# Blind SIMO FIR Channel Estimation by Utilizing Property of Companion Matrices

Jun Fang, A. Rahim Leyman, *Member, IEEE*, Yong Huat Chew, and Huiping Duan

***Abstract*—We present a closed-form solution for blind single-input multiple-output finite impulse response channel estimation driven by colored sources. The second-order statistics of the input source are known *a priori*. The uniqueness of the system solution is proved by exploiting the derived property of companion matrices that are constructed from the inherent structural relationship between the source autocorrelation matrices. Numerical simulation results are presented to illustrate the performance of the proposed algorithm.**

***Index Terms*—Blind channel estimation, colored source, companion matrix, single-input–multiple-output systems (SIMO).**

## I. Introduction

IN THIS letter, we consider the problem of blind single-input–multiple-output (SIMO) finite-impulse response (FIR) channel estimation driven by colored source signals. In particular, we are interested in the case where the second-order statistics (SOSs) of the input signals are known *a priori*. Colored sources with known statistics indeed occur in practice. For example, colored sources arise as a result of channel encoding [1], and the knowledge of the encoding scheme alone provides the required source statistics to the receiver. There are existing methods [2]–[4] that address the same problem as in this letter. Among them, [2] proposed a subspace-based method by exploiting the block Toeplitz structure of the channel convolution matrix and, thus, required no knowledge of input statistics whatsoever. Reference [4] imposes a somewhat restrictive condition on the source correlation, where an exponentially delaying autocorrelation function is assumed. Reference [3] constitutes a direct extension of the TXK method [5] by exploiting the inherent structural relationship between the source autocorrelation matrices $\mathbf{R}_s[0]$ and $\mathbf{R}_s[1]$. In this letter, we propose a new closed-form solution for blind channel estimation driven by colored sources. The contribution of this letter consists of the following three aspects. First, the inherent structural relationship between source autocorrelation matrices $\mathbf{R}_s[0]$ and $\mathbf{R}_s[\pm 1]$ is further exploited. Second, we derive a new property of a pair of constructed companion matrices, which plays a key role in devising and validating our algorithm. Third, unlike other methods [2]–[4], which have difficulties in extending to the multiuser's scenarios, the proposed algorithm has the potential to extend to the multiple-input multiple-output (MIMO) systems [6]. We include computer simulations to study the performance of the proposed algorithm.

Manuscript received August 4, 2004; revised November 21, 2004. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Zhi Ding.

J. Fang is with the Department of Electrical and Computer Engineering, Faculty of Engineering, National University of Singapore, Singapore 117576, Singapore (e-mail: g0202082@nus.edu.sg).

A. R. Leyman and Y. H. Chew are with the Digital Wireless Department, Institute for Infocomm Research, Singapore 119613, Singapore (e-mail: larahim@i2r.a-star.edu.sg; chewyh@i2r.a-star.edu.sg).

H. Duan is with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798, Singapore (e-mail: DUAN0002@ntu.edu.sg).

Digital Object Identifier 10.1109/LSP.2005.845605

## II. System Model and Basic Assumptions

We begin by considering the SIMO FIR channel model given as

$$\mathbf{x}_n \triangleq \mathbf{h}_n * s_n + \mathbf{w}_n \triangleq \sum_{l=0}^{L} \mathbf{h}_l s_{n-l} + \mathbf{w}_n \tag{1}$$

where $\{s_n\}$ is the zero mean, wide sense stationary sequence of transmitted symbols, $\{\mathbf{x}_n\}$ is the $q\times$ 1 channel output vector, $\{\mathbf{w}_n\}$ is the $q\times$ 1 white noise vector, and $\{\mathbf{h}_n\}$ represents the multichannel impulse response. By stacking the channel output vector $\{\mathbf{x}_n\}$ and defining $\vec{\mathbf{x}}_n \triangleq [\mathbf{x}_n^T\ \mathbf{x}_{n-1}^T\ \ldots\ \mathbf{x}_{n-N}^T]^T$, $\vec{\mathbf{s}}_n \triangleq [s_n\ s_{n-1}\ \ldots\ s_{n-N-L}]^T$, and $\vec{\mathbf{w}}_n \triangleq [\mathbf{w}_n^T\ \mathbf{w}_{n-1}^T\ \ldots\ \mathbf{w}_{n-N}^T]^T$, we can therefore re-express (1) as

