Adaptive Pre-distortion Techniques based on Orthogonal Polynomials

Robert Dallinger*, Henri Ruotsalainen†, Risto Wichman†, and Markus Rupp*

*Institute of Communications and Radio-Frequency Engineering, Vienna University of Technology,
Email: {rdalling, mrupp}@nt.tuwien.ac.at
†Department of Signal Processing and Acoustics, Aalto University School of Science and Technology,
Email: haruotsa@gmail.com, risto.wichman@tkk.fi

***Abstract*—Pre-distortion in digital baseband is a cost-effective method to linearise the transmit power amplifiers of spectrally efficient communication systems. From a signal processing point of view, the functional structure of the pre-distorter is primarily determined by the decision which model to choose, as well as by the selected adaptive algorithm. During the last two decades, in literature, a multitude of pre-distorter structures has been proposed and analysed. In this work, we focus on simple and commonly employed models (the Wiener model and the Hammerstein model) consisting of a linear filter and a static nonlinearity. The latter is represented using a basis of orthogonal polynomials. First, applying practical transmission signals, different orthogonal polynomial bases are compared with respect to the numerical condition of least squares estimation, and with respect to the convergence behaviour of gradient methods. In a second step, pre-distorters which employ orthogonal polynomials are adapted by the indirect learning structure, respectively, the nonlinear filtered-x least mean squares algorithm. Based on simulations and burst measurements with a commercial power amplifier, the real-world performance of such pre-distortion systems is investigated.**

I. Introduction

In the design of base stations for mobile communication networks of the third generation and higher, the engineer faces two oppositional constraints. On the one hand, the increasing demand on high data rates, which are expected to be ubiquitously available in stationary, as well as in mobile scenarios, accounts for modulation schemes with high spectral efficiency. These modulation schemes, e.g., orthogonal frequency division multiplexing (OFDM) as used in the long term evolution (LTE) [1], lead to transmission signals with high peak-to-average power ratio (PAPR) [2]. Therefore, in contrast to constant envelope modulation techniques, the transmission, and especially the power amplifier (PA), has to be highly linear over a wide dynamic range. On the other hand, power efficiency is a crucial requirement. Currently, more than 50% of the power consumed by a base station can be attributed to the radio frequency PAs. However, currently, such PAs show a power efficiency far below 50% if operated in their linear region [3], [4]. Consequently, in mobile communication networks which employ non-constant envelope modulation schemes, a large portion of the consumed energy is wasted due to the low efficiency of the PAs.

This work has been funded by the NFN SISE S10609 (National Research Network "Signal and Information Processing in Science and Engineering").

Aiming for an increased power efficiency of linearly operated PAs, a multitude of linearisation methods have been proposed [5], [6]. Out of these, the signal processing approach leads to the concept of digital pre-distortion (DPD), which can also be combined with one or even several of the other methods. Likewise, in DPD, plenty of different schemes have been developed (for an overview see, e.g., [6]). Most of them share the same principle, which is depicted in Fig. 1. Before up-conversion to the analogue passband, the baseband representation d_k of the desired transmit signal is additionally processed by a unit, commonly called the pre-distorter (PD), which aims to invert the detrimental effects of the actual PA, such that in the ideal case, the output $z(t)$ of the PA is simply a linearly amplified continuous passband version of d_k. In practice, the behaviour of the PA changes with time [7], therefore, usually, the PD needs to be adaptively adjusted, which is indicated by the feedback path in Fig. 1. Note that in this work, we focus on baseband signals. Hence, the term PA, in most cases refers to its baseband model, i.e., the whole boxed cascade in Fig. 1. In ambiguous cases, the actual meaning will be clarified specifically.

For small signal bandwidths, the behaviour of the PA can be very well approximated by a nonlinear saturating one-to-one mapping [8]. In this case, the PD is typically implemented by a look-up table [7]. For larger signal bandwidths [8], the output $z(t)$ also depends on the recent history of the input $x(t)$ [9]. Then, the distortions introduced by the PA cannot be compensated by a PD which simply maps the current input value to a corresponding output value. One of the most general approaches for DPD, which is additionally able to compensate for memory effects, is based on truncated Volterra series [8], [10]. However, due to the high number of parameters which is rapidly growing with the degree of the nonlinearity, as well as with the memory length, simplifications have been proposed, e.g., memory polynomials [11]. A further drastic reduction in the number of parameters can be achieved by the use of two-block models. These models consist of a block containing a linear filter which is in cascade with a second block that incorporates a static nonlinearity. If the nonlinearity precedes the linear filter, the structure is called a Hammerstein model. If the order of the blocks is reversed, the structure is commonly referred to as Wiener model [12]–[14].

This work is based on [15] which in turn strongly follows

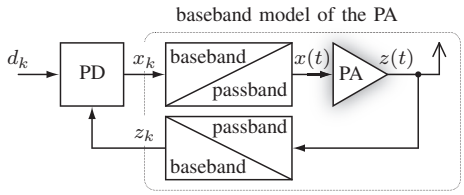

Fig. 1. Basic principle of an adaptive baseband pre-distortion system.

the lines of [16]–[18]. In Sec. II, an adequate parametrisation of the static nonlinearity is chosen. For this means, different orthogonal polynomial basis expansions are compared. In Sec. II-B1 this is done with respect to their numerical behaviour for least squares (LS) estimation. Sec. II-B2 investigates their influence to the learning performance of gradient type algorithms. Sec. III describes the two adaptive inverse system identification schemes, which are alternatively used for identification of the considered PDs. Finally, both identification schemes are compared by simulations in Sec. III-A and measurements in Sec. III-B.

