



Receiver-side nonlinearities mitigation using an extended iterative decision-based technique[☆]

Fernando H. Gregorio^{a,*}, Stefan Werner^b, Juan Cousseau^a, Jose Figueroa^a, Risto Wichman^b

^a CONICET-Department of Electrical and Computer Engineering, Universidad Nacional del Sur, Av. Alem 1253, Bahía Blanca 8000, Argentina

^b Aalto University School of Electrical Engineering, P.O. Box 13000, 00076 Aalto, Finland

ARTICLE INFO

Article history:

Received 12 July 2010

Received in revised form

10 December 2010

Accepted 16 March 2011

Available online 23 March 2011

Keywords:

Power amplifier

Nonlinear distortion

Mitigation

OFDM

ABSTRACT

This work presents an iterative receiver cancellation technique for mitigating the inband distortion introduced by a nonlinear wideband transmitter power amplifier (PA). The proposed decision-based technique employs a Wiener–Hammerstein model that accounts for the nonlinear transfer function and memory of the PA as well as for the wireless propagation channel. As such, the mitigation technique can be seen as a generalization of existing iterative decision-based techniques assuming memoryless PA nonlinearities. For successful distortion mitigation, the iterative technique requires an estimate of the nonlinear model that characterizes the PA. We propose to perform this model identification at the receiver, embedded in an iterative decision-based scheme, avoiding the nonideal analog-to-digital feedback loop associated with transmitter-based model identification. A stochastic algorithm is proposed for the model identification providing the necessary PA model parameters required for symbol detection. In addition, we analyze the convergence properties of the proposed technique. Simulation results confirm that the proposed mitigation technique provides distortion cancellation at almost the same level to the case of perfect knowledge of the PA model. These results enable us to employ power amplifiers with more relaxed linearity requirement, moving the operation point to a region with improved power efficiency while reducing the system overall degradation.

© 2011 Elsevier B.V. All rights reserved.

1. Introduction

Transmitter power amplifiers (PAs) should bring the desired signal to a suitable level with a limited power consumption and high power efficiency. The use of nonlinear PAs at base stations (BS) and mobile terminals (MT) can solve the efficiency problem, however, at the expense

of signal distortion in the form of signal compression and clipping. The resulting waveform distortion lead to undesired effects in communication systems by introducing adjacent channel interference and increased bit-error rate (BER). Considering OFDM-based systems, these effects are more pronounced due to the high peak-to-average-power-ratio (PAPR). Furthermore, the problems associated with nonlinear PAs are more severe if we need to account for, in addition to the nonlinear transfer function, PA memory effects. PA memory effects arise in wideband (e.g., IEEE 802.11 g, WiMAX) and high-power systems. The PA memory effects are more pronounced at the BS side, where the dissipated power changes with the signal level, creating temperature fluctuations in transistors and other components [1,2]. On the other side, in mobile terminals (MT); low-power PAs are required, and memory effects (appearing in the form of non-flat frequency

[☆] This work was partially funded by the Universidad Nacional del Sur, PGI 24/K044, the ANPCyT, PICT-2008-0182 and PICT-2008-00104, the CONICET, PIP 112-200801-01024, the Center of Excellence in Smart Radios and Wireless Research (SMARAD), the Academy of Finland, and by the Finnish Funding Agency for Technology and Innovation.

* Corresponding author.

E-mail addresses: fernando.gregorio@uns.edu.ar (F.H. Gregorio), stefan.werner@aalto.fi (S. Werner), jcousseau@uns.edu.ar (J. Cousseau), figueroa@uns.edu.ar (J. Figueroa), risto.wichman@aalto.fi (R. Wichman).

response) are due to electrical memory effects, originated from the bias and coupling impedance networks [3–5].

Power consumption is a key factor in both uplink and downlink transmission. Operating the power amplifier in the linear region requires a large back-off to prevent nonlinear distortion. In other words, the PA must operate in a power inefficient region. Considering a portable unit, a high power consumption reduces battery life time and, consequently, the system portability. Base stations with high power consumptions, on the other hand, is a concern when network costs should be minimized. By increasing the transmitter complexity, predistortion and peak-to-average power reduction techniques can be exploited to improve power efficiency and to reduce the nonlinear distortion effects. However, a system with predistortion is still subject to nonlinear effects because, even in the ideal case, the predistorter-PA cascade is a soft limiter. In other words, there is a clear trade-off between the allowed nonlinear distortion level, system complexity and power efficiency.

Nonlinear distortion can be compensated either at the transmitter side or the receiver side. For the former case, the signal to be transmitted is modified before the PA, and among the well-known methods for that purpose are: predistortion and PAPR reduction techniques. In uplink transmission, the implementation of transmitter-side compensation techniques at portable units is constrained by the limited processing capacity available where, due to cost and size reasons, additional hardware and signal processing required cannot be afforded. This problem motivates the implementation of receiver side compensation methods that can be justified at base stations where higher computational complexity, power consumption, and implementation cost are allowed while the portable unit is kept simple and power efficient.

The use of a receiver-based mitigation technique is a viable option to maintain transmitter complexity low, e.g., by removing residual interference in uplink when simple (memoryless) or no predistortion is used in the transmitter. It may also be used in combination with more complex predistortion techniques (in downlink) [8]. An early approach to receiver side compensation of nonlinearities appears in [6] where a traveling wave tube (TWT) amplifier is used in single-carrier QAM systems. The combination of equalization and a nonlinear distortion compensation for fixed wireless channels have been presented in [7]. A power amplifier nonlinearity cancellation (PANC) technique was proposed in [8,9] for removing the distortion originated at the transmitter due to a memoryless PA, at the expense of some increase in the receiver complexity. With an initial estimate of the transmitted OFDM symbols, the distortion effects can be estimated if the PA model is known. After that estimation, the nonlinear distortion can be removed from the received signal and new and improved symbol estimates can be obtained. This procedure can be repeated in an iterative manner to obtain almost undistorted estimates in two or three iterations. This concept was applied in [10] to a single-user wireline system using adaptive OFDM with a large number of carriers. More recently, novel receiver side compensation techniques have been presented in [11,12] for MIMO OFDM systems employing

memoryless PAs. In [13] a compensation technique for single-carrier frequency-division multiple access (SC-FDMA) is proposed to remove the interference of adjacent users. This implementation assumes that the receiver has perfect knowledge of transmitter power amplifier nonlinearities of the different users which are considered memoryless. To the best of our knowledge, earlier literature dealing with receiver-based (iterative) mitigation techniques is restricted to the case of memoryless PA. Considering the case of wideband PAs with memory, the above techniques [8–11] do not present adequate performance. This paper deals with more complicated memory structures, which will not only complicate the mitigation technique but also the associated model estimation.

Previous works with PANC techniques have assumed that the PA model is known, or that its parameters are estimated at the transmitter and sent to the receiver during the initialization. In this work we consider the more practical problem of estimating and tracking the PA model parameters at the receiver. As a consequence, we chose a composite Wiener–Hammerstein (WH) model to describe the corresponding dynamics of the parameter estimation problem at hand. The WH model consists of a static nonlinearity with a linear filter preceding its input and another linear filter at its output. The WH model has been found useful for describing many practical nonlinear systems, see, e.g., the exhaustive bibliography on nonlinear system identification [14]. Early approaches to WH model identification, see, e.g., [15–17], were based on correlation analysis. In these methods, the identification of the linear systems and the static nonlinearity is decoupled, requiring training sequences that restrict its application to a particular system. Other approaches employing stochastic gradient (SG) algorithms have been proposed [18,19]. When linear (FIR) filter is used, a conventional SG algorithm will identify a scaled version of the convolution of the input and output linear filters of the WH system [18]. This result was later used in a three-step identification scheme where the static nonlinearity was modeled using Hermite polynomials [19].

In [20,21] techniques for identifying the parameters of an AM/AM PA model, for an OFDM-based communication system, have been proposed. These techniques are constrained to the transmitter side and assume time-invariant linear filters. That is a very restrictive scenario. A recursive algorithm, where the linear and nonlinear blocks of the WH model are identified simultaneously, is presented in [22]. However, this approach has as drawback an estimation scheme with slow parameter convergence.

The concept of frequency-domain and time-domain-based WH model identification has been explored to obtain good initial conditions for elaborated identification algorithms [23–25]. Despite the use of similar conceptual information, these methods cannot be directly applied to the problem considered in this work, because they are not able to track time-varying wireless channels.

This paper generalizes the ideas in [8–10] and proposes a PANC technique that not only takes into account the nonlinear transfer function of the PA, but also its memory effects. For that purpose, an algorithm to

estimate the parameters of the nonlinear PA model embedded in PANC scheme is presented. Looking for algorithm simplicity and at the same time good performance, we propose a combined time- and frequency-domain identification approach that is especially suitable for implementation in OFDM systems, providing a good trade-off between algorithm convergence and computational burden. Our simplified estimation strategy consists of two steps that provide the necessary PA model parameters required for symbol detection with the general decision-directed technique. A stochastic algorithm is proposed for the model identification, and its convergence properties are studied.

In summary, the contributions of the paper are:

- The extension of the PANC technique in [10] to include a more general model for nonlinear wideband PA effects. The technique can serve as an alternative or in combination with a transmitter-based mitigation technique, e.g., a predistorter. Compared to a polynomial-based (Volterra) equalizer, the proposed PANC technique attains a lower bit-error rate and does not suffer from an error floor due to clipping noise.
- A new frequency- and time-domain strategy for WH model identification that is compatible with the PANC technique. The new approach allows us to develop a simple stochastic gradient algorithm for the estimation of the necessary PA model parameters. Compared to existing Wiener–Hammerstein algorithms [19,22], our identification approach features fast convergence, and low implementation complexity.
- The convergence properties of the proposed stochastic algorithm are studied.

The work is organized as follows. Section 2 introduces the OFDM system and PANC technique using a basic PA static model to put in evidence the decision-directed criterion concepts used. The new PANC technique for the extended wideband PA model, in addition to a two-step procedure to estimate the related parameters, is presented in Section 3. In that section we introduce a simplified time- and frequency-domain WH identification structure, which allows the parameter estimation to be carried out at the receiver. A simple stochastic gradient algorithm is proposed and tailored for the unknown variables of the PANC technique. We also provide, in Section 4, an analysis of the local convergence properties of this algorithm. In Section 5, the performance of the proposed adaptive WH identification scheme and its combination with PANC is validated via simulations. Comparison with other mitigation and identification techniques is also done in this section. Conclusions are given in Section 6.

2. System model and problem formulation

The OFDM transmission model used in this paper is illustrated in Fig. 1. The system under consideration has N subcarriers. Let $\{X(k)\}_{k=0}^{N-1} \in \mathbb{C}$ denote the modulated data symbols associated with carrier k . The transmitted time-domain OFDM symbols $\{x(n)\}_{n=0}^{N-1}$ are then obtained via the inverse discrete Fourier transform (IDFT)

$$x(n) = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} X(k) e^{j2\pi k n/N}, \quad n = 0, 1, \dots, N-1 \quad (1)$$

After IDFT, a cyclic prefix is added to the block $[x(0)x(1) \dots x(N-1)]$ to avoid intersymbol interference (ISI) and simplify equalization. The OFDM symbols $x(n)$ are then passed through a nonlinear PA with memory whose output is described by

$$z(n) = f[x(n), x(n-1), \dots, x(n-L_f)] \quad (2)$$

where $f(\cdot)$ is a nonlinear mapping of the current and L_f past input symbols.

