Finding Structural Information of RF Power Amplifiers using an Orthogonal Non-Parametric Kernel Smoothing Estimator

Efrain Zenteno, Student Member, IEEE, Zain Ahmed Khan, Student Member, IEEE, Magnus Isaksson, Senior Member, IEEE, and Peter Händel, Senior Member, IEEE

Abstract—A non-parametric technique to model the behavior of power amplifiers is presented. The proposed technique relies on the principles of density estimation using the kernel method and it is suited for the application of power amplifier modeling. The proposed methodology transforms the input domain to an orthogonal memory domain. In such a domain, non-parametric static functions are discovered using the kernel estimator. These orthogonal non-parametric functions can be fitted with any desired mathematical structure, e.g., facilitating its implementation. Furthermore, due to the orthogonality, the non-parametric functions can be analyzed and discarded individually, which eases pruning basis functions and trading off complexity and performance. The results show that the methodology can be employed to model power amplifiers yielding error performance similarly to state-of-art parametric models. Furthermore, a parameter efficient model structure with 6 coefficients was derived for a Doherty PA, significantly reducing the deployment computational complexity. Finally, the methodology can be as well exploited in digital linearization techniques.

Index Terms—Power amplifier, non-parametric model, kernel, basis functions, power amplifier linearization, Digital pre-distortion.

I. INTRODUCTION

Energy efficient power amplifiers (PAs) in wireless networks usually behave in a nonlinear fashion, producing significant nonlinear distortions that degrade the network performance. This raises the need of suitable behavioral models for PAs that provides simpler descriptions of nonlinear distortion mechanisms and tools for mitigation of these effects, such as, the digital pre-distortion (DPD) techniques [1].

Historically, behavioral models for PAs have been derived using the Volterra series [2]. The disadvantage of Volterra series is that it involves a large number of parameters which hinders of its practical implementation. Pruning Volterra series has been actively researched to provide low-complexity and high performance behavioral models to mitigate PA nonlinear distortions [3], [4], [5]. Although, pruning Volterra series has produced useful empirical models, these pruned models are general structures for smaller classes of nonlinear systems. This requires engineers to test different pruned model structures, and further select the nonlinear order and memory depths to meet certain performance requirements with a level of complexity that depends on application constraints. Hence, for a specific PA, trimming the pruned Volterra models may produce even lower complexity with the desired error performance [3], [5]. This raises the question whether there may exist techniques to give structural knowledge of a specific PA which in turn can be used to construct simpler model structures with a required error model performance. This paper presents a technique of this class.

Trimmed model structures of reduced complexity can be also obtained using sparse estimation techniques [6], [7]. However, sparse estimation techniques are usually computationally demanding and require choosing an initial model to be reduced. On one hand, a general model is desired as initial set preserving modeling properties. But, this involves a large set which worsens the complexity of the technique. On the other hand, starting from a small class of model structures, alleviating complexity, renders results that are dependent on this initial choice.

This paper presents a non-parametric method to discover PA structural information. Thus, it does not assume any a priori model structure for the PA. The proposed method considers static and dynamic distortion effects, and provides a tool for analyzing the PA transfer function. In particular, the tool can be used to tailor parametric models of simpler form. Thus, it effectively reduces the computational resources (complexity) in the model. For PA modeling other non-parametric techniques use statistical functions, as the cumulative distribution functions (CDF) [8], [9] and higher order statistics [10]. However, [8], [9], [10] consider solely memoryless distortion effects, and hence, they are ineffective to characterize or compensate PA distortion caused by memory effects.

The proposed technique is based on non-parametric density estimation [11] referred to as the kernel smoothing estimator or kernel method [12]. Compared to polynomial based PA models, the kernel method can estimate nonlinear functions of high nonlinear order without numerical problems [13]. Further, the kernel method uses window averages which are less computationally demanding than matrix (pseudo) inversions required in the least square methods. Finally, the kernel estimator has strong statistical properties: asymptotically convergent [14], optimal estimator in the square error sense given a limited
number of samples and robust to noise sources \cite{15}. All these properties make of the kernel method a suitable candidate for modeling PAs.

