

Behavioral Modeling and Linearization of Crosstalk and Memory Effects in RF MIMO Transmitters

Shoaib Amin, *Student Member, IEEE*, Per N. Landin, Peter Händel, *Senior Member, IEEE*,
and Daniel Rönnow, *Member, IEEE*

Abstract—This paper proposes three novel models for behavioral modeling and digital pre-distortion (DPD) of nonlinear 2×2 multiple-input multiple-output (MIMO) transmitters in the presence of crosstalk. The proposed models are extensions of the single-input single-output (SISO) generalized memory polynomial model. Three types of crosstalk effects were studied and characterized as linear, nonlinear, and nonlinear & linear crosstalk. A comparative study was performed with previously published models for the linearization of crosstalk in a nonlinear 2×2 MIMO transmitter. The experiments indicate that, depending on the type of crosstalk, the selection of the correct model in the transmitter is necessary for behavioral modeling and sufficient DPD performance. The effects of coherent and partially non-coherent signal generation on the performance of DPD were also studied. For crosstalk levels of -30 dB, the difference in the normalized mean square error and adjacent channel power ratio was found to be 3 to 4 dB between coherent and partially non-coherent signal generation.

Index Terms—Behavioral modeling, digital pre-distortion, multiple-input multiple-output (MIMO), nonlinear and linear crosstalk, radio frequency, power amplifiers.

I. INTRODUCTION

In wireless communication systems, the radio frequency (RF) power amplifier (PA) of the transmitter distorts the signal due to nonlinear dynamic effects. Numerous studies have been conducted to model and compensate for these effects. In behavioral modeling, black box models are used to relate input and output signals [1].

In digital pre-distortion (DPD), the input signal to the amplifier is distorted in the digital domain before up-converting to RF to compensate for nonlinear dynamic effects of the PA. The algorithms for DPD and behavioral

modeling that are used in most cases are memory polynomials [1]–[3], which are reduced forms of the general Volterra series that can be used to model any nonlinear dynamic system with fading memory [4]. The reduced forms are necessary because the number of coefficients of the Volterra series becomes large even for moderate nonlinear orders and memory depths [5]. In DPD, the inverse of the PAs nonlinear dynamic transfer function is used, which is also a nonlinear dynamic transfer function that can be modeled using a Volterra series.

Behavioral modeling of single-input single-output (SISO) RF PAs is important in DPD because of the relationship between a forward model and its inverse [1]. Behavioral models have also been used in system simulations [6], [7] and indirect learning algorithms for identifying DPD models [8]. Research on the behavioral modeling of RF PAs has primarily focused on SISO models. The generalized memory polynomial (GMP) is based on physical knowledge of PAs [9] and has been extensively studied for both behavioral modeling and DPD [9], [10].

In multiple-input multiple-output (MIMO) systems, the transceivers exhibit specific impairments due to crosstalk between different paths if they are implemented on the same chipset [11]. Therefore, the behavioral modeling and DPD of the MIMO transmitters must consider these effects, and the algorithms should be based on MIMO Volterra theory [12]. In [11], memory polynomial DPD was used to compensate for the crosstalk effects in nonlinear transmitters with two input and two output signals, i.e. a 2×2 system. In [13], neural-networks were used for a similar application.

A DPD of 2×2 systems has also been applied to concurrent dual band PAs in which two modulated RF signals at different center frequencies are amplified by the same PA. Memory polynomials of various types have been used for DPD [14]–[17] of concurrent dual band PAs and behavioral modeling with an enhanced Parallel Hammerstein model has been reported in [15].

In this paper, we derive different memory polynomials for behavioral modeling and DPD of a 2×2 MIMO transmitter. We combine the SISO GMP with different types of crosstalk to develop new 2×2 MIMO models. A full Volterra complex baseband model for a 2×2 MIMO transmitter is formulated and compared to the novel 2×2 MIMO memory polynomials. The proposed models are also compared to the models presented in [11] and [18]. The experimental results presented in this paper indicate

Shoaib Amin is with the Department of Electronics, Mathematics and Natural Sciences, University of Gävle, SE-80176, Gävle, Sweden, and also with the Department of Signal Processing, ACCESS Linnaeus Centre, KTH Royal Institute of Technology, SE-10044 Stockholm e-mail: (shoaib.amin@hig.se).

Per N Landin is with the GigaHertz Center, Department of Signals and Systems, Chalmers University of Technology, SE-41296, Gothenburg email: (perlan@chalmers.se).

Peter Händel is with the Department of Signal Processing, ACCESS Linnaeus Centre, KTH Royal Institute of Technology, SE-10044 Stockholm and also with the Department of Electronics, Mathematics and Natural Sciences University of Gävle, SE-80176, Gävle, Sweden email: (ph@kth.se).

Daniel Rönnow is with the Department of Electronics, Mathematics and Natural Sciences University of Gävle, SE-80176, Gävle, Sweden email: (daniel.ronnow@hig.se).

that the proposed models yield lower model errors in behavioral modeling and lower distortions in DPD compared to the existing models for both behavioral modeling and DPD. The effect of phase noise in signal generation on the DPD performance is also investigated. The phase noise results in bias terms in the estimated parameters for the behavioral modeling and DPD algorithm. This effect also occurs in SISO PAs but is not obvious. To the best of the authors' knowledge, this effect has not been studied for a 2×2 MIMO transmitter. This analysis also applies to other 2×2 systems, such as DPD of concurrent dual band PAs.

The paper is organized as follows. In Section II, the crosstalk models are introduced for a MIMO transmitter and according to the type of crosstalk, black box models are proposed for behavioral modeling and DPD. Additionally, a generalized framework is developed for the Volterra series for a 2×2 MIMO transmitter in Section II. The experimental setup is presented in Section III. The measurement results of the crosstalk and phase noise are presented and discussed in Section IV, and the conclusions are given in Section V. For a 2×2 MIMO transmitter an analysis of the effect of phase noise on model performance is address in the Appendix.

II. THEORY

A. Linear and Nonlinear crosstalk

Crosstalk in an electrical system occurs due to the coupling of a signal from one branch to another. In 2×2 MIMO transmitters, crosstalk can be classified as either linear or nonlinear [11], depending on where the crosstalk takes place. In wireless transmitters, PAs are the primary source of nonlinearity, and thus coupling before and after the PAs results in nonlinear and linear crosstalk respectively [11], [18], as shown in Fig. 1, where α , β , γ and δ are the impulse responses of the linear filters before and after the PAs, which indicates the amount of crosstalk. Linear and nonlinear crosstalk effects are also discussed in [19], [20] where 2×2 MIMO transceivers were implemented on the same package. The proposed models in this paper would also work for the situation studied in [19], [20].

If there is no crosstalk, i.e., α , β , γ and δ are equal to zero, then Fig. 1 can be viewed as two separate SISO systems, and the output of each PA can be modeled as in [9],

$$y(n) = \sum_{p=1}^P \sum_{m_1=0}^{M_1} \sum_{m_2=0}^{M_2} g_{p,(m_1,m_2)} x(n-m_1)|x(n-m_1-m_2)|^{2(p-1)}, \quad (1)$$

where (1) is defined as a generalized memory polynomial (GMP) in [9], P is the nonlinearity order defined as $(\text{degree}+1)/2$, M_1 and M_2 are memory depths. If $M_2 = 0$, then (1) reduces to the Parallel Hammerstein (PH) model. The models proposed in this paper are extended versions of the SISO GMP model. The SISO GMP model was chosen as the starting point because it has been extensively used and its performance has been evaluated in previous studies and is based on physical knowledge of the PA [9], [10].

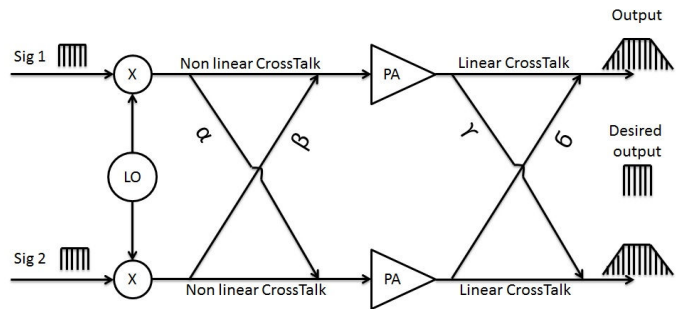


Fig. 1 Linear crosstalk (LC) and nonlinear crosstalk (NLC) in a MIMO transmitter.

In the presence of only nonlinear crosstalk ($\gamma = 0, \delta = 0$ and $\alpha \neq 0, \beta \neq 0$), the transmitter output can be modeled as in [11],

$$\begin{aligned} y_1(n) &= f_1(x_1(n) + \beta * x_2(n)) \\ y_2(n) &= f_2(\alpha * x_1(n) + x_2(n)), \end{aligned} \quad (2)$$

where $x_1(n)$ and $x_2(n)$ are the baseband input signals, $y_1(n)$ and $y_2(n)$ are the output signals, α and β are defined as above and $*$ indicates convolution. The functions $f_1(\cdot)$ and $f_2(\cdot)$ are nonlinear dynamic transfer functions that can be modeled using SISO memory polynomials. In this paper $x_1(n)$ and $x_2(n)$ are the generated input signals, $y_1(n)$ and $y_2(n)$ are the measured output signals. Similarly, in the case of linear crosstalk ($\alpha = 0, \beta = 0$ and $\gamma \neq 0, \delta \neq 0$), the output of the PAs can be modeled as

$$\begin{aligned} y_1(n) &= f_1(x_1(n)) + \delta * (f_2(x_2(n))) \\ y_2(n) &= \gamma * (f_1(x_1(n))) + f_2(x_2(n)), \end{aligned} \quad (3)$$

where γ, δ are defined previously, $y_1(n)$ and $y_2(n)$ are the output of the PAs after the linear crosstalk.

Equation (2) shows that the nonlinear transfer functions operate simultaneously on the input signals and (3) indicates that the PAs output is a linear combination of the nonlinear transfer functions $f_1(\cdot)$ and $f_2(\cdot)$. Therefore, conventional approaches used to model and linearize PAs are not sufficient for nonlinear MIMO transmitters in the presence of crosstalk.

In this paper, three different crosstalk effects appearing in a 2×2 MIMO transmitter are studied. These effects can be categorized into the following three cases: Linear crosstalk; $\alpha = 0, \beta = 0$ and $\gamma \neq 0, \delta \neq 0$, nonlinear crosstalk; $\alpha \neq 0, \beta \neq 0$ and $\gamma = 0, \delta = 0$ and nonlinear & linear crosstalk; $\alpha \neq 0, \beta \neq 0, \gamma \neq 0$ and $\delta \neq 0$. We also assumed that the crosstalk is memoryless i.e., frequency independent. Based upon these cases, three novel behavioral models are proposed in Section II-C. These novel models are developed such that in system identification, the model parameters take into account the level and type of crosstalk effect, i.e, the crosstalk is not an input parameter to the developed models. Also the developed models in Section II-C do not require prior knowledge of crosstalk effect and crosstalk level. Furthermore, any mismatch appearing between the crosstalk levels will also

be taken into account by all the models during system identification.

Volterra series are extensively used to understand the behavior of a nonlinear system(s). In the following section, complex baseband SISO Volterra series is extended to the 2×2 MIMO case.

