A New Closed-form Solution for Blind MIMO FIR Channel Estimation with Colored Sources

Jun Fang
Department of Electrical and Computer Engineering
National University of Singapore
Singapore 119260
Email: g0202082@nus.edu.sg

A. Rahim Leyman
Institute for Infocomm Research
A*STAR
21 Heng Mui Keng Terrace
Singapore 119613
Email: larahim@i2r.a-star.edu.sg

Yong Huat Chew
Institute for Infocomm Research
A*STAR
21 Heng Mui Keng Terrace
Singapore 119613
Email: chewyh@i2r.a-star.edu.sg

***Abstract*— In this paper, we present a closed-form solution for blind multiple-input multiple-output (MIMO) finite impulse response (FIR) channel estimation driven by colored sources whose second-order statistics (SOS) are assumed to be known *a priori*. The proof for the uniqueness of the system solution is provided. Numerical simulation results are presented to illustrate the performance of the proposed algorithm.**

I. Introduction

In this paper, we consider the problem of blind MIMO FIR channel estimation driven by colored signals. Of specific interest, we focus on the case where the second order statistics 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 will provide the required source statistics to the receiver. Moreover, the correlative filters can be utilized at the transmitters to induce distinct spectral patterns to the source signals [2]. Interestingly, there have been some works [3]–[5] on blind MIMO FIR channel identification when the input signals are colored but with unknown statistics. However, failing to utilize the information of input signals statistics severely affects the estimator's performance. Previous works that address the same problem as in this paper include [6], [7] for the SIMO case and [2] for the MIMO case. They all admit closed-form solutions. Recently, a frequency-domain nonlinear iterative method [8] was proposed for blind MIMO channel estimation with colored sources. Due to its nonlinear nature, the method may obtain better estimate as compared to existing linear methods, but it needs a good initialization to minimize the problem of local minima.

This paper proposes a new method for blind MIMO FIR system identification by utilizing the second-order statistics of the received data. By exploiting the new derived properties of the companion matrices that are constructed from the inherent structural relationship between source autocorrelation matrices, we provide an original proof for the uniqueness of the system solution, which servers as a theoretical basis for our new method. We adopt the following notations. 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 sub-diagonal 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} entry equal to one, and its other entries equal to zero. Note that the dimensions of $\mathbf{J}_1$, $\mathbf{J}^1$ and $\mathbf{e}_i$ are not specified here but dependent on the exact place where they are used. Let $\mathbb{C}^{n\times m}$ and $\mathbb{C}^n$ denote the set of $n\times m$ matrices and the set of n-dimensional column vectors with complex entries, respectively.

II. System Model and Basic Assumptions

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

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

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

$$\vec{\mathbf{x}}(n) = \sum_{i=1}^{p}\mathcal{H}_i\vec{\mathbf{s}}_i(n) + \vec{\mathbf{w}}(n) = \mathcal{H}\vec{\mathbf{s}}(n) + \vec{\mathbf{w}}(n) \tag{2}$$

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

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

$$\mathcal{H} \triangleq \left[\begin{array}{cccc} \mathcal{H}_1 & \mathcal{H}_2 & \cdots & \mathcal{H}_p \end{array} \right]$$

$$\vec{\mathbf{s}}(n) \triangleq \left[\begin{array}{cccc} \vec{\mathbf{s}}_1^T(n) & \vec{\mathbf{s}}_2^T(n) & \cdots & \vec{\mathbf{s}}_p^T(n) \end{array} \right]^T$$

Some basic assumptions are adopted as follows. 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 of each source is assumed to be known *a priori*. A4) The sources are zero-mean wide-sense stationary colored signals with the knowledge of input statistics being available. The sources are uncorrelated with each other. A5) Additive noises are spatially and temporally white noises, and they are statistically independent of the sources. As a consequence of A2, the MIMO channel matrix $\mathcal{H}$ is full column rank if the stack number N is chosen to satisfy $N + 1 \geq \sum_{i=1}^{p} L_i$ [9].