$$\vec{\mathbf{x}}_n = \mathcal{H}\vec{\mathbf{s}}_n + \vec{\mathbf{w}}_n \tag{2}$$

where the channel convolution matrix $\mathcal{H} \in \mathbb{C}^{(N+1)q\times d}$ is a block Toeplitz matrix written as follows, and $d \triangleq N + L + 1$:

$$\mathcal{H} \triangleq \begin{bmatrix} \mathbf{h}_0 & \ldots & \mathbf{h}_L & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{0} & \mathbf{h}_0 & \ldots & \mathbf{h}_L & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \ldots & \mathbf{0} & \mathbf{h}_0 & \ldots & \mathbf{h}_L \end{bmatrix}.$$

We adopt the following notations and assumptions throughout this letter. The notations $[\cdot]^T$, $[\cdot]^*$, $[\cdot]^H$, and $[\cdot]^\dagger$ stand for matrix transpose, complex conjugate, matrix Hermitian transpose, and matrix pseudo-inverse, respectively. $E[\cdot]$ represents the mathematical expectation. $\|\mathbf{X}\|$ ($\|\mathbf{x}\|$) denotes the Frobenius norm (vector 2-norm) of matrix $\mathbf{X}$ (vector $\mathbf{x}$). The symbol $\mathbf{J}_1$ ($\mathbf{J}^1$) stands for the one-lag down (up) shift square matrix whose first subdiagonal entries below (above) the main diagonal are unity, whereas all remaining entries are zero; $\mathbf{e}_i$ denotes the unit column vector, with its $i$th equal to one and its other entries equal to zero. $\mathbb{C}^{n\times m}$ and $\mathbb{C}^n$ denote the set of $n \times m$ matrices and the set of $n$-dimensional column vectors with complex entries, respectively. Some basic assumptions are adopted as follows: 1) $\mathcal{H}$ is full-column rank: a condition equivalent to requiring that the channel $\mathbf{h}(z)$ is irreducible. 2) Channel order $L$ is assumed to be known *a priori*. 3) Source signal is a zero mean, wide sense stationary colored signal with the knowledge that its input statistics are available. Its autocorrelation matrix with lag $k$ defined as $\mathbf{R}_s[k] \triangleq E[\vec{\mathbf{s}}_n \vec{\mathbf{s}}_{n-k}^H]$. 4) Additive noises are spatially and temporally white, and they are statistically independent of the source.

## III. Proposed Channel Identification Method

In order to simplify the presentation of the proposed channel identification method, we assume the noiseless case. Thus, the autocorrelation matrix of the received data $\vec{\mathbf{x}}(n)$ with lag $k$ can be expressed as

$$\mathbf{R}_x[k] = \mathcal{H}\mathbf{R}_s[k]\mathcal{H}^H. \tag{3}$$

Our goal is to find an estimate of $\mathcal{H}$ from (3) by using the knowledge of $\mathbf{R}_s[k]$. We commence by introducing the following lemma.