Under the assumption that the channel impulse response $c(n)$ remains constant over at least one OFDM block (block fading model) and can be modeled as an FIR filter of order L_c , the received signal can be written as

$$y(n) = \sum_{m=0}^{L_c} c_m z(n-m) + v(n) \quad (3)$$

where $v(n)$ is additive noise, assumed Gaussian circular symmetric with variance σ_v^2 . Finally, the frequency-domain version of (3) is obtained via the discrete Fourier transform (DFT), i.e.,

$$Y(k) = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} y(n) e^{-j2\pi n k/N}, \quad k = 0, 1, \dots, N-1 \quad (4)$$

In what follows we will consider an *iterative (decision-based) approach* to recover the transmitted symbols $X(k)$ from the received signal $Y(k)$ in (4). Our power amplifier nonlinearity cancellation (PANC) technique can be seen as an extension of the one in [10], which assumes a memoryless PA (i.e., $z(n) = f[x(n)]$), to the more general case when the PA has memory, see (2). In our derivation of PANC, we need to first choose a proper model for $f(\cdot)$ that not only enables iterative processing but also accurately describes the system in a compact form (small number of parameters). We shall see that for successful operation, PANC requires knowledge of the assumed model $f(\cdot)$ and, hence, we also develop a two-step estimation technique for obtaining the unknown parameters at the receiver.

Remark 1. A conceptually simpler approach would be to directly, from transmit–receive signals, design a nonlinear

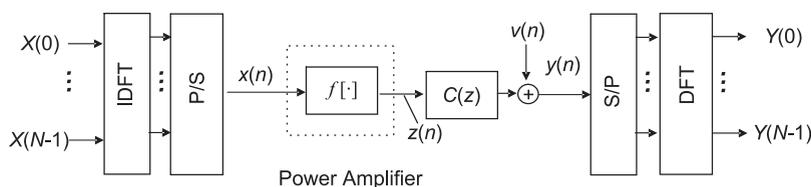


Fig. 1. Basic OFDM system model including a nonlinear PA with memory.

equalizer that is linear in the unknown parameters, e.g., based on memory polynomials or generalized polynomials. There are many advantages of using an iterative approach over a simple Volterra-based equalizer, e.g., a significant reduction in parameters (i.e., significant complexity reduction), convergence speed, robustness, and performance. More importantly, a simple polynomial-based (Volterra) equalizer is not able to mitigate clipping noise (from saturated signals) and, therefore, requires a large back-off to avoid problems associated with the equalization of a saturating (non-invertible) function. This is confirmed by our simulations provided in Section 5.

3. Power amplifier nonlinearity cancellation technique for general PA models

This section presents an iterative approach for removing the nonlinear PA effects at the receiver. It is an extension of the iterative ML-based technique in [10] to the case of nonlinear PAs with memory and time-varying transmission channels. For proper operation, the PANC technique, as presented in the following, requires knowledge of an assumed PA model. Receiver-based estimation of the PA model parameters are considered in Section 4.

In order to derive the PANC technique we model the nonlinear PA with memory using a Wiener–Hammerstein structure [1]. The Wiener–Hammerstein model is frequently used to model broadband PAs and is formed by a cascade of a linear filter $A(z)$, a nonlinear static function $g[\cdot]$, and another linear filter $B(z)$. The filters, $A(z)$ and $B(z)$, are here modeled as FIR filters of orders L_a and L_b , respectively. The static nonlinearity is here modeled with a polynomial (see Remark 3 below). The transmitted signal $z(n)$ can now be expressed as

$$z(n) = \sum_{m=0}^{L_b} b_m g[s(n-m)] \quad (5)$$

where

$$s(n) = \sum_{m=0}^{L_a} a_m x(n-m) \quad (6)$$

In order to detect the data symbols $X(k)$ and express the effects of the nonlinear distortion, it is common and useful to represent the output of the nonlinear static block $g[\cdot]$ as a sum of two uncorrelated components [26]

$$g[s(n)] = K_L s(n) + d(n) \quad (7)$$

The first term in (7) is just a scaled version of the input signal ($K_L \leq 1$), while $d(n)$ is an additive distortion term. This model is valid assuming that the input signal is a stationary Gaussian process, which is a reasonable assumption for an OFDM signal with a large number of subcarriers [26]. Substituting (7) in (5) and assuming the effective channel length $L=L_a+L_b+L_c$ falls within the cyclic prefix, (4) reduces to

$$Y(k) = C(k)B(k)[K_L A(k)X(k) + D(k)] + V(k), \quad k = 0, \dots, N-1 \quad (8)$$

where for subcarrier k , $V(k)$ is the additive noise, $D(k)$ is the nonlinear distortion (DFT of $\{d(n)\}_{n=0}^{N-1}$), and $A(k)$, $B(k)$ and $C(k)$ denote the responses of the linear filters of the PA model and the wireless channel, respectively.

In order to detect the uncoded $X(k)$ from the received signal in (8), we can consider the following Maximum Likelihood (ML) decision rule

$$\hat{X}(k) = \arg \min_{X(k)} \left\{ H_L(k) \left[X(k) - \frac{Y(k)}{H_L(k)} + \frac{D(k)}{A(k)K_L} \right]^2 \right\} \quad k = 0, 1, \dots, N-1 \quad (9)$$

where

$$H_L(k) = K_L A(k)B(k)C(k) \quad (10)$$

Assuming that $H_L(k)$, $A(k)$, and $D(k)$ are all known at the receiver, the ML estimate is given by the symbol $X(k)$ with the minimum distance to $Y(k)/H_L(k) - D(k)/(A(k)K_L)$. Knowing the static nonlinear function $g[\cdot]$ allows us to estimate $D(k)$ through (7) as

$$d(n) = g[s(n)] - K_L s(n), \quad n = 0, \dots, N-1$$

$$D(k) = \frac{1}{N} \sum_{n=0}^{N-1} d(n) e^{-j2\pi n k/N}, \quad k = 0, 1, \dots, N-1 \quad (11)$$

We see from (6) that $s(n)$ depends on both $\{X(k)\}_{k=0}^{N-1}$ and $A(z)$. Following the approach in [10], we can use an iterative technique that employs tentative decisions $\hat{X}(k)$ for mitigating the nonlinear distortion effects. The technique, depicted in Fig. 2, is summarized in Table 1. Our simulation study, which is presented in Section 5, suggests that two or three iterations (I_{\max} in Table 1) are usually sufficient. The PANC technique requires, in addition to the tentative decisions $\hat{X}(k)$, the estimates $\hat{H}_L(k)$, $\hat{A}(k)$ and $\hat{g}[\cdot]$, see Section 4.

Remark 2. Parameter K_L depends in general on the back-off and clipping levels used with the PA. For most practical cases $K_L \cong 1$, as it was assumed in Table 1. The specific value of K_L is in fact embedded in the parametrization of $g[\cdot]$ which is discussed below.

Remark 3. The static nonlinearity $g[\cdot]$ can be characterized by any nonlinear static model, as for example a polynomial model or even a piecewise linear model. For illustration purposes we choose $g[\cdot]$ to be modeled as a P th order polynomial with coefficients $\{g_{2k+1}\}_{k=0}^K$, where, to fit the nonlinear PA characteristics, the number of polynomial coefficients is given by $2K+1=P$ if P is even, or $2K+1=P-1$ if P is odd. We may consider also conventional or orthogonal (Legendre-based) polynomial basis functions. In case of conventional polynomials, $g[s(n)]$ is modeled as [27]

$$g[s(n)] = \sum_{k=0}^K g_{2k+1} \phi_{2k+1}[s(n)] \quad (12)$$

$$\phi_{2k+1}[s(n)] = s(n)|s(n)|^{2k} \quad (13)$$

Using Legendre-based polynomials, $g[s(n)]$ is modeled as [28]

$$g[s(n)] = \sum_{k=0}^K g_{2k+1} \psi_{2k+1}[s(n)] \quad (14)$$

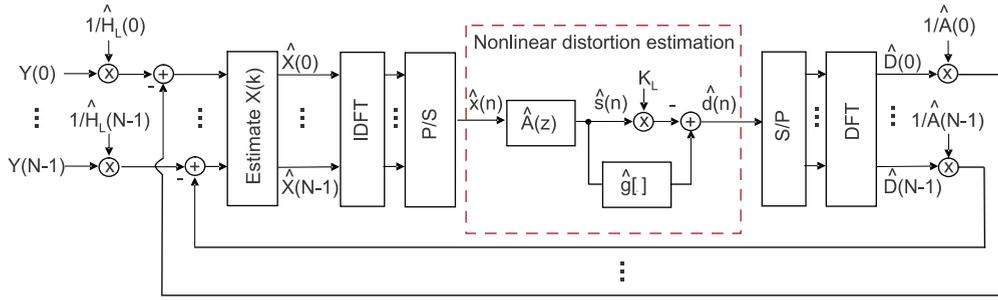


Fig. 2. Power amplifier nonlinearity cancellation (PANC) with the proposed PA distortion model.

Table 1

Power amplifier nonlinearity cancellation (PANC).

PANC
Initialization:
$\hat{D}^{(0)}(k) = 0, k = 0, \dots, N-1$
for $m=1$ to I_{\max}
{
Symbol decoding:
$\hat{X}^{(m)}(k) = \operatorname{argmin}_{X(k)} \left\{ \left H_L(k) \left[X(k) - \frac{Y(k)}{H_L(k)} + \frac{\hat{D}^{(m-1)}(k)}{A(k)} \right] \right ^2 \right\}$
Time domain:
$\hat{x}^{(m)}(n) = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} \hat{X}^{(m)}(k) e^{j2\pi k n/N}$
Estimate $\hat{s}(n)$:
$\hat{s}^{(m)}(n) = \sum_{l=0}^{L_a} a_l \hat{x}^{(m)}(n-l)$
Estimate distortion $\hat{d}(n)$:
$\hat{d}^{(m)}(n) = g[\hat{s}^{(m)}(n)] - K_L \hat{s}^{(m)}(n)$
Distortion in frequency domain $\hat{D}(k)$
$\hat{D}^{(m)}(k) = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} \hat{d}^{(m)}(n) e^{-j2\pi k n/N}$
}

$$\psi_{2k+1}[s(n)] = \sum_{l=0}^k (-1)^{k-l} \frac{\sqrt{k+1}}{(l+1)!} \binom{k}{l} \phi_{2l+1}[s(n)] \quad (15)$$

Orthogonal (Legendre-based) basis is particularly useful with OFDM signals, which are approximately complex Gaussian distributed [28]. There are many other advantages of employing orthogonal polynomials, e.g., improved numerical stability and robustness to quantization noise and finite-precision errors [29]. It will also facilitate the analysis of the adaptive PA model estimation algorithm that is carried out in Section 4.3.

Remark 4. We note that the Wiener model (linear filter followed by a static nonlinearity) is frequently used in literature to model the PA. The Wiener–Hammerstein model adopted in this work has different modelling capabilities than the Wiener model, hence, applies to a wider range of problems. Furthermore, even though we model the PA using a Wiener model, the end-to-end system would still be a Wiener–Hammerstein system due to the presence of the wireless channel. As a consequence, the new estimation technique, presented in Section 4, cover both Wiener and Wiener–Hammerstein PA models. Extensive studies available allows to consider these kind of models general enough to include PA

memory effects in addition to behavioral nonlinearities [30,20,21].

