



Digital Predistortion for Multi-Antenna Transmitters Affected by Antenna Crosstalk

Downloaded from: <https://research.chalmers.se>, 2025-06-18 02:36 UTC

Citation for the original published paper (version of record):

Hausmair, K., Landin, P., Gustavsson, U. et al (2018). Digital Predistortion for Multi-Antenna Transmitters Affected by Antenna Crosstalk. IEEE Transactions on Microwave Theory and Techniques, 66(3): 1524-1535. <http://dx.doi.org/10.1109/TMTT.2017.2748948>

N.B. When citing this work, cite the original published paper.

© 2018 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, or reuse of any copyrighted component of this work in other works.

Digital Predistortion for Multi-Antenna Transmitters Affected by Antenna Crosstalk

Katharina Hausmair, Per N. Landin, Ulf Gustavsson, Christian Fager, Thomas Eriksson

Abstract—In this paper, a digital predistortion (DPD) technique for wideband multi-antenna transmitters is proposed. The proposed DPD compensates for the combined effects of power amplifier (PA) nonlinearity, antenna crosstalk and impedance mismatch. The proposed technique consists of a linear crosstalk and mismatch model block shared by all transmit paths, and a dual-input DPD block in every transmit path. By avoiding the use of multi-input DPD blocks in every transmit path, the complexity of the proposed technique is kept low and scales more favorably with the number of antennas than competing techniques. It is shown that all blocks can be identified from measurements of the PA output signals using least-squares estimation. Measurement results of a four-path transmitter are presented and used to evaluate the proposed DPD technique against existing techniques. The results show that the performance of the proposed DPD technique is similar to those of existing techniques, while the complexity is lower.

Index Terms—MIMO transmitter, antenna crosstalk, power amplifier linearization, digital predistortion

I. INTRODUCTION

Multi-antenna systems are an important part of modern and future wireless telecommunication standards such as LTE, Wi-Fi, and 5G [1]. In such systems, each transmit path has its own power amplifier (PA) and antenna element, as shown in Fig. 1. Large-scale multi-antenna systems like massive MIMO comprise up to several hundreds of transmit paths [2]. Therefore, integrated system designs are used where expensive and bulky components like isolators between PAs and antennas are avoided to reduce system complexity and cost. However, such systems are vulnerable to antenna crosstalk due to mutual coupling and antenna mismatches [3]. As a consequence, integrated multi-antenna transmitters typically suffer from nonlinear distortion due to the mixing of the antenna crosstalk and mismatch with the PA output, in addition to the nonlinear distortion caused by the behavior of the PAs [4], [5]. To avoid

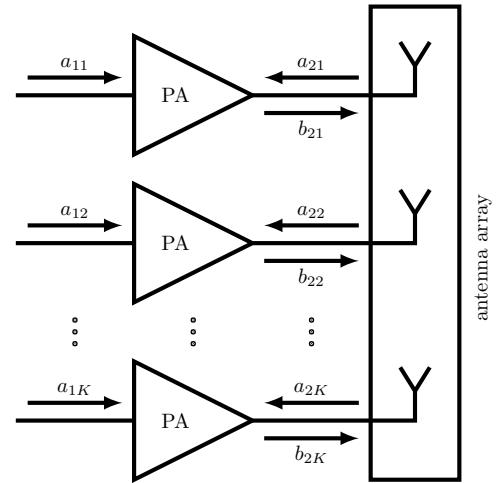


Fig. 1. Multi-antenna transmitter system model with K transmit paths. Each path consists of one PA connected to an antenna element. The antenna elements form the antenna array.

violating spectrum regulations and communication standard requirements, compensation techniques are needed to mitigate this distortion at the transmitter.

Digital predistortion (DPD) has been widely used to linearize PAs. Many algorithms have been proposed for single-input DPDs, e.g. [6], [7], which are designed for systems with only one transmit path. Single-input DPD can compensate for PA nonlinearity, but not for the effects caused by antenna crosstalk, and is therefore not suitable for multi-antenna systems [4], [5].

Several DPD techniques have been proposed to compensate for different types of nonlinear distortion in multi-antenna transmitters. The authors of [8], [9] and [10] propose a DPD technique that has a structure specifically designed to compensate for nonlinear distortion in systems where crosstalk is introduced before the PA. A similar technique is used in [11]. However, nonlinear effects of antenna crosstalk are not considered and cannot be compensated for using these DPD techniques. The crossover DPDs proposed in [12], [13] are designed to compensate for crosstalk before and after the PA. However, in those publications it is assumed that only crosstalk before the PAs is causing nonlinear effects, whereas the effects of crosstalk occurring after the PAs are assumed to be linear. Since the techniques do not consider necessary crossterms between signals of different transmit paths, they are not suitable for linearization of multi-antenna transmitters with antenna crosstalk. The authors of [14] include

This research has been carried out in GigaHertz Centre in a joint project financed by the Swedish Governmental Agency for Innovation Systems (VINNOVA), Chalmers University of Technology, Ericsson, Infineon Technologies Austria, Ampleon, National Instruments, Gotmic, and Saab. This project has received funding from the EMPIR programme co-financed by the Participating States and from the European Union's Horizon 2020 research and innovation programme

K. Hausmair and T. Eriksson are with the Department of Electrical Engineering, Communication Systems Group, Chalmers University of Technology, 41296 Gothenburg, Sweden (e-mail: {hausmair, thomase}@chalmers.se).

C. Fager is with the Department of Microtechnology and Nanoscience, Microwave Electronics Laboratory, Chalmers University of Technology, 41296 Gothenburg, Sweden (e-mail: christian.fager@chalmers.se).

P. N. Landin is with Ericsson Supply Kumla, Sweden (e-mail: per.landin@ericsson.com).

U. Gustavsson is with Ericsson Research Lindholmen, Sweden (e-mail: ulf.gustavsson@ericsson.com).

certain crossterms between different transmit paths in their augmented crossover DPD. However, also their system model does not consider mixing between the crosstalk and the PA output signals. Despite that, their approach can be applied to compensate for the combined effects of PA nonlinearity and antenna crosstalk in cases where the crosstalk can be considered small in power.

Also the authors of [15] use a system model without mixing of antenna crosstalk and PA output for their behavioral models and DPD of a two-path transmitter. Nevertheless, their DPD includes the necessary terms to compensate for the effects caused by any kind of antenna crosstalk. Another method that could be used to combat the nonlinear effects due to antenna crosstalk is presented in [16], where a method for joint nonlinearity and in-phase and quadrature imbalance in multi-antenna transmitters is proposed. However, these techniques do not scale well for systems with larger numbers of antennas, since they would require a multi-input DPD, i.e. a multi-variate polynomial or Volterra series, in every path of the transmitter. Due to complexity, such an approach is not feasible for emerging transmitters with many antennas. To reduce the complexity of such approaches, the authors of [17] propose a sparse estimation technique which can reduce the complexity but still requires the use of multi-input memory polynomial models.

In this manuscript, we propose a multi-antenna transmitter DPD technique that employs a completely different structure than existing techniques in order to reduce complexity. The basis for the proposed technique is a multi-antenna transmitter model first presented in [4], where dual-input PA models are combined with linear antenna array models. The proposed DPD technique consists of two main blocks: one linear block that models antenna crosstalk and mismatch and is shared by all paths of the transmitter, and a nonlinear dual-input DPD block in every transmit path, as shown in Fig. 2. Our solution is suitable for multi-antenna transmitters with any kind of crosstalk and mismatch at the PA outputs that can be described as a linear function of all transmit path outputs. The complexity of the antenna crosstalk and mismatch block increases linearly with the number of transmit paths, while the complexity of each dual-input DPD block is completely independent of the number of transmit paths. Hence, for transmitters with more than two paths, the complexity of the proposed DPD is lower compared to existing solutions. Just like existing solutions, the proposed DPD can be identified from measurements of the individual PA output signals using conventional linear least-squares estimation algorithms, and requires no other prior knowledge of the characteristics of the PAs or the antenna array. The proposed DPD technique is evaluated and compared to existing techniques in measurements of a four-path transmitter. To our knowledge, this is the first time that measurement results of a transmitter with more than two paths are presented for these kinds of DPDs.

The paper is organized as follows: In Section II we introduce the system model of a multi-antenna transmitter. In Section III, we present the proposed multi-antenna transmitter DPD including an identification procedure. The proposed technique is evaluated in measurements of a four-path transmitter. The

setup for the experiments is explained in Section IV, and the results are presented in Section V. Finally, we draw our conclusions in Section VI.

II. SYSTEM MODEL

In this section, the system model of the multi-antenna transmitter in equivalent discrete-time lowpass description is given. Note that, where applicable, the time dependency is omitted for better legibility, such that for example $a_{ik}(n)$ is written as a_{ik} . A block diagram of a multi-antenna transmitter with K transmit paths is shown in Fig. 1. Each transmit path consists of an RF PA connected to an antenna element. The antenna elements form the antenna array. All transmit paths operate in the same frequency band. The signal a_{1k} is the input to the PA of the k th path, and the signal b_{2k} is the PA output signal. Due to antenna crosstalk and mismatch, a signal a_{2k} is incident to the output of the PA. Each PA of the transmitter can be modeled as a dual-input system with one output. The crosstalk and mismatch signal a_{2k} is a function of the PA output signals of all paths, and the relation between a_{2k} and the output signals b_{2k} is determined by the characteristics of the antenna array. The system model of the multi-antenna transmitter can therefore be split in two parts: a crosstalk and mismatch model (CTMM), and a dual-input PA model [4], [5].

