2017 IEEE 7th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)

Performance analysis of least-squares Khatri-Rao factorization

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

[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 least-squares Khatri-Rao factorization is regarded as an important linear- and multilinear-algebraic tool and finds applications in, for instance, computation of the CP decomposition and channel estimation for two-way relaying systems. We conduct a "first-order" perturbation analysis for it, which is a crucial step towards establishing analytical performance evaluation of various schemes employing the least-squares Khatri-Rao factorization. Numerical results validating our analytical performance analysis are shown. Being new advance in perturbation analysis on matrix decompositions, the performance analysis of the least-squares Khatri-Rao factorization presented in this paper will also contribute to a promising enhancement of the SECSI-GU framework, which is able to estimate the loading matrices in a CP decomposition, both efficiently and accurately.**

I. Introduction

Linear and multilinear algebra play a significant role in a variety of research fields, including, e.g., signal processing, wireless communications, and image processing. Many problems have been modeled into and in the end successfully solved by matrix and tensor decompositions [1]–[8]. The least-squares Khatri-Rao factorization is one of such crucial linear- and multilinear-algebraic tools, which has applications in the semi-algebraic framework for approximate CP decompositions via simultaneous matrix diagonalizations (SECSI) [9] via generalized unfoldings (SECSI-GU) [10], dual-symmetric parallel factor analysis [11], tensor-based channel estimation for two-way relaying systems [12], R-D parameter estimation [13], etc.. Consider the Khatri-Rao product of R matrices $\boldsymbol{F}_r \in \mathbb{C}^{M_r \times d}$, $r = 1, 2, \ldots, R$

$$\boldsymbol{F} = \boldsymbol{F}_1 \diamond \boldsymbol{F}_2 \diamond \ldots \diamond \boldsymbol{F}_R \in \mathbb{C}^{M \times d}, \tag{1}$$

where $M = \prod_{r=1}^{R} M_r$, and $\diamond$ symbolizes the Khatri-Rao (column-wise Kronecker) product. In many applications, a least-squares Khatri-Rao factorization is applied on the perturbed version of $\boldsymbol{F}$

$$\hat{\boldsymbol{F}} = \boldsymbol{F} + \Delta\boldsymbol{F}, \tag{2}$$

resulting from some measurement data, e.g., contaminated by noise. To employ the least-squares Khatri-Rao factorization to obtain the R factor matrices, we first take the s-th column of $\hat{\boldsymbol{F}} \in \mathbb{C}^{M \times d}$ denoted by $\hat{\boldsymbol{f}}^{(s)}$, $s = 1, 2, \ldots, d$, and arrange it into an R-way rank-1 tensor $\hat{\boldsymbol{\mathcal{F}}}^{(s)} \in \mathbb{C}^{M_1 \times \cdots \times M_R}$. Fig. 1 illustrates a noiseless case, where $R = 3$, $M_1 = M_2 = M_3 = 2$, $d = 2$. We use $\boldsymbol{f}^{(s)}$ and $\boldsymbol{\mathcal{F}}^{(s)}$ $(s = 1, 2)$ to represent the s-th column of $\boldsymbol{F}$ and the corresponding rank-1 tensor constructed from $\boldsymbol{f}^{(s)}$, respectively. Then the truncated higher-order SVD (HOSVD) of $\hat{\boldsymbol{\mathcal{F}}}^{(s)}$ is computed such that an estimate of the s-th column of $\boldsymbol{F}_r$, $\boldsymbol{f}_r^{(s)} \in \mathbb{C}^{M_r}$, $r = 1, \ldots, R$, is estimated as

$$\hat{\boldsymbol{f}}_r^{(s)} = \begin{cases} \hat{\boldsymbol{u}}_{(r,s)}^{[\mathrm{s}]}, & r \neq 1 \\ \hat{s}_{(s)}^{[\mathrm{s}]} \cdot \hat{\boldsymbol{u}}_{(r,s)}^{[\mathrm{s}]}, & r = 1 \end{cases} \tag{3}$$

where $\hat{\boldsymbol{u}}_{(r,s)}^{[\mathrm{s}]}$ is the first left singular vector of the r-mode unfolding of $\hat{\boldsymbol{\mathcal{F}}}^{(s)}$, and the scalar $\hat{s}_{(s)}^{[\mathrm{s}]}$ is in fact the only element of the core tensor from the truncated HOSVD given by

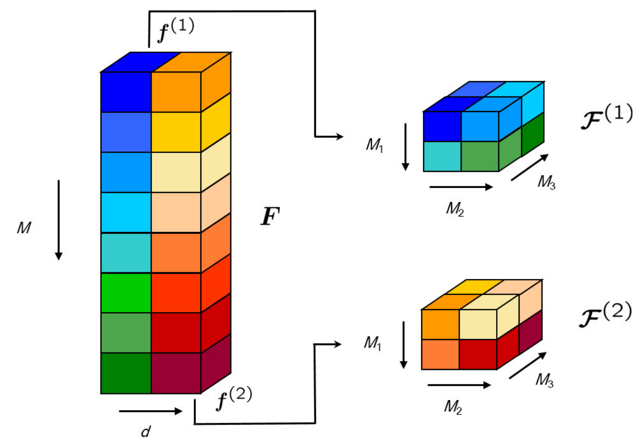

Fig. 1. Illustration of the first step of the least-squares Khatri-Rao factorization in an example, where $R = 3$, $M_1 = M_2 = M_3 = 2$, and $d = 2$

$$\hat{s}_{(s)}^{[\mathrm{s}]} = \hat{\boldsymbol{\mathcal{F}}}^{(s)} \times_1 \hat{\boldsymbol{u}}_{(1,s)}^{[\mathrm{s}]^{\mathrm{H}}} \times_2 \hat{\boldsymbol{u}}_{(2,s)}^{[\mathrm{s}]^{\mathrm{H}}} \times_3 \cdots \times_R \hat{\boldsymbol{u}}_{(R,s)}^{[\mathrm{s}]^{\mathrm{H}}}. \tag{4}$$

