
Volume 87 Issue 6 June 2007 ISSN 0165-1684
ELSEVIER
SIGNAL
PROCESSING
An International Journal
A publication of the European Association for Signal Processing (EURASIP)
EURASIP

Available online at www.sciencedirect.com
ScienceDirect
Signal Processing 87 (2007) 1434–1447
SIGNAL PROCESSING
www.elsevier.com/locate/sigpro

Some further results on blind identification of MIMO FIR channels via second-order statistics

Jun Fang[a,*], A. Rahim Leyman[b], Yong Huat Chew[b], Huiping Duan[c]

[a]*Department of Electrical and Computer Engineering, National University of Singapore 119260, Singapore*
[b]*Institute for Infocomm Research, A*STAR, 21 Heng Mui Keng Terrace, Singapore 119613, Singapore*
[c]*School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798, Singapore*

Received 1 May 2006; received in revised form 26 October 2006; accepted 13 December 2006
Available online 23 December 2006

Abstract

In this paper, we consider the problem of blind multiple-input multiple-output (MIMO) finite impulse response (FIR) channel identification driven by spatially correlated signals. The second-order statistics (SOS) of the input sources are assumed known *a priori*. It is shown that under certain specified conditions, the MIMO FIR channel can be completely identified using the second-order statistics of the channel output. A SOS-based method is proposed and the proof for the uniqueness of the system solution is provided. As a special case, our proposed method can still entirely identify the MIMO channel even if the input source signals are spatially and temporally uncorrelated, given that the channel orders corresponding to each pair of users are different from each other. Extensive numerical simulation results are included to illustrate the performance of the proposed algorithm.

Keywords: Blind channel estimation; Spatially correlated sources; Multiple-input multiple-output systems

1. Introduction

Blind identification of multiple-input multiple-output (MIMO) finite impulse response (FIR) channel arises in a wide variety of communication and signal processing applications, which include speech enhancement, wireless mobile communications and brain signal analysis. Thus far, numerous SOS-based techniques [1–7] have been proposed within such a framework. Nevertheless, they usually assume that the input sources are mutually independent or, at least, uncorrelated. In contrast, blind channel estimation driven by spatially correlated sources has not received much attention. Spatially correlated sources may indeed occur in practice. For example, the nonlinear single-input multiple-output (SIMO) channels can be reformulated into multiple-input linear systems in which the additional inputs are nonlinear functions of the signal of interest (the details of this reformulation can be referred to [8]). Clearly, in this case, the inputs of this reformulated linear MIMO channel may be correlated with each

*Corresponding author. Tel.: +65 94768196.
E-mail addresses: junfang@nus.edu.sg (J. Fang), larahim@i2r.a-star.edu.sg (A.R. Leyman), chewyh@i2r.a-star.edu.sg (Y.H. Chew), duan0002@ntu.edu.sg (H. Duan).

0165-1684/$ - see front matter
doi:10.1016/j.sigpro.2006.12.011

other. Specifically, this reformulated MIMO channel can be written as follows:

$$\mathbf{x}(n) = \sum_{i=1}^{p} \sum_{l=0}^{L_i} \mathbf{h}_i(l) s_i(n-l) + \mathbf{w}(n), \tag{1}$$

where $s_1(n) \triangleq a(n)$ is exactly the input signal to the nonlinear channels and also called as "linear kernel"; the terms $s_i(n) = f_i(a(n), a(n-1), \ldots)$ for $i \in \{2, \ldots, p\}$ are nonlinear functions of $a(\cdot)$ and also called as "nonlinear kernels"; $\{\mathbf{h}_i(l)\}$ are $q \times 1$ multichannel vectors; $\mathbf{x}(n)$ and $\mathbf{w}(n)$ represent the received data and the additive noise, respectively. In this case, if the statistical information of the input signal $a(n)$ and the functions $f_i(\cdot)$ generating the nonlinearities are known *a priori*, then the second-order statistics of the reformulated inputs to the MIMO channel are available. It is noted that both [8,9] were presented under such a framework described by Eq. (1). In [8], the authors proposed a deterministic method that exploited the channel order disparity between the linear kernel and nonlinear kernels. In fact, the techniques of [8] only resolve the kernel which has the largest channel order, irrespective of whether this kernel is a linear kernel or a nonlinear one. In the event that there are many kernels with maximum channel order, [8] has to resort to higher order methods to equalize the channel. On the other hand, the work in [9] shows that the linear kernel is resolvable under the right conditions imposed on the statistics of the signals $a(n)$ and $s_i(n)$, without the need to adhere to the particular channel order condition required by [8]. A SOS-based approach was put forward to determine the equalizers for the i.i.d. input signals $\{a(n)\}$.

In this paper, we address the problem of blind MIMO FIR channel identification driven by spatially correlated signals. A closed-form solution is proposed for blind MIMO FIR system identification by utilizing the estimated channel output autocorrelation matrices and the knowledge of the source autocorrelation matrices. As compared to [8,9], the problem we addressed here is more general because our goal is to identify the entire MIMO channel and recover all the source signals, which is different from the works [8,9] that focus on extracting and equalizing only one source signals. It is noted that, in our case, the terms $s_i(n)$ for $i \in \{2, \ldots, p\}$ in Eq. (1) may not necessarily be the functions of $s_1(n)$. In fact, even if we would only consider the MIMO models resulted from the nonlinear SIMO channels, our work has its own advantages over [8,9] in the following two aspects. Firstly, the particular channel order condition required by [8] is no longer necessary for our proposed method to identify and equalize the channel. Secondly, unlike the proposed algorithm in [9] which is only specified for the i.i.d. input signals $\{a(n)\}$, our proposed method applies to both i.i.d. and colored input signals $\{a(n)\}$. As a special case of our work, our proposed method can entirely identify the MIMO channel even when the input source signals are spatially and temporally uncorrelated, given that the channel orders corresponding to each pair of users are different. This result is different from most existing SOS-based methods that can only identify such a channel up to an unknown unitary matrix.

This paper is organized as follows. In Section 2, the MIMO system model and some basic assumptions are introduced. Next, in Section 3, we present our method for blind channel identification driven by spatially correlated sources. The channel identifiability conditions are investigated and an original proof for the uniqueness of the system solution is provided. In Section 4, we extend our method to the case of spatially and temporally uncorrelated input sources. The computational complexity of our proposed method is discussed in Section 5. Finally, in Section 6, numerical simulation results are presented to demonstrate the performance of the proposed algorithm.

We adopt the following notations throughout this paper. The notations $[\cdot]^{\mathrm{T}}$, $[\cdot]^{*}$, $[\cdot]^{\mathrm{H}}$ and $[\cdot]^{\dagger}$ stand for transpose, complex conjugate, Hermitian transpose and the Moore–Penrose pseudo-inverse, respectively. $E[\cdot]$ represents the mathematical expectation. $\|\mathbf{A}\|$ ($\|\mathbf{a}\|$) denotes the Frobenius norm (vector 2-norm) of matrix $\mathbf{A}$ (vector $\mathbf{a}$). $\otimes$ represents the Kronecker product; vec($\mathbf{A}$) is an operator creates a column vector from $\mathbf{A}$ by stacking the column vectors of $\mathbf{A}$ from left to right. The symbol $\mathbf{J}_n$ stands for the $n \times n$ one-lag down shift square matrix whose first sub-diagonal entries below the main diagonal are unity, whereas all remaining entries are zero; $\mathbf{I}_n$ denotes the $n \times n$ identity matrix. Let $\mathbf{A}[r_1 : r_2, c_1 : c_2]$ denote the sub-matrix of $\mathbf{A}$ from r_1th row to r_2th row and from c_1th column to c_2th column. $\mathbb{N}$, $\mathbb{I}$ and $\mathbb{C}$ denote the set of natural, integer and complex numbers, respectively. $\mathbb{C}^{n \times m}$ and $\mathbb{C}^n$ denote the set of $n \times m$ matrices and the set of n-dimensional column vectors with complex entries, respectively.

For clarity of this paper, we also summarize the dimensions and ranks of the matrices that will be used in this paper, where $d = ((N+1)p + \sum_{i=1}^{p} L_i)$.

Matrices	Dimensions	Ranks
$\mathscr{H}$	$(N+1)q \times d$	Full column rank
$\mathbf{R}_s[0]$	$d \times d$	Full rank
$\mathbf{R}_x[0]$	$(N+1)q \times (N+1)q$	d
$\mathbf{R}_s[k]$ for $k \neq 0$	$d \times d$	Depends on source statistics
$\mathbf{R}_x[k]$ for $k \neq 0$	$(N+1)q \times (N+1)q$	Depends on rank of $\mathbf{R}_s[k]$
$\mathbf{P}$	$d \times d$	Full rank
$\mathbf{G}$	$(N+1)q \times d$	Full column rank
$\mathbf{Q}$	$d \times d$	Unitary matrix to be estimated
$\bar{\mathbf{R}}_s[k]$	$d \times d$	Depends on rank of $\mathbf{R}_s[k]$
$\bar{\mathbf{R}}_x[k]$	$d \times d$	Depends on rank of $\mathbf{R}_s[k]$

2. System model and basic assumptions

Consider a noisy linear MIMO channel with p inputs, $s_i(n), i \in \{1, 2, \ldots, p\}$, and q outputs $\mathbf{x}(n) \triangleq [x_1(n) \ \cdots \ x_q(n)]^{\mathrm{T}}$ as follows:

$$\mathbf{x}(n) = \sum_{i=1}^{p} \sum_{l=0}^{L_i} \mathbf{h}_i(l) s_i(n-l) + \mathbf{w}(n), \tag{2}$$

where $\{\mathbf{h}_i(l)\}$ denotes the multichannel filter corresponding to the ith user, and L_i represents the channel order corresponding to the ith user. By stacking the channel output vector $\mathbf{x}(n)$ and defining $\vec{\mathbf{x}}(n) \triangleq [\mathbf{x}^{\mathrm{T}}(n) \ \mathbf{x}^{\mathrm{T}}(n-1) \ \ldots \ \mathbf{x}^{\mathrm{T}}(n-N)]^{\mathrm{T}}$, $\vec{\mathbf{s}}_i(n) \triangleq [s_i(n) \ s_i(n-1) \ \cdots \ s_i(n-N-L_i)]^{\mathrm{T}}$ and $\vec{\mathbf{w}}(n) \triangleq [\mathbf{w}^{\mathrm{T}}(n) \ \mathbf{w}^{\mathrm{T}}(n-1) \ \ldots \ \mathbf{w}^{\mathrm{T}}(n-N)]^{\mathrm{T}}$, we can rewrite Eq. (2) as

$$\vec{\mathbf{x}}(n) = \sum_{i=1}^{p} \mathscr{H}_i \vec{\mathbf{s}}_i(n) + \vec{\mathbf{w}}(n) = \mathscr{H} \vec{\mathbf{s}}(n) + \vec{\mathbf{w}}(n), \tag{3}$$

where $\mathscr{H}_i \in \mathbb{C}^{(N+1)q \times d_i}$ is a block Toeplitz matrix written as follows with $d_i \triangleq N + L_i + 1$

$$\mathscr{H}_i \triangleq \begin{bmatrix} \mathbf{h}_i(0) & \ldots & \mathbf{h}_i(L_i) & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{0} & \mathbf{h}_i(0) & \ldots & \mathbf{h}_i(L_i) & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \ldots & \mathbf{0} & \mathbf{h}_i(0) & \ldots & \mathbf{h}_i(L_i) \end{bmatrix},$$

$$\mathscr{H} \triangleq [\mathscr{H}_1 \ \mathscr{H}_2 \ \cdots \ \mathscr{H}_p],$$

$$\vec{\mathbf{s}}(n) \triangleq [\vec{\mathbf{s}}_1^{\mathrm{T}}(n) \ \vec{\mathbf{s}}_2^{\mathrm{T}}(n) \ \cdots \ \vec{\mathbf{s}}_p^{\mathrm{T}}(n)]^{\mathrm{T}}.$$

