TECHNISCHE UNIVERSITÄT WIEN
Vienna University of Technology

institute of telecommunications

Digital Pre-Distortion Algorithms for Envelope Tracking Power Amplifiers

Master's thesis
submitted in partial fulfillment of the requirements for the degree of
Diplom-Ingenieur
at the Vienna University of Technology,
to the Faculty of Electrical Engineering and Information Technology

submitted by
Stefan Zehetmayer
registration number 0626590

advisors:
Univ.Prof. Dipl.-Ing. Dr.techn. Markus Rupp
Dipl.-Ing. Robert Dallinger

Vienna, April

Stefan Zehetmayer

"The true delight is in the finding out rather than in the knowing".

~Isaac Asimov

I hereby certify that the work reported in this thesis is my own and that the work done by others is appropriately cited.

Stefan Zehetmayer
Vienna, April 2013

Acknowledgements

Before diving into the technical discussions, I would like to thank many people who made my thesis possible in so many ways.

First, I would like to thank my advisors, Dr. Markus Rupp and Dipl.-Ing. Robert Dallinger for their continuous support, expertise and excellent guidance throughout the development of this thesis. I feel very fortunate for the opportunity to have them as my research advisors.

The work was supported by Artesyn Austria GmbH & CoKG, department for advanced power technology. I am grateful to Dipl.-Ing. Andreas Stiedl and Dipl.-Ing. Harald Etlinger, whose fruitful cooperation and experience has been essential to the successful completion of my thesis.

I would like to thank my fellow students for their insightful discussions and all my buddies for their friendship.

Finally, I share the credit of my work to the most important people in my life, my family. Special thanks to my parents, Maria and Josef. Their love and unconditional support has always helped me in all my endeavors. Thanks to my brother, Christian, for his incessant encouragement and motivation when I was out of humor.

Abstract

Modern wireless communication systems have evolved to provide high data-rates to an increasing number of customers. As a part of this evolution, highly spectrally efficient modulation schemes are employed to achieve the desired capacity within the limited radio frequency (RF) resource. Unfortunately, these modulation schemes are sensitive to nonlinear distortions introduced by the transmitter. Digital pre-distortion (DPD) is one of the most cost effective solutions to compensate the system's nonlinear and dispersive behavior. Moreover, operating costs, transceiver portability, and environmental issues apparently impose a need for employing highly efficient power amplifier (PA) designs. The envelope tracking power amplifier (ETPA) is one promising efficiency enhancement technique. Its basic idea is to adapt the PA supply voltage according to the input signal envelope, thus allowing continuous optimal efficiency operation at the expense of linearity. In other words, due to the dynamic supply voltage changes the ETPA exhibits a rather different nonlinear behavior compared to stand-alone PAs. Therefore, conventional DPD techniques have troubles dealing with the nonlinear distortions introduced by the ETPA.

In this thesis, two in literature proposed DPD methods to mitigate the detrimental effects of the ETPA are discussed and compared: The *decomposed piecewise Volterra model* and the *vector-switched model*. Both models are based on several polynomial sub-models. The DPD algorithm based on the two proposed models is implemented in MATLAB for an 47W 2.14GHz ETPA system. Both algorithms have shown to successfully linearize the ETPA for a Wideband Code Division Multiple Access (WCDMA) signal with 10MHz bandwidth and 7.5dB peak-to-average power ratio (PAPR). Experimental results show that the nonlinear behavior and memory effects are effectively compensated and both DPD methods outperform classical models in terms of accuracy.

Kurzfassung

Um den immer weiter steigenden Ansprüchen von Telekommunikationssystemen gerecht zu werden, sind spektral höchst effiziente Modulationsverfahren unerlässlich. Diese Modulationsverfahren sind jedoch sehr anfällig auf nichtlineare Verzerrungen, welche vermehrt durch den Leistungsverstärker verursacht werden. Mittels digitaler Vorverzerrung kann dieses nichtlineare und dispersive Verhalten des Leistungsverstärkers kompensiert werden. Da außerdem die Wirtschaftlichkeit und der ökologische Betrieb von Basisstationen und die Akkulaufzeit portabler Geräte eine übergeordnete Rolle spielen, sind Methoden zur Effizienzsteigerung des Leistungsverstärkers unabdingbar. Mit Hilfe des Envelope Tracking Leistungsverstärkers wird eine merkliche Effizienzsteigerung erzielt. Hierbei wird die Versorgungsspannung des Leistungsverstärkers an die Einhüllende des Sendesignals angepasst, sodass der Leistungsverstärker permanent mit hohem Wirkungsgrad betrieben werden kann. Diese Effizienzsteigung geht jedoch auf Kosten der Linearität. Im Vergleich zu Leistungsverstärkern mit konstanter Versorgungsspannung, zeigt der Envelope Tracking Leistungsverstärker ein stark abweichendes Verhalten, sodass konventionelle Vorverzerrungsalgorithmen schwer im Stande sind dieses nichtlineare Verhalten ausreichend zu kompensieren.

In dieser Arbeit werden zwei aktuelle Methoden aus der Literatur zur Vorverzerrung von Envelope Tracking Leistungsverstärkern erörtert und verglichen. Das *decomposed piecewise Volterra Modell* und das *vector-switched Modell*. Beide Modelle verwenden mehrere Sub-Modelle die Polynome zur Beschreibung des Verhaltens verwenden. Der Vorverzerr Algorithmus ist implementiert in MATLAB und ausgelegt für einen 47W 2.14GHz Envelope Tracking Leistungsverstärker. Beide Algorithmen linearisierten erfolgreich das Verhalten des Envelope Tracking Leistungsverstärkers, für ein Wideband Code Division Multiple Access (WCDMA) Eingangssignal mit 10MHz Bandbreite und 7.5dB Peak-to-Average Power Ratio (PAPR). Weiters zeigten die Messergebnisse, dass beide Algorithmen konventionellen Vorverzerr Methoden bezüglich Genauigkeit überlegen sind.

Contents

1

Chapter 1

Introduction

Contents

1.1 Motivation

The demand for broadband wireless access has been booming for the last years [1]. To satisfy the market, modern wireless communication systems have been steadily meliorating in order to provide higher data-rates to an ever-growing number of customers. Obviously, high-speed data communication requires a broad modulation bandwidth, which is restricted by the limited and crowded radio frequency (RF) resource. Therefore, modern wireless systems have to allocate the valuable and sparse RF spectrum resource-efficient while offering reliable transmission. Hence, spectrally efficient modulation techniques, such as Wideband Code Division Multiple Access (WCDMA) or orthogonal frequency-division multiple access (OFDMA), e.g., in 3rd Generation (3G) mobile telecommunication networks and Long Term Evolution (LTE), respectively, have been employed to attain the desired system capacity and robustness. In general, these modulation schemes make use of non-constant envelope modulation techniques that usually exhibit signals having large envelope variations, resulting in a high peak-to-average power ratio (PAPR)[1].

Concerning signal integrity preservation and the spreading of the signal bandwidth due to intermodulation distortions, i.e., spectral re-growth, linear amplification for non-constant envelope modulation signals is a necessity because these techniques are very susceptible to nonlinear distortions. Specifically, nonlinear amplification causes out-of-band as well as in-band distortions. While in-band distortions can be treated by utilizing forward error correction (FEC) codes, out-of-band emissions for example in present 3G mobile telecommunication networks are handled by introducing so-called guard bands [2]. However, modern wireless broadband systems will impose tougher spectral efficiency requirements, so these guard bands will be shrinking [3], which in turn will restrict the spectral re-growth margin. One possibility to mitigate nonlinear distortions is to ensure quasi distortion-free amplification, i.e., linear amplification, by, e.g., backing-off the output power level in order not to excite the power amplifier (PA) nonlinearities. Yet, this degrades the power efficiency, since a PA is most efficient at high output power levels. Simply speaking, conventional PA designs are either very linear or very efficient, both cannot be reached at once. They show a trade-off between efficiency and linearity inherent to the PA design. In other words, efficiency enhancements are attained at the expense of linearity and vice versa [4].

Most energy of a base transceiver station (BTS) is consumed at the PA stage (cf. Figure 1.1). Therefore, the energy efficiency of a BTS can be considerably improved by using more efficient PA designs. Moreover, all the energy consumers listed in Figure 1.1 are coupled by temperature; using more efficient PAs and power supplies, less energy is dissipated as heat, which in turn reduces the energy-hungry cooling efforts. Therefore, the PA is the most critical part of the transmitter's final stage and in general optimized

[1] PAPR of a time-continuous signal $x(t)$ with length T

$$\mathrm{PAPR} \triangleq 10\log_{10}\left(\frac{\max\limits_{0\leq t\leq T}\left\{|s(t)|^2\right\}}{\frac{1}{T}\int_0^T |s(t)|^2\mathrm{d}t}\right) \quad \text{in dB}$$

for efficiency [5]. One promising technique for high efficiency operation for non-constant envelope signals is the envelope tracking power amplifier (ETPA), where the supply voltage of the PA is varied according to the transmit signal envelope. It allows to operate the PA continually in the compression regime [6]. To preserve the regained efficiency and to compensate for added detrimental intrinsic nonlinear effects, it is necessary to employ supplemental linearization techniques, e.g., digital pre-distortion (DPD). Where a pre-distorter is concatenated prior to the PA in order to compensate the nonlinear distortions and memory effects and thus, spectral re-growth [7].

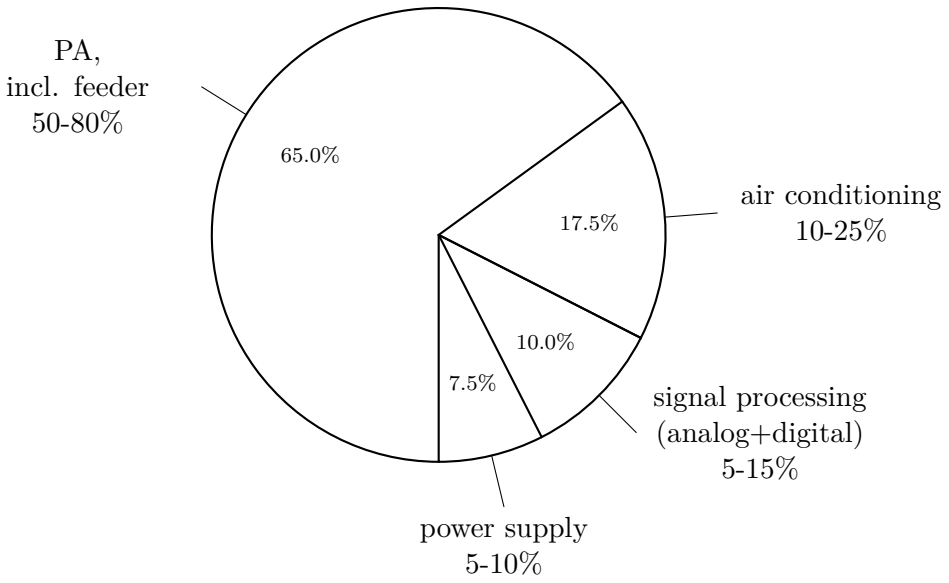


Figure 1.1.: Power consumption partitioning of a BTS [8].

1.2 Power Amplifier Linearity

For constant envelope modulation schemes like Global System for Mobile Communications (GSM), where Gaussian Minimum Shift Keying (GMSK) is utilized, linear amplification is not a crucial concernment. Since the RF signal amplitude conveys no information, amplitude distortions do not affect data transmission. In contrast, non-constant envelope modulation schemes require linear amplification, because the RF signal amplitude bears a part of the data to be transmitted.

Linear amplification implies that the output signal is a scalar multiple of the input signal, i.e.,

$$x_{\text{in}}(t) = G_0 x_{\text{out}}(t), \tag{1.1}$$

where $x_{\text{in}}(t)$ and $x_{\text{out}}(t)$ denote the input and output *baseband signals*, respectively, and G_0 is the *signal gain* (hereinafter denoted as *gain*) of the PA. Accordingly, no distortions are introduced, since the waveform of $x_{\text{out}}(t)$ is proportional to that of $x_{\text{in}}(t)$. In other words, the output is a scaled replica of the input [9].

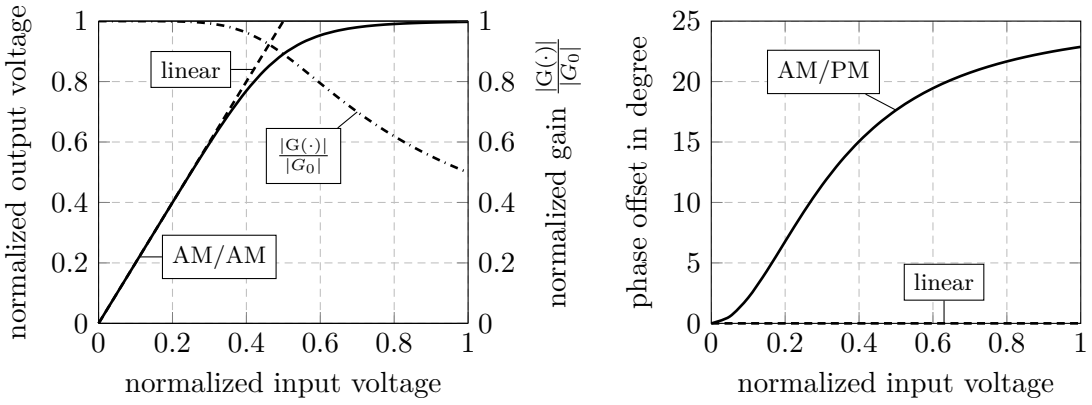
Pursuant to the law of conservation of energy that states that the total amount of energy remains constant entirely, the sum of powers delivered to the PA must equal the sum of emergent powers. As a matter of fact, the *input signal* is amplified with gain G_0, and analogously the *input power* with *power gain* G_P, i.e.,

$$G_\mathrm{P} \triangleq \frac{P_\mathrm{out}}{P_\mathrm{in}} = 1 + \frac{P_\mathrm{DC} - P_\mathrm{diss}}{P_\mathrm{in}}, \tag{1.2}$$

where P_out and P_in denote the RF output and input powers, respectively, P_diss the dissipated power and P_DC the supply power. Since P_DC is limited and P_diss cannot get negative, G_P cannot remain constant for any increasing input power. Thus, the PA shows gain compression above some input power level and the gain depends on $x_\mathrm{in}(t)$ [5], i.e.,

$$x_\mathrm{out}(t) = G(x_\mathrm{in}(t))x_\mathrm{in}(t), \tag{1.3}$$

where $G(\cdot)$ is a complex-valued nonlinear function. In fact, the PA can be characterized by the so-called amplitude modulation/amplitude modulation (AM/AM) and amplitude modulation/phase modulation (AM/PM) conversions. Where the AM/AM and AM/PM conversions are nonlinear functions mapping from $|x_\mathrm{in}(t)|$ to $|x_\mathrm{out}(t)|$ and from $|x_\mathrm{in}(t)|$ to the phase difference $\arg\{x_\mathrm{out}(t)\} - \arg\{x_\mathrm{in}(t)\}$, respectively. Typical AM/AM and AM/PM conversions are depicted in Figure 1.2 [6]. Accordingly, for small input power levels the PA habitually operates rather linear, while for increasing input power levels the AM/AM conversion starts to show nonlinear behavior because of gain compression as described before. However, if the PA behaves linear, the AM/AM conversion is linear with a positive slope given by $|G_0|$ (cf. dashed line in Figure 1.2a), while the AM/PM conversion is a constant $\arg\{G_0\}$ (cf. dashed line in Figure 1.2b).



(a) AM/AM conversion. **(b)** AM/PM conversion.

Figure 1.2.: A typical PA characteristic. The dashed line represents a perfect linear PA. The dash-dotted line indicates the normalized gain.

Heretofore, only static nonlinearities were present, i.e., the PA instantaneous complex-valued gain $G(\cdot)$ was a function of the actual input solely. However, PAs in general exhibit

memory effects, also known as dynamic distortions where the PA characteristics deviate from its static behavior. In other words, $G(\cdot)$ is as of now a function of the instantaneous input signal as well as its preceding history. Hence, the AM/AM and AM/PM conversions are no longer one-to-one functions, since the same input signal value possibly results in different output signal values depending on the history of the input signal. Among others, the three main causes for present memory effects are unintentional modulation on supply rails, dynamic thermal effects, and semiconductor trapping effects [6]. Thus, memory effects can be either thermal or electric. Thermal effects arise due to fluctuations of the dissipated energy caused by varying signal levels, therefore evoking volatile temperature changes that alter the electrical characteristics of the PA stage. Whereas these fluctuations are very slow and of lowpass type and its effects are present at bandwidths up to 1MHz, electric memory effects even more affect wide-band systems with bandwidths over 5MHz, because of the frequency dependent impedance of the bias- and matching-networks and trapping effects of the semiconductor. However, wide-band systems also include low frequency components that are affected by thermal memory effects [10]. On the whole, the nonlinear and non-static behavior of the PA is influenced by the transmit signal statistics (e.g., represented by the PAPR or the complementary cumulative distribution function (CCDF)) and the signal bandwidth.

Beside AM/AM and AM/PM conversions, other figures of merit exist that describe the linearity of PAs, e.g., adjacent channel power ratio (ACPR) (cf. Section 2.5.2) or the third order intercept point (IP3). They yield a fixed distortion quantity that eases analog front-ends comparability [9].

1.3 Power Amplifier Efficiency

The efficiency of PAs in wireless communication systems is very important. Not only in BTSs, where the PA is the major energy consumer (cf. Figure 1.1), also in mobile devices where battery life-time is crucial to customers. Therefore, designers tend to optimize the analog front-end for efficiency and accept the loss of linearity [5].

There are two widely used metrics for PA efficiency, the drain efficiency and the power added efficiency (PAE). The drain efficiency is defined as

$$\eta_{\mathrm{D}} \triangleq \frac{P_{\mathrm{out}}}{P_{\mathrm{DC}}}, \tag{1.4}$$

and can be interpreted as the capability of converting the supply power P_{DC} in RF output power P_{out}. However, the PAE is a richer indicator because it takes the input power P_{in} into account, i.e.,

$$\mathrm{PAE} \triangleq \frac{P_{\mathrm{out}} - P_{\mathrm{in}}}{P_{\mathrm{DC}}} = \eta_{\mathrm{D}} \left(1 - \frac{1}{G_p}\right). \tag{1.5}$$

Figure 1.3 delineates a typical PAE characteristic of a conventional solid state Class-A PA [6]. In general, the efficiency increases with increasing output power and reaches its

maximum at the saturation regime of the PA, before it starts to decline. Accordingly, for constant-envelope modulation schemes the PA is designed to operate consistently at maximum efficiency for the constant envelope level. Whereas for non-constant envelope modulation schemes, having high dynamic output power level ranges, i.e., high PAPR, the PA is incapable of permanent operation at its maximum efficiency. Figure 1.3 also depicts the envelope probability density function (PDF) of an LTE modulated signal with 20MHz bandwidth and 8.5dB PAPR [11]. Accordingly, lower output power levels are more likely and therefore, the PA operates in average at low efficiency. In other words, the probability that the PA operates far off to the saturation point, and thus, less efficient, is much higher than highly efficient amplification. Consequently, to improve the average efficiency for non-constant envelope modulation schemes, efficiency not only needs to be improved close to saturation but also for lower output power levels.

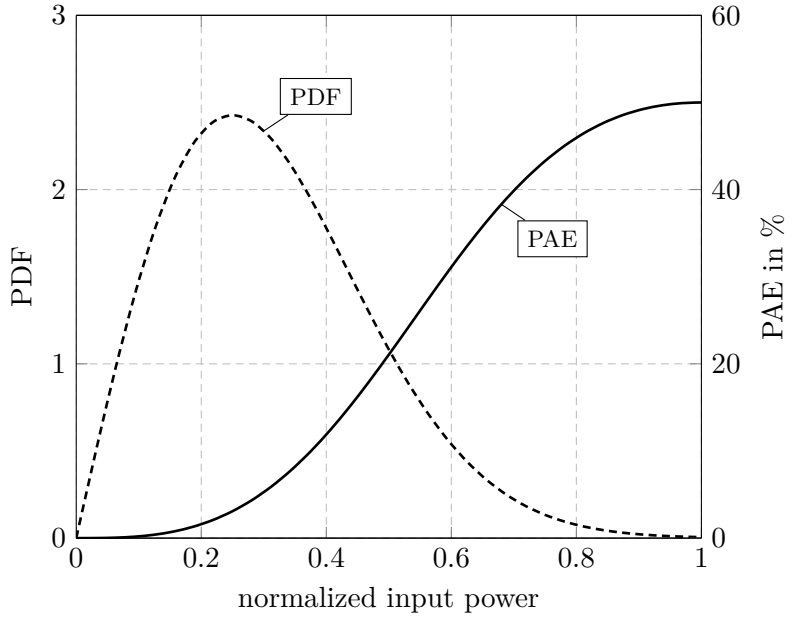


Figure 1.3.: PAE of a typical Class-A PA (solid line), versus the envelope PDF (dashed line) of an LTE modulated signal with 20MHz bandwidth and 8.5dB PAPR [11].

1.4 Efficiency Enhancement by Envelope Tracking

In average, conventional PAs with fixed supply voltage show low efficiency when amplifying signals with high PAPR. One way to improve the average efficiency is the envelope tracking power amplifier (ETPA) as depicted in Figure 1.4.

In the ETPA arrangement, the supply voltage of the PA is dynamically adapted by the envelope amplifier. In order to do so, the envelope signal is first either detected from the RF input signal, or immediately calculated in the baseband. Accordingly, the

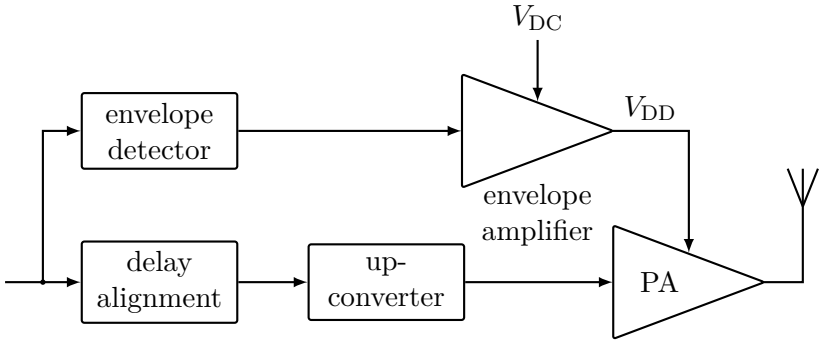
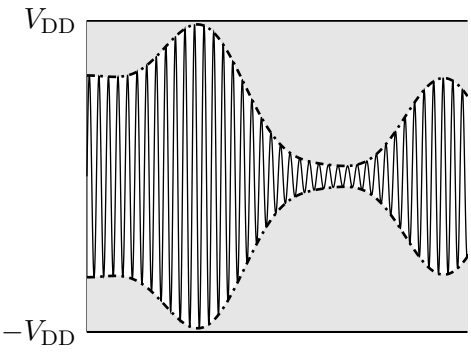
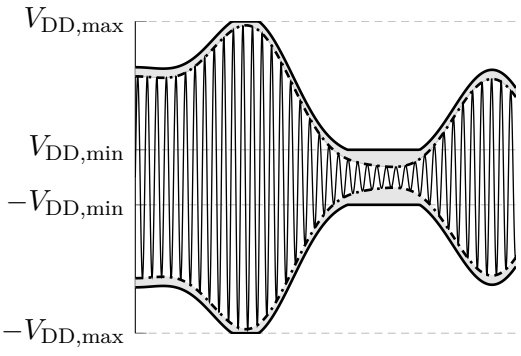


Figure 1.4.: Block diagram of an ETPA system.

envelope amplifier provides the PA with a dynamic supply voltage such that efficiency is kept near its maximum. In other words, the ETPA continuously operates close or near to saturation for all envelope levels due to the full rail-to-rail voltage swing as illustrated in Figure 1.5b. Moreover, Figure 1.5 provides a deeper insight in the operating principle of the ETPA compared to a conventional PA [6]. Again, while the supply voltage V_{DD} of a conventional PA remains constant throughout operation (cf. Figure 1.5a), the supply voltage of an ETPA is varied among the minimum and maximum supply voltage, $V_{DD,min}$ and $V_{DD,max}$, respectively, according to the instantaneous envelope of the RF signal (cf. Figure 1.5b). Hence, the ETPA continuously operates highly efficient, since less energy is dissipated as heat.



(a) Time behavior of a conventional PA.



(b) Time behavior of an ETPA.

Figure 1.5.: Operating principle of an ETPA compared to that of a conventional PA. The dash-dotted line represents the envelope of the signal to be transmitted. The solid line depicts the supply voltage. The gray area corresponds to the energy loss.

Whereas in general conventional linear amplifiers (e.g., Class-A, Class-AB or Class-B [6]) are utilized as PAs, the envelope amplifier is optimized for efficiency (e.g., Class-D, Class-E or Class-F [6]), since its only application is to provide the PA with sufficient

supply voltage for highly efficient amplification. Figure 1.6 illustrates the PAE for the same PA as in Figure 1.3, however, utilized in an envelope tracking system. Obviously, the efficiency is conspicuously improved for lower output power levels, which in turn exceedingly improves the average efficiency for the depicted LTE modulated signal with 20MHz bandwidth and 8.5dB PAPR [11]. Clearly, the ETPA operates at almost maximum efficiency, regardless of the input power level.

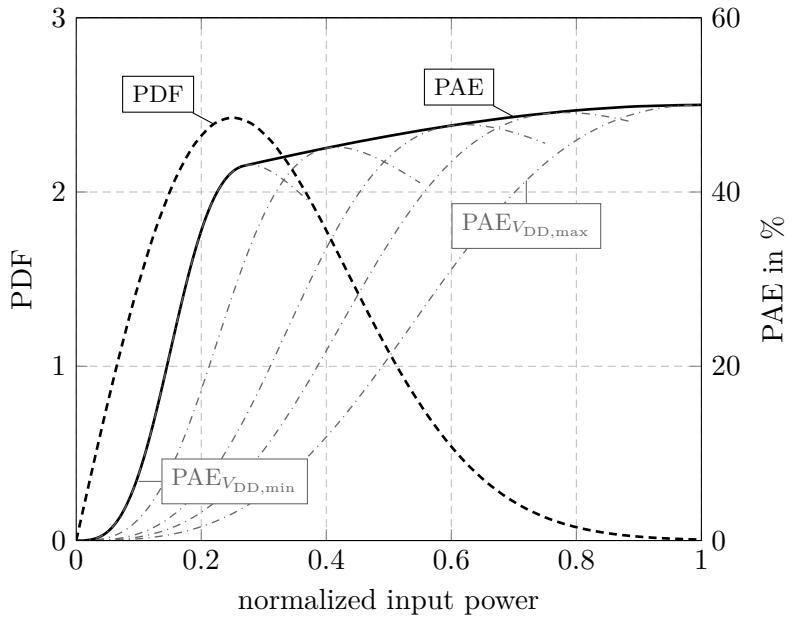


Figure 1.6.: A possible PAE trajectory of an ETPA (solid line), where the supply voltage is varied between $V_{\mathrm{DD,min}}$ and $V_{\mathrm{DD,max}}$, leading to PAE curves that vary between $\mathrm{PAE}_{V_{\mathrm{DD,min}}}$ and $\mathrm{PAE}_{V_{\mathrm{DD,max}}}$, respectively (cf. Figure 1.3 and Figure 1.5) [12], versus the envelope PDF (dashed line) of an LTE modulated signal with 20MHz bandwidth and 8.5dB PAPR [11].

Beside efficiency, linearity is an intrinsic design goal [5]. The nonlinear behavior of an ETPA is rather different to that of a conventional PA with fixed supply voltage, due to its dynamically changing supply voltage and the limited envelope tracking bandwidth. Moreover envelope tracking can be thought of as a nonlinear mixing between the supply voltage and the RF signal. In other words, the ETPA shows different detrimental behaviors at distinct output power levels. Additionally, whenever the input power approaches zero, nonlinearities are excited, because the supply voltage level is only varied down to a certain minimum supply voltage level $V_{\mathrm{DD,min}}$, in order to avoid supply voltage dropping to zero. Moreover, vast memory effects and nonlinearities arise because of a possible misalignment between the envelope tracked supply voltage and the RF signal. Even though both, the envelope tracking stage and the delay line are linear systems, they constitute a source of nonlinear distortion [13, 14]. Therefore, the use of linearizers, e.g., DPD, becomes an essential solution to mitigate nonlinear distortions arising from the

use of more efficient but highly nonlinear PAs.

1.5 Thesis Outline

This Thesis is organized as follows

Chapter 1 - Introduction, delineates the need of highly linear PAs for non-constant envelope modulation schemes. Moreover, it describes the principle of the ETPA as one way to enhance the power efficiency. Finally, the average efficiency of the ETPA is compared to conventional stand-alone PAs when amplifying signals with high PAPR.

Chapter 2 - Behavioral Modeling of Power Amplifiers, formulates the theory of the passband baseband transformation and ways to transform passband PA models to equivalent baseband models. The presented PA models include both, memoryless and models with memory. The classical Volterra series are introduced as the most common means for describing PAs. Moreover, ways to prune the Volterra series are discussed. Additionally, more sophisticated models like the decomposed piecewise Volterra model and the vector-switched models are described. Finally, figures of merit that evaluate the accuracy of the presented models are introduced.

Chapter 3 - Model Parameter Estimation, derives the least squares (LS) estimator for Volterra series based PA models. Moreover, the LS estimator is adapted to the decomposed piecewise Volterra model and the vector-switched model. An excursion to adaptive estimation algorithms concludes the chapter.

Chapter 4 - Linearization by Digital Pre-Distortion, introduces the concept of DPD as an efficient linearization technique. In order to linearize the PA, the direct-learning and the indirect-learning architectures are presented as methods for the identification of the inverse system. Finally, signal bandwidth considerations and means to choose the overall system gain are discussed.

Chapter 5 - Measurement Results, evaluates the performance of the discussed DPD models based on measurements of an prototype envelope tracking system operating at 2.14GHz with an average output power of 47W. The ETPA was driven with a 10MHz WCDMA signal with 7.5dB PAPR.