A. Crosstalk and Mismatch Model

The CTMM describes the crosstalk and mismatch signals a_{2k} as a function of the PA outputs b_{2k} . If the antennas are wideband compared to the signal bandwidth, a_{2k} can be described as a linear combination of the PA output signals of all transmit paths by [4], [5]

$$a_{2k} = \sum_{i=1}^K \lambda_{ki} b_{2i} = \mathbf{b}_2^T \boldsymbol{\lambda}_k \quad (1)$$

where $\mathbf{b}_2 = [b_{21}, \dots, b_{2K}]^T$, λ_{ki} are complex coefficients and $\boldsymbol{\lambda}_k = [\lambda_{k1}, \dots, \lambda_{kK}]^T$. The antenna array scattering parameters (S-parameters) can be used to describe the characteristics of an antenna array. The S-parameters measured at center frequency, i.e. the S-parameter matrix of the array, then correspond to a matrix $[\boldsymbol{\lambda}_1, \dots, \boldsymbol{\lambda}_K]$.

B. Dual-Input PA Model

The PA output signal b_{2k} of the k th PA is modeled as a function of the signals a_{1k} and a_{2k} by [5], [18]

$$b_{2k} = \sum_{p=0}^{(P-1)/2-1} \sum_{v=0}^p \sum_{u=0}^{p+1} \theta_{kpvu} a_{1k}^{p+1-u} a_{1k}^{*p-v} a_{2k}^u a_{2k}^{*v} \quad (2a)$$

$$= \left. \begin{aligned} &\sum_{p=0}^{(P-1)/2} \alpha_{kp} a_{1k}^{p+1} a_{1k}^{*p} \\ &+ \sum_{p=0}^{(P-1)/2} \beta_{kp} a_{1k}^p a_{1k}^{*p} a_{2k} \end{aligned} \right\} \quad (2b)$$

$$+ \left. \sum_{p=1}^{(P-1)/2} \gamma_{kp} a_{1k}^{p+1} a_{1k}^{*p-1} a_{2k}^* \right\} \quad (2c)$$

$$\left. \begin{aligned} & + \sum_{p=1}^{(P-1)/2} \sum_{v=0}^p \sum_{\substack{u=0 \\ u > 1-v}}^{p+1} \delta_{kpvu} \\ & \times a_{1k}^{p+1-u} a_{1k}^* a_{2k}^{p-v} a_{2k}^* \end{aligned} \right\} \quad (2d)$$

where θ_{kpvu} , α_{kp} , β_{kp} , γ_{kp} , δ_{kp} are complex coefficients. As can be seen, there are four types of basis functions, which all contain one more non-conjugate term than conjugate terms [7]: basis functions that depend only on a_{1k} in (2a), basis functions that depend on a_{1k} and linear terms of a_{2k} in (2b), basis functions that depend on a_{1k} and linear terms of a_{2k}^* in (2c), and basis functions that depend on a_{1k} and nonlinear terms of a_{2k} in (2d). If the crosstalk and mismatch signal a_{2k} can be considered relatively small in power, only linear terms of a_{2k} need to be considered in the dual-input PA model [19]. Then, all basis functions in (2d) become negligible and can be set to zero.

To make the dual-input PA model suitable for wideband signals, dynamic effects need to be considered [6]. This is done by introducing memory effects. In [15] and [18], several strategies for introducing memory effects in dual-input PA models are explained. Any of these strategies can be used in the models presented here. While the concept of introducing memory is rather simple, the equations become cumbersome to read. We therefore describe the memoryless equations in the main text here, and refer the reader to Appendix A, where we present the equations including memory effects.

In matrix form, the dual-input PA model can be written as

$$\begin{aligned} \mathbf{b}_{2k} &= \begin{bmatrix} \mathbf{G}^{(0)}(\mathbf{a}_{1k}) & \mathbf{G}^{(1)}(\mathbf{a}_{1k}, \mathbf{a}_{2k}) & \mathbf{G}^{(2)}(\mathbf{a}_{1k}, \mathbf{a}_{2k}) \\ & \mathbf{G}^{(3)}(\mathbf{a}_{1k}, \mathbf{a}_{2k}) \end{bmatrix} \begin{bmatrix} \boldsymbol{\alpha}_k^T & \boldsymbol{\beta}_k^T & \boldsymbol{\gamma}_k^T & \boldsymbol{\delta}_k^T \end{bmatrix}^T \\ &= \mathbf{G}(\mathbf{a}_{1k}, \mathbf{a}_{2k}) \boldsymbol{\theta}_k \end{aligned} \quad (3)$$

where \mathbf{a}_{1k} , \mathbf{a}_{2k} and \mathbf{b}_{2k} are vectors containing all time samples of the signals a_{1k} , a_{2k} and b_{2k} , e.g., $\mathbf{a}_{1k} = [a_{1k}(0), \dots, a_{1k}(N-1)]^T$ where N is the number of samples. The vectors $\boldsymbol{\alpha}_k$, $\boldsymbol{\beta}_k$, $\boldsymbol{\gamma}_k$, $\boldsymbol{\delta}_k$ and $\boldsymbol{\theta}_k$ contain the complex coefficients, e.g., $\boldsymbol{\alpha}_k = [\alpha_{k0}, \dots, \alpha_{k((P-1)/2-1)}]^T$. Furthermore, the matrix $\mathbf{G}^{(0)}(\mathbf{a}_{1k})$ contains the basis functions in (2a), $\mathbf{G}^{(1)}(\mathbf{a}_{1k}, \mathbf{a}_{2k})$ the basis functions in (2b), $\mathbf{G}^{(2)}(\mathbf{a}_{1k}, \mathbf{a}_{2k})$ the basis functions in (2c), $\mathbf{G}^{(3)}(\mathbf{a}_{1k}, \mathbf{a}_{2k})$ the basis functions in (2d), and $\mathbf{G}(\mathbf{a}_{1k}, \mathbf{a}_{2k})$ combines all basis functions in (2).

III. MULTI-ANTENNA TRANSMITTER DPD

We propose a DPD that consists of two main blocks: one linear CTMM block for the whole transmitter, and a dual-input DPD block in every transmit path. A block diagram of the proposed method is shown in Fig. 2. Each dual-input DPD is the inverse function of the respective dual-input PA, while the CTMM block emulates the behavior of the antenna array. One input of the k th dual-input DPD is the signal b_{dk} which is the desired output signal of the k th PA, i.e. in a perfectly linearised transmitter $b_{2k} = b_{dk}$. The second input to the k th dual-input DPD is an estimate \hat{a}_{2k} of the crosstalk and mismatch signal. The CTMM block creates the signals \hat{a}_{2k} from the signals b_{dk} . The output of the k th dual-input DPD, which is driving the k th PA, is the signal a_{1k} .

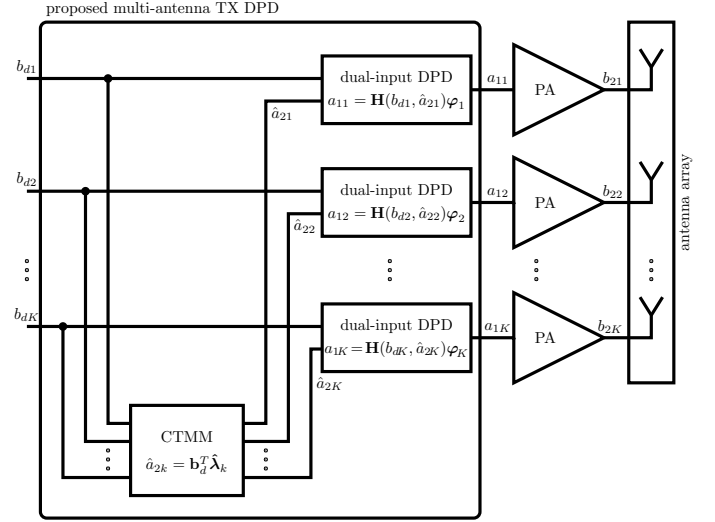


Fig. 2. Block diagram of a multi-antenna transmitter with the proposed DPD method. The method consists of two main blocks: one linear crosstalk and mismatch model (CTMM) block for the whole transmitter, and a dual-input DPD block in every transmit path.

The identification of the coefficients in the proposed technique is based on measurements of the individual PA output signals, and requires no other prior knowledge of the PA behavior or the characteristics of the antenna array. First, using the signals b_{dk} , which are known, the CTMM coefficients can be identified from measurements of the PA output signals b_{2k} . Then, with the signals b_{dk} and the output of the CTMM block \hat{a}_{2k} , the dual-input DPD coefficients can also be identified from measurements of the PA output signals b_{2k} . Note that for the identification of both the dual-input DPD and the CTMM coefficients, the input signals b_{dk} to the different transmit paths cannot be fully correlated. For the special case of multi-antenna systems with fully correlated signals like, e.g., beamforming systems, suitable training signals have to be used for the identification, which requires the transmission of user data to be interrupted. However, for applying the CTMM and the DPD, no such restrictions apply.

The CTMM block and dual-input DPD as well as the identification procedures are described below. Note that, without loss of generality, we assume amplifiers with unity gain to simplify the notation. When applying the proposed technique to PAs with non-unity gain, any conventional gain normalization concept can be used, e.g. [20], or [21].