III. Proposed Channel Estimation Method

We begin by defining the source autocorrelation matrices as

$$\mathbf{R}_{s_i}[k] \triangleq E[\vec{\mathbf{s}}_i(n)\vec{\mathbf{s}}_i^H(n-k)] \tag{3}$$

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

By invoking the assumption A4, we have

$$\mathbf{R}_s[k] = \mathrm{diag}(\mathbf{R}_{s_1}[k], \mathbf{R}_{s_2}[k], \cdots, \mathbf{R}_{s_p}[k]) \tag{5}$$

Also, 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

$$\mathbf{R}_x[k] = \mathcal{H}\mathbf{R}_s[k]\mathcal{H}^H \tag{6}$$

Our goal is to find an estimate of $\mathcal{H}$ from Eqn.(6) 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

$$\mathbf{R}_x[k]\mathbf{R}_x^{\dagger}[0] = \mathcal{H}\mathbf{R}_s[k]\mathbf{R}_s^{-1}[0]\mathcal{H}^{\dagger} \tag{7}$$

$$\mathbf{R}_x[k]\mathbf{R}_x^{\dagger}[0]\mathcal{H} = \mathcal{H}\mathbf{R}_s[k]\mathbf{R}_s^{-1}[0] \tag{8}$$

Proof: To admit lemma 1, we need to prove

$$\mathbf{R}_x^{\dagger}[0] = (\mathcal{H}^H)^{\dagger}\mathbf{R}_s^{-1}[0]\mathcal{H}^{\dagger}$$

Typically, $\mathbf{A}^{\dagger}$ is defined to be the unique matrix $\mathbf{T}$ that satisfies the four *Moore-Penrose conditions*: [10]

$$\begin{array}{llll} \text{(i)} & \mathbf{ATA} = \mathbf{A} & \text{(iii)} & (\mathbf{AT})^H = \mathbf{AT} \\ \text{(ii)} & \mathbf{TAT} = \mathbf{T} & \text{(iv)} & (\mathbf{TA})^H = \mathbf{TA} \end{array}$$

Therefore we only need to prove that $\mathbf{R}_x^{\dagger}[0]$ defined above satisfies the four *Moore-Penrose conditions*. This can be easily proved and thus omitted here. ■

For convenience, let

$$\begin{array}{ll} \Upsilon_{2k-1} \triangleq \mathbf{R}_x[k]\mathbf{R}_x^{\dagger}[0] & \Upsilon_{2k} \triangleq \mathbf{R}_x[-k]\mathbf{R}_x^{\dagger}[0] \\ \Theta_{2k-1} \triangleq \mathbf{R}_s[k]\mathbf{R}_s^{-1}[0] & \Theta_{2k} \triangleq \mathbf{R}_s[-k]\mathbf{R}_s^{-1}[0] \end{array}$$

We can therefore re-express Eqn.(8) (choose $K \geq k \geq 1$) as

$$\Upsilon_l \mathcal{H} = \mathcal{H}\Theta_l \qquad \forall\ l \in \{1, \ldots, 2K\} \tag{9}$$

and further, for every $l \in \{1, \ldots, 2K\}$, we have the following by exploiting the block diagonal structure of $\Theta_l \triangleq \mathrm{diag}(\Theta_{l,1}, \Theta_{l,2}, \cdots, \Theta_{l,p})$

$$\Upsilon_l \mathcal{H}_i = \mathcal{H}_i \Theta_{l,i} \qquad \forall\ i \in \{1, \ldots, p\} \tag{10}$$

where $\Theta_{l,i} \triangleq \mathbf{R}_{s_i}[\bar{k}]\mathbf{R}_{s_i}^{-1}[0]$, $\bar{k} = (l+1)/2$ if l is odd and $\bar{k} = -l/2$ if l is even. For each $i \in \{1, \ldots, p\}$, the above equation can be used to identify the channel convolution matrix of user i, i.e. $\mathcal{H}_i$, since the knowledge of $\Theta_{l,i}$ is known *a priori* and the information of Υ_l can be obtained from the second-order statistics of the observed data. By exploiting the block Toeplitz structure of $\mathcal{H}_i$, we can rewrite Eqn.(10) as

$$\mathcal{T}_1[\Upsilon_l]\mathbf{h}_i = \mathcal{T}_2[\Theta_{l,i}]\mathbf{h}_i \tag{11}$$

where $\mathbf{h}_i \triangleq \left[\mathbf{h}_i^T(0) \ \ldots \ \mathbf{h}_i^T(L_i)\right]^T$, $\mathcal{T}_1[\cdot]$ and $\mathcal{T}_2[\cdot]$ respectively represent a certain transformation on the bracketed matrix. Therefore we may estimate $\mathbf{h}_i$ by the following criterion

$$\hat{\mathbf{h}}_i = \arg\min_{\|\mathbf{u}\|=1} \sum_{l=1}^{2K} \| \left[\ \mathcal{T}_1[\Upsilon_l] - \mathcal{T}_2[\Theta_{l,i}] \ \right] \mathbf{u} \|^2 \tag{12}$$