2

Chapter 2

Behavioral Modeling of Power Amplifiers

Contents

2.1 Introduction and Outline

In the light of calculations and simulations, physical models are used to model the characteristics of telecommunication systems at circuit-level or transistor-level of description. However, sophisticated radio frequency (RF) and wireless communication systems are far too complex for complete real-time circuit simulations and theoretical calculations. Therefore, behavioral models (also referred to as black-box models) are used to model the system properties, in order to allow accurate simulations at a relatively high level of abstraction [15, 16]. The main advantage of modeling at high level of abstraction is that no specific and deep knowledge about the functionality and physics of the RF circuit is presupposed. Furthermore, the identification of the model parameters is based on output observations when to RF circuit is driven with a certain input signal. In other words, behavioral modeling simplifies the modeling to a mathematical input-output relation of favored complexity. Its performance is principally influenced by the model formulation and the accuracy of the behavioral observations [17]. Specifically, model formulation corresponds to the choice of an applicable mathematical representation that comprises all relevant input-output interactions of the RF circuit. Whereas model observation corresponds to accurate measurements of the RF circuit input and output signals. Additionally, the behavior of interest has to be properly excited in order to extract the behavioral model parameters. Accordingly, the behavioral model accuracy is highly affected by the model structure and the parameter extraction procedure. Although the behavioral model likely represents the RF circuit characteristics for the set of applied training signals, it is not guaranteed that the same behavioral model performs properly for a different set of signals. Obviously, different model formulations and distinct training signal sets will lead to diverse simulation results. Therefore, in contrast to physical models, the generalization of results obtained by simulations with behavioral models should always be viewed with circumspection [18].

This Chapter is organized as follows

Section 2.2 - Baseband Representation of Passband Signals, shows that real-valued passband signals can be equivalently represented as complex-valued baseband signals. Moreover, the transformation to pass back and forth between the passband and baseband domains is derived.

Section 2.3 - Memoryless Power Amplifier Models, gives insights into memoryless behavioral models. In fact, the memoryless baseband polynomial behavioral model is derived.

Section 2.4 - Power Amplifier Models with Memory, presents the Volterra series as one of the most important and general power amplifier (PA) model that incorporates memory effects. Moreover, baseband equivalent representations are derived and properties of the Volterra series are discussed. Furthermore, ways to prune the Volterra series, i.e., the memory polynomials, the Hammerstein model, and the

dynamic deviation reduction-based Volterra series, are introduced. Finally, more sophisticated models that make use of the afore derived pruned Volterra series are discussed.

Section 2.5 - Figures of Merit to Evaluate the Model Performance, derives and lists in-band performance measures, i.e., the normalized mean square error (NMSE), and out-of-band measures such as the adjacent channel power ratio (ACPR) and the adjacent channel error power ratio (ACEPR).

2.2 Baseband Representation of Passband Signals

PAs are passband devices [6]. In fact, behavioral simulations in the passband require high computational costs and are hardly feasible. Therefore, the most convenient workaround of representing passband systems is to work with equivalent baseband quantities [19]. Nevertheless, equivalent baseband behavioral models allow only simulations for the complex-valued envelope baseband signal, any arising effects from the carrier frequency must be individually incorporated [18].

In the passband, the transmit signal is a *real-valued* bandpass signal spectrally centered around the carrier frequency f_c with transmission bandwidth $B \ll f_\mathrm{c}$ and can be described by the *polar representation*

$$\tilde{x}(t) \triangleq \hat{x}(t)\cos(2\pi f_\mathrm{c} t + \varphi_{\tilde{x}}(t)) = \mathrm{Re}\Big\{\hat{x}(t)\mathrm{e}^{\jmath(2\pi f_\mathrm{c} t + \varphi_{\tilde{x}}(t))}\Big\}, \tag{2.1}$$

where $\hat{x}(t) \geq 0$ and $\varphi_{\tilde{x}}(t)$ denote the time-dependent *instantaneous envelope* and *instantaneous phase* of the passband signal $\tilde{x}(t)$, respectively. Moreover, the Fourier transform[1] $\tilde{X}(f)$ shows a symmetry around zero frequency and occupies the frequency band $\left[f_\mathrm{c} - \frac{B}{2}, f_\mathrm{c} + \frac{B}{2}\right]$ and the *symmetric* frequency band $\left[-f_\mathrm{c} - \frac{B}{2}, -f_\mathrm{c} + \frac{B}{2}\right]$ at negative frequencies (cf. Figure 2.1a). Since $\tilde{x}(t)$ is real-valued, without loss of generality, it can be uniquely reconstructed using only the positive frequency spectrum, which is the cisoidal (or analytic) signal [19]

$$\tilde{x}_\mathrm{A}(t) \triangleq \mathcal{A}\{\tilde{x}(t)\} = \tilde{x}(t) + \jmath\mathcal{H}\{\tilde{x}(t)\}, \tag{2.2}$$

where $\mathcal{H}\{\cdot\}$ denotes the Hilbert transform[2] that is a linear time-invariant (LTI) system with transfer function $H(f) = -\jmath\,\mathrm{sgn}(f)$. Therefore, the Fourier transform of the analytic signal can be written as

$$\tilde{X}_\mathrm{A}(f) = X(f) + \jmath H(f)X(f) = (1 + \mathrm{sgn}(f))\,X(f) = 2\,\mathrm{u}(f)X(f), \tag{2.3}$$

[1] Fourier transform $X(f)$ of a signal $x(t)$

$X(f) \triangleq \mathcal{F}\{x(t)\} = \int_{-\infty}^{\infty} x(t)\mathrm{e}^{-\jmath 2\pi f t}\mathrm{d}t$

[2] Hilbert transform of a signal $x(t)$

$\mathcal{H}\{x(t)\} \triangleq \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{x(\xi)}{t-\xi}\mathrm{d}\xi$

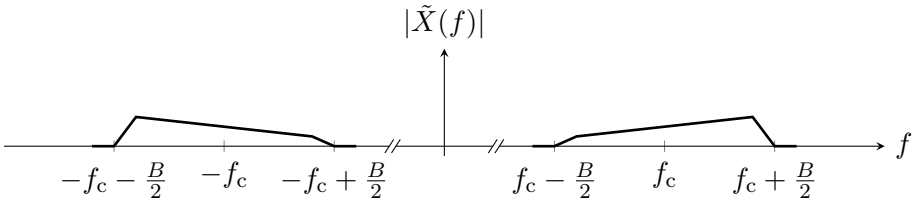
where $\mathrm{u}(\cdot)$ denotes the unit step function. Thus, the transfer function of the overall system $\mathcal{A}$ is $A(f) = 2\,\mathrm{u}(f)$ and the cisoidal signal $\tilde{x}_\mathrm{A}(t)$ is a *complex-valued* bandpass signal that includes the positive frequency band $\left[f_\mathrm{c} - \frac{B}{2}, f_\mathrm{c} + \frac{B}{2}\right]$ solely (cf. Figure 2.1b) [19]. In order to perform the entire passband baseband transformation (cf. Figure 2.1c), the carrier is suppressed and the equivalent complex-valued signal is

$$x(t) \triangleq \frac{1}{\sqrt{2}} \tilde{x}_\mathrm{A}(t) \mathrm{e}^{-\jmath 2\pi f_\mathrm{c} t}, \tag{2.4}$$

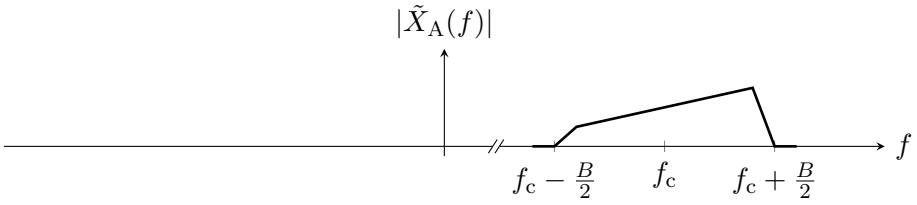
where the factor $\frac{1}{\sqrt{2}}$ is included for energy preservation considerations. The entire passband baseband transformation in the frequency domain, i.e.,

$$X(f) = \frac{1}{\sqrt{2}} \tilde{X}_\mathrm{A}(f + f_\mathrm{c}) = \sqrt{2}\,\mathrm{u}(f + f_\mathrm{c}) \tilde{X}(f + f_\mathrm{c}), \tag{2.5}$$

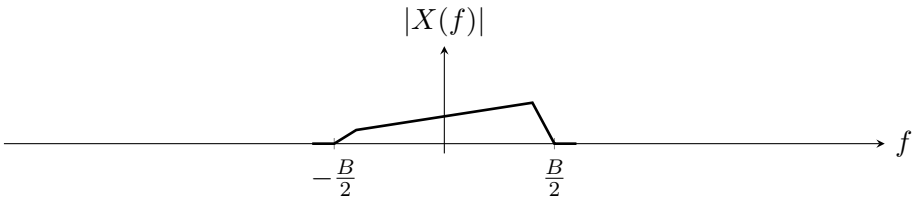
is summarized and illustrated in Figure 2.1.



(a) Spectrum of the passband signal $\tilde{x}(t)$.



(b) Spectrum of the analytical signal $\tilde{x}_A(t)$.



(c) Spectrum of the equivalent baseband signal $x(t)$.

Figure 2.1.: Obtaining the complex-valued baseband spectrum that is equivalent to the real-valued passband spectrum.

Conversely, the passband signal $\tilde{x}(t)$ can be recovered by simply taking the real part

of the frequency-shifted equivalent baseband signal $x(t)$ [20], i.e.,

$$\tilde{x}(t) \triangleq \sqrt{2}\,\mathrm{Re}\Big\{x(t)\mathrm{e}^{\jmath 2\pi f_c t}\Big\} = \frac{1}{\sqrt{2}}\left(x(t)\mathrm{e}^{\jmath 2\pi f_c t} + x^*(t)\mathrm{e}^{-\jmath 2\pi f_c t}\right), \tag{2.6}$$

again with coefficient $\sqrt{2}$ in order to force the energies of $\tilde{x}(t)$ and $x(t)$ to be equal. This baseband passband transformation can be easily verified [20], since

$$\sqrt{2}\,\mathrm{Re}\left\{x(t)\mathrm{e}^{\jmath 2\pi f_c t}\right\} = \sqrt{2}\,\mathrm{Re}\left\{\frac{1}{\sqrt{2}}\tilde{x}_{\mathrm{A}}(t)\right\} = \mathrm{Re}\left\{\tilde{x}(t) + \jmath\mathcal{H}\left\{\tilde{x}(t)\right\}\right\} = \tilde{x}(t). \tag{2.7}$$

Moreover, the passband signal can be written as

$$\tilde{x}(t) = \sqrt{2}x_{\mathrm{I}}(t)\cos(2\pi f_c t) - \sqrt{2}x_{\mathrm{Q}}(t)\sin(2\pi f_c t), \tag{2.8}$$

where

$$x_{\mathrm{I}}(t) \triangleq \mathrm{Re}\left\{x(t)\right\}, \tag{2.9}$$

and

$$x_{\mathrm{Q}}(t) \triangleq \mathrm{Im}\left\{x(t)\right\}. \tag{2.10}$$

This expresses the passband signal $\tilde{x}(t)$ with regard to the *inphase* and *quadrature components*, $x_{\mathrm{I}}(t)$ and $x_{\mathrm{Q}}(t)$, respectively [19]. Compared with the polar representation (2.1), the inphase and quadratur components can also be described as

$$x_{\mathrm{Q}}(t) = \frac{1}{\sqrt{2}}\hat{x}(t)\cos\left(\varphi_{\tilde{x}}(t)\right), \qquad x_{\mathrm{I}}(t) = \frac{1}{\sqrt{2}}\hat{x}(t)\sin\left(\varphi_{\tilde{x}}(t)\right), \tag{2.11}$$

and conversely,

$$\hat{x}(t) = \sqrt{2\left(x_{\mathrm{Q}}^2(t) + x_{\mathrm{I}}^2(t)\right)}, \qquad \varphi_{\tilde{x}}(t) = \tan^{-1}\left(\frac{x_{\mathrm{I}}(t)}{x_{\mathrm{Q}}(t)}\right). \tag{2.12}$$

2.3 Memoryless Power Amplifier Models

In the passband, a memoryless (or static) PA can be described as a nonlinear function that maps the real-valued input to a real-valued output [18]. In fact, the most intuitive way to describe this nonlinear mapping is to use a polynomial representation, i.e.,

$$\tilde{x}_{\mathrm{out}}(t) \triangleq \sum_{p=0}^{P} \tilde{a}_p \tilde{x}_{\mathrm{in}}^p(t), \tag{2.13}$$

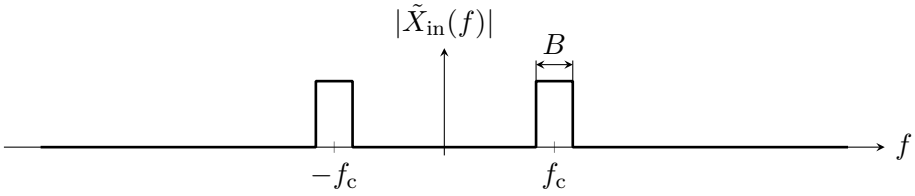
where $\tilde{x}_{\mathrm{in}}(t)$ and $\tilde{x}_{\mathrm{out}}(t)$ denote the passband input and output signals of the PA, respectively, and $\tilde{a}_p$ are the real-valued distortion coefficients. According to (2.6), the passband input signal $\tilde{x}_{\mathrm{in}}(t)$ can be written in the form

$$\tilde{x}_{\mathrm{in}}(t) = \sqrt{2}\,\mathrm{Re}\left\{x_{\mathrm{in}}(t)\mathrm{e}^{\jmath 2\pi f_c t}\right\} = \frac{1}{\sqrt{2}}\left(x_{\mathrm{in}}(t)\mathrm{e}^{\jmath 2\pi f_c t} + x_{\mathrm{in}}^*(t)\mathrm{e}^{-\jmath 2\pi f_c t}\right), \tag{2.14}$$

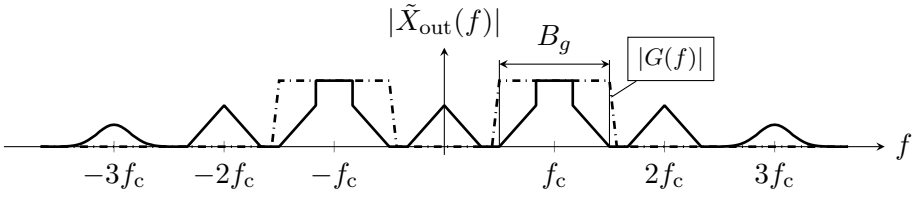
as illustrated in Figure 2.2a. Thus, by using the Binomial identity[1], the passband output signal is

$$\begin{aligned}\tilde{x}_{\text{out}}(t) &= \sum_{p=0}^{P} \frac{\tilde{a}_p}{\sqrt{2^p}} \left(x_{\text{in}}(t) \mathrm{e}^{\jmath 2\pi f_c t} + x_{\text{in}}^*(t) \mathrm{e}^{-\jmath 2\pi f_c t} \right)^p \\ &= \sum_{p=0}^{P} \sum_{k=0}^{p} \frac{\tilde{a}_p}{\sqrt{2^p}} \binom{p}{k} \left(x_{\text{in}}(t) \right)^{p-k} \left(x_{\text{in}}^*(t) \right)^k \mathrm{e}^{\jmath 2\pi (p-2k) f_c t}.\end{aligned} \tag{2.15}$$

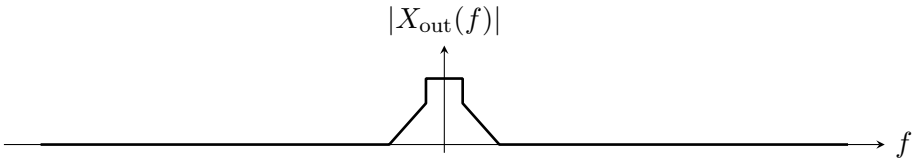
Therefore, $\tilde{x}_{\text{out}}(t)$ contains further frequency components that are located in the surroundings of f_c and are not present in $\tilde{x}_{\text{in}}(t)$. Moreover, $\tilde{x}_{\text{out}}(t)$ consists of spectral components around the harmonics of the carrier frequency f_c, i.e., $0, \pm f_c, \pm 2f_c, \ldots, \pm P f_c$, each of which with bandwidth larger than B due to the introduced intermodulation distortions.



(a) Spectrum of the passband input signal $\tilde{x}_{\text{in}}(t)$ [21].



(b) Spectrum of the intermediate passband output signal $\tilde{x}_{\text{out}}(t)$ before filtered by the zonal filter $g(t)$, in order to ensure bandpass signals [21].



(c) Spectrum of the distorted baseband output signal $x_{\text{out}}(t)$.

Figure 2.2.: An example of the frequency response of a memoryless PA and its equivalent baseband characterization [21].

As you can see in Figure 2.2b, $\tilde{x}_{\text{out}}(t)$ is no longer a bandpass signal. However, by

[1] Binomial identity

$$(x+y)^n \triangleq \sum_{k=0}^{n} \binom{n}{k} x^{n-k} y^k = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$$

filtering with a so-called zonal filter $g(t)$ all the frequency components other than those centered at $\pm f_\mathrm{c}$ are suppressed, which again leads to a bandpass signal. In other words, the zonal filter can be thought of as an ideal bandpass filter with bandwidth B_g and center frequency f_c after the polynomial model that removes all of its out-of-band frequency components. Precisely, B_g is as large as necessary to comprise the spectral broadening due to the nonlinear amplification, i.e., $B_g = \left(2\lceil\frac{P}{2}\rceil - 1\right)B \ll f_\mathrm{c}$. To demonstrate, let us focus on two frequency components in $\tilde{x}_\mathrm{in}(t)$, i.e., $\mathrm{e}^{\jmath 2\pi f_1 t}$ and $\mathrm{e}^{\jmath 2\pi f_2 t}$, which are distorted with p-th nonlinearity [22], i.e.,

$$\begin{aligned}\left(\mathrm{e}^{\jmath 2\pi f_1 t} + \mathrm{e}^{\jmath 2\pi f_2 t}\right)^p &= \sum_{k=0}^{p}\binom{p}{k}\left(\mathrm{e}^{\jmath 2\pi f_1 t}\right)^k\left(\mathrm{e}^{\jmath 2\pi f_2 t}\right)^{p-k} \\ &= \sum_{k=0}^{p}\binom{p}{k}\mathrm{e}^{\jmath 2\pi(kf_1+(p-k)f_2)t}.\end{aligned} \tag{2.16}$$

In fact, not all resulting frequency components $\mathrm{e}^{\jmath 2\pi(kf_1+(p-k)f_2)t}$ fall in the neighborhood of f_c and are thus, removed by the zonal filter $g(t)$. However, there are four cases to be considered [22, 21].

- $f_1, f_2 \in \left[-f_\mathrm{c} - \frac{B}{2}; -f_\mathrm{c} + \frac{B}{2}\right]$.
  It follows that $p\left(-f_\mathrm{c} - \frac{B}{2}\right) \leq kf_1 + (p-k)f_2 \leq p\left(-f_\mathrm{c} + \frac{B}{2}\right)$ is a frequency band in the neighborhood of $-pf_\mathrm{c}$ and is therefore suppressed by the zonal filter for $p \geq 2$.
- $f_1, f_2 \in \left[f_\mathrm{c} - \frac{B}{2}; f_\mathrm{c} + \frac{B}{2}\right]$.
  Similar to the first case, the resulting frequencies surround pf_c and are thus, suppressed by the zonal filter for $p \geq 2$.
- $f_1 \in \left[-f_\mathrm{c} - \frac{B}{2}; -f_\mathrm{c} + \frac{B}{2}\right] \wedge f_2 \in \left[f_\mathrm{c} - \frac{B}{2}; f_\mathrm{c} + \frac{B}{2}\right]$.
  Here, it follows that $(p-2k)f_\mathrm{c} - p\frac{B}{2} \leq kf_1 + (p-k)f_2 \leq (p-2k)f_\mathrm{c} + p\frac{B}{2}$. Since $kf_1 + (p-k)f_2$ falls in the neighborhood of $(p-2k)f_\mathrm{c}$, for p *even*, $p - 2k \neq 1$, thus $kf_1 + (p-k)f_2$ is out of band B_g for all even p. However, for p *odd*, it is entirely possible that $p - 2k = 1$ and the resulting frequencies can be in the passband of the zonal filter.
- $f_1 \in \left[f_\mathrm{c} - \frac{B}{2}; f_\mathrm{c} + \frac{B}{2}\right] \wedge f_2 \in \left[-f_\mathrm{c} - \frac{B}{2}; -f_\mathrm{c} + \frac{B}{2}\right]$.
  Similar to the third case, only for p *odd* the resulting frequencies $kf_1 + (p-k)f_2$ may fall into the passband of the zonal filter.

In general, for all frequency components in $\tilde{x}_\mathrm{in}(t)$, and assuming that the frequency shifted components of $\tilde{x}_\mathrm{out}(t)$ at the harmonics $\mathrm{e}^{\jmath 2\pi(p-2k)f_\mathrm{c}t}$ do not overlap, i.e., $f_\mathrm{c} > \frac{pB}{2}$ (cf. Figure 2.2b) [21], only terms with $p - 2k = 1$ contribute to the system output. Accordingly, the output of the zonal filter can only become nonzero when p is *odd* [20]. When $p - 2k = 1$, we have $k = \frac{p-1}{2}$ and $p - k = \frac{p-1}{2} + 1$, thus

$$\left(x_\mathrm{in}(t)\right)^{p-k}\left(x_\mathrm{in}^*(t)\right)^k = x_\mathrm{in}(t)|x_\mathrm{in}(t)|^{p-1}, \tag{2.17}$$

and the *baseband memoryless polynomial model* (cf. Figure 2.2c) is

$$x_{\text{out}}(t) \triangleq \sum_{p=1}^{\frac{P+1}{2}} a_p x_{\text{in}}(t)|x_{\text{in}}(t)|^{2(p-1)}, \tag{2.18}$$

with the corresponding baseband distortion coefficients

$$a_p \triangleq \frac{1}{\sqrt{2}^{p-1}} \binom{p}{\frac{p-1}{2}} \tilde{a}_p, \tag{2.19}$$

which are real-valued, since $\tilde{a}_p$ are real valued too. Figure 2.2 illustrates and summarizes the derivation of the baseband memoryless behavioral model.

2.4 Power Amplifier Models with Memory

Conventional memoryless behavioral models are frequency independent and yield acceptable performance for narrow-band systems. However, as the signal bandwidth increases, memory effects become more apparent (cf. Section 1.2). In other words, the PA shows a frequency-dependency [18]. These arising memory effect have to be incorporated in the behavioral models in order to model the PA for wide-band systems accurately.

2.4.1 The Truncated Volterra Series

In 1887, the Italian mathematician Vito Volterra was the first who introduced the representation of what now is commonly known as the Volterra series [18]. However, Norbert Wiener was the first who used the Volterra series in order to describe nonlinear systems [23, 24]. The Volterra series is a multivariate polynomial series that includes the actual and preceding signal values, i.e.,

$$\tilde{x}_{\text{out}}(t) \triangleq \tilde{h}_0 + \sum_{p=1}^{\infty} \iint \cdots \int \tilde{h}_p(t, \tau_1, \tau_2, \ldots, \tau_p) \prod_{i=1}^{p} \tilde{x}_{\text{in}}(\tau_i) \mathrm{d}\tau_1 \mathrm{d}\tau_2 \cdots \mathrm{d}\tau_p, \tag{2.20}$$

where $\tilde{x}_{\text{in}}(t)$ and $\tilde{x}_{\text{out}}(t)$ denote the real-valued passband input and output signals, respectively, $\tilde{h}_p(t, t_1, t_2, \ldots, t_p)$ for $p \geq 1$ are the so-called p-th order Volterra kernels, and $\tilde{h}_0$ is a constant, i.e., the zero order Volterra kernel. However, in this thesis, the truncated and time-invariant Volterra series

$$\tilde{x}_{\text{out}}(t) \triangleq \sum_{p=1}^{P} \iint \cdots \int \tilde{h}_p(\tau_1, \tau_2, \ldots, \tau_p) \prod_{i=1}^{p} \tilde{x}_{\text{in}}(t - \tau_i) \mathrm{d}\tau_1 \mathrm{d}\tau_2 \cdots \mathrm{d}\tau_p, \tag{2.21}$$

are used for PA modeling, where the kernels $\tilde{h}_p(t_1, t_2, \ldots, t_p)$ are assumed to be time-independent, and the nonlinearity order is truncated to a determined amount P, i.e., $\tilde{h}_p(t_1, t_2, \ldots, t_p) = 0$ for $p > P$, and $\tilde{h}_0$ is assumed to be zero.

2.4.2 Properties of the Volterra Series

Kernel Linearity

Any LTI system with memory can be described by the linear convolution [19]

$$x_{\text{out},1}(t) = \int_{-\infty}^{\infty} h_1(\tau) x_{\text{in}}(t-\tau) \mathrm{d}\tau, \tag{2.22}$$

where $x_{\text{out},1}(t)$ and $x_{\text{in}}(t)$ are the output and input, respectively, and $h_1(t)$ denotes the one-dimensional impulse response of the LTI system. Obviously, this representation corresponds to the first order Volterra functional.

Similarly, a two-dimensional LTI system can be described as

$$x_{\text{out},2}(t_1, t_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h_2(\tau_1, \tau_2) x_{\text{in}}(t_1 - \tau_1, t_2 - \tau_2) \mathrm{d}\tau_1 \mathrm{d}\tau_2, \tag{2.23}$$

again with input and output, $x_{\text{in}}(t_1, t_2)$ and $x_{\text{out},2}(t_1, t_2)$, respectively, and second order system impulse response $h_2(t_1, t_2)$. Moreover, assuming that $x_{\text{in}}(t_1, t_2)$ can be factorized, i.e., $x_{\text{in}}(t_1, t_2) = x_{\text{in}}(t_1) x_{\text{in}}(t_2)$ and $t = t_1 = t_2$, the output of the two-dimensional LTI system reads

$$x_{\text{out},2}(t) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h_2(\tau_1, \tau_2) x_{\text{in}}(t - \tau_1) x_{\text{in}}(t - \tau_2) \mathrm{d}\tau_1 \mathrm{d}\tau_2. \tag{2.24}$$

Again, this representation corresponds to the second order Volterra functional [21].

In a similar manner, the p-th order impulse response of an LTI system can described in the form

$$x_{\text{out},p}(t) = \iint_{\mathbb{R}^p} \cdots \int h_p(\tau_1, \tau_2, \ldots, \tau_p) \prod_{i=1}^{p} x_{\text{in}}(t - \tau_i) \mathrm{d}\tau_1 \mathrm{d}\tau_2 \cdots \mathrm{d}\tau_p. \tag{2.25}$$

In fact, combining all Volterra functional as

$$x_{\text{out}}(t) = \sum_{p=1}^{P} x_{\text{out},p}(t), \tag{2.26}$$

the truncated Volterra series representation (2.21) is obtained. Obviously, the output of the nonlinear Volterra system $x_{\text{out}}(t)$ is constructed by *linear operations* (summations and integrations) on the products $x_{\text{in}}(t_1) x_{\text{in}}(t_2) \cdots x_{\text{in}}(t_p)$. Although $x_{\text{out},p}(t)$ for $p > 1$ is nonlinear with respect to $x_{\text{in}}(t)$, the output $x_{\text{out},p}(t)$ is linear in relation to the Volterra kernel $h_p(\tau_1, \tau_2, \ldots, \tau_p)$ [21, 24].

Kernel Symmetry

It can be shown in [24] that the Volterra kernels can be assumed to be symmetric, e.g., $h_2(t_1, t_2) = h_2(t_2, t_1)$. For an asymmetric Volterra kernel, in general $p!$ different Volterra kernel permutations are possible. Therefore, any asymmetric Volterra kernel can be symmetrized by

$$h_{p,\mathrm{sym}}(t_1, t_2, \ldots, t_p) = \frac{1}{p!} \sum_{i=0}^{p!} h_{p,\mathrm{asymm}}(t_{\iota(1)_i}, t_{\iota(2)_i}, \ldots, t_{\iota(p)_i}), \tag{2.27}$$

where $\iota(k)_i$ denotes the k-th element in the i-th permutation of the set $\{1, 2, \ldots, p\}$. So there is no loss of generality suffered by considering the Volterra kernel to be symmetric.

Stability

A system is called bounded-input bounded-output (BIBO) stable if every system output is bounded for every bounded input. Thus, for a bounded input signal, i.e., $|x_{\mathrm{in}}(t)| < C$, the output of the p-th order Volterra functional is bounded as

$$\begin{aligned} |x_{\mathrm{out},p}(t)| &= \left| \iint_{\mathbb{R}^p} \cdots \int h_p(\tau_1, \tau_2, \ldots, \tau_p) \prod_{i=1}^{p} x_{\mathrm{in}}(t - \tau_i) \mathrm{d}\tau_1 \mathrm{d}\tau_2 \cdots \mathrm{d}\tau_p \right| \\ &\leq \iint_{\mathbb{R}^p} \cdots \int |h_p(\tau_1, \tau_2, \ldots, \tau_p)| \prod_{i=1}^{p} |x_{\mathrm{in}}(t - \tau_i)| \mathrm{d}\tau_1 \mathrm{d}\tau_2 \cdots \mathrm{d}\tau_p \\ &< C^p \iint_{\mathbb{R}^p} \cdots \int |h_p(\tau_1, \tau_2, \ldots, \tau_p)| \, \mathrm{d}\tau_1 \mathrm{d}\tau_2 \cdots \mathrm{d}\tau_p < \infty. \end{aligned} \tag{2.28}$$

Therefore, a Volterra system is said to be BIBO stable if for every Volterra kernel $h_p(\tau_1, \tau_2, \cdots, \tau_p)$ the condition

$$\iint_{\mathbb{R}^p} \cdots \int |h_p(\tau_1, \tau_2, \ldots, \tau_p)| \, \mathrm{d}\tau_1 \mathrm{d}\tau_2 \cdots \mathrm{d}\tau_p < \infty \qquad p = 1, \ldots, P, \tag{2.29}$$

is fulfilled. This is a sufficient but not necessary condition for BIBO stability [24].

