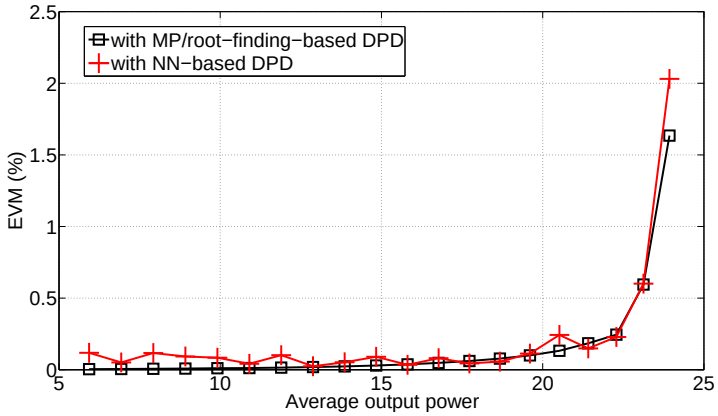


Figure 4.25: EVM of linearize PA output with NN-based DPD at different output power

DPD method	Time (us)
MP/root-finding-based DPD	181.67
NN-based DPD	9.12

Table 4.8: Consumed time with MP/root-finding-based DPD and NN-based DPD

5

Experimental setup and results

In this section, the proposed DPDs are validated experimentally. Some parameters which influence the linearization performance are discussed, such as the number of bits of ADC and bandwidth of instruments. The experimental results of ACPR and EVM are given to compare the different methods.

5.1 Experimental setup

In order to validate the proposed DPDs, an experimental testbed is built. Fig. 5.1 and Fig. 5.2 show the used experimental testbed. It consists of a PA (ZFL 2500), a PC, a Vector Signal Generator (VSG, Rohde & Schwarz SMU 200 A) and a Spectrum Analyzer (SA, Agilent E4440A, 3 Hz - 26.5 GHz). The PC provides the baseband signal to VSG, realizes DPD algorithms and controls the measurement system by using Matlab toolbox (Instrument Control Toolbox). The VSG generates a RF signal and sends it to the PA. The SA receives and analyzes the output signal of PA. It is worth noting that there is a 30 dB attenuator between the output of PA and SA because of the maximum input power limitation of SA. More details of this testbed are presented in [32, 95].

The bandwidth of PA ZFL2500 is 500-2500 MHz and its gain is about 31 dB [96]. Fig. 5.3 shows the gain characteristics of PA ZFL2500. The measured linear gain is 31.67 dB. The average output power at 1 dB compression point is around 17.40

dBm at frequency of 1.8 GHz. Fig. 5.4 shows the dynamic AM/AM and AM/PM characteristics of PA ZFL2500.

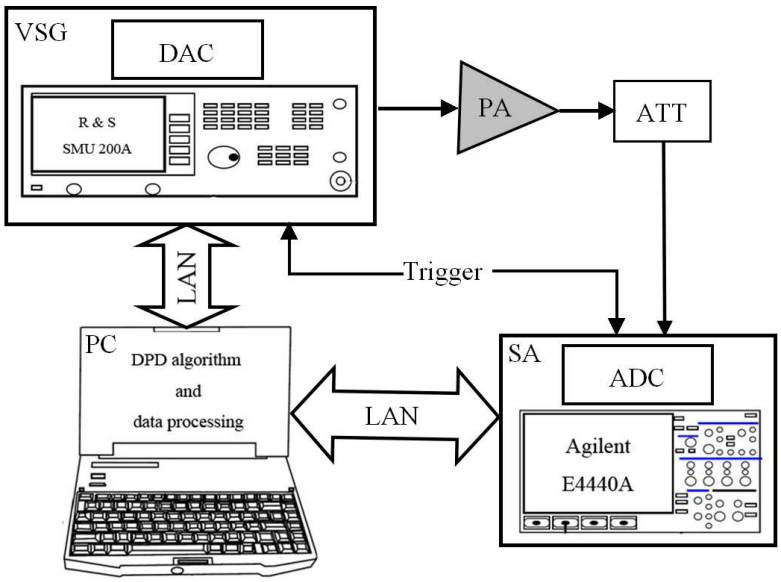


Figure 5.1: Block diagram of experimental setup

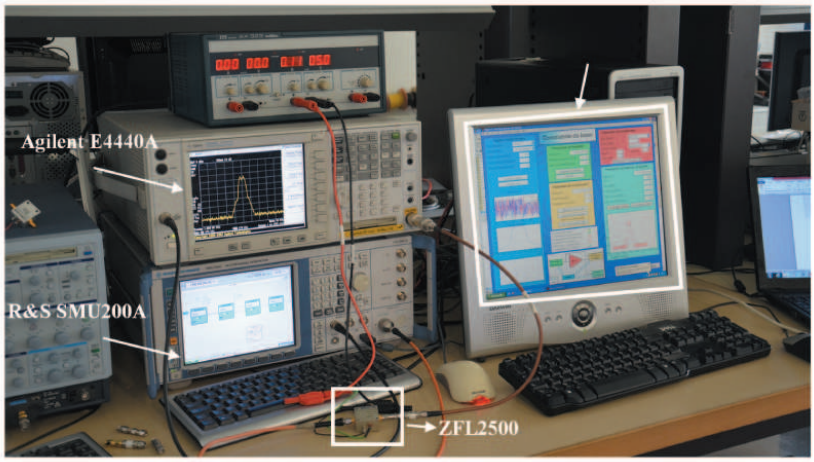


Figure 5.2: Photograph of experimental setup

5.2 Discussions

In practical measurements, hardware conditions can not be ignored, such as the number of bits of ADC (ADC is located is the SA) and instrument bandwidth. These

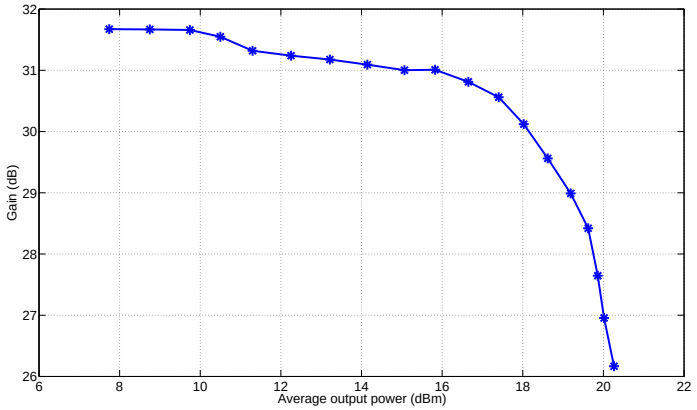


Figure 5.3: Gain characteristic of PA ZFL2500

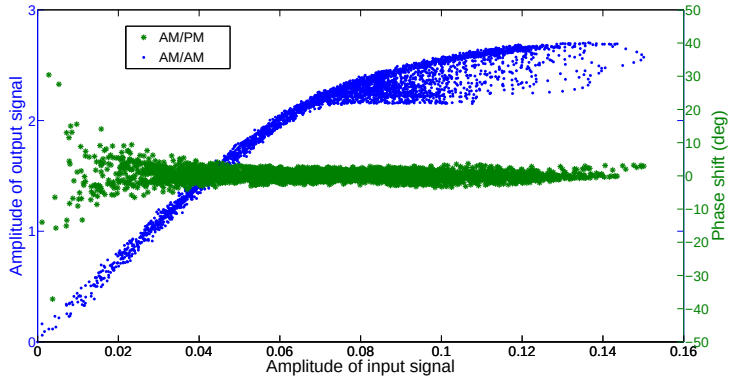


Figure 5.4: Dynamic AM/AM and AM/PM characteristics of PA ZFL2500

parameters and their influence on measurements are analyzed by some simulations. The simulated PA is still the Wiener model used in section 4.1. The original input signal is a 16QAM signal with 3.84 MHz bandwidth. The input sequence is 1000 symbols. And the same sequence is transmitted continuously. The input power is -5 dBm. The DPD method used is the MP/root-finding-based DPD.

