Subspace Projection Based Blind Channel Order Estimation of MIMO Systems

Diamantis Kotoulas, Student Member, IEEE, Panos Koukoulas, Member, IEEE, and Nicholas Kalouptsidis, Senior Member, IEEE

Abstract—In this paper, a novel algorithm based on subspace projections is developed for blindly estimating the discrete orders of a linear finite-impulse-response (FIR) multiple-input multiple-output (MIMO) system, the number of subsystems that attain each order as well as the total number of inputs. Furthermore, the proposed algorithm applies to single-input multiple-output (SIMO) system order estimation. Simulations in the context of blind channel order estimation show good performance in comparison to existing schemes developed for SIMO systems.

Index Terms—Blind channel identification, multiple-input multiple-output (MIMO) systems, order determination.

I. INTRODUCTION

Blind identification and equalization of linear time-invariant (LTI) systems with finite impulse response (FIR) has received a lot of attention in recent years. The problem is often encountered in digital communication systems where unknown signals are transmitted through unknown, multipath channels. Application areas include digital multiuser/multicore communication systems, digital television systems, multisensor sonar/radar systems, and speech systems [1]–[7].

In blind channel identification, the channel is identified without using a training sequence, as this may be a bandwidth consuming as well as throughput reducing technique. In general, the receiver uses the system output sequences in conjunction with some a priori statistical information about the system inputs.

To tackle the problem of blind identification for single-input multiple-output (SIMO) as well as multiple-input multiple-output (MIMO) systems, several algorithms have been developed using only (or at least to the maximum extent possible) the second-order statistics (SOS) of the data and exploiting the cyclostationarity of the received signal either through oversampling in rates higher than the baud rate or using multiple sensors (see [8] and [9] and references therein).

In principle, this type of algorithms can achieve satisfactory performance utilizing a smaller number of output samples and using either a direct or an indirect approach. The direct (equalization) approach extracts the desired source signals without explicitly identifying the unknown channel impulse response (see [10] and [11] and references therein). However, this approach may suffer from some drawbacks such as local minima and the need for a considerable amount of output data utilized before convergence.

On the other hand, according to the indirect approach, the channel is first identified and then equalized (see [2], [4], [8], [12]–[19], and references therein). Algorithms of this type may achieve superior performance but they are very sensitive to the exact knowledge of the channel order. When the later is known, estimation is satisfactory. When channel order is not known the performance of the channel estimation tends to be poor in the presence of noise. So we may argue that correct channel order estimation can be of significant importance for blind identification. However, research has proved that estimating the correct channel order is difficult for both SIMO and MIMO band limited channels [20]. Specifically, for the case of SIMO systems several blind order estimation algorithms have been developed (see [9] and [21]–[23] and references therein). These methods do not give totally satisfactory results yet, as their success depends highly on variations of SNR as well as the number of output data samples used for computations. The problem of blind order determination of MIMO systems was addressed for the first time in [24].

In this paper, order determination of MIMO systems is considered and a new algorithm based on subspace projections is developed. The proposed method is an extension of the algorithm presented in [21] for SIMO systems. It can also be applied for the SIMO case where simulations showed that it achieves better performance when compared to existing methods such as the ones proposed in [21]–[23]. In addition to blindly estimating channel orders, the algorithm can be used to compute the number of system inputs. Furthermore, it can be combined with identification methods that require the exact knowledge of channel orders. A typical representative of these methods is given in [8]. Even though the algorithm is developed for a noiseless setting, simulations showed good performance in the presence of noise.

The paper is organized as follows. Section II describes the mathematical model of the system, sets the underlying assumptions and defines the problem objectives. In Section III the theoretical basis on which the algorithm relies is established. Order determination is undertaken in Section IV. Data struc-
tures needed and computational steps taken when implementing the proposed algorithm are given in Section V. Simulations are provided in Section VI and the paper ends with Section VII where the conclusions are presented.

II. MODEL SPECIFICATION

A. Notations

The notation employed in this paper is standard. Signals are discrete-time and complex in general. Upper- and lower-case bold letters denote matrices and vectors respectively. $(\cdot)^t$ and $(\cdot)^\dagger$ are transpose and Hermitian operations, $\mathbf{0}_{m \times n}$ stands for the $m \times n$ zero matrix. Given a matrix $\mathbf{A}$, $\mathcal{R}(\mathbf{A})$ and $\mathcal{C}(\mathbf{A})$ stand for the row and column space of matrix $\mathbf{A}$ and $\|\mathbf{A}\|_\infty$ denotes the $\infty$ norm of the matrix. For a given matrix $\mathbf{X}$ having the same number of columns as $\mathbf{A}$, $\mathcal{P}_\mathbf{A}[\mathbf{X}]$ denotes the projection of $\mathbf{X}$ onto the row space of $\mathbf{A}$. A set of vectors $\mathbf{x}_1, \cdots, \mathbf{x}_n$, $sp[\mathbf{x}_1, \cdots, \mathbf{x}_n]$ is the linear subspace spanned by $\mathbf{x}_1, \cdots, \mathbf{x}_n$. If $A, B$ are vector spaces over the same field $F$, $A \oplus B$ is their direct sum and $A \cong B$ denotes that they are isomorphic, that is there exists an one to one and onto linear map from $A$ to $B$, (definition given for example in [26, p. 4]). If $\mathbf{x}, \mathbf{X}$ are a vector and a matrix, respectively, and $\mathcal{A}$ is a vector space, then $\mathbf{x}|_{\mathcal{A}}$ is the orthogonal projection of $\mathbf{x}$ onto $\mathcal{A}$ and $\mathbf{X}|_{\mathcal{A}}$ is the orthogonal projection of the rows of $\mathbf{X}$ onto $\mathcal{A}$. Finally, $\mathcal{N}$ denotes the set of positive integers.

B. The System Model

We shall be concerned with FIR MIMO systems of the form

$$\mathbf{x}(k) = \sum_{i=1}^{P} \sum_{j=0}^{L_i} \mathbf{h}_i(j) s_i(k-j).$$

(1)

The system has $P$ inputs and $M$ outputs. Thus, the output signal $\mathbf{x}(k)$ is an $M \times 1$ dimensional vector. The input sequences consist of the signals $s_1(k), s_2(k), \cdots, s_P(k)$. The orders of the $P$ subsystems are given by the integers $L_1, L_2, \cdots, L_P$. For each $1 \leq i \leq P$, and $0 \leq j \leq L_i$, $\mathbf{h}_i(j)$ is the corresponding $M \times 1$ kernel tap.

Equivalently, the above system can be described as

$$\mathbf{x}(k) = [\mathbf{H}(z)] \mathbf{s}(k)$$

(2)

where $[\mathbf{H}(z)]$ is the system transfer function and $\mathbf{s}(k) = [s_1(k) \cdots s_P(k)]^\dagger$.

If $L_q, 1 \leq q \leq P$, denotes the maximum of $L_1, L_2, \cdots, L_P$, then the channel polynomial matrix $\mathbf{H}(z)$ can be written as

$$\mathbf{H}(z) = \sum_{i=0}^{L_q} \mathbf{h}(i)z^{-i}$$

(3)

with

$$\mathbf{H}(i) = [\mathbf{h}_1(i) \mathbf{h}_2(i) \cdots \mathbf{h}_P(i)].$$

(4)

The following assumptions are made.

1) The input sequences $s_1(k), s_2(k), \cdots, s_P(k)$ consist of stationary independent identically distributed (i.i.d.) zero mean signals of finite variance that are mutually independent with each other.

2) An upper bound $I$ of the subsystems’ orders is known.