A. Crosstalk and Mismatch Model

The CTMM produces the signals \hat{a}_{2k} by

$$\hat{a}_{2k} = \mathbf{b}_d^T \hat{\boldsymbol{\lambda}}_k \quad (4)$$

where $\hat{\boldsymbol{\lambda}}_k$ is a vector with the CTMM coefficients of the k th transmit path, which have to be identified. The vector $\mathbf{b}_d = [b_{d1}, \dots, b_{dK}]^T$ contains the input signals b_{dk} to the transmitter, which are known. Also known are the PA output signals b_{2k} , which are obtained by measurements. Because no DPD is applied for the CTMM coefficient identification procedure, $a_{1k} = b_{dk}$.

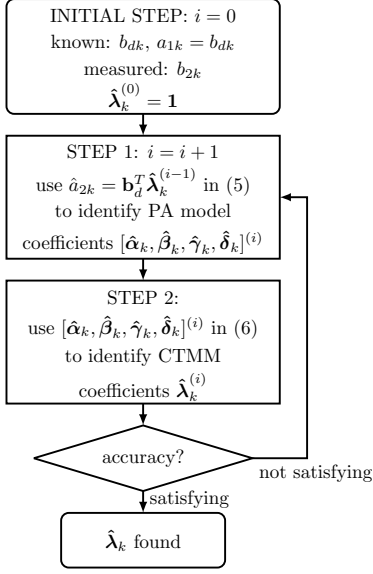


Fig. 3. Flowchart for the identification of the CTMM coefficients of the k th transmit path.

The CTMM coefficients $\hat{\lambda}_k$ are identified for every transmit path individually in a two-step procedure. First, the coefficients $\hat{\alpha}_k, \hat{\beta}_k, \hat{\gamma}_k, \hat{\delta}_k$ of the dual-input PA model in (3) are estimated. Then, using the estimated PA model coefficients, the CTMM coefficients $\hat{\lambda}_k$ are estimated. The two steps of the procedure are performed for several iterations until the result is satisfying. In the initial step, all CTMM coefficients $\hat{\lambda}_k$ are arbitrarily set to 1. A simple flowchart of the procedure is shown in Fig. 3.

In step 1 of the procedure, using (3), least-squares estimates $\hat{\alpha}_k, \hat{\beta}_k, \hat{\gamma}_k, \hat{\delta}_k$ of the PA model coefficients are obtained with

$$\begin{bmatrix} \hat{\alpha}_k^T & \hat{\beta}_k^T & \hat{\gamma}_k^T & \hat{\delta}_k^T \end{bmatrix}^T = \mathbf{G}(\mathbf{a}_{1k}, \hat{\mathbf{a}}_{2k})^+ \mathbf{b}_{2k} \quad (5)$$

where the pseudoinverse $\mathbf{G}^+ = (\mathbf{G}^H \mathbf{G})^{-1} \mathbf{G}^H$ is used.

In step 2 of the procedure, which is fully derived in Appendix B, an estimate of the CTMM coefficients $\hat{\lambda}_k = \Re\{\hat{\lambda}_k\} + j\Im\{\hat{\lambda}_k\}$ is found with

$$\begin{bmatrix} \Re\{\hat{\lambda}_k\} \\ \Im\{\hat{\lambda}_k\} \end{bmatrix} = \begin{bmatrix} \Re\{\mathbf{F}_k^{(1)} + \mathbf{F}_k^{(2)}\} & \Im\{-\mathbf{F}_k^{(1)} + \mathbf{F}_k^{(2)}\} \\ \Im\{\mathbf{F}_k^{(1)} + \mathbf{F}_k^{(2)}\} & \Re\{\mathbf{F}_k^{(1)} - \mathbf{F}_k^{(2)}\} \end{bmatrix}^+ \times \begin{bmatrix} \Re\{\mathbf{b}_{2k} - \mathbf{f}_k^{(0)} - \hat{\mathbf{f}}_k^{(3)}\} \\ \Im\{\mathbf{b}_{2k} - \mathbf{f}_k^{(0)} - \hat{\mathbf{f}}_k^{(3)}\} \end{bmatrix} \quad (6)$$

where $\Re\{\cdot\}$ and $\Im\{\cdot\}$ denote real part and imaginary part, respectively, and

$$\mathbf{F}_k^{(1)} = \text{diag}(\mathbf{f}_k^{(1)})\mathbf{B}_2 \quad \mathbf{F}_k^{(2)} = \text{diag}(\mathbf{f}_k^{(2)})\mathbf{B}_2^* \quad (7)$$

where $\text{diag}(\mathbf{f})$ denotes a diagonal matrix with the elements of the vector \mathbf{f} as diagonal entries and $\mathbf{B}_2 = [\mathbf{b}_{21}, \dots, \mathbf{b}_{2K}]$. The vectors $\mathbf{f}_k^{(0)}, \mathbf{f}_k^{(1)}, \mathbf{f}_k^{(2)}$, which contain values for all time samples, are obtained from

$$f_k^{(0)} = \sum_{p=0}^{(P-1)/2} \hat{\alpha}_{kp} a_{1k}^{p+1} a_{1k}^{*p}$$

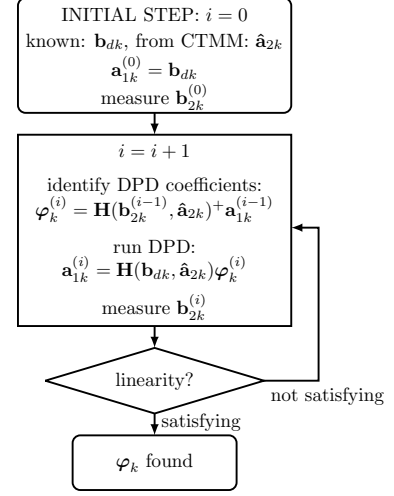


Fig. 4. Flowchart for the identification of the dual-input DPD coefficients of the k th transmit path.

$$\begin{aligned} f_k^{(1)} &= \sum_{p=0}^{(P-1)/2} \hat{\beta}_{kp} a_{1k}^p a_{1k}^{*p} \\ f_k^{(2)} &= \sum_{p=1}^{(P-1)/2} \hat{\gamma}_{kp} a_{1k}^{p+1} a_{1k}^{*p-1}. \end{aligned} \quad (8)$$

Furthermore,

$$\begin{aligned} f_k^{(3)}(\hat{\lambda}_k) &= \sum_{p=1}^{(P-1)/2} \sum_{v=0}^p \sum_{\substack{u=0 \\ u > 1-v}}^{p+1} \hat{\delta}_{kpuv} a_{1k}^{p+1-u} a_{1k}^{*p-v} \\ &\times \left(\mathbf{b}_2^T \hat{\lambda}_k \right)^u \left(\mathbf{b}_2^{*T} \hat{\lambda}_k^* \right)^v. \end{aligned} \quad (9)$$

As can be seen, $f_k^{(3)}(\hat{\lambda}_k)$ is nonlinear in the crosstalk coefficients $\hat{\lambda}_k$. However, it is not desirable to solve this nonlinear problem due to computational requirements. There are two alternatives to avoid this. The first alternative is to set $\hat{\lambda}_k = 0$ within $f_k^{(3)}$, such that $\hat{f}_k^{(3)} = f_k^{(3)}(\mathbf{0}) = 0$. The second option is to use a previous estimate $\hat{\lambda}_k$. Note that for a dual-input PA model using only linear terms of a_{2k} , $f_k^{(3)}(\hat{\lambda}_k) = 0$ inherently.

For a system model with memory effects, the equations for step 2 need be adapted. This is shown in Appendix C.

Note that another option of obtaining a set of coefficients for the CTMM block is to use antenna S-parameter measurements. We will show the results for both S-parameter measurements and the identification procedure outlined above in the experiment section.

B. Dual-Input DPD

Fig. 2 illustrates the proposed DPD method. The predistorted signal a_{1k} , i.e. the output of the dual-input DPD of the k th transmit path, is calculated for all time samples by

$$\mathbf{a}_{1k} = \mathbf{H}(\mathbf{b}_{dk}, \hat{\mathbf{a}}_{2k}) \boldsymbol{\varphi}_k \quad (10)$$

where $\boldsymbol{\varphi}_k$ are the dual-input DPD coefficients of the k th transmit path, which have to be identified. The signals b_{dk} to the DPD are known. The second input signals \hat{a}_{2k} to

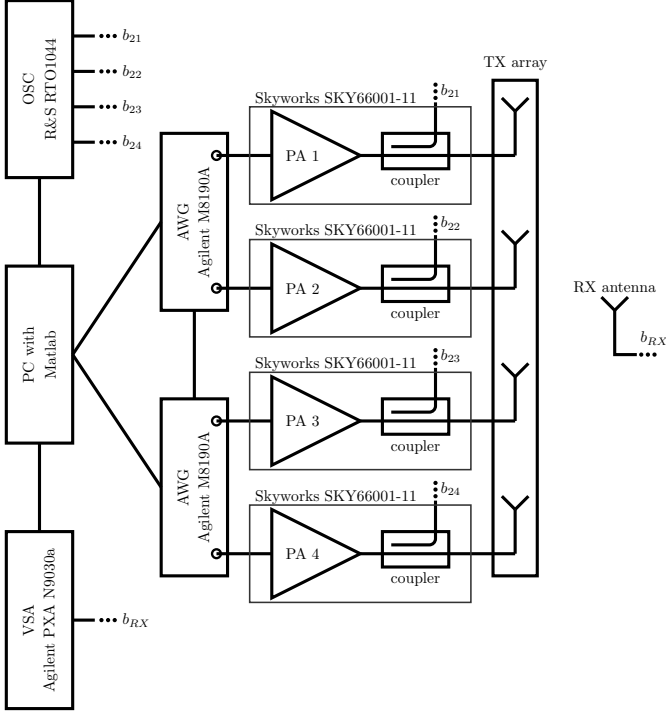


Fig. 5. Block diagram of the measurement setup.

the DPD are produced by the CTMM block, as discussed in Sec. III-A. The dual-input DPD matrix $\mathbf{H}(\mathbf{b}_{dk}, \hat{\mathbf{a}}_{2k})$ contains basis functions of the same kind as the dual-input PA model matrix $\mathbf{G}(\mathbf{a}_{1k}, \mathbf{a}_{2k})$ in (3). The PA output signals b_{2k} are obtained by measurements.