To evaluate the influence of number of bits of ADC, an ADC is assumed before the data acquisition. The samples are quantified and the number of bits of ADC varies from 3 to 16. The sampling frequency of the input signal is 8 samples/symbol. ACPR is used to quantify the linearization performance. Fig. 5.5 presents the evolution of ACPR versus the number of bits of ADC. It can be seen that the linearization performance improves as the number of bits of ADC increases. It indicates that if the number of bits of ADC is not big enough in the data acquisition, the quantization error is too large and degrades the linearization performance. In our experimental setup, the acquisition

of signals by Agilent E4440A is realized with ADC of 14 bits.

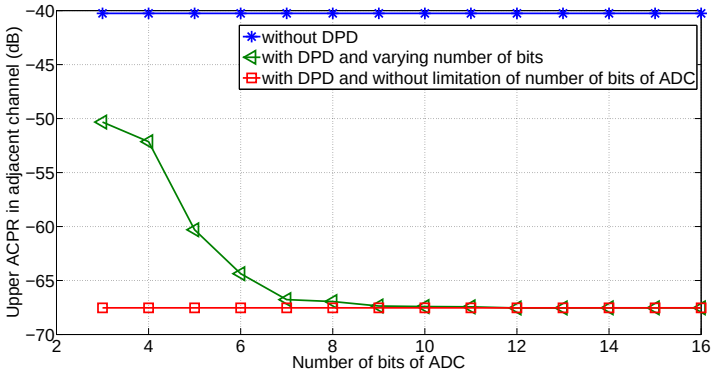


Figure 5.5: Evolution of adjacent channel ACPR versus number of bits of ADC

Another problem is the bandwidth limitation of instrument. In simulation, the signal bandwidth is only limited by the sampling frequency. In measurement, except for the influence of the sampling frequency, the signal bandwidth is also affected by the bandwidth limmitation of instrument. When the bandwidth of instrument is narrower than that of the signal, the signal passing the instrument will be distorted. Especially, the signal bandwidth will increase after the predistortion. The bandwidth of the predistorted signal is wider than that of the original input signal. Thus, if the bandwidth of instrument is not sufficiently wide to enable the transmission of the predistorted signal, the desired predistortion will be not achieved.

In order to evaluate the required bandwidth, the following simulation is made. It assumes that the sampling frequency of input signal is 50 samples/symbol. To simulate the instrument limitation, a lowpass linear-phase FIR filter, with order 200 and a varying cutoff frequency, is used to limit the bandwidth of the predistorted signal. The limited bandwidth signal is finally applied to the Wiener PA model. Fig. 5.6 presents the evolution of ACPR versus the cutoff frequency, where this cutoff frequency represents the bandwidth of instrument. It indicates that if the bandwidth of instrument is wide enough, the linearization performance will be perfect. Actually, to obtain good performance, 5 times is sufficient.

5.3 Experimental results analysis

Before testing the DPDs, the PA modeling is conducted firstly. A 16QAM signal with 3.84 MHz bandwidth is transmitted to the PA at frequency of 1.8 GHz. Then

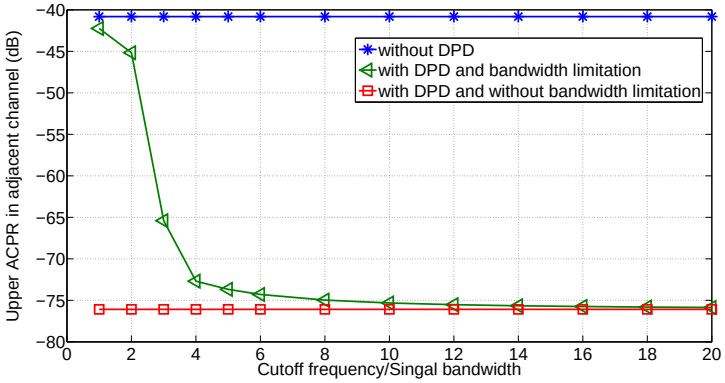


Figure 5.6: Evolution of adjacent channel ACPR versus cutoff frequency

the RF input and output signals of PA are saved. After obtaining two perfectly synchronous signals, these two RF signals are demodulated and processed in baseband. The baseband input and output are used for PA modeling. The average power of PA input is -15.93 dBm. The average power of PA output is 15 dBm. The input and output sequences both have 200 symbols (20 samples by each symbol). The MP model is chosen for digital predistortion. The best NMSE performance achieves -39.2 dB with a nonlinearity order 5 and a memory depth 3.

In the measurement, PA ZFL2500 is also driven by a similar 16QAM signal with different average power. The sampling frequency is 20 samples/symbol. The signal sequence is still 200 symbols. The proposed methods, MP/LUT-based DPD, MP/LILUT-based DPD, MP/QILUT-based DPD and NN-based DPD, are tested on the experimental testbed.

Fig. 5.7 shows the spectrums of the output of linearized PA with the proposed DPDs at average output power 14.3 dBm. It can be seen that MP/root-finding-based DPD has the best linearization performance. For MP/LUT-based DPD, the linearization performance is very low when the table size is small (size 10). The quantization effect related to the table size affects the linearization performance of MP/LUT-based DPD. It achieves a similar linearization performance as MP/root-finding-based DPD with table size of 1000 (the quantization error is small). While MP/LILUT-based DPD and MP/QILUT-based DPD achieve the same performance as MP/root-finding-based DPD, with table size of only 10. The table size required by MP/LUT-based DPD is far larger than that required by MP/LILUT-based DPD and MP/QILUT-based DPD. For MP/LILUT-based DPD and MP/QILUT-based DPD, their difference of ACPR linearization performance is very small with same table size in the measurement. Theoretically, MP/QILUT-based DPD is slightly better than MP/LILUT-based DPD. Because

the noise and measurement error, MP/QILUT-based DPD can not reflect its advantage perfectly. At least, MP/QILUT-based DPD is not worse than MP/LILUT-based DPD with same table size. For NN-based DPD, it can also achieve the same linearization performance as MP/root-finding-based DPD.

