Analytical Performance Assessment of Multi-Dimensional Matrix- and Tensor-Based ESPRIT-Type Algorithms

Florian Roemer, *Member, IEEE*, Martin Haardt, *Senior Member, IEEE*, and Giovanni Del Galdo, *Member, IEEE*

***Abstract*—In this paper we present a generic framework for the asymptotic performance analysis of subspace-based parameter estimation schemes. It is based on earlier results on an explicit first-order expansion of the estimation error in the signal subspace obtained via an SVD of the noisy observation matrix. We extend these results in a number of aspects. Firstly, we demonstrate that an explicit first-order expansion of the Higher-Order SVD (HOSVD)-based subspace estimate can be derived. Secondly, we show how to obtain explicit first-order expansions of the estimation error of arbitrary ESPRIT-type algorithms and provide the expressions for R-D Standard ESPRIT, R-D Unitary ESPRIT, R-D Standard Tensor-ESPRIT, as well as R-D Unitary Tensor-ESPRIT. Thirdly, we derive closed-form expressions for the mean square error (MSE) and show that they only depend on the second-order moments of the noise. Hence, to apply this framework we only need the noise to be zero mean and possess finite second order moments. Additional assumptions such as Gaussianity or circular symmetry are not needed.**

***Index Terms*—Direction-of-arrival estimation, tensors, subspace methods, performance analysis.**

I. Introduction

HIGH resolution parameter estimation from R-dimensional (R-D) signals is a task required for a variety of applications, such as estimating the multi-dimensional parameters of the dominant multipath components from MIMO channel measurements [13], which may be used for geometry-based channel modeling. Other applications include RADAR [26], wireless communications [22], sonar, seismology, and medical imaging. In [12], we have shown that in the R-D case ($R \geq 2$), tensors can be used to store and manipulate the R-D signals in a multidimensional form. Based on this idea, we have proposed an enhanced tensor-based signal subspace estimate as well as ESPRIT-type algorithms based on tensors in [12]. Their superior performance and the resulting tensor gain were shown based on Monte-Carlo simulations.

In this paper we present a framework for the analytical performance assessment of subspace-based parameter estimation schemes and apply it to derive a first-order perturbation expansion for the tensor-based subspace estimate. Moreover, we find first-order expansions for the estimation error of ESPRIT-type algorithms and derive generic mean square error (MSE) expressions which only depend on the second-order moments of the noise and hence do not require Gaussianity or circular symmetry. This approach allows to assess the gain obtained by using tensors instead of matrices analytically in order to determine in which scenarios it is particularly pronounced. We apply this framework for the analysis of R-D Standard ESPRIT, R-D Unitary ESPRIT, R-D Standard Tensor-ESPRIT and R-D Unitary Tensor-ESPRIT.

The analytical performance assessment of subspace-based parameter estimation schemes has a long standing history in signal processing. Analytical results on the performance of spectral MUSIC [37] and Standard ESPRIT [36] have appeared shortly after their original publication. The most frequently cited papers are [15] for the spectral MUSIC algorithm and [30] for Standard ESPRIT. However, many follow-up papers exist which extend the original results, e.g., [8], [24], [25], [28], [41], and many others. However, these results have in common that they all go more or less directly back to a result on the distribution of the eigenvectors of a sample covariance matrix from [1], [2].

In contrast to these results, in [19] an entirely different approach was proposed, which provides an explicit first-order expansion of the subspace of a desired signal component if observed superimposed by a small additive perturbation. This approach has a number of advantages compared to [2]. Firstly, [2] is asymptotic in the sample size N, i.e., the result becomes accurate only as the number of snapshots N is very large, whereas [19] is asymptotic in the effective SNR, i.e., it can be used even for $N = 1$ as long as the noise variance is sufficiently small. Secondly, [2] requires strong Gaussianity assumptions, not only on the perturbation (i.e., the additive noise), but also on the source symbols. Since [19] is explicit, no assumptions about the statistics of either the desired signal or the perturbation are needed. Note that it has recently been shown that [2] can be extended to the non-Gaussian and the non-circular case

Manuscript received February 25, 2013; revised January 13, 2014; accepted March 02, 2014. Date of publication March 25, 2014; date of current version April 24, 2014. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Jean Pierre Delmas. Parts of this paper was presented at the *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2010)*, Dallas, TX, USA, March 2010, and at the *Third International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP 2009)*, Aruba, Dutch Antilles, December 2009.

The authors are with Institute of Information Technology, Ilmenau University of Technology, D-98684 Ilmenau, Germany (e-mail: florian.roemer@tu-ilmenau.de; martin.haardt@tu-ilmenau.de; giovanni.delgaldo@tu-ilmenau.de).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSP.2014.2313530

in [6]. However, the large sample size assumption is still needed. Thirdly, the covariance expressions from [2] are much less intuitive than the expansion from [19] which shows directly how much of the noise subspace "leaks into" the signal subspace due to the erroneous estimate. Finally, the expressions involved in [2] are quite involved and hence tough to handle, whereas [19] requires only a few terms which appear directly as block matrices of the SVD of the noise-free observation matrix.

Due to these advantages we clearly favor [19] as a starting point. The authors in [19] have already shown that their results on the perturbation of the subspace can be used to find a first order expansion for spectral MUSIC, Root-MUSIC, Min-Norm, the State-Space-Realization, and even the Standard ESPRIT algorithm. However, they only considered 1-D Standard ESPRIT. We extend their work by considering multiple dimensions (R-D ESPRIT),[1] by incorporating forward-backward-averaging (for Unitary ESPRIT), by considering the tensor-based subspace estimate (for Standard and Unitary Tensor-ESPRIT), and by providing generic mean square error (MSE) expressions of the resulting estimation errors in these cases. Note that our MSE expressions depend on the second-order moments of the noise only. Hence we only assume it to be zero mean (due to the asymptotic nature of our performance analysis), but do not require it to be Gaussian distributed, white, or circularly symmetric. This is a particularly attractive feature of our approach with respect to different types of preprocessing which may alter the noise statistics, e.g., spatial smoothing (which yields spatially correlated noise) or forward-backward-averaging (which annihilates the circular symmetry of the noise). Since we do not require spatial whiteness or circular symmetry, our MSE expressions are directly applicable to a wide range of ESPRIT-type algorithms.

There have been other follow-up papers based on [19]. For instance, [40] provides a first-order and second-order perturbation expansion which can be seen as a generalization of [19]. In [21] the authors show that there is also a first-order contribution of the perturbation of the signal subspace which lies in the signal subspace (which is, however, irrelevant for subspace-based parameter estimation schemes). Note that perturbation expansions for the signal subspace and null space projectors based on the sample covariance matrix are provided in [16] where the expansion up to an arbitrary order is derived based on a recurrence relation. A mean square error expression for Standard ESPRIT is provided in [20]. However, it does not generalize easily to the tensor case and it assumes circular symmetry of the noise. Note that the latter assumption implies that it is not applicable to Unitary ESPRIT, since forward-backward-averaging annihilates the circular symmetry of the noise. Moreover, other authors have studied the asymptotical performance of Standard ESPRIT, e.g., [7], [39], where harmonic retrieval from time series is investigated and MSE expressions for a large number of snapshots as well as MSE expressions for a high SNR are derived. Note that we find MSE expressions compatible to [7] by only assuming a high effective SNR, i.e., either the number of snapshots or the SNR can tend to infinity. Some analytical results on the asymptotic efficiency of spectral MUSIC, Root-MUSIC, Standard ESPRIT using Least Squares (LS) and Total Least Squares (TLS) are, among others, presented in [29]–[31], and [27], respectively. However, these results are asymptotic in the number of snapshots N and sometimes even in the number of sensors M. The asymptotic equivalence of Standard ESPRIT using LS and TLS, Pro-ESPRIT, and the Matrix Pencil method has also been shown, see for instance [14]. Overall, in the matrix case, the number of existing results is quite large, since the underlying methods have been known for more than two decades. However, concerning the tensor case, existing results are much more scarce.

In the tensor case a first-order expansion for the HOSVD has been proposed in [5]. However, it is not suitable for our application since it does not consider the HOSVD-based subspace estimate [12] but the subspaces of the separate n-mode unfoldings and their singular values. A first-order expansion for the best rank-(R_1, R_2, R_3)-expansion is provided in [4]. However, again, it is not directly applicable for analyzing the HOSVD-based subspace estimate as it investigates the approximation error of the entire tensor.

In contrast to the existing literature, we consider a novel approach to analyze the HOSVD-based subspace estimate which is based on a link to the SVD-based subspace estimate via a structured projection. Thereby, we do not need to consider the core tensor or a perturbation for it, which significantly simplifies the development of the analytical performance results. To find a first-order expansion of the HOSVD-based subspace estimate via the structured projection we can then apply the product rule to expand it into first-order approximations of the individual factors. Moreover, we apply these results to find the analytical performance of Tensor-ESPRIT-type algorithms. It is a particular strength of this framework that many extensions and modifications of ESPRIT are easily incorporated, e.g., forward-backward-averaging or Structured Least Squares (SLS).

This manuscript extends the earlier conference versions [33], [34] in several aspects. Firstly, the assumptions about the additive noise are further relaxed so that neither circular symmetry nor spatial whiteness is required any longer. Secondly, the perturbation expansion for the HOSVD-based subspace estimate is generalized from the 2-D case presented in [34] to the general R-D case ($R \geq 2$). Thirdly, proofs for all theorems are given, which could not be included into the conference versions. Note that the closed-form MSE expressions for a single source shown in [33] as well as the extension to SLS discussed in [35] could not be accommodated into this manuscript due to the space limitation.

This paper is organized as follows: The notation and the data model are introduced in Sections II-A and II-B, respectively. The subsequent Section III reviews the first-order perturbation of the matrix-based subspace estimate and presents the extension to the tensor case. The performance analysis of ESPRIT-type algorithms is shown in Section IV. Numerical results are presented in Section V before drawing the conclusions in Section VI.

[1]Note that a performance analysis for 2-D Unitary ESPRIT based on [2] was presented in [24].

II. Notation and Data Model

A. Notation

In this paper, we use the terms "tensor" and "N-way array" interchangeably. In order to facilitate the distinction between scalars, matrices, and tensors, the following notation is used: Scalars are denoted as italic letters $(a, b, \ldots, A, B, \ldots, \alpha, \beta, \ldots)$, column vectors as lower-case bold-face letters $(\boldsymbol{a}, \boldsymbol{b}, \ldots)$, matrices as bold-face capitals $(\boldsymbol{A}, \boldsymbol{B}, \ldots)$, and tensors are written as bold-face calligraphic letters $(\boldsymbol{\mathcal{A}}, \boldsymbol{\mathcal{B}}, \ldots)$. Lower-order parts are consistently named: the (i,j)-element of the matrix $\boldsymbol{A}$, is denoted as $a_{i,j}$ and the (i,j,k)-element of a third order tensor $\boldsymbol{\mathcal{B}}$ as $b_{i,j,k}$.

We use the superscripts ${}^{\mathrm{T}}, {}^{\mathrm{H}}, {}^{*}, {}^{-1}, {}^{+}$ for transposition, Hermitian transposition, complex conjugation, matrix inversion, and the Moore-Penrose pseudo inverse of a matrix, respectively. The trace of a matrix $\boldsymbol{A}$ is written as $\mathrm{Tr}(\boldsymbol{A})$. Moreover, the Kronecker product of two matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ is denoted as $\boldsymbol{A} \otimes \boldsymbol{B}$ and the Khatri-Rao product (column-wise Kronecker product) as $\boldsymbol{A} \diamond \boldsymbol{B}$. The operator $\mathrm{vec}\{\boldsymbol{A}\}$ stacks the column of a matrix $\boldsymbol{A} \in \mathbb{C}^{M \times N}$ into a column vector of length $M \cdot N \times 1$. It satisfies the following property

$$\mathrm{vec}\{\boldsymbol{A} \cdot \boldsymbol{X} \cdot \boldsymbol{B}\} = (\boldsymbol{B}^{\mathrm{T}} \otimes \boldsymbol{A}) \cdot \mathrm{vec}\{\boldsymbol{X}\}. \tag{1}$$

We define an N-D tensor $\boldsymbol{\mathcal{A}} \in \mathbb{C}^{I_1 \times I_2 \times \ldots \times I_N}$ as an N-way array of size I_n along the n-th mode. An n-mode vector of $\boldsymbol{\mathcal{A}}$ is an I_n-dimensional vector obtained from $\boldsymbol{\mathcal{A}}$ by varying the index i_n and keeping the other indices fixed. Moreover, a matrix unfolding of the tensor $\boldsymbol{\mathcal{A}}$ along the n-th mode is denoted by $[\boldsymbol{\mathcal{A}}]_{(n)}$ and can be understood as a matrix containing all the n-mode vectors of the tensor $\boldsymbol{\mathcal{A}}$. The order of the columns is chosen in accordance with [5].

The outer product of the tensors $\boldsymbol{\mathcal{A}} \in \mathbb{C}^{I_1 \times I_2 \times \ldots \times I_N}$ and $\boldsymbol{\mathcal{B}} \in \mathbb{C}^{J_1 \times J_2 \times \ldots \times J_M}$ is given by

$$\boldsymbol{\mathcal{C}} = \boldsymbol{\mathcal{A}} \circ \boldsymbol{\mathcal{B}} \in \mathbb{C}^{I_1 \times \ldots \times I_N \times J_1 \times \ldots \times J_M}, \quad \text{where} \quad c_{i_1,i_2,\ldots,i_N,j_1,j_2,\ldots,j_M} = a_{i_1,i_2,\ldots,i_N} \cdot b_{j_1,j_2,\ldots,j_M}. \tag{2}$$

In other words, the tensor $\boldsymbol{\mathcal{C}}$ contains all possible combinations of pairwise products between the elements of $\boldsymbol{\mathcal{A}}$ and $\boldsymbol{\mathcal{B}}$. This operator is very closely related to the Kronecker product defined for matrices.

The n-mode product of a tensor $\boldsymbol{\mathcal{A}} \in \mathbb{C}^{I_1 \times I_2 \times \ldots \times I_N}$ and a matrix $\boldsymbol{U} \in \mathbb{C}^{J_n \times I_n}$ along the n-th mode is denoted as $\boldsymbol{\mathcal{B}} = \boldsymbol{\mathcal{A}} \times_n \boldsymbol{U}$ and defined via

$$\boldsymbol{\mathcal{B}} = \boldsymbol{\mathcal{A}} \times_n \boldsymbol{U} \quad \Leftrightarrow [\boldsymbol{\mathcal{B}}]_{(n)} = \boldsymbol{U} \cdot [\boldsymbol{\mathcal{A}}]_{(n)}, \tag{3}$$

i.e., it may be visualized by multiplying all n-mode vectors of $\boldsymbol{\mathcal{A}}$ from the left-hand side by the matrix $\boldsymbol{U}$. Note that the n-mode product satisfies

$$(\boldsymbol{\mathcal{A}} \times_n \boldsymbol{U}_n) \times_n \boldsymbol{V}_n = \boldsymbol{\mathcal{A}} \times_n (\boldsymbol{V}_n \cdot \boldsymbol{U}_n) \tag{4}$$

$$[\boldsymbol{\mathcal{A}} \times_1 \boldsymbol{U}_1 \ldots \times_N \boldsymbol{U}_N]_{(n)} = \boldsymbol{U}_n \cdot [\boldsymbol{\mathcal{A}}]_{(n)} \cdot (\boldsymbol{U}_{n+1} \otimes \ldots \otimes \boldsymbol{U}_N \otimes \boldsymbol{U}_1 \otimes \ldots \otimes \boldsymbol{U}_{n-1})^{\mathrm{T}} \tag{5}$$

for $\boldsymbol{\mathcal{A}} \in \mathbb{C}^{I_1 \times I_2 \times \ldots \times I_N}$, $\boldsymbol{U}_n \in \mathbb{C}^{J_n \times I_n}$ and $\boldsymbol{V}_n \in \mathbb{C}^{K_n \times J_n}$.

The higher-order SVD (HOSVD) [5] of a tensor $\boldsymbol{\mathcal{A}} \in \mathbb{C}^{I_1 \times I_2 \times \ldots \times I_N}$ is given by

$$\boldsymbol{\mathcal{A}} = \boldsymbol{\mathcal{S}} \times_1 \boldsymbol{U}_1 \times_2 \boldsymbol{U}_2 \ldots \times_N \boldsymbol{U}_N, \tag{6}$$

where $\boldsymbol{\mathcal{S}} \in \mathbb{C}^{I_1 \times I_2 \times \ldots \times I_N}$ is the core tensor which satisfies the all-orthogonality conditions [5] and $\boldsymbol{U}_n \in \mathbb{C}^{I_n \times I_n}$, $n = 1, 2, \ldots, N$, are the unitary matrices of n-mode singular vectors.

We also define the concatenation of two tensors along the n-th mode via the operator $[\boldsymbol{\mathcal{A}} \sqcup_n \boldsymbol{\mathcal{B}}]$ [12]. The Euclidean (vector) norm, the Frobenius (matrix) norm, and the Higher-Order Frobenius (tensor) norm are denoted by $\|\boldsymbol{a}\|_2$, $\|\boldsymbol{A}\|_{\mathrm{F}}$, and $\|\boldsymbol{\mathcal{A}}\|_{\mathrm{H}}$, respectively. All three norms are computed by taking the square-root of the sum of the squared magnitude of all the elements in their arguments.

The matrix $\boldsymbol{K}_{M \times N} \in \mathbb{R}^{MN \times MN}$ denotes the commutation matrix [23] which satisfies

$$\boldsymbol{K}_{M \times N} \cdot \mathrm{vec}\{\boldsymbol{A}^{\mathrm{T}}\} = \mathrm{vec}\{\boldsymbol{A}\} \tag{7}$$

$$\boldsymbol{K}_{M \times P}^{\mathrm{T}} \cdot (\boldsymbol{A} \otimes \boldsymbol{B}) \cdot \boldsymbol{K}_{N \times Q} = \boldsymbol{B} \otimes \boldsymbol{A} \tag{8}$$

for $\boldsymbol{A} \in \mathbb{C}^{M \times N}$, $\boldsymbol{B} \in \mathbb{C}^{P \times Q}$.

A $p \times p$ matrix $\boldsymbol{Q}_p$ is called left-$\boldsymbol{\Pi}$-real if $\boldsymbol{\Pi}_p \cdot \boldsymbol{Q}_p^* = \boldsymbol{Q}_p$, where $\boldsymbol{\Pi}_p$ is the $p \times p$ exchange matrix with ones on its antidiagonal and zeros elsewhere [18]. The special set of unitary sparse left-$\boldsymbol{\Pi}$ -real matrices used in [10] is denoted as $\boldsymbol{Q}_p^{(\mathrm{s})}$. Furthermore, a matrix $\boldsymbol{X} \in \mathbb{C}^{M \times N}$ is called centro-Hermitian $\boldsymbol{\Pi}_M \cdot \boldsymbol{X}^* \cdot \boldsymbol{\Pi}_N = \boldsymbol{X}$. The R-D identity tensor of size $d \times d \times \ldots \times d$ is denoted by $\boldsymbol{\mathcal{I}}_{R,d}$. It is equal to 1 if all its indices are equal and zero elsewhere.