Some basic assumptions are adopted throughout the paper: (A1) The number of sources is known *a priori*, and there are more outputs than inputs, i.e. $q > p$. (A2) Channel is irreducible and column-reduced. (A3) The channel order corresponding to each source is assumed to be known *a priori*. (A4) The sources are zero-mean wide-sense stationary colored signals or white signals and their input statistics (include autocorrelation of the single source signals and cross-correlation between any two source signals) are available. (A5) The source correlation matrix $\mathbf{R}_s[0]$ is positive definite, where $\mathbf{R}_s[0] \triangleq E[\vec{\mathbf{s}}(n)\vec{\mathbf{s}}^{\mathrm{H}}(n)]$. (A6) Additive noises $\mathbf{w}(n)$ are spatially and temporally white noises with same variance, and they are statistically independent of the sources. As a consequence of A2, the MIMO channel matrix $\mathscr{H}$ is full column rank if the stack number N is chosen to satisfy $N + 1 \geqslant \sum_{i=1}^{p} L_i$ (see [10]). In the sequel, we assume that $\mathscr{H}$ is full column rank.

3. Proposed channel identification method for spatially correlated sources

We begin by defining the source autocorrelation matrices with delay lag k as follows:

$$\mathbf{R}_s[k] \triangleq E[\vec{\mathbf{s}}(n)\vec{\mathbf{s}}^{\mathrm{H}}(n-k)]. \tag{4}$$

Since the additive noises are assumed spatially and temporally white with same variance, and statistically independent of the sources, the autocorrelation matrices of the channel output $\vec{\mathbf{x}}(n)$ can be written as

$$\mathbf{R}_x[k] \triangleq E[\vec{\mathbf{x}}(n)\vec{\mathbf{x}}^{\mathrm{H}}(n-k)] = \begin{cases} \mathscr{H}\mathbf{R}_s[k]\mathscr{H}^{\mathrm{H}} + \sigma_w^2 \mathbf{J}_{(N+1)q}^{kq}, & k>0, \\ \mathscr{H}\mathbf{R}_s[k]\mathscr{H}^{\mathrm{H}} + \sigma_w^2 \mathbf{I}_{(N+1)q}, & k=0, \\ \mathscr{H}\mathbf{R}_s[k]\mathscr{H}^{\mathrm{H}} + \sigma_w^2 (\mathbf{J}_{(N+1)q}^{kq})^{\mathrm{T}}, & k<0, \end{cases}$$

where σ_w^2 denotes the noise variance. In practice, the influence of the noise can be minimized by removing the noise contribution from the estimated autocorrelation matrices of the channel output. Considering $k=0$, we have

$$\mathbf{R}_x[0] = \mathscr{H}\mathbf{R}_s[0]\mathscr{H}^{\mathrm{H}} + \sigma_w^2 \mathbf{I}_{(N+1)q},$$

where the source correlation matrix $\mathbf{R}_s[0]$ is symmetric and positive definite under assumption A5. Hence the noise variance σ_w^2 can be estimated as the smallest eigenvalues of $\mathbf{R}_x[0]$ and then subtracted from any estimated autocorrelation matrix $\mathbf{R}_x[k]$ to provide our proposed algorithm with denoised autocorrelation estimates. The number of smallest eigenvalues of $\mathbf{R}_x[0]$ can be determined by applying the MDL criterion [11]. In doing in this way, we can ignore the noise effect in the following derivations. Also, this helps simplify the presentation of the proposed channel identification method. Henceforth, we can rewrite the autocorrelation matrices of the received data $\vec{\mathbf{x}}(n)$ as follows:

$$\mathbf{R}_x[k] \triangleq E[\vec{\mathbf{x}}(n)\vec{\mathbf{x}}^{\mathrm{H}}(n-k)] = \mathscr{H}\mathbf{R}_s[k]\mathscr{H}^{\mathrm{H}}. \tag{5}$$

In the following, we will show that, given that certain identifiability conditions are satisfied, the channel convolution matrix $\mathscr{H}$ can be completely identified up to a scalar factor by utilizing the estimated channel output autocorrelation matrices $\mathbf{R}_x[k], k \in \{0, \pm 1\}$ and the knowledge of $\mathbf{R}_s[k]$, $k \in \{0, \pm 1\}$.

It is clear that from A5, we have the following relationship

$$\mathbf{R}_s[0] = \mathbf{P}\mathbf{P}^{\mathrm{H}}, \tag{6}$$

where $\mathbf{P}$ is an invertible matrix. Utilizing the eigenvalue decomposition: $\mathbf{R}_s[0] \triangleq \mathbf{U}_s\mathbf{D}_s\mathbf{U}_s^{\mathrm{H}}$, we can write

$$\mathbf{P} = \mathbf{U}_s\mathbf{D}_s^{1/2}\mathbf{M}, \tag{7}$$

where $\mathbf{M}$ can be any arbitrary unitary matrix to satisfy Eq. (6). However, a properly chosen $\mathbf{M}$ (or $\mathbf{P}$) plays a key role in our later derivation. The choice of this unitary matrix $\mathbf{M}$ is illustrated as follows. Let $\bar{\mathbf{R}}_s[k] \triangleq \mathbf{P}^{-1}\mathbf{R}_s[k]\mathbf{P}^{-\mathrm{H}}$, we have the following by substituting Eq. (7) for $\mathbf{P}$:

$$\begin{aligned} \bar{\mathbf{R}}_s[k] &= \mathbf{P}^{-1}\mathbf{R}_s[k]\mathbf{P}^{-\mathrm{H}} \\ &= \mathbf{M}^{\mathrm{H}}\mathbf{D}_s^{-1/2}\mathbf{U}_s^{\mathrm{H}}\mathbf{R}_s[k]\mathbf{U}_s\mathbf{D}_s^{-1/2}\mathbf{M} \\ &= \mathbf{M}^{\mathrm{H}}\ddot{\mathbf{R}}_s[k]\mathbf{M}, \end{aligned} \tag{8}$$

where $\ddot{\mathbf{R}}_s[k] \triangleq \mathbf{D}_s^{-1/2}\mathbf{U}_s^{\mathrm{H}}\mathbf{R}_s[k]\mathbf{U}_s\mathbf{D}_s^{-1/2}$ can be computed since the input statistics are known *a priori*. The unitary matrix $\mathbf{M}$ is chosen to transform $\ddot{\mathbf{R}}_s[1]$ into an upper triangular matrix, i.e. $\bar{\mathbf{R}}_s[1]$ is upper triangular. This transformation is known as the Schur decomposition [12] in which the unitary matrix $\mathbf{M}$ is uniquely determined. The matrix $\mathbf{M}$ (or $\mathbf{P}$) selected in this way will provide us a key advantage in the proof of the solution uniqueness. We will show this in Sections 3.1 and 3.2.

Now we consider the eigenvalue decomposition of $\mathbf{R}_x[0]$:

$$\mathbf{R}_x[0] \triangleq \begin{bmatrix} \mathbf{U}_{x,1} & \mathbf{U}_{x,2} \end{bmatrix} \begin{bmatrix} \mathbf{D}_{x,1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{U}_{x,1}^{\mathrm{H}} \\ \mathbf{U}_{x,2}^{\mathrm{H}} \end{bmatrix}. \tag{9}$$

Let $\mathbf{G} \triangleq \mathbf{U}_{x,1}\mathbf{D}_{x,1}^{1/2}$. Since $\mathbf{R}_x[0] = \mathscr{H}\mathbf{P}\mathbf{P}^{\mathrm{H}}\mathscr{H}^{\mathrm{H}} = \mathbf{G}\mathbf{G}^{\mathrm{H}}$, it is clear that we have

$$\mathbf{G} = \mathscr{H}\mathbf{P}\mathbf{Q}, \tag{10}$$

where $\mathbf{Q}$ is an unknown unitary matrix to be determined. To resolve this unknown matrix, we have to further explore the relationship imposed on this unitary matrix $\mathbf{Q}$. Recalling Eq. (5) for $k \in \{\pm 1\}$, we have

$$\begin{aligned} \mathbf{G}^{\dagger}\mathbf{R}_x[k](\mathbf{G}^{\dagger})^{\mathrm{H}} &= \mathbf{Q}^{\mathrm{H}}\mathbf{P}^{-1}\mathscr{H}^{\dagger}\mathscr{H}\mathbf{R}_s[k]\mathscr{H}^{\mathrm{H}}(\mathscr{H}^{\dagger})^{\mathrm{H}}\mathbf{P}^{-\mathrm{H}}\mathbf{Q} \\ &= \mathbf{Q}^{\mathrm{H}}\mathbf{P}^{-1}\mathbf{R}_s[k]\mathbf{P}^{-\mathrm{H}}\mathbf{Q}. \end{aligned} \tag{11}$$

For notational convenience, let $\bar{\mathbf{R}}_x[k] \triangleq \mathbf{G}^{\dagger}\mathbf{R}_x[k](\mathbf{G}^{\dagger})^{\mathrm{H}}$. Thus Eq. (11) can be rewritten as

$$\bar{\mathbf{R}}_x[k] = \mathbf{Q}^{\mathrm{H}}\bar{\mathbf{R}}_s[k]\mathbf{Q}, \quad k \in \{\pm 1\} \tag{12}$$

and furthermore

$$\bar{\mathbf{R}}_x[k]\mathbf{Q}^{\mathrm{H}} = \mathbf{Q}^{\mathrm{H}}\bar{\mathbf{R}}_s[k], \quad k \in \{\pm 1\}. \tag{13}$$

