

Article

# Multiple-Antenna Emitters Identification Based on a Memoryless Power Amplifier Model

Jun Lu and Xiaodong Xu \* 

Department of Electrical Engineering and Information Science, University of Science and Technology of China, Hefei 230026, China; lujun052@mail.ustc.edu.cn

\* Correspondence: xdxu@ustc.edu.cn

Received: 22 October 2019; Accepted: 25 November 2019; Published: 28 November 2019



**Abstract:** Power amplifier (PA) nonlinearity is typically unique at the radio frequency (RF) front-end for particular emitters. It can play a crucial role in the application of specific emitter identification (SEI). In this paper, under the Multi-Input Multi-Output (MIMO) multipath communication scenario, two data-aided approaches are proposed to identify multi-antenna emitters using PA nonlinearity. Built upon a memoryless polynomial model, the first approach formulates a linear least square (LLS) problem and presents the closed-form solution of nonlinear coefficients in a MIMO system by means of singular value decomposition (SVD) operation. Another alternative approach estimates nonlinear coefficients of each individual PA through nonlinear least square (NLS) solved by the regularized Gauss–Newton iterative scheme. Moreover, there are some practical discussions of our proposed approaches about the mismatch of the order of PA model and the rank-deficient condition. Finally, the average misclassification rate is derived based on the minimum error probability (MEP) criterion, and the proposed approaches are validated to be effective through extensively numerical simulations.

**Keywords:** specific emitter identification; multiple antennas; power amplifier nonlinearity; nonlinear least square

## 1. Introduction

Specific emitter identification (SEI) is committed to distinguish individual radiation sources by using essential radio frequency fingerprint (RFF) features extracted from different emitters. It can be applied to military communication [1], physical layer authentication [2], and enhancement of the security in wireless network, such as very high frequency (VHF) radio networks, Wi-Fi networks, cognitive radios, cellular networks [3], and so on.

In general, based on different states of signals, the identifiable RFF features for SEI are usually extracted from the transient signal or the steady-state signal. The transient signal, actually the turn-on signal, carries unique and unintentional information that is advantageous to emitter identification, and the features underlying are mainly extracted from the instantaneous amplitude, phase, frequency, and energy envelope [4–6]. Nevertheless, it is difficult to capture the transient signal since the duration time is often too short to use. As for the steady-state signal, many researchers focus on extracting the distinguishable features that are generated by hardware imperfection of the components inside the radiation source through advanced signal processing techniques. On one hand, the statistical characteristics of the original RF signal, such as the high order spectrum, have been used as the features to identify different emitters [7,8]. On the other hand, the Time-Frequency Analysis (TFA) methods [9,10], Wavelet Transform (WT) [11,12], and Hilbert–Huang Transform (HHT) [13,14] are successively applied to extract the transform domain characteristics from the received RF signals. However, these methods have little knowledge of impairments inside the individual emitter, and the performance can be easily affected by the wireless channels. Hence, there are many additional works

in literature to model the characteristic of hardware imperfection of the internal component such as digital to analog converter (DAC) [15], modulator [16,17], and power amplifier (PA) [18–21], etc., and to extract the unique RF front-end feature for a particular emitter. This paper concentrates on the extraction of PA nonlinear features to identify communication emitters with multiple antennas under Multiple-Input Multiple-Output (MIMO) multipath channels.

It is well known that the MIMO transmission scheme can improve the spectral efficiency by introducing additional spatial diversity. In practice, transceivers with multiple transmit and receive antennas have shown their powerful merits and served as the major mechanism for current and future wireless communication systems. Meanwhile, due to the increasing number of antennas for MIMO emitters, there will be more diverse PA nonlinearities available at the RF front end for SEI. A lot of research efforts have been afforded to analyze the nonlinearity incurred by PAs with the Saleh model [22], polynomial model [23], and Volterra model [24], etc. in MIMO systems, whereas they are mainly devoted to implementing predistortion or PA linearization. However, to the authors' best knowledge, there are few open results of SEI in MIMO communication systems. In [21], Li uses a modified artificial bee colony (ABC) algorithm to estimate the coefficients of the Hammerstein model, a simplified version of the Volterra model, in a MIMO system. However, the impact of wireless propagation channels is not considered, and the ABC algorithm also appears complicated to obtain the optimal solution in the context of SEI. Recently, in [25,26], an estimation of signal parameters via rotational invariance technique (ESPRIT)-based approach, which takes advantage of the multiple antennas at the receiver to separate the RFF from wireless channel, is proposed for RFF estimation in orthogonal frequency division multiplexing (OFDM) systems, whereas it is only suitable for a Single-Input Multiple-Output (SIMO) system rather than the MIMO one. In [19,20], a memoryless polynomial model is used to characterize the nonlinearity of PA, and a data-aided iterative algorithm is proposed to estimate nonlinear coefficients of the PA model for SEI from the observations in both MIMO and single-input single-output (SISO) scenarios.

Treating the fact that all PAs of a multiple-antenna emitter are independent from each other and following a memoryless polynomial model, in this paper, we propose two data-aided solutions that are different from the idea of [19,20] in the MIMO multipath scenario. Given received signals, we extend the method in [27] to the MIMO multipath scenario; a closed-form solution of the nonlinear coefficients is thus obtained by combining the linear least square (LLS) and singular value decomposition (SVD) methods. An alternative but more effective approach is also presented through solving a nonlinear least square (NLS) problem with independent variables consisting of both channel coefficients and nonlinear coefficients of the PA model. Furthermore, we explicitly provide deep discussion on the parameter estimate bias for the general case of unknown order of the PA model. In particular, it is proved in this paper that the rank-deficient property of both the NLS and LLS problems are in association with the amplitude level of the training sequence. The average misclassification rate based on the minimum error probability (MEP) criterion is theoretically derived, and we finally verify the proposed approaches via extensive numerical simulations.

The rest of the paper is organized as follows. In Section 2, we introduce the memoryless nonlinear model of PA in the MIMO multipath system. In Section 3, we present the linear and nonlinear frameworks for SEI, respectively. Then, the practical discussions for the proposed approaches and the misclassification rate are separately given in Sections 4 and 5. In Section 6, numerical results are presented to demonstrate the effectiveness. Section 7 summarizes the paper.

## 2. Preliminaries and Problem Formulation

This paper mainly considers the scenario that  $K$  communication emitters equipped with multiple antennas are identified through a multiple-antenna receiver in the MIMO multipath environment. Since constant modulus modulation schemes such as phase-shift keying (PSK) may generally introduce less nonlinear distortion, we thus consider how to extract the underlying PA nonlinearity of each emitter from the received signal, with the help of a quadrature amplitude modulation (QAM) training

sequence. In other words, PA nonlinearity is treated as a unique RFF of the emitter to fulfill the identification task, and we assume QAM is used by emitters when communicating with the receiver.

### 2.1. Memoryless Nonlinear PA Modeling

Generally, the memoryless polynomial model is simple in expression and can describe the intermodulation distortion of an RF PA well, such as PA9440 amplifier [28], in the narrowband communication system. More specifically, the polynomial coefficients are corresponding to the intermodulation coefficient such as third-order interception point ( $IP_3$ ) and fifth-order interception point ( $IP_5$ ). In this work, we choose the memoryless polynomial model to characterize the nonlinear behavior of all PA units in a multi-antenna emitter; then, the relationship between baseband equivalent input and output of the PA [20] at  $j$ th antenna branch can be written as

$$x_j(n) = \sum_{p=1}^{(P+1)/2} \alpha_{2p-1,j} \cdot (s_j(n))^p \cdot (s_j^*(n))^{p-1}, \quad (1)$$

where  $P$  denotes the max order of the PA model and can be configured as  $P = 3, 5, 7, \dots$ . At the RF front-end of emitters, the frequency components that resulted from the even terms in the model can be removed by the bandpass filter, thus the even terms are ignored in Equation (1).  $s_j(n)$  is the baseband equivalent input of the nonlinear system, denoting the  $n$ th QAM symbol transmitted over the  $j$ th antenna. The superscript “\*” is the conjugate operator.  $\alpha_{2p-1,j}$  denotes the normalized  $(2p - 1)$ th order PA coefficient of the  $j$ th antenna and without loss of generality, we set  $\alpha_{1,j} = 1$  hereafter.  $x_j(n)$  stands for the response of the nonlinear system.

In this paper, we suppose that each emitter is equipped with  $J$  antennas. Therefore, the normalized discrete-time baseband equivalent form of the nonlinear distortion model for the multi-antenna system can be expressed as

$$\begin{bmatrix} x_1(n) \\ x_2(n) \\ \vdots \\ x_J(n) \end{bmatrix} = \begin{bmatrix} \bar{s}_1(n) \\ \bar{s}_2(n) \\ \vdots \\ \bar{s}_J(n) \end{bmatrix} + \sum_{p=2}^{(P+1)/2} \begin{bmatrix} \alpha_{2p-1,1} & 0 & \dots & 0 \\ 0 & \alpha_{2p-1,2} & \dots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ 0 & \dots & 0 & \alpha_{2p-1,J} \end{bmatrix} \begin{bmatrix} (\bar{s}_1(n))^p \cdot (\bar{s}_1^*(n))^{p-1} \\ (\bar{s}_2(n))^p \cdot (\bar{s}_2^*(n))^{p-1} \\ \vdots \\ (\bar{s}_J(n))^p \cdot (\bar{s}_J^*(n))^{p-1} \end{bmatrix}, \quad (2)$$

where  $\bar{s}_j(n)$  is the normalized version of  $s_j(n)$ .

### 2.2. MIMO Multipath Channel

The propagation channel considered in this paper is a linear discrete MIMO system with  $J$  transmit antennas and  $R$  receive antennas. Given the length of the training sequence  $N$ , the signal received at the antenna  $r$  can be represented by

$$\mathbf{y}_r = \mathbf{h}_{r1} \otimes \mathbf{x}_1 + \mathbf{h}_{r2} \otimes \mathbf{x}_2 + \dots + \mathbf{h}_{rJ} \otimes \mathbf{x}_J + \mathbf{v}_r, \quad (3)$$

where  $\otimes$  denotes the convolution operation.  $\mathbf{x}_j$  is the nonlinear distortion signal transmitted over the  $j$ th antenna,  $\mathbf{h}_{rj}$  represents the channel impulse response between the  $j$ th transmit antenna to the  $r$ th receive antenna and remains time-invariant during the data receiving process, and  $\mathbf{v}_r$  is the zero-mean additive white Gaussian noise received at the antenna  $r$ . Thus, we can unfold Equation (3) naturally in the form of

$$\mathbf{y}_r = \mathbf{X}^{(1)} \mathbf{h}_{r1} + \mathbf{X}^{(2)} \mathbf{h}_{r2} + \dots + \mathbf{X}^{(J)} \mathbf{h}_{rJ} + \mathbf{v}_r, \quad (4)$$

where the order of the channel is  $L$ ,  $\mathbf{X}^{(j)} \in \mathbb{C}^{(N+L-1) \times L}$  is a Toeplitz matrix populated by  $\mathbf{x}_j$  in Equation (3). Note that the nonlinear coefficients are cross-coupled with channel coefficients in Equation (4); therefore, the major target arising in this paper is to get the accurate estimations of the separate

nonlinear coefficient  $\{\alpha_{2p-1,j}\}$  via the received signal  $\{\mathbf{y}_r\}$  with the aid of training sequences in MIMO multipath scenarios, before we can identify a specific emitter.

### 3. The Proposed Estimation Approaches

As mentioned above, it has been reported in [19,20] that the nonlinear coefficients of PAs can be estimated with two stages. The first stage establishes the initial estimation of channel coefficients and nonlinear coefficients through some well-designed training sequences sorted by amplitude. Then, an iterative method is applied at the second stage to eliminate the estimate bias. In [27], a PA parameter estimator combined the best linear unbiased estimation (BLUE) and singular value decomposition (SVD) is proposed for the SISO system. Since the method in [27] can get a closed-form solution of the PA nonlinear coefficients and has no constraint on the ordering of the amplitude of training symbols, which is more practical compared to the one in [19], we extend it to the MIMO multi-path scenarios and mark it as linear method in MIMO (LMM). Furthermore, we propose an alternative method to extract PA parameters in a nonlinear least square (NLS) manner. It should be noted that both LMM and NLS approaches can be degraded into SISO systems.

#### 3.1. The LMM Approach

Note that, in [27], if the product terms of the channel coefficients and the nonlinear coefficients are substituted by some new integrated variables, we can obtain a set of linear equations with regard to the new variables, thus the product terms can be readily solved by LLS. Then, the only thing left to us is to extract  $\alpha_{2p-1,j}$  from the solution.

According to Equation (4), we first vectorize signals received by all  $R$  antennas into vector  $\mathbf{y}_{vec}$ , i.e.,  $\mathbf{y}_{vec} = [\mathbf{y}_1^T, \mathbf{y}_2^T, \dots, \mathbf{y}_R^T]^T \in \mathbb{C}^{R(N+L-1)}$ , where the superscript "T" is the transpose operator. In addition, the signal at the receiver side now is

$$\mathbf{y}_{vec} = \mathbf{D}_s \mathbf{h}_\alpha + \mathbf{v}_{vec}, \quad (5)$$

where  $\mathbf{v}_{vec}$  is the corresponding reshaped noise vector.  $\mathbf{h}_\alpha \in \mathbb{C}^{RJL(P+1)/2}$  is the integrated vector composed of all independent variables and can be represented as:

$$\mathbf{h}_\alpha = \left[ \mathbf{u}_{1,0}^T, \dots, \mathbf{u}_{1,L-1}^T \mid \dots \mid \mathbf{u}_{R,0}^T, \dots, \mathbf{u}_{R,L-1}^T \right]^T, \quad (6)$$

with

$$\mathbf{u}_{r,l_h} = \mathbf{W}_{r,l_h} \boldsymbol{\alpha}_{vec}, \quad (7)$$

$$\boldsymbol{\alpha}_j = [1, \alpha_{3,j}, \alpha_{5,j}, \dots, \alpha_{p,j}]^T, \quad (8)$$

$$\boldsymbol{\alpha}_{vec} = [\boldsymbol{\alpha}_1^T, \boldsymbol{\alpha}_2^T, \dots, \boldsymbol{\alpha}_J^T]^T, \quad (9)$$

$$\mathbf{W}_{r,l_h} = \text{blkdiag}(h_{r1}(l_h)\mathbf{I}_\alpha, h_{r2}(l_h)\mathbf{I}_\alpha, \dots, h_{rJ}(l_h)\mathbf{I}_\alpha), \quad (10)$$

in which  $l_h = 0, \dots, L-1$ ,  $\mathbf{I}_\alpha \in \mathbb{C}^{(P+1)/2 \times (P+1)/2}$  is a unit matrix and  $\text{blkdiag}(\cdot)$  is the block diagonalization function. In addition,  $\mathbf{D}_s$  is defined by

$$\mathbf{D}_s = \text{blkdiag} \left( \underbrace{\mathbf{d}_s, \mathbf{d}_s, \dots, \mathbf{d}_s}_R \right), \quad (11)$$

with  $\mathbf{d}_s \in \mathbb{C}^{(N+L-1) \times JL(P+1)/2}$  being constructed by the known training sequence. To elaborate the process, we provide a numerical example in Appendix A.