The work reported in this paper reviews the modeling methodology presented in \cite{16} and make the adaptations necessary for the PA set-up, characterized by band-limited complex-valued base-band signals. With the adaptations proposed in \cite{17} and our previous study \cite{13}, we obtain a methodology and method suitable for PA modeling. In contrast to traditional modeling PA techniques, the work reported here transforms the input sample domain to an orthogonal domain where the model structure is obtained using the kernel method. The orthogonal domain eases the analysis of the PA transfer function; allowing adding or removing basis functions to tradeoff between complexity and performance. This result can be transferred to the original sample domain reversing the orthogonalization process (linear combination) obtaining model structures that are comparable in performance with the traditional modeling PAs.

**II. PA MODELING**

**A. PA model**

Let \( u(n) \) and \( y(n) \) denote the \( n \)-th complex-valued sample of the base-band signals corresponding to the input and output of a PA, respectively. The PA nonlinear transfer function is approximated by \( g(\cdot) \)

\[
y(n) = \sum_{m_1} f_{m_1}(u(n - m_1)) + \sum_{m_1,m_2} f_{m_1,m_2}(u(n - m_1), u(n - m_2)) + \ldots + \sum_{m_1,m_2,\ldots,m_p} f_{m_1,\ldots,m_p}(u(n - m_1), \ldots, u(n - m_p)),
\]

where \( f_{m_1}(\cdot), f_{m_1,m_2}(\cdot, \cdot) \) and \( f_{m_1,\ldots,m_p}(\cdot, \ldots, \cdot) \) are nonlinear static functions, whose domain dimension is 1, 2 and \( p \), respectively. The summations go up \( M \), subject to \( 0 \geq m_1 > m_2 > \ldots > m_p \geq M \), where \( M \) is the maximum memory depth considered.

Volterra series is a special form of (1), which can be obtained when \( f_{m_1}(\cdot), f_{m_1,m_2}(\cdot, \cdot), \) and \( f_{m_1,\ldots,m_p}(\cdot, \ldots, \cdot) \) are defined as the scaled product of their arguments. The static functions in (1) represent high nonlinear orders of the Volterra series. In particular, high nonlinear orders coupled to different memory depths, which is the cause of the explosion in the number of parameters in the Volterra series. Despite of the different features of (1) compared to Volterra series, both suffer from high dimensionality. Considering \( 0 \geq m_1 > m_2 > \ldots > m_p \geq M \), the system in (1) has a total number of additive functions, \( \sum_{d=1}^{p} \binom{M+1}{d} = \sum_{d=1}^{p} \frac{(M+1)!}{d!(M+1-d)!} \), with \( p \) being the highest dimension of the functions in (1) and \( ! \) denoting the factorial operator. The high dimensionality increases the computational complexity in the identification and deployment for the models. Thus, we analyze the relationship (1) and study possible simplifications of it suitable for PA modeling.

![Fig. 1. Illustration of estimation of the function \( g(\cdot) \) at fixed grid \( x_i \) through a triangular kernel \( \varphi(\cdot) \). The value \( \hat{g}(x_i) \) is obtained as the weighted average of the output data \( z \) through the kernel \( \varphi(\cdot) \).](image)

**B. Kernel method brief**

The kernel estimator is briefly reviewed for the estimation of a static nonlinear input output relation \cite{12}. Consider the set of real-valued input data \( \{ x(n) \}_{n=0}^{N-1} \) passed through an unknown static nonlinear function \( g(\cdot) \) resulting in the output \( \{ z(n) \}_{n=0}^{N-1} \), that is \( z(n) = g(x(n)) \) for \( n = 0, \ldots, N-1 \). Then, the static function \( g(\cdot) \) can be estimated at a scalar point \( x_i \) as illustrated in Fig. 1 by the kernel (window) average \cite{12},

\[
\hat{g}(x_i) = \frac{1}{N-1} \sum_{n=0}^{N-1} \frac{\varphi(\frac{x(n)-x_i}{\delta})}{\int_{-\infty}^{\infty} \varphi(x) dx} z(n), \tag{2}
\]

where the grid of points \( x_i \) for \( i = 1, \ldots, T \) span the amplitude support of \( x(n) \), \( \varphi(\cdot) \) is the kernel with aperture \( \delta \). Here, the triangular kernel is preferred since it is the minimum mean square error estimator in a limited number of samples, and simultaneously is robust against noise \cite{15}, that is,

\[
\varphi(x) = \begin{cases} 1 - |x| & \text{if } |x| \leq 1 \\ 0 & \text{if } |x| > 1 \end{cases} \tag{3}
\]

with \( |\cdot| \) denoting the absolute value. Equation (2) is evaluated only if the denominator is different than zero. In this paper, a linear interpolation between the two nearest neighbors is employed to compute \( \hat{g}(\cdot) \) for an arbitrary input within the amplitude support of \( x(n) \).

