Blind Bilinear Approach for Nonlinear-Based System Identification

Abdulmajid Lawal, Karim Abed-Meraim, Azzedine Zerguine, and Ali Muqaibel

Abstract—In this paper, we develop an efficient nonlinear channel identification method for single input multiple output finite impulse response (FIR) channels. The developed algorithm utilizes the structure embedded in the columns and rows subspaces of the received signal matrix. Both the Toeplitz structure available in the signal matrix and the block Sylvester structure present in the channel matrix is used to develop a criterion that can be minimized to establish the optimal solution of the channel estimates. With nonlinearity in the system, the proposed bilinear nonlinear approach produces some extremely intriguing channel estimation findings.

***Keywords*—Blind identification, nonlinear SIMO channel, bilinear, subspace method.**

Nomenclature

This section includes the abbreviations used in the manuscript:

FIR	Finite Impulse Response.
SIMO	Single Input Multiple Output.
SCS	Structured Channel Subspace
BSCS	Nonlinear Biliear Structured Channel Subspace.
LBSCS	Linear Biliear Structured Channel Subspace.

I. Introduction

Many practical systems have inherent nonlinear behaviors which necessitate the use of dedicated processing especially when such nonlinearities can significantly impact the input signal restoration [1, 2]. The analysis and solutions to nonlinear problems have attracted different specialties such as engineers, mathematicians, and physicists. In particular, the transmission and reception of signals in communication systems involve the use of nonlinear devices such as power amplifiers and optical equipment [3]. Hence, communication channels may be corrupted as a result of nonlinear distortions caused by nonlinear multiple access interference, intersymbol interference, and inter-carrier interference just to mention a few. The signal obtained at the receiving end may deteriorate significantly as a result of these distortions. To tackle such problems, nonlinear models are deployed to accurately represent the channels and enhance the development of dedicated signal processing techniques that can effectively mitigate nonlinear distortions.

In system identification, the 'linear in parameter' nonlinear models are widely adopted. While the relationship between the input and output of the system is nonlinear, the identification problem is linear concerning the coefficients of the channel. Some typical examples are the Volterra filters [4] and polynomial filters which have been employed widely in several fields such as electrical engineering, mechanical engineering, and control engineering to mention a few [5]. These filters excellently model the real-life behavior of nonlinear phenomena and their memory effect. In nonlinear system identification, different Volterra filter-based approaches have been proposed. Some of these approaches are adaptive and use training symbols [6] and are typically based on least mean squares algorithms, recursive least mean square algorithms and affine projection algorithm [7]. Others are fully blind such as higher-order output cumulant-based method [8], and the subspace-based method [9, 10].

Moreover, it is worth pointing out here that the techniques presented in the recent studies of [11] and [12] were originally developed for, and successfully applied to, linear systems for which they exhibited an excellent performance. This has therefore provided us with ample encouragement to extend, in this work, these techniques to nonlinear systems so that their performance could be used as a baseline against which the performance of our newly-developed algorithm could be directly compared, thus providing us with a fair and reliable way to assess the level of improvement achieved by our proposed algorithm.

To emphasize more, the work in [13] is designed for linear systems and employs a linear model while the proposed work is designed for nonlinear systems and employs a nonlinear model. While this difference in approaches makes the comparison between both algorithms infeasible, it emphasizes the generality of our approach in that the nonlinear approach presented in our paper actually subsumes that of the linear approach presented in [13]. As such, our paper ought therefore to represent a useful and important extension of our previous work [13], which offers wider practical applications than can be afforded with the linear approach.

In this work, we propose a bilinear structure subspace method for blind channel estimation of a nonlinear SIMO system. The approach uses the intrinsic Toeplitz structure present in the signal matrix and the block Sylvester structure present in the channel matrix to construct a criterion that is minimized to estab-

Manuscript received December 16, 2024; revised June 12, 2025.

A. Lawal, A. Zerguine and A. Muqaibel are with the Department of Electrical Engineering and the Center for Communication Systems and Sensing, King Fahd University of Petroleum & Minerals, Dhahran, 31261, Saudi Arabia, (e-mail: abdulmajid.lawal@kfupm.edu.sa, azzedine@kfupm.edu.sa, and muqaibel@kfupm.edu.sa)

K. Abed-Meraim is a member of IUF affiliated with University of Orléans, PRISME Lab., France, (e-mail: karim.abed-meraim@univ-orleans.fr)

Digital Object Identifier (DOI): 10.53907/enpesj.v5i1.307

lish the estimates of the desired channel parameters. Simulation results reported in this work reveal the excellent performance of the proposed algorithm in a nonlinear environment.

Notations: The symbols $()^T$, $()^*$, $()^H$, $()^{-1}$, and $Tr()$, stand for the transpose, conjugate, conjugate transpose, inverse, and the trace operations respectively. A scalar is denoted by a, a vector by $\boldsymbol{a}$, and a matrix by $\boldsymbol{A}$. $\|\ .\ \|_{\mathrm{F}}^2$ represents the Frobenius norm operation. An $a \times a$ dimensional identity matrix is represented by $\boldsymbol{I}_a$, $a \times b$ dimensional zero matrix matrix is represented by $\mathbf{0}_{a,b}$. The entry of $\boldsymbol{A}$ at position (i, j) is denoted by $\boldsymbol{A}(i, j)$.

II. The System Model

Let us consider a SIMO nonlinear system consisting of a single transmit antenna, transmitting a signal $s(n)$, and multiple receiving antennas of size N_r as illustrated in Fig. 1. The received signal vector $\boldsymbol{y}(n)$, is given as follows [14]:

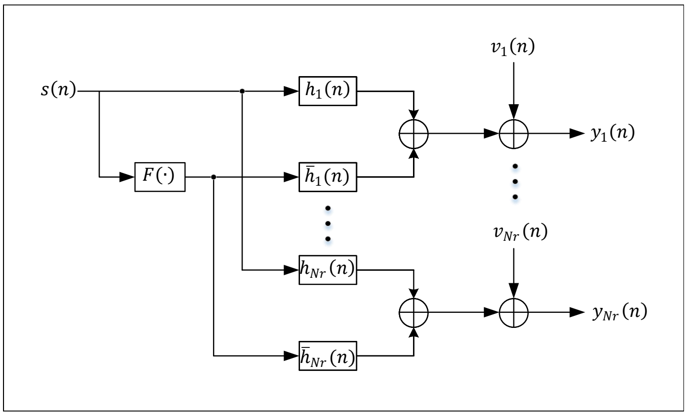

Fig. 1: The proposed SIMO system's block diagram.

