IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 13 December 2024, accepted 10 January 2025, date of publication 16 January 2025, date of current version 23 January 2025.

Digital Object Identifier 10.1109/ACCESS.2025.3530817

RESEARCH ARTICLE

Joint Channel and Nonlinearity Estimation for Memoryless Nonlinear Systems

ZAHRA MOKHTARI[1], RUI DINIS[2], (Senior Member, IEEE), SHA HU[3], AND DZEVDAN KAPETANOVIC[3]

[1]Instituto de Telecomunicações, 1040-001 Lisbon, Portugal
[2]Instituto de Telecomunicações, FCT-UNL, 2829-516 Caparica, Portugal
[3]Lund Research Center, Huawei Technologies Sweden AB, 223 63 Lund, Sweden

Corresponding author: Zahra Mokhtari (zahra.mokhtari@lx.it.pt)

This work was supported by Lund Research Center, Huawei Technologies Sweden AB, Sweden, and FCT (Fundação para a Ciência e Tecnologia) and Instituto de Telecomunicações under projects UIDB/50008/2020 (DOI identifier https://doi.org/10.54499/UIDB/50008/2020) and CELL-LESS6G (reference 2022.08786.PTDC and DOI identifier https://doi.org/10.54499/2022.08786.PTDC).

ABSTRACT System nonlinearity due to hardware impairments has always been a challenging issue. Distortion cancellation and iterative detection based receivers such as the Bussgang Noise Cancelling (BNC) receiver are used to detect the original data in the presence of strong nonlinear (NL) effects. However, these receivers require knowledge of the system nonlinearity which is usually unknown in practical systems. Bussgang decomposition and its general form denoted Generalized Bussgang decomposition (GBD), have been commonly used to model system nonlinearity. In GBD the nonlinearity output is decomposed as the sum of uncorrelated terms of increased orders and provides spectral characteristics of the useful and distortion terms. In this paper we consider nonlinearity at the transmitter side and model it with GBD. We aim to estimate the scalar weights in the GBD to later use them at the BNC receiver. However, knowledge of the channel is required to make a reliable estimate of the NL parameters. On the other hand the pilots for channel estimation are affected by the system nonlinearity, which can preclude reliable channel estimation. Therefore, in this paper we propose a joint channel and NL parameter estimation technique by designing appropriate training signals for each estimation phase (i.e. channel estimation and NL parameter estimation). We also derive a closed form expression for the average power of residual distortion in GBD with estimated parameters to see how well this model can characterize the nonlinearity. The results show that the proposed estimation technique has good accuracy and enables quasi-ideal performance for a BNC receiver.

INDEX TERMS Generalized Bussgang decomposition, nonlinear effects, nonlinear parameter estimation, OFDM.

I. INTRODUCTION

Nonlinear (NL) distortion is one of the main impairments in communications systems and can lead to a significant performance loss. This distortion can be intentional (e.g., digital clipping effects to reduce the envelope fluctuations of signals) or not (e.g., due to hardware imperfections such as low resolution quantizers, NL amplifiers, etc.) [1], [2]. This is particularly important for signals with large envelope fluctuations such as OFDM-like (Orthogonal Frequency Division Multiplexing) signals [3]. To understand the impact of NL effects and to develop appropriate counter measures, it is crucial to characterize the system nonlinearity with an accurate and suitable model.

The Bussgang decomposition is a widely used model for studying the performance of communication systems that involve nonlinearities [4], [5], [6], [7], [8], [9], [10], [11], [12]. This allows us to write the signal at the nonlinearity output as the sum of a useful term and a NL distortion term, with the two being uncorrelated. In the Gaussian case, the useful component is proportional to the input signal, although this is not necessarily true for other input distributions [13].

The associate editor coordinating the review of this manuscript and approving it for publication was Fang Yang.

The Bussgang decomposition has limitations, since it does not provide spectral characteristics of the useful and distortion terms. Moreover, the distortion term has information of the input signal, which can be exploited for performance enhancement. Therefore, a more complete model named the Generalized Bussgang Decomposition (GBD) has been introduced and used in the literature [14], [15], [16], [17], [18]. GBD provides spectral characteristics of the useful and distortion terms by decomposing the nonlinearity output as the sum of uncorrelated terms of increased orders. Orthogonal polynomials can be used as basis functions in the GBD to simplify the model and allow a modular approach, in the sense that increasing the number of parameters (to increase the accuracy of the nonlinearity model), does not affect the previous parameters [17].

Several receivers have been proposed for detecting desired signals in NL systems [19], [20], [21], [22]. The simplest approach is to employ the so-called Bussgang noise cancelling (BNC) receiver which iteratively attempts to estimate and cancel out the distortion factor present in the nonlinearly distorted signals [19]. A more sophisticated approach is to take advantage of the information on the signals to be transmitted that is inherent to the NL distortion term to improve the performance [11]. To achieve this, several receivers based on the generalized approximate message passing (GAMP) [23] were proposed to extract the information contained within the distortion term in NL signals [21], [22]. However, to be effective, both BNC and GAMP receivers need to know the NL characteristic employed in the transmitter with good accuracy.

Some works on NL parameter estimation have been done in the literature [24], [25], [26], [27], [28]. The authors in [24] and [25] address clipping level estimation and memoryless AM/PM nonlinearities estimation in an OFDM system with AWGN channel, respectively. Unlike [24] and [25], the authors in [26] have considered the channel effect on NL parameter estimation and proposed an iterative channel and clipping level estimation with block type pilots. However, the proposed technique has very high complexity and can only be applied for estimation of a specific nonlinearity type (i.e. clipping). More general NL functions have been considered in [27] and [28] by using polynomial based models for the nonlinearity. In [27] an iterative gradient descent based algorithm for NL parameter estimation has been proposed. However, the proposed technique needs several OFDM symbols for convergence. In [28] an iterative joint estimation of channel and NL parameters has been proposed using comb type pilots. However, for low training overheads (i.e., by increasing the distance between pilots) the performance of the technique is not accurate enough and, due to the iterations, the proposed technique has high complexity. To the best of our knowledge there has been no effort on designing training signals (TS) for joint channel and NL parameter estimation with high accuracy and low complexity.

In this paper, we consider the system nonlinearity at the transmitter side. Many nonlinearities in practice are memoryless or can be modeled with memoryless nonlinearity with high accuracy. For example, amplitude clipping for reducing PAPR, AM/AM and AM/PM conversion in nonlinear power amplifiers such as SSPA and TWTA model, cubic and multiplicative nonlinearities in mixers, etc. Also, for some nonlinearities with memory, techniques such as pre-distortion can mitigate the memory effects but the saturation is unavoidable. Therefore, in this paper we study memoryless nonlinearities and since we do not have knowledge of the system nonlinearity, we use the GBD model with orthogonal polynomials as basis functions to model the nonlinearity generally. We aim to estimate the scalar parameters in the GBD model to accurately model the NL function. For this purpose, we require knowledge of the channel to compensate the effect of the channel on the received signal before estimating the NL parameters. However, the pilots for channel estimation are affected by the nonlinearity, which results in an unreliable channel estimation. Therefore, we propose a joint channel and NL parameters estimation technique based on two TSs (one for channel estimation and the other for NL parameters estimation) to overcome this problem. We design the TSs regarding the requirements of each estimation phase to make an accurate and low complex estimation of the channel and the NL function. We should note that since the variation of the system nonlinearity is slow, the TS for NL parameters estimation is not transmitted frequently and therefore it does not reduce the spectral efficiency of the system. We also derive closed form expression for the average power of the residual distortion in GBD with the estimated parameters for both Cartesian and polar nonlinearities with complex Gaussian inputs.

Our performance results show that the channel estimation in a NL system with our designed TS can perform closely to channel estimation with demodulation reference signals (DMRS) [29] in a linear system. The results also indicate that the NL parameters estimated with the proposed technique have good accuracy and performance in BNC receivers. The residual distortion analysis shows that each additional term in the GBD can help to reduce the residual distortion, only if the power of estimation error is smaller than that of the weight itself.

The main contributions of this paper are summarized as follows:

- Present a new approach for jointly estimating the channel and GBD scalar weights with high accuracy and low complexity, leveraging training signal design rather than conventional pilot-based methods.
- Propose a two-phase training signal design tailored for low complex and accurate estimation of both the channel and multiple GBD terms, ensuring practical applicability.
- Derive closed form expression for the average power of the residual distortion in GBD with the estimated parameters, offering insights into system performance.

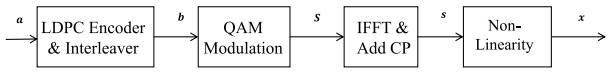

FIGURE 1. Block diagram of an OFDM transmitter with nonlinearity.

- Study the performance of the proposed joint nonlinearity and channel estimation technique in BNC receivers.

The rest of the paper is organized as follows. In Sec. II the system model of OFDM with nonlinearity at the transmitter is explained. Sec. III provides an overview of generalized Bussgang model and the orthogonal polynomials. Sec. IV introduces the proposed joint channel and nonlinearity estimation technique. In Sec. V the closed form expression for the average power of the residual distortion is derived. Sec. VI presents a set of perofmance results for the proposed technique and, finally, Sec. VII concludes the paper.

II. SYSTEM MODEL

In the rest of the paper, we present the vectors and matrices in bold. We use $(.)^*$, $(.)^T$, $(.)^H$ and $(.)^{-1}$ to denote conjugate, transpose, hermitian transpose and inverse of the argument, respectively.