3) The number of inputs $P$ is strictly less than the number of outputs $M$. Furthermore, the channel polynomial matrix $\mathbf{H}(z)$ is irreducible and column reduced.

As a result, (see [25]), there is a smallest $u_0 > 0$, $u_0 \in \mathcal{N}$ such that the matrix $\mathbf{H}_{u_0}^{-1}(\mathbf{h}) = [\mathbf{F}_{u_0}(\mathbf{h}) \cdots \mathbf{F}_{u_0}(\mathbf{h}_P)]$, where

$$\mathbf{F}_{u_0}(\mathbf{h}) = \begin{pmatrix}
\mathbf{h}_0(0) & \cdots & \mathbf{h}_r(0) \\
0_{M \times 1} & \ddots & \ddots & \ddots \\
\vdots & \ddots & \ddots & \ddots \\
0_{M \times 1} & \cdots & \mathbf{h}_0(L_r) & \cdots & \mathbf{h}_r(L_r)
\end{pmatrix}$$

has full column rank. $\mathbf{F}_{u_0}(\mathbf{h})$ is an $(M u_0) \times (L_r + u_0)$ matrix. As stated in [25], to guarantee the full column rank of $\mathbf{H}_{u_0}^{-1}(\mathbf{h})$ it suffices to select $u_0 \geq \sum_{i=1}^P L_i - 1$.

C. Objectives

Consider the framework of the previous subsection and suppose that assumptions A1)–A3) hold. Our objective is to establish the following:

a) the number of inputs $P$ can be computed;

b) $L_1, L_2, \cdots, L_P$ are uniquely determined.

Specifically, if $1 \leq i \leq r$, then for all $1 \leq i \leq r$, the number $m_i$ of subsystems that have order $L_i$ can be computed.

To establish both a) and b), we use rank properties of projection error matrices formed by the system’s output data. These properties are illustrated by Theorems 2 and 3 in Section IV, while the proposed algorithm is presented in Section V.

III. VARIABLES, SPACES, PROJECTIONS, AND ISOMORPHIC RELATIONS

The proposed approach relies on the work by Tong et al., described in [21]. Linear prediction and smoothing techniques that exploit the isomorphic relationship between input and output subspaces are employed. Precise definitions of relevant variables, spaces and projections are given next. The main result is Theorem 1, which establishes the isomorphic relationship between input and output subspaces. For our analysis, we shall use the following lemma, the validity of which is established in [25].

Lemma 1: For all $w \geq u_0$, $\mathbf{H}(w)$ has full column rank.

Next we proceed with some definitions. Assuming that $k_0 \in \mathcal{N}$ and $Q > k$. We collect successive output vectors in the matrix

$$\mathbf{T}(k) = [\mathbf{x}(k) \quad \mathbf{x}(k+1) \quad \cdots \quad \mathbf{x}(Q)]$$

(5)

Likewise, we combine successive input values in row vector form as follows:

$$\mathbf{s}_1(k) = [s_1(k) \quad s_1(k+1) \quad \cdots \quad s_1(Q)]$$

$$\vdots$$

$$\mathbf{s}_P(k) = [s_P(k) \quad s_P(k+1) \quad \cdots \quad s_P(Q)].$$

(6)
Using a window of length \( w \), we form the following array of output samples:

\[
X_{k,w} = \begin{pmatrix}
T(k) \\
(\cdots) \\
T(k-w+1)
\end{pmatrix}.
\]

(7)

\( X_{k,w} \) is the data matrix defined by stacking \( w \) consecutive such observations, starting with \( T(k) \) and going back to \( T(k-w+1) \). Similarly, for all \( j, 1 \leq j \leq P \), we have

\[
S^j_{k,w} = \begin{pmatrix}
s^j(k) \\
(\cdots) \\
s^j(k-w+1)
\end{pmatrix}.
\]

(8)

Finally, we consider the row spaces

\[
\mathcal{X}_{k,w} = \mathcal{R}(X_{k,w})
\]

(9)

\[
\mathcal{S}^j_{k,w} = \mathcal{R}(S^j_{k,w}).
\]

(10)

The following theorem establishes the isomorphism between input and output subspaces.

**Theorem 1:** For all \( w \geq w_0 \), \( \mathcal{X}_{k,w} \cong \mathcal{S}^1_{k,L_1+w} \oplus \cdots \oplus \mathcal{S}^P_{k,L_P+w} \).

**Proof:** The definitions preceding the theorem and (1) lead to

\[
X_{k,w} = H_w(h)
\]

\[
\begin{pmatrix}
S^1_{k,L_1+w} \\
S^2_{k,L_2+w} \\
(\cdots) \\
S^P_{k,L_P+w}
\end{pmatrix}
\]

Due to Lemma 1, \( H_w(h) \) has full column rank; therefore

\[
\mathcal{X}_{k,w} \cong \mathcal{R}\left(\begin{pmatrix}
S^1_{k,L_1+w} \\
S^2_{k,L_2+w} \\
(\cdots) \\
S^P_{k,L_P+w}
\end{pmatrix}\right).
\]

(11)

Notice that the rows of the matrix

\[
\begin{pmatrix}
S^1_{k,L_1+w} \\
S^2_{k,L_2+w} \\
(\cdots) \\
S^P_{k,L_P+w}
\end{pmatrix}
\]

(12)

are orthogonal with respect to the inner product induced by correlation in the Hilbert space of zero mean random vectors with finite second moments. This is due to the i.i.d. and mutual independence assumptions stated in A1). Therefore, the rows of the above matrix are linearly independent with probability one. This, together with (11) proves the theorem.

For any \( l \in \mathcal{N} \) we define the vector space

\[
\tilde{\mathcal{S}}_{k,l} = (S^1_{k-1,L_1+w} \oplus \cdots \oplus S^P_{k-1,L_P+w}) \cup (S^1_{k+l+w,L_1+w} \oplus \cdots \oplus S^P_{k+l+w,L_P+w}).
\]

(13)

If we think of \( I \) as a smoothing window, we see that \( \tilde{\mathcal{S}}_{k,l} \) is a subspace constructed from both past and future data observations.

According to Theorem 1, we have that

\[
\mathcal{X}_{k-1,w} \cong (S^1_{k-1,L_1+w} \oplus \cdots \oplus S^P_{k-1,L_P+w})
\]

and

\[
\mathcal{X}_{k+l+w} \cong (S^1_{k+l+w,L_1+w} \oplus \cdots \oplus S^P_{k+l+w,L_P+w}.)
\]

Using (13)–(15), we conclude that

\[
\tilde{\mathcal{S}}_{k,l} \cong \mathcal{X}_{k-1,w} \cup \mathcal{X}_{k+l+w}.
\]

(16)

IV. ORDER DETERMINATION—COMPUTATION OF THE NUMBER OF INPUTS

In this section, the proposed algorithm is established. It relies on subspace projections and rank determination of suitable matrices obtained by proper windowing and smoothing of output data to past and future observations.

We depart from (1), (5), and (6) and rewrite \( T(k) \) as follows:

\[
T(k) = \sum_{i=1}^{P} \sum_{j=0}^{L_i} h_i(j) s_i(k-j).
\]

(17)

Next we consider the projection error matrix \( E_{k,l} \) as

\[
E_{k,l} = \begin{pmatrix}
\frac{T(k+l) - T(k+l)_{\tilde{S}_{k,l}}}{T(k+l) - T(k+l)} & \cdots & 0 \\
\frac{T(k+l-1) - T(k+l-1)}{T(k+l-1)} & \cdots & 0 \\
\vdots & \ddots & \vdots \\
0 & \cdots & 0
\end{pmatrix}
\]

\( E_{k,l} \) is a \((M(l+1)) \times (Q-k+1)\) matrix. Each block entry \( T(k+m) - T(k+m)_{\tilde{S}_{k,l}} \) of \( E_{k,l} \) is formed by the error resulting from the projection of \( T(k+m) \) on the space \( \tilde{S}_{k,l} \) generated by past and future input values as defined in (13). The structure of the projection error matrix \( E_{k,l} \) is relevant to order information is elucidated by the following theorem.