B. Matrix-Based and Tensor-Based Data Model

The observations are modeled as a superposition of d undamped exponentials sampled on an R-dimensional grid of size $M_1 \times M_2 \times \ldots \times M_R$ at N subsequent time instants [11]. The measurement samples are given by[2]

$$x_{m_1,m_2,\ldots,m_R,t_n} = \sum_{i=1}^{d} s_i(t_n) \prod_{r=1}^{R} e^{\jmath \cdot (m_r - 1) \cdot \mu_i^{(r)}} + n_{m_1,m_2,\ldots,m_R,t_n}, \tag{9}$$

where $m_r = 1, 2, \ldots, M_r$, $n = 1, 2, \ldots, N$, $s_i(t_n)$ denotes the complex amplitude of the i-th exponential at time instant t_n, $\mu_i^{(r)}$ symbolizes the frequency of the i-th exponential in the r-th mode for $i = 1, 2, \ldots, d$ and $r = 1, 2, \ldots, R$. Moreover, $n_{m_1,m_2,\ldots,m_R,t_n}$ represents the zero mean additive noise component inherent in the measurement process. In the context of array signal processing, each of the R-dimensional exponentials represents one planar wavefront and the complex amplitudes $s_i(t_n)$ are the symbols. It is our goal to estimate the frequencies $\mu_i^{(r)}$ for $r = 1, 2, \ldots, R$, $i = 1, 2, \ldots, d$ and their correct pairing.

In order to arrive at a more compact formulation of the data model in (9) we collect the samples $x_{m_1,m_2,\ldots,m_R,t_n}$ into one

[2]Note that (9) assumes a uniform sampling in the spatial domain. However, this assumption can be relaxed to more generic geometries as long as they feature shift invariances and can be constructed as the outer product of R one-dimensional sampling grids.

array. As our signal is referenced by $R+1$ indices, the most natural way of formulating the model is to employ an $(R+1)$-way array $\boldsymbol{X} \in \mathbb{C}^{M_1 \times M_2 \ldots \times M_R \times N}$ which contains $x_{m_1,m_2,\ldots,m_R,t_n}$ for $m_r = 1,2,\ldots,M_r$, $r = 1,2,\ldots,R$, and $n = 1,2,\ldots,N$. We can then conveniently express $\boldsymbol{X}$ as [12]

$$\boldsymbol{\mathcal{X}} = \boldsymbol{\mathcal{A}} \times_{R+1} \boldsymbol{S}^{\mathrm{T}} + \boldsymbol{\mathcal{N}}, \tag{10}$$

where $\boldsymbol{S} \in \mathbb{C}^{d \times N}$ contains the amplitudes $s_i[n]$ and $\boldsymbol{\mathcal{N}} \in \mathbb{C}^{M_1 \times M_2 \ldots \times M_R \times N}$ collects all the noise samples $n_{m_1,m_2,\ldots,m_R,t_n}$ in the same manner as $\boldsymbol{\mathcal{X}}$. Finally, $\boldsymbol{\mathcal{A}} \in \mathbb{C}^{M_1 \times M_2 \ldots \times M_R \times d}$ is referred to as the "array steering tensor" [12]. It can be expressed by virtue of the concatenation operator via

$$\begin{aligned} \boldsymbol{\mathcal{A}} &= [\boldsymbol{\mathcal{A}}_1 \sqcup_{R+1} \boldsymbol{\mathcal{A}}_2 \sqcup_{R+1} \cdots \sqcup_{R+1} \boldsymbol{\mathcal{A}}_d] \\ \boldsymbol{\mathcal{A}}_i &= \boldsymbol{a}^{(1)}\left(\mu_i^{(1)}\right) \circ \boldsymbol{a}^{(2)}\left(\mu_i^{(2)}\right) \circ \ldots \circ \boldsymbol{a}^{(R)}\left(\mu_i^{(R)}\right) \\ &\in \mathbb{C}^{M_1 \times M_2 \times \ldots \times M_R} \end{aligned} \tag{11}$$

where $\boldsymbol{a}^{(r)}(\mu_i^{(r)}) \in \mathbb{C}^{M_r \times 1}$ represents the array steering vector of the i-th source in the r-th mode. An alternative expression for the array steering tensor is given by

$$\boldsymbol{\mathcal{A}} = \boldsymbol{\mathcal{I}}_{R+1,d} \times_1 \boldsymbol{A}^{(1)} \times_2 \boldsymbol{A}^{(2)} \ldots \times_R \boldsymbol{A}^{(R)}, \tag{12}$$

where $\boldsymbol{A}^{(r)} = [\boldsymbol{a}^{(r)}(\mu_1^{(r)}) \ \ldots \ \boldsymbol{a}^{(r)}(\mu_d^{(r)})] \in \mathbb{C}^{M_r \times d}$ is referred to as the array steering matrix in the r-th mode.

The strength of the data model in (10) is that it represents the signal in its natural multidimensional structure by virtue of the measurement tensor $\boldsymbol{X}$. Before tensor calculus was used in this area, a matrix-based formulation of (10) was needed. This requires stacking some of the dimensions into rows or columns. A meaningful definition of a measurement matrix $\boldsymbol{X}$ is to apply stacking to all "spatial" dimensions $1,2,\ldots,R$ along the rows and align the snapshots $n = 1,2,\ldots,N$ as the columns. Mathematically, we can write $\boldsymbol{X} = [\boldsymbol{\mathcal{X}}]^{\mathrm{T}}_{(R+1)} \in \mathbb{C}^{M \times N}$, where $M = \prod_{r=1}^{R} M_r$. Applying this stacking operation to (10), we arrive at the matrix-based data model [11]

$$\boldsymbol{X} = \boldsymbol{A} \cdot \boldsymbol{S} + \boldsymbol{N}, \quad \text{where} \tag{13}$$

$$\boldsymbol{A} = [\boldsymbol{\mathcal{A}}]^{\mathrm{T}}_{(R+1)} \in \mathbb{C}^{M \times d} \text{ and } \boldsymbol{N} = [\boldsymbol{\mathcal{N}}]^{\mathrm{T}}_{(R+1)} \in \mathbb{C}^{M \times N} \tag{14}$$

Note that $\boldsymbol{A}$ is highly structured since it satisfies

$$\boldsymbol{A} = [\boldsymbol{\mathcal{A}}]^{\mathrm{T}}_{(R+1)} = \boldsymbol{A}^{(1)} \diamond \boldsymbol{A}^{(2)} \diamond \ldots \diamond \boldsymbol{A}^{(R)}. \tag{15}$$

C. Subspace Estimation

The first step in all subspace-based parameter estimation schemes is the estimation of a basis for the signal subspace from the noisy observations. In the matrix case, this can, for instance, be achieved by a truncated SVD of $\boldsymbol{X}$. Let $\hat{\boldsymbol{U}}_{\mathrm{s}} \in \mathbb{C}^{M \times d}$ be the matrix containing the d dominant left singular vectors of $\boldsymbol{X}$. Then the column space of $\hat{\boldsymbol{U}}_{\mathrm{s}}$ is an estimate for the signal subspace spanned by the columns of $\boldsymbol{A}$ and we can write $\boldsymbol{A} \approx \hat{\boldsymbol{U}}_{\mathrm{s}} \cdot \boldsymbol{T}$ for a non-singular matrix $\boldsymbol{T} \in \mathbb{C}^{d \times d}$.

A tensor-based extension of this subspace estimate was proposed in [12]. To this end, let the truncated HOSVD of $\boldsymbol{\mathcal{X}}$ be given by

$$\boldsymbol{\mathcal{X}} \approx \hat{\boldsymbol{\mathcal{S}}}^{[\mathrm{s}]} \times_1 \hat{\boldsymbol{U}}_1^{[\mathrm{s}]} \ldots \times_R \hat{\boldsymbol{U}}_R^{[\mathrm{s}]} \times_{R+1} \hat{\boldsymbol{U}}_{R+1}^{[\mathrm{s}]}, \tag{16}$$

where $\hat{\boldsymbol{\mathcal{S}}}^{[\mathrm{s}]} \in \mathbb{C}^{p_1 \times \ldots \times p_R \times d}$ is the truncated core tensor and $\hat{\boldsymbol{U}}_r^{[\mathrm{s}]} \in \mathbb{C}^{M_r \times p_r}$ for $r = 1,2,\ldots,R$, $\hat{\boldsymbol{U}}_{R+1}^{[\mathrm{s}]} \in \mathbb{C}^{N \times d}$ are the matrices of dominant r-mode singular vectors. Moreover, $p_r = \min(M_r, d)$.[3] Based on (16), a tensor-based subspace estimate can be defined as

$$\hat{\boldsymbol{\mathcal{U}}}^{[\mathrm{s}]} = \hat{\boldsymbol{\mathcal{S}}}^{[\mathrm{s}]} \times_1 \hat{\boldsymbol{U}}_1^{[\mathrm{s}]} \ldots \times_R \hat{\boldsymbol{U}}_R^{[\mathrm{s}]} \times_{R+1} \hat{\boldsymbol{\Sigma}}_{R+1}^{[\mathrm{s}]^{-1}} \tag{17}$$

Note that the $(R+1)$-mode multiplication with $\hat{\boldsymbol{\Sigma}}_{R+1}^{[\mathrm{s}]^{-1}}$ is introduced in addition to its original definition in [12] since this normalization simplifies the notation we need at this point and it has no impact on the subspace estimation accuracy. Here, $\hat{\boldsymbol{\Sigma}}_{R+1}^{[\mathrm{s}]} \in \mathbb{R}^{d \times d}$ is the diagonal matrix containing the d dominant singular values of $[\boldsymbol{\mathcal{X}}]_{(R+1)}$ on its main diagonal. These are identical to the d dominant singular values of $\boldsymbol{X}$, since since $[\boldsymbol{\mathcal{X}}]^{\mathrm{T}}_{(R+1)} = \boldsymbol{X}$.

Note that $\hat{\boldsymbol{\mathcal{U}}}^{[\mathrm{s}]}$ satisfies $\hat{\boldsymbol{\mathcal{U}}}^{[\mathrm{s}]} \approx \boldsymbol{\mathcal{A}} \times_{R+1} \bar{\boldsymbol{T}}$ for a non-singular matrix $\bar{\boldsymbol{T}} \in \mathbb{C}^{d \times d}$. This approximation becomes asymptotically exact in the high effective SNR regime, i.e., for either a vanishing noise power or a growing number of snapshots N. Based on $\hat{\boldsymbol{\mathcal{U}}}^{[\mathrm{s}]}$, an improved signal subspace estimate is given by the matrix $[\hat{\boldsymbol{\mathcal{U}}}^{[\mathrm{s}]}]^{\mathrm{T}}_{(R+1)} \in \mathbb{C}^{M \times d}$. Its estimation accuracy is analyzed in the next section.

III. Perturbations of the Subspace Estimates

A. Review of Perturbation Results for the SVD

Let us first review the results from [19] which are relevant to the discussion in this section. As in the tensor case, let $\boldsymbol{X}_0 = \boldsymbol{A} \cdot \boldsymbol{S} \in \mathbb{C}^{M \times N}$ be a matrix containing the noise-free observations such that $\boldsymbol{X} = \boldsymbol{X}_0 + \boldsymbol{N}$ where $\boldsymbol{N}$ represents the undesired perturbation (noise).

The SVD of $\boldsymbol{X}_0$ can be expressed as

$$\boldsymbol{X}_0 = [\boldsymbol{U}_{\mathrm{s}} \, \boldsymbol{U}_{\mathrm{n}}] \cdot \begin{bmatrix} \boldsymbol{\Sigma}_{\mathrm{s}} & \boldsymbol{0}_{d \times (N-d)} \\ \boldsymbol{0}_{(M-d) \times d} & \boldsymbol{0}_{(M-d) \times (N-d)} \end{bmatrix} \cdot [\boldsymbol{V}_{\mathrm{s}} \, \boldsymbol{V}_{\mathrm{n}}]^{\mathrm{H}}, \tag{18}$$

where the columns of $\boldsymbol{U}_{\mathrm{s}} \in \mathbb{C}^{M \times d}$ provide an orthonormal basis for the signal subspace which we want to estimate. Moreover $\boldsymbol{\Sigma}_{\mathrm{s}} = \mathrm{diag}([\sigma_1, \sigma_2, \ldots, \sigma_d]) \in \mathbb{R}^{d \times d}$ contains the d non-zero singular values on its main diagonal. We find an estimate for $\boldsymbol{U}_{\mathrm{s}}$ by computing an SVD of the noisy observation matrix $\boldsymbol{X}$ which can be expressed as

$$\boldsymbol{X} = \begin{bmatrix} \hat{\boldsymbol{U}}_{\mathrm{s}} & \hat{\boldsymbol{U}}_{\mathrm{n}} \end{bmatrix} \cdot \begin{bmatrix} \hat{\boldsymbol{\Sigma}}_{\mathrm{s}} & \boldsymbol{0}_{d \times (N-d)} \\ \boldsymbol{0}_{(M-d) \times d} & \hat{\boldsymbol{\Sigma}}_{\mathrm{n}} \end{bmatrix} \cdot \begin{bmatrix} \hat{\boldsymbol{V}}_{\mathrm{s}} & \hat{\boldsymbol{V}}_{\mathrm{n}} \end{bmatrix}^{\mathrm{H}}, \tag{19}$$

[3]In the presence of noise we can estimate the model order d using tensor-based model order estimation algorithms [3].

where the "hat" denotes the estimated quantities. We can write $\hat{\boldsymbol{U}}_{\mathrm{s}} = \boldsymbol{U}_{\mathrm{s}} + \Delta\boldsymbol{U}_{\mathrm{s}}$, where $\Delta\boldsymbol{U}_{\mathrm{s}}$ represents the estimation error. At this point we can state the main result on the first order perturbation expansion of $\Delta\boldsymbol{U}_{\mathrm{s}}$ from [19]

$$\Delta\boldsymbol{U}_{\mathrm{s}} = \boldsymbol{U}_{\mathrm{n}} \cdot \boldsymbol{\Gamma}_{\mathrm{n}} + \mathcal{O}\{\Delta^2\}, \text{ where } \Delta = \|\boldsymbol{N}\| \text{ and } \boldsymbol{\Gamma}_{\mathrm{n}} = \boldsymbol{U}_{\mathrm{n}}^{\mathrm{H}} \cdot \boldsymbol{N} \cdot \boldsymbol{V}_{\mathrm{s}} \cdot \boldsymbol{\Sigma}_{\mathrm{s}}^{-1} \in \mathbb{C}^{(M-d)\times d}. \quad (20)$$

Here $\|\cdot\|$ represents an arbitrary sub-multiplicative[4] norm, e.g., the Frobenius norm. Equation (20) shows the first order expansion of the signal subspace estimation error $\Delta\boldsymbol{U}_{\mathrm{s}}$ in terms of the noise subspace (spanned by the columns of $\boldsymbol{U}_{\mathrm{n}}$), i.e., how much of the estimated signal subspace (spanned by the columns of $\hat{\boldsymbol{U}}_{\mathrm{s}}$) lies in the true noise subspace (spanned by the columns of $\boldsymbol{U}_{\mathrm{n}}$) due to the estimation errors from the perturbation $\boldsymbol{N}$. Since it is explicit in $\boldsymbol{N}$, it makes no assumptions about the *statistics* of $\boldsymbol{N}$. In fact, the first order expansion is purely deterministic.

The expansion (20) ignores the contribution of $\Delta\boldsymbol{U}_{\mathrm{s}}$ that lies in the true signal subspace spanned by the columns of $\boldsymbol{U}_{\mathrm{s}}$, which would measure the misalignment of the particular basis we choose to span the signal subspace. There are extensions of (20) to take the perturbation of the particular basis into account, e.g., [21]. However, we do not consider this additional term in this paper since our focus is on the perturbation analysis for subspace-based parameter estimation schemes, for which the choice of the particular basis is irrelevant.

B. Extension to the HOSVD-Based Subspace Estimate

As we have shown in Section II-C, in the multidimensional case, an improved signal subspace estimate can be computed via the HOSVD of the measurement tensor $\boldsymbol{\mathcal{X}}$. Since the HOSVD is computed via SVDs of the unfoldings, we can apply the same framework to find a perturbation expansion of the HOSVD-based subspace estimate. In order to distinguish unperturbed from estimated (perturbed) quantities we express $\boldsymbol{\mathcal{X}}$ as $\boldsymbol{\mathcal{X}} = \boldsymbol{\mathcal{X}}_0 + \boldsymbol{\mathcal{N}}$, where $\boldsymbol{\mathcal{X}}_0 = \boldsymbol{\mathcal{A}} \times_{R+1} \boldsymbol{S}^{\mathrm{T}}$ is the unperturbed observation tensor. The SVDs of the r-mode unfoldings of $\boldsymbol{\mathcal{X}}$ and $\boldsymbol{\mathcal{X}}_0$ are then given by

$$[\boldsymbol{\mathcal{X}}_0]_{(r)} = \begin{bmatrix} \boldsymbol{U}_r^{[\mathrm{s}]} & \boldsymbol{U}_r^{[\mathrm{n}]} \end{bmatrix} \cdot \begin{bmatrix} \boldsymbol{\Sigma}_r^{[\mathrm{s}]} & \boldsymbol{0}_{d\times(M\cdot N/M_r-d)} \\ \boldsymbol{0}_{(M_r-d)\times d} & \boldsymbol{0}_{(M_r-d)\times(M\cdot N/M_r-d)} \end{bmatrix} \cdot \begin{bmatrix} \boldsymbol{V}_r^{[\mathrm{s}]} & \boldsymbol{V}_r^{[\mathrm{n}]} \end{bmatrix}^{\mathrm{H}} \quad (21)$$

$$[\boldsymbol{\mathcal{X}}]_{(r)} = \begin{bmatrix} \hat{\boldsymbol{U}}_r^{[\mathrm{s}]} & \hat{\boldsymbol{U}}_r^{[\mathrm{n}]} \end{bmatrix} \cdot \begin{bmatrix} \hat{\boldsymbol{\Sigma}}_r^{[\mathrm{s}]} & \boldsymbol{0}_{d\times(M\cdot N/M_r-d)} \\ \boldsymbol{0}_{(M_r-d)\times d} & \hat{\boldsymbol{\Sigma}}_r^{[\mathrm{n}]} \end{bmatrix} \cdot \begin{bmatrix} \hat{\boldsymbol{V}}_r^{[\mathrm{s}]} & \hat{\boldsymbol{V}}_r^{[\mathrm{n}]} \end{bmatrix}^{\mathrm{H}}. \quad (22)$$

where $\boldsymbol{\Sigma}_r^{[\mathrm{s}]} = \mathrm{diag}([\sigma_1^{(r)}, \sigma_2^{(r)}, \ldots, \sigma_d^{(r)}])$ and $r = 1, 2, \ldots, R$. Note that since (21) and (22) are in fact SVDs, we can apply (20) and find $\hat{\boldsymbol{U}}_r^{[\mathrm{s}]} = \boldsymbol{U}_r^{[\mathrm{s}]} + \Delta\boldsymbol{U}_r^{[\mathrm{s}]}$ where

$$\begin{aligned} \Delta\boldsymbol{U}_r^{[\mathrm{s}]} &= \boldsymbol{U}_r^{[\mathrm{n}]} \cdot \boldsymbol{\Gamma}_r^{[\mathrm{n}]} + \mathcal{O}\{\Delta^2\}, \\ \boldsymbol{\Gamma}_r^{[\mathrm{n}]} &= \boldsymbol{U}_r^{[\mathrm{n}]^{\mathrm{H}}} \cdot [\boldsymbol{\mathcal{N}}]_{(r)} \cdot \boldsymbol{V}_r^{[\mathrm{s}]} \cdot \boldsymbol{\Sigma}_r^{[\mathrm{s}]^{-1}}, \end{aligned} \quad (23)$$

[4]A matrix norm is called submultiplicative if $\|\boldsymbol{A} \cdot \boldsymbol{B}\| \leq \|\boldsymbol{A}\| \cdot \|\boldsymbol{B}\|$ for arbitrary matrices $\boldsymbol{A}$ and $\boldsymbol{B}$.