$$\begin{aligned}\boldsymbol{y}(n) \ &= \sum_{k=0}^{L_L} \boldsymbol{h}(k)s(n-k) + \sum_{k=0}^{L_{NL}} \tilde{\boldsymbol{h}}(k)\tilde{s}(n-k) \\ &+ \ \boldsymbol{v}(n), \qquad (1)\end{aligned}$$

where $\boldsymbol{y}(n) = \left[y_1(n) \cdots y_{N_r}(n)\right]^T$, $\boldsymbol{h}(n) = \left[h_1(n) \cdots h_{N_r}(n)\right]^T$ and $\tilde{\boldsymbol{h}}(n) = \left[\tilde{h}_1(n) \cdots \tilde{h}_{N_r}(n)\right]^T$ denote the output signal vector, the channel vector of linear part and the channel vector of the nonlinear part of size $N_r \times 1$, respectively, $h_i(n)$ and $\tilde{h}_i(n)$ are the channel taps of the i^{th} receiving antenna and $\boldsymbol{v}(n) = \left[v_1(n) \cdots v_{N_r}(n)\right]^T$ is an additive white noise of covariance $\sigma_v^2 \boldsymbol{I}_{N_r}$, assumed to be independent of the transmitted signal. L_L and L_{NL}, respectively, represent the channel orders of the linear channel $\boldsymbol{h}(n)$ and the nonlinear channels $\tilde{h}(n)$ parts. The transmitted linear portion of the signal input, which is considered to be an i.i.d. complex random variable, is denoted by $s(n)$, $\tilde{s}(n)$ represents the nonlinear portion and F represents the nonlinear function so that $\tilde{s}(n) = F(s(n), s(n-1), ...)$ [6, 15]. In this work a quadratic nonlinearity is considered i.e., $F(s(n)) = s^2(n)$ due to the fact that many real-life applications have been modeled in this manner, amongst the popular example is the power amplifier and the optical devices [14, 16, 17]. Since the proposed model is linear with respect to the coefficients of the channel, a MIMO model that has two inputs, i.e., $\bar{\boldsymbol{s}}(n) = [s(n) \ \ \tilde{s}(n)]^T$ can be used to represent the proposed SIMO model [14]. As a result, the equivalent MIMO model representation is given as:

$$\boldsymbol{y}(n) = \sum_{k=0}^{L} \boldsymbol{H}(k)\bar{\boldsymbol{s}}(n-k) + \boldsymbol{v}(n), \ \ n = 0, \cdots, N-1, \quad (2)$$

where N is the signal size, the k^{th} channel matrix tap is denoted as $\boldsymbol{H}(k)$ can be expressed as:

$$\boldsymbol{H}(k) = \begin{bmatrix} h_1(k) & \tilde{h}_1(k) \\ \vdots & \vdots \\ h_{N_r}(k) & \tilde{h}_{N_r}(k) \end{bmatrix}. \quad (3)$$

and $L = max\{L_L, L_{NL}\}$. Assuming that N_w samples are successfully stacked into a $\boldsymbol{y}_{N_w}(n)$ vector of $M = N_w N_r$ dimension given as

$$\boldsymbol{y}_{N_w}(n) = [\boldsymbol{y}^T(n) \ \ \boldsymbol{y}^T(n-1) ... \boldsymbol{y}^T(n-N_w+1)]^T, \quad (4)$$

$$\boldsymbol{y}_{N_w}(n) = \boldsymbol{H}_{N_w}\bar{\boldsymbol{s}}_K(n) + \boldsymbol{v}_{N_w}(n), \quad (5)$$

$$\boldsymbol{H}_{N_w} = \begin{bmatrix} \boldsymbol{H}(0) & \cdots & \boldsymbol{H}(L) & & \mathbf{0} \\ & \ddots & & \ddots & \\ \mathbf{0} & & \boldsymbol{H}(0) & \cdots & \boldsymbol{H}(L) \end{bmatrix}, \quad (6)$$

where $\bar{\boldsymbol{s}}_K(n) = [\bar{\boldsymbol{s}}^T(n) \ \ \bar{\boldsymbol{s}}^T(n-1) ... \bar{\boldsymbol{s}}^T(n-K+1)]^T$ and $K = N_w + L$ and $\boldsymbol{H}_{N_w}$ represent the block Sylvester channel matrix. Finally, one can set up the data matrix as follows:

$$\begin{aligned}\boldsymbol{Y} &= [\boldsymbol{y}_{N_w}(N_w-1) \ \ \boldsymbol{y}_{N_w}(N_w), \cdots, \boldsymbol{y}_{N_w}(N-1)] \\ &= \boldsymbol{H}_{N_w}\bar{\boldsymbol{S}}_K + \boldsymbol{V}_{N_w},\end{aligned} \quad (7)$$

where

$$\bar{\boldsymbol{S}}_K = \begin{bmatrix} \bar{\boldsymbol{s}}(N_w-1) & \bar{\boldsymbol{s}}(N_w) & \cdots & \bar{\boldsymbol{s}}(N-1) \\ \bar{\boldsymbol{s}}(N_w-2) & \bar{\boldsymbol{s}}(N_w-1) & \cdots & \bar{\boldsymbol{s}}(N-2) \\ \vdots & \vdots & \vdots & \\ \bar{\boldsymbol{s}}(-L) & \bar{\boldsymbol{s}}(-L+1) & \cdots & \bar{\boldsymbol{s}}(N-K) \end{bmatrix} \quad (8)$$

In the following sections, the manuscript maintains the following assumptions: The input symbols are sufficiently complex to ensure that matrix $\bar{\boldsymbol{S}}_K$ has a full row rank. The block-Toeplitz matrix $\boldsymbol{H}_{N_w}$ has full column rank.