Fourier Transform

It can be shown that the multivariate Fourier transform of the p-th Volterra functional $x_{\mathrm{out},p}(t)$ is [24]

$$X_{\mathrm{out},p}(f_1, f_2, \ldots, f_p) \triangleq H_p(f_1, f_2, \ldots, f_p) \prod_{i=1}^{p} X_{\mathrm{in}}(f_i), \tag{2.30}$$

where $X_{\mathrm{in}}(f)$ is the Fourier transform of the input signal $x_{\mathrm{in}}(t)$, and $H_p(f_1, f_2, \ldots, f_p)$ is the multivariate Fourier transform Volterra kernel, i.e.,

$$H_p(f_1, f_2, \ldots, f_p) \triangleq \iint_{\mathbb{R}^p} \cdots \int h_p(\tau_1, \tau_2, \ldots, \tau_p) e^{-\jmath 2\pi \sum_{i=1}^{p} f_i \tau_i} \mathrm{d}\tau_1 \mathrm{d}\tau_2 \cdots \mathrm{d}\tau_p. \tag{2.31}$$

Note that since the Volterra kernel $h_p(t_1, t_2, \ldots, t_p)$ is symmetric, the corresponding Fourier transform $H_p(f_1, f_2, \ldots, f_p)$ is a symmetric function of its arguments.

The inverse Fourier transform can be described as

$$x_{\text{out},p}(t) \triangleq \frac{1}{(2\pi)^p} \iint_{\mathbb{R}^p} \cdots \int X_{\text{out},p}(f_1, f_2, \ldots, f_p) \mathrm{e}^{\jmath 2\pi \sum_{i=1}^{p} f_i t} \mathrm{d}f_1 \mathrm{d}f_2 \cdots \mathrm{d}f_p. \tag{2.32}$$

2.4.3 Baseband Representation of a Volterra System

For passband signals, the general theory of Volterra series can be transformed to the baseband in exact the same manner as for the memoryless case described in Section 2.3. For a real-valued bandpass signal with bandwidth B that is spectrally centered around the carrier frequency $f_c \gg B$, i.e.,

$$\tilde{x}_{\text{out}}(t) = \sqrt{2}\,\mathrm{Re}\left\{x_{\text{out}}(t)\mathrm{e}^{\jmath 2\pi f_c}\right\}, \tag{2.33}$$

it can be shown that the equivalent *baseband Volterra series* correspond to [21, 23]

$$x_{\text{out}}(t) \triangleq \sum_{p=1}^{\frac{P+1}{2}} \iint_{\mathbb{R}^p} \cdots \int h_{2p-1}(\tau_1, \tau_2, \ldots, \tau_p) \prod_{i=1}^{p} x_{\text{in}}(t-\tau_i) \prod_{i=p+1}^{2p-1} x_{\text{in}}^*(t-\tau_i) \mathrm{d}\tau_1 \mathrm{d}\tau_2 \cdots \mathrm{d}\tau_p, \tag{2.34}$$

where $x_{\text{out}}(t)$ and $x_{\text{in}}(t)$ denote the equivalent baseband output and input signals, respectively, P is the necessarily odd nonlinear order (cf. Section 2.3), and

$$h_p(t_1, t_2, \ldots, t_p) \triangleq \frac{1}{\sqrt{2}^{p-1}} \binom{p}{\frac{p-1}{2}} \tilde{h}_p(t_1, t_2, \ldots, t_p) \mathrm{e}^{\jmath 2\pi \left(\sum_{i1=1}^{\frac{p+1}{2}} t_{i_1} - \sum_{i_2=\frac{p+1}{2}+1}^{p} t_{i_2} \right)}, \tag{2.35}$$

are the equivalent baseband Volterra kernels that are in general complex-valued although the passband Volterra kernels are real-valued. Obviously, in (2.34) only odd order powers remain, since even order distortion products are suppressed by the zonal filter $g(t)$ (cf. Figure 2.2b) [18].

2.4.4 The Baseband Discrete-Time Volterra Model

In order to model nonlinear systems, discrete-time representations of the Volterra series are indispensable for signal processing. Therefore, the baseband Volterra model (2.34) has to be transformed to the discrete-time domain. Specifically, $x_{\text{out}}(t)$ has to be sampled. The following derivations are based on Dallinger's and Aschbacher's theses, [21] and [23], respectively.

According to Nyquist's sampling theorem for equidistant sampling, a time-continuous function $x(t)$ can be uniquely represented without loss of information by its sampling

values $x(nT_\mathrm{s})$ if it is band-limited with $B < \frac{1}{2T_\mathrm{s}}$ and sampling rate T_s. In other words, the corresponding discrete-time signal is

$$x[n] \triangleq x(nT_\mathrm{s}) \qquad n \in \mathbb{Z}, \tag{2.36}$$

with the corresponding discrete Fourier transform (DFT)[1]

$$X[\theta] \triangleq \frac{1}{T_\mathrm{s}} \sum_{-\infty}^{\infty} X\left(\frac{\theta + 2\pi n}{T_\mathrm{s}}\right). \tag{2.37}$$

The inverse relation reads

$$X(f) \triangleq \begin{cases} T_\mathrm{s} X[fT_\mathrm{s}], & -\frac{1}{2T_\mathrm{s}} \leq f < \frac{1}{2T_\mathrm{s}} \\ 0, & \text{otherwise} \end{cases} \tag{2.38}$$

With (2.30), the Fourier transform of the p-th baseband Volterra functional can be described as

$$X_{\mathrm{out},p}(f) = H_p(f_1, f_2, \ldots, f_p) \prod_{i=1}^{\frac{p+1}{2}} X_\mathrm{in}(f_i) \prod_{i=\frac{p+1}{2}+1}^{p} X_\mathrm{in}^*(-f_i). \tag{2.39}$$

Note, since the baseband Volterra series representation (2.34) contains conjugate complex versions of the input signal, i.e., $x_\mathrm{in}^*(t)$, the corresponding Fourier transform is $X_\mathrm{in}^*(-f)$. Additionally, the p-th order Volterra kernel $h_p(t_1, t_2, \ldots, t_p)$ can also be assumed to be band-limited with the spectrum of the p-dimensional hypercube

$$\mathcal{C}_p = \underbrace{[B;B) \times [B;B) \times \cdots \times [B;B)}_{p\text{-times}} = \left[-\frac{1}{2T_s}; \frac{1}{2T_s}\right)^p, \tag{2.40}$$

since the input signal $x_\mathrm{in}(t)$ is assumed to be band-limited with bandwidth B, the Volterra kernel is only excited within $\mathcal{C}_p$, eventhough it might be non-zero outside of $\mathcal{C}_p$. Therefore, the Fourier transform of the p-th baseband Volterra functional can be written as

$$\begin{aligned} x_{\mathrm{out},p}(t) = & \frac{1}{(2\pi)^p} \iint_{\mathcal{C}_p} \cdots \int H_p(f_1, f_2, \ldots, f_p) \mathrm{e}^{\jmath 2\pi \sum_{i=1}^{p} f_i t} \\ & \prod_{i=1}^{\frac{p+1}{2}} X_\mathrm{in}(f_i) \prod_{i=\frac{p+1}{2}+1}^{p} X_\mathrm{in}^*(-f_i) \mathrm{d}f_1 \mathrm{d}f_2 \cdots \mathrm{d}f_p. \end{aligned} \tag{2.41}$$

[1] DFT $X[\theta]$ of a discrete-time signal $x[n]$

$$X[\theta] \triangleq \mathcal{F}_D\{x[n]\} = \sum_{n=-\infty}^{\infty} x[n] \mathrm{e}^{-\jmath\theta n}$$

Moreover, $h_p(t_1, t_2, \ldots, t_p)$ can be sampled with a regular lattice of dimension p and grid-spacing T_s. The corresponding discrete-time Volterra kernel reads

$$h_p[n_1, n_2, \ldots, n_p] = h_p(n_1 T_\mathrm{s}, n_2 T_\mathrm{s}, \ldots, n_p T_\mathrm{s}), \tag{2.42}$$

and its multivariate DFT is

$$\begin{aligned} H_p[\theta_1, \theta_2, \ldots, \theta_p] &= \mathcal{F}_D \left\{ h_p[n_1, n_2, \ldots, n_p] \right\} \\ &= \sum_{n_1=-\infty}^{\infty} \sum_{n_2=-\infty}^{\infty} \cdots \sum_{n_p=-\infty}^{\infty} h_p[n_1, n_2, \ldots, n_p] \mathrm{e}^{-\jmath 2\pi \sum_{i=1}^{p} \theta_i n_i}, \end{aligned} \tag{2.43}$$

and corresponds to the time-continuous Volterra kernel as

$$H_p(f_1, f_2, \ldots, f_p) \triangleq \begin{cases} T_\mathrm{s}^p H_p[f_1 T_\mathrm{s}, f_2 T_\mathrm{s}, \ldots, f_p T_\mathrm{s}], & \{f_1, f_2, \ldots, f_p\} \in \mathcal{C}_p \\ 0, & \{f_1, f_2, \ldots, f_p\} \notin \mathcal{C}_p \end{cases}. \tag{2.44}$$

Therefore, the p-th time-continuous baseband Volterra functional reads

$$\begin{aligned} x_{\mathrm{out},p}(t) = &\frac{T_\mathrm{s}^p}{(2\pi)^p} \sum_{m_1=-\infty}^{\infty} \sum_{m_2=-\infty}^{\infty} \cdots \sum_{m_p=-\infty}^{\infty} h_p[m_1, m_2, \ldots, m_p] \\ &\prod_{i=1}^{\frac{p+1}{2}} x_\mathrm{in}[n - m_i] \prod_{i=\frac{p+1}{2}+1}^{p} x_\mathrm{in}^*[n - m_i] \iint_{\mathcal{C}_p} \cdots \int \mathrm{e}^{\jmath 2\pi \sum_{i=1}^{p} f_i t (m_i + n_i) T_s} \mathrm{d}f_1 \mathrm{d}f_2 \cdots \mathrm{d}f_p, \end{aligned} \tag{2.45}$$

as a function of the sampled input signal $x_\mathrm{in}[n]$ and the p-th sampled Volterra kernel $h_p[n_1, n_2, \ldots, n_p]$. The integral in (2.45) can be solved as follows

$$\iint_{\mathcal{C}_p} \cdots \int \mathrm{e}^{\jmath 2\pi \sum_{i=1}^{p} f_i t (m_i + n_i) T_s} \mathrm{d}f_1 \mathrm{d}f_2 \cdots \mathrm{d}f_p = \left(\frac{2\pi}{T_s} \right)^p \prod_{i=1}^{p} \operatorname{sinc} \left(\pi \left(\frac{t}{T_s} - m_i - n_i \right) \right). \tag{2.46}$$

Finally, (2.45) can be described as

$$\begin{aligned} x_\mathrm{out}(t) = &\sum_{p=1}^{\frac{P+1}{2}} \sum_{m_1=-\infty}^{\infty} \sum_{m_2=-\infty}^{\infty} \cdots \sum_{m_p=-\infty}^{\infty} h_{2p-1}[m_1, m_2, \ldots, m_p] \\ &\prod_{i=1}^{p} x_\mathrm{in}[n - m_i] \prod_{i=p+1}^{2p-1} x_\mathrm{in}^*[n - m_i] \prod_{i=1}^{p} \operatorname{sinc} \left(\pi \left(\frac{t}{T_s} - m_{2i-1} - n \right) \right). \end{aligned} \tag{2.47}$$

Since $x_\mathrm{out}(t)$ is sampled with sampling rate T_s, and $\operatorname{sinc}(\pi k)$ for $k \in \mathbb{Z}$ is a Nyquist pulse, the finally obtained *baseband discrete time Volterra model* reads

$$\begin{aligned} x_\mathrm{out}[n] \triangleq \mathcal{V} \left\{ x_\mathrm{in}[n] \right\} = &\sum_{p=1}^{\frac{P+1}{2}} \sum_{m_1=0}^{M} \sum_{m_2=0}^{M} \cdots \sum_{m_p=0}^{M} \sum_{m_{p+1}=0}^{M} \sum_{m_{p+2}=0}^{M} \cdots \sum_{m_{2p-1}=0}^{M} \\ &h_{2p-1}[m_1, m_2, \ldots, m_{2p-1}] \prod_{i=1}^{p} x_\mathrm{in}[n - m_i] \prod_{i=p+1}^{2p-1} x_\mathrm{in}^*[n - m_i], \end{aligned} \tag{2.48}$$

where M denotes the memory length, and P is the odd nonlinearity order. However, considering Volterra kernel symmetry, (2.48) simplifies to

$$x_{\text{out}}[n] = \sum_{p=1}^{\frac{P+1}{2}} \sum_{m_1=0}^{M} \sum_{m_2=m_1}^{M} \cdots \sum_{m_p=m_{p-1}}^{M} \sum_{m_{p+1}=0}^{M} \sum_{m_{p+2}=m_{p+1}}^{M} \cdots \sum_{m_{2p-1}=m_{2p-2}}^{M} h_{2p-1}[m_1, m_2, \ldots, m_{2p-1}] \prod_{i=1}^{p} x_{\text{in}}[n-m_i] \prod_{i=p+1}^{2p-1} x_{\text{in}}^*[n-m_i]. \tag{2.49}$$

2.4.5 Reduced Baseband Volterra Models

The number of Volterra kernel coefficients increases exponentially as the nonlinearity order P and the memory length M increase. Obviously, implementing the Volterra series in real-time applications can be expected to be unreasonable. Therefore, reducing the complexity of the Volterra series by simply considering its important terms solely reduces computational costs. Moreover, not all Volterra kernel coefficients have the same effect on the output of the model; some of them have less impact since they might be very small. Therefore, it is reasonable to set them to zero during Volterra kernel extraction (cf. Section 3.3). This complexity reduction is often named as reduced or pruned Volterra series modeling [25].

Memory Polynomials

Probably the simplest way to prune the Volterra series is to keep the diagonal terms of the Volterra kernels and set the off-diagonal coefficients to zero. This model is named *memory polynomials* [26], i.e.,

$$x_{\text{out}}[n] \triangleq \mathcal{V}_{\text{MP}}\{x_{\text{in}}[n]\} = \sum_{p=1}^{\frac{P+1}{2}} \sum_{m=0}^{M} h_{2p-1}[m, \ldots, m] x_{\text{in}}[n-m] |x_{\text{in}}[n-m]|^{2(p-1)}, \tag{2.50}$$

where $\mathcal{V}_{\text{MP}}\{\cdot\}$ denotes the memory polynomial operator, $h_p[n, \ldots, n]$ are the coefficients of the memory polynomial, M is the memory length, and P denotes the nonlinearity order.

The Hammerstein Model

The Hammerstein model is a two-box-model [17], i.e., a static nonlinear function prior to a dynamic linear filter, as depicted in Figure 2.3. In other words, the nonlinear behavior and memory effects of the PA are modeled separately and its output is given by

$$x_{\text{out}}[n] \triangleq \sum_{m=0}^{M} h[m]\, \mathrm{G}(x_{\text{in}}[n-m]), \tag{2.51}$$

where $\mathrm{G}(\cdot)$ denotes the memoryless nonlinear function and $h[n]$ is the impulse response of the linear finite impulse response (FIR) filter of length M.

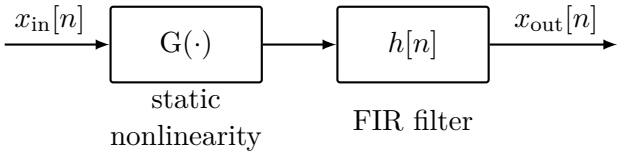


Figure 2.3.: Block-diagram of the Hammerstein model.

If the static nonlinearity is represented by a power series [27], the Hammerstein model reads

$$x_{\text{out}}[n] \triangleq \mathcal{V}_{\text{HS}}\left\{x_{\text{in}}[n]\right\} = \sum_{m=0}^{M} h[m] \sum_{p=1}^{\frac{P+1}{2}} a_{2p-1}|x_{\text{in}}[n-m]|^{2(p-1)} x_{\text{in}}[n-m], \tag{2.52}$$

where $\mathcal{V}_{\text{HS}}\{\cdot\}$ denotes the Hammerstein model operator. Compared to the discrete-time Volterra series (2.48), in the Hammerstein model, the p-th order Volterra kernel can be written as

$$h_p[n_1, n_2, \ldots, n_p] = \begin{cases} a_p h[n_1], & n_1 = n_2 = \cdots = n_p \\ 0, & \text{otherwise} \end{cases}. \tag{2.53}$$

Thus, the off-diagonal coefficients of the Volterra kernels are set to zero and the diagonals of each kernel correspond to the impulse response of the FIR filter, weighted by the coefficients a_p. However, the FIR filter coefficients $h[n]$ and the polynomial coefficients a_p are modeled separately and can be, therefore separately extracted (cf. Section 3.4).

Dynamic Deviation Reduction-Based Volterra Series

The dynamic deviation reduction-based Volterra series proposed by Zhu *et. al* in [28], is based on the fact that the discrete-time *passband* Volterra model

$$\tilde{x}_{\text{out}}[n] \triangleq \sum_{p=1}^{P} \sum_{m_1=0}^{M} \sum_{m_2=0}^{M} \cdots \sum_{m_p=0}^{M} \tilde{h}_p[m_1, m_2, \ldots, m_p] \prod_{i=1}^{p} \tilde{x}_{\text{in}}[n-m_i] \tag{2.54}$$

can be written as

$$\tilde{x}_{\text{out}}[n] \triangleq \tilde{x}_{\text{s}}[n] + \tilde{x}_{\text{d}}[n], \tag{2.55}$$

where $\tilde{x}_{\text{s}}[n]$ represents the static characteristics of the system, i.e.,

$$\tilde{x}_{\text{s}}[n] \triangleq \sum_{p=1}^{P} \tilde{h}_p[\underbrace{0, \ldots, 0}_{p}] \tilde{x}_{\text{in}}^{p}[n], \tag{2.56}$$

while

$$\tilde{x}_{\text{d}}[n] \triangleq \sum_{p=1}^{P} \sum_{r=1}^{p} \tilde{x}_{\text{in}}^{p-r}[n] \sum_{m_1=1}^{M} \sum_{m_2=m_1}^{M} \cdots \sum_{m_r=m_r-1}^{M} \tilde{h}_p[\underbrace{0, \ldots, 0}_{p-r}, \underbrace{m_1, m_2, \ldots, m_r}_{r}] \prod_{j=1}^{r} \tilde{x}_{\text{in}}[n-m_j], \tag{2.57}$$

is the purely dynamic part. In (2.57), $\tilde{h}_p[0,\ldots,0,n_1,\ldots,n_r]$ denotes the p-th order Volterra kernel where the first $p-r$ indices are zero, and r is the so-called *dynamic deviation reduction order* that represents the possible number of terms in the products of the delayed input samples $x_{\text{in}}[n-m_1]x_{\text{in}}[n-m_2]\cdot\ldots\cdot x_{\text{in}}[n-m_j]$. This product terms can be pruned to order $R \leq P$.

However, the dynamic deviation reduction-based Volterra model (2.55) is a passband representation. Therefore, in order to model PAs it has to be transformed to the baseband by assuming that $\tilde{x}_{\text{out}}(t)$ is a bandpass signal with carrier frequency f_{c} and bandwidth $B \ll f_{\text{c}}$. The *baseband deviation reduction-based Volterra model* can be derived as follows.

For $R = 0$, all delayed input terms are pruned and the equivalent baseband model is

$$x_{\text{out}}[n] \triangleq \mathcal{V}_{\text{DDR},0}\left\{x_{\text{in}}[n]\right\} = \sum_{p=1}^{\frac{P+1}{2}} h_{2p-1}[0\ldots,0]x_{in}[n]|x_{\text{in}}[n]|^{2(p-1)}, \tag{2.58}$$

where $\mathcal{V}_{\text{DDR},0}\{\cdot\}$ denotes the *zero order deviation reduction-based Volterra model*. Obviously, it coincides with the memory polynomials for $M = 0$ and the memoryless polynomial model (2.18), however, sampled.

For $R = 1$, only one delay of the input terms can deviate from zero. Therefore, the equivalent baseband model can be written as

$$\begin{aligned} x_{\text{out}}[n] &\triangleq \mathcal{V}_{\text{DDR},1}\left\{x_{\text{in}}[n]\right\} \\ &= \sum_{p=1}^{\frac{P+1}{2}} \sum_{m_1=0}^{M} g_{2p-1,1}[m_1]|x_{\text{in}}[n]|^{2(p-1)}x_{\text{in}}[n-m_1] \\ &+ \sum_{p=2}^{\frac{P+1}{2}} \sum_{m_1=1}^{M} g_{2p-1,2}[m_1]x_{\text{in}}^2[n]|x_{in}[n]|^{2(p-2)}x_{\text{in}}^*[n-m_1], \end{aligned} \tag{2.59}$$

where $\mathcal{V}_{\text{DDR},1}\{\cdot\}$ denotes the *first order deviation reduction-based Volterra model* with appropriate kernel coefficients $g_{2p-1,i}[\cdot]$. Whereas the first summation considers that the instantaneous input signal $x_{\text{in}}[n]$ is conjugated, i.e., $x_{\text{in}}^*[n]$, the last summation considers that the delayed input signal $x_{\text{in}}[n-m_1]$ is conjugated, i.e., $x_{\text{in}}^*[n-m_1]$. Since the second summation includes one conjugated product item, it occurs for nonlinearity orders $P \geq 3$.

For $R = 2$, the delay of two input terms can deviate from zero. Therefore, the

equivalent baseband model reads

$$\begin{aligned} x_{\text{out}}[n] \triangleq \mathcal{V}_{\text{DDR},2}\left\{x_{\text{in}}[n]\right\} = &\sum_{p=1}^{\frac{P+1}{2}} \sum_{m_1=0}^{M} g_{2p-1,1}[m_1]|x_{\text{in}}[n]|^{2(p-1)} x_{\text{in}}[n-m_1] \\ &+ \sum_{p=2}^{\frac{P+1}{2}} \sum_{m_1=1}^{M} \sum_{m_2=m_1}^{M} g_{2p-1,3}[m_1,m_2] x_{\text{in}}^*[n]|x_{\text{in}}[n]|^{2(p-2)} \\ &\quad x_{\text{in}}[n-m_1]x_{\text{in}}[n-m_2] \\ &+ \sum_{p=2}^{\frac{P+1}{2}} \sum_{m_1=0}^{M} \sum_{m_2=1}^{M} g_{2p-1,4}[m_1,m_2] x_{\text{in}}[n]|x_{\text{in}}[n]|^{2(p-2)} \\ &\quad x_{\text{in}}[n-m_1]x_{\text{in}}^*[n-m_2] \\ &+ \sum_{p=3}^{\frac{P+1}{2}} \sum_{m_1=1}^{M} \sum_{m_2=m_1}^{M} g_{2p-1,5}[m_1,m_2] x_{\text{in}}^3[n]|x_{\text{in}}[n]|^{2(p-3)} \\ &\quad x_{\text{in}}^*[n-m_1]x_{\text{in}}^*[n-m_2] \end{aligned} \tag{2.60}$$

where $\mathcal{V}_{\text{DDR},2}\{\cdot\}$ denotes the *second order deviation reduction-based Volterra model* with appropriate kernel coefficients $g_{2p-1,i}[\cdot]$. In fact, the first summation considers first order deviation, the second summation considers that the instantaneous input signal $x_{\text{in}}[n]$ is conjugated, i.e., $x_{\text{in}}^*[n]$, the third summation considers that either one of the two delayed input terms is conjugated, i.e., $x_{\text{in}}[n-m_1]x_{\text{in}}^*[n-m_2]$, and finally the last summation considers that both delayed input terms are conjugated, i.e., $x_{\text{in}}^*[n-m_1]x_{\text{in}}^*[n-m_2]$. Since the last summation includes two conjugated product terms, it occurs for nonlinearity orders $P \geq 5$.

More detailed derivations are present in Appendix A. Higher order models $\mathcal{V}_{\text{DDR},R}\{\cdot\}$ for $R > 2$ can be derived in the same manner. However, for real-time applications R should be limited to a small value, e.g., $R \leq 2$, in order to reduce model complexity and computational costs.

2.4.6 Decomposed Piecewise Volterra Model

To cope with the rather distinct nonlinear behavior of an envelope tracking power amplifier (ETPA) at different input power regions, Zhu *et. al* proposed in [29] a new approach for signal decomposition, named *vector threshold decomposition*, which is an extension from real-valued signal decomposition to complex-valued signals. In [29], a set of decomposition thresholds

$$\Lambda \triangleq \{\lambda_1, \lambda_2, \ldots, \lambda_S\}, \tag{2.61}$$

is defined such that the decomposition thresholds are real-valued and strictly monotonically increasing, i.e., $\lambda_1 < \lambda_2 < \cdots < \lambda_S$, and S is the total number of thresholds. The complex-valued signal $x[n]$ is decomposed with respect to its magnitude. Therefore,

each threshold λ_s represents the radius of the corresponding s-th threshold circle in the I/Q-plane, as depicted in Figure 2.4.

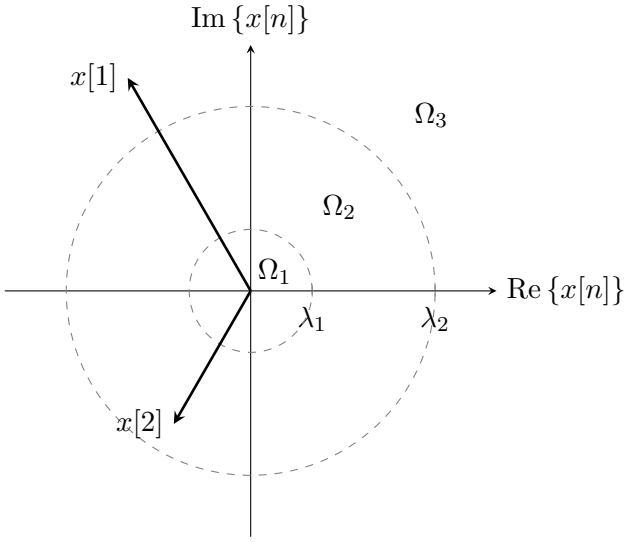


Figure 2.4.: The signal values $x[1]$ and $x[2]$ are decomposed in three sub-signals that correspond to the decomposition zones Ω_1, Ω_2, and Ω_3.

This allows to decompose the *complex-valued* signal $x[n]$ into $S+1$ sub-signals, which end up in the corresponding sub-signal region indicated by Ω_s in Figure 2.4. Specifically, the s-th sub-signal of $x[n]$ is

$$x_s[n] \triangleq \begin{cases} 0, & |x[n]| \leq \lambda_{s-1} \\ (|x[n]| - \lambda_{s-1})\, \mathrm{e}^{\jmath \arg\{x[n]\}}, & \lambda_{s-1} < |x[n]| \leq \lambda_s \\ (\lambda_s - \lambda_{s-1})\, \mathrm{e}^{\jmath \arg\{x[n]\}}, & |x[n]| > \lambda_s \end{cases} \tag{2.62}$$

where $\lambda_0 = 0$ and $1 \leq s \leq S+1$ is assumed. To illustrate this decomposition in more detail, consider the following examples. Let $\Lambda = \{0.2, 0.6\}$ be the set of two thresholds, i.e., $\lambda_1 = 0.2$ and $\lambda_2 = 0.6$. By this, the input space $\mathbb{C}$ is divided into three sub-spaces, namely $\Omega_1 = \{x[n] \colon |x[n]| \leq 0.2\}$, $\Omega_2 = \{x[n] \colon 0.2 < |x[n]| \leq 0.6\}$, $\Omega_3 = \{x[n] \colon 0.6 < |x[n]| < \infty\}$, and $\Omega_1 \cup \Omega_2 \cup \Omega_3 = \mathbb{C}$. A particular signal value such as $x[1] = 0.8\mathrm{e}^{\jmath\frac{2\pi}{3}}$, as depicted in Figure 2.4, can be decomposed into the three sub-signals $x_1[1] = 0.2\mathrm{e}^{\jmath\frac{2\pi}{3}}$, $x_2[1] = 0.4\mathrm{e}^{\jmath\frac{2\pi}{3}}$, and $x_3[1] = 0.2\mathrm{e}^{\jmath\frac{2\pi}{3}}$. Accordingly, the magnitude of the first sub-signal $x_1[1]$ is equal to λ_1, the magnitude of the second sub-signal $x_2[1]$ is equal to $\lambda_2 - \lambda_1$, and finally, the magnitude of $x_3[1]$ is equal to $|x[n]| - \lambda_2$. Since the complex-valued signal $x[n]$ is decomposed with respect to its magnitude, all phases of the sub-signals are identical to the phase of original signal. Moreover, if a signal value does not reach the next higher sub-spaces, the sub-signals corresponding to those zones are set to zero. For instance, $x[2] = 0.5\mathrm{e}^{-\jmath\frac{2\pi}{3}}$, as depicted in Figure 2.4, is decomposed into $x_1[2] = 0.2\mathrm{e}^{-\jmath\frac{2\pi}{3}}$, $x_2[2] = 0.3\mathrm{e}^{-\jmath\frac{2\pi}{3}}$, and since $x[2]$ does not reach Ω_3, $x_3[2] = 0$.

Figure 2.5 provides a deeper insight into the vector threshold decomposition; the original signal $x[n]$ is decomposed into three sub-signals, i.e., $x_1[n]$, $x_2[n]$, and $x_3[n]$, corresponding to the respective input power level. Moreover, the sum of all sub-signals equals the original signal, i.e.,

$$x[n] = \sum_{s=1}^{S+1} x_s[n]. \tag{2.63}$$

Obviously, the original signal can be easily recombined.

