Identification of the Linear Systems of the Wiener Hammerstein RF Power Amplifier Model Using DFT Analysis

Soner YESIL [1], Ali Ozgur YILMAZ [2]

[1] TUALCOM Inc., Ankara, Türkiye
[2] Dept. of Electrical and Electronics Eng., Middle East Tecnical University, Ankara, Türkiye

soner.yesil@tualcom.com.tr, aoyilmaz@metu.edu.tr

Submitted December 12, 2023 / Accepted March 25, 2024 / Online first April 24, 2024

Abstract. *This paper presents a novel method for identification of the sub-system parameters of a Wiener-Hammerstein Nonlinear (WHNL) system that is used for modeling RF Power Amplifier characteristics. The proposed method first isolates the overall linear system from the memoryless nonlinearity by exploiting the Bussgang decomposition method. Then, Discrete Fourier Transform (DFT) analysis is used for the estimation of the inner linear system. Finally, the outer linear system parameters are updated based on the inner system estimation. The estimated systems are then used to model the target system for an In-Band-Full-Duplex (IBFD) scenario. Performance of Self-Interferene Cancellation (SIC) has been evaluated under the existence of Signal-of-Interest (SoI). Error Vector Magnitude (EVM) metric of the SoI is used to compare with a Half-Duplex (HD) receiver under various inner linear system parameters. SIC performance has been examined with respect to the changing power levels of the SoI and self-interference signal for various delay and gain values of a practical two-tap inner linear system. The benefit of modeling the inner linear system has been revealed by comparing the SIC performance with Hammerstein nonlinear model. The performance has also been compared to well known black box models such as Generalized Memory Polynomial (GMP) and Artificial Neural Networks (ANN).*

Keywords

Wiener-Hammerstein nonlinear system, digital self-interference cancellation, in-band full-duplex communications

1. Introduction

RF Power Amplifiers (PA) may introduce substantial amount of nonlinear interference while amplifying the signal for transmission. PAs are often power-inefficient and need to be operated near the saturation point to obtain high signal powers at the output. Therefore, special cooling and mechanical techniques are applied in the hardware design of the devices including PA components, which can sometimes be met by designing external modules for this purpose. These requirements entail board-to-board or module-to-module connections over cables at the input and output of the PA components. Any wired or wireless reflections at the input and output of these nonlinear devices act as linear systems, which introduce memory effects in the nonlinear system characteristics and make the system very hard to identify.

Identification of nonlinear systems with memory effects has been one of the hottest topics in wireless communications for the last two decades especially after OFDM based signaling has become widely used in many standards. Having significant advantages in channel estimation and equalization compared to single-carrier modulation schemes, OFDM systems are very popular especially for non-line-of-sight (nLOS) and frequency selective wireless channel scenarios. However, these systems significantly suffer from the PA nonlinearity due to their multi-carrier nature that introduces high peak-to-average power ratios and inter-modulation product terms. The resulting in-band interference causes severe degradation in the transmit signal quality that inhibits using high modulation orders. In addition, nonlinearity also introduces out-of-band interference that cause spectral regrowth and hence failure in meeting the spectral mask requirements. Linearization of these devices by digital pre-distortion (DPD) is a popular and practical solution, however performance with DPD directly depends on the accuracy of the system identification [1], [2].

Another motivation for the accurate modeling of such systems has arisen from a rather new but emerging technology called in-band full-duplex communication (IBFD) [3]. IBFD requires that transmission and reception occur simultaneously using the same frequency resources. The challenge here is to cancel the transmitted signal, which is distorted by linear and nonlinear interference, down to the receiver noise floor so that the desired signal can be received as in

DOI: 10.13164/re.2024.0265

the traditional half-duplex systems. Since PA nonlinearity with memory effects constitutes the main source for the self-interference, accurate modeling of the PA system plays a crucial role in the digital self-interference cancellation (DSIC) process [4].

Wiener-Hammerstein nonlinear (WHNL) system modeling is widely used in the literature to model PA with memory effects in the digital base-band, where a memoryless nonlinear function is located between two cascaded linear systems [5], [6]. Extraction of the memoryless nonlinear characteristics is possible by offline processing of the input and output digital baseband samples of the target system. For example, in [7], authors exploit a dynamic moving average method, whereas in [8] an artificial neural network structure, which is trained by the magnitude and phase of the base-band samples, is used for this purpose. Spline-based interpolation is another popular approach that can achieve high accuracy in modeling the memoryless nonlinear behavior [9].

Although various techniques have been successful in estimating the memoryless nonlinearity, extracting the inner and outer linear systems is a much more stringent work that limits modeling performance. A common approach is to place finite impulse response filter structures to represent the memory effects. The coefficients are mostly estimated by using least squares (LS) or gradient search methods. Another method is to use black box structures, where base-band equivalent input/output signal pairs are used for estimating the unknown parameters of the WHNL system. Most common black box models are Generalized Memory Polynomial (GMP) and Artificial Neural Network (ANN) structures. GMP is a linear representation of the target system based on the simplified Volterra Series expansion of the nonlinear systems with memory. Nonlinear behavior is represented by the basis functions mostly defined by the power degrees of the input magnitude. The coefficients are estimated with the LS method by using the input/output sample pairs. ANN is another popular approach to model such systems as a black box structure. It has parallel computational blocks called neurons including activation functions to represent nonlinear behavior of the target system. Delayed samples of the input signal are fed to the layers of neurons and a weighted sum of the neuron outputs are collected at the output. The weights are calculated by back-propagation algorithm over a training period using the input/output samples. Although they are very popular and easy to use as black box models for an unknown nonlinear system, the performance of these common models gets saturated even with a high number of coefficients, and therefore they have limited use for the digital SIC process.

In this paper, a novel approach has been proposed for the identification of the inner and outer linear system parameters assuming that the memoryless nonlinear function is given. Our method first isolates the linear and nonlinear parts by exploiting the Bussgang decomposition method. Then the inner linear system is estimated by using the DFT representation of the system. Finally the outer linear system is updated accordingly by minimizing the model estimation error.