B. Volterra Series

The relationship between input $x(n)$ and output $y(n)$ in a nonlinear SISO system can be modeled using a Volterra series. In wireless systems, the input and output signals are written in complex baseband form because the bandwidth of the signal is on the order of MHz, whereas the carrier frequency is typically on the order of GHz. The signal of primary interest is close to the carrier frequency. A common method for representing a signal close to the carrier is by the use of a complex-valued low-pass equivalent signal [21]. The complex baseband representation of a truncated Q -order Volterra model is given as in [5],

$$y(n) = \sum_{m=0}^M g_{1(m)}x(n-m_1) + \sum_{m_1=0}^M \sum_{m_2=m_1}^M \sum_{m_3=0}^M g_{3(m_1,m_2,m_3)}x(n-m_1)x(n-m_2)x^*(n-m_3) + \sum_{m_1=0}^M \sum_{m_2=m_1}^M \dots \sum_{m_{(Q+1)/2}=m_{(Q-1)/2}}^M \dots \sum_{m_Q=m_{Q-1}}^M g_{Q(m_1,m_2,\dots,m_Q)} \times x(n-m_1)x(n-m_2)\dots x^*(n-m_{(Q+1)/2})\dots x^*(n-m_Q) \quad (4)$$

where $g_1(m)$, $g_3(m_1, m_2, m_3)$ and $g_Q(m_1, m_2, \dots, m_Q)$ are the linear, 3^{rd} and Q^{th} order kernels of a nonlinear system respectively and $(\cdot)^*$ denotes the complex conjugate. In (4), due to kernel symmetry, i.e., $g_3(m_1, m_2, m_3) = g_3(m_2, m_1, m_3)$, the redundant terms are removed. Additionally, even order kernels have been removed because their effect can be omitted in band limited modeling [22].

Equation (4) can be extended to a complex baseband low-pass equivalent odd order MIMO Volterra model. For two input signals, $x_1(n)$ and $x_2(n)$, with the same center frequency, the output signal $y_1(n)$ can be modeled as (5). In (5), (5a) represents the kernels for linear response, one for the linear filtering of $x_1(n)$, and the other for $x_2(n)$. Equations (5b)-(5c) represent the self kernels. These self kernels have the same symmetry properties as the kernels of a SISO system. In (5b), $g_{3,1,111}$ represents the 3^{rd} order kernel, where the subscripts 3, 1, 111 indicate the 3^{rd} order Volterra kernel, 1 indicates the channel and 111 indicates the combination of input signals. In the case with no crosstalk, all kernels except $g_{1,1,1}(m)$ and $g_{3,1,111}(m_1, m_2, m_3)$ in (5) are zero. The cross kernels (5d)-(5e) have symmetry properties different from the self kernels. For (5d), e.g., $g_{3,1,112}(m_1, m_2, m_3) = g_{3,1,112}(m_2, m_1, m_3)$, but not for other permutations of m_1 , m_2 and m_3 . For the real valued 2×2 MIMO Volterra series, there are two 3^{rd} order cross kernels, not four as in (5),

and these kernels have less symmetry than the self kernels [12], [23].

In the SISO Volterra series, the number of coefficients increases along with an increase in the model order [5]. From (5), the implementation of the Volterra series for a nonlinear MIMO system becomes even more complex due to more kernels and lower symmetry than SISO nonlinear systems. Therefore, the use of memory polynomials is proposed. In memory polynomials, many of the kernels are set equal to zero based on prior knowledge [1], [6].

In the next section, novel black box models are proposed for a 2×2 MIMO transmitter to compensate for the effects of crosstalk. The structure of the proposed models is compared to the MIMO Volterra. The comparison indicates that the proposed models are a subset of the MIMO Volterra series.

C. Generalized Memory Polynomial for MIMO transmitter

Conventional SISO models cannot model the crosstalk effects that occur in a nonlinear MIMO transmitter [11], [18]; therefore, new model structures are required. This type of model was formulated in [11] for DPD and in [18] for behavioral modeling. This paper proposes three novel models with improved performance to compensate for different crosstalk effects for behavioral modeling and in DPD, for the crosstalk effects discussed in Section II-A.

1) *Generalized Memory Polynomial for Linear Crosstalk (GMPLC)*: In the presence of linear crosstalk i.e., when $\alpha = 0$ and $\beta = 0$, as shown in (3), the output signal of each channel in a 2×2 MIMO transmitter is a combination of nonlinear transfer function of a PA that is linearly combined with the nonlinear transfer function of a PA in the other channel, where f_1 and f_2 in (3) are modeled by the SISO GMP given by (1).

For a 2×2 MIMO transmitter, output channel 1 can be modeled as

$$y_1(n) = \sum_{p=1}^P \sum_{m_1=0}^{M_1} \sum_{m_2=0}^{M_2} g_{1p,(m_1,m_2)} x_1(n-m_1) |x_1(n-m_1-m_2)|^{2(p-1)} + \sum_{p=1}^P \sum_{m_1=0}^{M_1} \sum_{m_2=0}^{M_2} g_{2p,(m_1,m_2)} x_2(n-m_1) |x_2(n-m_1-m_2)|^{2(p-1)}, \quad (6)$$

where P is the nonlinearity order defined as $(\text{degree}+1)/2$, M_1 and M_2 are the the memory depths, $y_1(n)$ is the output of channel 1, $x_1(n)$ and $x_2(n)$ are the input signals to channels 1 and 2, respectively. The crosstalk effects are included in the parameters. The GMPLC model contains linear combinations of the PA outputs in a 2×2 system.

2) *Generalized Memory Polynomial for Nonlinear Crosstalk (GMPNLC)*: In the presence of nonlinear crosstalk, the GMPLC model is not sufficient to characterize the effects of nonlinear crosstalk in a 2×2 MIMO transmitter, because the GMPLC model contains only SISO nonlinear terms of one channel linearly combined with SISO nonlinear terms of the second channel.

$$y_1(n) = \sum_{m=0}^M g_{1,1,1}(m)x_1(n-m) + \sum_{m=0}^M g_{1,1,2}(m)x_2(n-m) + \quad (5a)$$

$$\sum_{m_1=0}^M \sum_{m_2=m_1}^M \sum_{m_3=0}^M g_{3,1,111}(m_1, m_2, m_3)x_1(n-m_1)x_1(n-m_2)x_1^*(n-m_3) + \quad (5b)$$

$$\sum_{m_1=0}^M \sum_{m_2=m_1}^M \sum_{m_3=0}^M g_{3,1,222}(m_1, m_2, m_3)x_2(n-m_1)x_2(n-m_2)x_2^*(n-m_3) + \quad (5c)$$

$$\sum_{m_1=0}^M \sum_{m_2=m_1}^M \sum_{m_3=0}^M g_{3,1,112}(m_1, m_2, m_3)x_1(n-m_1)x_1(n-m_2)x_2^*(n-m_3) + \quad (5d)$$

$$\sum_{m_1=0}^M \sum_{m_2=0}^M \sum_{m_3=0}^M g_{3,1,121}(m_1, m_2, m_3)x_1(n-m_1)x_2(n-m_2)x_1^*(n-m_3) + \quad (5e)$$

$$\sum_{m_1=0}^M \sum_{m_2=0}^M \sum_{m_3=0}^M g_{3,1,122}(m_1, m_2, m_3)x_1(n-m_1)x_2(n-m_2)x_2^*(n-m_3) + \quad (5f)$$

$$\sum_{m_1=0}^M \sum_{m_2=m_1}^M \sum_{m_3=0}^M g_{3,1,221}(m_1, m_2, m_3)x_2(n-m_1)x_2(n-m_2)x_1^*(n-m_3) + \quad (5g)$$

...

Thus, in order to compensate the effects of nonlinear crosstalk where the nonlinear transfer function operates simultaneously on both input signals as shown in (2), the model should include crossterms between the input signals x_1 and x_2 , i.e., the nonlinear combinations of $x_1(n)$ and $x_2(n)$ along with nonlinear SISO combinations. The output of a 2×2 MIMO transmitter in presence of nonlinear crosstalk can be modeled as

$$\begin{aligned} y_1(n) = & \sum_{p=1}^P \sum_{q=1}^{P-p+1} \sum_{m_1=0}^{M_1} \sum_{m_2=0}^{M_2} g_{1,p,q}(m_1, m_2)x_1(n-m_1) \\ & |x_1(n-m_1-m_2)|^{2(p-1)} |x_2(n-m_1-m_2)|^{2(q-1)} + \\ & \sum_{p=1}^P \sum_{q=1}^{P-p+1} \sum_{m_1=0}^{M_1} \sum_{m_2=0}^{M_2} g_{2,p,q}(m_1, m_2)x_2(n-m_1) \\ & |x_2(n-m_1-m_2)|^{2(p-1)} |x_1(n-m_1-m_2)|^{2(q-1)}. \end{aligned} \quad (7)$$

For 3^{rd} order nonlinearity, (7) contains the following terms; $x_1(n-m_1)|x_1(n-m_1-m_2)|^2$, $x_1(n-m_1)|x_2(n-m_1-m_2)|^2$, $x_2(n-m_1)|x_1(n-m_1-m_2)|^2$ and $x_2(n-m_1)|x_2(n-m_1-m_2)|^2$. Compared to the GMPLC model, the GMPNLC model in (7) contains not only the SISO nonlinearity, but also the combinations of crossterms between the inputs $x_1(n)$ and $x_2(n)$.

3) *Extended Generalized Memory Polynomial for Non-linear Crosstalk (EGMPNLC)*: EGMPNLC model is an extension of the GMPNLC model. The difference between these models is that the former model contains more crossterm combinations between input signals $x_1(n)$ and $x_2(n)$ compared to (7). The basis functions in the EGMPNLC model are shown in Table I. The difference between the GMPNLC and EGMPNLC models is evident, where

the GMPNLC model for 3^{rd} order nonlinearity includes only the following basis functions; $x_1(n-m_1)|x_1(n-m_1-m_2)|^2$, $x_1(n-m_1)|x_2(n-m_1-m_2)|^2$, $x_2(n-m_1)|x_1(n-m_1-m_2)|^2$ and $x_2(n-m_1)|x_2(n-m_1-m_2)|^2$, whereas the EGMPNLC model contains four additional terms for the 3^{rd} order nonlinearity ($k=3$), as shown in Table I.

Note that the terms shown in Table I for $k=3$, are equal to the 3^{rd} order terms in (5) if $m_2=m_3$, except for the terms on rows 3 and 7 in Table I, which corresponds to $m_1=m_3$. Therefore, the basis functions are a subset of the MIMO Volterra basis functions: EGMPNLC \subset MIMO Volterra, where " \subset " denotes subset. Similarly, GMPLC \subset GMPNLC \subset EGMPNLC, and 2×2 PH \subset GMPNLC. Additionally, SISO GMP \subset GMPLC.

4) *Parallel Hammerstein for a MIMO transmitter*: The model used for comparison is an extension of the SISO Parallel Hammerstein (PH) [24], to 2×2 -PH [18]

$$\begin{aligned} y_1(n) = & \sum_{p=1}^P \sum_{q=1}^{P-p+1} \sum_{m=0}^M g_{1,p,q,m}x_1(n-m) \\ & |x_1(n-m)|^{2(p-1)} |x_2(n-m)|^{2(q-1)} + \\ & \sum_{p=1}^P \sum_{q=1}^{P-p+1} \sum_{m=0}^M g_{2,p,q,m}x_2(n-m) \\ & |x_2(n-m)|^{2(p-1)} |x_1(n-m)|^{2(q-1)}, \end{aligned} \quad (8)$$

where M is the maximum memory depth.

D. System identification

Let $x_1(n)$, $x_2(n)$, $y_1(n)$ and $y_2(n)$ be the input and output signals of the 2×2 MIMO transmitter. The output signal model of a 2×2 MIMO transmitter can be written

Table I Basis functions of the EGMPNLC models, where $L_1 = (n - m_1)$, $L_2 = (n - m_1 - m_2)$ and k is the order of nonlinearity.

