

# PARAFAC BASED BLIND ESTIMATION OF MIMO SYSTEMS WITH POSSIBLY MORE INPUTS THAN OUTPUTS

Yuanning Yu, Athina P. Petropulu

Electrical & Computer Eng. Dept., Drexel University, Philadelphia, PA 19104  
 {yuanning, athina}@cbis.ece.drexel.edu

## ABSTRACT

We consider the problem of frequency domain identification of a convolutive multiple-input multiple-output (MIMO) system driven by white, mutually independent unobservable inputs. In particular, we improve upon a method recently proposed in [1] that uses PARAFAC decomposition of a tensor that is formed based on third-order statistics of the system output. The approach of [1] utilizes only one slice of the output tensor to recover one row of the system response matrix. We here propose an approach that fully exploits the information in the output tensor. As a result, the proposed method not only achieves lower error values but also becomes applicable to MIMO systems with possibly more inputs than outputs. By combining two output tensors, we can make the approach applicable to more systems. We also extend the method to employ fourth-order statistics of the system output.

*keywords*-PARAFAC decomposition, high-order statistics, MIMO, blind estimation.

## 1. INTRODUCTION

We consider a linear time-invariant (LTI) convolutive multiple-input multiple-output (MIMO) system driven by  $N_i$  unobservable inputs. Our goal is to identify the system response based on the  $N_o$  system outputs. MIMO models arise frequently in speech processing, multi-access communications, multi track digital magnetic recording, and biomedical applications.

Among the possible approaches for identification of convolutive MIMO systems, frequency domain methods [4, 8], offer certain advantages over time domain ones; they do not require system length information, and also, their formulation can take advantage of existing results for the scalar MIMO problem. However, extra steps are required to resolve the frequency-dependent permutation, scaling and phase ambiguities.

Higher-order statistics (HOS) based MIMO methods [4, 7, 8], as opposed to second-order statistics based ones provide system information without requiring channel diversity, and also can deal with white inputs as long as they are non-Gaussian. The majority of HOS-based blind MIMO estimation methods require prewhitening in order to satisfy the unitary matrix requirement. However, this is a sensitive process as it tends to lengthen the global system response and thus increases complexity and estimation errors.

In [1], we showed how PARAFAC decomposition [2, 3, 5] of a HOS-based tensor of the system output can be applied to the frequency domain framework of [4] to avoid the need for prewhitening. However, the method in [1] is based on an iterative scheme that, in each iteration, uses only one slice of output tensor in order to recover one row of the channel response matrix. Here we propose an approach that fully exploits the information in the output tensor. As a

result, the proposed method not only achieves lower error values but also becomes applicable to MIMO systems with more inputs than outputs. We also extend the method to employ fourth-order statistics of the system output, thus making the proposed method applicable to a communications scenario.

## 2. PROBLEM FORMULATION

Let us consider a  $N_i$ -input  $N_o$ -output LTI system.

$$\mathbf{x}(k) = \sum_{l=0}^{L-1} \mathbf{h}(l)\mathbf{s}(k-l) + \mathbf{n}(k) \quad (1)$$

where  $\mathbf{s}(n)$  is a  $N_i$  by 1 source vector;  $\mathbf{x}(n)$  is  $N_o$  by 1 observation vector;  $\mathbf{n}(n)$  is the observation noise;  $\mathbf{h}(l)$  is the FIR MIMO system impulse response matrix whose  $(i, j)$  element is denoted by  $\{h_{ij}(n)\}_{1 \leq i \leq N_o, 1 \leq j \leq N_i}$ ;  $L$  is the length of the longest  $h_{ij}(n)$ .

Let  $\mathbf{H}(k)$  be an  $N_o \times N_i$  matrix defined as the  $N$ -point Discrete Fourier Transform of  $\mathbf{h}(n)$ , i.e.,  $\mathbf{H}(k) = \sum_{n=0}^{L-1} \mathbf{h}(n)e^{-j\frac{2\pi}{N}kn}$ ,  $k = 0, \dots, N-1$  where  $N > L$ .

