Extension of the “Sequentially Drilled” Joint Congruence Transformation (SeDJoCo) Problem

Yao Cheng[1], *Arie Yeredor*[2], *Amir Weiss*[2] and *Martin Haardt*[1]

[1]Communications Research Laboratory
Ilmenau University of Technology
P.O. Box 10 05 65, D-98684 Ilmenau, Germany

[2]School of Electrical Engineering
Tel-Aviv University
P.O. Box 39040, Tel-Aviv 69978, Israel

***Abstract*—The Sequentially Drilled Joint Congruence (SeDJoCo) transformation has been identified as an important tool in the context of both Blind Source Separation (BSS) and closed-form Coordinated Beamforming (CBF) for the multi-user MIMO downlink. It can be interpreted as a new tensor decomposition. In this contribution, we introduce an extension of the SeDJoCo problem and address its links to the field of Independent Vector Analysis (IVA) and to CBF in a multi-cast framework. A solution of the extended SeDJoCo problem is proposed by extending the LU decomposition-based concept that is frequently used for joint matrix diagonalization. Data sets in an IVA setup are used in numerical simulations which show that the proposed algorithm is very efficient in finding the exact solution for the extended SeDJoCo problem, at least in a two-data-set case.**

I. Introduction

Tensor decompositions and joint matrix transformations are fundamental multi-linear- and linear-algebraic tools, which are closely related and play key roles in diverse signal processing fields, encompassing Blind Source Separation (BSS), Independent Component Analysis (ICA), Independent Vector Analysis (IVA), Big Data Analysis, and Multi-User Multiple Input Multiple Output (MU-MIMO) systems in wireless communications [1], [2], [3]. In the context of data analysis, whenever the observed measurements (signals) or some sample-statistics thereof can be arranged in the form of several (possibly ordered) sets of matrices, or as multi-way tensors, interesting internal structures of the data can be revealed by attempting to find low-rank (or otherwise succinctly parameterized) representations (or approximations) of these matrix sets or tensors.

For example, in the context of blind separation of a static mixture of stationary independent sources, the Second-Order Blind Identification (SOBI) algorithm [4] estimates the mixing matrix by applying Approximate Joint Diagonalization (AJD, see, e.g., [5], [6] and references therein) to a set of sample-correlation matrices of the observed mixtures, taken at different time-lags. In fact, the AJD of such a set of N matrices, each of dimensions $K \times K$, can also be thought of as a symmetric canonical decomposition (CANDECOMP, PARAFAC, or CP, e.g., [7], [8], [9]), representing (or approximating) the respective three-way $K \times K \times N$ tensor as the sum of K rank-1 tensors.

A few years ago, a special form of joint diagonalization, which was initially termed a “Hybrid Exact-Approximate joint Diagonalization” (HEAD, [11], [12]) was derived. When HEAD is applied to a set of some $N = K$ spatial generalized correlation matrices of dimensions $K \times K$, it yields the Maximum Likelihood (ML) estimate of the mixing matrix when the independent source signals are Gaussian, each with a different temporal covariance matrix. More specifically, assume the classical static mixture model $\boldsymbol{X} = \boldsymbol{AS}$, in which $\boldsymbol{S} = [\boldsymbol{s}_1 \quad \boldsymbol{s}_2 \quad \cdots \quad \boldsymbol{s}_K]^{\mathrm{T}} \in \mathbb{R}^{K\times T}$ denotes a matrix of K statistically independent source signals ($\boldsymbol{s}_1, \ldots, \boldsymbol{s}_K \in \mathbb{R}^T$), $\boldsymbol{A} \in \mathbb{R}^{K\times K}$ denotes an unknown (but invertible) mixing-matrix and $\boldsymbol{X} \in \mathbb{R}^{K\times T}$ denotes the matrix of observed mixture signals, from which it is desired to estimate the mixing matrix and, subsequently, the source signals. When the source signals are z ero-mean Gaussian and each has its own temporal (and known[1]) covariance matrix $\boldsymbol{C}_k = E[\boldsymbol{s}_k\boldsymbol{s}_k^{\mathrm{T}}] \in \mathbb{R}^{T\times T}$ (for $k = 1, \ldots, K$) it can be shown (see [12], [13] (chapter 7) or [14] for more details) that the ML estimate $\widehat{\boldsymbol{A}}$ of $\boldsymbol{A}$ can be obtained (up to permutation and sign ambiguities) as follows. First, construct K symmetric “target-matrices” as

$$\boldsymbol{Q}_k = \boldsymbol{X}\boldsymbol{C}_k^{-1}\boldsymbol{X}^{\mathrm{T}} \in \mathbb{R}^{K\times K} \quad , \quad k = 1, \ldots, K. \tag{1}$$