Our goal is to use the perturbation of the r-mode unfoldings to find a corresponding expansion for the HOSVD-based subspace estimate introduced in Section II-C. This is facilitated by the following theorem:

Theorem 1: The HOSVD-based subspace estimate $[\hat{\boldsymbol{\mathcal{U}}}^{[\mathrm{s}]}]_{(R+1)}^{\mathrm{T}}$ defined in (17) is linked to the SVD-based subspace estimate $\hat{\boldsymbol{U}}_{\mathrm{s}}$ via the following algebraic relation

$$\left[\hat{\boldsymbol{\mathcal{U}}}^{[\mathrm{s}]}\right]_{(R+1)}^{\mathrm{T}} = \left(\hat{\boldsymbol{T}}_1 \otimes \hat{\boldsymbol{T}}_2 \otimes \ldots \otimes \hat{\boldsymbol{T}}_R\right) \cdot \hat{\boldsymbol{U}}_{\mathrm{s}}, \quad (24)$$

where $\hat{\boldsymbol{T}}_r \in \mathbb{C}^{M_r \times M_r}$ represent estimates of the projection matrices onto the r-spaces of $\boldsymbol{X}_0$, which are computed via $\hat{\boldsymbol{T}}_r = \hat{\boldsymbol{U}}_r^{[\mathrm{s}]} \hat{\boldsymbol{U}}_r^{[\mathrm{s}]^{\mathrm{H}}}$.

Proof: Relation (24) was shown in [34] for $R = 2$. The proof for $R > 2$ proceeds in an analogous manner and is presented in Appendix A.

Equation (24) shows that the HOSVD-based signal subspace estimate is related to the SVD-based signal subspace estimate via a projection matrix formed from the Kronecker product of projections onto the r-mode subspaces for $r = 1, 2, \ldots, R$. Thereby, the multidimensional structure is "imprinted" onto the matrix-based subspace estimate which shows how the exploitation of the R-D structure improves the signal subspace estimate. Moreover, we observe from (24) that a perturbation expansion for $[\hat{\boldsymbol{\mathcal{U}}}^{[\mathrm{s}]}]_{(R+1)}^{\mathrm{T}}$ can be developed based on the subspaces of all $R+1$ unfoldings, as the core tensor is not needed for its computation. The result is shown in the following theorem:

Theorem 2: The HOSVD-based signal subspace estimate can be written as $[\hat{\boldsymbol{\mathcal{U}}}^{[\mathrm{s}]}]_{(R+1)}^{\mathrm{T}} = \boldsymbol{U}_{\mathrm{s}} + [\Delta\hat{\boldsymbol{\mathcal{U}}}^{[\mathrm{s}]}]_{(R+1)}^{\mathrm{T}}$, where

$$\begin{aligned} &\left[\Delta\hat{\boldsymbol{\mathcal{U}}}^{[\mathrm{s}]}\right]_{(R+1)}^{\mathrm{T}} \\ &= (\boldsymbol{T}_1 \otimes \boldsymbol{T}_2 \otimes \ldots \otimes \boldsymbol{T}_R) \cdot \Delta\boldsymbol{U}_{\mathrm{s}} \\ &+ \left(\left[\Delta\boldsymbol{U}_1^{[\mathrm{s}]} \cdot \boldsymbol{U}_1^{[\mathrm{s}]^{\mathrm{H}}}\right] \otimes \boldsymbol{T}_2 \otimes \ldots \otimes \boldsymbol{T}_R\right) \cdot \boldsymbol{U}_{\mathrm{s}} \\ &+ \left(\boldsymbol{T}_1 \otimes \left[\Delta\boldsymbol{U}_2^{[\mathrm{s}]} \cdot \boldsymbol{U}_2^{[\mathrm{s}]^{\mathrm{H}}}\right] \otimes \ldots \otimes \boldsymbol{T}_R\right) \cdot \boldsymbol{U}_{\mathrm{s}} + \ldots \\ &+ \left(\boldsymbol{T}_1 \otimes \boldsymbol{T}_2 \otimes \ldots \otimes \left[\Delta\boldsymbol{U}_R^{[\mathrm{s}]} \cdot \boldsymbol{U}_R^{[\mathrm{s}]^{\mathrm{H}}}\right]\right) \cdot \boldsymbol{U}_{\mathrm{s}} + \mathcal{O}\{\Delta^2\}, \end{aligned} \quad (25)$$

the SVD-based signal subspace perturbation $\Delta\boldsymbol{U}_{\mathrm{s}}$ is given by (20) and the perturbations of the r-spaces can be computed via

$$\Delta\boldsymbol{U}_r^{[\mathrm{s}]} = \boldsymbol{U}_r^{[\mathrm{n}]} \cdot \boldsymbol{\Gamma}_r^{[\mathrm{n}]} = \boldsymbol{U}_r^{[\mathrm{n}]} \cdot \boldsymbol{U}_r^{[\mathrm{n}]^{\mathrm{H}}} \cdot [\boldsymbol{\mathcal{N}}]_{(r)} \cdot \boldsymbol{V}_r^{[\mathrm{s}]} \cdot \boldsymbol{\Sigma}_r^{[\mathrm{s}]^{-1}}. \quad (26)$$

Proof: cf. Appendix B.

IV. Asymptotical Analysis of the Parameter Estimation Accuracy

A. Review of Results for the 1-D Standard ESPRIT

In [19] the authors point out that once a first order expansion of the subspace estimation error is available it can be used to find a corresponding first order expansion of the estimation error of a suitable parameter estimation scheme. One of the examples the authors show is the 1-D Standard ESPRIT algorithm using

LS, which we use as a starting point to discuss various ESPRIT-type algorithms in this section. In the noise-free case, the shift invariance equation for 1-D Standard ESPRIT can be expressed as

$$\boldsymbol{J}_1 \cdot \boldsymbol{U}_{\rm s} \cdot \boldsymbol{\Psi} = \boldsymbol{J}_2 \cdot \boldsymbol{U}_{\rm s}. \tag{27}$$

where $\boldsymbol{J}_1, \boldsymbol{J}_2 \in \mathbb{R}^{M^{(\rm sel)} \times M}$ are the selection matrices that select the $M^{(\rm sel)}$ elements from the M antenna elements which correspond to the first and the second subarray, respectively. Moreover, $\boldsymbol{\Psi} = \boldsymbol{Q} \cdot \boldsymbol{\Phi} \cdot \boldsymbol{Q}^{-1}$, where $\boldsymbol{\Phi} = \operatorname{diag}([{\rm e}^{\jmath\mu_1}, \ldots, {\rm e}^{\jmath\mu_d}]) \in \mathbb{C}^{d \times d}$ contains the spatial frequencies μ_k, $k = 1, 2, \ldots, d$ that we want to estimate. Therefore, $\mu_k = \arg(\mathrm{EV}_k\{\boldsymbol{\Psi}\})$, i.e., the k-th spatial frequency is obtained from the phase of the k-th eigenvalue $(\mathrm{EV}_k\{\cdot\})$ of $\boldsymbol{\Psi}$.

In presence of noise, we only have an estimate $\hat{\boldsymbol{U}}_{\rm s}$ of the signal subspace $\boldsymbol{U}_{\rm s}$. Consequently, (27) in general does not have an exact solution anymore. A simple way of finding an approximate $\hat{\boldsymbol{\Psi}}$ is given by the LS solution which can be expressed as

$$\hat{\boldsymbol{\Psi}}_{\rm LS} = \left(\boldsymbol{J}_1 \cdot \hat{\boldsymbol{U}}_{\rm s}\right)^+ \cdot \boldsymbol{J}_2 \cdot \hat{\boldsymbol{U}}_{\rm s}. \tag{28}$$

To simplify the notation we skip the index "LS" for the remainder of this section. For the estimation error of the k-th spatial frequency corresponding to the LS solution (28), [19] provides the following expansion

$$\Delta\mu_k = \operatorname{Im}\left\{\boldsymbol{p}_k^{\rm T} \cdot (\boldsymbol{J}_1 \cdot \boldsymbol{U}_{\rm s})^+ \cdot [\boldsymbol{J}_2/\lambda_k - \boldsymbol{J}_1] \cdot \Delta\boldsymbol{U}_{\rm s} \cdot \boldsymbol{q}_k\right\} + \mathcal{O}\{\Delta^2\}, \tag{29}$$

where $\lambda_k = {\rm e}^{\jmath\mu_k}$ and $\boldsymbol{q}_k$ is the k-th column of $\boldsymbol{Q}$. Moreover, $\boldsymbol{p}_k^{\rm T}$ represents the k-th row vector of the matrix $\boldsymbol{P} = \boldsymbol{Q}^{-1}$. Note that $\Delta\boldsymbol{U}_{\rm s}$ can be expanded in terms of the perturbation term $\boldsymbol{N}$ by using the expansion (20).

B. Extension to R-D Standard (Tensor-)ESPRIT

The previous result from [19] on the first order perturbation expansion of 1-D Standard ESPRIT using LS is easily generalized to the R-D case. The reason is that for R-D LS-based ESPRIT, the R shift invariance equations are solved independently from each other. The resulting matrices $\hat{\boldsymbol{\Psi}}^{(r)}$ possess the following eigendecomposition

$$\hat{\boldsymbol{\Psi}}^{(r)} = \hat{\boldsymbol{Q}}^{(r)} \cdot \hat{\boldsymbol{\Phi}}^{(r)} \cdot \hat{\boldsymbol{Q}}^{(r)^{-1}}. \tag{30}$$

Note that $\hat{\boldsymbol{Q}}^{(r)} = \boldsymbol{Q} + \Delta\boldsymbol{Q}^{(r)}$ where $\boldsymbol{Q} = \boldsymbol{U}_{\rm s}^{\rm H} \cdot \boldsymbol{A} \in \mathbb{C}^{d \times d}$ is the noise-free eigenspace that is common to all matrices $\boldsymbol{\Psi}^{(r)}$ for $r = 1, 2, \ldots, R$. Moreover, $\hat{\boldsymbol{\Phi}}^{(r)} = \boldsymbol{\Phi}^{(r)} + \Delta\boldsymbol{\Phi}^{(r)}$, where $\boldsymbol{\Phi}^{(r)} = \operatorname{diag}([{\rm e}^{\jmath\mu_1^{(r)}} \ \ldots \ {\rm e}^{\jmath\mu_d^{(r)}}]) \in \mathbb{C}^{d \times d}$. It is shown in [17] that a first-order expansion of an eigendecomposition like (30) is given by

$$\hat{\boldsymbol{\Psi}}^{(r)} = \boldsymbol{\Psi}^{(r)} + \boldsymbol{Q} \cdot \Delta\boldsymbol{\Phi}^{(r)} \cdot \boldsymbol{Q}^{-1} + \mathcal{O}\{\Delta^2\}, \tag{31}$$

where $\boldsymbol{\Psi}^{(r)} = \boldsymbol{Q} \cdot \boldsymbol{\Phi}^{(r)} \cdot \boldsymbol{Q}^{-1}$. In other words, (31) shows that the perturbation of the eigenvectors $\Delta\boldsymbol{Q}^{(r)}$ is quadratic and hence asymptotically negligible with respect to the perturbation of the eigenvalues $\Delta\boldsymbol{\Phi}^{(r)}$.

Since (31) has the same form as in the 1-D case studied in [19], the same arguments are readily applied to all modes individually and we directly obtain a first order expansion for the estimation error of the k-th spatial frequency in the r-th mode

$$\Delta\mu_k^{(r)} = \operatorname{Im}\left\{\boldsymbol{p}_k^{\rm T} \cdot \left(\tilde{\boldsymbol{J}}_1^{(r)} \cdot \boldsymbol{\mathcal{U}}_{\rm s}\right)^+ \cdot \left[\tilde{\boldsymbol{J}}_2^{(r)}/\lambda_k^{(r)} - \tilde{\boldsymbol{J}}_1^{(r)}\right] \cdot \Delta\boldsymbol{\mathcal{U}}_{\rm s} \cdot \boldsymbol{q}_k\right\} + \mathcal{O}\{\Delta^2\}, \tag{32}$$

where $\tilde{\boldsymbol{J}}_1^{(r)}, \tilde{\boldsymbol{J}}_2^{(r)} \in \mathbb{R}^{\frac{M}{M_r} \cdot M_r^{(\rm sel)} \times M}$ are the effective R-D selection matrix for the first and the second subarray in the r-th mode, respectively. They can be expressed as $\tilde{\boldsymbol{J}}_\ell^{(r)} = \boldsymbol{I}_{\prod_{n=1}^{r-1} M_n} \otimes \boldsymbol{J}_\ell^{(r)} \otimes \boldsymbol{I}_{\prod_{n=r+1}^{R} M_n}$, for $\ell = 1,2$ and $r = 1, 2, \ldots, R$, where $\boldsymbol{J}_\ell^{(r)} \in \mathbb{R}^{M_r^{(\rm sel)} \times M_r}$ are the selection matrices which select the $M_r^{(\rm sel)}$ elements belonging to the first and the second subarray in the r-th mode, respectively.

Since this expansion for R-D Standard ESPRIT is explicit in the perturbation of the subspace estimate and R-D Standard Tensor-ESPRIT only differs in the fact that it uses the enhanced HOSVD-based subspace estimate, we can conclude that a first order perturbation expansion for R-D Standard Tensor-ESPRIT is given by

$$\Delta\mu_k^{(r)} = \operatorname{Im}\left\{\boldsymbol{p}_k^{\rm T} \cdot \left(\tilde{\boldsymbol{J}}_1^{(r)} \cdot \boldsymbol{U}_{\rm s}\right)^+ \cdot \left[\tilde{\boldsymbol{J}}_2^{(r)}/\lambda_k^{(r)} - \tilde{\boldsymbol{J}}_1^{(r)}\right] \cdot \left[\Delta\hat{\boldsymbol{\mathcal{U}}}^{[\rm s]}\right]_{R+1}^{\rm T} \cdot \boldsymbol{q}_k\right\} + \mathcal{O}\{\Delta^2\}. \tag{33}$$

An explicit expansion of $\Delta\mu_k^{(r)}$ in terms of the noise tensor $\boldsymbol{\mathcal{N}}$ is obtained by inserting the previous result (25).

Note that this discussion ignores the correct pairing of the parameters across the dimensions. In ESPRIT-type algorithms, this pairing is usually achieved by computing the eigenvalues of the estimated matrices $\hat{\boldsymbol{\Psi}}^{(r)}$ jointly via a joint approximate eigendecomposition [9] or a Simultaneous Schur Decomposition [11]. In addition to restoring the correct pairing, this step improves the estimation accuracy of the spatial frequencies due to an improved estimate of the common eigenspace of the matrices $\hat{\boldsymbol{\Psi}}^{(r)}$. However, it follows from (31) that the improvement in this eigenspace can be asymptotically neglected for high effective SNRs since the perturbation of the eigenvalues dominates. It is, therefore, not necessary to consider the joint approximate eigendecomposition in a first-order perturbation expansion of R-D ESPRIT-type algorithms.

C. Mean Square Errors

As discussed above, the advantage of the first order perturbation expansion is that it is explicit in the perturbation term $\boldsymbol{N}$ and hence makes no assumptions about its distribution. However, it is often also desirable to know the mean square error (MSE) given a specific distribution, i.e., the ensemble average of the squared estimation error over all noise realizations.

In the sequel, we show that the MSE only depends on the second-order moments of the noise samples. Hence, we can derive the MSE as a function of the covariance matrix and the pseudo-covariance matrix only assuming the noise to be zero mean. We neither need to assume Gaussianity nor circular symmetry. For notational simplicity, we consider the special case $R = 2$ for the tensor-based subspace estimate. The generalization to a larger number of dimensions is provided in Appendix D.

Theorem 3: Assume that the entries of the perturbation term $\boldsymbol{N}$ or $\boldsymbol{\mathcal{N}}$ are zero mean random variables with finite second-order moments described by the covariance matrix $\boldsymbol{R}_{\rm nn} = \mathbb{E}\{\boldsymbol{n} \cdot \boldsymbol{n}^{\rm H}\} \in \mathbb{C}^{MN \times MN}$ and the complementary covariance matrix $\boldsymbol{C}_{\rm nn} = \mathbb{E}\{\boldsymbol{n} \cdot \boldsymbol{n}^{\rm T}\} \in \mathbb{C}^{MN \times MN}$ for $\boldsymbol{n} = {\rm vec}\{\boldsymbol{N}\} = {\rm vec}\{[\boldsymbol{\mathcal{N}}]^{\rm T}_{(R+1)}\}$. Then, the first-order approximation of the mean square estimation error for the k-th spatial frequency in the r-th mode is given by

$$\mathbb{E}\left\{\left(\Delta\mu_k^{(r)}\right)^2\right\} = \frac{1}{2}\left(\boldsymbol{r}_k^{(r)^{\rm H}} \cdot \boldsymbol{W}_{\rm mat}^* \cdot \boldsymbol{R}_{\rm nn}^{\rm T} \cdot \boldsymbol{W}_{\rm mat}^{\rm T} \cdot \boldsymbol{r}_k^{(r)} - {\rm Re}\left\{\boldsymbol{r}_k^{(r)^{\rm T}} \cdot \boldsymbol{W}_{\rm mat} \cdot \boldsymbol{C}_{\rm nn} \cdot \boldsymbol{W}_{\rm mat}^{\rm T} \cdot \boldsymbol{r}_k^{(r)}\right\}\right) + \mathcal{O}\{{\rm Tr}(\boldsymbol{R}_{\rm nn})^2\} \quad (34)$$

for R-D Standard ESPRIT and

$$\mathbb{E}\left\{\left(\Delta\mu_k^{(r)}\right)^2\right\} = \frac{1}{2}\left(\boldsymbol{r}_k^{(r)^{\rm H}} \cdot \boldsymbol{W}_{\rm ten}^* \cdot \boldsymbol{R}_{\rm nn}^{\rm T} \cdot \boldsymbol{W}_{\rm ten}^{\rm T} \cdot \boldsymbol{r}_k^{(r)} - {\rm Re}\left\{\boldsymbol{r}_k^{(r)^{\rm T}} \cdot \boldsymbol{W}_{\rm ten} \cdot \boldsymbol{C}_{\rm nn} \cdot \boldsymbol{W}_{\rm ten}^{\rm T} \cdot \boldsymbol{r}_k^{(r)}\right\}\right) + \mathcal{O}\left\{{\rm Tr}(\boldsymbol{R}_{\rm nn})^2\right\} \quad (35)$$

for R-D Standard Tensor-ESPRIT The vector $\boldsymbol{r}_k^{(r)}$ and the matrices $\boldsymbol{W}_{\rm mat}$ and $\boldsymbol{W}_{\rm ten}$ are given by

$$\begin{aligned}
\boldsymbol{r}_k^{(r)} &= \boldsymbol{q}_k \otimes \left(\left[\left(\tilde{\boldsymbol{J}}_1^{(r)}\boldsymbol{U}_{\rm s}\right)^+\left(\tilde{\boldsymbol{J}}_2^{(r)}/{\rm e}^{\jmath\cdot\mu_k^{(r)}} - \tilde{\boldsymbol{J}}_1^{(r)}\right)\right]^{\rm T} \cdot \boldsymbol{p}_k\right)\\
\boldsymbol{W}_{\rm mat} &= \left(\boldsymbol{\Sigma}_{\rm s}^{-1} \cdot \boldsymbol{V}_{\rm s}^{\rm T}\right) \otimes \left(\boldsymbol{U}_{\rm n} \cdot \boldsymbol{U}_{\rm n}^{\rm H}\right) \quad (36)\\
\boldsymbol{W}_{\rm ten} &= \left(\boldsymbol{\Sigma}_3^{[{\rm s}]^{-1}}\boldsymbol{U}_3^{[{\rm s}]^{\rm H}}\right) \otimes \left([\boldsymbol{T}_1 \otimes \boldsymbol{T}_2]\boldsymbol{V}_3^{[{\rm n}]^*}\boldsymbol{V}_3^{[{\rm n}]^{\rm T}}\right)\\
&\quad + \left(\boldsymbol{U}_{\rm s}^{\rm T} \otimes \boldsymbol{I}_M\right)\bar{\boldsymbol{T}}_2\left(\boldsymbol{U}_1^{[{\rm s}]^*}\boldsymbol{\Sigma}_1^{[{\rm s}]^{-1}}\boldsymbol{V}_1^{[{\rm s}]^{\rm T}} \otimes \boldsymbol{U}_1^{[{\rm n}]}\boldsymbol{U}_1^{[{\rm n}]^{\rm H}}\right)\\
&\quad \cdot \boldsymbol{K}_{M_2 \times (M_1 \cdot N)}\\
&\quad + \left(\boldsymbol{U}_{\rm s}^{\rm T} \otimes \boldsymbol{I}_M\right)\bar{\boldsymbol{T}}_1\left(\boldsymbol{U}_2^{[{\rm s}]^*}\boldsymbol{\Sigma}_2^{[{\rm s}]^{-1}}\boldsymbol{V}_2^{[{\rm s}]^{\rm T}} \otimes \boldsymbol{U}_2^{[{\rm n}]}\boldsymbol{U}_2^{[{\rm n}]^{\rm H}}\right)
\end{aligned}$$

where

$$\bar{\boldsymbol{T}}_1 = \begin{bmatrix} \boldsymbol{I}_{M_2} \otimes \boldsymbol{t}_{1,1} \\ \vdots \\ \boldsymbol{I}_{M_2} \otimes \boldsymbol{t}_{1,M_1} \end{bmatrix} \otimes \boldsymbol{I}_{M_2}, \; \bar{\boldsymbol{T}}_2 = \boldsymbol{I}_{M_1} \otimes \begin{bmatrix} \boldsymbol{I}_{M_1} \otimes \boldsymbol{t}_{2,1} \\ \vdots \\ \boldsymbol{I}_{M_1} \otimes \boldsymbol{t}_{2,M_2} \end{bmatrix}, \quad (37)$$

$\boldsymbol{t}_{r,m}$ is the m-th column of $\boldsymbol{T}_r$, and $\boldsymbol{T}_r = \boldsymbol{U}_r^{[{\rm s}]} \cdot \boldsymbol{U}_r^{[{\rm s}]^{\rm H}} \in \mathbb{C}^{M_r \times M_r}$ is the projection matrix onto the column space of $[\boldsymbol{\mathcal{X}}_0]_{(r)}$ for $r = 1, 2$. Finally, $\boldsymbol{K}_{p \times q}$ is the commutation matrix from (7). Note that $\boldsymbol{W}_{\rm ten}$ is shown for $R = 2$ only for notational simplicity. The general expression for $R \geq 2$ is given in Appendix D. Due to the space limitation, the derivation for $\boldsymbol{W}_{\rm ten}$ is only shown for $R = 2$.