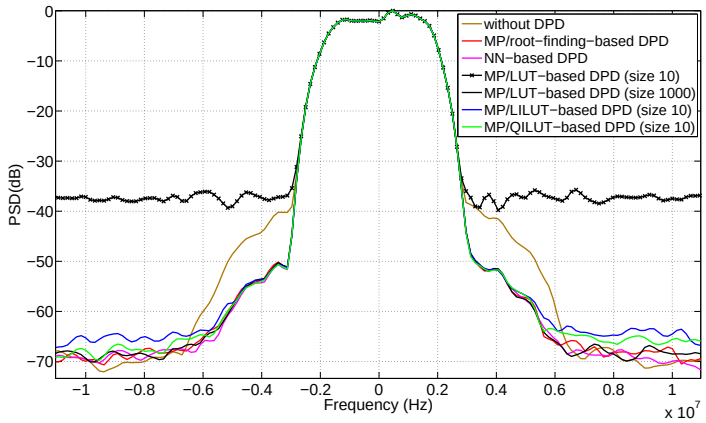
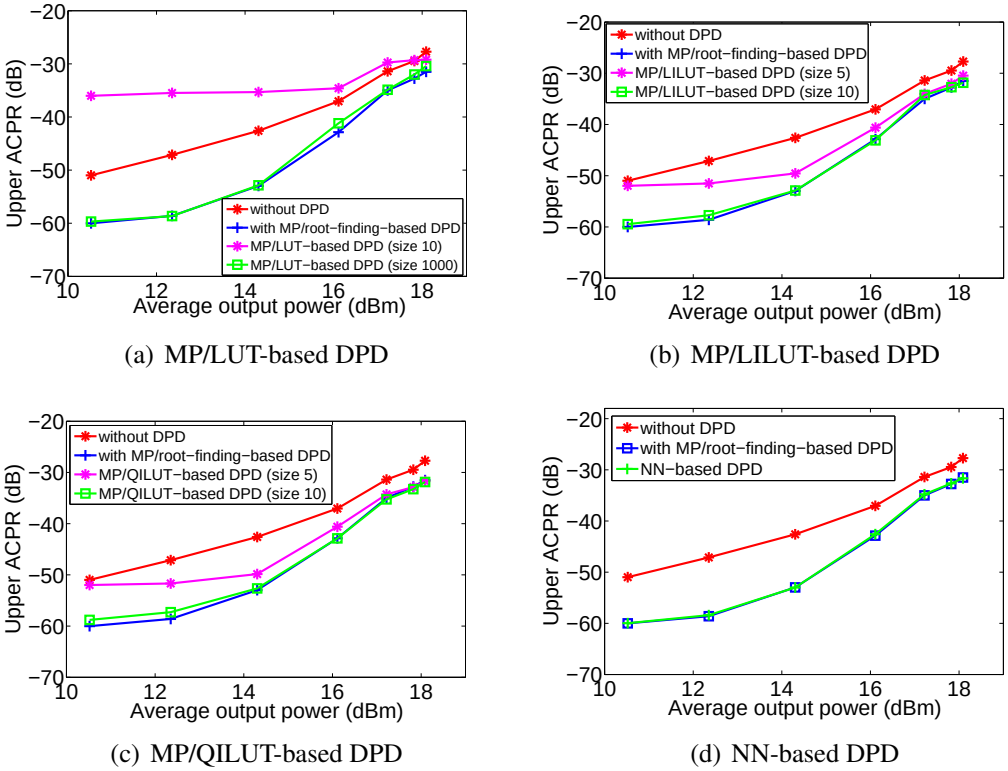


Figure 5.7: Spectrums of the output of linearized PA with proposed DPDs at average output power 14.3 dBm

The linearization performance is also evaluated in terms of ACPR and EVM of the output of linearized PA versus different average output power. Fig. 5.8 shows the evolution of ACPR of PA output versus average output power for the proposed DPD methods at the experimental testbed. The offset of adjacent channel is set to be 5 MHz. It indicates that MP/LUT-based DPD (size 1000), MP/LILUT-based DPD (size 10), MP/QILUT-based DPD (size 10) and NN-based DPD all can achieve the same linearization performance as MP/root-finding-based DPD.

Fig. 5.9 shows the evolution of EVM of PA output versus average output power for the proposed DPD methods. It further validates the results stated above. MP/LUT-based DPD (size 1000), MP/LILUT-based DPD (size 10), MP/QILUT-based DPD (size 10) and NN-based DPD all can reduce the distortion, especially at the high non-linear region.

About the comparison result of consumed time of the proposed DPDs, it is presented in section 4.3.4 and section 4.4.3. In the measurement, the number of signal sequence is 4000. Here the total time for obtaining 4000 samples of predistorted signal is compared. This consumed time only refers to the time of DPD process. For NN-based DPD, the configure parameter of NN is the same as shown in Table 4.7 and the consumed time of NN training is not included in DPD process. For a good accuracy, we run each DPD algorithm for 30 times and their average time is ob-



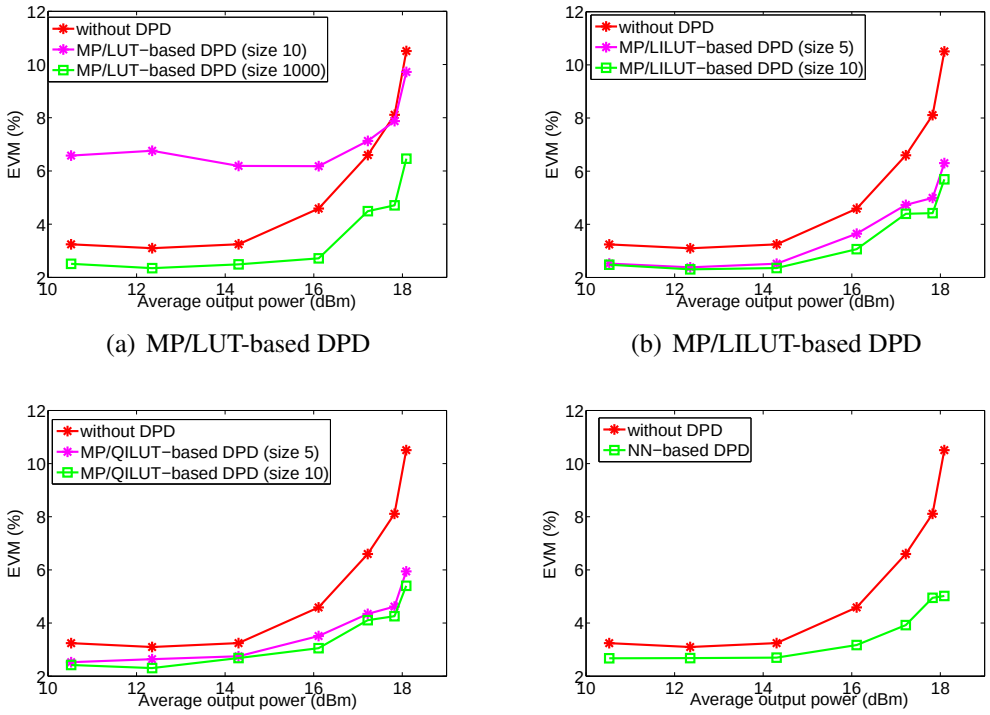
(a) MP/LUT-based DPD (b) MP/LILUT-based DPD

(c) MP/QILUT-based DPD (d) NN-based DPD

Figure 5.8: ACPR of linearized PA output with proposed DPDs

tained as the consumed time of DPD. The consumed time of MP/LUT-based DPD (size 1000), MP/LILUT-based DPD (size 10), MP/QILUT-based DPD (size 10) and NN-based DPD is summarizd in Table 5.1.

It shows MP/root-finding-based DPD requires about 699.8 ms for 4000 samples. It is the most consuming. The average consumed time of MP/LUT-based DPD (size 1000), MP/LILUT-based DPD (size 10), MP/QILUT-based DPD (size 10) and NN-based DPD are 112.5 ms, 189.3 ms, 559.4 ms and 38.4 ms, respectively. The average consumed time of NN-based DPD is the lowest, but the training of NN is very time-consuming.