Then, look for a matrix $\widehat{\boldsymbol{A}} \in \mathbb{R}^{K\times K}$ which decomposes these target matrices as

$$\boldsymbol{Q}_k = \widehat{\boldsymbol{A}}\boldsymbol{D}_k\widehat{\boldsymbol{A}}^{\mathrm{T}} \quad , \quad k = 1, \ldots, K, \tag{2}$$

such that the matrices $\boldsymbol{D}_1, \ldots, \boldsymbol{D}_K \in \mathbb{R}^{K\times K}$ are not necessarily diagonal, but satisfy the property

$$\boldsymbol{D}_k\boldsymbol{e}_k = \boldsymbol{e}_k \quad , \quad k = 1, \ldots, K, \tag{3}$$

where the pinning vector $\boldsymbol{e}_k$ denotes the k-th column of the $K \times K$ identity matrix. In other words, the k-th column of the k-th matrix $\boldsymbol{D}_k$ must equal $\boldsymbol{e}_k$, namely, must be all-zeros except for a ‘1’ in its k-th element (and, since each $\boldsymbol{D}_k$ must be symmetric by construction, this also applies to its k-th row). We term this structure a “drilled” structure. Note that substitution with $\boldsymbol{B} \triangleq \boldsymbol{A}^{-1}$ (the ML estimate of the unmixing matrices) is also possible by reformulating the decomposition as

$$\widehat{\boldsymbol{B}}\boldsymbol{Q}_k\widehat{\boldsymbol{B}}^{\mathrm{T}} = \boldsymbol{D}_k \quad , \quad k = 1, \ldots, K. \tag{4}$$

Such a transformation is termed a “Sequentially Drilled Joint Congruence” (SeDJoCo) transformation. Fig. 1 illustrates the “drilled”

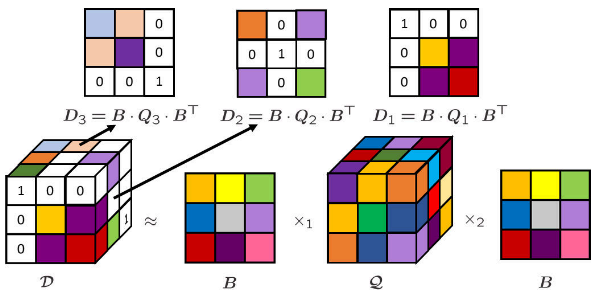

Fig. 1. Illustration of SeDJoCo as a tensor decomposition

structure of SeDJoCo and its interpretation as a tensor decomposition. Here $\boldsymbol{D}_1$, $\boldsymbol{D}_2$, and $\boldsymbol{D}_3$ are treated as the first, second, and third frontal slice of the tensor $\boldsymbol{\mathcal{D}}$, respectively. The target matrices $\boldsymbol{Q}_1$, $\boldsymbol{Q}_2$, and $\boldsymbol{Q}_3$ are associated with $\boldsymbol{\mathcal{Q}}$ in the same manner. Interestingly, in [15], [16] we have shown that the very same fundamental transformation is also useful in the context of closed-form (non-iterative) Coordinated

[1]Covariance matrices may be known in advance by studying statistical properties of unmixed sources, e.g., different musical instruments.

Beamforming (CBF), when a transmitter with K antennas transmits data to $N \leq K$ users, each user having K antennas as well, eliminating any inter-users interference. A very similar (complex-valued) formulation of the equation is obtained in this context, with the target matrices given by:

$$\boldsymbol{Q}_k = \boldsymbol{H}_k^{\mathrm{H}} \boldsymbol{H}_k \in \mathbb{C}^{K \times K} \quad , \quad k = 1, \ldots, K, \tag{5}$$

where the matrices $\boldsymbol{H}_k \in \mathbb{C}^{K \times K}$ denote the (flat fading) channel coefficients from each of the K transmit antennas to each of the K receive antennas of the k-th user ($k = 1, \ldots, K$). The solution matrix $\widehat{\boldsymbol{B}}$ in this context contains the desired transmission beamforming coefficients, such that its k-th row contains the coefficients required for transmission to the k-th user with no interference between users (see [15] for more details). Additionally, we obtained [15] important existence and uniqueness results, showing that for any set of K $K \times K$ symmetric (or Hermitian, in the complex-valued case), positive-definite (PD) target matrices a solution must exist, but may not be unique (even though the number of free parameters equals the number of equations; note that the equations are not linear in $\widehat{\boldsymbol{B}}$). However, when the target matrices are nearly jointly diagonalizable, namely admit an AJD solution with a good enough least squares (LS) fit, the SeDJoCo solution is unique. In addition, we considered (in [15]) several iterative algorithms for the solution of the generic fundamental SeDJoCo problem (for any such set of K $K \times K$ symmetric/Hermitian PD target-matrices, regardless of their origin or of the context of the problem).