4. Estimation of power amplifier model at the receiver

This section proposes a two-step procedure for obtaining estimates of $H_L(k)$, $A(k)$, and $g[\cdot]$ to be used with the proposed PANC in Section 3. The first step employs training symbols with low PAPR that allow us to estimate $H_L(k)$ without the influence of the nonlinear static block $g[\cdot]$. The obtained estimate of $H_L(k)$ is then used in a second step where the static nonlinearity $g[\cdot]$ and the linear block $A(k)$ are identified using another set of training symbols.

4.1. Step 1: Estimation of $H_L(k)$

It is important that training symbols are not affected by $g[\cdot]$ in order to obtain a reliable estimate of $H_L(k)$. If only a group $T < N$ carriers are active while the remaining $N - T$ subcarriers carry no data, the variance of the OFDM signal can be reduced (low PAPR) so that the PA mostly operates in the linear region.

Just to illustrate the general methodology, we employ the standard interpolation technique in [31], which is based on a comb-type pilot arrangement where a set of T dedicated pilot carriers are reserved for pilot data. The channel frequency response on the pilot subcarriers without nonlinear distortion $D(k)=0$ (i.e., low PAPR symbols) is obtained as

$$\hat{H}_L(k) = \frac{Y(k)}{X(k)}, \quad \forall k \in \mathcal{T} \quad (16)$$

where $\mathcal{T} = \{k_1, \dots, k_T\}$ denotes the index set specifying the T pilot carriers with $\{k_i\}_{i=1}^T$ taken from the set $\{1, \dots, N\}$. The whole frequency-domain channel response is obtained through interpolation using truncated DFT matrices. The obtained channel estimate $\hat{H}_L(k)$ is re-estimated periodically.

We note that, as an alternative to the interpolation-based technique presented above, low PAPR techniques may be applied for pilot symbol design. That allows (16) to be used with all N subcarriers, e.g., the use of Frank–Zadoff–Chu sequences [11]. Our simulation results, presented in Section 5, confirm that the number of pilot carriers is not critical. In case of higher mobility, the channel may need to be tracked on an OFDM block-by-block basis. In such situation, the low PAPR necessary for

the tracking of $H_L(k)$ cannot be guaranteed, and we need to consider alternative estimation strategies, see, e.g., [8].

4.2. Step 2: Estimation of $g(\cdot)$ and $A(z)$

The estimate $\hat{H}_L(k)$ is used in PANC to yield the time and frequency-domain signals

$$\hat{Y}(k) = \frac{Y(k)}{\hat{H}_L(k)} \quad (17)$$

where $\hat{y}(n) = (1/N) \sum_{k=0}^{N-1} \hat{Y}(k)e^{j2\pi kn/N}$. In practice, the zero-forcing (ZF) equalizer in (17) may be substituted by a minimum-mean-squared-error (MMSE) equalizer in order to reduce possible noise enhancement.

In the absence of noise and assuming a perfect channel estimate $\hat{H}_L(k) = H_L(k)$, the two sequences $x(n)$ and $\tilde{y}(n)$ are related via the cascaded system $A(z) : g[\cdot] : A^{-1}(z)$ (this equivalent system is obtained by cascading $B(z)C(z)$ with $H_L^{-1}(z)$). This observation enables us to adaptively estimate $A(z)$ and $g[\cdot]$ using the structure in the right scheme of Fig. 3. The proposed algorithm estimates the intermediate signal $u(n) = g[s(n)]$ by forming the following error

$$e(n) = \hat{u}_1(n) - \hat{u}_2(n) \quad (18)$$

where, following the scheme of Fig. 3,

$$\hat{u}_1(n) = \sum_{m=0}^{L_a} \hat{a}_m^*(n) \tilde{y}(n-m) \quad (19)$$

$$\hat{u}_2(n) = \sum_{k=0}^K \hat{g}_{2k+1}^*(n) \phi_{2k+1}[\hat{s}(n)] \quad (20)$$

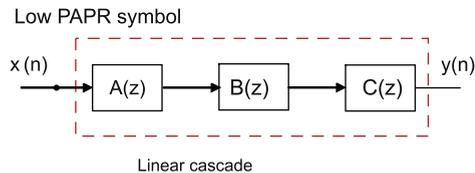
$$\hat{s}(n) = \sum_{m=0}^{L_a} \hat{a}_m(n) x(n-m) \quad (21)$$

To avoid ambiguity in the filter gain, $\hat{a}_0(n) \equiv \hat{a}_0$ is anchored to a fixed value [17,18]. The error to be minimized can now be written as

$$e(n) = \hat{u}_1(n) - \hat{u}_2(n) = \mathbf{a}_0^* \tilde{y}(n) - \boldsymbol{\theta}^H(n) \boldsymbol{\varphi}(n) \quad (22)$$

where $\boldsymbol{\theta}(n) \in \mathbb{C}^{(K+L_a+1) \times 1}$ and $\boldsymbol{\varphi}(n) \in \mathbb{C}^{(K+L_a+1) \times 1}$ are given by

$$\boldsymbol{\theta}(n) = [\hat{\mathbf{g}}^T(n) \ \hat{\mathbf{a}}^T(n)]^T \quad \boldsymbol{\varphi}(n) = \{-\phi[s(n)]^T \ \tilde{\mathbf{y}}^T(n)\}^T \quad (23)$$



with

$$\hat{\mathbf{g}}(n) = [\hat{g}_1(n) \ \hat{g}_3(n) \ \dots \ \hat{g}_{2K+1}(n)]^T$$

$$\hat{\mathbf{a}}(n) = [\hat{a}_1(n) \ \hat{a}_2(n) \ \dots \ \hat{a}_{L_a}(n)]^T$$

$$\boldsymbol{\phi}[\hat{s}(n)] = \{\phi_1[\hat{s}(n)] \ \phi_3[\hat{s}(n)] \ \dots \ \phi_{2K+1}[\hat{s}(n)]\}^T$$

$$\tilde{\mathbf{y}} = [\tilde{y}(n-1) \ \tilde{y}(n-2) \ \dots \ \tilde{y}(n-L_a)]^T \quad (24)$$

Using the instantaneous squared error $|e(n)|^2$ as an objective function, a stochastic gradient algorithm that updates $\boldsymbol{\theta}(n)$ is given by

$$\boldsymbol{\theta}(n+1) = \boldsymbol{\theta}(n) - \mu e^*(n) \nabla_{\boldsymbol{\theta}}[e(n)] \quad (25)$$

where μ is a step size controlling the convergence speed and algorithm stability. The gradient in (25) is given by

$$\nabla_{\boldsymbol{\theta}}[e(n)] = \begin{bmatrix} -\nabla_{\hat{\mathbf{g}}}[\hat{u}_2(n)] \\ \nabla_{\hat{\mathbf{a}}}[\hat{u}_1(n)] - \nabla_{\hat{\mathbf{a}}}[\hat{u}_2(n)] \end{bmatrix} \quad (26)$$

where

$$\nabla_{\hat{\mathbf{g}}}[\hat{u}_2(n)] = \boldsymbol{\phi}[\hat{s}(n)] \quad (27)$$

$$\nabla_{\hat{\mathbf{a}}}[\hat{u}_1(n)] = \tilde{\mathbf{y}}(n) \quad (28)$$

$$\nabla_{\hat{\mathbf{a}}}[\hat{u}_2(n)] = - \left(\hat{g}_1(n) + \sum_{k=0}^{K-1} (2k+3) \hat{g}_{2k+3}(n) \phi_{2k+1}[\hat{s}(n)] \right) \mathbf{x}(n) \quad (29)$$

and $\mathbf{x}(n) = [x(n-1) \ \dots \ x(n-L_1)]^T$. The verification of (29) is given in Appendix A.

The approach to define the error in (22), is similar to that used in [32] for the identification of a nonlinear Wiener-type system. Since the error, defined in (22), is linear in the parameters, a simple stochastic gradient algorithm (25) is proposed to obtain the estimates. Despite of that, other updating strategies can also be used with different trade-off in terms of computational complexity and convergence speed.

4.3. Local convergence analysis

In order to study the local convergence of the estimation algorithm proposed in Section 4, we first introduce the basic relationship between conventional

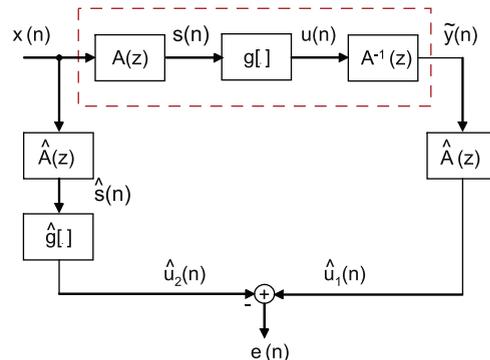


Fig. 3. Model identification providing estimation of the linear cascade $A(z):B(z):C(z)$ with a low PAPR symbol (left figure) and $A(z)$ and $g[\cdot]$ (right figure).

and orthogonal polynomial bases. Thereafter, we analyze the basic conditions for convergence of the proposed stochastic gradient algorithm.

4.3.1. Conventional and orthogonal polynomial basis

Let the vectors ϕ_p and ψ_p both of dimensions $P \times 1$ contain the respective P basis functions of the conventional and orthogonal polynomials in (13) and (15), i.e.,

$$\phi_p[z(n)] = \{\phi_1[z(n)], \phi_3[z(n)], \dots, \phi_{2K+1}[z(n)]\}^T \quad (30)$$

$$\psi_p[z(n)] = \{\psi_1[z(n)], \psi_3[z(n)], \dots, \psi_{2K+1}[z(n)]\}^T \quad (31)$$

Then from (13) and (15) we can relate both bases as

$$\psi_p[z(n)] = \Lambda_p \phi_p[z(n)] \quad (32)$$

where $\Lambda_p = [\lambda_{lk}]$ is a $P \times P$ lower triangular matrix whose non-zero elements are given by

$$\lambda_{lk} = (-1)^{k-l} \frac{\sqrt{k+1}}{(l+1)!} \binom{k}{l} \quad (33)$$

It is straightforward to show that for a unit-variance Gaussian $z(n)$ we have [28]

$$E\{\psi_p \psi_p^H\} = I \quad (34)$$

The one-to-one mapping in (32) and the unitary property in (34) will be exploited in the convergence study presented below.

4.3.2. Convergence study

Local convergence analysis of the proposed algorithm can be performed using the ordinary differential equation (ODE) method [33–35]. The general ODE approach comes from the field of stochastic approximation theory, and enables us to convert the study of convergence of a stochastic nonlinear equation into the study of the stability of the solutions of a deterministic differential equation.

The analysis is based on the following assumptions:

- (A1) Input signal $x(n)$ is a zero-mean unit-variance circular complex Gaussian-distributed process.
- (A2) The true nonlinear PA is described by a Wiener-Hammerstein model except for a measurement noise $v(n)$

Assumption A1 is not very restrictive, since the probability density function (pdf) of an OFDM symbol may be well approximated by a Gaussian pdf for a large number of subcarriers N . One might argue that choosing $x(n)$ to be uniformly distributed is better from the perspective of convergence speed $g[\cdot]$ [36]. However, for the application at hand, the training signals will be transmitted over a channel, causing interference on neighboring channels. Fig. 4 illustrates the power spectral density for an OFDM system when $x(n)$ is approximately uniformly and Gaussian distributed. We see that Gaussian training signal is preferable from an interference point-of-view. Uniformly distributed signals seem more appropriate when applying the identification algorithm for predistorter design.