The kernel method is not directly suited to mimic PA behavior. First, PA measurements show a significant correlation between different samples of the input signal \( u(n) \). The sample correlation makes the output of the functions \( f_{m_1}(\cdot), f_{m_1,m_2}(\cdot, \cdot) \) and \( f_{m_1,\ldots,m_p}(\cdot, \ldots, \cdot) \) jointly correlated. Thus, their estimation needs to account for all of them simultaneously.

Secondly, the kernel method is intended for real-value data. For complex-valued data widely available within PA instrumentation, the static functions in (1) are functions of complex-valued inputs and outputs \cite{13}. This turns to large computational and storage requirements for the method. In the following, a method tackling these drawbacks is outlined.
and discussed with simplifications (complexity reductions) of (1) suitable for PA modeling.

C. Removing correlation by orthogonalization

The measured PA input signal is band-limited and digitized with oversampling. This causes correlation between the samples of \( u(n) \). However, removing such a correlation in a band-limited signal can be viewed as orthogonalizing it \([13]\), which can be efficiently computed using the Gram-Schmidt (GS) process.

According to (1), the signal set to be orthogonalized lies in the space:

\[
\mathcal{U} = \{ u(n), \ldots, u(n-M) \}.
\]  

(4)

The GS process yields an orthogonal set \( \mathcal{U} = \{ \tilde{u}(n), \ldots, \tilde{u}(n-M) \} \), where \( \tilde{u}(n) = u(n) \) followed by an iterative process for \( k = 1, \ldots, M \),

\[
\tilde{u}(n-k) = u(n-k) - \sum_{\ell=0}^{k-1} P_{k,\ell} \tilde{u}(n-\ell).
\]  

(5)

The scalar \( P_{k,\ell} \) is a projection of the signal \( \tilde{u}(n-\ell) \) over \( u(n-k) \) defined by,

\[
P_{k,\ell} = \sum_n u^*(n-k) \tilde{u}(n-\ell),
\]  

(6)

where \( * \) denotes the complex conjugate operator. Note that the GS process is a linear combination, and thus, it can be reversed without any loss of information. Assuming \( u(n) \) being wide-sense stationary stochastic process with auto-correlation \( r_u(k) \), we note that the projections can be \( a \ priori \) calculated. This leads to a computationally preferable method. E.g., a rectangular shaped power spectral density of a Long Term Evolution (LTE) signal yields a sinc-shaped auto-correlation function.

D. Real-valued PA input signal

The distortion produced by PAs operating within wireless networks can be regarded as amplitude dependent \([4], [13]\). Thus, considering solely the amplitude of the signals in the orthogonal input set \( \mathcal{U} \) yields the set,

\[
\mathcal{U} = \{ |\tilde{u}(n)|, \ldots, |\tilde{u}(n-M)| \}.
\]  

(7)

which will be the input to our kernel estimator, e.g., \( x(n) = |\tilde{u}(n)| \). To compensate for the phase contribution, the output signal \( y(n) \) is transformed by,

\[
z(n) = y(n) e^{-j\varphi(n)}.
\]  

(8)

By applying the GS process to the input set \( \mathcal{U} \) followed by the real-value transformation, the system (1) turns in,

\[
z(n) = \sum_{m_1} g_{m_1}(x(n-m_1)) + \\
+ \sum_{m_1, m_2} g_{m_1,m_2}(x(n-m_1), x(n-m_2)) + \\
+ \sum_{m_1, \ldots, m_p} g_{m_1,m_2,\ldots,m_p}(x(n-m_1), \ldots, x(n-m_p)),
\]  

(9)

with complex-valued \( g_{m_1}(\cdot) \), \( g_{m_1,m_2}(\cdot, \cdot) \) and \( g_{m_1,m_2,\ldots,m_p}(\cdot, \ldots, \cdot) \) as the orthogonal counterparts of the functions in (1) but with real-valued arguments. This reduces the estimation dimension required in the kernel method. The system (9) has the similar features to model nonlinear behavior as (1). However, in contrast to (1), it has orthogonal basis functions. Thus, their estimation can be made individually and each basis contribution can be analyzed separately.