Fig. 4 shows a flowchart for the identification of the dual-input DPD coefficients φ_k . The least-squares estimate of the dual-input DPD coefficients φ_k is found as

$$\varphi_k = \mathbf{H}(\mathbf{b}_{2k}, \hat{\mathbf{a}}_{2k})^+ \mathbf{a}_{1k}. \quad (11)$$

In the initial step of the dual-input DPD identification, $a_{1k} = b_{dk}$. The dual-input DPD identification is iterated until the resulting linearity is satisfying. This procedure is exactly the same as a conventional indirect learning architecture (ILA) [21], with the extension that there is a known second input to the DPD.

The proposed dual-input DPD can be used as the regional model in a vector-switched DPD, which was proposed to improve the results of PA linearization techniques in [22]. In this approach, several switching regions are defined, and a separate DPD is calculated and applied for each region. The switching regions for the dual-input DPD are based on the signals b_{dk} .

C. Non-Uniqueness of Coefficients

The solution for CTMM coefficients $\hat{\lambda}_k$ and dual-input DPD coefficients φ_k is not unique. Multiplying the CTMM coefficients by an arbitrary scaling factor while reciprocally scaling dual-input DPD coefficients will result in the same DPD output signals, and therefore not change linearization performance. There is also ambiguity in the k th element of

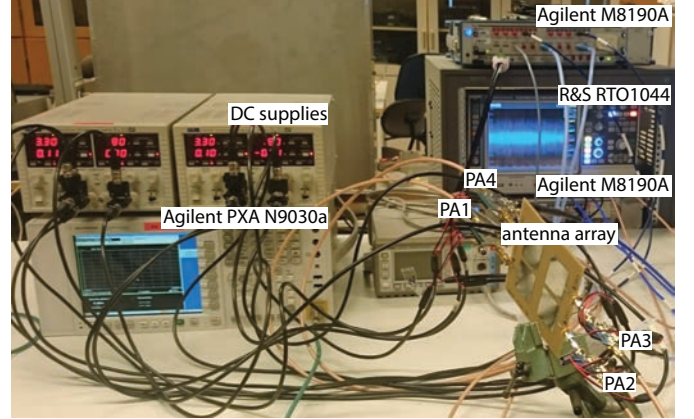


Fig. 6. Photo of the measurement setup in the lab showing measurement instruments, PAs and antenna array. PA1 is connected to antenna 1 of the array, PA2 to antenna 2 and so on.

each of the CTMM coefficient vector $\hat{\lambda}_k$, as can be seen from (17) where b_{2k} appears on both sides of the equation. This can cause numerical problems or unidentifiability. One way to prevent either CTMM coefficients or dual-input DPD coefficients from becoming arbitrarily large or small, is to keep the k th element of $\hat{\lambda}_k$ to a fixed value, and to normalize the other CTMM coefficients of the vector after each CTMM identification step 2 to a suitable value by dividing all of them by a scaling factor. For example, after each step 2 of the CTMM identification, perform the normalization $\hat{\lambda}_k / \rho_k$, where $\rho_k = \max_{j, j \neq k} (\hat{\lambda}_k(j))$, and set $\hat{\lambda}_k(k) = 1$. By doing so, CTMM coefficients with absolute values between zero and one are obtained. This is the approach we have taken in our studies.

IV. EXPERIMENTAL VALIDATION

In this section, the measurement setup of a four-element transmitter is presented, which is used to validate the proposed DPD technique and compare it to existing techniques. Measurement settings and performance evaluation measures are given as well. The results are presented in Section V.

A. Measurement Setup

A block diagram of the measurement setup of the four-element transmitter is shown in Fig. 5, and a photo of the setup is shown in Fig. 6.

The four driving signals for the PAs were created in MATLAB. The signals were different and independent orthogonal frequency-division multiplexing (OFDM) signals with 5 MHz bandwidth, and peak-to-average power ratios of around 8.5 dB. Two synchronized high-speed dual-channel arbitrary waveform generators (AWG, Agilent M8190A) were used to synthesize the four driving signals. The four PAs were identical GaAs PA evaluation boards from Skyworks (SKY66001-11), supplied with 3.3 V and operated at a center frequency of 2.12 GHz. The instantaneous gain of one of the PAs is shown in Fig. 7. The PAs have integrated couplers at their outputs, which were used to measure the individual PA output signals. The antenna array was a rectangular four-element array

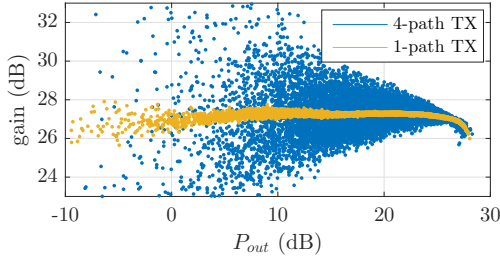


Fig. 7. PA gain vs. output power for one of the PAs (PA1). The PA gain is shown for the PA when used in a single-path transmitter (yellow), and for the PA when in the presented four-path transmitter (blue).

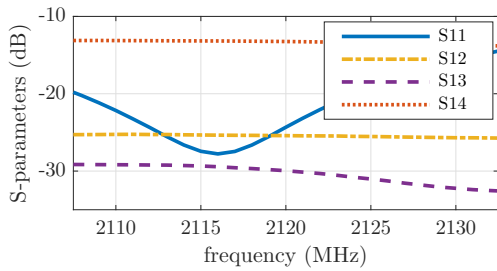


Fig. 8. Measured antenna array S-parameters vs. frequency. The characteristics are only shown for antenna 1, since they are similar for all antennas due to reciprocity. The figure shows the scattering parameters for: reflection (S11), adjacent element (S14), opposite element (S12), and diagonally opposite element (S13). The antenna array can be seen in the photo of the measurements setup in Figure 6.

with microstrip patch elements. The highest coupling factor between two elements of the array is around -12 dB. The measured array scattering parameters versus frequency are shown in Fig. 8 for one antenna in the array. The other S-parameters show similar behavior due to symmetry of the antenna array design, which can be seen in Fig. 6. A single antenna element of the same type as the transmitter elements was used as a receiver. The four individual PA output signals were measured with a four-channel 8-bit oscilloscope (OSC, Rohde&Schwarz RTO1044). The noise floor of the oscilloscope was the limit for the achievable linearization. The received signal, which is in the far field of the four-element transmitter, was measured with a vector signal analyzer (VSA, Agilent PXA N9030a). The received signal is a linear combination of the four transmitted signals including channel effects as well as the effect of the receiver antenna. Since independent signals are feeding the four antennas, the resulting relative distortion of the received signal becomes independent of direction. Hence, the nonlinear distortion observed at the receiver is independent of the receiver location, which is why measurement results are presented for only one position. Processing was done in MATLAB at a baseband sampling frequency of 25 MHz.

Dynamic effects are characteristic for PAs in a wideband scenario. A driving signal bandwidth of 5 MHz was chosen since it was wide enough to cause dynamic effects in the chosen PAs, while the antenna array characteristics were still approximately the same for all frequencies within the signal bandwidth.

B. Evaluated DPD Techniques and Performance Measures

The proposed DPD technique was tested and compared to single-input DPD and multi-input DPD, as well as the case where no DPD is used.

As mentioned, an alternative to identifying a set of CTMM coefficients from measurements of the PA output signals is to use measurements of the antenna array S-parameters at center frequency. Results of the proposed technique are presented for both these methods. A big disadvantage of using measured S-parameters as CTMM coefficients is that a separate calibrated setup is required to obtain the S-parameters. The S-parameter measurements were performed with a two-port vector network analyzer directly at the antenna ports. The measurements were performed pairwise, with the other ports terminated in 50 Ohm. Results of the proposed technique where the CTMM coefficients were identified from PA output measurements are indicated with DI-CTMM DPD, and results using S-parameters with DI-SParam DPD.

All multi-antenna DPD results, i.e., the proposed technique as well as the multi-input DPDs, are shown for two system models. The first system model, which is indicated by NLCT, is a memory polynomial dual-input PA model that considers nonlinear terms of the crosstalk signals a_{2k} as given in (15). The second system model, indicated by LCT, considers only linear terms of the crosstalk signals a_{2k} such that all terms in (15d) become zero. The LCT model is suitable if the crosstalk signals a_{2k} are relatively small in power. Both system models consider cross-products between PA output and antenna crosstalk signals. The advantage of the LCT model is that it leads to much lower complexity than the NLCT model. For the CTMM identification with NLCT, a previous estimate of the crosstalk coefficients λ_k was used in $f^{(3)}$ in (21).