(a) MP/LUT-based DPD (b) MP/LILUT-based DPD

(c) MP/QILUT-based DPD (d) NN-based DPD

Figure 5.9: EVM of linearized PA output with proposed DPDs

DPD method	Time (ms)
MP/root-finding-based DPD	699.8
MP/LUT-based DPD (size 1000)	112.5
MP/LILUT-based DPD (size 10)	189.3
MP/QILUT-based DPD (size 10)	559.4
NN-based DPD	38.4

Table 5.1: Consumed time of proposed DPDs

6

Conclusions and perspectives

In this section, we make a conclusion of this thesis. And a brief discussion is provided for the future work.

6.1 Conclusions

In this thesis, we studied the baseband digital predistortion method for linearizing power amplifiers with memory effects. In wireless communication system, power amplifier is an indispensable component. Its inherent nonlinearity and memory effects lead to the in-band and out-of-band distortions and affect the power efficiency. Baseband digital predistortion is one of the most promising methods for compensating the distortions.

In the beginning, this dissertation presents the background of linearization techniques. It includes the evolution of wireless communication systems, transceiver, power amplifier, nonlinearity characteristics of power amplifier, power amplifier behavior modeling and some criteria for evaluating the nonlinearity. After that, several classical linearization techniques are briefly introduced, such as power backoff, feedforward, feedback, LINC, EER and digital predistortion. Among these methods, digital predistortion is adopted in this thesis due to its ability of reconfiguration and simplicity of implementation. It exists lots of different digital predistortion methods in the literature. They can be classified into three categories: LUT based DPD, neu-

ral network based DPD and model based DPD. They have their own advantages and disadvantages. In this thesis, these are three types of DPDs all have been studied. In addition, the learning architecture and identification algorithm of digital predistortion are described. There are two different learning architectures (ILA and DLA) to implement digital predistortion. ILA and DLA are compared in terms of condition number of matrix, modeling performance and linearization performance. It indicates that DLA is more robust than ILA when considering the noise and quantification error caused by ADC.

Four DPD methods based on DLA, MP/LUT-based DPD, MP/LILUT-based DPD, MP/QILUT-based DPD and NN-based DPD are proposed in this thesis. In practice, they are implemented by two kinds of solutions. One solution is to combine MP and LUT. LUT based DPD has the advantage of simplicity, but its linearization performance depends on the LUT size. MP based DPD has good linearization performance but it is time-consuming, such as MP/root-finding-based DPD. Consequently, the method of combining LUT and MP is suggested. The other solution is to use neural network to realize digital predistortion due to its strong ability of learning and imitation.

MP/LUT-based DPD is proposed based on MP/root-finding-based DPD by combing MP and simple non-interpolated LUT. In MP/root-finding-based DPD, the amplitude and phase of the predistorted signal both are obtained by complicated mathematical operations. This process is very time-consuming. In MP/LUT-based DPD, the processes of calculating the amplitude and phase of predistorted signal are both replaced by using LUT. The consumed time of MP/LUT-based DPD decreases greatly. However, the linearization performance of MP/LUT-based DPD depends on the table size. Based on MP/LUT-based DPD, MP/LILUT-based DPD is proposed by introducing the linear interpolation technique. MP/LILUT-based DPD greatly reduces the table size at cost of a small increase of time. But the consumed time of MP/LILUT-based DPD is still less than that of MP/root-finding-based DPD. Compared with the linear interpolation, the quadratic interpolation has more accurate interpolation performance. In other words, there is still a room for improving MP/LILUT-based DPD. MP/QILUT-based DPD is proposed based on MP/LILUT-based DPD by using quadratic interpolation instead of linear interpolation. The simulation results show MP/QILUT-based DPD is slightly better than MP/LILUT-based DPD with same table size (small size). In addition, NN-based DPD is proposed based on a single layer feedforward neural network. The innovation is the training samples of neural network is obtained from MP/root-finding-based DPD. The trained neural network is taken as the predistorter. Although

the training process of neural network is time-consuming, the trained predistorter is less time-consuming.

6.2 Perspectives

To continue this work, some prospects are suggested for the future work.

1. PA behavioral modeling is very important for the performance of DPD. The generalized memory polynomial is a good model widely used in recent literature. The complexity of GMP is a little higher than that of MP, but far less than that of Volterra series model. How to design a predistorter by combining GMP and LUT should be investigated.
2. Nonuniform LUT technique is very promising for improving our work. It is important to distribute table size effectively in different regions of PA. Theoretically, at the high linearity region and low probability density region, it should distribute less table size. Conversely, at the nonlinear region and high probability density region, it should distribute more table size. How to adaptively distribute table size depending on the characteristic of PA and the probability density function of input signal is worthy of research.
3. In real applications, DPD is usually used with peak-to-average power ratio reduction technique. The combining of DPD and PAPR reduction can further improve the power efficiency of PA. In the future work, implementing a PAPR reduction technique and combining it with the proposed DPD methods are an interesting study orientation.

List of Tables

List of Figures

List of Abbreviations

0G Zero Generation
1G First Generation
2G Second Generation
3G Third Generation
4G Fourth Generation
5G Fifth Generation
ACPR Adjacent Channel Power Ratio
ADC Analog-to-Digital Converter
AM/AM Amplitude-to-Amplitude
AM/PM Amplitude-to-Phase
AMPS Advanced Mobile Phone System
ANN Artificial Neural Network
BB BaseBand
BP Back Propagation
BPF Band-Pass Filter
CDMA Code Division Multiple Access
CFR Crest Factor Reduction
DAC Digital-to-Analog Converter
DC Direct Current
DDR Dynamic Deviation Reduction
DLA Direct Learning Architecture
DO Digital Oscilloscope
DPD Digital PreDistortion
DSP Digital Signal Processing
DTMF Dual Tone Multi-Frequency
EDGE Enhanced Data GSM Environment
EER Envelope Elimination and Restoration
EVM Error Vector Magnitude

FDMA Frequency Division Multiple Access
FFNN Feed Forward Neural Network
FIR Finite Impulse Response
FM Frequency Modulation
FPGA Field-Programmable Gate Array
GMP Generalized Memory Polynomial
GPRS General Packet Radio Service
GSM Global System of Mobile communication
HSDPA High Speed Downlink Packet Access
IF Intermediate Frequency
ILA Indirect Learning Architecture
IMD InterModulation Distortion
IMTS Improved Mobile Telephone Service
IP Intellectual Property
IP3 Third-order Intercept Point
IQ In-phase and Quadrature
IS Interim Standard
ITU-R International Telecommunication Union-Radiocommunication
LDMOS Laterally Diffused Metal Oxide Semiconductor
LILUT Linear-Interpolated Look-Up Table
LINC LInear amplification with Nonlinear Components
LMS Least Mean Squares
LNA Low-Noise Amplifier
LO Local Oscillator
LS Least Squares
LTE Long Term Evolution
LUT Look-Up Table
MLP MultiLayer Perceptron
MP Memory Polynomial
MTS Mobile Telephone Service
MSE Mean Squared Error
NFxLMS Nonlinear Filtered-x LMS
NFxRLS Nonlinear Filtered-x RLS
NMSE Normalized Mean Squared Error
NMT Nordic Mobile Telephone
NN Neural Network