2. System Model

A general digital baseband representation of a WHNL system output is given in (1), where $\psi(\cdot)$ represents the memoryless NL function, $p[n]$ and $q[n]$ represent the inner and outer linear systems, respectively, and $z[n]$ represents the receiver noise that is assumed to have white Gaussian distribution (AWGN). Block diagram representation of the system is shown in Fig. 1.

$$r[n] = \psi\big(x[n] * p[n]\big) * q[n] + z[n] \tag{1}$$

The output of the memoryless NL function can be represented as the multiplication of the input samples by the corresponding gain samples as follows:

$$b[n] = a[n] g_a[n] \tag{2}$$

where $a[n]$ is the input sequence and $g_a[n]$ is the time sequence specific to the combined effect of the input and the nonlinearity. Using this, we can rewrite the system equation as:

$$r[n] = \overbrace{\Big[\underbrace{(x[n] * p[n])}_{a[n]}\, g_a[n]\Big]}^{b[n]} * q[n] + z[n]. \tag{3}$$

Note that, any circularity in time provided at the input is preserved at the output for the above relation. Hence, under circularity conditions, the analysis can be continued by taking the N-point DFT of both sides as in

$$R[k] = (1/N)\Big[(X[k]P[k]) \circledast_N G_a[k]\Big] Q[k] + Z[k] \tag{4}$$

where the capital letters represent the N-point DFT coefficients of the corresponding time domain signals for the frequency index $k \in \{0, 1, \ldots, N-1\}$. In the above expression, $\circledast_N$ is the N-point circular convolution operation and $G_a[k]$ is the N-point DFT of $g_a[n]$. This expression can be represented as:

$$\mathbf{R} = \frac{1}{N} \mathbf{D_Q} \mathbf{C}_{\mathbf{G}_a} \mathbf{D_P} \mathbf{X} + \mathbf{Z} \tag{5}$$

where $\mathbf{D_Q}$ and $\mathbf{D_P}$ are the diagonal matrices of the $\mathbf{Q}$ and $\mathbf{P}$, which are N-point DFT coefficient vectors of $q[n]$ and $p[n]$, respectively. Similarly $\mathbf{R}$, $\mathbf{X}$, and $\mathbf{Z}$ are N-point DFT vectors of $r[n]$, $x[n]$, and $z[n]$, respectively. $\mathbf{C}_{\mathbf{G}_a}$ (6) is the circulant matrix of $\mathbf{G}_a$, which is the DFT coefficient vector of the gain sequence, $g_a[n]$.

$$\mathbf{C}_{\mathbf{G}_a} \triangleq \begin{bmatrix} G_a[0] & G_a[N-1] & \ldots & G_a[1] \\ G_a[1] & G_a[0] & \ldots & G_a[2] \\ \vdots & \vdots & \ddots & \vdots \\ G_a[N-1] & G_a[N-2] & \ldots & G_a[0] \end{bmatrix}. \tag{6}$$

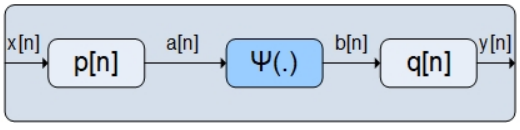


Fig. 1. Wiener-Hammerstein nonlinear model.

2.1 Estimation of the Overall Linear System

In order to identify the system parameters, we utilize the Bussgang Theorem [10], [11] to find out the input-output relation of the $\psi(\cdot)$ function. According to this theorem, the cross-correlation of the input and output of a memoryless NL system, when excited by a stationary Gaussian random process, is proportional to the input auto-correlation function by a scale factor which depends on the memoryless NL function and input distribution. That is,

$$\begin{aligned} R_{ba}[l] &= K_{\psi,a} \cdot R_{aa}[l], \\ K_{\psi,a} &= \mathbb{E}\{\psi'(a)\} \end{aligned} \tag{7}$$

where $R_{ba}[l]$ and $R_{aa}[l]$ represent the cross-correlation and auto correlation functions of time difference index, l, respectively. $\mathbb{E}\{.\}$ is the expectation operation with respect to input signal distribution, and $\psi'(a)$ is the first order derivative with respect to the input, $a[n]$. In other words, by using this theorem, the output of the nonlinearity can be decomposed into its fully correlated and uncorrelated components as follows:

$$b[n] = K_{\psi,a} \cdot a[n] + \phi[n] \tag{8}$$

where $\phi[n]$ is the uncorrelated signal component. By using this property, the system can be expressed as follows:

$$\begin{aligned} r[n] = \underbrace{K_{\psi,x}}_{\text{Scalar}} \cdot (x[n] * \underbrace{p[n] * q[n]}_{\triangleq h[n]}) \\ + \underbrace{\phi[n] * q[n]}_{\text{uncorrelated with x[n]}} + \underbrace{z[n]}_{\text{AWGN}}. \end{aligned} \tag{9}$$

Here, we define $h[n] \triangleq p[n] * q[n]$ as the overall linear system. Then, DFT representation of the above expression is

$$R[k] = K_{\psi,x} X[k] H[k] + \Phi[k] Q[k] + Z[k] \tag{10}$$

for $k = 0, 1, \ldots, N-1$. Assuming that the input is a zero-mean stationary Gaussian process, the correlation of the received vector with the input vector yields

$$\begin{aligned} \mathbb{E}\{R[k]X^*[k]\} &= K_{\psi,x} H[k] \mathbb{E}\{|X[k]|^2\} \\ &+ \underbrace{\mathbb{E}\{\Phi[k]X^*[k]\}}_{0} Q[k] + \underbrace{\mathbb{E}\{Z[k]X^*[k]\}}_{0}. \end{aligned} \tag{11}$$

Exciting the system by uncorrelated input sequences at each iteration, i, we propose a sub-optimum estimator for the overall linear channel as:

$$\hat{H}[k] = \frac{\sum_{i=1}^{n} R[k]^{(i)} X^*[k]^{(i)}}{\sum_{i=1}^{n} |X[k]^{(i)}|^2}. \tag{12}$$

Note that the complex scale factor, $K_{\psi,x}$, is assumed to be unity for this estimator. Resulting gain and phase offset for the overall linear system estimator based on this assumption is left to be compensated in the last stage of the estimation process, where the outer linear system parameters are updated adaptively as explained in Sec. 2.3.

2.2 Estimation of the Inner LTI System: $p[n]$

Having an estimator for the overall linear system, $\hat{H}[k] \triangleq P[k]Q[k]$, and the knowledge of the memoryless NL system, $\psi(\cdot)$, our first goal is to identify the DFT representation of the inner LTI system, $P[k]$. Since the nonlinearity is located in between the linear systems, $\hat{H}[k]$ and $\psi(\cdot)$ cannot be directly used in our system equation (4), because $P[k]$ and $Q[k]$ are not resolved, and $G_a[k]$ is not known. However, in order to exploit the $\psi(\cdot)$ function, we assume that the original signal has passed through it and obtain the following:

$$\begin{aligned} \psi(x[n]) &= x[n] g_x[n], \\ \Psi_x[k] &= (1/N)(X[k] \circledast_N G_x[k]), \\ \mathbf{\Psi}_x &= \frac{1}{N} \mathbf{C}_{\mathbf{G}_x} \mathbf{X} \end{aligned} \tag{13}$$