III. Bilinear-SIMO Nonlinear Channel Estimation Approach

This section derives the proposed nonlinear SIMO bilinear method. The method exploits the information from both the column and row subspaces of matrix $\boldsymbol{Y}$ which are used to develop a criterion that estimates the channel matrix $\boldsymbol{H}_{N_w}$ in an iterative manner. To implement the proposed method, the singular value decomposition (SVD) of the $\boldsymbol{Y}$ matrix is considered:

$$\boldsymbol{Y} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^H. \quad (9)$$

To start, let $\boldsymbol{U}_s$ represents a sub matrix of $\boldsymbol{U}$ that contains the first $2K$ columns of $\boldsymbol{U}$ with dimension $M \times 2K$, $\boldsymbol{\Sigma}_s$ is a sub matrix of $\boldsymbol{\Sigma}$ that has a dimension of $2K \times 2K$, and $\boldsymbol{V}_s$ is also

Note that in our study F(.) can be considered as a general nonlinear function. Only to test the validity of the proposed work in the sequel, we considered a second order nonlinearity.

formed by the first $2K$ columns of $\boldsymbol{V}$ and has a dimension of $(N-N_w+1)\times 2K$. The proposed bilinear method for nonlinear system is built by exploiting the structures of both the columns and row subspaces of matrix $\boldsymbol{Y}$ to estimate the channel matrix $\boldsymbol{H}_{N_w}$ (or equivalently $\boldsymbol{S}_K$ the signal matrix) in an iterative manner. In a noiseless case, the data matrix can be expressed exactly as follows:

$$\begin{aligned}\boldsymbol{Y} &= \boldsymbol{U}_s\boldsymbol{\Sigma}_s\boldsymbol{V}_s^H\\ &= \boldsymbol{U}_s\tilde{\boldsymbol{V}}_s\\ &= \boldsymbol{H}_{N_w}\bar{\boldsymbol{S}}_K,\end{aligned} \quad (10)$$

where $\tilde{\boldsymbol{V}}_s = \boldsymbol{\Sigma}_s\boldsymbol{V}_s^H$. It is obvious that the expression $\boldsymbol{H}_{N_w}\bar{\boldsymbol{S}}_K = (\boldsymbol{U}_s\boldsymbol{\mathcal{Q}})\left(\boldsymbol{\mathcal{Q}}^{-1}\tilde{\boldsymbol{V}}_s\right)$ can be satisfied by any non-singular matrix $\boldsymbol{\mathcal{Q}}$. Therefore, the aim is to search for the matrix $\boldsymbol{\mathcal{Q}}$ such that $\boldsymbol{H}_{N_w} = \boldsymbol{U}_s\boldsymbol{\mathcal{Q}}$ and $\bar{\boldsymbol{S}}_K = \boldsymbol{\mathcal{Q}}^{-1}\tilde{\boldsymbol{V}}_s$ (to the level of ambiguities inherent in blind processing).

The latter equalities are only approximately satisfied in the presence of noise by minimizing a composite criterion relative to the $\boldsymbol{H}_{N_w}$ and $\bar{\boldsymbol{S}}_K$ Sylvester and Toeplitz structure respectively as well as the nonlinear relationship between $\tilde{s}(n)$ and $s(n)$. A nonlinear matrix inversion (i.e., $\boldsymbol{\mathcal{Q}}^{-1}$) is required in such criterion. Hence, an iterative approach paired with a suitable linear approximation of matrix inverse update is proposed as

$$\boldsymbol{\mathcal{Q}}_{new} = \boldsymbol{\mathcal{Q}}_{old}(\boldsymbol{I}+\boldsymbol{\mathcal{E}}) \quad (11)$$

$$\boldsymbol{\mathcal{Q}}_{new}^{-1} \approx (\boldsymbol{I}-\boldsymbol{\mathcal{E}})\boldsymbol{\mathcal{Q}}_{old}^{-1}, \quad (12)$$

where $\boldsymbol{\mathcal{Q}}_{old}$ and $\boldsymbol{\mathcal{Q}}_{new}$ represent the current and updated value of $\boldsymbol{\mathcal{Q}}$, respectively. Here, $\boldsymbol{\mathcal{E}}$ is used to represent correction matrix, the matrix elements have values that are so small to permit the linear approximation considered. The expression for the combined cost function using(11) and (12) is given as follows:

$$J(\boldsymbol{\mathcal{E}}) = J_1 + J_2 + J_3 + J_4, \quad (13)$$

where J_1 represents the cost function responsible for minimizing the non zero portion of the Toeplitz structure of $\boldsymbol{U}_s\boldsymbol{\mathcal{Q}}_{new}$, J_2 minimizes the Toeplitz structure present in $\boldsymbol{\mathcal{Q}}_{new}^{-1}\tilde{\boldsymbol{V}}_s$, J_3 is the cost function responsible for minimizing the zero terms present in the first row and column blocks of $\boldsymbol{U}_s\boldsymbol{\mathcal{Q}}_{new}$ and J_4 is to enforce the nonlinear relation between entries $\tilde{s}(n)$ and $s(n)$ in $\boldsymbol{\mathcal{Q}}_{new}^{-1}\tilde{\boldsymbol{V}}_s$.

In the ensuing, the newly proposed cost function details are provided, starting with:

$$\begin{aligned}J_1 &= \| \boldsymbol{J}_a\boldsymbol{U}_s\boldsymbol{\mathcal{Q}}_{old}(\boldsymbol{I}+\boldsymbol{\mathcal{E}})\tilde{\boldsymbol{J}}_a - \boldsymbol{J}_b\boldsymbol{U}_s\boldsymbol{\mathcal{Q}}_{old}(\boldsymbol{I}+\boldsymbol{\mathcal{E}})\tilde{\boldsymbol{J}}_b \|_{\mathrm{F}}^2\\ &= \| \boldsymbol{A} + \boldsymbol{A}_1\boldsymbol{\mathcal{E}}\tilde{\boldsymbol{J}}_a - \boldsymbol{A}_2\boldsymbol{\mathcal{E}}\tilde{\boldsymbol{J}}_b \|_{\mathrm{F}}^2,\end{aligned}$$