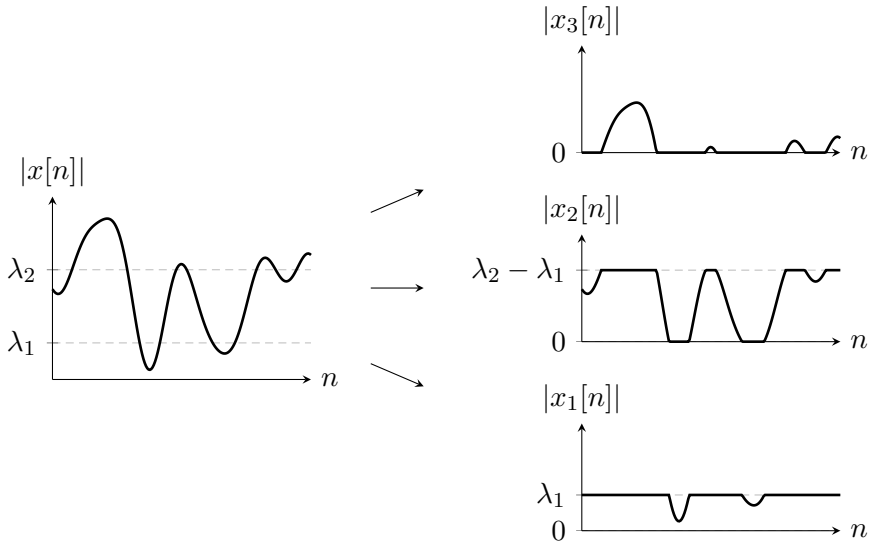


Figure 2.5.: Signal decomposition into three sub-signals, $x_1[n]$, $x_2[n]$, and $x_3[n]$.

As mentioned in Section 1.4, the output characteristics of an ETPA strongly depend on the input power level. Therefore, the threshold levels and sub-models of each region can be flexibly tailored to the ETPA characteristics. In other words, after the input signal $x_{\text{in}}[n]$ is decomposed, each sub-signal $x_{s;\text{in}}[n]$ is individually processed by the corresponding sub-model $\mathrm{G}_s(\cdot)$. If Volterra series are employed as sub-models, i.e., $\mathrm{G}_s(\cdot) = \mathcal{V}_s\{\cdot\}$, the model output $x_{\text{out}}[n]$ is still linear with respect to its Volterra kernel coefficients, although $x_{\text{in}}[n]$ is decomposed, i.e.,

$$x_{\text{out}}[n] = \sum_{s=1}^{S+1} \mathrm{G}_s\left(x_{s;\text{in}}[n]\right), \tag{2.64}$$

where the sub-model $\mathrm{G}_s(\cdot)$ is either the full blown Volterra model $\mathcal{V}\{\cdot\}$, the memory polynomial model $\mathcal{V}_{\text{MP}}\{\cdot\}$, the Hammerstein model $\mathcal{V}_{\text{HS}}\{\cdot\}$, or the dynamic deviation reduction-based Volterra model $\mathcal{V}_{\text{DDR},R}\{\cdot\}$, as depicted in Figure 2.6.

For instance, if memory polynomial sub-models are employed solely, the s-th sub-

model reads

$$\begin{aligned} x_{s;\mathrm{out}}[n] &= \mathcal{V}_{s;\mathrm{MP}}\left\{x_{s;\mathrm{in}}[n]\right\} \\ &= \sum_{p=1}^{\frac{P_s+1}{2}} \sum_{m=0}^{M_s} h_{s;2p-1}[m\ldots,m] x_{s;\mathrm{in}}[n-m]|x_{s;\mathrm{in}}[n-m]|^{2(p-1)}, \end{aligned} \tag{2.65}$$

where $x_{s;\mathrm{out}}[n]$ and $x_{s;\mathrm{in}}[n]$ are the output and input of the s-th sub-model $\mathcal{V}_{s;\mathrm{MP}}\{\cdot\}$, respectively, and $h_{s;p}[n,\ldots,n]$, P_s, and M_s are the memory polynomial coefficients, the odd nonlinearity order, and the memory length of the s-th sub-memory polynomial model $\mathcal{V}_{s;\mathrm{MP}}\{\cdot\}$, respectively. The decomposition permits to choose P_s and M_s according to the ETPA characteristics in the respective input power region. Therefore, the rather distinct behavior of the ETPA in the different zones can be accurately characterized. Finally, the output of the overall model reads

$$\begin{aligned} x_{\mathrm{out}}[n] &= \sum_{s=1}^{S+1} x_{s;\mathrm{out}}[n] = \sum_{s=1}^{S+1} \mathcal{V}_{s;\mathrm{MP}}\left\{x_{s;\mathrm{in}}[n]\right\} \\ &= \sum_{s=1}^{S+1} \sum_{p=1}^{\frac{P_s+1}{2}} \sum_{m=0}^{M_s} h_{s;2p-1}[m,m,\ldots,m] x_{s;\mathrm{in}}[n-m]|x_{s;\mathrm{in}}[n-m]|^{2(p-1)}, \end{aligned} \tag{2.66}$$

which is denoted as the *decomposed piecewise memory polynomials*.

However, employing the dynamic deviation reduction-based Volterra series with, e.g., $R = 1$, instead of the memory polynomials, the output of the overall model reads

$$\begin{aligned} x_{\mathrm{out}}[n] &= \sum_{s=1}^{S+1} \mathcal{V}_{s;\mathrm{DDR},1}\left\{x_{s;\mathrm{in}}[n]\right\} \\ &= \sum_{s=1}^{S+1} \sum_{p=1}^{\frac{P+1}{2}} \sum_{m_1=0}^{M} g_{s;2p-1,1}[m_1]|x_{s;\mathrm{in}}[n]|^{2(p-1)} x_{s;\mathrm{in}}[n-m_1] \\ &+ \sum_{s=1}^{S+1} \sum_{p=2}^{\frac{P+1}{2}} \sum_{m_1=1}^{M} g_{s;2p-1,2}[m_1] x_{s;\mathrm{in}}^2[n]|x_{s;\mathrm{in}}[n]|^{2(p-2)} x_{s;\mathrm{in}}^*[n-m_1], \end{aligned} \tag{2.67}$$

which is denoted as the *decomposed piecewise deviation reduction-based Volterra series* with deviation reduction $R = 1$.

Nevertheless, the overall model is not restricted in the use of the same model class for all sub-models. If it better fits the PA characteristics, one can use distinct model classes for the different sub-models. For instance, the first S_1 sub-models could be memory polynomials, i.e., $\mathcal{V}_{s;\mathrm{MP}}\{\cdot\}$, for $1 \leq s \leq S_1$, while the remaining $S_2 = S - S_1 + 1$ sub-models could be Volterra series, i.e, $\mathcal{V}_s\{\cdot\}$, for $S_1 < s \leq S + 1$. In fact, to obtain an overall model that is linear with respect to its coefficients, the sub-models have to be linear with respect to its coefficients as well. This is because the overall model output $x_{\mathrm{out}}[n]$ is composed of the sum of the $S + 1$ sub-models outputs $x_{s;\mathrm{out}}[n] = \mathrm{G}_s(x_{s;\mathrm{in}}[n])$.

Figure 2.6 provides an overview of the signal processing and recombination process. Each of the decomposed input signals $x_{s;\mathrm{in}}[n]$ is individually treated by the sub-model $\mathrm{G}_s(\cdot)$, and all outputs $x_{s;\mathrm{out}}[n]$ of the sub-models are recombined in order to produce the final model output $x_{\mathrm{out}}[n]$. Since multiple models are involved for a single output signal value, the number of coefficients of each sub-model should be kept reasonably small, e.g., the nonlinearity order P_s in each sub-model can be limited to fifth order to reduce the overall complexity.

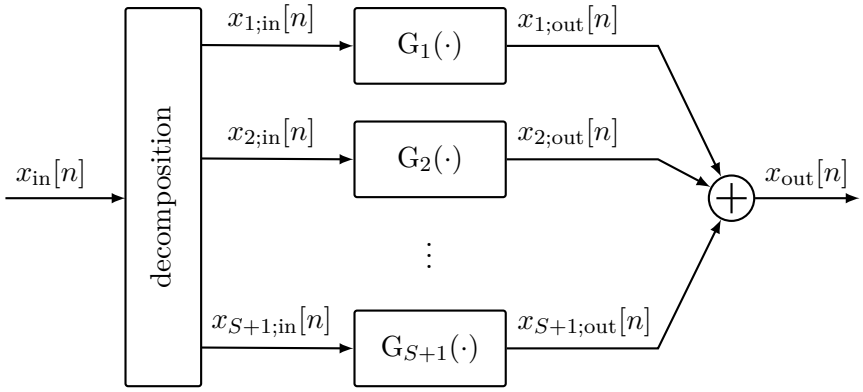


Figure 2.6.: Signal processing line of the decomposed piecewise Volterra model. The decomposed input signals $x_{s;\mathrm{in}}[n]$ are processed by the individual sub-models $\mathrm{G}_s(\cdot)$ and added up, in order to recombine to the overall output signal $x_{\mathrm{out}}[n]$.

2.4.7 Vector-Switched Model

Another solution to model the distinct behavior of the ETPA is to employ piecewise curve fitting models, where the nonlinear curve is divided into several segments, and then fit each segment separately by using different models. In fact, Afsardoost *et. al* proposed in [30] a novel switched behavioral model that can be thought of as a generalization of curve fitting models. Here, the *switching space* is divided into K regions. Specifically, the regions $\Omega_k, k = 1, \ldots, K$, are partitions of the J-dimensional space $\mathbb{C}^J$, such that

$$\bigcup_{k=1}^{K} \Omega_k = \mathbb{C}^J, \tag{2.68}$$

where the different spaces are disjoint, i.e., $\Omega_k \cap \Omega_l = \emptyset$ for $k \neq l$. The output of the overall model is calculated by the respective regional model, i.e.,

$$x_{\mathrm{out}}[n] = \mathrm{G}_k(x_{\mathrm{in}}[n]), \tag{2.69}$$

where $k \in \{1, \ldots, K\}$ is determined by the *switching function*, and $\mathrm{G}_k(\cdot)$ is the associated regional model as shown in Figure 2.7.

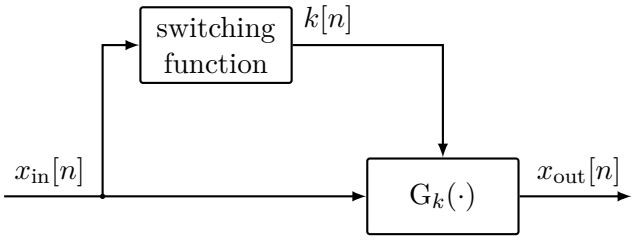


Figure 2.7.: Vector-switched behavioral model. The switching functions determines the sub-model.

Afsardoost *et. al* suggest to perform the vector switching over the amplitude space and to ignore the phases of the input signal samples. Thus, the switching space becomes $\mathbb{R}^J$. The design of the switching regions Ω_k is based on the instantaneous and its $J-1$ preceding samples of the training set, such that reasonable boundaries between the regions are determined. One way to achieve this, is the so-called *Voronoi* or *nearest neighbor partitioning* [31], where each region Ω_k contains all samples that are closest (in terms of the Euclidean distance) to the respective mean (or centroid) $\underline{\omega}_k \in \mathbb{R}^J$ of Ω_k, i.e.,

$$\Omega_k = \left\{ \underline{r}_{n,J} \colon \|\underline{r}_{n,J} - \underline{\omega}_k\|_2 < \|\underline{r}_{n,J} - \underline{\omega}_l\|_2,\ l = 1, \ldots, k-1, k+1, \ldots, K \right\}, \tag{2.70}$$

where $\underline{r}_{n,J} = \begin{bmatrix} |x_{\text{in}}[n]| & |x_{\text{in}}[n-1]| & \cdots & |x_{\text{in}}[n-J+1]| \end{bmatrix}^T$ contains the magnitudes of the instantaneous and its $J-1$ preceding input samples of the training set of length $N \gg J$.

The design of the switching regions is now limited to find the appropriate set of centroids $\underline{\omega}_k \in \mathbb{R}^J$. There are numerous different algorithms; a summary can be found in [32]. A broadly used algorithm is the K-means algorithm [33]. Figure 2.8 depicts a two-dimensional ($J = 2$) Voronoi partitioning of a Long Term Evolution (LTE) modulated signal with 20MHz bandwidth and 8.5dB peak-to-average power ratio (PAPR), where the switching space $\mathbb{R}^2$ is divided into eight ($K = 8$) regions. However, this vector-switching is not limited by using the input samples solely; other PA characteristics, e.g., the envelope tracked supply voltage or the PA temperature, can be used in order to select an adequate switching function.

As mentioned before, the vector-switched model is a generalization of curve fitting models. In fact, it describes a wide range of models too. For instance, if the number of regions is set to $K = 1$, it represents classical (non-switched) models. However, if $J = 1$ and K is chosen arbitrarily large, but the simplest possible regional model is employed, i.e., only a constant value in each region, the vector-switched model also represents look-up table (LUT) models. Obviously, each sub-model belonging to a certain sub-region Ω_k can be tailored to the ETPA characteristics in that region. Whereas one signal sample is processed by all $S+1$ sub-models in the decomposed piecewise Volterra model (cf. Section 2.4.6), only one of the K sub-model is needed in the vector-switched model (cf. Figure 2.7). Therefore, compared to classical (non-switched) models, for the

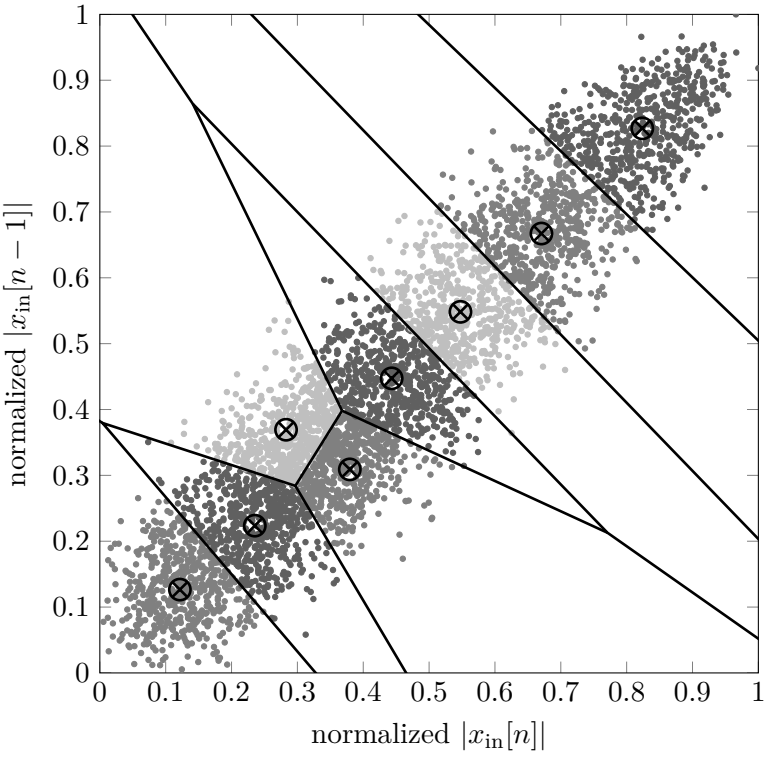


Figure 2.8.: Two-dimensional Voronoi partitioning of a LTE modulated signal with 20MHz bandwidth and 8.5dB PAPR. The switching space $\mathbb{R}^2$ is partitioned into eight regions. The centroids are marked via $\otimes$, and the corresponding samples with different grayscales.

vector-switched model, the only computational complexity is added by the switching function, since the partitioning into the regions Ω_k can be done offline[1]. On the other hand, since the nonlinear behavior is partitioned, the complexity of the sub-models can be kept reasonably small.

Beside complexity efforts, the performance of the vector-switched model is limited due to the so-called discontinuities distortion that occurs when different models are used for consecutive signal samples. The output signal may contain discontinuities because of possibly unsteady transition between the utilized models, thus evoking undesired high-frequency components in the output signal. Afsardoost *et. al* state in [30] that in most applications the added switching-noise is neglectable, because it is below the model error. However, there are cases for which the switching distortion can be suspected to be considerably larger than the modeling error. Specifically, when the training set is too small, some sub-models may get poorly excited since there are not enough signal values in those sub-model regions. The fact that the model performance suffers from a

[1]Note that if the statistical properties of the input signal change, the regions $\underline{\omega}_k$ must be retrained.

too small training set will be discussed in Section 3.6.

2.5 Figures of Merit to Evaluate the Model Performance

Suitable figures and characterizations of merit allow to compare the performance of different PA models and digital pre-distortion (DPD) algorithms, in order to favor one model over another. The most important evaluation criterion for PA behavioral models is accuracy. Qualitatively, modeling accuracy describes the ability to predict the output of the modeled PA, given a specific input and its operating conditions [18]. In literature, many figures of merit exist, a summary and comparison can be found in [18, 34].

In this thesis, the normalized mean square error (NMSE), the adjacent channel power ratio (ACPR), and the adjacent channel error power ratio (ACEPR) are used as a metric in order to evaluate the quality of the model and the DPD performance.

2.5.1 Normalized Mean Square Error (NMSE)

The NMSE is commonly described as

$$\mathrm{NMSE} \triangleq 10 \log_{10} \left(\frac{\sum\limits_{n} \left| x_{\mathrm{PA}}[n] - x_{\mathrm{model}}[n] \right|^2}{\sum\limits_{n} \left| x_{\mathrm{PA}}[n] \right|^2} \right) \qquad \text{in dB}, \tag{2.71}$$

where $x_{\mathrm{PA}}[n]$ and $x_{\mathrm{model}}[n]$ are the baseband output signals of the PA and the model, respectively. Accordingly, it yields the (normalized) total error of the model and is in general dominated by in-band errors [34].

2.5.2 Adjacent Channel Power Ratio (ACPR)

Since in real-life applications out-of-band errors, i.e., the adverse emission into other communication channels, is of more interest [34], the ACPR is used to describe the level of spectral re-growth, and is defined as the ratio between in-band and out-of-band signal powers. Assuming that all channels have the same bandwidth B and are spaced by $\Delta f \geq B$, the i-th adjacent channel occupies the frequency range

$$B_{\mathrm{adj},i} \triangleq \left\{ f \colon f_{\mathrm{c}} + i\Delta f - \frac{B}{2} \leq f \leq f_{\mathrm{c}} + i\Delta f + \frac{B}{2} \right\}. \tag{2.72}$$

Therefore, the total adjacent frequency band can be described as

$$B_{\mathrm{adj}} \triangleq \bigcup_{i=-N_{\mathrm{low}}}^{N_{\mathrm{hi}}} B_{\mathrm{adj},i}, \tag{2.73}$$

where N_{low} and N_{hi} are the number of adjacent channels to the left, respectively to the right of the transmission channel. Accordingly, the ACPR is calculated by

$$\text{ACPR} \triangleq 10\log_{10}\left(\frac{\int\limits_{B_{\text{adj}}} |S_{\mathbf{x}}(f)|^2\,\mathrm{d}f}{\int\limits_{B_{\text{ch}}} |S_{\mathbf{x}}(f)|^2\,\mathrm{d}f}\right) \qquad \text{in dB}, \tag{2.74}$$

where $S_{\mathbf{x}}(f)$ denotes the power spectral density (PSD)[1] of the analysed random signal $\mathbf{x}(t)$ and

$$B_{\text{ch}} = \left\{f\colon f_{\text{c}} - \frac{B}{2} \leq f \leq f_{\text{c}} + \frac{B}{2}\right\}, \tag{2.75}$$

is the RF transmission band. Figure 2.9 visualizes the definition of the ACPR, which can be thought of as the quotient of the light and the dark gray areas.

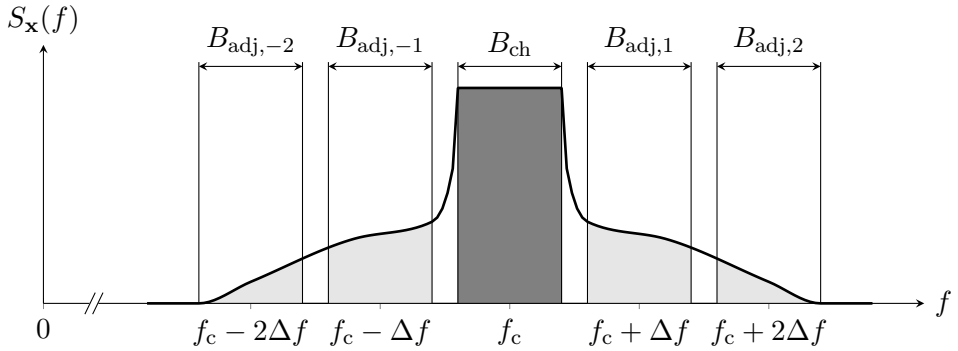


Figure 2.9.: Power emission in adjacent communication channels $B_{\text{adj},i}$ due to nonlinear distortions.

The ACPR cannot be used to evaluate the accuracy of a PA model, since the latter is not included. However, the ACPR ratio of the model and the PA can be used as a comparative measure, i.e.,

$$\Delta\text{ACPR} \triangleq \text{ACPR}_{\text{PA}} - \text{ACPR}_{\text{model}} \qquad \text{in dB}, \tag{2.76}$$

where ACPR_{PA} and $\text{ACPR}_{\text{model}}$ are the ACPRs of the PA and the model, respectively. The disadvantage of this representation is that only magnitude distortions are considered, i.e., the nonlinearities in the amplitude modulation/amplitude modulation (AM/AM) conversion (cf. Section 1.2). In contrast, the ACEPR also considers phase distortions

[1] The power spectral density (PSD) of a complex-valued stationary random signal $\mathbf{x}(t)$ with autocorrelation $r_{\mathbf{xx}}(t)$ is defined as
$S_{\mathbf{x}}(f) \triangleq \int\limits_{-\infty}^{\infty} r_{\mathbf{xx}}(\tau)\mathrm{e}^{\jmath 2\pi f\tau}\mathrm{d}\tau$.

and is defined as

$$\mathrm{ACEPR} \triangleq 10 \log_{10} \left(\frac{\int\limits_{B_{\mathrm{adj}}} |S_{\mathbf{x}_{\mathrm{PA}}}(f) - S_{\mathbf{x}_{\mathrm{model}}}(f)|^2 \, \mathrm{d}f}{\int\limits_{B_{\mathrm{ch}}} |S_{\mathbf{x}_{\mathrm{PA}}}(f)|^2 \mathrm{d}f} \right) \quad \text{in dB,} \tag{2.77}$$

where $S_{\mathbf{x}_{\mathrm{PA}}}(f)$ and $S_{\mathbf{x}_{\mathrm{model}}}(f)$ are the PSDs of the PA output and the model output, respectively [18].

3

Chapter 3

Model Parameter Estimation

Contents

3.1 Introduction and Outline

In this Chapter, parameter identification methods for the behavioral models presented in Chapter 2 are developed. The model $\mathrm{G}(\cdot)$ that is aimed to estimate the discrete-time equivalent baseband representation of the power amplifier (PA) can be extracted by the configuration shown in Figure 3.1. Accordingly, both, the PA and the model are driven by the same baseband input signal $x_{\mathrm{in}}[n]$. By an adaptive scheme, the model is trained with the goal to converge towards the "best" representation of the PA, in terms of a defined performance measure.

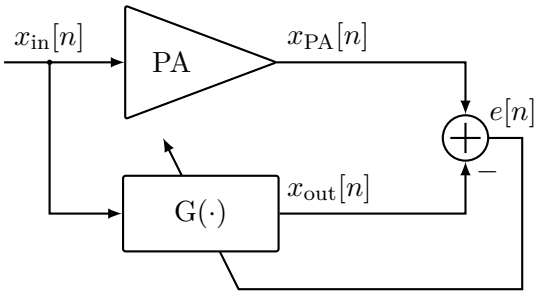


Figure 3.1.: PA identification; the PA and the model are discrete-time equivalent baseband representations.

Since the model parameters depend on the applied set of training signals [18], it is advantageous to apply the same kind of signal sets for the identification as used in the communication system, e.g., if the PA is designed for Long Term Evolution (LTE) signals, signals having LTE-like statistical properties should be used during parameter estimation. By this, it is guaranteed that the PA is excited with the pertinent frequency and amplitude range [17].

This Chapter is organized as follows

Section 3.2 - Parameter Estimation for the Volterra Series, derives the least squares (LS) estimator for the truncated Volterra series proposed in Section 2.4.4.

Section 3.3 - Parameter Estimation for Reduced Volterra Series, presents the implementation of the afore derived LS estimator for the pruned Volterra models proposed in Section 2.4.5.

Section 3.4 - Parameter Estimation for the Hammerstein Model, derives the LS estimator for the Hammerstein model proposed in Section 2.4.5

Section 3.5 - Parameter Estimation for the Decomposed Volterra Series, shows the adaption of the LS estimator to the decomposed piecewise Volterra series proposed in Section 2.4.6

Section 3.6 - Parameter Estimation for the Vector-Switched Model, presents the LS estimator for the vector-switched model proposed in Section 2.4.7

Section 3.7 - Adaptive Model Parameter Estimation, introduces two parameter estimation methods that are capable of estimating the model characteristics adaptively, namely the least mean squares (LMS) algorithm and the ϵ-normalized least mean squares (ϵ-NLMS) algorithm.

3.2 Parameter Estimation for the Volterra Series

As discussed in Section 2.4.2, the output of the Volterra model is linear with respect to its coefficients. Therefore, the discrete-time truncated Volterra model (2.48) can be described by the linear vector notation

$$x_{\text{out}}[n] \triangleq \underline{h}_{\mathcal{V}}^T \check{\underline{x}}_{\text{in},n}, \tag{3.1}$$

where $\underline{h}_{\mathcal{V}}$ contains all stacked Volterra kernel vectors $\underline{h}_p$ of odd order $p \leq P$, i.e.,

$$\underline{h}_{\mathcal{V}} \triangleq \begin{bmatrix} \underline{h}_1^T & \underline{h}_3^T & \underline{h}_5^T & \cdots & \underline{h}_P^T \end{bmatrix}^T. \tag{3.2}$$

Each Volterra kernel vector $\underline{h}_p$ contains the coefficients of the p-th order Volterra kernel arranged in a vector, i.e., $\underline{h}_p = \begin{bmatrix} h_p[0,\ldots,0] & \cdots & h_p[M,\ldots,M] \end{bmatrix}^T$. Likewise, $\check{\underline{x}}_{\text{in},n}$ contains the product terms of the delayed inputs in (2.48) and is constructed as follows [35]. Consider the Kronecker product[1],

$$\check{\underline{x}}_{p,\text{in},n} \triangleq \underbrace{\underline{x}_{\text{in},n,M} \otimes \underline{x}_{\text{in},n,M} \otimes \cdots \otimes \underline{x}_{\text{in},n,M}}_{p-\text{times}}, \tag{3.3}$$

p-times applied to the input vector $\underline{x}_{\text{in},n,M} \in \mathbb{C}^{M+1}$ that contains the instantaneous and its M preceding samples, i.e., $\underline{x}_{\text{in},n,M} = \begin{bmatrix} x_{\text{in}}[n] & x_{\text{in}}[n-1] & \cdots & x_{\text{in}}[n-M] \end{bmatrix}^T$. Accordingly, $\check{\underline{x}}_{p,\text{in},n}$ contains the p-th order product terms of the delayed inputs and $\check{\underline{x}}_{\text{in},n}$ can be written as

$$\check{\underline{x}}_{\text{in},n} \triangleq \begin{bmatrix} \check{\underline{x}}_{1,\text{in},n}^T & \check{\underline{x}}_{3,\text{in},n}^T & \check{\underline{x}}_{5,\text{in},n}^T & \cdots & \check{\underline{x}}_{P,\text{in},n}^T \end{bmatrix}^T. \tag{3.4}$$

From the point-of-view of system identification, the parameter extraction procedure tries to estimate the model parameters based on N input and output samples of the PA, which is referred to as the *training signal set*, $\underline{x}_{\text{in},n}$ and $\underline{x}_{\text{PA},n}$, respectively. The model that tries to represent the PA can be written in vector form over a block of N samples as

$$\underline{x}_{\text{out},n} = \check{X}_{\text{in},n} \underline{h}_{\mathcal{V}}, \tag{3.5}$$

[1] The Kronecker product of an $n \times p$ matrix A and an $m \times q$ matrix B is defined as the $mn \times pq$ matrix [36]

$$A \otimes B \triangleq \begin{bmatrix} a_{11}B & a_{12}B & \cdots & a_{1p}B \\ a_{21}B & a_{22}B & \cdots & a_{2p}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}B & a_{n2}B & \cdots & a_{np}B \end{bmatrix}, \text{ where } A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{np} \end{bmatrix}.$$

with the matrix

$$\check{X}_{\mathrm{in},n} = \left[\underline{\check{x}}_{\mathrm{in},n} \quad \underline{\check{x}}_{\mathrm{in},n-1} \quad \cdots \quad \underline{\check{x}}_{\mathrm{in},n-N+1}\right]^T. \tag{3.6}$$

As in Figure 3.1, the unmodeled error in vector form over a block of N samples can be written as

$$\underline{e}_n = \underline{x}_{\mathrm{PA},n} - \underline{x}_{\mathrm{out},n}. \tag{3.7}$$

The blockwise LS estimator that finds the global minimum of the cost function

$$\varepsilon^2(\underline{h}_{\mathcal{V}}) \triangleq \underline{e}_n^H \underline{e}_n = \|\underline{x}_{\mathrm{PA},n} - \underline{h}_{\mathcal{V}} \check{X}_{\mathrm{in},n}\|_2^2, \tag{3.8}$$

states that such an estimate can be given by [36]

$$\hat{\underline{h}}_{\mathcal{V}} \triangleq \arg\min_{\underline{h}_{\mathcal{V}}} \left\{\varepsilon^2(\underline{h}_{\mathcal{V}})\right\} = \left(\check{X}_{\mathrm{in},n}^H \check{X}_{\mathrm{in},n}\right)^{-1} \check{X}_{\mathrm{in},n}^H \, \underline{x}_{\mathrm{PA},n}. \tag{3.9}$$

Moreover, the number of training signal samples N should be reasonable large, in order to ensure that all nonlinearities of the PA are excited.