**Theorem 2:** The following statements are true.

i) For any \( l : l < J_1 \),

\[
E_{k,l} = 0.
\]

(18)

ii) For \( l : J_1 \leq l < J_2 \),

\[
E_{k,l} = H^i_I(G_1) \tilde{S}_{k,l-J_1}.
\]

(19)

iii) For \( n : 3 \leq n \leq r \) and \( l : J_{n-1} \leq l < J_n \),

\[
E_{k,l} = H^i_I(G_1) \tilde{S}_{k,l-J_1} + \cdots + H^i_I(G_{r-1} \tilde{S}_{k,l-J_{n-1}}.
\]

(20)

iv) For \( l : J_r \leq l \leq L \),

\[
E_{k,l} = H^i_I(G_1) \tilde{S}_{k,l-J_1} + \cdots + H^i_I(G_r) \tilde{S}_{k,l-J_r}.
\]
where
\[
\mathbf{D}_l^i \left( \mathbf{h}^{(i_1)} \right) = \begin{pmatrix}
\mathbf{h}_{i_1}(0) & \cdots & \mathbf{0}_{M \times 1} \\
\vdots & \ddots & \vdots \\
\mathbf{h}_{i_1}(J_i) & \cdots & \mathbf{h}_{i_1}(0) \\
\mathbf{0}_{M \times 1} & \cdots & \mathbf{h}_{i_1}(J_i)
\end{pmatrix}
\]

and
\[
\mathbf{H}_l^i(\mathbf{G}_j) = \left[ \mathbf{D}_l^i \left( \mathbf{h}^{(i_1)} \right) \cdots \mathbf{D}_l^i \left( \mathbf{h}^{(i_m)} \right) \right]
\]

and
\[
\mathbf{\tilde{S}}_{k,l,J_1}^i = \begin{pmatrix}
\mathbf{s}_1(k + l - J_i) - \mathbf{s}_1(k + l - J_i) | \mathbf{S}_{k,l,J_1} \\
\vdots \\
\mathbf{s}_i(k) - \mathbf{s}_i(k) | \mathbf{S}_{k,l,J_1} \\
\vdots \\
\mathbf{s}_{i_{m_1}}(k + l - J_i) - \mathbf{s}_{i_{m_1}}(k + l - J_i) | \mathbf{S}_{k,l,J_1} \\
\vdots \\
\mathbf{s}_{i_{m_1}}(k) - \mathbf{s}_{i_{m_1}}(k) | \mathbf{S}_{k,l,J_1}
\end{pmatrix}
\]

Proof: The proof is a direct extension of the proof of Theorem 1 in [21] for the case of MIMO systems. Indeed we first observe that the dimension of \( \mathbf{D}_l^i(\mathbf{h}^{(i_1)}) \) is \( M(l+1) \times (l-J_i+1) \).

We have the following.

i) \( l < J_1 \): By the definition of \( \mathbf{\tilde{S}}_{k,l,J_1} \) (see (13)) we see that for all \( i, j, 1 \leq i \leq P, 0 \leq j \leq L_i : \mathbf{s}_i(k + l - j) \in \mathbf{\tilde{S}}_{k,l,J_1} \). Therefore, (17) implies that \( \mathbf{T}(k + l) | \mathbf{S}_{k,l,J_1} = \mathbf{T}(k + l) \) and \( \mathbf{E}_{k,l,J_1} = 0 \).

ii) \( J_1 \leq l < J_2 \): Let \( 0 \leq l < J_2 \). All vectors of the form \( \mathbf{s}_i(k+q-j) \) where \( 1 \leq s \leq m_1 \) and \( q - l + J_1 \leq j \leq q \) do not belong to \( \mathbf{\tilde{S}}_{k,l,J_1} \). Therefore, using again (17), we have that
\[
\mathbf{T}(k + q) - \mathbf{T}(k + q) | \mathbf{S}_{k,l,J_1} = \sum_{1 \leq s \leq m_1} \sum_{j=q-l+1}^{q} \mathbf{h}_{i_s}(j) \\
\times \left( \mathbf{s}_{i_s}(k + q - j) - \mathbf{s}_{i_s}(k + q - j) | \mathbf{S}_{k,l,J_1} \right).
\]

The assertion follows if we take into account that, by definition, \( \mathbf{h}_{i_s}(j) = 0 \) for \( j \leq 0 \) or \( j > J_1 \) and write the previous equation in matrix form.

iii) For all \( n : 1 \leq n \leq r, \) we have \( \mathbf{J}_{n-1} \leq l < \mathbf{J}_n \) and \( 0 \leq q \leq l \). Similar arguments lead to
\[
\mathbf{T}(k + q) - \mathbf{T}(k + q) | \mathbf{S}_{k,l,J_1} = \sum_{1 \leq s \leq m_1} \sum_{j=q-l+1}^{q} \mathbf{h}_{i_s}(j) \\
\times \left( \mathbf{s}_{i_s}(k + q - j) - \mathbf{s}_{i_s}(k + q - j) | \mathbf{S}_{k,l,J_1} \right) + \cdots + \sum_{s=1}^{m_1} \sum_{j=q-l+1}^{q} \mathbf{h}_{i_s}(j) \\
\times \left( \mathbf{s}_{i_s}(k + q - j) - \mathbf{s}_{i_s}(k + q - j) | \mathbf{S}_{k,l,J_1} \right).
\]

Note that by definition for all \( i, s, j : 1 \leq i \leq r, \) \( 1 \leq s \leq m_i \) and \( j \leq 0 \) or \( j > J_i, \mathbf{h}_{i_s}(j) = 0 \). If we write (22) in matrix form, we prove (19).

iv) Proceeding in a similar fashion, we prove (20).

Next, we establish the ranks of the projection error matrices \( \mathbf{E}_{k,l,J_1} \).

**Theorem 3:** The following statements are true.

i) For any \( l : l < J_1 \).
\[
\text{rank}(\mathbf{E}_{k,l,J_1}) = 0.
\]

ii) For \( l : J_1 \leq l < J_2 \).
\[
\text{rank}(\mathbf{E}_{k,l,J_1}) = (l - J_1 + 1)m_1.
\]

iii) For \( n : 3 \leq n \leq r \) and \( l : J_{n-1} \leq l < J_n \).
\[
\text{rank}(\mathbf{E}_{k,l,J_1}) = \sum_{i=1}^{n-1} (l - J_i + 1)m_i.
\]

iv) For \( l : J_r \leq l \leq L \).
\[
\text{rank}(\mathbf{E}_{k,l,J_1}) = \sum_{i=1}^{r} (l - J_i + 1)m_i.
\]

Proof: For each \( n : 1 \leq n \leq r \), we define the matrix \( \mathbf{\Theta}_n \) as
\[
\mathbf{\Theta}_n = \left( \mathbf{H}_l(\mathbf{G}_1) \cdots | \mathbf{H}_l(\mathbf{G}_n) \right)
\]

Using assumption A3) and induction on \( n \) we easily prove that \( \mathbf{\Theta}_n \) has full column rank for each \( n : 1 \leq n \leq r \).