Proof: cf. Appendix C.

Note that for 1-D Standard ESPRIT, this MSE expression agrees with the one shown in [20] for the special case of circular symmetry. However, [20] does not directly generalize to the tensor case. In the MSE expressions (34) and (35) we only replace $\boldsymbol{W}_{\rm mat}$ by $\boldsymbol{W}_{\rm ten}$ to account for the enhanced signal subspace estimate. Furthermore, note that the special case of circularly symmetric white noise corresponds to $\boldsymbol{R}_{\rm nn} = \sigma_{\rm n}^2 \cdot \boldsymbol{I}_{MN}$ and $\boldsymbol{C}_{\rm nn} = \boldsymbol{0}_{MN \times MN}$.

D. Incorporation of Forward-Backward-Averaging

So far we have shown the explicit expansion of the estimation error of the spatial frequencies as well as the MSE expressions for R-D Standard ESPRIT and R-D Standard Tensor-ESPRIT. In order to extend these results to Unitary-ESPRIT-type algorithms we need to incorporate the mandatory preprocessing for Unitary ESPRIT which is given by forward-backward-averaging. The second step in Unitary ESPRIT is the transformation to the real-valued domain. However, it can be shown that this step has no impact on the performance for high SNRs.[5] Therefore, the asymptotic performance of Unitary-ESPRIT-type algorithms is found once forward-backward-averaging is taken into account. Forward-backward-averaging augments the N observations of the sampled R-D harmonics by N new "virtual" observations which are a conjugated and row-flipped as well as column-flipped version of the original ones [10]. This can be expressed in matrix form as

$$\boldsymbol{X}^{({\rm fba})} = [\boldsymbol{X} \quad \boldsymbol{\Pi}_M \cdot \boldsymbol{X}^* \cdot \boldsymbol{\Pi}_N] \in \mathbb{C}^{M \times 2N}. \quad (38)$$

Inserting $\boldsymbol{X} = \boldsymbol{X}_0 + \boldsymbol{N}$ we find

$$\begin{aligned}\boldsymbol{X}^{({\rm fba})} &= [\boldsymbol{X} \quad \boldsymbol{\Pi}_M \cdot \boldsymbol{X}_0^* \cdot \boldsymbol{\Pi}_N] + [\boldsymbol{N} \quad \boldsymbol{\Pi}_M \cdot \boldsymbol{N}^* \cdot \boldsymbol{\Pi}_N]\\ &= \boldsymbol{X}_0^{({\rm fba})} + \boldsymbol{N}^{({\rm fba})}. \end{aligned} \quad (39)$$

However, the latter relation shows that we need an expansion for the perturbation of the subspace of a matrix $\boldsymbol{X}_0^{({\rm fba})}$ superimposed by an additive perturbation $\boldsymbol{N}^{({\rm fba})}$, which is small. Since the explicit perturbation expansion we have used up to this point requires no additional assumptions, the surprisingly simple answer is that we do not need to change anything but we can apply the previous results directly. All we need is to replace all exact (noise-free) subspaces of $\boldsymbol{X}_0$ by the corresponding subspaces of $\boldsymbol{X}_0^{({\rm fba})}$. From (32), we immediately obtain the following explicit first-order expansion which is valid for R-D Unitary ESPRIT

$$\Delta\mu_k^{(r)} = {\rm Im}\left\{\boldsymbol{p}_k^{({\rm fba})^{\rm T}} \cdot \left(\tilde{\boldsymbol{J}}_1^{(r)} \cdot \boldsymbol{U}_{\rm s}^{({\rm fba})}\right)^+ \cdot \left[\tilde{\boldsymbol{J}}_2^{(r)}/\lambda_k^{(r)} - \tilde{\boldsymbol{J}}_1^{(r)}\right] \cdot \Delta\boldsymbol{U}_{\rm s}^{({\rm fba})} \cdot \boldsymbol{q}_k^{({\rm fba})}\right\} + \mathcal{O}\{\Delta^2\} \quad (40)$$

where $\Delta\boldsymbol{X}_{\rm s}^{({\rm fba})}$ is given by

$$\Delta\boldsymbol{U}_{\rm s}^{({\rm fba})} = \boldsymbol{U}_{\rm n}^{({\rm fba})} \cdot \boldsymbol{U}_{\rm n}^{({\rm fba})^{\rm H}} \cdot \boldsymbol{N}^{({\rm fba})} \cdot \boldsymbol{V}_{\rm s}^{({\rm fba})} \cdot \boldsymbol{\Sigma}_{\rm s}^{({\rm fba})^{-1}} \quad (41)$$

and $\boldsymbol{X}_{\rm s}^{({\rm fba})}, \boldsymbol{X}_{\rm n}^{({\rm fba})}, \boldsymbol{X}_{\rm s}^{({\rm fba})}, \boldsymbol{\Sigma}_{\rm s}^{({\rm fba})}$ correspond to the signal subspace, the noise subspace, the row space, and the singular values of $\boldsymbol{X}_0^{({\rm fba})}$, respectively. Likewise, $\boldsymbol{q}_k^{({\rm fba})}$ and $\boldsymbol{p}_k^{({\rm fba})}$ represent the corresponding versions of $\boldsymbol{q}_k$ and $\boldsymbol{p}_k$ if $\boldsymbol{U}_{\rm s}$ is replaced by $\boldsymbol{U}_{\rm s}^{({\rm fba})}$ in the shift invariance equations.

[5]The proof is omitted here due to the page limitation, however, it can be found in [32], Appendix D.13.

With the same reasoning, an explicit expansion for R-D Unitary Tensor-ESPRIT is obtained by consistently replacing $\boldsymbol{\mathcal{X}}_0$ by $\boldsymbol{\mathcal{X}}_0^{(\mathrm{fba})}$ in (33), i.e.,

$$\Delta\mu_k^{(r)} = \mathrm{Im}\left\{ \boldsymbol{p}_k^{(\mathrm{fba})^{\mathrm{T}}} \cdot \left(\tilde{\boldsymbol{J}}_1^{(r)} \cdot \boldsymbol{U}_{\mathrm{s}}^{(\mathrm{fba})} \right)^+ \cdot \left[\tilde{\boldsymbol{J}}_2^{(r)} / \lambda_k^{(r)} - \tilde{\boldsymbol{J}}_1^{(r)} \right] \cdot \left[\Delta \hat{\boldsymbol{\mathcal{U}}}^{[\mathrm{s}](\mathrm{fba})} \right]_{(R+1)}^{\mathrm{T}} \cdot \boldsymbol{q}_k^{(\mathrm{fba})} \right\} + \mathcal{O}\{\Delta^2\}. \quad (42)$$

Similarly, Theorem 3 can be applied to compute the MSE since we only assumed the noise to be zero mean and possess finite second order moments, which is still true after forward-backward-averaging. The following theorem summarizes the results for R-D Unitary ESPRIT and R-D Unitary Tensor-ESPRIT:

Theorem 4: For the case where $\boldsymbol{N}$ or $\boldsymbol{\mathcal{N}}$ contain zero mean random variables with finite second-order moments described by the covariance matrix $\boldsymbol{R}_{\mathrm{nn}} = \mathbb{E}\{\boldsymbol{n} \cdot \boldsymbol{n}^{\mathrm{H}}\}$ and the complementary covariance matrix $\boldsymbol{C}_{\mathrm{nn}} = \mathbb{E}\{\boldsymbol{n} \cdot \boldsymbol{n}^{\mathrm{T}}\}$ for $\boldsymbol{n} = \mathrm{vec}\{\boldsymbol{N}\} = \mathrm{vec}\{[\boldsymbol{\mathcal{N}}]_{(R+1)}^{\mathrm{T}}\}$, the MSE for R-D Unitary ESPRIT and R-D Unitary Tensor-ESPRIT are given by (34) and (35) if we replace $\boldsymbol{r}_k^{(r)}$ by $\boldsymbol{r}_k^{(r)^{(\mathrm{fba})}}$, $\boldsymbol{W}_{\mathrm{mat}}$ by $\boldsymbol{W}_{\mathrm{mat}}^{(\mathrm{fba})}$, $\boldsymbol{W}_{\mathrm{ten}}$ by $\boldsymbol{W}_{\mathrm{ten}}^{(\mathrm{fba})}$, and $\boldsymbol{R}_{\mathrm{nn}}$ as well as $\boldsymbol{C}_{\mathrm{nn}}$ by $\boldsymbol{R}_{\mathrm{nn}}^{(\mathrm{fba})}$ and $\boldsymbol{C}_{\mathrm{nn}}^{(\mathrm{fba})}$. Here, $\boldsymbol{r}_k^{(r)^{(\mathrm{fba})}}$, $\boldsymbol{W}_{\mathrm{mat}}^{(\mathrm{fba})}$, and $\boldsymbol{W}_{\mathrm{ten}}^{(\mathrm{fba})}$ are computed as in (34) and (35) by consistently replacing all quantities by their forward-backward-averaged equivalents. Moreover, $\boldsymbol{R}_{\mathrm{nn}}^{(\mathrm{fba})}$ and $\boldsymbol{C}_{\mathrm{nn}}^{(\mathrm{fba})}$ represent the covariance and the pseudo-covariance matrix of the forward-backward-averaged noise, which are given by

$$\boldsymbol{R}_{\mathrm{nn}}^{(\mathrm{fba})} = \begin{bmatrix} \boldsymbol{R}_{\mathrm{nn}} & \boldsymbol{C}_{\mathrm{nn}} \cdot \boldsymbol{\Pi}_{MN} \\ \boldsymbol{\Pi}_{MN} \cdot \boldsymbol{C}_{\mathrm{nn}}^* & \boldsymbol{\Pi}_{MN} \cdot \boldsymbol{R}_{\mathrm{nn}}^* \cdot \boldsymbol{\Pi}_{MN} \end{bmatrix},$$
$$\boldsymbol{C}_{\mathrm{nn}}^{(\mathrm{fba})} = \begin{bmatrix} \boldsymbol{C}_{\mathrm{nn}} & \boldsymbol{R}_{\mathrm{nn}} \cdot \boldsymbol{\Pi}_{MN} \\ \boldsymbol{\Pi}_{MN} \cdot \boldsymbol{R}_{\mathrm{nn}}^* & \boldsymbol{\Pi}_{MN} \cdot \boldsymbol{C}_{\mathrm{nn}}^* \cdot \boldsymbol{\Pi}_{MN} \end{bmatrix}.$$

Proof: cf. Appendix E.

Note that in the special case where the noise is circularly symmetric and white we have $\boldsymbol{R}_{\mathrm{nn}} = \sigma_{\mathrm{n}}^2 \cdot \boldsymbol{I}_{MN}$ and $\boldsymbol{C}_{\mathrm{nn}} = \boldsymbol{0}_{MN \times MN}$ and therefore $\boldsymbol{R}_{\mathrm{nn}}^{(\mathrm{fba})} = \sigma_{\mathrm{n}}^2 \cdot \boldsymbol{I}_{2MN}$ and $\boldsymbol{C}_{\mathrm{nn}}^{(\mathrm{fba})} = \sigma_{\mathrm{n}}^2 \cdot \boldsymbol{\Pi}_{2MN}$.

It is important to note that results on Unitary ESPRIT in this section relate to the LS solution only. If TLS is used instead, the equivalence of Standard ESPRIT with forward-backward-averaging and Unitary ESPRIT is shown in [10]. Note that the asymptotic equivalence of ESPRIT based on LS and TLS has been shown in [14].

V. Simulation Results

In this section we show numerical results to demonstrate the asymptotic behavior of the analytical performance assessment for ESPRIT-type algorithms that is derived in this paper. To this end, we compare the analytical ("ana") MSE expressions provided in Theorem 3, and Theorem 4, respectively, to the empirical ("emp") estimation error obtained by Monte-Carlo trials. The comparison is carried out for Standard ESPRIT (SE), Unitary ESPRIT (UE), Standard Tensor-ESPRIT (STE), and Unitary Tensor-ESPRIT (UTE). For all ESPRIT-type algorithms, LS is used to solve the shift invariance equations. The MSE is defined as

$$\mathrm{MSE} = \frac{1}{R \cdot d} \cdot \mathbb{E}\left\{ \sum_{r=1}^{R} \sum_{i=1}^{d} \left(\mu_i^{(r)} - \hat{\mu}_i^{(r)} \right)^2 \right\}, \quad (43)$$

where $\hat{\mu}_i^{(r)}$ is the estimate for the spatial frequency of the i-th source in the r-th mode.

For comparison, the deterministic Cramér-Rao Bound (CRB) [38] is also shown. Note that this bound also applies to the tensor-based estimation algorithms, since it only depends on the assumptions regarding the nature of the desired parameters, the nuisance parameters (stochastic according to a specific distribution or deterministic), and the measurement model. Since we do not change the parameters nor the model in the tensor case (we just rewrite the *same* model differently), the CRB remains unaffected.

For all the simulations we assume a known number of planar wavefronts impinging on a 2-D $M_1 \times M_2$ uniform rectangular array (URA) with uniform spacing in all dimensions and isotropic sensor elements. The sources emit narrow-band waveforms $s_i(t)$ modeled as complex Gaussian distributed symbols $s_i(t)$ and we observe N subsequent snapshots $t = t_1, t_2, \ldots, t_N$. All sources are assumed to have unit power, i.e., $\mathbb{E}\{|s_i(t)|^2\} = 1$. If source correlation is investigated we generate the symbols $s_i(t)$ such that $\mathbb{E}\{s_i^*(t) \cdot s_j(t)\} = \rho \cdot \mathrm{e}^{\jmath \varphi_{i,j}}$ for $i \neq j = 1, 2, \ldots, d$, where ρ is the magnitude of the correlation coefficient between each pair of sources and $\varphi_{i,j}$ is a uniformly distributed correlation phase. The additive noise is generated according to a circularly symmetric complex Gaussian distribution with zero mean and variance σ_{n}^2. The noise samples are assumed to be mutually independent. Therefore, the signal to noise ratio (SNR) is defined as $1/\sigma_{\mathrm{n}}^2$.

For Fig. 1 we employ a 5×6 URA and collect $N = 20$ snapshots from two sources located at $\mu_1^{(1)} = 1$, $\mu_2^{(1)} = -0.5$, $\mu_1^{(2)} = -0.5$, and $\mu_2^{(2)} = 1$. The sources are highly correlated with a correlation of $\rho = 0.9999$. On the other hand, for Fig. 2 we increase the number of sources to $d = 3$ and set the correlation coefficient to $\rho = 0.97$. Moreover, the spatial frequencies of the sources are given by $\mu_1^{(1)} = 0.7$, $\mu_2^{(1)} = 0.9$, $\mu_3^{(1)} = 1.1$, $\mu_1^{(2)} = -0.1$, $\mu_2^{(2)} = -0.3$, $\mu_3^{(2)} = -0.5$ and we use an 8×8 URA. In Fig. 3 we vary the number of snapshots N for a fixed SNR of 30 dB. We consider $d = 3$ sources with a mutual correlation of $\rho = 0.98$ impinging on an 8×8 URA. The spatial frequencies are given by $\mu_1^{(1)} = 1.0$, $\mu_2^{(1)} = 1.4$, $\mu_3^{(1)} = 1.8$, $\mu_1^{(2)} = -0.5$, $\mu_2^{(2)} = -0.1$, $\mu_3^{(2)} = 0.3$. These simulation results verify that the empirical estimation errors agree with the analytical ones for high effective SNRs, i.e., when either the noise variance is small or the number of samples is large. This is also expected as the performance analysis framework presented here is asymptotically accurate for high effective SNRs.

The effect of closely spaced sources is investigated in Fig. 4 for a fixed SNR of 40 dB by choosing the spatial frequencies of $d = 3$ sources according to $\mu_1^{(1)} = \mu_1^{(2)} = -0.5$, $\mu_2^{(1)} = \mu_2^{(2)} = -0.5 + \Delta\mu$, and $\mu_3^{(1)} = \mu_3^{(2)} = -0.5 + 2 \cdot \Delta\mu$ and varying $\Delta\mu$. We consider an 8×8 URA, $N = 20$ snapshots, and a source correlation of $\rho = 0.99$. The analytical results agree well

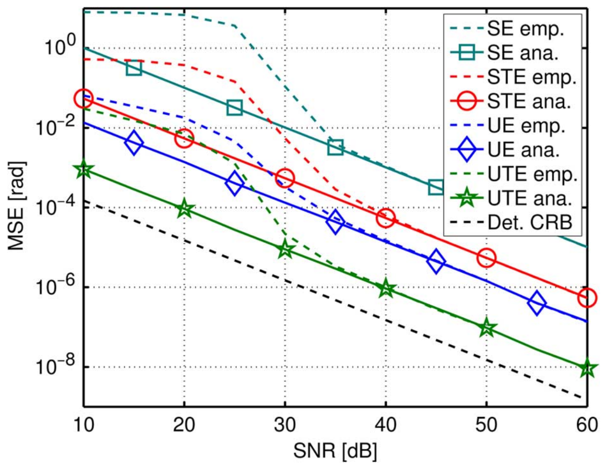

Fig. 1. Performance of 2-D SE, STE, UE, UTE for $d = 2$ highly correlated sources ($\rho = 0.9999$) located at $\mu_1^{(1)} = 1, \mu_2^{(1)} = -0.5, \mu_1^{(2)} = -0.5$, and $\mu_2^{(2)} = 1$, a 5×6 URA, and $N = 20$ snapshots.