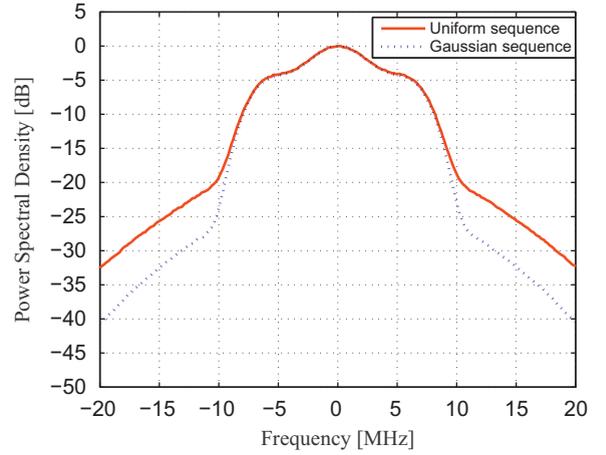


Fig. 4. Out-of-band distortion of an OFDM system with a nonlinear PA described by a Saleh model [37] with memory effects (see Section 5 for PA model details). The system is driven with Gaussian (dotted line) and Uniform (solid line) training sequences.

The ODE associated with the stochastic algorithm in (25) is given by

$$\frac{\partial \theta_s}{\partial t} = -E\{\nabla_{\theta}[e(n)]e^*(n)\} \quad (35)$$

where $e(n)$ is defined in (22) and $\nabla_{\theta}[e(n)]$ is given in (26). Under assumption A2, it can be verified that a stationary point $\bar{\theta}_s$ of the proposed algorithm corresponds to the solution of $E\{\nabla_{\theta}[e(n)]e^*(n)\} = 0$. Following the approach in [34], we now consider a linearization of (35) in a neighborhood of the stationary point, i.e.,

$$\begin{aligned} \frac{\partial \theta_s}{\partial t} &\cong -\left. \frac{\partial E\{\nabla_{\theta}[e(n)]e^*(n)\}}{\partial \theta_s} \right|_{\bar{\theta}_s} (\theta_s - \bar{\theta}_s) \\ &= -E\left\{ \frac{\partial \nabla_{\theta}[e(n)]}{\partial \theta_s} e^*(n) \right\} \Big|_{\bar{\theta}_s} (\theta_s - \bar{\theta}_s) \\ &\quad - E\left\{ \nabla_{\theta}[e(n)] \frac{\partial e^*(n)}{\partial \theta_s} \right\} \Big|_{\bar{\theta}_s} (\theta_s - \bar{\theta}_s) \\ &= -E\left\{ \nabla_{\theta}[e(n)] \frac{\partial e^*(n)}{\partial \theta_s} \right\} \Big|_{\bar{\theta}_s} (\theta_s - \bar{\theta}_s) = \mathbf{P}_s(\bar{\theta}_s)(\theta_s - \bar{\theta}_s) \end{aligned} \quad (36)$$

where the first term in (36) is zero due to the orthogonality principle and

$$\mathbf{P}_s(\bar{\theta}_s) = -E\{\nabla_{\theta}[e(n)]\nabla_{\theta}^H[e(n)]\} \Big|_{\bar{\theta}_s} \quad (37)$$

Thus, local convergence is guaranteed if the eigenvalues of $\mathbf{P}_s(\bar{\theta}_s)$ are negative or, equivalently, that $-\mathbf{P}_s(\bar{\theta}_s)$ is positive definite. Details of this verification can be found in Appendix B.

In the convergence analysis we assumed $A_1(z) = \sum_{n=0}^{N_1} a_1(n)z^{-n}$ to be an unbiased estimate of $A^{-1}(z)$. As mentioned above, the true $A^{-1}(z)$ in Fig. 3 depends both on how $H_L(z)$ is obtained and the type of equalizer used to produce $\hat{Y}(k)$ in (17). For example, using an MMSE equalizer will reduce noise enhancement at low SNR at the expense of bias. Although local convergence can be assured under general conditions, our simulations indicate that bias in this initial step leads to a biased

estimation of the second-step parameters $A(z)$ and $g[\cdot]$. On the other hand, when employing the obtained estimates with the proposed extended PANC technique, the final performance will be similar for ZF and MMSE equalizers. The reason is that PANC is a decision-based technique, which requires medium-to-high SNR. In that case the MMSE estimates approaches the ZF estimates.

5. Simulations

The performance of the proposed extended iterative decision-based technique is evaluated in an OFDM system with 16-QAM modulation on $N=512$ subcarriers. The carrier frequency is $f_0=5$ GHz and the bandwidth is $B=20$ MHz. The length of the cyclic prefix is equal to 64. The channel is Rayleigh fading with independent propagation paths, each generated according to a Jake's Doppler spectrum. The power loss and delay profiles of the channel are: $[0, -1, -3, -9]$ dB and $[0, 1, 2, 3]$ μ s corresponding to an urban scenario. The terminal speed is set to 2 km/h.

The unknown wideband PA that is to be estimated is modeled as a WH system, where the linear filters (IIR) are given by [30]

$$A(z) = \frac{1 + 0.1z^{-2}}{1 - 0.1z^{-1}} \quad \text{and} \quad B(z) = \frac{1 - 0.1z^{-1}}{1 - 0.2z^{-1}} \quad (38)$$

and the static nonlinearity is implemented as a solid state power amplifier (SSPA) represented by a Saleh model [37], i.e.,

$$g[x(n)] = \frac{|x(n)|}{\left[1 + \left(\frac{|x(n)|}{A_s}\right)^{2p}\right]^{1/p}} \exp[j\angle x(n)] \quad (39)$$

where the parameter p adjusts the smoothness of the transition from the linear region to the saturation region, and A_s is the amplifier input saturation point. The results are evaluated for different clipping levels γ defined as

$$\gamma = \frac{A_s}{\sqrt{E_n\{|x(n)|^2\}}} \quad (40)$$

where $\sqrt{E_n\{|x(n)|^2\}}$ is the RMS value of the OFDM signal. The algorithm in Section 4.2 used $L_a + 1 = 5$ taps in the FIR filter $\hat{A}(z)$ and the static nonlinearity $\hat{g}[\cdot]$ is tested using both conventional and orthogonal polynomial models of order $P=4$.

5.1. Estimation of $g[\cdot]$ and $A(z)$

The algorithm of Section 4 was used to estimate the prefilter $A(z)$ and static nonlinearity $g[\cdot]$. Before the identification algorithm can start, the received signal needs to be equalized by $\hat{H}_L(k)$, which is estimated as in (16). Results are shown for both ZF and MMSE equalizers. Initially, we used $T=32$ pilot carriers for estimating $H_L(k)$.

The MSE, $E[|e(n)|^2]$ with $e(n)$ defined in (22), is measured as a function of the SNR. The results are obtained by averaging the squared error magnitude $|e(n)|^2$ over three consecutive OFDM blocks after convergence of the algorithm. Fig. 5 shows the results for orthogonal and conventional polynomial models using ZF and MMSE

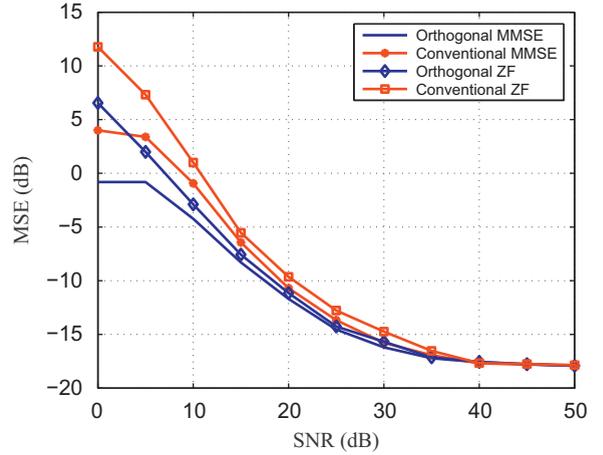


Fig. 5. MSE versus SNR for model identification employing conventional and orthogonal polynomials. The results are shown for the case when ZF and MMSE equalizers are used prior to the model identification. Wiener-Hammerstein type PA was considered where $g[\cdot]$ is given by (39) with $p=2$ and $\gamma = 4$ dB.

equalizers. The MSE values are obtained after the convergence of the algorithm. As expected, the MMSE equalizer outperforms the ZF equalizer at low SNR values. The results also confirm that model identification with orthogonal polynomials performs better than with conventional polynomials.

On the other hand, PANC techniques operate in medium-to-high SNRs due to the use of tentative decisions. Therefore, the performance of PANC will be similar for both equalizers or either polynomial type. This is also confirmed via simulations in the next section. Fig. 6 shows the relative parameter error $E(n)$ for vectors $\mathbf{a}(n)$ and $\mathbf{g}(n)$, employed to model the linear filter and the static nonlinearity, for SNR=15 dB, SNR=30 dB, and without channel noise, where for the linear part we have defined

$$E_a(n) = \frac{\|\mathbf{a}(n) - \mathbf{a}_\infty(\infty)\|}{\|\mathbf{a}_\infty(\infty)\|} \quad (41)$$

with \mathbf{a}_∞ being the converged vector in the absence of noise and with perfect knowledge of $H_L(k)$ and a similar expression $E_g(n)$ for the vector $\mathbf{g}(n)$. The error floor in the relative parameter error curves is set by the channel noise. For the noise-free case, undermodelling effects, i.e., insufficient memory and/or polynomial order, and model mismatch give the error floor in the relative parameter error curves.

Fig. 7 illustrates how the accuracy of the initial estimate $H_L(z)$ affects the estimation of $A(z)$ and $g[\cdot]$ by varying the number of pilot carriers T used in the estimation. We see that the proposed model identification technique is robust with respect to the initial channel estimation step. Increasing the number of pilot subcarriers leads to a small improvement in the identification performance. Thus, the number of pilot symbols is not critical to the identification as long as the number of pilots is larger than the length of the impulse response of the linear cascade.

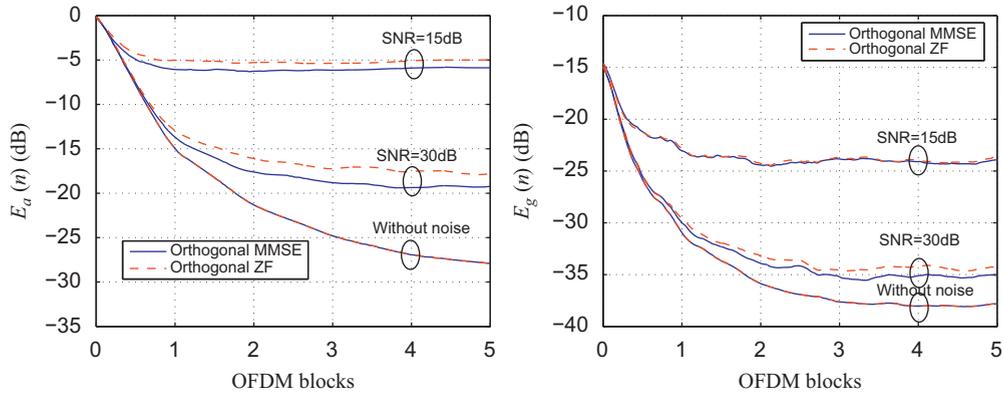


Fig. 6. Relative parameter error $E_a(n)$ and $E_g(n)$ as a function of time. Results are shown for orthogonal polynomial, employing ZF and MMSE equalization prior to the model identification. SNR is set to 15 dB, 30 dB, and without noise.