For notational purposes set \( J_{r+1} = L + 1 \). According to Theorem 2, for all \( n : 2 \leq n \leq r + 1 \) and all \( l : J_{n-1} \leq l < J_n \)
\[
\mathbf{E}_{k,l,J_1} = \mathbf{H}_l^1(\mathbf{G}_1)\mathbf{\tilde{S}}_{k,l,J_1-1}^1 + \cdots + \mathbf{H}_l^{n-1}(\mathbf{G}_n-1)\mathbf{\tilde{S}}_{k,l,J_1-1}^{n-1}.
\]

Equation (26) is written compactly as
\[
\mathbf{E}_{k,l,J_1} = \mathbf{\Theta}_n^{-1} \left( \begin{array}{c}
\mathbf{\tilde{S}}_{k,l,J_1-1}^1 \\
\vdots \\
\mathbf{\tilde{S}}_{k,l,J_1-1}^{n-1}
\end{array} \right).
\]

We know that \( \mathbf{\Theta}_n^{-1} \) has full column rank. Therefore
\[
\text{rank}(\mathbf{E}_{k,l,J_1}) = \text{rank} \left( \begin{array}{c}
\mathbf{\tilde{S}}_{k,l,J_1-1}^1 \\
\vdots \\
\mathbf{\tilde{S}}_{k,l,J_1-1}^{n-1}
\end{array} \right).
\]

Note that each block entry \( \mathbf{\tilde{S}}_{k,l,J_1-1}^i \) is a projection error. More precisely, it has the form
\[
\mathbf{\tilde{S}}_{k,l,J_1-1}^i = \begin{pmatrix}
\mathbf{s}_1(k + l - J_i) - \mathbf{s}_1(k + l - J_i) | \mathbf{S}_{k,l,J_1} \\
\vdots \\
\mathbf{s}_i(k) - \mathbf{s}_i(k) | \mathbf{S}_{k,l,J_1} \\
\vdots \\
\mathbf{s}_{i_{m_1}}(k + l - J_i) - \mathbf{s}_{i_{m_1}}(k + l - J_i) | \mathbf{S}_{k,l,J_1} \\
\vdots \\
\mathbf{s}_{i_{m_1}}(k) - \mathbf{s}_{i_{m_1}}(k) | \mathbf{S}_{k,l,J_1}
\end{pmatrix}
\]
Using arguments similar to the ones used in the proof of Theorem 1, we conclude that for all \( i, s, j : 1 \leq i \leq r, 1 \leq s \leq m_i \), \( 0 \leq j \leq l - J_i \), \( s_{m_i}(k + j) \) is orthogonal to the vectors belonging to \( S_{i,j} \). Therefore, \( s_{m_i}(k + j) S_{i,j} = 0 \).

As a consequence,

\[
\hat{S}_{i,k,l-J_i} = \begin{pmatrix}
S_{i,k}(k + l - J_i) \\
\vdots \\
S_{i,k}(k) \\
\vdots \\
S_{m_i}(k + l - J_i) \\
\vdots \\
S_{m_i}(k)
\end{pmatrix}, \quad (30)
\]

As it was the case in Theorem 1, we notice that the rows of the matrix

\[
\left( \begin{array}{c}
\hat{S}_{1,k,l-J_k} \\
\vdots \\
\hat{S}_{m_i,k,l-J_{k-m_i}}
\end{array} \right) \quad (31)
\]

are orthogonal with respect to the inner product induced by correlation in the Hilbert space of zero mean random vectors with finite second moments. This is due to the i.i.d. and mutual independence assumptions stated in A1). Therefore, the rows of the above matrix are linearly independent with probability one.

The theorem then follows from (28) and (30) by a straightforward counting argument.\(^2\)

The definition of the p-norm of a matrix and the previous theorem imply the following.

**Corollary 1:** The following statements are true.

i) For any \( l : J_1 \leq l \leq L \), \( \| E_{k,l} \|_p = 0 \).

ii) For any \( l : J_1 \leq l \leq L \), \( \| E_{k,l} \|_p > 0 \).

**Proof:** The proof is obvious for case i).

We shall prove case ii) by contradiction. Assume that there exists an \( l : J_1 \leq l \leq L \) such that \( \| E_{k,l} \|_p = 0 \). This implies that \( E_{k,l} = 0 \). The latter as well as (27) imply that the rows of the matrix

\[
\left( \begin{array}{c}
\hat{S}_{1,k,l-J_k} \\
\vdots \\
\hat{S}_{m_i,k,l-J_{k-m_i}}
\end{array} \right) \quad (32)
\]

are linearly dependent, something that cannot be true according to the proof of the previous theorem.\(\square\)

In the noiseless case, Theorem 3 indicates how to compute the system’s different orders, as well as the number of subsystems that attain it in a straightforward manner.

Indeed, starting with \( l = 0 \), we compute \( \text{rank}(E_{k,l}) \) and we increase \( l \) by one until \( \text{rank}(E_{k,l}) > 0 \). Equation (23) suggests that this value of \( l \) equals the smaller of the orders \( J_1 \). Moreover, it also gives the number \( m_1 \) of the subsystems that attain it.

Having determined \( m_1 \), we increase \( l \) in steps of one. As long as \( l < J_2 \), \( \text{rank}(E_{k,l}) \) remains a multiple of \( m_1 \). When this stops to hold, (24) suggests that \( l = J_2 \). At this point, having computed \( m_1 \), \( J_k \) we use (24) to determine the number \( m_2 \) of the subsystems that attain the order \( J_2 \). We continue increasing \( l \). As long as \( l < J_3 \), \( \text{rank}(E_{k,l}) \) increases by \( m_1 + m_2 \) each time \( l \) increases by one. Again, when this stops to hold, (24) suggests that \( l = J_3 \). At this point, having computed \( J_1, m_1, J_2, m_2 \) we use (24) to determine the number \( m_3 \) of the subsystems that have order equal to \( J_3 \).

We keep increasing \( l \) by one until we reach \( L \) and proceed in the same way, using (24) and (25) to determine \( J_k, m_k \) for all \( i : 1 \leq i \leq r \). When \( l = L \), we compute the number \( P \) of input signals as:

\[
P = \sum_{i=1}^{r} m_i. \quad (33)
\]

The above approach is effective in a noiseless setting, where the rank of a matrix can be computed in a straightforward manner simply by counting the number of its nonzero singular values. In the presence of noise, however, the problem of rank determination can be proved difficult to solve because singular values that should be zero in theory could become small, but not necessarily zero.

We address this issue in the next section, where we describe the implementation of our algorithm, using the concept of effective (numerical) matrix rank (Criterion 1) and a property implied by Corollary 1 (Criterion 2). Moreover, we provide a simpler version of the algorithm that applies when it is a priori known that the system is SIMO.

\section{Algorithm Implementation}

Two versions of the algorithm are supplied. The first applies in the general MIMO/SIMO case. It can be used without any information regarding the application type. The second version is preferable in cases where it is a priori known that the system is SIMO, as it reduces significantly computation time by avoiding unnecessary SVDs.

Noise can be managed via the use of two criteria. The first one is employed to determine the effective (numerical) rank of a matrix [27], [28, ch. 5].

Criterion 1: Let \( B \) be an \( m \times n \) matrix perturbed by noise and denote \( \beta_1, \beta_2, \ldots, \beta_r, \gamma \) as \( \text{rank}(m, n) \), its singular values with \( \beta_1 \geq \beta_2 \geq \cdots \geq \beta_r \). The effective rank of \( B \) is given by the integer \( t \) in the range \( 0 \leq t \leq r \) for which the ratio \( \gamma_t = \beta_t/\beta_{t+1} \) is maximized. The second criterion is suggested by Corollary 1.