Consequently, the least square (LS) estimation of  $\mathbf{h}_\alpha$  in Equation (5) is

$$\hat{\mathbf{h}}_\alpha = \mathbf{D}_s^\dagger \mathbf{y}_{vec}, \tag{12}$$

where the superscript “+” denotes the pseudo-inverse operation. It is worth noting that, compared to the BLUE method in [27], the least square solution of Equation (12) has no requirement on the estimation of noise power. Obviously, the condition of a unique solution to Equation (12) is natural that  $\mathbf{D}_s$  is full column rank and  $N \geq JL(P + 1)/2 - L + 1$  is satisfied.

Afterwards, we define

$$\mathbf{Q} = [\hat{\mathbf{u}}_{1,0}, \dots, \hat{\mathbf{u}}_{1,L-1} \mid \dots \mid \hat{\mathbf{u}}_{R,0}, \dots, \hat{\mathbf{u}}_{R,L-1}] \tag{13}$$

and further perform SVD on the matrix  $\mathbf{Q} \in \mathbb{C}^{J(P+1)/2 \times RL}$  to get a closed-form estimation of the PA parameters in the MIMO system. Since the normalized first-order nonlinear coefficients are assumed equal to 1, the execution steps of PA nonlinear coefficients estimator can be summarized as:

- (1) Reshape the observations according to Equation (5), and estimate  $\mathbf{h}_\alpha$  according to Equation (12).
- (2) Reshape the  $\hat{\mathbf{h}}_\alpha$  into the matrix  $\mathbf{Q}$ , then perform SVD on  $\mathbf{Q}_j = \mathbf{U}_j \Sigma_j \mathbf{V}_j^H$ , where  $\mathbf{Q}_j$  is a submatrix consisting of the  $((j - 1)(P + 1)/2 + 1)$ th to  $(j(P + 1)/2)$ th rows of the matrix  $\mathbf{Q}$ .
- (3) Estimate the nonlinear coefficients of  $j$ th transmitting antenna as follows:

$$\hat{\alpha}_j = \frac{1}{\mathbf{U}_j^{(1,1)}} \mathbf{U}_j^{(:,1)}, \tag{14}$$

where  $\mathbf{U}_j^{(:,1)}$  and  $\mathbf{U}_j^{(1,1)}$  are the first column and first element of  $\mathbf{U}_j$ , respectively.

### 3.2. The NLS Approach

Bearing in mind the received signal in Equation (4) for MIMO multipath transmission system, we can now further expand the expression with a series of nonlinear equations due to the existence of product term of the nonlinear coefficient  $\alpha_{2p-1,j}$  and the channel coefficient  $h_{rj}(l_h)$ . As a consequence, the problem of nonlinear coefficients estimation can thus be alternatively transformed into a NLS optimization one when introducing a training sequence with length  $N$ .

In order to get a more robust solution, we choose to solve a constrained NLS optimization problem with  $q = RJL + J(P - 1)/2$  independent variables, and the cost function can be given by:

$$\min_{\mathbf{z}} G_\gamma(\mathbf{z}) = \min_{\mathbf{z}} \frac{1}{2} \|\mathbf{g}(\mathbf{z})\|_2^2 + \frac{1}{2} \|\gamma \mathbf{z}\|_2^2, \tag{15}$$

where  $\|\cdot\|_2$  denotes 2-norm.  $\mathbf{z} \in \mathbb{C}^q$  is the vector consisting of the independent variables and can be described as  $\mathbf{z} = [\mathbf{z}_h^T, \mathbf{z}_\alpha^T]^T$ , in which  $\mathbf{z}_\alpha$  and  $\mathbf{z}_h$  are respectively as

$$\mathbf{z}_\alpha = [\alpha_{3,1}, \dots, \alpha_{P,1}, \dots, \alpha_{3,J}, \dots, \alpha_{P,J}]^T \tag{16}$$

$$\mathbf{z}_h = [\mathbf{H}_1^T, \mathbf{H}_2^T, \dots, \mathbf{H}_R^T]^T, \tag{17}$$

with

$$\mathbf{H}_r = [h_{r1}(0), \dots, h_{rj}(0), \dots, h_{r1}(L - 1), \dots, h_{rj}(L - 1)]^T. \tag{18}$$

$\|\gamma \mathbf{z}\|_2^2/2$  is the regularization.  $\mathbf{g}(\mathbf{z}) \in \mathbb{C}^{R(N+L-1)}$  denotes the residual function:  $\mathbb{C}^q \rightarrow \mathbb{C}^{R(N+L-1)}$  with  $R(N + L - 1) \geq q$ , that is,

$$\mathbf{g}(\mathbf{z}) = [\Delta \mathbf{y}_1^T, \Delta \mathbf{y}_2^T, \dots, \Delta \mathbf{y}_R^T]^T, \tag{19}$$

with

$$\Delta \mathbf{y}_r = \mathbf{y}_r - \left( \mathbf{X}^{(1)} \mathbf{h}_{r1} + \mathbf{X}^{(2)} \mathbf{h}_{r2} + \dots + \mathbf{X}^{(J)} \mathbf{h}_{rJ} \right). \quad (20)$$

According to [29], the regularized Gauss–Newton iterative method is introduced here to figure out  $\mathbf{z}$ , i.e.,

$$\mathbf{z}_{i+1} = \mathbf{z}_i + \Delta \mathbf{z}_i \quad (21)$$

with

$$\Delta \mathbf{z}_i = - \left( (\mathbf{J}(\mathbf{z}_i))^H \mathbf{J}(\mathbf{z}_i) + \gamma^2 \mathbf{I}_q \right)^{-1} \begin{bmatrix} \mathbf{J}(\mathbf{z}_i) \\ \gamma \mathbf{I}_q \end{bmatrix}^H \begin{bmatrix} \mathbf{g}(\mathbf{z}_i) \\ \gamma \mathbf{z}_i \end{bmatrix}. \quad (22)$$

The superscript “ $H$ ” denotes conjugate transposition, and  $\mathbf{I}_q$  is a  $q \times q$  unit matrix.  $\mathbf{J}(\mathbf{z})$  is the Jacobian matrix of  $\mathbf{g}(\mathbf{z})$ , that is, the first-order derivative of  $\mathbf{g}(\mathbf{z})$  with respect to  $\mathbf{z}$ .

In general, given the Jacobian matrix  $\mathbf{J}(\mathbf{z})$  with full column rank, one can apply the regularized Gauss–Newton method in Equation (21) to optimize problem Equation (15) and can eventually obtain the optimal solution  $\mathbf{z}_{opt}$ . In addition, the appropriate regularization factor  $\gamma$  can improve the condition number of the inverse matrix in Equation (22), which guarantees the robustness of the NLS optimization.

#### 4. Practical Discussions on the Proposed Approaches

As mentioned in Section 3, we can get a closed-form solution for the problem of nonlinear coefficient estimation through the LMM approach. Alternatively, an iterative NLS approach can also be applicable to extract PA parameters from observations. However, in practice, there are two main factors that affect the accuracy of our proposed algorithms. The first factor is the case where the order of PA model is mismatched between the transmitting and receiving ends. The second one is the case where the matrix  $\mathbf{D}_s$  and the Jacobian matrix  $\mathbf{J}(\mathbf{z})$  are rank-deficient. In the next Sections 4.1 and 4.2, we first theoretically analyze the impact of the mismatched model. After that, we present the rank-deficient condition of the  $\mathbf{D}_s$  and  $\mathbf{J}(\mathbf{z})$  matrix in Section 4.3.

##### 4.1. Overdetermined Order of the PA Model

In this subsection, we consider the impact of the overdetermined order of the PA model on the estimation accuracy of the nonlinear coefficients. Assume that all the PAs of an emitter actually have the same model order as  $P$ . If we obtain an overdetermined order of the PA model beforehand, e.g.,  $P_0$  with ( $P_0 > P$ ), the expected observation of  $\mathbf{y}_{vec} = \mathbf{D}_s^{(1, \dots, P)} \mathbf{h}_\alpha^{(1, \dots, P)} + \mathbf{v}_{vec}$  may be formulated as  $\mathbf{y}_{vec} = \mathbf{D}_s^{(1, \dots, P_0)} \mathbf{h}_\alpha^{(1, \dots, P_0)} + \mathbf{v}_{vec}$  instead. Since  $\mathbf{D}_s^{(1, \dots, P)} \mathbf{h}_\alpha^{(1, \dots, P)}$  equals  $\left( \mathbf{D}_s^{(1, \dots, P_0)} \mathbf{h}_\alpha^{(1, \dots, P_0)} - \mathbf{D}_s^{(P+2, \dots, P_0)} \mathbf{h}_\alpha^{(P+2, \dots, P_0)} \right)$ , the estimation of  $\mathbf{h}_\alpha$  can thus be expressed as follows:

$$\begin{aligned} \hat{\mathbf{h}}_\alpha &= \left( \mathbf{D}_s^{(1, \dots, P_0)} \right)^\dagger \mathbf{y}_{vec} \\ &= \left( \mathbf{D}_s^{(1, \dots, P_0)} \right)^\dagger \left( \mathbf{D}_s^{(1, \dots, P)} \mathbf{h}_\alpha^{(1, \dots, P)} + \mathbf{v}_{vec} \right) \\ &= \mathbf{h}_\alpha^{(1, \dots, P_0)} - \left( \mathbf{D}_s^{(1, \dots, P_0)} \right)^\dagger \left( \mathbf{D}_s^{(P+2, \dots, P_0)} \mathbf{h}_\alpha^{(P+2, \dots, P_0)} \right) + \left( \mathbf{D}_s^{(1, \dots, P_0)} \right)^\dagger \mathbf{v}_{vec}, \end{aligned} \quad (23)$$

where  $\mathbf{D}_s^{(a, \dots, b)}$  denotes a matrix  $\mathbf{D}_s$  that is populated by the  $a$ th to the  $b$ th order versions of training sequence, and  $\mathbf{h}_\alpha^{(a, \dots, b)}$  denotes a vector  $\mathbf{h}_\alpha$  that is populated by the corresponding  $a$ th to the  $b$ th order nonlinear coefficients. Some numerical examples for the construction process are also provided in Appendix A.

Note that the  $(P+2)$ th to the  $P_0$ th order the nonlinear coefficients and  $\mathbf{h}_\alpha^{(P+2, \dots, P_0)}$  should be all zeros in this case, then, we have  $\hat{\mathbf{h}}_\alpha = \mathbf{h}_\alpha^{(1, \dots, P_0)} + \left( \mathbf{D}_s^{(1, \dots, P_0)} \right)^\dagger \mathbf{v}_{vec}$ , which is equivalent to the unbiased estimation of the  $\mathbf{h}_\alpha^{(1, \dots, P)}$ , and the additive white noise has no effect on the unbiasedness of

the estimation results. Therefore, we can conclude that, if the order of the PA model is overdetermined, the estimation of the nonlinear coefficients obtained by the LMM approach is still unbiased.

#### 4.2. Underdetermined Order of the PA Model

In the sequel, we attempt to analyze the case that the order of the PA model is underdetermined, i.e.,  $P_0$  is lower than the actual  $P$ . It is easy to understand that  $\mathbf{D}_s^{(1,\dots,P)} \mathbf{h}_\alpha^{(1,\dots,P)}$  now equals  $(\mathbf{D}_s^{(1,\dots,P_0)} \mathbf{h}_\alpha^{(1,\dots,P_0)} + \mathbf{D}_s^{(P_0+2,\dots,P)} \mathbf{h}_\alpha^{(P_0+2,\dots,P)})$ , so that the estimation of  $\mathbf{h}_\alpha$  can be written as:

$$\hat{\mathbf{h}}_\alpha = \mathbf{h}_\alpha^{(1,\dots,P_0)} + (\mathbf{D}_s^{(1,\dots,P_0)})^\dagger (\mathbf{D}_s^{(P_0+2,\dots,P)} \mathbf{h}_\alpha^{(P_0+2,\dots,P)}) + (\mathbf{D}_s^{(1,\dots,P_0)})^\dagger \mathbf{v}_{vec}. \quad (24)$$

Attention must be paid to non-zero  $\mathbf{h}_\alpha^{(P_0+2,\dots,P)}$  in Equation (24) due to the non-zero nonlinear PA coefficients through the  $(P_0 + 2)$ th to the  $P$ th order. The resulting  $\hat{\mathbf{h}}_\alpha$  is a biased estimation w.r.t.  $\mathbf{h}_\alpha^{(1,\dots,P)}$  and the estimate bias equals  $(\mathbf{D}_s^{(1,\dots,P_0)})^\dagger (\mathbf{D}_s^{(P_0+2,\dots,P)} \mathbf{h}_\alpha^{(P_0+2,\dots,P)})$ .