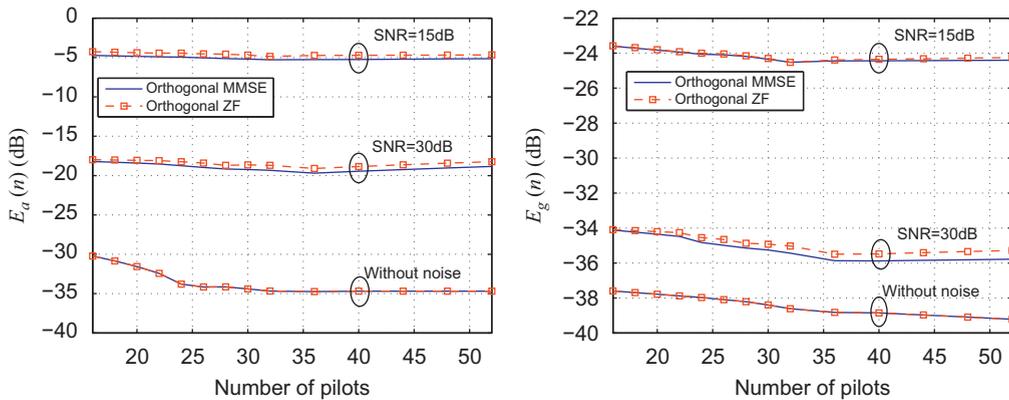


Fig. 7. Relative parameter error $E_a(n)$ and $E_g(n)$ versus number of pilots for the adaptive model identification employing orthogonal polynomials.

5.2. Comparison with existing Wiener–Hammerstein identification techniques

Considering the system in Fig. 1(b), we may apply standard Wiener–Hammerstein techniques where $A(z)$, $g(\cdot)$, and $B(z)C(z)$ are identified from the input–output signals $\{x(n), y(n)\}$. After forming $H_L(z)$, the identified blocks can be used with PANC in Table 1. In the following the proposed two-step estimation algorithm in Section 4 is compared with two state-of-the-art algorithms [19,22]. The first algorithm [22] is a single-step recursive method based on the output estimation error, where the linear and nonlinear blocks are identified simultaneously. The second algorithm [19] identifies the blocks in three steps: (1) identification of $g(\cdot)$, (2) identification of a scaled version of the cascade of the linear filters $A(z)B(z)C(z)$, and, (3) sequential identification of the individual filter taps of $A(z)$ and $B(z)C(z)$. For a more detailed description of the algorithms, see [19,22].

The algorithms were applied to the identification of two different Wiener–Hammerstein models. The first model is described by (38) and (39) (Model 1), and the

second, defined in [19] (Model 2), has linear blocks given by

$$A(z) = 0.9184 + 0.3674z^{-1} + 0.1469z^{-2} \tag{42}$$

$$B(z) = 0.9184 + 0.5510z^{-1} + 0.3306z^{-2} \tag{43}$$

and a static nonlinearity taken as an Hermite polynomial with coefficients [0,1.5,0.3,0.6]. Comparing the proposed algorithm with the one in [22], Figs. 8 and 9 show the learning curves, i.e., $E[|e(n)|^2]$ with $e(n)$ defined in (22) and convergence of the tap estimates of linear filter $A(z)$. We see that the proposed estimation algorithm achieves faster convergence to a lower steady-state estimation error for the first PA model. Therefore, the proposed approach may employ a shorter OFDM training sequence for the PA model identification. For the second PA model, our method shows lower convergence speed, four OFDM symbols are required to reach the steady-state, than the one in [22]. However, the method proposed [22] gives a large steady-state estimation error.

Fig. 10 illustrates the convergence behavior of the three-step method in [19]. Steps 1 and 2, which can be

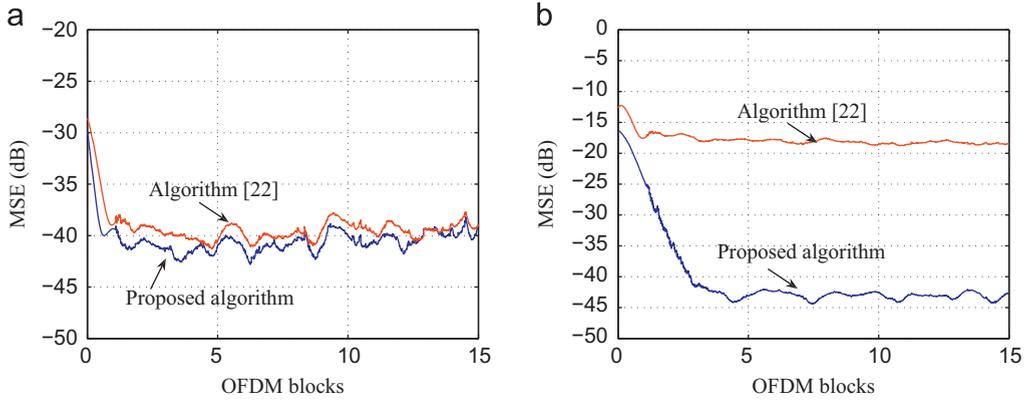


Fig. 8. MSE versus number of OFDM blocks: estimation algorithm in Section 4 and [22]. (a) PA model 1, (b) PA model 2.

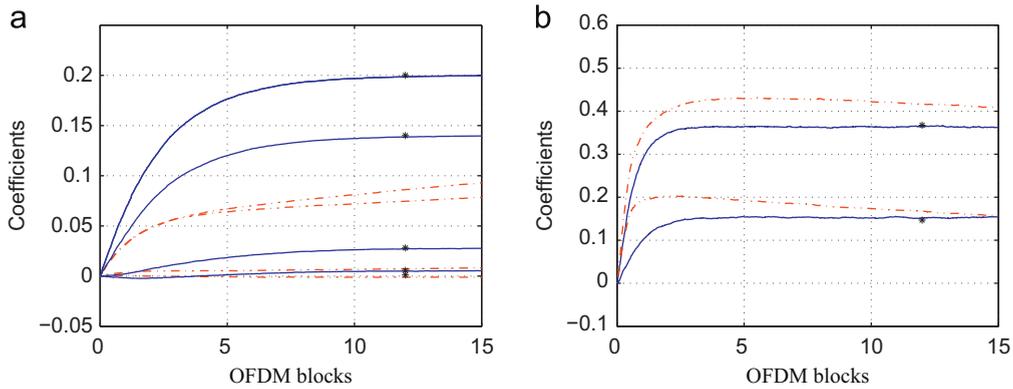


Fig. 9. Learning curves for the linear filter $A(z)$ identification: estimation algorithm in Section 4 (solid line) and [22] (dotted line). To avoid ambiguity in the filter gain, $\hat{a}_0(n) \equiv \hat{a}_0$ is anchored to a fixed value [17,18]. (a) PA model 1, (b) PA model 2. * marks refer to the optimum coefficients.

executed simultaneously, require approximately eight OFDM blocks to reach steady-state. After that, a third step is executed to obtain the coefficients of the individual filter (here $A(z)$ and $B(z)C(z)$) via three intermediate variables W_0, W_1, W_2 using a number of algebraic operations (see [19], Section 4-E). This last step requires more than 20 OFDM blocks to reach convergence as seen from Fig. 10. Therefore, the method in [19] is not feasible for the application at hand due to its slow convergence. Moreover, the set of algebraic operations that follows the third step, the conversion from the intermediate variables to $A(z)$ and $B(z)C(z)$, will increase the overall implementation complexity.

5.3. Performance of PANC with receiver-based model identification

In order to evaluate the extended PANC technique employing model estimation, we consider an initial training phase for carrying out the model identification: one OFDM block for estimating $H_L(k)$ and three OFDM blocks for estimating $A(z)$ and $g[\cdot]$ (as was done above). After the initial training phase, the estimated PA model is employed with PANC to mitigate the nonlinear distortion. The wireless channel variations are tracked through the

re-estimation of $H_L(k)$ once in every seven OFDM blocks resulting in a signaling overhead of 12.5%. In case of higher mobility, alternative channel estimation techniques can be employed that track $H_L(k)$ for each OFDM block using dedicated pilot symbols [8,38].

Fig. 11(a) and (b) show the BER versus SNR for two different clipping levels, $\gamma = 3$ and 4 dB. We see that the PANC technique employing receiver-based model estimation performs almost identically to a solution having perfect knowledge of the PA model. Curves for a linear PA with channel estimation and channel state information (CSI), nonlinear PA without PANC, and PANC with perfect PA model estimation are included for reference. The channel estimation technique used with the linear PA was the same as the one used for estimating $H_L(z)$. We see that the performance with MMSE and ZF equalizers (for removing the effect of $H_L(k)$) are similar and reach BER close to that of a linear PA. BER results are plotted for the case when $H_L(z)$ is estimated using $T=16, 32$, and 64 pilot subcarriers. We see that the BER results for $T=64$ pilot subcarriers is only slightly better than for $T=16$ and 32 pilot subcarriers.

Simulation results (not included) confirm that BER results using approximately Gaussian training signals (i.e., normal OFDM symbols) of four, five and six OFDM blocks and uniformly distributed training sequence of

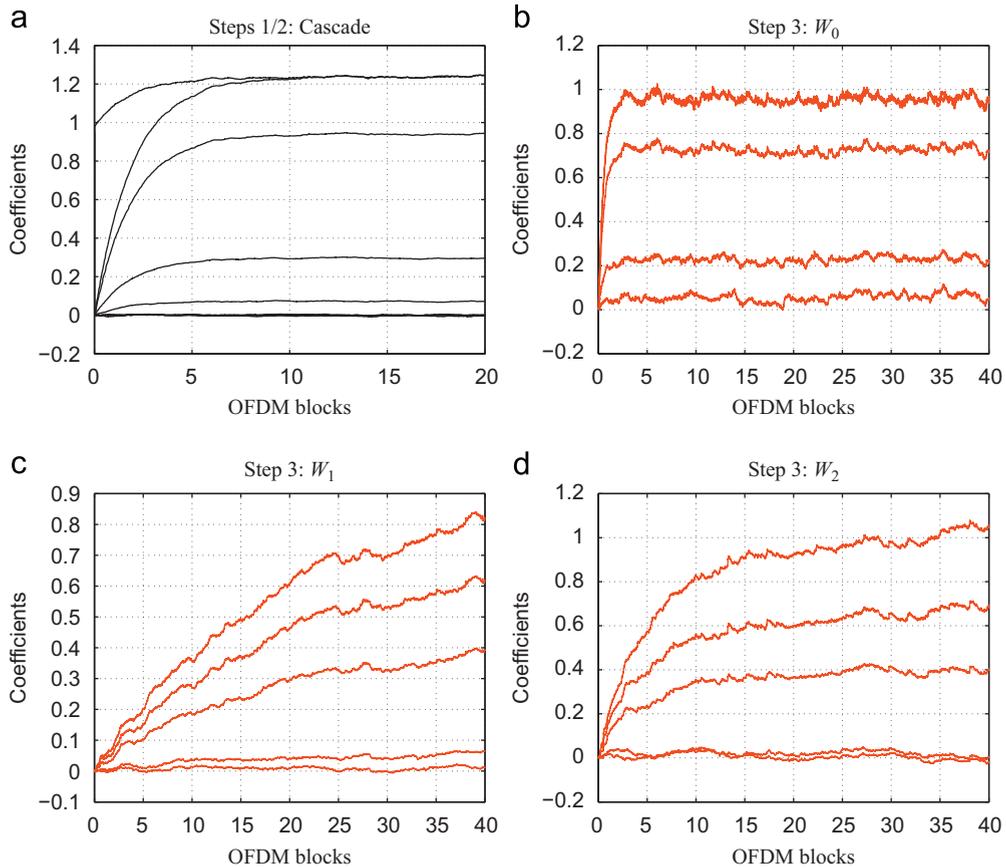


Fig. 10. Three-step method in [19]. (a) Learning curves for the linear cascade and NL block estimation executed in the steps 1 and 2 of the algorithm, (b) Learning curve for intermediate variable W_0 executed in the step 3, (c) Learning curve for intermediate variable W_1 executed in the step 3, (d) Learning curve for intermediate variable W_2 executed in the step 3. The coefficients of the linear filters, $A(z)$ and $B(z)C(z)$, are obtained using algebraic operations from the intermediate variables.