where $\boldsymbol{J}_a$, $\tilde{\boldsymbol{J}}_a$, $\boldsymbol{J}_b$, and $\tilde{\boldsymbol{J}}_b$ are all selection matrices that contain ones and zeros and are used to pick the desired portion. These are, respectively, defined as $\boldsymbol{J}_a = [\boldsymbol{I}_{M-N_r}\ \boldsymbol{0}_{M-N_r,N_r}]$, $\tilde{\boldsymbol{J}}_a = [\boldsymbol{I}_{2K-1}\ \boldsymbol{0}_{2(K-1),2}]^T$, $\boldsymbol{J}_b = [\boldsymbol{0}_{M-N_r,N_r}\ \boldsymbol{I}_{M-N_r}]$, $\tilde{\boldsymbol{J}}_b = [\boldsymbol{0}_{2(K-1),2}\ \boldsymbol{I}_{2(K-1)}]^T$. Also, the matrices $\boldsymbol{A}$, $\boldsymbol{A}_1$, and $\boldsymbol{A}_2$ are found to be expressed as $\boldsymbol{A} = \boldsymbol{J}_a\boldsymbol{U}_s\boldsymbol{\mathcal{Q}}_{old}\tilde{\boldsymbol{J}}_a - \boldsymbol{J}_b\boldsymbol{U}_s\boldsymbol{\mathcal{Q}}_{old}\tilde{\boldsymbol{J}}_b$, $\boldsymbol{A}_1 = \boldsymbol{J}_a\boldsymbol{U}_s\boldsymbol{\mathcal{Q}}_{old}$, and $\boldsymbol{A}_2 = \boldsymbol{J}_b\boldsymbol{U}_s\boldsymbol{\mathcal{Q}}_{old}$.

The tightness of the proposed approximation is because the subspace method in [13] (we used for initialization) provides already a good channel estimate that is further refined by our proposed method using (typically) just few iterations.

A first-order approximation of J_1 can be shown to be:

$$J_1 \approx \| \boldsymbol{A} \|_{\mathrm{F}}^2 + 2Re\left\{Tr\left((\tilde{\boldsymbol{J}}_a\boldsymbol{A}^H\boldsymbol{A}_1 - \tilde{\boldsymbol{J}}_b\boldsymbol{A}^H\boldsymbol{A}_2)\boldsymbol{\mathcal{E}}\right)\right\},$$

Similarly, the second part of the criterion (13) is given as follows

$$\begin{aligned}J_2 &\approx \| \boldsymbol{J}_c(\boldsymbol{I}-\boldsymbol{\mathcal{E}})\boldsymbol{\mathcal{Q}}_{old}^{-1}\tilde{\boldsymbol{V}}_s\tilde{\boldsymbol{J}}_c - \boldsymbol{J}_d(\boldsymbol{I}-\boldsymbol{\mathcal{E}})\boldsymbol{\mathcal{Q}}_{old}^{-1}\tilde{\boldsymbol{V}}_s\tilde{\boldsymbol{J}}_d \|_{\mathrm{F}}^2\\ &\approx \| \boldsymbol{B} \|_{\mathrm{F}}^2 + 2Re\left\{Tr\left((\boldsymbol{B}_2\boldsymbol{B}^H\boldsymbol{J}_d - \boldsymbol{B}_1\boldsymbol{B}^H\boldsymbol{J}_c)\boldsymbol{\mathcal{E}}\right)\right\},\end{aligned}$$

where $\boldsymbol{J}_c = \left[\boldsymbol{I}_{2(K-1)}\ \boldsymbol{0}_{2(K-1),2}\right]$, $\tilde{\boldsymbol{J}}_c = \left[\boldsymbol{I}_{N-N_w}\ \boldsymbol{0}_{1,(N-N_w)}\right]^T$, $\boldsymbol{J}_d = \left[\boldsymbol{0}_{2(K-1),2}\ \boldsymbol{I}_{2(K-1)}\right]$, and $\tilde{\boldsymbol{J}}_d = \left[\boldsymbol{0}_{(N-N_w),1}\ \boldsymbol{I}_{N-N_w}\right]^T$, with $\boldsymbol{B} = \boldsymbol{J}_c\boldsymbol{\mathcal{Q}}_{old}^{-1}\tilde{\boldsymbol{V}}_s\tilde{\boldsymbol{J}}_c - \boldsymbol{J}_d\boldsymbol{\mathcal{Q}}_{old}^{-1}\tilde{\boldsymbol{V}}_s\tilde{\boldsymbol{J}}_d$, $\boldsymbol{B}_1 = \boldsymbol{\mathcal{Q}}_{old}^{-1}\tilde{\boldsymbol{V}}_s\tilde{\boldsymbol{J}}_c$ and $\boldsymbol{B}_2 = \boldsymbol{\mathcal{Q}}_{old}^{-1}\tilde{\boldsymbol{V}}_s\tilde{\boldsymbol{J}}_d$.

Finally, J_3 can be expressed as follows:

$$\begin{aligned}J_3 &= \| \boldsymbol{J}_{rw}\boldsymbol{U}_s\boldsymbol{\mathcal{Q}}_{old}(\boldsymbol{I}+\boldsymbol{\mathcal{E}})\tilde{\boldsymbol{J}}_{rw} \|_{\mathrm{F}}^2\\ &\quad + \| \boldsymbol{J}_{cl}\boldsymbol{U}_s\boldsymbol{\mathcal{Q}}_{old}(\boldsymbol{I}+\boldsymbol{\mathcal{E}})\tilde{\boldsymbol{J}}_{cl} \|_{\mathrm{F}}^2\\ &\approx \| \boldsymbol{C} \|_{\mathrm{F}}^2 + \| \boldsymbol{D} \|_{\mathrm{F}}^2\\ &\quad + 2Re\left\{Tr\left((\tilde{\boldsymbol{J}}_{rw}\boldsymbol{C}^H\boldsymbol{C}_1 + \tilde{\boldsymbol{J}}_{cl}\boldsymbol{D}^H\boldsymbol{D}_1)\boldsymbol{\mathcal{E}}\right)\right\},\end{aligned}$$

where $\boldsymbol{J}_{rw} = [\boldsymbol{I}_{N_r}\ \boldsymbol{0}_{N_r,N_wN_r-N_r}]$, $\tilde{\boldsymbol{J}}_{rw} = [\boldsymbol{0}_{2(N_w-1),2(L+1)}\ \boldsymbol{I}_{2(N_w-1)}]^T$, $\boldsymbol{J}_{cl} = [\boldsymbol{0}_{M-N_r,N_r}\ \boldsymbol{I}_{M-N_r}]$ and $\tilde{\boldsymbol{J}}_{cl} = [\boldsymbol{I}_2\ \boldsymbol{0}_{2,2(K-1)}]^T$, $\boldsymbol{C} = \boldsymbol{J}_{rw}\boldsymbol{U}_s\boldsymbol{\mathcal{Q}}_{old}\tilde{\boldsymbol{J}}_{rw}$, $\boldsymbol{C}_1 = \boldsymbol{J}_{rw}\boldsymbol{U}_s\boldsymbol{\mathcal{Q}}_{old}$, $\boldsymbol{D} = \boldsymbol{J}_{cl}\boldsymbol{U}_s\boldsymbol{\mathcal{Q}}_{old}\tilde{\boldsymbol{J}}_{cl}$, and $\boldsymbol{D}_1 = \boldsymbol{J}_{cl}\boldsymbol{U}_s\boldsymbol{\mathcal{Q}}_{old}$.