OFDM Orthogonal Frequency Division Multiplexing
PA Power Amplifier
PAE Power Added Efficiency
PAPR Peak to Average Power Ratio
PC Personal Computer
PD PreDistorter
PDC Personal Digital Cellular
PSD Probability Density Function
QAM Quadrature Amplitude Modulation
QILUT Quadratic-Interpolated Look-Up Table
QPSK Quadrature Phase Shift Keying
RBFNN Radial Basis Function artificial Neural Network
RF Radio Frequency
RLS Recursive Least Squares
RNN Recurrent artificial Neural Network
Rx Receiver
SA Spectrum Analyzer
SoC System On Chip
SSPA Solid-State Power Amplifier
TACS Total Access Communication Systems
TD-SCDMA Time Division-Synchronous Code Division Multiple Access
TDMA Time Division Multiple Access
TRx Transceiver
TWTA Traveling-Wave Tube Amplifier
Tx Transmitter
VCO Voltage-Controlled Oscillator
VS Volterra Series
VSG Vector Signal Generator
WCDMA Wide Code Division Multiple Access
WiMAX Worldwide Interoperability for Microwave Access
WWWW World Wild Wireless Web

Publications

Journal

1. **X. Feng**, B. Feuvrie, A.S. Descamps, Y. Wang. Digital predistortion method combining memory polynomial and feed-forward neural network, Electronics Letters, 2015, 51, (12), pp. 943-945.
2. **X. Feng**, B. Feuvrie, A.S. Descamps, Y. Wang. Digital predistortion technique based on non-uniform MP model and interpolated LUT for linearizing PAs with memory effects, Electronics Letters, 2014, 50, (24), pp. 1882-1884.
3. **X. Feng**, B. Feuvrie, A.S. Descamps, Y. Wang. Improved baseband digital predistortion for linearizing PAs with nonlinear memory effects using linearly interpolated LUT, Electronics Letters, 2013, 49, (22), pp. 1389-1391.

Conference

1. **X. Feng**, B. Feuvrie, A.S. Descamps, Y. Wang. A digital predistortion method based on nonuniform memory polynomial model using interpolated LUT, IEEE Radio Wireless Week (RWW2015), Topical Conference on RF/Microwave Power Amplifiers for Radio and Wireless Applications (PAWR2015), 25-28 janvier 2015, San Diego, CA, USA.
2. B. Feuvrie, **X. Feng**, A. S. Descamp, Y. Wang. Une technique efficace de linéarisation d'un amplificateur de puissance à effet mémoire modélisé basé sur un modèle MP, 18èmes Journées Nationales Microondes, 15-17 Mai 2013, Paris, France.
3. **X. Feng**, B. Feuvrie, A.S. Descamps, Y. Wang. An improved digital predistortion method under noisy environment, Second Sino-French Workshop on Education and Research collaborations in Information and Communication Technologies (SIFWICT 2013), 3-4 juin 2013, Guangzhou, China.

Bibliography

[1] E. COTTAIS, Y. WANG, AND S. TOUTAIN. **A new adaptive baseband digital predistortion technique**. *European Microwave Association*, **2**:154–159, 2006. 10, 59, 69, 70, 73, 76, 78

[2] F. LI, B. FEUVRIE, Y. WANG, AND W. CHEN. **MP/LUT baseband digital predistorter for wideband linearisation**. *Electronics Letters*, **47**(19):1096–1098, 2011. 10, 79

[3] X. FENG, B. FEUVRIE, A. S. DESCAMPS, AND Y. WANG. **An improved baseband digital predistortion for linearizing PAs with nonlinear memory effects using linearly interpolated LUT**. *Electronics Letters*, **49**(22):1389–1391, 2013. 10, 56

[4] X. FENG, B. FEUVRIE, A.-S. DESCAMPS, AND Y. WANG. **Digital predistortion technique based on non-uniform MP model and interpolated LUT for linearising PAs with memory effects**. *Electronics Letters*, **50**(24):1882–1884, 2014. 10, 56

[5] N. SRIRATTANA, A. RAGHAVAN, D. HEO, P.E. ALLEN, AND J. LASKAR. **Analysis and design of a high-efficiency multistage Doherty power amplifier for wireless communications**. *IEEE Transactions on Microwave Theory and Techniques*, **53**(3):852–860, March 2005. 16

[6] K. MOON, Y. CHO, J. KIM, S. JIN, B. PARK, D. KIM, AND B. KIM. **Investigation of Intermodulation Distortion of Envelope Tracking Power Amplifier for Linearity Improvement**. *IEEE Transactions on Microwave Theory and Techniques*, **PP**(99):1–10, 2015. 16

[7] F.H. RAAB, P. ASBECK, S. CRIPPS, P.B. KENINGTON, Z.B. POPOVIC, N. POTHECARY, J.F. SEVIC, AND N.O. SOKAL. **Power amplifiers and transmitters for RF and microwave**. *IEEE Transactions on Microwave Theory and Techniques*, **50**(3):814–826, Mar 2002. 16, 44

[8] Y.Y. HUANG, W. WOO, H. JEON, C.H. LEE, AND J.S. KENNEY. **Compact Wideband Linear CMOS Variable Gain Amplifier for Analog-Predistortion Power Amplifiers**. *IEEE Transactions on Microwave Theory and Techniques*, **60**(1):68–76, Jan 2012. 16

[9] J. YI, Y. YANG, M. PARK, W. KANG, AND BUMMAN KIM. **Analog predistortion linearizer for high-power RF amplifiers**. *IEEE Transactions on Microwave Theory and Techniques*, **48**(12):2709–2713, Dec 2000. 16

[10] B. AI, Z. X. YANG, C. Y. PAN, S. G. TANG, AND T. T. ZHANG. **Analysis on LUT Based Predistortion Method for HPA with Memory**. *IEEE Transactions on Broadcasting*, **53**(1):127–131, 2007. 16, 54, 56

[11] F.M. BARRADAS, T.R. CUNHA, P.M. LAVRADOR, AND J.C. PEDRO. **Polynomials and LUTs in PA Behavioral Modeling: A Fair Theoretical Comparison**. *IEEE Transactions on Microwave Theory and Techniques*, **62**(12):3274–3285, Dec 2014. 16, 54, 56

[12] L. DING, G. T. ZHOU, D. R. MORGAN, Z. MA, J. S. KENNEY, J. KIM, AND C. R. GIARDINA. **A robust digital baseband predistorter constructed using memory polynomials**. *IEEE Transactions on Communications*, **52**(1):159–165, 2004. 16, 17, 41, 58, 59, 69, 70

[13] D.R. MORGAN, Z. MA, J. KIM, M.G. ZIERDT, AND J. PASTALAN. **A Generalized Memory Polynomial Model for Digital Predistortion of RF Power Amplifiers**. *IEEE Transactions on Signal Processing*, **54**(10):3852–3860, Oct 2006. 16, 17, 40, 58

[14] F. MKADEM AND S. BOUMAIZA. **Physically Inspired Neural Network Model for RF Power Amplifier Behavioral Modeling and Digital Predistortion**. *IEEE Transactions on Microwave Theory and Techniques*, **59**(4):913–923, 2011. 16, 56, 57

[15] L.M. AGUILAR-LOBO, A. GARCIA-OSORIO, J.R. LOO-YAU, S. ORTEGA-CISNEROS, P. MORENO, J.E. RAYAS-SANCHEZ, AND J.A. REYNOSO-HERNANDEZ. **A digital predistortion technique based on a NARX network to linearize GaN class F power amplifiers**. In *IEEE 57th International Midwest Symposium on Circuits and Systems*, pages 717–720, Aug 2014. 16

[16] W.J. RUGH. *Nonlinear system theory*. The Johns Hopkins University Press, 1981. 17

[17] T. LIU, S. BOUMAIZA, AND F.M. GHANNOUCHI. **Augmented hammerstein predistorter for linearization of broad-band wireless transmitters**. *IEEE*

Transactions on Microwave Theory and Techniques, **54**(4):1340–1349, June 2006. 17