Furthermore, if we define a guess as:

**Guess 1.** If the training sequences are discrete amplitude communication symbols such as PSK and QAM, then

$$(\mathbf{D}_s^{(1,\dots,P_0)})^\dagger \mathbf{D}_s^{(P_0+2,\dots,P)} = \text{blkdiag}(\underbrace{\mathbf{c}, \mathbf{c}, \dots, \mathbf{c}}_{R/L}), \quad (25)$$

where  $\mathbf{c} \in \mathbb{C}^{m \times n}$  is a constant matrix with  $m = \frac{P_0+1}{2}$  and  $n = \frac{P-P_0}{2}$ .

Then, the estimate bias of the nonlinear PA coefficients can be given as the following Proposition 1.

**Proposition 1.** If Guess 1 holds true, the estimate bias of the nonlinear PA coefficients obtained by SVD according to Equation (14) is shown as

$$\left\{ \begin{array}{l} \Delta\alpha_{3,j} = \frac{(c_{21}\alpha_{P_0+2,j} + \dots + c_{2n}\alpha_{P,j}) - (c_{11}\alpha_{P_0+2,j} + \dots + c_{1n}\alpha_{P,j})\alpha_{3,j}}{1 + (c_{11}\alpha_{P_0+2,j} + \dots + c_{1n}\alpha_{P,j})} \\ \vdots \\ \Delta\alpha_{P_0,j} = \frac{(c_{m1}\alpha_{P_0+2,j} + \dots + c_{mn}\alpha_{P,j}) - (c_{m1}\alpha_{P_0+2,j} + \dots + c_{mn}\alpha_{P,j})\alpha_{P_0,j}}{1 + (c_{11}\alpha_{P_0+2,j} + \dots + c_{1n}\alpha_{P,j})} \end{array} \right. \quad (26a)$$

$$\quad (26b)$$

$$\quad (26c)$$

**Proof of Proposition 1.** See Appendix B.  $\square$

Therefore, we can conclude that, when the order of the PA model is underdetermined, the bias term of each nonlinear coefficient is related to training sequence and the higher-order nonlinear coefficients. In addition, the high-order nonlinear coefficients (i.e.,  $\alpha_{P_0+2,j}$  to  $\alpha_{P,j}$ ) are still unknown.

#### 4.3. Rank Deficiency Condition of the Proposed Approaches

As mentioned earlier, it is impossible to determine a unique solution for the LLS problem in Equation (12) when the matrix  $\mathbf{D}_s$  is rank-deficient. In addition, for the NLS problem in Equation (15), it is also hard for the regularized Gauss–Newton method in Equation (21) to find the global optimal solution when the Jacobian matrix  $\mathbf{J}(\mathbf{z})$  is rank-deficient at every  $\mathbf{z}$ . However, since QAM is employed by all emitters, the training sequence has only a limited level of amplitude. We find that the rank attributes of  $\mathbf{D}_s$  and  $\mathbf{J}(\mathbf{z})$  are both associated with the maximum order of the PA model. The specific relationship can be revealed by Proposition 2 as follows.

**Proposition 2.** Given QAM symbols with  $M$  modulus values as the training sequence and the maximum  $P$  order of the PA model, the matrix  $\mathbf{D}_s$  is rank-deficient and the Jacobian matrix  $\mathbf{J}(\mathbf{z})$  is also rank-deficient at every  $\mathbf{z}$ , if  $M < (P + 1)/2$ .

**Proof of Proposition 2.** See Appendix C.  $\square$

According to the Proposition 2, the  $\mathbf{D}_s$  and  $\mathbf{J}(\mathbf{z})$  are full rank as long as the amplitude type of signal is sufficient. Therefore, our proposed approach can be applied in some single-carrier communication systems with higher order QAM modulation such as 16-QAM, 64-QAM and 256-QAM, which have 3, 9, and 32 different modulus values, respectively. Actually, it appears that our proposed approaches are readily suitable for other popular wireless communication systems, such as MIMO-OFDM systems, where there will be no rank-deficient problem of  $\mathbf{D}_s$  and  $\mathbf{J}(\mathbf{z})$  in nature since the amplitude of the transmitted signal is generally continuous. Moreover, it is not necessary to estimate the channel order in a MIMO-OFDM system due to its ability to resist multipath effects, which will make the proposed approach more practical.

### 5. Error Rate Analysis For Classification

In this paper, we apply a minimum error probability (MEP) criterion [30] based on Bayesian theory to classify different emitters, and the RFF feature of each emitter is composed of the estimated nonlinear coefficients and can be expressed as follows:

$$\mathbf{a} = [\mathbf{a}_3^T, \mathbf{a}_5^T, \dots, \mathbf{a}_P^T]^T, \tag{27}$$

with

$$\mathbf{a}_p = [Re(\hat{\alpha}_{p,1}), \dots, Re(\hat{\alpha}_{p,J}), Im(\hat{\alpha}_{p,1}), \dots, Im(\hat{\alpha}_{p,J})]^T, \tag{28}$$

where  $p = 3, 5, \dots, P$ .

Here, we take the case where there are two emitters as an example to give the derivation of the decision criteria, and the binary hypothesis test model can be considered as

$$\begin{cases} C_1 : \mathbf{a} = \mathbf{m}_1 + \mathbf{v}_\alpha, & (29a) \\ C_2 : \mathbf{a} = \mathbf{m}_2 + \mathbf{v}_\alpha, & (29b) \end{cases}$$

where  $C_i$  denotes the category  $i, i = 1, 2$ ;  $\mathbf{v}_\alpha$  is composed of the residual additive Gaussian noise in the estimations of the nonlinear coefficients and we assume that each element of  $\mathbf{v}_\alpha$  obeys a Gaussian distribution with a zero mean and a variance of  $\delta^2$ ;  $\mathbf{m}_1 \in \mathbb{C}^{J(P-1)}$  and  $\mathbf{m}_2 \in \mathbb{C}^{J(P-1)}$  are respectively the mean vectors of the estimated feature vector  $\mathbf{a}$  for the two emitters, which can be obtained from the samples collected offline. Thus, the decision rule can be derived based on MEP criterion as follows:

$$\begin{cases} (\mathbf{m}_2 - \mathbf{m}_1)^T \mathbf{a} > \frac{\mathbf{m}_2^T \mathbf{m}_2 - \mathbf{m}_1^T \mathbf{m}_1}{2}, \mathbf{a} \in C_2, & (30a) \\ (\mathbf{m}_2 - \mathbf{m}_1)^T \mathbf{a} < \frac{\mathbf{m}_2^T \mathbf{m}_2 - \mathbf{m}_1^T \mathbf{m}_1}{2}, \mathbf{a} \in C_1, & (30b) \end{cases}$$

where the test statistic is  $a_{ts} = (\mathbf{m}_2 - \mathbf{m}_1)^T \mathbf{a}$  and the decision threshold is  $thr = (\mathbf{m}_2^T \mathbf{m}_2 - \mathbf{m}_1^T \mathbf{m}_1)/2$ .

For simplicity, if we assume that the variables in  $\mathbf{a}$  are independent of each other, then the test statistic  $a_{ts} \sim \mathcal{N}(m_{ts}, \delta_{ts}^2)$ , for the  $i$ -th emitter, the mean is as  $m_{ts}^{(i)} = (\mathbf{m}_2 - \mathbf{m}_1)^T \mathbf{m}_i, i = 1, 2$ , and the variance is as  $\delta_{ts}^2 = \sum_{k=1}^{J(P-1)} \delta^2 \beta_k^2$  with  $\beta_k$  being the  $k$ -th element of  $(\mathbf{m}_2 - \mathbf{m}_1)$ . As a result, with the

assumption of equally probable hypotheses, the average misclassification rate based on MEP criterion can be derived as

$$\begin{aligned}\bar{P}_e &= \frac{1}{2} [P_r(a_{ts} > thr | C_1) + P_r(a_{ts} < thr | C_2)] \\ &= \Phi\left(-\frac{(\mathbf{m}_2 - \mathbf{m}_1)^T(\mathbf{m}_2 - \mathbf{m}_1)}{2\delta_{ts}}\right) \\ &= \Phi\left(-\frac{\sqrt{(\mathbf{m}_2 - \mathbf{m}_1)^T(\mathbf{m}_2 - \mathbf{m}_1)}}{2\delta}\right),\end{aligned}\quad (31)$$

where  $\Phi(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx$  and  $\Phi(-x) = 1 - \Phi(x)$ . When  $x$  increases,  $\Phi(-x)$  decreases.

As for the case where the estimations of the nonlinear coefficients are unbiased, the mean values of the RFF feature for the two emitters are readily as

$$\mathbf{m}_i = [\mathbf{m}_i^{(3)T}, \mathbf{m}_i^{(5)T}, \dots, \mathbf{m}_i^{(P)T}]^T \quad (32)$$

with

$$\mathbf{m}_i^{(p)} = [Re(\alpha_{p,1}^{(i)}), \dots, Re(\alpha_{p,J}^{(i)}), Im(\alpha_{p,1}^{(i)}), \dots, Im(\alpha_{p,J}^{(i)})]^T, \quad (33)$$

where  $\alpha_{p,j}^{(i)}$  is the nonlinear coefficient of the  $i$ -th emitter. However, for the case of biased estimations,  $\tilde{\mathbf{m}}_i = \mathbf{m}_i + \Delta\mathbf{m}_i$ , where

$$\Delta\mathbf{m}_i = [\Delta\mathbf{m}_i^{(3)T}, \Delta\mathbf{m}_i^{(5)T}, \dots, \Delta\mathbf{m}_i^{(P)T}]^T \quad (34)$$

with

$$\Delta\mathbf{m}_i^{(p)} = [Re([\Delta\alpha_{p,1}^{(i)}, \dots, \Delta\alpha_{p,J}^{(i)}]), Im([\Delta\alpha_{p,1}^{(i)}, \dots, \Delta\alpha_{p,J}^{(i)}])]^T \quad (35)$$

being the mean of the biases with  $i = 1, 2$  in the estimation of RF fingerprint features. Therefore, according to Equation (31), the average misclassification rate for this case can be obtained by:

$$\begin{aligned}\bar{P}_e &= \Phi\left(-\frac{(\tilde{\mathbf{m}}_2 - \tilde{\mathbf{m}}_1)^T(\tilde{\mathbf{m}}_2 - \tilde{\mathbf{m}}_1)}{2\sqrt{\sum_{k=1}^{J(P-1)} \delta^2 \tilde{\beta}_k^2}}\right) \\ &= \Phi\left(-\frac{\sqrt{\theta_1 + \theta_2 + \theta_3}}{2\delta}\right),\end{aligned}\quad (36)$$

where  $\Delta\beta_k$  is the  $k$ -th element of  $(\Delta\mathbf{m}_2 - \Delta\mathbf{m}_1)$  and

$$\left\{ \begin{aligned} \theta_1 &= (\mathbf{m}_2 - \mathbf{m}_1)^T(\mathbf{m}_2 - \mathbf{m}_1) & (37a) \\ \theta_2 &= 2(\mathbf{m}_2 - \mathbf{m}_1)^T(\Delta\mathbf{m}_2 - \Delta\mathbf{m}_1) & (37b) \\ \theta_3 &= (\Delta\mathbf{m}_2 - \Delta\mathbf{m}_1)^T(\Delta\mathbf{m}_2 - \Delta\mathbf{m}_1). & (37c) \end{aligned} \right.$$

Therefore, we can conclude as follows:

- (1) The  $\bar{P}_e$  decreases as the variance of the additive noise decrease.
- (2) The  $\bar{P}_e$  decreases as the difference between the mean values of the RF fingerprint feature for the two emitters increases, which indicates that the PA parameters of each emitter should be designed to be as different as possible in a bid to achieve better performance.
- (3) According to Equation (31), as for a fixed  $P$ , more nonlinear coefficients are used as features will make the  $\sqrt{(\mathbf{m}_2 - \mathbf{m}_1)^T(\mathbf{m}_2 - \mathbf{m}_1)}$  larger, which obviously leads to better classification performance.
- (4) According to Equation (36), when there are biases in the estimations, the additional terms  $\theta_2$  and  $\theta_3$  may cause  $\bar{P}_e$  to decrease compared to the case where there is no bias.

More generally, as for  $K$  ( $K > 2$ ) emitters, with the assumption of equally probable hypotheses, the average misclassification rate based on MEP criterion can be represented as

$$\bar{P}_e = 1 - \bar{P}_c = 1 - \frac{1}{K} [P_r(D_1 | C_1) + P_r(D_2 | C_2) + \dots + P_r(D_K | C_K)], \quad (38)$$

where  $\bar{P}_c$  is the average correct classification rate;  $D_k(k = 1, 2, \dots, K)$  indicates the discriminant domain of the  $k$ th class and it shows as Equation (39)

$$\left\{ \begin{array}{l} (\mathbf{m}_k - \mathbf{m}_1)^T \mathbf{a} > (\mathbf{m}_k^T \mathbf{m}_k - \mathbf{m}_1^T \mathbf{m}_1) / 2 \\ \vdots \\ (\mathbf{m}_k - \mathbf{m}_{k-1})^T \mathbf{a} > (\mathbf{m}_k^T \mathbf{m}_k - \mathbf{m}_{k-1}^T \mathbf{m}_{k-1}) / 2 \\ (\mathbf{m}_k - \mathbf{m}_{k+1})^T \mathbf{a} > (\mathbf{m}_k^T \mathbf{m}_k - \mathbf{m}_{k+1}^T \mathbf{m}_{k+1}) / 2 \\ \vdots \\ (\mathbf{m}_k - \mathbf{m}_K)^T \mathbf{a} > (\mathbf{m}_k^T \mathbf{m}_k - \mathbf{m}_K^T \mathbf{m}_K) / 2. \end{array} \right. \quad (39)$$

Therefore, the  $P_r(D_k | C_k)$  can be obtained by solving the multiple integrals in Equation (40)

$$P_r(D_k | C_k) = \int_{\frac{\mathbf{m}_k^T \mathbf{m}_k - \mathbf{m}_K^T \mathbf{m}_K}{2}}^{+\infty} \dots \int_{\frac{\mathbf{m}_k^T \mathbf{m}_k - \mathbf{m}_1^T \mathbf{m}_1}{2}}^{+\infty} f(w_1) \dots f(w_{k-1}) f(w_{k+1}) \dots f(w_K) d_{w_1} \dots d_{w_{k-1}} d_{w_{k+1}} \dots d_{w_K}, \quad (40)$$

where  $f(w_i) = \frac{1}{\sqrt{2\pi}\delta_{w_i}} \exp(-\frac{(w_i - (\mathbf{m}_k - \mathbf{m}_i)^T \mathbf{a})^2}{2\delta_{w_i}^2})$  with  $w_i = (\mathbf{m}_k - \mathbf{m}_i)^T \mathbf{a}$  ( $i = 1, 2, \dots, K$  and  $i \neq k$ ), and  $\delta_{w_i}$  is a standard deviation similar to the  $\delta_{ts}$  in Equation (31).