Note that the LCT multi-input DPD is based on the technique in [14], and the NLCT multi-input DPD is based on the technique in [15]. Both techniques have been extended towards a four-path transmitter. For all DPDs, a vector-switched memory-polynomial DPD structure [22] with four switching regions based on the desired output signals b_{dk} was used.

For the evaluation of results, the normalized mean square error (NMSE) and the adjacent channel leakage ratio (ACLR) are used. The NMSE is calculated by

$$\text{NMSE} = \frac{\sum_{n=0}^{N-1} |x(n) - \hat{x}(n)|^2}{\sum_{n=0}^{N-1} |x(n)|^2} \quad (12)$$

where N is number of time samples, $x(n)$ is the desired output signal and $\hat{x}(n)$ is the measured output signal. The ACLR is calculated as

$$\text{ACLR} = \max_{c=1,2} \left\{ \frac{\sum_{f_{(adj)_c}} |X(f)|^2}{\sum_{f_{ch}} |X(f)|^2} \right\} \quad (13)$$

where $X(f)$ is the measured amplitude spectrum, f_{ch} denotes inband frequencies and f_{adj} frequencies in the adjacent channel.

V. RESULTS

In this section, we first evaluate the CTMM identification procedure, since its reliability is integral for the proposed

TABLE I
ACLR AND NMSE FOR DIFFERENT DPD TECHNIQUES.

	DPD technique	ACLR (dB)					NMSE (dB)			
		PA1	PA2	PA3	PA4	RX	PA1	PA2	PA3	PA4
NLCT ^a	none	-40.0	-36.0	-36.6	-38.2	-36.3	-21.8	-20.6	-23.4	-22.1
	single-input	-45.7	-44.2	-45.2	-44.2	-45.9	-22.2	-20.8	-25	-22.3
	proposed DI-CTMM	-51.6	-50.5	-49.9	-52.4	-50.7	-41.2	-41.1	-39.7	-40.5
	proposed DI-SParam	-50.8	-51.0	-50.8	-52.6	-51.2	-33.7	-33.0	-34.1	-39.5
	multi-input	-50.1	-49.1	-49.0	-50.0	-47.2	-41.9	-40.8	-40.2	-41.1
LCT ^b	proposed DI-CTMM	-50.4	-51.6	-50.2	-52.2	-51.3	-40.9	-40.9	-40.0	-41.1
	proposed DI-SParam	-51.1	-51.3	-50.1	-52.9	-50.9	-34.8	-32.2	-33.5	-39.8
	multi-input	-51.1	-51.5	-50.5	-53.0	-51.7	-41.3	-41.4	-40.5	-41.8

^a NLCT indicates DPDs based on models with linear and nonlinear crosstalk terms

^b LCT indicates DPDs based on models with only linear crosstalk terms

DPD technique. Then, the results for the proposed DPD technique are presented, evaluated and compared to existing DPD techniques.

A. Evaluation of CTMM Identification

The CTMM coefficients are an essential part of the proposed DPD technique. Therefore, it is important that the identification procedure is reliable. The reliability of the CTMM identification method using PA output measurements as proposed in Section III-A is evaluated by extracting CTMM coefficients for different initial values and comparing the results. The initial values were complex numbers with real and imaginary parts that were randomly chosen according to a uniform distribution with interval $[-1, 1]$.

We have used the procedure suggested in Section III-C to avoid numerical and identification problems. Using this procedure, the k th CTMM coefficient and the maximum CTMM coefficient of the k th path assume the fixed value 1. The results for identification of the remaining CTMM coefficients for all paths are shown in Fig. 9(a) for LCT, and in Fig. 9(b) for NLCT. Each blue dot represents an initial value, and the yellow circles show the results after the first iteration. The black plus signs show the results after the second iteration. As can be seen, the results converge to the same value for all initial values. After the second iteration, no difference between the results for different initial values can be noticed. Furthermore, the final results for LCT and NLCT agree.

The accuracy of the identified CTMM coefficients becomes evident from the performance evaluation of the proposed DPD technique, which is presented in the following subsection.

B. Performance of the DPDs

In Table I, the ACLRs for the received signal and each of the individual PA output signals are given for different DPD methods. For the case without any DPD the ACLR is between -40 dB and -36 dB. Using single-input DPD improves the ACLR to around -45 dB. The ACLR results for the proposed DI-CTMM and DI-Sparam DPDs and the multi-input DPDs are very similar, around -50 dB, for both NLCT and LCT. The multi-input LCT DPD reaches the best result. As can be seen, linearizing each transmit path via the measured PA

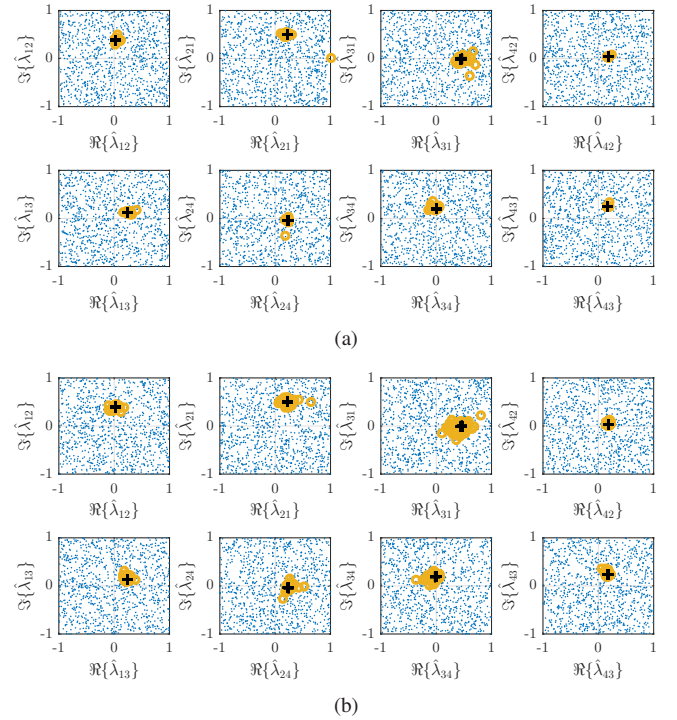


Fig. 9. Results of the CTMM coefficient identification based on (a) LCT and (b) NLCT. The figures show the initial values in blue dots, the results after the first iteration in yellow circles, and the results after the second iteration in black plus signs. As can be seen, the results converge such that after the second iteration no difference can be seen. The results for LCT and NLCT agree.

output signals, improves the ACLR of the received signal to a similar degree as the ACLR of the individual PA output signals.

The NMSEs for different DPD techniques are given in Table I. Results are shown for each of the individual PA output signals. Since we have no knowledge of the channel, it is not relevant to evaluate the NMSE for the received signal. Compared to the case without any DPD, the single-input DPD does not significantly improve the NMSE. The proposed DI-SParam DPD achieves an improvement of around 12 dB for both NLCT and LCT. The proposed DI-CTMM DPD and the multi-input DPD improve the NMSE by around 19 dB for both

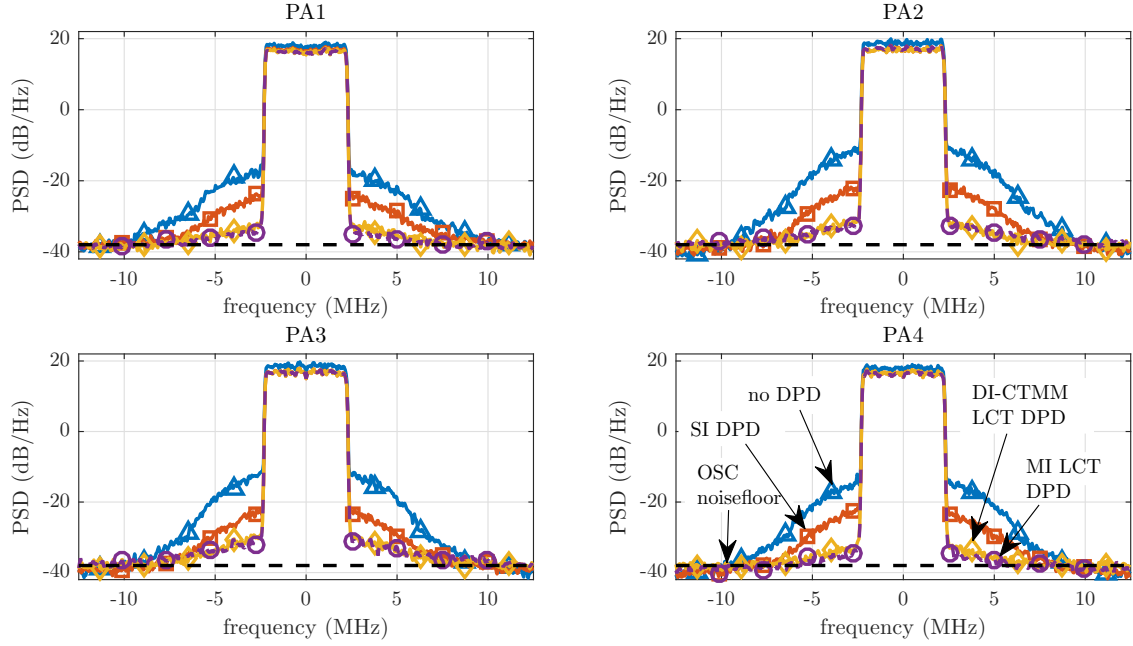


Fig. 10. Spectra of the PA output signals. The figure shows the results without DPD (blue \triangle), with single-input DPD (SI DPD, red \square), multi-input DPD (MI LCT DPD, purple \circ) and the proposed DPD (DI-CTMM LCT, yellow \diamond). The multi-input DPD and the proposed DPD are based on models using only linear terms of the crosstalk signals.