II. Nonlinear basis expansion using orthogonal polynomials

This work focuses on Hammerstein PDs. However, for the direct learning approach in Sec. III, also Wiener models need to be considered. For both of these two-block models, in the sequel, the linear part is designed as finite impulse response (FIR) filter. The representation of a static nonlinearity can be done in many ways. One approach is to approximate it by piecewise linear functions. This method has been applied to the real-valued case [19], as well as to the complex domain [20]. It allows for computationally very efficient implementations. However, due to the inherently occurring edges in the nonlinear characteristic, by principle, smooth (in the sense of differentiable) nonlinear functions can never be represented perfectly with a finite number of partitions. In this work, an alternative approach is chosen. The nonlinearity is represented by a polynomial, or more precisely, by a sum of polynomials, which will become clear in Sec. II-B. Consequently, the actual static nonlinearity is approximated by a truncated Taylor series about the origin [21].

A. Polynomial baseband models

As explained, e.g., in [22, p. 69], for a polynomial passband nonlinearity of degree $P > 0$, the output z_k of the equivalent (static) nonlinear baseband model in Fig. 1, can be expressed as a function of the input x_k by

$$z_k = \sum_{p=0}^{\lfloor \frac{P-1}{2} \rfloor} a_{2p+1} x_k \left| x_k \right|^{2p}, \tag{1}$$

with the complex-valued coefficients a_i. However, in [23] it was argued that for a PA, the static nonlinearity itself as well as its inverse, can be approximated more precisely if the baseband model in (1) also incorporates even order terms. Since this also represents the more general approach, in this work, the following baseband is adopted,

$$z_k = \sum_{p=1}^{P} a_p x_k \left| x_k \right|^{p-1}. \tag{2}$$

Eqn. (2) can be seen as basis expansion with the monomial basis functions $\phi_p(x_k) = x_k \left| x_k \right|^{p-1}$. As a further generalisation, the basis functions can be allowed to be polynomials themselves. This leads to the representation of the static nonlinearity which is used throughout this work,

$$z_k = \sum_{m=1}^{M_h} h_{m,k} \psi_m(x_k), \quad \psi_m(x_k) = \sum_{p=1}^{m} b_{m,p} x_k \left| x_k \right|^{p-1}, \tag{3}$$

where $h_{m,k}$ are the *expansion coefficients* for the nonlinear basis composed by M_h basis polynomials $\psi_m(x_k)$. The actual basis polynomials $\psi_m(x_k)$ are characterised by the coefficients $b_{m,p}$, to which we will refer by the term *polynomial coefficients*. Introducing the row vector of monomials $\boldsymbol{\phi}(x_k) = [x_k, x_k|x_k|, x_k|x_k|^2, \ldots, x_k|x_k|^{M_h-1}]$, and the upper triangular matrix $\mathbf{B}$ of the polynomial coefficients,

$$\mathbf{B} = \begin{pmatrix} b_{1,1} & b_{1,2} & \ldots & b_{1,M_h} \\ 0 & b_{2,2} & \ddots & b_{2,M_h} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \ldots & b_{M_h,M_h} \end{pmatrix}, \tag{4}$$

the row vector of basis functions $\boldsymbol{\psi}(x_k)$ can be compactly written as $\boldsymbol{\psi}(x_k) = [\psi_1(x_k), \psi_2(x_k), \ldots, \psi_{M_h}(x_k)] = \boldsymbol{\phi}(x_k)\mathbf{B}$. Consequently, with $\mathbf{h}_k = [h_{1,k}, h_{2,k}, \ldots, h_{M_h,k}]^{\mathsf{T}}$, the output

$$z_k = \boldsymbol{\psi}(x_k)\mathbf{h}_k = \boldsymbol{\phi}(x_k)\mathbf{B}\mathbf{h}_k. \tag{5}$$

B. Orthogonal polynomial bases

In this section, the choice of the polynomial coefficients $b_{m,p}$ in (3) and (4) is considered. From an algebraic point of view, they can be chosen rather arbitrarily. However, in the context of nonlinear system identification, one choice may outperform another. In the following, we consider the common problem of an unknown reference system that has to be identified based on the observations of its input samples x_k and its output samples z_k (cf. Fig. 2). If a polynomial model is used according to Sec. II-A, with $N \geq M_h$ observations of the pair (x_k, z_k), the LS estimate $\hat{\mathbf{h}}_k$ of the unknown parameter vector $\mathbf{h}$ is given by [24, p. 664]

$$\hat{\mathbf{h}}_k = \left(\boldsymbol{\Psi}^{\mathsf{H}}(\mathbf{x}_k)\boldsymbol{\Psi}(\mathbf{x}_k) \right)^{-1} \boldsymbol{\Psi}^{\mathsf{H}}(\mathbf{x}_k)\mathbf{z}_k, \tag{6}$$