In this section we introduce the system model that will be considered in this paper. We consider an OFDM transmission scheme in a single input single output (SISO) scenario. OFDM is a multicarrier transmission technique with orthogonal subcarriers and can be implemented digitally by means of the fast Fourier transform (FFT). Fig. 1 shows the block diagram of an OFDM transmitter with nonlinearity due to hardware impairments, which will be assumed to be memoryless.

We assume a vector of P data bits as $\boldsymbol{a} = [a_0, a_1, \ldots, a_{P-1}]$ that needs to be transmitted. The bits are encoded and interleaved with an encoder with rate r_c, where $r_c = P/Q$ and Q is the number of output bits of the encoder. The coded bits $\boldsymbol{b} = [b_0, b_1, \ldots, b_{Q-1}]$ are passed through an M-QAM modulator that generate the set of complex data symbols $\boldsymbol{S} = [S_0, S_1, \ldots, S_{N-1}]$, where $N = Q/\log_2(M)$ is the number of subcarriers. The modulated data symbols are passed through an IFFT block to generate the time-domain samples as

$$s_n = \sum_{l=0}^{N-1} S_l e^{\frac{j2\pi ln}{N}}, \quad n = 0, \ldots, N-1. \tag{1}$$

Next, a cyclic prefix (CP) is added to the time domain signal. The OFDM signal is submitted to a NL function $f(\cdot)$ before transmission, since we are considering a NL transmitter. Two classes of nonlinearities are considered [30]:

- Cartesian nonlinearities where $x_n = f(\mathcal{R}\{s_n\})) + jf(\mathcal{I}\{s_n\}))$, where $\mathcal{R}\{s_n\}$ and $\mathcal{I}\{s_n\}$ denote the real and imaginary part of s_n, respectively.
- Polar nonlinearities where $x_n = f(s_n) = A(|s_n|) \times \exp(j(\Theta(|s_n|) + \arg(s_n)))$, and $A(|s_n|)$ and $\Theta(|s_n|)$ denote the so-called AM/AM and AM/PM NL conversions. These nonlinearities are also denoted as bandpass memoryless nonlinearites.

These models include many common nonlinearities like clipping, quantization and NL power amplifiers, among others. The studies in this paper are done on sampled signals. However, we should note that the discrete-time signals can be obtained from underlying continuous-time signals and, therefore, the studies are valid for nonlinearities operating on continuous signals when the oversampling order is large enough.

After passing through the channel, the received signal in the frequency domain is

$$Y_k = H_k X_k + W_k \tag{2}$$

where H_k, X_k, and W_k are the channel frequency response, the transmitted signal affected with hardware impairments, and the noise component associated with the k-th subcarrier, respectively.

III. OVERVIEW OF GENERALIZED BUSSGANG DECOMPOSITION

Bussgang decomposition is commonly used for modeling system nonlinearity such as NL power amplifiers. In fact, the Bussgang decomposition is a consequence of the Bussgang Theorem [4] which states that the output of the NL function can be decomposed as

$$f(s) = \alpha s + d, \tag{3}$$

where α ($\alpha = \frac{E[xs^*]}{E[|s|^2]}$) is a scalar factor and equal to the correlation between the input and output signal and d is the distortion which is uncorrelated to s.

The Bussgang decomposition can be viewed as the minimum mean square error (MMSE) estimate of the output signal given the input and d is the estimation error. Therefore, although the results in the Bussgang Theorem only hold for Gaussian distributed input signals, the Bussgang decomposition is valid for input signals with arbitrary distribution [5].

Although we can use the Bussgang decomposition to obtain the total power of the useful and NL distortion terms (see, e.g., [13]), it does not provide the spectral characterization of the signals at the nonlinearity output, which means we do not know which part of the NL distortion falls in-band and out-of-band, and which are the NL distortion levels at each subcarrier. Moreover, it provides a characterization of the NL operation that is oversimplified and not enough for enhanced receivers like the BNC and the GAMP.

The GBD provides a more detailed characterization of the nonlinearity. GBD aims to characterize the NL operation with a finite number of terms. In GBD the nonlinearity output is decomposed as

$$f(s) = \sum_{m=0}^{\mu-1} \beta_m g_m(s) + \tilde{d} \tag{4}$$

where β_m and $g_m(s)$ for $m = 0, 1, .., \mu - 1$ are the scalar factors and the basis functions modeling different orders of nonlinearity, respectively, and μ is the number of basis functions and $\tilde{d}$ is the remaining distortion signal. It can be shown that the optimum vector of the scalar factors in GBD, $\beta = [\beta_0, \beta_1, \ldots, \beta_{\mu-1}]$, that minimizes the power of the remaining distortion signal is [31]

$$\beta_{opt} = \boldsymbol{C}_{f(s),\boldsymbol{g}(s)}\boldsymbol{C}^{-1}_{\boldsymbol{g}(s),\boldsymbol{g}(s)}, \tag{5}$$

where $\boldsymbol{C}_{f(s),\boldsymbol{g}(s)} = E[f(s)\boldsymbol{g}(s)^H]$, $\boldsymbol{C}_{\boldsymbol{g}(s),\boldsymbol{g}(s)} = E[\boldsymbol{g}(s)\boldsymbol{g}(s)^H]$ and $\boldsymbol{g}(s) = [g_0(s), g_1(s), \ldots, g_{\mu-1}(s)]^T$. We should note that β_{opt} in (5) can be seen as the linear MMSE estimator of $f(s)$ given $\boldsymbol{g}(s)$. By substituting β_{opt} in (4) and multiplying both sides with $\boldsymbol{g}(s)^H$, it follows that $\tilde{d}$ is uncorrelated with $\boldsymbol{g}(s)$, i.e., $E[\tilde{d}\boldsymbol{g}(s)^H] = \boldsymbol{0}$.

From the expression in (5), we can see that if orthogonal basis functions are used, then β_{opt} can be simplified. In other words if $g_m(s)$ and $g_i(s)$ for $m \neq i$ are orthogonal with respect to the distribution of s, then we have $\boldsymbol{C}_{\boldsymbol{g}(s),\boldsymbol{g}(s)} = I$, where I is the identity matrix. Consequently, we have

$$\beta_{opt} = \boldsymbol{C}_{f(s),\boldsymbol{g}(s)}. \tag{6}$$

One common option for orthogonal basis functions in GBD is orthogonal polynomials which can be derived from the *Gram-Schmidt* process for any arbitrary distribution [32]. Since the input signal of the NL function in the OFDM system is complex Gaussian and its amplitude has Rayleigh distribution, in the following we will introduce the Hermite and generalized Laguerre polynomials [33], which are orthogonal with respect to the Gaussian and exponential distributions, respectively, and, from Laguerre polynomials, we define polynomials orthogonal under the Rayleigh distribution.

A. HERMITE POLYNOMIALS

The Hermite polynomial of m-th order is defined as

$$\tilde{H}_m(s) = \frac{(-1)^m}{\sqrt{m!}} e^{\frac{s^2}{2}} \frac{d^m}{ds^m} e^{\frac{-s^2}{2}}, \tag{7}$$

where $\frac{d^m}{ds^m}$ is the m-th derivative with respect to s. The Hermite polynomials are orthogonal with respect to the function $e^{\frac{-s^2}{2}}$, i.e.,

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \tilde{H}_m(s)\tilde{H}_i(s) e^{\frac{-s^2}{2}} ds = \delta_{m-i}, \tag{8}$$

where δ_i is the discrete Dirac delta function. From (8) it can be interpreted that if s is real and has a Gaussian distribution with zero mean and variance 1, then $E[\tilde{H}_m(s)\tilde{H}_i(s)] = 0$ for $m \neq i$. With a simple change of the variable, this property holds for general Gaussian distributed s with arbitrary mean and variance. The Hermite polynomials can be used as orthogonal basis functions for OFDM signals submitted to Cartesian nonlinearities, since the real and imaginary parts of the OFDM signal are approximately Gaussian. Therefore we have

$$\begin{aligned} g_m(\mathcal{R}\{s_n\}) &= \tilde{H}_m(\mathcal{R}\{s_n\}), \\ g_m(\mathcal{I}\{s_n\}) &= \tilde{H}_m(\mathcal{I}\{s_n\}). \end{aligned} \tag{9}$$

B. GENERALIZED LAGUERRE POLYNOMIALS

The generalized Laguerre polynomial of m-th order is defined as

$$L_m^{(a)}(s) = \frac{s^{-a}}{\sqrt{(m+a)!m!}} e^s \frac{d^m}{ds^m}(e^{-s}s^{m+a}). \tag{10}$$

The generalized Laguerre polynomials are orthogonal with respect to function $s^a e^{-s}$, i.e.,

$$\int_0^{\infty} L_m^{(a)}(s)L_i^{(a)}(s)s^a e^{-s} ds = \delta_{m-i}, \tag{11}$$

Substituting the variable $s = \frac{z^2}{2\sigma^2}$ in (11) and setting $a = 1$, we have

$$\int_0^{\infty} \tilde{L}_{2m+1}(z)\tilde{L}_{2i+1}(z)\frac{z}{\sigma^2} e^{\frac{-z^2}{2\sigma^2}} dz = \delta_{m-i}, \tag{12}$$

where

$$\tilde{L}_{2m+1}(z) = \frac{z}{\sqrt{2\sigma^2}} L_m^{(1)}(\frac{z^2}{2\sigma^2}). \tag{13}$$

From (12), it can be interpreted that if z is real and has a Rayleigh distribution with parameter σ, then $E[\tilde{L}_{2m+1}(z)\tilde{L}_{2i+1}(z)] = 0$ for $m \neq i$. In the remaining of the paper we denote $\tilde{L}_m(s)$ as Laguerre-based polynomial and it is defined as

$$\tilde{L}_m(s) = \begin{cases} \frac{s}{\sqrt{2\sigma^2}} L^{(1)}_{\frac{m-1}{2}}(\frac{s^2}{2\sigma^2}) & \text{if } m \quad \text{is odd} \\ 0 & \text{if } m \quad \text{is even} \end{cases}. \tag{14}$$

Since the input samples to the NL function in OFDM are complex Gaussian, their amplitude has a Rayleigh distribution, and the generalized Laguerre-based polynomials defined in (14) can be used as orthogonal basis in the GBD such that

$$g_m(s_n) = e^{j\arg(s_n)}\tilde{L}_m(|s_n|). \tag{15}$$

We should note that the orthogonal polynomials defined in (7) and (14) are also orthonormal.