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 is trivial if the solution of Eqn.(11) is not unique, i.e. there exist other non-zero vectors, $\mathbf{g}_i$, that are *linearly independent of* $\mathbf{h}_i$ and also satisfy $\mathcal{T}_1[\Upsilon_l]\mathbf{g}_i = \mathcal{T}_2[\Theta_{l,i}]\mathbf{g}_i$ for any $l \in \{1, \ldots, 2K\}$. Hence we are faced with the following two problems. First, whether or not the solution of Eqn.(11) is unique (up to a scalar factor). Second, under what conditions the solution of Eqn.(11) will be unique. These two problems are studied in the following and we will establish the uniqueness of the solution to Eqn.(11) by utilizing only $\mathbf{R}_x[0]$ and $\mathbf{R}_x[\pm 1]$ and the knowledge of $\mathbf{R}_s[0]$ and $\mathbf{R}_s[\pm 1]$, i.e. the uniqueness of the solution can be guaranteed by choosing $l = 1, 2$ in Eqn.(11).

We begin by exploiting the inherent structural relationship between $\mathbf{R}_{s_i}[0]$ and $\mathbf{R}_{s_i}[\pm 1]$ for any $i \in \{1, \ldots, p\}$. It can be readily seen that for each source s_i, the last $d_i - 1$ rows of $\mathbf{R}_{s_i}[1]$ are the first $d_i - 1$ rows of $\mathbf{R}_{s_i}[0]$, and the first $d_i - 1$ rows of $\mathbf{R}_{s_i}[-1]$ are the last $d_i - 1$ rows of $\mathbf{R}_{s_i}[0]$. Hence we can establish the following relationship

$$\mathbf{R}_{s_i}[1] = \mathbf{J}_1\mathbf{R}_{s_i}[0] + \mathbf{e}_1\mathbf{r}_{i1}^H \tag{13}$$

$$\mathbf{R}_{s_i}[-1] = \mathbf{J}^1\mathbf{R}_{s_i}[0] + \mathbf{e}_{d_i}\mathbf{r}_{i2}^H \tag{14}$$

where $\mathbf{r}_{i1}^H \triangleq \mathbf{e}_1^H\mathbf{R}_{s_i}[1]$ and $\mathbf{r}_{i2}^H \triangleq \mathbf{e}_{d_i}^H\mathbf{R}_{s_i}[-1]$. Using Eqn.$(13-14)$, we can re-express $\Theta_{1,i}$ and $\Theta_{2,i}$ as follows

$$\Theta_{1i} \triangleq \mathbf{R}_{s_i}[1]\mathbf{R}_{s_i}^{-1}[0] \triangleq \mathbf{J}_1 - \mathbf{e}_1\vec{\alpha}_i^H \tag{15}$$

$$\Theta_{2i} \triangleq \mathbf{R}_{s_i}[-1]\mathbf{R}_{s_i}^{-1}[0] \triangleq \mathbf{J}^1 - \mathbf{e}_{d_i}\vec{\beta}_i^H \tag{16}$$

where $\Theta_{1,i}, \Theta_{2,i} \in \mathbb{C}^{d_i \times d_i}$ and $\vec{\alpha}_i$ and $\vec{\beta}_i$ can be obtained as

$$\vec{\alpha}_i = \begin{bmatrix} \alpha_{i,1} & \cdots & \alpha_{i,d_i} \end{bmatrix}^T = -\mathbf{R}_{s_i}^{-1}[0]\mathbf{r}_{i1} \tag{17}$$

$$\vec{\beta}_i = \begin{bmatrix} \beta_{i,1} & \cdots & \beta_{i,d_i} \end{bmatrix}^T = -\mathbf{R}_{s_i}^{-1}[0]\mathbf{r}_{i2} \tag{18}$$

Observe that for each $i \in \{1, \ldots, p\}$, both Θ_{1i} and Θ_{2i} are companion matrices. Due to their special structures, these companion matrices have some important properties we shall exploit in the following.