E. Complexity reduction of (1)

Despite using real-valued input signals, the complexity of the estimation of a multi-variable function in (9) is still large. E.g., a \( p \)-th variable function estimated at \( T \) points for each variable gives a total of \( T^p \) estimation points. Thus, the memory requirements and data manipulation grows exponentially with \( p \) leading to the well-known curse of dimensionality problem. Aiming to alleviate this, a \( p \)-variable function \( g_{m_1,\ldots,m_p}(\cdot, \ldots, \cdot) \) is approximated as a sum of single variable functions,

\[
g_{m_1,\ldots,m_p}(x(n-m_1), \ldots, x(n-m_p)) \approx \sum_{k=1}^{p} h_{mk}(x(n-m_k)) \prod_{d=1, d\neq k}^{p} x(n-m_d). \hspace{1cm} (10)
\]

Thus, the single-variable functions \( h_{mk}(\cdot) \) can be estimated using (2). In PA modeling, (10) has been motivated from a physical \([19]\) and a signal processing perspective \([17]\). Note that the new single-variables \( x(n-m_k) \prod_{d=1, d\neq k}^{p} x(n-m_d) \) can similarly be considered in the non-orthogonal domain \( \mathcal{U} \) by augmenting it as,

\[
\mathcal{U}' = \left[ \mathcal{U}, \{ u(n-m_k) \prod_{d=1, d\neq k}^{p} u(n-m_d) \}_{m_k=0}^{M} \right].
\]  

(11)

for \( p = 2, \ldots, M \). The orthogonalization of this data set is performed using the GS procedure.

F. Summary and Implementation

Consider the data set of complex-valued input and output measurements \( \{ u(n) \} \) and \( \{ y(n) \} \), respectively. The non-parametric modeling approach begins by creating the input space \( \mathcal{U}' \) as indicated by (11) for the chosen maximum memory depth \( M \). The implementation of the method can be made storing \( \mathcal{U}' \) as a matrix which columns are the basis in (11). This matrix is column-wise orthogonalized with the GS process yielding an orthogonal matrix. Only magnitude entries of the orthogonal matrix are retained according to (7) rendering a magnitude matrix. Finally, each column of the magnitude matrix is used as a domain to estimate single-variable functions with the kernel method in (2). The results show that significant contribution to the model output remains within few single-variable functions. Thus, due to the orthogonality, the non-contributing functions can be eliminated from the model structure (remove the corresponding columns) while retaining the obtained model performance.
III. EXPERIMENTAL

A. Measurement setup

The measurement setup comprises of an R&S SMU 200A vector signal generator used to excite the PA. The PA output is measured using a wide band down converter and a high performance analog-to-digital converter (ADC) with 14 bits resolution and a sampling rate of 400 MHz. The amplifier being tested is the MRF8S21120HS Doherty amplifier with 14 dB linear gain, an operation frequency in the 2.1-2.2 GHz band and rated at 46 dBm output power and operated at around 3 dB of compression.

Two independent excitation signals are generated with bandwidths of 12 and 24 MHz. These excitations are noise-like signals with peak-to-average power ratio of 11.2 and 11.4 dB, respectively. The excitations were created in a PC using 10\(^5\) complex-valued samples uploaded to the generator and up-converted to 2.14 GHz exciting the PA. The measurements consist on 10\(^5\) complex-valued samples for the input and the output of the amplifier with post-processing time and phase delay compensation. The non-parametric structure is obtained using 10\% (estimation phase) of the measured data while the remaining 90\% (validation phase) is used to evaluate the modeling error.

B. Results

1) User-defined parameters: In the proposed method, the number of grid points \(T\) and the kernel aperture \(\delta\) are user-defined parameters. Fig. 2 shows the normalized mean square error (NMSE) contours over both \(\delta\) and \(T\) in a linearly spaced grid. The kernel aperture \(\delta\) is shown as the percentage of the span of the input signal.

The number of grid points \(T\) sets the resolution of the static function and the kernel aperture \(\delta\) sets the size of input neighborhood to perform the average (estimation) (c.f. to Fig. 1). Thus, a large value of \(T\) and small \(\delta\) are desired for accurate estimation. However, for a fixed number \(N\) of measurement samples, reducing \(\delta\) may degrade the performance since the number of measurements samples in each kernel function reduces and hence its average (estimate) has increased variance (less reliability). That is the reason for the loss in performance for small values of \(\delta\) in Fig. 2. Since \(T\) is the number of entries to be stored, \(T\) can be chosen depending on the available memory resources and, as a rule of thumb, the kernel aperture can be set \(\delta = 1/T\) avoiding performance detrimental effects cf. Fig. 2. The choice \(\delta = 1/T\) has the advantage of an efficient use of all training data for estimating the non-parametric model.