As for the last part J_4, since $\bar{\boldsymbol{S}}_K \approx \boldsymbol{\mathcal{Q}}_{new}^{-1}\tilde{\boldsymbol{V}}_s$, the linear and nonlinear parts of the matrix can be extracted using appropriate selection matrix $\boldsymbol{J}_e$ and $\boldsymbol{J}_f$. $\boldsymbol{J}_e$ is a matrix of size $K\times 2K$ formed from the odd indexed rows of an identity matrix of size $2K\times 2K$, while $\boldsymbol{J}_f$ is formed from the even indexed rows of the same matrix. Let $\boldsymbol{S}_1$ and $\boldsymbol{S}_2$ represent the linear and nonlinear signal matrices, respectively. Hence, they can be written as follows:

$$\begin{aligned}\boldsymbol{S}_1 &= \boldsymbol{J}_e\boldsymbol{\mathcal{Q}}_{new}^{-1}\tilde{\boldsymbol{V}}_s \approx \boldsymbol{J}_e(\boldsymbol{I}-\boldsymbol{\mathcal{E}})\boldsymbol{\mathcal{Q}}_{old}^{-1}\tilde{\boldsymbol{V}}_s\\ \boldsymbol{S}_2 &= \boldsymbol{J}_f\boldsymbol{\mathcal{Q}}_{new}^{-1}\tilde{\boldsymbol{V}}_s \approx \boldsymbol{J}_f(\boldsymbol{I}-\boldsymbol{\mathcal{E}})\boldsymbol{\mathcal{Q}}_{old}^{-1}\tilde{\boldsymbol{V}}_s\end{aligned} \quad (14)$$

Hence, J_4 can be expressed element-wise as follows:

$$\begin{aligned}J_4 &= \sum_{i,j}|F(\boldsymbol{S}_1(i,j)) - \boldsymbol{S}_2(i,j)|^2\\ &= \sum_{i,j}|\boldsymbol{S}_1^2(i,j) - \boldsymbol{S}_2(i,j)|^2\\ &\approx \sum_{i,j}|(\boldsymbol{J}_e(\boldsymbol{I}-\boldsymbol{\mathcal{E}})\bar{\boldsymbol{S}}_{old}(i,j))^2 - \boldsymbol{J}_f(\boldsymbol{I}-\boldsymbol{\mathcal{E}})\bar{\boldsymbol{S}}_{old}(i,j)|^2\end{aligned} \quad (15)$$

where $\bar{\boldsymbol{S}}_{old} = \boldsymbol{\mathcal{Q}}_{old}^{-1}\tilde{\boldsymbol{V}}_s$. After straightforward derivations, the first order approximation of J_4, can be written in matrix form as:

$$J_4 \approx J_{4,old} + 2Re\left\{Tr((\bar{\boldsymbol{S}}_{old}\boldsymbol{F}\boldsymbol{J}_e + \bar{\boldsymbol{S}}_{old}\boldsymbol{E}\boldsymbol{J}_f)\boldsymbol{\mathcal{E}})\right\} \quad (16)$$

where $J_{4,old} = |(\boldsymbol{J}_e\bar{\boldsymbol{S}}_{old}(i,j))^2 - \boldsymbol{J}_f\bar{\boldsymbol{S}}_{old}(i,j)|^2$ and

$$\boldsymbol{E}(j,i) = [(\boldsymbol{J}_e\bar{\boldsymbol{S}}_{old}(i,j))^2 - \boldsymbol{J}_f\bar{\boldsymbol{S}}_{old}(i,j)]^*$$

$$\boldsymbol{F}(j,i) = -2\boldsymbol{E}(j,i)(\boldsymbol{J}_e\bar{\boldsymbol{S}}_{old}(i,j)).$$

Finally, the first order expansion of $J(\boldsymbol{\mathcal{E}})$ is expressed as follows:

$$\begin{aligned} J(\boldsymbol{\mathcal{E}}) \approx &\parallel \boldsymbol{A} \parallel_F^2 + \parallel \boldsymbol{B} \parallel_F^2 + \parallel \boldsymbol{C} \parallel_F^2 + \parallel \boldsymbol{D} \parallel_F^2 \\ &+ J_{4,old} + 2Re\left\{Tr\left((\boldsymbol{M}_a + \boldsymbol{M}_b + \boldsymbol{M}_c + \boldsymbol{M}_d)\boldsymbol{\mathcal{E}}\right)\right\}, \end{aligned} \quad (17)$$

where $\boldsymbol{M}_a = \tilde{\boldsymbol{J}}_a\boldsymbol{A}^H\boldsymbol{A}_1 - \tilde{\boldsymbol{J}}_b\boldsymbol{A}^H\boldsymbol{A}_2$, $\boldsymbol{M}_b = \boldsymbol{B}_2\boldsymbol{B}^H\boldsymbol{J}_d - \boldsymbol{B}_1\boldsymbol{B}^H\boldsymbol{J}_c$, $\boldsymbol{M}_c = \tilde{\boldsymbol{J}}_{rw}\boldsymbol{C}^H\boldsymbol{C}_1 + \tilde{\boldsymbol{J}}_{cl}\boldsymbol{D}^H\boldsymbol{D}_1$ and $\boldsymbol{M}_d = \boldsymbol{S}_{old}\boldsymbol{F}\boldsymbol{J}_e + \boldsymbol{S}_{old}\boldsymbol{E}\boldsymbol{J}_f$. Here, to ensure that the cost function decreases, $\boldsymbol{\mathcal{E}}$ is selected such that it moves in a direction opposite to that of the gradient, according to:

$$\boldsymbol{\mathcal{E}} = -\rho\left(\boldsymbol{M}_a + \boldsymbol{M}_b + \boldsymbol{M}_c + \boldsymbol{M}_d\right)^H, \quad (18)$$

where ρ represents a small positive constant.