Lemma 2: Given that $\mathbf{Y} \in \mathbb{C}^{d_i \times d_j}$ satisfies the following two equations

$$(a)\ \Theta_{1i}\mathbf{Y} = \mathbf{Y}\Theta_{1j} \quad (b)\ \Theta_{2i}\mathbf{Y} = \mathbf{Y}\Theta_{2j} \tag{19}$$

where $i, j \in \{1, \ldots, p\}$ and the modulus of the last entry in $\vec{\alpha}_j$ is not equal to one, i.e. $|\alpha_{j,d_j}| \neq 1$, we have

- If $d_i = d_j$, $\Theta_{1i} = \Theta_{1j}$ and $\Theta_{2i} = \Theta_{2j}$, then $\mathbf{Y} = \lambda\mathbf{I}$, where λ could be any complex scalar including zero.
- If $d_i = d_j$, $\Theta_{1i} \neq \Theta_{1j}$ and $\Theta_{2i} \neq \Theta_{2j}$, then $\mathbf{Y} = \mathbf{0}$.
- If $d_i > d_j$, then $\mathbf{Y} = \mathbf{0}$.
- If $d_i < d_j$, and $|\alpha_{i,m_i}| \neq |\alpha_{j,d_j - t_i}|$, where $t_i \triangleq d_i - m_i$, α_{i,m_i} denotes the last non-zero entry in $\vec{\alpha}_i$, then $\mathbf{Y} = \mathbf{0}$. The condition $|\alpha_{i,m_i}| \neq |\alpha_{j,d_j - t_i}|$ can be removed if there exists a non-zero entry for $\alpha_{j,k}, k \in \{d_j - t_i + 1, \ldots, d_j\}$.

Proof: This can be proved from first principles by considering the entries of $\mathbf{G}_1 \triangleq \Theta_{1i}\mathbf{Y} = \mathbf{Y}\Theta_{1j}$ and $\mathbf{G}_2 \triangleq \Theta_{2i}\mathbf{Y} = \mathbf{Y}\Theta_{2j}$, and using the relationships in Eqn.$(13-18)$. The details can be found in [11] and will not be given here. Interested readers can obtain this technique report by contacting the authors. ∎

The significance of Lemma 2 not only lies in the fact that it provides a theoretical basis for our original proof for the uniqueness of the system solution, but also it establishes the identifiability conditions imposed on the input colored sources. We describe these identifiability conditions as follows

- *IC*1) The modulus of the last entry in each $\vec{\alpha}_i$ is not equal to one, i.e. $|\alpha_{i,d_i}| \neq 1$ for each $i \in \{1, \ldots, p\}$. This condition can be guaranteed if for each user s_i, the source autocorrelation matrix $\mathbf{R}_{s_i}[0]$ is positive definite. This is because $|\alpha_{i,d_i}|$ will be strictly less than one under the assumption that $\mathbf{R}_{s_i}[0]$ is positive definite [6].
- *IC*2) Sources have distinct second order statistics (power spectrum) such that for each pair of sources $\{s_i, s_j\}$, where $d_j \geq d_i$, the corresponding $\{\vec{\alpha}_i, \vec{\alpha}_j\}$ *does not* satisfy the following two conditions simultaneously

$$\begin{array}{ll} \text{(i)} & |\alpha_{i,m_i}| = |\alpha_{j,d_j - t_i}| \\ \text{(ii)} & \alpha_{j,k} = 0 \qquad \forall\ k \in \{d_j - t_i + 1, \ldots, d_j\} \end{array}$$

  This identifiability condition can be assured with probability one in practice when the sources have distinct second order statistics.