IV. JOINT CHANNEL AND NONLINEARITY ESTIMATION

In a NL system, to obtain the original data at the receiver side, we need to estimate both the channel and the system nonlinearity. The estimated channel will be used at the equalizer to compensate the channel effect and the estimated NL function will be used in NL receivers such as the BNC or GAMP receiver to compensate the NL effect on the signal.

In conventional systems, the channel is estimated by transmission of pilots and/or training sequences. However, when we have nonlinearity in the system the pilots/training sequences are affected by the nonlinearity and the channel estimation will no longer be accurate and reliable. On the

other hand, to estimate the NL function we need to remove the channel effects first. Therefore, in this section we aim to jointly estimate the channel and the NL function. For this we design two TSs (one for the channel estimation and the other for the nonlinearity estimation) in a manner that effectively addresses the problem while also maintaining low system complexity and ensuring spectral efficiency.

A. DESIGNING TS FOR CHANNEL ESTIMATION IN A NL SYSTEM

To reduce the effect of nonlinearity on channel estimation we can design a TS for the channel estimation phase in a way that it works in the linear region of the system. Designing a constant envelope signal will be a suitable choice. For this, we consider an OFDM signal that passes through a hard limiter (HL) (in both time and frequency domains) as the TS. In other words, the HL is applied both in the time and frequency domains to reduce the PAPR of the signal and obtain a near-constant envelope for the TS to avoid effects of nonlinearity on the TS and consequently have a reliable channel estimation. If we consider $\boldsymbol{s^{(0)}} = [s_0^{(0)}, \ldots, s_{N-1}^{(0)}]$ as the time domain OFDM block for the training phase, then the HL is applied as follows [34]:

- The time domain samples are passed through the HL as
$$\bar{s}_n^{(0)} = \exp(j \arg(s_n^{(0)})). \tag{16}$$
- The FFT of the resulting signal is calculated, leading to the frequency block $\boldsymbol{S_k^{(0)}} = [\bar{S}_0^{(0)}, \ldots, \bar{S}_{N-1}^{(0)}]$
- The resulting frequency domain samples, are passed through the HL as
$$S_k^{(1)} = \exp(j \arg(\bar{S}_k^{(0)})). \tag{17}$$
- The IFFT of the resulting signal is calculated to obtain the new time domain TS block, $\boldsymbol{s^{(1)}} = [s_0^{(1)}, \ldots, s_{N-1}^{(1)}]$.
- This procedure is repeated several times.

We should note that the disadvantage of this technique with respect to pilots is that we are using a complete block for channel estimation. However, the block length in the training phase can be reduced depending on the number of multipath components of the channel to reduce training overhead and maintain spectral efficiency. In the rest of the paper we call this TS as HL-TS.

B. NL PARAMETER ESTIMATION

In this section, we assume to have no knowledge of the system nonlinearity. Therefore, we consider the GBD introduced in Section III to model the system nonlinearity and our aim is to estimate the scalar parameters in this model.

1) ESTIMATING NL PARAMETERS AFTER EQUALIZATION

Lets define the time domain OFDM TS as $\boldsymbol{s} = [s_0, \ldots, s_{N-1}]$. Then, by taking into account the system characterization in Section II, we can write the received, frequency domain, OFDM signal in (2) in matrix form as

$$\boldsymbol{Y} = \Sigma \boldsymbol{F} \boldsymbol{x}(\boldsymbol{s}) + \boldsymbol{W}, \tag{18}$$

where $\boldsymbol{Y} = [Y_0, \ldots, Y_{N-1}]^T$, $\Sigma = diag[H_0, \ldots, H_{N-1}]$, $\boldsymbol{x}(\boldsymbol{s}) = [x_0, \ldots, x_{N-1}]^T$, $\boldsymbol{W} = [W_0, \ldots, W_{N-1}]^T$, and $\boldsymbol{F}$ is the $N \times N$ Fourier transform matrix. Considering the GBD in Section III, $\boldsymbol{x}(\boldsymbol{s})$ can be defined as

$$\boldsymbol{x}(\boldsymbol{s}) = \boldsymbol{G}(\tilde{s})^T \beta + \tilde{\boldsymbol{d}}, \tag{19}$$

where $\boldsymbol{G}(\tilde{s}) = \boldsymbol{G}(\mathcal{R}\{\boldsymbol{s}\}) + j\boldsymbol{G}(\mathcal{I}\{\boldsymbol{s}\})$ and $\boldsymbol{G}(\tilde{s}) = \boldsymbol{G}(\boldsymbol{s})$ for the Cartesian and polar case, respectively, where $\boldsymbol{G}(\boldsymbol{s}) = [\boldsymbol{g}(s_0), \ldots, \boldsymbol{g}(s_{N-1})]$, $\boldsymbol{g}(s) = [g_0(s), \ldots, g_{\mu-1}(s)]^T$, $\beta = [\beta_0, \ldots, \beta_{\mu-1}]^T$ and $\tilde{\boldsymbol{d}} = [\tilde{d}_0, \ldots, \tilde{d}_{N-1}]^T$.

Therefore, we have

$$\boldsymbol{Y} = \Sigma \boldsymbol{F} \tilde{\boldsymbol{G}}(\boldsymbol{s})^T \beta + \tilde{\boldsymbol{W}}, \tag{20}$$

where $\tilde{\boldsymbol{W}} = \boldsymbol{W} + \tilde{\boldsymbol{d}}$. Having the CSI knowledge, the received training block for NL parameter estimation is equalized as

$$\tilde{Y}_k = \frac{Y_k}{H_k}. \tag{21}$$

Then, the NL parameters are estimated using the equalized signal as

$$\hat{\beta} = \left(\mathbf{G}(\tilde{\mathbf{s}})^* \mathbf{G}(\tilde{\mathbf{s}})^T\right)^{-1} \mathbf{G}(\tilde{\mathbf{s}})^* \mathbf{F}^H \tilde{\mathbf{Y}}, \tag{22}$$

where $\tilde{\boldsymbol{Y}} = [\tilde{Y}_0, \ldots, \tilde{Y}_{N-1}]^T$.

2) DESIGNING TS FOR NL PARAMETER ESTIMATION

Unlike the channel estimation case, for the NL parameter estimation, we need to design the TS in a way that it has enough samples in different working regions of the NL system. Therefore, we attempt to design a TS that has the same amount of samples in different working regions of the NL system. We call this TS as uniformly distributed TS (UD-TS).

If we consider that the training block has length N, then the UD-TS in time domain is designed as follows:

- Divide the training block into K equivalent sub-blocks. This is to design the signal for one sub-block and then repeat it in other sub-blocks to overcome possible bad channel conditions on some samples in a sub-block.
- Since we have $L = N/K$ samples in each sub-block, we divide the period $[0, 2A_m]$, where A_m is the saturation level of the NL function,[1] into L equally spaced levels. These levels will be the amplitude of the samples in a sub-block.
- To produce the samples in a sub-block, L random QPSK samples with the different amplitudes obtained in the previous step are produced.

However, we should note that since β_1 is the scalar weight of the input signal itself (the linear term) in the GBD model, it depends highly on the distribution of the input signal of the nonlinearity. And, since we want to estimate the NL parameters for a system where the data blocks are OFDM

[1] Since A_m varies slowly and having a rough estimate of it in our case is sufficient, it can be estimated once at the beginning of transmission with techniques such as back-off testing or input-output measurement.

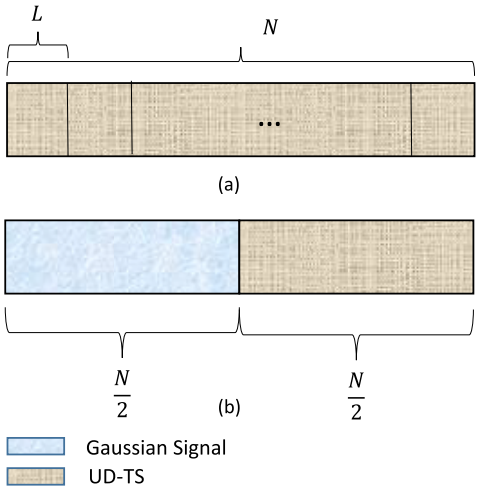

FIGURE 2. Designed TSs for NL parameter estimation: (a) UD-TS, (b) C-TS.

blocks (i.e., the distribution of the data blocks is complex Gaussian), then we need the TS to be complex Gaussian to have a good estimate of β_1. On the other hand, we need a uniformly distributed TS in terms of the amplitude of the samples to have a good estimate of the higher-order terms. Therefore, we designed another TS for the NL parameter estimation by using a combination of the UD-TS and a Gaussian TS. In other words, half of the block is complex Gaussian and half is designed based on the UD-TS to have sufficient samples in different regions of NL function. Fig. 2 shows the mixed designed TS for the NL parameter estimation. In the rest of the paper we name this TS as combination TS (C-TS).

