Constrained Cramér-Rao Bound for Higher-Order Singular Value Decomposition

Calis, Metin; Mischi, Massimo; Veen, Alle-Jan van der; Rajan, Raj Thilak; Hunyadi, Borbàla

DOI
10.1109/OJSP.2025.3607278

Publication date
2025

Document Version
Final published version

Published in
IEEE Open Journal of Signal Processing

Citation (APA)
Calis, M., Mischi, M., Veen, A.-J. V. D., Rajan, R. T., & Hunyadi, B. (2025). Constrained Cramér-Rao Bound for Higher-Order Singular Value Decomposition. *IEEE Open Journal of Signal Processing*, *6*, 1048 - 1055. https://doi.org/10.1109/OJSP.2025.3607278

IEEE Open Journal of Signal Processing

Received 12 May 2025; revised 8 August 2025; accepted 25 August 2025. Date of publication 8 September 2025; date of current version 18 September 2025. The review of this article was arranged by Associate Editor L. Balzano.

Digital Object Identifier 10.1109/OJSP.2025.3607278

Short Paper

Constrained Cramér-Rao Bound for Higher-Order Singular Value Decomposition

METIN CALIS [1], MASSIMO MISCHI [2] (Senior Member, IEEE), ALLE-JAN VAN DER VEEN [1] (Fellow, IEEE), RAJ THILAK RAJAN [1] (Senior Member, IEEE), AND BORBÀLA HUNYADI [1] (Senior Member, IEEE)

[1]Faculty of Electrical Engineering, Mathematics and Computer Science, Delft University of Technology, 2628 CD Delft, The Netherlands
[2]Department of Electrical Engineering, Eindhoven University of Technology, 5600 MB Eindhoven, The Netherlands

CORRESPONDING AUTHOR: METIN CALIS (email: m.calis@tudelft.nl).

This work was supported in part by Angiogenesis Analytics and in part by Holland High Tech with a Publieke-Privaat Samenwerking [Public-Private Partnership (PPS)] supplement for Research and Development in the Topsector High Tech Systemen en Materialen [High Tech Systems and Material (HTSM)] under Project 20.091.

ABSTRACT Tensor decomposition methods for signal processing applications are an active area of research. Real data are often low-rank, noisy, and come in a higher-order format. As such, low-rank tensor approximation methods that account for the high-order structure of the data are often used for denoising. One way to represent a tensor in a low-rank form is to decompose the tensor into a set of orthonormal factor matrices and an all-orthogonal core tensor using a higher-order singular value decomposition. Under noisy measurements, the lower bound for recovering the factor matrices and the core tensor is unknown. In this paper, we exploit the well-studied constrained Cramér-Rao bound to calculate a lower bound on the mean squared error of the unbiased estimates of the components of the multilinear singular value decomposition under additive white Gaussian noise, and we validate our approach through simulations.

INDEX TERMS Cramér-Rao bound, performance analysis and bounds, tensor-based signal processing, higher-order singular value decomposition, low multilinear rank.

I. INTRODUCTION

The study and application of tensors has been an active area of research in various domains, ranging from machine learning [1], biomedical signal processing [2] to large-scale optimization and data compression [3]. In many applications, the data tensor exhibits low-rank characteristics and is frequently contaminated by noise [4]. In this paper, we consider the problem of estimating the components that make up a low multilinear rank tensor from noisy measurements.

A standard generalization of matrix rank for tensors does not exist. Many standard matrix rank definitions lead to different notions of rank for tensors. For example, the rank related to the canonical polyadic decomposition (CPD) is the smallest number of rank-one tensors that sum up to the tensor, similar to the smallest number of rank-one matrices that sum up to a matrix. Border rank generalizes the description of the set of ranks less than or equal to R matrices as an algebraic variety [5]. Different matricizations of a tensor will lower bound the tensor rank, such as grouping the first and the last two orders of a fourth-order tensor or the Koszul-Young flattenings [6]. The N different matricizations for an Nth-order tensor, commonly known as the mode-n unfoldings, can be considered as a generalization of modal ranks, i.e., the column and row rank of a matrix. The multilinear rank of a tensor is then an N-tuple with elements defined by the rank of each mode-n unfolding [7]. We focus on such low multilinear rank tensors, the root of which can be traced back to the Tucker decomposition [8].

Any tensor can be expressed as the product of a core tensor and a set of factor matrices, which together constitute the Tucker decomposition [8]. This decomposition is not unique in its unconstrained form: the factor matrices can be right-multiplied by any invertible matrices, provided the core tensor is adjusted accordingly, i.e., multiplied by their inverse, to yield the same reconstructed tensor. To resolve this ambiguity, the authors in [7] proposed the higher-order singular value decomposition (HOSVD) by introducing structural constraints, specifically, the all-orthogonality of the core tensor and the orthonormality of the factor matrices. A numerical procedure for computing the HOSVD components is the multilinear singular value decomposition (MLSVD). The factor matrices are obtained by applying the singular value decomposition (SVD) to each mode-n unfolding of the tensor, taking the left singular vectors as the factor matrices. The all-orthogonal core tensor is then computed by multiplying the tensor by the transposes of the factor matrices along the corresponding modes. When the singular values in each mode-n

unfolding are distinct and ordered, the HOSVD of a real-valued tensor is unique up to a sign ambiguity [7].

According to the Eckart-Young theorem [9], the best low-rank approximation of a matrix can be obtained through the truncated SVD if we consider the Frobenius norm on the residual error. An analogous extension to the high-order case is not possible for HOSVD. The truncated multilinear singular value decomposition (tr-MLSVD) does not guarantee the best $(R_1, R_2, \ldots, R_N)$ multilinear rank approximation of a tensor for both noisy and noiseless cases [10]. Higher-order orthogonal iteration (HOOI) improves the estimation by alternating updates of the factor matrices. Note that in both the matrix and the tensor case, noise can perturb the singular values, singular vectors, and the HOSVD factors, respectively [11], [12]. An important question remains: How well can any unbiased estimator recover the true HOSVD components?

The Cramér-Rao bound (CRB) [13] is a statistical tool to study the performance of an unbiased estimator. Under certain regularity conditions, the CRB provides a lower bound on the asymptotic performance of locally unbiased estimators. In the presence of additional constraints, the authors in [14] introduced the constrained Cramér-Rao bound (CCRB) that lower bounds the error covariance for constrained, unbiased estimators when the unconstrained model has a nonsingular Fisher information matrix (FIM). This is further generalized in [15], where the authors introduced the CCRB without assuming a full rank FIM [16]. In [17], the trace of the pseudo-inverse of the singular FIM is shown to be the minimum achievable MSE given the optimal constraints. We call this bound the oracle bound, since it requires the optimal constraints that make the FIM invertible in the reduced space [18]. The OB is used as a performance bound in various applications, for example, blind channel estimation [19], reference-free clock synchronization [20], and phase estimation [21]. We incorporate the constraints and assumptions that create a unique HOSVD and aim to calculate a lower bound for estimating the noiseless factor matrices and the all-orthogonal core tensor.

The CRB is calculated for a CPD in [22] and further explored in [1]. In [23], the CCRB of a coupled CPD is calculated. In [24], the CCRB of the delayed exponential fitting problem using HOSVD is explored. In this paper, we present the CCRB for the HOSVD of a real tensor under additive white Gaussian noise (WGN). We assume the multilinear ranks are known and there is a unique HOSVD decomposition. We compare our proposed bound with HOOI and tr-MLSVD, which are well-suited for signals in the presence of WGN.

The layout of the paper is as follows. In Section II, tensor notation and tensor preliminaries are introduced. In Section III, the signal model is introduced, and the problem is formulated. In Section IV and Section V, the CRB and the CCRB are introduced. The supplementary material for calculating the bounds can be found in the Appendix. In Section VI, we compare the CRB and CCRB with HOOI and tr-MLSVD through simulation. Finally, we conclude the chapter in Section VIII.

II. TENSOR NOTATION AND PRELIMINARIES

Vectors are represented by lowercase boldface letters such as $\mathbf{b}$. Matrices are represented by boldface letters such as $\mathbf{U}^{(1)}$ and $\mathbf{I}$. The numbers given as superscripts in parentheses are used to refer to the different matrices that share a similar property. For example, the three-factor matrices of the HOSVD for a third-order tensor are denoted by $\mathbf{U}^{(1)}$, $\mathbf{U}^{(2)}$, and $\mathbf{U}^{(3)}$ [7]. The (i_1, i_2)th entry of $\mathbf{A} \in \mathbb{R}^{I_1 \times I_2}$ is represented with $(\mathbf{A})_{i_1,i_2}$, while the i_2th column is represented with $(\mathbf{A})_{i_2}$. Tensors are represented by underlined boldface letters such as $\underline{\mathbf{Y}}$. We use several operations and decompositions using tensors, which are explained in the following paragraph.