We now prove the uniqueness of the system solution to Eqn.(11) by utilizing the above lemma. We, firstly, prove that the solution to Eqn.(10) is unique (up to a scalar factor). The problem is formulated as follows: Given that (note that the following two equations are directly from Eqn.(7))

$$(a)\ \Upsilon_1 = \mathcal{H}\Theta_1\mathcal{H}^\dagger \quad (b)\ \Upsilon_2 = \mathcal{H}\Theta_2\mathcal{H}^\dagger \tag{20}$$

If $\mathcal{H}$ is full column rank and the input colored sources satisfy the identifiability conditions IC1–IC2, we need to prove that for each $i \in \{1, \ldots, p\}$, any non-zero matrix $\mathcal{G}_i$ that has the same block Toeplitz structure as $\mathcal{H}_i$ and also satisfies Eqn.(10) for $l = 1, 2$, i.e. $\Upsilon_1\mathcal{G}_i = \mathcal{G}_i\Theta_{1,i}$ and $\Upsilon_2\mathcal{G}_i = \mathcal{G}_i\Theta_{2,i}$, can be written as $\mathcal{G}_i = \lambda_i\mathcal{H}_i$, where λ_i is a non-zero complex scalar.

Proof: Suppose a non-zero matrix $\mathcal{G}_i \in \mathbb{C}^{(N+1)q \times d_i}$ with the same block Toeplitz structure as $\mathcal{H}_i$ also satisfies Eqn.(10) for $l = 1, 2$, then we have

$$\Upsilon_1\mathcal{G}_i = \mathcal{G}_i\Theta_{1i} \Rightarrow \mathcal{H}\Theta_1\mathcal{H}^\dagger\mathcal{G}_i = \mathcal{G}_i\Theta_{1i} \Rightarrow \Theta_1\mathcal{H}^\dagger\mathcal{G}_i = \mathcal{H}^\dagger\mathcal{G}_i\Theta_{1i}$$

$$\Upsilon_2\mathcal{G}_i = \mathcal{G}_i\Theta_{2i} \Rightarrow \mathcal{H}\Theta_2\mathcal{H}^\dagger\mathcal{G}_i = \mathcal{G}_i\Theta_{2i} \Rightarrow \Theta_2\mathcal{H}^\dagger\mathcal{G}_i = \mathcal{H}^\dagger\mathcal{G}_i\Theta_{2i}$$

Let $\mathbf{X} \triangleq \mathcal{H}^\dagger\mathcal{G}_i \triangleq \begin{bmatrix} \mathbf{X}_1^T & \cdots & \mathbf{X}_p^T \end{bmatrix}^T$, where $\mathbf{X}_k \in \mathbb{C}^{d_k \times d_i}$, then we have

$$\Theta_{1k}\mathbf{X}_k = \mathbf{X}_k\Theta_{1i}\ \ \forall\ k \in \{1, \ldots, p\} \tag{21}$$

$$\Theta_{2k}\mathbf{X}_k = \mathbf{X}_k\Theta_{2i}\ \ \forall\ k \in \{1, \ldots, p\} \tag{22}$$

Since the input sources satisfy the identifiability conditions IC1–IC2, by applying the results in Lemma 2, we know that $\mathbf{X}_k = \mathbf{0}$ for any $k \neq i$ and $\mathbf{X}_k = \lambda_i\mathbf{I}_i$ for $k = i$, i.e.

$$\mathcal{H}^\dagger\mathcal{G}_i = \begin{bmatrix} \mathbf{0} & \cdots & \lambda_i\mathbf{I}_i & \cdots & \mathbf{0} \end{bmatrix}^T \triangleq \lambda_i\mathbf{E}_i \tag{23}$$

Therefore we only need to prove that the solution of $\mathcal{G}_i$ that satisfies Eqn.(23) is unique and $\mathcal{G}_i = \lambda_i\mathcal{H}_i$. Notice that $\mathcal{G}_i$ has the same block Toeplitz structure as $\mathcal{H}_i$. If we write $\mathcal{H}^\dagger \triangleq [\mathbf{V}_0\ \cdots\ \mathbf{V}_N]$, we can transform $\mathcal{H}^\dagger\mathcal{G}_i = \lambda_i\mathbf{E}_i$ as

$$\mathcal{V}\begin{bmatrix} \mathbf{g}_i(0) \\ \vdots \\ \mathbf{g}_i(L_i) \end{bmatrix} = \mathrm{vec}(\lambda_i\mathbf{E}_i) \tag{24}$$

