First-Order Perturbation Analysis of Low-Rank Tensor Approximations Based on the Truncated HOSVD

Emilio Rafael Balda, Sher Ali Cheema*, Jens Steinwandt*, Martin Haardt*, Amir Weiss†, and Arie Yeredor†*
*Communications Research Laboratory, Ilmenau University of Technology, Germany
†School of Electrical Engineering, Tel-Aviv University, Israel
Email: *{emilio.balda, martin.haardt}@tu-ilmenau.de, †{arie@eng.tau.ac.il}

***Abstract*—The truncated version of the higher-order singular value decomposition (HOSVD) has a great significance in multi-dimensional tensor-based signal processing. It allows to extract the principal components from noisy observations in order to find a low-rank approximation of the multi-dimensional data. In this paper, we address the question of how good the approximation is by analytically quantifying the tensor reconstruction error introduced by the truncated HOSVD. We present a first-order perturbation analysis of the truncated HOSVD to obtain analytical expressions for the signal subspace error in each dimension as well as the tensor reconstruction error induced by the low-rank approximation of the noise corrupted tensor. The results are asymptotic in the signal-to-noise ratio (SNR) and expressed in terms of the second-order moments of the noise, such that apart from a zero mean, no assumptions on the noise statistics are required. Empirical simulation results verify the obtained analytical expressions.**

***Index Terms*—Perturbation analysis, higher-order singular value decomposition (HOSVD), tensor signal processing.**

I. Introduction

The problem of extracting information and parameters of multi-dimensional signals from noisy observations plays an important role in a broad variety of applications in signal processing. In order to exploit the multi-dimensional structure of the signals, tensor-based algorithms are often used. Many of these algorithms require a low-rank approximation of the measurement tensor as a preprocessing step. Such an approximation is usually obtained by the higher-order singular value decomposition (HOSVD) [1], which preserves the multi-dimensional nature inherent in the data. Its truncated version enables the retrieval of the principal components to form a low-rank approximation of the measurement tensor. In contrast to the matrix case, the truncated HOSVD is not necessarily the best low-rank approximation of a tensor in the Frobenius norm. Thus, [2] proposes an iterative algorithm based on higher-order orthogonal iterations (HOOI) to compute the best rank-$(r_1, r_2, \ldots, r_N)$ approximation. However, as shown in [3], the improvement in terms of the reconstruction error from the HOOI algorithm over the truncated HOSVD is only marginal in the low signal-to-noise ratio (SNR) regime and negligible for high SNRs. Hence, the truncated HOSVD is usually preferred. Applications where the truncated HOSVD is commonly used are, for instance, image processing [4]–[6], object/pattern recognition [7]–[11], parameter estimation [12]–[14], control engineering [15], [16], data analysis [17]–[22], and others. Therefore, a performance analysis to assess the reconstruction error introduced by the low-rank approximation based on the truncated HOSVD is of major importance when analyzing the performance of several algorithms that are based on such a low-rank approximation.

A first-order perturbation analysis for the n-mode singular vectors and singular values obtained from the HOSVD was first presented in [1]. Therein, however, only the full HOSVD without the truncation was investigated. In [23], a first-order performance analysis for the best low-rank approximation of a tensor [2] in the least squares sense was presented. However, both perturbation analyses do not provide explicit analytical expressions for the reconstruction error in terms of the noise statistics. In the case of the multidimensional harmonic retrieval problem, a first-order expansion of the HOSVD-based subspace estimation error of the noisy measurement tensor has been proposed in [24], which generalizes the first-order performance analysis framework for the subspace estimation error from the SVD of the measurement matrix [25] to the tensor case. This framework is applicable whenever the signal component is superimposed by a small noise contribution. In [24], the derived analytical mean square error (MSE) expressions merely depend on the second-order moments of the noise and hence, only require the noise to be zero mean. However, a perturbation analysis of the truncated HOSVD and its associated low-rank approximation of the tensor has so far, to the best of our knowledge, not been reported in the literature.

In this work, we further extend [24] and propose a first-order perturbation analysis of the truncated HOSVD. Specifically, we provide analytical closed-form MSE expressions for the tensor reconstruction error in terms of the second-order moments of the noise. Thus, apart from a zero mean and finite second-order moments, no assumptions on the statistics of the noise are required, i.e., the derived expressions are even valid for non-Gaussian or colored noise. Moreover, the expressions are asymptotic in the signal-to-noise ratio (SNR). In order to provide further insights into the truncated HOSVD, we also provide analytical MSE expressions for the subspace estimation error in the n-th mode, which may be of interest in some of the aforementioned applications. Additionally, we derive simplified versions of the analytical MSE expressions for the reconstruction error and the subspace estimation error for the special case of uncorrelated noise with equal variance. Simulations show that the analytical results provide an excellent match to the empirical ones.

II. Notation

For the sake of notation, we will use the symbols a, $\boldsymbol{a}$, $\boldsymbol{A}$, and $\boldsymbol{\mathcal{A}}$ for a scalar, column vector, matrix, and tensor variables respectively. The superscripts $^{-1}$, *, $^{\mathrm{T}}$, $^{\mathrm{H}}$ denote the matrix inverse, complex conjugate, transposition, and complex conjugate transposition, respectively. The notations $\mathbb{E}\{\cdot\}$, $\mathrm{Tr}[\cdot]$, $\otimes$, $\|\cdot\|_{\mathrm{F}}$, and $\|\cdot\|_2$ are used for expectation, trace, Kronecker product, Frobenius norm, and 2-norm operators, respectively. In this work, we will use the property $\mathrm{Tr}[\boldsymbol{A} \cdot \boldsymbol{B}] = \mathrm{Tr}[\boldsymbol{B} \cdot \boldsymbol{A}]$ for any $\boldsymbol{A} \in \mathbb{C}^{N \times M}$ and $\boldsymbol{B} \in \mathbb{C}^{M \times N}$. For a matrix $\boldsymbol{A} = [\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_M] \in \mathbb{C}^{N \times M}$, where $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_M$ denotes its columns, the operator $\mathrm{vec}\{\cdot\}$ defines the vectorization operation as $\mathrm{vec}\{\boldsymbol{A}\}^{\mathrm{T}} = [\boldsymbol{a}_1^{\mathrm{T}}, \boldsymbol{a}_2^{\mathrm{T}}, \ldots, \boldsymbol{a}_M^{\mathrm{T}}]$. This operator has the

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

property $\text{vec}\{\boldsymbol{A} \cdot \boldsymbol{X} \cdot \boldsymbol{B}\} = (\boldsymbol{B}^{\text{T}} \otimes \boldsymbol{A}) \cdot \text{vec}\{\boldsymbol{X}\}$, where $\boldsymbol{A}$, $\boldsymbol{X}$, $\boldsymbol{B}$ are matrices with proper dimensions.

Let $\boldsymbol{\mathcal{A}} \in \mathbb{C}^{M_1 \times M_2 \times \cdots \times M_R}$ be a tensor of order R (an R-way array), where M_r is its size along the r-th mode. Furthermore, $[\boldsymbol{\mathcal{A}}]_{(r)}$ denotes the r-mode unfolding of $\boldsymbol{\mathcal{A}}$ which is performed according to [1]. Additionally, the r-mode product of a tensor $\boldsymbol{\mathcal{A}}$ with a matrix $\boldsymbol{B} \in \mathbb{C}^{N \times M_r}$ (i.e., $\boldsymbol{\mathcal{A}} \times_r \boldsymbol{B}$) is defined as $\boldsymbol{\mathcal{C}} = \boldsymbol{\mathcal{A}} \times_r \boldsymbol{B} \iff [\boldsymbol{\mathcal{C}}]_{(r)} = \boldsymbol{B} \cdot [\boldsymbol{\mathcal{A}}]_{(r)}$, where $\boldsymbol{\mathcal{C}}$ is a tensor with the corresponding dimensions. For the sake of notational simplicity, we define the following notation for multiple Kronecker and r-mode products

$$\bigotimes_{r=1}^{R} \boldsymbol{B}_r = \boldsymbol{B}_1 \otimes \boldsymbol{B}_2 \otimes \cdots \otimes \boldsymbol{B}_R$$

$$\boldsymbol{\mathcal{A}} \bigtimes_{r=1}^{R}{}_r \boldsymbol{B}_r = \boldsymbol{\mathcal{A}} \times_1 \boldsymbol{B}_1 \times_2 \boldsymbol{B}_2 \times_3 \cdots \times_R \boldsymbol{B}_R.$$