The mode-n unfolding of $\underline{\mathbf{Y}} \in \mathbb{R}^{I_1 \times \ldots \times I_N}$ is $\mathbf{Y}_{(n)} \in \mathbb{R}^{I_n \times I_1 \ldots I_{n-1} I_{n+1} \ldots I_N}$. The mode-$n$ product of tensor $\underline{\mathbf{Y}} \in \mathbb{R}^{I_1 \times \ldots \times I_N}$ and a matrix $\mathbf{U} \in \mathbb{R}^{J \times I_n}$ yields a tensor $\underline{\mathbf{Y}} \times_n \mathbf{U} = \underline{\mathbf{C}} \in \mathbb{R}^{I_1 \times \ldots \times I_{n-1} \times J \times I_{n+1} \times \ldots I_N}$ with the property $\mathbf{C}_{(n)} = \mathbf{U}\,\mathbf{Y}_{(n)}$. The grouping of indices $\overline{i_1 \ldots i_N} = i_1 + (i_2 - 1)I_1 + \cdots + (i_N - 1)I_1 \ldots I_{N-1}$ results in $(\mathrm{vec}(\underline{\mathbf{Y}}))_{\overline{i_1 \ldots i_N}} = (\underline{\mathbf{Y}})_{i_1, i_2, \ldots, i_N}$. For more information regarding tensor notations, we refer to [3].

The HOSVD of a low multilinear rank tensor $\underline{\mathbf{Y}} \in \mathbb{R}^{I_1 \times I_2 \times \ldots \times I_N}$ yields $\underline{\mathbf{S}} \times_1 \mathbf{U}^{(1)} \cdots \times_N \mathbf{U}^{(N)}$ where $\underline{\mathbf{S}} \in \mathbb{R}^{R_1 \times R_2 \times \cdots \times R_N}$ is called the all-orthogonal core tensor and $\mathbf{U}^{(n)} \in \mathbb{R}^{I_n \times R_n}$ are column-wise orthonormal factor matrices. The column-wise orthonormality of the factor matrices can be expressed as $\mathbf{U}^{(n)^T}\mathbf{U}^{(n)} = \mathbf{I}$. The all-orthogonality property can be expressed using the diagonal structure of the mode-n unfoldings of the core tensor multiplied by its transpose for $n \in \{1, \ldots, N\}$. Define $\mathbf{\Sigma}^{(n)}$ as the diagonal matrix that holds the squared and ordered mode-n singular values on the diagonal and zeros elsewhere. The all-orthogonality of the core tensor can then be expressed as the diagonal

$$\mathbf{S}_{(n)}\,\mathbf{S}_{(n)}{}^T = \mathbf{\Sigma}^{(n)}, \tag{1}$$

for $n \in \{1, \ldots N\}$. The multilinear ranks $(R_1 \ldots, R_N)$ are the number of columns of the column-wise orthonormal factor matrices $\mathbf{U}^{(n)}$ for $n \in \{1, \ldots, N\}$.

Commonly used mathematical notations are as follows. The Euclidean norm is the square root of the dot product of a vector with itself, which is shown by $\|\cdot\|$. The Kronecker product is shown with $\otimes$. The Kronecker delta $\delta_{i_1 i_2}$ is 1 when $i_1 = i_2$ and zero otherwise. The trace of a matrix $\mathrm{Tr}(\cdot)$ is the sum of the diagonals of that matrix.

We use selection matrices, permutation matrices, and matrices that define indexing relations. $\mathbf{E}^{(n)} \in \mathbb{R}^{R_n(R_n+1)/2 \times R_n^2}$ and $\mathbf{K}^{(n)} \in \mathbb{R}^{R_n(R_n-1)/2 \times R_n^2}$ for $n \in \{1, \ldots, N\}$ are binary-valued matrices. For $n \in \{1, \ldots, N\}$, the products $\mathbf{E}^{(n)}\,\mathrm{vec}(\mathbf{C}^{(n)})$ and $\mathbf{K}^{(n)}\,\mathrm{vec}(\mathbf{C}^{(n)})$ extract the lower triangle and the strict lower triangle (excluding the diagonal) of $\mathbf{C}^{(n)} \in \mathbb{R}^{R_n \times R_n}$, respectively. Permutation matrices $\mathbf{P}^{(n)} \in \mathbb{R}^{D \times D}$ for $n \in \{1, \ldots, N\}$ define a row-wise permutation between the vectorizations of mode-n unfoldings and the vectorization of mode-1 unfolding, where $D = \prod_{n=1}^{N} I_n$. An example is the permutation matrix $\mathbf{P}^{(2)}$ that permutes the vectorization of the mode-2 unfolding of $\underline{\mathbf{Y}}$, such that vectorization of the mode-1 unfolding of $\underline{\mathbf{Y}}$ is obtained, i.e. $\mathrm{vec}(\mathbf{Y}_{(1)}) = \mathbf{P}^{(2)}\,\mathrm{vec}(\mathbf{Y}_{(2)})$. The matrix $\mathbf{P}^{(2)}$ can be generated by first creating a square matrix of 0s of size $D \times D$ and assigning $(\mathbf{P}^{(2)})_{\overline{i_1 i_2 \ldots i_N}, \overline{i_2 i_1 \ldots i_N}} = 1$. As an example, we illustrate the structure of the permutation matrix $\mathbf{P}^{(2)}$ in Fig. 1 for a tensor $\underline{\mathbf{M}} \in \mathbb{R}^{2 \times 2 \times 2}$.

An entry in the tensor $\underline{\mathbf{Y}}$ is placed in different rows or columns according to the indexing rules of the mode-n unfolding and vectorization. We use the matrix $\mathbf{G}^{(n;i_n,k;i_k)} \in \mathbb{R}^{D/I_n \times D/I_k}$ to define an indexing relation between the i_nth row of $\mathbf{Y}_{(n)}$ and i_kth row of $\mathbf{Y}_{(k)}$ for $n \in \{1, \ldots, N\}$ and $k \in \{1, \ldots, N\}$. Additionally, we use $\mathbf{G}^{(n;i_n)} \in \mathbb{R}^{D/I_n \times D}$ to define the indexing relation between the elements of the i_nth row of $\mathbf{Y}_{(n)}$ and $\mathrm{vec}(\underline{\mathbf{Y}})$. To achieve this, we introduce a new notation to describe a grouping with stride, which is formulated as $i_j \overline{i_1 \ldots i_{j-1} i_{j+1} \ldots i_N} = i_j + i_1 + (i_2 - 1)I_1 + \cdots + (i_{j-1} - 1)I_1 \ldots I_{j-2} + (i_{j+1} - 1)I_1 \ldots I_{j-1} + \cdots + (i_N - 1)I_1 \ldots I_{j-1} I_{j+1} \ldots I_{N-1}$ for $j \in \{1, \ldots, N\}$. For illustration, consider the matrix $\mathbf{G}^{(1;i_1,2;i_2)} \in \mathbb{R}^{I_2 I_3 \times I_1 I_3}$ that defines an indexing correspondence between the i_1th row of $\mathbf{Y}_{(1)}$, and the i_2th row of $\mathbf{Y}_{(2)}$. The element $(\mathbf{G})_{i_1 \overline{i_2 i_3}, i_2 \overline{i_1 i_3}}$ is 1, if $i_1 \overline{i_2 i_3} = i_2 \overline{i_1 i_3}$ for i_1, i_2 and

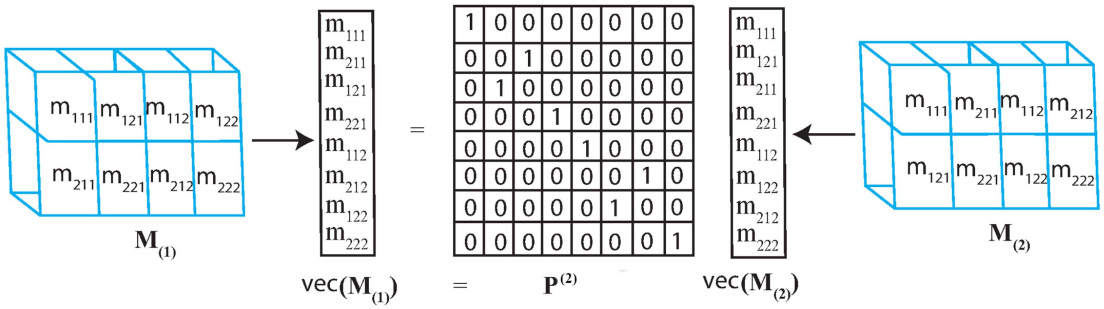

FIGURE 1. The structure of the $\mathbf{P}^{(2)}$ matrix for a tensor $\underline{\mathbf{M}} \in \mathbb{R}^{2\times2\times2}$. P permutes the vectorization of the second mode unfolding of $\underline{\mathbf{M}}$ to the vectorization of the first mode unfolding of $\underline{\mathbf{M}}$.