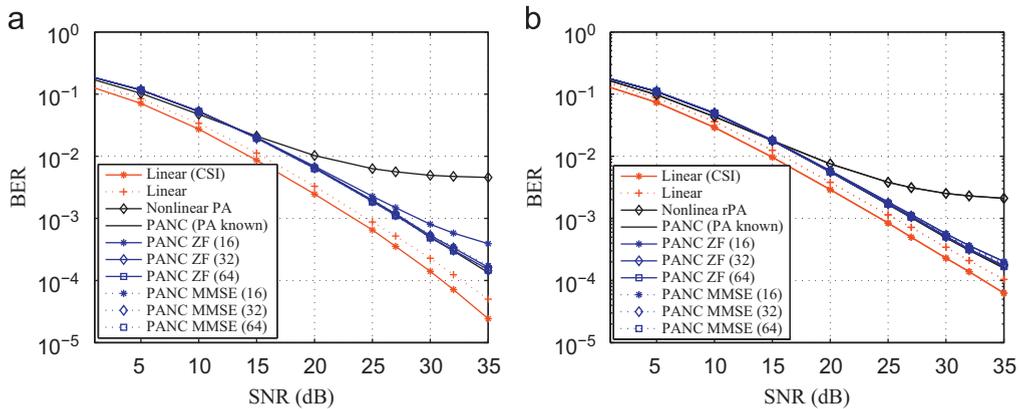


Fig. 11. BER versus SNR for PANC with model estimation in an OFDM system with 16-QAM modulation on $N=512$ subcarriers, for the case of linear PA and NL PA. The mobile speed is set to $v=2$ km/h. A Wiener–Hammerstein type nonlinear PA was considered where $g[\cdot]$ emulates an SSPA model with $p=2$. Results are shown for ZF and MMSE equalization of $H_i(z)$ which was estimated during one OFDM block using $T=16$, $T=32$ or $T=64$ pilot subcarriers. The model identification was performed during 4 OFDM blocks. BER curves for linear PA with perfect channel state information (Linear (CSI)), and PANC assuming perfect knowledge of PA model (PANC (PA known)), are included for comparison purposes. (a) Clipping level $\gamma = 3$ dB, (b) Clipping level $\gamma = 4$ dB.

four OFDM blocks are almost identical. It verified that there is no performance loss due to the use of Gaussian training signals.

To compare the performance of PANC with existing methods, we implemented a nonlinear equalizer [39,40] using memory polynomials [2]. That is, the output of the

nonlinear equalizer is given by

$$\hat{x}(n) = \sum_{k=1}^K \sum_{q=0}^Q c_{qk} y(n-q) |y(n-q)|^{k-1} \quad (44)$$

where c_{qk} are the equalizer coefficients, $\hat{x}(n)$ and $y(n)$ are the respective output and input of the equalizer, K and Q define the polynomial order and the memory length, respectively. Given a training sequence, the equalizer coefficients are easily obtained using a least-squares approach, see, e.g., [41].

Fig. 12 shows the resulting BER curves for the nonlinear equalizer (44) and PANC both using a training sequence of four OFDM symbols. The memory polynomial was implemented using $Q=5$ and $K=5$ (only odd terms) which gives an equalizer with 15 coefficients. The PA is a Wiener-Hammerstein model $g[\cdot]$ defined by (39) with $p=2$ and clipping level $\gamma=3$ dB. We see that the PANC technique presents a significant BER improvement when compared with the nonlinear equalizer. The error floor associated with the nonlinear equalizer is due to the inability to mitigate clipping distortion.

In order to reach tolerable BER levels when a nonlinear PA is employed, power back-off can be applied to linearize the system, at the expense of a decreased power efficiency. The input back-off (IBO) is defined as the ratio of the average power at the PA input and the input saturation power.

Total degradation (TD) quantifies the allowed level of nonlinear distortion and power efficiency. It is defined as $TD_{dB} = SNR_{dB}^{NL} - SNR_{dB}^L + IBO_{dB}$, where SNR_{dB}^{NL} is the SNR required to obtain a fixed BER target in presence of PA nonlinearities with a fixed IBO, and SNR_{dB}^L expresses the SNR required in the case of linear PA. The BER target is typically set to $BER=10^{-4}$. Total degradation curves for a nonlinear PA with and without PANC shown in Fig. 13. We see that the best operating point for a system with PANC is $IBO=1$ dB with a $TD=1$ dB. Without applying PANC, the

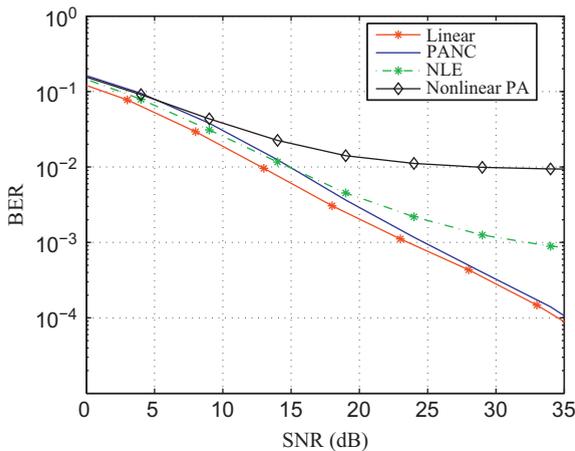


Fig. 12. BER versus SNR for a nonlinear equalizer (NLE), with $Q=5$ and $K=5$, and PANC with model estimation in an OFDM system with 16-QAM modulation on $N=512$ subcarriers, for the case of linear PA and NL PA. The mobile speed is set to $v=2$ km/h. Wiener-Hammerstein type nonlinear PA was considered where $g[\cdot]$ is given by (39). Clipping level γ was set to 3 dB.

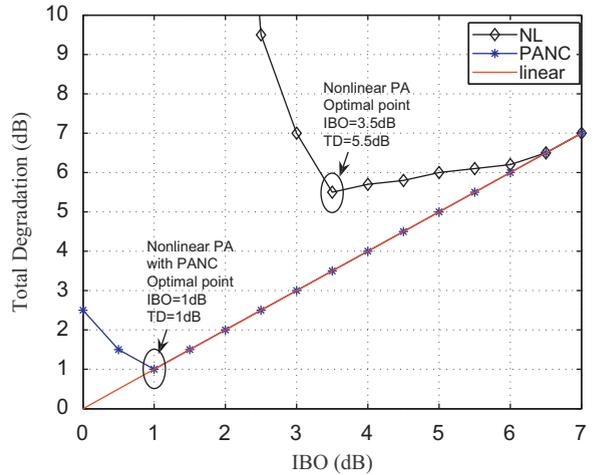


Fig. 13. Total degradation versus IBO for PANC with model estimation in an OFDM system with 16-QAM modulation on $N=512$ subcarriers, for the case of linear PA and nonlinear PA. The mobile speed is set to $v=2$ km/h. Wiener-Hammerstein type nonlinear PA was considered where $g[\cdot]$ was an SSPA model with $p=2$. Clipping level γ was set to 4 dB.

optimal point is $IBO=3.5$ dB with a $TD=3.5$ dB. Assuming a Class-A PA, the optimal points define a power efficiency of $\eta_{PANC}=30\%$ with PANC and $\eta_{NL}=15\%$ without PANC.

Finally, in order to evaluate the effects of model mismatch between the PA employed and the model assumed by the identification algorithm (W-H model), we use a PA modeled with a three-branch parallel Wiener model [44]. This structure was successfully employed to model class AB (low-power amplifier) and class B (high efficiency) amplifiers [45]. The linear blocks of the parallel Wiener model are given by [44]

$$H_1(z) = 1, \quad H_2(z) = \frac{1+0.3z^{-1}}{1-0.1z^{-1}}, \quad H_3(z) = \frac{1-0.2z^{-1}}{1-0.4z^{-1}} \quad (45)$$

The memoryless nonlinearity of branch i is modeled using a polynomial given

$$y_i(n) = \sum_{k=1}^K d_{ki} v_i(n) |v_i(n)|^{k-1} \quad (46)$$

where $v_i(n)$ is the output of the linear filter $H_i(z)$. The polynomial coefficients are

$$\begin{aligned} d_{11} &= 1.0018 + 0.0858j, & d_{31} &= 0.0879 - 0.1583j, \\ d_{51} &= -1.0992 - 8891j \\ d_{12} &= 0.1179 + 0.0004j, & d_{32} &= -0.1818 + 0.0391j, \\ d_{52} &= 0.1684 + 0.0034j \\ d_{13} &= 0.0473 - 0.0058j, & d_{33} &= 0.0395 + 0.0283j, \\ d_{53} &= -0.1015 - 0.0196j \end{aligned} \quad (47)$$

BER curves for the parallel Wiener model are shown in Fig. 14. We see that the PANC technique presents a significant BER improvement when compared to the case without compensation.

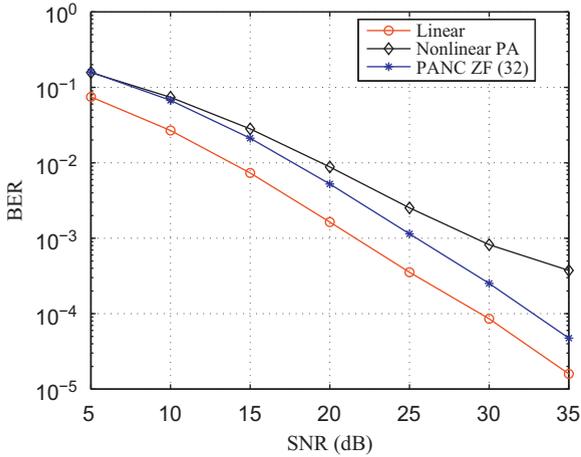


Fig. 14. BER versus SNR for PANC with model estimation in an OFDM system with 16-QAM modulation on $N=512$ subcarriers, for the case of linear PA and NL PA. The mobile speed is set to $v=2$ km/h. A parallel Wiener type nonlinear PA was considered that emulates a class AB PA with $I\text{BO}=6$ dB. Results are shown for ZF equalization of $H_i(z)$ which was estimated during one OFDM block using $T=32$. The model identification was performed during 4 OFDM blocks.