Eq. (13) defines the relationship the unknown unitary matrix $\mathbf{Q}$ must satisfy. Since $\bar{\mathbf{R}}_x[k]$ can be estimated from the second-order statistics of the channel output, $\bar{\mathbf{R}}_s[k]$ can also be computed from the knowledge of the input statistics, the above equation can thus be used to estimate the unknown unitary matrix $\mathbf{Q}$. By utilizing the property of

Kronecker product, we rewrite Eq. (13) as

$$\mathbf{I}_d \otimes \bar{\mathbf{R}}_x[k] \cdot \mathrm{vec}(\mathbf{Q}^{\mathrm{H}}) = \bar{\mathbf{R}}_s^{\mathrm{T}}[k] \otimes \mathbf{I}_d \cdot \mathrm{vec}(\mathbf{Q}^{\mathrm{H}}), \quad k \in \{\pm 1\}, \tag{14}$$

where $d = \sum_{i=1}^{p} d_i$ is the dimension of the matrices $\mathbf{Q}$, $\bar{\mathbf{R}}_s[k]$ and $\bar{\mathbf{R}}_x[k]$. By defining $\mathbf{q} \triangleq \mathrm{vec}(\mathbf{Q}^{\mathrm{H}})$, we may estimate the unknown unitary matrix $\mathbf{Q}$ by the following criterion:

$$\hat{\mathbf{q}} = \arg\min_{\|\mathbf{q}\|=1} \left\| \begin{bmatrix} \mathbf{I}_d \otimes \bar{\mathbf{R}}_x[1] - \bar{\mathbf{R}}_s^{\mathrm{T}}[1] \otimes \mathbf{I}_d \\ \mathbf{I}_d \otimes \bar{\mathbf{R}}_x[-1] - \bar{\mathbf{R}}_s^{\mathrm{T}}[-1] \otimes \mathbf{I}_d \end{bmatrix} \mathbf{q} \right\|^2. \tag{15}$$

The above optimization has a closed-form solution which can be obtained as the right singular vector associated with the smallest singular value. However, this criterion fails to provide the true channel estimation if the solution to Eq. (14) is not unique, i.e. there exist other non-zero vectors, $\mathbf{g}$, that are *linearly independent of* $\mathbf{q}$ and also satisfy $\mathbf{I}_d \otimes \bar{\mathbf{R}}_x[k] \cdot \mathbf{g} = \bar{\mathbf{R}}_s^{\mathrm{T}}[k] \otimes \mathbf{I}_d \cdot \mathbf{g}$ for $k \in \{\pm 1\}$. Hence we are faced with the following problem, that is, whether or not the solution to Eq. (14) is unique (up to an unknown scalar factor) and under what conditions the solution to Eq. (14) will be unique. This problem is studied in the following and we will show that, under certain identifiability conditions, the uniqueness of the solution to Eq. (14) can be established.

3.1. Property of triangular matrix

In the previous discussion, the matrix $\mathbf{M}$ (or $\mathbf{P}$) is determined via Schur decomposition to make $\bar{\mathbf{R}}_s[1]$ upper triangular. Also, it is easy to know that $\bar{\mathbf{R}}_s[-1]$ is a lower triangular matrix by exploiting the symmetry relationship: $\bar{\mathbf{R}}_s[-1] = \bar{\mathbf{R}}_s^{\mathrm{H}}[1]$. In the following, we derive an important property related to the triangular matrix. This property serves as a theoretical basis in the proof of the solution uniqueness.

Lemma 1. *Given that* $\mathbf{T} \in \mathbb{C}^{n\times n}$ *is an upper triangular matrix*, $\mathbf{Y} \in \mathbb{C}^{n\times n}$ *and we have*

$$\mathbf{TY} = \mathbf{YT} \tag{16}$$

if any pair of diagonal elements in $\mathbf{T}$ *are different from each other, i.e.* $t_{i,i} \neq t_{j,j}$ *for any* $i \neq j$, *then* $\mathbf{Y}$ *is also an upper triangular matrix.*

Proof. See Appendix A. □

3.2. Proof of the solution uniqueness and the proposed algorithm

We now proceed to prove the uniqueness of the system solution based on Lemma 1. We, firstly, prove that the solution to Eq. (13) is unique (up to a scalar factor). The problem is formulated in the following theorem.

Theorem 1. *Given the following from Eq.* (12)

$$\bar{\mathbf{R}}_x[k] = \mathbf{Q}^{\mathrm{H}} \bar{\mathbf{R}}_s[k] \mathbf{Q}, \quad k \in \{\pm 1\} \tag{17}$$

if the upper triangular matrix $\bar{\mathbf{R}}_s[1]$ *satisfies the following two identifiability conditions*: (IC1) *the diagonal entries are all unequal*; (IC2) *for any* $i \in \{2, \ldots, d\}$, *there exists at least one non-zero entry* $r_{j_1,i}$ *for* $j_1 < i$, *or for any* $i \in \{1, \ldots, d-1\}$, *there exists at least one non-zero entry* r_{i,j_2} *for* $j_2 > i$, *where* $r_{i,j}$ *denotes the* (i,j)*th element of* $\bar{\mathbf{R}}_s[1]$, *then any non-zero matrix* $\mathbf{C}$ *that satisfies the following relationship*:

$$\bar{\mathbf{R}}_x[k]\mathbf{C} = \mathbf{C}\bar{\mathbf{R}}_s[k], \quad k \in \{\pm 1\} \tag{18}$$

can be written as: $\mathbf{C} = \lambda \mathbf{Q}^{\mathrm{H}}$, *where* λ *can be any non-zero complex scalar*.

Proof. Replacing $\bar{\mathbf{R}}_x[k]$ with the right term of Eq. (17), it is easy to obtain the following from Eq. (18)

$$\mathbf{Q}^{\mathrm{H}} \bar{\mathbf{R}}_s[k] \mathbf{QC} = \mathbf{C}\bar{\mathbf{R}}_s[k] \Rightarrow \bar{\mathbf{R}}_s[k]\mathbf{QC} = \mathbf{QC}\bar{\mathbf{R}}_s[k]. \tag{19}$$

Let $\mathbf{Z} \triangleq \mathbf{QC}$, we can rewrite the above equation as

$$\bar{\mathbf{R}}_s[k]\mathbf{Z} = \mathbf{Z}\bar{\mathbf{R}}_s[k], \quad k \in \{\pm 1\}. \tag{20}$$

We now only need to prove that $\mathbf{Z} = \lambda \mathbf{I}$. Since we have $\bar{\mathbf{R}}_s[1]\mathbf{Z} = \mathbf{Z}\bar{\mathbf{R}}_s[1]$, where $\bar{\mathbf{R}}_s[1]$ is an upper triangular matrix whose diagonal entries are all unequal, recalling Lemma 1, it is easy to conclude that $\mathbf{Z}$ is an upper triangular matrix. For $\bar{\mathbf{R}}_s[-1]\mathbf{Z} = \mathbf{Z}\bar{\mathbf{R}}_s[-1]$, by exploiting the symmetry $\bar{\mathbf{R}}_s[-1] = \bar{\mathbf{R}}_s^{\mathrm{H}}[1]$, we have

$$\bar{\mathbf{R}}_s[-1]\mathbf{Z} = \mathbf{Z}\bar{\mathbf{R}}_s[-1] \Rightarrow \bar{\mathbf{R}}_s^{\mathrm{H}}[1]\mathbf{Z} = \mathbf{Z}\bar{\mathbf{R}}_s^{\mathrm{H}}[1] \Rightarrow \bar{\mathbf{R}}_s[1]\mathbf{Z}^{\mathrm{H}} = \mathbf{Z}^{\mathrm{H}}\bar{\mathbf{R}}_s[1]. \tag{21}$$

Thus $\mathbf{Z}^{\mathrm{H}}$ is also proved to be an upper triangular matrix. Therefore, we can conclude that $\mathbf{Z}$ must be a diagonal matrix. Obviously, under the identifiability condition IC2, the diagonal elements in $\mathbf{Z}$ must be equal in order to satisfy $\bar{\mathbf{R}}_s[1]\mathbf{Z} = \mathbf{Z}\bar{\mathbf{R}}_s[1]$. Hence we have $\mathbf{Z} = \lambda \mathbf{I}$ and $\mathbf{C} = \lambda \mathbf{Q}^{\mathrm{H}}$. The proof is completed here. □

Notice that Eqs. (13) and (14) can be derived from each other. This implies that the solution to

Table 1
Channel identification Algorithm I

1. Compute the eigenvalue decomposition of $\mathbf{R}_s[0]=\mathbf{U}_s\mathbf{D}_s\mathbf{U}_s^{\mathrm{H}}$. Let $\ddot{\mathbf{R}}_s[1]\triangleq\mathbf{D}_s^{-1/2}\mathbf{U}_s^{\mathrm{H}}\mathbf{R}_s[1]\mathbf{U}_s\mathbf{D}_s^{-1/2}$.
2. Compute the Schur decomposition of $\ddot{\mathbf{R}}_s[1]=\mathbf{M}\bar{\mathbf{R}}_s[1]\mathbf{M}^{\mathrm{H}}$, where $\bar{\mathbf{R}}_s[1]$ is an upper triangular matrix, $\mathbf{M}$ is a unitary matrix. Let $\mathbf{P}\triangleq\mathbf{U}_s\mathbf{D}_s^{1/2}\mathbf{M}$.
3. Compute the eigenvalue decomposition of $\mathbf{R}_x[0]$ and let $\mathbf{G}\triangleq\mathbf{U}_{x,1}\mathbf{D}_{x,1}^{1/2}$.
4. Compute $\bar{\mathbf{R}}_x[k]=\mathbf{G}^{\dagger}\mathbf{R}_x[k](\mathbf{G}^{\dagger})^{\mathrm{H}}$ for $k\in\{\pm1\}$.
5. Estimate $\mathbf{q}$ by using the criterion in Eq. (15). The unknown unitary matrix $\mathbf{Q}$ can be easily obtained by doing the inverse vec operation on $\mathbf{q}$. The channel is then estimated as $\hat{\mathscr{H}}=\mathbf{G}\hat{\mathbf{Q}}^{\mathrm{H}}\mathbf{P}^{-1}$.

Eq. (14) is also unique and the solution is a scaling constant of the "true" vector $\mathbf{q}$. Therefore $\mathbf{q}$ can be estimated by the criterion of Eq. (15). For clarity, we enumerate the steps for our channel identification procedure in Table 1.