Another property that is often used in this work is $\boldsymbol{\mathcal{C}} = \boldsymbol{\mathcal{A}} \bigtimes_{r=1}^{R}{}_r \boldsymbol{B}_r \iff [\boldsymbol{\mathcal{C}}]_{(r)} = \boldsymbol{B}_r \cdot [\boldsymbol{\mathcal{A}}]_{(r)} \cdot \left(\bigotimes_{j=r+1}^{R} \boldsymbol{B}_j^{\text{T}} \otimes \bigotimes_{i=1}^{r-1} \boldsymbol{B}_i^{\text{T}} \right)$. Moreover, the higher order norm of a tensor is defined as $\|\boldsymbol{\mathcal{A}}\|_{\text{H}} = \|[\boldsymbol{\mathcal{A}}]_{(r)}\|_{\text{F}} = \|\text{vec}\{[\boldsymbol{\mathcal{A}}]_{(r)}\}\|_2 \quad \forall r = 1, 2, \ldots, R$.

III. Signal Model

Let $\boldsymbol{\mathcal{X}}_0 \in \mathbb{C}^{M_1 \times M_2 \times \cdots \times M_R}$ be a noiseless tensor of order R. In addition, let $M = \prod_{r=1}^{R} M_r$ be the number of elements of $\boldsymbol{\mathcal{X}}_0$. Let $p_1, p_2, \ldots, p_R$ denote the r-ranks of such a tensor that are defined as $p_r = \text{rank}\left([\boldsymbol{\mathcal{X}}_0]_{(r)}\right)$ for $r = 1, 2, \ldots, R$. By calculating the HOSVD of $\boldsymbol{\mathcal{X}}_0$ we obtain

$$\boldsymbol{\mathcal{X}}_0 = \boldsymbol{\mathcal{S}} \bigtimes_{r=1}^{R}{}_r \boldsymbol{U}_r = \boldsymbol{\mathcal{S}}^{[\text{s}]} \bigtimes_{r=1}^{R}{}_r \boldsymbol{U}_r^{[\text{s}]},$$

where $\boldsymbol{U}_r = \begin{bmatrix} \boldsymbol{U}_r^{[\text{s}]} & \boldsymbol{U}_r^{[\text{n}]} \end{bmatrix} \in \mathbb{C}^{M_r \times M_r}$ are the unitary matrices obtained from the Singular Value Decomposition (SVD) of $[\boldsymbol{\mathcal{X}}_0]_{(r)} \in \mathbb{C}^{M_r \times \frac{M}{M_r}}$, and $\boldsymbol{U}_r^{[\text{s}]} \in \mathbb{C}^{M_r \times p_r}$ and $\boldsymbol{U}_r^{[\text{n}]} \in \mathbb{C}^{M_r \times (M_r - p_r)}$ have unitary columns $\forall r = 1, 2, \ldots, R$, i.e.,

$$[\boldsymbol{\mathcal{X}}_0]_{(r)} = \boldsymbol{U}_r \cdot \boldsymbol{\Sigma}_r \cdot \boldsymbol{V}_r^{\text{H}}$$
$$= \begin{bmatrix} \boldsymbol{U}_r^{[\text{s}]} & \boldsymbol{U}_r^{[\text{n}]} \end{bmatrix} \begin{bmatrix} \boldsymbol{\Sigma}_r^{[\text{s}]} & \boldsymbol{0}_{p_r \times M/M_r} \\ \boldsymbol{0}_{(M_r - p_r) \times p_r} & \boldsymbol{0}_{(M_r - p_r) \times M/M_r} \end{bmatrix} \begin{bmatrix} \boldsymbol{V}_r^{[\text{s}]} & \boldsymbol{V}_r^{[\text{n}]} \end{bmatrix}^{\text{H}},$$

where $\boldsymbol{\Sigma}_r^{[\text{s}]} = \text{diag}\left\{\sigma_r^{(i)}\right\}_{i=1}^{p_r}$ contains the singular values of $[\boldsymbol{\mathcal{X}}_0]_{(r)}$, and $\boldsymbol{V}_r^{[\text{s}]} \in \mathbb{C}^{\frac{M_r}{M} \times p_r}$ and $\boldsymbol{V}_r^{[\text{n}]} \in \mathbb{C}^{\frac{M_r}{M} \times (\frac{M_r}{M} - p_r)}$ have unitary columns. Furthermore, let $\boldsymbol{\mathcal{X}} \in \mathbb{C}^{M_1 \times M_2 \times \cdots \times M_R}$ be a noisy tensor from which we want to estimate $\boldsymbol{\mathcal{X}}_0$. Therefore, we have

$$\boldsymbol{\mathcal{X}} = \boldsymbol{\mathcal{X}}_0 + \boldsymbol{\mathcal{N}}, \tag{1}$$

where $\boldsymbol{\mathcal{N}} \in \mathbb{C}^{M_1 \times M_2 \times \cdots \times M_R}$ is a random additive noise tensor. Moreover, we define the r-mode correlation matrices $\boldsymbol{R}_r$ of $\boldsymbol{\mathcal{N}}$ as[1]

$$\boldsymbol{R}_r \triangleq \mathbb{E}\left\{\text{vec}\left\{[\boldsymbol{\mathcal{N}}]_{(r)}\right\} \cdot \text{vec}\left\{[\boldsymbol{\mathcal{N}}]_{(r)}\right\}^{\text{H}}\right\} \in \mathbb{C}^{M \times M}.$$

Note that the signal to be represented by $\boldsymbol{\mathcal{X}}$ and $\boldsymbol{\mathcal{X}}_0$ depends on the application. To illustrate this point, some examples are shown in Table I. As before, we can compute the HOSVD of the noisy tensor $\boldsymbol{\mathcal{X}}$ as $\boldsymbol{\mathcal{X}} = \hat{\boldsymbol{\mathcal{S}}} \bigtimes_{r=1}^{R}{}_r \hat{\boldsymbol{U}}_r$, where

$$[\boldsymbol{\mathcal{X}}]_{(r)} = \hat{\boldsymbol{U}}_r \cdot \hat{\boldsymbol{\Sigma}}_r \cdot \hat{\boldsymbol{V}}_r^{\text{H}}$$
$$= \begin{bmatrix} \hat{\boldsymbol{U}}_r^{[\text{s}]} & \hat{\boldsymbol{U}}_r^{[\text{n}]} \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{\Sigma}}_r^{[\text{s}]} & \boldsymbol{0}_{p_r \times M/M_r} \\ \boldsymbol{0}_{(M_r - p_r) \times p_r} & \hat{\boldsymbol{\Sigma}}_r^{[\text{n}]} \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{V}}_r^{[\text{s}]} & \hat{\boldsymbol{V}}_r^{[\text{n}]} \end{bmatrix}^{\text{H}}.$$

[1]Note that $\boldsymbol{R}_1, \boldsymbol{R}_2, \ldots, \boldsymbol{R}_R$ are permuted versions of each other.