*Lemma 1:* Given $\mathbf{R}_x[k] = \mathcal{H}\mathbf{R}_s[k]\mathcal{H}^H$, $\mathcal{H}$ is full-column rank, and $\mathbf{R}_s[0]$ is invertible, we have

$$\mathbf{R}_x[k]\mathbf{R}_x^\dagger[0] = \mathcal{H}\mathbf{R}_s[k]\mathbf{R}_s^{-1}[0]\mathcal{H}^\dagger \tag{4}$$

$$\mathbf{R}_x[k]\mathbf{R}_x^\dagger[0]\mathcal{H} = \mathcal{H}\mathbf{R}_s[k]\mathbf{R}_s^{-1}[0]. \tag{5}$$

*Proof:* This lemma can be easily proven since we have $\mathbf{R}_x^\dagger[0] = (\mathcal{H}^H)^\dagger \mathbf{R}_s^{-1}[0]\mathcal{H}^\dagger$, which satisfies the four *Moore–Penrose conditions*. ∎

For convenience, let

$$\Upsilon_{2k-1} \triangleq \mathbf{R}_x[k]\mathbf{R}_x^\dagger[0] \quad \Upsilon_{2k} \triangleq \mathbf{R}_x[-k]\mathbf{R}_x^\dagger[0]$$
$$\Phi_{2k-1} \triangleq \mathbf{R}_s[k]\mathbf{R}_s^{-1}[0] \quad \Phi_{2k} \triangleq \mathbf{R}_s[-k]\mathbf{R}_s^{-1}[0].$$

We can therefore re-express (5) (choose $K \geq k \geq 1$) as

$$\Upsilon_i\mathcal{H} = \mathcal{H}\Phi_i, \qquad 2K \geq i \geq 1. \tag{6}$$

The above set of equations can be used to identify the channel $\mathcal{H}$ since the knowledge of $\Phi_i$ is known *a priori*, and the information of $\Upsilon_i$ can be obtained from the second-order statistics of the observed data. By exploiting the block Toeplitz structure of $\mathcal{H}$, we can rewrite (6) as

$$\mathcal{T}_1[\Upsilon_i]\mathbf{h} = \mathcal{T}_2[\Phi_i]\mathbf{h}, \qquad 2K \geq i \geq 1 \tag{7}$$

where $\mathbf{h} \triangleq \begin{bmatrix}\mathbf{h}_0^T & \cdots & \mathbf{h}_L^T\end{bmatrix}^T$, $\mathcal{T}_1[\cdot]$, and $\mathcal{T}_2[\cdot]$, respectively, represent a certain transformation on the bracketed matrix. The transformed matrices $\mathcal{T}_1[\Upsilon_i]$ and $\mathcal{T}_2[\Phi_i]$ are all of the same dimension $\mathbb{C}^{(N+L+1)(N+1)q\times(L+1)q}$. Therefore, we may estimate $\mathbf{h}$ by the following criterion:

$$\hat{\mathbf{h}} = \arg\min_{\|\mathbf{u}\|=1} \sum_{k=1}^{2K} \|\left[\mathcal{T}_1[\Upsilon_k] - \mathcal{T}_2[\Phi_k]\right]\mathbf{u}\|^2. \tag{8}$$

The above optimization has a closed-form solution that can be obtained as the right singular vector associated with the smallest singular value. However, this criterion is trivial if the solution of (7) is not unique, i.e., there exist other nonzero vectors that are *linearly independent of* $\mathbf{h}$ and also satisfy (7). Hence, two fundamental problems arise. First, we would like to ascertain whether or not the solution of (7) is unique (up to a scalar factor). Second, we would like to determine under what conditions the solution of (7) will be unique. These two problems are studied in the following, and we will establish the uniqueness of the solution to (7) by using only the autocorrelation matrices $\mathbf{R}_x[0]$ and $\mathbf{R}_x[\pm 1]$, i.e., the uniqueness of the solution can be guaranteed by choosing $K = 1$ in (7).