3.3. Discussions

To guarantee that our proposed channel identification method works, two identifiability conditions IC1–IC2 are proposed and stated in Theorem 1. Since the conditions are only related to the second-order statistics of the input sources, they can be checked *a priori* to determine whether the channel identifiability conditions are satisfied or not. Also, we emphasize that IC1–IC2 are sufficient but not necessary identifiability conditions for the channel identification. For the linear MIMO channels with independent or spatially uncorrelated input signals, we usually require the spectra diversity condition, that is, the sources have sufficiently distinct power spectra or second-order statistics, to identify the channel (see [13–15]). For the case of spatially correlated input signals, the identifiability conditions IC1–IC2 do not have a very clear physical meaning. However, these two identifiability conditions IC1–IC2 imposed on the source autocorrelation matrices still suggest that, (i) the sources should have sufficiently diverse power spectra, and (ii) the sources, though spatially correlated, cannot be highly relevant with each other. To illustrate the former point, we can assume the case where the sources are independent or spatially uncorrelated, then $\bar{\mathbf{R}}_s[1]$ turns into a block diagonal matrix with each matrix on the diagonal being upper triangular. In this case, the diagonal entries of the ith matrix on the diagonal are exactly the eigenvalues of the autocorrelation matrix $\ddot{\mathbf{R}}_{s_i}[1]\triangleq\mathbf{D}_{s_i}^{-1/2}\mathbf{U}_{s_i}^{\mathrm{H}}\mathbf{R}_{s_i}[1]\mathbf{U}_{s_i}\mathbf{D}_{s_i}^{-1/2}$, where $\mathbf{R}_{s_i}[1]\triangleq E[\vec{\mathbf{s}}_i(n)\vec{\mathbf{s}}_i^{\mathrm{H}}(n-1)]$. We can see that, if only the delay lag $k\in\{\pm1\}$ is considered, the identifiability condition IC1 is almost equivalent to the spectra diversity condition proposed in [13] (the spectral diversity condition in [13] requires that for each pair of sources $\{s_i,s_j\}$, there is a correlation lag k such that $\sigma(\ddot{\mathbf{R}}_{s_i}(k))\cap\sigma(\ddot{\mathbf{R}}_{s_j}(k))=\emptyset$, where $\sigma(\mathbf{A})$ denotes the set of eigenvalues of matrix $\mathbf{A}$). The latter point that the sources should not be highly correlated can be explained as follows. If there are two or more sources highly relevant, then the autocorrelation matrix $\bar{\mathbf{R}}_s[1]$ would be rank deficient and its diagonal entries must have more than one zeros, which violates our identifiability condition IC1.

In the following, the case where the input sources are spatially and temporally uncorrelated is considered. We will show that our proposed method can still entirely identify the channel given that the channel orders corresponding to each pair of users are different from each other.

4. Channel identification for spatially and temporally uncorrelated inputs

For the spatially and temporally uncorrelated input sources, we have $\mathbf{R}_s[0]=\mathbf{I}_d$ and

$$\mathbf{R}_s[k]=\mathrm{diag}(\mathbf{R}_{s_1}[k],\mathbf{R}_{s_2}[k],\ldots,\mathbf{R}_{s_p}[k]),\quad k\in\{\pm1\} \tag{22}$$

with

$$\mathbf{R}_{s_i}[k]\triangleq E[\vec{\mathbf{s}}_i(n)\vec{\mathbf{s}}_i^{\mathrm{H}}(n-k)]=\begin{cases}\mathbf{J}_{d_i} & k=1,\\ \mathbf{J}_{d_i}^{\mathrm{T}} & k=-1.\end{cases} \tag{23}$$

We choose $\mathbf{P}$ as the identity matrix. Hence we have $\bar{\mathbf{R}}_s[k]=\mathbf{R}_s[k]$. Notice that $\mathbf{R}_s[1]$ and $\mathbf{R}_s[-1]$ are lower and upper triangular matrices, respectively. However, they do not satisfy IC1–IC2 because their diagonal entries are all zeros. Nevertheless, we will show that the channel can still be identified unambiguously from the second-order statistics of the channel output provided that another identifiability condition is satisfied. The results are formulated in the following theorem.

Theorem 2. *Given the following from Eq.* (12)

$$\bar{\mathbf{R}}_x[k] = \mathbf{Q}^{\mathrm{H}}\mathbf{R}_s[k]\mathbf{Q}, \quad k \in \{\pm 1\} \tag{24}$$

if the channel orders corresponding to each pair of sources $\{s_i, s_j\}$ *are different, i.e.* $L_i \neq L_j$, *then any non-zero matrix* $\mathbf{C}$ *that satisfies*

$$\bar{\mathbf{R}}_x[k]\mathbf{C} = \mathbf{C}\mathbf{R}_s[k], \quad k \in \{\pm 1\} \tag{25}$$

can be written as $\mathbf{C} = \mathbf{Q}^{\mathrm{H}}\mathbf{D}$, *where* $\mathbf{D} \triangleq \mathrm{diag}(\lambda_1 \mathbf{I}_{d_1}, \ldots, \lambda_p \mathbf{I}_{d_p})$, λ_i *for* $i \in \{1, \ldots, p\}$ *can be any complex scalar including zero.*

Before we proceed to prove Theorem 2, we first introduce the following lemma that exploits the properties of the one-lag down and up shift square matrices.

Lemma 2. *Given that* $\mathbf{Y} \in \mathbb{C}^{m\times n}$ *satisfies the following two equations*:

$$\text{(a) } \mathbf{J}_m\mathbf{Y} = \mathbf{Y}\mathbf{J}_n, \quad \text{(b) } \mathbf{J}_m^{\mathrm{T}}\mathbf{Y} = \mathbf{Y}\mathbf{J}_n^{\mathrm{T}} \tag{26}$$

then we have

- *If* $m = n$, *then* $\mathbf{Y} = \lambda\mathbf{I}$, *where* λ *could be any complex scalar including zero.*
- *If* $m \neq n$, *then* $\mathbf{Y} = \mathbf{0}$.

Proof. See Appendix B. □

Now we proceed to prove Theorem 2.

Proof. Substituting the right term of Eq. (24) for Eq. (25), we have

$$\mathbf{Q}^{\mathrm{H}}\mathbf{R}_s[k]\mathbf{Q}\mathbf{C} = \mathbf{C}\mathbf{R}_s[k] \Rightarrow \mathbf{R}_s[k]\mathbf{Q}\mathbf{C} = \mathbf{Q}\mathbf{C}\mathbf{R}_s[k]. \tag{27}$$

Let $\mathbf{Z} \triangleq \mathbf{Q}\mathbf{C}$, we can rewrite the above equation as

$$\mathbf{R}_s[k]\mathbf{Z} = \mathbf{Z}\mathbf{R}_s[k], \quad k \in \{\pm 1\}. \tag{28}$$

We now prove that $\mathbf{Z} = \mathrm{diag}(\lambda_1\mathbf{I}_{d_1}, \ldots, \lambda_p\mathbf{I}_{d_p})$. We partition matrix $\mathbf{Z}$ as follows:

$$\mathbf{Z} = \begin{bmatrix} \mathbf{Z}_{11} & \mathbf{Z}_{12} & \cdots & \mathbf{Z}_{1p} \\ \mathbf{Z}_{21} & \mathbf{Z}_{22} & \cdots & \mathbf{Z}_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{Z}_{p1} & \mathbf{Z}_{p2} & \cdots & \mathbf{Z}_{pp} \end{bmatrix},$$

where $\mathbf{Z}_{ij} \in \mathbb{C}^{d_i \times d_j}$. Since the matrices $\mathbf{R}_s[k]$ for $k \in \{\pm 1\}$ are block diagonal matrices, it is straightforward for us to obtain the following from Eq. (28)

$$\text{(a) } \mathbf{J}_{d_i}\mathbf{Z}_{ij} = \mathbf{Z}_{ij}\mathbf{J}_{d_j}, \quad \text{(b) } \mathbf{J}_{d_i}^{\mathrm{T}}\mathbf{Z}_{ij} = \mathbf{Z}_{ij}\mathbf{J}_{d_j}^{\mathrm{T}}. \tag{29}$$

Obviously, for the case where $d_i \neq d_j$ (note that $L_i \neq L_j$ is equivalent to $d_i \neq d_j$ since $d_i = N + L_i + 1$) for each pair of $\{i, j\}$, we can conclude that $\mathbf{Z} = \mathrm{diag}(\lambda_1\mathbf{I}_{d_1}, \ldots, \lambda_p\mathbf{I}_{d_p})$ by utilizing Lemma 2. Hence we have $\mathbf{C} = \mathbf{Q}^{\mathrm{H}}\mathbf{Z} = \mathbf{Q}^{\mathrm{H}}\mathbf{D}$. The proof is completed here. □

We now develop the corresponding algorithm for the channel identification. Let

$$\bar{\mathbf{Q}} \triangleq \mathbf{Q}^{\mathrm{H}} \triangleq [\bar{\mathbf{Q}}_1 \;\; \bar{\mathbf{Q}}_2 \;\; \cdots \;\; \bar{\mathbf{Q}}_p],$$

where $\bar{\mathbf{Q}}_i \in \mathbb{C}^{d\times d_i}$. By exploiting the block diagonal structure of $\mathbf{R}_s[k]$, the set of equations $\bar{\mathbf{R}}_x[k]\mathbf{Q}^{\mathrm{H}} = \mathbf{Q}^{\mathrm{H}}\mathbf{R}_s[k]$ for $k \in \{\pm 1\}$ can be decoupled into the following p sets of equations:

$$\bar{\mathbf{R}}_x[k]\bar{\mathbf{Q}}_i = \bar{\mathbf{Q}}_i\mathbf{R}_{s_i}[k], \quad i \in \{1, \ldots, p\}. \tag{30}$$

By invoking the property of Kronecker product, $\bar{\mathbf{Q}}_i$ is estimated as a closed-form minimizer of the following criterion by defining $\mathbf{q}_i \triangleq \mathrm{vec}(\bar{\mathbf{Q}}_i)$

$$\hat{\mathbf{q}}_i = \arg\min_{\|\mathbf{q}_i\|=1} \left\| \begin{bmatrix} \mathbf{I}_{d_i} \otimes \bar{\mathbf{R}}_x[1] - \mathbf{R}_{s_i}^{\mathrm{T}}[1] \otimes \mathbf{I}_d \\ \mathbf{I}_{d_i} \otimes \bar{\mathbf{R}}_x[-1] - \bar{\mathbf{R}}_{s_i}^{\mathrm{T}}[-1] \otimes \mathbf{I}_d \end{bmatrix} \mathbf{q}_i \right\|^2. \tag{31}$$

The unknown unitary matrix $\bar{\mathbf{Q}}$ is estimated in p parallel threads with the ith thread leading to the estimation of $\bar{\mathbf{Q}}_i$. From Theorem 2, we know that this unknown unitary matrix $\bar{\mathbf{Q}}$ is identified up to a diagonal matrix $\mathbf{D} = \mathrm{diag}(\lambda_1\mathbf{I}_{d_1}, \ldots, \lambda_p\mathbf{I}_{d_p})$, where λ_i for $i \in \{1, \ldots, p\}$ can be any *non-zero* complex scalar (note that $\lambda_i = 0$ is excluded because of the constraint $\|\mathbf{q}_i\| = 1$). Hence we have

$$\hat{\mathscr{H}} = \mathbf{G}\hat{\bar{\mathbf{Q}}}^{\mathrm{H}} = \mathbf{G}\mathbf{Q}^{\mathrm{H}}\mathbf{D}^{-1} = \mathscr{H}\mathbf{D}^{-1}. \tag{32}$$

The channel per user, $\mathscr{H}_i$ for $i \in \{1, \ldots, p\}$, is identified up to an unknown complex scalar factor. This result, of course, is valid under the condition that the channel orders corresponding to every pair of users $\{s_i, s_j\}$ are different, i.e. $L_i \neq L_j$. This identifiability condition can be further relaxed if only the desired user channels rather than the MIMO channels for all users are identified and equalized. This result is formulated in Theorem 3 and for simplicity, we consider the case where only one user channel is desired.