[18] S. CHEN, X. HONG, Y. GONG, AND C.J. HARRIS. **Digital Predistorter Design Using B-Spline Neural Network and Inverse of De Boor Algorithm**. *IEEE Transactions on Circuits and Systems I: Regular Papers*, **60**(6):1584–1594, June 2013. 17

[19] A. ZHU, J.C. PEDRO, AND T.J. BRAZIL. **Dynamic Deviation Reduction-Based Volterra Behavioral Modeling of RF Power Amplifiers**. *IEEE Transactions on Microwave Theory and Techniques*, **54**(12):4323–4332, Dec 2006. 17, 39, 43

[20] M. ABI HUSSEIN, V.A. BOHARA, AND O. VENARD. **On the system level convergence of ILA and DLA for digital predistortion**. In *IEEE International Symposium on Wireless Communication Systems*, pages 870–874, Aug 2012. 17, 18, 59, 63, 98

[21] P. DUHAMEL, M. MONTAZERI, AND K. HILAL. **Classical adaptive algorithms (LMS, RLS, CMA, decision directed) seen as recursive structures**. In *IEEE International Conference on Acoustics, Speech, and Signal Processing*, **3**, pages 496–499 vol.3, April 1993. 17

[22] M. ISAKSSON, D. WISELL, AND D. RONNOW. **A comparative analysis of behavioral models for RF power amplifiers**. *IEEE Transactions on Microwave Theory and Techniques*, **54**(1):348–359, Jan 2006. 17, 63

[23] D. ZHOU AND V.E. DEBRUNNER. **Novel Adaptive Nonlinear Predistorters Based on the Direct Learning Algorithm**. *IEEE Transactions on Signal Processing*, **55**(1):120–133, Jan 2007. 17, 63, 64

[24] O. HAMMI, S. CARICHNER, B. VASSILAKIS, AND F.M. GHANNOUCHI. **Synergetic Crest Factor Reduction and Baseband Digital Predistortion for Adaptive 3G Doherty Power Amplifier Linearizer Design**. *IEEE Transactions on Microwave Theory and Techniques*, **56**(11):2602–2608, Nov 2008. 18

[25] XILINX. **Xilinx PB006 Logicore IP Digital Pre-Distortion v7.0**. *www.xilinx.com*, 2014. 18

[26] H. PAASO AND A. MAMMELA. **Comparison of direct learning and indirect learning predistortion architectures**. In *IEEE International Symposium on Wireless Communication Systems*, pages 309–313, Oct 2008. 18, 59

[27] M. STEER. **Beyond 3G**. *IEEE Microwave Magazine*, **8**(1):76–82, 2007. 22, 23

[28] A. KUMAR, Y.F. LIU, J. SENGUPTA, AND DIVYA. **Evolution of mobile wireless communication networks: 1G to 4G**. *International Journal of Electronics & Communication Technology*, **1**(1):68–72, Dec. 2010. 22, 23

[29] T. MSHVIDOBADZE. **Evolution mobile wireless communication and LTE networks**. *Application of Information and Communication Technologies (AICT), 2012 6th International Conference on*, pages 1–7, Oct. 2012. 22, 23

[30] AKHTAR SHAKIL. **Evolution of Technologies, Standards, and Deployment of 2G-5G Networks**. *White papers, Clyton state university*, 2008. 22, 23

[31] PANKAJ SHARMA. **Evolution of Mobile Wireless Communication Networks-1G to 5G as well as Future Prospective of Next Generation Communication Network**. *International Journal of Computer Science and Mobile Computing*, **2**(8):47–53, 2013. 23

[32] F. LI. *Linéarisation des amplificateurs de puissance dans les systèmes de communication large bande par prédistorsion numérique en bande de base*. PhD thesis, Université de Nantes, Nantes,French, 2012. 23, 29, 65, 75, 107

[33] N. POTHECARY. *Feedforward Linear Power Amplifiers*. Artech House Microwave Library, Hardcover, 1999. 23, 27, 30, 45

[34] A. LUZZATTO AND G. SHIRAZI. *Wireless Transceiver Design: Mastering the Design of Modern Wireless Equipment and Systems*. John Wiley & Sons, Inc, 2007. 23, 25, 26

[35] M. YAJIMA, Y. HISADA, Y. ISHIDA, Y. SAITO, K. YAMAMOTO, S. MURAKAMI, S. TANAKA, N. GOTO, AND K. HONJO. **Trial manufacture of 26 GHz band solid state power amplifier (SSPA) module**. In *High Performance Electron Devices for Microwave and Optoelectronic Applications Workshop*, pages 172–175, Nov 1996. 27

[36] A. SALEH. **Frequency-Independent and Frequency-Dependent Nonlinear Models of TWT Amplifiers**. *IEEE Transactions on Communications*, **29**(11):1715–1720, November 1981. 27, 38

[37] A. GHORBANI AND M. SHEIKHAN. **The effect of solid state power amplifiers (SSPAs) nonlinearities on MPSK and M-QAM signal transmission**. In *Sixth International Conference on Digital Processing of Signals in Communications*, pages 193–197, Sep 1991. 27, 39

[38] F. MKADEM. *Behavioral Modeling and Digital Predistortion of Wide- and Multi-Band Transmitter Systems*. PhD thesis, University of Waterloo, Ontario,Canada, 2014. 28

[39] F. M. GHANNOUCHI AND O. HAMMI. **Behavioral modeling and predistortion**. *IEEE Microwave Magazine*, **10**(7):52–64, 2009. 30, 41

[40] T. LIU, S. BOUMAIZA, A.B. SESAY, AND F.M. GHANNOUCHI. **Quantitative Measurements of Memory Effects in Wideband RF Power Amplifiers Driven by Modulated Signals**. *IEEE Microwave and Wireless Components Letters*, **17**(1):79–81, Jan 2007. 30

[41] L. DING. *Digital Predistortion of Power Amplifiers for Wireless Applications*. PhD thesis, Georgia Institute of Technology, Atlanta,USA, 2004. 34

[42] S.P. STAPLETON AND F.C. COSTESCU. **An adaptive predistorter for a power amplifier based on adjacent channel emissions**. *IEEE Transactions on Vehicular Technology*, **41**(1):49–56, Feb 1992. 39

[43] N. MESSAOUDI, M.C. FARES, S. BOUMAIZA, AND J. WOOD. **Complexity reduced odd-order memory polynomial pre-distorter for 400-watt multi-carrier Doherty amplifier linearization**. In *IEEE MTT-S International Microwave Symposium Digest*, pages 419–422, June 2008. 39

[44] A. ZHU, P.J. DRAXLER, J.J. YAN, T.J. BRAZIL, D.F. KIMBALL, AND P.M. ASBECK. **Open-Loop Digital Predistorter for RF Power Amplifiers Using Dynamic Deviation Reduction-Based Volterra Series**. *IEEE Transactions on Microwave Theory and Techniques*, **56**(7):1524–1534, July 2008. 39, 43

[45] S.L. CHANG AND T. OGUNFUNMI. **LMS/LMF and RLS Volterra system identification based on nonlinear Wiener model**. In *Proceedings of IEEE International Symposium on Circuits and Systems*, **5**, pages 206–209, May 1998. 40

[46] J. WOOD. *Behavioral Modeling and Linearization of RF Power Amplifiers*. Artech House Publishers, 2014. 40

[47] J. TSIMBINOS AND K.V. LEVER. **Computational complexity of Volterra based nonlinear compensators**. *Electronics Letters*, **32**(9):852–854, Apr 1996. 40