II. The Extended Problem

In the context of Independent Vector Analysis (IVA, sometimes also termed "Joint BSS", see, e.g., [17], [18], [19], [20] and references therein), the SeDJoCo problem can be extended as follows. Consider M sets of static mixtures

$$\boldsymbol{X}^{(m)} = \boldsymbol{A}^{(m)} \boldsymbol{S}^{(m)} \quad , \quad m = 1, 2, \ldots, M \tag{6}$$

where $\boldsymbol{S}^{(m)} = [\boldsymbol{s}_1^{(m)} \ \cdots \ \boldsymbol{s}_K^{(m)}]^{\mathrm{T}} \in \mathbb{R}^{K \times T}$ denotes (for $m = 1, \ldots, M$) a matrix of K source signals of length T, belonging to the m-th set out of M such sets. In each set the source signals are mixed with an unknown respective mixing matrix $\boldsymbol{A}^{(m)} \in \mathbb{R}^{K \times K}$, and the observed mixture signals are given by $\boldsymbol{X}^{(m)} \in \mathbb{R}^{K \times T}$. Based on the observed mixtures sets $\boldsymbol{X}^{(1)}, \ldots, \boldsymbol{X}^{(M)}$, it is desired to estimate all M mixing matrices and thereby recover the source signals.

Like in standard ICA, the source signals in each set are statistically independent. When the source signals in different sets are also all statistically independent, this problem merely amounts to a set of M distinct, standard ICA problems. However, IVA allows statistical dependence between respective sources in different sets, such that the k-th row of $\boldsymbol{S}^{(m_1)}$ may depend on the k-th row of $\boldsymbol{S}^{(m_2)}$ (for all $1 \leq m_1, m_2 \leq M$). In other words, source signals with the same indices from different sets may be statistically dependent on each other (but still may not depend on source signals with a different index, either from their own set or from another set). This occurs, for example, in analyzing electroencephalogram (EEG) signals from multiple (M) subjects reacting to the same stimuli [21]: The sources within each subject's brain are assumed to be independent of each other, but when reacting to the same stimuli, dependence between corresponding source s associated with different subjects can be assumed. Another example is the separation of mixtures of color images [19], where there are $M = 3$ sets, corresponding to the red, green and blue image layers. While the mixed images are independent, different color layers of the same image are definitely dependent, giving rise to dependence between sets.

It turns out that in the (zero-mean) Gaussian case the likelihood equations for obtaining the ML estimate in the IVA problem require the solution of an "extended" SeDJoCo problem as follows. Let us first denote by $\boldsymbol{C}_k^{(m_1,m_2)} = E[\boldsymbol{s}_k^{(m_1)} (\boldsymbol{s}_k^{(m_2)})^{\mathrm{T}}] \in \mathbb{R}^{T \times T}$ the temporal covariance matrices between the k-th source at the m_1-th set and the k-th source at the m_2-th sets (for all $k = 1, \ldots, K$ and $m_1, m_2 = 1, \ldots, M$). Now define an augmented blocks-structured covariance matrix $\boldsymbol{C}_k \in \mathbb{R}^{MT \times MT}$

$$\boldsymbol{C}_k \triangleq \begin{bmatrix} \boldsymbol{C}_k^{(1,1)} & \cdots & \boldsymbol{C}_k^{(1,M)} \\ \vdots & \ddots & \vdots \\ \boldsymbol{C}_k^{(M,1)} & \cdots & \boldsymbol{C}_k^{(M,M)} \end{bmatrix} \tag{7}$$

and denote the respective block-partition of its inverse as:

$$\begin{bmatrix} \boldsymbol{P}_k^{(1,1)} & \cdots & \boldsymbol{P}_k^{(1,M)} \\ \vdots & \ddots & \vdots \\ \boldsymbol{P}_k^{(M,1)} & \cdots & \boldsymbol{P}_k^{(M,M)} \end{bmatrix} \triangleq \boldsymbol{C}_k^{-1} \quad k = 1, \ldots, K \tag{8}$$

thereby defining the KM^2 matrices $\boldsymbol{P}_k^{(m_1,m_2)} \in \mathbb{R}^{T \times T}$. Using the observed mixtures and these matrices, construct a set of KM^2 "target matrices" as

$$\boldsymbol{Q}_k^{(m_1,m_2)} = \frac{1}{T} \boldsymbol{X}^{(m_1)} \boldsymbol{P}_k^{(m_1,m_2)} (\boldsymbol{X}^{(m_2)})^{\mathrm{T}} \in \mathbb{R}^{K \times K} \quad , \qquad k = 1, \ldots, K, \ \ \forall m_1, m_2 = 1, \ldots, M. \tag{9}$$