TABLE I: Different interpretations for the low-rank tensor $\boldsymbol{\mathcal{X}}_0$, and the noisy tensor $\boldsymbol{\mathcal{X}}$ according to equation (1)

Application	$\boldsymbol{\mathcal{X}}_0$	$\boldsymbol{\mathcal{X}}$
De-noising	True data	Noisy data
Compression	Compressed data	Uncompressed data
Parameter Estimation	Ideal measurements	Noisy measurements

In this case, $\hat{\boldsymbol{U}}_r^{[\text{s}]} \in \mathbb{C}^{M_r \times p_r}$ is an estimate of $\boldsymbol{U}_r^{[\text{s}]}$. Moreover, we assume that the r-ranks of the noiseless tensor $\boldsymbol{\mathcal{X}}_0$ are either known, fixed or estimated before performing the truncated HOSVD. Thereby, using these $\hat{\boldsymbol{U}}_r^{[\text{s}]}$ matrices, we can estimate $\boldsymbol{\mathcal{X}}_0$ from $\boldsymbol{\mathcal{X}}$ as $\boldsymbol{\mathcal{X}}_0 \approx \hat{\boldsymbol{\mathcal{X}}} = \hat{\boldsymbol{\mathcal{S}}}^{[\text{s}]} \bigtimes_{r=1}^{R}{}_r \hat{\boldsymbol{U}}_r^{[\text{s}]}$, where the truncated core tensor $\hat{\boldsymbol{\mathcal{S}}}^{[\text{s}]}$ is computed as

$$\hat{\boldsymbol{\mathcal{S}}}^{[\text{s}]} = \boldsymbol{\mathcal{X}} \bigtimes_{r=1}^{R}{}_r \hat{\boldsymbol{U}}_r^{[\text{s}]^{\text{H}}} \in \mathbb{C}^{p_1 \times p_2 \times \cdots p_R}. \tag{2}$$

With this setup, our goal is to find an analytical expression for $\mathbb{E}\left\{\|\Delta \boldsymbol{\mathcal{X}}\|_{\text{H}}^2\right\}$, where $\Delta \boldsymbol{\mathcal{X}} \triangleq \hat{\boldsymbol{\mathcal{X}}} - \boldsymbol{\mathcal{X}}_0$.

IV. Signal Subspace Perturbation

Let us first analyze the perturbation of the r-mode signal subspaces, which are defined as the column spaces spanned by $\boldsymbol{U}_r^{[\text{s}]}$ for $r = 1, 2, \ldots, R$.

A. General Expression

Let $\Delta \boldsymbol{U}_r^{[\text{s}]}$ be the perturbation present in $\hat{\boldsymbol{U}}_r^{[\text{s}]}$. Therefore, motivated by [25], we can write $\hat{\boldsymbol{U}}_r^{[\text{s}]}$ as $\hat{\boldsymbol{U}}_r^{[\text{s}]} = \boldsymbol{U}_r^{[\text{s}]} + \Delta \boldsymbol{U}_r^{[\text{s}]}$. For the truncated HOSVD tensor estimation, we are only interested in the perturbation on the column space spanned by $\hat{\boldsymbol{U}}_r^{[\text{s}]}$. Therefore, we can use the result obtained in [24], which is

$$\Delta \boldsymbol{U}_r^{[\text{s}]} = \boldsymbol{U}_r^{[\text{n}]} \cdot \boldsymbol{U}_r^{[\text{n}]^{\text{H}}} \cdot [\boldsymbol{\mathcal{N}}]_{(r)} \cdot \boldsymbol{V}_r^{[\text{s}]} \cdot \boldsymbol{\Sigma}_r^{[\text{s}]^{-1}} + \mathcal{O}(\Delta^2). \tag{3}$$

Note that $\Delta \boldsymbol{U}_r^{[\text{s}]^{\text{H}}} \cdot \boldsymbol{U}_r^{[\text{s}]} = \boldsymbol{0}$, which means that the perturbation is orthogonal to the column space of $\boldsymbol{U}_r^{[\text{s}]}$. To simplify this expression let $\widetilde{\boldsymbol{N}}_r \in \mathbb{C}^{(M_r - p_r) \times p_r}$ be the transformed version of r-mode unfolding of the noise, defined as

$$\widetilde{\boldsymbol{N}}_r \triangleq \boldsymbol{U}_r^{[\text{n}]^{\text{H}}} \cdot [\boldsymbol{\mathcal{N}}]_{(r)} \cdot \boldsymbol{V}_r^{[\text{s}]}, \tag{4}$$

with its corresponding correlation matrix denoted as $\widetilde{\boldsymbol{R}}_r \triangleq \mathbb{E}\{\text{vec}\{\widetilde{\boldsymbol{N}}_r\} \cdot \text{vec}\{\widetilde{\boldsymbol{N}}_r\}^{\text{H}}\}$. Note that we can express $\widetilde{\boldsymbol{R}}_r$ in terms of the r-mode correlation matrix of the noise $\boldsymbol{R}_r$ as

$$\widetilde{\boldsymbol{R}}_r = \mathbb{E}\left\{\text{vec}\left\{\widetilde{\boldsymbol{N}}_r\right\} \cdot \text{vec}\left\{\widetilde{\boldsymbol{N}}_r\right\}^{\text{H}}\right\}$$
$$= \left(\boldsymbol{V}_r^{[\text{s}]^{\text{T}}} \otimes \boldsymbol{U}_r^{[\text{n}]^{\text{H}}}\right) \cdot \boldsymbol{R}_r \cdot \left(\boldsymbol{V}_r^{[\text{s}]^*} \otimes \boldsymbol{U}_r^{[\text{n}]}\right). \tag{5}$$

Therefore, by applying the notation proposed in equation (4), we rewrite equation (3) as

$$\Delta \boldsymbol{U}_r^{[\text{s}]} = \boldsymbol{U}_r^{[\text{n}]} \cdot \widetilde{\boldsymbol{N}}_r \cdot \boldsymbol{\Sigma}_r^{[\text{s}]^{-1}} + \mathcal{O}(\Delta^2). \tag{6}$$

Note that the choice of $\boldsymbol{U}_r^{[\text{s}]}$ is not unique when the HOSVD of $\boldsymbol{\mathcal{X}}_0$ is calculated, but the subspace spanned by the columns of $\boldsymbol{U}_r^{[\text{s}]}$ is always unique. Therefore, we focus on the subspace perturbation only. To that end, let $\boldsymbol{T}_r = \boldsymbol{U}_r^{[\text{s}]} \cdot \boldsymbol{U}_r^{[\text{s}]^{\text{H}}}$ be the projection matrix onto the column space spanned by $\boldsymbol{U}_r^{[\text{s}]}$. Since the matrices $\boldsymbol{T}_r$ are unique for $r = 1, 2, \ldots, R$, regardless of the choice of $\boldsymbol{U}_r^{[\text{s}]}$, we will investigate the perturbations of those matrices (denoted as $\Delta \boldsymbol{T}_r$). Furthermore, motivated by [24], the estimated signal subspace projection matrices (defined as $\hat{\boldsymbol{T}}_r = \hat{\boldsymbol{U}}_r^{[\mathrm{s}]} \cdot \hat{\boldsymbol{U}}_r^{[\mathrm{s}]^{\mathrm{H}}}$) can be expressed as $\hat{\boldsymbol{T}}_r = \boldsymbol{T}_r + \Delta\boldsymbol{T}_r$, where

$$\Delta\boldsymbol{T}_r = \boldsymbol{U}_r^{[\mathrm{s}]} \cdot \Delta\boldsymbol{U}_r^{[\mathrm{s}]^{\mathrm{H}}} + \Delta\boldsymbol{U}_r^{[\mathrm{s}]} \cdot \boldsymbol{U}_r^{[\mathrm{s}]^{\mathrm{H}}} + \mathcal{O}(\Delta^2). \quad (7)$$

Note that $\Delta\boldsymbol{U}_r^{[\mathrm{s}]} \cdot \Delta\boldsymbol{U}_r^{[\mathrm{s}]^{\mathrm{H}}}$ is a second order term and, therefore, it is considered inside $\mathcal{O}(\Delta^2)$. In this section the goal is to obtain an analytical expression for the expected value of the Frobenius norm of the r-mode subspace estimation error, i.e., $\mathbb{E}\left\{\|\Delta\boldsymbol{T}_r\|_{\mathrm{F}}^2\right\}$. To that end, we first analyze the following expression

$$\|\Delta\boldsymbol{T}_r\|_{\mathrm{F}}^2 = \mathrm{Tr}\left[\Delta\boldsymbol{T}_r \cdot \Delta\boldsymbol{T}_r^{\mathrm{H}}\right] \quad (8)$$