Criterion 2: For fixed \( k \in N \), the ratio \( \gamma_k = \| E_{k,l} \|_p/\| E_{k,l+1} \|_p \) is minimized for \( l = J_k - 1 \), where \( J_k \) is the smallest of the system orders.

Next, we present the two versions of our algorithm. We denote by \( L \) and \( Q \) an upper bound of the system orders and the number of output data vectors used, respectively.
A. MIMO and SIMO Systems

The proposed algorithm consists of the following steps.

1) For a fixed \( w \geq w_0 \) and for all \( l, 0 \leq l \leq L \), define the overall data matrices \( Z_{w,l} \) as

\[
Z_{w,l} = \begin{pmatrix}
x(2w + l + 1) & \cdots & x(Q) \\
x(w + l + 2) \\
x(w + l + 1) & \cdots \\
x(w + 1) & \cdots \\
x(w) & \cdots \\
x(1)
\end{pmatrix}.
\]

2) For all \( l, 0 \leq l \leq L \), use SVD and Criterion 1 to compute the effective rank of \( Z_{w,l} \). Reconstruct \( Z_{w,l} \), using the singular values that correspond to its effective rank and the associated left and right singular vectors to compute the matrix \( \tilde{Z}_{w,l} \). From \( \tilde{Z}_{w,l} \), compute the matrices \( \tilde{Y}_{w,l} \) and \( \tilde{D}_{w,l} = \left( \frac{\tilde{F}_{w,l}}{\tilde{P}_{w,l}} \right) \), that correspond to the matrices \( Y_{w,l} \) and \( D_{w,l} = \left( \frac{F_{w,l}}{P_{w,l}} \right) \) defined in the previous step. Step 2) will be referred as the denoising step.

3) For all \( l, 0 \leq l \leq L \), compute an orthogonal base of \( \mathcal{R}(\tilde{D}_{w,l}) \), \( \{v_1, \ldots, v_f\} \). Use Theorems 1 and 2 to compute \( E_{w,l} \) in a way analogous to that given in [21], that is

\[
E_{w,l} = \tilde{Y}_{w,l} - \tilde{Y}_{w,l} V V^\top, \quad V = \begin{pmatrix} v_1 \\ \vdots \\ v_f \end{pmatrix}.
\]

4) For all \( 0 \leq l \leq L \), compute \( \hat{\delta}_{w,l} \) and use Criterion 2 to find \( \hat{J}_l \). Then use Criterion 1 to compute the effective rank of \( E_{w,J_l} \). Conclude that this is the number \( m_1 \) of subsystems that have order \( J_l \).

5) For all \( J_l < l \leq L \), use Criterion 1 to find the effective rank of the matrices \( E_{w,l} \). Having determined \( J_l, m_1 \), use Theorem 3 to compute the pairs \( J_l, m_1 \), for all \( 1 < n \leq r \).

6) Determine the number \( P \) of system inputs from (33).

B. SIMO Systems

If it is a priori known that the system is SIMO the algorithm is simplified as follows.

Steps 1–3) Steps 1–3) remain the same as with the version described above.

Step 4) For all \( 0 \leq l \leq L \), compute \( \hat{\delta}_{w,l} \) and use Criterion 2 to find the order \( J_l \) of the system.

Remark: In [21], it is proved that for a SIMO system, in the noiseless case, \( \operatorname{rank}(Z_{w,l}) = 2w + l + L_\omega + 1 \), where \( L_\omega \) is the true order of the system. Simulations showed that for some cases of low SNR values the estimated effective rank of \( Z_{w,l} \) could be either too low or two high introducing significant error when \( \hat{Z}_{w,l} \) is computed. Therefore, in our simulations we slightly modified Step 2) for SIMO systems, so that if for a given \( l \) the effective rank \( \hat{J}_l \) of \( Z_{w,l} \) is less than \( 2w \) or greater than \( 2w + L \), the following actions are taken:

- if \( 0 > l \), \( \hat{J}_l \) is set to \( \hat{J}_l - 1 \), where \( \hat{J}_l - 1 \) is the effective rank of \( Z_{w,J-1} \);
- if \( 0 = l \), \( \hat{J}_l \) is set to \( 2w + 1 \).

VI. SIMULATION EXAMPLES

In this section, we present the simulation experiments that we have performed for the SIMO as well as the MIMO case. We assume that the system’s output is given by the equation

\[
y(k) = x(k) + \eta(k)
\]

where \( x(k) \) is given by (1) and \( \eta(k) = [\eta_1(k) \cdots \eta_M(k)]^\top \) is an \( M \times 1 \) spatial-temporal white noise vector sequence. Furthermore, the additive noise \( \eta(k) \) is a zero mean stationary sequence uncorrelated to the inputs. For all \( i, 1 \leq i \leq M, \sigma_i^2 \) is the finite variance of \( \eta_i(k) \). Let \( Q \) denote the number of output vectors used. The signal-to-noise ratio is defined as

\[
\text{SNR} = 10\log_{10} \frac{\frac{1}{Q} \sum_{k=1}^{Q} \|y(k)\|^2 - M\sigma^2}{M\sigma^2}.
\]

Even though Corollary 1 is true for any matrix norm, \( \|E_{w,k}\|_\infty \) was used in simulations, because the experiments showed that it achieves better performance over other known matrix norms. Finally, we mention that whenever the term denoising is used in the following, it refers to Step 2) of versions A and B of the proposed algorithm, as given in the previous section. Moreover, we shall use the terms output vector and output sample interchangeably to refer to the \( M \times 1 \) output data vector collected at the receiver.

A. MIMO Systems

Simulations entailing three different systems have been performed. In all cases, the input sequences are quadrature phase-shift keyed (QPSK), while 100 independent Monte Carlo runs were obtained per system, number of output samples, and SNR value. We depict the percentage of successful order detection related to SNR and the number of output vectors used. In all examples, the predictor’s size \( w \) equals to the upper bound of the system’s orders \( L = 10 \).

Example 1: In this case we have simulated a two-input six-output MIMO system. The two subsystems \( h_1, h_2 \), with orders \( L_1 = 3 \) and \( L_2 = 5 \), respectively, are shown in Tables I and II. According to Criterion 2 and Theorem 3 the minimum order \( J_1 \) of the system should be equal to 3, while for \( l = 3, 4, \ldots, 10 \) the corresponding ranks of \( E_{w,l} \) should be detected as 1, 2, 4, 6, 8, 10, 12, 14.

Simulation results are presented in Table III. The algorithm’s success rate exceeds 90% using 1000 output vectors for \( \text{SNR} \geq 20 \) dB. For SNR values between 20 and 25 dB, increasing the amount of output samples from 400 to 800 or 1000 improves significantly the algorithm’s performance. For SNR values greater or equal to 27-dB high success rates are achieved even using 400 output samples.
Example 2: In this example, another two-input six-output MIMO system was tested. Subsystems $h_1$, $h_2$ comprising the total system are shown in Tables IV and V, respectively. Both their orders $L_1$, $L_2$ are equal to 4. According to Criterion 2 and Theorem 3, the minimum order $J_1$ of the system should be equal to 4, while for $l = 4, 5, \ldots, 10$ the corresponding ranks of $E_{u,l}$ should be detected as 2, 4, 7, 10, 13, 16, 19, 22. The results are shown in Table VIII.

In comparison to the previous examples, we observe a performance degradation that can be explained by the extra amount of information that has to be detected by the algorithm due to the introduction of the new subsystem.