with the matrix $\boldsymbol{\Psi}^{\mathsf{T}}(\mathbf{x}_k) = [\boldsymbol{\psi}^{\mathsf{T}}(x_k), \ldots, \boldsymbol{\psi}^{\mathsf{T}}(x_{k-N+1})]$, $\mathbf{x}_k = [x_k, \ldots, x_{k-N+1}]^{\mathsf{T}}$, and with $\mathbf{z}_k = [z_k, \ldots, z_{k-N+1}]^{\mathsf{T}}$. Hence, for the matrix inversion in (6), the numerical condition of the $(M_h \times M_h)$-matrix $\boldsymbol{\Psi}^{\mathsf{H}}(\mathbf{x}_k)\boldsymbol{\Psi}(\mathbf{x}_k)$ is of special interest. Under the assumption that x_k is a random sequence, a point of reference can be found by observing the condition number[1] of the expectation of $\boldsymbol{\Psi}^{\mathsf{H}}(\mathbf{x}_k)\boldsymbol{\Psi}(\mathbf{x}_k)$,

$$\kappa(\mathbf{P}) = \frac{\lambda_{\max}(\mathbf{P})}{\lambda_{\min}(\mathbf{P})}, \quad \text{with} \quad \mathbf{P} = \mathrm{E}\left\{ \boldsymbol{\Psi}^{\mathsf{H}}(\mathbf{x}_k)\boldsymbol{\Psi}(\mathbf{x}_k) \right\}, \tag{7}$$

where $\lambda_{\min}(\mathbf{P})$ and $\lambda_{\max}(\mathbf{P})$ are the minimum and the maximum eigenvalue of the matrix $\mathbf{P}$, respectively [21]. Assuming that x_k is stationary, with the magnitude $|x_k|$ obeying the

[1] Here, we use the condition number based on the induced l_2 matrix norm.

probability density function (PDF) $f_{|x|}$, (5) together with the matrix $\mathbf{\Phi}^{\mathsf{T}}(\mathbf{x}_k) = [\boldsymbol{\phi}^{\mathsf{T}}(x_k), \ldots, \boldsymbol{\phi}^{\mathsf{T}}(x_{k-N+1})]$, leads to

$$\mathbf{P} = \mathbf{B}^{\mathsf{H}} \, \mathrm{E}\left\{ \mathbf{\Phi}^{\mathsf{H}}(\mathbf{x}_k) \mathbf{\Phi}(\mathbf{x}_k) \right\} \mathbf{B} = N \mathbf{B}^{\mathsf{H}} \mathbf{M}_{f_{|x|}} \mathbf{B}, \quad (8)$$

where the element at row i, and column j, of the matrix $\mathbf{M}_{f_{|x|}}$ is given by the $(i+j)$th statistical moment of $|x_k|$. Obviously, if the upper triangular matrix $\mathbf{B}$ can be chosen such that $\mathbf{P}$ becomes diagonal, the polynomials characterised by $\mathbf{B}$ are orthogonal in the sense,

$$\mathrm{E}\left\{ \psi_m(x_k) \, \psi_n^*(x_k) \right\} = 0, \quad \forall \, m \neq n. \quad (9)$$

Each of the well-known polynomials, such as Hermite, Chebyshev, Laguerre, and Legendre [25], is orthogonal with respect to a certain PDF. These classical polynomials differ from the baseband model in (3) in two ways. First, their orthogonality is defined with respect to a real-valued variable. Consequently, the PDF which they are based on, does not necessarily vanish for negative values. In contrast, the PDF $f_{|x|}$ in (9) always satisfies $f_{|x|}(\xi) = 0, \xi < 0$. Second, the classical polynomials include a zero-order term, i.e., a constant, which does not agree with (3). In the remainder of this section, we compare five kinds of orthogonal polynomials[2]:

1) The *conventional polynomials* represent the basis of the baseband model in (2). Hence, the row vector of basis functions is given by $\boldsymbol{\phi}(x_k)$ as introduced in Sec. II-A.
2) The *Hermite polyomials* are classical polynomials [25]. They are orthogonal with respect to a zero-mean real-valued Gaussian distributed random variable.
3) The *complex Gaussian polynomials* have been proposed in [17]. They are orthogonal with respect to a (spherically invariant) complex Gaussian PDF. Consequently, the corresponding magnitude is Rayleigh distributed [26].
4) The *complex uniform polynomials* are orthogonal with respect to a complex-valued random variable, with a magnitude uniformly distributed on $[0, 1]$ and with an angle uniformly distributed on $[0, 2\pi)$. See [16] for details.
5) The *modified Chebyshev polynomials* are based on the shifted Chebyshev polynomials of the 2nd kind. Latter are orthogonal for a real-valued argument that is distributed by a Wigner semicircle distribution shifted to the interval $[0, 1]$ [25]. However, here, they are modified such, that no zero-order component is contained in the basis functions. Moreover, the PDF refers to the magnitude of the now complex-valued argument. The derivation is skipped here for brevity reasons and will be presented elsewhere.