where $\mathcal{V} \in \mathbb{C}^{d_i(d_1 + \cdots + d_p) \times (L_i+1)q}$ 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 Eqn.(24) we know that $\mathcal{G}_i$ can be uniquely determined if $\mathcal{V}$ has full column rank. Recalling Theorem 1 in [12], $\mathcal{V}$ has full column rank if the following condition holds, i.e. there exists a nonzero z_0 (including ∞) such that the polynomial matrix $V(z_0)$ has full column rank, where $V(z_0) \triangleq \mathbf{V}_0 + \mathbf{V}_1 z^{-1} + \cdots + \mathbf{V}_N z^{-N}$. When channel order $L \geq 1$, $\mathcal{H}^\dagger$ can be considered to be a randomly generated matrix. We can assure that this mild condition can be satisfied

TABLE I: Symbol Error Rates (SER) Versus SNR

	$T_s = 1600$				$T_s = 1200$				$T_s = 800$			
SNR(dB)	Proposed Algorithm		SIF		Proposed Algorithm		SIF		Proposed Algorithm		SIF	
	user 1	user 2	user 1	user 2	user 1	user 2	user 1	user 2	user 1	user 2	user 1	user 2
25	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0016	0.0013
22	0.0000	0.0000	0.0009	0.0006	0.0001	0.0000	0.0037	0.0028	0.0009	0.0000	0.0159	0.0132
19	0.0018	0.0000	0.0220	0.0169	0.0017	0.0000	0.0296	0.0228	0.0034	0.0000	0.0490	0.0341
16	0.0043	0.0001	0.0769	0.0571	0.0074	0.0002	0.0874	0.0664	0.0221	0.0045	0.1007	0.0663
13	0.0299	0.0119	0.1472	0.1074	0.0522	0.0210	0.1637	0.1089	0.0994	0.0657	0.1725	0.1192
10	0.1313	0.1145	0.2132	0.1566	0.1924	0.1684	0.2280	0.1643	0.2865	0.2667	0.2475	0.1798

with probability one. Thus we can conclude that the solution of $\mathcal{G}_i$ is unique and $\mathcal{G}_i = \lambda_i \mathcal{H}_i$. Note that λ_i can not be zero because $\mathcal{G}_i$ would be zero under the condition $\lambda_i = 0$, which contradicts our previously made assumption $\mathcal{G}_i \neq \mathbf{0}$. The proof is completed here. ∎

Since Eqn.(10) and Eqn.(11) can be derived from each other, it implies that the solution to Eqn.(11) is unique up to a scaling constant of the "true" channel $\mathbf{h}_i$. Therefore $\mathbf{h}_i$ can be estimated by the criterion in Eqn.(12) with $K = 1$.

IV. Simulation Results

We now present simulation results to illustrate the performance of our proposed algorithm. We compare our method to the *second-order statistics isometry fitting* (SIF) method proposed in [2]. Our method directly estimate $\mathcal{H}$ by matching $\mathbf{R}_x[k]\mathbf{R}_x^{\dagger}[0]$ and $\mathcal{H}\mathbf{R}_s[k]\mathbf{R}_s^{-1}[0]\mathcal{H}^{\dagger}$, whilst [2] involves a two-step estimation algorithm. In our simulations, we consider $p = 2$ sources which are independent and identically distributed (i.i.d.) information sequences with 4-QPSK digital modulation format. To generate the colored sources, we pass these two i.i.d. information sequences through correlative filters prior to transmission. The correlative filters are chosen to be $f_1(z) = k_1(1 - \frac{1}{4}e^{-i\pi/2}z^{-1})$ and $f_2(z) = k_2(1 - \frac{1}{2}e^{i\pi/4}z^{-1})$ respectively for user 1 and user 2, where k_1 and k_2 are normalizing constants to ensure unit-power outputs. A wireless communication channel with these two colored user signals arriving at $q = 3$ antennas is randomly generated and given as

$$[\mathbf{h}(0) \ \cdots \ \mathbf{h}(3)] = \begin{bmatrix} -0.38 & 0.26 & -0.06 & -0.02 \\ 0.09 & 0.02 & -0.11 & 0.03 \\ -0.26 & -0.17 & 0.13 & 0.20 \\ -0.08 & 0.41 & 0.09 & -0.08 \\ 0.04 & -0.26 & -0.04 & 0.26 \\ -0.47 & 0.03 & -0.01 & 0.24 \end{bmatrix}$$