[48] B. KIM, J. KIM, I KIM, AND J. CHA. **The Doherty power amplifier**. *IEEE Microwave Magazine*, **7**(5):42–50, Oct 2006. 43

[49] F.F. TAFURI, D. SIRA, T.S. NIELSEN, O.K. JENSEN, J.H. MIKKELSEN, AND J. LARSEN. **Memory models for behavioral modeling and digital predistortion of envelope tracking power amplifiers**. *Microprocessors and Microsystems*, 2015. 43

[50] L. LARSON, D. KIMBALL, AND P. ASBECK. **Wideband envelope tracking power amplifiers for wireless communications**. In *IEEE 14th Topical Meeting on Silicon Monolithic Integrated Circuits in RF Systems*, pages 32–37, Jan 2014. 43

[51] GANG LIU, P. HALDI, TSU-JAE KING LIU, AND A.M. NIKNEJAD. **Fully Integrated CMOS Power Amplifier With Efficiency Enhancement at Power Back-Off**. *Solid-State Circuits, IEEE Journal of*, **43**(3):600–609, Mar. 2008. 43

[52] F. TABATABAI AND H.S. AL-RAWESHIDY. **Feedforward Linearization Technique for Reducing Nonlinearity in Semiconductor Optical Amplifier**. *Journal of Lightwave Technology*, **25**(9):2667–2674, Sept 2007. 44

[53] C.L. LAROSE AND F.M. GHANNOUCHI. **Optimization of feedforward amplifier power efficiency on the basis of drive statistics**. *IEEE Transactions on Microwave Theory and Techniques*, **51**(1):41–54, 2003. 44

[54] G. ZHAO, F.M. GHANNOUCHI, F. BEAUREGARD, AND A.B. KOUKI. **Digital implementations of adaptive feedforward amplifier linearization techniques**. In *IEEE MTT-S International Microwave Symposium Digest*, **2**, pages 543–546 vol.2, 1996. 44

[55] I. MEIER AND J. B. DE SWARDT. **Error-feedback for amplifier linearization**. In *Communications and Signal Processing, 1998. COMSIG '98. Proceedings of the 1998 South African Symposium on*, pages 381–386, 1998. 45

[56] P. GARCIA, A. ORTEGA, J. DE MINGO, AND A. VALDOVINOS. **Nonlinear distortion cancellation using LINC transmitters in OFDM systems**. *IEEE Transactions on Broadcasting*, **51**(1):84–93, 2005. 48

[57] B. SHI AND L. SUNDSTROM. **A novel design using translinear circuit for linear LINC transmitters**. *IEEE International Symposium on Circuits and Systems*, **1**:64–67, 2000. 48

[58] L.R. KAHN. **Single-Sideband Transmission by Envelope Elimination and Restoration**. *Proceedings of the IRE*, **40**(7):803–806, July 1952. 48

[59] F.H. RAAB. **Intermodulation distortion in Kahn-technique transmitters**. *IEEE Transactions on Microwave Theory and Techniques*, **44**(12):2273–2278, Dec 1996. 48

[60] J. MACDONALD. **Nonlinear Distortion Reduction by Complementary Distortion**. *IRE Transactions on Audio*, **7**:128–133, September-October 1959. 52

[61] G. SATOH. **Linearizer for high-power traveling-wave tube amplifier**. *Electronics and Communications in Japan*, **62**(10):702–703, October 1979. 52

[62] T. NOJIMA AND Y. OKAMOTO. **Predistortion Nonlinear Compensator for Microwave SSB-AM System**. *Electronics and Communications in Japan*, **67**(5):57–66, 1984. 52

[63] J.K. CAVERS. **Optimum table spacing in predistorting amplifier linearizers**. *IEEE Transactions on Vehicular Technology*, **48**(5):1699–1705, Sep 1999. 54, 55, 56

[64] K. F. LIANG, J. H. CHEN, AND Y. J.E. CHEN. **A Quadratic-Interpolated LUT-Based Digital Predistortion Technique for Cellular Power Amplifiers**. *IEEE Transactions on Circuits and Systems II: Express Briefs*, **61**(3):133–137, 2014. 54, 56, 84

[65] S. BOUMAIZA, J. LI, M. JAIDANE-SAIDANE, AND F. M. GHANNOUCHI. **Adaptive digital/RF predistortion using a nonuniform LUT indexing function with built-in dependence on the amplifier nonlinearity**. *IEEE Transactions on Microwave Theory and Techniques*, **52**(12):2670–2677, 2004. 54, 55, 56

[66] Y. NAGATA. **Linear amplification technique for digital mobile communications**. In *IEEE Vehicular Technology Conference*, pages 159–164, May 1989. 54

[67] K. FINNERTY, J. DOOLEY, AND R. FARRELL. **Cartesian pre-distortion using a sigma delta modulator for multi-standard RF power amplifiers**. In *IET Irish Signals and Systems Conference*, pages 1–5, June 2012. 54

[68] M. FAULKNER AND M. JOHANSSON. **Adaptive linearization using predistortion-experimental results**. *IEEE Transactions on Vehicular Technology*, **43**(2):323–332, May 1994. 54

[69] S.N. BA, K. WAHEED, AND G.T. ZHOU. **Optimal spacing for a polar look-up table predistorter**. In *IEEE Northeast Workshop on Circuits and Systems*, pages 189–192, Aug 2007. 54

[70] G. COVIELLO, F. CANNONE, AND G. AVITABILE. **Robust behavioral non uniform look-up table spacing in adaptive digital baseband predistortion technique for RF power amplifier**. In *AFRICON, 2013*, pages 1–5, Sept 2013. 54, 55

[71] J.K. CAVERS. **Amplifier linearization using a digital predistorter with fast adaptation and low memory requirements**. *IEEE Transactions on Vehicular Technology*, **39**(4):374–382, Nov 1990. 54

[72] G.E. CORAZZA. *Digital Satellite Communications*. Springer US, 2007. 54

[73] W. MAO, L. RAN, AND K. CHEN. **Adaptive predistortion for RF power amplifier based on new look-up table indexing method**. In *International Conference on Microwave and Millimeter Wave Technology*, pages 932–935, Aug 2002. 55

[74] I. TEIKARI, J. VANKKA, AND K. HALONEN. **Baseband digital predistorter with quadrature error correction**. In *Proceedings of Norchip Conference*, pages 159–162, Nov 2004. 56, 80

[75] G. NORRIS, J. STAUDINGER, J. CHEN, C. REY, P. PRATT, R. SHERMAN, AND H. FRAZ. **Application of Digital Adaptive Pre-distortion to Mobile Wireless Devices**. In *IEEE Radio Frequency Integrated Circuits Symposium*, pages 247–250, June 2007. 56, 80

[76] G. CYBENKO. **Approximation by superpositions of a sigmoid function**. *Mathematics of Control, Signals and Systems*, **2**(4):303–314, 1989. 56, 99

[77] D.R. HUSH AND B.G. HORNE. **Progress in supervised neural networks**. *IEEE Signal Processing Magazine*, **10**(1):8–39, 1993. 56, 99

[78] N. NASKAS AND Y. PAPANANOS. **Neural-network-based adaptive baseband predistortion method for RF power amplifiers**. *IEEE Transactions on Circuits and Systems II: Express Briefs*, **51**(11):619–623, 2004. 56, 58

[79] A. AHMED, E.R. SRINIDHI, AND G. KOMPA. **Efficient PA modeling using neural network and measurement setup for memory effect characterization in the power device**. In *IEEE MTT-S International Microwave Symposium Digest*, pages 473–476, June 2005. 56, 57