$k = 1$	$k = 3$	$k = 5$	\dots	$k = \infty$
$x_1(L_1)$	$x_1(L_1) x_1(L_2) ^2$	$x_1(L_1) x_1(L_2) ^4$	\dots	$x_1(L_1) x_1(L_2) ^{(k-1)}$
$x_2(L_1)$	$x_1(L_1) x_2(L_2) ^2$	$x_1(L_1) x_2(L_2) ^4$	\dots	$x_1(L_1) x_2(L_2) ^{(k-1)}$
	$x_2(L_1) x_1(L_2) ^2$	$x_2(L_1) x_1(L_2) ^4$	\dots	$x_2(L_1) x_1(L_2) ^{(k-1)}$
	$x_2(L_1) x_2(L_2) ^2$	$x_2(L_1) x_2(L_2) ^4$	\dots	$x_2(L_1) x_2(L_2) ^{(k-1)}$
	$x_1(L_1)x_1(L_2)x_2^*(L_2)$	$x_1(L_1) x_1(L_2) ^2x_1(L_2)x_2^*(L_2)$	\dots	$x_1(L_1) x_1(L_2) ^{(k-3)}x_1(L_2)x_2^*(L_2)$
	$x_1(L_1)x_1^*(L_2)x_2(L_2)$	$x_1(L_1) x_1(L_2) ^2x_1^*(L_2)x_2(L_2)$	\dots	$x_1(L_1) x_1(L_2) ^{(k-3)}x_1^*(L_2)x_2(L_2)$
	$x_2(L_1)x_1(L_2)x_2^*(L_2)$	$x_1(L_1) x_2(L_2) ^2x_1(L_2)x_2^*(L_2)$	\dots	$x_1(L_1) x_2(L_2) ^{(k-3)}x_1(L_2)x_2^*(L_2)$
	$x_2(L_1)x_1^*(L_2)x_2(L_2)$	$x_1(L_1) x_2(L_2) ^2x_1^*(L_2)x_2(L_2)$	\dots	$x_1(L_1) x_2(L_2) ^{(k-3)}x_1^*(L_2)x_2(L_2)$
		$x_1(L_1) x_1(L_2) ^2 x_2(L_2) ^2$	\dots	$x_1(L_1) x_1(L_2) ^{(k-1)/2} x_2(L_2) ^{(k-1)/2}$
		$x_2(L_1) x_1(L_2) ^2 x_2(L_2) ^2$	\dots	$x_2(L_1) x_1(L_2) ^{(k-1)/2} x_2(L_2) ^{(k-1)/2}$
		$x_2(L_1) x_1(L_2) ^2x_1(L_2)x_2^*(L_2)$	\dots	$x_2(L_1) x_1(L_2) ^{(k-3)}x_1(L_2)x_2^*(L_2)$
		$x_2(L_1) x_1(L_2) ^2x_1^*(L_2)x_2(L_2)$	\dots	$x_2(L_1) x_1(L_2) ^{(k-3)}x_1^*(L_2)x_2(L_2)$
		$x_2(L_1) x_2(L_2) ^2x_1(L_2)x_2^*(L_2)$	\dots	$x_2(L_1) x_2(L_2) ^{(k-3)}x_1(L_2)x_2^*(L_2)$
		$x_2(L_1) x_2(L_2) ^2x_1^*(L_2)x_2(L_2)$	\dots	$x_2(L_1) x_2(L_2) ^{(k-3)}x_1^*(L_2)x_2(L_2)$

as

$$\begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{H}_1 & 0 \\ 0 & \mathbf{H}_2 \end{bmatrix} \begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}, \quad (9)$$

where \mathbf{H}_1 and \mathbf{H}_2 are the regression matrices for channels 1 and 2, respectively, and θ_1 and θ_2 are the complex valued model parameters. The signals $v_1(n)$ and $v_2(n)$ are noise in channels 1 and 2, respectively, and are assumed to be mutually uncorrelated and to have zero-mean. As shown in (9), the parameter estimation for output channels 1 and 2 is decoupled, i.e., the parameters can be independently estimated. For simplicity, considering channel 1 of a 2×2 nonlinear system, the regression matrix \mathbf{H}_1 is

$$\mathbf{H}_1 = \begin{bmatrix} \phi_1(1) & \phi_2(1) & \dots & \phi_O(1) \\ \phi_1(2) & \phi_2(2) & \dots & \phi_O(2) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_1(N) & \phi_2(N) & \dots & \phi_O(N) \end{bmatrix}, \quad (10)$$

where $\phi_i(\cdot)$ are the basis functions of the model, O is the number of basis functions and N is the number of samples.

The behavior of the PA is captured in the construction of the basis functions, because the models proposed in Section II-C are linear in the parameters, linear least square estimation (LSE) [25] can be used to estimate the model parameters θ_1 by minimizing the cost function $S(\theta_1)$,

$$S(\theta_1) = \arg \min_{\theta_1} \|\mathbf{y}_1 - \mathbf{H}_1 \theta_1\|. \quad (11)$$

The LSE solution can be written in matrix form as,

$$\hat{\theta}_1 = (\mathbf{H}_1^* \mathbf{H}_1)^{-1} \mathbf{H}_1^* \mathbf{y}_1, \quad (12)$$

where \mathbf{H}_1 is the regression matrix and $\hat{\theta}_1$ contains the estimated parameters for channel 1 in a 2×2 MIMO transmitter. In a realistic scenario, the measurement process is also affected by I/Q imbalance [26] [27], measurement noise [28] and phase noise. Other estimators such as BLUE [25] can be used to reduce the variance of the estimated parameters if the color of the noise is known. However, the estimated parameters under zero-mean noise is unbiased with the LSE.

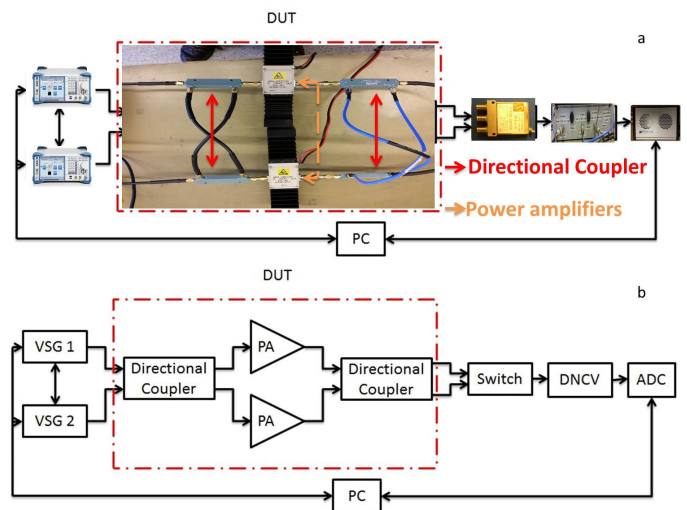


Fig. 2 a) Measurement setup. b) Outline of the measurement setup consisting of vector signal generators, a DUT with nonlinear and linear crosstalk, down converter, and an ADC.

Note that in partially non-coherent nonlinear MIMO transmitter (see Appendix), phase noise results in a bias term that scales the estimated parameters $\hat{\theta}$ if a standard linear least-squares estimation is used. To the best of the authors' knowledge, no methods that eliminates this bias have been proposed.

III. EXPERIMENTAL SETUP

The measurement setup is shown in Fig. 2. The setup consists of two Rohde & Schwartz SMBV100A signal generators, that are baseband synchronized and can be used with either a common external LO for coherent operation, or independently generated LO signals based on a common 10 MHz reference. To introduce crosstalk, directional couplers together with two Mini-Circuits ZVE-8G+ amplifiers were used as the devices under test (DUT). The measured RF signals operating at a carrier frequency of 2.14 GHz were downconverted to intermediate frequency (IF) signals using a wideband downconverter. An SP Devices ADQ214 analog-to-digital converter (ADC) was used for sampling

the IF signals. The ADC had a sampling rate of 400 MHz and a resolution of 14 bits. The sampling frequency and the number of samples were chosen such that an integer number of repeated periods were captured which is known as coherent sampling [29]. Note that the receiver was locked to the transmitter using the 10 MHz instrument reference. When effects of phase noise were tested, only one of the transmitters' 10 MHz reference was locked to the receiver.

Two separate sets of wideband code division multiple access (WCDMA) test signals were used, and each set consisted of 40,960 samples at a sampling rate of 30.72 MHz and a crest factor of 8.5 dB. Separate signal sets were used for identification and validation of the model performance. The measurements were taken with 100 coherent averages to minimize the impact of measurement noise. The performance limit of the measurement setup was limited to -68 dB adjacent channel power ratio.

The performance was evaluated in terms of normalized mean-square error (NMSE), adjacent channel error power ratio (ACEPR) and adjacent channel power ratio (ACPR). NMSE is defined as [30]

$$\text{NMSE} = \frac{\int \Phi_e(f) df}{\int \Phi_y(f) df}, \quad (13)$$

where $\Phi_y(f)$ is the power spectrum of the measured output signal and $\Phi_e(f)$ is the power spectrum of the difference between measured and the desired signal; integration is carried out across the available bandwidth. The ACEPR is defined as [30]

$$\text{ACEPR} = \frac{\int_{\text{adj. ch.}} \Phi_e(f) df}{\int_{\text{ch.}} \Phi_y(f) df}, \quad (14)$$

where the integration in the numerator is performed over the adjacent channel with maximum error power and in the denominator, integration is performed over the input channel. The ACPR is defined as [30]

$$\text{ACPR} = \frac{\int_{\text{adj. ch.}} \Phi_y(f) df}{\int_{\text{ch.}} \Phi_y(f) df}, \quad (15)$$

where in the numerator integration is performed over the adjacent channel with the largest amount of power; in the denominator, integration is performed over the input channel band.

IV. RESULTS

The performance of the proposed models was evaluated against the 2×2 PH model and the SISO GMP model in terms of behavioral modeling and DPD. For evaluation purposes, the following cases were studied; a) linear crosstalk (LC) case with -20 dB and -30 dB crosstalk levels, b) nonlinear crosstalk (NLC) case with -20 dB and -30 dB crosstalk levels, and c) nonlinear & linear crosstalk (NL&LC) case with -20 dB and -30 dB crosstalk levels. The proposed models were also evaluated without any crosstalk and are denoted as NC. In Section IV-C the effects of phase noise on the DPD performance are

presented for the transmitters in coherent and partially non-coherent modes.

For all of the measurements, the performance evaluation was conducted by setting the model order to be constant, i.e., a nonlinearity order of 9 was used, and the memory depths M_1 and M_2 were 2 and 5, respectively, for the proposed models and the SISO GMP. For the 2×2 PH model, the memory depth M was equal to 5. This selection of model order results in the smallest errors for the proposed models. Increasing the memory depth or nonlinearity order did not result in improved performance of the investigated models. Different model parameter pruning techniques e.g., least absolute shrinkage and selection operator (LASSO) [31] or principal component analysis [32] could be used to reduce the large number of model parameters while maintaining the performance. These techniques are out of the scope of this paper and thus are not discussed. Table II shows a comparison of the number of coefficients of the considered models.

Table II Total number of coefficients of respective models at nonlinearity order of 9 and a maximum memory depth of 5.

Model	Number of coefficients
SISO GMP	63
2×2 PH	180
GMPLC	126
GMPNLC	242
EGMPNLC	486
SISO Volterra	40116
MIMO Volterra	960924

A. Behavioral Modeling

Table III NMSE [dB] for given behavioral models and crosstalk type, the crosstalk level was -20 dB.

Model	LC	NLC	NL&LC	NC
SISO GMP	-23.8	-22.1	-22.9	-50.4
2×2 PH	-41.5	-40.1	-40.0	-41.8
GMPLC	-50.1	-42.3	-41.1	-50.3
GMPNLC	-50.3	-45.4	-45.1	-50.4
EGMPNLC	-50.2	-50.4	-50.3	-50.6

Table IV ACEPR [dB] for given behavioral models and crosstalk type, the crosstalk level was -20 dB.