C. SYSTEM WITH JOINT CHANNEL AND NL PARAMETER ESTIMATION

In this subsection we assume that we have no knowledge of both the channel and the NL parameters. To estimate them we consider 2 blocks for the training phase. We denote the first block as $\boldsymbol{s}^{(TS_1)}$ and is used for channel estimation and the second one is denoted as $\boldsymbol{s}^{(TS_2)}$ and is considered for NL parameter estimation. The first block is designed based on the approach in Section IV-A to estimate the channel with no NL effect. The first received TS at subcarrier k can be written as

$$Y_k^{(1)} = \alpha H_k S_k^{(TS_1)} + H_k D_k^{(1)} + W_k^{(1)}, \tag{23}$$

where $\alpha = E[Y_k^{(1)} S_k^{(TS_1)*}]/E[|S_k^{(TS_1)}|^2]$ is the attenuation factor and $D_k^{(1)}$ is the NL distortion at subcarrier k. Since the first TS is designed to work in the linear region, $D_k^{(1)}$ can be considered almost zero. Now, to estimate the channel we can have two cases. In the first case, we assume to have knowledge of α and in the second case, we assume that we do not know the value of α. In the first case the channel coefficient at subcarrier k is estimated as

$$\hat{H}_k = \frac{Y_k^{(1)}}{\alpha S_k^{(1)}}. \tag{24}$$

Then the second received training block is equalized using the estimated channel as

$$\tilde{Y}_k^{(2)} = \frac{Y_k^{(2)}}{\hat{H}_k}, \tag{25}$$

where $Y_k^{(2)}$ is the second received TS at subcarrier k. Next, the NL parameters are estimated using the equalized signal. We should note that for estimating the NL parameters, either the whole second TS can be used to estimate β as below

$$\hat{\beta} = \left(\mathbf{G}\left(\tilde{\mathbf{s}^{(TS_2)}}\right)^* \mathbf{G}\left(\tilde{\mathbf{s}^{(TS_2)}}\right)^T\right)^{-1} \mathbf{G}\left(\tilde{\mathbf{s}^{(TS_2)}}\right)^* \mathbf{F}^H \tilde{\mathbf{Y}}^{(2)}, \tag{26}$$

or we can use the first Gaussian half of the TS to estimate β_1 and then the whole TS to estimate the other β_m parameters as follows

$$\hat{\beta}_1 = \frac{\mathbf{g}_1\left(\tilde{\mathbf{s}}^{(TS_2)}_{[0:\frac{N}{2}-1]}\right)^* (\mathbf{F}^H \tilde{\mathbf{Y}}^{(2)})_{[0:\frac{N}{2}-1]}}{\mathbf{g}_1\left(\tilde{\mathbf{s}}^{(TS_2)}_{[0:\frac{N}{2}-1]}\right)^* \mathbf{g}_1\left(\tilde{\mathbf{s}}^{(TS_2)}_{[0:\frac{N}{2}-1]}\right)^T},$$

$$\hat{\beta}_{\notin 1} = \left(\mathbf{G}_{\notin 1}\left(\tilde{\mathbf{s}}^{(TS_2)}\right)^* \mathbf{G}_{\notin 1}\left(\tilde{\mathbf{s}}^{(TS_2)}\right)^T\right)^{-1} \mathbf{G}_{\notin 1}\left(\tilde{\mathbf{s}}^{(TS_2)}\right)^* \times \left(\mathbf{F}^H \tilde{\mathbf{Y}}^{(2)} - \beta_1 \tilde{\mathbf{g}}_1\left(\mathbf{s}^{(TS_2)}\right)\right). \tag{27}$$

In (27) $\tilde{\boldsymbol{g}_1}$ is the first row of matrix $\tilde{\boldsymbol{G}}$ and $\boldsymbol{s}^{(TS_2)}_{[0:\frac{N}{2}-1]}$ indicates the first $N/2$ samples of TS $\boldsymbol{s}^{(TS_2)}$. Also, $\hat{\beta}_{\notin 1}$ and $\tilde{\boldsymbol{G}_{\notin 1}}$ are $\hat{\beta}$ and $\tilde{\boldsymbol{G}}$ without their first row, respectively. Finally, the estimated NL parameters can be used in the BNC receiver to obtain the transmitted data. Fig. 3 shows the block diagram of the receiver with channel and NL parameter estimation.

In the second case where we have no knowledge of α, we can estimate a scaled version of the channel. In other words we estimate αH_k instead of H_k as follows

$$\hat{H}_k^{(\alpha)} = \frac{Y_k^{(1)}}{S_k^{(1)}}, \tag{28}$$

where $\hat{H}_k^{(\alpha)}$ is the estimation of H_k scaled by α. Therefore, when we equalize the second TS we have

$$\tilde{Y}_k^{(2)} = \frac{Y_k^{(2)}}{\hat{H}_k^{(\alpha)}}, \tag{29}$$

and since $\tilde{Y}_k^{(2)}$ in this case is scaled by $1/\alpha$, the estimated $\hat{\beta}$ in (26) and (27) will also be the estimation of β scaled by $1/\alpha$. However, we should note that this scaling does not affect the performance of the BNC receiver since both the channel estimation and the NL parameters are scaled and the effect of α vanishes when using them at the receiver.

Although we are using two TSs for the channel and NL function estimation in this technique, we should note that the changes of the NL function are usually very slow. As a result, the second training signal for estimating the nonlinearity function is used infrequently, ensuring that spectral efficiency remains intact. The period required for updating the NL

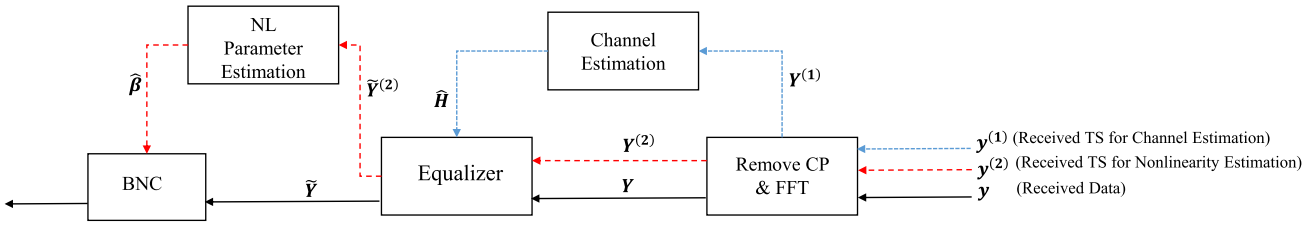

FIGURE 3. Block diagram of the receiver with joint channel and NL parameter estimation.

parameters depends on the system. Therefore, in our technique, whenever the period for estimating the NL parameters reaches, we use the designed HL-TS and C-TS to jointly estimate the channel and the NL parameters. And in other times we use the previous estimated NL parameter in the BNC receiver and only estimate and update the channel with the designed HL-TS.

V. ANALYSIS OF RESIDUAL DISTORTION

As stated in Sec. III, the idea of GBD is to characterize the NL operation with a finite number of terms, or, which is equivalent, to characterize the distortion term in the conventional Bussgang decomposition. This characterization can then be employed to define suitable BNC or GAMP receivers. In this section, we aim to analyze how well these estimated (finite number) terms model the distortion term in the conventional Bussgang decomposition. In general, when using polynomials as basis functions (orthogonal or not), a NL function can be expanded as

$$f(s) = \sum_{m=0}^{\infty} \beta_m g_m(s). \tag{30}$$

Considering (5) and the definition of orthonormal polynomials, it can be shown that α in the conventional Bussgang decomposition is a scaled version of β_1 in (30) such that we have $\alpha s = \beta_1 g_1(s)$. Therefore, the distortion signal in the conventional Bussgang decomposition in (3) is

$$d = \sum_{m=0,m\neq 1}^{\infty} \beta_m g_m(s). \tag{31}$$

We define $\Delta d(s)$ as the difference between d and the terms in the GBD that try to characterize d. Thus, we have

$$\begin{aligned}\Delta d(s) &= d - \sum_{m=0,m\neq 1}^{\mu-1} \hat{\beta}_m g_m(s) \\ &= \sum_{m=0,m\neq 1}^{\mu-1} (\beta_m - \hat{\beta}_m) g_m(s) + \sum_{m=\mu}^{\infty} \beta_m g_m(s).\end{aligned} \tag{32}$$

In the following we derive the normalized average power of $\Delta d(s)$ for OFDM system with Cartesian and polar nonlinearity where we use Hermite polynomials and Laguerre-based polynomials as basis functions, respectively. Without loss of generality, and for the sake of simplicity. we only derive the normalized average power of $\Delta d(s)$ for the real (or imaginary) part of the signal in the Cartesian nonlinearity case. To have a unified notation for both the Cartesian and the polar case, we define

$$\tilde{s}_n = \begin{cases} \mathcal{R}\{s_n\} \text{ or } \mathcal{I}\{s_n\} & \text{for Cartesian} \\ s_n & \text{for polar} \end{cases} \tag{33}$$

Considering (32), the average normalized power of $\Delta d(\tilde{s})$ for an OFDM block is defined as

$$P_{\Delta d} = \frac{E\left[\frac{1}{N}\sum_{n=0}^{N-1}\left|\sum_{\substack{m=0,\\ m\neq 1}}^{\mu-1}(\beta_m-\hat{\beta}_m)g_m(\tilde{s}_n)+\sum_{m=\mu}^{\infty}\beta_m g_m(\tilde{s}_n)\right|^2\right]}{P_d}, \tag{34}$$

where

$$P_d = E\left|\sum_{m=0,m\neq 1}^{\infty} \beta_m g_m(\tilde{s})\right|^2 \tag{35}$$

is the average power of the distortion signal in the conventional Bussgang decomposition. By taking advantage of the orthogonality of the basis functions, (34) can be simplified as

$$\begin{aligned} P_{\Delta d} &= \frac{\sum_{m=0,m\neq 1}^{\mu-1} E|\beta_m - \hat{\beta}_m|^2 E|g_m(\tilde{s})|^2 + \sum_{m=\mu}^{\infty} |\beta_m|^2 E|g_m(\tilde{s})|^2}{\sum_{m=0,m\neq 1}^{\infty} |\beta_m|^2 E|g_m(\tilde{s})|^2} \\ &\overset{(a)}{=} \frac{\sum_{m=0,m\neq 1}^{\mu-1} E|\beta_m - \hat{\beta}_m|^2 + \sum_{m=\mu}^{\infty} |\beta_m|^2}{\sum_{m=0,m\neq 1}^{\infty} |\beta_m|^2}, \end{aligned} \tag{36}$$

where (a) results form the use of normalized orthogonal polynomials.