B. SIMO Systems

In this subsection, simulations of SIMO systems are performed. The proposed method is compared with existing...
SIMO system order estimation techniques, namely the ones presented in [21] (J-LSS method), [22] (Liavas et al. method), and the MDL algorithm [23]. In order to study the effect of denoising, both J-LSS and the proposed algorithm were tested using denoised as well as non denoised data matrices. To make fair comparisons, all the algorithms were tested against the same data sequence. That is, in each Monte Carlo run, all the algorithms worked on the same received data. A success rate threshold of 90% was employed for comparison.

Four systems were simulated the first two of which are artificial. The remaining are real microwave channels given in http://spib.rice.edu/spib/microwave.html. In all cases, the input sequences were assumed to be QPSK and 200 output samples were used. One hundred independent Monte Carlo runs were run per system and SNR value. Finally, in all the experiments the upper bound of system order $L$ was set equal to 15 and the predictor’s size $w$ was set equal to $L$.

### Table V

**MIMO Systems: Example 2. Subchannel Coefficients of Subsystem 2**

<table>
<thead>
<tr>
<th>$h_2(0)$</th>
<th>$h_2(1)$</th>
<th>$h_2(2)$</th>
<th>$h_2(3)$</th>
<th>$h_2(4)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.04 + 0.017i</td>
<td>0.256 + 0.107i</td>
<td>0.226 + 0.120i</td>
<td>-0.147 + 0.110i</td>
<td>-0.113 - 0.141i</td>
</tr>
<tr>
<td>0.10 + 0.12i</td>
<td>0.265 + 0.208i</td>
<td>0.204 + 0.119i</td>
<td>-0.143 + 0.106i</td>
<td>0.110 - 0.162i</td>
</tr>
<tr>
<td>0.07 + 0.13i</td>
<td>0.274 + 0.210i</td>
<td>0.182 + 0.118i</td>
<td>-0.139 + 0.101i</td>
<td>0.134 - 0.184i</td>
</tr>
<tr>
<td>0.04 + 0.13i</td>
<td>0.282 + 0.212i</td>
<td>0.161 + 0.117i</td>
<td>-0.134 - 0.119i</td>
<td>0.258 - 0.267i</td>
</tr>
<tr>
<td>0.04 + 0.127i</td>
<td>0.126 + 0.127i</td>
<td>0.226 + 0.220i</td>
<td>-0.247 + 0.210i</td>
<td>-0.213 - 0.241i</td>
</tr>
<tr>
<td>0.20 + 0.12i</td>
<td>0.365 + 0.308i</td>
<td>0.104 + 0.119i</td>
<td>-0.243 + 0.206i</td>
<td>0.210 - 0.262i</td>
</tr>
</tbody>
</table>

### Table VI

**MIMO Systems: Example 2. Percentage of Successful Order Detection**

<table>
<thead>
<tr>
<th>Number of Samples</th>
<th>25 dB</th>
<th>27 dB</th>
<th>28 dB</th>
<th>30 dB</th>
<th>35 dB</th>
<th>40 dB</th>
<th>50 dB</th>
</tr>
</thead>
<tbody>
<tr>
<td>400</td>
<td>0%</td>
<td>8%</td>
<td>17%</td>
<td>88%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>800</td>
<td>12%</td>
<td>62%</td>
<td>98%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>1000</td>
<td>20%</td>
<td>87%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
</tbody>
</table>

### Table VII

**MIMO Systems: Example 3. Subchannel Coefficients of Subsystem 3**

<table>
<thead>
<tr>
<th>$h_3(0)$</th>
<th>$h_3(1)$</th>
<th>$h_3(2)$</th>
<th>$h_3(3)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.15 + 0.1187i</td>
<td>0.1747 + 0.1825i</td>
<td>0.1241 + 0.1469i</td>
<td>-0.1148 - 0.1151i</td>
</tr>
<tr>
<td>0.124 + 0.1261i</td>
<td>0.1934 + 0.1844i</td>
<td>0.1905 + 0.1099i</td>
<td>-0.1160 - 0.1175i</td>
</tr>
<tr>
<td>0.199 + 0.1336i</td>
<td>0.1120 + 0.1136i</td>
<td>0.1169 + 0.1192i</td>
<td>-0.1147 - 0.1099i</td>
</tr>
<tr>
<td>0.173 + 0.1411i</td>
<td>0.0307 + 0.0255i</td>
<td>0.0233 + 0.0285i</td>
<td>-0.1134 - 0.0423i</td>
</tr>
<tr>
<td>0.15 + 0.0187i</td>
<td>0.1747 + 0.1825i</td>
<td>0.1241 + 0.1469i</td>
<td>-0.1448 - 0.1751i</td>
</tr>
<tr>
<td>0.124 + 0.0261i</td>
<td>0.1934 + 0.1844i</td>
<td>0.0905 + 0.0099i</td>
<td>-0.1560 - 0.1975i</td>
</tr>
</tbody>
</table>

### Table VIII

**MIMO Systems: Example 3. Percentage of Successful Order Detection**

<table>
<thead>
<tr>
<th>Number of Samples</th>
<th>25 dB</th>
<th>27 dB</th>
<th>30 dB</th>
<th>33 dB</th>
<th>35 dB</th>
<th>40 dB</th>
<th>50 dB</th>
</tr>
</thead>
<tbody>
<tr>
<td>400</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>16%</td>
<td>73%</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>800</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>60%</td>
<td>97%</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>1000</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>4%</td>
<td>73%</td>
<td>99%</td>
<td>100%</td>
</tr>
</tbody>
</table>

### Table IX

**MIMO Systems: Example 1. Subchannel Coefficients**

<table>
<thead>
<tr>
<th>$h(0)$</th>
<th>$h(1)$</th>
<th>$h(2)$</th>
<th>$h(3)$</th>
<th>$h(4)$</th>
<th>$h(5)$</th>
<th>$h(6)$</th>
<th>$h(7)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.4</td>
<td>0.18</td>
<td>0.22</td>
<td>0.5</td>
<td>0.4</td>
<td>0.16</td>
<td>0.43</td>
<td>0.38</td>
</tr>
<tr>
<td>0.25</td>
<td>0.31</td>
<td>0.16</td>
<td>0.43</td>
<td>0.20</td>
<td>0.61</td>
<td>0.43</td>
<td>0.22</td>
</tr>
</tbody>
</table>

**Example 1:** A two-output SIMO system is considered. The subchannel coefficients are shown in Table IX, while simulation results are presented in Table X. The first and the last taps of the system concentrate 11% and 9.6% of the total power, respectively. Therefore, the estimated system order should equal 7.

Simulations showed that the proposed algorithm achieves a success rate greater than 90% for SNR $\geq$ 20 dB. The corresponding SNR level for the Liavas et al. method is 26 dB, for MDL 21 dB, and for J-LSS’s 35 dB. Denoising improved J-LSS, giving an over 90% success rate for an SNR values equal to or
greater 28 dB. Applying no denoising slightly improves the proposed method giving a success rate of over 90% for an SNR value of 19 dB.

Example 2: A two-output SIMO system is considered. The subchannel coefficients are shown in Table XI. It is easily verified that even though the last tap of the second subchannel is zero, the power of the corresponding tap of the first subchannel approximates 24% of the total channel energy. Therefore, blind order estimation should detect the maximum order of the subchannels, which equals 4. The results are shown in Table XII.

Simulations showed that the proposed algorithm achieves a success rate greater or equal to 90% for SNR $\geq 12$ dB. The corresponding SNR level for Liavas et al. method is 20 dB, for MDL 15 dB and for J-LSS is 26 dB. Denoising improves J-LSS, giving an over 90% success rate for SNR values greater than 20 dB. Applying no denoising improves the proposed method giving a success rate of over 90% for SNR values equal to or greater than 8 dB.