2) Modeling performance: Using \(T = 70\) and \(\delta = 1/70\), a non-parametric model of the PA is obtained for the two input signals under consideration. Table I shows the individual contribution of the basis functions to the NMSE and to the adjacent channel error power ratio (ACEPR) for these two signals. As seen in Table I the function \(\hat{g}_0(x)\) contributes with -38.3 and -32.5 dB of the total NMSE in the 12 and 24 MHz signals, respectively. This function is the largest model contributor as it captures linear and nonlinear static effects.

In the 12 MHz signal, the NMSE is dominated by the contribution of \(\hat{g}_0(x)\) and \(\hat{g}_1(x)\) providing a combined NMSE of -43.3 dB. However, in the 24 MHz input signal, the function \(\hat{g}_2(x)\) increases its contribution to the NMSE from -0.3 dB to -2.3 dB revealing the impact of memory effects caused by the increase in signal bandwidth. Notice that, the contribution from the 2 and 3-variable functions is negligible in the 12 MHz case, while \(\hat{g}_{0,1}(\cdot)\) rises its contribution significantly when increasing the signal band from 12 to 24 MHz from -0.1 dB to -0.5 dB. In terms of the ACEPR, it is only \(\hat{g}_0(x)\) that provides a significant contribution for both signal bandwidths. The static nonlinear distortion is modeled by \(\hat{g}_0(\cdot)\) while the linear and nonlinear dynamics are described by \(\hat{g}_1(\cdot)\) and \(\hat{g}_2(\cdot)\). Since these functions are arbitrary, the nonparametric structure can

---

**TABLE I**

**INDIVIDUAL CONTRIBUTION OF THE FUNCTIONS FOR TWO DIFFERENT SIGNAL BANDWIDTHS**

| Basis | 12 MHz (dB) | 24 MHz (dB) |
|-------|-------------|-------------|
| \(\hat{g}_0(\cdot)\) | -38.3 | -32.5 |
| \(\hat{g}_1(\cdot)\) | -5.0 | -6.7 |
| \(\hat{g}_2(\cdot)\) | -0.3 | -2.3 |
| \(\hat{g}_3(\cdot)\) | 0.0 | 0.0 |
| \(\hat{g}_{0,1}(\cdot)\) | -0.1 | -0.5 |
| \(\hat{g}_{0,2}(\cdot)\) | -0.1 | 0.0 |
| \(\hat{g}_{0,3}(\cdot)\) | 0.0 | 0.0 |
| \(\hat{g}_{1,2}(\cdot)\) | 0.0 | 0.0 |
| \(\hat{g}_{1,3}(\cdot)\) | 0.0 | 0.0 |
| \(\hat{g}_{2,3}(\cdot)\) | 0.0 | 0.0 |
| Total | -43.9 | -51.0 |
model memory effects coupled to strong nonlinearities, which is one of the causes of poor behavior in polynomial-based model methodologies.

The functions contributing more than -2 dB to the NMSE are shown in Fig. 3 for both signals bandwidths. Despite that the two signals were independently created and with different bandwidths, the estimated functions are similar to each other, which suggest that the method renders structural information of the modeled PA.

The power spectral density (PSD) of the input, output and model error evolution are plotted in Fig. 4 for the 24 MHz signal. The model was updated sequentially to include the first six basis functions of Table I. It is observed that the in-band error spectrum decreases with the addition of basis functions, while the out-of-band error is suppressed by the use of \( \hat{g}_0(x) \). These two observations are in accordance with Table I.

An advantage of the proposed method is that we can utilize the estimated basis functions (cf. Fig. 3) to build a parametric model the PA. These parametric models can be of any form, they can be chosen in order to ease implementation, identification or to maximize performance. We seek a parameter efficient representation using a polynomial family as an example. Thus, the function \( \hat{g}_0(x) \) is modeled with a seventh order polynomial and \( \hat{g}_1(\cdot) \) and \( \hat{g}_2(\cdot) \) with linear polynomials (cf. Fig. 3), that is,

\[
z(n) = \sum_{p=1}^{4} \gamma_p \tilde{u}(n)|\tilde{u}(n)|^{2(p-1)} + \gamma_2 \tilde{u}(n-1) + \gamma_6 \tilde{u}(n-2).
\]

Due to the orthogonality the rest of the functions in Table I can be discarded without affecting the model performance. Further, replacing the orthogonal variables for the linear combination of their non-orthogonal counterparts given in the GS process

\[
y(n) = \sum_{p=1}^{4} \alpha_p u(n)|u(n)|^{2(p-1)} + \alpha_5 u(n-1) + \alpha_6 u(n-2).
\]

This 6 parameter model \([\alpha_1, \ldots, \alpha_6]\) can be identified using linear regression techniques; commonly used in PA modeling.