Note that throughout this paper, the r-mode product between an R-way tensor with size I_r along mode $r = 1, 2, \ldots, R$ represented as $\boldsymbol{\mathcal{A}} \in \mathbb{C}^{I_1 \times I_2 \times \ldots \times I_R}$ and a matrix $\boldsymbol{U} \in \mathbb{C}^{P_r \times I_r}$ is written as $\boldsymbol{\mathcal{A}} \times_r \boldsymbol{U}$. It is computed by multiplying all r-mode vectors of $\boldsymbol{\mathcal{A}}$ with $\boldsymbol{U}$, whereas the r-mode vectors of $\boldsymbol{\mathcal{A}}$ are obtained by varying the r-th index from 1 to I_r and keeping all other indices fixed. Aligning all r-mode vectors as the columns of a matrix yields the r-mode unfolding of $\boldsymbol{\mathcal{A}}$ which is denoted by $[\boldsymbol{\mathcal{A}}]_{(r)} \in \mathbb{C}^{I_r \times I_{r+1} \cdot \ldots \cdot I_R \cdot I_1 \cdot \ldots \cdot I_{r-1}}$. In other words, $[\boldsymbol{\mathcal{A}} \times_r \boldsymbol{U}]_{(r)} = \boldsymbol{U} \cdot [\boldsymbol{\mathcal{A}}]_{(r)}$. Here the reverse cyclical ordering of the columns, as proposed in [14], is used for the r-mode unfoldings. The tensor $\boldsymbol{\mathcal{I}}_{R,d}$ is an R-dimensional identity tensor of size $d \times d \times \ldots \times d$, which is equal to one if all R indices are equal and zero otherwise.

In recent years, the analytical performance analysis of matrix and tensor decompositions has attracted increased attention, and its value has been widely acknowledged [15]–[20]. For instance, the perturbation analysis of SVD and its application in the performance evaluation of subspace-based DOA estimation algorithms have been addressed in [15], [16], etc.. The prosperous development of multilinear signal processing, on the other hand, has sparked interest in the performance analysis of tensor decompositions, including the truncated HOSVD [19] and the approximate CP decomposition [20]. It is of theoretical and practical interest to conduct an analytical performance analysis on the least-squares Khatri-Rao factorization, which predicts the accuracy of the factor matrices estimation in the presence of perturbation. This will also serve as an indispensable part of an analytical performance evaluation for schemes, where the least-squares Khatri-Rao factorization is employed, such as SECSI-

978-1-5386-1251-4/17/$31.00

GU [10]. In this contribution, we present a "first-order" perturbation analysis of the least-squares Khatri-Rao factorization and address its key role in the design of an efficient version of SECSI-GU. As we show via numerical simulations, it leads to a very accurate prediction for the relative Mean Square Factor Error (rMSFE) of the least-squares Khatri-Rao factorization for a wide range of SNR values.

II. Closed-form expression of the relative Mean Square Factor Error (rMSFE)

In this section, we derive the closed-form expression of the rMSFE with respect to the r-th factor matrix $\boldsymbol{F}_r$. Let us assume that the covariance matrix of $\boldsymbol{n} = \text{vec}\{\Delta \boldsymbol{F}\} \in \mathbb{C}^{M \cdot d}$ is known. It is denoted as $\boldsymbol{R}_{\text{nn}} = \mathbb{E}\left\{\boldsymbol{n} \cdot \boldsymbol{n}^{\text{H}}\right\}$. Here $\text{vec}\{\cdot\}$ symbolizes the vectorization operation of its input matrix. The rMSFE with respect to the r-th factor matrix $\boldsymbol{F}_r$ is written as

$$\text{rMSFE}_{(r)} = \mathbb{E}\left\{ \frac{\left\| \boldsymbol{F}_r - \hat{\boldsymbol{F}}_r \cdot \boldsymbol{P}_r^{(\text{opt})} \right\|_{\text{F}}^2}{\|\boldsymbol{F}_r\|_{\text{F}}^2} \right\}, \tag{5}$$

where $\hat{\boldsymbol{F}}_r$ denotes an estimate of $\boldsymbol{F}_r$ obtained via the least-squares Khatri-Rao factorization, and $\boldsymbol{P}_r^{(\text{opt})}$ is a diagonal matrix that corrects the scaling ambiguity inherent in the estimation of the factor matrices. In addition, we use $\|\cdot\|_{\text{F}}$ to denote the Frobenius norm of a matrix. Note that the least-squares Khatri-Rao factorization estimates the factor matrices up to inevitable permutation and scaling ambiguities. For the performance analysis only the scaling ambiguity is relevant. A performance evaluation of the least-squares Khatri-Rao factorization, being either empirical or analytical as we present in this paper, uses the knowledge of the "true" factor matrices to resolve the aforementioned ambiguities such that relative errors in estimating the factor matrices are computed either numerically or predicted theoretically.

As shown in [20], we approximate $\left(\boldsymbol{F}_r - \hat{\boldsymbol{F}}_r \cdot \boldsymbol{P}_r^{(\text{opt})}\right)$ in the expression of the rMSFE given in (5) as

$$\begin{aligned} &\boldsymbol{F}_r - \hat{\boldsymbol{F}}_r \cdot \boldsymbol{P}_r^{(\text{opt})} \\ \approx &\boldsymbol{F}_r \cdot \text{Ddiag}\left(\boldsymbol{F}_r^{\text{H}} \cdot \Delta \boldsymbol{F}_r\right) \cdot \boldsymbol{D}^{-1} - \Delta \boldsymbol{F}_r, \end{aligned} \tag{6}$$

where $\text{Ddiag}\{\cdot\}$ denotes the operation of extracting the diagonal elements of the input matrix and then using them as the diagonal of a diagonal matrix, and $\boldsymbol{D} = \text{Ddiag}\{\boldsymbol{F}_r^{\text{H}} \cdot \boldsymbol{F}_r\}$. In other words, the s-th diagonal element of $\boldsymbol{D}$, where $s = 1, 2, \ldots, d$, can be expressed as ${\boldsymbol{f}_r^{(s)}}^{\text{H}} \cdot \boldsymbol{f}_r^{(s)}$ Consequently, the approximation of the s-th column of $\left(\boldsymbol{F}_r - \hat{\boldsymbol{F}}_r \cdot \boldsymbol{P}_r^{(\text{opt})}\right)$, represented by a short-hand notation $\Delta \widetilde{\boldsymbol{f}}_r^{(s)}$ is written as