Substituting equation (7) into (8) and neglecting the terms that contain $\mathcal{O}(\Delta^2)$, we obtain

$$\begin{aligned}\|\Delta\boldsymbol{T}_r\|_{\mathrm{F}}^2 \approx\ & \mathrm{Tr}\left[\boldsymbol{U}_r^{[\mathrm{s}]} \cdot \Delta\boldsymbol{U}_r^{[\mathrm{s}]^{\mathrm{H}}} \cdot \boldsymbol{U}_r^{[\mathrm{s}]} \cdot \Delta\boldsymbol{U}_r^{[\mathrm{s}]^{\mathrm{H}}}\right] \\ &+ \mathrm{Tr}\left[\boldsymbol{U}_r^{[\mathrm{s}]} \cdot \Delta\boldsymbol{U}_r^{[\mathrm{s}]^{\mathrm{H}}} \cdot \Delta\boldsymbol{U}_r^{[\mathrm{s}]} \cdot \boldsymbol{U}_r^{[\mathrm{s}]^{\mathrm{H}}}\right] \\ &+ \mathrm{Tr}\left[\Delta\boldsymbol{U}_r^{[\mathrm{s}]} \cdot \boldsymbol{U}_r^{[\mathrm{s}]^{\mathrm{H}}} \cdot \boldsymbol{U}_r^{[\mathrm{s}]} \cdot \Delta\boldsymbol{U}_r^{[\mathrm{s}]^{\mathrm{H}}}\right] \\ &+ \mathrm{Tr}\left[\Delta\boldsymbol{U}_r^{[\mathrm{s}]} \cdot \boldsymbol{U}_r^{[\mathrm{s}]^{\mathrm{H}}} \cdot \Delta\boldsymbol{U}_r^{[\mathrm{s}]} \cdot \boldsymbol{U}_r^{[\mathrm{s}]^{\mathrm{H}}}\right].\end{aligned}$$

Since $\Delta\boldsymbol{U}_r^{[\mathrm{s}]^{\mathrm{H}}} \cdot \boldsymbol{U}_r^{[\mathrm{s}]} = \left(\boldsymbol{U}_r^{[\mathrm{s}]^{\mathrm{H}}} \cdot \Delta\boldsymbol{U}_r^{[\mathrm{s}]}\right)^{\mathrm{H}} = \boldsymbol{0}_{p_r \times p_r}$ and $\boldsymbol{U}_r^{[\mathrm{s}]^{\mathrm{H}}} \cdot \boldsymbol{U}_r^{[\mathrm{s}]} = \boldsymbol{I}_{p_r}$, we can use the properties of the trace operator to simply this further to

$$\|\Delta\boldsymbol{T}_r\|_{\mathrm{F}}^2 \approx 2 \cdot \mathrm{Tr}\left[\Delta\boldsymbol{U}_r^{[\mathrm{s}]^{\mathrm{H}}} \cdot \Delta\boldsymbol{U}_r^{[\mathrm{s}]}\right] \quad (9)$$

Now, by using equation (6) and neglecting the second order term, we can see that

$$\begin{aligned}&\mathrm{Tr}\left[\Delta\boldsymbol{U}_r^{[\mathrm{s}]^{\mathrm{H}}} \cdot \Delta\boldsymbol{U}_r^{[\mathrm{s}]}\right] \\ &\approx \mathrm{Tr}\left[\boldsymbol{\Sigma}_r^{[\mathrm{s}]^{-1}} \cdot \widetilde{\boldsymbol{N}}_r^{\mathrm{H}} \cdot \boldsymbol{U}_r^{[\mathrm{n}]^{\mathrm{H}}} \cdot \boldsymbol{U}_r^{[\mathrm{n}]} \cdot \widetilde{\boldsymbol{N}}_r \cdot \boldsymbol{\Sigma}_r^{[\mathrm{s}]^{-1}}\right] \\ &= \mathrm{Tr}\left[\boldsymbol{\Sigma}_r^{[\mathrm{s}]^{-1}} \cdot \widetilde{\boldsymbol{N}}_r^{\mathrm{H}} \cdot \widetilde{\boldsymbol{N}}_r \cdot \boldsymbol{\Sigma}_r^{[\mathrm{s}]^{-1}}\right] \\ &= \mathrm{Tr}\left[\left(\widetilde{\boldsymbol{N}}_r \cdot \boldsymbol{\Sigma}_r^{[\mathrm{s}]^{-1}}\right)^{\mathrm{H}} \cdot \left(\widetilde{\boldsymbol{N}}_r \cdot \boldsymbol{\Sigma}_r^{[\mathrm{s}]^{-1}}\right)\right] = \|\widetilde{\boldsymbol{N}}_r \cdot \boldsymbol{\Sigma}_r^{[\mathrm{s}]^{-1}}\|_{\mathrm{F}}^2 \\ &= \mathrm{Tr}\left[\mathrm{vec}\left\{\widetilde{\boldsymbol{N}}_r \cdot \boldsymbol{\Sigma}_r^{[\mathrm{s}]^{-1}}\right\} \cdot \mathrm{vec}\left\{\widetilde{\boldsymbol{N}}_r \cdot \boldsymbol{\Sigma}_r^{[\mathrm{s}]^{-1}}\right\}^{\mathrm{H}}\right] \\ &= \mathrm{Tr}\Big[\left(\boldsymbol{\Sigma}_r^{[\mathrm{s}]^{-2}} \otimes \boldsymbol{I}_{(M_r - p_r)}\right) \cdot \mathrm{vec}\{\widetilde{\boldsymbol{N}}_r\} \cdot \mathrm{vec}\{\widetilde{\boldsymbol{N}}_r\}^{\mathrm{H}}\Big]. \quad (10)\end{aligned}$$

Next, we take the expected value of equation (10) and obtain

$$\begin{aligned}&\mathbb{E}\left\{\mathrm{Tr}\left[\Delta\boldsymbol{U}_r^{[\mathrm{s}]^{\mathrm{H}}} \cdot \Delta\boldsymbol{U}_r^{[\mathrm{s}]}\right]\right\} \\ &\approx \mathbb{E}\left\{\mathrm{Tr}\left[\left(\boldsymbol{\Sigma}_r^{[\mathrm{s}]^{-2}} \otimes \boldsymbol{I}_{(M_r - p_r)}\right) \cdot \mathrm{vec}\{\widetilde{\boldsymbol{N}}_r\} \cdot \mathrm{vec}\{\widetilde{\boldsymbol{N}}_r\}^{\mathrm{H}}\right]\right\} \\ &= \mathrm{Tr}\left[\left(\boldsymbol{\Sigma}_r^{[\mathrm{s}]^{-2}} \otimes \boldsymbol{I}_{(M_r - p_r)}\right) \cdot \widetilde{\boldsymbol{R}}_r\right]\end{aligned}$$

We can now take the expected value of $\|\Delta\boldsymbol{T}_r\|_{\mathrm{F}}^2$ (from equation (9)) and use this relation to obtain the desired closed-form expression

$$\begin{aligned}\mathbb{E}\left\{\|\Delta\boldsymbol{T}_r\|_{\mathrm{F}}^2\right\} &\approx \mathbb{E}\left\{2 \cdot \mathrm{Tr}\left[\Delta\boldsymbol{U}_r^{[\mathrm{s}]^{\mathrm{H}}} \cdot \Delta\boldsymbol{U}_r^{[\mathrm{s}]}\right]\right\} \\ &= 2 \cdot \mathrm{Tr}\left[\left(\boldsymbol{\Sigma}_r^{[\mathrm{s}]^{-2}} \otimes \boldsymbol{I}_{(M_r - p_r)}\right) \cdot \widetilde{\boldsymbol{R}}_r\right]. \quad (11)\end{aligned}$$