Our goal is to obtain an estimate of  $\mathbf{H}(k)$  within a column permutation ambiguity  $\mathbf{P}$ , a constant diagonal scalar ambiguity  $\mathbf{\Lambda}$ , and a linear phase term  $e^{j\frac{2\pi}{N}k\mathbf{M}}$  ( $\mathbf{M}$  is a diagonal matrix with integer elements), i.e.:

$$\hat{\mathbf{H}}(k) = \mathbf{H}(k)\mathbf{P}\mathbf{\Lambda}e^{j\frac{2\pi}{N}k\mathbf{M}} \quad (2)$$

*Notation/Definitions:* Let  $\mathcal{C}^n$  denote the  $n$ -way tensor;  $\mathbf{C}$  denotes a matrix;  $k_A$  denotes the  $k$ -rank of matrix  $\mathbf{A}$ . The matrix  $\mathbf{A}$  of size  $I \times J$  has  $k$ -rank  $k_A = l$  if every  $l \leq J$  columns of  $\mathbf{A}$  are linearly independent, but either  $l = J$ , or there exist a collection of  $l+1$  linearly dependent columns in  $\mathbf{A}$ .

## 3. PROPOSED METHODS

### 3.1. Third-order statistics based approach

*Assumptions:* (A1) Each  $s_i(\cdot)$  is a zero mean, non-symmetrically distributed, independent identically distributed (i.i.d.), stationary process with nonzero skewness. The  $s_i$ 's are mutually independent. (A2)  $n_i(\cdot)$ ,  $i = 1, \dots, N_o$  are zero mean Gaussian stationary random processes with variance  $\sigma_n^2$ , mutually independent and independent of the inputs. (A3) The  $k$ -rank of  $\mathbf{H}(k)$  satisfies:  $3k_H \geq 2N_i + 2$  for every  $k$ .

Assumption (A1) and (A2) are common in HOS based methods that involve third-order statistics. To get a sense of (A3), if the elements of the channel matrix are drawn independently from an absolutely continuous distribution, the  $\mathbf{H}(k)$  has both rank and  $k$ -rank equal to  $\min(N_i, N_o)$  with probability one [9]. For such channel, assumption (A3) is satisfied if  $3\min(N_i, N_o) \geq 2N_i + 2$ . Under the

assumption  $N_o \geq N_i$  that was made in [1], (A3) is easily satisfied as long as  $N_o \geq 2$ .

The  $N \times N$  discrete-frequency cross-bispectrum of the outputs  $x_l(k)$ ,  $x_i^*(k)$ ,  $x_j(k)$  is the two-dimensional Discrete Fourier transform of the third order cross-cumulant[10], and equals:

$$C_{lij}^3(k_1, k_2) = \sum_{p=1}^{N_i} \gamma_{s_p}^3 H_{lp}(-k_1 - k_2) H_{ip}^*(-k_1) H_{jp}(k_2) \quad (3)$$

where  $k_1, k_2 = 0, \dots, N-1$ . For  $k_1 = -m + r\delta, k_2 = \delta$ , where  $m, r, \delta$  integers,  $C_{lij}^3(-m + r\delta, \delta)$  can be viewed as the  $(l, i, j)$ -th element of tensor  $\mathcal{C}^3(-m + r\delta, \delta)$  ( $N_o \times N_o \times N_o$ ). Let us define:

$$\mathbf{A}_r \triangleq \mathbf{H}(m - r\delta - \delta) \quad \mathbf{B}_r \triangleq \mathbf{H}^*(m - r\delta)\mathbf{\Gamma}^3 \quad \mathbf{C}_r \triangleq \mathbf{H}(\delta) \quad (4)$$

By fixing the index  $l$  we can rewrite (3) in matrix form as:

$$\mathbf{C}_l^3(-m + r\delta, \delta) = \mathbf{B}_r \mathbf{D}_l(\mathbf{A}_r) \mathbf{C}_r^T \quad (5)$$

In [1] it was shown that the  $l$ -th row of  $\mathbf{A}_r$  can be recovered from  $\mathbf{C}_l^3(-m + r\delta, \delta)$  in an iterative way. Here we propose the following approach for recovering the entire matrix  $\mathbf{A}_r$ .

Let us slice the tensor  $\mathcal{C}^3(-m + r\delta, \delta)$  differently by fixing the index  $i$  as:

$$\mathbf{C}_i^3(-m + r\delta, \delta) = \mathbf{C}_r \mathbf{D}_i(\mathbf{B}_r) \mathbf{A}_r^T \quad (6)$$

we can stack the matrices  $\mathbf{C}_i^3(-m + r\delta, \delta)$  for  $i = 1, \dots, N_o$  to form a tall matrix  $\mathbf{U}_A(r)$ . It holds:

$$\mathbf{U}_A(r) = (\mathbf{B}_r \odot \mathbf{C}_r) \mathbf{A}_r^T, \quad (N_o^2 \times N_o) \quad (7)$$

where  $\odot$  is the Khatri-Rao (column-wise Kronecker) Product.