Examples 3 and 4: Two FIR microwave radio channels given in http://spib.rice.edu/spib/microwave.html were tested, namely channel 10 and channel 13. Channels were oversampled by a factor of 2. Their characteristic is that their true impulse response is very long, but they can be partitioned into a “significant part” that contains the “large” channel terms and concentrates the most of the channel power and a “nonsignificant” part that comprises of “small” leading and/or trailing terms. The order of the significant part is the “effective channel order” as defined in [22]. Intuitively, the satisfactory effective order estimate of these channels is two, giving three taps per subchannel. Simulation results are shown in Table XIII and Table XIV.

Interpreting the results for channel 10, we should notice first that the 14th, 15th, and 16th taps concentrate approximately 99% of the total channel power, while all the remaining leading and trailing taps account only for approximately 1%. Therefore, the effective order of channel 10 should equal 2.

Simulations showed that the proposed method’s success percentage exceeds 90% when SNR $\geq 14$ dB, while the corresponding SNR level for Liavas et al. method is 22 dB and for MDL 15 dB. The J-LSS success rate was not satisfactory. However, applying denoising in J-LSS dramatically improved its performance, giving a success rate exceeding 90% for SNR $\geq 22$ dB. MDL’s performance proved poor for SNR $\geq 35$ dB, verifying that it tends to overmodel systems at high SNR values, as opposed to the proposed algorithm that kept a 100% success rate even for high SNR levels. For the same range of SNR values, the proposed algorithm proved superior to the Liavas et al. method. Finally, we notice that when the proposed algorithm failed to determine the channel’s effective order it undermodeled it by one, giving one as the estimated order.

In case of channel 13, we mention that the 22nd, 23rd, and 24th taps concentrate approximately 99% of the total channel power, while all the remaining leading and trailing taps concentrate approximately the remaining 1%. Therefore, the effective order of channel 13 should equal 2.

Simulations showed that the proposed algorithm achieves a success rate greater than 90% for SNR $\geq 17$ dB. The corresponding SNR level for Liavas et al. method is 27 dB and for MDL 16 dB. As it was the case with channel 10, J-LSS’s performance was not satisfactory, but again it was improved by denoising. For SNR $\geq 30$ dB MDL showed poor performance, due to overmodeling. On the contrary, the proposed method kept the 100% success rate even for high SNR levels. Similar overmodeling behavior proved true for J-LSS, that is for SNR $\geq 40$ dB its performance dropped significantly even
though denoising was used. For SNR values ranging between 10 and 20 dB the proposed method achieved superior performance compared to Liavas et al. algorithm. Furthermore, for SNR values between 10 and 20 dB, denoising improves the performance of the new algorithm, gaining 4 dB in achieving a success rate greater than or equal to 90%. As with the previous case, when the algorithm failed to determine the channel’s effective order it undermodeled it by one, computing it as 1 rather than 2.

C. Comments

The performance of the proposed algorithm in the presence of noise relies heavily on the validity of Criteria 1 and 2. Among them, Criterion 1, that is the determination of the effective rank of a matrix perturbed by noise, is the most sensitive. In [27], various criteria are given that estimate the effective rank of a perturbed matrix using its singular values. All but one out of these use threshold values that either do not appear to be based on any explicit analytical expressions but are selected on an ad hoc basis, or lower and upper bounds for them can be derived analytically assuming known noise statistics. The only criterion among those presented in [27] that does not use either an ad hoc threshold value or a threshold with upper and lower bounds based on known noise statistics, determines the numerical rank $t$ of a perturbed matrix $B$ from its singular values $\beta_1 \geq \beta_2 \geq \cdots \geq \beta_t$ as the index $t$ for which $\beta_t \gg \beta_{t+1}$. This criterion has been considered for the first time in [28, ch. 5]. It is straightforward to see that it is equivalent to Criterion 1, used by the proposed method. That is, it is equivalent to determining the effective rank $t$ of a perturbed matrix $B$ as the index $t$ for which the ratio $\beta_t / \beta_{t+1}$ is maximized. This is the reason it is also referred as the “ratio test” in [27], [28], and [30].

It is well known that the squares of the singular values of a matrix $B$ are the corresponding eigenvalues of the matrix $BB^T$. Therefore, maximizing the ratio of successive singular values of $B$ is equivalent to maximizing the ratio of the corresponding eigenvalues of the matrix $BB^T$. Thus, Criterion 1 is equivalent to those given in [9, vol. 1, p. 101–102] and [31]. It is also equivalent to the one given in [22] by Liavas et al., provided that there exists at least one $i$ for which the successive eigenvalue ratio $\lambda_i / \lambda_{i+1} \geq 3$. As it is stated in [9, vol. 1, p. 102], there are cases for low SNR values for which the “ratio test” proves more convenient to apply than the criterion given in [22].

Simulation experiments showed the following.

- For MIMO systems, the validity of Criterion 2 was established even for low SNR values. On the other hand,
the rank pattern suggested by Theorem 3 is harder to
detect in the presence of noise.
• For SIMO systems the proposed method outperforms
Liavas et al. algorithm in all cases at least by 6 dB.
MDL achieves the 90% success rate at about the same
SNR levels with the proposed method, with an excep-
tion occurring in Example 2, where it proved worse by
3 dB. For the microwave channels tested in Examples
3 and 4 MDL’s performance proved poor for high SNR
values due to overmodeling, while the proposed algo-
rithm achieved a 100% success rate. The new method
outperforms the J-LSS algorithm by at least 14 dB, while
J-LSS performed poorly in the case of true microwave
channels.
• Numerous simulations conducted for MIMO systems
showed that increasing the number of outputs improves
the algorithm’s performance. This implies additional
hardware complexity and increases computational cost.
On the contrary, two outputs proved sufficient for SIMO
systems.
• Denoising improved dramatically the algorithm’s per-
f ormance for MIMO systems. In Examples 2 and 3,
using 1000 output vectors and no denoising, the method
achieves a success rate greater or equal to 90% for
SNR levels of 37–38 and 41–42 dB, respectively. The
 corresponding levels when denoising is applied drop
down to 27–28 and 32–33 dB, respectively.
• For SIMO systems, denoising had a mixed effect. In case
of Examples 1 and 2, where all taps of the system have
significant power, the performance was reduced because
finding the effective rank of $Z_{w,x}$ took away significant
data information. In the case of Examples 3 and 4, that
are real microwave channels where the system’s energy
is concentrated in a few taps but there also exist nonzero
head and tail taps, denoising improved the method. Fi-
nally, we should mention the performance boost that
was achieved when denoising was applied at the J-LSS
method.