B. Special Case of Uncorrelated Noise

Equation (11) can be simplified even further if we assume that the noise is zero-mean and uncorrelated with variance σ_N^2, i.e., $\boldsymbol{R}_r = \sigma_N^2 \cdot \boldsymbol{I}_M$. Therefore, $\widetilde{\boldsymbol{R}}_r$ in equation (5) is simplified to

$$\begin{aligned}\widetilde{\boldsymbol{R}}_r &= \left(\boldsymbol{V}_r^{[\mathrm{s}]^{\mathrm{T}}} \otimes \boldsymbol{U}_r^{[\mathrm{n}]^{\mathrm{H}}}\right) \cdot \sigma_N^2 \cdot \boldsymbol{I}_M \cdot \left(\boldsymbol{V}_r^{[\mathrm{s}]^*} \otimes \boldsymbol{U}_r^{[\mathrm{n}]}\right) \\ &= \sigma_N^2 \cdot \left(\boldsymbol{V}_r^{[\mathrm{s}]^{\mathrm{T}}} \cdot \boldsymbol{V}_r^{[\mathrm{s}]^*} \otimes \boldsymbol{U}_r^{[\mathrm{n}]^{\mathrm{H}}} \cdot \boldsymbol{U}_r^{[\mathrm{n}]}\right) \\ &= \sigma_N^2 \cdot \left(\boldsymbol{I}_{p_r} \otimes \boldsymbol{I}_{M_r - p_r}\right) = \sigma_N^2 \cdot \boldsymbol{I}_{p_r \cdot (M_r - p_r)}. \quad (12)\end{aligned}$$

Thus, the desired expression in equation (11) becomes

$$\begin{aligned}\mathbb{E}\left\{\|\Delta\boldsymbol{T}_r\|_{\mathrm{F}}^2\right\} &= 2 \cdot \mathrm{Tr}\left[\left(\boldsymbol{\Sigma}_r^{[\mathrm{s}]^{-2}} \otimes \sigma_N^2 \cdot \boldsymbol{I}_{(M_r - p_r)}\right)\right] \\ &= 2 \cdot (M_r - p_r) \cdot \sigma_N^2 \cdot \sum_{i=1}^{p_r} \frac{1}{\left(\sigma_r^{(i)}\right)^2}. \quad (13)\end{aligned}$$

V. Truncated HOSVD estimate perturbation

A. General Expression

In this section we investigate the overall perturbation of the estimated tensor $\hat{\boldsymbol{\mathcal{X}}}$, which is calculated as

$$\hat{\boldsymbol{\mathcal{X}}} = \hat{\boldsymbol{\mathcal{S}}}^{[\mathrm{s}]} \mathop{\times}_{r=1}^{R}{}_r \hat{\boldsymbol{U}}_r^{[\mathrm{s}]}.$$

To this end, we can use equation (2) to obtain

$$\hat{\boldsymbol{\mathcal{X}}} = \hat{\boldsymbol{\mathcal{S}}}^{[\mathrm{s}]} \mathop{\times}_{r=1}^{R}{}_r \hat{\boldsymbol{U}}_r^{[\mathrm{s}]} = \boldsymbol{\mathcal{X}} \mathop{\times}_{r=1}^{R}{}_r \hat{\boldsymbol{T}}_r.$$

Moreover, by inserting the relations $\boldsymbol{\mathcal{X}} = \boldsymbol{\mathcal{X}}_0 + \boldsymbol{\mathcal{N}}$ and $\hat{\boldsymbol{T}} = \boldsymbol{T} + \Delta\boldsymbol{T}$, this expression expands to

$$\begin{aligned}\hat{\boldsymbol{\mathcal{X}}} &= \boldsymbol{\mathcal{X}}_0 \mathop{\times}_{r=1}^{R}{}_r \hat{\boldsymbol{T}}_r + \boldsymbol{\mathcal{N}} \mathop{\times}_{r=1}^{R}{}_r \hat{\boldsymbol{T}}_r \\ &= \boldsymbol{\mathcal{X}}_0 \mathop{\times}_{r=1}^{R}{}_r (\boldsymbol{T}_r + \Delta\boldsymbol{T}_r) + \boldsymbol{\mathcal{N}} \mathop{\times}_{r=1}^{R}{}_r (\boldsymbol{T}_r + \Delta\boldsymbol{T}_r) \quad (14) \\ &= \boldsymbol{\mathcal{X}}_0 \mathop{\times}_{r=1}^{R}{}_r \boldsymbol{T}_r + \boldsymbol{\mathcal{N}} \mathop{\times}_{r=1}^{R}{}_r \boldsymbol{T}_r + \sum_{r=1}^{R}\left(\boldsymbol{\mathcal{X}}_0 \times_r \Delta\boldsymbol{T}_r \mathop{\times}_{\substack{i=1 \\ i \neq r}}^{R}{}_i \boldsymbol{T}_i\right) + \mathcal{O}(\Delta^2)\end{aligned}$$

Note that, all the terms with more than one error variable (such as $\boldsymbol{\mathcal{N}}$, $\Delta\boldsymbol{T}_1, \Delta\boldsymbol{T}_2, \ldots, \Delta\boldsymbol{T}_R$) are included inside $\mathcal{O}(\Delta^2)$. For the sake of simplicity, let us define the noise tensor projected on the signal subspaces $\boldsymbol{\mathcal{N}}^{[\mathrm{s}]}$ as $\boldsymbol{\mathcal{N}}^{[\mathrm{s}]} \triangleq \boldsymbol{\mathcal{N}} \mathop{\times}_{r=1}^{R}{}_r \boldsymbol{T}_r$. Since we know that $\boldsymbol{\mathcal{X}}_0 \times_r \boldsymbol{T}_r = \boldsymbol{\mathcal{X}}_0$ for all $r = 1, 2, \ldots, R$, we can further simplify equation (14) to $\hat{\boldsymbol{\mathcal{X}}} = \boldsymbol{\mathcal{X}}_0 + \sum_{r=1}^{R}(\boldsymbol{\mathcal{X}}_0 \times_r \Delta\boldsymbol{T}_r) + \boldsymbol{\mathcal{N}}^{[\mathrm{s}]} + \mathcal{O}(\Delta^2)$. Therefore, $\hat{\boldsymbol{\mathcal{X}}}$ can be expressed as $\hat{\boldsymbol{\mathcal{X}}} = \boldsymbol{\mathcal{X}}_0 + \Delta\boldsymbol{\mathcal{X}}$, where

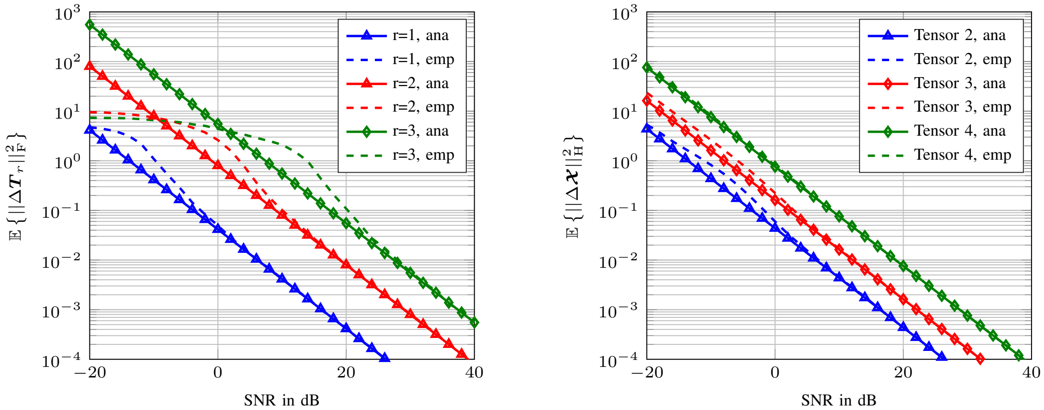
$$\Delta\boldsymbol{\mathcal{X}} = \sum_{r=1}^{R}(\boldsymbol{\mathcal{X}}_0 \times_r \Delta\boldsymbol{T}_r) + \boldsymbol{\mathcal{N}}^{[\mathrm{s}]} + \mathcal{O}(\Delta^2).$$