## 6. Numerical Results

### 6.1. Simulation Setting

In the following simulations, if not specified, we use a  $2 \times 2$  MIMO channel with Rayleigh multipath fading model, and the number of paths  $L$  is set to 2. The training sequences are QAM symbols with  $N = 35$  for proper complexity and performance trade-off. The regularization factor  $\gamma$  is set to 0.05 for the trade-off between the variance and bias in the estimation of NLS approach. The complex nonlinear coefficients lead to AM-AM and AM-PM distortion, where the AM-AM distortion is mainly caused by the  $IP_3$  and  $IP_5$ , and higher order intermodulation distortion is usually ignored [28,31]. Generally, the actual communication system has specific requirements for the out-of-band spectral emission level of RF signals. Therefore, for the rationality of PA parameters in reality, we select the PA parameters that AM-AM characteristic obeys the method in [28,31] and the AM-PM is obtained by slightly adjusting the phase of the parameter in [20]. The nonlinear PA coefficients are displayed in Table 1, and it has been verified that the out-of-band emission level of the amplified signals is about 40 dB<sub>r</sub>, which meets the requirement of general agreement.

**Table 1.** Nonlinear coefficients of the separate PA for each RF chain.

	$\alpha_{3,1}$	$\alpha_{3,2}$	$\alpha_{5,1}$	$\alpha_{5,2}$
Emitter1	0.0210 - 0.0365i	0.0290 - 0.0503i	-0.0099 - 0.0011i	-0.0158 - 0.0018i
Emitter2	0.0182 + 0.0028i	0.0131 + 0.0020i	0.0014 - 0.0008i	0.0009 - 0.0005i
Emitter3	0.0780 - 0.0488i	0.0566 - 0.0354i	-0.0111 - 0.0384i	-0.0070 - 0.0242i
Emitter4	0.0270 - 0.0247i	0.0196 - 0.0179i	0.0001 - 0.0063i	0.0000 - 0.0040i
Emitter5	0.0241 + 0.0393i	0.0174 + 0.0285i	0.0023 - 0.0097i	0.0014 - 0.0061i
Emitter6	0.0514 - 0.0520i	0.0372 - 0.0377i	0.0112 - 0.0226i	0.0070 - 0.0143i
Emitter7	0.0129 - 0.0167i	0.0102 - 0.0132i	-0.0006 - 0.0024i	-0.0004 - 0.0015i
Emitter8	0.0246 + 0.0226i	0.0195 + 0.0179i	0.0011 - 0.0062i	0.0007 - 0.0039i
Emitter9	0.0335 + 0.0410i	0.0266 + 0.0326i	0.0081 - 0.0137i	0.0051 - 0.0086i
Emitter10	0.0256 - 0.0069i	0.0204 - 0.0055i	0.0022 - 0.0033i	0.0014 - 0.0021i

In all experiments, we use the Normalized Mean Squared Error (NMSE) to evaluate the estimation accuracy of the nonlinear coefficients, i.e.,

$$err_{\alpha} = 10 \log_{10} \left( E \left\{ \frac{\|\alpha_{est} - \alpha_{true}\|_2^2}{\|\alpha_{true}\|_2^2} \right\} \right), \quad (41)$$

where  $\alpha_{true} = [\alpha_{3,1}, \alpha_{3,2}, \alpha_{5,1}, \alpha_{5,2}, \dots, \alpha_{P,1}, \alpha_{P,2}]^T$ ,  $\alpha_{est}$  is the estimation of the  $\alpha_{true}$ , and  $E\{\cdot\}$  denotes the Mathematical Expectation.

## 6.2. Simulation Results

### 6.2.1. The Impact of SNR

Note that Liu's algorithm in [19] assumes all PAs in one emitter are the same. To facilitate the comparison of the proposed approaches with Liu's algorithm, we extend Liu's algorithm to be suitable for our MIMO scenario. For brevity, we use the name Modified Liu Algorithm (MLA) to represent the modified version in the following. Here, we set  $P$  to 5, and the training sequence is 16-QAM symbols. At first, we give comparisons among the MLA, LMM, and NLS in Figures 1 and 2, where the third-order and fifth-order coefficients are combined as a classification feature, from which one can see that the performance of the LMM and NLS are apparently better than that of MLA, especially at low signal to noise ratio (SNR) regime. One possible reason lies in that when  $P$  equals 5 and the training sequence is composed of 16-QAM symbols, the  $D_s$  and  $J(\mathbf{z})$  are both full column rank, whereas the matrix populated by the training sequence for the initial estimation of MLA method is rank-deficient regardless of how long the training sequence is. In addition, we notice that the NLS approach also performs better than the LMM one, which is because the introduction of regularization during the process of optimization. Furthermore, in order to explore the impact of iterations on performance of NLS approach, when SNR is 25 dB, we give some results in Figures 3 and 4 to demonstrate the convergence speed of NLS approach, where 19 iterations are needed to obtain the optimal solution.

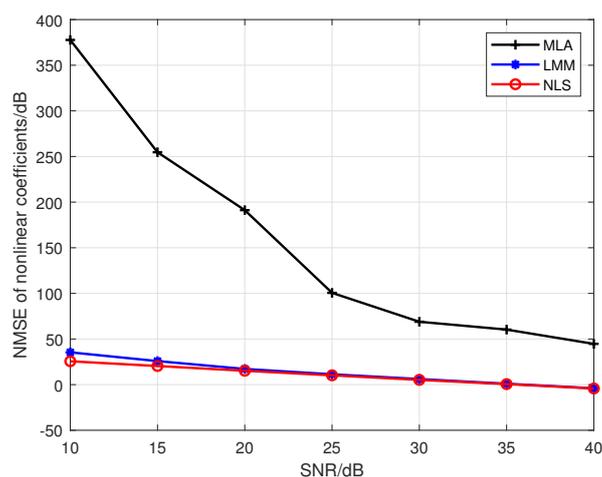


Figure 1. The comparison of the estimation accuracy among the MLA, LMM, and NLS.

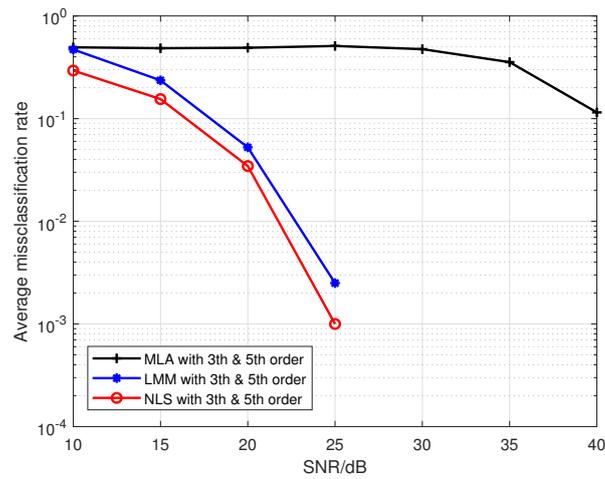


Figure 2. The comparison of the average misclassification rate among the MLA, LMM, and NLS.

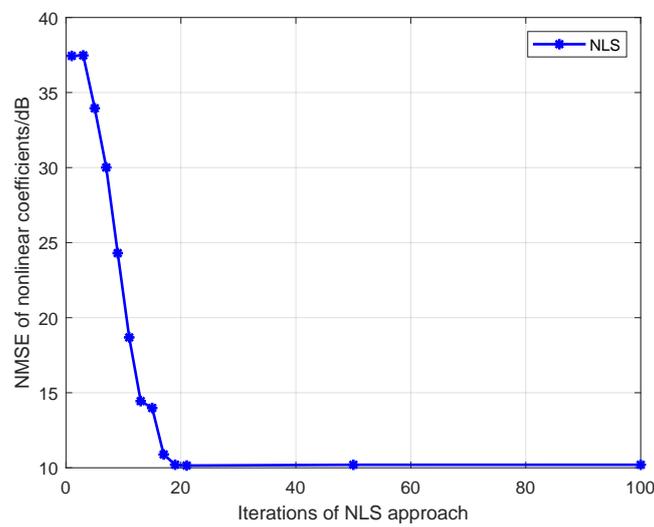


Figure 3. The impact of iterations on the estimation accuracy for the NLS approach.

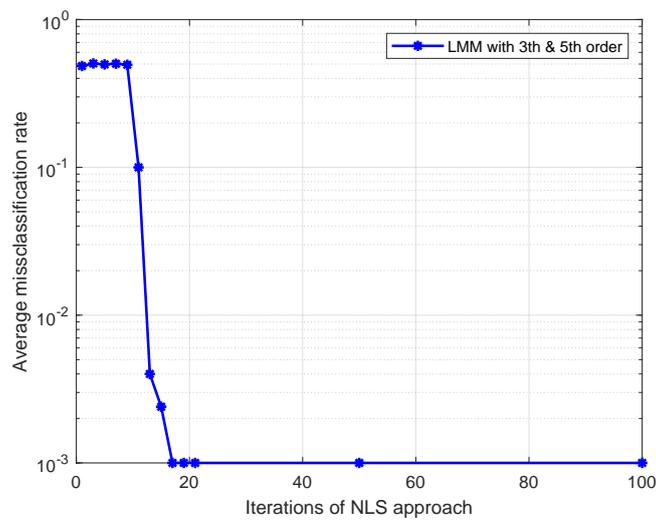
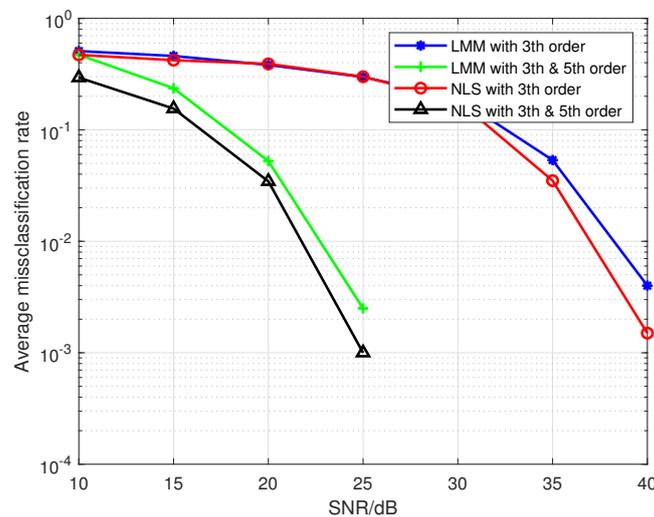


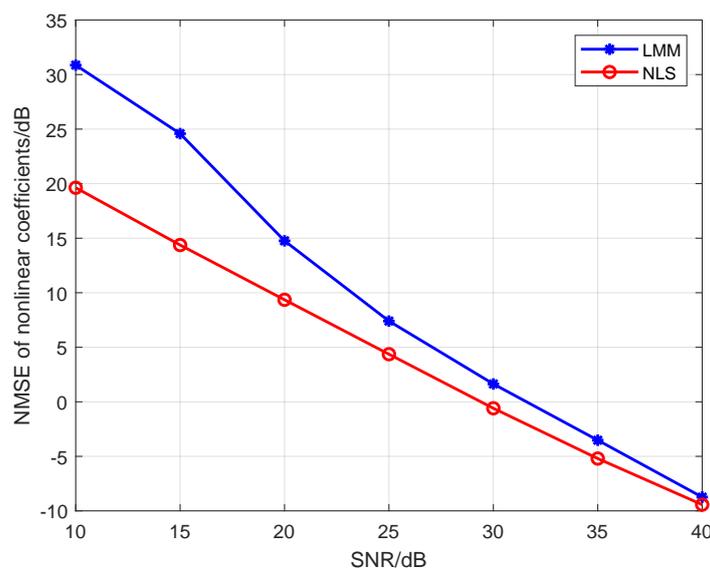
Figure 4. The impact of iterations on the average misclassification rate for the NLS approach.

In Figure 5, we fit a PA with  $P = 5$ , when third-order and fifth-order coefficients are combined as a classification feature, the classification performance is better than when only the third-order coefficient is used. This confirms the analysis that, for a fixed  $P$ , features with more nonlinear coefficients can effectively improve classification performance in Section 5. In the next simulations, if not specified, we use the estimated third-order and fifth-order coefficients together as classification features.

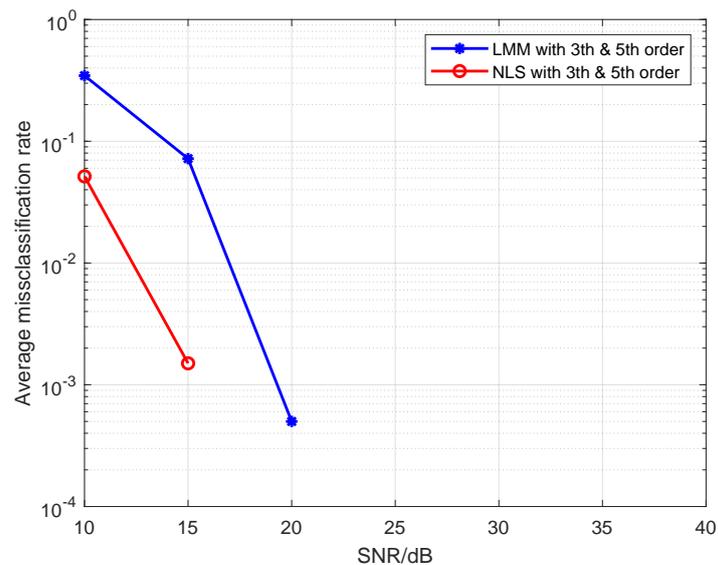


**Figure 5.** The impact of SNR on the average misclassification rate for two emitters.

Moreover, in Figures 6 and 7, we present results of the proposed approaches under a  $4 \times 4$  MIMO scenario; as you can see, more RF channels bring more nonlinear coefficients, which is beneficial to improve classification performance. Here, for convenience, we take the nonlinear coefficients of Emitter2 and Emitter4 in Table 1 as those of the first 4-antenna emitter, and take the nonlinear coefficients of Emitter3 and Emitter6 in Table 1 as those of the second 4-antenna emitter.