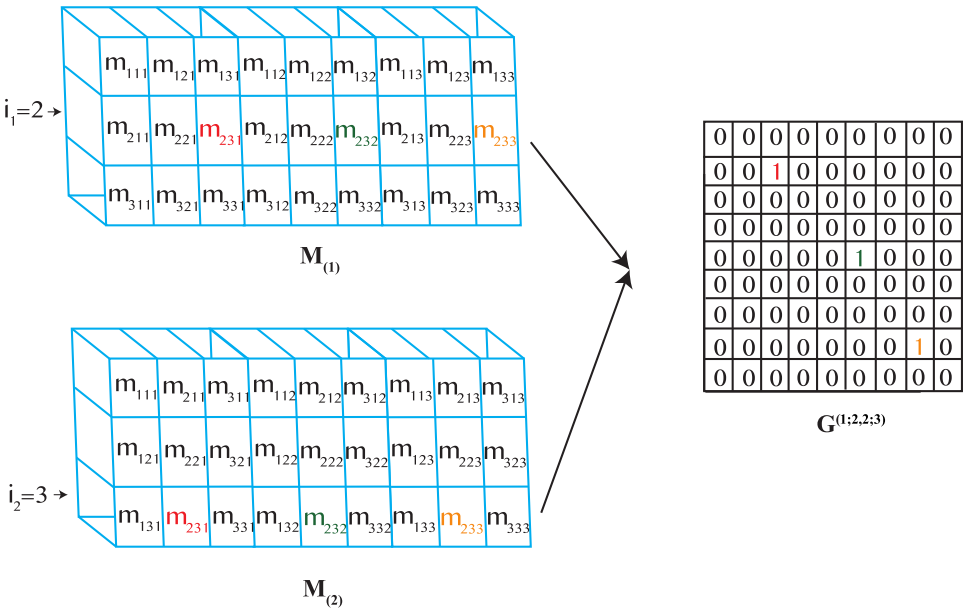

FIGURE 2. The structure of the $\mathbf{G}^{(1;2,2;3)}$ matrix for a tensor $\underline{\mathbf{M}} \in \mathbb{R}^{3\times3\times3}$. G has 1 at indices that are the same between the 2nd row of the mode-1 unfolding of $\underline{\mathbf{M}}$ and the 3rd row of the mode-2 unfolding of the same tensor and 0 s elsewhere.

$i_3 \in \{1, \ldots, I_3\}$ and zero otherwise. This is generated by creating matrix of 0s of size $\mathbf{G} \in \mathbb{R}^{I_2 I_3 \times I_1 I_3}$ and assigning $(\mathbf{G})_{i_1\overline{i_2 i_3}, i_2\overline{i_1 i_3}} = 1$. An example is $\mathbf{G}^{(1;2,2;3)}$ which has 1 at indices that are the same between the 2nd row of the mode-1 unfolding of a 3rd-order tensor of size (3,3,3) and the 3rd row of the mode-2 unfolding of the same tensor and 0s elsewhere. Note that i_1 and i_2 are fixed and i_3 belongs to the set $\{1, \ldots, I_3\}$. The matrix structure is illustrated in Fig. Finally, 2. $\mathbf{G}^{(n;i_n)} \in \mathbb{R}^{D/I_n \times D}$ can be created by initializing a matrix of 0s and assigning $(\mathbf{G})_{i_n\overline{i_1 \ldots i_{n-1} i_{n+1} \ldots i_N}, \overline{i_1 \ldots i_N}} = 1$.

III. PROBLEM FORMULATION

In this paper, we aim to find the CCRB of the core tensor and the factor matrices of a low multilinear rank tensor from a measurement model with additive noise, which can be formulated by

$$\underline{\mathbf{Y}} = \underline{\mathbf{L}} + \underline{\mathbf{M}}, \tag{2}$$

where $\underline{\mathbf{L}} \in \mathbb{R}^{I_1 \times \cdots \times I_N}$ is a deterministic low multilinear rank tensor, and $\underline{\mathbf{M}} \in \mathbb{R}^{I_1 \times \cdots \times I_N}$ is assumed to be a zero-mean white Gaussian noise tensor where each entry is independent and identically distributed with variance σ^2. The low multilinear rank tensor can be written as

$$\underline{\mathbf{L}} = \underline{\mathbf{S}} \times_1 \mathbf{U}^{(1)} \cdots \times_N \mathbf{U}^{(N)}, \tag{3}$$

with orthonormal factor matrices $\mathbf{U}^{(n)} \in \mathbb{R}^{I_n \times R_n}$ for $n \in \{1, \ldots, N\}$ and an all-orthogonal core tensor $\underline{\mathbf{S}} \in \mathbb{R}^{R_1 \times \cdots \times R_N}$, where we assume that the ranks R_n are known for $n \in \{1, \ldots, N\}$.

We can now write the deterministic parameters of interest as $\boldsymbol{\theta} \in \mathbb{R}^{N_\theta}$

$$\boldsymbol{\theta} = [(\mathrm{vec}(\mathbf{U}^{(1)}))^T, \ldots, (\mathrm{vec}(\mathbf{U}^{(N)}))^T, (\mathrm{vec}(\underline{\mathbf{S}}))^T]^T, \tag{4}$$

where $N_\theta = \sum_{n=1}^{N} I_n R_n + \prod_{n=1}^{N} R_n$ is the total size of the factor matrices and the core tensor.

In the following sections, we introduce the CRB and calculate the CCRB of $\boldsymbol{\theta}$ (4) from the signal model (2), given the tensor structure (3).

IV. CRAMÉR-RAO BOUND

Let $f(\underline{\mathbf{Y}}; \boldsymbol{\theta})$ be the likelihood of the observed data $\underline{\mathbf{Y}}$, given the model parameters $\boldsymbol{\theta}$. The log-likelihood $\ln f(\underline{\mathbf{Y}}; \boldsymbol{\theta})$ can be written in three

different ways. For each order $n \in \{1, \ldots, N\}$, we will have

$$\ln f(\underline{\mathbf{Y}}; \boldsymbol{\theta}) = -\left(\prod_{n=1}^{N} I_n\right) \ln(\sigma\sqrt{2\pi}) - \frac{1}{2\sigma^2} \left\| \mathbf{Y}_{(n)} - \mathbf{L}_{(n)} \right\|^2 , \quad (5)$$

which is the logarithm of the multiplication of zero-mean Gaussian probability distribution functions. We take the partial derivative of the log-likelihood with respect to each element of (4) and create the FIM that is defined as

$$\boldsymbol{\Omega}(\boldsymbol{\theta}) = \mathbb{E}\left\{ \left(\frac{\partial \ln f(\underline{\mathbf{Y}}; \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right) \left(\frac{\partial \ln f(\underline{\mathbf{Y}}; \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right)^T \right\}, \quad (6)$$

with size $\mathbb{R}^{N_{\boldsymbol{\theta}} \times N_{\boldsymbol{\theta}}}$.