Model	LC	NLC	NL&LC	NC
SISO GMP	-47.6	-47.3	-47.1	-60.6
2×2 PH	-56.6	-54.7	-54.3	-57.2
GMPLC	-58.9	-52.7	-51.9	-60.4
GMPNLC	-59.2	-58.9	-58.3	-60.7
EGMPNLC	-60.2	-60.1	-60.3	-60.7

1) *-20 dB crosstalk*: Tables III and IV summarize the performance of each model. In the absence of crosstalk (NC), all of the proposed models except 2×2 PH exhibited the same performance in terms of NMSE and ACEPR. A

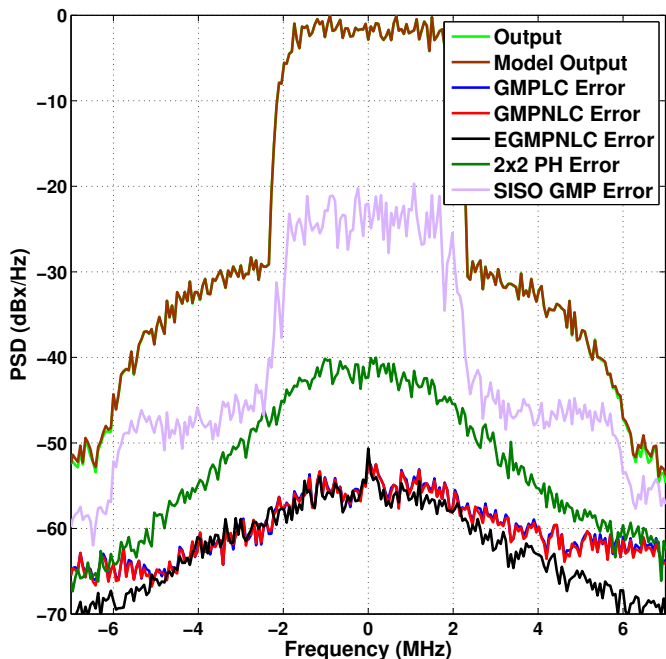


Fig. 3 Measured and modeled power spectral density (PSD) vs frequency for DUT with -20 dB *linear crosstalk* (LC). Also shown are the error spectra for the given models.

similar trend was observed in channel 2, and thus only the results from channel 1 are presented.

In the presence of LC, the proposed models had approximately the same NMSE of -50 dB and ACEPR of -59 to -60 dB, whereas the GMPLC model exhibited the largest errors and the EGMPNLC model exhibited the smallest errors. The 2×2 PH model had an NMSE of -41.5 dB, i.e., nearly 9 dB higher than the GMPLC model, which is expected considering that the model is not an adequate model for these PAs, as demonstrated by the NC case. As expected, the SISO GMP model was not able to model the LC, and is henceforth not further commented but only included for reference. Fig. 3 shows the error spectra of models for -20 dB LC.

In the presence of NLC, the difference between the performances of the proposed models is shown in Fig. 4. The GMPLC model has the largest NMSE compared to the EGMPNLC model, which had an NMSE of -50 dB, whereas the GMPNLC model had an NMSE of -45 dB. The difference in the performance of the proposed models is because the GMPLC model lacks crossterm combinations of input signals $x_1(n)$ and $x_2(n)$ that are required to model the NLC effect. The difference in performance between the GMPNLC and EGMPNLC models is also evident from Sections II-C2 and II-C3, in which EGMPNLC contains more combinations of crossterms, especially the conjugate crossterms i.e., $x_1(L_1)x_1(L_2)x_2^*(L_2)$, between the input signals compared to the GMPNLC model. A similar trend is observed when the performance of the proposed models is evaluated as measured by the ACEPR. The measurement results are summarized in Tables III and IV for the NMSE and ACEPR, respectively. As

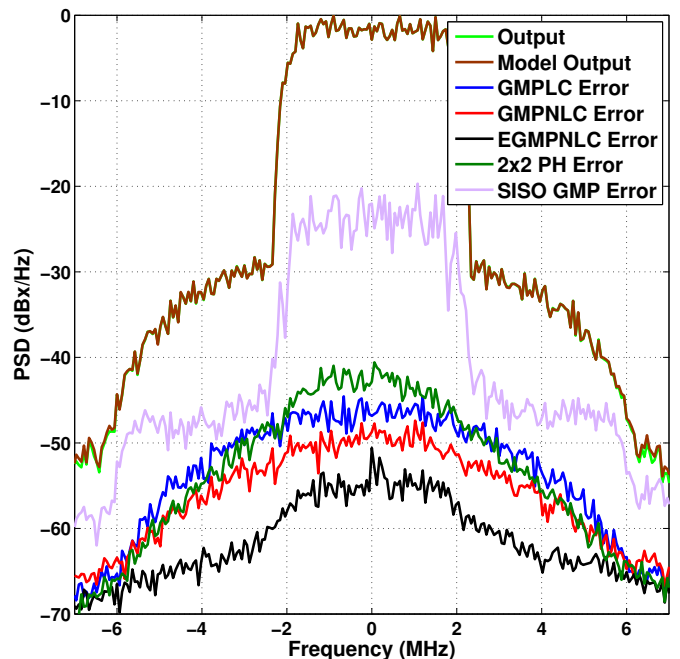


Fig. 4 Measured and modeled power spectral density (PSD) vs frequency for DUT with -20 dB *nonlinear crosstalk* (NLC). Also shown are the error spectra for the given models.

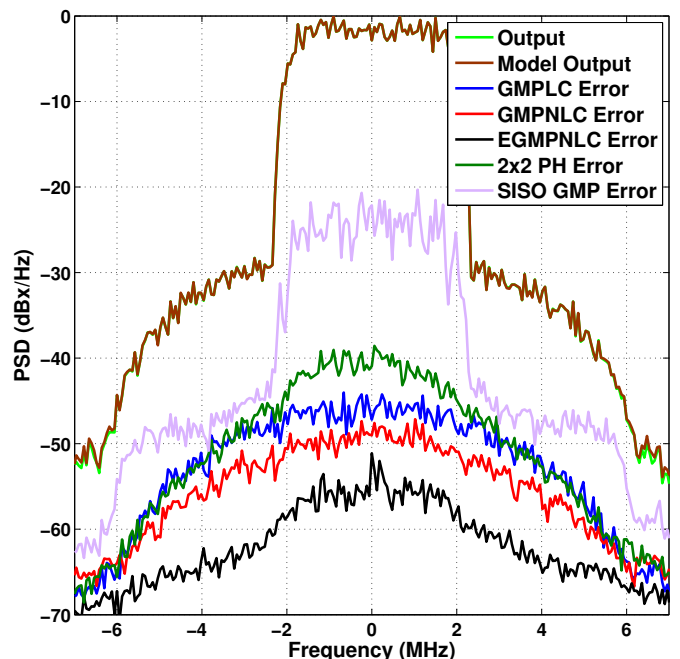


Fig. 5 Measured and modeled power spectral density (PSD) vs frequency for DUT with -20 dB *nonlinear & linear crosstalk* (NL&LC). Also shown are the error spectra for the given models.

measured by the ACEPR, the GMPLC model had slightly inferior performance compared to 2×2 PH model due to the lack of crossterms between the input signals. The crossterms included in the 2×2 PH model are $x_1(n-1)|x_1(n-m)|^{2(p-1)}|x_2(n-m)|^{2(q-1)}$ and $x_2(n-1)|x_2(n-m)|^{2(p-1)}|x_1(n-m)|^{2(q-1)}$.

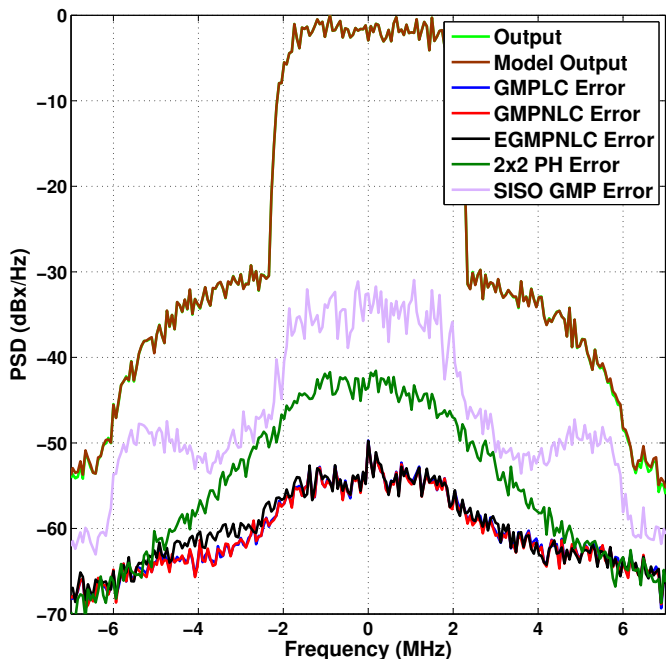


Fig. 6 Measured and modeled power spectral density (PSD) vs frequency for DUT with -30 dB *linear crosstalk* (LC). Also shown are the error spectra for the given models.

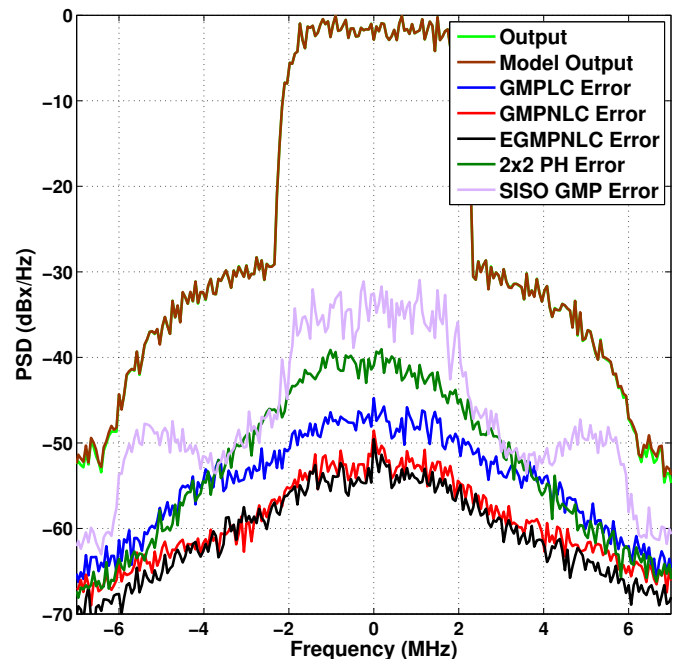


Fig. 7 Measured and modeled power spectral density (PSD) vs frequency for DUT with -30 dB *nonlinear crosstalk* (NLC). Also shown are the error spectra for the given models.

In the presence of NL&LC, Tables III and IV, indicate that the EGMPNLC model resulted in lower model error in terms of both the NMSE and ACEPR compared to the other proposed models, whereas the GMPNLC model exhibited higher performance compared to the GMPLC and 2×2 PH models. Compared to the GMPLC model, the 2×2 PH model resulted in slightly smaller model error when measured by the ACEPR. The improved performance of the EGMPNLC model compared to the GMPLC and GMPNLC models was described previously. The error spectra of the models in the presence of NL&LC are shown in Fig. 5. From Tables III and IV, under various crosstalk conditions, the EGMPNLC model in terms of NMSE and ACEPR resulted in approximately the same performance compared to the case in which no crosstalk is present.

2) *-30 dB crosstalk*: In the presence of -30 dB crosstalk and under all crosstalk conditions, the GMPNLC and EGMPNLC models exhibited the same performance in terms of both NMSE and ACEPR, as shown in Figs. 6, 7, and 8. For the NLC and NL&LC cases, the GMPLC model had an NMSE of -43 dB which is approximately 2 dB lower than the 2×2 PH model, whereas the GMPLC model had an ACEPR of approximately -54 dB which is 2 dB higher than the 2×2 PH model due to the lack of crossterms between the input signals. The results are summarized in Tables V and VI.