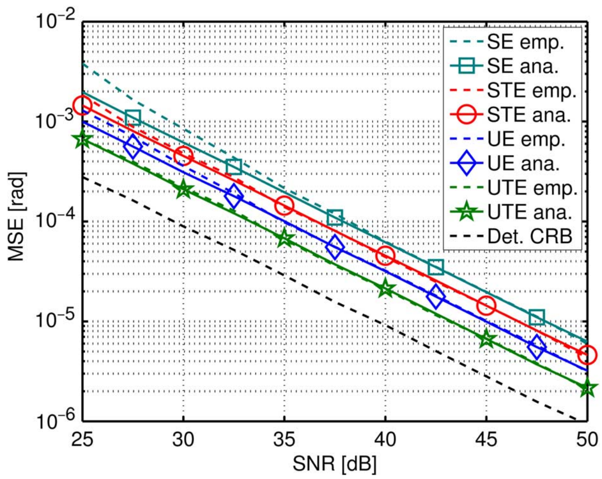

Fig. 2. Performance of 2-D SE, STE, UE, UTE for $d = 3$ correlated sources ($\rho = 0.97$) positioned at $\mu_1^{(1)} = 0.7, \mu_2^{(1)} = 0.9, \mu_3^{(1)} = 1.1, \mu_1^{(2)} = -0.1$, $\mu_2^{(2)} = -0.3, \mu_3^{(2)} = -0.5$, an 8×8 URA, and $N = 20$ snapshots.

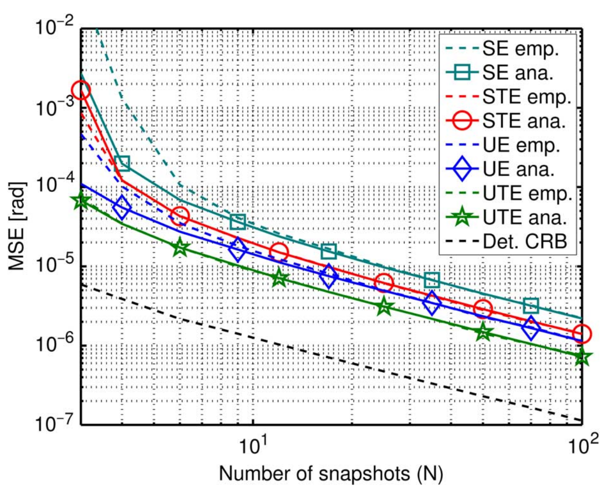

Fig. 3. Performance of 2-D SE, STE, UE, UTE for $d = 3$ correlated sources ($\rho = 0.98$) positioned at $\mu_1^{(1)} = 1.0, \mu_2^{(1)} = 1.4, \mu_3^{(1)} = 1.8, \mu_1^{(2)} = -0.5$, $\mu_2^{(2)} = -0.1, \mu_3^{(2)} = 0.3$, an 8×8 URA, and an SNR of 30 dB as a function of the number of snapshots N.

with the empirical estimation errors except for very small source separations. In these cases, the threshold SNR where the first-

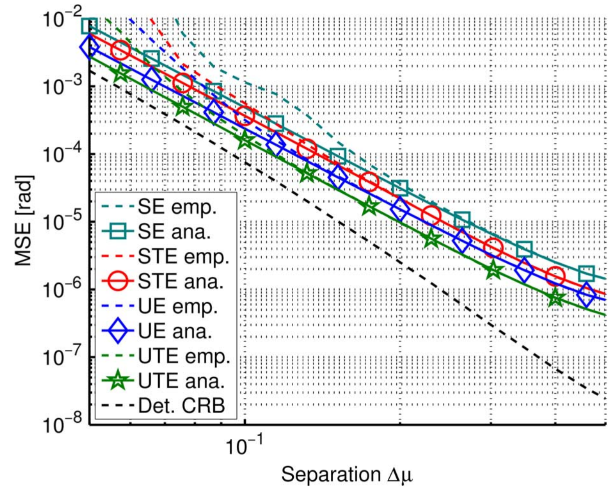

Fig. 4. Performance of 2-D SE, STE, UE, UTE for $d = 3$ correlated sources ($\rho = 0.99$) positioned at $\mu_1^{(1)} = \mu_1^{(2)} = -0.5, \mu_1^{(1)} = \mu_1^{(2)} = -0.5 + \Delta\mu$, and $\mu_3^{(1)} = \mu_3^{(2)} = -0.5 + 2 \cdot \Delta\mu$, an 8×8 URA, $N = 20$ snapshots, and an SNR of 40 dB as a function of the source separation $\Delta\mu$.

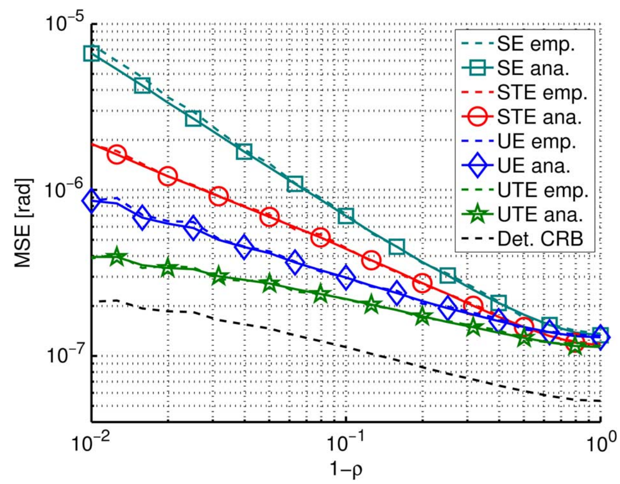

Fig. 5. Performance of 2-D SE, STE, UE, UTE for $d = 2$ positioned at $\mu_1^{(1)} = 1, \mu_1^{(2)} = 0.85, \mu_2^{(1)} = 0.85, \mu_2^{(2)} = 1$, an 8×8 URA, $N = 20$ snapshots, and an SNR of 40 dB as a function of $1 - \rho$ where ρ is the source correlation.

order perturbation expansion becomes accurate is even higher. Therefore, if the SNR is chosen larger, the empirical and the analytical results converge as well.

Finally, the impact of the source correlation is analyzed in Fig. 5 for the same fixed SNR of 40 dB. The spatial frequencies of $d = 2$ sources are given by $\mu_1^{(1)} = 1$, $\mu_1^{(2)} = 0.85$, $\mu_2^{(1)} = 0.85$, $\mu_2^{(2)} = 1$. As before, we employ an 8×8 URA and $N = 20$ snapshots. We depict the MSE as a function of $1 - \rho$ where ρ is the source correlation. Again, we observe a very good match between the analytical MSE expressions and the empirical results.

VI. Conclusions

In this paper we have discussed a framework for the analytical performance assessment of multi-dimensional subspace-based parameter estimation schemes. It is based on earlier results on an explicit first-order expansion of the SVD and its application to 1-D versions of subspace-based parameter estimation schemes, e.g., ESPRIT. We have extended this framework in a number of ways. Firstly, we have derived an explicit first-order expansion of the HOSVD-based subspace estimate which is the basis for Tensor-ESPRIT-type algorithms. Secondly, we have shown that the first-order expansion for 1-D Standard ESPRIT can be extended to other ESPRIT-type algorithms, e.g., R-D Standard ESPRIT, R-D Unitary ESPRIT, R-D Standard Tensor-ESPRIT, or R-D Unitary Tensor-ESPRIT.

All these expansions have in common that they are explicit, i.e., no assumption about the statistics of either desired signal or additive perturbation need to be made. We only require the perturbation to be small compared to the desired signal. We also do not need the number of snapshots to be large, i.e., they even apply to the single snapshot case ($N = 1$). We have also shown that the mean square error can readily be computed in closed form and that it depends only on the second-order moments of the noise. Consequently, for the MSE expressions we only need the noise to be zero mean and its second order moments to be finite. Neither Gaussianity nor circular symmetry is required. This is a particularly attractive feature of our approach with respect to different types of preprocessing which alters the noise statistics, e.g., spatial smoothing (which yields spatially correlated noise) or forward-backward-averaging (which annihilates the circular symmetry of the noise). Since we do not require spatial whiteness or circular symmetry, our MSE expressions are directly applicable. The resulting MSE expressions are asymptotic in the effective SNR, i.e., they become accurate as either the noise variance goes to zero or the number of observations goes to infinity.

APPENDIX A
PROOF OF THEOREM 1

As shown in (17), the estimated signal subspace tensor can be computed via

$$\hat{\boldsymbol{\mathcal{U}}}^{[\mathrm{s}]} = \hat{\boldsymbol{\mathcal{S}}}^{[\mathrm{s}]} \times_1 \hat{\boldsymbol{U}}_1^{[\mathrm{s}]} \ldots \times_R \hat{\boldsymbol{U}}_R^{[\mathrm{s}]} \times_{R+1} \hat{\boldsymbol{\Sigma}}_{\mathrm{s}}^{-1}. \tag{44}$$

Here, $\hat{\boldsymbol{\mathcal{S}}}^{[\mathrm{s}]}$ represents the truncated version core tensor $\hat{\boldsymbol{\mathcal{S}}}$ from the HOSVD of $\boldsymbol{X}$. In order to eliminate $\hat{\boldsymbol{\mathcal{S}}}^{[\mathrm{s}]}$ in (44), we require the following Lemma:

Lemma 1: The truncated core tensor $\hat{\boldsymbol{\mathcal{S}}}^{[\mathrm{s}]}$ can be computed from $\boldsymbol{X}$ directly via

$$\hat{\boldsymbol{\mathcal{S}}}^{[\mathrm{s}]} = \boldsymbol{\mathcal{X}} \times_1 \hat{\boldsymbol{U}}_1^{[\mathrm{s}]^{\mathrm{H}}} \ldots \times_{R+1} \hat{\boldsymbol{U}}_{R+1}^{[\mathrm{s}]^{\mathrm{H}}}. \tag{45}$$

Proof: To show (45) we insert the HOSVD of $\boldsymbol{\mathcal{X}}$ given by $\boldsymbol{\mathcal{X}} = \hat{\boldsymbol{\mathcal{S}}} \times_1 \hat{\boldsymbol{U}}_1 \ldots \times_{R+1} \hat{\boldsymbol{U}}_{R+1}$. Using (4), we obtain

$$\hat{\boldsymbol{\mathcal{S}}}^{[\mathrm{s}]} = \hat{\boldsymbol{\mathcal{S}}} \times_1 \left(\hat{\boldsymbol{U}}_1^{[\mathrm{s}]^{\mathrm{H}}} \cdot \hat{\boldsymbol{U}}_1\right) \ldots \times_{R+1} \left(\hat{\boldsymbol{U}}_{R+1}^{[\mathrm{s}]^{\mathrm{H}}} \cdot \hat{\boldsymbol{U}}_{R+1}\right). \tag{46}$$

However, since the matrices of r-mode singular vectors $\hat{\boldsymbol{U}}_r$ are unitary, they satisfy $\hat{\boldsymbol{U}}_r^{[\mathrm{s}]^{\mathrm{H}}} \cdot \hat{\boldsymbol{U}}_r = [\boldsymbol{I}_{p_r},\ \boldsymbol{0}_{p_r \times (M_r - p_r)}]$. Therefore, $\hat{\boldsymbol{\mathcal{S}}}^{[\mathrm{s}]}$ computed via (45) contains the first p_r elements of $\hat{\boldsymbol{\mathcal{S}}}$ in the r-th mode, which shows that it is indeed the truncated core tensor. ■

Next, we use Lemma 1 to eliminate $\hat{\boldsymbol{\mathcal{S}}}$ in (44). We obtain

$$\begin{aligned}\hat{\boldsymbol{\mathcal{U}}}^{[\mathrm{s}]} &= \boldsymbol{\mathcal{X}} \times_1 \left(\hat{\boldsymbol{U}}_1^{[\mathrm{s}]} \cdot \hat{\boldsymbol{U}}_1^{[\mathrm{s}]^{\mathrm{H}}}\right) \ldots \times_R \left(\hat{\boldsymbol{U}}_R^{[\mathrm{s}]} \cdot \hat{\boldsymbol{U}}_R^{[\mathrm{s}]^{\mathrm{H}}}\right) \quad (47)\\ &\quad \times_{R+1} \left(\hat{\boldsymbol{\Sigma}}_{\mathrm{s}}^{-1} \cdot \hat{\boldsymbol{U}}_{R+1}^{[\mathrm{s}]^{\mathrm{H}}}\right)\\ &= \boldsymbol{\mathcal{X}} \times_1 \hat{\boldsymbol{T}}_1 \ldots \times_R \hat{\boldsymbol{T}}_R \times_{R+1} \left(\hat{\boldsymbol{\Sigma}}_{\mathrm{s}}^{-1} \cdot \hat{\boldsymbol{U}}_{R+1}^{[\mathrm{s}]^{\mathrm{H}}}\right), \quad (48)\end{aligned}$$

where we have introduced the short hand notation $\hat{\boldsymbol{T}}_r = \hat{\boldsymbol{U}}_r^{[\mathrm{s}]} \cdot \hat{\boldsymbol{U}}_r^{[\mathrm{s}]^{\mathrm{H}}}$. The next step is to compute the matrix $[\hat{\boldsymbol{\mathcal{U}}}^{[\mathrm{s}]}]_{(R+1)}^{\mathrm{T}}$. Inserting (48) and using (5), we obtain

$$\left[\hat{\boldsymbol{\mathcal{U}}}^{[\mathrm{s}]}\right]_{(R+1)}^{\mathrm{T}} = (\hat{\boldsymbol{T}}_1 \otimes \ldots \otimes \hat{\boldsymbol{T}}_R) \cdot [\boldsymbol{\mathcal{X}}]_{(R+1)}^{\mathrm{T}} \cdot \hat{\boldsymbol{U}}_{R+1}^{[\mathrm{s}]^*} \cdot \hat{\boldsymbol{\Sigma}}_{\mathrm{s}}^{-1}. \tag{49}$$

As pointed out in Section II-B, the link between the measurement matrix $\boldsymbol{X}$ and the measurement tensor $\boldsymbol{\mathcal{X}}$ is given by $\boldsymbol{X} = [\boldsymbol{\mathcal{X}}]_{(R+1)}^{\mathrm{T}}$. Therefore, their SVDs (cf. (19) and (22)) are linked through the following identities

$$\hat{\boldsymbol{U}}_{\mathrm{s}} = \hat{\boldsymbol{V}}_{R+1}^{[\mathrm{s}]^*},\ \hat{\boldsymbol{U}}_{\mathrm{n}} = \hat{\boldsymbol{V}}_{R+1}^{[\mathrm{n}]^*},\ \hat{\boldsymbol{V}}_{\mathrm{s}} = \hat{\boldsymbol{U}}_{R+1}^{[\mathrm{s}]^*},\ \hat{\boldsymbol{V}}_{\mathrm{n}} = \hat{\boldsymbol{U}}_{R+1}^{[\mathrm{n}]}.$$

Consequently we can write

$$\begin{aligned}[\boldsymbol{\mathcal{X}}]_{(R+1)}^{\mathrm{T}} \cdot \hat{\boldsymbol{U}}_{R+1}^{[\mathrm{s}]} \cdot \hat{\boldsymbol{\Sigma}}_{\mathrm{s}}^{-1} &= \boldsymbol{X} \cdot \hat{\boldsymbol{V}}_{\mathrm{s}} \cdot \hat{\boldsymbol{\Sigma}}_{\mathrm{s}}^{-1} = \hat{\boldsymbol{U}} \cdot \hat{\boldsymbol{\Sigma}} \cdot \hat{\boldsymbol{V}}^{\mathrm{H}} \cdot \hat{\boldsymbol{V}}_{\mathrm{s}} \cdot \hat{\boldsymbol{\Sigma}}_{\mathrm{s}}^{-1}\\ &= \hat{\boldsymbol{U}}_{\mathrm{s}} \cdot \hat{\boldsymbol{\Sigma}}_{\mathrm{s}} \cdot \hat{\boldsymbol{\Sigma}}_{\mathrm{s}}^{-1} = \hat{\boldsymbol{U}}_{\mathrm{s}}. \quad (50)\end{aligned}$$

Finally, inserting (50) into (49) yields

$$\left[\hat{\boldsymbol{\mathcal{U}}}^{[\mathrm{s}]}\right]_{(R+1)}^{\mathrm{T}} = (\hat{\boldsymbol{T}}_1 \otimes \ldots \otimes \hat{\boldsymbol{T}}_R) \cdot \hat{\boldsymbol{U}}_{\mathrm{s}}, \tag{51}$$

which is the desired result. □

Corollary 1: A corollary which follows from this theorem is that the exact subspace $\boldsymbol{U}_{\mathrm{s}}$ satisfies the following identity

$$\boldsymbol{U}_{\mathrm{s}} = (\boldsymbol{T}_1 \otimes \ldots \otimes \boldsymbol{T}_R) \cdot \boldsymbol{U}_{\mathrm{s}}. \tag{52}$$

Proof: The corollary follows by considering the special case where $\boldsymbol{\mathcal{X}} = \boldsymbol{\mathcal{X}}_0$ and hence $\hat{\boldsymbol{T}}_r = \boldsymbol{T}_r$ as well as $\hat{\boldsymbol{U}}_{\mathrm{s}} = \boldsymbol{U}_{\mathrm{s}}$. For this case we also have $[\hat{\boldsymbol{\mathcal{U}}}^{[\mathrm{s}]}]_{(R+1)}^{\mathrm{T}} = [\boldsymbol{\mathcal{U}}^{[\mathrm{s}]}]_{(R+1)}^{\mathrm{T}} = \boldsymbol{U}_{\mathrm{s}}$, where the last identity comes from the fact that in the noise-free case, the HOSVD-based subspace estimate coincides with the SVD-bases subspace estimate. ■

APPENDIX B
PROOF OF THEOREM 2

We start by inserting $\hat{\boldsymbol{U}}_{\mathrm{s}} = \boldsymbol{U}_{\mathrm{s}} + \Delta\boldsymbol{U}_{\mathrm{s}}$ and $\hat{\boldsymbol{T}}_r = \boldsymbol{T}_r + \Delta\boldsymbol{T}_r$ into (24). Then we obtain

$$\begin{aligned}\left[\hat{\boldsymbol{\mathcal{U}}}^{[\mathrm{s}]}\right]_{(R+1)}^{\mathrm{T}} &= [(\boldsymbol{T}_1 + \Delta\boldsymbol{T}_1) \otimes \ldots \otimes (\boldsymbol{T}_R + \Delta\boldsymbol{T}_R)] \cdot (\boldsymbol{U}_{\mathrm{s}} + \Delta\boldsymbol{U}_{\mathrm{s}})\\ &= \underbrace{[\boldsymbol{T}_1 \otimes \ldots \otimes \boldsymbol{T}_R] \cdot \boldsymbol{U}_{\mathrm{s}}}_{\boldsymbol{U}_{\mathrm{s}}} + [\boldsymbol{T}_1 \otimes \ldots \otimes \boldsymbol{T}_R] \cdot \Delta\boldsymbol{U}_{\mathrm{s}}\\ &\quad + [\Delta\boldsymbol{T}_1 \otimes \boldsymbol{T}_2 \otimes \ldots \otimes \boldsymbol{T}_R] \cdot \boldsymbol{U}_{\mathrm{s}} + \ldots\\ &\quad + [\boldsymbol{T}_1 \otimes \boldsymbol{T}_2 \otimes \ldots \otimes \Delta\boldsymbol{T}_R] \cdot \boldsymbol{U}_{\mathrm{s}} + \mathcal{O}\{\Delta^2\}, \quad (53)\end{aligned}$$

since all terms that contain more than one perturbation term can be absorbed into $\mathcal{O}\{\Delta^2\}$. The first term in (53) represents the exact signal subspace (cf. Corollary 1), hence the remaining terms are the first order expansion of $[\Delta\hat{\boldsymbol{\mathcal{U}}}^{[\mathrm{s}]}]_{(R+1)}^{\mathrm{T}}$. As the first term of this expansion already agrees with Theorem 2, we still need to show that for the remaining terms we have for $r = 1, 2, \ldots, R$

$$[\boldsymbol{T}_1 \otimes \ldots \otimes \Delta\boldsymbol{T}_r \otimes \ldots \boldsymbol{T}_R] \cdot \boldsymbol{U}_{\mathrm{s}} = \left[\boldsymbol{T}_1 \otimes \ldots \otimes \left(\boldsymbol{U}_r^{[\mathrm{n}]} \cdot \boldsymbol{\Gamma}_r^{[\mathrm{n}]} \cdot \boldsymbol{U}_r^{[\mathrm{s}]^{\mathrm{H}}}\right) \otimes \ldots \boldsymbol{T}_R\right] \cdot \boldsymbol{U}_{\mathrm{s}} + \mathcal{O}\{\Delta^2\}. \tag{54}$$