All of the above polynomials comply with the baseband model in (3) except for the Hermite polynomials and the complex Gaussian polynomials. The former are not conform with the baseband model, since they also include zero-order terms; the latter only include odd order terms and thus represent a basis for (1). As a consequence, for the complex Gaussian polynomials, a basis with dimension M_h has a polynomial degree of $P = 2M_h - 1$, in contrast to $P = M_h$ for all of the other considered polynomial bases.

[2]Note that in this work we introduce short names for the considered polynomials to facilitate readability. They are neither necessarily common in literature, nor are they mathematically precise.

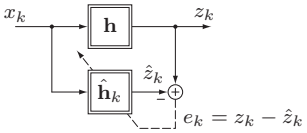

Fig. 2. Nonlinear system identification using a polynomial basis.

The simulations described further down, use either of the following two excitation signals d_k (cf. Fig. 1):

1) A complex-valued zero-mean white random sequence, with a magnitude uniformly distributed on $[0, 1]$ and with a phase, uniformly distributed on $[0, 2\pi)$. Thus, for this signal the complex uniform polynomials are orthogonal. In brief we will refer to this signal as *complex uniform*.
2) A second signal is generated by the publicly available LTE physical layer simulator [27], [28]. The signal represents a downlink OFDM signal of 10MHz bandwidth, with subcarriers modulated by a 16QAM (for details, see [1]). It is normalised to unit maximum magnitude $|d_k|_{\max} = 1$. The PDF of such a signal can be very well approximated by a zero-mean complex Gaussian random signal [29]. Thus, almost perfect orthogonality is expected if used with the complex Gaussian polynomials. In brief we will term this signal the *OFDM signal*.

1) Evaluation of the condition number: In a first step, for all of the previously listed polynomials, the condition number $\kappa(\mathbf{P})$ in (7) is evaluated by simulations for $M_h = 1, 2, \ldots, 7$. Two cases are considered. In one case, the complex uniform signal is used for excitation, in the other case, the OFDM signal. Fig. 3 depicts the results for $N = 10\,000$, using 10 000 averaging passes for the estimation of $\mathbf{P}$. As expected, the classical Hermite polynomials lead to a very poor numerical condition. Also the conventional polynomials do not perform very well for both input signals. The numerical condition is eminent for the complex Gaussian polynomials excited by the OFDM signal, as well as for the complex uniform polynomials excited by the complex uniform signal. For a lower number of basis functions, up to $M_h = 4$, compared to the complex uniform polynomials, the complex Gaussian polynomials seem to be more robust with respect to a deviating input PDF. For $M_h > 4$, the contrary is the case. Possibly due to the qualitative similarity among the uniform PDF and the shifted Wigner semicircle distribution, the modified Chebyshev polynomials behave very similar to the complex uniform polynomials.

2) Learning behaviour of gradient type algorithms: In a second step, the system identification in Fig. 2 is performed by a gradient method. The update equation reads,

$$\hat{\mathbf{h}}_{k+1} = \hat{\mathbf{h}}_k + \frac{\mu_h}{\|\boldsymbol{\psi}^{\mathsf{H}}(x_k)\|_{\mathbf{D}}^2} e_k \mathbf{D} \boldsymbol{\psi}^{\mathsf{H}}(x_k), \quad (10)$$

$$e_k = \boldsymbol{\psi}(x_k) \left(\mathbf{h} - \hat{\mathbf{h}}_k \right) + v_k, \quad (11)$$

where the vectors $\mathbf{h}$ and $\hat{\mathbf{h}}_k$ contain the expansion coefficients of the unknown reference system and the estimated system, respectively. For the simulations, the reference system is first modelled by a Saleh model with the parameters $\alpha_a = 2, \beta_a = 1, \alpha_\phi = 2, \beta_\phi = 1$ (cf. [30]). Then, for this model the expansion coefficients $\mathbf{h}$ are estimated by LS. These

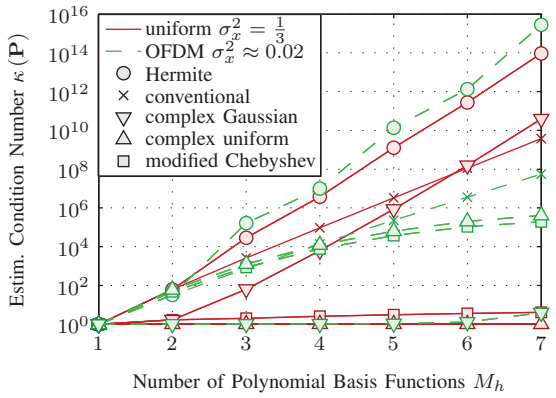

Fig. 3. Condition number $\kappa(\mathbf{P})$ for different polynomial bases, obtained by simulations with $N = 10\,000$, averaged over 10 000 simulation passes.