B. Digital Pre-distortion

Linearization of the system was performed by the use of an indirect learning architecture (ILA) [33]. The goal of the ILA is to determine the post-inverse of the DUT and use it as a pre-inverse, i.e., as a DPD algorithm. Three ILA

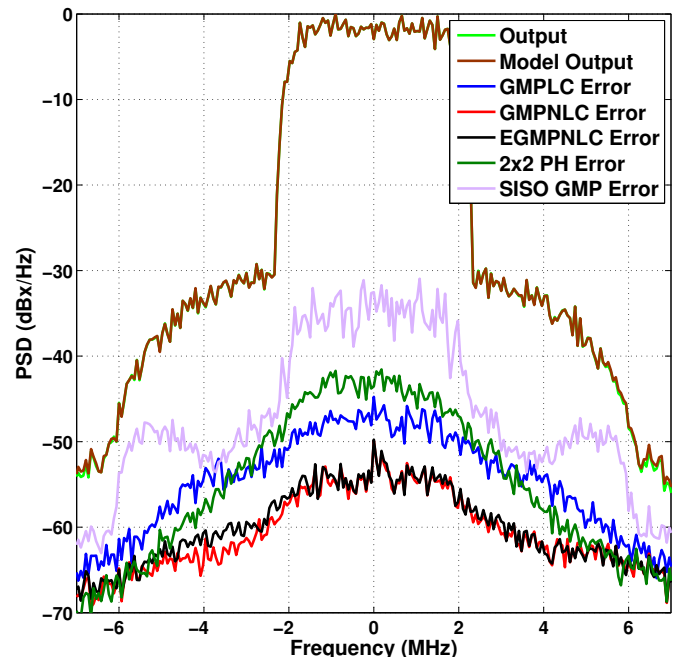


Fig. 8 Measured and modeled power spectral density (PSD) vs frequency for DUT with -30 dB *nonlinear & linear crosstalk* (NL&LC). Also shown are the error spectra for the given models.

iterations were performed for the DPD algorithm to fully converge to the operational point of the PAs. Linearization of the system can also be done by using the direct learning architecture (DLA), but it requires nonlinear least mean squares algorithms for estimating the DPD parameters [34] in contrast to ILA which is often used for linearization

Table V NMSE [dB] for given behavioral models and crosstalk type, the crosstalk level was -30 dB.

Model	LC	NLC	NL&LC	NC
SISO GMP	-32.4	-31.6	-31.4	-50.4
2×2 PH	-41.0	-41.3	-41.0	-41.8
GMPLC	-50.1	-43.3	-42.9	-50.3
GMPNLC	-50.3	-50.2	-49.9	-50.4
EGMPNLC	-50.5	-50.1	-50.2	-50.6

Table VI ACEPR [dB] for given behavioral models and crosstalk type, the crosstalk level was -30 dB.

Model	LC	NLC	NL&LC	NC
SISO GMP	-51.0	-51.6	-51.0	-60.6
2×2 PH	-57.2	-57.3	-56.1	-57.2
GMPLC	-60.2	-54.5	-54.7	-60.4
GMPNLC	-60.4	-60.0	-60.1	-60.7
EGMPNLC	-60.6	-60.7	-60.5	-60.7

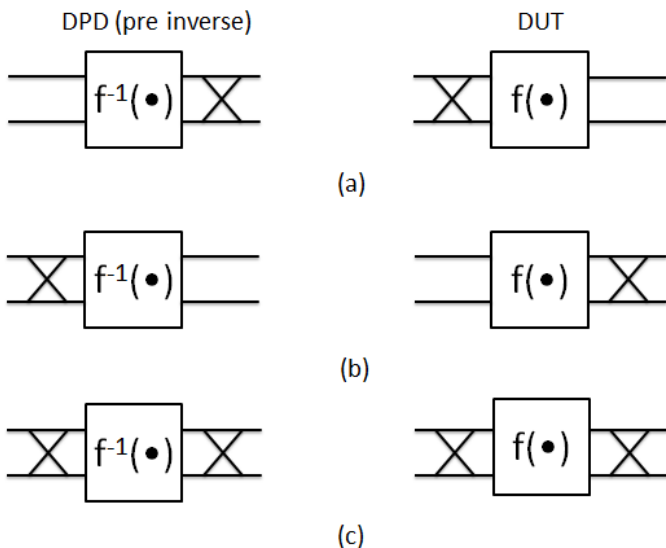


Fig. 9 The relationship between the DUT with different crosstalk types and the corresponding DPD structure.

of PA [11], [18]. Fig. 9 illustrates the DPD structures that were used for the DUTs under different crosstalk conditions.

1) *Digital predistortion for -20 dB crosstalk*: For a DUT with NLC, the proposed models exhibited the same DPD performance in terms of NMSE and ACEPR, as shown in Fig. 10. In presence of NLC, the DUT has the structure of crosstalk followed by a nonlinearity $f(\cdot)$. Therefore, by using the ILA for linearization, the current DPD algorithm exhibits the structure shown in Fig. 9a, i.e., the inverse of nonlinearity $f^{-1}(\cdot)$ followed by crosstalk. Furthermore, in Section IV-A, indicates that under the LC condition, the proposed models exhibited similar performance when used as direct models. Therefore, when used as an inverse model for a DUT with NLC, these models exhibit approximately the same DPD performance in terms of both NMSE and ACEPR. The measurement results are summarized in

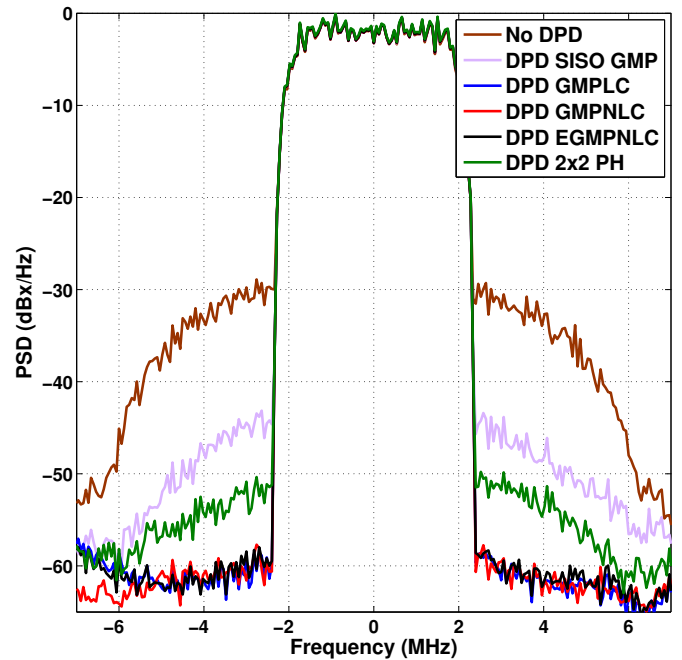


Fig. 10 Measured power spectral density (PSD) vs frequency for DPD of a DUT with -20 dB *nonlinear crosstalk* (NLC). The different inverse models used for DPD are explained in the legend.

Tables VII and VIII, which are in agreement with the previous discussion and the DPD structure shown in Fig. 9a. Compared to the proposed models, the 2×2 PH model had an NMSE of -40 dB which was 5 dB higher than the proposed model, and an ACEPR of -52 dB. The proposed models had an ACEPR of approximately -58 dB.

Table VII NMSE [dB] for given inverse models used for DPD and given crosstalk types of DUT. The crosstalk level was -20 dB.

Inverse Model	LC	NLC	NL&LC	NC
SISO GMP	-24.0	-24.5	-24.0	-49.4
2×2 PH	-39.7	-40.1	-38.1	-42.0
GMPLC	-41.2	-45.3	-37.3	-49.4
GMPNLC	-45.3	-45.1	-45.8	-49.5
EGMPNLC	-45.1	-45.4	-45.6	-49.3

Table VIII ACEPR [dB] for given inverse models used for DPD and given crosstalk types of DUT. The crosstalk level was -20 dB.

Inverse Model	LC	NLC	NL&LC	NC
SISO GMP	-45.0	-46.0	-44.7	-59.8
2×2 PH	-54.3	-52.1	-51.3	-58.2
GMPLC	-52.1	-58.6	-48.3	-59.8
GMPNLC	-58.2	-58.5	-58.7	-59.7
EGMPNLC	-58.4	-58.8	-57.9	-59.9

In the presence of LC, the DUT and its corresponding DPD structure are shown in Fig. 9b, where $f(\cdot)$ represents the nonlinear DUT function followed by crosstalk. Thus,

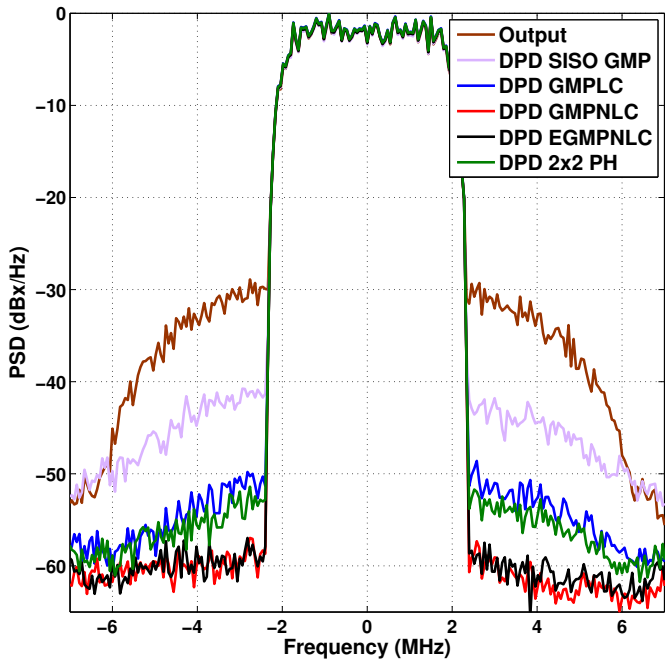


Fig. 11 Measured power spectral density (PSD) vs frequency for DPD of a DUT with -20 dB *linear crosstalk* (LC). The different inverse models used for DPD are explained in the legend.

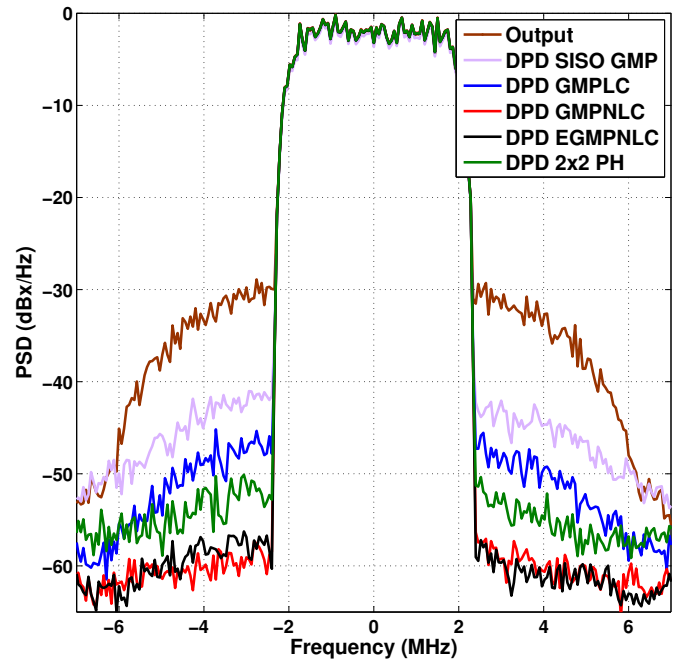


Fig. 12 Measured power spectral density (PSD) vs frequency for DPD of a DUT with -20 dB *nonlinear & linear crosstalk* (NL&LC). The different inverse models used for DPD are explained in the legend.

for a DUT with LC, the DPD structure can be described as crosstalk followed by an inverse of the nonlinear function $f^{-1}(\cdot)$. As shown in Fig. 11, the GMPNLC and EGMPNLC models exhibited similar DPD performance in terms of both NMSE and ACPR. The GPLC model had an NMSE of -41 dB, which was approximately 4 dB higher than the GMPNLC and EGMPNLC models. In terms of ACPR, the GPLC model had an ACPR that was 6 dB higher than those of GMPNLC and EGMPNLC. As shown in Section IV-A, model errors exhibited by the GMPNLC and EGMPNLC models when used as direct models for the DUT with NLC, are lower compared to the GPLC model. Therefore, for the DUT with LC, the use of the GMPNLC and EGMPNLC models as inverse models resulted in better performance than the GPLC model. The results are summarized in Tables VII and VIII. The inferior performance of GPLC compared to GMPNLC and EGMPNLC as an inverse model is due to the lack of nonlinear crossterms.