Note that Volterra kernel symmetry is not considered here; an extension including Volterra kernel symmetry can be found in [21].

3.3 Parameter Estimation for Reduced Volterra Series

The LS estimator (3.9) for the pruned Volterra models discussed in Section 2.4.5 can be easily determined by simply neglecting the appropriate Volterra kernel coefficients, i.e., setting them to zero. For example, the p-th order Volterra kernel vector $\underline{h}_p$ for the memory polynomial model $\mathcal{V}_{\mathrm{MP}}\{\cdot\}$ can be written as

$$\underline{h}_p = \left[h_p[0,\ldots,0] \quad h_p[1,\ldots,1] \quad \cdots \quad h_p[M,\ldots,M]\right]^T, \tag{3.10}$$

which contains only the diagonal terms of the p-th order Volterra kernel $h_p[n_1, n_2, \ldots, n_p]$, whereas the off-diagonal terms are set to zero and are therefore ignored. The corresponding input vector to (3.4) reads

$$\underline{\check{x}}_{p,\mathrm{in},n} = \begin{bmatrix} |x_{\mathrm{in}}[n]|^{p-1} x_{\mathrm{in}}[n] \\ |x_{\mathrm{in}}[n-1]|^{p-1} x_{\mathrm{in}}[n-1] \\ \vdots \\ |x_{\mathrm{in}}[n-M]|^{p-1} x_{\mathrm{in}}[n-M] \end{bmatrix}, \tag{3.11}$$

where the appropriate product terms are likewise neglected.

Similarily, $\underline{h}_p$ and $\underline{\check{x}}_{p,\mathrm{in},n}$, for the dynamic deviation reduction-based Volterra series $\mathcal{V}_{\mathrm{DDR},R}\{\cdot\}$ can be determined by simply pruning the number of Volterra kernel coefficients and product terms as discussed in Section 2.4.5. For instance, $\underline{h}_p$ and $\underline{\check{x}}_{p,\mathrm{in},n}$ for

the first order deviation reduction-based Volterra series with nonlinearity order $p \geq 3$, can be written as

$$\underline{h}_p = \begin{bmatrix} g_{p,1}[0] \\ g_{p,2}[1] \\ \vdots \\ g_{p,1}[M] \\ g_{p,2}[1] \\ g_{p,2}[2] \\ \vdots \\ g_{p,2}[M] \end{bmatrix} \qquad \check{\underline{x}}_{p,\mathrm{in},n} = \begin{bmatrix} |x_{\mathrm{in}}[n]|^{p-1} x_{\mathrm{in}}[n] \\ |x_{\mathrm{in}}[n-1]|^{p-1} x_{\mathrm{in}}[n-1] \\ \vdots \\ |x_{\mathrm{in}}[n-M]|^{p-1} x_{\mathrm{in}}[n-M] \\ x_{\mathrm{in}}^2[n-1]|x_{\mathrm{in}}[n-1]|^{p-3} x_{\mathrm{in}}^*[n-1] \\ x_{\mathrm{in}}^2[n-2]|x_{\mathrm{in}}[n-2]|^{p-3} x_{\mathrm{in}}^*[n-2] \\ \vdots \\ x_{\mathrm{in}}^2[n-M]|x_{\mathrm{in}}[n-M]|^{p-3} x_{\mathrm{in}}^*[n-M] \end{bmatrix}. \tag{3.12}$$

Accordingly, as R increases, not only the vector size increases, but also vector construction effort and thus computational complexity. Therefore, in real-life applications R should be limited to a small value, e.g., $R \leq 2$, in order keep the complexity low.

3.4 Parameter Estimation for the Hammerstein Model

Since the Hammerstein model can be thought of as a pruned Volterra model, the parameters can be estimated by neglecting the appropriate Volterra kernel coefficients and input product terms. However, the polynomial and linear time-invariant (LTI) filter coefficients, a_p and $h[n]$ in (2.52), respectively, are modeled separately. Thus, they can be estimated separately in a two-step estimation procedure [23].

First, the vector $\underline{a} = \begin{bmatrix} a_1 & a_3 & \cdots & a_P \end{bmatrix}^T$ that contains all distortion coefficients a_p for $p \leq P$ and characterizes the static nonlinearity of the PA, is estimated using the LS estimator (3.9) such that

$$\hat{\underline{a}} = \left(\check{X}_{\mathrm{in},n}^H \check{X}_{\mathrm{in},n} \right)^{-1} \check{X}_{\mathrm{in}}^H \, \underline{x}_{\mathrm{PA},n}, \tag{3.13}$$

with input matrix

$$\check{X}_{\mathrm{in},n} = \begin{bmatrix} x_{\mathrm{in}}[n] & |x_{\mathrm{in}}[n]|^2 x_{\mathrm{in}}[n] & \cdots & |x_{\mathrm{in}}[n]|^{P-1} x_{\mathrm{in}}[n] \\ x_{\mathrm{in}}[n-1] & |x_{\mathrm{in}}[n-1]|^2 x_{\mathrm{in}}[n-1] & \cdots & |x_{\mathrm{in}}[n-1]|^{P-1} x_{\mathrm{in}}[n-1] \\ \vdots & \vdots & & \vdots \\ x_{\mathrm{in}}[n-N+1] & |x_{\mathrm{in}}[n-N+1]|^2 x_{\mathrm{in}}[n-N+1] & \cdots & |x_{\mathrm{in}}[n-N+1]|^{P-1} x_{\mathrm{in}}[n-N+1] \end{bmatrix}. \tag{3.14}$$

Afterwards, the input signal $x_{\mathrm{in}}[n]$ is passed through the estimated nonlinearity, producing the intermediate signal

$$\hat{x}_{\mathrm{int}}[n] = \sum_{p=1}^{\frac{P+1}{2}} \hat{a}_{2p-1} |x_{\mathrm{in}}[n]|^{2(p-1)} x_{\mathrm{in}}[n]. \tag{3.15}$$

This intermediate signal $\hat{x}_{\text{int}}[n]$ is then used as the input of the LTI filter, and the output of the model reads

$$x_{\text{out}}[n] = \sum_{m=0}^{M} h[m]\hat{x}_{\text{int}}[n-m]. \tag{3.16}$$

Therefore, the LS estimator (3.9) can be utilized to identify the LTI filter coefficients $h[n]$, i.e.,

$$\underline{\hat{h}} = \left(X_{\text{int},n}^H X_{\text{int},n}\right)^{-1} X_{\text{int},n}^H \underline{x}_{\text{PA},n}, \tag{3.17}$$

with input matrix

$$X_{\text{int},n} = \begin{bmatrix} \hat{x}_{\text{int}}[n] & \hat{x}_{\text{int}}[n-1] & \cdots & \hat{x}_{\text{int}}[n-M] \\ \hat{x}_{\text{int}}[n-1] & \hat{x}_{\text{int}}[n-2] & \cdots & \hat{x}_{\text{int}}[n-M-1] \\ \vdots & \vdots & & \vdots \\ \hat{x}_{\text{int}}[n-N+1] & \hat{x}_{\text{int}}[n-N] & \cdots & \hat{x}_{\text{int}}[n-N-M+1] \end{bmatrix}, \tag{3.18}$$

and $\underline{\hat{h}} = \begin{bmatrix} \hat{h}[0] & \hat{h}[1] & \cdots & \hat{h}[M] \end{bmatrix}^T$ that contains the estimated LTI filter coefficients.

The handicap of this two-step estimation procedure is that the estimation of the static nonlinearity is not accurate due to the present memory effects in the measured PA output signal. In other words, the memory effects can be seen as a disturbance for the estimation of the nonlinearity. If the memory effects are not too strong, which is the case for narrow-band systems (cf. Section 1.2), the estimation of the nonlinearity is not exceedingly disturbed [23]. However, this systematic error can be avoided when treating the Hammerstein model as reduced Volterra series and estimate the nonlinearity and the LTI filter coefficients at one go (cf. Section 3.3).

3.5 Parameter Estimation for the Decomposed Volterra Series

Although multiple sub-models are involved to produce one single output signal value, as proposed by Zhu *et. al* in [29] and discussed in Section 2.4.6, the overall system can be estimated by using only one LS estimator. The vector

$$\underline{h}_{\mathcal{V}} = \begin{bmatrix} \underline{h}_{1;\mathcal{V}}^T & \underline{h}_{2;\mathcal{V}}^T & \cdots & \underline{h}_{S+1;\mathcal{V}}^T \end{bmatrix}^T, \tag{3.19}$$

containing all Volterra kernel coefficients of the $S+1$ sub-models can be estimated as

$$\underline{\hat{h}}_{\mathcal{V}} = \left(\check{X}_{\text{in},n}^H \check{X}_{\text{in},n}\right)^{-1} \check{X}_{\text{in},n}^H \, \underline{x}_{\text{PA},n}, \tag{3.20}$$

with the input matrix $\check{X}_{\text{in},n}$ that contains the product terms of the $S+1$ decomposed input signals of the training set, i.e.,

$$\check{X}_{\text{in},n} = \begin{bmatrix} \check{X}_{1;\text{in},n} & \check{X}_{2;\text{in},n} & \cdots & \check{X}_{S+1;\text{in},n} \end{bmatrix}. \tag{3.21}$$

If, however, the sub-models are based on pruned Volterra series, as it is the case for, e.g., the decomposed piecewise memory polynomials (2.66), or the decomposed piecewise reduction-based Volterra series (2.67), the overall model is estimated in the same manner, but neglecting the appropriate Volterra kernel coefficients for each sub-model. In other words, each $\underline{h}_{s;\mathcal{V}}$ and $\check{X}_{s;\mathrm{in},n}$ for $s = 1, \ldots, S+1$ in (3.19) and (3.21), respectively, are pruned as discussed in Section 3.3.

3.6 Parameter Estimation for the Vector-Switched Model

The vector-switched model, proposed by Afsardoost *et. al* in [30] and discussed in Section 2.4.7, is highly flexible in terms of the number of switching regions K and the respective regional models $\mathrm{G}_k(\cdot)$. In fact, once the switching regions are determined, the regional models $\mathrm{G}_k(\cdot)$ can be separately estimated by partitioning the training signal set into the regions according to the switching function. Precisely, the set that contains the time indices addressing the training signal samples, i.e., $I = \{n, n-1, \ldots, n-N+1\}$, is firstly split such that

$$I = \bigcup_{k=1}^{K} I_k, \tag{3.22}$$

where I_k denotes the sub set containing the time indices that correspond to the k-th region according to the switching function, i.e.,

$$I_k = \{n_{k,1}, n_{k,2}, \ldots, n_{k,N_k}\}. \tag{3.23}$$

The I_k are disjoint, i.e., $I_k \cap I_l = \emptyset$ for $k \neq l$. For instance, consider that the switching is based on the magnitude of the current and its $J-1$ previous samples of the training set, then the k-th sub set for a given set of training signals at time instant n can be written as

$$I_k = \left\{ l \colon \underline{r}_{n-l,J} \in \Omega_k,\ l = J, \ldots, N+1 \right\}, \tag{3.24}$$

where $\underline{r}_{m,J} = \begin{bmatrix} |x_{\mathrm{in}}[m]| & |x_{\mathrm{in}}[m-1]| & \cdots & |x_{\mathrm{in}}[m-J+1]| \end{bmatrix}^T$ contains the magnitudes of the m-th and its $J-1$ preceding samples of the training set. Accordingly, the cardinalities $|I_k| = N_k$ of the K sub sets are probably dissimilar and $N = \left(\sum_{k=1}^{K} N_k\right) + J$.

Since every model utilized in this thesis can be described by the Volterra series (cf. Section 2.4.5), it is reasonable to assume that Volterra series are employed for the k-th sub-model, i.e., $G_k(\cdot) = \mathcal{V}_k\{\cdot\}$. Moreover, the sub-models are estimated by performing K separate LS model estimations. For instance, the k-th sub-model is estimated by collecting all N_k vectors $\underline{x}_{\mathrm{in},m,M_k} \in \mathbb{C}^{M_k}$ for $m \in I_k$, i.e.,

$$\underline{x}_{\mathrm{in},n_{k,i},M_k} = \begin{bmatrix} x_{\mathrm{in}}[n_{k,i}] & x_{\mathrm{in}}[n_{k,i}-1] & \cdots & x_{\mathrm{in}}[n_{k,i}-M_k] \end{bmatrix}^T, \qquad i = 1, \ldots, N_k, \tag{3.25}$$

where M_k denotes the memory length of the k-th sub-model. Clearly, because of M_k, these vectors contain signal samples $x_{\mathrm{in}}[n_{k,i}-l]$ for $0 < l \leq M_k$ that possibly correspond to other sub sets. In other words, if models with memory are employed, every signal

sample is not only used to estimate the corresponding sub-model, but multiple sub-models. The LS estimator (3.9) for the Volterra kernel vector $\underline{h}_{k;\mathcal{V}}$ of $\mathcal{V}_k\{\cdot\}$, can be written as

$$\hat{\underline{h}}_{k;\mathcal{V}} = \left(\check{X}^H_{k;\mathrm{in},n}\check{X}_{k;\mathrm{in},n}\right)^{-1}\check{X}^H_{k;\mathrm{in},n}\,\underline{x}_{k;\mathrm{PA},n}, \tag{3.26}$$

with input matrix

$$\check{X}_{k;\mathrm{in},n} = \begin{bmatrix}\check{\underline{x}}_{k;\mathrm{in},n_{k,1}} & \check{\underline{x}}_{k;\mathrm{in},n_{k,2}} & \cdots & \check{\underline{x}}_{k;\mathrm{in},n_{k,N_k}}\end{bmatrix}^T, \tag{3.27}$$

and

$$\underline{x}_{k;\mathrm{PA},n} = \begin{bmatrix}x_{\mathrm{PA}}[n_{k,1}] & x_{\mathrm{PA}}[n_{k,2}] & \cdots & x_{\mathrm{PA}}[n_{k,N_k}]\end{bmatrix}^T. \tag{3.28}$$

Since the length of the training set N is limited, the number of sub-models K and the respective model parameters should be chosen such that enough data for each sub-model is available for accurate LS estimation.

3.7 Adaptive Model Parameter Estimation

Up to now, the PA characteristics were assumed to remain constant throughout operation. Whereas extrinsic and intrinsic influences, e.g., self-heating effects, could alter the behavior of the PA over time [6], the PA characteristics can be assumed to be time-varying and the model has to be estimated adaptively. Consequently, as from now the Volterra kernel vector of the model $\underline{h}_{\mathcal{V}}$ is time-dependent, i.e., $\underline{h}_{\mathcal{V}}[n]$.

Stochastic-gradient algorithms are widely used for adaptive system identification, since they are computationally efficient, simple, and show good performance under varying estimation parameters [37]. In fact, stochastic-gradient algorithms employ the method of *steepest descent optimization*, which localizes the minimum of a multivariate quadratic cost function $\varepsilon^2(\underline{h}_{\mathcal{V}})$. The steepest descent direction at any point of $\varepsilon^2(\underline{h}_{\mathcal{V}})$ gives the best direction to move in order to find a point with lower cost. If this such steps are performed repeatedly, the minimum of $\varepsilon^2(\underline{h}_{\mathcal{V}})$ is reached. In other words, the cost is narrowed down by moving in the direction of the steepest descent, i.e., the negative gradient of the cost function, with adequate step-size. The principle of stochastic-gradient algorithms is to start with an initial guess for the estimation parameters $\hat{\underline{h}}_{\mathcal{V}}[-1]$, and then improve the estimate in an iterative manner in the form of [38]

$$\{\text{new guess}\} = \{\text{old guess}\} + \{\text{correction term}\},$$

specifically,

$$\hat{\underline{h}}_{\mathcal{V}}[n] \triangleq \hat{\underline{h}}_{\mathcal{V}}[n-1] - \mu\nabla\varepsilon^2(\hat{\underline{h}}_{\mathcal{V}}[n-1]), \tag{3.29}$$

where $\mu > 0$ is the so-called *step-size*, since it affects the impact of the correction term on the "new guess".

Obviously, different choices for the cost function generally lead to diverse estimation results, each of which leading to a different sense of optimum. One such cost function is

the mean square error (MSE)

$$\varepsilon^2(\hat{\underline{h}}_{\mathcal{V}}[n]) \triangleq \mathrm{E}\left\{\mathbf{e}^*[n]\mathbf{e}[n]\right\} = \mathrm{E}\left\{\left|\mathbf{x}_{\mathrm{PA}}[n] - \hat{\underline{\mathbf{h}}}_{\mathcal{V}}^T[n]\check{\underline{\mathbf{x}}}_{\mathrm{in},n}\right|^2\right\}. \tag{3.30}$$

In order to approximate the in general unknown expectation in (3.30), it can be simply replaced by the instantaneous value of its argument. This leads to the so-called LMS recursion, also sometimes referred to as the Widrow-Hoff algorithm in honor of its originators [38, 39]

$$\hat{\underline{h}}_{\mathcal{V}}[n] \triangleq \hat{\underline{h}}_{\mathcal{V}}[n-1] + \mu\check{\underline{x}}_{\mathrm{in},n}^* e[n] \qquad n > 0. \tag{3.31}$$

Accordingly, the choice of the step-size μ is critical for achieving a good performance. A too small step-size requires an excessive number of iterations in order to minimize the MSE. Additionally, the LMS possibly can hardly track PA behavior for the time varying $\hat{h}_{\mathcal{V}}[n]$ due to the small step-size. However, a too large step-size might slow down convergence, and a far too large μ even causes the algorithm to diverge, where the subsequent cost is greater than the actual one, i.e., $\varepsilon^2(\underline{h}_{\mathcal{V}}[n]) > \varepsilon^2(\underline{h}_{\mathcal{V}}[n-1])$. Therefore, a proper choice of the step-size is indispensable for the correct functionality of the LMS and guidelines for selecting a proper step-size exist [40]. A necessary condition to ensure the LMS algorithm to converge in the mean-square sense is [38]

$$0 < \mu < \frac{2}{\lambda_{\max}}, \tag{3.32}$$

and the fastest convergence is achieved with the choice

$$0 < \mu < \frac{2}{\lambda_{\max} + \lambda_{\min}}, \tag{3.33}$$

where $\lambda_{\min}$ and $\lambda_{\max}$ denote the minimum and maximum eigenvalue of the autocorrelation matrix $R_{\check{\underline{\mathbf{x}}}_{\mathrm{in},n},\check{\underline{\mathbf{x}}}_{\mathrm{in},n}}$, respectively. However, in most cases the statistics, especially the higher moments of $\check{\underline{\mathbf{x}}}_{\mathrm{in},n}$ are unknown, e.g., for the memory polynomial model (2.50) with $P = 5$ and $M = 0$, the autocorrelation matrix $R_{\check{\underline{\mathbf{x}}}_{\mathrm{in},n},\check{\underline{\mathbf{x}}}_{\mathrm{in},n}}$ reads

$$R_{\check{\underline{\mathbf{x}}}_{\mathrm{in},n},\check{\underline{\mathbf{x}}}_{\mathrm{in},n}} = \mathrm{E}\left\{\check{\underline{\mathbf{x}}}_{\mathrm{in},n}\check{\underline{\mathbf{x}}}_{\mathrm{in},n}^H\right\} = \begin{bmatrix} \mathrm{E}\left\{|\mathbf{x}_{\mathrm{in}}[n]|^2\right\} & \mathrm{E}\left\{|\mathbf{x}_{\mathrm{in}}[n]|^4\right\} & \mathrm{E}\left\{|\mathbf{x}_{\mathrm{in}}[n]|^6\right\} \\ \mathrm{E}\left\{|\mathbf{x}_{\mathrm{in}}[n]|^4\right\} & \mathrm{E}\left\{|\mathbf{x}_{\mathrm{in}}[n]|^6\right\} & \mathrm{E}\left\{|\mathbf{x}_{\mathrm{in}}[n]|^8\right\} \\ \mathrm{E}\left\{|\mathbf{x}_{\mathrm{in}}[n]|^6\right\} & \mathrm{E}\left\{|\mathbf{x}_{\mathrm{in}}[n]|^8\right\} & \mathrm{E}\left\{|\mathbf{x}_{\mathrm{in}}[n]|^{10}\right\} \end{bmatrix}. \tag{3.34}$$

Accordingly, for nonlinearity order P, moments up to order $2P$ are required to determine the eigenvalues and thus the convergence bound (3.32) [23]. For this reason, a more practical bound is [38]

$$0 < \mu < \frac{2}{\mathrm{tr}\left\{R_{\check{\underline{\mathbf{x}}}_{\mathrm{in},n},\check{\underline{\mathbf{x}}}_{\mathrm{in},n}}\right\}}, \tag{3.35}$$

where $\mathrm{tr}\{\cdot\}$ denotes the trace operator[1] that can be approximated

$$\mathrm{tr}\left\{R_{\underline{\check{\mathbf{x}}}_{\mathrm{in},n},\underline{\check{\mathbf{x}}}_{\mathrm{in},n}}\right\} = \mathrm{E}\left\{\|\underline{\check{\mathbf{x}}}_{\mathrm{in},n}\|_2^2\right\} \approx \underline{\check{\mathbf{x}}}_{\mathrm{in},n}^H \underline{\check{\mathbf{x}}}_{\mathrm{in},n}. \tag{3.36}$$

Inserting this approximation in the LMS recursion (3.31), leads to the ϵ-NLMS recursion [38]

$$\underline{\hat{h}}_{\mathcal{V}}[n] \triangleq \underline{\hat{h}}_{\mathcal{V}}[n-1] + \frac{\bar{\mu}}{\epsilon + \|\underline{\check{x}}_{\mathrm{in},n}\|_2^2} \underline{\check{x}}_{\mathrm{in},n}^* e[n] \qquad n > 0, \tag{3.37}$$

where $\bar{\mu}$ is the step-size and ϵ is an arbitrarily small constant, included in order to avoid division by zero when $\underline{\check{x}}_{\mathrm{in},n}$ is zero or close to the zero vector. The convergence bound for the ϵ-NLMS is found to be

$$0 < \bar{\mu} < 2. \tag{3.38}$$

[1] The trace of an $m \times m$ square matrix A is

$$\mathrm{tr}\{A\} \triangleq a_{11} + a_{22} + \cdots + a_{mm} = \sum_{i=1}^{m} a_{ii}, \text{ where } A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mm} \end{bmatrix}.$$

4

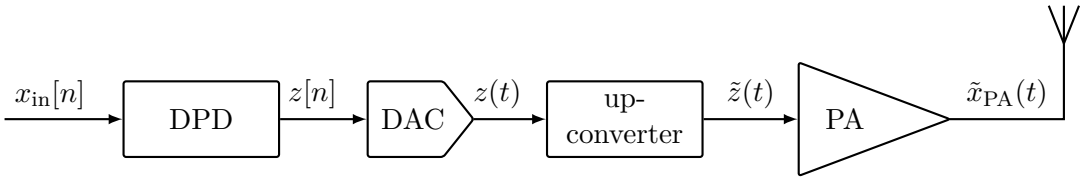
Chapter 4

Linearization by Digital Pre-Distortion

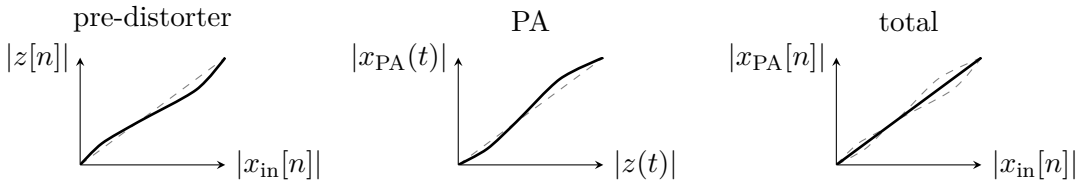
Contents

4.1 Introduction and Outline

As discussed in Chapter 1, in order to meet the spectral requirements under the constraint of high power efficiency, highly linear and highly efficient power amplification is required. To mitigate the discrepancy between linearity and power efficiency, linearization techniques are indispensable. Out of these digital pre-distortion (DPD) is one of the most preferable choices, since it is highly flexible, features excellent linearization capability, and keeps the computational costs moderate [6]. As illustrated in Figure 4.1a, the basic idea of DPD is to distort the baseband representation of the desired transmit signal $x_{\text{in}}[n]$ prior to the up-converter, in order to mitigate nonlinear distortions arising from the use of highly nonlinear but more efficient power amplifiers (PAs). The pre-distorter is characterized by the *scaled complementary inverse* of the PA output characteristics (cf. Figure 4.1b) and aims to compensate the detrimental effects of the transmitter and its spectral re-growth. As Figure 4.1 shows, together with the pre-distorter, the total DPD-PA catenation is intended to behave approximately as a linear system, concealing the nonlinear behavior of the PA itself. Figure 4.1b illustrates amplitude modulation/amplitude modulation (AM/AM) (cf. Section 1.2) correction done by the pre-distorter solely. However, in general, the PA amplitude modulation/phase modulation (AM/PM) conversion needs to be compensated too. Moreover, for broadband systems, e.g., Long Term Evolution (LTE), *dynamic distortions* of the PA come into effect (cf. Section 1.2). Therefore, the pre-distorter should also be capable of eliminating the dispersive behavior of the PA, such that the total response of the DPD-PA catenation is not only linear, but also memoryless.



(a) Schematic of the DPD-PA catenation.



(b) Typical AM/AM correction done by the pre-distorter [18].

Figure 4.1.: Principle of DPD. The pre-distorter is characterized by the scaled complementary inverse of the PA characteristics in order to mitigate the introduced distortions.

The pre-distorter can be implemented at different transmitter stages, e.g., radio frequency (RF)-stage, intermediate frequency (IF)-stage, or at the baseband [18]. However, as the name implies, DPD is performed in the digital baseband. As illustrated in

Figure 4.2, the equivalent baseband PA characteristics are estimated by means of the baseband observations of $z[n]$ and $x_{\mathrm{PA}}[n]$. Likewise, the pre-distorter is characterized by the inverse of the estimated equivalent baseband PA model. Obviously, the DPD structure, depicted in Figure 4.2, uses an *open loop* estimation, since the DPD function itself does not affect the estimation, i.e., the pre-distorter is outside of the estimation loop [6]. After modeling the PA by employing one of the models proposed in Chapter 2, and estimating the model parameters by means of the methods proposed in Chapter 3, the subsequent task is to identify the inverse of the PA characteristics, i.e., the pre-distorter.

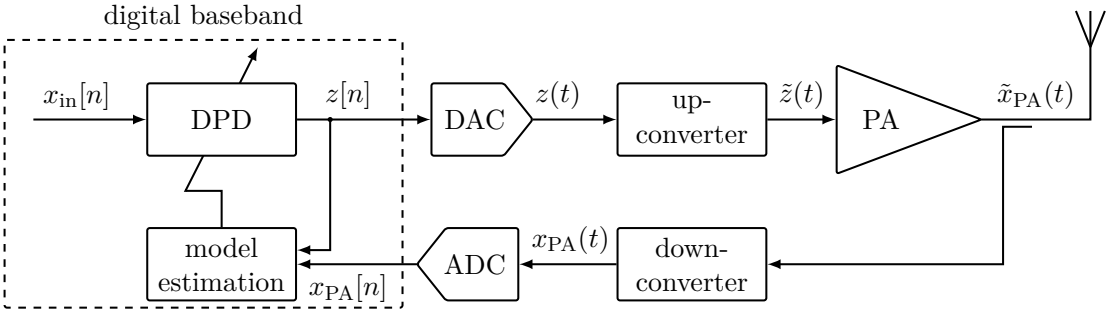


Figure 4.2.: Pre-distorter estimation approach at the digital baseband.

This Chapter is organized as follows

Section 4.2 - Inverse System Identification, gives insights into the inverse modeling of the PA characteristics in order to perform DPD. It introduces two strategies, namely the direct-learning architecture and the indirect-learning architecture.

Section 4.3 - Gain Selection, shows the influence of the desired overall gain.

Section 4.4 - Feedback Bandwidth Requirements, explains the motivation for the requirements of the DPD bandwidth.

4.2 Inverse System Identification

In the context of pre-distortion, inverse system identification is the estimation of the PA inverse characteristics. There are two learning strategies: direct-learning and indirect-learning, which are special techniques of inverse control theory [41].

4.2.1 Direct-Learning Architecture

For the direct-learning architecture, the transfer function of the PA $\mathrm{G}_{\mathrm{PA}}(\cdot)$ itself is first identified, and then, the DPD function is obtained by directly inverting the estimated PA function, as depicted in Figure 4.3. This strategy is commonly used for memoryless

DPD, e.g., look-up table (LUT)-DPD, since the AM/AM and AM/PM conversions are one-to-one mappings, where the inverse can be easily identified, e.g., inverse coordinate mapping [42]. However, for systems where memory effects are not neglectable in order to achieve adequate linearization results, complicated procedures in general are involved to invert the PA model [9].

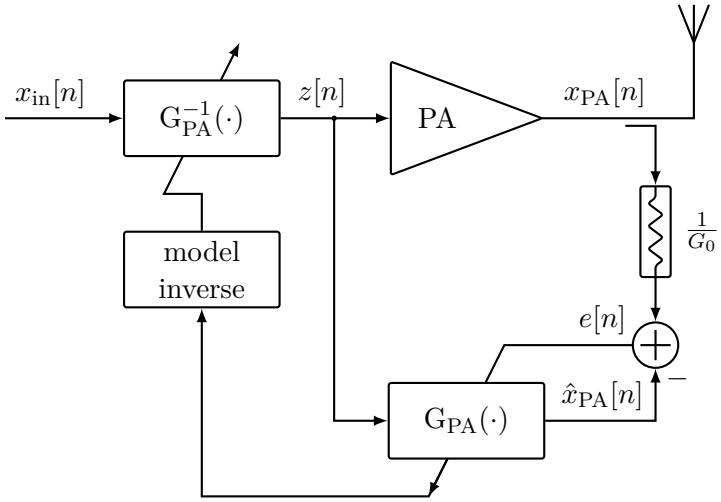


Figure 4.3.: Schematic of the direct-learning architecture [41].