**Lemma 1:** The Khatri-Rao product  $(\mathbf{B}_r \odot \mathbf{C}_r)$  has a left inverse for all  $r = 0, \dots, N-1$  under (A3).

*Proof:* It has been proved in [9] that  $(\mathbf{B}_r \odot \mathbf{C}_r)$  has full column rank  $N_i$  once  $k_B + k_C \geq N_i + 1$ . By noting that  $k_A = k_H \leq N_i$  and  $k_A + k_B + k_C \geq 2N_i + 2$  (A3), we get  $k_B + k_C \geq N_i + 1$  and thus  $(\mathbf{B}_r \odot \mathbf{C}_r)$  has left inverse.

We can then solve (7) for  $\mathbf{A}_r$  as:

$$\mathbf{A}_r = ((\mathbf{B}_r \odot \mathbf{C}_r)^{-1} \mathbf{U}_A(r))^T \quad (8)$$

where  $(\mathbf{B}_r \odot \mathbf{C}_r)^{-1}$  is the left pseudo inverse of  $(\mathbf{B}_r \odot \mathbf{C}_r)$ .

By using the above structure, we avoid the need for the existence of the inverse of  $\mathbf{B}_r$  that was required in [1]. As it will be seen later, this makes the method applicable to systems with  $N_i > N_o$  once (A3) is satisfied.

To find the system response we need to compute  $\mathbf{A}_r$  for  $r = 0, \dots, N-1$ , we apply PARAFAC decomposition once to the tensor  $\mathcal{C}^3(-m, \delta)$ , and then use an iterative approach to find all the  $\mathbf{A}_r$ 's. Under assumption (A3), PARAFAC decomposition of  $\mathcal{C}^3(-m, \delta)$ , yields [2]:

$$\hat{\mathbf{A}}_0 \triangleq \mathbf{H}(m - \delta) \mathbf{P} \mathbf{\Lambda}_1 \quad \hat{\mathbf{B}}_0 \triangleq \mathbf{H}^*(m) \mathbf{\Gamma}^3 \mathbf{P} \mathbf{\Lambda}_2 \quad \hat{\mathbf{C}}_0 \triangleq \mathbf{H}(\delta) \mathbf{P} \mathbf{\Lambda}_3 \quad (9)$$

where  $\mathbf{P}$  is a permutation matrix and  $\mathbf{\Lambda}_i$  are complex diagonal matrices that satisfy:  $\mathbf{\Lambda}_2 \mathbf{\Lambda}_1 \mathbf{\Lambda}_3 = \mathbf{I}$ . Subsequently, by noting that  $\mathbf{B}_r = \mathbf{A}_r^* \mathbf{\Gamma}^3$ , let us define an iteration for  $r = 1, \dots, N$  as:

$$\begin{aligned} \hat{\mathbf{A}}(r) &= ((\hat{\mathbf{A}}^*(r-1) \odot \hat{\mathbf{C}}_0)^{-1} \mathbf{U}_A(r))^T \\ \hat{\mathbf{A}}(0) &= \hat{\mathbf{A}}_0 \end{aligned} \quad (10)$$

It can be shown that (see Appendix I):

$$\hat{\mathbf{A}}(r) = \mathbf{H}(m - r\delta - \delta) \mathbf{P} \mathbf{K}_{((r))_2} e^{j(\Phi_1 + r\Phi_2)}, \quad r = 0, \dots, N$$

where  $\Phi_1, \Phi_2$  are diagonal matrices,  $\mathbf{K}_1, \mathbf{K}_0$  are diagonal matrices with positive elements, and  $((\cdot))_2$  denotes modulo 2 operation.