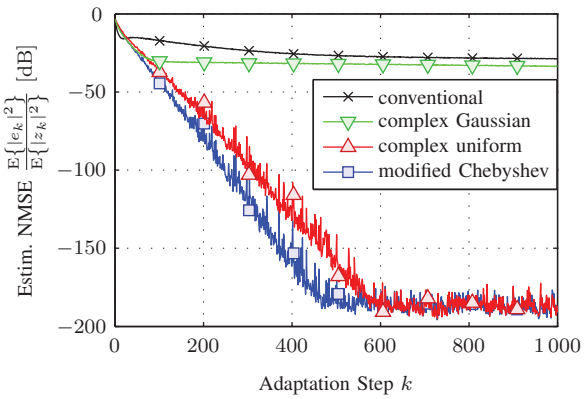

Fig. 4. Estim. normalised mean square error (NMSE) of the PNLMS in (10) with $\mu_h = \frac{1}{2}$, used for adaptive identification of a nonlinearity with complex uniform excitation of variance $\sigma_x^2 = \frac{1}{3}$ (averaged over 10 000 simulations).

expansion coefficients are then used as reference system $\mathbf{h}$ for the adaptive system identification. By this, it is ensured that the system $\hat{\mathbf{h}}_k$, identified by the gradient method, can in principle perfectly resemble the reference system. Hence, an error floor due to inadequate modelling can be excluded. Note that the update error e_k in (11) is additionally disturbed by additive zero-mean complex-valued white Gaussian noise v_k (not depicted in Fig. 2), to ensure a well defined steady state error. The matrix $\mathbf{D}$ is constant and given by

$$\mathbf{D} = \mathrm{diag}_{m=1}^{M_h} \left\{ \mathrm{E}\left\{ \psi_m^2(x_k) \right\}^{-1} \right\}. \tag{12}$$

Hence, the gradient method in (10) can be interpreted as a proportionate normalised LMS (PNLMS) with additionally normalised step-size [31]. Fig. 4 shows the simulation results of the adaptive system identification with constant step-size $\mu_h = \frac{1}{2}$ using the complex uniform excitation signal, averaged over 10 000 passes. The variance of v_k is chosen to be 10^{-20} times smaller than the variance of the output z_k. The simulation demonstrates that with the complex Gaussian polynomials and with the conventional polynomials, the adaptation stalls at a rather high error level[3]. Fast convergence to the minimum steady state error is obtained with both, the complex uniform polynomials and the modified Chebyshev polynomials. What surprises, is that the modified Chebyshev polynomials perform even slightly better than the complex uniform polynomials. Based on the investigations of the condition number in Sec. II-B1, this seems to be counter intuitive.

Repeating the simulations with the complex uniform signal replaced by the OFDM signal, leads to the results depicted in Fig. 5. In this case, the complex uniform polynomials seem to allow for a marginally faster adaptation than the modified Chebyshev polynomials. Interestingly, the complex Gaussian polynomials do not lead to the fastest convergence.

Based on the simulations performed in this section, we conclude that for an OFDM signal, the complex Gaussian polynomials are well suited if the LS method is employed. However, somehow unexpectedly, for gradient methods, the complex uniform polynomials and the modified Chebyshev polynomials lead to significantly faster convergence. Eventually, for the static nonlinearities in Sec. III, the modified Chebyshev polynomials are selected.

[3]Note that even simulations with $N = 150\,000$ do not show any further reduction of the NMSE in Fig. 4.

III. Digital Pre-Distortion Using Orthogonal Polynomials

Sec. II focuses on the representation and estimation of the static nonlinearity based on specific polynomial bases. In this section, we return to the original problem of DPD, as introduced in Sec. I. It is common practice to model the PA by a Wiener system [13], [32]–[34]. Then, its inverse has the structure of a Hammerstein model. Sticking to these design constraints, in the sequel, two well known inverse system identification schemes are compared based on simulations and measurements. The first identification method is a scheme, frequently referred to by indirect learning (IL) [35], [36]. The basic structure of the IL is obtained, when in Fig. 2, the input x_k and the output z_k are interchanged. Hence, the direction of the signal flow in the upper branch is reversed, the rest remains unchanged. The reader is referred to literature for further details. The actual algorithm used to identify the Hammerstein PD is a straightforwardly modified version of the Wiener system identification presented in [14]. The advantage of the IL is its computational simplicity. After adaptation, the estimated parameters approximate the post-inverse of the reference system. Consequently, these parameters only need to be copied to the PD. As mentioned in [18], the drawback of this approach is the increased vulnerability to noise, and the fact that for time varying systems, especially if they are additionally nonlinear, the post-inverse cannot be expected to be equivalent to the pre-inverse. The second method considered here, is a direct identification scheme based on the filtered-x LMS (FxLMS) [37], [38], modified for the identification of dispersive nonlinear systems. It was first proposed as nonlinear FxLMS (NFxLMS) in [39] employing truncated Volterra series. Recently, it has been revisited and generalised in [18]. In this work, we adopt the results in [18, pp. 127-128] for the Hammerstein PD adapted by the NFxLMS. This scheme additionally requires an estimate of the reference system, which is obtained by the algorithm in [14]. It is conjectured to be robust based on the results in [40].