Theorem 3. *Given the following from Eq.* (12)

$$\bar{\mathbf{R}}_x[k] = \mathbf{Q}^{\mathrm{H}}\mathbf{R}_s[k]\mathbf{Q}, \quad k \in \{\pm 1\} \tag{33}$$

if for a desired user s_l, *we have* $L_l \neq L_i$ *for* $i \in \{1, \ldots, l-1, l+1, \ldots, p\}$, *then any non-zero matrix* $\mathbf{C}_l$ *that satisfies*

$$\bar{\mathbf{R}}_x[k]\mathbf{C}_l = \mathbf{C}_l \mathbf{R}_{s_l}[k], \quad k \in \{\pm 1\} \tag{34}$$

can be written as $\mathbf{C}_l = \lambda_l \bar{\mathbf{Q}}_l$.

Proof. See Appendix C. □

Based on Theorem 3, we can estimate $\bar{\mathbf{Q}}_l$ using the criterion of Eq. (31) with i replaced by l and the desired user channel $\mathscr{H}_l$ can be estimated as $\hat{\mathscr{H}}_l = \mathbf{G}\hat{\bar{\mathbf{Q}}}_l$. It is noted that identifying $\mathscr{H}_l$ alone allows us to compute the MMSE equalizers to recover the transmitted signals s_l by removing the intersymbol interference and canceling the multiuser interference [16]. For clarity, we also enumerate the steps for channel identification for the spatially and temporally uncorrelated sources in Table 2.

Remark. As we can see, for the spatially and temporally uncorrelated sources, the identifiability conditions IC1–IC2 proposed in previous section are no longer necessary for the complete channel identification. This point can also be corroborated by [13,14] for the spatially uncorrelated but temporally correlated sources. Since the identifiability conditions IC1–IC2 are proposed for generally correlated sources, i.e. the sources can be spatially and temporally correlated, it is no surprise that these conditions can be relaxed for the special cases where the sources are spatially and temporally uncorrelated or spatially uncorrelated but temporally correlated. This is because, for the special cases, the source autocorrelation matrices and their revised forms possess some special structure other than upper triangular structure that can be better utilized.

5. Computational complexity

In this section, we compare the computational complexity of our proposed method with that of [8,9]. Since the works [8,9] only consider the equalization of one user's signals, for fair comparison, we also consider the computational complexity of equalizing the desired one signals. For our proposed method and the method [9], the required computations consist of the following two aspects: the computation of the second-order statistics and the linear algebraic operations (eigenvalue decomposition (EVD) or singular value decomposition (SVD)) involved in algorithms implementation. The computation of the second-order statistics requires about NT_sq^2 flops for both methods, where N is the stack number, q is the number of channel outputs and T_s is the number of data samples used for statistics estimation. Now considering the linear algebraic operations, for our method, the computations are dominated by the eigenvalue decomposition of $\mathbf{R}_x[0]$ and the estimation of $\mathbf{Q}_i$, which requires computing the EVD of a matrix with dimension $(N+1)q \times (N+1)q$ and SVD of a matrix with dimension $d_id \times d_id$, respectively. Hence the flops required by our method in linear algebraic operations are of order $\mathcal{O}(((N+1)q)^3)+\mathcal{O}((d_id)^3)$. For the method [9], it needs to compute the EVD of an $(N+1)q \times (N+1)q$ matrix and a matrix power R^{d_i}, where R is an $d \times d$ matrix. Therefore the flops required by [9] in linear algebraic operations are of order $\mathcal{O}(((N+1)q)^3)+\mathcal{O}(d_id^3)$. We can see that the computational complexity of our proposed method is greater than that of [9] in the second term ($\mathcal{O}((d_id)^3)$ for our method and $\mathcal{O}(d_id^3)$ for [9]). However, for most cases, $d_i = N + L_i + 1$ is a small value and usually not greater than 10, especially when the receiver diversity is of high dimension, that is, $(N+1)q \gg d$, these two methods have similar computational complexity.

The method proposed in [8] is a deterministic method, hence do not need to compute the second-order statistics. This method requires to compute the right singular vectors of a matrix with dimension about $T_s \times q(K+1)$, where K is a parameter to be determined in the work [8]. The flops required by the method [8] are of order $\mathcal{O}(T_s(q(K+1))^2)$. Clearly, for the case where K is not large, the

Table 2
Channel identification Algorithm II

1. Compute the eigenvalue decomposition of $\mathbf{R}_x[0]$ and let $\mathbf{G} \triangleq \mathbf{U}_{x,1}\mathbf{D}_{x,1}^{1/2}$.
2. Compute $\bar{\mathbf{R}}_x[k] = \mathbf{G}^{\dagger}\mathbf{R}_x[k](\mathbf{G}^{\dagger})^{\mathrm{H}}$ for $k \in \{\pm 1\}$.
3. Estimate $\mathbf{q}_i$ by using the criterion in Eq. (31), $\bar{\mathbf{Q}}_i$ can be obtained by doing the inverse vec operation on $\mathbf{q}_i$. And the desired user channel is estimated as $\hat{\mathscr{H}}_i = \mathbf{G}\hat{\bar{\mathbf{Q}}}_i$.

Table 3
Computational complexity comparison

Algorithms	Approximate required flops
Our proposed method	$NT_sq^2 + \mathcal{O}(((N+1)q)^3) + \mathcal{O}((d_id)^3)$
The method proposed in [9]	$NT_sq^2 + \mathcal{O}(((N+1)q)^3) + \mathcal{O}(d_id^3)$
The method proposed in [8]	$\mathcal{O}(T_s(q(K+1))^2)$

method [8] has a less computational complexity than that of our method and the method [9]. We summarize the computational complexity of these three methods in Table 3.

6. Simulation results

We now present simulation results to validate the performance of our proposed algorithm. Four examples are studied in the paper. In the first example, we show the equalization performance of our proposed algorithm for the case where the sources are spatially and temporally uncorrelated, and consequently we investigate how the equalization performance hinges on the following parameters: equalization delays d_e, number of samples used for statistics estimation T_s and signal-to-noise ratio (SNR). In the rest of the examples, we consider the SIMO nonlinear channels which have been adopted by the work [8,9], and we compare our method to the SOS-based method proposed in [9] and the deterministic method presented in [8], which are named as "RS" method (R. López-Valcarce and S. Dusgupta) and "GE" method (G.B. Giannakis and E. Serpedin), respectively. Both the cases of i.i.d. input signals and colored input signals to the nonlinear channels are investigated in our examples. Also, in our simulations, the additive noise $\mathbf{w}(n)$ is taken as spatial–temporal white complex Gaussian noise with zero mean and variance σ_w^2. The SNR is defined as

$$\mathrm{SNR} = 10\log\frac{E[\|\mathscr{H}\vec{\mathbf{s}}(n)\|^2]}{E[\|\vec{\mathbf{w}}(n)\|^2]}.$$

6.1. Example one

We consider $p=2$ sources arriving at $q=3$ sensors via a multipath channel. The source signals are independent and identically distributed (i.i.d.) information sequences drawn from a 4-QPSK constellation $\mathscr{S}=\{1,-1,i,-i\}$. The channel is randomly generated as

$$\{\mathbf{h}_1(l)\} = \begin{bmatrix} 0.0572 & 0.2074 & -0.0466 & 0.1085 \\ 0.2475 & -0.1004 & 0.0213 & -0.2331 \\ 0.0968 & -0.2527 & -0.3888 & 0.2701 \end{bmatrix},$$

$$\{\mathbf{h}_2(l)\} = \begin{bmatrix} 0.2885 & 0.4926 & 0.2480 & 0 \\ 0.1714 & -0.2387 & 0.1945 & 0 \\ 0.0455 & -0.0463 & -0.0256 & 0 \end{bmatrix}.$$

It can be seen that the channel orders corresponding to these two users are different and this suffices for the complete channel identification. Once the channel has been estimated by our algorithm, we can compute the zero-forcing equalizers and the minimum mean-squared error (MMSE) equalizers, respectively, as

$$\mathscr{E}_{\mathrm{ZF}} = \hat{\mathscr{H}}^{\dagger},$$

$$\mathscr{E}_{\mathrm{MMSE}} = \mathscr{E}_{\mathrm{ZF}}(\mathbf{I} - \sigma_w^2\hat{\mathbf{R}}_x^{-1}[0]),$$

where $\hat{\mathbf{R}}_x[0]$ is the undenoised estimated autocorrelation matrix. The above expression for the MMSE equalizers was derived in [17], which is applicable for (spatially and temporally) uncorrelated sources and correlated sources. The inherent phase ambiguity of equalizers per user is removed before we perform the equalization. In the simulations, we choose stack number $N=5$. The channel order of each user is assumed known *a priori*. Results are averaged over 500 Monte Carlo runs. Fig. 1 shows the symbol error rates (SER) as a function of SNR for the MMSE equalizers with delays 1, 3, 5 and 7. $T_s = 2000$ data samples are used for the estimation of the autocorrelation matrices of the received data. Clearly, we can see that, the SER decreases as SNR increases. And the equalizers of intermediate delays are superior to those of extremal delays in performance. Fig. 2 depicts the SER of the MMSE equalizer with $d_e = 5$ using different number of data samples. As expected, the performance improves with an increasing T_s. From these results, we can see that, even for the spatially and temporally uncorrelated sources, the channel can still be completely identified by exploiting the channel order disparity.