The proposed bilinear method for nonlinear systems uses the structured channel subspace method in [13] for initialization and few iterations are used for channel and signal matrix refinement in order to improve estimation quality. The step-by-step procedure involved in the proposed bilinear method is described in Algorithm 1.

Algorithm 1: Summary of the proposed Bilinear method.

initialization;
$\boldsymbol{\mathcal{Q}}_{old}$, $\bar{\boldsymbol{S}}_{old}$
while *Number of iterations* $\leq N_{max}$ **do**
  $\boldsymbol{\mathcal{Q}}_{old} = \boldsymbol{\mathcal{Q}}_{new}$
  $\boldsymbol{M}_a = \tilde{\boldsymbol{J}}_a\boldsymbol{A}^H\boldsymbol{A}_1 - \tilde{\boldsymbol{J}}_b\boldsymbol{A}^H\boldsymbol{A}_2$
  $\boldsymbol{M}_b = \boldsymbol{B}_2\boldsymbol{B}^H\boldsymbol{J}_d - \boldsymbol{B}_1\boldsymbol{B}^H\boldsymbol{J}_c$
  $\boldsymbol{M}_c = \tilde{\boldsymbol{J}}_{rw}\boldsymbol{C}^H\boldsymbol{C}_1 + \tilde{\boldsymbol{J}}_{cl}\boldsymbol{D}^H\boldsymbol{D}_1$
  $\boldsymbol{M}_d = \bar{\boldsymbol{S}}_{old}\boldsymbol{F}\boldsymbol{J}_e + \bar{\boldsymbol{S}}_{old}\boldsymbol{E}\boldsymbol{J}_f$.
  $\boldsymbol{\mathcal{E}} = -\rho\left(\boldsymbol{M}_a + \boldsymbol{M}_b + \boldsymbol{M}_c + \boldsymbol{M}_d\right)^H$
  $\boldsymbol{\mathcal{Q}}_{new} = \boldsymbol{\mathcal{Q}}_{old}(\boldsymbol{I} + \boldsymbol{\mathcal{E}})$
  $\boldsymbol{\mathcal{Q}}_{new}^{-1} \approx (\boldsymbol{I} - \boldsymbol{\mathcal{E}})\boldsymbol{\mathcal{Q}}_{old}^{-1}$
  $\bar{\boldsymbol{S}}_{new} \approx (\boldsymbol{I} - \boldsymbol{\mathcal{E}})\bar{\boldsymbol{S}}_{old}$
  $\boldsymbol{\mathcal{H}}_{N_w} = \boldsymbol{U}_s\boldsymbol{\mathcal{Q}}_{new}$
end

IV. Computational Complexity

The proposed method's computation complexity is compared to the complexity of the subspace (SS) method [11] and the structured channel subspace (SCS) method [12]. The proposed bilinear method has the heaviest computational cost with a total complexity of $O((N_rN_w)(N-N_w)^2) + O((K(N-N_w))^2) + O((N_rN_wK)^2)$ due to the data matrix SVD and the fact that it is initialized with $\boldsymbol{\mathcal{Q}}$ from the SCS method. The next in terms of computational cost is the SCS method with a complexity of $O((N_rN_w)^2(N-N_w)) + O((N_rN_wK)^2)$. Finally, the SS method has the least computational cost with $O((N_rN_w)^2(N-N_w)) + O((N_r(L+1))^2)$. Here, O represents the order of complexity.

This method estimates the $N_r \times 2$ MIMO channel in (3) up to a 2×2 unknown matrix (see [13] for details).

V. Simulation Results

In this section, the normalized mean squared error (NMSE) is used as performance metric for the developed nonlinear bilinear method (BSCS), the SS method and the SCS method are investigated via simulations experiments. In fact, the SCS and SS methods provide partial channel estimate of the $N_r \times 2$ MIMO system in (2). More precisely, the estimated channel taps $\hat{\boldsymbol{H}}(k)$, $k = 0, \cdots, L$ satisfy

$$\hat{\boldsymbol{H}}(k) \approx \boldsymbol{H}(k)\mathbf{D}$$

where $\mathbf{D}$ is a 2×2 unknown matrix (see details in [18]). In comparison, our bilinear method fully estimates the SIMO nonlinear channel up to a scalar factor (which is the inherent ambiguity of the blind processing methods for SIMO systems) [19]- [20]. In the sequel, for comparison purpose, we remove these ambiguities from the channel estimates, using a least squares fitting criterion with the exact channel matrix, before their use in criterion (19).

The effectiveness of the proposed channel estimator is firstly verified by the normalized mean squares error (NMSE) of channel estimation, which is given in dB as:

$$\text{NMSE} = 20\log_{10}\left(\frac{\sqrt{\frac{1}{N_{mc}}\sum_{i=1}^{N_{mc}} \parallel \hat{\boldsymbol{H}}_i - \boldsymbol{H} \parallel_F^2}}{\parallel \boldsymbol{H} \parallel_F^2}\right), \quad (19)$$

where $N_{mc} = 100$ is the number of Monte Carlo runs, $\boldsymbol{H} = [\boldsymbol{H}(0) \cdots \boldsymbol{H}(L)]$ represents the true channel employed in the simulations, and $\hat{\boldsymbol{H}}_i$ represents the estimated channel at the i^{th} run.

The input stream $s(n)$ is drawn from a QAM 16 constellation, the nonlinear part is obtained by passing the input stream through a nonlinear function $\tilde{s}(n) = F(s(n)) = s^2(n)$ and the additive noise is white Gaussian with zero mean which is generated for each Monte Carlo run. Throughout the simulations, the input signal length of $N = 100$ is considered, $N_r = 4$ receiver antennas, a window size of $N_w = 5$, and the channel is generated randomly with an order of $L_L = L_{NL} = L = 3$ in all experiments except when otherwise specified. It is crucial to note that the nonlinear bilinear technique is iterative, with the step size set to $\rho = 1 \times 10^{-6}$ in all experiments.Here, in this work our aim was to test the superiority of our algorithm in fair comparison with the rest of the algorithms.