where $\mathbf{h}(l) \triangleq \left[\mathbf{h}_1^T(l) \ \ \mathbf{h}_2^T(l)\right]^T$. In the simulations, the channel order is assumed known *a priori* and the stack number (smoothed factor) N is chosen to be 6. Once the channel has been estimated, the minimum mean-square-error (MMSE) equalizers can be computed as $\mathcal{E}_{\mathrm{MMSE}} = \hat{\mathcal{H}}^{\dagger}(\mathbf{I} - \sigma_w^2 \hat{\mathbf{R}}_x^{-1}[0])$. The equalizers per user with equalization delay equal to 3 are used. The scalar ambiguity of equalizers is removed before we perform the equalization. After channel equalization, the *filtered information sequences* (the outputs after the information sequences passing through the correlative filters) of each source are recovered and we can further detect the information sequences by adopting the Viterbi algorithm-based maximum likelihood detector. We present the equalization performance of the algorithms in the following table. The results are averaged over 200 Monte Carlo runs. In Table I, we show the symbol error rates (SER) associated to the two sources as a function of SNR and the number of data samples T_s used for channel estimation. It can be seen that our proposed algorithm presents a clear advantage over SIF in terms of SER.

V. Conclusion

We present a new SOS-based method that admits a closed-form solution for blind MIMO FIR channel estimation driven by colored sources. An original proof for the uniqueness of the closed-form system solution is provided by exploiting the inherent structural relationship between $\mathbf{R}_s[0]$ and $\mathbf{R}_s[\pm 1]$ and the derived properties of the companion matrices. Simulation results show that the new method compares favorably with the existing SOS-based method [2].

References

[1] J. Mannerkoski and V. Koivunen, "Autocorrelation properties of channel encoded sequences – Applicability to blind equaliztion," *IEEE Trans. Signal Processing*, vol. 43, pp. 3501–3506, Dec. 2000.

[2] J. Xavier, V. Barroso, and J. Moura, "Closed-form correlative coding blind identification of MIMO channels: Isometry fitting to second order statistics," *IEEE Trans. Signal Processing*, vol. 49, pp. 1073–1086, May 2001.

[3] Y. Hua, S. An, and Y. Xiang, "Blind identification of FIR channels by decorrelating subchannels," *IEEE Trans. Signal Processing*, vol. 51, pp. 1143–1155, May 2003.

[4] A. Gorokhov and P. Loubaton, "Subspace-based techniques for blind separation of convolutive mixtures with temporally correlated sources," *IEEE Trans. Circuits Syst. I*, vol. 44, pp. 813–820, Sept. 1997.

[5] T. Ma, Z. Ding, and S. F. Yau, "A two-stage algorithm for MIMO blind deconvolution of colored input signals," *IEEE Trans. Signal Processing*, vol. 48, pp. 1187–1192, Apr. 2000.

[6] R. Valcarce and S. Dasgupta, "Blind channel equalization with colored sources based on second-order statistics: A linear prediction approach," *IEEE Trans. Signal Processing*, vol. 49, pp. 2050–2059, Sept. 2001.

[7] K. H. Afkhamie and Z.-Q. Luo, "Blind identification of FIR systems driven by markov-like input signals," *IEEE Trans. Signal Processing*, vol. 48, pp. 1726–1736, June 2000.

[8] I. Bradaric, A. Petropulu, and K. Diamantaras, "Blind MIMO FIR channel identification based on second-order spectra correlations," *IEEE Trans. Signal Processing*, vol. 51, pp. 1668–1674, June 2003.

[9] K. A. Meraim, P. Loubaton, and E. Moulines, "A subspace algorithm for certain blind identification problems," *IEEE Trans. Inform. Theory*, vol. 43, pp. 499–511, Mar. 1997.

[10] G. Golub and C. van Loan, *Matrix Computations, 3rd ed.* Baltimore, MD: Johns Hopkins Univ. Press, 1996.

[11] J. Fang and A. R. Leyman, "Technical report no. 2004TR2/09/0221060041, I2R," Sept. 2004.

[12] J. Liang and Z. Ding, "Blind MIMO system identification based on cumulant subspace decomposition," *IEEE Trans. Signal Processing*, vol. 51, pp. 1457–1468, June 2003.