$$\Delta \widetilde{\boldsymbol{f}}_r^{(s)} \approx \boldsymbol{f}_r^{(s)} \cdot \frac{{\boldsymbol{f}_r^{(s)}}^{\text{H}} \cdot \Delta \boldsymbol{f}_r^{(s)}}{{\boldsymbol{f}_r^{(s)}}^{\text{H}} \cdot \boldsymbol{f}_r^{(s)}} - \Delta \boldsymbol{f}_r^{(s)}, \tag{7}$$

where we use $\Delta \boldsymbol{f}_r^{(s)} \in \mathbb{C}^{M_r}$ to denote the s-th column of the perturbation of the r-th factor matrix $\Delta \boldsymbol{F}_r$. For a matrix $\boldsymbol{A}$ having d columns denoted by $\boldsymbol{a}^{(s)}$, $s = 1, 2, \ldots, d$

$$\|\boldsymbol{A}\|_{\text{F}}^2 = \sum_{s=1}^{d} \text{trace}\left\{\boldsymbol{a}^{(s)} \cdot {\boldsymbol{a}^{(s)}}^{\text{H}}\right\}. \tag{8}$$

We apply this to $\left\| \boldsymbol{F}_r - \hat{\boldsymbol{F}}_r \cdot \boldsymbol{P}_r^{(\text{opt})} \right\|_{\text{F}}^2$ such that the rMSFE with respect to the r-th factor matrix $\boldsymbol{F}_r$ can be computed as

$$\text{rMSFE}_{(r)} = \mathbb{E}\left\{ \frac{\sum_{s=1}^{d} \text{trace}\left\{\Delta \widetilde{\boldsymbol{f}}_r^{(s)} \cdot \Delta {\widetilde{\boldsymbol{f}}_r^{(s)}}^{\text{H}}\right\}}{\|\boldsymbol{F}_r\|_{\text{F}}^2} \right\}. \tag{9}$$

Therefore, in the following, we derive the s-th column of the $\Delta \boldsymbol{F}_r$, $\Delta \boldsymbol{f}_r^{(s)}$, then $\Delta \widetilde{\boldsymbol{f}}_r^{(s)}$, and finally the closed-form expression of the rMSFE, differentiating two cases, where $r \neq 1$ and $r = 1$, respectively.

For $r \neq 1$, based on the perturbation analysis on the SVD [15] of $\left[\hat{\boldsymbol{\mathcal{F}}}^{(s)}\right]_{(r)} \in \mathbb{C}^{M_r \times M_r'}$ with $M_r' = \frac{M}{M_r}$, $\Delta \boldsymbol{f}_r^{(s)}$ takes the following form

$$\Delta \boldsymbol{f}_r^{(s)} = \Delta \boldsymbol{u}_{(r,s)}^{[\text{s}]} = \boldsymbol{\Gamma}_{(r,s)}^{[\text{n}]} \cdot \left[\Delta \boldsymbol{\mathcal{F}}^{(s)}\right]_{(r)} \cdot \boldsymbol{v}_{(r,s)}^{[\text{s}]} \cdot {\sigma_{(r,s)}^{[\text{s}]}}^{-1} + \mathcal{O}\left(\Delta^2\right), \tag{10}$$

where $\boldsymbol{\Gamma}_{(r,s)}^{[\text{n}]} = \boldsymbol{U}_{(r,s)}^{[\text{n}]} \cdot {\boldsymbol{U}_{(r,s)}^{[\text{n}]}}^{\text{H}}$. Here the columns of $\boldsymbol{U}_{(s,r)}^{[\text{n}]} \in \mathbb{C}^{M_r \times (M_r - 1)}$ are the orthonormal basis of the null space of $\left[\boldsymbol{\mathcal{F}}^{(s)}\right]_{(r)}$, $\boldsymbol{v}_{(r,s)}^{[\text{s}]} \in \mathbb{C}^{M_r'}$ denotes the first right singular vector of $\left[\boldsymbol{\mathcal{F}}^{(s)}\right]_{(r)}$, and $\sigma_{(r,s)}^{[\text{s}]}$ is the corresponding singular value. Throughout the paper, we use $\mathcal{O}\left(\Delta^2\right)$ to represent all terms with an order higher than one, i.e., second-order, third-order, and so forth.

The r-mode unfolding of the perturbation of $\boldsymbol{\mathcal{F}}^{(s)}$ denoted by $\left[\Delta \boldsymbol{\mathcal{F}}^{(s)}\right]_{(r)}$ can be expressed with $\Delta \boldsymbol{f}^{(s)} \in \mathbb{C}^M$ representing the s-th column of $\Delta \boldsymbol{F}$ as

$$\left[\Delta \boldsymbol{\mathcal{F}}^{(s)}\right]_{(r)} = \boldsymbol{J}_r \cdot \left(\boldsymbol{I}_{M_r'} \otimes \Delta \boldsymbol{f}^{(s)}\right) \in \mathbb{C}^{M_r \times M_r'}, \tag{11}$$

where $\otimes$ denotes the Kronecker product, $\boldsymbol{I}_{M_r'}$ represents an $M_r' \times M_r'$ identity matrix, and $\boldsymbol{J}_r \in \mathbb{R}^{M_r \times \left(M \cdot M_r'\right)}$ is a selection matrix employed to rearrange $\boldsymbol{f}^{(s)}$ into the r-mode unfolding of $\boldsymbol{\mathcal{F}}^{(s)}$ formed by $\boldsymbol{f}^{(s)}$ as illustrated in Fig. 1.