We begin by observing the structural relationship between $\mathbf{R}_s[0]$ and $\mathbf{R}_s[\pm 1]$. It can be seen that the last $d-1$ rows of $\mathbf{R}_s[1]$ are the first $d-1$ rows of $\mathbf{R}_s[0]$, and the first $d-1$ rows of $\mathbf{R}_s[-1]$ are the last $d-1$ rows of $\mathbf{R}_s[0]$. Hence, we can establish the following relationship:

$$\mathbf{R}_s[1] = \mathbf{J}_1\mathbf{R}_s[0] + \mathbf{e}_1\mathbf{r}_1^H \tag{9}$$

$$\mathbf{R}_s[-1] = \mathbf{J}^1\mathbf{R}_s[0] + \mathbf{e}_d\mathbf{r}_2^H \tag{10}$$

where

$$\mathbf{r}_1^H \triangleq \mathbf{e}_1^H\mathbf{R}_s[1] = E[s_n\vec{\mathbf{s}}_{n-1}^H] \tag{11}$$

$$\mathbf{r}_2^H \triangleq \mathbf{e}_d^H\mathbf{R}_s[-1] = E[s_{n-d+1}\vec{\mathbf{s}}_{n+1}^H]. \tag{12}$$

In addition, if we define $r_{1,i}$ and $r_{2,i}$ as the $i$th entries of the vectors $\mathbf{r}_1$ and $\mathbf{r}_2$, respectively; then, the entries in these vectors are related as follows:

$$r_{1,i} = r_{2,d+1-i}^*, \qquad \forall\, i \in \{1,\ldots,d\}. \tag{13}$$

Using (9) and (10), we can express $\Phi_i$, $i = 1, 2$ as follows:

$$(a)\ \Phi_1 \triangleq \mathbf{J}_1 - \mathbf{e}_1\vec{\alpha}_1^H, \quad (b)\ \Phi_2 \triangleq \mathbf{J}^1 - \mathbf{e}_d\vec{\alpha}_2^H \tag{14}$$

where

$$\vec{\alpha}_1 = \begin{bmatrix}\alpha_{1,1} & \cdots & \alpha_{1,d}\end{bmatrix}^T = -\mathbf{R}_s^{-1}[0]\mathbf{r}_1 \tag{15}$$

$$\vec{\alpha}_2 = \begin{bmatrix}\alpha_{2,1} & \cdots & \alpha_{2,d}\end{bmatrix}^T = -\mathbf{R}_s^{-1}[0]\mathbf{r}_2. \tag{16}$$

It is clear that the entries in $\vec{\alpha}_1$ are the coefficients of the $d$th-order optimum forward prediction error filter for the process $\{s_n\}$ and the entries in $\vec{\alpha}_2$ are exactly the coefficients of the $d$th-order optimum backward prediction error filter for the process $\{s_n\}$ [8]. Moreover, the relationship of $\vec{\alpha}_1$ and $\vec{\alpha}_2$ can be formalized in the following lemma.

*Lemma 2:* Given that $\vec{\alpha}_1$ and $\vec{\alpha}_2$ are defined in (15) and (16), respectively, there holds

$$\alpha_{1,i} = \alpha_{2,d+1-i}^*, \qquad \forall\, i \in \{1,\ldots,d\}. \tag{17}$$

*Proof:* See [8, Sec. 6.2]. ∎

Observe that both $\Phi_1 = \mathbf{J}_1 - \mathbf{e}_1\vec{\alpha}_1^H$ and $\Phi_2 = \mathbf{J}^1 - \mathbf{e}_d\vec{\alpha}_2^H$ are companion matrices that have some important properties to be investigated and exploited. We highlight one of the exploited properties as follows

*Lemma 3:* If matrix $\mathbf{Y}$ commutes with $\Phi_1$ and $\Phi_2$, respectively, i.e.,

$$(a)\ \Phi_1\mathbf{Y} = \mathbf{Y}\Phi_1, \quad (b)\ \Phi_2\mathbf{Y} = \mathbf{Y}\Phi_2 \tag{18}$$

where $\mathbf{Y} \in \mathbb{C}^{d\times d}$ and the modulus of $\alpha_{1,d}$ in $\Phi_1$ is not equal to one, i.e., $|\alpha_{1,d}| \neq 1$, then $\mathbf{Y} = \lambda\mathbf{I}$, where $\lambda$ could be any complex scalar, including zero.