Fig. 12, shows the DPD performance of the proposed models in the presence of NL&LC. The DUT and its corresponding DPD structure are shown in Fig. 9c. From measurement results summarized in Tables VII and VIII, the GMPNLC and EGMPNLC models had 8 dB better performance as a measure of NMSE compared to the GPLC model. As measured by ACPR, the GMPNLC and EGMPNLC models exhibited 10 dB higher DPD performance than the GPLC model.

In order to compensate a DUT with LC and, NL&LC (cf. Fig. 9 b and c), a model should contain both linear and nonlinear crossterms. As shown in (6), the GPLC

model lacks such crossterms, therefore, when it was used as the inverse model for a DUT with LC and NL&LC, its performance as measured by NMSE and ACPR was lower than the GMPNLC and EGMPNLC models. Similarly due to the lack of crossterms, the GPLC model has slightly inferior performance compared to the 2×2 PH model for the cases with crosstalk.

2) *Digital predistortion for -30 dB crosstalk*: For the -30 dB crosstalk level, the DPD performance of the proposed models is summarized in Tables IX and X. For the NLC, the DUT and its corresponding DPD structure is shown in Fig. 9a. For such a DPD structure, the proposed models exhibited similar DPD performance in terms of NMSE and ACPR. Compared to the proposed models, the 2×2 PH model exhibited inferior performance in terms of both NMSE and ACPR. Even in the case with no crosstalk, the 2×2 PH model performance was approximately 7 dB and 2 dB higher in terms of NMSE and ACPR, respectively, compared to the other models. This result indicates that for these PAs, the 2×2 PH model showed inferior performance both as a direct and an indirect model.

In the presence of -30 dB LC, the GMPNLC and EGMPNLC models exhibited a similar DPD performance trend as the DUT with -20 dB LC. The GPLC model had an NMSE that was 9 dB higher than the GMPNLC and EGMPNLC models, whereas the GPLC models had an ACPR of -55 dB which was 4 dB higher compared to those of GMPNLC and EGMPNLC models. Compared to the 2×2 PH model, the GPLC model exhibited slightly lower DPD performance in terms of NMSE and ACPR.

In the presence of NL&LC, the GMPNLC and

EGMPNLC models had an NMSE and ACPR lower than -48 dB and -59 dB, respectively. As described in the previous section, the 2×2 PH model had slightly better performance as measured by NMSE and ACPR compared to the GMPLC model.

Table IX NMSE [dB] for given inverse models used for DPD and given crosstalk types of DUT. The crosstalk level was -30 dB.

Inverse Model	LC	NLC	NL&LC	NC
SISO GMP	-32.0	-32.6	-32.1	-49.4
2×2 PH	-41.6	-41.2	-39.8	-42.0
GMPLC	-40.2	-48.6	-41.0	-49.4
GMPNLC	-49.4	-48.4	-48.6	-49.5
EGMPNLC	-49.2	-48.6	-48.4	-49.3

Table X ACPR [dB] for given inverse models used for DPD and given crosstalk types of DUT. The crosstalk level was -30 dB.

Inverse Model	LC	NLC	NL&LC	NC
SISO GMP	-50.0	-50.4	-50.6	-59.8
2×2 PH	-56.4	-57.8	-57.1	-58.2
GMPLC	-55.4	-59.1	-54.0	-59.8
GMPNLC	-59.6	-59.5	-59.7	-59.7
EGMPNLC	-59.8	-59.7	-59.4	-59.9

C. Effect of phase noise on digital predistortion

A brief study was made to analyze the effect of coherent and partially non-coherent measurement on the identification of the digital predistortion parameters. As described in appendix, in partially non-coherent measurements, the transmitters do not share a common RF LO; therefore, the phase relationship between the transmitters does not remain constant. This phase noise affects the measurement and introduces bias terms in the LSE as analyzed in appendix. For this test, the measurement setup shown in Fig. 2 was modified such that the signal sources did not share a common RF LO but were only linked through the 10 MHz frequency reference. The crosstalk level was -30 dB and NLC was used. The measured signals were averaged 10 to 100 times to determine the effect of phase noise on coherent averaging. For coherent and partially non-coherent measurements, the GMPNLC model was used because, as shown in Tables IX and X, the GMPNLC and EGMPNLC models exhibited similar performance when used as a pre-inverse.

Fig. 13 shows the impact of phase noise on the performance of DPD. Partially non-coherent measurement resulted in an NMSE of -45.1 dB and ACPR of -56.1 dB whereas coherent measurement resulted in an NMSE of -48.51 dB and an ACPR of -59.4 dB for channel 1 when the coherent averaging was equals to 100. For this measurement setup, the difference in NMSE was 2.5-3 dB and 3-4 dB in ACPR depending on the additive noise level.

Fig. 13 indicates that even by increasing the number of averages, the performance of DPD as a measurement of

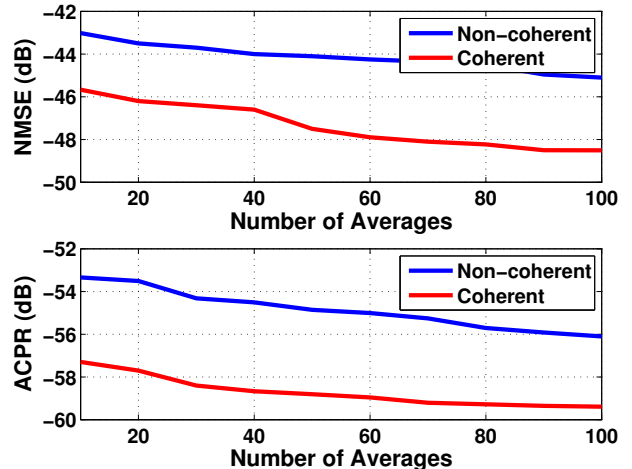


Fig. 13 Measured NMSE and ACPR vs number of averages of the output signal for DPD with coherent and partially non-coherent signal generation, respectively. The DUT had -30 dB nonlinear crosstalk and the inverse model for DPD was GMPNLC.

NMSE and ACPR for partially non-coherent measurement is lower compared to coherent measurement. This is due to the fact that when coherent averaging is done, it is assumed that the phase of the signal is the same at the beginning of each measured sample set but due to phase noise this assumption is not met, consequently partially non-coherent averaging produces a biased estimate as mentioned in appendix. Note that Fig. 13 illustrate that in presence of partially non-coherent measurement i.e., when the signal generators are only coupled with 10 MHz reference clock (see Appendix), the DPD performance will degrade compared to coherent measurement.

V. CONCLUSION

Three novel models for the direct modeling and linearization in the form of DPD of a nonlinear transmitter were considered in three crosstalk cases in a MIMO transmitter, and the results were compared to an existing PH based model. The considered crosstalk cases were linear crosstalk, nonlinear crosstalk and linear & nonlinear crosstalk. The proposed models that were used for the linearization of a DUT in the presence of nonlinear crosstalk exhibited similar performance and had lower model errors compared to the 2×2 PH models. Similarly, for linear and nonlinear & linear crosstalk, the GMPNLC and EGMPNLC models exhibited improved performance in both NMSE and ACPR than the GMPLC and 2×2 PH models. From the measurement results for the DPD, it can be concluded that for a DUT with nonlinear crosstalk only, GMPLC is sufficient for linearization. However for a DUT with linear and nonlinear & linear crosstalk, GMPLC lacks the necessary nonlinear crossterm combinations that are essential for the linearization.

The importance of phase coherency in MIMO transmitters' linearity when DPD is used is also illustrated. The ef-

fects on DPD performance for coherent and partially non-coherent RF LOs on the signal sources were considered for varying averages. For the used measurement setup, the difference in performance was 2.3 dB in NMSE for less than 40 averages but increases to more than 3.5 dB for 90 averages. The difference in ACPR was between 3-4 dB and also limited by the dynamic range of the measurement.

APPENDIX ANALYSIS OF PHASE NOISE IN PARTIALLY NON-COHERENT MIMO TRANSMITTER

In coherent MIMO systems, transmitters and receivers can share a common local oscillator (LO), which results in improved measurement accuracy and relaxed phase noise requirements in the system [20]. In this paper, measurements were taken with transmitters that did not use the same LO, rather, the RF LOs' coupled through the 10 MHz instrument frequency reference, which made the system partially non-coherent. Therefore, the relative phase between the transmitters did not remain constant, which resulted in phase noise and produced additional measurement errors. This section addresses the impact of phase noise on parameter estimation in behavioral modeling and DPD when standard LSE is used to estimate the parameters of the proposed models.

The analysis in [35] concluded that the phase noise of a free-running oscillator has a white Gaussian distribution with a variance that linearly increases with time. However, the oscillators in this case were not free-running but were coupled through phase-locked loops (PLL;s) to a common low frequency reference. Nevertheless, we assume that the phase noise in each sample can be modeled as additive white Gaussian noise $\Psi_{i,n}$ in channel i and sample n .

In the following the bias terms due to phase noise in the model parameters when using a linear LSE are derived. This derivation assumes that the receiver is the phase reference, i.e., there is no phase noise in the receiver; instead, all of phase noise contributions originate from the signal sources. The resulting equations are similar if one of the signal sources is used as a reference which is discussed at the end of this section.

Let $\mathbf{A}_1 = \text{diag}([e^{j\Psi_{1,1}} e^{j\Psi_{1,2}} \dots e^{j\Psi_{1,N}}])$ be the diagonal matrix that consist of the phase differences between each sample in all regressors involving the terms $x_1(n-m_1)|x_1(n-m_1-m_2)|^{2(p-1)}|x_2(n-m_1-m_2)|^{2(q-1)}$. Similarly, let \mathbf{A}_2 be a diagonal matrix similar to \mathbf{A}_1 that consist of all the regressors involving the terms $x_2(n-m_1)|x_2(n-m_1-m_2)|^{2(p-1)}|x_1(n-m_1-m_2)|^{2(q-1)}$. Note that under the current assumption, the transmitter noise is only observed in the $x_1(n-m_1)$ and $x_2(n-m_1)$ terms because these are the only terms without an absolute value.

The regression matrix \mathbf{H}_1 that was defined in (10) can be decomposed into two sub-matrices as, $\mathbf{H}_1 = [\mathbf{h}_1 \mathbf{h}_2]$. Each sub-matrix can be defined as $\mathbf{h}_1 = [\phi_{1,1} \phi_{1,2} \dots \phi_{1,o}]$ and $\mathbf{h}_2 = [\phi_{2,1} \phi_{2,2} \dots \phi_{2,o}]$, where $\phi_{1,1} = [\phi_{1,1}(1) \dots \phi_{1,1}(N)]^T$ and $\phi_{2,1} =$

$[\phi_{2,1}(1) \dots \phi_{2,1}(N)]^T$ are the column vectors in \mathbf{h}_1 and \mathbf{h}_2 respectively.