4.2.2 Indirect-Learning Architecture

In the indirect-learning architecture, a post-distorter $\text{G}_{\text{post}}^{-1}(\cdot)$, i.e., a nonlinear system following the PA in order to linearize the transmitter, first estimates the post-inverse of the PA characteristics, and then $\text{G}_{\text{post}}^{-1}(\cdot)$ is used as a pre-distorter, i.e., $\text{G}_{\text{pre}}^{-1}(\cdot)$ [41], as shown in Figure 4.4. This strategy is based on the assumption that $\text{G}_{\text{post}}^{-1}(\cdot)$ and $G_{\text{pre}}^{-1}(\cdot)$ are identical, which has been proven by Schetzen in [24].

The post-inverse $\text{G}_{\text{post}}^{-1}(\cdot)$ can be estimated by employing one of the estimation methods proposed in Chapter 3, where the input and output signals of the estimation method are represented by the PA output and input, $x_{\text{PA}}[n]$ and $x_{\text{in}}[n]$, respectively; the input and output signals of the PA can be thought of being swapped. Assuming that $G_{\text{post}}^{-1}(\cdot)$ is modeled using Volterra series, i.e., $G_{\text{post}}^{-1}(\cdot) = \mathcal{V}\{\cdot\}$, the least squares (LS) estimator for the Volterra kernel vector $\underline{h}_{\mathcal{V},\text{post}}$ reads

$$\hat{\underline{h}}_{\mathcal{V},\text{post}} = \left(\check{X}_{\text{PA},n}^{H}\check{X}_{\text{PA},n}\right)^{-1}\check{X}_{\text{PA},n}^{H}\underline{z}_n. \tag{4.1}$$

Afterwards, $\text{G}_{\text{pre}}^{-1}(\cdot)$ is directly cloned from $\text{G}_{\text{post}}^{-1}(\cdot)$, i.e., $\hat{\underline{h}}_{\mathcal{V},\text{pre}} = \hat{\underline{h}}_{\mathcal{V},\text{post}}$. Clearly, when employing the blockwise LS estimator, $\text{G}_{\text{pre}}^{-1}(\cdot)$ initially represents a simple linear amplification, i.e., $\text{G}_{\text{pre}}^{-1}(\cdot) = 1$, while gathering N pairs of samples of $z[n]$ and $x_{\text{PA}}[n]$, in order to estimate $\text{G}_{\text{post}}^{-1}(\cdot)$. For subsequent blocks, the previous estimate of $\text{G}_{\text{post}}^{-1}$ is then used

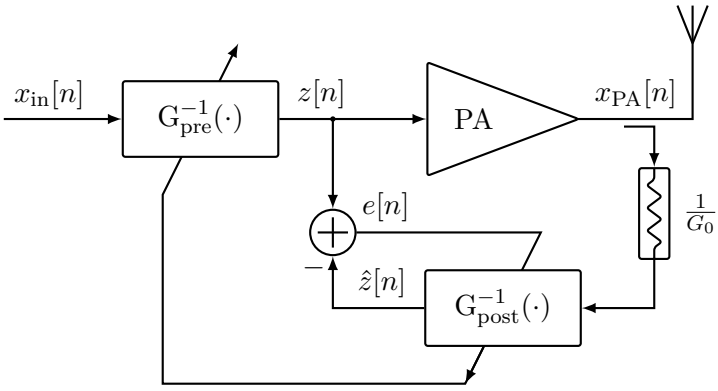


Figure 4.4.: Schematic of the indirect-learning architecture [43].

to generate $z[n]$. However, adaptive estimation methods, as discussed in Section 3.7, are capable of online estimation, e.g., the ϵ-normalized least mean squares (ϵ-NLMS) reads

$$\hat{\underline{h}}_{\mathcal{V},\text{post}}[n] = \hat{\underline{h}}_{\mathcal{V},\text{post}}[n-1] + \frac{\bar{\mu}}{\epsilon + \|\check{\underline{x}}_{\text{PA},n}\|_2^2} \check{\underline{x}}_{\text{PA},n}^* \left(z[n] - \hat{\underline{h}}_{\mathcal{V},\text{post}}^T[n-1] \check{\underline{x}}_{\text{PA},n} \right). \tag{4.2}$$

Note that not all nonlinear systems possess an exact inverse, specially when the AM/AM and AM/PM conversions are no one-to-one mappings [36], as it is the case for PAs with strong memory effects (cf. Section 1.2). In other words, the inversion of a nonlinear system my not be possible and the inverse modeling may diverge [9].

4.3 Gain Selection

If the PA characteristics are invertible, the composition of the exact inverse and the original function would completely compensate each other [36], i.e., $\text{G}\left(\text{G}^{-1}\left(x_{\text{in}}[n]\right)\right) = x_{\text{in}}[n]$. Accordingly, the output of the composition equals the input, and the overall gain would be exactly $G = 1$. Therefore, as shown in Figure 4.3 and 4.4, the feedback path has to be attenuated by the desired overall gain G_0, in order to normalize the output power level to the same power level as the original input signal. Due to this normalization, the extracted parameters can be directly used in the DPD. Otherwise the parameters, i.e., the Volterra kernel, must be nonlinear scaled after model extraction [44].

There are numerous options for choosing the desired gain G_0. Figure 4.5 shows three arbitrary choices, which are G_1, G_2, and G_3. Moreover, Figure 4.5 shows that the pre-distorter can successfully correct the distortions only up to the full saturation level of the PA [44]. Additionally, for high input power levels, the required pre-distorter gain must skyrocket, in order to compensate the strong signal compression (cf. Figure 4.1b). On the other hand, the output power level of the pre-distorter cannot follow this shoot up in order not to overdrive the PA input. One way to choose G_0 is, in such way that the magnitude of the input signal spans the maximum linearizable magnitude range [44],

i.e.,

$$G_0 = \frac{\max\limits_{0 \le n < N} \{|x_{\mathrm{PA}}[n]|\}}{\max\limits_{0 \le n < N} \{|x_{\mathrm{in}}[n]|\}}, \tag{4.3}$$

where the AM/AM conversion is assumed as a strictly monotonically increasing mapping in order to ensure that the output power peak level occurs at the maximum input power level. This choice for G_0 represents G_3 in Figure 4.5. Accordingly, all input and output signals can be normalized to unity. Additionally, normalizing at the peak eases system implementation, since all signals can be normalized by the same scaling factor that facilitates power control [44].

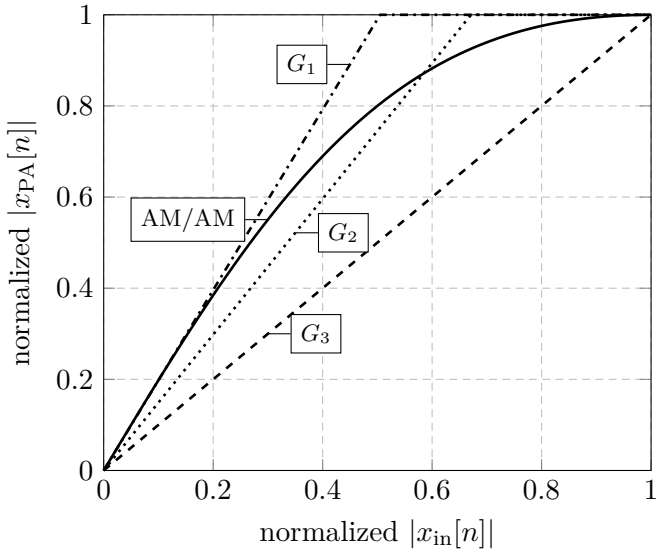


Figure 4.5.: Desired gain selection [44]. G_1 and G_2 represent arbitrary choice for the desired gain, and G_3 respresents the saturation gain.

4.4 Feedback Bandwidth Requirements

In order to compensate the adverse out-of-band spectral emissions due to nonlinear amplification, the bandwidth of the feedback path in the learning architectures (cf. Section 4.2), has to be sufficiently wide. As discussed in Section 2.3, the feedback bandwidth must be P-times the transmission bandwidth B, where P denotes the highest order of significant nonlinearity of the PA to be linearized. For instance, if the feedback bandwidth is $5B$, the pre-distorter can only compensate first, third, and fifth order distortions.

5

Chapter 5

Measurement Results

Contents

5.1 Introduction and Outline

This Chapter investigates the performance of linearizing an envelope tracking power amplifier (ETPA), employing the indirect-learning architecture discussed in Section 4.2.2 with various digital pre-distortion (DPD) models. Namely, pruned Volterra series discussed in Section 2.4.5, the decomposed piecewise Volterra model proposed in Section 2.4.6, and the vector-switched model proposed in Section 2.4.7. The performance is evaluated based on measurements of a prototype envelope tracking system.

This Chapter is organized as follows

Section 5.2 - Measuring Configuration, shows the measurement setup for the prototype envelope tracking system and the utilized instruments. Moreover, it lists the input signal characteristics and the transmission parameters. Finally, the evaluation criteria are illustrated.

Section 5.3 - Reduced Volterra Models Experimental Results, analyses the linearization performance of the reduced Volterra models discussed in Section 2.4.5, based on measurements.

Section 5.4 - Decomposed Piecewise Volterra Model Experimental Results, reveals the DPD performance of the decomposed piecewise Volterra model discussed in Section 2.4.6, based on measurements.

Section 5.5 - Vector-Switched Model Experimental Results, evaluates the DPD performance of the vector-switched model discussed in Section 2.4.7, based on measurements

Section 5.6 - Discussion, compares the performance of the three afore evaluated models and give thoughts about model complexity.

5.2 Measuring Configuration

The prototype envelope tracking system and the measurement setup was provided by Artesyn Austria GmbH & CoKG. The schematic of the provided test-bench is depicted in Figure 5.2 and the employed instruments are listed in Table 5.1. It was designed to be fully controlled by MATLAB. All digital signal processing that was not required for real-time operation was implemented in MATLAB. These were, e.g., test signal generation, parameter estimation for the employed pre-distorter, envelope calculation, parameter evaluation, and some auxiliary functions, like signal synchronization and phase alignment. The generated baseband test-signal was a Wideband Code Division Multiple Access (WCDMA)-like signal comprised of two channels side-by-side and 7.5dB peak-to-average power ratio (PAPR). Figure 5.1 shows the power spectral density (PSD). The transmit signal was sent via Ethernet to the vector signal generator (VSG), which directly up-converted it to the radio frequency (RF) with carrier-frequency $f_c = 2.14\text{GHz}$

and sampling frequency $f_s = 125\text{MHz}$. The envelope pattern was directly calculated in MATLAB based on the baseband transmit signal $x_{\text{in}}[n]$ and sent to the envelope tracking unit. The VSG *periodically* outputted the passband transmit signal in order to avoid a variation of the power amplifier (PA) characteristics due to power level fluctuations (cf. Section 1.2). Moreover the VSG triggered the envelope tracking unit. Since there was a signal propagation delay between the VSG and the ETPA, MATLAB allows to adjust the trigger delay of the envelope tracking unit in order to ensure the alignment of the transmit signal and the envelope signal. Moreover, the envelope signal was monitored by the oscilloscope in order to screen the delay alignment with the transmit signal. The pre-amplifier was a highly linear broadband amplifier with 45dB gain and protected against misuse by the isolator. The PA was operated at 2.14GHz with an average output power level of 47W. The attenuated output signal of the ETPA was converted down to digital baseband by the vector signal analyser (VSA) with 100MHz signal analysis bandwidth and read back by MATLAB via Ethernet. Additionally, the RF output power level was monitored. All instruments were synchronized via the 10MHz reference clock, generated by the VSG. The instruments after the ETPA (attenuator, directional coupler) had an overall insertion loss of 51.6dB.

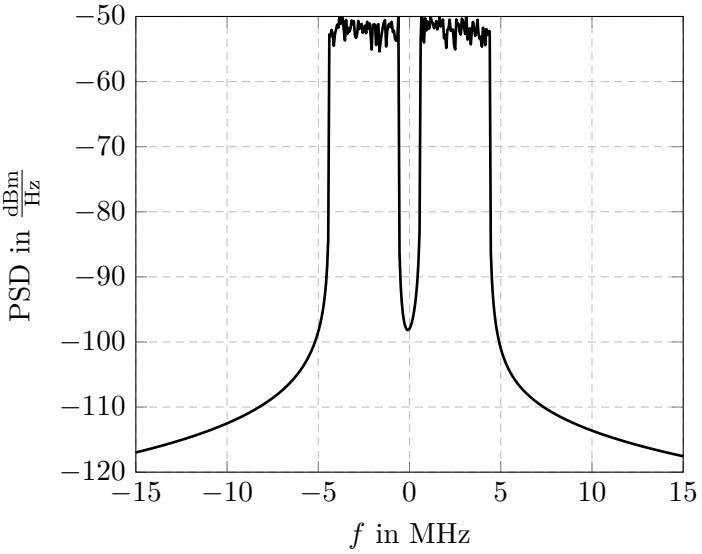


Figure 5.1.: PSD of the transmit signal $x_{\text{in}}[n]$ without any DPD.

As mentioned before, the MATLAB script allowed to select different DPD algorithms and to freely chose the appropriate model parameters. Since the test-bench was no real-time system, the measurements were performed in burst mode. In other words, each measurement required two steps. First, the envelope tracking system amplified the desired transmit signal $x_{\text{in}}[n]$ without any DPD, and the attenuated ETPA output was read back. The transmit signal had a block length of 2^{17} I/Q samples. The gathered input and output samples were normalized to the acquired RF output power level in order

to adjust the PA gain (cf. section 4.3). 20 000 of these 2^{17} I/Q input and output samples were used to estimate the parameters for the selected DPD algorithm (cf. Chapter 3). The remaining samples were used to evaluate the performance. Afterwards, the desired input signal is pre-distorted and transmitted. These measurements were repeated five times.

Three assessment were considered, namely, the normalized mean square error (NMSE), the adjacent channel power ratio (ACPR), and the adjacent channel error power ratio (ACEPR) (cf. Section 2.5). The ACPR was calculated for three higher and lower adjacent channels, respectively. Precisely, the transmission band was set to

$$B_{\text{ch}} = \{f\colon\ f_{\text{c}} - B \leq f \leq f_{\text{c}} + B\}\,, \tag{5.1}$$

with channel bandwidth $B = 5\text{MHz}$. The i-th *higher* adjacent channel was set to

$$B_{\text{adj,hi},i} = \{f\colon\ f_{\text{c}} + iB \leq f \leq f_{\text{c}} + (i+1)\,B\}\,, \tag{5.2}$$

and analogously the i-th *lower* adjacent channel occupied the frequency band

$$B_{\text{adj,lo},i} = \{f\colon\ f_{\text{c}} - (i+1)\,B \leq f \leq f_{\text{c}} - iB\}\,. \tag{5.3}$$

However, to evaluate the ACEPR all six adjacent channels were united, i.e.,

$$B_{\text{adj}} = \bigcup_{i=1}^{3} \left(B_{\text{adj,lo,i}} \cup B_{\text{adj,hi,i}}\right). \tag{5.4}$$

As mentioned before, since the DPD feedback bandwidth was 100MHz, the pre-distorter was only capable of countering ninth order nonlinear distortions for the employed input signal with 10MHz bandwidth (cf. Section 4.4).

instrument	type
vector signal generator (VSG)	Agilent N5182A, Option 654
vector signal analyser (VSA)	Rohde & Schwarz FSQ26
power meter	Rohde & Schwarz NRP-Z11
oscilloscope	Lecroy 454
pre-amplifier	AR 30S1G4
PA	Freescale AFT21S232S
isolator	Pasternack PE8300
directional coupler	Pasternack PE2209-20
attenuator	30dB, 80W

Table 5.1.: List of the employed instruments in the measurement setup illustrated in Figure 5.2.

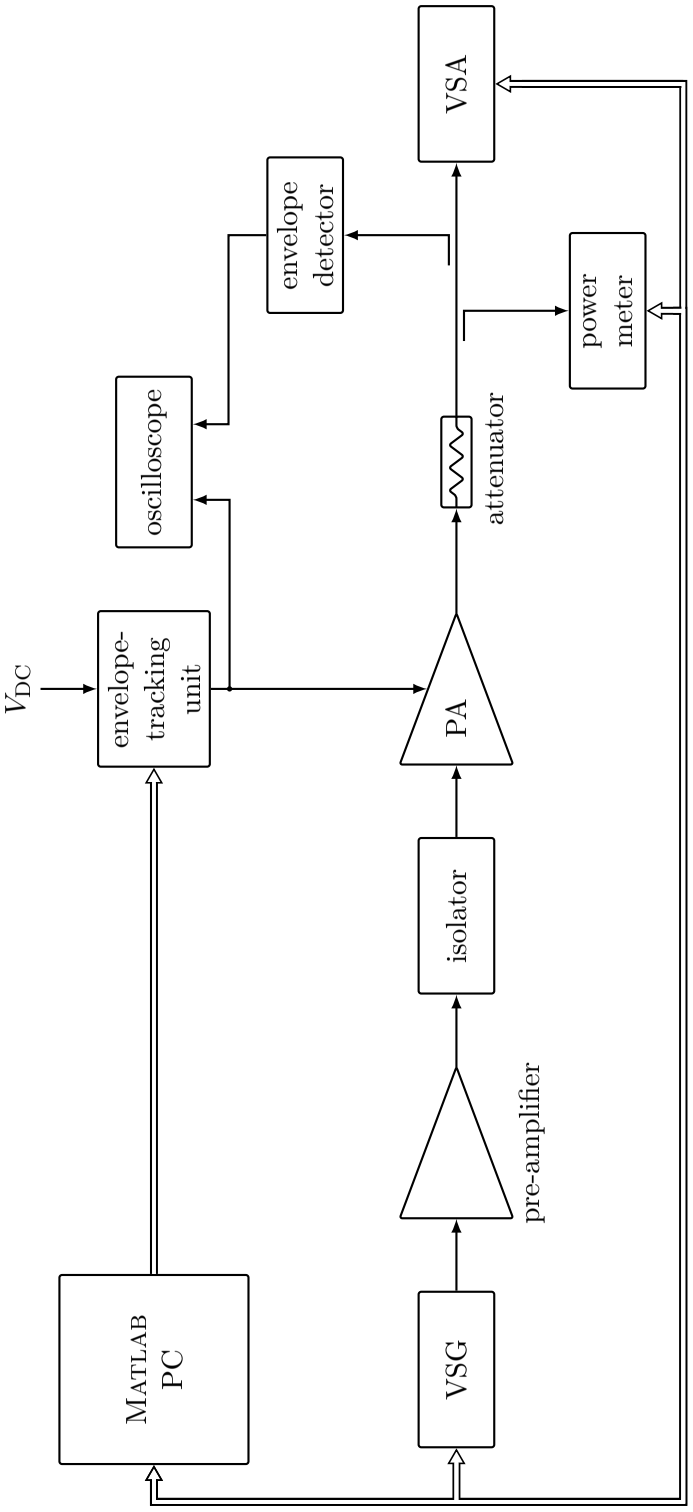


Figure 5.2.: Outline of the measurement setup. The corresponding instruments are listed in Table 5.1.

5.3 Reduced Volterra Models Experimental Results

In order to evaluate the performance of the reduced Volterra models discussed in Section 2.4.5, the *memory polynomials* and the *dynamic deviation reduction-based Volterra series* were considered. Table 5.2 lists the NMSE, ACPR, and ACEPR of the first model, Table 5.3 of the latter one. The nonlinearity order P was varied from third up to ninth order with either $M = 0$, i.e., the static case[1], or with two memory taps, i.e., $M = 2$.

order		NMSE	ACPR						ACEPR
P	M		low,3	low,2	low,1	hi,1	hi,2	hi,3	
w/o DPD		-21.1	-39.5	-34.1	-28.8	-29.4	-33.0	-37.8	-21.8
3	0	-23.1	-39.4	-35.2	-32.5	-30.5	-33.2	-37.7	-23.9
3	2	-23.7	-39.1	-34.9	-32.6	-31.2	-33.6	-37.7	-24.2
5	0	-24.4	-39.4	-35.5	-33.3	-32.7	-34.8	-38.1	-25.5
5	2	-25.4	-39.1	-35.7	-34.1	-33.7	-35.3	-38.4	-26.1
7	0	-26.7	-41.6	-39.0	-36.0	-35.7	-37.9	-40.3	-28.7
7	2	-28.7	-41.5	-39.6	-37.9	-37.8	-39.6	-41.1	-30.3
9	0	-27.7	-43.6	-40.3	-37.1	-36.9	-39.6	-42.1	-30.0
9	2	-30.4	-43.8	-42.2	-40.1	-39.8	-41.4	-42.9	-32.4

Table 5.2.: NMSE, ACPR, and ACEPR in dB of the memory polynomials as DPD model.

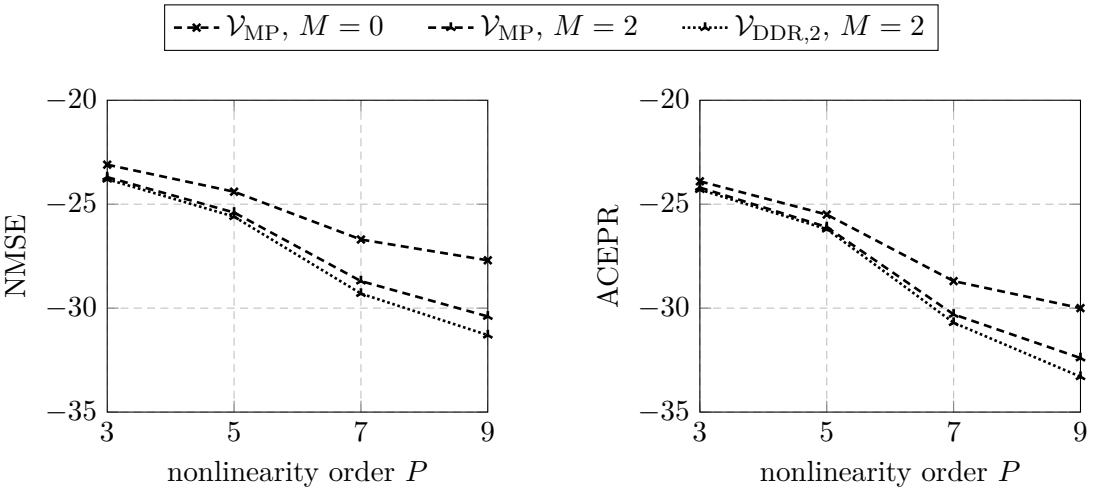
order			NMSE	ACPR						ACEPR
P	M	R		low,3	low,2	low,1	hi,1	hi,2	hi,3	
w/o DPD			-21.1	-39.5	-34.1	-28.8	-29.4	-33.0	-37.8	-21.8
3	2	1	-23.8	-38.9	-34.8	-32.6	-31.4	-33.8	-37.9	-24.3
5	2	1	-25.6	-38.9	-35.5	-34.1	-34.0	-35.5	-38.5	-26.2
7	2	1	-29.1	-41.1	-39.5	-38.3	-38.2	-39.5	-40.9	-30.6
9	2	1	-31.3	-43.7	-42.1	-40.5	-40.5	-41.6	-43.1	-33.2
3	2	2	-23.8	-38.7	-34.7	-32.5	-31.5	-33.9	-37.7	-24.3
5	2	2	-25.6	-38.8	-35.5	-34.0	-34.0	-35.5	-38.3	-26.2
7	2	2	-29.3	-41.5	-39.6	-37.9	-38.5	-39.8	-41.2	-30.7
9	2	2	-31.3	-43.6	-42.2	-40.6	-40.9	-41.9	-43.1	-33.3

Table 5.3.: NMSE, ACPR, and ACEPR in dB of the dynamic deviation reduction-based Volterra series as DPD model.

Figure 5.3 compares the performance of the memory polynomials with $M = 0$, $M = 2$,

[1] Note that the static case for the memory polynomials coincide with that of the dynamic deviation reduction-based Volterra series.

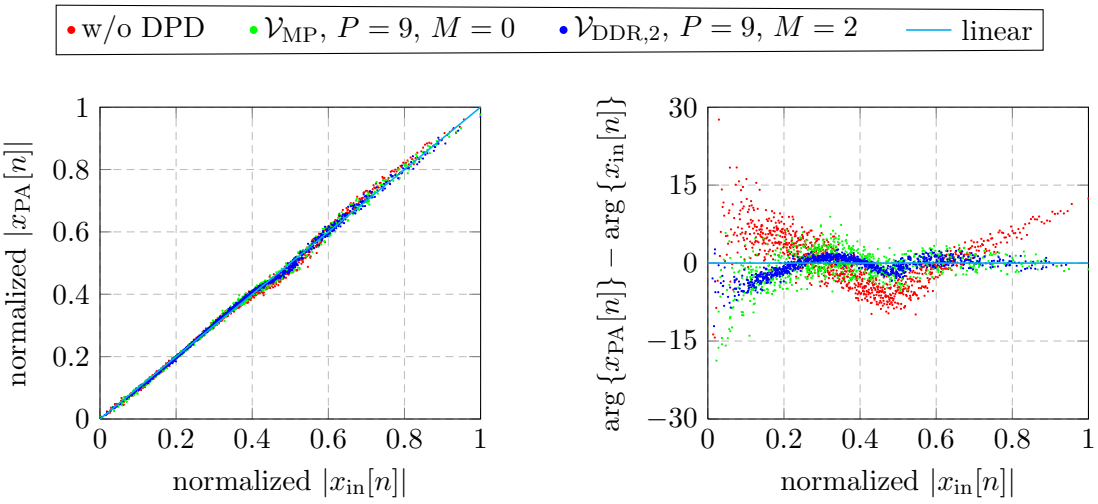
and the second order deviation reduction-based Volterra series with $M = 2$, with respect to the varying nonlinearity order P of the pre-distorter. Figure 5.3a shows the NMSE and Figure 5.3b the ACEPR. Accordingly, for increasing P the performance of the pre-distorter increased. Beside nonlinear distortions, the ETPA exhibited a dispersive behavior (cf. Section 1.2). Compared to the static case, the pre-distorter that incorporated memory effects, improved the NMSE and ACEPR by about 3dB for $P = 9$. Nevertheless, the first order deviation-reduction based Volterra series showed a slightly better performance than the memory polynomials. Moreover, the second order dynamic deviation reduction-based Volterra series even more improved the performance, especially in the second and third adjacent channel.



(a) NMSE comparison. **(b)** ACEPR comparison.

Figure 5.3.: Comparison of the NMSE and the ACEPR of the pre-distorter that employs either the static polynomials, the memory polynomials, or the second order deviation reduction-based Volterra series, both of the latter with $M = 2$.

Figure 5.4 shows the nonlinear behavior of the ETPA compared to the specific DPD configuration that showed the best NMSE performance, namely the second order deviation reduction-based Volterra series with $P = 9$ and $M = 2$. Additionally, the nonlinear behavior when employing static polynomials as DPD model is illustrated in order to demonstrate the achievable performance gain due to the elimination of memory effects. Figure 5.4a and Figure 5.4b show the amplitude modulation/amplitude modulation (AM/AM) and amplitude modulation/phase modulation (AM/PM) conversions, respectively. Although most memory effects were effectively eliminated after DPD, the linearization performance was rather limited, especially in the AM/PM conversion.
Figure 5.5 shows the PSDs of the PA output signal $x_{\mathrm{PA}}[n]$ for three different DPD models with $P = 9$, namely the static polynomials, the memory polynomials, and the second order deviation reduction-based Volterra series, both of the latter with $M = 2$. Accordingly, the ACPR suffered from not considering memory effects, especially for the



(a) AM/AM conversion. **(b)** AM/PM conversion.

Figure 5.4.: Measured AM/AM and AM/PM characteristics when employing either second order deviation reduction-based Volterra series with $P = 9$ and $M = 2$ or static polynomials with $P = 9$.

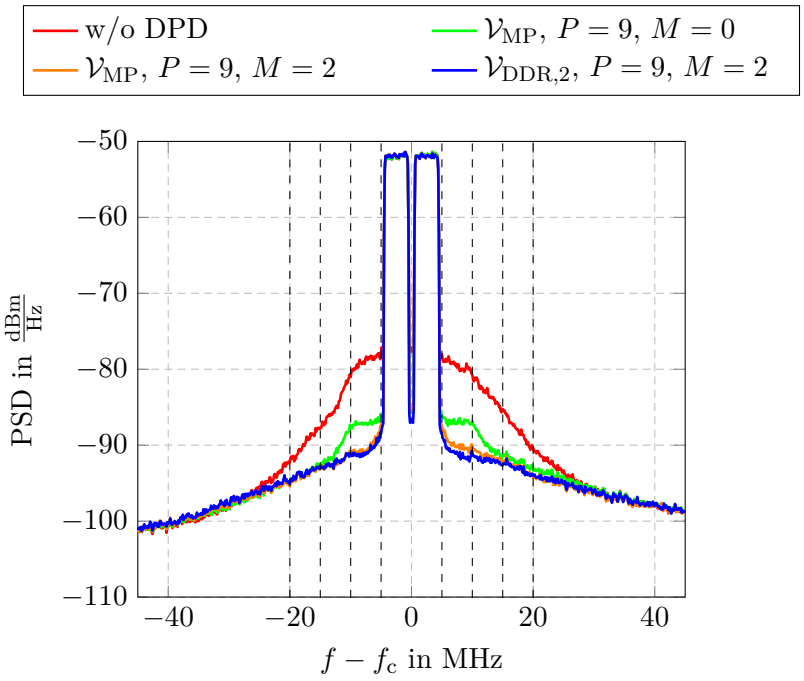


Figure 5.5.: Measured PSDs of $x_{\mathrm{PA}}[n]$ for either employing static polynomials, memory polynomials with $M = 2$, or the second order deviation reduction-based Volterra series with $M = 2$, as DPD models all with $P = 9$.

first and second adjacent channels. Moreover, one can see a slight performance improvement of the second order deviation reduction-based Volterra series compared to memory polynomials.