The Likelihood Equations for the ML estimates $\widehat{\boldsymbol{B}}^{(m)}$ of the unmixing matrices (the inverses of the mixing matrices) turn out[2] to be:

$$\left[\sum_{m_2=1}^{M} \widehat{\boldsymbol{B}}^{(m_1)} \boldsymbol{Q}_k^{(m_1,m_2)} (\widehat{\boldsymbol{B}}^{(m_2)})^{\mathrm{T}} \right] \boldsymbol{e}_k = \boldsymbol{e}_k \quad , \qquad \forall k = 1, \ldots, K, \ \ \forall m_1 = 1, \ldots, M. \tag{10}$$

Like in standard SeDJoCo, this means that the k-th column of the matrix in the brackets should be "drilled" (namely be all-zeros except for its k-th element equaling 1), but unlike standard SeDJoCo, here these matrices are not necessarily symmetric, in general, so for the extended SeDJoCo the rows are generally not "drilled".

When the source signals in different sets are independent, the augmented covariance matrices $\boldsymbol{C}_k$, as well as their inverses, become block-diagonal, so that all $\boldsymbol{P}_k^{(m_1,m_2)}$, and therefore also all $\boldsymbol{Q}_k^{(m_1,m_2)}$, vanish for all $m_1 \neq m_2$. Consequently (as expected), the problem reduces to a set of M "standard" (and decoupled) SeDJoCo problems, since only one element (with $m_2 = m_1$) is left in each sum. In a "true" IVA setup, however, sources from different sets are correlated, giving rise to a non-trivial extended SeDJoCo problem.

Interestingly, a rather similar formulation can again be linked to a CBF problem, but in an extended multi-cast framework as follows. Assume a set of M inter-connected transmitters, each has K antennas and K associated subscribers (users). The users of each transmitter are indexed as $k = 1, 2, \ldots, K$, and there are K different, similarly indexed data streams to be multi-cast to the respective users (only). Namely, the k-th data stream is intended for user k of each transmitter (so there are M users receiving each data stream). Each user has L receive antennas.

We denote by $\boldsymbol{H}_k^{(m_1,m_2)} \in \mathbb{C}^{L \times K}$ the channel matrix (assuming flat fading) from the m_2-th transmitter to user k of the m_1-th transmitter. We further denote by $\boldsymbol{b}_k^{(m)} \in \mathbb{C}^K$ the transmission

[2]Due to space limitations the derivation is deferred to future publications.

weights (beamformer) placed by the m-th transmitter for transmitting the k-th data stream. We denote the resulting beamforming matrix as $\boldsymbol{B}^{(m)} \triangleq [\boldsymbol{b}_1^{(m)} \;\; \boldsymbol{b}_2^{(m)} \;\; \cdots \;\; \boldsymbol{b}_K^{(m)}]^{\mathrm{H}} \in \mathbb{C}^{K\times K}$. Thus, if we denote by $\boldsymbol{x} \in \mathbb{C}^K$ the data vector of K symbols (belonging to the K different data streams) to be multi-cast at each time-instance, the transmission of the m-th transmitter is given by $\boldsymbol{y}^{(m)} = (\boldsymbol{B}^{(m)})^{\mathrm{H}}\boldsymbol{x} \in \mathbb{C}^K$. Note that we use the same $\boldsymbol{x}$ for all transmitters.

An additional underlying assumption is that in addition to sharing the data, all M transmitters share all the Channel State Information (CSI) matrices, namely all $\boldsymbol{H}_k^{(m_1,m_2)}$ for all $m_1, m_2 = 1,\ldots,M$ and $k = 1,\ldots,K$. However, each user only knows its own CSI to its own transmitter, namely user k of transmitter m knows only $\boldsymbol{H}_k^{(m,m)}$. Accordingly, we assume that each user sets its receive antennas weights as a maximum ratio (matched filter) tuned to its own transmitter, namely as

$$\boldsymbol{w}_k^{(m)} = \boldsymbol{H}_k^{(m,m)}\boldsymbol{b}_k^{(m)} = \boldsymbol{H}_k^{(m,m)}(\boldsymbol{B}^{(m)})^{\mathrm{H}}\boldsymbol{e}_k \in \mathbb{C}^L. \tag{11}$$