**Figure 6.** The impact of SNR on the estimation accuracy under a  $4 \times 4$  MIMO scenario.

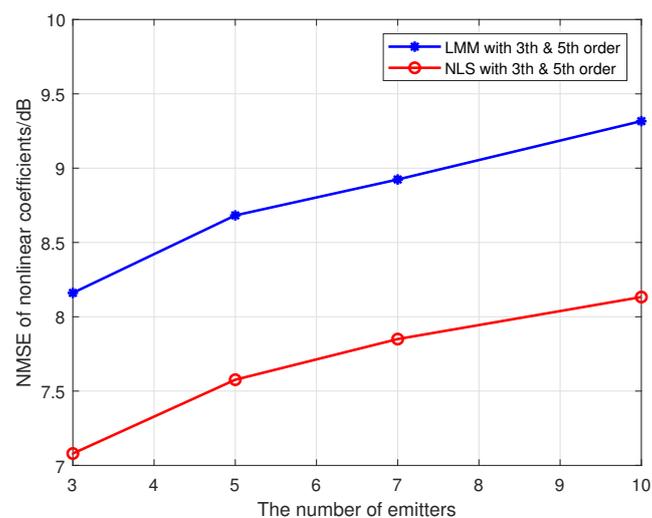


**Figure 7.** The impact of SNR on the average misclassification rate under a  $4 \times 4$  MIMO scenario.

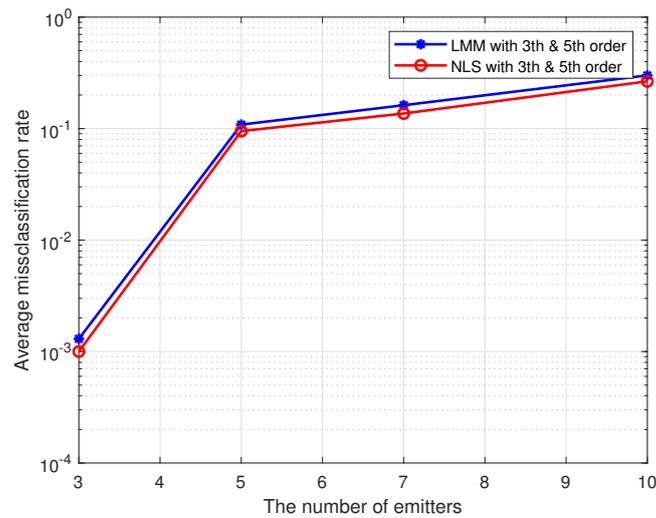
### 6.2.2. Identification for Multiple Emitters

In this simulation, we test the performance of proposed methods under the multi-user case. Here, we set the maximum order of PAs in individual emitters to 5, and 16-QAM symbols are used as the training sequence. The detailed nonlinear coefficients of each emitter are displayed in Table 1. The results are shown in Figures 8 and 9, where the SNR is set to 25 dB, and each method is performed by 1000 Monte Carlo simulations.

As we can see from the Figure 8 and Figure 9, our proposed approaches are also robust and applicable to the multi-user case, and average misclassification rates of them are both increased as the number of emitters increased. Obviously, the greater the difference in nonlinear characteristics between emitters, the higher the resolution of the proposed approaches. In this simulation, as far as the nonlinear coefficients in Table 1 are concerned, the proposed approaches can resolve up to seven emitters at most.



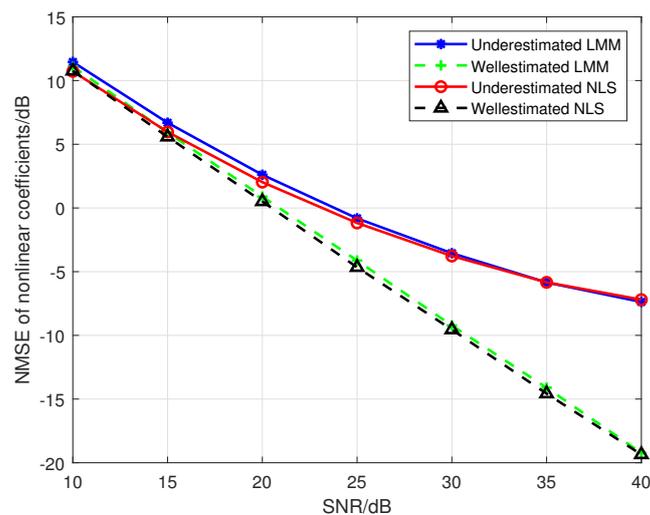
**Figure 8.** The impact of the number of emitters on the estimation accuracy for multiple transmitters.



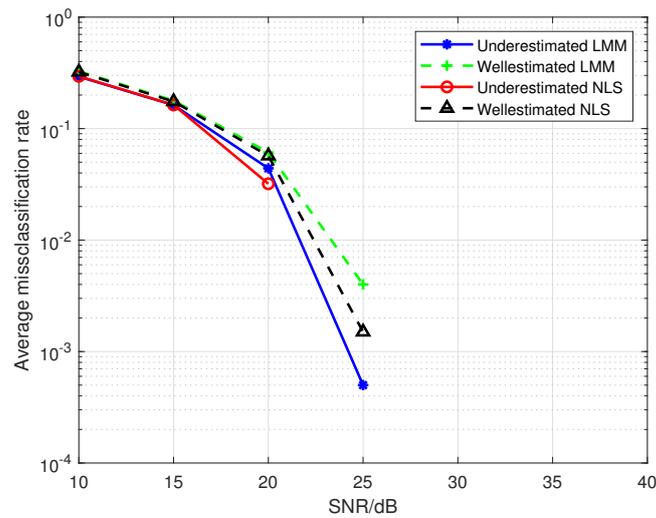
**Figure 9.** The impact of the number of emitters on the average misclassification rate for multiple transmitters.

### 6.2.3. The Confirmation of Practical Discussions

Hereafter, we set the number of emitters to 2 (i.e., emitter1 and emitter2), the maximum order of the PA model  $P$  to 5. The first to fifth order parameters are selected from Table 1. When we use a PA model with  $P_0 = 3$  to fit the nonlinearity of the emitters, the results of Figure 10 confirm the conclusion in the Section 4.2 that the underestimated order of the PA model indeed leads to the biases in the estimations of PA parameters. However, the results in Figure 11 show that the classification performance can be improved by using the biased estimations as features, which reveals that the biases in Equation (36) can make the difference between the classification features of emitters larger. Where the training sequence is 16-QAM symbols, the estimated third-order are used as classification features, and the legends “Well-estimated LMM(NLS)” in Figures 10 and 11 denote that results obtained from a well-determined order of PA model with  $P_0 = P = 3$ .

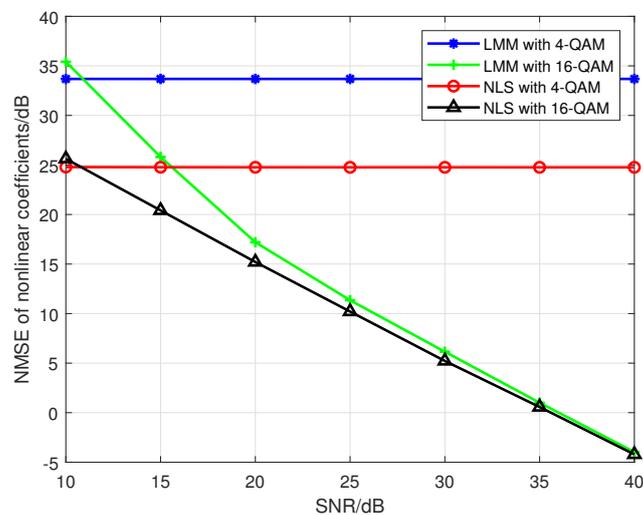


**Figure 10.** The impact of SNR on the estimation accuracy for a mismatched PA model.



**Figure 11.** The impact of SNR on the average misclassification rate for a mismatched PA model.

Then, according to the Proposition 2, when we use a PA model with  $P_0 = P = 5$  to fit the nonlinearity of emitters, if 4-QAM, which is a constant modulus modulation, is adopted by each emitter, then it is clear that the  $\mathbf{D}_s$  and  $\mathbf{J}(\mathbf{z})$  are both rank-deficient, and there is no unique solution in this case. Here, we compare the estimation accuracy of when taking 4-QAM and 16-QAM symbols as training sequence; the results are shown in Figure 12, which obviously corroborates the correctness of Proposition 2.

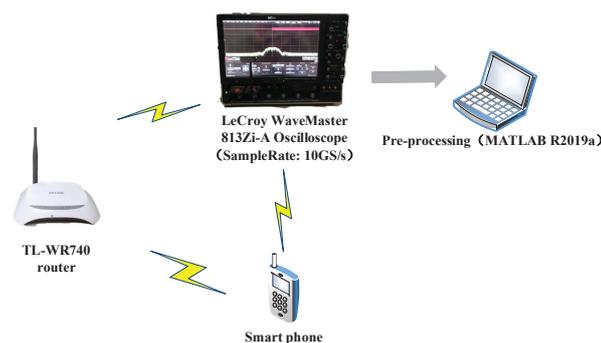


**Figure 12.** The impact of SNR on the estimation accuracy for rank-deficient approaches.

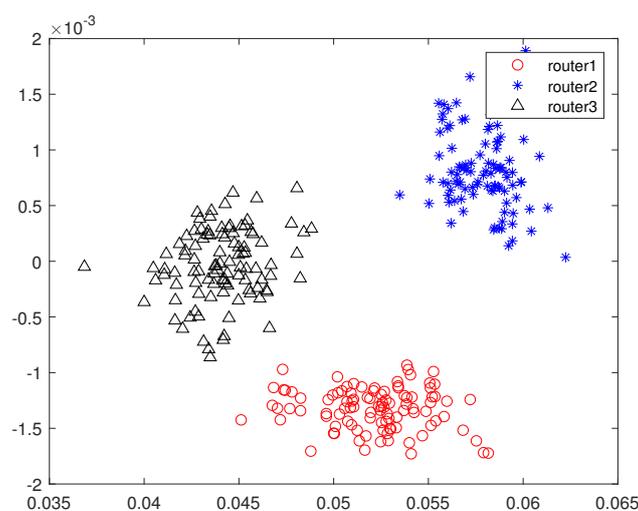
### 6.3. Experimental Results

In this subsection, we design a preliminary verification experiment to explore the effectiveness of our proposed approaches in reality. Due to the limitation of single-channel acquisition, we limit the proposed approaches in the SISO-OFDM system and validate them in a 802.11.g-based wireless local area network (WLAN) that is very common in real life. Therefore, we build an experiment platform, which is shown in Figure 13, to collect the measured router data. In this platform, we first use a LeCroy WaveMaster 813Zi-A Oscilloscope (Chestnut Ridge, NY, USA) equipped with a single antenna to acquire the RF signals from three TL-WR740 routers communicating with a smart phone on the 2.412 GHz channel, respectively. In this experiment, we use the extended NLS approach as an example to

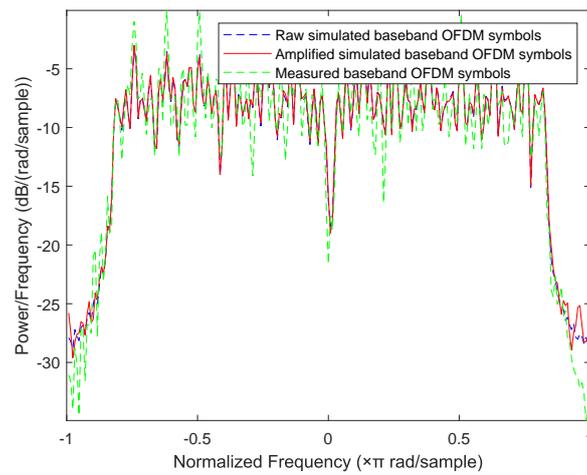
estimate nonlinear coefficients of each router, here we use the “wlan toolbox” in MATLAB R2019a (MathWorks, Natick, Massachusetts, USA) to perform pre-processing such as timing, synchronization, and de-frequency offset on the acquired RF signals. Since the nonlinearity of PA mainly caused by  $IP_3$ , we set  $P = 3$ . Finally, a MEP-based classifier is used to identify the individual router. Note that the memoryless polynomial model may not be able to describe the nonlinearity of the PA in a broadband WLAN system, whereas Table 2 indicates that, according to the estimated PA coefficients in Figure 14, the mean is far greater than the variance for each router, thus the three routers are identifiable based on the extended NLS approach. Moreover, the average misclassification rates of the three routers are all 0. Therefore, we can conclude that the mismatch of PA model does not affect classification performance. In addition, we also compare the power spectral density among measured data and simulated data for three routers in Figures 15–17, respectively, where the legend “Measured baseband OFDM symbols” denotes downconverted acquired RF signals, the legend “Amplified simulated baseband OFDM symbols” denotes simulated baseband OFDM signals amplified by the PA with measured nonlinear coefficients, and the legend “Raw simulated baseband OFDM symbols” denotes raw simulated baseband OFDM signals, all of their power being normalized. In order to explain the results in Figures 15–17, we calculate the NMSE between the PSDs of “Amplified simulated baseband OFDM symbols” and that of “Measured baseband OFDM symbols” for each router, and the NMSEs of three routers are, respectively,  $-7.6472$ ,  $-7.5301$  and  $-7.3589$  dB, which reveal that the PSD of the signal reconstructed by using the estimated PA coefficient can well fit that of the measured signal.