5.4 Decomposed Piecewise Volterra Model Experimental Results

The performance of the decomposed piecewise Volterra model discussed in Section 2.4.6 was evaluated for a particular set of thresholds $\Lambda = \{0.2, 0.4, 0.6, 0.8\}$[1]. Accordingly, five models, i.e., $S + 1 = 5$, were included in order to pre-distort the input signal. Table 5.4 and Table 5.5 list the NMSE, ACPR, and ACEPR of the *decomposed piecewise memory polynomials* and the *decomposed piecewise dynamic deviation reduction-based Volterra series*, respectively. All regional models had the same configuration, i.e., $P_s = P_1 = P_2 = \cdots = P_5$ and $M_s = M_1 = M_2 = \cdots = M_5$; the nonlinearity order P_s was varied from third up to seventh order with either $M_s = 0$, $M_s = 1$, or $M_s = 2$.

order		NMSE	ACPR						ACEPR
P_s	M_s		low,3	low,2	low,1	hi,1	hi,2	hi,3	
w/o DPD		-21.1	-39.5	-34.1	-28.8	-29.4	-33.0	-37.8	-21.8
5	0	-28.7	-46.9	-42.1	-37.9	-37.8	-41.4	-45.2	-31.5
3	2	-32.2	-48.2	-45.9	-42.1	-41.8	-44.7	-46.8	-35.3
5	2	-32.2	-48.3	-45.9	-42.0	-41.8	-44.6	-46.6	-35.2
7	2	-32.2	-48.3	-45.7	-42.0	-41.5	-44.3	-46.7	-35.1

Table 5.4.: NMSE, ACPR, and ACEPR in dB of the decomposed piecewise memory polynomials as DPD model.

order			NMSE	ACPR						ACEPR
P_s	M_s	R_s		low,3	low,2	low,1	hi,1	hi,2	hi,3	
w/o DPD			-21.1	-39.5	-34.1	-28.8	-29.4	-33.0	-37.8	-21.8
3	1	1	-33.7	-47.6	-45.1	-42.7	-42.4	-44.2	-46.5	-36.6
5	1	1	-33.8	-48.0	-45.5	-42.8	-42.7	-44.5	-46.7	-36.9
7	1	1	-34.1	-48.1	-45.6	-42.9	-43.0	-44.6	-46.7	-37.1
3	2	2	-33.9	-48.1	-46.1	-43.4	-42.9	-45.1	-47.0	-37.3
5	2	2	-34.3	-48.4	-46.7	-43.6	-43.6	-45.9	-47.2	-37.9

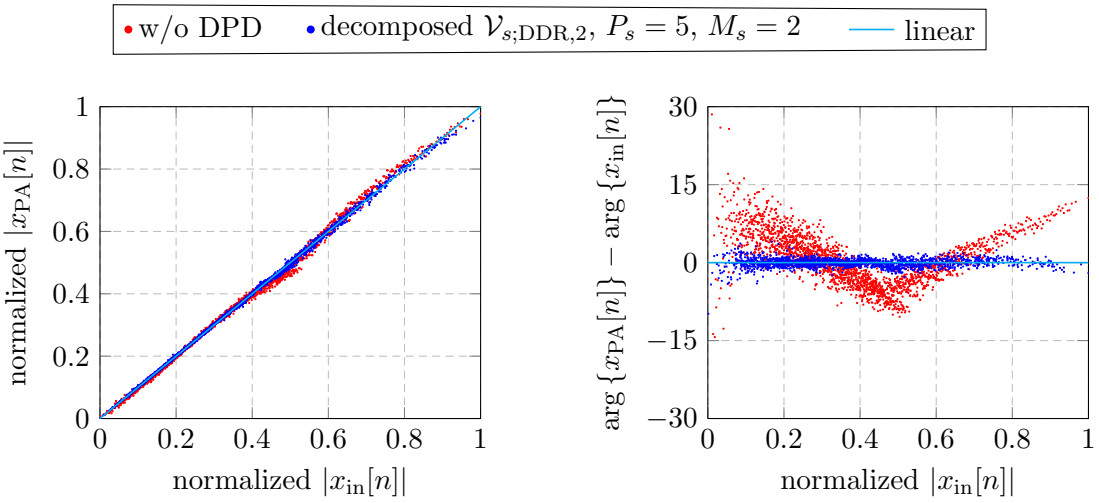
Table 5.5.: NMSE, ACPR, and ACEPR in dB of the decomposed piecewise dynamic deviation reduction-based Volterra series as DPD model.

Accordingly, the performance gain of increasing the nonlinearity order P_s was rather

[1] Note that the thresholds are normalized.

limited. This was because the five sub-models have already well adapted to the ETPA inverse characteristics even for $P_s = 3$. However, comparing the decomposed piecewise memory polynomials to the decomposed piecewise dynamic deviation reduction-based Volterra series, one can observe that the latter model showed a slightly better performance. For instance, the ACEPR improved by around 2dB for deviation order $R_s = 1$, and by roughly 2.5dB for $R_s = 2$, while the ACPR improved by almost 1dB and roughly 1.5dB for $R_s = 1$ and $R_s = 2$, respectively for the first and second adjacent channel.

Figure 5.6 shows the nonlinear characteristics of the ETPA when the decomposed piecewise dynamic deviation reduction based Volterra series with $P_s = 5$ and $M_s = 2$ were employed. Figure 5.6a and Figure 5.6b show the AM/AM and AM/PM conversions, respectively. From these two Figures, it can be seen that the decomposed piecewise Volterra model featured an excellent linearization capability. The memory effects were significantly compensated after DPD. Additionally, without DPD, AM/PM distortions spread over a range of 30 °, while introducing DPD reduced this spread to a range of 5 °.



(a) AM/AM conversion. **(b)** AM/PM conversion.

Figure 5.6.: Measured AM/AM and AM/PM characteristics when employing decomposed piecewise dynamic deviation reduction-based Volterra series with $R_s = 2$, $P_s = 5$, and $M_s = 2$ as DPD model.

Figure 5.7 compares the PSD of the ETPA output signal $x_{\mathrm{PA}}[n]$ when the decomposed piecewise static polynomials, the decomposed piecewise memory polynomials with $M_s = 2$, the decomposed piecewise dynamic deviation reduction-based Volterra series with $R_s = 2$ and $M_s = 2$, as DPD models all with $P_s = 5$ were employed. Additionally, it shows the PSD of $x_{\mathrm{PA}}[n]$ when the input signal was not decomposed and the second order deviation reduction-based Volterra series with $P = 5$ and $M = 2$ were employed as pre-distorter, as it is the case in Section 5.3. Since five sub-models were employed, it is not fair to compare the performance of the sub-model to the overall model. Nev-

ertheless, it is depicted in order to illustrate the ACPR gain when employing the signal decomposition. Moreover, one can see that the decomposed piecewise dynamic deviation reduction-based Volterra series performed slightly better than the decomposed piecewise memory polynomials, especially for the first and second adjacent channel.

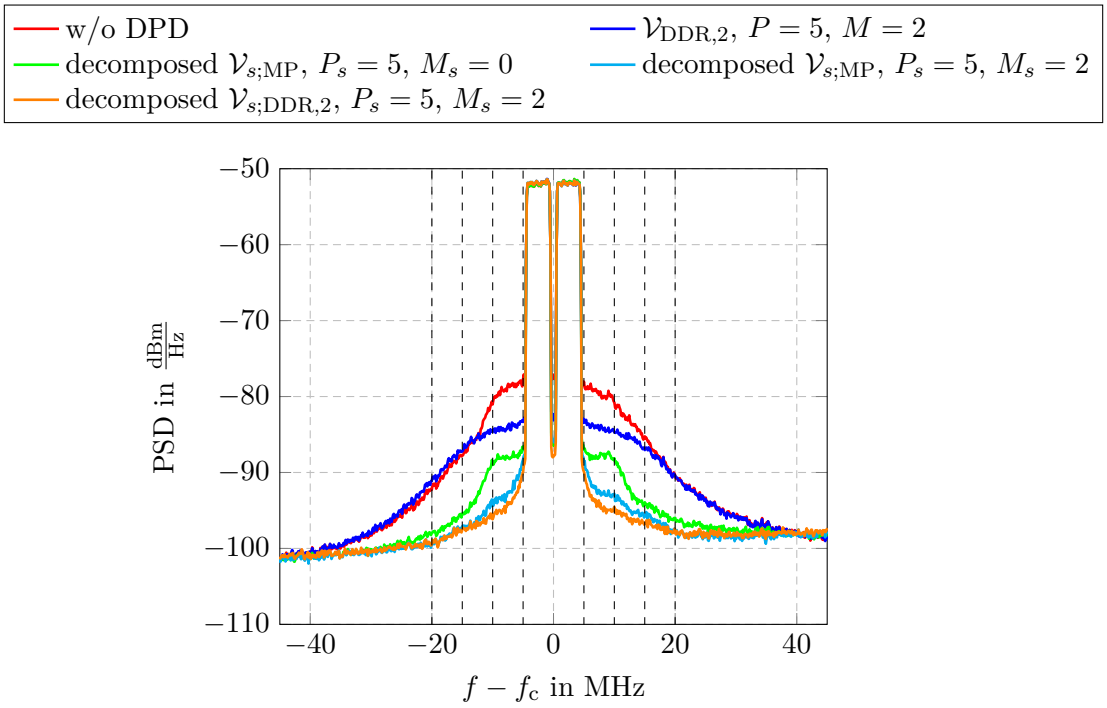


Figure 5.7.: Measured PSDs of $x_\mathrm{PA}[n]$ of the second order dynamic deviation reduction-based Volterra series with $P = 5$ and $M = 2$, the decomposed piecewise static polynomials, the decomposed piecewise memory polynomials, and the decomposed piecewise dynamic deviation reduction-based Volterra series with $R_s = 2$, as DPD models all with $P_s = 5$ and $M_s = 2$.

5.5 Vector-Switched Model Experimental Results

The performance of the vector-switched model, discussed in Section 2.4.7 was evaluated for $J = 1$. This means that the switching space had dimension $\mathbb{R}$ and the magnitude of the instantaneous input signal $|x_\mathrm{in}[n]|$ selected the sub-model to be employed. Moreover, experiments have shown that no performance gain was noticeable when employing more than 30 sub-models ($K > 30$). Thus, the number of sub-models was set to $K = 30$. Since in total 20 000 samples were used for parameter estimation (cf. Section 5.2), each region had roughly 670 samples in order to estimate the sub-model. Table 5.6 and Table 5.7 list the NMSE, ACPR, and ACEPR of the *vector-switched memory polynomials* and the *vector-switched dynamic deviation reduction-based Volterra series*, respectively. All

of the 30 sub-models had the same configuration, i.e., $P_k = P_1 = P_2 = \cdots = P_{30}$ and $M_k = P_1 = M_2 = \cdots = M_{30}$. The sub-model nonlinearity order P_k was either $P_k = 3$ or $P_k = 5$, all other sub-model parameters were chosen like for the decomposed piecewise Volterra series (cf. Section 5.4).

order		NMSE	ACPR						ACEPR
P_k	M_k		low,3	low,2	low,1	hi,1	hi,2	hi,3	
w/o DPD		-21.1	-39.5	-34.1	-28.8	-29.4	-33.0	-37.8	-21.8
3	2	-34.3	-48.2	-45.9	-42.9	-43.0	-45.1	-47.2	-37.2
5	0	-28.7	-46.7	-42.0	-38.1	-37.5	-41.2	-45.1	-31.3
5	2	-34.5	-48.5	-46.3	-43.5	-43.5	-45.7	-47.6	-37.8

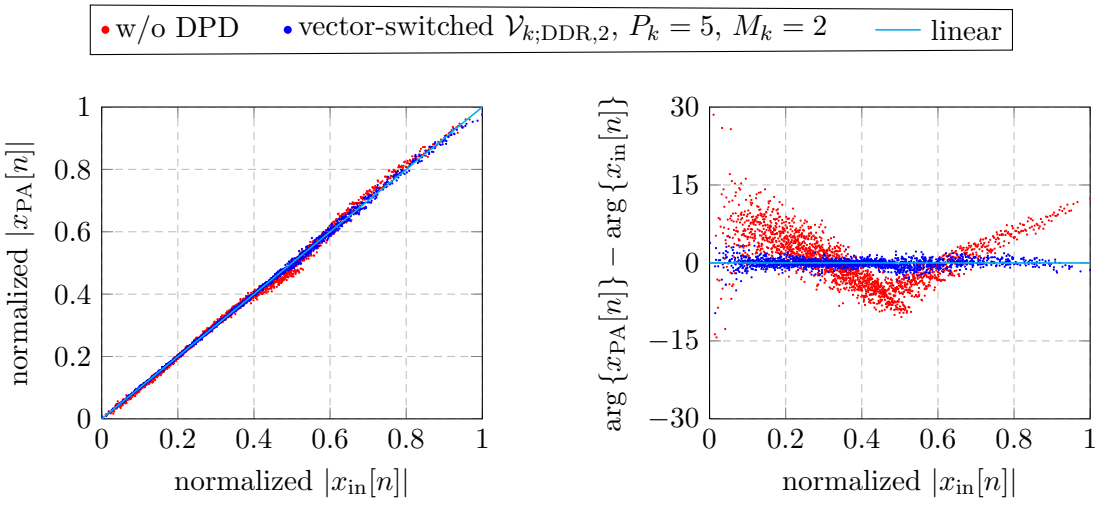
Table 5.6.: NMSE, ACPR, and ACEPR in dB of the vector-switched memory polynomials as DPD model.

order			NMSE	ACPR						ACEPR
P_k	M_k	R_k		low,3	low,2	low,1	hi,1	hi,2	hi,3	
w/o DPD			-21.1	-39.5	-34.1	-28.8	-29.4	-33.0	-37.8	-21.8
3	1	1	-34.1	-47.7	-45.3	-42.8	-42.7	-44.4	-46.5	-36.8
5	1	1	-34.1	-47.7	-45.3	-42.8	-42.7	-44.4	-46.5	-36.8
3	2	2	-34.5	-48.5	-46.3	-43.4	-43.4	-45.5	-47.2	-37.6
5	2	2	-34.6	-48.6	-46.5	-43.5	-43.6	-45.8	-47.5	-38.0

Table 5.7.: NMSE, ACPR, and ACEPR in dB of the vector-switched dynamic deviation reduction-based Volterra series as DPD model.

Figure 5.8 illustrates the nonlinear characteristics of the ETPA for the vector-switched dynamic deviation reduction-based Volterra series with $R_k = 2$, $P_k = 5$, and $M_k = 2$ as DPD model. Figure 5.8a shows the AM/AM conversion, Figure 5.8b the AM/PM conversion. As one can see, the distortions in both, the phase and the amplitude were considerably reduced. Whereas without DPD, the phase distortions spread over a range of 30 °, after introducing DPD only phase distortions up to 5 ° were noticeable. Moreover, memory effects were significantly reduced.

Figure 5.9 shows the PSD of $x_{\mathrm{PA}}[n]$ for a pre-distorter using the vector-switched static polynomials, the vector-switched memory polynomials with $M_k = 2$, the vector-switched dynamic deviation reduction-based Volterra series with $R_k = 2$ and $M_k = 2$, all with $P_k = 5$. Additionally, it shows the PSD of $x_{\mathrm{PA}}[n]$ for the second order dynamic deviation reduction-based Volterra series with $P = 5$ and $M = 2$ as DPD model, as it is the case in Section 5.3. It is illustrated in order to demonstrate the performance gain when 30 sub-models with the same configuration were employed.



(a) AM/AM conversion.

(b) AM/PM conversion.

Figure 5.8.: Measured AM/AM and AM/PM characteristics of the vector-switched dynamic deviation reduction-based Volterra series with $R_k = 2$, $P_k = 5$, and $M_k = 2$ as DPD model.

5.6 Discussion

Comparing the three analyzed models, it can observed that classical models (non-switched, non-decomposed) are impractical for the distinct behavior of the ETPA, especially they had troubles dealing with phase distortions (cf. Figure 5.4b), resulting in a poor ACEPR performance. Moreover, it is revealed that the dynamic deviation reduction based Volterra series are more suitable than the memory polynomials in all aspects. Generally speaking, the decomposed piecewise Volterra series and the vector-switched model showed comparable performance; they improved the ACEPR in average by about 10dB, compared to classical models with the same configuration as the sub-models. Moreover, comparing the best results of the reduced Volterra series, the decomposed piecewise Volterra series, and the vector-switched model, the latter two models improved the ACEPR by about 5dB and the NMSE by roughly 3dB.

Computational complexity is not considered in this thesis. The decomposed piecewise Volterra series and the vector-switched model in general feature a higher complexity. Whereas the complexity of the vector switched model is only increased by the switching function, the complexity of the decomposed piecewise Volterra series is $S + 1$ times the complexity of the sub-model; not mentioning the decomposition and recombination. On the other hand, the vector-switched model requires a lot more training samples in order to accurately estimate all sub-models.

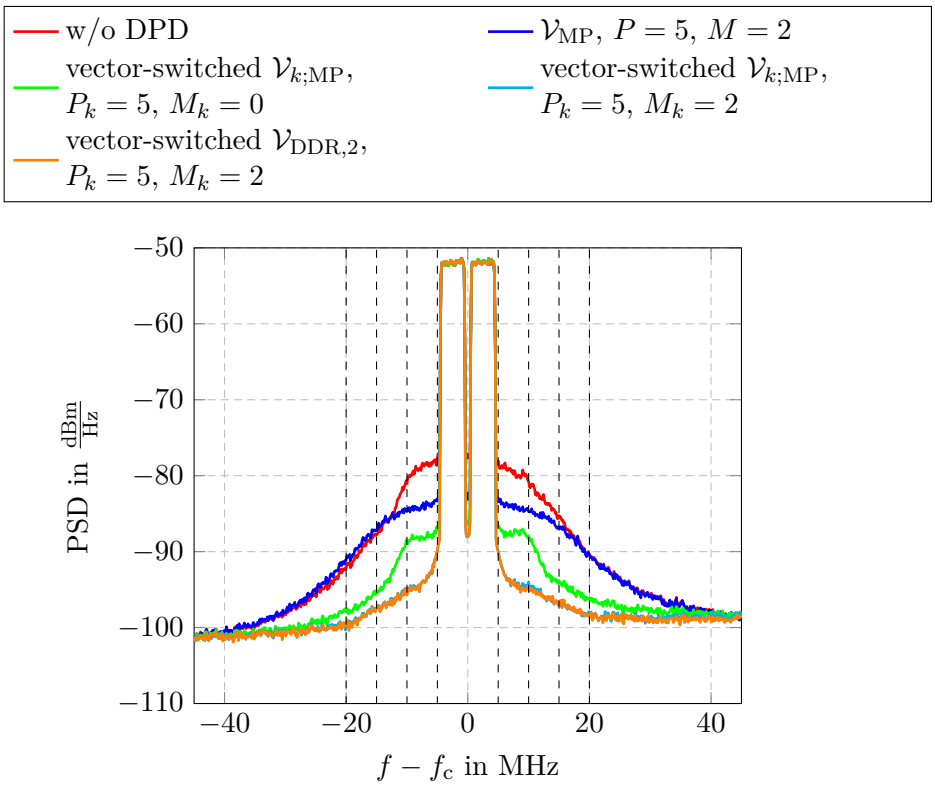


Figure 5.9.: Measured PSDs of $x_{\mathrm{PA}}[n]$ for the second order dynamic deviation reduction-based Volterra series with $P = 5$ and $M = 2$, the vector-switched static polynomials, the vector-switched memory polynomials, and the vector-switched dynamic deviation reduction-based Volterra series with $R_k = 2$, all with $P_k = 5$ and $M_k = 2$ as DPD models.

Conclusions and Outlook

This thesis considered digital pre-distortion (DPD) algorithms to linearize envelope tracking power amplifiers (ETPAs). In this regard, two in literature proposed highly developed DPD algorithms that are capable of handling the distinct behavior of the ETPA were discussed and compared by measurements.

The decomposed piecewise Volterra model decomposes the input signal into several sub-signals, each of which separately processed by a sub-model, and finally recombined in order to produce the pre-distorted signal.

The vector-switched model allocates each input signal sample to an appropriate sub-model that computes the pre-distorted signal.

As sub-models either memory polynomials, first or, second order deviation reduction-based Volterra series, respectively were utilized. To evaluate the performance, both models were implemented in MATLAB and tested for a Wideband Code Division Multiple Access (WCDMA)-like signal with 10MHz bandwidth and 7.5dB peak-to-average power ratio (PAPR). Compared to traditional models, experimental results show that both models outperformed classical polynomial models, even when using a small number of parameters. However, both models showed comparable performance when using the same sub-model. Moreover, the dynamic deviation reduction-based Volterra series were more accurate than the memory polynomials in every aspect.

Open questions for future research:

- Measurements with Long Term Evolution (LTE) signals that have higher bandwidth.
- Finding the optimum parameters for the decomposed piecewise Volterra series and the vector switched model.
- Comparing the computational complexity and investigate the feasibility of implementation in real-time systems.

A

Appendix A

Baseband Dynamic Deviation Reduced-Based Volterra Series Derivations

Contents

A.1 Introduction

In this Chapter, the passband representation for the *dynamic deviation reduction-based Volterra series* proposed by Zhu *et. al* in [28], i.e.,

$$\begin{aligned}\tilde{x}_{\text{out}}[n] \triangleq & \sum_{p=1}^{P} \tilde{h}_p[0,\ldots,0]\tilde{x}_{\text{in}}^p[n] \\ & + \sum_{p=1}^{P} \sum_{r=1}^{\min\{p,R\}} \tilde{x}_{\text{in}}^{p-r}[n] \sum_{m_1=1}^{M} \cdots \sum_{m_r=m_{r-1}}^{M} \tilde{h}_p[0,\ldots,0,m_1,\ldots,m_r] \prod_{j=1}^{r} \tilde{x}_{\text{in}}[n-m_j],\end{aligned} \tag{A.1}$$

is transformed to the equivalent baseband representation. The first $p-r$, $r=1,\ldots,R$ indices of the Volterra kernel $\tilde{h}_p[n_1,\ldots,n_p]$ are set to zero and the *dynamic deviation reduction order* R indicates the highest possible number of product terms of the delayed inputs.

A.2 Zero Order Derivations

Setting $R = 0$, means that the dynamic part, i.e., the second summation in (A.1) vanishes. Thus, the passband model reads

$$\tilde{x}_{\text{out}}[n] = \sum_{p=1}^{P} h_p[0,\ldots,0]\tilde{x}_{\text{in}}^p[n], \tag{A.2}$$

which can be easily transformed to the baseband (cf. Section 2.3) and the baseband *zero order deviation reduction-based Volterra series* reads

$$x_{\text{out}}[n] = \sum_{p=1}^{\frac{P+1}{2}} h_p[0,\ldots,0]x_{\text{in}}[n]|x_{\text{in}}[n]|^{2(p-1)}. \tag{A.3}$$

A.3 First Order Derivations

Starting with (A.1) for $R = 1$, the model output is

$$\begin{aligned}\tilde{x}_{\text{out}}[n] &= \sum_{p=1}^{P} \tilde{h}_p[0,\ldots,0]\tilde{x}_{\text{in}}^p[n] + \sum_{p=1}^{P} \sum_{m_1=1}^{M} \tilde{h}_p[0,\ldots,0,m_1]\tilde{x}_{\text{in}}^{p-1}[n]\tilde{x}_{\text{in}}[n-m_1] \\ &= \sum_{p=1}^{P} \sum_{m_1=0}^{M} \tilde{h}_p[0,\ldots,0,m_1]\tilde{x}_{\text{in}}^{p-1}[n]\tilde{x}_{\text{in}}[n-m_1],\end{aligned} \tag{A.4}$$

where $\tilde{x}_{\mathrm{in}}[n]$ and $\tilde{x}_{\mathrm{in}}[n-m_1]$ can be expressed using their corresponding equivalent baseband representation (cf. Section 2.2), which leads to

$$\begin{aligned}\tilde{x}_{\mathrm{out}}[n] = \sum_{p=1}^{P}\sum_{m_1=0}^{M}\frac{1}{\sqrt{2^p}}\tilde{h}_p[0,\ldots,0,m_1]\left(x_{\mathrm{in}}[n]\mathrm{e}^{\jmath 2\pi f_{\mathrm{c}} n T_{\mathrm{s}}} + x_{\mathrm{in}}^*[n]\mathrm{e}^{-\jmath 2\pi f_{\mathrm{c}} n T_{\mathrm{s}}}\right)^{p-1} \\ \times\left(x_{\mathrm{in}}[n-m_1]\mathrm{e}^{\jmath 2\pi f_{\mathrm{c}}(n-m_1)T_{\mathrm{s}}} + x_{\mathrm{in}}^*[n-m_1]\mathrm{e}^{-\jmath 2\pi f_{\mathrm{c}}(n-m_1)T_{\mathrm{s}}}\right).\end{aligned} \tag{A.5}$$

Using the binomial identity, the model output can be written as

$$\begin{aligned}\tilde{x}_{\mathrm{out}}[n] = \sum_{p=1}^{P}\sum_{m_1=0}^{M}\frac{1}{\sqrt{2^p}}\tilde{h}_p[0,\ldots,0,m_1] \\ \times\left[\sum_{k_1=0}^{p-1}\binom{p-1}{k_1}x_{\mathrm{in}}^{p-1-k_1}[n]\left(x_{\mathrm{in}}^*[n]\right)^{k_1}x_{\mathrm{in}}[n-m_1]\mathrm{e}^{\jmath 2\pi f_{\mathrm{c}} n T_{\mathrm{s}}(p-2k_1)}\mathrm{e}^{-\jmath 2\pi f_{\mathrm{c}} m_1 T_{\mathrm{s}}}\right. \\ \left.+\sum_{k_2=0}^{p-1}\binom{p-1}{k_2}x_{\mathrm{in}}^{p-1-k_2}[n]\left(x_{\mathrm{in}}^*[n]\right)^{k_2}x_{\mathrm{in}}^*[n-m_1]\mathrm{e}^{\jmath 2\pi f_{\mathrm{c}} n T_{\mathrm{s}}(p-2-2k_2)}\mathrm{e}^{\jmath 2\pi f_{\mathrm{c}} m_1 T_{\mathrm{s}}}\right].\end{aligned} \tag{A.6}$$

Obviously, components of the model output fall in the surroundings of f_{c} for *odd* p, $p-2k_1=1$, and $p-2-2k_2=1$ for the first and second summation, respectively. All other harmonic components are suppressed by the zonal filter (cf. Section 2.3). Therefore, we have $k_1=\frac{p-1}{2}>0$, $k_2=\frac{p-3}{2}>0$, and the model output reads

$$\begin{aligned}\tilde{x}_{\mathrm{out}}[n] = \sum_{p=1}^{P}\sum_{m_1=0}^{M}\frac{1}{\sqrt{2^p}}\tilde{h}_p[0,\ldots,0,m_1] \\ \times\left[\sum_{k_1=0}^{p-1}\binom{p-1}{k_1}x_{\mathrm{in}}^{\frac{p-1}{2}}[n]\left(x_{\mathrm{in}}^*[n]\right)^{\frac{p-1}{2}}x_{\mathrm{in}}[n-m_1]\mathrm{e}^{\jmath 2\pi f_{\mathrm{c}}(n-m_1)T_{\mathrm{s}}}\right. \\ \left.+\sum_{k_2=0}^{p-1}\binom{p-1}{k_2}x_{\mathrm{in}}^{\frac{p+1}{2}}[n]\left(x_{\mathrm{in}}^*[n]\right)^{\frac{p-3}{2}}x_{\mathrm{in}}^*[n-m_1]\mathrm{e}^{\jmath 2\pi f_{\mathrm{c}}(n+m_1)T_{\mathrm{s}}}\right].\end{aligned} \tag{A.7}$$

Which can be simplified and the model output of the baseband *first order deviation reduced-based Volterra series* can be written as

$$\begin{aligned}x_{\mathrm{out}}[n] = \sum_{p=1}^{\frac{P+1}{2}}\sum_{m_1=0}^{M}g_{2p-1,1}[m_1]|x_{\mathrm{in}}[n]|^{2(p-1)}x_{\mathrm{in}}[n-m_1] \\ +\sum_{p=2}^{\frac{P+1}{2}}\sum_{m_1=1}^{M}g_{2p-1,2}[m_1]x_{\mathrm{in}}^2[n]|x_{in}[n]|^{2(p-2)}x_{\mathrm{in}}^*[n-m_1],\end{aligned} \tag{A.8}$$

where $g_{p,i}[\cdot]$ denotes the transformed complex-valued Volterra kernel of the system.