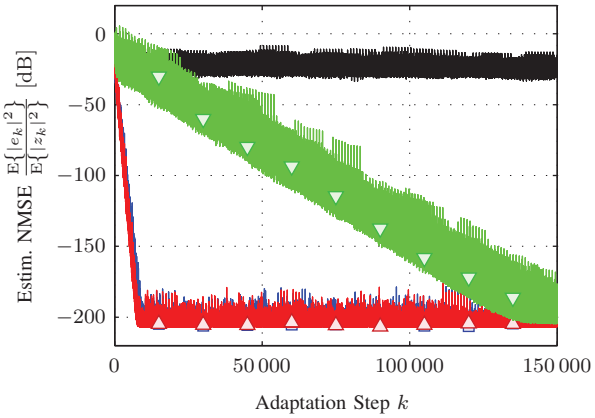

Fig. 5. Estimated NMSE (averaged over 10 000 simulations) of the PNLMS in (10) with $\mu_h = \frac{1}{2}$, used for adaptive identification of a static nonlinearity excited by the OFDM signal with variance $\sigma_x^2 \approx 0.02$ (legend, see Fig. 4).

For all of the DPD schemes used in the simulations and measurements, the occurring two-block models consist of the same blocks, of course with different parameters. The linear part is implemented as an FIR filter with five weights. The nonlinear part utilises the first five modified Chebyshev polynomials described in Sec. II-B. Additional simulations which are not presented here, have shown that the actual choice of the polynomials affects the quality of the pre-distortion only marginally, as long as the adaptation is performed long enough, such that the steady state can be reached. Of course, only if the adaptation does not stall at a local minimum.

A. Simulations

For the simulations, the reference system is represented by a Wiener model consisting of a 5-tap FIR filter followed by a Saleh model. The weights of the linear filter are generated randomly and normalised such that the l_2-norm of the weight vector equals one. However, previous to the normalisation, the stable invertibility of the filter is ensured by projecting any zero, which lies outside of the unit circle, to the inside. The Saleh model has the same parameters as described in Sec. II-B2. Its maximum excitation level is approximately 2dB above the 1-dB gain compression point [41]. The OFDM signal described in Sec. II-B, is used as (desired) transmission signal d_k. Fig. 6 depicts the power spectral densities (PSDs) of the PA output, for the case with:

- no DPD,
- a static PD (without linear filter), adapted by the IL,
- a Hammerstein PD adapted by the IL,
- a Hammerstein PD adapted by the NFxLMS,

and compares all of them to the adequately normalised PSD of the signal d_k. As expected, the Hammerstein PDs achieve a higher (approximately 5dB) suppression of the spectral regrowth compared to the static PD. While the Hammerstein PD identified by the IL shows a slightly better compensation of intermodulation products in the near vicinity of the transmission band, it is outperformed by the Hammerstein PD trained by the NFxLMS for more distant frequency ranges.

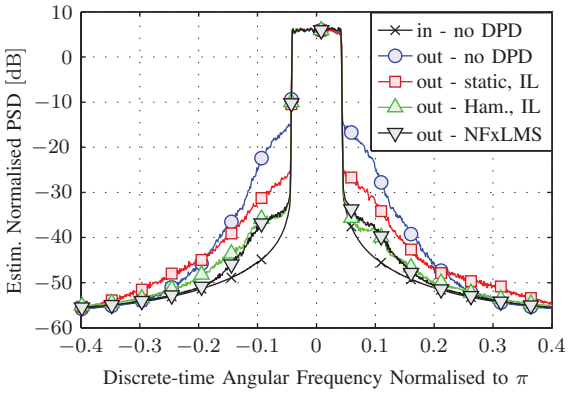

Fig. 6. Simulation results for different DPD schemes, using a reference model consisting of an FIR filter of length five followed by a Saleh model.

B. Measurements

The measurements are based on the same MATLAB® code like the simulations in Sec. III-A. However, referring to Fig. 1, the input signal x_k is not fed through a baseband Wiener model, but amplified by an actual commercially available PA (Mini-Circuits® ZVE-8G+). This PA has a frequency range of 2GHz–8GHz and a nominal output power of 1W at the 1-dB gain compression point. The conversion from the digital baseband to the transmission band at a centre frequency of $f_c = 2.1$GHz is done by the vector signal generator R&S® SMU200A. After attenuating the output signal of the PA by 30dB, it is converted down to digital baseband using the signal analyser R&S® FSQ26. Since the algorithms are implemented in MATLAB®, the measurements have to be performed in burst mode. Consequently, each measurement requires two phases. First, the desired transmit signal is transmitted without any DPD, i.e., $x_k = d_k$, and the attenuated output of the PA is captured, yielding a block of output samples z_k. Based on the blocks of x_k and z_k, the PD parameters are estimated. In the second phase, the desired signal d_k is first pre-distorted and then again transmitted. From the PA output, captured during this second phase, the PSDs depicted in Fig. 7 are estimated. To prevent a variation of the PA characteristics due to changes of the input power, the PA is permanently provided with an input signal. The required gain adjustment in the PD is set manually. Fig. 7 presents the results of the measurements. In contrast to the simulations in Sec. III-A, all DPD schemes achieve almost the same performance. Since the correct operation of the PDs is ensured by the simulations, this result indicates that the memory effects of the used PA are negligibly small for the bandwidth of the applied OFDM signal. It is assumed that with a PA imposing stronger memory effects, the results of the measurements would better conform to the simulation results.