The FIM is a matrix consisting of submatrices in the following form

$$\boldsymbol{\Omega}(\boldsymbol{\theta}) = \begin{bmatrix} \boldsymbol{\Omega}_{\mathbf{U}^{(1)}\mathbf{U}^{(1)}} & \cdots & \boldsymbol{\Omega}_{\mathbf{U}^{(1)}\mathbf{U}^{(N)}} & \boldsymbol{\Omega}_{\mathbf{U}^{(1)}\underline{\mathbf{S}}} \\ \vdots & \ddots & \vdots & \vdots \\ \boldsymbol{\Omega}^T_{\mathbf{U}^{(1)}\mathbf{U}^{(N)}} & \cdots & \boldsymbol{\Omega}_{\mathbf{U}^{(N)}\mathbf{U}^{(N)}} & \boldsymbol{\Omega}_{\mathbf{U}^{(N)}\underline{\mathbf{S}}} \\ \boldsymbol{\Omega}^T_{\mathbf{U}^{(1)}\underline{\mathbf{S}}} & \cdots & \boldsymbol{\Omega}^T_{\mathbf{U}^{(N)}\underline{\mathbf{S}}} & \boldsymbol{\Omega}_{\underline{\mathbf{S}}\underline{\mathbf{S}}} \end{bmatrix}, \quad (7)$$

where the submatrices divide the FIM according to the parts of $\boldsymbol{\theta}$. The FIM is a symmetric matrix. Therefore, only the upper or lower triangle needs to be calculated. The partial derivatives and the elements of the FIM are given in the Appendix from (21)–(24).

The CRB is the trace of the inverse of the Fisher information matrix. Asymptotically, the mean squared error (MSE) of any locally unbiased estimator is greater than or equal to the CRB, i.e.,

$$\mathbb{E}\left[\left\| \hat{\boldsymbol{\theta}} - \boldsymbol{\theta} \right\|^2 \right] \geq \mathrm{Tr}\left(\boldsymbol{\Omega}(\boldsymbol{\theta})^{-1} \right) , \quad (8)$$

where $\hat{\boldsymbol{\theta}}$ is an estimate of the true parameter vector $\boldsymbol{\theta}$. The FIM (7) is observed to be typically singular, and hence the inverse does not exist [25]. In such cases, constraints can be incorporated into the FIM to effectively reduce the parameter space and obtain a non-singular information matrix over the constrained subspace. The minimum achievable MSE under such optimal constraints is given by the Oracle Bound (OB), which is defined as the trace of the pseudo-inverse of the FIM [17], [18]

$$\mathbb{E}\left[\left\| \hat{\boldsymbol{\theta}} - \boldsymbol{\theta} \right\|^2 \right] \geq \mathrm{Tr}\left(\boldsymbol{\Omega}(\boldsymbol{\theta})^{\dagger} \right) , \quad (9)$$

where $\hat{\boldsymbol{\theta}}$ is an estimate of the true parameter vector $\boldsymbol{\theta}$. Since the optimal constraints are not known a priori, we can impose known structural constraints, such as the orthonormality of factor matrices and the all-orthogonality of the core tensor. This modified CRB or constrained CRB can then be readily derived.

V. CONSTRAINED CRAMÉR-RAO BOUND

We have two constraints on the unknown parameters $\boldsymbol{\theta}$. First, the factor matrices are orthonormal, which can be expressed for $n \in \{1, \ldots, N\}$ as

$$\mathbf{E}^{(n)} \mathrm{vec}(\mathbf{U}^{(n)T}\, \mathbf{U}^{(n)} - \mathbf{I}) = \mathbf{0} , \quad (10)$$

which has $R_n(R_n + 1)/2$ independent equations. Second, the core tensor is all-orthogonal. This condition implies that all off-diagonal elements of $\mathbf{S}_{(n)}\, \mathbf{S}^T_{(n)}$ for $n \in \{1, \ldots, N\}$ are zero, that is,

$$\mathbf{K}^{(n)} \mathrm{vec}\left(\mathbf{S}_{(n)}\, \mathbf{S}^T_{(n)} \right) = \mathbf{0} , \quad (11)$$

which has $R_n(R_n - 1)/2$ independent equations. Since both $\mathbf{S}_{(n)}\, \mathbf{S}^T_{(n)}$ and $\mathbf{U}^{(n)T}\, \mathbf{U}^{(n)}$ are symmetric matrices, we extract the strictly triangular and the triangular parts, respectively, using the selection matrices. The selection matrices $\mathbf{E}^{(n)}$ and $\mathbf{K}^{(n)}$ are introduced in Section II. The constraints in (10) and (11) are quadratic in the unknowns $\boldsymbol{\theta}$, since each involves a bilinear form of either the factor matrices or the unfolded core tensor. Nevertheless, these quadratic constraints can be linearized around $\boldsymbol{\theta}$ through partial differentiation, which yields $\mathbf{C}(\boldsymbol{\theta})$ that is used in the computation of the CCRB. The matrix $\mathbf{C}(\boldsymbol{\theta})$ defines a linear map acting on perturbations of $\boldsymbol{\theta}$ and enables projection onto the constraint-consistent subspace. In total, there are $N_c = \sum_{n=1}^{N}(R_n(R_n+1))/2 + \sum_{n=1}^{N}(R_n(R_n-1))/2 = \sum_{n=1}^{N} R_n^2$ number of constraints.

The partial derivative of $\mathbf{c}(\boldsymbol{\theta})$ with respect to the unknown $\boldsymbol{\theta}$ yields

$$\mathbf{C}(\boldsymbol{\theta}) = \frac{\partial \mathbf{c}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^T} , \quad (12)$$

where $\mathbf{C}(\boldsymbol{\theta}) \in \mathbb{R}^{N_c \times N_\theta}$. The matrix $\mathbf{C}(\boldsymbol{\theta})$ consists of several submatrices in the following form

$$\mathbf{C}(\boldsymbol{\theta}) = \begin{bmatrix} \mathbf{C}_{\mathbf{U}^{(1)}} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \ddots & \mathbf{0} & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{C}_{\mathbf{U}^{(N)}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{C}_{\mathbf{S}_{(1)}} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{C}_{\mathbf{S}_{(2)}}\, \mathbf{P}^{(2)} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{C}_{\mathbf{S}_{(N)}}\, \mathbf{P}^{(N)} \end{bmatrix} \quad (13)$$

The elements of the submatrices in (13) are described in (26) in the Appendix. The submatrix $\mathbf{C}_{\mathbf{U}^{(n)}} \in \mathbb{R}^{R_n(R_n+1)/2 \times I_n R_n}$ is the partial derivative of (10) with respect to $(\mathrm{vec}(\mathbf{U}^{(n)}))^T$ for $n \in \{1, \ldots, N\}$ and the submatrix $\mathbf{C}_{\mathbf{S}_{(n)}} \in \mathbb{R}^{R_n(R_n-1)/2 \times \prod_{n=1}^{N} R_n}$ is the partial derivative of (11) with respect to $(\mathrm{vec}(\mathbf{S}_{(n)}))^T$ for $n \in \{1, \ldots, N\}$. The partial derivatives $\mathbf{C}_{\mathbf{S}_{(n)}}$ are calculated in (27), however the $\mathrm{vec}(\mathbf{S}_{(n)})$ for $n \in \{2, \ldots, N\}$ are not consistent with the $\mathrm{vec}(\mathbf{S}_{(1)}) = \mathrm{vec}(\underline{\mathbf{S}})$. Due to the indexing difference between the vectorization of the different unfoldings of $\underline{\mathbf{S}}$, we use the permutation matrices $\mathbf{P}^{(\cdot,\cdot)}$.

We define $\mathbf{V}$ as a matrix whose columns form an orthonormal basis for the null space of $\mathbf{C}(\boldsymbol{\theta})$, which can be found through SVD [26]. The CCRB [15] can then be calculated by

$$\mathbb{E}\left[\left\| \hat{\boldsymbol{\theta}} - \boldsymbol{\theta} \right\|^2 \right] \geq \mathrm{Tr}\left(\left(\mathbf{V}^{\top} \boldsymbol{\Omega}(\boldsymbol{\theta}) \mathbf{V} \right)^{-1} \right) , \quad (14)$$

where $\boldsymbol{\Omega}(\boldsymbol{\theta})$ is the Fisher information matrix defined in (6). The CCRB only exists if $N_c \geq N_\theta - \mathrm{rank}(\boldsymbol{\Omega}(\boldsymbol{\theta}))$ and the constraints are linearly independent, which are sufficient conditions for the inverse of $\mathbf{V}^T\, \boldsymbol{\Omega}(\boldsymbol{\theta})\, \mathbf{V}$ to exist as the constrained FIM then becomes full-rank. The matrix on the right-hand side in (14) becomes the OB if the columns of $\mathbf{V}$ form an orthonormal basis for the null space of the FIM in (9). With unknown parameters, we do not know constraints that give such a $\mathbf{V}$. Since OB has a larger parameter space, including those provided in (10) and (11), the OB is smaller than the CCRB [17], that is,

$$\mathrm{Tr}\left(\boldsymbol{\Omega}(\boldsymbol{\theta})^{\dagger} \right) \leq \mathrm{Tr}\left(\left(\mathbf{V}^{\top} \boldsymbol{\Omega}(\boldsymbol{\theta}) \mathbf{V} \right)^{-1} \right) . \quad (15)$$