The proposed bilinear method significantly outperforms the SS method, the SCS method, and the Bayesian method as shown in Fig. 2, with a gain of over 2dB in our algorithm's favor.

In the next experiment, the performance of the algorithms, including the previous linear-based algorithm (LBSCS) [13], and the Bayesian method is investigated in terms of the symbol error rate (SER) a more discerning and a more reliable metric

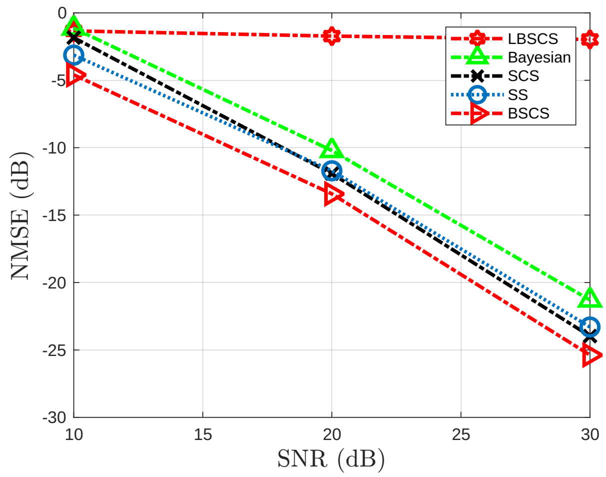

Fig. 2: The NMSE performance versus SNR.

for this type of steady-state analysis. The results of this comparison are all shown in Fig. 3, and clearly demonstrate the strong consistency between the performance of our proposed algorithm and other previously published and related algorithms. As shown in Fig 3, the proposed bilinear method outperforms all the other algorithms, including our algorithm in [13] with the worst performance, and gives it an edge in steady-state performance over the other algorithms. For a 6×10^{-3} SER, a gain of 2 dB in favor of the proposed algorithm against the SCS method is achieved. Consistency in performance of the proposed algorithm is maintained in all the experiments performed in this study even the Bayesian method was outperformed by the proposed algorithm. The proposed method requires the longest computation time among the evaluated techniques, primarily due to the joint estimation of signal and channel subspaces and the iterative refinement process. The Bayesian method follows, exhibiting moderate computational time. In contrast, both the SCS and SS methods are computationally efficient, with execution times significantly shorter than those of the proposed and Bayesian approaches.

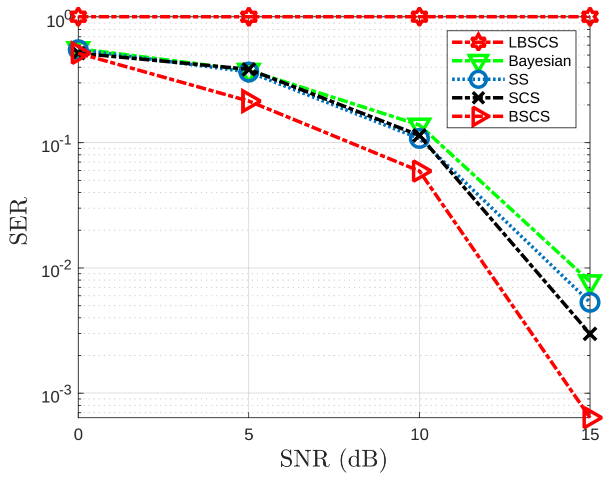

Fig. 3: The SER performance versus the SNR.

VI. Conclusion

In this paper, an iterative and efficient bilinear structure subspace method for blind channel identification is developed which exploits the column and row subspaces. In addition, the algorithm was designed to take care of any nonlinearity in the system. This was achieved by introducing a new term in the minimization procedure as to tackle the system's nonlieanrity. The algorithm performs better than the considered methods in terms of NMSE and SER. To examine how well the proposed algorithm performed, several scenarios were examined.

Overall, the proposed method outperformed other methods at the expense of a moderate computational load which is mainly due to the involvement of the iterative procedure.

References

[1] V. J. Mathews, "Adaptive polynomial filters," *Jul*, vol. 1, pp. 10–26, 1991.

[2] J. Schoukens and L. Ljung, "Nonlinear system identification: A user-oriented road map," *IEEE Control Syst. Mag.*, vol. 39, no. 6, pp. 28–99, 2019.

[3] A. Amari, O.A. Dobre, R. Venkatesan, O.S. Kumar, P. Ciblat, Y. Jaouen, "A survey on fiber nonlinearity compensation for 400 Gb/s and beyond optical communication systems," *IEEE Commun. Surv. Tutor.*, vol. 19, no. 4, pp. 3097–3113, 2017.

[4] T. Ogunfunmi, *Adaptive nonlinear system identification: The Volterra and Wiener model approaches*. Springer Science & Business Media, 2007.

[5] C. Cheng, Z. Peng, W. Zhang, and G. Meng, "Volterra-series-based nonlinear system modeling and its engineering applications: A state-of-the-art review," *Mechanical Systems and Signal Processing*, vol. 87, pp. 340–364, 2017.

[6] T. Ogunfunmi and S.-L. Chang, "Second-order adaptive volterra system identification based on discrete nonlinear wiener model," *IEE Proceedings-Vision, Image and Signal Processing*, vol. 148, no. 1, pp. 21–29, 2001.

[7] Y. V. Zakharov, G. P. White, and J. Liu, "Low-complexity RLS algorithms using dichotomous coordinate descent iterations," *IEEE Trans. Signal Process.*, vol. 56, no. 7, pp. 3150–3161, 2008.

[8] N. Kalouptsidis and P. Koukoulas, "Blind identification of volterra-hammerstein systems," *IEEE Trans. Signal Process.*, vol. 53, no. 8, pp. 2777–2787, 2005.

[9] J. C. Gómez and E. Baeyens, "Subspace-based identification algorithms for hammerstein and wiener models," *European Journal of Control*, vol. 11, no. 2, pp. 127–136, 2005.

[10] W. Qiu and Y. Hua, "Performance analysis of the subspace method for blind channel identification," *Signal Processing*, vol. 50, no. 1-2, pp. 71–81, 1996.