6. Conclusions

This paper proposed a receiver-based approach for modeling and compensating the distortion from nonlinear power amplifiers (PAs) with memory. A power amplifier nonlinearity cancellation (PANC) technique was proposed to reduce the harmful effects of wideband PAs in OFDM systems. Our technique is very useful to minimize the impact of wideband nonlinear amplifiers on the system performance, that it is the special relevance for mobile terminals, allowing for a more efficient usage of the power amplifier. The proposed technique estimates the nonlinear distortion by passing tentative decisions through a wideband PA model. The memory effects of the PA are taken into account by modeling the PA as a Wiener–Hammerstein system. The PANC technique is suitable for systems where reduced implementation complexity is required in the transmitter and can be combined with other distortion mitigation techniques, e.g., predistortion. We, thereafter, addressed the problem of identifying and tracking the Wiener–Hammerstein PA model parameters at the receiver. The approach greatly simplifies implementation complexity as the need for a computationally complex PA identification step at the transmitter is avoided. The proposed PA model identification scheme consists of two steps, where an initial channel estimation and equalization step in frequency-domain is followed by a time-domain identification of a simplified Wiener–Hammerstein model. A stochastic algorithm was derived for identifying the PA model parameters and its convergence properties were analyzed. Comparison with other methods showed that the proposed identification technique exhibits faster convergence speed and lower complexity.

Our simulation results verify a minimal performance loss, in terms of BER, when comparing the performance of

the PANC technique using estimated and perfect PA models. BER close to those observed for the case of a linear PA are obtained which allows for a more power efficient transmission. Furthermore, the PANC technique avoids the error floor due to clipping noise associated with polynomial-based nonlinear equalizers. Moreover, the robustness of our method against model mismatch was verified given very good results. The proposed technique can be extended for a system including channel coding in order to get reliable results in a low SNR scenario. The combination of PANC and channel coding has been addressed with very good results in [46].

Appendix A. Verification of $\nabla_{\hat{\mathbf{a}}}[u_2(n)]$

Verification of $\nabla_{\hat{\mathbf{a}}}[u_2(n)]$ in (29) is straightforward, as follows: Using (23) and (24), it is possible to obtain

$$\nabla_{\hat{\mathbf{a}}}[u_2(n)] = \frac{\partial \hat{u}_2(n)}{\partial \hat{\mathbf{s}}(n)} \cdot \frac{\partial \hat{\mathbf{s}}(n)}{\partial \hat{\mathbf{a}}} = \frac{\partial \hat{u}_2(n)}{\partial \hat{\mathbf{s}}(n)} \mathbf{x}(n) \quad (48)$$

To proceed, we employ the basic passband representation of the conventional polynomial model that is described by

$$\tilde{u}_2(n) = \sum_{k=0}^P \tilde{g}_k \tilde{s}^k(n) \quad (49)$$

where $\tilde{(\cdot)}$ is the notation for passband signal and coefficients, and P is the order of the polynomial model [42]. To obtain $\partial \hat{u}_2(n)/\partial \hat{\mathbf{s}}(n)$ and avoid the problem of baseband model differentiation, we first differentiate (49) and then transform the result to baseband. Differentiation of (49) with respect to \tilde{s} gives

$$\frac{\partial \tilde{u}_2(n)}{\partial \tilde{s}(n)} = \sum_{k=1}^P k \tilde{g}_k \tilde{s}^{k-1}(n) \quad (50)$$

The baseband form of (50), assuming s to be a narrow-band signal with respect to the carrier frequency, is then given by [43]

$$\begin{aligned} \frac{\partial u_2(n)}{\partial s(n)} &= g_1 + \sum_{k=0}^{K-1} g'_{2k+3} \{s^{k+1}(n)[(s^*(n))^k]\} \\ &= g_1 + \sum_{k=0}^{K-1} (2k+3) g_{2k+3} |s(n)|^{2k} s(n) \end{aligned} \quad (51)$$

We see that, except for the constant g_1 and the order of the polynomial, the derivative has the same form as the memoryless polynomial model. Since we have used an arbitrary bandpass memoryless polynomial model for $\tilde{u}_2(n)$ [42], we have avoided the difficulties with differentiation that arise when adopting an odd-order polynomial approximation in baseband. Finally, replacing coefficients g_k by their estimates $\hat{g}_k(n)$, the gradient $\nabla_{\hat{\mathbf{a}}}[u_2(n)]$ is obtained.

Appendix B. Local convergence proof

By definition, the matrix $-\mathbf{P}_s(\bar{\theta}_s)$ is positive semidefinite. For this purpose we use the orthogonal basis description of the memoryless polynomial model. Let us

rewrite $\mathbf{P}_s(\bar{\theta}_s)$ as

$$-\mathbf{P}_s(\bar{\theta}_s) = \begin{bmatrix} \mathbf{R}_1 & \mathbf{R}_3 \\ \mathbf{R}_3^H & \mathbf{R}_2 \end{bmatrix} \quad (52)$$

To show that $-\mathbf{P}_s(\bar{\theta}_s) > 0$, it is sufficient to show that \mathbf{R}_1 and \mathbf{R}_2 are positive definite matrices, see, e.g., [47, Lemma A.3] or [34, Lemma G.1]. Using the orthogonal basis $\phi_p = \Lambda_p^{-1} \psi_p$ and (34), we get

$$\mathbf{R}_1 = \Lambda_p^{-1} E\{\psi_p \psi_p^H\} \Lambda_p^{-H} = \Lambda_p^{-1} \Lambda_p^{-H} \quad (53)$$

Therefore, \mathbf{R}_1 is positive definite because Λ_p matrix is lower triangular, i.e., non singular. Λ_p in Eq. (53) acts as a similarity transformation on \mathbf{R}_1 . i.e., $\Lambda_p \mathbf{R}_1 \Lambda_p^H = \mathbf{I}$. Then, \mathbf{R}_1 is positive definite. We now turn to matrix \mathbf{R}_2 which is given by

$$\mathbf{R}_2 = E\{\nabla_{\hat{\mathbf{a}}}[u_1] \nabla_{\hat{\mathbf{a}}}^H[u_1]\} + E\{\nabla_{\hat{\mathbf{a}}}[u_2] \nabla_{\hat{\mathbf{a}}}^H[u_2]\} - E\{\nabla_{\hat{\mathbf{a}}}[u_1] \nabla_{\hat{\mathbf{a}}}^H[u_2]\} - E\{\nabla_{\hat{\mathbf{a}}}[u_2] \nabla_{\hat{\mathbf{a}}}^H[u_1]\} \quad (54)$$

Element (ij) of the second term in (54) is obtained using (15), (21) and (21) as

$$\begin{aligned} [E\{\nabla_{\hat{\mathbf{a}}}[u_2] \nabla_{\hat{\mathbf{a}}}^H[u_2]\}]_{i,j} &= [\hat{g}_1]^2 E\{x(n-i)x^*(n-j)\} \\ &+ E\left\{ \left[\sum_{k=0}^{K-1} (2k+3) \hat{g}_{2k+3} \phi_{2k+1}(\hat{s}(n)) \right] x(n-i) \right. \\ &\left. \times x^*(n-j) \left[\sum_{k=0}^{K-1} (2k+3) \hat{g}_{2k+3} \phi_{2k+1}^*(\hat{s}(n)) \right] \right\} \end{aligned} \quad (55)$$

where the cross-terms in $[E\{\nabla_{\hat{\mathbf{a}}}[u_2] \nabla_{\hat{\mathbf{a}}}^H[u_2]\}]_{i,j}$ are zero because they can be written in terms of odd-order moments of zero-mean Gaussian random variables. Using the orthogonal basis representation for the derivative polynomial Model,¹ $\phi_{p-1} = \Lambda_{p-1}^{-1} \psi_{p-1}$, we obtain

$$\begin{aligned} [E\{\nabla_{\hat{\mathbf{a}}}[u_2] \nabla_{\hat{\mathbf{a}}}^H[u_2]\}]_{i,j} &= [\hat{g}_1]^2 E\{x(n-i)x^*(n-j)\} \\ &+ \sum_{k=0}^{K-1} |\bar{\lambda}_{kk}|^2 E\{\psi_{2k+1}(\hat{s}(n))x(n-i)x^*(n-j)\psi_{2k+1}^*(\hat{s}(n))\} \end{aligned} \quad (56)$$

where $\bar{\lambda}_{kk}$ are the diagonal values of the matrix Λ_{p-1}^{-1} . By employing the following property of high-order moments of Gaussian random variables [48]

$$E\{z_1 z_2 z_3^* z_4^*\} = E\{z_1 z_3^*\} E\{z_2 z_4^*\} + E\{z_2 z_3^*\} E\{z_1 z_4^*\} \quad (57)$$

the second expectation in (56) is simplified to

$$\begin{aligned} E\{\psi_{2k+1}(\hat{s}(n))x(n-i)x^*(n-j)\psi_{2k+1}^*(\hat{s}(n))\} \\ = \underbrace{E\{\psi_{2k+1}(\hat{s}(n))\psi_{2k+1}^*(\hat{s}(n))\}}_{=1} E\{x(n-i)x^*(n-j)\} \\ + E\{\psi_{2k+1}(\hat{s}(n))x(n-i)\} E\{x^*(n-j)\psi_{2k+1}^*(\hat{s}(n))\} \end{aligned} \quad (58)$$

Note that the second term $E\{\psi_{2k+1}(\hat{s}(n))x(n-i)\} = E\{s^{k+1}(s^*)^k x(n-i)\}$ in (58) is zero because it is an odd-order moment of a zero-mean Gaussian variable [48]. Thus, $E\{\nabla_{\hat{\mathbf{a}}}[u_2] \nabla_{\hat{\mathbf{a}}}^H[u_2]\} > 0$ if the input correlation matrix $\mathbf{R}_x = [E\{x(n-i)x^*(n-j)\}]_{i,j}$ is positive definite, which is naturally fulfilled.

Let us now consider the first term of \mathbf{R}_2 in (54). In order to proceed we let $\{a_i(n)\}_{n=0}^{n_{a_i}}$ be the unbiased

estimates of the coefficients associated with the inverse system $A^{-1}(z)$, which appears in Fig. 3 as a result of (17). Component (ij) is then given by

$$\begin{aligned} [E\{\nabla_{\hat{\mathbf{a}}}[u_1] \nabla_{\hat{\mathbf{a}}}^H[u_1]\}]_{i,j} &= E\left\{ \left(\sum_{p=0}^{n_{a_i}} \sum_{k=0}^K a_i(p) \hat{g}_{2k+1} \phi_{2k+1}[s(n-i+p)] \right) \right. \\ &\left. \times \left(\sum_{q=0}^{n_{a_i}} \sum_{m=0}^K a_i^*(q) \hat{g}_{2m+1} \phi_{2m+1}^*[s(n-j+q)] \right) \right\} \end{aligned} \quad (59)$$

Replacing the conventional polynomial bases with the orthogonal basis, we get

$$\begin{aligned} [E\{\nabla_{\hat{\mathbf{a}}}[u_1] \nabla_{\hat{\mathbf{a}}}^H[u_1]\}]_{i,j} &= \sum_{p=0}^{n_{a_i}} \sum_{k=0}^K \sum_{q=0}^{n_{a_i}} \sum_{m=0}^K a_i(p) a_i^*(q) \lambda_{kk} \lambda_{mm}^* \\ &\times E\{\psi_{2k+1}[s(n-i+p)] \psi_{2m+1}^*[s(n-j+q)]\} \end{aligned} \quad (60)$$

Furthermore, making use of (34) and (60) reduces to

$$[E\{\nabla_{\hat{\mathbf{a}}}[u_1] \nabla_{\hat{\mathbf{a}}}^H[u_1]\}]_{i,i} = \sum_{k=0}^K \sum_{p=i}^{n_{a_i}} |a_i(p)|^2 |\lambda_{kk}|^2 \quad (61)$$

which is positive definite. Finally, using also the orthogonal basis description for the conventional and the derivative polynomial models, the last term of \mathbf{R}_2 in (54), $E\{\nabla_{\hat{\mathbf{a}}}[u_1] \nabla_{\hat{\mathbf{a}}}^H[u_2]\}$ can be written in terms of odd moments of Gaussian variables and, as a consequence, its contribution to \mathbf{R}_2 is zero. Therefore, \mathbf{R}_2 is positive definite. This concludes the local convergence proof.