The rank of the constrained FIM is equal to the dimension of the manifold of Tucker tensors as provided in Theorem 3.6 in [27],

TABLE 1. The Summary of the Scenarios and the Condition Number of Constrained FIM Introduced in Section VI

	Multilinear Ranks	Mode-n Singular Value Separation	Condition Number of $\mathbf{V}^T\,\boldsymbol{\Omega}(\boldsymbol{\theta})\,\mathbf{V}$
Scenario 1	(3,3,3,3,3,3,3)	d=2	$2\cdot 10^3$
	(3,3,3,3,3,3)	d=2	$3\cdot 10^3$
	(3,3,3,3,3)	d=2	$4\cdot 10^3$
	(3,3,3,3)	d=2	$6\cdot 10^3$
Scenario 2	(5,5,5,5)	d=2	$2\cdot 10^3$
	(5,5,5,5)	d=1	$8\cdot 10^3$
	(5,5,5,5)	d=0.5	$20\cdot 10^3$
	(5,5,5,5)	d=0.25	$72\cdot 10^3$

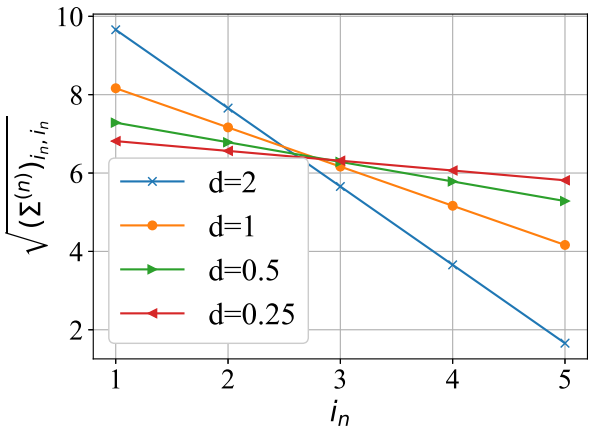

FIGURE 3. Mode-*n* singular values of the 4th-order tensors with multilinear rank (5,5,5,5) generated in Scenario 2 as described in Section VI.

that is,

$$N_\theta - N_c = \sum_{n=1}^{N} I_n R_n + \prod_{n=1}^{N} R_n - \sum_{n=1}^{N} R_n^2\,, \tag{16}$$

which confirms that $(\mathbf{V}^T\,\boldsymbol{\Omega}(\boldsymbol{\theta})\,\mathbf{V})$ is of full rank and the inverse exists.

VI. SIMULATION

In this section, we analyze the performance of existing low multilinear rank tensor estimation methods, tr-MLSVD [7] and HOOI [10], against the computed lower bounds. We make sure that the simulated tensors have unique mode-n singular values by following [28], where a tensor that satisfies the prescribed mode-n singular values is generated. Specifically, we select mode-n singular values that are separated by a $d \in \{2, 1, 0.5, 0.25\}$ in each mode, that is,

$$\sqrt{(\boldsymbol{\Sigma}^{(n)})_{i_n-1,i_n-1}} - \sqrt{(\boldsymbol{\Sigma}^{(n)})_{i_n,i_n}} = d\,, \tag{17}$$

for $n \in \{1,\ldots,N\}$ and $i_n \in \{2,\ldots,I_n\}$. We investigate two scenarios. For both scenarios, we fix the squared sum of the mode-n singular values to 200, that is,

$$\mathrm{Tr}(\boldsymbol{\Sigma}^{(n)}) = \mathrm{Tr}(\mathbf{S}_{(n)}\,\mathbf{S}_{(n)}{}^T) = 200\,. \tag{18}$$

for $n \in \{1,\ldots,N\}$ and generate equally spaced singular values according to the desired multilinear ranks defined therein.

In the first scenario, the effect of the order on CCRB, OB, and MSE is observed. We increase the order of the tensor from 4 to 7, without changing the rank or the singular value distribution in each order. The setup of Scenario 1 is summarized in Table 1, and the mode-n singular values are plotted in Fig. 3. In the second scenario, the effect of the distribution of the mode-n singular values on CCRB, OB, and MSE is analyzed. Therefore, we fixed the multilinear ranks and the order of the tensor, and changed the distance between each consecutive mode-n singular values using $d \in \{2, 1, 0.5, 0.25\}$.

The CCRB and OB are calculated using (14) and (9) for an SNR range of {0, 5, 10, 15, 20, 25, 30} dB. The following steps are taken to achieve a given SNR. Each element of the noise tensor $\underline{\mathbf{M}}$ is sampled from a normal distribution. The scaling factor is then obtained by the ratio of the norm of $\underline{\mathbf{L}}$ to the norm $\underline{\mathbf{M}}$ scaled by an SNR-related parameter, i.e.,

$$\sigma = \frac{1}{\sqrt{10^{\mathrm{SNR}/10}}} \cdot \frac{\left\|\underline{\mathbf{L}} - \bar{\underline{\mathbf{L}}}\right\|}{\left\|\underline{\mathbf{M}} - \bar{\underline{\mathbf{M}}}\right\|}\,, \tag{19}$$

where $\bar{\underline{\mathbf{M}}}$ and $\bar{\underline{\mathbf{L}}}$ are tensors where each element is assigned as the mean of $\underline{\mathbf{M}}$ and $\underline{\mathbf{L}}$, respectively. The resulting σ is used consecutively to scale each entry of $\underline{\mathbf{M}}$. This is a standard practical approach for SNR control, though it introduces a small theoretical discrepancy with the unconstrained Gaussian model (2) as each entry of $\underline{\mathbf{M}}$ is now dependent on each other through the denominator in (19).

The CCRB and OB are compared with the estimates from HOOI [10] and tr-MLSVD [7], which are averaged over 10^4 Monte Carlo simulations. The tolerance of HOOI is set to 10^{-6}, the maximum iteration is set to 2000, and it is initialized with the tr-MLSVD. The true ranks are used to truncate in tr-MLSVD. Define the estimate of the true vector at the tth Monte Carlo iteration as $\hat{\boldsymbol{\theta}}^{(t)}$. The mean squared error is calculated by

$$\mathrm{MSE} = \frac{1}{10^4}\sum_{t=1}^{10^4}\left\|\hat{\boldsymbol{\theta}}^{(t)} - \boldsymbol{\theta}\right\|^2\,, \tag{20}$$

for both estimators after fixing the trivial ambiguities and enforcing all-orthogonality of the core tensor estimates of tr-MLSVD and HOOI. We fix the sign ambiguity and permutation ambiguity by using the Hungarian algorithm. The sign of the estimated singular vector and the corresponding slice of the estimated core tensor are flipped, accordingly. We enforce all-orthogonality of the estimated core tensor by applying an additional MLSVD decomposition and absorbing the rotation into the estimated factor matrices.

VII. RESULTS

The results are plotted for the two scenarios in Figs. 4 and 5, respectively. In Fig. 4, we can observe that as the orders increase, the distance between the OB and the CCRB decreases, and a tight connection between the estimates and the CCRB is observed for all SNRs. As the order decreases, a deviation from the CCRB is observed for both estimators, which can be seen in the bottom right subplots of Figs. 4 and 5. The effect of the gap between the consecutive mode-n singular values on the CCRB is shown using gaps from the set $d \in \{2, 1, 0.5, 0.25\}$. We see that as the gaps decrease, a deviation of the MSE of estimates from the CCRB is observed, and the distance between OB and the CCRB increases.

We tabularized the summary of the two scenarios in Table 1 along with the condition number of the constrained FIM (14) for the tensors that are generated. Increasing the tensor order leads to a reduction in the condition number, indicating that the estimation problem becomes better conditioned. Similarly, greater separation between consecutive mode-n singular values also results in a lower condition number. In well-conditioned cases, such as those shown in the top left subplots of Figs. 4 and 5, the CCRB approaches the oracle bound (OB) closely.