[11] C. Yu, L. Xie, M. Verhaegen, and J. Chen, "Blind identification of MIMO systems," in *Blind Identification of Structured Dynamic Systems*, pp. 167–198, Springer, 2022.

[12] Q. Mayyala, K. Abed-Meraim, and A. Zerguine, "Structure-based subspace method for multichannel blind system identification," *IEEE Signal Process. Lett.*, vol. 24, no. 8, pp. 1183–1187, 2017.

[13] A. Lawal, Q. Mayyala, K. Abed-Meraim, N. Iqbal, and A. Zerguine, "Toeplitz structured subspace for multi-channel blind identification methods," *Signal Processing*, vol. 188, p. 108152, 2021.

[14] O. Rekik, K. Abed-Meraim, M. Nait-Meziane, A. Mokraoui, and N. L. Trung, "Maximum likelihood based identification for nonlinear multichannel communications systems," *Signal Processing*, vol. 189, p. 108297, 2021.

[15] L. Tao, H. Tan, C. Fang, and N. Chi, "Volterra series based blind equalization for nonlinear distortions in short reach optical cap system," *Optics Communications*, vol. 381, pp. 240–243, 2016.

[16] A. H. Abdelhafiz, O. Hammi, A. Zerguine, A. T. Al-Awami, and F. M. Ghannouchi, “A pso based memory polynomial predistorter with embedded dimension estimation,” *IEEE Transactions on Broadcasting*, vol. 59, no. 4, pp. 665–673, 2013.

[17] A. E. Abdelrahman, O. Hammi, A. K. Kwan, A. Zerguine, and F. M. Ghannouchi, “A novel weighted memory polynomial for behavioral modeling and digital predistortion of nonlinear wireless transmitters,” *IEEE Transactions on Industrial Electronics*, vol. 63, no. 3, pp. 1745–1753, 2015.

[18] K. Abed-Meraim, P. Loubaton, and E. Moulines, “A subspace algorithm for certain blind identification problems,” *IEEE Trans. Inf. Theory*, vol. 43, no. 3, pp. 499–511, 1997.

[19] G. Xu, H. Liu, L. Tong, and T. Kailath, “A least-squares approach to blind channel identification,” *IEEE Transactions on Signal Processing*, vol. 43, no. 12, pp. 2982–2993, 1995.

[20] W. Qiu and Y. Hua, “Performance analysis of the subspace method for blind channel identification,” *Signal Processing*, vol. 50, no. 1, pp. 71–81, 1996.

ABDULMAJID LAWAL earned a B.Eng. in electrical engineering from Bayero University in Kano in 2012, a Master’s in electrical engineering from King Fahd University of Petroleum and Minerals (KFUPM) in Saudi Arabia in 2016, and a PhD in electrical engineering from the same institution in 2021. He is presently a postdoctoral fellow student at the KFUPM, Saudi Arabia, center for communication systems and sensing. He focuses on issues with blind source separation, blind system/channel identification, sparse signal representations, and artificial intelligence. Statistical signal processing, modeling, estimation, and classification of signals, blind system identification, blind source separation, statistical performance analysis, and sparse representations are some of his areas of interest in study.

KARIM ABED-MERAIM was born in 1967. He received the State Engineering Degree from the École Polytechnique, Palaiseau, France, in 1990, the State Engineering Degree from the École Nationale Supérieure des Télécommunications (ENST), Paris, France, in 1992, the M.Sc. degree from Paris XI University, Orsay, France, in 1992, and the Ph.D. degree in the field of signal processing and communications from ENST, in 1995. From 1995 to 1998, he was a Research Staff with the Electrical Engineering Department, The University of Melbourne, where he worked on several research project related to “Blind System Identification for Wireless Communications,” "Blind Source Separation," and "Array Processing for Communications." From 1998 to 2012, he has been an Assistant Professor and then as an Associate Professor with the Signal and Image Processing Department, Télécom ParisTech. His research interests include signal processing for communications, adaptive filtering and tracking, array processing, and statistical performance analysis. In September 2012, he joined the University d’Orléans (PRISME Laboratory), as a Full Professor. He is the author of over 550 scientific publications, including book chapters, international journal and conference papers, and patents. Dr. Abed-Meraim is a past member of the IEEE SAM-TC and past Associate Editor (and Senior Area Editor) of the IEEE-TSP.

AZZEDINE ZERGUINE received the B.Sc. degree in electrical engineering from Case Western Reserve University, Cleveland, OH, USA, in 1981, the M.Sc. degree in electrical engineering from the King Fahd University of Petroleum and Minerals (KFUPM), Dhahran, Saudi Arabia, in 1990, and the Ph.D. degree in electrical engineering from Loughborough University, Loughborough, U.K., in 1996. He is currently a Professor with the Department of Electrical Engineering, KFUPM, working in the areas of signal processing and communications. His current research interests include signal processing for communications, adaptive filtering, neural networks, the applications of artificial intelligence, blind source separation, and blind equalization. He was a recipient of the three Best Teaching Awards at KFUPM, in 2000, 2005, and 2011, and the Best Research Award at KFUPM, in 2017. Actually, he serves as an Associate Editor for the *Signal Processing* Journal and the IEEE COMMUNICATIONS LETTERS.

ALI HUSSEIN MUQAIBEL received the Ph.D. degree from the Virginia Polytechnic Institute and State University, Blacksburg, VA, USA, in 2003. He is currently a Professor with the Electrical Engineering Department, King Fahd University of Petroleum and Minerals (KFUPM). He is also the Director of the Interdisciplinary Research Center for Communication Systems and Sensing (IRC-CSS). During his study with Virginia Tech, he was with the Time Domain and RF Measurements Laboratory and the Mobile and Portable Radio Research Group. He was a Visiting Associate Professor with the Center of Advanced Communications, Villanova University, Villanova, PA, USA, in 2013, a Visiting Professor with Georgia Institute of Technology, in 2015, and a Visiting Scholar with the King Abdullah University for Science and Technology (KAUST), Thuwal, Saudi Arabia, in 2018 and 2019. He has authored three book chapters and more than 150 articles. His research interests include communications and sensing applications, including the direction of arrival estimation, through-wall-imaging, localization, channel characterization, and ultra-wideband signal processing. He was a recipient of many awards for excellence in teaching, advising, and instructional technology.