The signal received by the k-th user of the m_2-th transmitter is then given by

$$\boldsymbol{y}_k^{(m_2)} = (\boldsymbol{w}_k^{(m_2)})^{\mathrm{H}}\left[\sum_{m_1=1}^{M}\boldsymbol{H}_k^{(m_2,m_1)}\boldsymbol{y}^{(m_1)} + \boldsymbol{v}_k^{(m_2)}\right] \in \mathbb{C}, \tag{12}$$

where $\boldsymbol{v}_k^{(m_2)} \in \mathbb{C}^L$ denotes the additive noise received by user k of transmitter m_2. Substituting the receive beamformer weights and the transmitted signals, this expression can also be written as

$$\begin{aligned}\boldsymbol{y}_k^{(m_2)} = \boldsymbol{e}_k^{\mathrm{T}}\boldsymbol{B}^{(m_2)}\sum_{m_1=1}^{M}(\boldsymbol{H}_k^{(m_2,m_2)})^{\mathrm{H}}\boldsymbol{H}_k^{(m_2,m_1)}(\boldsymbol{B}^{(m_1)})^{\mathrm{H}}\boldsymbol{x}\\ + (\boldsymbol{w}_k^{(m_2)})^{\mathrm{H}}\boldsymbol{v}_k^{(m_2)}.\end{aligned} \tag{13}$$

Ignoring the noise-related term, perfect interference cancellation is attained if and only if the term multiplying the data vector $\boldsymbol{x}$ equals $\boldsymbol{e}_k^{\mathrm{T}}$ (or a nonzero multiple thereof), which requires:

$$\begin{aligned}\sum_{m_1=1}^{M}\boldsymbol{B}^{(m_1)}(\boldsymbol{H}_k^{(m_2,m_1)})^{\mathrm{H}}\boldsymbol{H}_k^{(m_2,m_2)}(\boldsymbol{B}^{(m_2)})^{\mathrm{H}}\boldsymbol{e}_k = \gamma\boldsymbol{e}_k\\ \forall k = 1,\ldots,K \;\; \forall m_2 = 1,\ldots,M\end{aligned} \tag{14}$$

where γ accounts for possible scaling.

By defining target-matrices as

$$\boldsymbol{Q}_k^{(m_1,m_2)} \triangleq (\boldsymbol{H}_k^{(m_2,m_1)})^{\mathrm{H}}\boldsymbol{H}_k^{(m_2,m_2)} \in \mathbb{C}^{K\times K} \tag{15}$$

for $k = 1,\ldots,K$, it becomes evident that these equations resemble the extended SeDJoCo equations (10) above (with a change of the summation index; with a conjugate transpose replacing the transpose; and with the allowed scaling factor γ). The solution of the resulting extended SeDJoCo problem would enable interference free delivery of the intended data streams to all users.

III. To find solutions of extended SeDJoCo

Addressing the CBF version (14) of the equations and focusing on the real-valued case, we aim at minimizing the following cost function (a scaling constraint is imposed on all $\boldsymbol{B}^{(m)}$ by keeping their determinants constant throughout the update process, which consists of multiplications by unit-determinant matrices):

$$J = \sum_{k=1}^{K}\sum_{m_1=1}^{M}\sum_{m_2=m_1}^{M}\sum_{\substack{p=1\\p\neq k}}^{K}\boldsymbol{D}_k^{(m_1,m_2)^2}(p,k), \tag{16}$$

where $\boldsymbol{D}_k^{(m_1,m_2)} = \boldsymbol{B}^{(m_1)}\cdot\boldsymbol{Q}_k^{(m_1,m_2)}\cdot\boldsymbol{B}^{(m_2)\mathrm{T}}$. Similarly as in the standard SeDJoCo problem [15], we propose an LU-based scheme where triangular Jacobi matrices are used for the minimization of the cost function. Let us denote $\boldsymbol{L}^{(m)} \in \mathbb{R}^{K\times K}$ and $\boldsymbol{U}^{(m)} \in \mathbb{R}^{K\times K}$ as unit-determinant lower and upper triangular matrices with diagonal elements of one, respectively. Then the matrices $\boldsymbol{B}^{(m)}$ ($m = 1,2,\ldots,M$) are updated in an iterative manner

$$\boldsymbol{B}^{(m)} \leftarrow \boldsymbol{L}^{(m)}\cdot\boldsymbol{U}^{(m)}\cdot\boldsymbol{B}^{(m)}. \tag{17}$$