6.2. Example two

In this example, we consider the following SIMO nonlinear channel which was adopted by the third

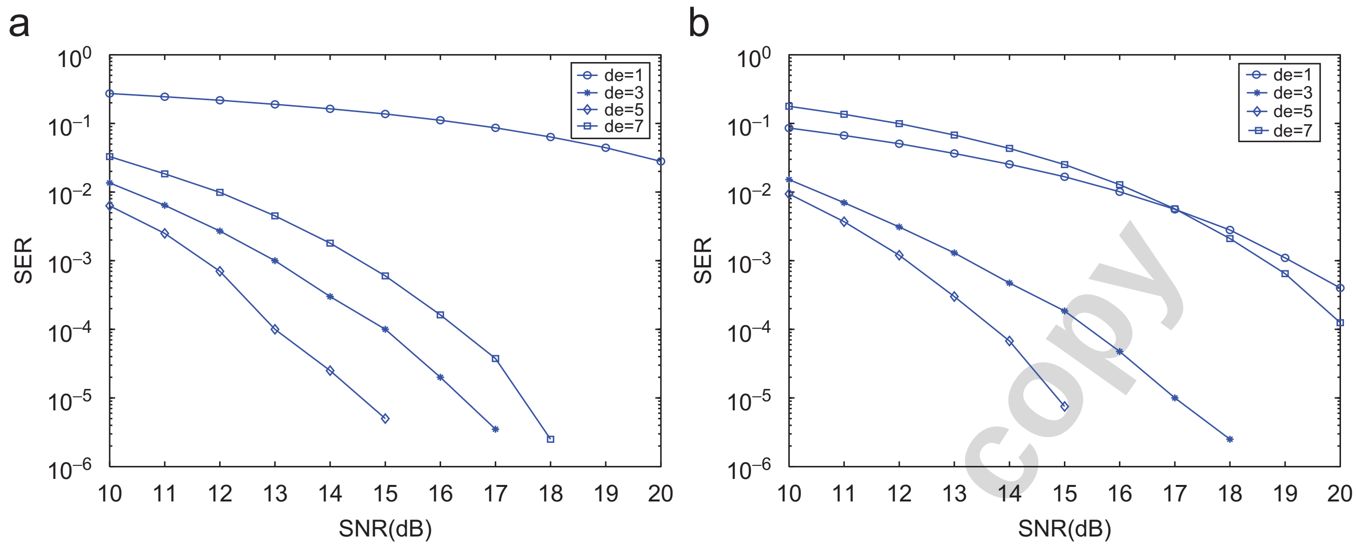

Fig. 1. Example 1: SER versus SNR for different equalization delays d_e; $T_s = 2000$.

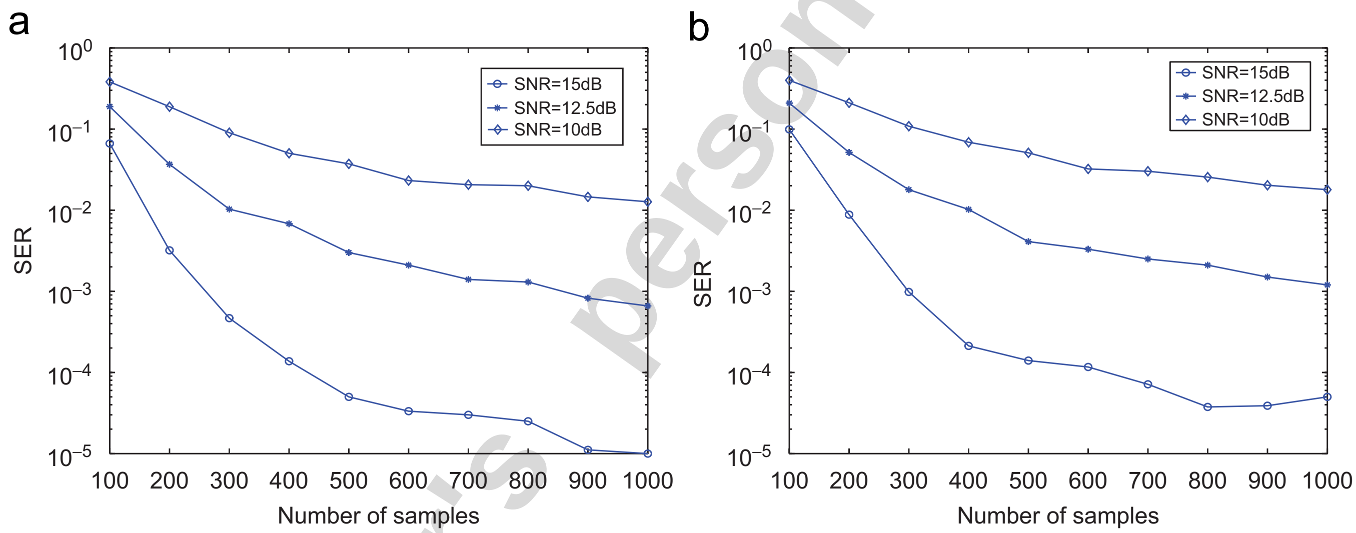

Fig. 2. Example 1: SER versus T_s for different SNR; $d_e = 5$. (a) User 1, (b) User 2.

example of [8]

$$\mathbf{x}(n) = \sum_{l=0}^{3} \mathbf{h}_1(l)a(n-l) + \sum_{l=0}^{1} \mathbf{h}_2(l)s_2(n-l) + \mathbf{w}(n),$$

where now, $s_2(n) \triangleq a(n)a(n-1)a^*(n-2)$, $a(n)$ are i.i.d. symbols drawn from the 4-QPSK constellation $\mathscr{S} = \{1, -1, i, -i\}$. It can be easily verified that $s_2(n)$ is also a temporally uncorrelated sequence, and the "two sources" $a(n)$ and $s_2(n)$ are spatially uncorrelated with different channel order. Thus our proposed method can be applied to this nonlinear channel example. We compare our proposed algorithm with the RS method [9] and the GE method [8]. In our simulations, the stack number N is chosen to be 4 for our method and the RS method, and we use the equalizer with $d_e = 7$ which achieves the best performance. For the GE method, it only provides equalizers with minimal and maximal delays, and here we employ the maximal delay $d_e = 6$ which has a better performance. Fig. 3 displays the equalization performance (for the nonlinear channel input $a(n)$) of the three algorithms as a function of SNR and T_s, respectively. From the figures, we can see that our proposed algorithm presents a clear performance advantage over RS

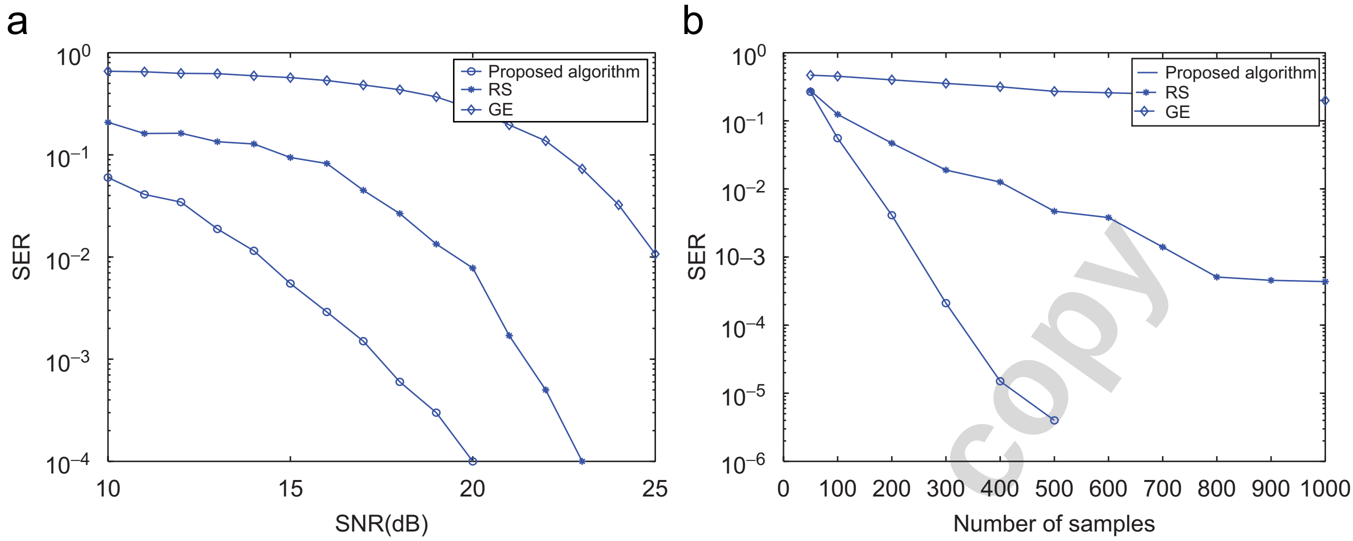

Fig. 3. Example 2: performance of respective algorithms. (a) SER versus SNR, $T_s = 500$, (b) SER versus T_s, SNR = 20 dB.

and GB methods. Also, we observed that the second-order statistics methods (RS and our proposed method) are more favorable than the deterministic method (GB) to obtain an accurate symbol estimation. It seems that using the source statistics can help to gain a stronger robustness to the noise.