It can be concluded from (36) that each term in the GBD with the estimated parameters can contribute in reducing the power of Δd only if $E\left[|\beta_m - \hat{\beta}_m|^2\right] < |\beta_m|^2$. In other words if the variance of the estimation error of β_m is larger than the power of β_m, the associated term in the GBD cannot help reduce the distortion power and therefore, removing that term in the GBD will give a more accurate characterization of the NL operation.

VI. PERFORMANCE RESULTS

In this section we will first study the suitability of the GBD with different number of terms in modeling the system nonlinearity through simulations. Then we will analyze the accuracy of the proposed joint channel and NL parameter estimation technique and see how well is the system performance when these estimated parameters are used at the BNC receiver. We will also investigate how well can the higher order terms of the GBD, estimated with our proposed technique, model the distortion term in the conventional Bussgang decomposition.

Unless otherwise stated, for the simulations, we have assumed an LDPC code with rate $r_c = 0.5$, 256-QAM modulation and $N = 512$ subcarriers. Since in the following simulations we are considering GDB model with polynomial order of up to 9, we have considered an oversampling factor of 16 to have an accurate analysis of the behaviour of different terms of the GBD model without too much aliasing. However, in practice a much smaller oversampling can be employed by the receiver without changing the BER performance. We would like to point out that this oversampling is only for simulation purposes and the oversampling factor for generating the time-domain OFDM samples through the IFFT operation is naturally much smaller. For the channel, we have assumed an exponential decaying multipath Rayleigh fading channel with 15 paths and normalized delay spread of 3 and for the equalization we utilize a ZF equalizer. We test both the Cartesian and the polar nonlinearity cases in an OFDM system. For the Cartesian case we have considered Hermite polynomials as basis functions and the nonlinearity considered at the transmitter is the clipping function defined as

$$f(\tilde{s}_n) = \begin{cases} -A_{\text{clip}} & \tilde{s}_n \le -A_{\text{clip}} \\ \tilde{s}_n & -A_{\text{clip}} < \tilde{s}_n < A_{\text{clip}} \\ A_{\text{clip}} & \tilde{s}_n \ge A_{\text{clip}} \end{cases} \tag{37}$$

where, A_{clip} is the clipping level. In the polar case, we used Laguerre-based polynomials defined in (14) for the basis functions and considered solid state power amplifier (SSPA) and/or traveling wave tube amplifier (TWTA) as the system nonlinearity. The relation between the input and output of the SSPA model is

$$f(\tilde{s}_n) = \frac{\tilde{s}_n}{\sqrt[2p]{1 + \left(\frac{|\tilde{s}_n|}{A_{\text{sat}}}\right)^{2p}}}, \tag{38}$$

where, A_{sat} and p are the saturation level and the smoothing factor, respectively. The AM/AM and AM/PM conversions for the TWTA model are as

$$A(|\tilde{s}_n|) = \frac{|\tilde{s}_n|}{1 + \left(\frac{|\tilde{s}_n|}{A_{\text{sat}}}\right)^2},$$
$$\Theta(|\tilde{s}_n|) = \frac{\Theta_{\text{sat}}|\tilde{s}_n|^2}{|\tilde{s}_n|^2 + A_{\text{sat}}{}^2}, \tag{39}$$

where A_{sat} and Θ_{sat} are the saturation level and the saturation phase shift, respectivley. We should note that, due to the nature of these nonlinearities, we only consider the odd power terms of the polynomials in the GBD [13], i.e.,

$$f(\tilde{s}_n) = \sum_{\gamma=0}^{\frac{\mu}{2}-1} \beta_{2\gamma+1} g_{2\gamma+1}(\tilde{s}_n) + \tilde{d} \tag{40}$$

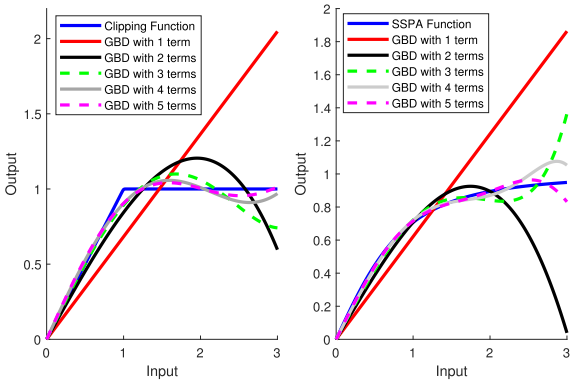

FIGURE 4. Output versus input for clipping and SSPA function and their associated GBD with different number of terms.

A. GBD SIMULATION ANALYSIS FOR CLIPPING AND SSPA

Fig. 4 shows the relation between the input and output signal for the clipping and the SSPA functions and their associated GBD with different number of terms. In this figure we have assumed $A_{\text{clip}} = 1$, $A_{\text{sat}} = 1$ and $p = 1$. It can be seen that as the number of terms in the GBD increase, we can better model the clipping and the SSPA function. We can also see that the SSPA nonlinearity can be modeled with the GBD with fewer number of terms than the clipping function (the GBD with 5 terms can model the SSPA function very well), which is related to the non-differential nature of the clipping at the clipping level.

Fig. 5 presents the power of the residual distortion ($P_{\tilde{d}}$) versus $A_{\text{clip}}/\sigma_{\tilde{s}}$ for the GBD with different number of terms. $\sigma_{\tilde{s}}^2$ is the power of the input signal of the NL function. It can be seen that as $A_{\text{clip}}/\sigma_{\tilde{s}}$ increases, the effect of the nonlinearity on the input signal and consequently $P_{\tilde{d}}$ decreases. It can be also seen that in general as the number of GBD terms increase, the power of the residual distortion signal decreases. However, some terms are dominant and have more contribution in decreasing $P_{\tilde{d}}$. For example, in $A_{\text{clip}}/\sigma_{\tilde{s}} = 1.5$ the second GBD term has higher effect in decreasing the residual distortion power since by adding the third, fourth and the fifth term, $P_{\tilde{d}}$ does not decrease much compared to the black curve. However, for $A_{\text{clip}}/\sigma_{\tilde{s}} = 3$ the third term of the GBD has higher effect in decreasing $P_{\tilde{d}}$.

Fig. 6 shows the power of the residual distortion versus $A_{\text{sat}}/\sigma_{\tilde{s}}$ for the GBD with different number of terms in an OFDM system. It can be seen that as $A_{\text{sat}}/\sigma_{\tilde{s}}$ increases and as the number of GBD terms increase, the power of the residual distortion signal decreases. As seen, the reduction rate of $P_{\tilde{d}}$ decreases as the number of GBD terms increase. For example, by increasing the number of terms in the GBD from 1 to 2

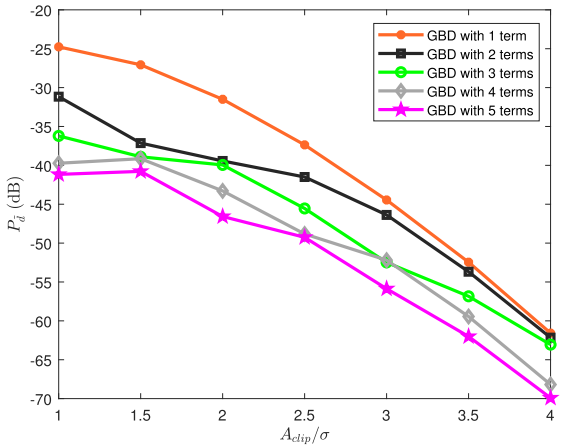

FIGURE 5. Power of residual distortion ($P_{\tilde{d}}$) versus $A_{\text{clip}}/\sigma_{\tilde{s}}$ for GBD with different number of terms in OFDM system with clipping nonlinearity.