Fig. 5 plots the NMSE performance against the complexity incurred by the feed forward model usage. The complexity is measured in floating point operations per second (FLOPS) [20]. Fig. 5 compares the proposed method to several parametric models as the static nonlinear, memory polynomial, generalized memory polynomial [4], Multi-LUT [21], Volterra [2], Kautz-Volterra [22] and non-parametric models such as, Histogram [8]. Different points correspond to several model settings (nonlinear order and memory depth) being tested.

Although, in general is possible to reduce the NMSE with an increasing complexity, these settings need to be chosen with care to produce optimal performance for the level of complexity incurred, as observed by the performance dispersion in Fig. 5. In fact, some model settings with increased model complexity degrade the NMSE performance as seen in Fig. 5 that is due to an unsuitable model chosen. E.g., a static model of high nonlinear order may have large complexity but it is still unable to model dynamic effects rendering limited performance. Similar arguments can be drawn for memory depths. Moreover, these detrimental effects have been discussed in previous studies [20]. The proposed kernel method has good performance/complexity when compared to state-of-art parametric models. Finally, the proposed method was used to construct a parameter-efficient structure (6-parameter model) which performs with the best NMSE for its reduced level of complexity as it was tailored specifically for the PA.

Fig. 3. Amplitude and phase (in degrees) of the estimated single-variable functions \( \hat{g}_0(\cdot) \), \( \hat{g}_1(\cdot) \), and \( \hat{g}_2(\cdot) \) for the 24 MHz signal (solid - blue) and \( \hat{g}_0(\cdot) \), \( \hat{g}_1(\cdot) \) for the 12 MHz (dashed - red) signal. The function \( \hat{g}_2(\cdot) \) not presented for the 12 MHz signal since its contribution is negligible.

Fig. 4. Power spectral density of the output and error model obtained by including sequentially the first 6 basis functions indicated in Table I in the nonparametric model for the 24 MHz signal.
IV. DIGITAL PRE-DISTORTION (DPD)

The proposed method is tested as a predistorter compensating for nonlinear distortions at the PA output. The non-parametric structure is obtained using the inverse learning architecture in which input and output are interchanged. To enhance efficiency, a clipping technique has been applied to the 24 MHz input signal reducing its PAPR from 11.4 to 8.8 dB. However, care must be exercised as clipping techniques introduce in-band and out-of-band errors.

From (9), the function $g_0(\cdot)$ of the DPD model has to be the inverse of the same function in the feed-forward model. For the rest of the functions in the DPD model, they have to be the negative of their counterparts in the feed-forward model (same amplitude with phase shifted by 180 degrees). This is depicted in Fig. 6 where the feed-forward and inverse (DPD) estimated non-parametric functions are plotted.

The PA operates at -25 dB of NMSE and -36 dB of ACPR, respectively, without DPD. The predistorted PA with the outlined method performs at a NMSE and ACPR of -42 dB and -49.5 dB, respectively, which shows its effectiveness to compensate nonlinear distortion.

V. CONCLUSIONS

A non-parametric method to model RF power amplifiers is presented. The method does not assume an a priori model structure of the PA. Thus, basis functions that describe its behavior are found out during the identification process, leading to the development of tailored parametric models. These tailored models can be fitted with any desired structure that eases its implementation. In particular, parameter-efficient models with low level of error can be obtained reducing the implementation and deployment computational costs.

The method presented is based on the kernel estimator which performs solely sample averages and hence does not suffer from numerical instabilities. Further, adaptive schemes of can be made using running averages which have low computational resources and feature real time implementations. The proposed methodology can lead to low computational resource implementation of look up tables (LUTs) for adaptive digital predistortion (linearization).