Inserting (11) into (10) followed by some manipulations yields

$$\begin{aligned} \Delta \boldsymbol{f}_r^{(s)} &= \boldsymbol{K}_0^{(r,s)} \cdot \left(\boldsymbol{I}_{M_r'} \otimes \boldsymbol{n}\right) \cdot \boldsymbol{v}_{(r,s)}' + \mathcal{O}\left(\Delta^2\right) \\ &= \boldsymbol{K}_0^{(r,s)} \cdot \boldsymbol{w}_r^{(s)} + \mathcal{O}\left(\Delta^2\right), \end{aligned} \tag{12}$$

where

$$\boldsymbol{K}_0^{(r,s)} = \boldsymbol{\Gamma}_{(r,s)}^{[\text{n}]} \cdot \boldsymbol{J}_r \cdot \left(\boldsymbol{I}_{M_r'} \otimes \boldsymbol{J}^{(s)}\right) \in \mathbb{C}^{M_r \times (M \cdot d \cdot M_r')}, \tag{13}$$

$$\boldsymbol{v}_{(r,s)}' = \boldsymbol{v}_{(r,s)}^{[\text{s}]} \cdot {\sigma_{(r,s)}^{[\text{s}]}}^{-1} \in \mathbb{C}^{M_r'}, \tag{14}$$

$$\boldsymbol{w}_r^{(s)} = \boldsymbol{v}_{(r,s)}' \otimes \boldsymbol{n} \in \mathbb{C}^{M \cdot d \cdot M_r'}. \tag{15}$$

Here $\boldsymbol{J}^{(s)} \in \mathbb{R}^{M \times M \cdot d}$ denotes a selection matrix leading to

$$\Delta \boldsymbol{f}^{(s)} = \boldsymbol{J}^{(s)} \cdot \boldsymbol{n}. \tag{16}$$

With $\Delta \boldsymbol{f}_r^{(s)}$ in (12), $\Delta \widetilde{\boldsymbol{f}}_r^{(s)}$ given in (7) is now written as

$$\Delta \widetilde{\boldsymbol{f}}_r^{(s)} \approx \boldsymbol{K}_1^{(r,s)} \cdot \boldsymbol{w}_r^{(s)}, \tag{17}$$

where

$$\boldsymbol{K}_1^{(r,s)} = \boldsymbol{f}_r^{(s)} \cdot {\boldsymbol{f}_r^{(s)}}^{\text{H}} \cdot \boldsymbol{K}_0^{(r,s)} - \boldsymbol{K}_0^{(r,s)}. \tag{18}$$

Note that $\boldsymbol{\Gamma}_{(r,s)}^{[\text{n}]}$ in the expression of $\boldsymbol{K}_0^{(r,s)}$ given by (13) is a projection matrix into the null space of $\left[\boldsymbol{\mathcal{F}}^{(s)}\right]_{(r)}$, whereas $\boldsymbol{f}_r^{(s)}$ spans the column space of $\left[\boldsymbol{\mathcal{F}}^{(s)}\right]_{(r)}$. Due to this observation, the first term in $\boldsymbol{K}_1^{(r,s)}$, i.e., $\boldsymbol{f}_r^{(s)} \cdot \boldsymbol{f}_r^{(s)^{\mathrm{H}}} \cdot \boldsymbol{K}_0^{(r,s)}$, is a zero matrix. Thus, we obtain

$$\boldsymbol{K}_1^{(r,s)} = -\boldsymbol{K}_0^{(r,s)}. \tag{19}$$

Based on the fact that

$$\boldsymbol{R}_{\mathrm{ww}}^{(r,s)} = \mathbb{E}\left\{\boldsymbol{w}_r^{(s)} \cdot \boldsymbol{w}_r^{(s)^{\mathrm{H}}}\right\} = \left(\boldsymbol{v}'_{(r,s)} \cdot \boldsymbol{v}'^{\mathrm{H}}_{(r,s)}\right) \otimes \boldsymbol{R}_{\mathrm{nn}},$$

the rMSFE of $\boldsymbol{F}_r$ in (9) is obtained as

$$\mathrm{rMSFE}_{(r)} = \frac{\sum_{s=1}^{d} \mathrm{trace}\left\{\boldsymbol{K}_1^{(r,s)} \cdot \boldsymbol{R}_{\mathrm{ww}}^{(r,s)} \cdot \boldsymbol{K}_1^{(r,s)^{\mathrm{H}}}\right\}}{\|\boldsymbol{F}_r\|_{\mathrm{F}}^2}, \quad r \neq 1 \tag{20}$$

where $\mathrm{Re}\{\cdot\}$ symbolizes the real part of the input argument.

When $r = 1$, with the perturbed version of $\boldsymbol{f}_1^{(s)}$ given by

$$\hat{\boldsymbol{f}}_1^{(s)} = \left(s_{(s)}^{[\mathrm{s}]} + \Delta s_{(s)}^{[\mathrm{s}]}\right) \cdot \left(\boldsymbol{u}_{(1,s)}^{[\mathrm{s}]} + \Delta \boldsymbol{u}_{(1,s)}^{[\mathrm{s}]}\right), \tag{21}$$

$\Delta \boldsymbol{f}_r^{(s)}$ is approximated as

$$\Delta \boldsymbol{f}_1^{(s)} \approx s_{(s)}^{[\mathrm{s}]} \cdot \Delta \boldsymbol{u}_{(1,s)}^{[\mathrm{s}]} + \Delta s_{(s)}^{[\mathrm{s}]} \cdot \boldsymbol{u}_{(1,s)}^{[\mathrm{s}]}, \tag{22}$$

neglecting the second-order terms. Note that $\Delta \boldsymbol{u}_{(1,s)}^{[\mathrm{s}]}$ is already obtained through (12). Thus, based on the expression of $s_{(s)}^{[\mathrm{s}]}$ in (4), we approximate its perturbation $\Delta s_{(s)}^{[\mathrm{s}]}$ as