NLCT and LCT. The NMSEs for the proposed DI-CTMM DPDs and the multi-input DPDs are very similar, with the multi-input LCT DPD achieving the lowest values.

The lowest values for ACLR and NMSE are achieved by the multi-input LCT DPD and the proposed DI-CTMM LCT DPD. In Fig. 10 the spectra of the individual PA outputs are shown, and in Fig. 11 the normalized spectra of the received signals are shown for these two DPD techniques. Also shown for comparison are the results for single-input DPD and for the case without any DPD. As can be seen, the single-input DPD reduces the out-of-band distortion. However, both DI-CTMM LCT DPD and multi-input LCT DPD reach a much better result. The results for the received signal, which is the far field of the transmitter, are equal to the results of the PA output signals. Note that the performance of all linearization techniques is limited by the noise floor of the oscilloscope, which was used to measure the PA output signals. The noise floor is indicated in the figures.

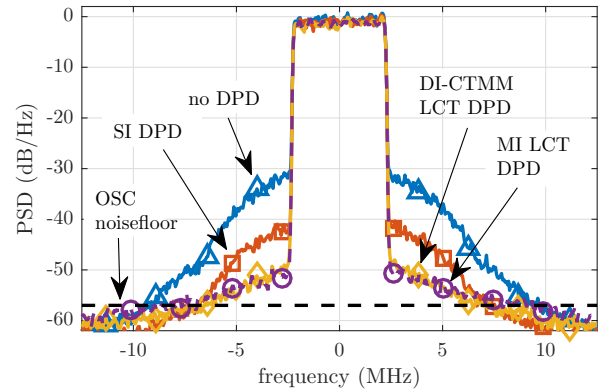


Fig. 11. Normalized spectra of the received signals. The figure shows the results without DPD (blue \triangle), with single-input DPD (SI DPD, red \square), multi-input DPD (MI LCT DPD, purple \circ) and the proposed DPD (DI-CTMM LCT, yellow \diamond). The multi-input DPD and the proposed DPD are based on models using only linear terms of the crosstalk signals.

C. Complexity

The issue of complexity has many aspects. There are different types of complexity that should be considered [23]:

- Run-time complexity is the complexity to execute the DPD on the input signals. It depends on the number of calculations that are necessary for each input signal sample to obtain the predistorted signal. Therefore, run-time complexity also depends on signal and evaluation bandwidth, i.e., the required sampling rate of the system.
- Identification complexity is the complexity required to find the initial version of the predistorter. This is typically done using least-squares techniques [30]. Identification

complexity depends on the amount of parameters that needs to be identified. The initial identification is only performed once, usually in the lab or factory. After that, adaptation is used to adjust the predistorter to changes in the system. For this reason, identification complexity is negligible and the focus is put on adaptation complexity.

- Adaptation complexity is the complexity to adjust the identified predistorter to changes in system behavior while the system is running. Adaptation is commonly done using algorithms like least mean squares (LMS), recursive least squares (RLS), or similar [26]. Using these techniques, every coefficient of the predistorter is updated

TABLE II
NUMBER OF DPD COEFFICIENTS PER TRANSMIT PATH PER SWITCHING REGION AND NUMBER OF CTMM COEFFICIENTS PER PATH REQUIRED FOR THE DIFFERENT DPD TECHNIQUES.

	DPD technique	DPD coeff.	CTMM coeff.
4-path	single-input	12	-
	proposed LCT	21	4
	proposed NLCT	30	4
	multi-input LCT	39	-
	multi-input NLCT	135	-
K -path	single-input	12	-
	proposed LCT	21	K
	proposed NLCT	30	K
	multi-input LCT	$12 + 9(K - 1)$	-
	multi-input NLCT	$12 + (9 + 8K)(K - 1)$	-

individually at run-time. Adaptation complexity depends on how much and how fast the systems changes over time due to, e.g., temperature drift.

The exact complexity and related measures, such as power consumption, cost, and space, always depend on a specific implementation, i.e., implementation concept [24], [25], [26], used hardware, necessity and frequency of adaptation, training algorithm, adaptation algorithm, matrix inversion algorithm, bandwidth requirements etc.

We have chosen to use the number of predistorter coefficients as a basis for our complexity analysis. Comparing the number of predistorter coefficients is a simple way of comparing the complexity of the different DPDs. It can also be used to investigate how complexity scales when increasing the numbers of transmit paths, which serves as an indicator of whether a DPD technique is feasible for larger multi-antenna systems or not. For predistorters based on the Volterra series, as the ones considered in this work, all types of complexity depend on the number of predistorter coefficients [27]. Therefore, when considering such DPD structures with similar types of basis functions and the same requirements for evaluation bandwidth, complexity comparison is commonly based on comparing the number of DPD coefficients. Thus, reducing the number of coefficients is often used as a technique to reduce complexity [14], [17], [27], [28], [29].

For single-input DPD and the multi-input DPDs, all types of complexity scale with the number of DPD coefficients. For the proposed DPDs, run-time and adaptation complexity scale with the total number of predistorter coefficients, which includes the dual-input DPD coefficients and the CTMM coefficients.

Table II compares the number of DPD coefficients per transmit path per switching region and the number of CTMM coefficients. Note that these numbers are specific to the multi-antenna transmitter that is used in the measurements. The proposed DI-CTMM and DI-Sparam DPDs require fewer coefficients than the multi-input DPDs. The number of coefficients required for multi-input NLCT DPD is excessively higher than for the other DPDs. Except for the single-input DPD, the lowest number of coefficients is required for the proposed DPD based on LCT. The table also shows how the number of

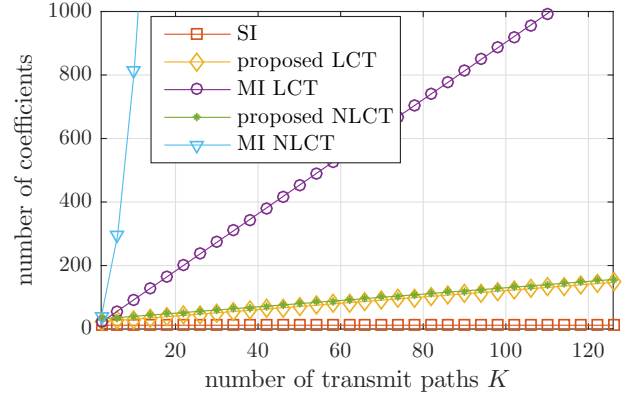


Fig. 12. Illustration of how the number of coefficients per path would scale with the number of transmit paths K . The figure shows the number of DPD coefficients per switching region plus the number of CTMM coefficients. Numbers are shown for: single-input DPD (SI DPD), the proposed DPD (proposed LCT) and multi-input DPD (MI LCT DPD) based on models with only linear crosstalk terms, the proposed DPD (proposed NLCT) and multi-input DPD (MI NLCT DPD) based on models using linear and nonlinear terms of the crosstalk signals.

coefficients would scale for a K -path transmitter with the same type of components. This is also illustrated in Fig. 12, where the number of DPD coefficients per switching region plus the number of CTMM coefficients are plotted versus number of transmit paths K . It is important to realize that for all multi-input DPDs the number of coefficients increases rapidly with the number of paths, while for the proposed DPDs only the number of CTMM coefficients increases and the dual-input DPD is not affected. For multi-antenna systems with many antennas, the difference in total coefficient numbers becomes huge, as can be seen in Fig. 12. The major advantage of the proposed technique is how it scales for increasing numbers of transmit paths.

The numbers given in Table II and Fig. 12 are specific to the multi-antenna transmitter used in the measurements and can therefore not give an exact prediction of the number of coefficients for other systems. However, due to its structure, the proposed technique will inherently have lower run-time complexity than the existing approaches for any transmitter where antenna crosstalk from more than one transmit path needs to be considered in the DPD.

D. Discussion

Analyzing the presented results, several things can be noticed. It is obvious that single-input DPD is not suitable for a multi-antenna transmitter. While the ACLR can be improved with such a DPD, the NMSE is almost the same as when using no DPD at all.

The ACLR results for the proposed DI-CTMM and DI-Sparam DPDs are very similar. However, the NMSE results are worse for the DI-Sparam DPD. The reason is that the reference plane for the measurements of the S-parameters is not exactly the same as the reference plane for measurements of the PA output signals. Observations at different reference planes can have different phase shifts, gain shifts, and delays. Such issues can lead to a degradation of performance, or even failure of

the DPD to linearize the system. With careful calibration, the phase, gain and delay offsets were kept very small, such that a performance degradation when using DI-SParam DPDs is only noticeable in terms of NMSE while the ACLR stays unaffected. The CTMM coefficients estimated directly from PA output signals allow for better DPD results since they are obtained from measurements in the exact reference plane. In addition to the worse NMSE results and the complicated calibration, another disadvantage of using S-parameters as CTMM coefficients is that a separate measurement setup is required. In highly integrated transmitters these measurements might not be feasible. Furthermore, for a system implementation, is not possible to use adaptive algorithms to update the CTMM when using the S-parameters instead of the proposed identification method.