*Proof:* See the Appendix. ∎

We now prove the uniqueness of the system solution to (7) by using the above lemma. Notice that (6) and (7) can be derived from each other. Therefore, we only need to prove that the solution of (6) is unique (up to a scalar factor). Thus, the problem can be formulated as follows: Given that the following two equations hold

$$\text{(a) } \Upsilon_1 = \mathcal{H}\Phi_1\mathcal{H}^\dagger \quad \text{(b) } \Upsilon_2 = \mathcal{H}\Phi_2\mathcal{H}^\dagger \tag{19}$$

and $\mathcal{H}$ is full-column rank, we need to prove that $\mathcal{H}$ can be uniquely determined up to a complex scalar by the following two equations:

$$\text{(a) } \Upsilon_1\mathcal{H} = \mathcal{H}\Phi_1 \quad \text{(b) } \Upsilon_2\mathcal{H} = \mathcal{H}\Phi_2. \tag{20}$$

It implies that, if any nonzero matrix $\mathcal{G}$ that has the same structure as $\mathcal{H}$ also satisfies (20a) and (20b), then $\mathcal{G} = \lambda\mathcal{H}$, where $\lambda$ is a nonzero complex scalar.

*Proof:* Suppose a nonzero matrix $\mathcal{G}$ that has the same Toeplitz structure as $\mathcal{H}$ also satisfies (20a) and (20b); then, we have

$$\begin{aligned}\Upsilon_1\mathcal{G} = \mathcal{G}\Phi_1 &\Rightarrow \mathcal{H}\Phi_1\mathcal{H}^\dagger\mathcal{G} = \mathcal{G}\Phi_1\\ &\Rightarrow \Phi_1\mathcal{H}^\dagger\mathcal{G} = \mathcal{H}^\dagger\mathcal{G}\Phi_1 \end{aligned}\tag{21}$$

$$\begin{aligned}\Upsilon_2\mathcal{G} = \mathcal{G}\Phi_2 &\Rightarrow \mathcal{H}\Phi_2\mathcal{H}^\dagger\mathcal{G} = \mathcal{G}\Phi_2\\ &\Rightarrow \Phi_2\mathcal{H}^\dagger\mathcal{G} = \mathcal{H}^\dagger\mathcal{G}\Phi_2. \end{aligned}\tag{22}$$

By invoking Lemma 3, we know that $\mathcal{H}^\dagger\mathcal{G} = \lambda\mathbf{I}$. Therefore, we only need to prove that the solution of $\mathcal{G}$ that satisfies $\mathcal{H}^\dagger\mathcal{G} = \lambda\mathbf{I}$ is unique and $\mathcal{G} = \lambda\mathcal{H}$. Note that $\mathcal{G}$ has the same block Toeplitz structure as $\mathcal{H}$. If we write $\mathcal{H}^\dagger \triangleq [\mathbf{V}_0 \cdots \mathbf{V}_N]$, we can rewrite $\mathcal{H}^\dagger\mathcal{G} = \lambda\mathbf{I}$ as

$$\mathcal{V}\begin{bmatrix}\mathbf{g}_0\\ \vdots\\ \mathbf{g}_L\end{bmatrix} = \text{vec}(\lambda\mathbf{I}) \tag{23}$$

where $\mathcal{V} \in \mathbb{C}^{d^2\times(L+1)q}$ is a block Toeplitz matrix written as

$$\mathcal{V} \triangleq \begin{bmatrix}\mathbf{V}_0 & \mathbf{0} & \cdots & \mathbf{0}\\ \vdots & \mathbf{V}_0 & \ddots & \vdots\\ \mathbf{V}_N & \ddots & \ddots & \mathbf{0}\\ \mathbf{0} & \mathbf{V}_N & \ddots & \mathbf{V}_0\\ \vdots & \ddots & \ddots & \vdots\\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{V}_N\end{bmatrix}.$$