As a first step, we expand the left-hand side of (54) by applying Corollary 1

$$\begin{aligned}[\boldsymbol{T}_1 \otimes \ldots \otimes \Delta\boldsymbol{T}_r \otimes \ldots \boldsymbol{T}_R] \cdot \boldsymbol{U}_{\mathrm{s}} &= [\boldsymbol{T}_1 \otimes \ldots \otimes \Delta\boldsymbol{T}_r \otimes \ldots \boldsymbol{T}_R] \cdot [\boldsymbol{T}_1 \otimes \ldots \otimes \boldsymbol{T}_r \otimes \ldots \boldsymbol{T}_R] \cdot \boldsymbol{U}_{\mathrm{s}} \\ &= [(\boldsymbol{T}_1 \cdot \boldsymbol{T}_1) \otimes \ldots \otimes (\Delta\boldsymbol{T}_r \cdot \boldsymbol{T}_r) \otimes \ldots (\boldsymbol{T}_R \cdot \boldsymbol{T}_R)] \cdot \boldsymbol{U}_{\mathrm{s}} \\ &= [\boldsymbol{T}_1 \otimes \ldots \otimes (\Delta\boldsymbol{T}_r \cdot \boldsymbol{T}_r) \otimes \ldots \boldsymbol{T}_R] \cdot \boldsymbol{U}_{\mathrm{s}},\end{aligned} \tag{55}$$

where we have used the fact that the matrices $\boldsymbol{T}_r$ are projection matrices and hence idempotent, i.e., $\boldsymbol{T}_r \cdot \boldsymbol{T}_r = \boldsymbol{T}_r$. What remains to be shown is that $\Delta\boldsymbol{T}_r \cdot \boldsymbol{T}_r = \boldsymbol{U}_r^{[\mathrm{n}]} \cdot \boldsymbol{\Gamma}_r^{[\mathrm{n}]} \cdot \boldsymbol{U}_r^{[\mathrm{s}]^{\mathrm{H}}} + \mathcal{O}\{\Delta^2\}$. Since $\hat{\boldsymbol{T}}_r = \hat{\boldsymbol{U}}_r^{[\mathrm{s}]} \cdot \hat{\boldsymbol{U}}_r^{[\mathrm{s}]^{\mathrm{H}}}$ and $\hat{\boldsymbol{U}}_r^{[\mathrm{s}]} = \boldsymbol{U}_r^{[\mathrm{s}]} + \Delta\boldsymbol{U}_r^{[\mathrm{s}]}$, a first order expansion for $\Delta\boldsymbol{T}_r$ is obtained via

$$\begin{aligned}\hat{\boldsymbol{T}}_r &= \left(\boldsymbol{U}_r^{[\mathrm{s}]} + \Delta\boldsymbol{U}_r^{[\mathrm{s}]}\right) \cdot \left(\boldsymbol{U}_r^{[\mathrm{s}]^{\mathrm{H}}} + \Delta\boldsymbol{U}_r^{[\mathrm{s}]^{\mathrm{H}}}\right) \\ &= \boldsymbol{T}_r + \boldsymbol{U}_r^{[\mathrm{s}]} \cdot \Delta\boldsymbol{U}_r^{[\mathrm{s}]^{\mathrm{H}}} + \Delta\boldsymbol{U}_r^{[\mathrm{s}]} \cdot \boldsymbol{U}_r^{[\mathrm{s}]^{\mathrm{H}}} + \mathcal{O}\{\Delta^2\} \\ \Rightarrow \Delta\boldsymbol{T}_r &= \boldsymbol{U}_r^{[\mathrm{s}]} \cdot \Delta\boldsymbol{U}_r^{[\mathrm{s}]^{\mathrm{H}}} + \Delta\boldsymbol{U}_r^{[\mathrm{s}]} \cdot \boldsymbol{U}_r^{\mathrm{H}} + \mathcal{O}\{\Delta^2\},\end{aligned} \tag{56}$$

where in general we have $\Delta\boldsymbol{U}_r^{[\mathrm{s}]} = \boldsymbol{U}_r^{[\mathrm{n}]} \cdot \boldsymbol{\Gamma}_r^{[\mathrm{n}]} + \mathcal{O}\{\Delta^2\}$ (cf. (23)). Using this expansion in (56) we obtain

$$\begin{aligned}\Delta\boldsymbol{T}_r \cdot \boldsymbol{T}_r &= \boldsymbol{U}_r^{[\mathrm{s}]} \cdot \boldsymbol{\Gamma}_r^{[\mathrm{n}]^{\mathrm{H}}} \cdot \underbrace{\boldsymbol{U}_r^{[\mathrm{n}]^{\mathrm{H}}} \cdot \boldsymbol{T}_r}_{\boldsymbol{0}_{M_r - p_r \times M_r}} + \boldsymbol{U}_r^{[\mathrm{n}]} \cdot \boldsymbol{\Gamma}_r^{[\mathrm{n}]} \\ &\quad \cdot \underbrace{\boldsymbol{U}_r^{[\mathrm{s}]^{\mathrm{H}}} \cdot \boldsymbol{T}_r}_{\boldsymbol{U}_r^{[\mathrm{s}]^{\mathrm{H}}}} + \mathcal{O}\{\Delta^2\} \\ &= \boldsymbol{U}_r^{[\mathrm{n}]} \cdot \boldsymbol{\Gamma}_r^{[\mathrm{n}]} \cdot \boldsymbol{U}_r^{[\mathrm{n}]^{\mathrm{H}}} + \mathcal{O}\{\Delta^2\},\end{aligned} \tag{57}$$

which is the desired result. This completes the proof of the theorem. □

APPENDIX C
PROOF OF THEOREM 3

For R-D Standard ESPRIT, the explicit first-order expansion of the estimation error for the k-th spatial frequency in the r-th mode in terms of the signal subspace estimation error $\Delta\boldsymbol{U}_{\mathrm{s}}$ is given by (29). This error can be expressed in terms of the perturbation (noise) matrix $\boldsymbol{N}$ by inserting (20). We obtain

$$\begin{aligned}\Delta\mu_k^{(r)} &= \mathrm{Im}\left\{\boldsymbol{r}_k^{(r)^{\mathrm{T}}} \cdot \mathrm{vec}\{\Delta\boldsymbol{U}_{\mathrm{s}}\}\right\} + \mathcal{O}\{\Delta^2\} \\ &= \mathrm{Im}\left\{\boldsymbol{r}_k^{(r)^{\mathrm{T}}} \cdot \boldsymbol{W}_{\mathrm{mat}} \cdot \mathrm{vec}\{\boldsymbol{N}\}\right\} + \mathcal{O}\{\Delta^2\} \\ \boldsymbol{r}_k^{(r)^{\mathrm{T}}} &= \boldsymbol{q}_k^{(r)^{\mathrm{T}}} \otimes \left(\boldsymbol{p}_k^{(r)^{\mathrm{T}}} \cdot \left(\tilde{\boldsymbol{J}}_1^{(r)} \cdot \boldsymbol{U}_{\mathrm{s}}\right)^+ \cdot \left[\tilde{\boldsymbol{J}}_2^{(r)} / \lambda_k^{(r)} - \tilde{\boldsymbol{J}}_1^{(r)}\right]\right) \\ \boldsymbol{W}_{\mathrm{mat}} &= \left(\boldsymbol{\Sigma}_{\mathrm{s}}^{-1} \cdot \boldsymbol{V}_{\mathrm{s}}^{\mathrm{T}}\right) \otimes \left(\boldsymbol{U}_{\mathrm{n}} \cdot \boldsymbol{U}_{\mathrm{n}}^{\mathrm{H}}\right)\end{aligned} \tag{58}$$

which follows directly by applying property (1) to (29) and to (20). In order to expand $\mathbb{E}\{(\Delta\mu_k^{(r)})^2\}$ using (58), we observe that for arbitrary complex vectors $\boldsymbol{z}_1, \boldsymbol{z}_2$ we have $\mathrm{Im}\{\boldsymbol{z}_1^{\mathrm{T}} \cdot \boldsymbol{z}_2\} = \mathrm{Im}\{\boldsymbol{z}_1\}^{\mathrm{T}} \cdot \mathrm{Re}\{\boldsymbol{z}_2\} + \mathrm{Re}\{\boldsymbol{z}_1\}^{\mathrm{T}} \cdot \mathrm{Im}\{\boldsymbol{z}_2\}$ and hence

$$\begin{aligned}\mathrm{Im}\left\{\boldsymbol{z}_1^{\mathrm{T}} \cdot \boldsymbol{z}_2\right\}^2 &= \mathrm{Im}\{\boldsymbol{z}_1\}^{\mathrm{T}} \cdot \mathrm{Re}\{\boldsymbol{z}_2\} \cdot \mathrm{Re}\{\boldsymbol{z}_2\}^{\mathrm{T}} \cdot \mathrm{Im}\{\boldsymbol{z}_1\} \\ &\quad + \mathrm{Re}\{\boldsymbol{z}_1\}^{\mathrm{T}} \cdot \mathrm{Im}\{\boldsymbol{z}_2\} \cdot \mathrm{Im}\{\boldsymbol{z}_2\}^{\mathrm{T}} \cdot \mathrm{Re}\{\boldsymbol{z}_1\} \\ &\quad + \mathrm{Im}\{\boldsymbol{z}_1\}^{\mathrm{T}} \cdot \mathrm{Re}\{\boldsymbol{z}_2\} \cdot \mathrm{Im}\{\boldsymbol{z}_2\}^{\mathrm{T}} \cdot \mathrm{Re}\{\boldsymbol{z}_1\} \\ &\quad + \mathrm{Re}\{\boldsymbol{z}_1\}^{\mathrm{T}} \cdot \mathrm{Im}\{\boldsymbol{z}_2\} \cdot \mathrm{Re}\{\boldsymbol{z}_2\}^{\mathrm{T}} \cdot \mathrm{Im}\{\boldsymbol{z}_1\}\end{aligned} \tag{59}$$

Using (58) in $\mathbb{E}\{(\Delta\mu_k^{(r)})^2\}$ and applying (59) for $\boldsymbol{z}_1^{\mathrm{T}} = \boldsymbol{r}_k^{(r)^{\mathrm{T}}} \cdot \boldsymbol{W}_{\mathrm{mat}}$ and $\boldsymbol{z}_2 = \mathrm{vec}\{\boldsymbol{N}\} = \boldsymbol{n}$ we find

$$\begin{aligned}\mathbb{E}\left\{\left(\Delta\mu_k^{(r)}\right)^2\right\} &= \mathbb{E}\left\{\mathrm{Im}\left\{\boldsymbol{z}_1^{\mathrm{T}}\right\} \cdot \mathrm{Re}\{\boldsymbol{n}\} \cdot \mathrm{Re}\{\boldsymbol{n}\}^{\mathrm{T}} \cdot \mathrm{Im}\{\boldsymbol{z}_1\}\right\} \\ &\quad + \mathbb{E}\left\{\mathrm{Re}\left\{\boldsymbol{z}_1^{\mathrm{T}}\right\} \cdot \mathrm{Im}\{\boldsymbol{n}\} \cdot \mathrm{Im}\{\boldsymbol{n}\}^{\mathrm{T}} \cdot \mathrm{Re}\{\boldsymbol{z}_1\}\right\} \\ &\quad + \mathbb{E}\left\{\mathrm{Im}\left\{\boldsymbol{z}_1^{\mathrm{T}}\right\} \cdot \mathrm{Re}\{\boldsymbol{n}\} \cdot \mathrm{Im}\{\boldsymbol{n}\}^{\mathrm{T}} \cdot \mathrm{Re}\{\boldsymbol{z}_1\}\right\} \\ &\quad + \mathbb{E}\left\{\mathrm{Re}\left\{\boldsymbol{z}_1^{\mathrm{T}}\right\} \cdot \mathrm{Im}\{\boldsymbol{n}\} \cdot \mathrm{Re}\{\boldsymbol{n}\}^{\mathrm{T}} \cdot \mathrm{Im}\{\boldsymbol{z}_1\}\right\}\end{aligned} \tag{60}$$

Since the only random quantity in (60) is the vector of noise samples $\boldsymbol{n}$, we can move $\boldsymbol{z}_1$ out of the expectation operator. We are then left with the covariance matrices of the real and the imaginary part of the noise, respectively, as well as with the cross-covariance matrix between the real and the imaginary part. To proceed, we require the following lemma:

Lemma 2: Let $\boldsymbol{n}$ be a zero mean random vector with covariance matrix $\boldsymbol{R}_{\mathrm{nn}} = \mathbb{E}\{\boldsymbol{n} \cdot \boldsymbol{n}^{\mathrm{H}}\} \in \mathbb{C}^{MN \times MN}$ and pseudo-covariance matrix $\boldsymbol{C}_{\mathrm{nn}} = \mathbb{E}\{\boldsymbol{n} \cdot \boldsymbol{n}^{\mathrm{T}}\} \in \mathbb{C}^{MN \times MN}$. Then, the covariance matrices of the real part of $\boldsymbol{n}$, the imaginary part of $\boldsymbol{n}$ and the cross-covariance between the real and the imaginary part of $\boldsymbol{n}$ are given by

$$\begin{aligned}\boldsymbol{R}_{\mathrm{nn}}^{(\mathrm{R,R})} &\doteq \mathbb{E}\left\{\mathrm{Re}\{\boldsymbol{n}\} \cdot \mathrm{Re}\{\boldsymbol{n}\}^{\mathrm{T}}\right\} = \frac{1}{2}\mathrm{Re}\{\boldsymbol{R}_{\mathrm{nn}} + \boldsymbol{C}_{\mathrm{nn}}\} \\ \boldsymbol{R}_{\mathrm{nn}}^{(\mathrm{I,I})} &\doteq \mathbb{E}\left\{\mathrm{Im}\{\boldsymbol{n}\} \cdot \mathrm{Im}\{\boldsymbol{n}\}^{\mathrm{T}}\right\} = \frac{1}{2}\mathrm{Re}\{\boldsymbol{R}_{\mathrm{nn}} - \boldsymbol{C}_{\mathrm{nn}}\} \\ \boldsymbol{R}_{\mathrm{nn}}^{(\mathrm{R,I})} &\doteq \mathbb{E}\left\{\mathrm{Re}\{\boldsymbol{n}\} \cdot \mathrm{Im}\{\boldsymbol{n}\}^{\mathrm{T}}\right\} = -\frac{1}{2}\mathrm{Im}\{\boldsymbol{R}_{\mathrm{nn}} - \boldsymbol{C}_{\mathrm{nn}}\} \\ \boldsymbol{R}_{\mathrm{nn}}^{(\mathrm{I,R})} &\doteq \mathbb{E}\left\{\mathrm{Im}\{\boldsymbol{n}\} \cdot \mathrm{Re}\{\boldsymbol{n}\}^{\mathrm{T}}\right\} = \frac{1}{2}\mathrm{Im}\{\boldsymbol{R}_{\mathrm{nn}} + \boldsymbol{C}_{\mathrm{nn}}\}.\end{aligned}$$

Proof: To prove this lemma we expand $\boldsymbol{R}_{\mathrm{nn}}$ and $\boldsymbol{C}_{\mathrm{nn}}$ by inserting $\boldsymbol{n} = \mathrm{Re}\{\boldsymbol{n}\} + \jmath\mathrm{Im}\{\boldsymbol{n}\}$. We then obtain

$$\boldsymbol{R}_{\mathrm{nn}} = \boldsymbol{R}_{\mathrm{nn}}^{(\mathrm{R,R})} + \boldsymbol{R}_{\mathrm{nn}}^{(\mathrm{I,I})} + \jmath\left(\boldsymbol{R}_{\mathrm{nn}}^{(\mathrm{I,R})} - \boldsymbol{R}_{\mathrm{nn}}^{(\mathrm{R,I})}\right) \tag{61}$$

$$\boldsymbol{C}_{\mathrm{nn}} = \boldsymbol{R}_{\mathrm{nn}}^{(\mathrm{R,R})} - \boldsymbol{R}_{\mathrm{nn}}^{(\mathrm{I,I})} + \jmath\left(\boldsymbol{R}_{\mathrm{nn}}^{(\mathrm{I,R})} + \boldsymbol{R}_{\mathrm{nn}}^{(\mathrm{R,I})}\right). \tag{62}$$

Since $\boldsymbol{R}_{\mathrm{nn}}^{(\mathrm{R,R})}$, $\boldsymbol{R}_{\mathrm{nn}}^{(\mathrm{R,I})}$, $\boldsymbol{R}_{\mathrm{nn}}^{(\mathrm{I,R})}$, and $\boldsymbol{R}_{\mathrm{nn}}^{(\mathrm{I,I})}$ are real-valued, inserting (61) and (62) into the equations above immediately shows the lemma. ■

Using Lemma 2 in (60), we obtain for the mean square error

$$\begin{aligned}\mathbb{E}\left\{\left(\Delta\mu_k^{(r)}\right)^2\right\} &= \frac{1}{2}\left(\mathrm{Im}\left\{\boldsymbol{z}_1^{\mathrm{T}}\right\} \cdot \mathrm{Re}\{\boldsymbol{R}_{\mathrm{nn}} + \boldsymbol{C}_{\mathrm{nn}}\} \cdot \mathrm{Im}\{\boldsymbol{z}_1\}\right. \\ &\quad + \mathrm{Re}\left\{\boldsymbol{z}_1^{\mathrm{T}}\right\} \cdot \mathrm{Re}\{\boldsymbol{R}_{\mathrm{nn}} - \boldsymbol{C}_{\mathrm{nn}}\} \cdot \mathrm{Re}\{\boldsymbol{z}_1\} \\ &\quad + \mathrm{Im}\left\{\boldsymbol{z}_1^{\mathrm{T}}\right\} \cdot \mathrm{Im}\{-\boldsymbol{R}_{\mathrm{nn}} + \boldsymbol{C}_{\mathrm{nn}}\} \cdot \mathrm{Re}\{\boldsymbol{z}_1\} \\ &\quad \left. + \mathrm{Re}\left\{\boldsymbol{z}_1^{\mathrm{T}}\right\} \cdot \mathrm{Im}\{\boldsymbol{R}_{\mathrm{nn}} + \boldsymbol{C}_{\mathrm{nn}}\} \cdot \mathrm{Im}\{\boldsymbol{z}_1\}\right).\end{aligned} \tag{63}$$

Finally, via straightforward algebraic manipulations, (63) can be expressed in more compact form as

$$\mathbb{E}\left\{\left(\Delta\mu_k^{(r)}\right)^2\right\} = \frac{1}{2}\left(\boldsymbol{z}_1^{\mathrm{H}}\cdot\boldsymbol{R}_{\mathrm{nn}}^{\mathrm{T}}\cdot\boldsymbol{z}_1 - \mathrm{Re}\left\{\boldsymbol{z}_1^{\mathrm{T}}\cdot\boldsymbol{C}_{\mathrm{nn}}^{\mathrm{T}}\cdot\boldsymbol{z}_1\right\}\right) \tag{64}$$

for $\boldsymbol{z}_1 = \boldsymbol{W}_{\mathrm{mat}}^{\mathrm{T}}\cdot\boldsymbol{r}_k^{(r)}$, which is the desired result. □