REFERENCES

[1] M. Isaksson, D. Wisell, and D. Rönnow, “A comparative analysis of behavioral models for RF power amplifiers,” IEEE Trans. Microw. Theory Tech., vol. 54, no. 1, pp. 348–359, Jan. 2006.
[2] M. Schetzen, The Volterra and Wiener Theories of Nonlinear Systems. New York: Wiley & Sons, 1980.
[3] Y.-J. Liu, J. Zhou, W. Chen, and B.-H. Zhou, “A robust augmented complexity-reduced generalized memory polynomial for wideband RF power amplifiers,” IEEE Trans. Ind. Electron., vol. 61, no. 5, pp. 2389–2401, May 2014.
[4] D. Morgan, Z. Ma, J. Kim, M. Zierdt, and J. Pastalan, “A generalized memory polynomial model for digital predistortion of RF power amplifiers,” IEEE Trans. Signal Process., vol. 54, no. 10, pp. 3852–3860, Oct. 2006.
[5] X. Yu and H. Jiang, “Digital predistortion using adaptive basis functions,” IEEE Trans. Circuits Syst. I, Reg. Papers, vol. 60, no. 12, pp. 3317–3327, Dec. 2013.
[6] J. Reina-Tosina, M. Allegue-Martínez, C. Crespo-Cadenas, C. Yu, and S. Cruces, “Behavioral modeling and predistortion of power amplifiers under sparsity hypothesis,” IEEE Trans. Microw. Theory Tech., vol. 63, no. 2, pp. 745–753, Feb. 2015.
[7] E. Zenteno, S. Amin, M. Isaksson, P. Händel, and D. Rönnow, “Combating the dimensionality of nonlinear MIMO amplifier predistortion by basis pursuit,” European Microwave Conf. (EuMC), pp. 833–836, Oct. 2014.
[8] D. Huang, X. Huang, and H. Leung, “Nonlinear compensation of high power amplifier distortion for communication using a histogram-based method,” IEEE Trans. Signal Process., vol. 54, no. 11, pp. 4343–4351, Nov. 2006.
[9] Z. Zhu, X. Huang, and M. Caron, “Theoretical and experimental studies of a probabilistic-based memoryless PA linearization technique,” Circuits, Systems, and Signal Processing, vol. 32, no. 6, pp. 3031–3057, 2013.
[10] Z. Zhu, X. Huang, M. Caron, and H. Leung, “A blind AM/PM estimation method for power amplifier linearization,” IEEE Signal Process. Lett., vol. 20, no. 11, pp. 1042–1045, Nov. 2013.
[11] M. Rosenblatt, “Remarks on some nonparametric estimates of a density function,” Ann. of Math. Stat., vol. 27, no. 3, pp. 832–837, 1956.
A. A. Georgiev, “Nonparametric system identification by kernel methods,” IEEE Trans. Autom. Control, vol. 29, no. 4, pp. 356–358, Apr 1984.

Z. Khan, E. Zenteno, M. Isaksson, and P. Händel, “Density estimation models for strong nonlinearities in RF power amplifiers,” in Asia Pacific Microw. Conf., Sendai, APMC, Dec 2014, pp.–.

V. Epanechnikov, “Non-parametric estimation of a multivariate probability density,” Theory of Probability & Its Applications, vol. 14, no. 1, pp. 153–158, 1969.

E.-W. Bai and Y. Liu, “Recursive direct weight optimization in nonlinear system identification: A minimal probability approach,” IEEE Trans. Autom. Control, vol. 52, no. 7, pp. 1218–1231, Jul. 2007.

E.-W. Bai, “Non-parametric nonlinear system identification: A data-driven orthogonal basis function approach,” IEEE Trans. Autom. Control, vol. 53, no. 11, pp. 2615–2626, Dec. 2008.

H. Jiang and P. A. Wilford, “Digital predistortion for power amplifiers using separable functions,” IEEE Trans. Signal Process., vol. 58, no. 8, pp. 4121–4130, Aug 2010.

S. J. Leon and W. Gander, “Gram-schmidt orthogonalization: 100 years and more,” Numerical Linear Algebra with Applications, vol. 20, no. 3, pp. 492 – 532, May 2013.

T. Cunha, E. Lima, and J. Pedro, “Validation and physical interpretation of the power-amplifier polar volterra model,” IEEE Trans. Microw. Theory Tech., vol. 58, no. 12, pp. 4012–4021, Dec. 2010.

A. Tehrani, H. Cao, S. Afsardoost, T. Eriksson, M. Isaksson, and C. Fager, “A comparative analysis of the complexity/accuracy tradeoff in power amplifier behavioral models,” IEEE Trans. Microw. Theory Tech., vol. 58, no. 6, pp. 1510–1520, Jun. 2010.