Let us consider  $N$  to be even and co-prime to  $\delta$ . The phase ambiguity in eq. (11) can be solved in the same manner as in [1], which is included here for the readers' convenience. Define:

$$\begin{aligned} \hat{\mathbf{H}}(m - r\delta - \delta) &\triangleq \hat{\mathbf{A}}(r) [\hat{\mathbf{A}}^{-1}(N) \hat{\mathbf{A}}(0)]^{r/N} \\ &= \mathbf{H}(m - r\delta - \delta) \mathbf{P} \mathbf{K}_{((r))_2} e^{j(\Phi_1 + \frac{2\pi}{N} kr)} \end{aligned} \quad (11)$$

where  $k$  is an integer. Applying an  $N/2$  point IDFT on the even samples of  $\hat{\mathbf{H}}(k)$  of (11), we get an upsampled by  $\delta$  version of  $\mathbf{h}$ , circularly shifted by  $k$  and modulated due to the  $m - \delta$  term in (11). Once we extract the  $L$ -samples long segment (modulo  $N$ ) with the maximum energy based on its absolute value, we can cancel the modulating factor, and compute the amount of circular shift.

The steps consisting of applying equations (9),(10) and(11) define the proposed method, which we will refer to as the *Improved Single PARAFAC decomposition (ISPD) method*.

### 3.2. Extension to fourth-order statistics

**Assumptions:** In addition to (A2) we assume: (A4) Each  $s_i(\cdot)$  is a zero mean, i.i.d., stationary process with nonzero kurtosis. The  $s_i$ 's are mutually independent. (A5) The  $k$ -rank of  $\mathbf{H}(k)$  satisfies:  $4k_H \geq 2N_i + 3$  for every  $k$ .

Assumption (A4) requires that the fourth-order cumulants of the inputs are not identically zero. Assumption (A4), unlike assumption (A1), is satisfied by most communication signals. For a channel matrix with independent taps, assumption (A5) is satisfied if  $4\min(N_i, N_o) \geq 2N_i + 3$ . We can see that, for  $N_o \geq N_i$ , (A3) and (A6) are satisfied as long as  $N_i \geq 2$ .

Based on (A2) and (A4) the 4<sup>th</sup>-order discretized Trispectrum  $C_{ijlm}^4(k_1, k_2, k_3)$ , defined as the three dimensional DFT of the fourth order cross-cumulants, and equals [10]:

$$\sum_{p=1}^{N_i} \gamma_{s_i}^4 H_{ip}(-k_3 - (k_1 + k_2)) H_{jp}^*(-k_1) H_{lp}(k_2) H_{mp}^*(-k_3) \quad (12)$$

Let  $k_1 = -\delta, k_2 = 0, k_3 = -m + r\delta$ , and  $\mathcal{C}^4(-\delta, 0, -m + r\delta)$  denote the four-way tensor constructed by elements  $C_{ijlm}^4(-\delta, 0, -m + r\delta)$ . Let us define:

$$\mathbf{A}_r \triangleq \mathbf{H}(m - (r-1)\delta) \mathbf{\Gamma}^4 \quad (13)$$

$$\mathbf{B}_r \triangleq \mathbf{H}^*(\delta) \quad (14)$$

$$\mathbf{C}_r \triangleq \mathbf{H}(0) \quad (15)$$

$$\mathbf{D}_r \triangleq \mathbf{H}^*(m - r\delta) \quad (16)$$

For fixed  $i, j$ , let us stack the  $N_o^2$  matrices ( $N_o \times N_o$ )  $\mathbf{C}_{ij}^4(-\delta, 0, -m + r\delta)$ ,  $i, j = 1..N_o$ ) to form a  $N_o^3 \times N_o$  tall matrix  $\mathbf{U}_D^4(r)$ . It holds [2]:

$$\mathbf{U}_D^4(r) = (\mathbf{A}_r \odot \mathbf{B}_r \odot \mathbf{C}_r) \mathbf{D}_r^T \quad (17)$$

**Lemma 2:** The Khatri-Rao Product  $(\mathbf{A}_r \odot \mathbf{B}_r \odot \mathbf{C}_r)$  has a left inverse for all  $r = 0, \dots, N-1$  under (A5).

*Proof:* It is easy to shown that  $k_B + k_C \geq N_i + 1$  from (A5), thus  $(\mathbf{B}_r \odot \mathbf{C}_r)$  has full column rank, also full k-rank  $N_i$ . By noting that  $k_A = k_H \geq 1$ , we get  $k_A + k_{(B \odot C)} \geq N_i + 1$ , and thus  $(\mathbf{A}_r \odot \mathbf{B}_r \odot \mathbf{C}_r)$  has left inverse.