where, $g_x[n]$ is the time sequence that depends on the input, and $G_x[k]$ is the corresponding DFT coefficients for $k = 0, 1, \ldots, N-1$. The last equality represents the linear algebraic representation where $\mathbf{C}_{\mathbf{G}_x}$ is the circulant matrix obtained by the $G_x[k]$ values similar to the one shown in (6). With the above assumption an error is introduced that can be defined as:

$$\mathbf{C}_{\mathbf{G}_\Delta} = \mathbf{C}_{\mathbf{G}_a} - \mathbf{C}_{\mathbf{G}_x} \tag{14}$$

where $\mathbf{C}_{\mathbf{G}_\Delta}$ is the circulant matrix for the difference vector, $G_\Delta[k] = G_a[k] - G_x[k]$ for $k = 0, 1, \ldots, N-1$. Replacing $\mathbf{C}_{\mathbf{G}_a} = \mathbf{C}_{\mathbf{G}_x} + \mathbf{C}_{\mathbf{G}_\Delta}$ and $\mathbf{D_Q} = \mathbf{D}_{\hat{\mathbf{H}}} \mathbf{D}_{\mathbf{P}}^{-1}$ in the system equation (5), we get

$$\mathbf{D}_{\mathbf{P}}^{-1} \mathbf{C}_{\mathbf{G}_x} \mathbf{D_P X} + \underbrace{\mathbf{D}_{\mathbf{P}}^{-1} \mathbf{C}_{\mathbf{G}_\Delta} \mathbf{D_P X} + N \cdot \mathbf{D}_{\hat{\mathbf{H}}}^{-1} \mathbf{Z}}_{\triangleq \mathbf{E} \in \mathbb{C}^{N \times 1}} = \underbrace{N \cdot \mathbf{D}_{\hat{\mathbf{H}}}^{-1} \mathbf{R}}_{\triangleq \mathbf{V} \in \mathbb{C}^{N \times 1}}. \tag{15}$$

Here we have grouped the noise and error terms together and renamed as $\mathbf{E} \in \mathbb{C}^{N \times 1}$ vector. Similarly the known vector term on the right-hand-side is renamed as $\mathbf{V} \in \mathbb{C}^{N \times 1}$. Then equation (15) can be rewritten as

$$\mathbf{D}_{\mathbf{P}}^{-1} \left[\mathbf{G}_x \circledast_N \mathbf{D_P X} \right] + \mathbf{E} = \mathbf{V} \tag{16}$$

where $\mathbf{G}_x$ is the DFT coefficient vector of the $g_x[n]$ sequence defined in (13). Focusing on the k^{th} row of the vector equation (16) gives

$$P[k]^{-1} \mathbf{P}^{\mathrm{T}} \underbrace{\left\{ X[j] G_x\left[((k-j))_N\right] \right\}_{j=0}^{N-1}}_{\triangleq \mathbf{M}_k \in \mathbb{C}^{N \times 1}} + E[k] = V[k] \tag{17}$$

where $(\cdot)^{\mathrm{T}}$ is the vector transpose and $((\cdot))_N$ is the modulo operation with respect to N. Note that, although the newly defined $\mathbf{M}_k \in \mathbb{C}^{N\times 1}$ is a known vector, equation (17) is not sufficient alone for solving $P[k]$'s due to the unknown $\mathbf{P} \in \mathbb{C}^{N\times 1}$ vector. However, multiple observations of the system output for uncorrelated $\mathbf{X}$ vectors will provide a solution. Let the system be excited by L ($L \geq N$) such input sequences, then equation (17) can be extended as

$$P[k]^{-1}\mathbf{P}^{\mathrm{T}} \underbrace{\left[\mathbf{M}_k^{(1)},\ldots,\mathbf{M}_k^{(L)}\right]}_{\triangleq \mathbf{M}_k\in\mathbb{C}^{N\times L}} + \underbrace{\left[E[k]^{(1)},\ldots,E[k]^{(L)}\right]}_{\triangleq \mathbf{E}_k\in\mathbb{C}^{1\times L}} = \underbrace{\left[V[k]^{(1)},\ldots,V[k]^{(L)}\right]}_{\triangleq \mathbf{V}_k\in\mathbb{C}^{1\times L}}. \tag{18}$$

For each k index, the least squares (LS) estimate for the term $(P[k]^{-1}\mathbf{P}^{\mathrm{T}}) \in \mathbb{C}^{1\times N}$ can be obtained as

$$\widehat{(P[k]^{-1}\mathbf{P}^{\mathrm{T}})}_{\mathrm{LS}} = \mathbf{V}_k\mathbf{M}_k^{\dagger} \tag{19}$$

where $\mathbf{M}_k^{\dagger} \in \mathbb{C}^{L\times N}$ is the pseudo-inverse, $\mathbf{M}_k^{\mathrm{H}}(\mathbf{M}_k\mathbf{M}_k^{\mathrm{H}})^{-1}$, of $\mathbf{M}_k$ with $\mathbf{M}_k^{\mathrm{H}}$ denoting the Hermitian transpose. Letting the estimation error vector be $\boldsymbol{\delta}_{\mathrm{LS}} \in \mathbb{C}^{1\times N}$, equation (19) can be written as

$$P[k]^{-1}\mathbf{P}^{\mathrm{T}} + \boldsymbol{\delta}_{\mathrm{LS}} = \mathbf{V}_k\mathbf{M}_k^{\dagger}. \tag{20}$$

Taking the average of both sides over the vector entries, we get

$$\underbrace{P[k]^{-1}\frac{1}{N}\sum_{i=0}^{N-1}P[i]}_{\mathcal{F}^{-1}\{\mathbf{P}\}\big|_{n=0}} + \overline{\boldsymbol{\delta}_{\mathrm{LS}}} = \overline{\mathbf{V}_k\mathbf{M}_k^{\dagger}} \tag{21}$$

and

$$P[k]^{-1}p[0] + \overline{(\boldsymbol{\delta}_{\mathrm{LS}})} = \overline{(\mathbf{V}_k\mathbf{M}_k^{\dagger})} \tag{22}$$

where $\overline{(\cdot)}$ represents the averaging operation over the terms of a vector. Note that $\mathbf{P}$ is the DFT coefficient vector, hence, averaging over the entries of $\mathbf{P}$ gives the inverse DFT of $\mathbf{P}$ at $n = 0$, that is, $p[0]$. This helps us extract the target $P[k]$'s alone for all k. Equation (22) can be interpreted as an estimation problem where $P[k]$ is to be estimated over a single observation, $\overline{\mathbf{V}_k\mathbf{M}_k^{\dagger}}$, under an additive noise, $\overline{(\boldsymbol{\delta}_{\mathrm{LS}})}$, with unknown statistics. Since $p[0]$ is the main tap of the inner linear system that represents the direct path to the memoryless nonlinearity, it can be normalized to $p[0] = 1$ during the estimation of $P[k]$'s. Under these assumptions, the estimator for the inner memory parameters can simply be obtained as

$$\hat{P}[k] = \left[\overline{(\mathbf{V}_k\mathbf{M}_k^{\dagger})}\right]^{-1}. \tag{23}$$

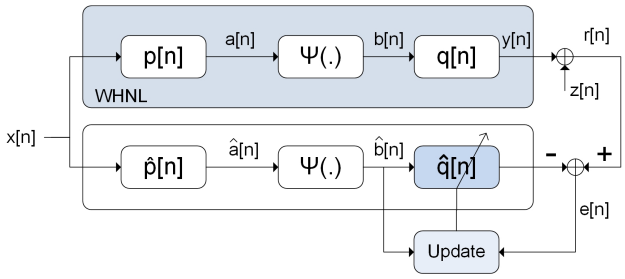


Fig. 2. Update for the outer linear system.