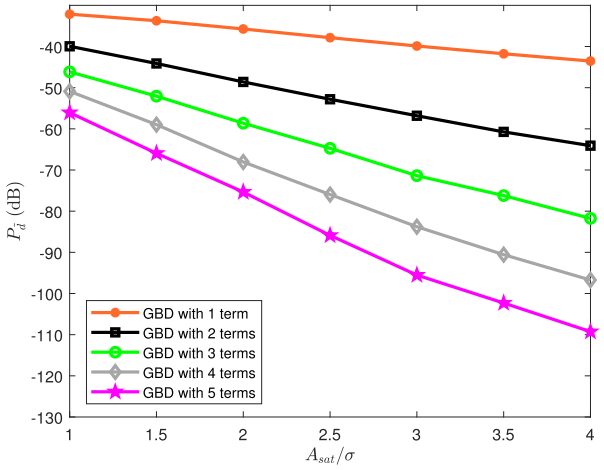

FIGURE 6. Power of residual distortion ($P_{\tilde{d}}$) versus $A_{\text{sat}}/\sigma_{\tilde{s}}$ for GBD with different number of terms in OFDM system with SSPA nonlinearity.

for $A_{\text{sat}}/\sigma_{\tilde{s}} = 1$, $P_{\tilde{d}}$ decreases with about 9 dB, while by increasing the number of terms from 4 to 5 it decreases approximately 5 dB. We should note that the same general conclusions, as in the SSPA case, can be drawn for the TWTA case and therefore we have not included the results for the TWTA case in this subsection.

B. JOINT CHANNEL AND NONLINEAIRTY ESTIMATION TECHNIQUE PERFORMANCE ANALYSIS

In the following, we study the accuracy of the proposed joint channel and GBD parameter estimation technique. Since the same conclusions can be drawn for the Cartesian and polar nonlinearity cases, we only study the polar nonlinearity in the following figures. For the results of this subsection we assumed a GBD with 5 terms and Laguerre-based polynomials to characterize the system nonlinearity.

Fig. 7 indicates the Inphase-Quadrature diagram of an OFDM block and the designed HL-TS. This figure supports the fact that the designed HL-TS has lower envelope

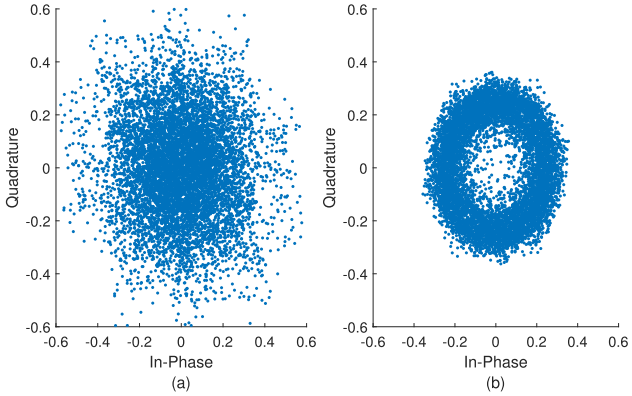

FIGURE 7. Inphase-Quadrature diagram of (a) an OFDM signal and (b) the designed HL-TS.

fluctuation and therefore the training signal for channel estimation is less effected by the system nonlinearity. Also, due to the lower dynamic range of the HL-TS, we can use this signal with higher power for a given amplifying characteristic without entering the nonlinear region of the amplifier.

In the following we will first study the SSPA model and later the TWTA model. For the SSPA model we have considered $A_{\text{sat}}/\sigma_{\tilde{s}} = 3$ and $p = 1$ in (38).

Fig. 8 presents the normalized MSE (NMSE) versus SNR for estimation of the effective channel ($\hat{H}_k^{(\alpha)}$) with both DMRS and the designed HL-TS in Sec. IV-A. NMSE of $\hat{H}_k^{(\alpha)}$ is defined as

$$NMSE_{\hat{H}_k^{(\alpha)}} = \frac{E|\hat{H}_k^{(\alpha)} - \alpha H_k|^2}{E|\alpha H_k|^2}. \tag{41}$$

DMRS are the pilots used in 5G to estimate the channel. The generation and allocation of these reference signals is defined in [29]. For this figure we considered the first configuration of DMRS in the standard. In other words, every other subcarrier is allocated to the DMRS in frequency domian (i.e. half of the OFDM block is allocated to the pilots). It can be seen from Fig. 8 that, when using DMRS reference signals in a NL system, the accuracy of the channel estimation degrades significantly with respect to the linear system, specially for high SNR. We can also see that our designed HL-TS with lenght $N = 512$ outperforms the DMRS reference signal case even in a linear system. To have a fair comparison and same training overhead as the DMRS case, we have also analyzed the performance of our designed HL-TS with length 256. It is seen that even with the same training overhead the HL-TS out performs the DMRS case in a NL system and has a performance close to the linear system case.

Fig. 9 shows the MSE of estimation of different terms of the considered GBD in two cases. The MSE for $\hat{\beta}_m$ is defined as

$$MSE_{\hat{\beta}_m} = E|\hat{\beta}_m - \beta_m|^2. \tag{42}$$

In the first case we have used an OFDM block for estimating the NL parameters and in the second case we have used the designed UD-TS for estimating the NL parameters. As it can

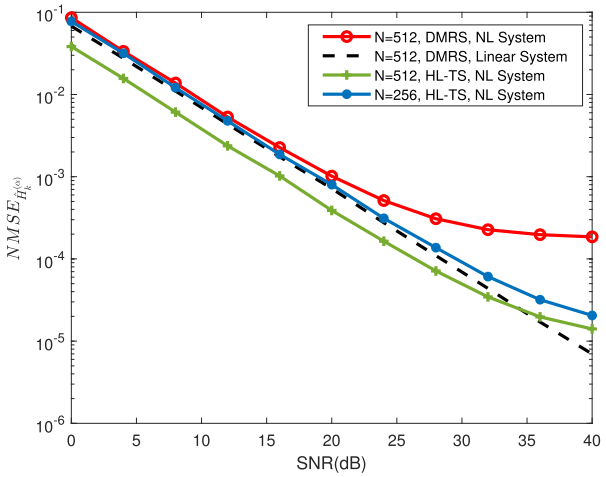

FIGURE 8. NMSE versus SNR for channel estimation with DMRS reference signals and the designed HL-TS.

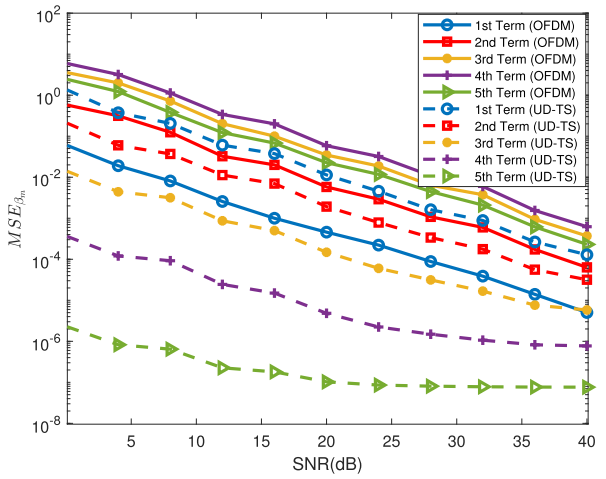

FIGURE 9. MSE versus SNR for NL parameter estimation with perfect CSI for OFDM block TS and UD-TS in SSPA case.

be seen in the figure, the estimation of β_1 is better with the OFDM block than the UD-TS, since β_1 is sensitive to the distribution of the input signal of the NL function and the OFDM block has a Gaussian distribution. However, it can be concluded that the estimation of the parameter related to the higher order terms is more accurate with the UD-TS. This is because we need to have enough samples in different working regions of the NL function.

Fig. 10 shows the MSE of estimation of different terms of the GBD with the designed C-TS in section IV-B2. By comparing Fig. 9 and Fig. 10, we can see that with the designed C-TS, the estimation of β_1 is as good as the OFDM TS case and the estimation of the rest of the β_ms is slightly better than the UD-TS case. This is because with the C-TS the estimation of β_1 is better than the UD-TS case and therefore we can make a better estimate of the rest of the terms. We should note that in Fig. 9 and Fig. 10 we have considered perfect knowledge of CSI to solely study the capability of different TSs in estimating the NL parameters.

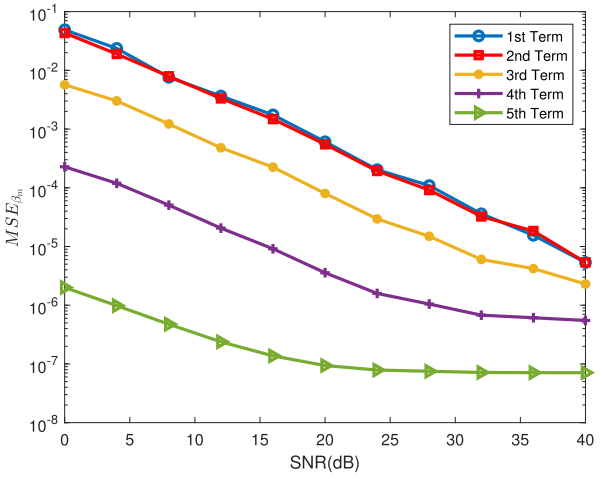

FIGURE 10. MSE versus SNR for NL parameter estimation with perfect CSI for designed C-TS in SSPA case.

Fig. 11 studies the BER performance of the BNC receiver for different TSs. We have considered a BNC receiver with 2 iterations where in each iteration the NL distortion is re-produced based on the estimated NL parameters and approximation of the transmitted signal. To improve the BNC receiver we have taken advantage of the decoder to better approximate the transmitted signal at each Bussgang iteration. This requires re-encoding and modulating the decoded bits at the end of each Bussgang iteration to re-produce the NL distortion and remove it from the received signal. As it is seen in this figure, the performance of the proposed joint channel and NL parameter estimation with the designed TS in Sec. IV-A and C-TS, significantly outperforms the case where the NL parameters are estimated with an OFDM block or UD-TS, even with the assumption of perfect CSI. The OFDM block and UD-TS cases do not reach BER of 10^{-4}, while the proposed joint channel and NL parameter estimation with the designed TS in Sec. IV-A and C-TS reaches BER of 10^{-4} at 5dB more gain with respect to the perfect estimation case. This figure also confirms that not having knowledge of α does not effect the performance of the receiver with the proposed joint channel and NL parameter estimation technique.