Obviously, from (23), we know that $\mathcal{G}$ can be uniquely determined if $\mathcal{V}$ has full-column rank. Recalling Theorem 1 in [8], $\mathcal{V}$ has full-column rank if the following condition holds, i.e., there exists a nonzero $z_0$ (including $\infty$) such that the polynomial matrix $V(z_0)$ has full-column rank, where $V(z_0) \triangleq \mathbf{V}_0 + \mathbf{V}_1z^{-1} + \cdots + \mathbf{V}_Nz^{-N}$. This mild condition can be assured with probability one since generally, when $L \geq 1$, the entries of matrix $\mathcal{H}^\dagger$ can be considered as randomly generated. Thus, we can conclude that the solution of $\mathcal{G}$ is unique and $\mathcal{G} = \lambda\mathcal{H}$. Note that $\lambda$ cannot be zero because $\mathcal{G}$ would be zero under the condition $\lambda = 0$, which contradicts our previously made assumption $\mathcal{G} \neq \mathbf{0}$. The proof is completed here. ■

TABLE I
SER VERSUS SNR AND $T_s$, RESPECTIVELY

| SNR(dB) | Proposed Method | LP | SS |
|---|---|---|---|
| 20 | 0.0000 | 0.0000 | 0.0000 |
| 17.5 | 0.0000 | 0.0000 | 0.0000 |
| 15 | 0.0006 | 0.0006 | 0.0023 |
| 12.5 | 0.0268 | 0.0101 | 0.0366 |
| 10 | 0.1368 | 0.0744 | 0.1417 |
| $T_s$ | Proposed Method | LP | SS |
| 2000 | 0.0087 | 0.0055 | 0.0121 |
| 1600 | 0.0099 | 0.0077 | 0.0186 |
| 1200 | 0.0224 | 0.0129 | 0.0292 |
| 800 | 0.0598 | 0.0277 | 0.0605 |
| 400 | 0.1368 | 0.0744 | 0.1417 |

## IV. SIMULATION RESULTS

We now present simulation results to illustrate the performance of our proposed algorithm. We compare our method to the other two methods, namely, the *subspace* (SS) method proposed in [2] and the so-called *linear prediction* (LP) approach presented in [3]. In our simulations, as an approximation of a two-ray multipath environment, the channel impulse response is obtained from the two delayed raised cosine pulses with its coefficients given by

$$[\,\mathbf{h}_0 \cdots \mathbf{h}_3\,] = \begin{bmatrix}-0.1470 & 0.4461 & 0.1126 & -0.2233\\ 0.0213 & 0.5356 & -0.2911 & 0.0660\end{bmatrix}.$$

The colored source is induced in the same way as the simulation example in [3]. The channel order is assumed known *a priori*, and the stack number (smoothed factor) $N$ is chosen to be three. For our proposed method, we only use the autocorrelation matrices $\mathbf{R}_x[0]$ and $\mathbf{R}_x[\pm1]$, i.e., $K = 1$ in criterion (8). Once the channel has been estimated, the minimum mean-square-error (MMSE) equalizers can be computed. The equalizer with equalization delay $d_e$, which is equal to 3, is used in our simulations. We present the equalization performance of the respective algorithms in Table I. The results are averaged over 500 Monte Carlo runs. In the first part of Table I, we show the symbol error rate (SER) as a function of signal-to-noise ratio (SNR) with the number of samples used to estimate signal statistics $T_s = 400$. Next, in the latter part of Table I, the SER is shown to be a function of $T_s$ for SNR $= 10$ dB. From the following table, we can see that the three algorithms perform similarly, with the performance of LP slightly better than that of the other two algorithms. In addition, our proposed method seems to lie somewhere between LP and SS.