$$\begin{aligned}\Delta s_{(s)}^{[\mathrm{s}]} \approx\ & \boldsymbol{\mathcal{F}}^{(s)} \times_1 \Delta \boldsymbol{u}_{(1,s)}^{[\mathrm{s}]^{\mathrm{H}}} \times_2 \boldsymbol{u}_{(2,s)}^{[\mathrm{s}]^{\mathrm{H}}} \times_3 \cdots \times_R \boldsymbol{u}_{(R,s)}^{[\mathrm{s}]^{\mathrm{H}}} \\ & + \boldsymbol{\mathcal{F}}^{(s)} \times_1 \boldsymbol{u}_{(1,s)}^{[\mathrm{s}]^{\mathrm{H}}} \times_2 \Delta \boldsymbol{u}_{(2,s)}^{[\mathrm{s}]^{\mathrm{H}}} \times_3 \cdots \times_R \boldsymbol{u}_{(R,s)}^{[\mathrm{s}]^{\mathrm{H}}} \\ & \vdots \\ & + \boldsymbol{\mathcal{F}}^{(s)} \times_1 \boldsymbol{u}_{(1,s)}^{[\mathrm{s}]^{\mathrm{H}}} \times_2 \boldsymbol{u}_{(2,s)}^{[\mathrm{s}]^{\mathrm{H}}} \times_3 \cdots \times_R \Delta \boldsymbol{u}_{(R,s)}^{[\mathrm{s}]^{\mathrm{H}}} \\ & + \Delta \boldsymbol{\mathcal{F}}^{(s)} \times_1 \boldsymbol{u}_{(1,s)}^{[\mathrm{s}]^{\mathrm{H}}} \times_2 \boldsymbol{u}_{(2,s)}^{[\mathrm{s}]^{\mathrm{H}}} \times_3 \cdots \times_R \boldsymbol{u}_{(R,s)}^{[\mathrm{s}]^{\mathrm{H}}}. \end{aligned} \tag{23}$$

Now let us take a close look at the first R items in expression of $\Delta s_{(s)}^{[\mathrm{s}]}$ shown above. The r-mode unfolding of the r-th term ($r = 1, 2, \ldots, R$) takes the following form

$$\begin{aligned}\Delta \boldsymbol{u}_{(r,s)}^{[\mathrm{s}]^{\mathrm{H}}} \cdot \left[\boldsymbol{\mathcal{F}}^{(s)}\right]_{(r)} \cdot \Big(& \boldsymbol{u}_{(r+1,s)}^{[\mathrm{s}]^{\mathrm{H}}} \otimes \boldsymbol{u}_{(r+2,s)}^{[\mathrm{s}]^{\mathrm{H}}} \otimes \cdots \otimes \boldsymbol{u}_{(R,s)}^{[\mathrm{s}]^{\mathrm{H}}} \\ & \otimes \boldsymbol{u}_{(1,s)}^{[\mathrm{s}]^{\mathrm{H}}} \otimes \cdots \otimes \boldsymbol{u}_{(r-1,s)}^{[\mathrm{s}]^{\mathrm{H}}}\Big)^{\mathrm{T}}. \end{aligned} \tag{24}$$

Due to the fact that $\boldsymbol{\Gamma}_{(r,s)}^{[\mathrm{n}]}$ in the expression of $\Delta \boldsymbol{u}_{(r,s)}^{[\mathrm{s}]}$ given by (12) represents a projection matrix into the null space of $\left[\boldsymbol{\mathcal{F}}^{(s)}\right]_{(r)}$, the first R terms of $\Delta s_{(s)}^{[\mathrm{s}]}$ are zeros. This observation enables us to simplify $\Delta s_{(s)}^{[\mathrm{s}]}$ into

$$\Delta s_{(s)}^{[\mathrm{s}]} \approx \Delta \boldsymbol{\mathcal{F}}^{(s)} \times_1 \boldsymbol{u}_{(1,s)}^{[\mathrm{s}]^{\mathrm{H}}} \times_2 \boldsymbol{u}_{(2,s)}^{[\mathrm{s}]^{\mathrm{H}}} \times_3 \cdots \times_R \boldsymbol{u}_{(R,s)}^{[\mathrm{s}]^{\mathrm{H}}}. \tag{25}$$

Further manipulations on (25) lead to

$$\Delta s_{(s)}^{[\mathrm{s}]} \approx \boldsymbol{g}_{(s,\mathrm{n})}^{\mathrm{T}} \cdot \boldsymbol{P} \cdot \Delta \boldsymbol{f}^{(s)}, \tag{26}$$

where

$$\boldsymbol{g}_{(s,\mathrm{n})}^{\mathrm{T}} = \boldsymbol{u}_{(2,s)}^{[\mathrm{s}]^{\mathrm{H}}} \otimes \boldsymbol{u}_{(3,s)}^{[\mathrm{s}]^{\mathrm{H}}} \otimes \cdots \otimes \boldsymbol{u}_{(R,s)}^{[\mathrm{s}]^{\mathrm{H}}} \otimes \boldsymbol{u}_{(1,s)}^{[\mathrm{s}]^{\mathrm{H}}}, \tag{27}$$

and $\boldsymbol{P} \in \{0,1\}^{M \times M}$ is a permutation matrix satisfying

$$\mathrm{vec}\left\{\left[\Delta \boldsymbol{\mathcal{F}}^{(s)}\right]_{(1)}\right\} = \boldsymbol{P} \cdot \Delta \boldsymbol{f}^{(s)}. \tag{28}$$

With $\Delta \boldsymbol{u}_{(r,s)}^{[\mathrm{s}]}$, $r = 1, \ldots, R$, and $\Delta s_{(s)}^{[\mathrm{s}]}$ obtained via (12) and (26), respectively, $\Delta \boldsymbol{f}_1^{(s)}$ is written as

$$\Delta \boldsymbol{f}_1^{(s)} \approx s_{(s)}^{[\mathrm{s}]} \cdot \boldsymbol{K}_0^{(1,s)} \cdot \boldsymbol{w}_1^{(s)} + \left(\boldsymbol{g}_{(s,\mathrm{n})}^{\mathrm{T}} \cdot \boldsymbol{P} \cdot \boldsymbol{J}^{(s)} \cdot \boldsymbol{n}\right) \cdot \boldsymbol{u}_{(1,s)}^{[\mathrm{s}]}, \tag{29}$$

where (16) is used to substitute $\Delta \boldsymbol{f}^{(s)}$ in (26). In the subsequent derivations, we use the following simplified expression of $\Delta \boldsymbol{f}_1^{(s)}$

$$\Delta \boldsymbol{f}_1^{(s)} \approx s_{(s)}^{[\mathrm{s}]} \cdot \boldsymbol{K}_0^{(1,s)} \cdot \boldsymbol{w}_1^{(s)} + \boldsymbol{u}_{(1,s)}^{[\mathrm{s}]} \cdot a_{(1,s)}^{[\mathrm{w}]}, \tag{30}$$

where

$$a_{(1,s)}^{[\mathrm{w}]} = \boldsymbol{g}_{(s,\mathrm{n})}^{\mathrm{T}} \cdot \boldsymbol{P} \cdot \boldsymbol{J}^{(s)} \cdot \boldsymbol{n}. \tag{31}$$