Instead of directly finding $\boldsymbol{L}^{(m)}$ and $\boldsymbol{U}^{(m)}$ that minimize J, the minimization problem can be solved by sequentially finding unit lower and upper triangular Jacobi matrices of dimensions $K\times K$, each minimizing J [15]. Let $\boldsymbol{L}_{i,j}(a^{(m)})$ represent a unit lower-triangular matrix that has $a^{(m)}$ at the position (i,j). On the other hand, $\boldsymbol{U}_{i,j}(a^{(m)})$ denotes a unit upper-triangular matrix that has $a^{(m)}$ at the position (i,j). The cost function with respect to $\boldsymbol{L}_{i,j}(a^{(m)})$ ($m = 1,2,\ldots,M$) or $\boldsymbol{U}_{i,j}(a^{(m)})$ ($m = 1,2,\ldots,M$) can be expressed as

$$J = \sum_{m=1}^{M}J^{(m)}(a^{(m)}) \tag{18}$$

where $J^{(m)}(a^{(m)}) = a^{(m)^2}\cdot b_2^{(m)} + a^{(m)}\cdot b_1^{(m)} + b_0^{(m)}$ with

$$\begin{aligned}b_2^{(m)} =& \sum_{\substack{p=1\\p\neq i}}^{K}\sum_{m_1=1}^{m}\boldsymbol{Q}_i^{(m_1,m)^2}(p,j) + \sum_{\substack{k=1\\k\neq i}}^{K}\sum_{m_2=m}^{M}\boldsymbol{Q}_k^{(m,m_2)^2}(j,k)\\ b_1^{(m)} =& 2\cdot\sum_{\substack{p=1\\p\neq i}}^{K}\sum_{m_1=1}^{m}\boldsymbol{Q}_i^{(m_1,m)}(p,j)\cdot\boldsymbol{Q}_i^{(m_1,m)}(p,i)\\ &+ 2\cdot\sum_{\substack{k=1\\k\neq i}}^{K}\sum_{m_2=m}^{M}\boldsymbol{Q}_k^{(m,m_2)}(j,k)\cdot\boldsymbol{Q}_k^{(m,m_2)}(i,k)\\ b_0^{(m)} =& \sum_{\substack{p=1\\p\neq i}}^{K}\sum_{m_1=1}^{m}\boldsymbol{Q}_i^{(m_1,m)^2}(p,i) + \sum_{\substack{k=1\\k\neq i}}^{K}\sum_{m_2=m}^{M}\boldsymbol{Q}_k^{(m,m_2)^2}(i,k).\end{aligned}$$

Note that the parameter $a^{(m)}$ is only associated with $J^{(m)}(a^{(m)})$, where $m = 1,2,\ldots,M$. Therefore, $a^{(m)}$ is obtained via the minimization of $J^{(m)}(a^{(m)})$ ($m = 1,2,\ldots,M$) [22]. Taking the derivative of $J^{(m)}(a^{(m)})$ and setting it to zero yield

$$a^{(m)} = -\frac{b_1^{(m)}}{-2\cdot b_2^{(m)}}. \tag{19}$$

To this end, we briefly summarize the proposed scheme. For the initialization, set $\boldsymbol{B}^{(m)} = \boldsymbol{I}_K$, $\boldsymbol{L}^{(m)} = \boldsymbol{I}_K$, and $\boldsymbol{U}^{(m)} = \boldsymbol{I}_K$, where $m = 1,2,\ldots,M$. In each iteration, for $i = 1,2,\ldots,K$ and $j = 1,2,\ldots,K$ ($i \neq j$), $a^{(m)}$ ($m = 1,2,\ldots,M$) are found via (19) that minimizes (18). Consequently, $\boldsymbol{L}_{i,j}(a^{(m)})$ (for $j < i$) or $\boldsymbol{U}_{i,j}(a^{(m)})$ (for $i < j$) is obtained. If $j < i$, the unit lower triangular matrix $\boldsymbol{L}^{(m)}$ is updated as

$$\boldsymbol{L}^{(m)} \leftarrow \boldsymbol{L}_{i,j}(a^{(m)})\cdot\boldsymbol{L}^{(m)},\;\; m = 1,2,\ldots,M, \tag{20}$$

whereas $\boldsymbol{Q}_k^{(m_1,m_2)}$ is updated via

$$\begin{aligned}\boldsymbol{Q}_k^{(m_1,m_2)} \leftarrow \boldsymbol{L}_{i,j}(a^{(m_1)})\cdot\boldsymbol{Q}_k^{(m_1,m_2)}\cdot\boldsymbol{L}_{i,j}^{\mathrm{T}}(a^{(m_2)})\\ k = 1,2,\ldots,K,\;\; m_1,\;\; m_2 = 1,2,\ldots,M.\end{aligned} \tag{21}$$

For the case where $i < j$, the lower triangular matrices in (20) and (21) are replaced by their upper triangular counterparts. At the end of each iteration, $\boldsymbol{B}^{(m)}$ ($m = 1,2,\ldots,M$) are updated via (17).