2.3 Estimation of the Outer LTI System: $q[n]$

Having the estimate of the inner LTI system, coefficients of the outer LTI system are updated to match the desired system output as shown in Fig. 2.

Time domain analysis of the update process is given in (24)–(26) as

$$\underbrace{\psi(x[n]*\hat{p}[n])}_{\hat{b}[n]} * \hat{q}[n] + e[n] = r[n], \tag{24}$$

$$\hat{\mathbf{B}}\hat{\mathbf{q}} + \mathbf{e} = \mathbf{r}, \tag{25}$$

$$\begin{aligned} \hat{\mathbf{B}}^{(i)}(\hat{\mathbf{q}}^{(i)} + \Delta\mathbf{q}^{(i+1)}) &= \mathbf{r}^{(i)}, \\ \hat{\mathbf{B}}^{(i)}\Delta\mathbf{q}^{(i+1)} &= \underbrace{\mathbf{r}^{(i)} - \hat{\mathbf{B}}^{(i)}\hat{\mathbf{q}}^{(i)}}_{\mathbf{e}^{(i)}}, \\ \Delta\mathbf{q}^{(i+1)} &= [\hat{\mathbf{B}}^{(i)}]^{\dagger}\mathbf{e}^{(i)} \end{aligned} \tag{26}$$

where $\hat{\mathbf{B}}$ represents the convolution matrix of $\hat{b}[n]$, whereas $\hat{\mathbf{q}}$, $\mathbf{e}$, and $\mathbf{r}$ represent vector forms for the estimated filter coefficients, error signal, and received signal, respectively. In (26), i shows the iteration number and $[.]^{\dagger}$ represents the pseudo-inverse operation. This update process initially sets $\hat{\mathbf{q}}^{(0)}$ vector as the estimated coefficient values from the previous step by using $\mathcal{F}^{-1}\{\hat{\mathbf{H}}/\hat{\mathbf{P}}\}$ relation, where $\hat{\mathbf{H}}$ and $\hat{\mathbf{P}}$ are the estimated DFT coefficient vectors of the overall and inner linear systems, respectively, and $\mathcal{F}^{-1}\{\cdot\}$ represents the inverse DFT operation. Then, the coefficient vector is updated by $\Delta\mathbf{q}$ so that $\mathbf{e}$ is forced to be a zero-vector for the next iteration iteration.

3. Verification by Simulation

The proposed method has been evaluated using MATLAB simulation environment in an IBFD communication scenario as shown in Fig. 3. In order to evaluate the estimation performance, a target WHNL system has been obtained with a sample set of the inner and outer linear system parameters (Tab. 1), and a Saleh type memoryless nonlinear function [12] (Tab. 2). During the simulations, it has been

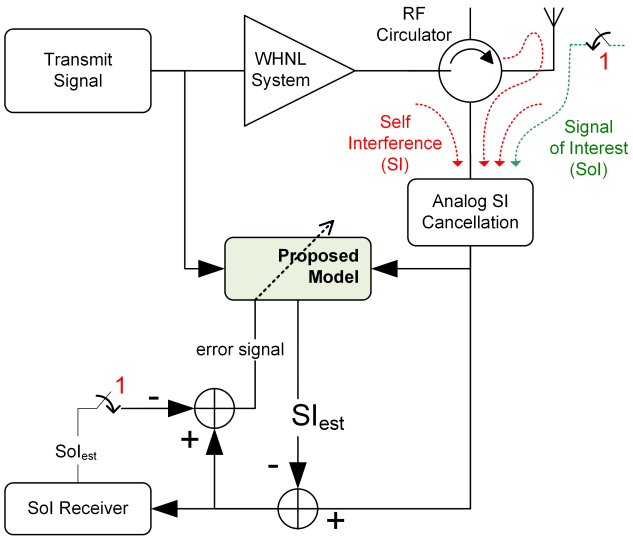


Fig. 3. IBFD simulation scenario.

Channel	Tap	Level [dB]	Delay ($\times T_s$)
$p[n]$	1 (main)	0	0.05
	2	–11	2.17
	3	–20	5.33
$q[n]$	1 (main)	0	0.23
	2	–7	1.87
	3	–13	6.73

Tab. 1. Linear channel parameters.

AM-AM Characteristics	AM-PM Characteristics
$f_{\mathrm{AM}}(x) = \frac{\alpha \cdot \lvert x\rvert}{1+\beta \cdot \lvert x\rvert^2}$	$f_{\mathrm{PM}}(x) = \frac{\alpha \cdot \lvert x\rvert^2}{1+\beta \cdot \lvert x\rvert^2}$
$(\alpha = 1.20, \beta = 0.05)$	$(\alpha = 0.2, \beta = 10)$

Tab. 2. Saleh model for memoryless nonlinearity (input power normalized to $||\mathbf{x}||^2 = 1$).

assumed that analog SI cancellation has already been done so that the SI signal power level is reduced down to the practical analog-to-digital converter dynamic range. It has also been assumed that WHNL system characteristics don't change in the time period of estimation and SoI receiving processes. The base-band sampling frequency has been set to 61.44 Msps and the delay values have been chosen to be fractional in terms of the sampling period, $T_s \cong 16.27$ ns, to simulate the reflections in the analog domain. Please note that delay values that are integer multiples of T_s usually lead to better performance yet do not reveal the performance in real operation. OFDM signals with FFT size (N) of 256 have been used for the channel estimation and SI cancellation steps.

3.1 Calibration

In the first stage of the simulations, the system shown in Fig. 3 has been calibrated over the SI signal without any SoI component by setting switch-1 as as open. During this process, the inner and outer linear systems are estimated over the SI channel by applying the equations (12)–(23). Figure 4 shows the target and the estimated linear system parameters obtained by this process. Using these estimation results and the memoryless nonlinear function, WHNL system model has been extracted. In the following stages SoI is injected by setting switch-1 closed and the receiver performance is compared to a traditional half-duplex communication scenario where SoI is solely received either in a separate time or frequency band without any SI signal component on it.