For a non-coherent 2×2 MIMO system, the output signal for channel 1 can be modeled as

$$\mathbf{y}_1 = [\mathbf{A}_1 \mathbf{h}_1, \mathbf{A}_2 \mathbf{h}_2] \begin{bmatrix} \theta_{1,1} \\ \theta_{2,1} \end{bmatrix} + \mathbf{v}, \quad (16)$$

where \mathbf{y}_1 is the output signal of channel 1, $\theta_1 = [\theta_{1,1}; \theta_{2,1}]$ are the complex valued parameters and \mathbf{v} is the additive zero mean noise. A standard linear LSE gives,

$$\begin{bmatrix} \hat{\theta}_{1,1} \\ \hat{\theta}_{2,1} \end{bmatrix} = (\mathbf{H}_1^* \mathbf{H}_1)^{-1} \mathbf{H}_1^* \mathbf{y}_1 = (\mathbf{H}_1^* \mathbf{H}_1)^{-1} \mathbf{H}_1^* \left[[\mathbf{A}_1 \mathbf{h}_1 \mathbf{A}_2 \mathbf{h}_2] \begin{bmatrix} \theta_{1,1} \\ \theta_{2,1} \end{bmatrix} + \mathbf{v} \right] = (\mathbf{H}_1^* \mathbf{H}_1)^{-1} \mathbf{H}_1^* [\mathbf{A}_1 \mathbf{h}_1 \mathbf{A}_2 \mathbf{h}_2] \begin{bmatrix} \theta_{1,1} \\ \theta_{2,1} \end{bmatrix} + \mathbf{w}, \quad (17)$$

where $\mathbf{H}_1 = [\mathbf{h}_1 \mathbf{h}_2]$ is the regression matrix composed of \mathbf{h}_1 and \mathbf{h}_2 and $\mathbf{w} = (\mathbf{H}_1^* \mathbf{H}_1)^{-1} \mathbf{H}_1^* \mathbf{v}$ is the weighted noise. By neglecting the noise term \mathbf{w} for simplicity, it follows that

$$\begin{bmatrix} \hat{\theta}_{1,1} \\ \hat{\theta}_{2,1} \end{bmatrix} = (\mathbf{H}_1^* \mathbf{H}_1)^{-1} \mathbf{H}_1^* [\mathbf{A}_1 \mathbf{h}_1 \mathbf{A}_2 \mathbf{h}_2] \begin{bmatrix} \theta_{1,1} \\ \theta_{2,1} \end{bmatrix} = (\mathbf{H}_1^* \mathbf{H}_1)^{-1} \begin{bmatrix} \mathbf{h}_1^* \mathbf{A}_1 \mathbf{h}_1 & \mathbf{h}_1^* \mathbf{A}_2 \mathbf{h}_2 \\ \mathbf{h}_2^* \mathbf{A}_1 \mathbf{h}_1 & \mathbf{h}_2^* \mathbf{A}_2 \mathbf{h}_2 \end{bmatrix} \begin{bmatrix} \theta_{1,1} \\ \theta_{2,1} \end{bmatrix}. \quad (18)$$

Using the properties of the respective terms of $\mathbf{h}_1^* \mathbf{A}_1 \mathbf{h}_1$ gives,

$$\mathbf{h}_1^* \mathbf{A}_1 \mathbf{h}_1 = \begin{bmatrix} \phi_{1,1}^* \mathbf{A}_1 \phi_{1,1} & \phi_{1,1}^* \mathbf{A}_1 \phi_{1,2} & \dots & \phi_{1,1}^* \mathbf{A}_1 \phi_{1,o} \\ \phi_{1,2}^* \mathbf{A}_1 \phi_{1,1} & \phi_{1,2}^* \mathbf{A}_1 \phi_{1,2} & \dots & \phi_{1,2}^* \mathbf{A}_1 \phi_{1,o} \\ \vdots & \ddots & \dots & \vdots \\ \phi_{1,o}^* \mathbf{A}_1 \phi_{1,1} & \phi_{1,o}^* \mathbf{A}_1 \phi_{1,2} & \dots & \phi_{1,o}^* \mathbf{A}_1 \phi_{1,o} \end{bmatrix}. \quad (19)$$

Each entry $\phi_{1,o}^* \mathbf{A}_1 \phi_{1,k}$ in (20) can now be computed as a weighted sum

$$\phi_{1,o}^* \mathbf{A}_1 \phi_{1,k} = \sum_{n=1}^N \phi_{1,o}^*(n) e^{j\Psi_{1,n}} \phi_{1,k}(n), \quad (20)$$

where $\Psi_{1,n}$ are the random variables. The random variables $\Psi_{1,n}$ are assumed to be Gaussian with zero mean due to the use of a common frequency reference and variance σ_1^2 . The probability density function (PDF) of $\Psi_{1,n}$ is

$$p(\Psi_{1,n}) = \frac{1}{\sigma_1 \sqrt{2\pi}} e^{-\frac{\Psi_{1,n}^2}{2\sigma_1^2}}, \quad (21)$$

the expected value of $e^{j\Psi_{1,n}}$ can be calculated as

$$E\{e^{\Psi_{1,n}}\} = \int_{-\infty}^{\infty} e^{j\Psi_{1,n}} p(\Psi_{1,n}) d\Psi = \quad (22a)$$

$$\int_{-\infty}^{\infty} e^{j\Psi_{1,n}} \frac{1}{\sigma_1 \sqrt{2\pi}} e^{-\frac{\Psi_{1,n}^2}{2\sigma_1^2}} d\Psi = \quad (22b)$$

$$\int_{-\infty}^{\infty} (\cos(\Psi_{1,n}) + j \sin(\Psi_{1,n})) \frac{1}{\sigma_1 \sqrt{2\pi}} e^{-\frac{\Psi_{1,n}^2}{2\sigma_1^2}} d\Psi = \quad (22c)$$

$$\int_{-\infty}^{\infty} \cos(\Psi_{1,n}) \frac{1}{\sigma_1 \sqrt{2\pi}} e^{-\frac{\Psi_{1,n}^2}{2\sigma_1^2}} d\Psi = \quad (22d)$$

$$2 \int_0^{\infty} \cos(\Psi_{1,n}) \frac{1}{\sigma_1 \sqrt{2\pi}} e^{-\frac{\Psi_{1,n}^2}{2\sigma_1^2}} d\Psi = \quad (22e)$$

$$\frac{2}{\sigma_1 \sqrt{2\pi}} \frac{1}{2} \sqrt{\frac{\pi}{\frac{1}{2\sigma_1^2}}} e^{-\frac{1}{2} \cdot \frac{1}{2\sigma_1^2}} = \quad (22f)$$

$$e^{-\frac{\sigma_1^2}{2}}, \quad (22g)$$

where (22d) follows from the fact that the Gaussian function is even and sine is odd over $[-\infty, \infty]$. Equation (22e) follows from that the argument is even over the interval and (22f) is from [36]: $\int_0^{\infty} e^{-at^2} \cos(2xt) dt = \frac{1}{2} \sqrt{\frac{\pi}{a}} e^{-\frac{x^2}{a}}$.

Using the result in (22g), the expected value of (19) is

$$E\{\mathbf{h}_1^* \mathbf{A}_1 \mathbf{h}_1\} = N e^{-\frac{\sigma_1^2}{2}} \mathbf{h}_1^* \mathbf{h}_1, \quad (23)$$

which is the bias term for all terms that contain \mathbf{A}_1 . Using (23) in (17) and computing the expected value gives,

$$E\left\{\begin{bmatrix} \hat{\theta}_{1,1} \\ \hat{\theta}_{2,1} \end{bmatrix}\right\} = (\mathbf{H}_1^* \mathbf{H}_1)^{-1} \begin{bmatrix} e^{-\frac{\sigma_1^2}{2}} \mathbf{I} & 0 \\ 0 & e^{-\frac{\sigma_2^2}{2}} \mathbf{I} \end{bmatrix} \begin{bmatrix} \theta_{1,1} \\ \theta_{2,1} \end{bmatrix} \mathbf{H}_1^* \mathbf{H}_1. \quad (24)$$

If $\sigma = \sigma_1 = \sigma_2$, it is possible to move out the term $e^{-\frac{\sigma^2}{2}}$ in front of (24), which shows that the estimated parameters are scaled by the constant $e^{-\frac{\sigma^2}{2}}$. However, in general this is not possible. Considering the effects when $\sigma_1 \neq \sigma_2$, let

$$\mathbf{G} = \begin{bmatrix} \mathbf{G}_{11} & \mathbf{G}_{12} \\ \mathbf{G}_{21} & \mathbf{G}_{22} \end{bmatrix} = (\mathbf{H}_1^* \mathbf{H}_1)^{-1} = \mathbf{F}^{-1}, \quad (25)$$

$$\mathbf{F} = \begin{bmatrix} \mathbf{F}_{11} & \mathbf{F}_{12} \\ \mathbf{F}_{21} & \mathbf{F}_{22} \end{bmatrix}, \quad (26)$$

and

$$\mathbf{D}_i = e^{-\frac{\sigma_i^2}{2}} \mathbf{I}, \quad (27)$$

where \mathbf{I} is the identity matrix of appropriate size; then

$$(\mathbf{H}_1^* \mathbf{H}_1)^{-1} \begin{bmatrix} e^{-\frac{\sigma_1^2}{2}} \mathbf{I} & 0 \\ 0 & e^{-\frac{\sigma_2^2}{2}} \mathbf{I} \end{bmatrix} \mathbf{H}_1^* \mathbf{H}_1 = \begin{bmatrix} \mathbf{G}_{11} & \mathbf{G}_{12} \\ \mathbf{G}_{21} & \mathbf{G}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{D}_1 \mathbf{F}_{11} & \mathbf{D}_1 \mathbf{F}_{12} \\ \mathbf{D}_2 \mathbf{F}_{21} & \mathbf{D}_2 \mathbf{F}_{22} \end{bmatrix} = \begin{bmatrix} \mathbf{D}_1 \mathbf{G}_{11} \mathbf{F}_{11} + \mathbf{D}_2 \mathbf{G}_{12} \mathbf{F}_{21} & \mathbf{D}_1 \mathbf{G}_{11} \mathbf{F}_{12} + \mathbf{D}_2 \mathbf{G}_{12} \mathbf{F}_{22} \\ \mathbf{D}_1 \mathbf{G}_{21} \mathbf{F}_{11} + \mathbf{D}_2 \mathbf{G}_{22} \mathbf{F}_{21} & \mathbf{D}_1 \mathbf{G}_{21} \mathbf{F}_{12} + \mathbf{D}_2 \mathbf{G}_{22} \mathbf{F}_{22} \end{bmatrix}. \quad (28)$$

Unless $\mathbf{D}_1 = \mathbf{D}_2$, this equation cannot be a scaled version of the identity matrix. Therefore, for a system with more than one input, phase noise not only results in a bias term that scales the estimated parameters $\hat{\theta}$ depending on the variance of the phase noise increments, but also mixes the parameters from the two parts $\theta_{1,1}$ and $\theta_{2,1}$ if a standard linear least-squares estimation is used. To the best of the authors' knowledge, no methods that eliminates this bias have been proposed.

As previously mentioned, the result is similar if one of the sources is considered to be a phase reference, i.e., there is no phase noise. We assume that the receiver and second signal source are noisy and we let \mathbf{B}_0 and \mathbf{A}_2 be the phase noise matrices as previously described. Since the first signal source is the phase reference, i.e., $\mathbf{A}_1 = \mathbf{I}$, the LSE can be written as

$$\begin{bmatrix} \hat{\theta}_{1,1} \\ \hat{\theta}_{2,1} \end{bmatrix} = (\mathbf{H}_1^* \mathbf{H}_1)^{-1} \mathbf{H}_1^* [\mathbf{B}_0 \mathbf{h}_1 \quad \mathbf{B}_0 \mathbf{A}_2 \mathbf{h}_2] \begin{bmatrix} \theta_{1,1} \\ \theta_{2,1} \end{bmatrix} = (\mathbf{H}_1^* \mathbf{H}_1)^{-1} \begin{bmatrix} \mathbf{h}_1^* \mathbf{B}_0 \mathbf{h}_1 & \mathbf{h}_1^* \mathbf{B}_0 \mathbf{A}_2 \mathbf{h}_2 \\ \mathbf{h}_2^* \mathbf{B}_0 \mathbf{h}_1 & \mathbf{h}_2^* \mathbf{B}_0 \mathbf{A}_2 \mathbf{h}_2 \end{bmatrix} \begin{bmatrix} \theta_{1,1} \\ \theta_{2,1} \end{bmatrix}, \quad (29)$$

which is the same form as the LSE in (18). In (29), \mathbf{B}_0 and \mathbf{A}_2 are two different phase noise matrices since for the partially non-coherent case, always two different LOs are used.