Our proposed method estimates the channel matrix $\mathcal{H}$ by matching $\mathcal{H}\mathbf{R}_s[k]\mathbf{R}_s^{-1}[0]\mathcal{H}^\dagger$ and $\mathbf{R}_x[k]\mathbf{R}_x^\dagger[0]$ for $k = \pm1$. The accuracy of our estimated channel is subject to the estimation errors of $\mathbf{R}_x[k]\mathbf{R}_x^\dagger[0]$. This accounts for the lack of performance improvement of our proposed algorithm as compared to LP. Despite the slightly degraded performance, our algorithm shows an advantage over [2] and [3] since its extension to

MIMO systems is straightforward. For the multiuser's scenarios, (6) still holds, and under the assumption that all sources are uncorrelated with each other, we can further decompose (6) into $\Upsilon_i \mathcal{H}_l = \mathcal{H}_l \Phi_{i,l}$, where $\mathcal{H}_l$ denotes the channel convolution matrix corresponding to the $l$th user $\Phi_{i,l} \triangleq \mathbf{R}_{s_l}[\bar{k}]\mathbf{R}_{s_l}^{-1}[0]$, $s_l$ represents the $l$th source, $\bar{k} = (i+1)/2$ is odd, and $\bar{k} = -i/2$ if $i$ is even. Each user's channel convolution matrix $\mathcal{H}_l$ can then be identified according to the described algorithm. However, the proof for the uniqueness of (6) has to further exploit additional properties on companion matrices and also to impose a spectral diversity identifiability condition on the input colored sources [6].

## V. Conclusion

In this letter, we present a new SOS-based method that admits a closed-form solution for blind SIMO FIR channel estimation driven by colored source signals. The uniqueness of the closed-form solution is proven by exploiting the inherent structural relationship between $\mathbf{R}_s[0]$ and $\mathbf{R}_s[\pm 1]$ and the derived property of one pair of companion matrices. The proposed method is valid under a very mild condition on the source correlation. In fact, our method still works even if the source signals are white (this can be easily proven by following the procedure in this letter). Simulation results show that our proposed algorithm achieves a better performance than the classical subspace method [2].

## Appendix
## Proof of Lemma 3

We present our proof in the following three steps.

### A. Step 1

For notational convenience, let $\mathbf{G}_1 \triangleq \Phi_1 \mathbf{Y} = \mathbf{Y}\Phi_1$ and $\mathbf{G}_2 \triangleq \Phi_2 \mathbf{Y} = \mathbf{Y}\Phi_2$, $\mathbf{y}_i$ denote the $i$th column of $\mathbf{Y}$. We consider the last column of $\mathbf{G}_1$, denoted by $\mathbf{G}_1[:, d]$, and the first column of $\mathbf{G}_2$, denoted by $\mathbf{G}_2[:, 1]$. Thus, we have

$$\begin{aligned}\mathbf{G}_1[:, d] &= \begin{bmatrix} -\vec{\alpha}_1^H \mathbf{y}_d & y_{1,d} & \cdots & y_{d-1,d} \end{bmatrix}^T \\ &= -\alpha_{1,d}^* \begin{bmatrix} y_{1,1} & y_{2,1} & \cdots & y_{d,1} \end{bmatrix}^T \\ \mathbf{G}_2[:, 1] &= \begin{bmatrix} y_{2,1} & \cdots & y_{d,1} & -\vec{\alpha}_2^H \mathbf{y}_1 \end{bmatrix}^T \\ &= -\alpha_{1,d} \begin{bmatrix} y_{1,d} & y_{2,d} & \cdots & y_{d,d} \end{bmatrix}^T \end{aligned}$$

and we can obtain

$$y_{k,d} = |\alpha_{1,d}|^2 y_{k,d}, \quad \forall\, k \in \{1, \ldots, d-1\} \tag{24}$$

$$y_{k+1,1} = |\alpha_{1,d}|^2 y_{k+1,1}, \quad \forall\, k \in \{1, \ldots, d-1\}. \tag{25}$$