IV. Conclusion

In the context of nonlinear system identification, parameter estimation can be improved substantially by the use of orthogonal polynomials. While for LS, their use leads to a numerically better conditioned matrix inversion, for gradient type algorithms, the convergence can be accelerated and the risk of stalling at a local minimum can be reduced. In the

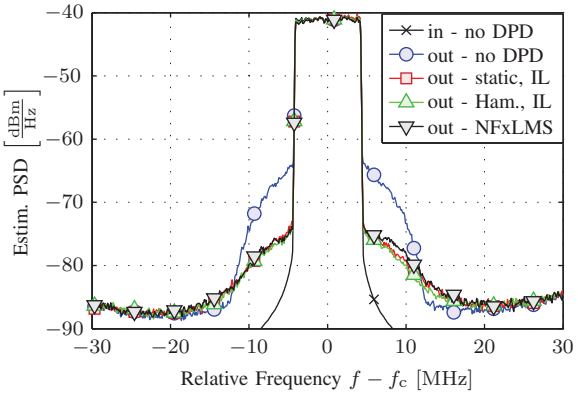

Fig. 7. Results for different DPD schemes, measured at $f_c = 2.1$GHz.

simulations, these improvements are observed for the complex uniform polynomials and for the modified Chebyshev polynomials. Although, the complex Gaussian polynomials show the expected low condition numbers for LS, the effect on the gradient method was rather moderate. This may be caused by the fact that for the same number of parameters, the complex Gaussian polynomials have nearly twice the degree compared to the other polynomials. For the considered PDs, as long as convergence to the steady state was ensured, the applied polynomials only affect the convergence rate, the influence on the performance of the PD is negligible. In the simulations, the Hammerstein PD trained by the NFxLMS seems to outperform the other DPD schemes slightly. The measurements lead to almost the same results for all DPD methods.

REFERENCES

[1] 3GPP, Technical Specification Group Radio Access Network, "E-UTRA and E-UTRAN, Overall description, Stage 2, (Release 10)," Sep. 2010.

[2] S. Shepherd, J. Orriss, and S. Barton, "Asymptotic limits in peak envelope power reduction by redundant coding in orthogonal frequency-division multiplex modulation," *IEEE Trans. Commun.*, vol. 46, no. 1, pp. 5–10, Jan. 1998.

[3] H. Karl (ed.), "An overview of energy efficiency techniques for mobile communication systems," Telecommunication Networks Group, Technische Universität Berlin, Technical Report TKN-03-017, Sep. 2003.

[4] M. ODroma, J. Portilla, E. Bertran, S. Donati, T. Brazil, M. Rupp, and R. Quay, "Linearisation issues in microwave amplifiers," in *Proc. European Microwave Week, 12th GAAS*, 2004, pp. 199–202.

[5] P. Kenington, *High linearity RF amplifier design*. Boston (MA), USA: Artech House, Inc., 2000.

[6] S. C. Cripps, *Advanced techniques in RF power amplifier design*. Boston (MA), USA: Artech House, Inc., 2002.

[7] K. Muhonen, M. Kavehard, and R. Krishnamoorthy, "Look-up table techniques for adaptive digital predistortion: A development and comparison," *IEEE Trans. Veh. Technol.*, vol. 49, no. 5, pp. 1995–2002, Sep. 2000.

[8] A. Zhu and T. Brazil, "An adaptive Volterra predistorter for the linearization of RF high power amplifiers," in *IEEE MTT-S International Microwave Symposium Digest*, vol. 1, 2002, pp. 461–464.

[9] J. Vuolevi, T. Rahkonen, and J. Manninen, "Measurement technique for characterizing memory effects in RF power amplifiers," *IEEE Trans. Microw. Theory Tech.*, vol. 49, no. 8, pp. 1383–1389, Aug. 2001.

[10] M. Schetzen, *The Volterra and Wiener theories of nonlinear systems*. New York (NY), USA: John Wiley & and Sons, Inc., 1980.

[11] J. Kim and K. Konstantinou, "Digital predistortion of wideband signals based on power amplifier model with memory," *Electronics Letters*, vol. 37, no. 23, pp. 1417–1418, Nov. 2001.

[12] M. Cheong, E. Aschbacher, P. Brunmayr, M. Rupp, and T. Laakso, "Comparison and experimental verification of two low-complexity digital predistortion methods," in *Rec. 39th ACSSC*, Pacific Grove (CA), USA, Nov. 2005, pp. 432–436.

[13] D. Schreurs, M. O'Droma, A. Goacher, and M. Gadringer, Eds., *RF power amplifier behavioural modeling*. Cambridge, UK: Cambridge University Press, 2009.

[14] R. Dallinger and M. Rupp, "Stability analysis of an adaptive Wiener structure," in *Proc. IEEE ICASSP*, Dallas (TX), USA, Mar. 2010, pp. 3718–3721.

[15] H. Ruotsalainen, "Investigation of orthogonal basis expansions for adaptive Wiener models," Master's thesis, Aalto University School of Science and Technology, Espoo, Finland, Apr. 2010.

[16] R. Raich, H. Qian, and G. Zhou, "Orthogonal polynomials for power amplifier modeling and predistorter design," *IEEE Trans. Veh. Technol.*, vol. 53, no. 5, pp. 1468–1479, Sep. 2004.