Note that the phase noise also causes the coherent averaging to produce a biased estimate that does not improve with the number of averages compared to additive noise [28], in which the bias terms tend to zero by increasing the number of averages.

REFERENCES

- [1] M. Isaksson and D. Rönnow, "A parameter-reduced Volterra model for dynamic RF power amplifier modeling based on orthonormal basis functions," *Int. J. RF and Microw Comp Aid Eng*, 17, pp. 542–551, 2007.
- [2] A. Zhu and T. Brazil, "Behavioral modeling of RF power amplifiers based on pruned Volterra series," *IEEE Microw. Wireless Compon. Lett.*, vol. 14, no. 12, pp. 563–565, 2004.
- [3] F. Ghannouchi, M. Younes, and M. Rawat, "Distortion and impairments mitigation and compensation of single- and multi-band wireless transmitters," *IET Microw. Antennas Propag.*, vol. 7, no. 7, 2013.
- [4] M. Schetzen, "Theory of P^{th} -order inverses of nonlinear systems," *IEEE Trans Circuits Syst.*, vol. 23, no. 5, pp. 285–291, 1976.
- [5] M. Isaksson, D. Wisell, and D. Rönnow, "A comparative analysis of behavioral models for RF power amplifiers," *IEEE Trans. Microw. Theory Techn.*, vol. 54, no. 1, pp. 348–359, 2006.
- [6] S. Maas and J. Pedro, "Implementation of a Volterra behavioral model for system simulation," in *IEEE MTT-S Int. Microw. Symp.*, 2008, pp. 943–946.
- [7] A. Bennadji, E. Ngoya, and R. Quere, "Implementation of behavioral models for amplifiers in system level simulation and RF circuit-system co-simulation," *Eng. Lett. Int. Accoc. Eng.*, 2009.
- [8] D. Zhou and V. DeBrunner, "Novel adaptive nonlinear predistorters based on the direct learning algorithm," *IEEE Trans. Signal Process.*, vol. 55, no. 1, pp. 120–133, 2007.
- [9] T. Cunha, E. Lima, and J. Pedro, "Validation and physical interpretation of the power-amplifier polar Volterra model," *IEEE Trans. Microw. Theory Techn.*, vol. 58, no. 12, pp. 4012–4021, 2010.

- [10] M. Herman, B. Miller, and J. Goodman, "The cube coefficient subspace architecture for nonlinear digital predistortion," in *42nd Asilomar Conference on Signals, Systems and Computers, 2008*, 2008, pp. 1857–1861.
- [11] S. Bassam, M. Helaoui, and F. Ghannouchi, "Crossover digital predistorter for the compensation of crosstalk and nonlinearity in MIMO transmitters," *IEEE Trans. Microw. Theory Techn.*, vol. 57, no. 5, pp. 1119–1128, 2009.
- [12] A. Swain and S. Billings, "Generalized frequency response function matrix for MIMO non-linear systems," *Int. J. Control*, pp. 829–844, 2001.
- [13] R. Zayani, R. Bouallegue, and D. Roviras, "Crossover neural network predistorter for the compensation of crosstalk and nonlinearity in MIMO OFDM systems," in *IEEE 21st International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC)*, 2010, pp. 966–970.
- [14] S. Bassam, W. Chen, M. Helaoui, F. Ghannouchi, and Z. Feng, "Linearization of concurrent dual-band power amplifier based on 2D-DPD technique," *IEEE Microw. Wireless Compon. Lett.*, vol. 21, no. 12, pp. 685–687, 2011.
- [15] J. Moon, P. Saad, J. Son, C. Fager, and B. Kim, "2-D enhanced Hammerstein behavior model for concurrent dual-band power amplifiers," in *42nd European Microwave Conference (EuMC)*, 2012, pp. 1249–1252.
- [16] L. Ding, Z. Yang, and H. Gandhi, "Concurrent dual-band digital predistortion," in *IEEE MTT-S Int. Microw. Symp.*, 2012, pp. 1–3.
- [17] Y.-J. Liu, W. Chen, J. Zhou, B.-H. Zhou, and F. Ghannouchi, "Digital predistortion for concurrent dual-band transmitters using 2-D modified memory polynomials," *IEEE Trans. Microw. Theory Techn.*, vol. 61, no. 1, pp. 281–290, 2013.
- [18] D. Saffar, N. Boulejfen, F. Ghannouchi, A. Gharsallah, and M. Helaoui, "Behavioral modeling of MIMO nonlinear systems with multivariable polynomials," *IEEE Trans. Microw. Theory Techn.*, vol. 59, no. 11, pp. 2994–3003, 2011.
- [19] F. Dai, Y. Shi, J. Yan, and X. Hu, "MIMO RFIC Transceiver Designs for WLAN Applications," in *7th International Conference on ASIC, ASICON '07.*, Oct 2007, pp. 348–351.
- [20] Y. Palaskas, A. Ravi, S. Pellerano, B. Carlton, M. Elmala, R. Bishop, G. Banerjee, R. Nicholls, S. Ling, N. Dinur, S. Taylor, and K. Soumyanath, "A 5-GHz 108-Mb/s 2x2 MIMO Transceiver RFIC With Fully Integrated 20.5-dBm P_{1dB} power amplifiers in 90-nm CMOS," *IEEE J. Solid-State Circuits*, vol. 41, no. 12, pp. 2746–2756, 2006.
- [21] T. Öberg, *Modulation Detection and Coding*. West Sussex, England: John Wiley & Sons, 2001.
- [22] A. Zhu and T. Brazil, "An overview of Volterra series based behavioral modeling of RF/microwave power amplifiers," in *Wireless and Microwave Technology Conference, 2006. WAMICON '06. IEEE Annual*, 2006, pp. 1–5.
- [23] H.-W. Chen, "Modeling and identification of parallel nonlinear systems: structural classification and parameter estimation methods," *Proc. IEEE*, vol. 83, no. 1, pp. 39–66, 1995.
- [24] 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, 2001.
- [25] S. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Upper Saddle River, New Jersey, USA: Prentice-Hall PTR, 1998.
- [26] L. Anttila, P. Handel, and M. Valkama, "Joint mitigation of power amplifier and I/Q modulator impairments in broadband direct-conversion transmitters," *IEEE Trans. Microw. Theory Techn.*, vol. 58, no. 4, pp. 730–739, April 2010.
- [27] F. Gregorio, J. Cousseau, S. Werner, T. Riihonen, and R. Wichman, "Power amplifier linearization technique with IQ imbalance and crosstalk compensation for broadband MIMO-OFDM transmitters," *EURASIP J. Adv. Signal Process*, 2011.
- [28] S. Amin, E. Zenteno, P. Landin, D. Ronnow, M. Isaksson, and P. Handel, "Noise impact on the identification of digital predistorter parameters in the indirect learning architecture," in *Swedish Communication Technologies Workshop (Swe-CTW)*, 2012, pp. 36–39.
- [29] "IEEE standard for terminology and test methods for analog-to-digital converters," *IEEE Std 1241-2010 (Revision of IEEE Std 1241-2000)*, pp. 1–139, January 2011.
- [30] P. Landin, M. Isaksson, and P. Handel, "Comparison of evaluation criteria for power amplifier behavioral modeling," in *IEEE MTT-S Int. Microw. Symp.*, 2008, pp. 1441–1444.
- [31] D. Wisell, J. Jalden, and P. Handel, "Behavioral Power Amplifier Modeling Using the Lasso," in *IEEE IMTC 2008.*, May 2008, pp. 1864–1867.
- [32] P. Gilabert, G. Montoro, D. Lopez, N. Bartzoudis, E. Bertran, M. Payaro, and A. Hourtane, "Order reduction of wideband digital predistorters using principal component analysis," in *IEEE MTT-S Int. Microw. Symp.*, June 2013, pp. 1–7.
- [33] 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.
- [34] M. Abi Hussein, V. Bohara, and O. Venard, "On the system level convergence of 1la and dla for digital predistortion," in *Int. Symp. on Wireless Commun. Syst.*, Aug 2012, pp. 870–874.
- [35] A. Demir, A. Mehrotra, and J. Roychowdhury, "Phase noise in oscillators: a unifying theory and numerical methods for characterization," *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, vol. 47, no. 5, pp. 655–674, 2000.
- [36] L. Råde and B. Westergren, *Mathematics Handbook for Science and Engineering*. Lund, Sweden: Studentlitteratur AB, 2003.



Shoaib Amin (S'12) received the B.E degree in avionics engineering from College of Aeronautical Engineering, National University of Sciences and Technology, Islamabad, Pakistan in 2008 and the M.Sc. degree in electronics and telecommunications from the University of Gävle, Gävle, Sweden in 2011.

He is currently working toward the Ph.D degree at University of Gävle, Gävle, Sweden, and the KTH Royal Institute of Technology. His main interest is signal processing techniques for modeling and linearization of MIMO power amplifiers.



Per N. Landin (S'07-M'12) received the M.Sc. degree from Uppsala University, Uppsala, Sweden in 2007 and the Ph.D degree jointly from KTH Royal Institute of Technology, Stockholm, Sweden and Vrije Universiteit Brussel, Belgium, in 2012.

He is currently working as a post-doc researcher at Chalmers University of Technology, Gothenburg, Sweden. His main research interests are system identification and linearization of nonlinear high frequency devices, specially MIMO power amplifiers.



Peter Händel (S'88-M'94-SM'98) received the Ph.D. degree from Uppsala University, Uppsala, Sweden, in 1993.

From 1987 to 1993, he was with Uppsala University. From 1993 to 1997, he was with Ericsson AB, Kista, Sweden. From 1996 to 1997, he was also with the Tampere University, Finland. Since 1997, he has been with the Royal Institute of Technology, Stockholm, Sweden, where he is currently a Professor of Signal Processing and Head of the Department of Signal Processing. From 2000 to 2006, he worked part time with Swedish Defence Research Agency. In 2010, he spent time at the Indian Institute of Science (IISc), Bangalore, as Guest Professor. He is currently a Guest Professor with the University of Gävle, Gävle, Sweden.

Dr. Händel has served as an associate editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING.



Daniel Rönnow received the M.Sc. degree in engineering physics in 1991 and the Ph.D. degree in solid-state physics in 1996 from Uppsala University, Uppsala, Sweden.

From 1996 to 1998, he was with Max-Planck-Institut für Festkörperforschung, Stuttgart, Germany, working on semiconductor physics. From 1998 to 2000, he was with Acreo AB, Stockholm, Sweden, working on infrared sensors and systems. From 2000 to 2004, he was a

Technical Consultant and the Head of Research of Racomna AB, Uppsala, where he worked on power amplifier linearization and "smart" materials for microwave applications. From 2004 to 2006, he was a University Lecturer with the University of Gävle, Gävle, Sweden. From 2006 to 2011, he was a Senior Sensor Engineer with Westerngeco, Oslo, Norway, where he was involved with signal processing and seismic sensors. In 2011, he became a Professor of electronics with the University of Gävle. Since 2000, he has been an Associate Professor with Uppsala University. He has authored or coauthored more than 45 peer-reviewed papers. He holds eight patents. His current research interests are RF measurement techniques and linearization of nonlinear RF circuits and systems.