References

- [1] D.R. Morgan, Z. Ma, J. Kim, M.G. Zierdt, J. Pastalan, A generalized memory polynomial model for digital predistortion of RF power amplifiers, *IEEE Trans. Signal Process.* 54 (10) (2006) 3852–3860.
- [2] L. Ding, G.T. Zhou, D.R. Morgan, Z. Ma, J.S. Kenney, J. Kim, C.R. Giardina, A robust digital baseband predistorter constructed using memory polynomials, *IEEE Trans. Commun.* 52 (1) (2004) 159–165.
- [3] J. Vuolevi, Analysis, measurement and cancellation of the bandwidth and amplitude dependence of intermodulation distortion in RF power amplifiers, Ph.D. Thesis, University of Oulu, Finland, October 2001.
- [4] J. Vuolevi, T. Rahkonen, *Distortion in RF Power Amplifiers*, Norwood Artech House, 2003.
- [5] H. Ku, M.D. McKinley, J.S. Kenney, Quantifying memory effects in RF power amplifiers, *IEEE Trans. Microw. Theory Technol.* 50 (12) (2002) 2843–2849.
- [6] S. Serfaty, J.L. LoCicero, G.E. Atkin, Cancellation of nonlinearities in bandpass QAM systems, *IEEE Trans. Commun.* 38 (10) (1990) 1835–1843.
- [7] J.-H. Ryu, Y.-H. Lee, Combined equalization and nonlinear distortion cancellation for transmission of QAM signal in fixed wireless channel, in: *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2003*, vol. IV, April 2003, pp. 489–492.
- [8] F. Gregorio, S. Werner, J. Cousseau, T. Laakso, Receiver cancellation technique for nonlinear power amplifier distortion in SDMAOFDM systems, *IEEE Trans. Veh. Technol.* 56 (5) (2007) 2499–2516 Part I.
- [9] F. Gregorio, T. Laakso, J. Cousseau, Receiver cancellation of nonlinear power amplifier distortion in SDMA-OFDM systems, in: *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2006*, May 2006, pp. IV–IV.
- [10] J. Tellado, L. Hoo, J. Cioffi, Maximum-likelihood detection of nonlinearly distorted multicarrier symbols by iterative decoding, *IEEE Trans. Commun.* 51 (February) (2003) 218–228.
- [11] T. Schenk, C. Dehos, D. Morche, E. Fledderus, Receiver-based compensation of transmitter-incurred nonlinear distortion in multiple-antenna OFDM systems, in: *Proceedings of the IEEE Vehicular Technology Conference, VTC-Fall 2007*, Baltimore, 2007, vol. 11, pp. 1346–1350.

¹ Same as for the conventional polynomial model except for the order, see (51).

- [12] P. Drotr, J. Gazda, P. Galajda, D. Kocur, P. Pavelka, Receiver technique for iterative estimation and cancellation of nonlinear distortion in MIMO SFBC-OFDM systems, *IEEE Trans. Consum. Electron.* 56 (2) (2010) 471–475.
- [13] A.S. Tehrani, H. Cao, A. Behravan, T. Eriksson, C. Fager, Successive cancellation of power amplifier distortion for multiuser detection, in: *Proceedings of the IEEE Vehicular Technology Conference, VTC-Fall 2010, Ottawa, September 2010*, pp. 1–5.
- [14] G. Giannakis, E. Serpendin, A bibliography on nonlinear system identification, *Signal Processing* 81 (March) (2001) 533–580.
- [15] S. Fakhouri, Identification of a class of non-linear systems with Gaussian non-white inputs, *Int. J. Syst. Sci.* 1 (May) (1980) 541–550.
- [16] S.A. Billings, S. Fakhouri, Identification of systems containing linear dynamic and static nonlinear elements, *Automatica* 18 (January) (1982) 15–26.
- [17] H. Boutayeb, M. Darouach, Recursive identification method for MISO Wiener–Hammerstein model, *IEEE Trans. Autom. Control* 40 (2) (2001) 287–291.
- [18] N. Bershad, S. Bouchired, F. Castanie, Stochastic analysis of adaptive gradient identification of Wiener–Hammerstein systems for Gaussian inputs, *IEEE Trans. Signal Process.* 48 (2) (2000) 557–560.
- [19] N. Bershad, P. Celka, S. McLaughlin, Analysis of stochastic gradient identification of Wiener–Hammerstein systems for nonlinearities with hermite polynomial expansions, *IEEE Trans. Signal Process.* 49 (5) (2001) 1060–1072.
- [20] E. Aschbacher, M. Rupp, Robust identification of an L–N–L system, *Conference Record of the Thirty-Seventh Asilomar Conference on Signals, Systems and Computers*, vol. 2, 2003, pp. 1298–1302.
- [21] E. Aschbacher, M. Rupp, Identification of a nonlinear power-amplifier L–N–L structure for pre-distortion purposes, *Proceedings of the IEEE 14th International Symposium on Personal, Indoor and Mobile Radio Communications*, vol. 3, 2003, pp. 2102–2106.
- [22] H.E. Emara-Shabaik, M.S. Ahmed, K.H. Al-Ajmi, Wiener–Hammerstein model identification-recursive algorithms, *JSMIE Int. J. Ser. C Mech. Syst., Mach. Elem. Manuf.* 45 (2) (2002) 606–613.
- [23] P. Crama, J. Schoukens, Initial estimates of Wiener and Hammerstein systems using multisine excitation, *IEEE Trans. Instrum. Meas.* 50 (6) (2001) 1791–1795.
- [24] P. Crama, J. Schoukens, Wiener–Hammerstein system estimator initialisation using a random multisine excitation, *Proceedings of the 58th ARFTG Conference Digest-Fall*, vol. 40, 2001, pp. 1–6.
- [25] J. Schoukens, R. Pintelon, J. Paduart, G. Vandersteen, Nonparametric initial estimates for Wiener–Hammerstein systems, *Proceedings of the 14th IFAC Symposium on System Identification (SYSID 2006)*, vol. 3, 2006.
- [26] D. Dardari, V. Tralli, A. Vaccari, A theoretical characterization of nonlinear distortion effects in OFDM systems, *IEEE Trans. Commun.* 48 (10) (2000) 1755–1764.
- [27] G.T. Zhou, H. Qian, L. Ding, R. Raich, On the baseband representation of a bandpass nonlinearity, *IEEE Trans. Signal Process.* 53 (8) (2005) 2953–2957.
- [28] R. Raich, G.T. Zhou, Orthogonal polynomials for complex Gaussian processes, *IEEE Trans. Signal Process.* 52 (10) (2004) 2788–2797.
- [29] L. Ding, H. Qian, N. Chen, G.T. Zhou, A memory polynomial predistorter implemented using TMS320C67, in: *Proceedings of the Texas Instruments Developer Conference, February 2004, Houston, USA*.
- [30] R. Raich, H. Qian, G.T. Zhou, Orthogonal polynomials for power amplifier modeling and predistorter design, *IEEE Trans. Veh. Technol.* 53 (5) (2004) 1468–1479.
- [31] L. Deneire, P. Vandenameele, L. van der Perre, B. Gyselinckx, M. Engels, A low-complexity ML channel estimator for OFDM, *IEEE Trans. Commun.* 51 (2) (2003) 135–140.
- [32] A.D. Kalafatis, L. Wang, W.R. Cluett, Identification of Wiener-type nonlinear systems in noisy environment, *Int. J. Control* 66 (April) (1997) 923–941.
- [33] L. Ljung, T. Soderstrom, *Theory and Practice of Recursive Identification*, MIT Press, Cambridge, MA, 1983.
- [34] T. Wigren, Recursive identification based on the nonlinear Wiener model, Ph.D. Thesis, Uppsala University, Sweden, October 1990.
- [35] J. Figueroa, J. Cousseau, R. de Figueiredo, A simplicial canonical piecewise linear adaptive filter, *Int. J. Circuits, Syst., Signal* 23 (2004) 365–386.
- [36] W.G. Jeon, K.H. Chang, Y.S. Cho, An adaptive data predistorter for compensation of nonlinear distortion in OFDM systems, *IEEE Trans. Commun.* 45 (10) (1997) 1167–1172.
- [37] A.A.M. Saleh, Frequency-independent and frequency-dependent nonlinear models of TWT amplifiers, *IEEE Trans. Commun.* 29 (11) (1981) 1715–1720.
- [38] F. Gregorio, S. Werner, J. Cousseau, T. Laakso, Iterative channel estimation for multiuser OFDM systems in the presence of power amplifier nonlinearities, *Proceedings of the IEEE 17th International Symposium on Personal, Indoor and Mobile Radio Communications*, vol. 1, 2006, pp. 1–5.
- [39] S. Benedetto, E. Biglieri, Nonlinear equalization of digital satellite channels, *IEEE J. Sel. Areas Commun.* 1 (1) (1983) 57–62.
- [40] A. Behravan, T. Eriksson, Baseband compensation techniques for bandpass nonlinearities [RF front-ends], in: *Proceedings of the IEEE Vehicular Technology Conference, VTC Fall, 2003*, vol. 1, pp. 279–283.
- [41] L. Ding, Z. Ma, D.R. Morgan, M.G. Zierdt, J. Pastalan, A least-squares/Newton method for digital predistortion of wideband signals, *IEEE Trans. Commun.* 54 (5) (2006) 833–840.
- [42] L. Ding, G.T. Zhou, Effects of even-order nonlinear terms on power amplifier modeling and predistortion linearization, *IEEE Trans. Veh. Technol.* 53 (5) (2004) 156–162.
- [43] S. Benedetto, E. Biglieri, R. Daffara, Modeling and performance evaluation of nonlinear satellite links—a volterra series approach, *IEEE Trans. Aerosp. Electron. Syst.* 15 (4) (1979) 494–507.
- [44] L. Ding, Digital predistortion of power amplifiers for wireless applications, Ph.D. Thesis, School of Electrical and Computer Engineering, Georgia Institute of Technology, March 2004.
- [45] H. Ku, M.D. McKinley, J.S. Kenney, Quantifying memory effects in RF power amplifiers, *IEEE Trans. Microw. Theory and Technol.* 50 (12) (2002) 2843–2849.
- [46] L.G. Baltar, S. Dierks, F. Gregorio, J. Cousseau, J.A. Nossek, OFDM receivers with iterative nonlinear distortion cancellation, in: *Proceedings of the IEEE Signal Processing Advances in Wireless Communications, SPAWC 2010, Marrakech, June 2010*.
- [47] T. Soderstrom, P. Stoica, *System Identification*, Prentice-Hall, 1989.
- [48] I. Reed, On the theorem for complex Gaussian processes, *IRE Trans. Inf. Theory* 8 (3) (1962) 194–195.