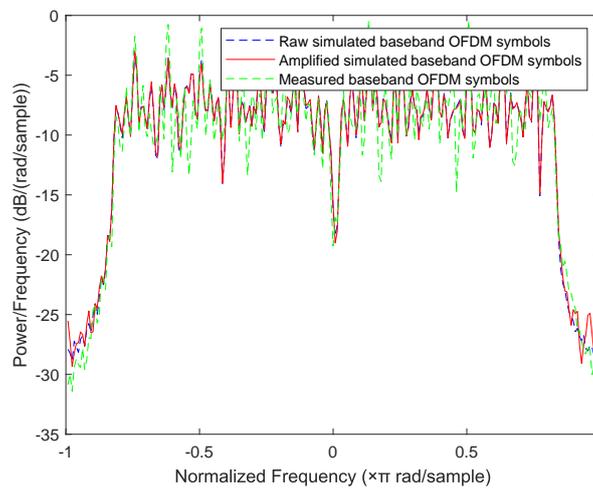
**Figure 13.** The 802.11.g-based experiment platform.



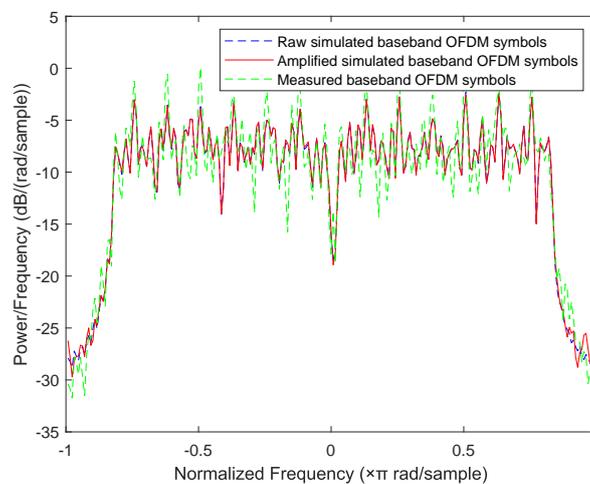
**Figure 14.** The scatter plot of the estimated 3rd order coefficients for three routers.



**Figure 15.** Comparison of power spectral density between measured data and simulated data for router 1.



**Figure 16.** Comparison of power spectral density between measured data and simulated data for router 2.



**Figure 17.** Comparison of power spectral density between measured data and simulated data for router 3.

**Table 2.** The mean and variance of the estimated nonlinear coefficients in Figure 14 for the three routers.

	Router1	Router2	Router3
Mean	0.0522 − 0.0013i	0.0579 + 0.0008i	0.0441 − 0.0000i
Variance	7.1501e <sup>−6</sup>	2.6937e <sup>−6</sup>	3.9458e <sup>−6</sup>

## 7. Conclusions

This paper investigates the SEI scheme for multiple-antenna communication emitters, using PA nonlinearity as RFF features with the assumption that all PAs of a multiple-antenna emitter are independent from each other. Both the LMM and the NLS approaches are proposed to estimate the nonlinear coefficients in association with the memoryless polynomial PA model, where a closed-form solution can be obtained by the LMM approach, and the alternative NLS approach achieves better performance by adopting a regularized Newton–Gauss scheme. Practical discussion on the PA model mismatch is presented, and some theoretical results about the estimate bias and rank-deficient condition are provided to guide the design and implementation of the SEI over MIMO channels. In addition, an error rate analysis is also introduced for the MEP classifier. Simulation results demonstrate that the proposed approaches outperform the other existing schemes, especially in the rank-deficient case, and are effective to deal with SEI in MIMO communication systems. Moreover, the proposed approaches are verified to be effective on a 802.11.g-based experiment platform.

**Author Contributions:** Conceptualization, X.X.; methodology, J.L.; software, J.L.; validation, J.L., X.X.; formal analysis, J.L.; resources, X.X.; data curation, J.L.; writing—original draft preparation, J.L.; writing—review and editing, X.X.; supervision, X.X.; project administration, X.X.; funding acquisition, X.X..

**Funding:** This research was funded in part by ZTE Corporation (Shenzhen, China) (WiFi Interference Source Recognition) and the National Natural Science Foundation of China (No. 61271272).

**Acknowledgments:** The authors would like to thank Yingke Lei from the National University of Defense Technology for his valuable discussion on PA modeling introduced in this paper. The authors also appreciate anonymous reviewers for their helpful comments.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Appendix A

In this appendix, we give a simplified example for the composition of the matrix in the proposed LMM approach. Here, the parameters are assigned as  $R = 2, J = 2, L = 2$  and  $P = 7$ . Therefore, the matrix  $\mathbf{d}_s$  should be as

$$\mathbf{d}_s = \begin{bmatrix} \mathbf{d}_{temp} & \mathbf{0}_{1 \times 8} \\ \mathbf{0}_{1 \times 8} & \mathbf{d}_{temp} \end{bmatrix}, \quad (\text{A1})$$

where  $\mathbf{0}_{1 \times 8}$  is a  $1 \times 8$  zero vector,  $\mathbf{d}_{temp} = [\mathbf{G}_1, \mathbf{G}_2]$ , and  $\mathbf{G}_j \in \mathbb{C}^{N \times (P+1)/2}, j = 1, 2$ , is as follows:

$$\mathbf{G}_j = \begin{bmatrix} \bar{s}_j(0) & \dots & |\bar{s}_j(0)|^{P-1} \bar{s}_j(0) \\ \bar{s}_j(1) & \dots & |\bar{s}_j(1)|^{P-1} \bar{s}_j(1) \\ \vdots & \ddots & \vdots \\ \bar{s}_j(N-1) & \dots & |\bar{s}_j(N-1)|^{P-1} \bar{s}_j(N-1) \end{bmatrix}. \quad (\text{A2})$$

Then,  $\mathbf{D}_s$  is

$$\mathbf{D}_s = \text{blkdiag}(\mathbf{d}_s, \mathbf{d}_s), \quad (\text{A3})$$

which is actually the same as  $\mathbf{D}_s^{(1,3,5,7)}$ .

As for the case where the order of the PA order is overdetermined, note that the parameters are assigned as  $R = 2, J = 2, L = 2$  and  $P = 5$ , whereas  $P_0$  equals 7, then  $\mathbf{D}_s^{(1, \dots, P_0)}$  is equivalent to the matrix  $\mathbf{D}_s^{(1,3,5,7)}$  in Equation (A3).  $\mathbf{D}_s^{(1, \dots, P)}$  and  $\mathbf{D}_s^{(P+2, \dots, P_0)}$  are respectively written as

$$\begin{cases} \mathbf{D}_s^{(1,3,5)} = \text{blkdiag}(\mathbf{d}_s^{(1,3,5)}, \mathbf{d}_s^{(1,3,5)}), & \text{(A4a)} \\ \mathbf{D}_s^{(7)} = \text{blkdiag}(\mathbf{d}_s^{(7)}, \mathbf{d}_s^{(7)}), & \text{(A4b)} \end{cases}$$

where

$$\begin{cases} \mathbf{d}_s^{(1,3,5)} = [\mathbf{e}_{10}^{(1)}, \mathbf{e}_{10}^{(2)}, \mathbf{e}_{01}^{(1)}, \mathbf{e}_{01}^{(2)}], & \text{(A5a)} \\ \mathbf{d}_s^{(7)} = [\mathbf{f}_{10}^{(1)}, \mathbf{f}_{10}^{(2)}, \mathbf{f}_{01}^{(1)}, \mathbf{f}_{01}^{(2)}], & \text{(A5b)} \end{cases}$$

and  $\mathbf{e}_{10}^{(j)}, \mathbf{e}_{01}^{(j)}, \mathbf{f}_{10}^{(j)}$ , and  $\mathbf{f}_{01}^{(j)}$  ( $j = 1, 2$ ) are respectively expressed by:

$$\mathbf{e}_{10}^{(j)} = \begin{bmatrix} \bar{s}_j(0) & \dots & |\bar{s}_j(0)|^4 \bar{s}_j(0) \\ \bar{s}_j(1) & \dots & |\bar{s}_j(1)|^4 \bar{s}_j(1) \\ \vdots & \ddots & \vdots \\ \bar{s}_j(N-1) & \dots & |\bar{s}_j(N-1)|^4 \bar{s}_j(N-1) \\ 0 & \dots & 0 \end{bmatrix}, \quad \text{(A6)}$$

$$\mathbf{e}_{01}^{(j)} = \begin{bmatrix} 0 & \dots & 0 \\ \bar{s}_j(0) & \dots & |\bar{s}_j(0)|^4 \bar{s}_j(0) \\ \bar{s}_j(1) & \dots & |\bar{s}_j(1)|^4 \bar{s}_j(1) \\ \vdots & \ddots & \vdots \\ \bar{s}_j(N-1) & \dots & |\bar{s}_j(N-1)|^4 \bar{s}_j(N-1) \end{bmatrix}, \quad \text{(A7)}$$

$$\mathbf{f}_{10}^{(j)} = \begin{bmatrix} |\bar{s}_j(0)|^6 \bar{s}_j(0) \\ |\bar{s}_j(1)|^6 \bar{s}_j(1) \\ \vdots \\ |\bar{s}_j(N-1)|^6 \bar{s}_j(N-1) \\ 0 \end{bmatrix} \quad \text{(A8)}$$

$$\mathbf{f}_{01}^{(j)} = \begin{bmatrix} 0 \\ |\bar{s}_j(0)|^6 \bar{s}_j(0) \\ |\bar{s}_j(1)|^6 \bar{s}_j(1) \\ \vdots \\ |\bar{s}_j(N-1)|^6 \bar{s}_j(N-1) \end{bmatrix}. \quad \text{(A9)}$$

Similarly, as for the case that the order of the PA model is underdetermined, be aware that the parameters are assigned as  $R = 2, J = 2, L = 2$  and  $P = 7$ , whereas  $P_0$  equals 5, then  $\mathbf{D}_s^{(1, \dots, P)}$  is equivalent to the matrix  $\mathbf{D}_s^{(1,3,5,7)}$  in Equation (A3).  $\mathbf{D}_s^{(1, \dots, P_0)}$  and  $\mathbf{D}_s^{(P_0+2, \dots, P)}$  are equivalent to the matrix  $\mathbf{D}_s^{(1,3,5)}$  and  $\mathbf{D}_s^{(7)}$  in Equation (A4a) and Equation (A4b), respectively.

The examples for  $\mathbf{h}_\alpha^{(1,3,5)}$  and  $\mathbf{h}_\alpha^{(7)}$  are shown as

$$\begin{cases} \mathbf{h}_\alpha^{(1,3,5)} = \left[ \left( \mathbf{w}_{1,0}^{(3)} \boldsymbol{\alpha}_{vec}^{(1,3,5)} \right)^T, \left( \mathbf{w}_{1,1}^{(3)} \boldsymbol{\alpha}_{vec}^{(1,3,5)} \right)^T, \left( \mathbf{w}_{2,0}^{(3)} \boldsymbol{\alpha}_{vec}^{(1,3,5)} \right)^T, \left( \mathbf{w}_{2,1}^{(3)} \boldsymbol{\alpha}_{vec}^{(1,3,5)} \right)^T \right]^T & \text{(A10a)} \end{cases}$$

$$\begin{cases} \mathbf{h}_\alpha^{(7)} = \left[ \left( \mathbf{w}_{1,0}^{(1)} \boldsymbol{\alpha}_{vec}^{(7)} \right)^T, \left( \mathbf{w}_{1,1}^{(1)} \boldsymbol{\alpha}_{vec}^{(7)} \right)^T, \left( \mathbf{w}_{2,0}^{(1)} \boldsymbol{\alpha}_{vec}^{(7)} \right)^T, \left( \mathbf{w}_{2,1}^{(1)} \boldsymbol{\alpha}_{vec}^{(7)} \right)^T \right]^T & \text{(A10b)} \end{cases}$$

and

$$\begin{cases} \mathbf{w}_{r,l_h}^{(1)} = \text{diag}(h_{r1}(l_h), h_{r2}(l_h)) & \text{(A11a)} \end{cases}$$

$$\begin{cases} \mathbf{w}_{r,l_h}^{(3)} = \text{blkdiag}(h_{r1}(l_h)\mathbf{I}_3, h_{r2}(l_h)\mathbf{I}_3) & \text{(A11b)} \end{cases}$$

$$\begin{cases} \boldsymbol{\alpha}_{vec}^{(1,3,5)} = [1, \alpha_{3,1}, \alpha_{5,1}, 1, \alpha_{3,2}, \alpha_{5,2}]^T & \text{(A11c)} \end{cases}$$

$$\begin{cases} \boldsymbol{\alpha}_{vec}^{(7)} = [\alpha_{7,1}, \alpha_{7,2}]^T, & \text{(A11d)} \end{cases}$$

where  $\text{diag}(\cdot)$  denotes a diagonal function,  $\mathbf{I}_3$  is a  $3 \times 3$  unit matrix,  $r = 1, 2$  and  $l_h = 0, 1$ .

## Appendix B

In this appendix, we first give the basis of Guess 1.