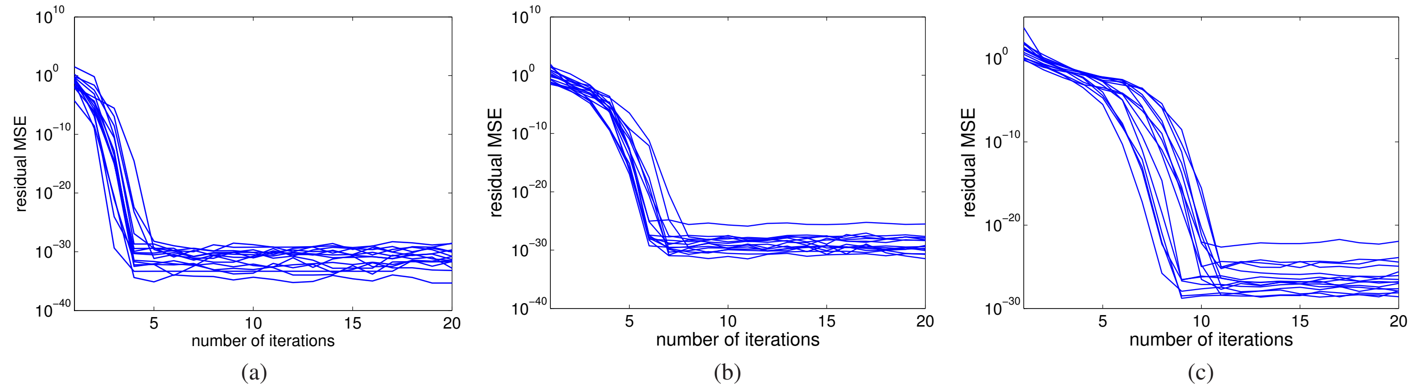

Fig. 2. Performance of the proposed solution to the extended SeDJoCo, with the intra-set target matrices removed (15 independent trials). (a): $M = 2$, $K = 3$; (b): $M = 2$, $K = 5$; (c): $M = 2$, $K = 10$.

IV. Simulaton reults

Now we evaluate the performance of the proposed solution of the extended SeDJoCo problem. Gaussian stationary sources groups with different spectra and cross-spectra in an IVA setup are used, and the set of target matrices are constructed via (9). In Fig. 2, it can be observed that when $M = 2$ and the intra-set statistical target matrices ($\boldsymbol{Q}_k^{(1,1)}$, $\boldsymbol{Q}_k^{(2,2)}$, $k = 1, 2, \ldots, K$) are removed, even with $K = 10$, the proposed solution of the extended SeDJoCo problem leads to very good performance in terms of convergence. The exact solution is always obtained within 15 iterations. Numerical simulations have also shown that as the number of sets M increases, a much larger number of iterations is required for the proposed scheme to find the exact solution of extended SeDJoCo. In addition, for a relatively small value of M, e.g., $M = 2$ as in the example shown here, when keeping the intra-set statistical target matrices, the proposed algorithm fails to provide satisfactory performance. Note that a similar phenomenon has been also observed for the generalized non-orthogonal joint diagonalization scheme in [22]. These observations motivate further enhancements of the proposed scheme for future work.

Acknowledgment

The authors gratefully acknowledge the financial support by the German-Israeli Foundation (GIF), grant number I-1282-406.10/2014.

References

[1] Q. H. Spencer, A. L. Swindlehurst and M. Haardt, "Zero-Forcing Method for Downlink Spatial Multiplexing in Multi-User MIMO Channels," *IEEE Transactions on Signal Processing*, vol. 52, no. 2, pp. 461-471, 2004.

[2] Q. H. Spencer, C. B. Peel, A. L. Swindlehurst and M. Haardt, "An Introduction to the Multi-User MIMO Downlink," *IEEE Communications Magazine*, pp. 60-67, 2004.

[3] V. Stankovic and M. Haardt, "Generalized Design of Multi-User MIMO Precoding Matrices," *IEEE Trans. on Wireless Communications*, vol. 7, pp. 953-961, 2008.