Thus, the key ingredient for the closed-form expression of the rMSFE of $\boldsymbol{F}_1$ is obtained as

$$\begin{aligned}\Delta \widetilde{\boldsymbol{f}}_1^{(s)} \approx\ & \boldsymbol{f}_1^{(s)} \cdot \frac{\boldsymbol{f}_1^{(s)^{\mathrm{H}}} \cdot s_{(s)}^{[\mathrm{s}]} \cdot \boldsymbol{K}_0^{(1,s)} \cdot \boldsymbol{w}_1^{(s)}}{\boldsymbol{f}_1^{(s)^{\mathrm{H}}} \cdot \boldsymbol{f}_1^{(s)}} - s_{(s)}^{[\mathrm{s}]} \cdot \boldsymbol{K}_0^{(1,s)} \cdot \boldsymbol{w}_1^{(s)} \\ & + \boldsymbol{f}_1^{(s)} \cdot \frac{\boldsymbol{f}_1^{(s)^{\mathrm{H}}} \cdot \boldsymbol{u}_{(1,s)}^{[\mathrm{s}]} \cdot a_{(1,s)}^{[\mathrm{w}]}}{\boldsymbol{f}_1^{(s)^{\mathrm{H}}} \cdot \boldsymbol{f}_1^{(s)}} - \boldsymbol{u}_{(1,s)}^{[\mathrm{s}]} \cdot a_{(1,s)}^{[\mathrm{w}]}. \end{aligned} \tag{32}$$

The first term in (32) is a zero vector owing to a similar observation as previously that $\boldsymbol{\Gamma}_{(1,s)}^{[\mathrm{n}]}$ in the expression of $\boldsymbol{K}_0^{(1,s)}$ is a projection matrix into the null space of $\left[\boldsymbol{\mathcal{F}}^{(s)}\right]_{(1)}$, whereas $\boldsymbol{f}_1^{(s)}$ spans the column space of $\left[\boldsymbol{\mathcal{F}}^{(s)}\right]_{(1)}$. In addition, with the fact that

$$\boldsymbol{f}_1^{(s)} = b_{(1,s)} \cdot s_{(s)}^{[\mathrm{s}]} \cdot \boldsymbol{u}_{(1,s)}^{[\mathrm{s}]}, \tag{33}$$

where $b_{(1,s)}$ is a short-hand notation for scaling, the third term in (32) is further written as

$$\begin{aligned}\boldsymbol{f}_1^{(s)} \cdot \frac{\boldsymbol{f}_1^{(s)^{\mathrm{H}}} \cdot \boldsymbol{u}_{(1,s)}^{[\mathrm{s}]} \cdot a_{(1,s)}^{[\mathrm{w}]}}{\boldsymbol{f}_1^{(s)^{\mathrm{H}}} \cdot \boldsymbol{f}_1^{(s)}} &= \boldsymbol{f}_1^{(s)} \cdot \frac{\boldsymbol{f}_1^{(s)^{\mathrm{H}}} \cdot \boldsymbol{f}_1^{(s)} \cdot a_{(1,s)}^{[\mathrm{w}]}}{b_{(1,s)} \cdot s_{(s)}^{[\mathrm{s}]} \cdot \boldsymbol{f}_1^{(s)^{\mathrm{H}}} \cdot \boldsymbol{f}_1^{(s)}} \\ &= \frac{\boldsymbol{f}_1^{(s)}}{b_{(1,s)} \cdot s_{(s)}^{[\mathrm{s}]}} \cdot a_{(1,s)}^{[\mathrm{w}]} \\ &= \boldsymbol{u}_{(1,s)}^{[\mathrm{s}]} \cdot a_{(1,s)}^{[\mathrm{w}]}. \end{aligned} \tag{34}$$

Hence, the expression of $\Delta \widetilde{\boldsymbol{f}}_1^{(s)}$ is now simplified as

$$\Delta \widetilde{\boldsymbol{f}}_1^{(s)} \approx \boldsymbol{K}_1^{(1,s)} \cdot \boldsymbol{w}_1^{(s)}, \tag{35}$$

where

$$\boldsymbol{K}_1^{(1,s)} = -s_{(s)}^{[\mathrm{s}]} \cdot \boldsymbol{K}_0^{(1,s)}. \tag{36}$$

The rMSFE in (20) obtained for the case where $r \neq 1$ also applies here for $r = 1$. To summarize, the closed-form expression of the rMSFE for the r-th factor matrix obtained via the least-squares Khatri-Rao factorization is given by

$$\mathrm{rMSFE}_{(r)} = \frac{\sum_{s=1}^{d} \mathrm{trace}\left\{\boldsymbol{K}_1^{(r,s)} \cdot \boldsymbol{R}_{\mathrm{ww}}^{(r,s)} \cdot \boldsymbol{K}_1^{(r,s)^{\mathrm{H}}}\right\}}{\|\boldsymbol{F}_r\|_{\mathrm{F}}^2}, \tag{37}$$

where

$$\boldsymbol{K}_1^{(r,s)} = \begin{cases} -\boldsymbol{K}_0^{(r,s)}, & r \neq 1 \\ -s_{(s)}^{[\mathrm{s}]} \cdot \boldsymbol{K}_0^{(r,s)}, & r = 1. \end{cases} \tag{38}$$