[17] R. Raich and G. Zhou, "Orthogonal polynomials for complex Gaussian processes," *IEEE Trans. Signal Process.*, vol. 52, no. 10, pp. 2788–2797, Oct. 2004.

[18] D. Zhou and V. DeBrunner, "Novel adaptive nonlinear predistorters based on the direct learning algorithm," *IEEE Trans. Signal Process.*, vol. 55, no. 1, pp. 120–133, Jan. 2007.

[19] J. Figueroa, J. Cousseau, and R. de Figueiredo, "A simplicial canonical piecewise linear adaptive filter," *Circuits, Systems, and Signal Processing*, vol. 23, no. 5, pp. 365–386, 2004.

[20] J. Cousseau, J. Figueroa, S. Werner, and T. Laakso, "Efficient nonlinear Wiener model identification using a complex-valued simplicial canonical piecewise linear filter," *IEEE Trans. Signal Process.*, vol. 55, no. 5, pp. 1780–1792, May 2007.

[21] T. Moon and W. Stirling, *Mathematical methods and algorithms for signal processing*. Upper Saddle River (NJ), USA: Prentice-Hall, 2000.

[22] S. Benedetto and E. Biglieri, *Principles of digital transmission*. New York (NY), USA: Kluwer Academic/Plenum Publishers, 1999.

[23] L. Ding and G. Zhou, "Effects of even-order nonlinear terms on power amplifier modeling and predistortion linearization," *IEEE Trans. Veh. Technol.*, vol. 53, no. 1, pp. 156–162, Jan. 2004.

[24] A. Sayed, *Fundamentals of adaptive filtering*. Hoboken (NJ), USA: John Wiley & Sons, Inc., 2003.

[25] M. Abramowitz, *Handbook of mathematical functions*, 10th ed. Washington (DC), USA: National Bureau of Standards, 1972.

[26] J. G. Proakis, *Digital communications*, 3rd ed. New York (NY), USA: McGraw-Hill, Inc., 1995.

[27] C. Mehlführer, M. Wrulich, J. Colom Ikuno, D. Bosanska, and M. Rupp, "Simulating the long term evolution physical layer," in *Proc. EUSIPCO*, Glasgow, Scotland, 2009.

[28] LTE physical layer simulator.
[Online]. Available: http://www.nt.tuwien.ac.at/ltesimulator

[29] S. Wei, D. Goeckel, and P. Kelly, "Convergence of the complex envelope of bandlimited OFDM signals," *IEEE Trans. Inf. Theory*, vol. 56, no. 10, pp. 4893–4904, Oct. 2010.

[30] A. Saleh, "Frequency-independent and frequency-dependent nonlinear models of TWT amplifiers," *IEEE Trans. Commun.*, vol. 29, no. 11, pp. 1715–1720, Nov. 1981.

[31] D. Duttweiler, "Proportionate normalized least-mean-squares adaptation in echo cancelers," *IEEE Trans. Speech Audio Process.*, vol. 8, no. 5, pp. 508–518, Sep. 2000.

[32] C. Clark, G. Chrisikos, M. Muha, A. Moulthrop, and C. Silva, "Time-domain envelope measurement technique with application to wideband power amplifier modeling," *IEEE Trans. Microw. Theory Tech.*, vol. 46, no. 12, pp. 2531–2540, Dec. 1998.

[33] L. Ding, R. Raich, and G. Zhou, "A Hammerstein predistortion linearization design based on the indirect learning architecture," in *Proc. IEEE ICASSP*, vol. 3, Orlando (FL), USA, May 2002, pp. 2689–2692.

[34] E. Aschbacher, M. Steinmair, and M. Rupp, "Iterative linearization methods suited for digital pre-distortion of power amplifiers," in *Rec. 38th ACSSC*, vol. 2, Nov. 2004, pp. 2198 – 2202.

[35] B. Widrow, J. McCool, and B. Medoff, "Adaptive control by inverse modeling," in *Rec. 12th ACSSC*, Nov. 1978, pp. 90–94.

[36] C. Eun and E. Powers, "A new Volterra predistorter based on the indirect learning architecture," *IEEE Trans. Signal Process.*, vol. 45, no. 1, pp. 223–227, Jan. 1997.

[37] B. Widrow, D. Shur, and S. Shaffer, "On adaptive inverse control," in *Rec. 15th ACSSC*, Nov. 1981, pp. 185–189.

[38] B. Widrow and S. Stearns, *Adaptive signal processing*. Englewood Cliffs (NJ), USA: Prentice-Hall, 1985.

[39] Y. Lim, Y. Cho, I. Cha, and D. Youn, "An adaptive nonlinear prefilter for compensation of distortion in nonlinear systems," *IEEE Trans. Signal Process.*, vol. 46, no. 6, pp. 1726–1730, Jun. 1998.

[40] R. Dallinger and M. Rupp, "A strict stability limit for adaptive gradient type algorithms," in *Rec. 42nd ACSSC*, Pacific Grove (CA), USA, Nov. 2009, pp. 1370–1374.

[41] G. Gonzalez, *Microwave transistor amplifiers*, 2nd ed. Upper Saddle River (NJ), USA: Prentice-Hall, 1997.