P. Gilabert, A. Cesari, G. Montoro, E. Bertran, and J.-M. Dilliac, “Multi-lookup table fpga implementation of an adaptive digital predistorter for linearizing RF power amplifiers with memory effects,” IEEE Trans. Microw. Theory Tech., vol. 56, no. 2, pp. 372–384, Feb 2008.

M. Isaksson and D. Rönnö, “A parameter-reduced Volterra model for dynamic RF power amplifier modeling based on orthonormal basis functions,” Int. J. RF and Microw. Comput.-Aided Eng., vol. 17, no. 6, pp. 542–551, 2007.

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, 1997.

T. Jiang and Y. Wu, “An overview: Peak-to-average power ratio reduction techniques for OFDM signals,” IEEE Trans. Broadcast., vol. 54, no. 2, pp. 257–268, Jun. 2008.

M. Isaksson (S’98-M’07-SM’12) received the M.Sc. degree in microwave engineering from the University of Gävle, Gävle, Sweden, in 2000, the Licentiate degree from Uppsala University, Uppsala, Sweden, in 2006, and the Ph.D. degree from the Royal Institute of Technology, Stockholm, Sweden, in 2007. In 2012, he was appointed Docent in Telecommunications in the Royal Institute of Technology (KTH), Stockholm, Sweden. During 1989-1999, he was with the Televerket, Sweden, working on communication products. Since 1999, he is with the Department of Electronics, Mathematics, and Natural Sciences at the University of Gävle, Gävle, Sweden, where he is currently a professor of electronics and department head. His research activities are in signal processing algorithms for radio-frequency measurements and characterization, modeling, and compensation of nonlinear microwave devices and systems. Dr. Isaksson is the author or co-author of many published peer-review journal articles, books, and conference proceedings in the area, and is currently Head of research within the fields of mathematics, and natural sciences at the University of Gävle, Gävle, Sweden.

Peter Händel (S’88-M’94-SM’98) received the M.Sc. degree in engineering physics and the Ph.D. degree in automatic control from Uppsala University, Uppsala, Sweden, in 1987 and 1993, respectively. In 1987 and 1988, he was with the Svedberg Laboratory, Uppsala University. From 1988 to 1993, he was with the Systems and Control Group, Uppsala University. In 1996, he was appointed as Docent at Uppsala University. From 1993 to 1997, he was with the Research and Development Division, Ericsson AB, Kista, Sweden. During 1996 - 1997, he was a visiting Scholar at the Signal Processing Laboratory, Tampere University of Technology, Tampere, Finland. In 1998, he was appointed as Docent at Tampere University of Technology. Since 1997, he has been with the Signal Processing Laboratory, Royal Institute of Technology, Stockholm, Sweden, where he is currently a Professor of signal processing. He has authored approximately 35 journal articles and 75 conference papers and is the holder of 11 patents. He has conducted research on the design and analysis of digital filters and adaptive filters, measurement and estimation theory (particularly, temporal frequency estimation), system identification, speech processing, and measurement methods for characterization of analogdigital converters and power amplifiers. Dr. Händel is a Registered Engineer (EUR ING). He is a member of the editorial board of the EURASIP Journal of Applied Signal Processing. He is a former President of the IEEE Finland Joint Signal Processing and Circuits and Systems Chapter. Currently, he is the President of the IEEE Sweden Signal Processing Chapter. He is an Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING.

Efrain Zenteno (S’10) received the B.S. degree from San Agustin University, Arequipa, Peru, in 2004 and the M.Sc. degree in electronics/telecommunications engineering from the University of Gävle, Gävle, Sweden, in 2008. From 2008 to 2010, he was with the Universidad Católica San Pablo, program of telecommunications engineering, Arequipa, Peru. He is currently with the Department of Electronics, Mathematics, and Natural Sciences, University of Gävle, and with the Department of Signal Processing, Royal Institute of Technology KTH, Stockholm, Sweden. His main interests are instrumentation, measurements, and signal processing algorithms.

Zain Ahmed Khan (S’13) is a PhD student at KTH and University of Gävle in the field of Signal Processing. Currently, He started his PhD in October 2013 and is working on the behavioral modeling of nonlinear systems and design of digital predistorters for RF power amplifiers. He completed his Masters Degree in Communication and Signal Processing from Technical University Ilmenau, Germany in September 2013. His Masters Thesis with the Electronic Measurements Lab at TU Ilmenau focused on the comparison between Beam Pattern Shaping and Channel Synthesis for MIMO Channel Sounding Data Analysis.