III. Least-squares Khatri-Rao factorization and its performance analysis for SECSI-GU

The CP decomposition of an R-way rank-d tensor $\boldsymbol{\mathcal{X}} \in \mathbb{C}^{M_1 \times M_2 \times \cdots \times M_R}$ written as

$$\boldsymbol{\mathcal{X}} = \boldsymbol{\mathcal{I}}_{R,d} \times_1 \boldsymbol{F}_1 \times_2 \cdots \times_R \boldsymbol{F}_R \tag{39}$$

can be efficiently computed via SECSI which algebraically formulates the CP decomposition into a set of simultaneous matrix diagonalization (SMD) problems [9]. Combining generalized unfoldings with the idea of considering all possible unfoldings to obtain multiple candidate CP models as in SECSI leads to SECSI-GU [10]. It outperforms SECSI for tensors with $R > 3$ dimensions in terms of the estimation accuracy, and it is very flexible in controlling the complexity-accuracy trade-off. Dividing the set of indices $(1, 2, \ldots, R)$ into a P-dimensional subset $\boldsymbol{\alpha}^{(1)} = [\alpha_1, \alpha_2, \ldots, \alpha_P]$ and an $(R-P)$-dimensional subset $\boldsymbol{\alpha}^{(2)} = [\alpha_{P+1}, \alpha_{P+2}, \ldots, \alpha_R]$ with $1 \leq P < R$, SECSI-GU considers generalized unfoldings

$$[\boldsymbol{\mathcal{X}}]_{\boldsymbol{\alpha}^{(1)}\boldsymbol{\alpha}^{(2)}} = (\boldsymbol{F}_{\alpha_1} \diamond \cdots \diamond \boldsymbol{F}_{\alpha_P}) \cdot \left(\boldsymbol{F}_{\alpha_{P+1}} \diamond \cdots \diamond \boldsymbol{F}_{\alpha_R}\right)^{\mathrm{T}}, \tag{40}$$

where the first P indices are arranged into the rows and the rest $R-P$ indices into the columns. Once the two Khatri-Rao products in (40), $\boldsymbol{F}_{\mathrm{A}} = \boldsymbol{F}_{\alpha_1} \diamond \cdots \diamond \boldsymbol{F}_{\alpha_P}$ and $\boldsymbol{F}_{\mathrm{B}} = \boldsymbol{F}_{\alpha_{P+1}} \diamond \cdots \diamond \boldsymbol{F}_{\alpha_R}$, are obtained via SMDs in SECSI-GU [10], the least-squares Khatri-Rao factorization is employed to recover estimates of the loading matrices $\boldsymbol{F}_r$ $(r = 1, 2, \ldots, R)$.

In light of the large number of SMDs each associated with one possible partitioning of the tensor modes, several heuristic selection criteria have been presented in [10] as examples to decide which SMDs to solve and how to select the final estimates of the loading matrices. It is thus important to predict the performance of SECSI-GU analytically so that the generalized unfolding is selected leading to the "best" solution, in the sense of minimal total rMSFE. In other words, an analytical performance evaluation of SECSI-GU is the key to its enhancement with a significantly reduced computational complexity. This is a crucial application of the performance analysis on the least-squares Khatri-Rao factorization presented in this paper, and hence our strong motivation for this work.

Inspired by the "first-order" perturbation analysis of the SECSI framework conducted in [20] for three-way tensors, $\Delta\boldsymbol{F}_{\mathrm{A}}$ and $\Delta\boldsymbol{F}_{\mathrm{B}}$ will be first derived. Then the results shown in Section II will be used to finally obtain closed-form expressions of the rMSFE for all loading matrices, which completes the analytical performance analysis of SECSI-GU.

IV. Simulation results

To demonstrate the validity of our analytical performance evaluation of the least-squares Khatri-Rao factorization, we present comparisons between predicted estimation errors and empirical ones obtained via Monte Carlo simulations. The factor matrices $\boldsymbol{F}_r$ $(r = 1, 2, \ldots, R)$ contain zero-mean uncorrelated Gaussian entries with unit variance, while elements of the perturbation $\Delta\boldsymbol{F}$ are drawn similarly with variance σ_n^2. Accordingly, we define $\mathrm{SNR} = 1/\sigma_n^2$. Fig. 2 and 3 depict the results for the cases where $R = 3$, $M_1 = M_2 = M_3 = 3$, $d = 2$, and $R = 4$, $M_1 = 3$, $M_2 = 4$, $M_3 = 5$, $M_4 = 3$, and $d = 2$, respectively. In both cases, a good match between the predicted analytical and empirical results is observed, especially in the higher SNR regime.

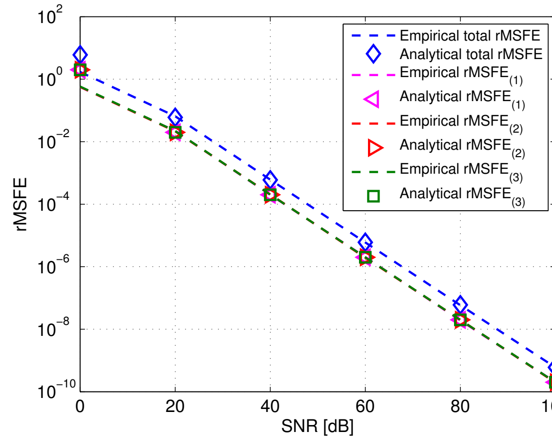

Fig. 2. Empirical and analytical rMSFE versus SNR for $R = 3$, $M_1 = M_2 = M_3 = 3$, and $d = 2$

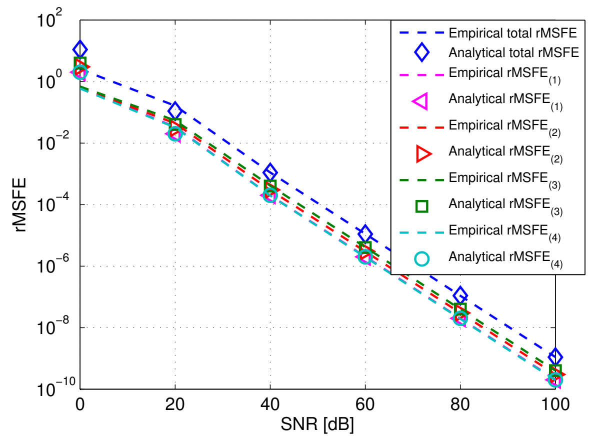

Fig. 3. Empirical and analytical rMSFE versus SNR for $R = 4$, $M_1 = 3$, $M_2 = 4$, $M_3 = 5$, $M_4 = 3$, and $d = 2$