All results obtained with the proposed DI-CTMM DPD are almost the same as the results of the multi-input DPDs. For both these techniques, the DPDs based on NLCT have much higher complexity than the DPDs based on LCT. Therefore, for the multi-antenna transmitter used in these measurements, there is no advantage in using the NLCT based DPDs. Hence, for a maximum coupling between antennas of -12 dB it is not necessary to consider a dual-input PA model with nonlinear crosstalk and mismatch terms. However, for systems with higher coupling it can be expected that NLCT based DPDs are more suitable [18].

Considering both NMSE, ACLR, and the complexity, the best results are obtained by the proposed DI-CTMM LCT DPD. Even though the DI-CTMM LCT DPD is narrowly outperformed by the multi-input LCT DPD, the number of DPD coefficients is reduced. For emerging multi-antenna systems with large numbers of transmit paths, the complexity of the proposed technique scales much better than for existing solutions, as shown in Table II and Fig. 12. The presented results also indicate that in cases where the coupling is so high that NLCT based DPDs are required, existing solutions will quickly reach the limits of what is feasible to implement, while the complexity of the proposed technique still rises only relatively slowly with the number of transmit paths. Hence, the proposed DPD technique is an extremely attractive alternative for the linearization of emerging multi-antenna transmitters.

VI. CONCLUSIONS

In this paper, we present a technique to jointly compensate for the effects of PA nonlinearity, antenna crosstalk and mismatch in wideband multi-antenna transmitters. Dual-input DPDs in every transmit path are combined with a linear model of the crosstalk and mismatch characteristics of the antenna array that is shared by all transmit paths. By using this structure, the use of multi-input DPDs is avoided. Therefore, the complexity of the proposed technique scales very favorably with the number of transmit paths: the CTMM block increases linearly with the number of paths, while the complexity of the dual-input DPDs is not affected by the number of paths. This is a huge benefit of our technique compared to existing approaches, especially considering the trend towards larger-scale multi-antenna systems.

As we have shown, no prior knowledge of the system components is required to identify the dual-input DPDs or the CTMM. Similar to conventional DPD approaches, all necessary information can be identified from measurements of the PA output signals using least-squares estimation techniques. A potential disadvantage of the proposed technique is that, for the identification procedure, the input signals b_{dk} to the different transmit paths cannot be fully correlated, such that for systems with fully correlated signals like, e.g., beamforming systems, special training signals have to be used.

Results are shown for the linearization of a four-path transmitter. When comparing the proposed technique to existing approaches, it can be seen that the performances are similar, while our technique has lower complexity. With the presented measurements we show that our technique is suitable for the linearization of wideband PAs. A topic for future work is to include also dynamic antenna behavior in our algorithms. This can be necessary for systems where the antenna characteristics vary strongly within the input signal bandwidth.

APPENDIX

A. Dual-Input PA Models Including Memory Effects

To account for dynamic effects in PAs driven by wideband signals, memory has to be considered in the dual-input PA model. The most general form of a dual-input PA model with memory effects is a dual-input model according to the Volterra series [31]:

$$\begin{aligned}
 b_{2k}(n) = & \sum_{q_1=0}^1 \sum_{m_1=0}^M \theta_{k0q_10m_1} (a_{1k}(n-m_1))^{1-q_1} \\
 & \times (a_{2k}(n-m_1))^{q_1} \\
 & + \sum_{p=1}^{(P-1)/2} \sum_{q_1=0}^{p+1} \sum_{q_2=0}^p \sum_{m_1=0}^M \cdots \sum_{m_{p+1-q_1}=m_{p-q_1}}^M \\
 & \sum_{m_{p+2+q_1}=0}^M \cdots \sum_{m_{p+1}=m_p}^M \sum_{m_{p+2}=0}^M \cdots \sum_{m_{2p+1-q_2}=m_{2p-q_2}}^M \\
 & \sum_{m_{2p+2-q_2}=0}^M \cdots \sum_{m_{2p+1}=m_{2p}}^M \theta_{kpq_1q_2m_1m_2\dots m_{2p+1}} \\
 & \times \prod_{i=1}^{p+1-q_1} a_{1k}(n-m_i) \prod_{l=p+2-q_1}^{p+1} a_{2k}(n-m_l) \\
 & \times \prod_{s=p+2}^{2p+1-q_2} a_{1k}^*(n-m_s) \prod_{r=2p+2-q_2}^{2p+1} a_{2k}^*(n-m_r). \quad (14)
 \end{aligned}$$

Because of the high model complexity, the full Volterra series approach is usually infeasible. Therefore, many models with reduced complexity have been proposed, such as the memory polynomial [6] and the generalized memory polynomial [7]. For the evaluation of the proposed DPD in measurements, we used the memory polynomial approach. The dual-input memory polynomial PA model is given in the same structure

as (2) by

$$\begin{aligned}
 b_{2k}(n) &= \sum_{m_1=0}^M \sum_{p=0}^{(P-1)/2} \alpha_{kp m_1} a_{1k}(n-m_1) \\
 &\quad \times |a_{1k}(n-m_1)|^{2p} \quad (15a) \\
 &+ \sum_{m_2=0}^M \beta_{k0 m_1} a_{2k}(n-m_2) \\
 &+ \sum_{m_1=0}^M \sum_{m_2=0}^M \sum_{p=1}^{(P-1)/2} \beta_{kp m_1 m_2} \\
 &\quad \times a_{2k}(n-m_2) |a_{1k}(n-m_1)|^{2p} \quad (15b) \\
 &+ \sum_{m_1=0}^M \sum_{m_2=0}^M \sum_{p=1}^{(P-1)/2} \gamma_{kp m_1 m_2} a_{2k}^*(n-m_2) \\
 &\quad \times (a_{1k}(n-m_1))^{p+1} (a_{1k}^*(n-m_1))^{p-1} \quad (15c) \\
 &+ \sum_{m_1=0}^M \sum_{m_2=0}^M \sum_{p=1}^{(P-1)/2} \sum_{v=0}^p \sum_{\substack{u=0 \\ u>1-v}}^{p+1} \delta_{kp uv m_1 m_2} \\
 &\quad \times (a_{1k}(n-m_1))^{p+1-u} (a_{1k}^*(n-m_1))^{p-v} \\
 &\quad \times (a_{2k}(n-m_2))^u (a_{2k}^*(n-m_2))^v \quad (15d)
 \end{aligned}$$

The dual-input memory polynomial PA model can be written in matrix form exactly as (3).

B. Derivation of CTMM Coefficient Identification

To find the equations for step 2 of the CTMM coefficient identification, (1) is introduced into (2)

$$\begin{aligned}
 b_{2k} &= \sum_{p=0}^{(P-1)/2} \alpha_{kp} a_{1k}^{p+1} a_{1k}^*{}^p \\
 &+ \sum_{p=0}^{(P-1)/2} \beta_{kp} a_{1k}^p a_{1k}^*{}^p \mathbf{b}_2^T \hat{\lambda}_k \\
 &+ \sum_{p=1}^{(P-1)/2} \gamma_{kp} a_{1k}^{p+1} a_{1k}^*{}^{p-1} \mathbf{b}_2^{*T} \hat{\lambda}_k^* \\
 &+ \sum_{p=1}^{(P-1)/2} \sum_{v=0}^p \sum_{\substack{u=0 \\ u>1-v}}^{p+1} \delta_{kp uv} \\
 &\quad \times a_{1k}^{p+1-u} a_{1k}^*{}^{p-v} \left(\mathbf{b}_2^T \hat{\lambda}_k \right)^u \left(\mathbf{b}_2^{*T} \hat{\lambda}_k^* \right)^v \quad (16)
 \end{aligned}$$

which is then rewritten as

$$b_{2k} = f_k^{(0)} + f_k^{(1)} \mathbf{b}_2^T \hat{\lambda}_k + f_k^{(2)} \mathbf{b}_2^{*T} \hat{\lambda}_k^* + f_k^{(3)}(\hat{\lambda}_k). \quad (17)$$

Using (7), this is rearranged and expressed for all time samples in matrix form as

$$\mathbf{b}_{2k} - \mathbf{f}_k^{(0)} - \hat{\mathbf{f}}_k^{(3)} = \mathbf{F}_k^{(1)} \lambda_k + \mathbf{F}_k^{(2)} \lambda_k^* \quad (18)$$

where $\hat{\mathbf{f}}_k^{(3)}$ is $\mathbf{f}_k^{(3)}(\hat{\lambda}_k)$ with $\hat{\lambda}_k$ either chosen equal to zero or set to a value from a previous identification step as explained in Section III-A. Finally, by splitting it into real and imaginary parts, (18) can be solved for the CTMM coefficients as given in (6).