**Proof.** At first, we discuss the case where the parameters are set to be  $J = 1, L = 2$  and  $P = 7$ , respectively. As mentioned in Appendix A,  $\mathbf{d}_s^{(1,3,5)}$  and  $\mathbf{d}_s^{(7)}$  can be expressed as block matrices as follows:

$$\begin{cases} \mathbf{d}_s^{(1,3,5)} = [\mathbf{e}_{10}^{(1)}, \mathbf{e}_{01}^{(1)}], & \text{(A12a)} \end{cases}$$

$$\begin{cases} \mathbf{d}_s^{(7)} = [\mathbf{f}_{10}^{(1)}, \mathbf{f}_{01}^{(1)}]. & \text{(A12b)} \end{cases}$$

We thus have

$$\begin{aligned} & \left( \mathbf{d}_s^{(1,3,5)} \right)^\dagger \mathbf{d}_s^{(7)} \\ &= \left( \begin{bmatrix} \left( \mathbf{e}_{10}^{(1)} \right)^H \\ \left( \mathbf{e}_{01}^{(1)} \right)^H \end{bmatrix} \begin{bmatrix} \mathbf{e}_{10}^{(1)} \\ \mathbf{e}_{01}^{(1)} \end{bmatrix} \right)^{-1} \begin{bmatrix} \left( \mathbf{e}_{10}^{(1)} \right)^H \\ \left( \mathbf{e}_{01}^{(1)} \right)^H \end{bmatrix} \begin{bmatrix} \mathbf{f}_{10}^{(1)} \\ \mathbf{f}_{01}^{(1)} \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{A} & \mathbf{V} \\ \mathbf{V}^H & \mathbf{A} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{E} & \mathbf{F}_1 \\ \mathbf{F}_2 & \mathbf{E} \end{bmatrix}, \end{aligned} \quad \text{(A13)}$$

where

$$\begin{cases} \mathbf{A} = \left( \mathbf{e}_{10}^{(1)} \right)^H \mathbf{e}_{10}^{(1)} = \left( \mathbf{e}_{01}^{(1)} \right)^H \mathbf{e}_{01}^{(1)}, & \text{(A14a)} \end{cases}$$

$$\begin{cases} \mathbf{V} = \left( \mathbf{e}_{10}^{(1)} \right)^H \mathbf{e}_{01}^{(1)}, & \text{(A14b)} \end{cases}$$

$$\begin{cases} \mathbf{E} = \left( \mathbf{e}_{10}^{(1)} \right)^H \mathbf{f}_{10}^{(1)} = \left( \mathbf{e}_{01}^{(1)} \right)^H \mathbf{f}_{01}^{(1)}, & \text{(A14c)} \end{cases}$$

$$\begin{cases} \mathbf{F}_1 = \left( \mathbf{e}_{10}^{(1)} \right)^H \mathbf{f}_{01}^{(1)}, & \text{(A14d)} \end{cases}$$

$$\begin{cases} \mathbf{F}_2 = \left( \mathbf{e}_{01}^{(1)} \right)^H \mathbf{f}_{10}^{(1)}, & \text{(A14e)} \end{cases}$$

with  $(\cdot)^{-1}$  being the inverse of a matrix. Then, according to the mathematic formula for the inverse of the block matrix, we get

$$\begin{bmatrix} \mathbf{A} & \mathbf{V} \\ \mathbf{V}^H & \mathbf{A} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{\Pi}_1 & -\mathbf{\Pi}_1 \mathbf{V} \mathbf{A}^{-1} \\ -\mathbf{\Pi}_2 \mathbf{V}^H \mathbf{A}^{-1} & \mathbf{\Pi}_2 \end{bmatrix}, \tag{A15}$$

with

$$\mathbf{\Pi}_1 = (\mathbf{A} - \mathbf{V} \mathbf{A}^{-1} \mathbf{V}^H)^{-1}, \tag{A16}$$

$$\mathbf{\Pi}_2 = (\mathbf{A} - \mathbf{V}^H \mathbf{A}^{-1} \mathbf{V})^{-1}. \tag{A17}$$

Therefore, if we guess that

$$\begin{cases} \mathbf{F}_1 = \mathbf{V} \mathbf{A}^{-1} \mathbf{E}, \\ \mathbf{F}_2 = \mathbf{V}^H \mathbf{A}^{-1} \mathbf{E}, \end{cases} \tag{A18a}$$

$$\tag{A18b}$$

which is actually valid for 16-QAM symbols in practice, then the matrix  $(\mathbf{d}_s^{(1,3,5)})^\dagger \mathbf{d}_s^{(7)}$  is a block diagonal matrix. As a result, for arbitrary  $J$  and  $L$ , we can guess that

$$(\mathbf{D}_s^{(1,\dots,P_0)})^\dagger \mathbf{D}_s^{(P_0+2,\dots,P)} = \text{blkdiag} \left( \underbrace{\mathbf{c}, \dots, \mathbf{c}}_{R/L} \right). \tag{A19}$$

□

Next, we give the proof of Proposition 1.

**Proof.** When the order of the PA model is underdetermined (i.e.,  $P_0 < P$ ), according to Equation (24), ignoring the white noise term, if the Guess 1 holds true, then the  $j$ th submatrix of  $\mathbf{Q}$  in Equation (13) can be expanded into

$$\begin{aligned} \mathbf{Q}_j &= \mathbf{Q}_j^{(1,\dots,P_0)} + \mathbf{Q}_j^{(P_0+2,\dots,P)} \\ &= [h_{1j}(0)(\boldsymbol{\alpha}_j^{(1,\dots,P_0)} + \boldsymbol{\lambda}_j), \dots, h_{R_j}(L-1)(\boldsymbol{\alpha}_j^{(1,\dots,P_0)} + \boldsymbol{\lambda}_j)], \end{aligned} \tag{A20}$$

where  $\mathbf{Q}_j^{(a,\dots,b)}$  and  $\boldsymbol{\alpha}_j^{(a,\dots,b)}$  denote the  $\mathbf{Q}_j$  and  $\boldsymbol{\alpha}_j$  that are populated by the  $a$ th to the  $b$ th order PA parameters, respectively.  $\boldsymbol{\lambda}_j \in \mathbb{C}^m$  is populated by the  $(P_0 + 2)$ th to the  $P$ th order PA parameters and the matrix  $\mathbf{c}$  in Equation (A19), whose  $i$ th ( $i = 1, 2, \dots, m$ ) element is as  $\lambda_j(i) = c_{i1} \alpha_{P_0+2,j} + \dots + c_{im} \alpha_{P,j}$ .

Since the  $\mathbf{Q}_j$  is a matrix with rank 1, the first left singular vector  $\mathbf{U}_j^{(:,1)}$  of  $\mathbf{Q}_j$  obtained by SVD is equivalent to the eigenvector corresponding to the non-zero eigenvalue of the matrix  $\mathbf{Q}_j \mathbf{Q}_j^H$ . Thus, the  $\mathbf{U}_j^{(:,1)}$  of Equation (A20) can be computed by the eigenvalue decomposition (EVD) of  $\mathbf{Q}_j \mathbf{Q}_j^H$ , which is as

$$\mathbf{U}_j^{(:,1)} = [1 + \lambda_j(1), \alpha_{3,j} + \lambda_j(2), \dots, \alpha_{P_0,j} + \lambda_j(m)]^T, \tag{A21}$$

then, according to Equation (14), we have biased  $\hat{\boldsymbol{\alpha}}_j$  as

$$\hat{\boldsymbol{\alpha}}_j = \left[ 1, \frac{\alpha_{3,j} + \lambda_j(2)}{1 + \lambda_j(1)}, \dots, \frac{\alpha_{P_0,j} + \lambda_j(m)}{1 + \lambda_j(1)} \right]^T. \tag{A22}$$

Eventually, we can obtain the biases as shown in Equations (26a)–(26c). □

## Appendix C

In this appendix, we give the proof for Proposition 2.

At first, we prove that, when the training sequence, consisting of QAM symbols, has  $M$  different modulus values and the maximum order of the PA model is  $P$ , if  $M < (P + 1)/2$ , then the matrix  $\mathbf{D}_s$  is rank deficient.

**Proof.** At first, we consider a matrix  $\mathbf{B} \in \mathbb{C}^{l_B \times c_B}$ , which is

$$\mathbf{B} = \begin{bmatrix} 1 & b_1 & b_1^2 & \dots & b_1^{c_B-1} \\ 1 & b_2 & b_2^2 & \dots & b_2^{c_B-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & b_{l_B} & b_{l_B}^2 & \dots & b_{l_B}^{c_B-1} \end{bmatrix}. \quad (\text{A23})$$

When  $l_B = c_B$ ,  $\mathbf{B}$  is actually a Vandermonde matrix and its determinant can be expressed as

$$\det(\mathbf{B}) = \prod_{1 \leq k < i \leq c_B} (b_i - b_k). \quad (\text{A24})$$

Therefore, if the level of the amplitude of  $b_i$  ( $i = 1, \dots, l_B$ ) is less than  $c_B$ , then  $\det(\mathbf{B})$  equals 0 and  $\mathbf{B}$  is definitely singular.

When  $l_B > c_B$ , if the level of the amplitude of  $b_i$  ( $i = 1, \dots, l_B$ ) is less than  $c_B$ , then any  $c_B \times c_B$  sub-matrix of  $\mathbf{B}$  is a singular Vandermonde matrix, that is, any  $c_B \times c_B$  sub-matrix of  $\mathbf{B}$  is column linear correlation, then  $\mathbf{B}$  is also column linear correlation, which means  $\mathbf{B}$  is a column rank-deficient matrix.

Now, we consider the matrix  $\mathbf{G}_j \in \mathbb{N}^{N \times (P+1)/2}$  in Equation (A2), here  $j = 1, \dots, J$ , and it can be rewritten by

$$\begin{aligned} \mathbf{G}_j &= \mathbf{G}_j^{(1)} \mathbf{G}_j^{(2)} \\ &= \mathbf{G}_j^{(1)} \begin{bmatrix} 1 & |\bar{s}_j(0)|^2 & \dots & |\bar{s}_j(0)|^{P-1} \\ 1 & |\bar{s}_j(1)|^2 & \dots & |\bar{s}_j(1)|^{P-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & |\bar{s}_j(N-1)|^2 & \dots & |\bar{s}_j(N-1)|^{P-1} \end{bmatrix}, \end{aligned} \quad (\text{A25})$$

where  $\mathbf{G}_j^{(1)} = \text{diag}(\bar{s}_j(0), \bar{s}_j(1), \dots, \bar{s}_j(N-1))$  is a reversible diagonal matrix. In addition,  $\mathbf{G}_j^{(2)}$  is equivalent to the matrix  $\mathbf{B}$  when we set  $b_i = |\bar{s}_j(i-1)|^2$ ,  $i = 1, 2, \dots, N$ . Since the training sequence is composed of QAM symbols that have only  $M$  different moduli values, then any  $\frac{P+1}{2} \times \frac{P+1}{2}$  sub-matrix of  $\mathbf{G}_j^{(2)}$  is a singular Vandermonde matrix when  $M < (P + 1)/2$ . Thus,  $\mathbf{G}_j$  is column rank-deficient, and, according to Appendix A,  $\mathbf{d}_{temp} = [\mathbf{G}_1, \mathbf{G}_2, \dots, \mathbf{G}_J]$  is also column rank-deficient. Eventually,  $\mathbf{d}_s$  and  $\mathbf{D}_s$  are both column rank-deficient due to the linear correlation among their column vectors. However, when  $M \geq (P + 1)/2$ , as long as  $N$  is large enough, then at least one of  $\frac{P+1}{2} \times \frac{P+1}{2}$  sub-matrix of the  $\mathbf{G}_j^{(2)}$  is a non-singular Vandermonde matrix. Thus, the  $\mathbf{G}_j^{(2)}$  is a column full-rank matrix, and  $\mathbf{D}_s$  is also a column full-rank matrix.  $\square$

Next, we prove that when the training sequence, QAM symbols, has  $M$  different modulus values and the maximum order of the PA model is  $P$ , if  $M < (P + 1)/2$ , then the matrix  $\mathbf{J}(\mathbf{z})$  is rank-deficient at every  $\mathbf{z}$ .

**Proof.** When the number of transmit antennas is  $J$ , then the Jacobian matrix  $\mathbf{J}(\mathbf{z})$  can be computed from  $\mathbf{g}(\mathbf{z})$  in Equation (15) as

$$\mathbf{J}(\mathbf{z}) = \begin{bmatrix} \boldsymbol{\kappa} & \mathbf{0} & \dots & \mathbf{0} & \chi_1 \\ \mathbf{0} & \boldsymbol{\kappa} & \vdots & \vdots & \chi_2 \\ \vdots & \vdots & \ddots & \mathbf{0} & \vdots \\ \mathbf{0} & \dots & \mathbf{0} & \boldsymbol{\kappa} & \chi_R \end{bmatrix} \quad (\text{A26})$$

with

$$\boldsymbol{\kappa} = [\boldsymbol{\mathcal{O}}^{(0)}, \dots, \boldsymbol{\mathcal{O}}^{(L-1)}], \quad (\text{A27})$$

$$\chi_r = [\boldsymbol{\mathcal{W}}_r^{(0)} + \dots + \boldsymbol{\mathcal{W}}_r^{(L-1)}], \quad (\text{A28})$$

where  $\boldsymbol{\mathcal{O}}^{(l_h)}$  and  $\boldsymbol{\mathcal{W}}_r^{(l_h)}$  ( $l_h = 0, 1, \dots, L-1$ ) are respectively shown as

$$\boldsymbol{\mathcal{O}}^{(l_h)} = [\mathbf{O}_1^{(l_h)}, \mathbf{O}_2^{(l_h)}, \dots, \mathbf{O}_J^{(l_h)}], \quad (\text{A29})$$

$$\boldsymbol{\mathcal{W}}_r^{(l_h)} = [\mathbf{W}_{r1}^{(l_h)}, \mathbf{W}_{r2}^{(l_h)}, \dots, \mathbf{W}_{rj}^{(l_h)}]. \quad (\text{A30})$$