It is known that (see Theorem 1 in [3]) $|\alpha_{1,d}|$ will be less than one under the assumption that the $(d+1) \times (d+1)$ source autocorrelation matrix $\mathbf{R}_s[0]$ is positive definite. In fact, even if this assumption does not hold, the probability of $|\alpha_{1,d}| = 1$ is still almost equal to zero. Therefore, we can conclude that

$$\begin{aligned} y_{k,d} &= 0, \quad \forall\, k \in \{1, \ldots, d-1\} \\ y_{k+1,1} &= 0, \quad \forall\, k \in \{1, \ldots, d-1\}. \end{aligned}$$

### B. Step 2

We consider the submatrix of $\mathbf{G}_1$ from second row to $d$th row and from first column to $(d-1)$th column, denoted by $\mathbf{G}_1[2:d, 1:d-1]$. This submatrix can be easily computed if we write $\Phi_1$ and $\mathbf{Y}$ as follows:

$$\Phi_1 = \begin{bmatrix} -\vec{\alpha}_1^H[1:d-1] & -\alpha_{1,d}^* \\ \mathbf{I} & \mathbf{0} \end{bmatrix}$$

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

obviously, from $\mathbf{G}_1 = \Phi_1 \mathbf{Y}$, we have

$$\mathbf{G}_1[2:d, 1:d-1] = \mathbf{Y}[1:d-1, 1:d-1]. \tag{26}$$

On the other hand, if we rewrite $\mathbf{Y}$ as

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

then from $\mathbf{G}_1 = \mathbf{Y}\Phi_1$, we have (note that $\mathbf{Y}[2:d, 1] = \mathbf{0}$ from step 1)

$$\mathbf{G}_1[2:d, 1:d-1] = \mathbf{Y}[2:d, 2:d]. \tag{27}$$

By combining (26) and (27), we conclude that

$$y_{i,j} = y_{i+1,j+1} \tag{28}$$

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

### C. Step 3

Based on the results of previous steps, we know that all entries of $\mathbf{Y}$ on the main diagonal are equal, and all entries of $\mathbf{Y}$ off the main diagonal are zero. Therefore, we can write $\mathbf{Y} = \lambda \mathbf{I}$, where $\lambda$ could be any complex scalar, including zero.

## References

[1] J. Mannerkoski and V. Koivunen, "Autocorrelation properties of channel encoded sequences – Applicability to blind equaliztion," *IEEE Trans. Signal Process.*, vol. 48, no. 12, pp. 3501–3506, Dec. 2000.
[2] E. Moulines, P. Duhamel, J. F. Cardoso, and S. Mayrargue, "Subspace methods for the blind identification of multichannel FIR filters," *IEEE Trans. Signal Process.*, vol. 43, no. 2, pp. 516–525, Feb. 1995.
[3] R. Valcarce and S. Dasgupta, "Blind channel equalization with colored sources based on second-order statistics: A linear prediction approach," *IEEE Trans. Signal Process.*, vol. 49, no. 9, pp. 2050–2059, Sep. 2001.
[4] K. H. Afkhamie and Z.-Q. Luo, "Blind identification of FIR systems driven by Markov-like input signals," *IEEE Trans. Signal Process.*, vol. 48, no. 6, pp. 1726–1736, Jun. 2000.
[5] L. Tong, G. Xu, and T. Kailath, "Blind identification and equalization based on second-order statistics: A time domain approach," *IEEE Trans. Inf. Theory*, vol. 40, no. 3, pp. 340–349, Mar. 1994.
[6] J. Fang, A. R. Leyman, Y. H. Chew, and H. Duan, "A new closed-form solution for blind MIMO FIR channel estimation with colored sources," in *Proc. ICAS*, 2005, to be published.
[7] S. Haykin, *Adaptive Filter Theory*, 3rd ed. Englewood Cliffs, NJ: Prentice-Hall, 1996.
[8] J. Liang and Z. Ding, "Blind MIMO system identification based on cumulant subspace decomposition," *IEEE Trans. Signal Processing*, vol. 51, no. 6, pp. 1457–1468, Jun. 2003.
[9] G. Golub and C. van Loan, *Matrix Computations*, 3rd ed. Baltimore, MD: The Johns Hopkins Univ. Press, 1996.