Under assumption (A5), the tensor  $\mathcal{C}^4(-\delta, 0, -m)$  can be decomposed by the PARAFAC algorithm into [2]:

$$\begin{aligned}\hat{\mathbf{A}}_0 &= \mathbf{A}_0 \mathbf{P} \mathbf{\Lambda}_{10} & \hat{\mathbf{B}}_0 &= \mathbf{B}_0 \mathbf{P} \mathbf{\Lambda}_{20} \\ \hat{\mathbf{C}}_0 &= \mathbf{C}_0 \mathbf{P} \mathbf{\Lambda}_{30} & \hat{\mathbf{D}}_0 &= \mathbf{D}_0 \mathbf{P} \mathbf{\Lambda}_{40}\end{aligned}\quad (18)$$

where  $\mathbf{\Lambda}_{10} \mathbf{\Lambda}_{20} \mathbf{\Lambda}_{30} \mathbf{\Lambda}_{40} = \mathbf{I}$ .

For  $r = 0, 1, \dots, N - 1$ , let us define the iteration:

$$\begin{aligned}\hat{\mathbf{D}}^4(r) &= ((\hat{\mathbf{D}}^{4*}(r-1) \odot \hat{\mathbf{B}}_0 \odot \hat{\mathbf{C}}_0)^{-1} \mathbf{U}_D^4(r))^T \\ \hat{\mathbf{D}}^4(0) &= \hat{\mathbf{D}}_0\end{aligned}\quad (19)$$

Similar to the third-order case, it holds:

$$\hat{\mathbf{D}}^{4*}(r) = \mathbf{H}(m - r\delta) \mathbf{P} \mathbf{S}_{((r)_2)} e^{j(-\Phi_1^4 + (r+1)\Phi_2^4)} \quad (20)$$

where  $\Phi_1^4, \Phi_2^4$  are constant diagonal matrices, and  $\mathbf{S}_1, \mathbf{S}_0$  are diagonal matrices with positive elements.

Similar to (11), equation (20) provides the even- or odd-indexed samples of the system frequency response within a phase and constant permutation and scalar ambiguities. We need to take  $\delta, N$  co-prime, and  $m$  as integer multiple of  $\delta$  to facilitate solving the trivial phase ambiguity. The phase ambiguity can be handled in exactly the same manner as in (11).

### 3.3. MIMO systems with more inputs than outputs

Even if  $N_i > N_o$ , it is still possible to apply the proposed ISPD method for certain  $N_i$  and  $N_o$  combinations as long as the relevant assumptions are satisfied, i.e., (A3) for the  $3^{rd}$  order case and (A5) for the  $4^{th}$  order case.

For the  $3^{rd}$  order case, if the channel matrix elements are random and independent, then the  $k$ -rank of  $\mathbf{H}(k)$  equals  $k_H = \min\{N_o, N_i\} = N_o$  and assumption (A4) becomes:

$$3k_H = 3N_o \geq 2N_i + 2, \quad (21)$$

Thus, the third-order based ISPD method can be applied to system in which  $N_o = 4, N_i = 5$ , or  $N_o = 6, N_i = 8$ .

Similar arguments apply to the  $4^{th}$  order case. MIMO system with  $N_o = 3, N_i = 4$  or  $N_o = 4, N_i = 6$  can be solved by the fourth-order based ISPD method.

The main constraint to get the system estimate is the assumptions to guarantee the uniqueness of the initial PARAFAC decomposition, i.e., (A3) for  $3^{rd}$  order case and (A5) for  $4^{th}$  order case.

Next we show for the  $3^{rd}$  order case how to combine two tensors  $\mathcal{C}^3(-m, \delta)$  and  $\mathcal{C}^3(-m + \delta, \delta)$  to construct a  $2N_o \times N_o \times N_o$  tensor, which can give us unique initial PARAFAC decomposition under the more relaxed condition  $2N_o \geq N_i + 2$  (instead of (A3)).

Along the lines of the definition of  $\mathbf{U}_A(r)$  (7), let us define  $\mathbf{U}_B(r)$  by fixing index  $j$  and stacking the matrices corresponding to  $j = 1, \dots, N_o$ . Also, let us define  $\mathbf{U}_C(r)$  by fixing index  $l$  and stacking the matrices corresponding to  $l = 1, \dots, N_o$ .