For the TWTA case we have considered $A_{\text{sat}}/\sigma_{\tilde{s}} = 3$ and $\Theta_{\text{sat}} = \pi/3$ in (39). In the TWTA case, unlike the SSPA case, the scalar weights of the GBD terms are complex numbers. This is due to having AM/PM conversion as well as AM/AM conversion in TWTA model.

Fig. 12 shows the MSE of estimation of scalar parameters of different terms of the GBD with the designed C-TS for the TWTA case. As it is seen the scalar weights of different terms are estimated with good accuracy, specially at high SNRs. We should note that since for the TWTA case similar conclusions can be drawn as the SSPA case (i.e. C-TS outperforms OFDM TS and UD-TS), we have only presented the MSE performance for the C-TS.

Fig. 13 indicates the BER performance of the BNC receiver for different TSs in the TWTA case. As it is seen the proposed

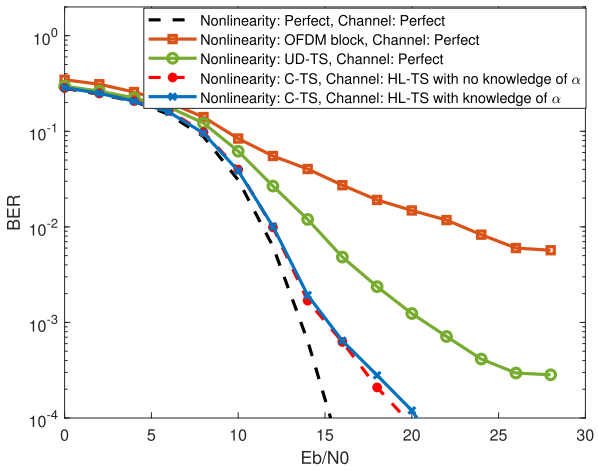

FIGURE 11. BER versus E_b/N_0 for BNC receiver with the proposed joint channel and NL parameter estimation technique in SSPA case.

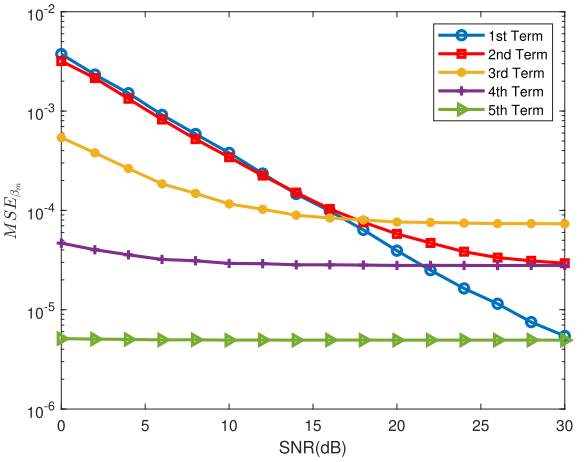

FIGURE 12. MSE versus SNR for NL parameter estimation with perfect CSI for designed C-TS in TWTA case.

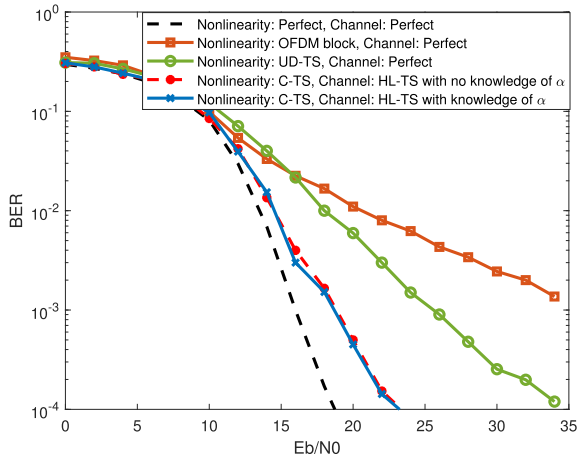

FIGURE 13. BER versus E_b/N_0 for BNC receiver with the proposed joint channel and NL parameter estimation technique in TWTA case.

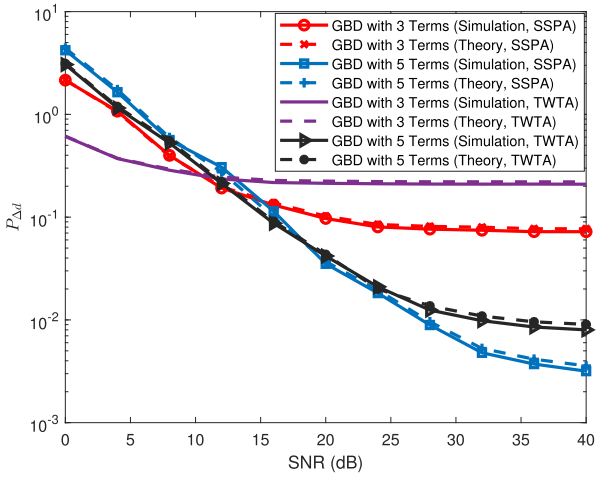

FIGURE 14. Normalized average power of residual distortion versus SNR for different number of terms in the GBD.

joint channel and NL parameter estimation technique with the designed HL-TS and C-TS has the best performance compared to the case where the NL parameters are estimated with an OFDM block or UD-TS, even with the assumption of perfect CSI. The proposed technique with the designed HL-TS and C-TS reaches BER of 10^{-4} at 4dB more gain with respect to the perfect estimation case, while the UD-TS with perfect channel estimation reaches BER of 10^{-4}dB at about 16 dB more gain with respect to the perfect estimation case. The figure also shows that not having knowledge of α in the proposed joint channel and NL parameter estimation technique does not effect the performance of the receiver.

C. RESIDUAL DISTORTION SIMULATION ANALYSIS

In the following figures we aim to analyze how well can the estimated higher order terms of the GBD, model the distortion term in the conventional Bussgang decomposition.

Fig. 14 shows $P_{\Delta d}$ versus SNR for different number of terms in the GBD through both simulation and theory analysis. This figure indicates the consistency of the closed form expression derived in (36) with the simulation results. Clearly, as SNR increases and, consequently, we have better estimation of $\beta_{2\gamma+1}$, $P_{\Delta d}$ decreases and therefore we have a better estimate of the distortion signal. This figure also shows that only at higher SNRs, increasing the number of terms in the GBD from 3 to 5, decreases $P_{\Delta d}$. This is because at low SNRs the variance of the estimation error of β_9 is larger than the power of β_9.

VII. CONCLUSION

In this paper we considered transmitter side nonlinearity and employed the GBD with orthogonal polynomials as basis functions to model the system nonlinearity. We proposed a joint channel and NL parameter estimation technique based on designing two complementary TSs. We also derived a

closed form expression for the average power of the residual distortion in the GBD with the estimated parameters. Our performance results indicate that the proposed estimation technique can estimate the channel coefficients accurately, close to the channel estimation performance in a linear system. The results also show that the scalar parameters in the GBD can be estimated accurately enough with the designed TS to have a good performance at the BNC receiver with the estimated parameters. The analysis shows that in general, as both the number of terms in the GBD and the SNR increase, the higher order terms of the GBD can better characterize the nonlinearity distortion compared to the conventional Bussgang decomposition. However, at low SNR, where the estimation of parameters is not accurate enough, increasing the number of terms in GBD does not improve the distortion characterization and can actually degrade it.

REFERENCES

[1] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin, and R. Zhang, "An overview of massive MIMO: Benefits and challenges," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 742–758, Oct. 2014.

[2] Z. Mokhtari, M. Sabbaghian, and R. Dinis, "A survey on massive MIMO systems in presence of channel and hardware impairments," *Sensors*, vol. 19, no. 1, p. 164, Jan. 2019.

[3] X. Zhang, X. Zhu, and Y. Tang, "Evaluating and reducing the envelope fluctuations of OFDM signals based on distortion prediction," *Frequenz*, vol. 71, nos. 1–2, pp. 73–81, Jan. 2017, doi: 10.1515/freq-2016-0009.

[4] J. J. Bussgang, "Crosscorrelation functions of amplitude-distorted Gaussian signals," Res. Lab. Electron., Massachusetts Inst. Technol., Cambridge, MA, USA, Tech. Rep., 216, 1952.

[5] O. T. Demir and E. Bjornson, "The Bussgang decomposition of nonlinear systems: Basic theory and MIMO extensions," *IEEE Signal Process. Mag.*, vol. 38, no. 1, pp. 131–136, Jan. 2021.

[6] Z. Mokhtari and R. Dinis, "Sum-rate of cell free massive MIMO systems with power amplifier non-linearity," *IEEE Access*, vol. 9, pp. 141927–141937, 2021.

[7] M. Bashar, H. Q. Ngo, K. Cumanan, A. G. Burr, P. Xiao, E. Björnson, and E. G. Larsson, "Uplink spectral and energy efficiency of cell-free massive MIMO with optimal uniform quantization," *IEEE Trans. Commun.*, vol. 69, no. 1, pp. 223–245, Jan. 2021.

[8] E. Björnson, L. Sanguinetti, and J. Hoydis, "Hardware distortion correlation has negligible impact on UL massive MIMO spectral efficiency," *IEEE Trans. Commun.*, vol. 67, no. 2, pp. 1085–1098, Feb. 2019.