V. Conclusions

We have presented a "first-order" perturbation analysis of the least-squares Khatri-Rao factorization. The procedures of the factorization entail the tensor representation of each column of the Khatri-Rao product and the use of the truncated HOSVD. Though this fact has allowed us to start with the existing perturbation analysis of SVD, in the meantime, it renders the subsequent derivation of the closed-form expressions of the rMSFE for all factor matrices a non-trivial task. To accomplish it, we have proposed to express the rMSFE in an elegant form and have devised convenient matrix-vector reformulations, transforming and decoupling the possibly complicated derivation into tractable and comprehensible procedures. The validity of the analytical performance analysis established in this paper has been corroborated by numerical simulations. In addition, its use in the analytical performance evaluation and a further enhancement of SECSI-GU has been highlighted. The concept of generalized unfoldings being more widely adopted in the computation and also the tracking of the CP decomposition of tensors with $R > 3$ will open up to an even larger number of applications of the least-squares Khatri-Rao factorization and its analytical performance analysis.

Acknowledgment

The authors acknowledge the financial support by the Carl-Zeiss-Foundation (http://carl-zeiss-stiftung.de/) and the German-Israeli Foundation (GIF), grant number I-1282-406.10/2014.

References

[1] A. L. F. de Almeida, G. Favier, and J. C. M. Mota, "PARAFAC-based unified tensor modeling for wireless communication systems with application to blind multiuser equalization," *Signal Processing*, vol. 87, no. 2, pp. 337 – 351, 2007.

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

[3] 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.

[4] G. Favier, C. A. R. Fernandes, and A. L. F. de Almeida, "Nested Tucker tensor decomposition with application to MIMO relay systems using tensor spacetime coding (TSTC)," *Signal Processing*, vol. 128, pp. 318 – 331, 2016.

[5] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *SIAM Review*, vol. 51, no. 3, pp. 455 – 500, Sept. 2009.

[6] M. Haardt, F. Roemer, and G. Del Galdo, "Higher-order SVD based subspace estimation to improve the parameter estimation accuracy in multi-dimensional harmonic retrieval problems," *IEEE Transactions on Signal Processing*, vol. 56, pp. 3198 – 3213, July 2008.

[7] Y. Cheng, S. Li, J. Zhang, F. Roemer, M. Haardt, Y. Zhou, and M. Dong, "Linear precoding-based geometric mean decomposition (LP-GMD) for multi-user MIMO systems," in *Proc. 9-th International Symposium on Wireless Communications Systems (ISWCS 2012)*, Aug. 2012.

[8] K. Naskovska, S. A. Cheema, M. Haardt, B. Valeev, and Y. Evdokimov, "Iterative GFDM Receiver based on the PARATUCK2 tensor decomposition," in *Proc. 21-st International ITG Workshop on Smart Antennas (WSA)*, Mar. 2017.

[9] F. Roemer and M. Haardt, "A semi-algebraic framework for approximate CP decompositions via simultaneous matrix diagonalizations (SECSI)," *Elsevier Signal Processing*, vol. 93, pp. 2722 – 2738, Sept. 2013.

[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. of the 46th Asilomar Conference on Signals, Systems, and Computers*, Nov. 2012.

[11] M. Weis, F. Roemer, M. Haardt, and P. Husar, "Dual-symmetric Parallel Factor analysis using Procrustes estimation and Khatri-Rao factorization," in *Proc. 20-th European Sig. Proc. Conf. (EUSIPCO 2012)*, Aug. 2012.

[12] F. Roemer and M. Haardt, "Tensor-based channel estimation (TENCE) and iterative refinements for two-way relaying with multiple antennas and spatial reuse," *IEEE Transactions on Signal Processing*, vol. 58, no. 10, pp. 5720 – 5735, Nov. 2010.

[13] J. P. C. L. da Costa, F. Roemer, M. Weis, and M. Haardt, "Robust R-D parameter estimation via closed-form PARAFAC," in *Proc. ITG Workshop on Smart Antennas (WSA'10)*, Feb. 2010.

[14] L. de Lathauwer, B. de Moor, and J. Vanderwalle, "A multilinear singular value decomposition," *SIAM J. Matrix Anal. Appl.*, vol. 21, no. 4, pp. 1253 – 1278, 2000.

[15] F. Li, H. Liu, and R. J. Vaccaro, "Performance analysis for DOA estimation algorithms: Unification, simplifications, observations," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 29, no. 4, pp. 1170 – 1184, Oct. 1993.

[16] J. Liu, X. Liu, and X. Ma, "First-Order Perturbation Analysis of Singular Vectors in Singular Value Decomposition," in *Proc. of SSP 2007*, Aug. 2007.

[17] F. Roemer, H. Becker, M. Haardt, and M. Weis, "Analytical performance evaluation for HOSVD-based parameter estimation schemes," in *Proc. of the IEEE Int. Workshop on Comp. Adv. in Multi-Sensor Adaptive Proc. (CAMSAP 2009)*, Dec. 2009.

[18] F. Roemer, M. Haardt, and G. Del Galdo, "Analytical performance assessment of multi-dimensional matrix- and tensor-based ESPRIT-type algorithms," *IEEE Transactions on Signal Processing*, vol. 62, no. 10, pp. 2611 – 2625, May 2014.

[19] E. R. Balda, S. A. Cheema, J. Steinwandt, M. Haardt, A. Weiss, and A. Yeredor, "First-order perturbation analysis of low-rank tensor approximations based on the truncated HOSVD," in *Proc. of 50th Asilomar Conf. Signals, Systems, and Computers (Pacific Grove, CA)*, Nov. 2016.

[20] S. A. Cheema, E. R. Balda, Y. Cheng, A. Weiss, A. Yeredor, and M. Haardt, "First-Order Perturbation Analysis of the SECSI Framework for the Approximate CP Decomposition of 3-D Noise-Corrupted Low-Rank Tensors," submitted to IEEE Transactions on Signal Processing, [Online]. Available: http://www2.tu-ilmenau.de/nt/generic/paper_pdfs/CBCHWY_2017.pdf.