Similar to (8), and noting that  $\mathbf{B}_1 = \mathbf{A}_0^*$  and  $\mathbf{C}_0 = \mathbf{C}_1$ , we can get two  $N_o^2 \times N_o$  equations for  $\mathbf{B}_0$  and  $\mathbf{A}_1$  as:

$$\mathbf{B}_0 = ((\mathbf{C}_0 \odot \mathbf{A}_0)^{-1} \mathbf{U}_B(0))^T \quad (22)$$

$$\mathbf{A}_1 = ((\mathbf{A}_0^* \odot \mathbf{C}_0)^{-1} \mathbf{U}_A(1))^T \quad (23)$$

Noting that  $\mathbf{B}_1 = \mathbf{A}_0^*$  and  $\mathbf{C}_0 = \mathbf{C}_1$ , we can also get two  $2N_o^2 \times N_o$  equations for  $\mathbf{C}_0$  and  $\mathbf{A}_0$  as:

$$\mathbf{C}_0 = \left( \begin{pmatrix} \mathbf{A}_0 \odot \mathbf{B}_0 \\ \mathbf{A}_1 \odot \mathbf{A}_0^* \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{U}_C(0) \\ \mathbf{U}_C(1) \end{pmatrix} \right)^T \quad (24)$$

and

$$\mathbf{A}_0 = \left( \begin{pmatrix} \mathbf{B}_0 \odot \mathbf{C}_0 \\ \mathbf{C}_0^* \odot \mathbf{A}_1 \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{U}_A(0) \\ \mathbf{U}_B(1)^* \end{pmatrix} \right)^T \quad (25)$$

Now we have four equations (22-25) for the four unknowns:  $\mathbf{A}_0, \mathbf{B}_0, \mathbf{C}_0$  and  $\mathbf{A}_1$ . We can solve them in an Alternative Least Square (ALS) [9] manner as: every step update the four unknowns according to the four equations, and compute the fitting of the decomposed matrices and the tensors, stop when the convergence is slow. This is equivalent to solving a PARAFAC decomposition problem for a 3-way tensor of dimensions  $2N_o \times N_o \times N_o$ . As long as  $2N_o \geq N_i$ , we can reduce the assumption (A3) to  $2N_o \geq N_i + 2$ .

Thus, by combining two tensors, the third-order based ISPD can be applied to more systems with more inputs than outputs, i.e.,  $N_o = 3, N_i = 4$  or  $N_o = 4, N_i = 6$ .

The same idea can be applied to the  $4^{th}$  order case as well. By combining two consecutive tensors, we can have three  $2N_o^3 \times N_o$  updating equations for  $\hat{\mathbf{B}}_0, \hat{\mathbf{C}}_0$  and  $\hat{\mathbf{D}}_0$  (which is only different with  $\hat{\mathbf{A}}_1^*$  in a diagonal matrix  $\mathbf{\Gamma}^4$ ), and two  $N_o^3 \times N_o$  updating equations for  $\hat{\mathbf{A}}_0$  and  $\hat{\mathbf{D}}_1$ . This is equal to solving a PARAFAC decomposition for a 4-way tensor of dimension  $2N_o \times N_o \times N_o \times N_o$ . As long as  $2N_o \geq N_i$ , we can reduce the assumption (A5) to  $3N_o \geq N_i + 3$ , which can be satisfied with  $N_o = 2, N_i = 3$  or  $N_o = 3, N_i = 5$ .

### 3.4. Simulations

We next demonstrate the performance of the proposed approach.

We still consider the  $2 \times 2$  bandpass MIMO channels as in [1]:  $h_{ij}(n) = r_1 \cdot c(0.25(n - 10)) + r_2 \cdot c(0.25(n - 6)) + r_3 \cdot c(0.25(n - 8))$  where  $c(n)$  is a sinc function with delay  $n$  and the  $r_i$ 's are zero-mean Gaussian random variables. By varying the  $r_i$ 's we can generate multiple bandpass channels. We tested the performance for the proposed  $3^{rd}$  order ISPD method for 50  $2 \times 2$  channels of length  $L = 6$ , generated by varying the  $r_i$ 's randomly for each channel. We also provide comparison results between the proposed method against the SPD method of [1], the methods of [4], and [7]. The method of [7] is a deflation-type approach, where the input sequences are extracted and removed one by one and then the system is estimated based on the system output and the estimated input.

For each channel we performed 50 Monte Carlo runs. We took  $L_e = 10, T = 8000, SNR = 20dB$  in all methods. In all methods the output cross-cumulants were estimated using the same parameters. We use  $N = 128, m = \delta$  for the SPD and ISPD method, and  $\delta$  is chosen from the peak of the power-spectra.