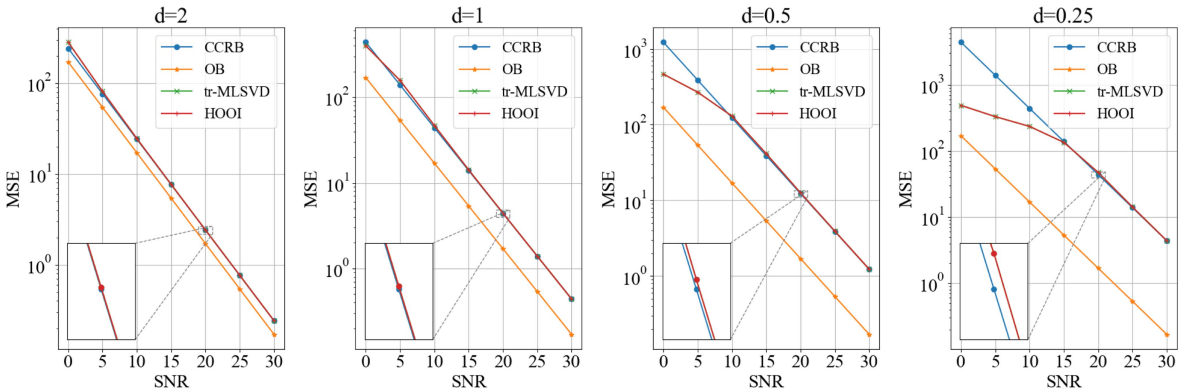

FIGURE 4. The CCRB, OB, and the MSE between the true parameters of interest θ and $\hat{\theta}$ obtained by HOOI and tr-MLSVD for tensors in Scenario 1, where the distribution between the mode-*n* singular values is fixed and the order is changed. The CCRB, OB, and MSE are defined in (14), (9), and (20), respectively. The setup of Scenario 1 is described in Table 1.

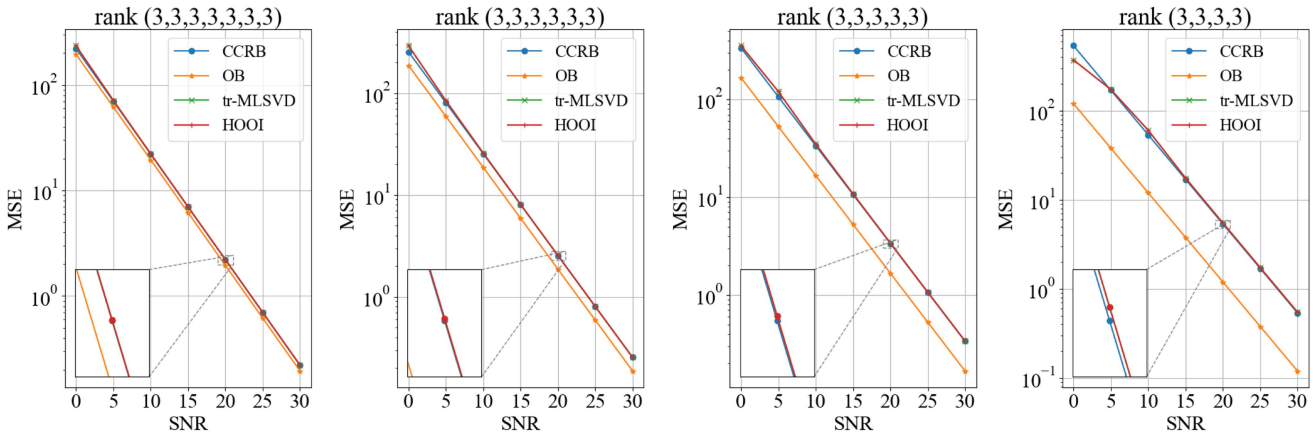

FIGURE 5. The CCRB, OB, and the MSE between the true parameters of interest θ and $\hat{\theta}$ obtained by HOOI and tr-MLSVD for tensors in Scenario 2, where the order is fixed but the gap between the consecutive mode-*n* singular values is changed. The CCRB, OB, and MSE are defined in (14), (9), and (20), respectively. The setup of Scenario 2 is described in Table 1.

VIII. DISCUSSION AND CONCLUSION

In this paper, we introduced a lower bound on the mean squared error of the unbiased estimates of the components of HOSVD under additive WGN using CCRB. Asymptotically, for high SNRs and high Monte Carlo simulations, we showed through simulation that the HOOI and tr-MLSVD converge fairly close to the CCRB. In the non-asymptotic case, a deviation is observed from the CCRB. The SNR in which the convergence occurs is dependent on the condition number of $\mathbf{V}^T\,\boldsymbol{\Omega}(\boldsymbol{\theta})\,\mathbf{V}$, which is shown in Table 1. If the condition number is high, both estimators get close to the CCRB in higher SNRs. Such a case occurs when the singular values of the true tensor are close to each other, and the order of the true tensor is low. For well-conditioned scenarios, the OB and CCRB get close to each other, which suggests that the theoretical gain provided by any constraints beyond those presented in (10) and (11) will be marginal.

We have observed that the MSE of the estimates can lie between CCRB and OB in low SNRs if the condition number is high. This can be seen in Fig. 5. Both tr-MLSVD and HOOI fail to hold the constraint (11) in the non-asymptotic region. In the noisy setting, the off-diagonals of the truncated core tensor are not zero. The lack of this restriction can explain their better performance compared to CCRB in Fig. 5. The authors in [29] introduce the CCRB using Lehman-unbiasedness, which is a weaker restriction than the unbiasedness definition in [15], which might further improve the bound in the non-asymptotic region.

The two estimators that are used in this paper do not enforce all-orthogonality into the low-rank approximation. Note that although the HOSVD core is all-orthogonal, the truncation step, which removes columns, rows, and fibers of the core, may render the new trMLSVD core not all-orthogonal. Although we apply an additional MLSVD operation such that the tr-MLSVD core is all-orthogonal, inherently the estimators does not incorporate such a constraint. The HOOI estimates $\underline{\mathbf{L}}$ by minimizing $||\underline{\mathbf{Y}}-\underline{\mathbf{L}}||^2$, with the constraint that $\underline{\mathbf{L}}$ has a low multilinear rank, which is the mean squared estimate of a low-rank approximation. Similar to the tr-MLSVD, the all-orthogonality of the core tensor is not enforced during estimation.

Additionally, the unbiasedness assumption of the CCRB is violated for both HOOI and tr-MLSVD. When a matrix is perturbed, its eigenvectors and eigenvalues change, and the magnitude of this change depends on the noise power and the spectral gap between the eigenvalues [30]. While classical perturbation theory, such as in [31] and references therein, shows that subspace estimates (e.g., eigenvectors) may be asymptotically unbiased under high SNR and well-separated eigenvalues, this assumption does not always hold in finite-sample scenarios. In particular, when eigenvalues are closely spaced or repeated, the SVD becomes

non-unique due to rotational ambiguity among the corresponding eigenvectors.

In the multilinear case, although the mode-n unfoldings are processed independently, the resulting components (factor matrices and the core tensor) are assembled jointly. Therefore, noise-induced perturbations in each unfolding can influence the overall decomposition and estimation accuracy. This effect is more difficult to characterize analytically than in the matrix case and has been noted in multilinear perturbation studies [12]. Our simulations confirm this, showing that the MSE of the estimates can fall below the CCRB, which is indicative of estimator bias. For well-separated mode-n singular values and high SNR scenarios, the estimates from tr-MLSVD and HOOI lie fairly close to the CCRB, indicating an unbiased estimation regime. Thus, while asymptotic unbiasedness is theoretically valid under ideal conditions, in general, tensor decomposition can often deviate from this assumption.

Another reason the MSE of the estimates can lie between the CCRB and OB could be due to the sign ambiguity resolution explained in Section VI. In cases where the mode-n singular values are not well-separated and the SNR is low, using the sign of the true factor matrices to fix the sign ambiguity decreases the MSE. The convention of HOSVD is to order the mode-n singular values according to the norm of the slices. After the clean tensor is corrupted with noise, such an ordering of the estimated factors may not be the best way to align with the true factors. Especially in low SNR scenarios, the ordering of the noisy mode-n singular values might change, and the corresponding resolution to fix the sign ambiguity might fail.

We propose additional directions for future work. The constrained and the unconstrained FIM given in (14) and (7) can get fairly large with increasing ranks and orders. We have not conducted a study to invert the matrices efficiently. Matrix inversion lemma, such as the one used in [1], can be used to invert the FIM in a less expensive way. The CCRB is calculated assuming that the noise is WGN. Only the log-likelihood in (21) depends on the noise type. Given that the unbiasedness and the regularity conditions hold, the constrained Cramér-Rao bound can be calculated for other additive noise types. The subspace spanned by the factor matrices is commonly used in signal processing applications [32], [33]. If the subspaces are of interest, the intrinsic Cramér-Rao bound [34] can be readily extended to a higher-order case.