6.3. Example three

We investigate the case where the input signal $\{a(n)\}$ to the nonlinear channel is colored. The SIMO nonlinear channel used by the first example of [9] is adopted.

$$\mathbf{x}(n) = \sum_{l=0}^{2} \mathbf{h}_1(l)a(n-l) + \sum_{l=0}^{1} \mathbf{h}_2(l)s_2(n-l) + \mathbf{w}(n),$$

where $s_2(n) \triangleq a(n)a(n-1)$. The input signals $\{a(n)\}$ are generated from the 4-QAM constellation $S = \{-1-i, -1+i, 1-i, 1+i\}$ according to the model which simulates a *Markov* source by implementing the transition probabilities of Table 4. The autocorrelation function of this source is given in Table 5. Clearly, these two reformulated "sources" $a(n)$ and $s_2(n)$ are spatially and temporally correlated. Also, it can be verified that the identifiability conditions IC1–IC2 required by our proposed algorithm in Section 3 are satisfied. Since the input signals are correlated, the RS method no longer applies in this example. We compare our proposed algorithm with the GE method. In our simulations, we choose the stack number $N = 3$ for our method,

Table 4
Transition probabilities for *Markov* source

$p(s_k \mid s_{k-1})$	$s_k = -1-i$	$s_k = -1+i$	$s_k = 1-i$	$s_k = 1+i$
$s_{k-1} = -1-i$	0.5	0.3	0.1	0.1
$s_{k-1} = -1+i$	0.2	0.4	0.2	0.2
$s_{k-1} = 1-i$	0.2	0.1	0.4	0.3
$s_{k-1} = 1+i$	0.1	0.2	0.2	0.5

Table 5
Autocorrelation function of the *Markov* source

$r(0)$	2.0000	$r(5)$	0.0173	$r(10)$	0.0085	$r(15)$	0.0020
$r(1)$	0.6584	$r(6)$	0.0088	$r(11)$	0.0095	$r(16)$	0.0047
$r(2)$	0.2343	$r(7)$	0.0017	$r(12)$	−0.0029	$r(17)$	0.0047
$r(3)$	0.0893	$r(8)$	0.0053	$r(13)$	−0.0068	$r(18)$	0.0027
$r(4)$	0.0384	$r(9)$	0.0082	$r(14)$	0.0031	$r(19)$	0.0039

and the equalizer with $d_e = 4$ is used for both methods. Fig. 4(a) shows the SER (for the nonlinear input signals) of the respective algorithms as a function of SNR using $T_s = 500$ data samples. Fig. 4(b) shows the variation of SER (for the nonlinear input signals) with the number of data samples T_s for SNR = 10 dB. We can see that, for both cases, our proposed algorithm clearly outperforms the GE method.

6.4. Example four

In this example, we consider the nonlinear channel which was used by the third example

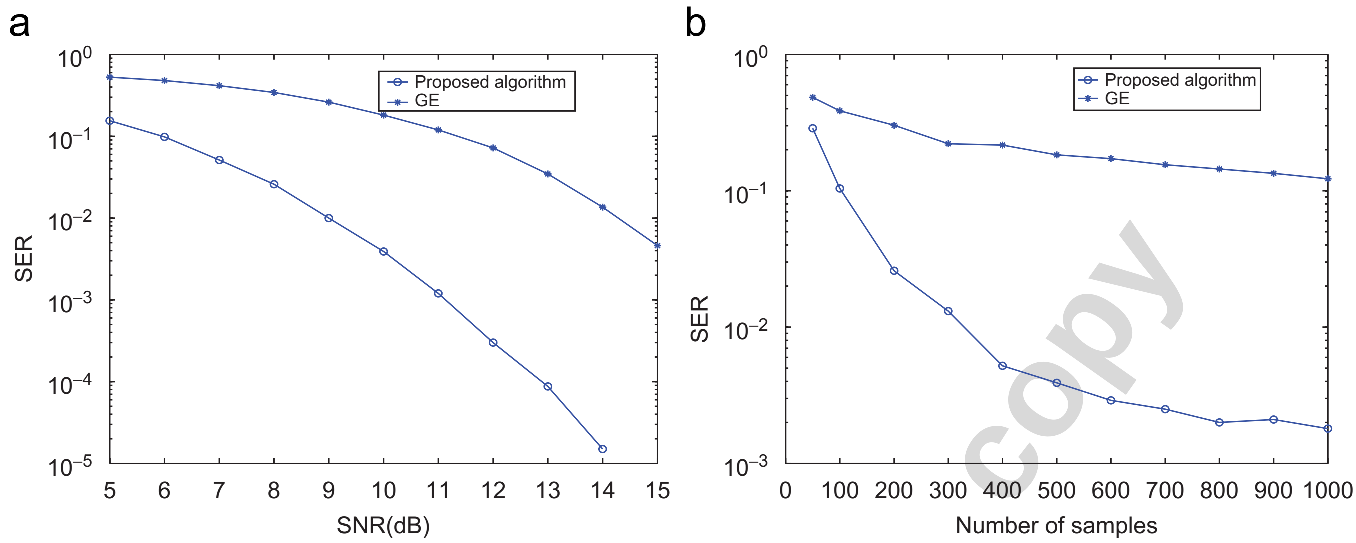

Fig. 4. Example 3: performance of respective algorithms. (a) SER versus SNR, $T_s = 500$. (b) SER versus T_s, SNR = 10 dB.

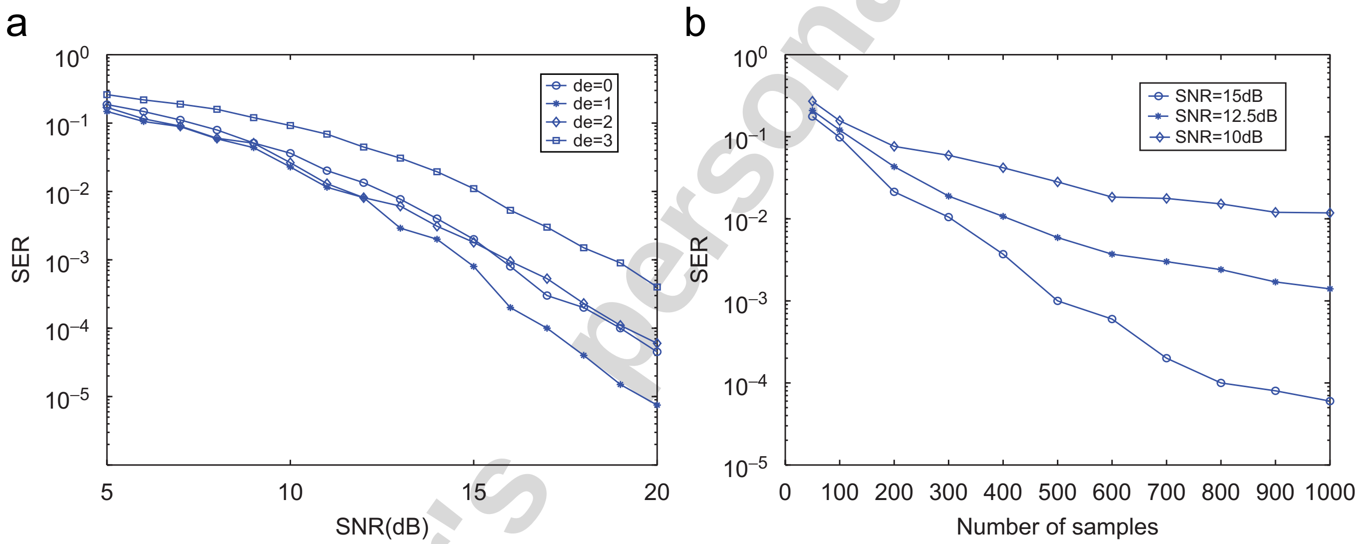

Fig. 5. Example 4: performance of the proposed algorithm. (a) SER versus SNR, $T_s = 500$. (b) SER versus T_s, $d_e = 1$.

of [9].

$$\mathbf{x}(n) = \sum_{l=0}^{1} \mathbf{h}_1(l)a(n-l) + \sum_{l=0}^{1} \mathbf{h}_2(l)s_2(n-l) + \mathbf{w}(n).$$

The two "sources" $a(n)$ and $s_2(n)$ are spatially, temporally correlated and generated in the same way as the above example. Observe that in this case, the linear and nonlinear kernels have the same channel order. Nevertheless, given that the conditions IC1–IC2 are satisfied, our proposed method still applies. In our simulations, the stack number N is chosen to be 2. Fig. 5(a) shows the SER (for the nonlinear input signals) versus SNR for the different equalization delays. It can be seen that the equalizer with $d_e = 1$ yields the best performance. The poorest results are obtained for $d_e = 3$. In Fig. 5(b), the variation of the SER (for the nonlinear input signals) with the number of data samples T_s for SNR = 10, 12.5, 15 dB is displayed. The equalizer with $d_e = 1$ is used.

7. Conclusion

In this paper, we consider the problem of blind MIMO FIR channel identification driven by spatially correlated sources whose second-order statistics are known *a priori*. A SOS-based method that admits a closed-form solution is proposed and its corresponding identifiability conditions are investigated. As a further result, we show that our method still applies to the spatially and temporally uncorrelated sources given that a certain channel order disparity condition is satisfied. Simulation results show that our method can be successfully employed for blind nonlinear SIMO channel equalization. As compared to other existing methods [8,9] for blind nonlinear SIMO channel equalization, our method renders a wider applicability for the input sources than [9] and exhibits better performance than [8].

Appendix A. Proof of Lemma 1

For notational convenience, let $\mathbf{G}_1 \triangleq \mathbf{TY}$ and $\mathbf{G}_2 \triangleq \mathbf{YT}$. Also let $g^1_{i,j}$ denote the entry of ith row and jth column in $\mathbf{G}_1$, $g^2_{i,j}$ denote the entry of ith row and jth column in $\mathbf{G}_2$.

Step 1: We first consider the first column of $\mathbf{G}_1$ and $\mathbf{G}_2$. Clearly, we have

$$g^1_{n,1} = t_{n,n}y_{n,1}, \quad g^2_{n,1} = y_{n,1}t_{1,1}.$$

Since $g^1_{n,1} = g^2_{n,1}$ and $t_{n,n} \neq t_{1,1}$, we have $y_{n,1} = 0$. By utilizing this result, we can further derive that

$$g^1_{n-1,1} = t_{n-1,n-1}y_{n-1,1}, \quad g^2_{n-1,1} = y_{n-1,1}t_{1,1}.$$

Also, we have $y_{n-1,1} = 0$ because $g^1_{n-1,1} = g^2_{n-1,1}$ and $t_{n-1,n-1} \neq t_{1,1}$. In this iterative way, by comparing $g^1_{k,1}$ with $g^2_{k,1}$ for $k \in \{2, \ldots, n-2\}$, it is not hard to conclude that

$$y_{k,1} = 0 \quad \forall k \in \{2, \ldots, n\}. \tag{35}$$

Step 2: Now we proceed to consider the second column of $\mathbf{G}_1$ and $\mathbf{G}_2$. We have the following by using the derived results in step 1

$$g^1_{n,2} = t_{n,n}y_{n,2}, \quad g^2_{n,2} = y_{n,2}t_{2,2}.$$

Thus we have $y_{n,2} = 0$. By a similar iterative way as in step 1, we can conclude that

$$y_{k,2} = 0 \quad \forall k \in \{3, \ldots, n\}. \tag{36}$$

Step 3: Now we assume that $y_{i,j} = 0$ if $j \leqslant m$ and $i > j$, where m is a certain value between 2 and $n-2$, we need to prove that $y_{k,m+1} = 0$ for $k \in \{m+2, \ldots, n\}$. Similarly as the previous steps, it is easy to derive that

$$g^1_{n,m+1} = t_{n,n}y_{n,m+1}, \quad g^2_{n,m+1} = y_{n,m+1}t_{m+1,m+1}.$$

Since $g^1_{n,m+1} = g^2_{n,m+1}$ and $t_{n,n} \neq t_{m+1,m+1}$, we have $y_{n,m+1} = 0$. And also in an iterative way, we can conclude that

$$y_{k,m+1} = 0 \quad \forall k \in \{m+2, \ldots, n\}. \tag{37}$$

The proof is completed here.