As in the previous section, we derive an analytical expression for the expected value of Frobenius norm of the error term of interest (i.e., $\mathbb{E}\left\{\|\Delta\boldsymbol{\mathcal{X}}\|_{\mathrm{H}}^2\right\}$). Let us assume that $\Delta\boldsymbol{T}_1, \Delta\boldsymbol{T}_2, \ldots, \Delta\boldsymbol{T}_R$ and $\boldsymbol{\mathcal{N}}^{[\mathrm{s}]}$ are independent from each other. Therefore, we approximate $\mathbb{E}\left\{\|\Delta\boldsymbol{\mathcal{X}}\|_{\mathrm{H}}^2\right\}$ to

$$\mathbb{E}\left\{\|\Delta\boldsymbol{\mathcal{X}}\|_{\mathrm{H}}^2\right\} \approx \mathbb{E}\left\{\sum_{r=1}^{R}\|\boldsymbol{\mathcal{X}}_0 \times_r \Delta\boldsymbol{T}_r\|_{\mathrm{H}}^2 + \|\boldsymbol{\mathcal{N}}^{[\mathrm{s}]}\|_{\mathrm{H}}^2\right\}. \quad (15)$$

Let us analyze the term $\Delta\boldsymbol{T}_r^{\mathrm{H}} \cdot \Delta\boldsymbol{T}_r$. By applying the result obtained in equation (7), we can approximate this term (by neglecting the terms that contain $\mathcal{O}(\Delta^2)$ in (7)) to

$$\Delta\boldsymbol{T}_r^{\mathrm{H}} \cdot \Delta\boldsymbol{T}_r \approx \boldsymbol{U}_r^{[\mathrm{s}]} \cdot \Delta\boldsymbol{U}_r^{[\mathrm{s}]^{\mathrm{H}}} \cdot \Delta\boldsymbol{U}_r^{[\mathrm{s}]} \cdot \boldsymbol{U}_r^{[\mathrm{s}]^{\mathrm{H}}} + \Delta\boldsymbol{U}_r^{[\mathrm{s}]} \cdot \Delta\boldsymbol{U}_r^{[\mathrm{s}]^{\mathrm{H}}}.$$



(a) Subspace estimation error $\mathbb{E}\left\{||\Delta \boldsymbol{T}_r||_{\mathrm{F}}^2\right\}$.

(b) Tensor estimation error $\mathbb{E}\left\{||\Delta \boldsymbol{\mathcal{X}}||_{\mathrm{H}}^2\right\}$.

Fig. 1: (a): Subspace estimation errors corresponding to each of the subspaces of tensor 1, described in Table II. (b): Tensor estimation error for tensors 2, 3 and 4, described in Table III.

We can now use this relation to obtain an expression for the term $\|\boldsymbol{\mathcal{X}}_0 \times_r \Delta \boldsymbol{T}_r\|_{\mathrm{H}}^2$ contained in equation (15) as

$$\begin{aligned}
&\|\boldsymbol{\mathcal{X}}_0 \times_r \Delta \boldsymbol{T}_r\|_{\mathrm{H}}^2 \\
&= \mathrm{Tr}\left[[\boldsymbol{\mathcal{X}}_0]_{(r)}^{\mathrm{H}} \cdot \Delta \boldsymbol{T}_r^{\mathrm{H}} \cdot \Delta \boldsymbol{T}_r [\boldsymbol{\mathcal{X}}_0]_{(r)}\right] \\
&\approx \mathrm{Tr}\Big[(\boldsymbol{U}_r^{[\mathrm{s}]} \cdot \Delta \boldsymbol{U}_r^{[\mathrm{s}]^{\mathrm{H}}} \cdot \Delta \boldsymbol{U}_r^{[\mathrm{s}]} \cdot \boldsymbol{U}_r^{[\mathrm{s}]^{\mathrm{H}}} + \Delta \boldsymbol{U}_r^{[\mathrm{s}]} \cdot \Delta \boldsymbol{U}_r^{[\mathrm{s}]^{\mathrm{H}}}) \\
&\quad \cdot [\boldsymbol{\mathcal{X}}_0]_{(r)} \cdot [\boldsymbol{\mathcal{X}}_0]_{(r)}^{\mathrm{H}}\Big] \\
&= \mathrm{Tr}\left[\boldsymbol{U}_r^{[\mathrm{s}]} \cdot \Delta \boldsymbol{U}_r^{[\mathrm{s}]^{\mathrm{H}}} \cdot \Delta \boldsymbol{U}_r^{[\mathrm{s}]} \cdot \boldsymbol{U}_r^{[\mathrm{s}]^{\mathrm{H}}} \cdot [\boldsymbol{\mathcal{X}}_0]_{(r)} \cdot [\boldsymbol{\mathcal{X}}_0]_{(r)}^{\mathrm{H}}\right].
\end{aligned}$$

Now, we use equation (6) and again neglect the terms that contain $\mathcal{O}(\Delta^2)$ to simplify this further to

$$\begin{aligned}
&\|\boldsymbol{\mathcal{X}}_0 \times_r \Delta \boldsymbol{T}_r\|_{\mathrm{H}}^2 \\
&\approx \mathrm{Tr}\Big[\boldsymbol{U}_r^{[\mathrm{s}]} \cdot \boldsymbol{\Sigma}_r^{[\mathrm{s}]^{-1}} \cdot \widetilde{\boldsymbol{N}}_r^{\mathrm{H}} \cdot \boldsymbol{U}_r^{[\mathrm{n}]^{\mathrm{H}}} \cdot \boldsymbol{U}_r^{[\mathrm{n}]} \cdot \widetilde{\boldsymbol{N}}_r \cdot \boldsymbol{\Sigma}_r^{[\mathrm{s}]^{-1}} \\
&\quad \cdot \boldsymbol{U}_r^{[\mathrm{s}]^{\mathrm{H}}} \cdot [\boldsymbol{\mathcal{X}}_0]_{(r)} \cdot [\boldsymbol{\mathcal{X}}_0]_{(r)}^{\mathrm{H}}\Big] \\
&= \mathrm{Tr}\Big[\boldsymbol{U}_r^{[\mathrm{s}]} \cdot \boldsymbol{\Sigma}_r^{[\mathrm{s}]^{-1}} \cdot \widetilde{\boldsymbol{N}}_r^{\mathrm{H}} \cdot \widetilde{\boldsymbol{N}}_r \cdot \boldsymbol{\Sigma}_r^{[\mathrm{s}]^{-1}} \cdot \boldsymbol{U}_r^{[\mathrm{s}]^{\mathrm{H}}} \\
&\quad \cdot \boldsymbol{U}_r^{[\mathrm{s}]} \cdot \boldsymbol{\Sigma}_r^{[\mathrm{s}]} \cdot \boldsymbol{V}_r^{[\mathrm{s}]^{\mathrm{H}}} \cdot \boldsymbol{V}_r^{[\mathrm{s}]} \cdot \boldsymbol{\Sigma}_r^{[\mathrm{s}]} \cdot \boldsymbol{U}_r^{[\mathrm{s}]^{\mathrm{H}}}\Big] \\
&= \mathrm{Tr}\Big[\boldsymbol{U}_r^{[\mathrm{s}]} \cdot \boldsymbol{\Sigma}_r^{[\mathrm{s}]^{-1}} \cdot \widetilde{\boldsymbol{N}}_r^{\mathrm{H}} \cdot \widetilde{\boldsymbol{N}}_r \cdot \boldsymbol{\Sigma}_r^{[\mathrm{s}]^{-1}} \cdot \boldsymbol{\Sigma}_r^{[\mathrm{s}]} \cdot \boldsymbol{\Sigma}_r^{[\mathrm{s}]} \cdot \boldsymbol{U}_r^{[\mathrm{s}]^{\mathrm{H}}}\Big] \\
&= \mathrm{Tr}\Big[\widetilde{\boldsymbol{N}}_r^{\mathrm{H}} \cdot \widetilde{\boldsymbol{N}}_r \cdot \boldsymbol{\Sigma}_r^{[\mathrm{s}]} \cdot \boldsymbol{U}_r^{[\mathrm{s}]^{\mathrm{H}}} \cdot \boldsymbol{U}_r^{[\mathrm{s}]} \cdot \boldsymbol{\Sigma}_r^{[\mathrm{s}]^{-1}}\Big] \\
&= \mathrm{Tr}\Big[\widetilde{\boldsymbol{N}}_r^{\mathrm{H}} \cdot \widetilde{\boldsymbol{N}}_r\Big] = \mathrm{Tr}\Big[\mathrm{vec}\{\widetilde{\boldsymbol{N}}_r\} \cdot \mathrm{vec}\{\widetilde{\boldsymbol{N}}_r\}^{\mathrm{H}}\Big].
\end{aligned}$$