Many low-rank tensor decomposition methods, such as HOOI [10] tr-MLSVD [7], or algorithms based on optimization on manifolds [35] [27], provide theoretical upper bounds to the low-rank approximation problem. We introduced the CCRB as an information-theoretic lower bound for the low multilinear rank approximation problem. More significantly, we identify the condition number of $\mathbf{V}^T\,\boldsymbol{\Omega}(\boldsymbol{\theta})\,\mathbf{V}$ as a metric that explains how the locally unbiased estimators closely approach the CCRB, particularly in non-asymptotic and high condition number regimes. In such regions, the OB and CCRB are separated, suggesting that there are constraints other than (10) and (11), which could yield a lower MSE. We leave the investigation of more optimal estimation techniques in the non-asymptotic regions as future work.

APPENDIX

Let $\mathbf{e}_{i_n} \in \mathbb{R}^{I_n}$, $\mathbf{e}_{r_n} \in \mathbb{R}^{R_n}$, and $\mathbf{e}_{\overline{r_1 \dots r_N}} \in \mathbb{R}^{\prod_{n=1}^{N} R_n}$ be the basis vectors that have 1 located at i_n, r_n and $r_1 + \cdots + (r_N - 1)R_1 \dots R_{N-1}$, respectively, and 0s elsewhere. We can then write the partial derivative of $\ln f(\underline{\mathbf{Y}}; \boldsymbol{\theta})$ with respect to the elements of $\boldsymbol{\theta}$ as

$$\begin{aligned}
\frac{\partial \ln f(\underline{\mathbf{Y}}; \boldsymbol{\theta})}{\partial (\mathbf{U}^{(n)})_{i_n, r_n}} &= \frac{1}{\sigma^2} \mathbf{e}_{i_n}^T (\mathbf{Y}_{(n)} - \mathbf{L}_{(n)})(\mathbf{U}^{(N)} \otimes \cdots \otimes \mathbf{U}^{(n+1)} \\
&\quad \otimes \mathbf{U}^{(n-1)} \otimes \cdots \otimes \mathbf{U}^{(1)})\, \mathbf{S}_{(n)}^T\, \mathbf{e}_{r_n}\,, \\
\frac{\partial \ln f(\underline{\mathbf{Y}}; \boldsymbol{\theta})}{\partial (\underline{\mathbf{S}})_{r_1, \dots, r_N}} &= \frac{1}{\sigma^2} \mathbf{e}_{\overline{r_1 \dots r_N}}^T (\mathbf{U}^{(N)} \otimes \cdots \otimes \mathbf{U}^{(1)})^T \\
&\quad (\mathrm{vec}(\underline{\mathbf{Y}}) - \mathrm{vec}(\underline{\mathbf{L}}))\,. \qquad (21)
\end{aligned}$$

The submatrices of the FIM in (7) are defined for $n \in \{1, \dots, N\}$ as

$$\begin{aligned}
(\boldsymbol{\Omega}_{\mathbf{U}^{(n)}\mathbf{U}^{(n)}})_{\overline{i_n r_n}, \overline{i_n r_n}} &= \mathbb{E}\left[\frac{\partial \ln f(\underline{\mathbf{Y}}; \boldsymbol{\theta})}{\partial (\mathbf{U}^{(n)})_{i_n, r_n}} \frac{\partial \ln f(\underline{\mathbf{Y}}; \boldsymbol{\theta})}{\partial (\mathbf{U}^{(n)})_{i_n, r_n}} \right], \\
(\boldsymbol{\Omega}_{\mathbf{U}^{(n)}\underline{\mathbf{S}}})_{\overline{i_n r_n}, \overline{r_1 \dots r_N}} &= \mathbb{E}\left[\frac{\partial \ln f(\underline{\mathbf{Y}}; \boldsymbol{\theta})}{\partial (\mathbf{U}^{(n)})_{i_n, r_n}} \frac{\partial \ln f(\underline{\mathbf{Y}}; \boldsymbol{\theta})}{\partial (\underline{\mathbf{S}})_{r_1, \dots, r_N}} \right], \\
(\boldsymbol{\Omega}_{\underline{\mathbf{S}}\underline{\mathbf{S}}})_{\overline{r_1 \dots r_N}, \overline{r_1 \dots r_N}} &= \mathbb{E}\left[\frac{\partial \ln f(\underline{\mathbf{Y}}; \boldsymbol{\theta})}{\partial (\underline{\mathbf{S}})_{r_1, \dots, r_N}} \frac{\partial \ln f(\underline{\mathbf{Y}}; \boldsymbol{\theta})}{\partial (\underline{\mathbf{S}})_{r_1, \dots, r_N}} \right]. \qquad (22)
\end{aligned}$$

For the submatrices in the diagonal of the FIM in (7), we have the following expressions after simplifications for $n \in \{1, \dots, N\}$,

$$\begin{aligned}
(\boldsymbol{\Omega}_{\mathbf{U}^{(n)}\mathbf{U}^{(n)}})_{\overline{i_n r_n}, \overline{i_n r_n}} &= \frac{1}{\sigma^2} \mathbf{e}_{r_n}^T\, \mathbf{S}_{(n)}\, \mathbf{S}_{(n)}^T\, \mathbf{e}_{r_n} \delta_{i_n i_n}\,, \\
(\boldsymbol{\Omega}_{\underline{\mathbf{S}}\underline{\mathbf{S}}})_{\overline{r_1 \dots r_N}, \overline{r_1 \dots r_N}} &= \frac{1}{\sigma^2} \delta_{r_1 r_1} \dots \delta_{r_n r_n}\,. \qquad (23)
\end{aligned}$$

For $n \in \{1, \dots, N\}$, $k \in \{1, \dots, N\}$, and $k > n$, we have the cross partial derivatives of the factor matrices in the sub-diagonals of the FIM in (7) with the elements

$$\begin{aligned}
&(\boldsymbol{\Omega}_{\mathbf{U}^{(n)}\mathbf{U}^{(k)}})_{\overline{i_n r_n}, \overline{i_k r_k}} \\
&= \frac{1}{\sigma^2} \mathbf{e}_{r_n}^T\, \mathbf{S}_{(n)} (\mathbf{U}^{(N)} \otimes \cdots \otimes \mathbf{U}^{(n+1)} \otimes \mathbf{U}^{(n-1)} \cdots \otimes \mathbf{U}^{(1)})^T \\
&\quad \mathbf{G}^{(n; i_n, k; i_k)} \\
&\quad (\mathbf{U}^{(N)} \otimes \cdots \otimes \mathbf{U}^{(k+1)} \otimes \mathbf{U}^{(k-1)} \otimes \cdots \otimes \mathbf{U}^{(1)})\, \mathbf{S}_{(k)}^T\, \mathbf{e}_{r_k}\,. \qquad (24)
\end{aligned}$$

Additionally, we have the cross partial derivatives of factor matrices and the core tensor of the FIM (7) for $n \in \{1, \dots, N\}$ as

$$\begin{aligned}
&(\boldsymbol{\Omega}_{\mathbf{U}^{(n)}\underline{\mathbf{S}}})_{\overline{i_n r_n}, \overline{r_1 \dots r_N}} \\
&= \frac{1}{\sigma^2} \mathbf{e}_{r_n}^T\, \mathbf{S}_{(n)} (\mathbf{U}^{(N)} \otimes \cdots \mathbf{U}^{(n+1)} \otimes \mathbf{U}^{(n-1)} \cdots \otimes \mathbf{U}^{(1)})^T \\
&\quad \mathbf{G}^{(n; i_n)} (\mathbf{U}^{(N)} \otimes \cdots \otimes \mathbf{U}^{(1)}) \mathbf{e}_{\overline{r_1 \dots r_N}}\,. \qquad (25)
\end{aligned}$$

Finally, the partial derivatives of the constraints given in (13) with respect to the factor matrices have the columns described with

$$(\mathbf{C}_{\mathbf{U}^{(n)}})_{\overline{i_n r_n}} = \mathbf{E}^{(n)} \mathrm{vec}\left(\mathbf{U}^{(n)T} \mathbf{e}_{i_n} \mathbf{e}_{r_n}^T + \mathbf{e}_{r_n} \mathbf{e}_{i_n}^T \mathbf{U}^{(n)} \right), \qquad (26)$$

and the partial derivatives of the constraints given in (13) with respect to the mode-n unfoldings of the core tensor have the columns described with

$$\begin{aligned}
(\mathbf{C}_{\mathbf{S}_{(n)}})_{\overline{r_1 \dots r_{n-1} r_{n+1} \dots r_N}} = \mathbf{K}^{(n)} \mathrm{vec}\Big(&\mathbf{S}_{(n)}\, \mathbf{e}_{\overline{r_1 \dots r_{n-1} r_{n+1} \dots r_N}} \mathbf{e}_{r_n}^T \\
&+ \mathbf{e}_{r_n} \mathbf{e}_{\overline{r_1 \dots r_{n-1} r_{n+1} \dots r_N}}^T \mathbf{S}_{(n)}^T \Big)\,, \qquad (27)
\end{aligned}$$

for $n \in \{1, \dots, N\}$. The matrices $\mathbf{E}^{(n)}$, $\mathbf{K}^{(n)}$, $\mathbf{P}$, and $\mathbf{G}$ are described in Section II.