C. CTMM Identification for Dual-Input Memory Polynomial PA Model

Using the dual-input memory polynomial PA model for the identification of the CTMM coefficients, the equation given in (6) can be applied with the following adaptations: (7) changes to

$$\begin{aligned}
 \mathbf{F}_k^{(1)} &= \sum_{m=0}^M \text{diag}(\mathbf{f}_{km}^{(1)}) \mathbf{B}_2(n-m) \\
 \mathbf{F}_k^{(2)} &= \sum_{m=0}^M \text{diag}(\mathbf{f}_{km}^{(2)}) \mathbf{B}_2^*(n-m) \quad (19)
 \end{aligned}$$

with $\mathbf{B}_2(n-m) = [\mathbf{b}_{21}(n-m), \dots, \mathbf{b}_{2K}(n-m)]$ where $\mathbf{b}_{2k}(n-m) = [b_{2k}(0-m), \dots, b_{2k}(N-1-m)]^T$ and

$$\begin{aligned}
 f_{km}^{(1)} &= \hat{\beta}_{k0m} + \sum_{m_1=0}^M \sum_{p=1}^{(P-1)/2} \hat{\beta}_{kp m_1 m} |a_{1k}(n-m_1)|^{2p} \\
 f_{km}^{(2)} &= \sum_{m_1=0}^M \sum_{p=1}^{(P-1)/2} \hat{\gamma}_{kp m_1 m} (a_{1k}(n-m_1))^{p+1} \\
 &\quad \times (a_{1k}^*(n-m_1))^{p-1}. \quad (20)
 \end{aligned}$$

Furthermore,

$$\begin{aligned}
 f_k^{(0)} &= \sum_{m_1=0}^M \sum_{p=0}^{(P-1)/2} \hat{\alpha}_{kp m_1} a_{1k}(n-m_1) \\
 &\quad \times |a_{1k}(n-m_1)|^{2p} \\
 f_k^{(3)}(\hat{\lambda}_k) &= \sum_{m_1=0}^M \sum_{m_2=0}^M \sum_{p=1}^{(P-1)/2} \sum_{v=0}^p \sum_{\substack{u=0 \\ u>1-v}}^{p+1} \hat{\delta}_{kp uv m_1 m_2} \\
 &\quad \times (a_{1k}(n-m_1))^{p+1-u} (a_{1k}^*(n-m_1))^{p-v} \\
 &\quad \times \left((\mathbf{b}_2(n-m_2))^T \hat{\lambda}_k \right)^u \\
 &\quad \times \left((\mathbf{b}_2^*(n-m_2))^T \hat{\lambda}_k^* \right)^v \quad (21)
 \end{aligned}$$

where $\mathbf{b}_2(n-m) = [b_{21}(n-m), \dots, b_{2K}(n-m)]^T$.

ACKNOWLEDGMENT

The authors would like to thank Skyworks Solutions, Inc. for donating the PA test boards used in the experiments.

REFERENCES

- [1] "5G Radio Access," Ericsson White Paper, Uen 284 23-3204 Rev C, April 2016, accessed on 2017-03-02. [Online]. Available: <https://www.ericsson.com/res/docs/whitepapers/wp-5g.pdf>
- [2] F. Rusek, D. Persson, B. K. Lau, E. Larsson, T. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, Jan 2013.
- [3] M. Romier, A. Barka, H. Aubert, J.-P. Martinaud, and M. Soiron, "Load-pull effect on radiation characteristics of active antennas," *IEEE Antennas Wireless Propag. Lett.*, vol. 7, pp. 550–552, 2008.
- [4] C. Fager, X. Bland, K. Hausmair, J. Cahuana, and T. Eriksson, "Prediction of smart antenna transmitter characteristics using a new behavioral modeling approach," in *IEEE MTT-S Int. Microw. Symp.*, June 2014, pp. 1–4.

- [5] K. Hausmair, S. Gustafsson, C. Sánchez-Pérez, P. N. Landin, U. Gustavsson, T. Eriksson, and C. Fager, "Prediction of nonlinear distortion in wideband active antenna arrays," *IEEE Trans. Microw. Theory Tech.*, accepted for publication.
- [6] J. Kim and K. Konstantinou, "Digital predistortion of wideband signals based on power amplifier model with memory," *Electron. Lett.*, vol. 37, no. 23, pp. 1417–1418, Nov 2001.
- [7] 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.
- [8] P. Suryasarman, M. Hoflehner, and A. Springer, "Digital pre-distortion for multiple antenna transmitters," in *European Microw. Conf.*, Oct 2013, pp. 412–415.
- [9] P. Suryasarman and A. Springer, "Adaptive digital pre-distortion for multiple antenna transmitters," in *2013 IEEE Global Conference on Signal and Information Processing*, Dec 2013, pp. 1146–1149.
- [10] P. M. Suryasarman and A. Springer, "A comparative analysis of adaptive digital predistortion algorithms for multiple antenna transmitters," *IEEE Trans. Circuits Syst. I*, vol. 62, no. 5, pp. 1412–1420, May 2015.
- [11] Z. Zhang, Y. Shen, S. Shao, W. Pan, and Y. Tang, "An improved cross talk cancelling digital predistortion for MIMO transmitters," *Mobile Information Systems*, vol. 2016, Article ID 5626495, 2016.
- [12] S. Bassam, M. Helaoui, and F. Ghannouchi, "Crossover digital predistorter for the compensation of crosstalk and nonlinearity in MIMO transmitters," *IEEE Trans. Microw. Theory Tech.*, vol. 57, no. 5, pp. 1119–1128, May 2009.
- [13] M. Amiri, S. Bassam, M. Helaoui, and F. Ghannouchi, "Matrix-based orthogonal polynomials for MIMO transmitter linearization," in *IEEE International Workshop on Computer Aided Modeling, Analysis and Design of Communication Links and Networks*, 2010, pp. 57–60.
- [14] A. Abdelhafiz, L. Behjat, F. M. Ghannouchi, M. Helaoui, and O. Hammi, "A high-performance complexity reduced behavioral model and digital predistorter for MIMO systems with crosstalk," *IEEE Transactions on Communications*, vol. 64, no. 5, pp. 1996–2004, May 2016.
- [15] S. Amin, P. Landin, P. Händel, and D. Rönnnow, "Behavioral modeling and linearization of crosstalk and memory effects in RF MIMO transmitters," *IEEE Trans. Microw. Theory Tech.*, vol. 62, no. 4, pp. 810–823, April 2014.
- [16] Z. A. Khan, E. Zenteno, P. Händel, and M. Isaksson, "Digital predistortion for joint mitigation of I/Q imbalance and MIMO power amplifier distortion," *IEEE Trans. Microw. Theory Tech.*, vol. 65, no. 1, pp. 322–333, Jan 2017.
- [17] E. Zenteno, S. Amin, M. Isaksson, D. Rönnnow, and P. Händel, "Combating the dimensionality of nonlinear MIMO amplifier predistortion by basis pursuit," in *European Microw. Conf.*, Oct 2014, pp. 833–836.
- [18] H. Zargar, A. Banai, and J. Pedro, "A new double input-double output complex envelope amplifier behavioral model taking into account source and load mismatch effects," *IEEE Trans. Microw. Theory Tech.*, vol. 63, no. 2, pp. 766–774, Feb 2015.
- [19] G. Z. El Nashef, F. Torres, S. Mons, T. Reveyrand, T. Monediere, E. N'Goya, and R. Quere, "Second order extension of power amplifiers behavioral models for accuracy improvements," in *European Microw. Conf.*, Sept 2010, pp. 1030–1033.
- [20] A. Zhu, P. J. Draxler, J. J. Yan, T. J. Brazil, D. F. Kimball, and P. M. Asbeck, "Open-loop digital predistorter for RF power amplifiers using dynamic deviation reduction-based volterra series," *IEEE Trans. Microw. Theory Tech.*, vol. 56, no. 7, pp. 1524–1534, July 2008.
- [21] J. Chani-Cahuana, C. Fager, and T. Eriksson, "A new variant of the indirect learning architecture for the linearization of power amplifiers," in *European Microw. Conf.*, Sept 2015, pp. 1295–1298.
- [22] S. Afsardoost, T. Eriksson, and C. Fager, "Digital predistortion using a vector-switched model," *IEEE Trans. Microw. Theory Tech.*, vol. 60, no. 4, pp. 1166–1174, 2012.
- [23] 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, June 2010.
- [24] A. Zhu, J. Pedro, and T. Brazil, "Dynamic deviation reduction-based volterra behavioral modeling of RF power amplifiers," *IEEE Trans. Microw. Theory Tech.*, vol. 54, no. 12, pp. 4323–4332, Dec 2006.
- [25] Y. Ma, Y. Yamao, Y. Akaiwa, and C. Yu, "FPGA implementation of adaptive digital predistorter with fast convergence rate and low complexity for multi-channel transmitters," *IEEE Trans. Microw. Theory Tech.*, vol. 61, no. 11, pp. 3961–3973, Nov 2013.
- [26] P. M. Suryasarman and A. Springer, "A comparative analysis of adaptive digital predistortion algorithms for multiple antenna transmitters," *IEEE Trans. Microw. Theory Tech.*, vol. 62, no. 5, pp. 1412–1420, May 2015.
- [27] F. M. Ghannouchi and O. Hammi, "Behavioral modeling and predistortion," *IEEE Microw. Mag.*, vol. 10, no. 7, pp. 52–64, Dec 2009.
- [28] N. Kelly and A. Zhu, "Low-complexity stochastic optimization-based model extraction for digital predistortion of RF power amplifiers," *IEEE Trans. Microw. Theory Tech.*, vol. 64, no. 5, pp. 1373–1382, May 2016.
- [29] Z. Wang, W. Chen, G. Su, F. M. Ghannouchi, Z. Feng, and Y. Liu, "Low computational complexity digital predistortion based on direct learning with covariance matrix," *IEEE Trans. Microw. Theory Tech.*, vol. PP, no. 99, pp. 1–11, 2017.
- [30] L. Guan and A. Zhu, "Optimized low-complexity implementation of least squares based model extraction for digital predistortion of RF power amplifiers," *IEEE Trans. Microw. Theory Tech.*, vol. 60, no. 3, pp. 594–603, March 2012.
- [31] M. Schetzen, *The Volterra and Wiener theories of nonlinear systems*, 2nd ed. Krieger Publishing Company, Malabar, Florida, 2006.