Therefore, $\mathbb{E}\{\|\boldsymbol{\mathcal{X}}_0 \times_r \Delta \boldsymbol{T}_r\|_{\mathrm{H}}^2\}$ is reduced to

$$\mathbb{E}\{\|\boldsymbol{\mathcal{X}}_0 \times_r \Delta \boldsymbol{T}_r\|_{\mathrm{F}}^2\} \approx \mathrm{Tr}[\widetilde{\boldsymbol{R}}_r]. \tag{16}$$

Next, we need to obtain an expression for $\mathbb{E}\left\{\|\boldsymbol{\mathcal{N}}^{[\mathrm{s}]}\|_{\mathrm{H}}^2\right\}$. To that end, let us express $\boldsymbol{\mathcal{N}}^{[\mathrm{s}]}$ and $\boldsymbol{\mathcal{N}}$ as $(R+1)$-order tensors $\boldsymbol{\mathcal{N}}^{[\mathrm{s}]} = \boldsymbol{\mathcal{N}} \mathop{\times_r}\limits_{r=1}^{R} \boldsymbol{T}_r \times_{R+1} 1$ and $\boldsymbol{\mathcal{N}} = \boldsymbol{\mathcal{N}} \times_{R+1} 1$. Using this formulation we can take the $R+1$-mode unfolding of $\boldsymbol{\mathcal{N}}^{[\mathrm{s}]}$ (i.e., $[\boldsymbol{\mathcal{N}}^{[\mathrm{s}]}]_{(R+1)} \in \mathbb{C}^{1\times M}$) to obtain $[\boldsymbol{\mathcal{N}}^{[\mathrm{s}]}]_{(R+1)} = 1 \cdot [\boldsymbol{\mathcal{N}}]_{(R+1)} \cdot \left(\bigotimes_{r=1}^{R} \boldsymbol{T}_r\right)^{\mathrm{T}} = \mathrm{vec}\left\{[\boldsymbol{\mathcal{N}}]_{(R)}\right\}^{\mathrm{T}} \cdot \left(\bigotimes_{r=1}^{R} \boldsymbol{T}_r\right)^{*}$. Using this expression we get

$$\begin{aligned}
\mathbb{E}\left\{\|\boldsymbol{\mathcal{N}}^{[\mathrm{s}]}\|_{\mathrm{H}}^2\right\} &= \mathbb{E}\left\{\|[\boldsymbol{\mathcal{N}}^{[\mathrm{s}]}]_{(R+1)}^{\mathrm{T}}\|_2^2\right\} \\
&= \mathbb{E}\left\{\mathrm{Tr}\left[[\boldsymbol{\mathcal{N}}^{[\mathrm{s}]}]_{(R+1)}^{\mathrm{T}} \cdot [\boldsymbol{\mathcal{N}}^{[\mathrm{s}]}]_{(R+1)}^{*}\right]\right\} \\
&= \mathrm{Tr}\left[\left(\bigotimes_{r=1}^{R} \boldsymbol{T}_r\right) \cdot \boldsymbol{R}_R\right].
\end{aligned} \tag{17}$$

Finally, using equations (16) and (17), $\mathbb{E}\left\{\|\Delta \boldsymbol{\mathcal{X}}\|_{\mathrm{F}}^2\right\}$ from equation (15) is approximated to

$$\mathbb{E}\left\{\|\Delta \boldsymbol{\mathcal{X}}\|_{\mathrm{H}}^2\right\} \approx \sum_{r=1}^{R} \mathrm{Tr}[\widetilde{\boldsymbol{R}}_r] + \mathrm{Tr}\left[\left(\bigotimes_{r=1}^{R} \boldsymbol{T}_r\right) \cdot \boldsymbol{R}_R\right]. \tag{18}$$

B. Special Case of Uncorrelated Noise

As in Section IV, we simplify this expression for the special case of uncorrelated noise with variance σ_N^2. It can be easily shown that, in this case, $\mathbb{E}\left\{\|\boldsymbol{\mathcal{N}}^{[\mathrm{s}]}\|_{\mathrm{H}}^2\right\} = \sigma_N^2 \cdot \prod_{r=1}^{R} p_r$. Using this property and equation (12), we simplify equation (18) further to $\mathbb{E}\left\{\|\Delta \boldsymbol{\mathcal{X}}\|_{\mathrm{H}}^2\right\} \approx \sum_{r=1}^{R} \mathrm{Tr}\left[\sigma_N^2 \cdot \boldsymbol{I}_{p_r(M_r - p_r)}\right] + \sigma_N^2 \cdot \prod_{r=1}^{R} p_r$. Finally, we reach the desired $\mathbb{E}\left\{\|\Delta \boldsymbol{\mathcal{X}}\|_{\mathrm{H}}^2\right\}$ in terms of the noise variance for the uncorrelated noise case, i.e.,

$$\mathbb{E}\left\{\|\Delta \boldsymbol{\mathcal{X}}\|_{\mathrm{H}}^2\right\} \approx \sigma_N^2 \cdot \left(\sum_{r=1}^{R} (M_r - p_r) \cdot p_r + \prod_{r=1}^{R} p_r\right). \tag{19}$$

VI. Simulation Results

To validate the analytical results obtained in the previous sections, we perform empirical simulations. To this end we first define the noiseless tensor(s) and the noise characteristics. For all the simulations, the noiseless low-rank tensor $\boldsymbol{\mathcal{X}}_0$ is $\boldsymbol{\mathcal{X}}_0 \in \mathbb{R}^{20\times 20\times 20}$ and norm $\|\boldsymbol{\mathcal{X}}_0\|_{\mathrm{H}}^2 = 1$. Note that, a tensor $\boldsymbol{\mathcal{B}} \in \mathbb{R}^{20\times 20\times 20}$ with r-ranks (p_1, p_2, p_3) can be generated as $\boldsymbol{\mathcal{B}} = \boldsymbol{\mathcal{A}} \times_1 \boldsymbol{W} \times_2 \boldsymbol{Y} \times_3 \boldsymbol{Z}$, where $\boldsymbol{\mathcal{A}} \in \mathbb{R}^{p_1\times p_2\times p_3}$, $\boldsymbol{W} \in \mathbb{R}^{p_1\times 20}$, $\boldsymbol{Y} \in \mathbb{R}^{p_2\times 20}$ and $\boldsymbol{Z} \in \mathbb{R}^{p_3\times 20}$ are randomly generated using independent zero-mean Gaussian distributions with equal variance. Furthermore, we use uncorrelated zero-mean Gaussian distributed noise with variance σ_N^2 for the realizations of the noise tensor $\boldsymbol{\mathcal{N}} \in \mathbb{R}^{20\times 20\times 20}$, where σ_N^2 is calculated as

$$\sigma_N^2 = \frac{\|\boldsymbol{\mathcal{X}}_0\|_{\mathrm{H}}^2}{\mathrm{SNR} \cdot M}$$

TABLE II: Parameters of tensor 1, used in Figure 1(a)

	p_r	$\mathrm{Tr}\left[\boldsymbol{\Sigma}_r^{[s]^{-2}}\right]$
$r = 1$	3	9.7420
$r = 2$	9	290.0909
$r = 3$	15	4418.2000

TABLE III: r-ranks of tensors 2, 3, and 4, used in Figure 1(b)

$\boldsymbol{\mathcal{X}}_0$	p_1	p_2	p_3
Tensor 2	5	5	5
Tensor 3	10	10	10
Tensor 4	15	20	20

for the different signal-to-noise ratio (SNR) values [2].