Note that Gaussianity is not needed for these properties to hold. Consequently, the MSE expressions are still valid if the noise is not Gaussian. We only need it to be zero mean. Also, note that for the special case of circularly symmetric white noise we have $\boldsymbol{R}_{\mathrm{nn}} = \sigma_{\mathrm{n}}^2\cdot\boldsymbol{I}_{MN}$ and $\boldsymbol{C}_{\mathrm{nn}} = \boldsymbol{0}_{MN\times MN}$ and hence the MSE simplifies into

$$\mathbb{E}\left\{\left(\Delta\mu_k^{(r)}\right)^2\right\} = \frac{\sigma_{\mathrm{n}}^2}{2}\cdot\|\boldsymbol{z}_1\|^2 = \frac{\sigma_{\mathrm{n}}^2}{2}\cdot\left\|\boldsymbol{W}_{\mathrm{mat}}^{\mathrm{T}}\cdot\boldsymbol{r}_k^{(r)}\right\|^2. \tag{65}$$

The procedure for 2-D Standard Tensor-ESPRIT is in fact quite similar. The first step is to express the estimation error in $\mu_k^{(r)}$ in terms of the perturbation $\boldsymbol{n} = \mathrm{vec}\{\boldsymbol{N}\} = \mathrm{vec}\{[N]_{(3)}^{\mathrm{T}}\}$ (cf. (14)). This expression takes the form

$$\begin{aligned}\Delta\mu_k^{(r)} &= \mathrm{Im}\left\{\boldsymbol{r}_k^{(r)^{\mathrm{T}}}\cdot\mathrm{vec}\left\{\left[\Delta\hat{\boldsymbol{\mathcal{U}}}^{[\mathrm{s}]}\right]_{(R+1)}^{\mathrm{T}}\right\}\right\} + \mathcal{O}\{\Delta^2\}\\ &= \mathrm{Im}\left\{\boldsymbol{r}_k^{(r)^{\mathrm{T}}}\cdot\boldsymbol{W}_{\mathrm{ten}}\cdot\mathrm{vec}\{\boldsymbol{N}\}\right\} + \mathcal{O}\{\Delta^2\}\end{aligned} \tag{66}$$

since $[\Delta\hat{\boldsymbol{\mathcal{U}}}^{[\mathrm{s}]}]_{(R+1)}^{\mathrm{T}}$ depends linearly on $\mathrm{vec}\{\boldsymbol{N}\}$. The explicit structure of $\boldsymbol{W}_{\mathrm{ten}}$ is derived below. Due to the fact that (66) has the same form as (58), the second step to expand the MSE expressions follows the same lines as for R-D Standard ESPRIT, which immediately shows that the MSE becomes

$$\mathbb{E}\left\{\left(\Delta\mu_k^{(r)}\right)^2\right\} = \frac{1}{2}\left(\boldsymbol{z}_1^{\mathrm{H}}\cdot\boldsymbol{R}_{\mathrm{nn}}^{\mathrm{T}}\cdot\boldsymbol{z}_1 - \mathrm{Re}\left\{\boldsymbol{z}_1^{\mathrm{T}}\cdot\boldsymbol{C}_{\mathrm{nn}}^{\mathrm{T}}\cdot\boldsymbol{z}_1\right\}\right) \tag{67}$$

for $\boldsymbol{z}_1 = \boldsymbol{W}_{\mathrm{ten}}^{\mathrm{T}}\cdot\boldsymbol{r}_k^{(r)}$. Therefore, the final step is finding an explicit expression for $\boldsymbol{W}_{\mathrm{ten}}$ which satisfies

$$\left[\Delta\hat{\boldsymbol{\mathcal{U}}}^{[\mathrm{s}]}\right]_{(3)}^{\mathrm{T}} = \boldsymbol{W}_{\mathrm{ten}}\cdot\mathrm{vec}\{\boldsymbol{N}\} + \mathcal{O}\{\Delta^2\}. \tag{68}$$

Recall from Theorem 2 that for $R = 2$, the HOSVD-based signal subspace estimation error $[\Delta\hat{\boldsymbol{\mathcal{U}}}^{[\mathrm{s}]}]_{(R+1)}$ can be expanded into

$$\begin{aligned}\left[\Delta\hat{\boldsymbol{\mathcal{U}}}^{[\mathrm{s}]}\right]_{(3)}^{\mathrm{T}} &= (\boldsymbol{T}_1\otimes\boldsymbol{T}_2)\cdot\Delta\boldsymbol{U}_{\mathrm{s}} + \left(\left[\Delta\boldsymbol{U}_1^{[\mathrm{s}]}\cdot\boldsymbol{U}_1^{[\mathrm{s}]^{\mathrm{H}}}\right]\otimes\boldsymbol{T}_2\right)\cdot\boldsymbol{U}_{\mathrm{s}}\\ &\quad + \left(\boldsymbol{T}_1\otimes\left[\Delta\boldsymbol{U}_2^{[\mathrm{s}]}\cdot\boldsymbol{U}_2^{[\mathrm{s}]^{\mathrm{H}}}\right]\right)\cdot\boldsymbol{U}_{\mathrm{s}} + \mathcal{O}\{\Delta^2\},\end{aligned} \tag{69}$$

where $\Delta\boldsymbol{U}_{\mathrm{s}}$, $\Delta\boldsymbol{U}_1^{[\mathrm{s}]}$, and $\Delta\boldsymbol{U}_2^{[\mathrm{s}]}$ are given by

$$\begin{aligned}\Delta\boldsymbol{U}_{\mathrm{s}} &= \boldsymbol{U}_{\mathrm{n}}\cdot\boldsymbol{U}_{\mathrm{n}}^{\mathrm{H}}\cdot\boldsymbol{N}\cdot\boldsymbol{V}_{\mathrm{s}}\cdot\boldsymbol{\Sigma}_{\mathrm{s}}^{-1}\\ &= \boldsymbol{V}_3^{[\mathrm{n}]^*}\cdot\boldsymbol{V}_3^{[\mathrm{n}]^{\mathrm{T}}}\cdot\boldsymbol{N}\cdot\boldsymbol{U}_3^{[\mathrm{s}]^*}\cdot\boldsymbol{\Sigma}_3^{[\mathrm{s}]^{-1}} \quad \text{and}\\ \Delta\boldsymbol{U}_r^{[\mathrm{s}]} &= \boldsymbol{U}_r^{[\mathrm{n}]}\cdot\boldsymbol{U}_r^{[\mathrm{n}]^{\mathrm{H}}}\cdot[\boldsymbol{\mathcal{N}}]_{(r)}\cdot\boldsymbol{V}_r^{[\mathrm{s}]}\cdot\boldsymbol{\Sigma}_r^{[\mathrm{s}]^{-1}} \quad \text{for } r = 1, 2.\end{aligned} \tag{70}$$

The first term in (69) is easily vectorized by applying property (1) which yields the first term of $\boldsymbol{W}_{\mathrm{ten}}$ as

$$\begin{aligned}&\mathrm{vec}\left\{(\boldsymbol{T}_1\otimes\boldsymbol{T}_2)\cdot\Delta\boldsymbol{U}_{\mathrm{s}}\right\}\\ &\quad = \mathrm{vec}\left\{(\boldsymbol{T}_1\otimes\boldsymbol{T}_2)\cdot\boldsymbol{V}_3^{[\mathrm{n}]^*}\cdot\boldsymbol{V}_3^{[\mathrm{n}]^{\mathrm{T}}}\cdot\boldsymbol{N}\cdot\boldsymbol{U}_3^{[\mathrm{s}]^*}\cdot\boldsymbol{\Sigma}_3^{[\mathrm{s}]^{-1}}\right\}\\ &\quad = \left(\boldsymbol{U}_3^{[\mathrm{s}]^*}\cdot\boldsymbol{\Sigma}_3^{[\mathrm{s}]^{-1}}\right)^{\mathrm{T}}\otimes\left[(\boldsymbol{T}_1\otimes\boldsymbol{T}_2)\cdot\boldsymbol{V}_3^{[\mathrm{n}]^*}\cdot\boldsymbol{V}_3^{[\mathrm{n}]^{\mathrm{T}}}\right]\cdot\mathrm{vec}\{\boldsymbol{N}\}\\ &\quad = \left(\boldsymbol{\Sigma}_3^{[\mathrm{s}]^{-1}}\cdot\boldsymbol{U}_3^{[\mathrm{s}]^{\mathrm{H}}}\right)\otimes\left[(\boldsymbol{T}_1\otimes\boldsymbol{T}_2)\cdot\boldsymbol{V}_3^{[\mathrm{n}]^*}\cdot\boldsymbol{V}_3^{[\mathrm{n}]^{\mathrm{T}}}\right]\cdot\mathrm{vec}\{\boldsymbol{N}\}.\end{aligned}$$

However, for the second term in (69) we obtain

$$\begin{aligned}&\mathrm{vec}\left\{\left(\left[\boldsymbol{U}_1^{[\mathrm{n}]}\cdot\boldsymbol{U}_1^{[\mathrm{n}]^{\mathrm{H}}}\cdot[\boldsymbol{\mathcal{N}}]_{(1)}\cdot\boldsymbol{V}_1^{[\mathrm{s}]}\cdot\boldsymbol{\Sigma}_1^{[\mathrm{s}]^{-1}}\cdot\boldsymbol{U}_1^{[\mathrm{s}]^{\mathrm{H}}}\right]\otimes\boldsymbol{T}_2\right)\cdot\boldsymbol{U}_{\mathrm{s}}\right\}\\ &\quad = \left(\boldsymbol{U}_{\mathrm{s}}^{\mathrm{T}}\otimes\boldsymbol{I}_M\right)\cdot\mathrm{vec}\left\{\left[\boldsymbol{U}_1^{[\mathrm{n}]}\cdot\boldsymbol{U}_1^{[\mathrm{n}]^{\mathrm{H}}}\cdot[\boldsymbol{\mathcal{N}}]_{(1)}\right.\right.\\ &\qquad\qquad \left.\left.\cdot\boldsymbol{V}_1^{[\mathrm{s}]}\cdot\boldsymbol{\Sigma}_1^{[\mathrm{s}]^{-1}}\cdot\boldsymbol{U}_1^{[\mathrm{s}]^{\mathrm{H}}}\right]\otimes\boldsymbol{T}_2\right\}\end{aligned}$$

by inserting (70) for $\Delta\boldsymbol{U}_1^{[\mathrm{s}]}$. To proceed we need to rewrite the vectorization of a Kronecker product. After straightforward calculations we obtain

$$\begin{aligned}&\left(\boldsymbol{U}_{\mathrm{s}}^{\mathrm{T}}\otimes\boldsymbol{I}_M\right)\cdot\mathrm{vec}\left\{\left[\boldsymbol{U}_1^{[\mathrm{n}]}\cdot\boldsymbol{U}_1^{[\mathrm{n}]^{\mathrm{H}}}\cdot[\boldsymbol{\mathcal{N}}]_{(1)}\right.\right.\\ &\qquad\qquad \left.\left.\cdot\boldsymbol{V}_1^{[\mathrm{s}]}\cdot\boldsymbol{\Sigma}_1^{[\mathrm{s}]^{-1}}\cdot\boldsymbol{U}_1^{[\mathrm{s}]^{\mathrm{H}}}\right]\otimes\boldsymbol{T}_2\right\}\\ &\quad = \left(\boldsymbol{U}_{\mathrm{s}}^{\mathrm{T}}\otimes\boldsymbol{I}_M\right)\cdot\bar{\boldsymbol{T}}_2\\ &\qquad \cdot\mathrm{vec}\left\{\left[\boldsymbol{U}_1^{[\mathrm{n}]}\cdot\boldsymbol{U}_1^{[\mathrm{n}]^{\mathrm{H}}}\cdot[\boldsymbol{\mathcal{N}}]_{(1)}\cdot\boldsymbol{V}_1^{[\mathrm{s}]}\cdot\boldsymbol{\Sigma}_1^{[\mathrm{s}]^{-1}}\cdot\boldsymbol{U}_1^{[\mathrm{s}]^{\mathrm{H}}}\right]\right\}\\ &\quad = \left(\boldsymbol{U}_{\mathrm{s}}^{\mathrm{T}}\otimes\boldsymbol{I}_M\right)\cdot\bar{\boldsymbol{T}}_2\\ &\qquad \cdot\left[\left(\boldsymbol{V}_1^{[\mathrm{s}]}\cdot\boldsymbol{\Sigma}_1^{[\mathrm{s}]^{-1}}\cdot\boldsymbol{U}_1^{[\mathrm{s}]^{\mathrm{H}}}\right)^{\mathrm{T}}\otimes\left(\boldsymbol{U}_1^{[\mathrm{n}]}\cdot\boldsymbol{U}_1^{[\mathrm{n}]^{\mathrm{H}}}\right)\right]\mathrm{vec}\left\{[\boldsymbol{\mathcal{N}}]_{(1)}\right\}\end{aligned} \tag{71}$$

where the matrix $\bar{\boldsymbol{T}}_2$ is constructed from the columns of $\boldsymbol{T}_2$ given by $\boldsymbol{t}_{2,m}$ for $m = 1, 2, \ldots, M_2$ in the following manner

$$\bar{\boldsymbol{T}}_2 = \boldsymbol{I}_{M_1}\otimes\begin{bmatrix}\boldsymbol{I}_{M_1}\otimes\boldsymbol{t}_{2,1}\\ \vdots\\ \boldsymbol{I}_{M_1}\otimes\boldsymbol{t}_{2,M_2}\end{bmatrix}. \tag{72}$$

The final step is to rearrange the elements of $\mathrm{vec}\{[\boldsymbol{\mathcal{N}}]_{(1)}\}$ so that they appear in the same order as in $\mathrm{vec}\{\boldsymbol{N}\}$. However, since $\boldsymbol{N} = [\boldsymbol{\mathcal{N}}]_{(3)}^{\mathrm{T}}$, this can easily be achieved in the following manner

$$\mathrm{vec}\left\{[\boldsymbol{\mathcal{N}}]_{(1)}\right\} = \boldsymbol{K}_{M_2\times(M_1\cdot N)}\cdot\mathrm{vec}\{\boldsymbol{N}\} \tag{73}$$

where $\boldsymbol{K}_{M_2\times(M_1\cdot N)} \in \mathbb{R}^{MN\times MN}$ is the commutation matrix (cf. (7)). This completes the derivation of the second term of $\boldsymbol{W}_{\mathrm{ten}}$. The third term is obtained in a similar manner. In this case, no permutation is needed, since $\mathrm{vec}\{[\boldsymbol{\mathcal{N}}]_{(2)}\} = \mathrm{vec}\{[\boldsymbol{\mathcal{N}}]_{(3)}^{\mathrm{T}}\} = \mathrm{vec}\{\boldsymbol{N}\}$. □

Appendix D

General R-D Form of Theorem 3

The expression for $\boldsymbol{W}_{\mathrm{ten}}$ is shown in Theorem 3 only for the special case $R = 2$ for notational simplicity. It can easily be generalized to an arbitrary $R \geq 2$. There are two main differences in the general R-D case. Firstly, while for $R = 2$ the vectorization of Kronecker products of two matrices is required (i.e., expanding $\mathrm{vec}\{\boldsymbol{A} \otimes \boldsymbol{X}\}$ and $\mathrm{vec}\{\boldsymbol{X} \otimes \boldsymbol{A}\}$ in terms of $\mathrm{vec}\{\boldsymbol{X}\}$), the general R-D case requires the vectorization of Kronecker products of three matrices (i.e., expanding $\mathrm{vec}\{\boldsymbol{A} \otimes \boldsymbol{X} \otimes \boldsymbol{B}\}$ in terms of $\mathrm{vec}\{\boldsymbol{X}\}$). This is achieved by the matrices $\bar{\boldsymbol{T}}_{a:b}$, which are defined below. Secondly, the noise samples appear through vectorized versions of all unfoldings of the noise tensor. In order to permute these vectors such that they are in a consistent order with the vector $\boldsymbol{n} = \mathrm{vec}\{\boldsymbol{N}\} = \mathrm{vec}\{[\boldsymbol{\mathcal{N}}]^{\mathrm{T}}_{(R+1)}\}$ we employ the permutation matrices $\boldsymbol{P}^{(r)}_{I_1,I_2,\ldots,I_R}$, $r = 1, 2, \ldots, R$, which are defined via

$$\boldsymbol{P}^{(r)}_{I_1,\ldots,I_R} \cdot \mathrm{vec}\left\{[\boldsymbol{\mathcal{Z}}]_{(r)}\right\} = \mathrm{vec}\{\boldsymbol{\mathcal{Z}}\} \tag{74}$$

for arbitrary tensors $\boldsymbol{\mathcal{Z}} \in \mathbb{C}^{I_1 \times I_2 \ldots \times I_R}$. For details on the computation and the properties of these permutation matrices, the reader is referred to [32].

Following similar steps as in the proof for $R = 2$ shown in Appendix C we obtain

$$\begin{aligned}
\boldsymbol{W}_{\mathrm{ten}} =& \left(\boldsymbol{\Sigma}_{\mathrm{s}}^{-1} \cdot \boldsymbol{V}_{\mathrm{s}}^{\mathrm{T}}\right) \otimes \left(\boldsymbol{T}_{1:R} \cdot \boldsymbol{U}_{\mathrm{n}} \cdot \boldsymbol{U}_{\mathrm{n}}^{\mathrm{H}}\right) + \sum_{r=1}^{R} \left(\boldsymbol{U}_{\mathrm{s}}^{\mathrm{T}} \otimes \boldsymbol{I}_M\right) \\
&\cdot (\bar{\boldsymbol{T}}_{1:r-1} \otimes \boldsymbol{I}_{M_{r:R}}) \cdot (\boldsymbol{I}_{M_r} \otimes \bar{\boldsymbol{T}}_{r+1:R}) \\
&\cdot \left(\left[\boldsymbol{U}_r^{[\mathrm{s}]^*} \cdot \boldsymbol{\Sigma}_r^{[\mathrm{s}]^{-1}} \cdot \boldsymbol{V}_r^{[\mathrm{s}]^{\mathrm{T}}}\right] \otimes \left[\boldsymbol{U}_r^{[\mathrm{n}]} \cdot \boldsymbol{U}_r^{[\mathrm{n}]^{\mathrm{H}}}\right] \right) \\
&\cdot \boldsymbol{P}^{(r)^{\mathrm{T}}}_{M_1,\ldots,M_R,N} \cdot \boldsymbol{P}^{(R)}_{M_1,\ldots,M_R,N} \quad \text{where} \\
\bar{\boldsymbol{T}}_{1:r-1} =& \begin{bmatrix} \boldsymbol{I}_{M_{r:R}} \otimes \boldsymbol{t}_{1:r-1,1} \\ \boldsymbol{I}_{M_{r:R}} \otimes \boldsymbol{t}_{1:r-1,2} \\ \vdots \\ \boldsymbol{I}_{M_{r:R}} \otimes \boldsymbol{t}_{1:r-1,M_{1:r-1}} \end{bmatrix}, \\
\bar{\boldsymbol{T}}_{r+1:R} =& \begin{bmatrix} \boldsymbol{I}_{M_r} \otimes \boldsymbol{t}_{r+1:R,1} \\ \boldsymbol{I}_{M_r} \otimes \boldsymbol{t}_{r+1:R,2} \\ \vdots \\ \boldsymbol{I}_{M_r} \otimes \boldsymbol{t}_{r+1:R,M_{r+1:R}} \end{bmatrix},
\end{aligned} \tag{75}$$

and we have introduced the short-hand notation $\boldsymbol{t}_{a:b,n}$ as the n-th column of the matrix $\boldsymbol{T}^{\otimes}_{a:b}$ defined as

$$\boldsymbol{T}^{\otimes}_{a:b} = \begin{cases} \boldsymbol{T}_a \otimes \boldsymbol{T}_{a+1} \otimes \ldots \otimes \boldsymbol{T}_b & a \leq b \\ 1 & a > b \end{cases}. \tag{76}$$

Similarly, $M_{a:b}$ is defined as $M_{a:b} = \prod_{r=a}^{b} M_r$ for $a \leq b$ and equal to 1 for $a > b$.