3.2 IBFD Communication Performance by SI Power

In the second stage of the simulations, randomly generated OFDM-based transmit signals have continuously been passed through the original system and the estimated model. Similarly, randomly generated OFDM-based SoI component existed on the receiver. SI cancellation has been done by adaptively subtracting the model output from the received signal. The performance of the proposed digital SI cancellation (DSIC) method has been measured with the error-vector-magnitude (EVM) metric of the SoI by the formula

$$\mathrm{EVM}_{\mathrm{SoI}} = \frac{1}{S_{\max}} \sqrt{\frac{1}{N_{\mathrm{sc}}} \sum_{i \in I} \left| X_{\mathrm{SoI}}[i] - \hat{X}_{\mathrm{SoI}}[i] \right|^2} \qquad (27)$$

where X_{SoI} are the message symbols, $\hat{X}_{\mathrm{SoI}}$ are the estimated symbols, I is the set of nonzero subcarriers, N_{sc} is the total number of nonzero subcarriers, and $S_{\max}$ is the maximum amplitude on the constellation diagram of SoI depending on the modulation order. In order to examine the DSIC performance under various SI power levels, an example IBFD scenario has been built. In this scenario, the average power of SoI is kept at a constant level at −69 dBm and receiver noise power is set to −74 dBm. Under these circumstances, EVM of the SoI has been measured as −12.3 dB for the half-duplex case where no SI signal existed. This value has been set as the target for the proposed DSIC process. Figure 5 depicts DSIC performance for such a scenario under two different residual SI signal power levels remained after analog cancellation stages. In Fig. 5(a), residual SI signal with 0 dBm power can only be reduced down to −66 dBm, making the IBFD communication impractical with −2.4 dB EVM value. On the other hand, a better analog cancellation giving 20 dB decrease in the SI power (Fig. 5(b)) can achieve the target EVM value with only 0.63 dB loss compared to the half-duplex communication. As shown in the figures, the proposed method can achieve up to 66 dB digital cancellation performance and can reduce the residual SI signal down to noise floor provided that effective analog cancellation steps are applied. Figure 6 shows that the SI power dramatically affects the DSIC performance and needs to be kept as low as possible for an acceptable EVM loss compared to half-duplex communication.

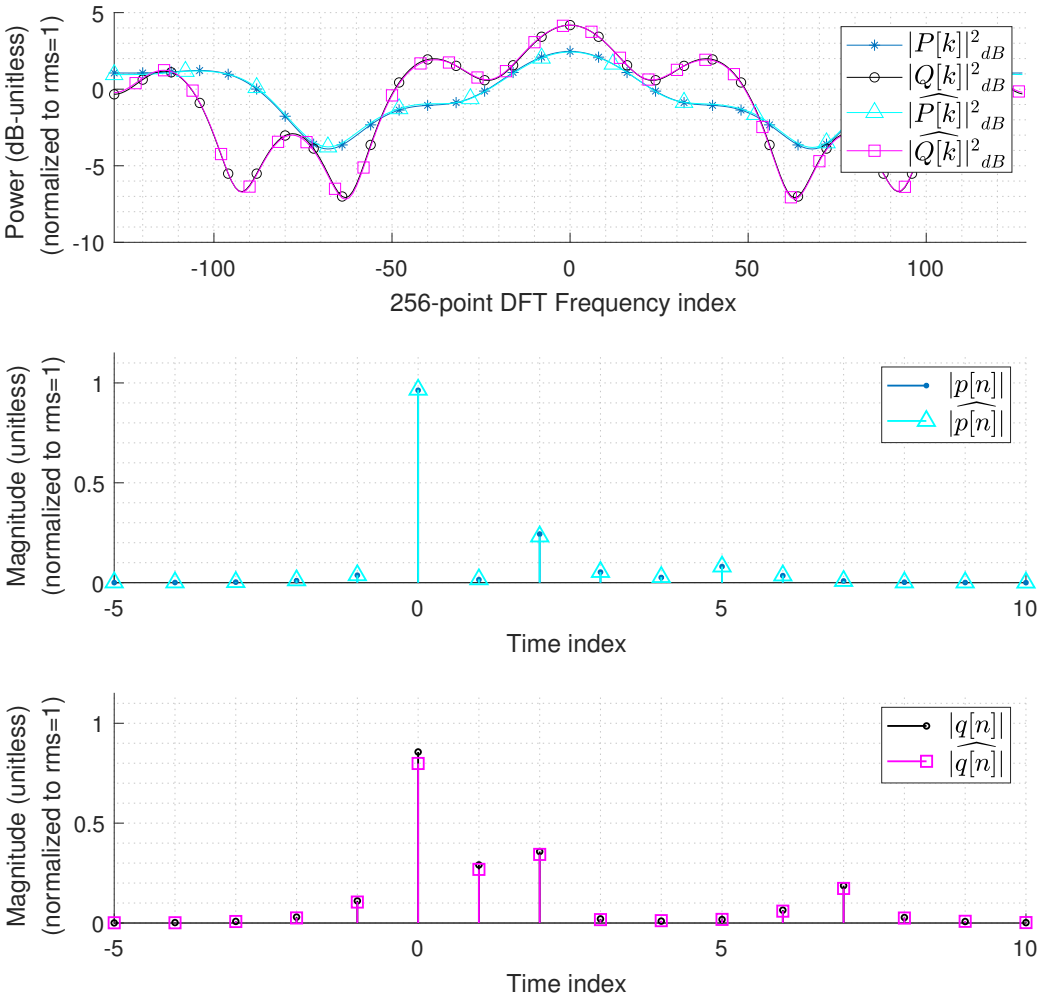
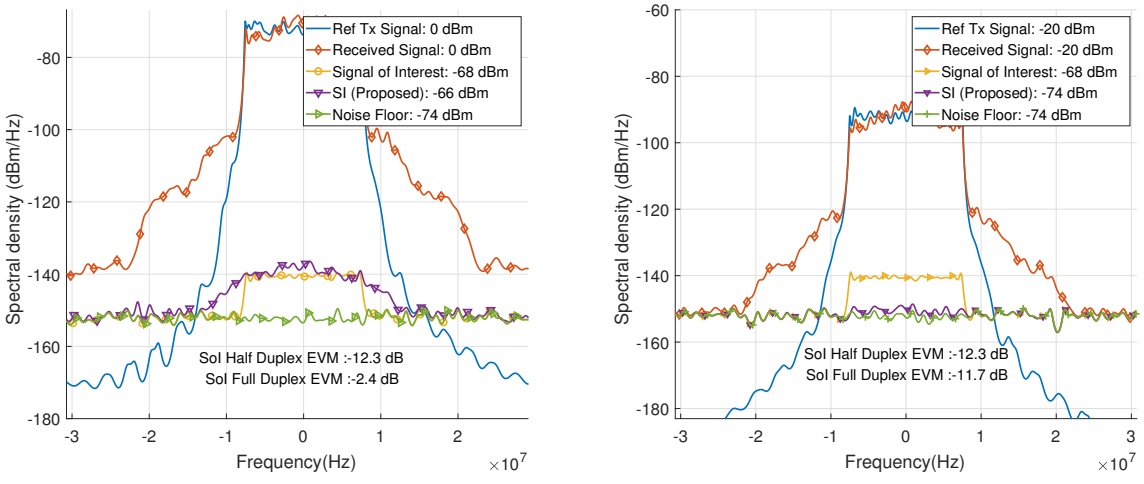


Fig. 4. Linear channel characteristics and estimation results.