Appendix B. Proof of Lemma 2

We present our proof in the following three steps.

Step 1: For notational convenience, let $\mathbf{G}_1 \triangleq \mathbf{J}_m\mathbf{Y} = \mathbf{YJ}_n$, $\mathbf{G}_2 \triangleq \mathbf{J}_m^{\mathrm{T}}\mathbf{Y} = \mathbf{YJ}_n^{\mathrm{T}}$. Considering the relationship of $\mathbf{G}_1$, we have

$$\begin{aligned}\mathbf{G}_1[2:m,n] &= [y_{1,n} \;\; y_{2,n} \;\; \cdots \;\; y_{m-1,n}]^{\mathrm{T}} \\ &= [0 \;\; 0 \;\; \cdots \;\; 0]^{\mathrm{T}}\end{aligned} \tag{38}$$

$$\begin{aligned}\mathbf{G}_1[1,1:n-1] &= [0 \;\; 0 \;\; \cdots \;\; 0] \\ &= [y_{1,2} \;\; y_{1,3} \;\; \cdots \;\; y_{1,n}]\end{aligned} \tag{39}$$

Similarly, considering the relationship of $\mathbf{G}_2$, we have

$$\begin{aligned}\mathbf{G}_2[1:m-1,1] &= [y_{2,1} \;\; y_{3,1} \;\; \cdots \;\; y_{m-1,1}]^{\mathrm{T}} \\ &= [0 \;\; 0 \;\; \cdots \;\; 0]^{\mathrm{T}},\end{aligned} \tag{40}$$

$$\begin{aligned}\mathbf{G}_2[m,2:n] &= [0 \;\; 0 \;\; \cdots \;\; 0] \\ &= [y_{m,1} \;\; y_{m,2} \;\; \cdots \;\; y_{m,n-1}].\end{aligned} \tag{41}$$

Therefore, we can conclude that all entries located at the edges of the matrix $\mathbf{Y}$ are zero except the entries $y_{1,1}$ and $y_{m,n}$.

Step 2: Now we consider the sub-matrix of $\mathbf{G}_1$ from second row to mth row and from first column to $(n-1)$th column, denoted by $\mathbf{G}_1[2:m,1:n-1]$. This sub-matrix can be easily computed as if we write $\mathbf{J}_m$ and $\mathbf{Y}$ as follows:

$$\mathbf{J}_m = \begin{bmatrix} \mathbf{0}_{1\times(m-1)} & 0 \\ \mathbf{I}_{(m-1)\times(m-1)} & \mathbf{0}_{(m-1)\times 1} \end{bmatrix},$$

$$\mathbf{Y} = \begin{bmatrix} \mathbf{Y}[1:m-1,1:n-1] & \mathbf{Y}[1:m-1,n] \\ \mathbf{Y}[m,1:n-1] & y_{m,n} \end{bmatrix}.$$

Obviously from $\mathbf{G}_1 = \mathbf{J}_m\mathbf{Y}$ we have

$$\mathbf{G}_1[2:m,1:n-1] = \mathbf{Y}[1:m-1,1:n-1]. \tag{42}$$

On the other hand, we can write $\mathbf{J}_n$ and $\mathbf{Y}$ as

$$\mathbf{Y} = \begin{bmatrix} y_{1,1} & \mathbf{Y}[1,2:n] \\ \mathbf{Y}[2:m,1] & \mathbf{Y}[2:m,2:n] \end{bmatrix},$$

$$\mathbf{J}_n = \begin{bmatrix} \mathbf{0}_{1\times(n-1)} & 0 \\ \mathbf{I}_{(n-1)\times(n-1)} & \mathbf{0}_{(n-1)\times 1} \end{bmatrix}.$$

Then from $\mathbf{G}_1 = \mathbf{Y}\mathbf{J}_n$ we have

$$\mathbf{G}_1[2:m, 1:n-1] = \mathbf{Y}[2:m, 2:n]. \tag{43}$$

By combining Eqs. (42) and (43), we can conclude that

$$y_{i,j} = y_{i+1,j+1} \tag{44}$$

for $i \in \{1, \ldots, m-1\}, j \in \{1, \ldots, n-1\}$, which shows that $\mathbf{Y}$ has a Toeplitz form.

Step 3: If $m = n$, based on the above derived results, it is easy to know that all entries on the main diagonal are equal, and all entries off the main diagonal are zero. Therefore, we conclude that $\mathbf{Y} = \lambda\mathbf{I}$, where λ could be any complex scalar including zero. If $m \neq n$, since $\mathbf{Y}$ has a Toeplitz form and all entries located at the edges of the matrix $\mathbf{Y}$ are zero (note that $y_{1,1}$ and $y_{m,n}$ can be easily proved to be zero by utilizing the Toeplitz form when $m \neq n$), hence $\mathbf{Y} = \mathbf{0}$. The proof is completed here.

Appendix C. Proof of Theorem 3

We can derive the following:

$$\mathbf{Q}^{\mathrm{H}}\mathbf{R}_s[k]\mathbf{Q}\mathbf{C}_l = \mathbf{C}_l\mathbf{R}_{s_l}[k] \Rightarrow \mathbf{R}_s[k]\mathbf{Q}\mathbf{C}_l = \mathbf{Q}\mathbf{C}_l\mathbf{R}_{s_l}[k]. \tag{45}$$

Let $\mathbf{Z}_l \triangleq \mathbf{Q}\mathbf{C}_l$, we can rewrite the above equation as

$$\mathbf{R}_s[k]\mathbf{Z}_l = \mathbf{Z}_l\mathbf{R}_{s_l}[k], \quad k \in \{\pm 1\}. \tag{46}$$

If we partition matrix $\mathbf{Z}_l$ as $\mathbf{Z}_l \triangleq [\mathbf{Z}_{1l}^{\mathrm{T}}\ \mathbf{Z}_{2l}^{\mathrm{T}}\ \cdots\ \mathbf{Z}_{pl}^{\mathrm{T}}]^{\mathrm{T}}$, where $\mathbf{Z}_{il} \in \mathbb{C}^{d_i \times d_l}$, then, by exploiting the block diagonal structure of $\mathbf{R}_s[k]$, we have the following:

$$\text{(a) } \mathbf{J}_{d_i}\mathbf{Z}_{il} = \mathbf{Z}_{il}\mathbf{J}_{d_l}, \quad \text{(b) } \mathbf{J}_{d_i}^{\mathrm{T}}\mathbf{Z}_{il} = \mathbf{Z}_{il}\mathbf{J}_{d_l}^{\mathrm{T}}. \tag{47}$$

By utilizing the results of Lemma 2, we can conclude that $\mathbf{Z}_l = [\mathbf{0}\ \cdots\ \mathbf{0}\ \lambda_l\mathbf{I}_{d_l}\ \mathbf{0}\ \cdots\ \mathbf{0}]^{\mathrm{T}}$. Hence we have $\mathbf{C}_l = \mathbf{Q}^{\mathrm{H}}\mathbf{Z}_l = \lambda_l\bar{\mathbf{Q}}_l$. The proof is completed here.

References

[1] Y. Hua, S. An, Y. Xiang, Blind identification of FIR channels by decorrelating subchannels, IEEE Trans. Signal Process. 51 (May 2003) 1143–1155.

[2] T. Ma, Z. Ding, S.F. Yau, A two-stage algorithm for MIMO blind deconvolution of colored input signals, IEEE Trans. Signal Process. 48 (April 2000) 1187–1192.

[3] Z. Ding, L. Qiu, Blind MIMO channel identification from second order statistics using rank deficient channel convolution matrix, IEEE Trans. Signal Process. 51 (February 2003) 535–544.

[4] J.K. Tugnait, B. Huang, On a whitening approach to partial channel estimation and blind equalization of FIR/IIR multiple-input multiple-output channels, IEEE Trans. Signal Process. 48 (March 2000) 832–845.

[5] K.A. Meraim, Y. Hua, Blind identification of multi-input multi-output system using minimum noise subspace, IEEE Trans. Signal Process. 45 (January 1997) 254–258.

[6] J.K. Tugnait, B. Huang, Multistep linear predictors based blind identification and equalization of multiple-input multiple-output channels, IEEE Trans. Signal Process. 48 (January 2000) 26–38.

[7] J. Shen, Z. Ding, Direct blind MMSE channel equalization based on second order statistics, IEEE Trans. Signal Process. 48 (April 2000) 1015–1022.

[8] G.B. Giannakis, E. Serpedin, Linear multichannel blind equalizers of nonlinear FIR volterra channels, IEEE Trans. Signal Process. 45 (January 1997) 67–81.

[9] R. López-Valcarce, S. Dasgupta, Blind equalization of nonlinear channels from second-order statistics, IEEE Trans. Signal Process. 49 (December 2001) 3084–3097.

[10] K.A. Meraim, P. Loubaton, E. Moulines, A subspace algorithm for certain blind identification problems, IEEE Trans. Inform. Theory 43 (March 1997) 499–511.

[11] M. Wax, T. Kailath, Detection of signals by information theoretic criterion, IEEE Trans. Acoust. Speech Signal Process. 33 (April 1985) 387–392.

[12] G. Golub, C. van Loan, Matrix Computations, third ed., Johns Hopkins University Press, Baltimore, MD, 1996.

[13] J. Xavier, V. Barroso, J. Moura, Closed-form correlative coding blind identification of MIMO channels: Isometry fitting to second order statistics, IEEE Trans. Signal Process. 49 (May 2001) 1073–1086.

[14] J. Fang, A.R. Leyman, Y.H. Chew, H. Duan, Blind channel identification with colored sources by exploiting properties of companion matrices, IEEE Trans. Signal Process. 54 (March 2006) 894–906.

[15] H. Luo, R.-W. Liu, A closed-form solution to blind MIMO FIR channel equalization for wireless communication systems based on autocorrelation matching, Proc. ICASSP 4 (2003) 293–296.

[16] R. López-Valcarce, Z. Ding, S. Dasgupta, Equalization and interference cancellation in linear multiuser systems based on second-order statistics, IEEE Trans. Signal Process. 49 (September 2001) 2042–2049.

[17] C.B. Papadias, D.T.M. Slock, Fractionally spaced equalization of linear polyphase channels and related blind techniques based on multichannel linear prediction, IEEE Trans. Signal Process. 47 (March 1999) 641–654.