[9] Z. Mokhtari, M. Sabbaghian, and R. Dinis, "Massive MIMO downlink based on single carrier frequency domain processing," *IEEE Trans. Commun.*, vol. 66, no. 3, pp. 1164–1175, Mar. 2018.

[10] M. Fozooni, M. Matthaiou, E. Bjornson, and T. Q. Duong, "Performance limits of MIMO systems with nonlinear power amplifiers," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2015, pp. 1–7.

[11] J. Guerreiro, R. Dinis, and P. Montezuma, "On the optimum multicarrier performance with memoryless nonlinearities," *IEEE Trans. Commun.*, vol. 63, no. 2, pp. 498–509, Feb. 2015.

[12] L. Xu, X. Lu, S. Jin, F. Gao, and Y. Zhu, "On the uplink achievable rate of massive MIMO system with low-resolution ADC and RF impairments," *IEEE Commun. Lett.*, vol. 23, no. 3, pp. 502–505, Mar. 2019.

[13] T. Araujo and R. Dinis, *Analytical Evaluation of Nonlinear Distortion Effects on Multicarrier Signals*. Boca Raton, FL, USA: CRC Press, 2015, doi: 10.1201/b18370.

[14] K. J. Keesman, *System Identification: An Introduction*. Berlin, Germany: Springer, 2011.

[15] D. Dardari, "Joint clip and quantization effects characterization in OFDM receivers," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 53, no. 8, pp. 1741–1748, Aug. 2006.

[16] D. Dardari, V. Tralli, and A. Vaccari, "A theoretical characterization of nonlinear distortion effects in OFDM systems," *IEEE Trans. Commun.*, vol. 48, no. 10, pp. 1755–1764, Oct. 2000, doi: 10.1109/26.871400.

[17] R. Raich and G. T. Zhou, "Orthogonal polynomials for complex Gaussian processes," *IEEE Trans. Signal Process.*, vol. 52, no. 10, pp. 2788–2797, Oct. 2004, doi: 10.1109/TSP.2004.834400.

[18] C. Mollén, U. Gustavsson, T. Eriksson, and E. G. Larsson, "Spatial characteristics of distortion radiated from antenna arrays with transceiver nonlinearities," *IEEE Trans. Wireless Commun.*, vol. 17, no. 10, pp. 6663–6679, Oct. 2018, doi: 10.1109/TWC.2018.2861872.

[19] J. Tellado, L. M. C. Hoo, and J. M. Cioffi, "Maximum-likelihood detection of nonlinearly distorted multicarrier symbols by iterative decoding," *IEEE Trans. Commun.*, vol. 51, no. 2, pp. 218–228, Feb. 2003.

[20] J. Madeira, J. Guerreiro, and R. Dinis, "Iterative frequency-domain detection and compensation of nonlinear distortion effects for MIMO systems," *Phys. Commun.*, vol. 37, Dec. 2019, Art. no. 100869.

[21] S. V. Zhidkov, "Detection of nonlinearly distorted OFDM signals via generalized approximate message passing," 2017, *arXiv:1703.01562*.

[22] S. V. Zhidkov and R. Dinis, "Belief propagation receivers for near-optimal detection of nonlinearly distorted OFDM signals," in *Proc. IEEE 89th Veh. Technol. Conf. (VTC-Spring)*, Apr. 2019, pp. 1–6.

[23] S. Rangan, P. Schniter, A. Fletcher, and S. Sarkar, "On the convergence of approximate message passing with arbitrary matrices," *IEEE Trans. Inf. Theory.*, vol. 65, no. 9, pp. 5339–5351, Sep. 2019.

[24] E. Olfat and M. Bengtsson, "Estimation of the clipping level in OFDM systems," in *Proc. 49th Asilomar Conf. Signals, Syst. Comput.*, Nov. 2015, pp. 1169–1173.

[25] J. Dohl and G. Fettweis, "Blind estimation of memoryless AM/PM nonlinearities in OFDM systems," in *Proc. IEEE 78th Veh. Technol. Conf. (VTC Fall)*, Sep. 2013, pp. 1–5.

[26] E. Olfat and M. Bengtsson, "Joint channel and clipping level estimation for OFDM in IoT-based networks," *IEEE Trans. Signal Process.*, vol. 65, no. 18, pp. 4902–4911, Sep. 2017.

[27] F. H. Gregorio, S. Werner, J. Cousseau, J. Figueroa, and R. Wichman, "Receiver-side nonlinearities mitigation using an extended iterative decision-based technique," *Signal Process.*, vol. 91, no. 8, pp. 2042–2056, Aug. 2011.

[28] E. Olfat and M. Bengtsson, "A general framework for joint estimation-detection of channel, nonlinearity parameters and symbols for OFDM in IoT-based 5G networks," *Signal Process.*, vol. 167, Feb. 2020, Art. no. 107298.

[29] *5G; NR; Physical Channels and Modulation*, document (TS) 38.211. Version 17.4.0, 3GPP, Jan. 2023.

[30] R. Pasricha, S. Sharma, and S. Pasricha, "Memoryless non linear modeling of power amplifier," *Int. J. Electron.*, vol. 2, no. 1, pp. 55–58, 2010.

[31] A. Papoulis and S. U. Pillai, *Probability, Random Variables, and Stochastic Processes*, 4th ed., New York, NY, USA: McGraw-Hill, 2002.

[32] L. E. Ghaoui, "Orthogonalization: The Gram-Schmidt procedure," in *Linear Algebra and Applications*. Hanoi, Vietnam: VinUniversity, 2022. [Online]. Available: https://ecampusontario.pressbooks.pub/linearalgebraandapplications/

[33] M. I. Falcão, I. Cação, and H. R. Malonek, "Monogenic generalized Laguerre and Hermite polynomials and related functions," in *Proc. 9th Int. Conf. Clifford Algebras Appl. Math. Phys.*, Jan. 2011, pp. 1–10.

[34] R. Dinis, R. Kalbasi, D. Falconer, and A. Banihashemi, "Channel estimation for MIMO systems employing single-carrier modulations with iterative frequency-domain equalization," in *Proc. IEEE 60th Veh. Technol. Conf. (VTC Fall)*, Sep. 2004, pp. 1–5.

[35] J. Guerreiro, "Analytical characterization and optimum detection of nonlinear multicarrier schemes," Ph.D. dissertation, Dept. Elect. Comput. Eng., Faculdade de Ciências e Tecnologia, Univ. Nova de Lisboa, Caparica, Portugal, 2016. [Online]. Available: http://hdl.handle.net/10362/19916

ZAHRA MOKHTARI received the B.Sc. degree from the University of Yazd, Yazd, Iran, in 2011, and the M.Sc. and Ph.D. degrees from the University of Tehran, Tehran, Iran, in 2013 and 2019, respectively, all in electrical engineering. In 2017, she was a Visiting Researcher with the Chalmers University of Technology, Gothenburg, Sweden. From 2018 to 2022, she was a Researcher with Yazd Science and Technology Park, Yazd. She has been a Researcher with the Instituto de Telecomunicações, Lisbon, Portugal, since 2022. Her research interests include wireless communications, 5G, cell free and cellular massive MIMO, and hardware impairments.

RUI DINIS (Senior Member, IEEE) received the Ph.D. degree from the Instituto Superior Técnico (IST), Technical University of Lisbon, Portugal, in 2001, and the Habilitation degree in telecommunications from the Faculdade de Ciências e Tecnologia (FCT), Universidade Nova de Lisboa (UNL), in 2010. From 2001 to 2008, he was an Assistant Professor with IST. He was a Researcher with the Centro de Analise e Processamento de Sinal (CAPS), IST, from 1992 to 2005, and a Researcher with the Instituto de Sistemas e Robotica (ISR), from 2005 to 2008. Since 2009, he has been a Senior Researcher with the Instituto de Telecomunicações (IT). In 2003, he was an Invited Researcher with Carleton University, Ottawa, Canada. He is currently a Full Professor with FCT, UNL. His research interests include transmission, estimation, and detection techniques for wireless communications. He is a VTS Distinguished Speaker and a ComSoc Distinguished Lecturer. He is or was an Editor of IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, IEEE TRANSACTIONS ON COMMUNICATIONS, IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, IEEE OPEN JOURNAL OF THE COMMUNICATIONS SOCIETY, and *Physical Communication* (Elsevier). He was also the guest editor of several special issues.

SHA HU received the B.S. and M.S. degrees in pure mathematics from Wuhan University, China, and the Ph.D. degree in electrical engineering from Lund University, Sweden. He has been working in academia and industry over 15 years in wireless communication and modem algorithm designs. He is currently an Algorithm Expert with Lund Research Center, Huawei Technologies Sweden AB, and leads a research group in exploring potentials and addressing challenges in 6G. He is one of the pioneers in large intelligent surface (LIS) and the recipients of the Best Journal Paper Award of SWE-CTW in 2019, the Best Paper Award of IEEE VTC-Fall in 2022, and the IEEE Fred W. Ellersick Prize Paper Award in 2023.

DZEVDAN KAPETANOVIC received the M.Sc. degree in computer science and the Ph.D. degree in electrical engineering from Lund University, Lund, Sweden, in 2007 and 2012, respectively. From 2012 to 2013, he was a Research Associate with the Security and Trust (SnT) Center, University of Luxembourg, Luxembourg. From 2014 to 2018, he was a Senior Researcher with Ericsson Research, Lund. Since June 2018, he has been with Huawei Technologies Sweden AB, Lund, where his current role is that of a Principal Engineer. His work is mainly within the fields of communication theory, signal processing, and artificial intelligence.