(a) $P_{\mathrm{SI}} = 0\,\mathrm{dBm}$ (b) $P_{\mathrm{SI}} = -20\,\mathrm{dBm}$

Fig. 5. Power spectrum performance under two different residual SI power levels after analog SIC.

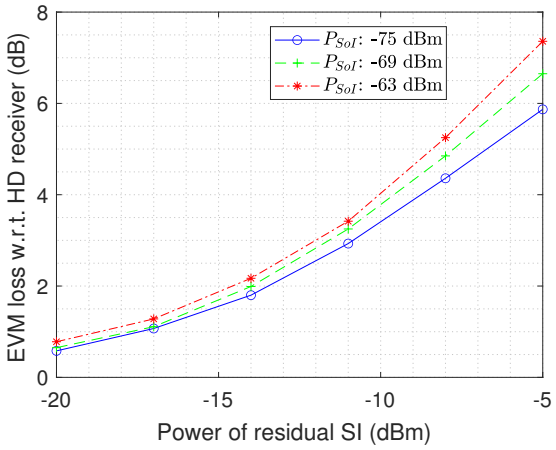


Fig. 6. EVM loss compared to half-duplex receiver versus residual SI power after analog SIC.

3.3 IBFD Communication Performance by Inner Linear System Parameters and SoI Power

The performance of the proposed method has also been examined under varying inner linear system parameters of the target WHNL system. Figure 7 gives the performance results for different delay and gain values of the second tap of a two-tap inner linear system. In Fig. 7(a), the delay is fixed to $1.5 \times T_s$ and the gain is changed from -1 dB to -15 dB. In Fig. 7(b), the gain is set to 7 dB less than the main tap and the delay is changed from $0.1 \times T_s$ to $2.9 \times T_s$. In both figures, the horizontal axis shows the increasing SoI power. It has been observed that EVM performance decreases with the increasing delay and power of the second tap, which result in severe in-band frequency selective channel effects. Note also that the performance decreases as the SoI power increases. This is because the SoI signal introduces an uncorrelated interference for the proposed estimation processes. The effect of SoI and an iterative approach to cancel the SI under SoI contamination is examined more comprehensively in [13].

3.4 IBFD Communication Performance Comparison by Hammerstein Model

The effect of the inner linear system has been examined in Fig. 8 by giving a comparison between the proposed method (Fig. 8(b)) with the Hammerstein model (Fig. 8(a)), where the inner linear system is ignored but the outer linear system is estimated adaptively. The contours in the figures represent the EVM metric values of the models for a 2-tap inner linear system. The horizontal and vertical axes represent the increasing delay and power of the 2nd tap with respect to the main tap, respectively. The rectangular regions in both figures represent the practical delay and power levels for a real PA hardware setup. The EVM performance improvement by estimating the inner linear system is highlighted in this region. As seen in Fig. 8(a), Hammerstein model suffers from the inner linear system even for very low delay and gain values and EVM value deviates significantly in this region. However, in Fig. 8(b), the proposed model behaves much more robustly and gives acceptable EVM values comparable with the half-duplex case even for the upper-right corner of the rectangular region showing more drastic delay and gain values.

3.5 Performance Comparison by Black Box Models

The proposed method has been compared to SI cancellation results obtained by the two well known black box system models, namely generalized memory polynomial (GMP) and neural networks (NN). GMP model [14] function and parameter estimation procedure is given in (28a) and (28b), respectively.

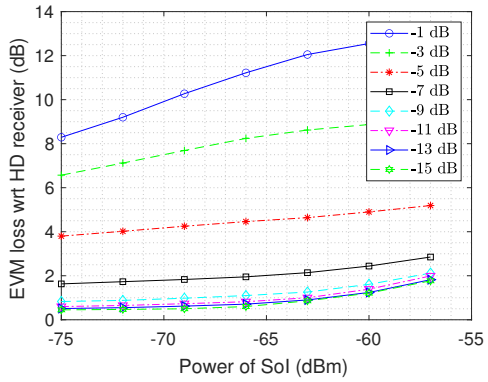
$$\Psi_{\mathrm{GMP}}(x_n) = \sum_{k=0}^{K} \sum_{l=-L}^{L} \sum_{m=-M}^{M} \alpha_{k,l,m} x_{n-l} |x_{n-l-m}|^k, \quad (28a)$$

$$\boldsymbol{\alpha} = (\Phi^{\mathrm{H}}\Phi)^{-1}\Phi^{\mathrm{H}}\boldsymbol{\Psi}_{\mathbf{x}}. \quad (28b)$$

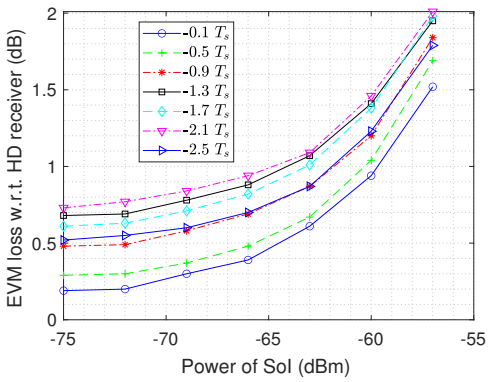
Here, $\boldsymbol{\Psi}_{\mathbf{x}} \in \mathbb{C}^{N\times 1}$ represents the output vector of the NL system excited by the input vector $\mathbf{x} \triangleq [x_0 x_1 \ldots x_{N-1}]^{\mathrm{T}}$. $\boldsymbol{\alpha} \in \mathbb{C}^{P\times 1}$ is the parameter vector with $P \triangleq (K+1)(2L+1)(2M+1)$. $\boldsymbol{\Phi} \in \mathbb{C}^{N\times P}$ is the range space matrix, columns of which form the basis vectors obtained by computing $x_{n-l}|x_{n-l-m}|^k$ for each $k \in \{0, 1, \ldots, K\}$, $l \in \{-L, \ldots, L\}$, and $m \in \{-M, \ldots, M\}$. The variables for the GMP model have been chosen as $K = 3$, $L = 5$, and $M = 4$ during the simulations (396 coefficients). As for the neural network model, Real Valued Time Delay Neural Network (RVTDNN) structure has been used [15]. Delayed samples of the real and imaginary components of the input sequence are fed in parallel to the input layer of the network. Linear combinations of these parallel inputs are then level-shifted by a bias parameter and exposed to a nonlinear activation function at each neuron of the hidden layer, which form the second stage of the network. Finally, the outputs of these neurons are again linearly combined and biased to give the real and imaginary components of the desired signal as summarized in (29):