• When the algorithm misses the correct value of the
system order for the case of SIMO systems this results
in system undermodeling in the vast majority of cases.
Misses in the MIMO case can be interpreted as the false
detection either of systems attaining nonexistent orders
or of more systems that attain a specific order than they
really are.
• In all our simulation experiments, $w$ was chosen equal to
$L$, the upper bound of system’s orders. This is because,
as noticed in [21], for a fixed data length, large values of
$w$ imply fewer columns in the algorithm’s data matrices
that correspond to smaller sample size when for example
projections on subspaces are evaluated. Therefore, we
tried to keep $w$ as small as possible.
• We have conducted extensive simulation ex-
periments on the microwave channels given in
http://spib.rice.edu/spib/microwave.html. Most of these
channels, (12 out of a total of 15), are
“good shaped.” This means that their significant part (the part that
contains most of the channel power) consists of a small
number of successive taps. Applying the proposed
method to all these channels, we saw that the estimated
effective channel length varied, depending on the
channel, from two to four taps that correspond to over
95% of the total channel power. As it was the case for
channels 10 and 13, simulations veri-
fied the superiority
of the proposed method’s performance against all the
other methods.
However, there are a few channels in
http://spib.rice.edu/spib/microwave.html such as
3, 14, and 15 (3 out of a total of 15) that are “bad
shaped.” This means that most of the channel’s power
is not contained in a number of successive taps as it
was the case with “good shaped” channels. Rather,
there exist groups of successive taps that concentrate
a significant amount of the channel’s power as well as
sporadic taps of significant power. An example of a “bad
shaped” channel is channel 3. For this channel, kernel

<table>
<thead>
<tr>
<th>SNR (dB)</th>
<th>LIAVAS ET AL.</th>
<th>MDL</th>
<th>J-LSS WITH NO DENOISING</th>
<th>PROPOSED ALGORITHM WITH NO DENOISING</th>
<th>PROPOSED ALGORITHM WITH NO DENOISING</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>47%</td>
</tr>
<tr>
<td>15</td>
<td>0%</td>
<td>80%</td>
<td>0%</td>
<td>14%</td>
<td>82%</td>
</tr>
<tr>
<td>16</td>
<td>0%</td>
<td>99%</td>
<td>0%</td>
<td>23%</td>
<td>85%</td>
</tr>
<tr>
<td>17</td>
<td>0%</td>
<td>100%</td>
<td>1%</td>
<td>30%</td>
<td>90%</td>
</tr>
<tr>
<td>20</td>
<td>0%</td>
<td>99%</td>
<td>3%</td>
<td>59%</td>
<td>95%</td>
</tr>
<tr>
<td>21</td>
<td>0%</td>
<td>100%</td>
<td>7%</td>
<td>60%</td>
<td>92%</td>
</tr>
<tr>
<td>25</td>
<td>39%</td>
<td>100%</td>
<td>18%</td>
<td>88%</td>
<td>100%</td>
</tr>
<tr>
<td>27</td>
<td>97%</td>
<td>94%</td>
<td>22%</td>
<td>98%</td>
<td>100%</td>
</tr>
<tr>
<td>30</td>
<td>100%</td>
<td>4%</td>
<td>26%</td>
<td>94%</td>
<td>100%</td>
</tr>
<tr>
<td>40</td>
<td>100%</td>
<td>0%</td>
<td>0%</td>
<td>14%</td>
<td>100%</td>
</tr>
<tr>
<td>50</td>
<td>100%</td>
<td>0%</td>
<td>0%</td>
<td>6%</td>
<td>100%</td>
</tr>
</tbody>
</table>
taps 0 to 9 are approximately zero, tap 10 accounts for 29% of the total channel power, tap 11 for 58.5%, tap 12 for 2.6%, tap 13 for 1%, while taps 13 to 37 are "small." However, tap 38 is a significant, sporadic tap, accounting for approximately 5% of the total channel power while the remaining terms of the channel impulse response are again "small" tail taps. Simulation experiments that were run using the proposed as well as all the other methods indicated one as the estimated effective channel order. The two taps found (taps 10 and 11) account for 87.5% of the total channel power, while the remaining 12.5% missed makes equalization poor for this channel, when $\text{SNR} \leq 40 \text{ dB}$, as it is stated in [22].

- The proposed algorithm performs $L+1$ SVD operations in Step 2), $L+1$ SVD operations in Step 3), and $L-J+1$ SVD operations in Steps 4) and 5) in case of MIMO systems. For SIMO systems, it performs $L+1$ SVD operations in Step 2) and $L+1$ SVD operations in Step 3), since SVD calculations associated with Steps 4) and 5) are not performed. To calculate its complexity we take into account the cost of an SVD decomposition in flops as given in [29, p. 254], as well as the number of flops required for matrix multiplications in Steps 2) (due to the reconstruction of $Z_{ud}$ and 3). Thus, we conclude that the computational complexity of the proposed algorithm is $O(M^3Q^2L(L+1)^4+M^3Q^2L^3(L+1)+M^3Q^2(L+1)^3(L-J+1))$ for MIMO systems. This reduces to $O(M^3Q^2L(L+1)^4+M^3Q^2L^3(L+1)+M^3Q^2L^3(L+1))$ for SIMO systems. In the above, $M$ is the number of outputs and $Q$ is the number of output data samples used. This suggests that the oversampling rate or the number of diverse sensors $M$ used at the receiver should be kept small to avoid heavy computations. In addition, the less output samples used and the tighter the estimate of the system orders’ upper bound $L$ is, the less expensive computations become.

VII. Conclusion

In this paper, a novel algorithm is proposed for blindly estimating the order as well as the number of inputs of FIR MIMO systems. In addition, the method can be applied to SIMO systems. It is shown by simulations that satisfactory results are achieved even when a small number of output data vectors is used. The proposed method can be combined with several identification algorithms that benefit from the knowledge of the system’s exact order(s) (such as those given, for example, in [2], [8], and [19]) to identify the corresponding channels. Work in progress aims toward reducing the method’s complexity.

### References


Diamantis Kotoulas (S’00) was born in Larissa, Greece, on February 20, 1967. He received the Diploma of Engineering degree in computer engineering and information science from the Polytechnic School of the University of Patras, Patras, Greece, in 1989 and the M.Sc. degree in computer and information science from the Department of Computing and Information Science, Queen’s University, Kingston, ON, Canada, in 1993.

From 1996 to 1997, he was with the Access Network and Wireless Communications Department, Intracom S.A., Athens, Greece, where he worked on wireless systems design. He is currently with the Department of Access Network Planning, Division of Telecommunication Network Planning, Hellenic Telecommunications Organization, Maroussi, Athens, Greece. His research interests include blind system identification, higher order statistics, and digital signal processing.

Panos Koukoulas (M’02) was born in Athens, Greece, on November 17, 1965. He received the B.Sc. degree in physics, the M.S. degree in electronics and communications, and the Ph.D. degree in informatics all from the University of Athens, Greece, in 1989, 1993, and 1997, respectively.

From 1997 to 2000, he was with the Access Network and Wireless Communications Department, Intracom S.A., Athens, Greece, where he worked on aspects of digital modem design. He is currently with the Department of Radar and Automation of the Hellenic Civil Aviation Authority, Athens, Greece. His research interests include nonlinear system identification, higher order statistics, and digital signal processing.

Nicholas Kalouptsidis (M’82–SM’85) was born in Athens, Greece, on September 13, 1951. He received the B.Sc. degree in mathematics (with highest honors) from the University of Athens, Athens, Greece, in 1973 and the M.S. and Ph.D. degrees in systems science and mathematics from Washington University, St. Louis, MO, in 1975 and 1976, respectively.

He has held visiting positions at Washington University, St. Louis, MO; Politecnico di Torino, Turin, Italy; Northeastern University, Boston, MA; and CNET Lannion, France. He has been an Associate Professor and Professor with the Department of Physics, University of Athens. In fall 1998, he was a Clyde Chair Professor with the School of Engineering, University of Utah, Salt Lake City. He is currently a Professor with the Department of Informatics and Telecommunications, University of Athens. He is the author of the textbook *Signal Processing Systems: Theory and Design* (New York: Wiley, 1997) and coeditor, with S. Theodoridis, of the book *Adaptive System Identification and Signal Processing Algorithms* (Englewood Cliffs, NJ: Prentice-Hall, 1993). His research interests are in system theory and signal processing.