REFERENCES

[1] N. D. Sidiropoulos, L. De Lathauwer, X. Fu, K. Huang, E. E. Papalexakis, and C. Faloutsos, "Tensor decomposition for signal processing and machine learning," *IEEE Trans. Signal Process.*, vol. 65, no. 13, pp. 3551–3582, Jul. 2017.

[2] B. Hunyadi, S. Van Huffel, and M. De Vos, "The power of tensor decompositions in biomedical applications," *Proc. Mach. Learn. Healthcare Technol.*, 2016, pp. 83–109.

[3] A. Cichocki et al., "Tensor networks for dimensionality reduction and large-scale optimization: Part 2 applications and future perspectives," *Found. Trends Mach. Learn.*, vol. 9, no. 6, pp. 431–673, 2017.

[4] A. Cichocki et al., "Tensor decompositions for signal processing applications: From two-way to multiway component analysis," *IEEE Signal Process. Mag.*, vol. 32, no. 2, pp. 145–163, Mar. 2015.

[5] J. Zuiddam, "A note on the gap between rank and border rank," *Linear Algebra Appl.*, vol. 525, pp. 33–44, 2017.

[6] P. K. Kothari, A. Moitra, and A. S. Wein, "Overcomplete tensor decomposition via koszul-young flattenings," 2024, *arXiv:2411.14344*.

[7] 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.

[8] L. R. Tucker et al., "The extension of factor analysis to three-dimensional matrices," *Contributions Math. Psychol.*, vol. 110119, pp. 110–182, 1964.

[9] C. Eckart and G. Young, "The approximation of one matrix by another of lower rank," *Psychometrika*, vol. 1, no. 3, pp. 211–218, 1936.

[10] L. De Lathauwer, B. De Moor, and J. Vandewalle, "On the best rank-1 and rank-($r_1, r_2, \ldots, r_n$) approximation of higher-order tensors," *SIAM J. Matrix Anal. Appl.*, vol. 21, no. 4, pp. 1324–1342, 2000.

[11] G. W. Stewart, "Perturbation theory for the singular value decomposition," Univ. Maryland, Inst. Adv. Comput. Stud., College Park, MD, USA, Tech. Rep. CS-TR-2539, Sep. 1990.

[12] W. Hackbusch, D. Kressner, and A. Uschmajew, "Perturbation of higher-order singular values," *SIAM J. Appl. Algebra Geometry*, vol. 1, no. 1, pp. 374–387, 2017.

[13] C. R. Rao, "Information and the accuracy attainable in the estimation of statistical parameters," in *Breakthroughs in Statistics: Foundations and Basic Theory*. Berlin, Germany: Springer, 1992, pp. 235–247.

[14] J. D. Gorman and A. O. Hero, "Lower bounds for parametric estimation with constraints," *IEEE Trans. Inf. Theory*, vol. 36, no. 6, pp. 1285–1301, Nov. 1990.

[15] P. Stoica and B. C. Ng, "On the Cramér-Rao bound under parametric constraints," *IEEE Signal Process. Lett.*, vol. 5, no. 7, pp. 177–179, Jul. 1998.

[16] T. J. Moore, B. M. Sadler, and R. J. Kozick, "Maximum-likelihood estimation, the Cramér–Rao bound, and the method of scoring with parameter constraints," *IEEE Trans. Signal Process.*, vol. 56, no. 3, pp. 895–908, Mar. 2008.

[17] Y.-H. Li and P.-C. Yeh, "An interpretation of the Moore-Penrose generalized inverse of a singular Fisher information matrix," *IEEE Trans. Signal Process.*, vol. 60, no. 10, pp. 5532–5536, Oct. 2012.

[18] R. T. Rajan, "Relative space-time kinematics of an anchorless network," Ph.D. dissertation, (TU Delft), Delft Univ. Technol., Delft, The Netherlands, 2016.

[19] E. D. Carvalho, J. Cioffi, and D. Slock, "Cramer-Rao bounds for blind multichannel estimation," in *Proc. IEEE. Glob. Telecommun. Conf.. Conf. Rec.*, 2000, vol. 2, pp. 1036–1040.

[20] R. T. Rajan and A.-J. van der Veen, "Joint ranging and synchronization for an anchorless network of mobile nodes," *IEEE Trans. Signal Process.*, vol. 63, no. 8, pp. 1925–1940, Apr. 2015.

[21] S. J. Wijnholds and A.-J. van der Veen, "Effects of parametric constraints on the CRLB in gain and phase estimation problems," *IEEE Signal Process. Lett.*, vol. 13, no. 10, pp. 620–623, Oct. 2006.

[22] X. Liu and N. D. Sidiropoulos, "Cramér-Rao lower bounds for low-rank decomposition of multidimensional arrays," *IEEE Trans. Signal Process.*, vol. 49, no. 9, pp. 2074–2086, Sep. 2001.

[23] C. Prévost, K. Usevich, M. Haardt, P. Comon, and D. Brie, "Constrained Cramér–Rao bounds for reconstruction problems formulated as coupled canonical polyadic decompositions," *Signal Process.*, vol. 198, 2022, Art. no. 108573.

[24] R. Boyer, L. De Lathauwer, and K. Abed-Meraim, "Higher order tensor-based method for delayed exponential fitting," *IEEE Trans. Signal Process.*, vol. 55, no. 6, pp. 2795–2809, Jun. 2007.

[25] Z. Ben-Haim and Y. C. Eldar, "The Cramér-Rao bound for estimating a sparse parameter vector," *IEEE Trans. Signal Process.*, vol. 58, no. 6, pp. 3384–3389, Jun. 2010.

[26] G. H. Golub and C. F. Van Loan, *Matrix Computations*. Baltimore, MD, USA: JHU Press, 2013.

[27] M. M. Steinlechner, "Riemannian optimization for solving high-dimensional problems with low-rank tensor structure," Ph.D. thesis, EPFL, Lausanne, Switzerland, 2016.

[28] M. Calis and B. Hunyadi, "Generating a tensor with prescribed mode-n singular values," *TechRxiv*, 2025.

[29] E. Nitzan, T. Routtenberg, and J. Tabrikian, "Cramér–Rao bound for constrained parameter estimation using Lehmann-unbiasedness," *IEEE Trans. Signal Process.*, vol. 67, no. 3, pp. 753–768, Feb. 2019.

[30] C. Davis and W. M. Kahan, "The rotation of eigenvectors by a perturbation. III," *SIAM J. Numer. Anal.*, vol. 7, no. 1, pp. 1–46, 1970.

[31] M. Viberg, B. Ottersten, and T. Kailath, "Detection and estimation in sensor arrays using weighted subspace fitting," *IEEE Trans. Signal Process.*, vol. 39, no. 11, pp. 2436–2449, Nov. 1991.

[32] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propag.*, vol. TAP-34, no. 3, pp. 276–280, Mar. 1986.

[33] A.-J. van der Veen, E. F. Deprettere, and A. L. Swindlehurst, "Subspace-based signal analysis using singular value decomposition," *Proc. IEEE*, vol. 81, no. 9, pp. 1277–1308, Sep. 1993.

[34] S. T. Smith, "Covariance, subspace, and intrinsic Cramér-Rao bounds," *IEEE Trans. Signal Process.*, vol. 53, no. 5, pp. 1610–1630, May 2005.

[35] M. Ishteva, P.-A. Absil, S. Van Huffel, and L. De Lathauwer, "Best low multilinear rank approximation of higher-order tensors, based on the riemannian trust-region scheme," *SIAM J. Matrix Anal. Appl.*, vol. 32, no. 1, pp. 115–135, 2011.