$$\Psi_{\mathrm{NN}}(\mathbf{x}_{\mathrm{IN}}) = \mathbf{W}_{\mathrm{OL}} \cdot \psi(\mathbf{W}_{\mathrm{HL}} \cdot \mathbf{x}_{\mathrm{IN}} + \mathbf{b}_{\mathrm{HL}}) + \mathbf{b}_{\mathrm{OL}} \quad (29)$$

where $\mathbf{x}_{\mathrm{IN}} \triangleq [\Re(\mathbf{x})\ \Im(\mathbf{x})]^{\mathrm{T}}$ for the input vector defined as $\mathbf{x} \triangleq [x_{n+M} \ldots x_n \ldots x_{n-M}]$. NN has M input delay elements both in the positive and negative time axis, and H neurons in the hidden layer. $\mathbf{W}_{\mathrm{OL}} \in \mathbb{R}^{2\times H}$ and $\mathbf{W}_{\mathrm{HL}} \in \mathbb{R}^{H\times(4M+2)}$ represent the weight matrices, whereas $\mathbf{b}_{\mathrm{OL}} \in \mathbb{R}^{2\times 1}$ and $\mathbf{b}_{\mathrm{HL}} \in \mathbb{R}^{H\times 1}$ represent the bias vectors for the output and hidden layers, respectively. Hyperbolic tangent sigmoid function, $\psi(x) = (1 - \mathrm{e}^{-2x})/(1 + \mathrm{e}^{-2x})$, has been used as the nonlinear activation function inside the hidden layer neurons. The coefficient estimation process used the Levenberg-Marquardt back-propagation algorithm. The size of the NN structure has been set by choosing $H = 18$ and $M = 4$ during the simulations (380 coefficients). Figure 9 shows the SI cancellation performance of these two black box models

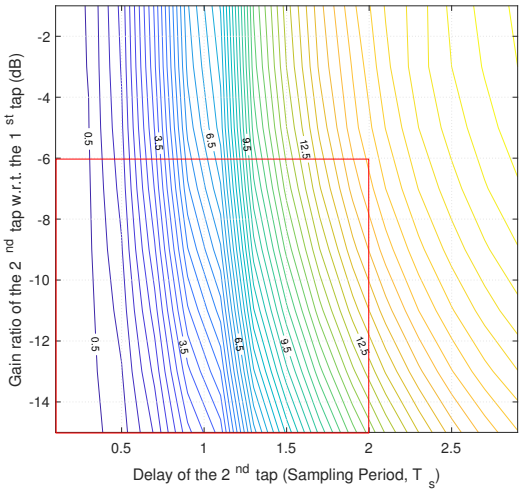


(a) DSIC performance compared to HD receiver for a fixed second-tap delay of $1.5 \times T_s$, where T_s is the sampling period

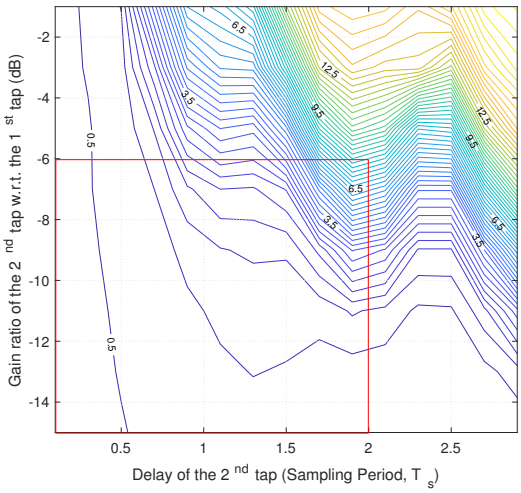


(b) DSIC performance compared to HD receiver for a fixed second-tap gain ratio of -7 dB

Fig. 7. IBFD EVM loss with respect to inner linear system parameters under increasing SoI power levels.

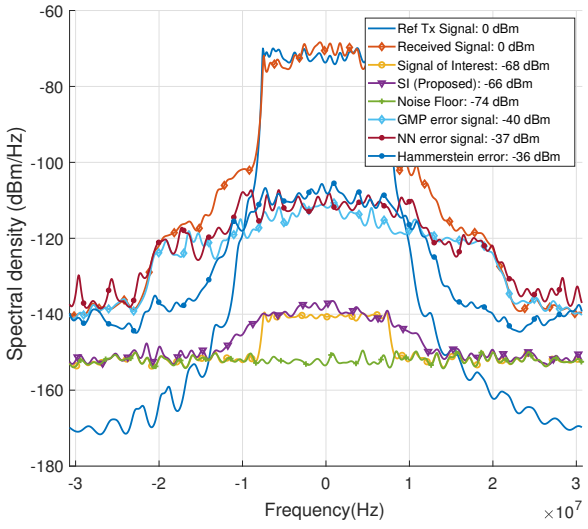


(a) EVM loss by Hammerstein model (memoryless nonlinearity followed by outer memory)

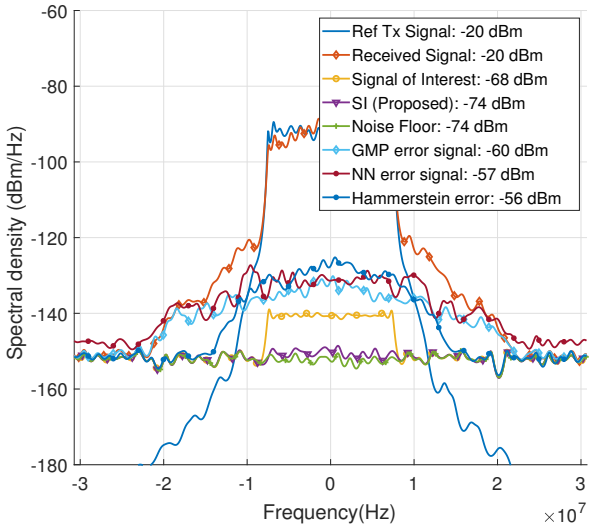


(b) EVM loss by Wiener-Hammerstein model (inner memory estimated by the proposed method)

Fig. 8. Contour plots of the EVM loss compared to HD communication under the effect of gain and delay parameters of the inner system.