[80] T. LIU, S. BOUMAIZA, AND F.M. GHANNOUCHI. **Dynamic behavioral modeling of 3G power amplifiers using real-valued time-delay neural networks**. *IEEE Transactions on Microwave Theory and Techniques*, **52**(3):1025–1033, March 2004. 56, 57

[81] M. ISAKSSON, D. WISELL, AND D. RONNOW. **Wide-band dynamic modeling of power amplifiers using radial-basis function neural networks**. *IEEE Transactions on Microwave Theory and Techniques*, **53**(11):3422–3428, Nov 2005. 56, 57

[82] Q. ZHANG, K.C. GUPTA, AND V.K. DEVABHAKTUNI. **Artificial neural networks for RF and microwave design - from theory to practice**. *IEEE Transactions on Microwave Theory and Techniques*, **51**(4):1339–1350, Apr 2003. 56, 58

[83] A.S. CIMINSKI. **Recurrent neural networks usefulness in digital predistortion of power amplifiers**. In *15th International Conference on Microwaves, Radar and Wireless Communications*, **1**, pages 249–252, May 2004. 56, 58

[84] R. RAICH, H. QIAN, AND G.T. ZHOU. **Orthogonal polynomials for power amplifier modeling and predistorter design**. *IEEE Transactions on Vehicular Technology*, **53**(5):1468–1479, Sept 2004. 58

[85] Y. LIU, J. ZHOU, W. CHEN, AND B. ZHOU. **A Robust Augmented Complexity-Reduced Generalized Memory Polynomial for Wideband RF Power Amplifiers**. *IEEE Transactions on Industrial Electronics*, **61**(5):2389–2401, May 2014. 58

[86] J. KIM AND K. KONSTANTINOU. **Digital predistortion of wideband signals based on power amplifier model with memory**. *Electronics Letters*, **37**(23):1417–1418, 2001. 59

[87] D. ZHOU AND V.E. DEBRUNNER. **Novel Adaptive Nonlinear Predistorters Based on the Direct Learning Algorithm**. *IEEE Transactions on Signal Processing*, **55**(1):120–133, Jan 2007. 59

[88] Y.H. LIM, Y.S. CHO, I.W. CHA, AND H.D. YOUN. **An adaptive nonlinear prefilter for compensation of distortion in nonlinear systems**. *IEEE Transactions on Signal Processing*, **46**(6):1726–1730, Jun 1998. 61, 63

[89] H.W. KANG, Y.S. CHO, AND D.H. YOUN. **On compensating nonlinear distortions of an OFDM system using an efficient adaptive predistorter**. *IEEE Transactions on Communications*, **47**(4):522–526, Apr 1999. 64

[90] J. KWON AND C. EUN. **Digital feedforward compensation scheme for the nonlinear power amplifier with memory**. In *International Conference on Information Technology and Applications*, pages 169–172, 2009. 65

[91] E. WESTESSON AND L. SUNDSTROM. **A complex polynomial predistorter chip in CMOS for baseband or IF linearization of RF power amplifiers**. In *IEEE International Symposium on Circuits and Systems*, **1**, pages 206–209, 1999. 75

[92] R.C. DE LAMARE AND R. SAMPAIO-NETO. **Adaptive reduced-rank MMSE filtering with interpolated FIR filters and adaptive interpolators**. *IEEE Signal Processing Letters*, **12**(3):177–180, March 2005. 80

[93] R.C. DE LAMARE AND R. SAMPAIO-NETO. **Adaptive Reduced-Rank Processing Based on Joint and Iterative Interpolation, Decimation, and Filtering**. *IEEE Transactions on Signal Processing*, **57**(7):2503–2514, July 2009. 80

[94] K. FUNAHASHI. **On the approximate realization of continuous mappings by neural networks**. *Neural Networks*, **2**(3):183–192, 1989. 99

[95] M. ABI HUSSEIN. *Linéarisation des amplificateurs de puissance: prédistorsion numérique adaptative en bande de base*. PhD thesis, Université de Nantes, Nantes,French, 2009. 107

[96] **ZFL-2500**. *Data sheet*. 107

l'université nantes angers le mans
COMMUNAUTÉ D'UNIVERSITÉS ET ÉTABLISSEMENTS

UNIVERSITÉ DE NANTES

Thèse de Doctorat

Xiaowen FENG

Techniques efficaces en bande de base pour linéariser un amplificateur de puissance avec effet mémoire

Efficient baseband digital predistortion techniques for linearizing power amplifier by taking into account nonlinear memory effect

Résumé

Les techniques de pré-distorsion numérique (DPD) en bande de base permettant de linéariser un amplificateur de puissance (PA) avec effets mémoires non linéaires sont étudiées. Des éléments de base liés à la linéarisation du PA sont introduits, tels que le comportement non-linéaire du PA, son influence sur les systèmes de communication, sa modélisation et caractérisation... Puis certaines techniques de linéarisation existantes sont décrites, telles que le recul de puissance, le feedforward, la boucle de retour, l'amplification linéaire avec des composants non linéaires et la DPD. La DPD est la technique de linéarisation la plus prometteuse. L'architecture d'implémentation et les algorithmes d'identification de la DPD sont décrits. Dans cette thèse, quatre méthodes originales de DPD sont proposées. La première méthode (MP/LUT DPD) combine un modèle polynomial à mémoire (MP) avec une simple LUT. L'amplitude et la phase du signal prédistordu sont calculées par cette LUT. La deuxième méthode consiste à ajouter une interpolation linéaire à cette méthode. La troisième méthode consiste à améliorer la technique d'interpolation par une interpolation quadratique. La quatrième méthode exploite conjointement la MP DPD et un réseau de neurones. L'innovation principale dans cette dernière, réside dans l'apprentissage du réseau de neurones avec les échantillons du signal pré-distordu par la MP DPD. Enfin, les résultats de simulation et expérimentaux en termes de linéarisation, d'amélioration spectrale, de rotation de constellation et de temps de calculs sont fournis. Les méthodes proposées permettent différents compromis entre les performances de linéarisation et la complexité.

Mots clés

Amplificateur de puissance, Prédistorsion numérique, Techniques de linéarisation, Effets mémoires non linéaire, Look-up tables, Réseau de neurones.

Abstract

Baseband digital predistortion (DPD) techniques for linearizing power amplifiers (PA) with memory effects are investigated in this thesis. Firstly, the relevant elements concerning PA linearization are introduced, such as PA nonlinearity behavior, its influence on the communication systems, its modeling and characterization... Then some existing linearization techniques are presented, such as power backoff, feedforward, feedback, linear amplification with nonlinear components and DPD. DPD is the most promising linearization technique. After that, the implementation architecture and identification algorithms of DPD are described. In this thesis, four DPD methods are proposed. The first method (MP/LUT DPD) is to combine a memory polynomial (MP) model and a simple non-interpolated LUT. Both the amplitude and phase of the predistorted signal are calculated by LUT. The second method is to add linear interpolation technique to MP/LUT DPD. The third method improves the second one by using a quadratic interpolation technique to MP/LUT DPD. The fourth method is to combine MP DPD and feed forward neural network. The principal innovation is that the training samples of the neural networks are the predistorded signal obtained by MP DPD. Finally, simulation results and experimental results are given and analyzed. The proposed methods provide different trade-off between the linearization performance, time efficiency and complexity.

Key Words

Power amplifier, Digital predistortion, Linearization techniques, Memory effects, Look-up tables, Neural network.