Note that  $\mathbf{O}_j^{(l_h)}$  and  $\mathbf{W}_{rj}^{(l_h)}$  ( $l_h = 0, 1, \dots, L-1$ ) are

$$\mathbf{O}_j^{(l_h)} = \boldsymbol{\Lambda}_{l_h} \begin{bmatrix} \underbrace{0, \dots, 0}_{l_h}, \varphi(\bar{s}_j(0)), \dots, \varphi(\bar{s}_j(N-1)), \underbrace{0, \dots, 0}_{L-1-l_h} \end{bmatrix}^T, \quad (\text{A31})$$

$$\mathbf{W}_{rj}^{(l_h)} = h_{rj}(l_h) \boldsymbol{\Lambda}_{l_h} \begin{bmatrix} \mathbf{0}_{l_h \times (P-1)/2} \\ \boldsymbol{\psi}_s \\ \mathbf{0}_{(L-1-l_h) \times (P-1)/2} \end{bmatrix}, \quad (\text{A32})$$

respectively, and

$$\varphi(\bar{s}_j(n)) = 1 + \alpha_{3,j} |\bar{s}_j(n)|^2 + \alpha_{5,j} |\bar{s}_j(n)|^4 + \dots + \alpha_{P,j} |\bar{s}_j(n)|^{P-1}, \quad (\text{A33})$$

$$\boldsymbol{\Lambda}_{l_h} = \text{diag} \left( \underbrace{1, \dots, 1}_{l_h}, \bar{s}_j(0), \dots, \bar{s}_j(N-1), \underbrace{1, \dots, 1}_{L-1-l_h} \right), \quad (\text{A34})$$

$$\boldsymbol{\psi}_s = \begin{bmatrix} |\bar{s}_j(0)|^2 & |\bar{s}_j(0)|^4 & \dots & |\bar{s}_j(0)|^{P-1} \\ |\bar{s}_j(1)|^2 & |\bar{s}_j(1)|^4 & \dots & |\bar{s}_j(1)|^{P-1} \\ \vdots & \vdots & \ddots & \vdots \\ |\bar{s}_j(N-1)|^2 & |\bar{s}_j(N-1)|^4 & \dots & |\bar{s}_j(N-1)|^{P-1} \end{bmatrix}. \quad (\text{A35})$$

If we define a matrix  $\mathcal{J}_r$  as

$$\mathcal{J}_r = [\boldsymbol{\mathcal{O}}^{(0)}, \dots, \boldsymbol{\mathcal{O}}^{(L-1)}, \boldsymbol{\mathcal{W}}_r^{(0)} + \dots + \boldsymbol{\mathcal{W}}_r^{(L-1)}], \quad (\text{A36})$$

then, as for  $\boldsymbol{\Theta}_j$ , a sub-matrix of  $\mathcal{J}_r$ , composed of the elements corresponding to the  $j$ th transmit antenna, which is

$$\boldsymbol{\Theta}_j = [\mathbf{O}_j^{(0)}, \mathbf{O}_j^{(1)}, \dots, \mathbf{O}_j^{(L-1)}, \mathbf{W}_{rj}^{(0)} + \mathbf{W}_{rj}^{(1)} + \dots + \mathbf{W}_{rj}^{(L-1)}]. \quad (\text{A37})$$

Similar to the discussion of the matrix  $\mathbf{B}$  in Equation (A23), as for the matrix  $\mathbf{\Omega}$ ,

$$\mathbf{\Omega} = \begin{bmatrix} \varphi(\bar{s}_j(0)) & h_{1j}(l_h)|\bar{s}_j(0)|^2 & \dots & h_{1j}(l_h)|\bar{s}_j(0)|^{P-1} \\ \varphi(\bar{s}_j(1)) & h_{1j}(l_h)|\bar{s}_j(1)|^2 & \dots & h_{1j}(l_h)|\bar{s}_j(1)|^{P-1} \\ \vdots & \vdots & \ddots & \vdots \\ \varphi(\bar{s}_j(N-1)) & h_{1j}(l_h)|\bar{s}_j(N-1)|^2 & \dots & h_{1j}(l_h)|\bar{s}_j(N-1)|^{P-1} \end{bmatrix} \quad (\text{A38})$$

when the training sequence is composed of QAM symbols that have only  $M$  different moduli values and  $N \geq (P+1)/2$ , if  $M < (P+1)/2$ , then there are at least  $(P+1)/2 - M + 1$  identical row vectors in any  $\frac{P+1}{2} \times \frac{P+1}{2}$  sub-matrix of  $\mathbf{\Omega}$ , which means that any  $\frac{P+1}{2} \times \frac{P+1}{2}$  sub-matrix of  $\mathbf{\Omega}$  is a singular matrix, and the column vectors of it are also considered to be linear correlation. As a result,  $\mathbf{\Omega}$  is also column linear correlation, that is, the  $\mathbf{\Omega}$  is a column rank-deficient matrix.

Since the matrix  $\mathbf{\Omega}$  has been proved to be a column rank-deficient matrix, the matrix  $[\mathbf{O}_j^{(l_h)}, \mathbf{W}_{rj}^{(l_h)}]$  is obviously column rank-deficient, and there must be a set of coefficients  $\{\eta_1, \eta_2, \dots, \eta_{(P+1)/2}\}$  that are not all equal to 0. Consequently, the following formula holds

$$\eta_1 \mathbf{O}_j^{(l_h)} + \sum_{c_W=1}^{(P-1)/2} \eta_{c_W+1} \mathbf{W}_{rj}^{(l_h)}(:, c_W) = \mathbf{0}, \quad (\text{A39})$$

where  $\mathbf{W}_{rj}^{(l_h)}(:, c_W)$  denotes the  $c_W$ -th column of the matrix  $\mathbf{W}_{rj}^{(l_h)}$ .

Therefore, for the matrix  $\mathbf{\Theta}_j$ , there are also a set of coefficients  $\underbrace{\{\eta_1, \eta_1, \dots, \eta_1, \eta_2, \dots, \eta_{(P+1)/2}\}}_L$  that are not all equal to 0 to make the following expression valid, i.e.,

$$\eta_1 \sum_{l_h=0}^{L-1} \mathbf{O}_j^{(l_h)} + \sum_{c_W=1}^{(P-1)/2} \eta_{c_W+1} \mathbf{\Gamma}_{c_W} = \mathbf{0} \quad (\text{A40})$$

with

$$\mathbf{\Gamma}_{c_W} = \sum_{l_h=0}^{L-1} \mathbf{W}_{rj}^{(l_h)}(:, c_W). \quad (\text{A41})$$

Then, we can conclude that the matrix  $\mathbf{\Theta}_j$  is rank-deficient and the matrix  $[\mathbf{\Theta}_1, \mathbf{\Theta}_2, \dots, \mathbf{\Theta}_J]$  is also rank-deficient. Thus, the Jacobian matrix  $\mathcal{J}_r$  is also rank-deficient because  $\mathcal{J}_r$  is the elementary column transformation of the matrix  $[\mathbf{\Theta}_1, \mathbf{\Theta}_2, \dots, \mathbf{\Theta}_J]$  and the elementary column transformation does not change the rank of the matrix. Eventually, the the Jacobian matrix  $\mathbf{J}(\mathbf{z})$  in Equation (A26) is rank-deficient.  $\square$

## References

1. Rehman, S.U.; Sowerby, K.W.; Coghill, C. Radio-frequency fingerprinting for mitigating primary user emulation attack in low-end cognitive radios. *IET Commun.* **2014**, *8*, 1274–1284. [[CrossRef](#)]
2. Suski II, W.C.; Temple, M.A.; Mendenhall, M.J. Radio frequency fingerprinting commercial communication devices to enhance electronic security. *Int. J. Electron. Secur. Digit. Forensics* **2008**, *1*, 301–322.
3. Ureten, O.; Serinken, N. Wireless security through RF fingerprinting. *Can. J. Electr. Comput. Eng.* **2007**, *32*, 27–33. [[CrossRef](#)]
4. Ureten, O.; Serinken, N. Bayesian detection of Wi-Fi transmitter RF fingerprints. *Electron. Lett.* **2005**, *41*, 373–374. [[CrossRef](#)]
5. Choe, H.C.; Poole, C.E.; Andrea, M.Y. Novel identification of intercepted signals from unknown radio transmitters. In Proceedings of the Wavelet Applications II, Orlando, FL, USA, 6 April 1995; Volume 2491, pp. 504–518.

6. Hawkes, K.D. Transient Analysis System for Characterizing RF Transmitters by Analyzing Transmitted RF Signals. U.S. Patent 5,758,277, 26 May 1998.
7. Du, L.; Liu, H.; Bao, Z. Radar HRRP target recognition based on higher order spectra. *IEEE Trans. Signal Process.* **2005**, *53*, 2359–2368.
8. Zhang, X.D.; Shi, Y.; Bao, Z. A new feature vector using selected bispectra for signal classification with application in radar target recognition. *IEEE Trans. Signal Process.* **2001**, *49*, 1875–1885. [[CrossRef](#)]
9. Reising, D.R.; Temple, M.A. WiMAX mobile subscriber verification using Gabor-based RF-DNA fingerprints. In Proceedings of the 2012 IEEE International Conference on Communications (ICC), Ottawa, ON, Canada, 10–15 June 2012; pp. 1005–1010.
10. Reising, D.R.; Temple, M.A.; Oxley, M.E. Gabor-based RF-DNA fingerprinting for classifying 802.16 e WiMAX mobile subscribers. In Proceedings of the 2012 International Conference on Computing, Networking and Communications (ICNC), Maui, HI, USA, 30 January–2 February 2012; pp. 7–13.
11. Klein, R.W.; Temple, M.A.; Mendenhall, M.J. Application of wavelet-based RF fingerprinting to enhance wireless network security. *J. Commun. Netw.* **2009**, *11*, 544–555. [[CrossRef](#)]
12. Klein, R.W.; Temple, M.A.; Mendenhall, M.J. Application of wavelet denoising to improve OFDM-based signal detection and classification. *Secur. Commun. Netw.* **2010**, *3*, 71–82. [[CrossRef](#)]
13. Yuan, Y.; Huang, Z.; Wu, H. Specific emitter identification based on Hilbert-Huang transform-based time-frequency-energy distribution features. *IET Commun.* **2014**, *8*, 2404–2412. [[CrossRef](#)]
14. Xie, Y.; Wang, S.; Zhang, E. Specific Emitter Identification based on nonlinear complexity of signal. In Proceedings of the 2016 IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC), Hong Kong, China, 5–8 August 2016; pp. 1–6.
15. Polak, A.C.; Dolatshahi, S.; Goeckel, D.L. Identifying wireless users via transmitter imperfections. *IEEE J. Sel. Areas Commun.* **2011**, *28*, 1469–1479. [[CrossRef](#)]
16. Brik, V.; Banerjee, S.; Gruteser, M. Wireless device identification with radiometric signatures. In Proceedings of the 14th ACM International Conference on Mobile Computing and Networking, San Francisco, CA, USA, 14–19 September 2008; pp. 116–127.
17. Huang, Y.; Zheng, H. Radio frequency fingerprinting based on the constellation errors. In Proceedings of the 2012 18th Asia-Pacific Conference on Communications (APCC), Jeju Island, Korea, 15–17 October 2012; pp. 900–905.
18. Schubert, B.; Liszewski, J.; Keusgen, W. Identification of the volterra kernels of nonlinear power amplifiers. In Proceedings of the 2009 International Conference on Communications, Circuits and Systems, Milpitas, CA, USA, 23–25 July 2009; pp. 767–771.
19. Liu, M.W.; Doherty, J.F.; Keusgen, W. Wireless device identification in MIMO channels. In Proceedings of the 2009 43rd Annual Conference on Information Sciences and Systems, Baltimore, MD, USA, 18–20 March 2009; pp. 563–567.
20. Liu, M.W.; Doherty, J.F. Nonlinearity estimation for specific emitter identification in multipath channels. *IEEE Trans. Inf. Forensics Secur.* **2011**, *6*, 1076–1085.
21. Li, D.; Jia, Y.; Li, Q. Identification and nonlinear model predictive control of MIMO Hammerstein system with constraints. *J. Cent. South Univ.* **2017**, *24*, 448–458. [[CrossRef](#)]
22. Belkacem, O.B.H.; Ammari, M.L.; Bouallegue, R. Effect of Power Amplifier Nonlinearity on the Physical Layer Security of MIMO Systems. *Wirel. Pers. Commun.* **2017**, *96*, 5587–5601. [[CrossRef](#)]
23. Zou, Y.; Raeesi, O.; Antilla, L. Impact of power amplifier nonlinearities in multi-user massive MIMO downlink. In Proceedings of the 2015 IEEE Globecom Workshops (GC Wkshps), San Diego, CA, USA, 6–10 December 2015; pp. 1–7.
24. Fernandes, C.A.R. Nonlinear MIMO Communication Systems: Channel Estimation and Information Recovery Using Volterra Models. Ph.D. Thesis, Universit de Nice Sophia Antipolis, Nice, France, 2010.
25. Wang, D.; Hu, A.; Chen, Y. An esprit-based approach for rf fingerprint estimation in multi-antenna ofdm systems. *IEEE Wirel. Commun. Lett.* **2017**, *6*, 702–705. [[CrossRef](#)]
26. Wang, D.; Hu, A.; Peng, L. Energy Selected Transmitter RF Fingerprint Estimation in Multi-Antenna OFDM Systems. In Proceedings of the 2018 10th International Conference on Wireless Communications and Signal Processing (WCSP), Hangzhou, China, 18–20 October 2018; pp. 1–6.

27. Cai, K.; Li, H.; Mitola, J. Channel Identification of communication system with nonlinear power amplifier. In Proceedings of the 2013 47th Annual Conference on Information Sciences and Systems (CISS), Baltimore, MD, USA, 20–22 March 2013; pp. 1–5.
28. Wu, Q.; Testa, M.; Larkin, R. On design of linear RF power amplifier for CDMA signals. *Int. J. RF Microw. Comput.-Aided Eng.* **2015**, *8*, 283–292. [[CrossRef](#)]
29. Eriksson, J.; Wedin, P.A.; Gulliksson, M.E. Regularization methods for uniformly rank-deficient nonlinear least-squares problems. *J. Optim. Theory Appl.* **2005**, *127*, 1–26. [[CrossRef](#)]
30. Duda, R.O.; Hart, P.E.; Stork, D.G. *Pattern Classification*; Wiley-Interscience: Hoboken, NJ, USA, 2000.
31. Liu, C.; Xiao, H.; Wu, Q. Spectrum Design of RF Power Amplifier for Wireless Communication Systems. *IEEE Trans. Consum. Electron.* **2002**, *48*, 72–80.



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).