(a) $P_{\mathrm{SI}} = 0$ dBm



(b) $P_{\mathrm{SI}} = -20$ dBm

Fig. 9. Comparison with black box and Hammerstein models under two different residual SI power levels after analog SIC.

compared to the proposed algorithm under to SI analog cancellation levels. It can be observed that GMP and RVTDNN models can only cancel the SI signal about 40 dB, whereas the proposed method suppress the SI signal about 65 dB for this scenario. Although GMP and NN models have inner and outer memory registers to imitate the target WHNL system, they suffer from the lack of memoryless nonlinear function. GMP uses linear combination of polynomial basis functions and neural network tries to fit the target system by nonlinear activation functions operating in parallel. Although the proposed method assumes to have the memoryless nonlinear function, comparison with the Hammerstein model indicates that superiority in cancellation performance is mainly due to the estimation of the inner memory coefficients.

4. Conclusion

This paper presents a contribution for the identification problem of the nonlinear PA devices affected by linear systems with memory. A novel method has been proposed for estimating the linear components of the Wiener-Hammerstein nonlinear system that is widely used for modeling such systems. In this method, the overall linear system has first been derived and obtained by exploiting the Bussgang Decomposition method. Then, a novel analytical approach by using DFT technique has been applied to estimate the inner linear system coefficients. Finally, the outer linear system has been estimated in an adaptive manner to compensate with the amplitude and phase offsets introduced in the earlier stages of the identification process.

The proposed model has been verified by simulations by using OFDM based calibration and test signals under stringent linear and nonlinear system parameters. The performance limits of the proposed model has been examined under increasing residual self interference power leves and existence of the signal-of-interest. The importance of the inner memory effects has been emphasized by comparing the proposed technique with the Hammerstein model. SI cancellation performance has also been compared to well known GMP and NN black box models. Up to 65 dB of digital SI cancellation ratio with the proposed method has been recorded, whereas the performance of the black box methods had saturated around 40 dB of cancellation ratio. This study will be continued by including the memoryless nonlinearity extraction stage and modeling a real PA device on hardware.

References

[1] DING, L., RAICH, R., ZHOU, G. T. A Hammerstein predistortion linearization design based on the indirect learning architecture. In *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*. Orlando (FL, USA), 2002, p. 2689–2692. DOI: 10.1109/ICASSP.2002.5745202

[2] GILABERT, P., MONTORO, G., BERTRAN, E. On the Wiener and Hammerstein models for power amplifier predistortion. In *Asia-Pacific Microwave Conference Proceedings*. Suzhou (China), 2005, p. 1–4. DOI: 10.1109/APMC.2005.1606491

[3] KOLODZIEJ, K. E., PERRY, B. T., HERD, J. S. In-band full-duplex technology: Techniques and systems survey. *IEEE Transactions on Microwave Theory and Techniques*, 2019, vol. 67, no. 7, p. 3025–3041. DOI: 10.1109/TMTT.2019.2896561

[4] TAPIO, V., JUNTTI, M. Non-linear self-interference cancelation for full-duplex transceivers based on Hammerstein-Wiener model. *IEEE Communications Letters*, 2021, vol. 25, no. 11, p. 3684–3688. DOI: 10.1109/LCOMM.2021.3109669

[5] KIBANGOU, A. Y., FAVIER, G. Wiener-Hammerstein systems modeling using diagonal Volterra kernels coefficients. *IEEE Signal Processing Letters*, 2006, vol. 13, no. 6, p. 381–384. DOI: 10.1109/LSP.2006.871705

[6] SAFARI, N., FEDORENKO, P., KENNEY, J. S., et al. Spline-based model for digital predistortion of wide-band signals for high power amplifier linearization. In *IEEE/MTT-S International Microwave Symposium*. Honolulu (HI, USA), 2007, p. 1441–1444. DOI: 10.1109/MWSYM.2007.380504

[7] TAIJUN, L., BOUMAIZA, S., GHANNOUCHI, F. M. Augmented Hammerstein predistorter for linearization of broad-band wireless transmitters. *IEEE Transactions on Microwave Theory and Techniques*, 2006, vol. 54, no. 4, p. 1340–1349. DOI: 10.1109/TMTT.2006.871230

[8] MKADEM, F., BOUMAIZA, S. Extended Hammerstein behavioral model using artificial neural networks. *IEEE Transactions on Microwave Theory and Techniques*, 2009, vol. 57, no. 4, p. 745–751. DOI: 10.1109/TMTT.2009.2015092

[9] PASCUAL CAMPO, P., ANTTILA, L., KORPI, D., et al. Cascaded spline-based models for complex nonlinear systems: Methods and applications. *IEEE Transactions on Signal Processing*, 2021, vol. 69, p. 370–384. DOI: 10.1109/TSP.2020.3046355

[10] DEMIR, O. T., BJORNSON, E. The Bussgang decomposition of nonlinear systems: Basic theory and MIMO extensions [Lecture notes]. *IEEE Signal Processing Magazine*, 2021, vol. 38, no. 1, p. 131–136. DOI: 10.1109/MSP.2020.3025538

[11] ENQVIST, M., LJUNG, L. Linear models of nonlinear FIR systems with Gaussian inputs. In *13th IFAC Symposium on System Identification (SYSID)*. Rotterdam (Netherlands), 2003, p. 1873–1878. DOI: 10.1016/S1474-6670(17)35033-4

[12] SALEH, A. A. M. Frequency-independent and frequency-dependent nonlinear models of TWT amplifiers. *IEEE Transactions on Communications*, 1981, vol. 29, no. 11, p. 1715–1720. DOI: 10.1109/TCOM.1981.1094911

[13] KURT, A., SALMAN, M. B., SARAC, U. B., et al. An adaptive-iterative nonlinear interference cancellation in time-varying full-duplex channels. *IEEE Transactions on Vehicular Technology*, 2023, vol. 72, no. 2, p. 1862–1878. DOI: 10.1109/TVT.2022.3208766

[14] MORGAN, D. R., MA, Z., KIM, J., et al. A generalized memory polynomial model for digital predistortion of RF power amplifiers. *IEEE Transactions on Signal Processing*, 2006, vol. 54, no. 10, p. 3852–3860. DOI: 10.1109/TSP.2006.879264

[15] BOUMAIZA, S., MKADEM, F. Wideband RF power amplifier predistortion using real-valued time-delay neural networks. *European Microwave Conference (EuMC)*. Rome (Italy), 2009, p. 1449–1452. DOI: 10.23919/EUMC.2009.5296072