A.4 Second Order Derivations

Starting with (A.1) for $R = 2$, the model output is

$$\begin{aligned}\tilde{x}_{\mathrm{out}}[n] = &\sum_{p=1}^{P} \tilde{h}_p[0,\ldots,0]\tilde{x}_{\mathrm{in}}^{p}[n] + \sum_{p=1}^{P} \tilde{x}_{\mathrm{in}}^{p-1}[n] \sum_{m_1=1}^{M} \tilde{h}_p[0,\ldots,0,m_1]\tilde{x}_{\mathrm{in}}[n-m_1] \\ &+ \sum_{p=2}^{P} \tilde{x}_{\mathrm{in}}^{p-2}[n] \sum_{m_1=1}^{M} \sum_{m_2=m_1}^{M} \tilde{h}_p[0,\ldots,0,m_1,m_2]\tilde{x}_{\mathrm{in}}[n-m_1]\tilde{x}_{\mathrm{in}}[n-m_2],\end{aligned} \tag{A.9}$$

where $\tilde{x}_{\mathrm{in}}[n]$, $\tilde{x}_{\mathrm{in}}[n-m_1]$, and $\tilde{x}_{\mathrm{in}}[n-m_2]$ can be represented by the respective equivalent baseband signal (cf. Section 2.2). Therefore, the model output reads

$$\begin{aligned}\tilde{x}_{\mathrm{out}}[n] = &\sum_{p=1}^{P} \frac{1}{\sqrt{2^p}} \tilde{h}_p[0,\ldots,0] \left(x_{\mathrm{in}}[n]\mathrm{e}^{\jmath 2\pi f_{\mathrm{c}} n T_{\mathrm{s}}} + x_{\mathrm{in}}^{*}[n]\mathrm{e}^{-\jmath 2\pi f_{\mathrm{c}} n T_{\mathrm{s}}}\right) \\ &+ \sum_{p=1}^{P} \left(x_{\mathrm{in}}[n]\mathrm{e}^{\jmath 2\pi f_{\mathrm{c}} n T_{\mathrm{s}}} + x_{\mathrm{in}}^{*}[n]\mathrm{e}^{-\jmath 2\pi f_{\mathrm{c}} n T_{\mathrm{s}}}\right)^{p-1} \\ &\times \sum_{m_1=1}^{M} \tilde{h}_p[0,\ldots,0,m_1] \left(x_{\mathrm{in}}[n-m_1]\mathrm{e}^{\jmath 2\pi f_{\mathrm{c}}(n-m_1) T_{\mathrm{s}}} + x_{\mathrm{in}}^{*}[n-m_1]\mathrm{e}^{-\jmath 2\pi f_{\mathrm{c}}(n-m_1) T_{\mathrm{s}}}\right) \\ &+ \sum_{p=2}^{P} \left(x_{\mathrm{in}}[n]\mathrm{e}^{\jmath 2\pi f_{\mathrm{c}} n T_{\mathrm{s}}} + x_{\mathrm{in}}^{*}[n]\mathrm{e}^{-\jmath 2\pi f_{\mathrm{c}} n T_{\mathrm{s}}}\right)^{p-2} \sum_{m_1=1}^{M} \sum_{m_2=m_1}^{M} \frac{1}{\sqrt{2^p}} \tilde{h}_p[0,\ldots,0,m_1,m_2] \\ &\times \left(x_{\mathrm{in}}[n-m_1]\mathrm{e}^{\jmath 2\pi f_{\mathrm{c}}(n-m_1) T_{\mathrm{s}}} + x_{\mathrm{in}}^{*}[n-m_1]\mathrm{e}^{-\jmath 2\pi f_{\mathrm{c}}(n-m_1) T_{\mathrm{s}}}\right) \\ &\times \left(x_{\mathrm{in}}[n-m_2]\mathrm{e}^{\jmath 2\pi f_{\mathrm{c}}(n-m_2) T_{\mathrm{s}}} + x_{\mathrm{in}}^{*}[n-m_2]\mathrm{e}^{-\jmath 2\pi f_{\mathrm{c}}(n-m_2) T_{\mathrm{s}}}\right).\end{aligned} \tag{A.10}$$

Using the binomial identity, the model output can be written as

$$
\begin{aligned}
\tilde{x}_{\mathrm{out}}[n] = &\sum_{p=1}^{P}\sum_{k_1=0}^{p}\binom{p}{k_1}\frac{1}{\sqrt{2^p}}\tilde{h}_p[0,\ldots,0]x_{\mathrm{in}}^{p-k_1}[n]\left(x_{\mathrm{in}}^{*}[n]\right)^{k_1}\mathrm{e}^{\jmath 2\pi f_{\mathrm{c}} n T_{\mathrm{s}}(p-2k_1)}\\
&+\sum_{p=1}^{P}\sum_{k_2=0}^{p-1}\sum_{m_1=1}^{M}\binom{p-1}{k_2}\frac{1}{\sqrt{2^p}}\tilde{h}_p[0,\ldots,0,m_1]x_{\mathrm{in}}^{p-1-k_2}[n]\left(x_{\mathrm{in}}^{*}[n]\right)^{k_2}\\
&\times x_{\mathrm{in}}[n-m_1]\mathrm{e}^{\jmath 2\pi f_{\mathrm{c}} n T_{\mathrm{s}}(p-2k_2)}\mathrm{e}^{-\jmath 2\pi f_{\mathrm{c}} m_1 T_{\mathrm{s}}}\\
&+\sum_{p=1}^{P}\sum_{k_3=0}^{p-1}\sum_{m_1=1}^{M}\binom{p-1}{k_3}\frac{1}{\sqrt{2^p}}\tilde{h}_p[0,\ldots,0,m_1]x_{\mathrm{in}}^{p-1-k_3}[n]\left(x_{\mathrm{in}}^{*}[n]\right)^{k_3}\\
&\times x_{\mathrm{in}}^{*}[n-m_1]\mathrm{e}^{\jmath 2\pi f_{\mathrm{c}} n T_{\mathrm{s}}(p-2-2k_3)}\mathrm{e}^{\jmath 2\pi f_{\mathrm{c}} m_1 T_{\mathrm{s}}}\\
&+\sum_{p=2}^{P}\sum_{k_4=0}^{p-2}\sum_{m_1=1}^{M}\sum_{m_2=m_1}^{M}\binom{p-2}{k_4}\frac{1}{\sqrt{2^p}}\tilde{h}_p[0,\ldots,0,m_1,m_2]x_{\mathrm{in}}^{p-2-k_4}[n]\left(x_{\mathrm{in}}^{*}[n]\right)^{k_4}\\
&\times x_{\mathrm{in}}[n-m_1]x_{\mathrm{in}}[n-m_2]\mathrm{e}^{\jmath 2\pi f_{\mathrm{c}} n T_{\mathrm{s}}(p-2k_4)}\mathrm{e}^{\jmath 2\pi f_{\mathrm{c}}(-m_1-m_2)T_{\mathrm{s}}}\\
&+\sum_{p=2}^{P}\sum_{k_5=0}^{p-2}\sum_{m_1=1}^{M}\sum_{m_2=m_1}^{M}\binom{p-2}{k_5}\frac{1}{\sqrt{2^p}}\tilde{h}_p[0,\ldots,0,m_1,m_2]x_{\mathrm{in}}^{p-2-k_5}[n]\left(x_{\mathrm{in}}^{*}[n]\right)^{k_5}\\
&\times x_{\mathrm{in}}[n-m_1]x_{\mathrm{in}}^{*}[n-m_2]\mathrm{e}^{\jmath 2\pi f_{\mathrm{c}} n T_{\mathrm{s}}(p-2-k_5)}\mathrm{e}^{\jmath 2\pi f_{\mathrm{c}}(-m_1+m_2)}\\
&+\sum_{p=2}^{P}\sum_{k_6=0}^{p-2}\sum_{m_1=1}^{M}\sum_{m_2=m_1}^{M}\binom{p-2}{k_6}\frac{1}{\sqrt{2^p}}\tilde{h}_p[0,\ldots,0,m_1,m_2]x_{\mathrm{in}}^{p-2-k_6}[n]\left(x_{\mathrm{in}}^{*}[n]\right)^{k_6}\\
&\times x_{\mathrm{in}}^{*}[n-m_1]x_{\mathrm{in}}[n-m_2]\mathrm{e}^{\jmath 2\pi f_{\mathrm{c}} n T_{\mathrm{s}}(p-2-k_6)}\mathrm{e}^{\jmath 2\pi f_{\mathrm{c}}(m_1-m_2)}\\
&+\sum_{p=2}^{P}\sum_{k_7=0}^{p-2}\sum_{m_1=1}^{M}\sum_{m_2=m_1}^{M}\binom{p-2}{k_7}\frac{1}{\sqrt{2^p}}\tilde{h}_p[0,\ldots,0,m_1,m_2]x_{\mathrm{in}}^{p-2-k_7}[n]\left(x_{\mathrm{in}}^{*}[n]\right)^{k_7}\\
&\times x_{\mathrm{in}}^{*}[n-m_1]x_{\mathrm{in}}^{*}[n-m_2]\mathrm{e}^{\jmath 2\pi f_{\mathrm{c}} n T_{\mathrm{s}}(p-4-2k_7)}\mathrm{e}^{\jmath 2\pi f_{\mathrm{c}}(m_1+m_2)T_{\mathrm{s}}}.
\end{aligned}
\tag{A.11}
$$

Obviously, components of the model output fall in the surroundings of f_{c} when *odd* p, $p-2k_1=1$, $p-2k_2=1$, $p-2-2k_3=1$, $p-2k_4=1$, $p-2-2k_5=1$, $p-2-2k_6=1$, and $p-4-2k_6=1$. All other harmonic components are suppressed by the zonal filter (cf. Section 2.3). Therefore, we have $k_1=\frac{p-1}{2}>0$, $k_2=\frac{p-1}{2}>0$, $k_3=\frac{p-3}{2}>0$, $k_4=\frac{p-1}{2}>0$, $k_5=\frac{p-3}{2}>0$, $k_6=\frac{p-3}{2}>0$, and $k_7=\frac{p-5}{2}>0$ for the respective

summation and the model output reads

$$
\begin{aligned}
\tilde{x}_{\mathrm{out}}[n] = & \sum_{p=1}^{P} \sum_{k_1=0}^{p} \binom{p}{k_1} \frac{1}{\sqrt{2^p}} \tilde{h}_p[0,\ldots,0] x_{\mathrm{in}}[n]^{\frac{p+1}{2}} \left(x_{\mathrm{in}}^*[n]\right)^{\frac{p-1}{2}} \mathrm{e}^{\jmath 2\pi f_{\mathrm{c}} n T_{\mathrm{s}}} \\
& + \sum_{p=1}^{P} \sum_{k_2=0}^{p-1} \sum_{m_1=1}^{M} \binom{p-1}{k_2} \frac{1}{\sqrt{2^p}} \tilde{h}_p[0,\ldots,0,m_1] x_{\mathrm{in}}^{\frac{p-1}{2}}[n] \left(x_{\mathrm{in}}^*[n]\right)^{\frac{p-1}{2}} \\
& \times x_{\mathrm{in}}[n-m_1] \mathrm{e}^{\jmath 2\pi f_{\mathrm{c}} (n-m_1) T_{\mathrm{s}}} \\
& + \sum_{p=1}^{P} \sum_{k_3=0}^{p-1} \sum_{m_1=1}^{M} \binom{p-1}{k_3} \frac{1}{\sqrt{2^p}} \tilde{h}_p[0,\ldots,0,m_1] x_{\mathrm{in}}^{\frac{p+1}{2}}[n] \left(x_{\mathrm{in}}^*[n]\right)^{\frac{p-3}{2}} \\
& \times x_{\mathrm{in}}^*[n-m_1] \mathrm{e}^{\jmath 2\pi f_{\mathrm{c}} (n+m_1) T_{\mathrm{s}}} \\
& + \sum_{p=2}^{P} \sum_{k_4=0}^{p-2} \sum_{m_1=1}^{M} \sum_{m_2=m_1}^{M} \binom{p-2}{k_4} \frac{1}{\sqrt{2^p}} \tilde{h}_p[0,\ldots,0,m_1,m_2] x_{\mathrm{in}}^{\frac{p-3}{2}}[n] \left(x_{\mathrm{in}}^*[n]\right)^{\frac{p-1}{2}} \\
& \times x_{\mathrm{in}}[n-m_1] x_{\mathrm{in}}[n-m_2] \mathrm{e}^{\jmath 2\pi f_{\mathrm{c}} (n-m_1-m_2) T_{\mathrm{s}}} \\
& + \sum_{p=2}^{P} \sum_{k_5=0}^{p-2} \sum_{m_1=1}^{M} \sum_{m_2=m_1}^{M} \binom{p-2}{k_5} \frac{1}{\sqrt{2^p}} \tilde{h}_p[0,\ldots,0,m_1,m_2] x_{\mathrm{in}}^{\frac{p-1}{2}}[n] \left(x_{\mathrm{in}}^*[n]\right)^{\frac{p-3}{2}} \\
& \times x_{\mathrm{in}}[n-m_1] x_{\mathrm{in}}^*[n-m_2] \mathrm{e}^{\jmath 2\pi f_{\mathrm{c}} (n-m_1+m_2) T_{\mathrm{s}}} \\
& + \sum_{p=2}^{P} \sum_{k_6=0}^{p-2} \sum_{m_1=1}^{M} \sum_{m_2=m_1}^{M} \binom{p-2}{k_6} \frac{1}{\sqrt{2^p}} \tilde{h}_p[0,\ldots,0,m_1,m_2] x_{\mathrm{in}}^{\frac{p-1}{2}}[n] \left(x_{\mathrm{in}}^*[n]\right)^{\frac{p-3}{2}} \\
& \times x_{\mathrm{in}}^*[n-m_1] x_{\mathrm{in}}[n-m_2] \mathrm{e}^{\jmath 2\pi f_{\mathrm{c}} (n+m_1-m_2) T_{\mathrm{s}}} \\
& + \sum_{p=2}^{P} \sum_{k_6=0}^{p-2} \sum_{m_1=1}^{M} \sum_{m_2=m_1}^{M} \binom{p-2}{k_6} \frac{1}{\sqrt{2^p}} \tilde{h}_p[0,\ldots,0,m_1,m_2] x_{\mathrm{in}}^{\frac{p+1}{2}}[n] \left(x_{\mathrm{in}}^*[n]\right)^{\frac{p-5}{2}} \\
& \times x_{\mathrm{in}}^*[n-m_1] x_{\mathrm{in}}^*[n-m_2] \mathrm{e}^{\jmath 2\pi f_{\mathrm{c}} (n+m_1+m_2) T_{\mathrm{s}}},
\end{aligned}
\tag{A.12}
$$

which can be simplified and symmetric terms, i.e., cases when $x_{\mathrm{in}}[n-m_1]x_{\mathrm{in}}^*[n-m_2]$ equals $x_{\mathrm{in}}^*[n-m_1]x_{\mathrm{in}}[n-m_2]$ can be considered. Therefore, the baseband representation

of (A.12) can be written as

$$
\begin{aligned}
x_{\text{out}}[n] = &\sum_{p=1}^{\frac{P+1}{2}} g_{2p-1,0} x_{\text{in}}[n]|x_{\text{in}}[n]|^{2(p-1)} \\
&+ \sum_{p=1}^{\frac{P+1}{2}} \sum_{m_1=1}^{M} g_{2p-1,1}[m_1]|x_{\text{in}}[n]|^{2(p-1)} x_{\text{in}}[n-m_1] \\
&+ \sum_{p=2}^{\frac{P+1}{2}} \sum_{m_1=1}^{M} g_{2p-1,2}[m_1] x_{\text{in}}^2[n]|x_{\text{in}}[n]|^{2(p-2)} x_{\text{in}}^*[n-m_1] \\
&+ \sum_{p=2}^{\frac{P+1}{2}} \sum_{m_1=1}^{M} \sum_{m_2=m_1}^{M} g_{2p-1,3}[m_1,m_2] x_{\text{in}}^*[n]|x_{\text{in}}[n]|^{2(p-2)} x_{\text{in}}[n-m_1] x_{\text{in}}[n-m_2] \\
&+ \sum_{p=2}^{\frac{P+1}{2}} \sum_{m_1=1}^{M} \sum_{m_2=m_1}^{M} g_{2p-1,4}[m_1,m_2] x_{\text{in}}[n]|x_{\text{in}}[n]|^{2(p-2)} x_{\text{in}}[n-m_1] x_{\text{in}}^*[n-m_2] \\
&+ \sum_{p=3}^{\frac{P+1}{2}} \sum_{m_1=1}^{M} \sum_{m_2=m_1}^{M} g_{2p-1,5}[m_1,m_2] x_{\text{in}}^3[n]|x_{\text{in}}[n]|^{2(p-3)} x_{\text{in}}^*[n-m_1] x_{\text{in}}^*[n-m_2],
\end{aligned}
\tag{A.13}
$$

where $g_{p,i}[\cdot]$ denotes the transformed complex-valued Volterra kernel of the system. Moreover, the first and second, and the third and fifth summation can be merged such that the model output of the baseband *second order deviation reduced-based Volterra series* can be written as

$$
\begin{aligned}
x_{\text{out}}[n] = &\sum_{p=1}^{\frac{P+1}{2}} \sum_{m_1=0}^{M} g_{2p-1,1}[m_1]|x_{\text{in}}[n]|^{2(p-1)} x_{\text{in}}[n-m_1] \\
&+ \sum_{p=2}^{\frac{P+1}{2}} \sum_{m_1=1}^{M} \sum_{m_2=m_1}^{M} g_{2p-1,3}[m_1,m_2] x_{\text{in}}^*[n]|x_{\text{in}}[n]|^{2(p-2)} x_{\text{in}}[n-m_1] x_{\text{in}}[n-m_2] \\
&+ \sum_{p=2}^{\frac{P+1}{2}} \sum_{m_1=0}^{M} \sum_{m_2=1}^{M} g_{2p-1,4}[m_1,m_2] x_{\text{in}}[n]|x_{\text{in}}[n]|^{2(p-2)} x_{\text{in}}[n-m_1] x_{\text{in}}^*[n-m_2] \\
&+ \sum_{p=3}^{\frac{P+1}{2}} \sum_{m_1=1}^{M} \sum_{m_2=m_1}^{M} g_{2p-1,5}[m_1,m_2] x_{\text{in}}^3[n]|x_{\text{in}}[n]|^{2(p-3)} x_{\text{in}}^*[n-m_1] x_{\text{in}}^*[n-m_2]
\end{aligned}
\tag{A.14}
$$

Nomenclature

The following notation is used throughout the thesis whenever possible. If not, it is explicitly stated.

a	scalar constant
$\mathbf{a}$	random variable with PDF $f_{\mathbf{a}}(a)$, mean $\mu_{\mathbf{a}}$, variance $\sigma^2_{\mathbf{a}}$, and auto-correlation function $r_{\mathbf{aa}}(\tau)$
$a(t)$	function with independent variable t
$a[n]$	sequence with independent variable n
$A(f)$	Fourier transform of $a(t)$, i.e., $A(f) = \mathcal{F}\{a(t)\}$
$A[\theta]$	discrete Fourier transform (DFT) of $a[n]$, i.e., $A[\theta] = \mathcal{F}_D\{a[n]\}$
$\underline{a}$	(column) vector
$\underline{\mathbf{a}}$	(column) random vector with joint-PDF $f_{\underline{\mathbf{a}}}(\underline{a})$, mean $\underline{\mu}_{\underline{\mathbf{a}}}$, variance $\underline{\sigma}^2_{\underline{\mathbf{a}}}$, and auto-correlation matrix $R_{\underline{\mathbf{aa}}} = \mathrm{E}\left\{\underline{\mathbf{aa}}^H\right\}$
$\tilde{a}(t)$	passband representation of $a(t)$
$\hat{a}$	estimate of a
A	constant matrix of defined dimension
$\mathbf{A}$	random matrix of defined dimension

*	complex conjugate
^{T}	transpose of a vector or matrix
^{H}	Hermitian (complex conjugate transpose) of a vector or matrix
$\mathrm{E}\{\cdot\}$	expectation operator
$\mathrm{Re}\{\cdot\}$	real part
$\mathrm{Im}\{\cdot\}$	imaginary part
$\arg\{\cdot\}$	angle of a complex-valued entity
$\mathcal{F}\{\cdot\}$	Fourier transform
$\mathcal{F}_D\{\cdot\}$	discrete Fourier transform (DFT)
$\mathcal{H}\{\cdot\}$	Hilbert transform
$\mathrm{tr}\{\cdot\}$	trace of a square matrix
$\mathcal{V}\{\cdot\}$	Volterra operator
$\mathcal{V}_{MP}\{\cdot\}$	memory polynomials operator

$\mathcal{V}_{HS}\{\cdot\}$	Hammerstein model operator
$\mathcal{V}_{DDR,R}\{\cdot\}$	dynamic deviation reduction-based Volterra series operator, with deviation reduction order R
$\lvert\cdot\rvert$	absolute value
$\lVert\cdot\rVert_p$	ℓ_p-vector norm
$\lceil\cdot\rceil$	ceiling function
∇	gradient

List of Abbreviations

3G	3rd Generation
ACEPR	adjacent channel error power ratio
ACPR	adjacent channel power ratio
ADC	analog to digital converter
AM/AM	amplitude modulation/amplitude modulation
AM/PM	amplitude modulation/phase modulation
BIBO	bounded-input bounded-output
BTS	base transceiver station
CCDF	complementary cumulative distribution function
DAC	digital to analog converter
DFT	discrete Fourier transform
DPD	digital pre-distortion
ETPA	envelope tracking power amplifier
FEC	forward error correction
FIR	finite impulse response
GMSK	Gaussian Minimum Shift Keying
GSM	Global System for Mobile Communications
IF	intermediate frequency
IP3	third order intercept point

LMS	least mean squares
LTE	Long Term Evolution
LTI	linear time-invariant
LS	least squares
LUT	look-up table
MSE	mean square error
ϵ-**NLMS**	ϵ-normalized least mean squares
NMSE	normalized mean square error
OFDMA	orthogonal frequency-division multiple access
PA	power amplifier
PAE	power added efficiency
PDF	probability density function
PSD	power spectral density
PAPR	peak-to-average power ratio
RF	radio frequency
VSA	vector signal analyser
VSG	vector signal generator
WCDMA	Wideband Code Division Multiple Access

Bibliography

[1] D. Raychaudhuri and N. B. Mandayam, "Frontiers of Wireless and Mobile Communications," *Proceedings of the IEEE*, vol. 100, no. 4, pp. 824–840, April 2012.

[2] A. Loureiro, D. Gallegos, and G. Caldwell, "Interference Analysis on UMTS-2100 Co-Existence with GSM-1900," in *Proceedings of the IEEE Vehicular Technology Conference (VTC Fall)*, September 2011.

[3] S. Kapp, "802.11a. More Bandwidth Without The Wires," *IEEE Internet Computing*, vol. 6, no. 4, pp. 75–79, July 2002.

[4] Y. Louët and A. O. Gouba, "Theoretical Analysis of the Trade-Off Between Efficiency and Linearity of the High Power Amplifier in OFDM Context," in *Proceedings of the 18th European Wireless Conference*, April 2012.

[5] P. M. Lavrador, T. R. Cunha, P. M. Cabral, and J. C. Pedro, "The Linearity-Efficiency Compromise," *IEEE Microwave Magazine*, vol. 11, no. 5, pp. 44–58, August 2010.

[6] S. C. Cripps, *RF Power Amplifiers for Wireless Communications*, 2nd ed. Artech House, Inc., 2006.

[7] P. J. Draxler, J. J. Yan, D. F. Kimball, and P. M. Asbeck, "Digital Predistortion for Envelope Tracking Power Amplifiers," in *Proceedings of the 13th Annual IEEE Wireless and Microwave Technology Conference (WAMICON)*, April 2012.

[8] Alcatel-Lucent and Vodafone Chair on Mobile Communication Systems, *Study on Energy Efficient Radio Access Network (EERAN) Technologies.* Project Report of TU Dresden, 2009.

[9] P. B. Kennigton, *High-Linearity RF Amplifier Design.* Artech House, Inc., 2000.

[10] J. H. K. Vuolevi, T. Rahkonen, and J. P. A. Manninen, "Measurement Technique for Characterizing Memory Effects in RF Power Amplifiers," *IEEE Transactions on Microwave Theory and Techniques*, vol. 49, no. 8, pp. 1383–1389, August 2001.

[11] Z. Wang, R. Ma, X. Yang, and S. Lanfranco, "Driver Amplifier Design Considerations for High Efficiency Envelope Tracking Power Amplifier Line-Up," in *Proceedings of the 13th Annual IEEE Wireless and Microwave Technology Conference (WAMICON)*, April 2012.

[12] M. Hassan, M. Kwak, V. W. Leung, C. Hsia, J. J. Yan, D. F. Kimball, L. E. Larson, and P. M. Asbeck, "High Efficiency Envelope Tracking Power Amplifier with Very Low Quiescent Power for 20 MHz LTE," in *Proceedings of the IEEE Radio Frequency Integrated Circuits Symposium (RFIC)*, June 2011.

[13] F. Wang, A. Hueiching, D. F. Kimball, L. E. Larson, and P. M. Asbeck, "Design of Wide-Bandwidth Envelope-Tracking Power Amplifiers for OFDM Applications," *IEEE Transactions on Microwave Theory and Techniques*, vol. 53, no. 4, pp. 1244–1255, April 2005.

[14] M. Hassan, L. E. Larson, V. W. Leung, and P. M. Asbeck, "Effect of Envelope Amplifier Nonlinearities on the Output Spectrum of Envelope Tracking Power Amplifiers," in *Proceedings of the IEEE 12th Topical Meeting on Silicon Monolithic Integrated Circuits in RF Systems (SiRF)*, Januar 2012.

[15] M. Isaksson, D. Wisell, and D. Rönnow, "A Comparative Analysis of Behavioral Models for RF Power Amplifiers," *IEEE Transactions on Microwave Theory and Techniques*, vol. 54, no. 1, pp. 348–359, Januar 2006.

[16] J. Wood, D. E. Root, and N. B. Tufillaro, "A Behavioral Modeling Approach to Nonlinear Model-Order Reduction for RF/Microwave ICs and Systems," *IEEE Transactions on Microwave Theory and Techniques*, vol. 52, no. 9, pp. 2274–2284, September 2004.

[17] F. M. Ghannouchi and O. Hammi, "Behavioral Modeling and Predistortion," *IEEE Microwave Magazine*, vol. 10, no. 7, pp. 52–64, December 2009.

[18] D. Schreurs, M. O'Droma, A. Goacher, and M. Gadringer, *RF Power Amplifier Behavioral Modeling*. Cambridge University Press, 2008.

[19] I. A. Glover and P. M. Grant, *Digital Communications*, 2nd ed. Prentice Hall, Inc., 2004.

[20] S. Benedetto and E. Biglieri, *Principles of Digital Transmission: With Wireless Applications*, 2nd ed. Kluwer Academic Publishers, 1999.

[21] R. Dallinger, "Pre-Distortion Algorithms for Power Amplifiers," Master's thesis, Vienna University of Technology, November 2007.

[22] L. Ding and G. T. Zhou, "Effects of Even-Order Nonlinear Terms on Power Amplifier Modeling and Predistortion Linearization," *IEEE Transactions on Vehicular Technology*, vol. 53, no. 1, pp. 156–162, January 2004.

[23] E. Aschbacher, "Digital Pre-Distortion of Microwave Power Amplifiers," Ph.D. dissertation, Vienna University of Technology, September 2005.

[24] M. Schetzen, *The Volterra & Wiener Theories of Nonlinear Systems.* John Wiley & Sons, Inc., 1980.

[25] A. Zhu and T. J. Brazil, "Behavioral Modeling of RF Power Amplifiers Based on Pruned Volterra Series," *IEEE Microwave and Wireless Components Letters*, vol. 14, no. 12, pp. 563–565, December 2004.

[26] H. Ku and J. S. Kenney, "Behavioral Modeling of Nonlinear RF Power Amplifiers Considering Memory Effects," *IEEE Transactions on Microwave Theory and Techniques*, vol. 51, no. 12, pp. 2495–2504, December 2003.

[27] T. Sadeghpour, H. Karkhaneh, R. Abd-Alhameed, A. Ghorbani, I. T. E. Elfergani, and Y. A. S. Dama, "Hammerstein Predistorter for High Power RF Amplifiers in OFDM Transmitters," in *Proceedings of the 30th URSI General Assembly and Scientific Symposium*, August 2011.

[28] 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*, vol. 54, no. 12, pp. 4323–4332, December 2006.

[29] A. Zhu, P. J. Draxler, C. Hsia, T. J. Brazil, D. F. Kimball, and P. M. Asbeck, "Digital Predistortion for Envelope-Tracking Power Amplifiers Using Decomposed Piecewise Volterra Series," *IEEE Transactions on Microwave Theory and Techniques*, vol. 56, no. 10, pp. 2237–2247, October 2008.

[30] S. Afsardoost, T. Eriksson, and C. Fager, "Digital Predistortion Using a Vector-Switched Model," *IEEE Transactions on Microwave Theory and Techniques*, vol. 60, no. 4, pp. 1166–1174, April 2012.

[31] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression.* Kluwer Academic Publishers, 1991.

[32] R. Xu and I. Wunsch, D., "Survey of Clustering Algorithms," *IEEE Transactions on Neural Networks*, vol. 16, no. 3, pp. 645–678, May 2005.

[33] S. Lloyd, "Least Squares Quantization in PCM," *IEEE Transactions on Information Theory*, vol. 28, no. 2, pp. 129–137, March 1982.

[34] P. Ladin, M. Isaksson, and P. Händel, "Comparison of Evaluation Criteria for Power Amplifier Behavioral Modeling," in *IEEE MTT-S International Microwave Symposium Digest*, June 2008.

[35] M. J. Reed and M. O. Hawksford, "Efficient Implementation of the Volterra Filter," *IEEE Proceedings on Vision, Image and Signal Processing*, vol. 147, no. 2, pp. 109–114, April 2000.

[36] T. K. Moon and W. C. Stirling, *Mathematical Methods and Algorithms for Signal Processing*. Prentice Hall, Inc., 2000.

[37] A. H. Sayed and M. Rupp, "Error-Energy Bounds for Adaptive Gradient Algorithms," *IEEE Transactions on Signal Processing*, vol. 44, no. 8, pp. 1982–1989, 1996.

[38] A. H. Sayed, *Adaptive Filters*. John Wiley & Sons, Inc., 2008.

[39] B. Widrow and M. E. Hoff, "Adaptive Switching Circuits," *IRE WESCON Convention Record*, vol. 4, no. 1, pp. 96–104, August 1960.

[40] R. Dallinger and M. Rupp, "A Strict Stability Limit for Adaptive Gradient Type Algorithms," in *Conference Record of the Forty-Third Asilomar Conference on Signals, Systems and Computers*, November 2009.

[41] H. Paaso and A. Mammela, "Comparison of Direct Learning and Indirect Learning Predistortion Architectures," in *Proceedings of the IEEE International Symposium on Wireless Communication Systems (ISWCS)*, October 2008.

[42] M. Y. Cheong, E. Aschbacher, P. Brunmayr, M. Rupp, and T. Laakso, "Comparison and Experimental Verification of Two Low-Complexity Digital Predistortion Methods," in *Conference Record of the Thirty-Ninth Asilomar Conference on Signals, Systems and Computer*, November 2005.

[43] E. Changsoo and E. J. Powers, "A New Volterra Predistorter Based on the Indirect Learning Architecture," *IEEE Transactions on Signal Processing*, vol. 45, no. 1, pp. 223–227, January 1997.

[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*, vol. 56, no. 7, pp. 1524–1534, July 2008.