[4] A. Belouchrani, K. Abed-Meraim, J.-F. Cardoso and E. Moulines, "A blind source separation technique using second-order statistics," *IEEE Trans. on Signal Processing*, vol. 45, no. 2, pp. 434-444, 1997.

[5] A. Yeredor, "Non-Orthogonal Joint Diagonalization in the Least-Squares Sense with Application in Blind Source Separation," *IEEE Trans. on Signal Processing*, vol. 50, no. 7, pp. 1545-1553, 2002.

[6] P. Tichavský and A. Yeredor, "Fast Approximate Joint Diagonalization Incorporating Weight Matrices," *IEEE Trans. on Signal Processing*, vol. 57, no. 3, pp. 878-891, 2009.

[7] T. G. Kolda and B. W. Bader, "Tensor Decompositions and Applications," *Siam Review*, vol. 51, no. 3, pp. 455-500, 2009.

[8] F. Roemer and M. Haardt, "A Semi-Algebraic Framework for Approximate CP Decomposition via Simultaneous Matrix Diagonalization (SECSI)," *Signal Processing*, vol. 93, pp. 2722-2738, 2013.

[9] L. de Lathauwer, "A Link Between the Canonical Decomposition in Multilinear Algebra and Simultaneous Matrix Diagonalization," *SIAM Journal on Matrix Analysis and Applications*, vol. 28, no. 3, pp. 642-666, 2006.

[10] F. Roemer, C. Schroeter and M. Haardt, "A Semi-Algebraic Framework for Approximate CP Decompositions via Joint Matrix Diagonalization and Generalized Unfoldings," in *Proc. 46th Asilomar Conference on Signal, Systems and Compputers*, Pacific Grove, CA, USA, 2012.

[11] A. Yeredor, "On Hybrid Exact-Approximate Joint Diagonalization," in *Proc. IEEE CAMSAP 2009*, Aruba, Dutch Antilles, 2009.

[12] A. Yeredor, "Blind Separation of Gaussian Sources with General Covariance Structures: Bounds and Optimal Estimation," *IEEE Trans. on Signal Processing*, vol. 58, no. 10, pp. 5057-5068, 2010.

[13] P. Comon and C. Jutten, Eds., Handbook of Blind Source Separation, Academic Press, 2010.

[14] D.-T. Pham, "Blind Separation of Instantaneous Mixture of Sources via the Gaussian Mutual Information Criterion," *Signal Processing*, vol. 81, pp. 855-870, 2001.

[15] A. Yeredor, B. Song, F. Roemer and M. Haardt, "A "Sequentially Drilled" Joint Congruence (SeDJoCo) Transformation with Applications in Blind Source Separation and Multi-User MIMO Communication," *IEEE Transactions on Signal Processing*, vol. 60, no. 6, pp. 2744-2757, 2012.

[16] B. Song, F. Roemer and M. Haardt, "Using a New Structured Joint Congruence (STJOCO) Transformation of Hermitian Matrices for Precoding in Multi-User MIMO Systems," in *Proc. IEEE ICASSP'10*, Dallas, TX, 2010.

[17] T. Kim, I. Lee and T.-W. Lee, "Independent Vector Analysis: Definition and Algorithms," in *Proc. 40th Asilomar Conference on Signals, Systems and Computers*, pp. 1393-1396, 2006.

[18] J.-H. Lee, T.-W. Lee, F. A. Jolesz and S.-S. Yoo, "Independent Vector Analysis (IVA): Multivariate Approach for fMRI Group Study," *NeuroImage*, vol. 40, no. 1, pp. 86-109, 2008.

[19] J. Via, M. Anderson, X.-L. Li and T. Adali, "A Maximum Likelihood Approach for Independent Vector Analysis of Gaussian Data Sets," in *Proc. IEEE International Workshop on Machine Learning for Signal Processing (MLSP'11)*, pp. 1-6, 2011.

[20] M. Anderson, T. Adali and X.-L. Li, "Joint Blind Source Separation with Multivariate Gaussian Model: Algorithms and Performance Analysis," *IEEE Transactions on Signal Processing*, vol. 60, no. 4, pp. 1672-1683, 2012.

[21] J. Chatel-Goldman, M. Congedo and R. Phlypo, "Joint BSS as a natural analysis framework for EEG-hyperscanning," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'13)*, pp. 1212-1216, 2013.

[22] X. Gong, X. Wang, and Q. Lin, "Generalized non-orthogonal joint diagonalization with LU decomposition and successive rotations," *IEEE Trans. Signal Process.*, vol. 63, no. 5, pp. 1322 - 1334, 2015.