In Fig. 1, we show the cumulative probability function for 50 random channel runs (50 MC runs each). The lower bound was generated for the same channel assuming that the input is known and obtaining the channel estimate by cross-correlating input and output. We can see the proposed ISPD has the best performance. For example, the probability that the ONMSE is less than 0.1 is significantly higher in ISPD.

In Table 1, we show the performance for the proposed  $4^{th}$  order ISPD method for 50  $2 \times 2$  channels of length  $L = 4$ . For each channel we performed 30 Monte Carlo runs. We took  $L_e = 6, T = 16000, N = 128, SNR = 20dB, m = \delta$ , and  $\delta$  is chosen from the peak of the power-spectra. Table 1 shows the mean, variance and the

cumulative distribution of the ONMSE for both the proposed method and also the lower bound. The ONMSE of the 4<sup>th</sup> order ISPD is less than 0.08 with probability higher than 0.9.

In Table. 2, we showed the performance for the proposed 3<sup>rd</sup> order ISPD method for 50  $N_o = 3, N_i = 4$  channels of length  $L = 6$ . For each channel we performed 50 Monte Carlo runs. We took  $L_e = 10, T = 8000, N = 128, SNR = 20dB, m = \delta$ , and  $\delta$  is chosen from the peak of the power-spectra.

**Table 1.** ONMSE Dist. for 4th-order ISPD for 50  $2 \times 2$  channels

ONMSE	mean	std	50%	90%	100%
ISPD 4 <sup>th</sup>	<b>0.0354</b>	<b>0.0012</b>	<b>0.019</b>	<b>0.08</b>	<b>0.11</b>
lower bound	<b>0.0016</b>	$0.4 \times 10^{-6}$	<b>0.0016</b>	<b>0.0018</b>	<b>0.002</b>

**Table 2.** ONMSE Dist. for 3rd-order ISPD for 50  $3 \times 4$  channels

ONMSE	mean	std	50%	90%	100%
ISPD 3 <sup>rd</sup>	<b>0.214</b>	<b>0.009</b>	<b>0.18</b>	<b>0.29</b>	<b>0.36</b>
lower bound	<b>0.005</b>	$0.6 \times 10^{-6}$	<b>0.005</b>	<b>0.0055</b>	<b>0.006</b>

#### 4. CONCLUSION

We presented a robust iterative scheme (ISPD) under the frequency domain framework of [1] for the identification of a multiple-input multiple-output (MIMO) system driven by white, mutually independent unobservable inputs. The proposed ISPD method, can achieve lower ONMSE value compare with the SPD method of [1], the methods of [4] and [7]. And the ISPD can be applied to certain systems with more inputs than outputs. Also we extended the proposed ISPD method to the 4<sup>th</sup> order case for communication scenario.

#### Appendix I

$$\begin{aligned}
\mathbf{A}(1) &= ((\mathbf{A}^*(0) \odot \hat{\mathbf{C}}_0)^{-1} \mathbf{U}_A(1))^T \\
&= ((\mathbf{A}^*_0 \mathbf{P}^*_0 \mathbf{\Lambda}^*_{10} \odot \mathbf{C}_0 \mathbf{P}_0 \mathbf{\Lambda}_{30})^{-1} \mathbf{U}_A(1))^T \\
&= ((\mathbf{B}_1(\Gamma_3)^{-1} \mathbf{P}_0 \mathbf{\Lambda}^*_{10} \odot \mathbf{C}_1 \mathbf{P}_0 \mathbf{\Lambda}_{30})^{-1} \mathbf{U}_A(1))^T \\
&= ((\mathbf{B}_1 \mathbf{P}_0 (\Gamma_p^3)^{-1} \mathbf{\Lambda}^*_{10} \odot \mathbf{C}_1 \mathbf{P}_0 \mathbf{\Lambda}_{30})^{-1} \mathbf{U}_A(1))^T \\
&= (((\mathbf{B}_1 \odot \mathbf{C}_1) \mathbf{P}_0 (\Gamma_p^3)^{-1} \mathbf{\Lambda}^*_{10} \mathbf{\Lambda}_{30})^{-1} \mathbf{U}_A(1))^T \\
&= ((\mathbf{P}_0 (\Gamma_p^3)^{-1} \mathbf{\Lambda}^*_{10} \mathbf{\Lambda}_{30})^{-1} (\mathbf{B}_1 \odot \mathbf{C}_1)^{-1} \mathbf{U}_A(1))^T \\
&= (\mathbf{\Lambda}_{30}^{-1} (\mathbf{\Lambda}_{10}^*)^{-1} \Gamma_p^3 \mathbf{P}_0^T \mathbf{A}_1^T)^T \\
&= \mathbf{A}_1 \mathbf{P}_0 | \mathbf{\Lambda}_{20} | \Gamma_p^3 e^{j(\Phi_{10} - \Phi_{30})}
\end{aligned}$$