In order to show that for $R = 2$, expression (76) simplifies to (37) we note that $\bar{\boldsymbol{T}}_{1:r-1} = \boldsymbol{I}_M$ for $r = 1$ (since $M_{1:r-1} = M_{1:0} = 1$) and $\bar{\boldsymbol{T}}_{r+1:R} = \boldsymbol{I}_{M_R}$ for $r = R$ (since $M_{r+1:R} = M_{R+1:R} = 1$). Moreover, it is shown in [32] that $\boldsymbol{P}^{(1)^{\mathrm{T}}}_{M_1,M_2,N} \cdot \boldsymbol{P}^{(2)}_{M_1,M_2,N} = \boldsymbol{K}_{M_2 \times M_1 \cdot N}$ which is used to eliminate the projection matrices for $r = 1$. Finally, for $r = R = 2$ we have $\boldsymbol{P}^{(2)^{\mathrm{T}}}_{M_1,M_2,N} \cdot \boldsymbol{P}^{(2)}_{M_1,M_2,N} = \boldsymbol{I}_{M_2 \cdot M_1 \cdot N}$ since permutation matrices are always unitary.

Appendix E

Proof of Theorem 4

As pointed out in Section IV-D, the inclusion of forward-backward-averaging leads to a very similar model, where all quantities originating from the noise-free observation $\boldsymbol{X}_0$ (or $\boldsymbol{\mathcal{X}}_0$) are replaced by the corresponding quantities for $\boldsymbol{X}_0^{(\mathrm{fba})}$ (or $\boldsymbol{\mathcal{X}}_0^{(\mathrm{fba})}$). Since for Theorem 3 it is only assumed that the desired signal component is superimposed by a zero mean noise contribution, it is directly applicable.

The only point we need to derive are the covariance matrix and the pseudo-covariance matrix of the forward-backward-averaged noise $\boldsymbol{n}^{(\mathrm{fba})} \doteq \mathrm{vec}\{\boldsymbol{N}^{(\mathrm{fba})}\}$, which are needed for the MSE expressions. To this end, we can express $\boldsymbol{n}^{(\mathrm{fba})}$ as

$$\begin{aligned}
\mathrm{vec}\left\{\boldsymbol{N}^{(\mathrm{fba})}\right\} &= \mathrm{vec}\left\{[\boldsymbol{N}, \ \boldsymbol{\Pi}_M \cdot \boldsymbol{N}^* \cdot \boldsymbol{\Pi}_N]\right\} \\
&= \begin{bmatrix} \mathrm{vec}\{\boldsymbol{N}\} \\ (\boldsymbol{\Pi}_N \otimes \boldsymbol{\Pi}_M) \cdot \mathrm{vec}\{\boldsymbol{N}^*\} \end{bmatrix} = \begin{bmatrix} \boldsymbol{n} \\ \boldsymbol{\Pi}_{NM} \cdot \boldsymbol{n}^* \end{bmatrix}
\end{aligned} \tag{77}$$

Equation (77) allows us to express the covariance matrix and the pseudo-covariance matrix of $\boldsymbol{n}^{(\mathrm{fba})}$ via the covariance matrix and the pseudo-covariance matrix of $\boldsymbol{n}$. We obtain

$$\begin{aligned}
&\mathbb{E}\left\{\boldsymbol{n}^{(\mathrm{fba})} \cdot \boldsymbol{n}^{(\mathrm{fba})^{\mathrm{H}}}\right\} \\
&= \begin{bmatrix} \mathbb{E}\{\boldsymbol{n} \cdot \boldsymbol{n}^{\mathrm{H}}\} & \mathbb{E}\{\boldsymbol{n} \cdot \boldsymbol{n}^{\mathrm{T}}\} \cdot \boldsymbol{\Pi}_{MN} \\ \boldsymbol{\Pi}_{MN} \cdot \mathbb{E}\{\boldsymbol{n}^* \cdot \boldsymbol{n}^{\mathrm{H}}\} & \boldsymbol{\Pi}_{MN} \cdot \mathbb{E}\{\boldsymbol{n}^* \cdot \boldsymbol{n}^{\mathrm{T}}\} \cdot \boldsymbol{\Pi}_{MN} \end{bmatrix} \\
&= \begin{bmatrix} \boldsymbol{R}_{\mathrm{nn}} & \boldsymbol{C}_{\mathrm{nn}} \cdot \boldsymbol{\Pi}_{MN} \\ \boldsymbol{\Pi}_{MN} \cdot \boldsymbol{C}_{\mathrm{nn}}^* & \boldsymbol{\Pi}_{MN} \cdot \boldsymbol{R}_{\mathrm{nn}}^* \cdot \boldsymbol{\Pi}_{MN} \end{bmatrix} \\
&\mathbb{E}\left\{\boldsymbol{n}^{(\mathrm{fba})} \cdot \boldsymbol{n}^{(\mathrm{fba})^{\mathrm{T}}}\right\} \\
&= \begin{bmatrix} \mathbb{E}\{\boldsymbol{n} \cdot \boldsymbol{n}^{\mathrm{T}}\} & \mathbb{E}\{\boldsymbol{n} \cdot \boldsymbol{n}^{\mathrm{H}}\} \cdot \boldsymbol{\Pi}_{MN} \\ \boldsymbol{\Pi}_{MN} \cdot \mathbb{E}\{\boldsymbol{n}^* \cdot \boldsymbol{n}^{\mathrm{T}}\} & \boldsymbol{\Pi}_{MN} \cdot \mathbb{E}\{\boldsymbol{n}^* \cdot \boldsymbol{n}^{\mathrm{H}}\} \cdot \boldsymbol{\Pi}_{MN} \end{bmatrix} \\
&= \begin{bmatrix} \boldsymbol{C}_{\mathrm{nn}} & \boldsymbol{R}_{\mathrm{nn}} \cdot \boldsymbol{\Pi}_{MN} \\ \boldsymbol{\Pi}_{MN} \cdot \boldsymbol{R}_{\mathrm{nn}}^* & \boldsymbol{\Pi}_{MN} \cdot \boldsymbol{C}_{\mathrm{nn}}^* \cdot \boldsymbol{\Pi}_{MN} \end{bmatrix}.
\end{aligned}$$

This completes the proof of the theorem. □

References

[1] T. W. Anderson, "Asymptotic theory for principal component analysis," *Ann. Math. Statist.*, vol. 34, pp. 122–148, 1963.

[2] D. R. Brillinger, *Time Series: Data Analysis and Theory*. New York: Holt, Rhinehart and Winston, 1975.

[3] J. P. C. L. Da Costa, F. Roemer, M. Haardt, and R. T. de Sousa, Jr., "Multi-Dimensional model order selection," *EURASIP J. Adv. Signal Process.*, vol. 26, July 2011, review article.

[4] L. De Lathauwer, "First-order perturbation analysis of the best rank-(R_1, R_2, R_3) approximation in multilinear algebra," *J. Chemometr.*, vol. 18, no. 1, pp. 2–11, 2004.

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

[6] J. P. Delmas and Y. Meurisse, "On the second-order statistics of the EVD of sample covariance matrices—Application to the detection of noncircular or/and NonGaussian components," *IEEE Trans. Signal Process.*, vol. 59, no. 8, pp. 4017–4023, Aug. 2011.

[7] A. Eriksson, P. Stoica, and T. Soderstrom, "Second-order properties of MUSIC and ESPRIT estimates of sinusoidal frequencies in high SNR scenarios," *IEE Proc.-F*, vol. 140, no. 4, Aug. 1993.

[8] B. Friedlander, “A sensitivity analysis of MUSIC algorithm,” *IEEE Trans. Acoustics, Speech, Signal Process.*, vol. ASSP-38, pp. 1740–1751, 1990.

[9] T. Fu and X. Gao, “Simultaneous diagonalization with similarity transformation for non-defective matrices,” in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP 2006)*, Toulouse, France, May 2006.

[10] M. Haardt and J. A. Nossek, “Unitary ESPRIT: How to obtain increased estimation accuracy with a reduced computational burden,” *IEEE Trans. Signal Process.*, vol. 43, no. 5, pp. 1232–1242, May 1995.

[11] M. Haardt and J. A. Nossek, “Simultaneous Schur decomposition of several non-symmetric matrices to achieve automatic pairing in multidimensional harmonic retrieval problems,” *IEEE Trans. Signal Process.*, vol. 46, no. 1, pp. 161–169, Jan. 1998.

[12] 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 Trans. Signal Process.*, vol. 56, pp. 3198–3213, Jul. 2008.

[13] M. Haardt, R. S. Thomä, and A. Richter, “Multidimensional high-resolution parameter estimation with applications to channel sounding,” in *High-Resolution and Robust Signal Processing*, Y. Hua, A. Gershman, and Q. Chen, Eds. New York: Marcel Dekker, 2004, ch. 5, pp. 255–338.

[14] Y. Hua and T. K. Sarkar, “On SVD for estimating generalized eigenvalues of singular matrix pencil in noise,” *IEEE Trans. Signal Process.*, vol. 39, pp. 892–900, Apr. 1991.

[15] M. Kaveh and A. J. Barabell, “The statistical performance of the MUSIC and the minimum-norm algorithms in resolving plane waves in noise,” *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-34, pp. 331–341, Apr. 1986.

[16] H. Krim, P. Forster, and J. G. Proakis, “Operator approach to performance analysis of root-MUSIC and root-min-norm,” *IEEE Trans. Signal Process.*, vol. 40, no. 7, pp. 1687–1696, Jul. 1992.

[17] P. Lancaster and M. Tismentsky, *The Theory of Matrices*, 2nd ed. New York: Academic , 1978.

[18] A. Lee, “Centrohermitian and skew-centrohermitian matrices,” *Linear Algebra Its Appl.*, vol. 29, pp. 205–210, 1980.

[19] 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.

[20] F. Li and R. J. Vaccaro, “Performance degradation of DOA estimators due to unknown noise fields,” *IEEE Trans. Signal Process.*, vol. 40, no. 3, pp. 686–690, Mar. 1992.

[21] J. Liu, X. Liu, and X. Ma, “First-order perturbation analysis of singular vectors in singular value decomposition,” *IEEE Trans. Signal Process.*, vol. 56, no. 7, pp. 3044–3049, Jul. 2008.

[22] X. Liu, N. D. Sidiropoulos, and A. Swami, “Blind high-resolution localization and tracking of multiple frequency hopped signals,” *IEEE Trans. Signal Process.*, vol. 50, no. 4, pp. 889–901, Apr. 2002.

[23] J. R. Magnus and H. Neudecker, *Matrix Differential Calculus With Applications in Statistics and Econometrics*. New York: Wiley , 1995.

[24] C. P. Mathews, M. Haardt, and M. D. Zoltowski, “Performance analysis of closed-form, ESPRIT based 2-D angle estimator for rectangular arrays,” *IEEE Signal Process. Lett.*, vol. 3, pp. 124–126, Apr. 1996.

[25] C. P. Mathews and M. D. Zoltowski, “Performance analysis of the UCA-ESPRIT algorithm for circular ring arrays,” *IEEE Trans. Signal Process.*, vol. 42, no. 9, Sep. 1994.

[26] D. Nion and N. D. Sidiropoulos, “Tensor algebra and multidimensional harmonic retrieval in signal processing for MIMO radar,” *IEEE Trans. Signal Process.*, vol. 58, no. 11, pp. 5693–5705, Nov. 2010.

[27] B. Ottersten, M. Viberg, and T. Kailath, “Performance analysis of the total least squares ESPRIT algorithm,” *IEEE Trans. Signal Process.*, vol. 39, no. 5, pp. 1122–1135, May 1991.

[28] S. U. Pillai and B. H. Kwon, “Performance analysis of MUSIC-Type high resolution estimators for direction finding in correlated and coherent scenes,” *IEEE Trans. Acoust., Speech, Signal Process.*, vol. AASP-37, no. 8, Aug. 1989.

[29] B. Porat and B. Friedlander, “Analysis of the asymptotic relative efficiency of the MUSIC algorithm,” *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 36, no. 4, pp. 532–544, Apr. 1988.

[30] B. D. Rao and K. V. S. Hari, “Performance analysis of ESPRIT and TAM in determining the direction of arrival of plane waves in noise,” *IEEE Trans. Acoust., Speech, Signal Process.*, vol. AASP-37, no. 12, Dec. 1989.

[31] B. D. Rao and K. V. S. Hari, “Performance analysis of Root-MUSIC,” *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 12, pp. 1939–1948, Dec. 1989.

[32] F. Roemer, “Advanced Algebraic Concepts for Efficient Multi-Channel Signal Processing” Ph.D. thesis, Ilmenau Univ. Technol., Ilmenau, Germany, Oct. 2012 [Online]. Available: http://www.db-thueringen.de/servlets/DocumentServlet?id=21627

[33] F. Roemer, H. Becker, and M. Haardt, “Analytical performance analysis for multi-dimensional Tensor-ESPRIT-type parameter estimation algorithms,” in *Proc. IEEE Int. Conf. Acoust, Speech Signal Process. (ICASSP 2010)*, Dallas, TX, Mar. 2010.

[34] F. Roemer, H. Becker, M. Haardt, and M. Weis, “Analytical performance evaluation for HOSVD-based parameter estimation schemes,” in *Proc. IEEE Int. Workshop on Comp. Adv. in Multi-Sens. Adapt. Process. (CAMSAP 2009)*, Aruba, Dutch Antilles, Dec. 2009.

[35] F. Roemer and M. Haardt, “Analytical performance assessment of 1-D structured least squares,” in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP 2011)*, Prague, Czech Republic, May 2011.

[36] R. Roy, A. Paulraj, and T. Kailath, “ESPRIT—A subspace rotation approach to estimation of parameters of cisoids in noise,” *IEEE Trans. Acoust., Speech, Signal Process.*, vol. AASP-34, pp. 1340–1342, Oct. 1986.

[37] R. O. Schmidt, “Multiple emitter location and signal parameter estimation,” *IEEE Trans. Antennas Propag.*, vol. AP-34, no. 3, pp. 243–258, Mar. 1986, first published in Proc. RADC Spectrum Estimation Workshop, Griffiss Air Force Base, NY, Oct. 1979.

[38] P. Stoica and A. Nehorai, “MUSIC, maximum likelihood, Cramér-Rao bound,” *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 5, pp. 720–741, May 1989.

[39] P. Stoica and T. Soderstrom, “Statistical analysis of MUSIC and subspace rotation estimates of sinusoidal frequencies,” *IEEE Trans. Signal Process.*, vol. 39, no. 8, Aug. 1991.

[40] Z. Xu, “Perturbation analysis for subspace decomposition with applications in subspace-based algorithms,” *IEEE Trans. Signal Process.*, vol. 50, no. 11, pp. 2820–2830, Nov. 2002.

[41] M. D. Zoltowski, G. M. Kautz, and C. P. Mathews, “Performance analysis of eigenstructure based DOA estimators employing conjugate symmetric beamformers,” in *Proc. IEEE Sixth SP Workshop on Statist. Signal Array Process.*, Victoria, BC, Canada, Oct. 1992.

Florian Roemer (S'04–M'13) studied computer engineering at the Ilmenau University of Technology, Germany, and McMaster University, Hamilton, ON, Canada. He received the Diplom-Ingenieur degree in communications engineering and the doctoral (Dr.-Ing.) degree in electrical engineering from the Ilmenau University of Technology in October 2006 and October 2012, respectively.

From December 2006 until September 2012, he was a Research Assistant in the Communications Research Laboratory at Ilmenau University of Technology. In October 2012 he joined the Digital Broadcasting Research Group, a joint research activity between the Fraunhofer Institute for Integrated Circuits IIS and Ilmenau University of Technology, as a postdoctoral research fellow. His research interests include multi-dimensional signal processing, high-resolution parameter estimation, MIMO communications as well as compressive sensing.

Dr. Roemer received the Siemens Communications Academic Award in 2006 for his diploma thesis and the best dissertation award from the Förder- und Freundeskreis (FFK) of the Ilmenau University of Technology for his dissertation.

Martin Haardt (S'90–M'98–SM'99) received his Diplom-Ingenieur (M.S.) degree from the Ruhr-University Bochum in 1991 and his Doktor-Ingenieur (Ph.D.) degree from Munich University of Technology in 1996 after studying electrical engineering at the Ruhr-University Bochum, Germany, and at Purdue University, USA.

He has been a Full Professor in the Department of Electrical Engineering and Information Technology and Head of the Communications Research Laboratory at Ilmenau University of Technology, Germany, since 2001. Since 2012, he

has also served as an Honorary Visiting Professor in the Department of Electronics at the University of York, UK. In 1997 he joined Siemens Mobile Networks, Munich, Germany, where he was responsible for strategic research for third generation mobile radio systems. From 1998 to 2001, he was the Director for International Projects and University Cooperations in the mobile infrastructure business of Siemens in Munich, where his work focused on mobile communications beyond the third generation. During his time at Siemens, he also taught in the international Master of Science in Communications Engineering program at Munich University of Technology. In fall 2006 and fall 2007, he was a Visiting Professor at the University of Nice in Sophia-Antipolis, France, and at the University of York, UK, respectively. His research interests include wireless communications, array signal processing, high-resolution parameter estimation, as well as numerical linear and multi-linear algebra.

Dr. Haardt has received the 2009 Best Paper Award from the IEEE Signal Processing Society, the Vodafone (formerly Mannesmann Mobilfunk) Innovations-Award for outstanding research in mobile communications, the ITG Best Paper award from the Association of Electrical Engineering, Electronics, and Information Technology (VDE), and the Rohde & Schwarz Outstanding Dissertation Award. Prof. Haardt has served as an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING (2002–2006 and since 2011), the IEEE SIGNAL PROCESSING LETTERS (2006–2010), the *Research Letters in Signal Processing* (2007–2009), the Hindawi *Journal of Electrical and Computer Engineering* (since 2009), the EURASIP *Signal Processing Journal* (since 2011), and as a Guest Editor for the EURASIP *Journal on Wireless Communications and Networking*. He has also served as an elected member of the Sensor Array and Multichannel (SAM) technical committee of the IEEE Signal Processing Society (since 2011), as the Technical Co-Chair of the IEEE International Symposiums on Personal Indoor and Mobile Radio Communications (PIMRC) 2005 in Berlin, Germany, as the Technical Program Chair of the IEEE International Symposium on Wireless Communication Systems (ISWCS) 2010 in York, UK, as the General Chair of ISWCS 2013 in Ilmenau, Germany, and as the General-Co Chair of the 5th IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP) 2013 in Saint Martin, French Caribbean.

Giovanni Del Galdo (M'12) was born in 1977 in Merano, Italy. He studied telecommunications engineering at Politecnico di Milano and received his Master's degree in 2002. He then joined the Communications Research Laboratory at Ilmenau University of Technology (TUIL), where he received his doctoral degree in 2007 on the topic of channel modeling for MIMO wireless communication systems.

He has been a Full Professor in the Department of Electrical Engineering and Information Technology and Head of the DVT Research Group at Ilmenau University of Technology since 2012. Shortly thereafter, he joined the Fraunhofer Institute for Integrated Circuits IIS and the International Audio Laboratories Erlangen as senior scientist. Here, he worked on audio watermarking, parametric representations of spatial sound, and sound field analysis via microphone arrays. In 2012 he was appointed full professor at TUIL as head of the DVT Research Group, a joint activity between TUIL and Fraunhofer IIS on the research area of wireless distribution systems and digital broadcasting. The group manages and develops the Fraunhofer IIS Facility for Over-the-air Research and Testing (FORTE) further. His current research interests include the analysis, modeling, and manipulation of multi-dimensional signals, over-the-air testing for terrestrial and satellite communication systems, and sparsity promoting reconstruction methods.