In Figure 1(a), we validate the results obtained in Section IV. Here, the noiseless tensor $\boldsymbol{\mathcal{X}}_0$ (referred to as tensor 1 in Figure 1(a)), has the characteristics shown in Table II. Then, after 1000 trials of uncorrelated noise realizations for each SNR point, the empirical error curve for $\mathbb{E}\left\{\|\Delta \boldsymbol{T}_r\|_{\mathrm{F}}^2\right\}$ is computed for $r = 1, 2, 3$. Furthermore, we can observe how the analytical expressions obtained using equation (13) asymptotically match the empirical error curve for each of the subspace estimates of the noiseless tensor.

Likewise, in Figure 1(b) we validate the results obtained in Section V. Here, the simulations are conducted for 3 different noiseless tensors (i.e., tensors 2, 3 and 4) of the same sizes, but with different r-ranks. Furthermore, the r-ranks (p_1, p_2, p_3) of this tensors are shown in Table III. As before, 1000 trials of uncorrelated noise realizations for each SNR point are simulated to obtain the empirical error curve for $\mathbb{E}\left\{\|\Delta \boldsymbol{\mathcal{X}}\|_{\mathrm{H}}^2\right\}$. Furthermore, we can see how the analytical expressions obtained using equation (19) match the empirical curves as expected.

VII. Conclusion

In this work, a first-order perturbation analysis of the truncated HOSVD is presented, where we provide closed-form expressions for the tensor reconstruction error. The derived expressions are formulated in terms of the second-order moments of the noise, such that apart from a zero mean, no assumptions on the noise statistics are required. In addition, the obtained general expressions have been simplified for the special case of uncorrelated noise with equal variance. This simplification provides better insights into the truncated HOSVD performance. The simulation results show that the proposed solution achieves an excellent match to the empirical results.

References

[1] L. De Lathauwer, B. De Moor, and J. Vandewalle, “A multilinear singular value decomposition,” *SIAM Journal on Matrix Analysis and Applications*, vol. 21, no. 4, pp. 1253–1278, 2000. [Online]. Available: http://dx.doi.org/10.1137/S0895479896305696

[2] ——, “On the best rank-1 and rank-(R1,R2, . . . , Rn) approximation of higher-order tensors,” *SIAM Journal of Matrix Analysis and Applications*, vol. 21, no. 4, pp. 1324–1342, 2000.

[3] F. Roemer and M. Haardt, “A semi-algebraic framework for approximate CP decompositions via simultaneous matrix diagonalizations (SECSI),” *Signal Processing*, vol. 93, no. 9, pp. 2722 – 2738, 2013. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0165168413000704

[4] C. Liu, J. Zhou, and K. He, “Image compression based on truncated HOSVD,” in *Proceedings of International Conference on Information Engineering and Computer Science*, Dec 2009, pp. 1–4.

[5] W. Dong, G. Li, G. Shi, X. Li, and Y. Ma, “Low-rank tensor approximation with Laplacian scale mixture modeling for multiframe image denoising,” in *Proceedings of IEEE International Conference on Computer Vision (ICCV)*, Dec 2015, pp. 442–449.

[6] S. Afra and E. Gildin, “Permeability parametrization using higher order singular value decomposition (HOSVD),” in *Proceedings of 12th International Conference on Machine Learning and Applications (ICMLA)*, vol. 2, Dec 2013, pp. 188–193.

[7] F. Janoos, S. Singh, O. Irfanoglu, R. Machiraju, and R. Parent, “Activity analysis using spatio-temporal trajectory volumes in surveillance applications,” in *Proceedings of IEEE Symposium on Visual Analytics Science and Technology (VAST)*, Oct 2007, pp. 3–10.

[8] Z. H. Feng, J. Kittler, W. Christmas, X. J. Wu, and S. Pfeiffer, “Automatic face annotation by multilinear AAM with missing values,” in *Proceedings of 21st International Conference on Pattern Recognition (ICPR)*, Nov 2012, pp. 2586–2589.

[9] H. Y. Mesri, M. K. Najafabadi, and T. McKelvey, “A multidimensional signal processing approach for classification of microwave measurements with application to stroke type diagnosis,” in *Proceedings of Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC*, Aug 2011, pp. 6465–6469.

[10] M. R. Ameri, M. Haji, A. Fischer, D. Ponson, and T. D. Bui, “A feature extraction method for cursive character recognition using higher-order singular value decomposition,” in *Proceedings of 14th International Conference on Frontiers in Handwriting Recognition (ICFHR)*, Sept 2014, pp. 512–516.

[11] B. Cyganek and K. Socha, “Novel parallel algorithm for object recognition with the ensemble of classifiers based on the higher-order singular value decomposition of prototype pattern tensors,” in *Proceedings of International Conference on Computer Vision Theory and Applications (VISAPP)*, vol. 2, Jan 2014, pp. 648–653.

[12] A. Csapo and P. Baranyi, “A tensor algebraic framework for the intuitive exploration of sensory substitution spaces,” in *Proceedings of 8th International Symposium on Intelligent Systems and Informatics (SISY)*, Sept 2010, pp. 547–552.

[13] M. Haardt, F. Roemer, and G. Del Galdo, “Higher-order SVD-based subspace estimation to improve the parameter estimation accuracy in multidimensional harmonic retrieval problems,” *IEEE Transactions on Signal Processing*, vol. 56, no. 7, pp. 3198–3213, July 2008.

[14] M. Boizard, G. Ginolhac, F. Pascal, and P. Forster, “A new tool for multidimensional low-rank STAP filter: Cross HOSVDs,” in *Proceedings of the 20th European Signal Processing Conference (EUSIPCO)*, Aug 2012, pp. 1324–1328.

[15] G. Orbán, A. Rövid, and P. Várlaki, “Control model for loading systems using higher order singular value decomposition,” in *Proceedings of IEEE International Conference on Computer Science and Automation Engineering (CSAE)*, vol. 3, June 2011, pp. 754–758.

[16] A. Hajiloo and W. F. Xie, “Multi-objective control design of the nonlinear systems using genetic algorithm,” in *Proceedings of IEEE International Symposium on Innovations in Intelligent Systems and Applications (INISTA) Proceedings*, June 2014, pp. 27–34.

[17] J. Li, Y. Yan, W. Duan, S. Song, and M. H. Lee, “Tensor decomposition of Toeplitz Jacket matrices for big data processing,” in *Proceedings of International Conference on Big Data and Smart Computing (BigComp)*, Feb 2015, pp. 11–14.

[18] A. Rövid, L. Szeidl, and P. Várlaki, “Data representation in HOSVD-DCT based domain,” in *Proceedings of IEEE 17th International Conference on Intelligent Engineering Systems (INES)*, June 2013, pp. 103–106.

[19] S. Manna, Z. Petres, and T. Gedeon, “Tensor term indexing: An application of HOSVD for document summarization,” in *Proceedings of 4th International Symposium on Computational Intelligence and Intelligent Informatics*, Oct 2009, pp. 135–141.

[20] A. Rövid, L. Szeidl, S. Sergyán, and P. Várlaki, “HOSVD-wavelet based framework for multidimensional data approximation,” in *Proceedings of IEEE 9th International Conference on Computational Cybernetics (ICCC)*, July 2013, pp. 29–33.

[21] D. Rafailidis and P. Daras, “The TFC model: Tensor factorization and tag clustering for item recommendation in social tagging systems,” *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 43, no. 3, pp. 673–688, May 2013.

[22] N. Liu, B. Zhang, J. Yan, Z. Chen, W. Liu, F. Bai, and L. Chien, “Text representation: from vector to tensor,” in *Proceedings of Fifth IEEE International Conference on Data Mining*, Nov 2005, pp. 4 pp.–.

[23] L. De Lathauwer, “First-order perturbation analysis of the best rank-(R1,R2,R3) approximation in multilinear algebra,” *Journal of Chemometrics*, vol. 18, pp. 2–11, 2004.

[24] 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.

[25] F. Li, H. Liu, and R. J. Vaccaro, “Performance analysis for DOA estimation algorithms: unification, simplification, and observations,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 29, no. 4, pp. 1170–1184, Oct 1993.

[2]The SNR here is defined on a linear scale (not in dB).