Similarly, based on  $\mathbf{C}_l^3(-m + r\delta, \delta)$  we can get:

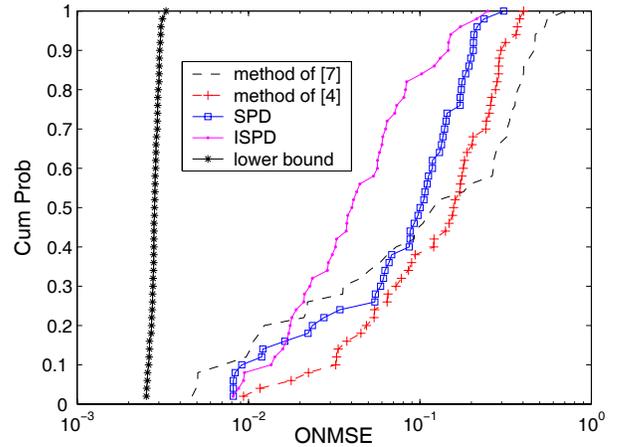
$$\mathbf{A}(r) = \mathbf{H}(m - r\delta - \delta) \mathbf{P}_0 \mathbf{K}_r e^{j(\Phi_{10} + r(\Phi_{\Gamma_p^3} - \Phi_{30}))}, \quad (26)$$

where

$$\mathbf{K}_r = \begin{cases} |\mathbf{\Lambda}_{20}| |\Gamma_p^3| & \text{for } r \text{ odd} \\ |\mathbf{\Lambda}_{10}| & \text{for } r \text{ even} \end{cases} \quad (27)$$

#### A. REFERENCES

- [1] Y. Yu, A.P. Petropulu, "Blind MIMO system estimation based on PARAFAC decomposition of HOS tensors" *IEEE Workshop on Statistical Signal Processing 2005*, July, 2005.
- [2] R. A. Harshman and M. E. Lundy, "The PARAFAC model for three-way factor analysis and multidimensional scaling," in *Res. Methods Multimode Data Anal.*, New York: Praeger, 1984, pp 122-215.
- [3] R. B. Cattell, "Parallel Proportional Profiles" and other principles for determining the choice of factors by rotation. *Psychometrika*, vol.9, pp.267-283, 1944.
- [4] B. Chen, A. P. Petropulu, "Frequency Domain Blind MIMO System Identification Based on Second and Higher Order Statistics," *IEEE Trans. Signal Processing*, vol. 49, pp. 1677-1688, Aug. 2001.
- [5] J. B. Kruskal, "Three-Way Arrays: Rank and Uniqueness of Trilinear Decompositions, with Application to Arithmetic Complexity and Statistics," *Linear Algebra and Its Applications*, 18:95-138, 1977.
- [6] G. B. Giannakis, Y. Hua, P. Stoica, L. Tong, "Signal Processing Advances in Wireless & Mobile Communications" vol. 2. Upper Saddle River, NJ: Prentice-Hall, 2001.
- [7] J. K. Tugnait, "Identification and deconvolution of multichannel linear non-Gaussian processes using higher order statistics and inverse filter criteria," *IEEE Trans. Signal Processing*, vol. 45, pp. 658-672, Mar. 1997.
- [8] M. Castella, J.C. Pesquet and A.P. Petropulu, "Family of Frequency and Time-Domain Contrasts for Blind Separation of Convolutional Mixtures of Temporally Dependent Signals," *IEEE Trans. on Signal Processing*, December 2004.
- [9] N. D. Sidiropoulos, G. B. Giannakis, R. Bro, "Parallel Factor Analysis in Sensor Array Processing" *IEEE Trans. Signal Processing*, vol. 48, No.8, pp. 2377-2388, Aug. 2000.
- [10] C. L. Nikias and A. P. Petropulu, *Higher Order Spectra Analysis*. Englewood Cliffs, NJ: Prentice-Hall, 1993.



**Fig. 1.** Comparison of Cumulative distribution of the ONMSEs of 50 channels with 2 sources and 2 sensors (3<sup>rd</sup> order)